Credit Risk and Nonlinear Filtering: Computational Aspects and Empirical Evidence

Thesis by

Agostino Capponi

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended May 27, 2009)

To my family.

Acknowledgements

I would like to express my gratitude to the outstanding Caltech environment which allowed me to conduct my research in a stimulating and very constructive way. I consider myself very fortunate for having the opportunity to work in this highly stimulating intellectual atmosphere. I would like to express my gratitude to my advisor, Prof. Jakša Cvitanić, for providing guidance and assistance throughout the period of my PhD research. He has always been available and supportive and helped me to become familiar with the field of credit risk. I would also like to thank Prof. K. Mani Chandy for his great participation in my research work during all my stay at Caltech and for serving on my committee. I would like to thank Prof. John O. Ledyard for serving on my committee and giving me the possibility to present the results of my research work at SISL seminars on a regular basis. Those seminars have been sources of ideas and constructive criticism. They had a significant impact on my PhD work and improved substantially my presentation skills. I would also like to thank the other member of my committee, Prof. Yaser Abu-Mostafa for reading my work. I would like to thank the SISL grants that supported my PhD research during all these years at Caltech. I would also like to thank those people who made life at Caltech more enjoyable; thanks to all colleagues and friends for their company and for sharing with me some of the best moments in my life. Finally but most importantly, I would like to express my deepest gratitude to my parents and Concetta for their endless love and unconditional support; without their sacrifices, it would have been impossible for me to reach this point of my life.

Abstract

This thesis proposes a novel credit risk model which deals with incomplete information on the firm's asset value. Such incompleteness is due to reporting bias deliberately introduced by insider managers and executives of the firm and unobserved by outsiders.

The pricing of corporate securities and the evaluation of default measures in our credit risk framework requires the solution of a computationally unfeasible nonlinear filtering problem. We propose a polynomial time-sequential Bayesian approximation scheme which employs convex optimization methods to iteratively approximate the optimal conditional density of the state on the basis of received market observations. We also provide an upper bound on the total variation distance between the actual filter density and our approximate estimator. We use the filter estimator to derive analytical expressions for the price of corporate securities (bond and equity) as well as for default measures (default probabilities, recovery rates, and credit spreads) under our credit risk framework. We propose a novel statistical calibration method to recover the parameters of our credit risk model from market price of equity and balance sheet indicators. We apply the method to the Parmalat case, a real case of misreporting and show that the model is able to successfully isolate the misreporting component. We also provide empirical evidence that the term structure of credit default swaps quotes exhibits special patterns in cases of misreporting by using three well known cases of accounting irregularities in US history: Tyco, Enron, and WorldCom.

We conclude the thesis with a study of bilateral credit risk, which accommodates the case in which both parties of the financial contract may default on their payments. We introduce the general arbitrage-free valuation framework for counterparty risk adjustments in presence of bilateral default risk. We illustrate the symmetry in the valuation and show that the adjustment involves a long position in a put option plus a short position in a call option, both with zero strike and written on the residual net value of the contract at the relevant default times. We allow for correlation between the default times of each party of the contract and the underlying portfolio risk factors. We introduce stochastic intensity models and a trivariate copula function on the default times exponential variables to model default dependence. We provide evidence that both default correlation and credit spread volatilities have a relevant and structured impact on the adjustment. We also study a case involving British Airways, Lehman Brothers, and Royal Dutch Shell, illustrating the bilateral adjustments in concrete crisis situations.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Credit risk is the risk of default or of reductions in market value caused by changes in the credit quality of issuer or counterparties. The distribution of credit losses is complex. At its center is the probability of default, by which we mean any type of failure to honor a financial agreement. The estimation of the probability of default requires specifying a model of investor uncertainty, a model of the available information and its evolution over time, and a model for the default event. However, default probabilities alone would not be sufficient to price credit-sensitive securities. We would need to know how much the marketplace charges for holding risky assets, meaning assets subject to risk of default. This in turn requires the specification of a model of recovery upon default, and most importantly a model of the premium that investors require as compensation for bearing credit risk.

Although credit denotes the extension of access to any liquid assets today in return for a promise to pay in the future, when we think of credit we typically think of the debt that one party owes to another and this will also be the case for this thesis. In a debt transaction, there is usually a lender and a borrower. A common form of credit is a bond, which is an interest-bearing certificate issued by a government or business promising to pay the holder a specified sum on a specified date.

Bonds can be issued by the government, the so called Treasury bonds, or by corporations, thereby named corporate bonds. Treasury bonds are typically assumed to be risk-free in countries with stable economies, and this is supported by the fact that the government can always either raise taxes or print out more money as a last resort to pay back debt. Under the simplest assumptions and using semi-annual compounding, the price of a treasury coupon paying bond with face value 100 dollars would be given by

$$B_{treas} = \sum_{n=1}^{2T} \frac{c/2}{(1+\frac{y}{2})^{n/2}} + \frac{100}{(1+y)^T} \tag{1.1.1}$$

where c is the paid coupon and y the treasury yield. Corporate bonds instead are exposed to a default risk with magnitude depending on the financial health of the particular issuer and its ability to raise revenue. Under the same scenario as above, the price of a corporate coupon paying bond is given by:

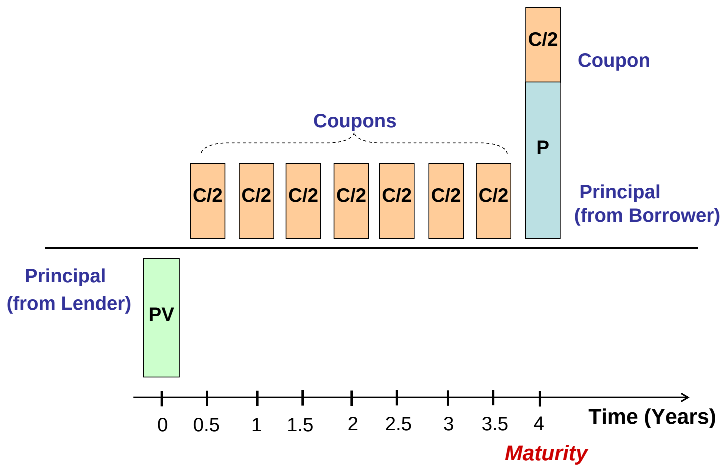

Figure 1.1: Schematic diagram of a bond's cash flow

$$B_{corp} = \sum_{n=1}^{2T} \frac{c/2}{(1+\frac{y+s}{2})^{n/2}} + \frac{100}{(1+\frac{y+s}{2})^{T}} \tag{1.1.2}$$

It appears from Eq. (1.1.1) and Eq. (1.1.2) that the price difference between the two bonds depends on the term s, which is called the *credit spread*. The larger the credit spread, the smaller the relative price of corporate bond with respect to treasury bond. The credit spread is a key concept and denotes the additional compensation that the holders of risky assets are demanding for bearing the default risk. If no additional compensation is offered to holders of the corporate bonds, then they would be better off buying the treasury bond.

Intuitively, we can decompose the yield spread into the three categories, see Figure 1.2. The *default risk* is the risk due to the potential loss associated with a single default and depends on the amount recovered. If the default time and recovery value were known, then it would be possible to diversify the risk and construct a portfolio consisting of names having specified likelihood of losses which guarantees on average the risk-free rate. In years when default did not occur, the bond would return a little more due to the expected loss premium. However, in the event of default, it would return much less. Since an investor could obtain the risk-free base rate not just “on average”, but all the time by buying the risk free bond, the risky bond must provide additional compensatory return. This indicates that there is an amount of nondiversifiable credit risk and that the market provides compensation for this unavoidable risk bearing. Such component is called *risk premium* and includes uncertainty regarding default timing, recovery value, and accuracy of revealed information for which investors ought to be compensated. The final source of risk is *liquidity risk*, and refers to the chance that the company may have insufficient cash flow to meet its obligations due to the lack of marketability of an investment that cannot be bought or sold quickly enough by the company to fulfill its obligation.

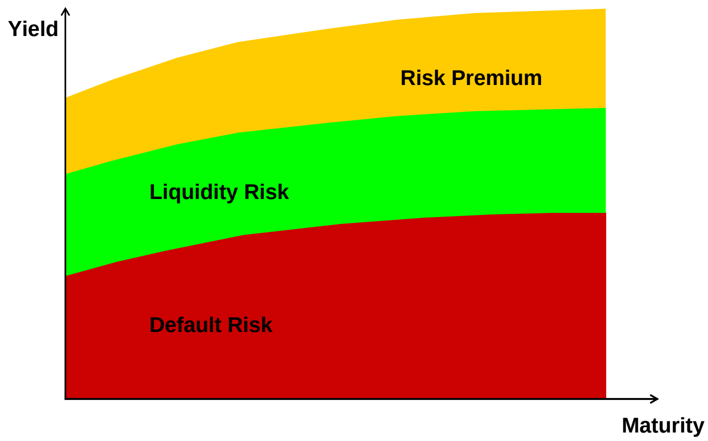

Figure 1.2: Credit risk decomposition

Much research is devoted to separate out the effects of all of these variables when computing the spread of corporate bond yields to US Treasury. Default modeling is becoming a major problem nowadays due to the credit crunch crisis which we are experiencing, and it is also an important tool which is daily used for hedging credit exposures. This importance is confirmed by Figure 1.3 which shows the distribution of defaulted debt as a percentage of the total amount of outstanding debt.

It is evident from Figure 1.3 that defaulted debt accounts for more than 5% of the total outstanding debt, thus motivating the enormous amount of academic and industrial research devoted to this topic.

1.2 Summary of Contributions and Overview of the Thesis

The main contributions of this thesis are

- Development of a novel credit risk model. Such model accounts for the possibility that the firm is misrepresenting its accounting data. This model adds to the branch of credit risk literature focusing on the role of incomplete information on pricing. The level of incompleteness derives from the fact that the true asset value observed by outsiders may be biased and thus needs to be filtered out from market- available information. We provide analytical expressions for corporate security prices and default measures under the proposed modeling framework. These results are presented in Chapter 3 and have been published in the finance journal *International Journal of Theoretical and Applied Finance* (see Capponi and Cvitanić, 2009).

- Development of a stochastic Bayesian filtering algorithm for jump linear jump systems with state-dependent transitions. The mathematical framework of our credit risk model turns out to be a Hidden Markov model, and more specifically a Markovian jump linear system with state-dependent transitions.

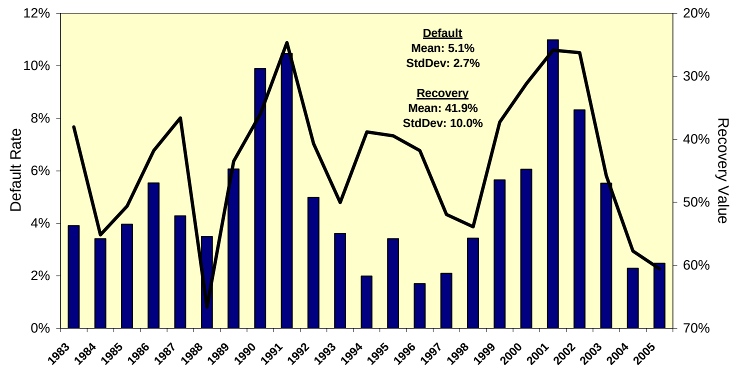

Figure 1.3: Defaulted bonds as a percentage of face value outstanding. Source: E. Altman and S. Jha. *Market Size and Investment Performance of Defaulted Bonds and Bank Loans: 1987-2006.* NYU Salomon Center Publication, 2006.

We show that the problem of evaluating the optimal filtering density is computationally intractable and then propose a polynomial time approximation method with guaranteed error bound to compute the density. These results are presented in Chapter 5 and have been published in the journal *Automatica* (see Capponi, 2009a).

- Development of a novel calibration algorithm for structural models of credit risk with incomplete information. The algorithm combines maximum likelihood estimation methods along with option price inversion approaches to recover drift, volatility, accounting noise variance, and reporting bias. These results are presented in Chapter 4 and have been published in the proceedings of the *IEEE Conference on Computational Intelligence for Financial Engineering*, where they were granted with the best student paper award (see Capponi, 2009b).

- Framework for bilateral counterparty risk valuation. We introduce the general arbitrage-free valuation framework for counterparty risk adjustments in presence of bilateral default risk, including default of both investor and her counterparty. We show that the adjustment involves a long position in a put option plus a short position in a call option, both with zero strike and written on the residual net value of the contract at the relevant default times. We then specialize our analysis to Credit Default Swaps and show the impact of credit spread volatility and default correlation on the bilateral credit risk value adjustment. These results are presented in Chapter 6. A subset of these results has been submitted for peer review publication to *Risk* magazine (see Brigo and Capponi) and a more extended version of the work is currently in preparation for *Finance and Stochastics*.

Chapter 7 draws conclusions and discusses future problems of interest. More results and technical proofs are provided in the appendices.

Chapter 2

Preliminaries and Definitions

In this chapter we provide notation and terminology which will be heavily used in the remainder of this thesis. We also briefly describe the general taxonomy of credit risk models which is needed to understand the results of this thesis. Those include structural models which use option pricing theory to evaluate credit risk, reduced-form models using term structure theory to explain credit spread behavior and models with incomplete information which provide a trade-off between the two. This is not intended to be a complete survey and the interested reader is referred to books on the topic (see Lando, 2004).

2.1 Notation and Terminology

This section presents the basic notation and terminology.

- $n(x; \mu, \sigma)$: Gaussian density with mean μ and standard deviation σ
- $n(\mathbf{x}; \mu, \mathbf{\Sigma})$: multivariate Gaussian density with mean μ and covariance $\mathbf{\Sigma}$
- $D(t, T)$: price of a risk-free zero coupon bond maturing at T, as seen at time t
- $B(t, T)$: price of a defaultable zero coupon bond maturing at T, as seen at time t
- $\tau := T - t$: time to maturity
- $y(t, T)$: yield of a corporate zero coupon bond maturing at T, as seen at time t
- r: risk-free rate
- $CS(t, T) := y(t, T) - r$: credit spread at maturity T as seen at time t. It is defined as the difference between the yield of a defaultable zero coupon bond and a corresponding risk-free zero coupon bond
- ς: the default time indicator
- $PD(t, T)$: probability of defaulting at T as seen at time t
- $RR(t, T)$: recovery rate at time T as seen at time t

- $BS_{call}(t, X_t, K, T, \sigma)$: time t price of a Black Scholes call option with strike price K, volatility σ, maturity T, and initial asset vale X_t
- $BS_{put}(t, X_t, K, T, \sigma)$: time t price of a Black Scholes put option with strike price K, volatility σ, maturity T, and initial asset vale X_t
- $\lambda^{i,j}(\mathbf{y}) = P(\theta_k = j|\theta_{k-1} = i, \mathbf{x}_k = \mathbf{y})$: mode switching probabilities
- $\lambda^{j}(\mathbf{y}) = P(\theta_0 = j|\mathbf{x}_0 = \mathbf{y})$: prior mode probability
- $p(\mathbf{x}|\mathbf{y}, l) = P(\mathbf{x}_k = \mathbf{x}|\theta_{k-1} = l, \mathbf{x}_{k-1} = \mathbf{y})$: mode dependent transition density
- $p(\mathbf{x}) = P(\mathbf{x}_0 = \mathbf{x})$: the initial density on $\mathbf{x}_0$
- $p^l_{k|k}(x) := P(x_k = x, \theta_{k-1} = l|\mathscr{F}^z_k)$: the joint state-mode posterior hybrid density
- $p^l_{k,k}(\mathbf{x}) := P(x_k = x, \theta_{k-1} = l, \mathscr{F}^z_k)$: the unnormalized posterior density
- $p_{k|k}(\mathbf{x}) = P(\mathbf{x}_k = \mathbf{x}|\mathscr{F}^z_k)$: the posterior density
- $L^l_{k|k}(\mathbf{x}) = p(\mathbf{z}_k|\mathbf{x}_k = \mathbf{x}, \theta_{k-1} = l)$: the mode conditioned measurement likelihood
- $E_f[g] = \int_{\mathbb{R}^n} f(\mathbf{w})g(\mathbf{w})d\mathbf{w}$: the expectation of g with respect to the density f, where f and g are functions
- $\mathscr{F}^x_k = \sigma(x_1, x_2, \ldots, x_k)$: the filtration generated by the state process x up to time k
- $\mathscr{F}^z_k = \sigma(z_1, z_2, \ldots, z_k)$: the filtration generated by the observation process z up to time k
- $\hat{x}_{l|k} := E[x_l|\mathscr{F}^z_k]$: the conditional expectation of x at time l given the observation filtration at time k
- $\sigma^2_{l|k} := E[(x_l - \hat{x}_{l|k})^2|\mathscr{F}^z_k] \qquad l > k$: the conditional variance of x at time l given the observation filtration at time k

2.2 Structural Models

In structural models, corporate liabilities are evaluated by decomposing their pay-offs in linear and nonlinear products, and using standard option pricing theory to price them. Those models make explicit assumptions about the dynamics of a firm's assets and its capital structure, which are then used to determine the occurrence of default. The literature on structural models goes back to Merton (1974), where the firm defaults if, at the time of servicing the debt, its assets are below its outstanding debt. A more general approach was introduced by Black and Cox (1976) who relax the Merton's assumption and model default as the first passage time of the firm's asset value below a certain threshold. Further generalizations treat coupon bonds, the effect of bond indenture provisions (see Geske, 1977), stochastic interest rates (see Longstaff and Schwartz, 1995, Collin-Dufresne et al., 2004), and endogenous default barriers optimally triggered by equity owners when the asset fall to a sufficiently low level (see Leland and Toft, 1996). Since this thesis builds

	Assets	Bonds	Equity
No default	$V_T \geq K$	K	$V_T - K$
Default	$V_T < K$	V_T	0

Table 2.1: Payoffs at maturity in Merton model

upon the Merton model, we present it in some detail in this section and refer the interested reader to the above references for a complete overview of structural models.

2.2.1 The Merton Model

Merton assumes that the firm is financed by equity and a zero coupon bond with face value K and maturity date T. The firm's contractual obligation is to repay the amount K to the bond investors at time T. Debt covenants grant bond investors absolute priority: if the firm cannot fulfill its payment obligation, then bond holders will immediately take over the firm. The payoff to equity and bond holders are summarized in Table 2.1.

If the asset value V_T exceeds or equals the face value K of the bond, the bond holders will receive their promised payment K and the shareholders will get the remaining $V_T - K$. However, if the value of assets V_T is less than K, the ownership of the firm will be transferred to the bondholders, who lose the amount $K - V_T$. Equity is worthless because of limited liability. The asset value of the firm is assumed to follow a log-normal diffusion process

$$dV_t = \mu V_t dt + \sigma V_t dW_t \tag{2.2.1}$$

where μ is a drift parameter, $\sigma > 0$ is a volatility parameter, and W_t is a standard Brownian motion. Setting $m = \mu - 0.5\sigma^2$, Ito's lemma implies that

$$V_t = V_0 e^{mt+\sigma W_t} \tag{2.2.2}$$

Figure 2.1 depicts the situation graphically. The bond payoff in Table 2.1 can be rewritten as

$$B(T,T) = K - \max[K - V_T, 0] \tag{2.2.3}$$

thus implying that the value of the defaultable zero coupon bond at any time t is given by

$$B(t,T) = Ke^{-r(T-t)} - BS_{put}(t, V_t, K, T, \sigma) \tag{2.2.4}$$

where r denotes the risk-free discount factor. The $BS_{put}(.)$ is also referred in the literature as the *default put option*, since it measures the default risk of the bond.

As for equity, it follows from Table 2.1 that the final payoff can be expressed as

$$E(T,T) = \max[V_T - K, 0] \tag{2.2.5}$$

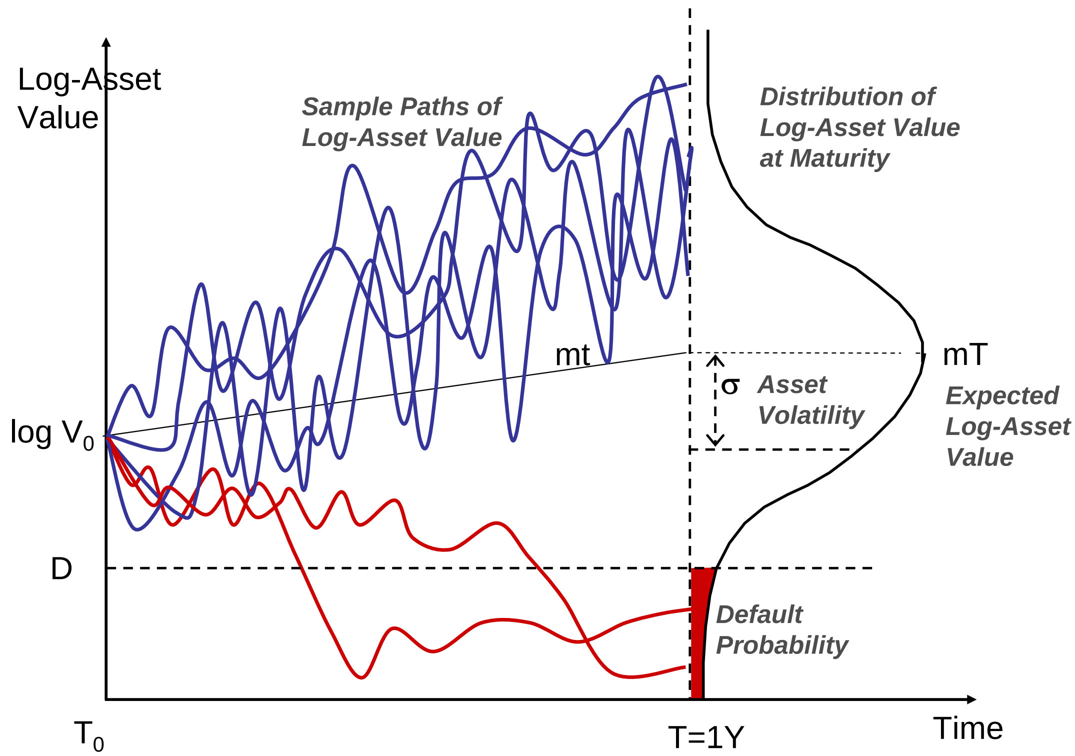

Figure 2.1: The asset paths in the figure represent the uncertain future states of the firm's assets, with the distribution shown at time T representing the density of final asset states at that time. The probability of default, PD, is given as the density of the distribution below the level of the debt (i.e., the proportion of states in which the firm is insolvent at time T).

thus implying that the equity value at any time t is given by

$$E(t,T) = BS_{call}(t, V_t, K, T, \sigma) \tag{2.2.6}$$

Since W_T is normally distributed with mean zero and variance T, the probability $PD(t,T)$ of defaulting at time T as seen at time t is given by

$$\begin{aligned} PD(t,T) &= P(V_T < K) \\ &= P(\sigma(W_T - W_t) < L - m\tau) \\ &= N\left(\frac{L - m\tau}{\sigma\sqrt{\tau}}\right) \end{aligned} \tag{2.2.7}$$

where $L = \log(\frac{K}{V_t})$, and $\tau = T - t$ denotes the time to maturity.

Structural models are particularly elegant and informative, mainly because they impose an arbitrage relationship between equity and debt. From this point of view, they provide a natural guideline for relative value trades between the stock market and the credit derivatives market which is of utmost practical interest. Moreover, models are forward looking and allow incorporating investors expectations of a firm's future performance. Unfortunately, these models result in a generally poor fit to market data due to several limitations, some of which are outlined next.

- Typically, reasonable values for leverage of the firm and volatility of assets produce lower credit spreads than those observed on the market.
- Undervaluation is particularly relevant for short term maturities: a typical credit term structure in structural models shows a hump and zero intercept.
- Undervaluation is particularly relevant for high credit standing obligors.
- Bond prices play no role in estimating the value of the firm.

The first two items above represent an important concern for all structural models and have motivated a lot of research, including this thesis. Those concerns arise from the fact that all structural models are based on the assumption that the firm's dynamic is regulated by a diffusion process and that the value of the firm can be observed directly. Since diffusion processes have continuous sample paths and default is the first hitting time of a barrier, then default is a predictable stopping time, and this leads to an underestimation of the short-term credit spread as outlined above. From a mathematical point of view, this can be explained using a key result from diffusion theory (see Karatzas and Shreve, 1995), stating that for any diffusion process X_t

$$\lim_{h \to 0} \frac{P(|X_{t+h} - X_t| \geq \epsilon)}{h} = 0 \tag{2.2.8}$$

We next show why this fact implies zero spread in the short term. Let us consider a risky zero coupon bond paying the fact value K at time T if the issuing firm does not default, and zero otherwise. Differently from the bond in the Merton model, such bond does not pay anything in case of default, thus it must have a lower price and consequently a higher yield than the corresponding bond in the Merton model. The price of such bond is obtained as

$$\begin{aligned} B(t,T) &:= e^{-r(T-t)}E[K\mathbf{1}_{V_T>K}] \\ &= Ke^{-r(T-t)}P(V_T>K) \end{aligned} \tag{2.2.9}$$

The yield at time t of a bond maturing at T with principal K is defined as

$$y(t,T) = \frac{1}{T-t}\log\left(\frac{K}{B(t,T)}\right) \tag{2.2.10}$$

and since r is the yield of a zero-coupon risk free bond, we obtain that the time t credit spread for a bond maturing at $t+h$ is given by

$$CS(t,t+h) = -\frac{1}{h}\log\frac{B(t,t+h)}{K} - r \tag{2.2.11}$$

If we start at time t from a non-default state ($V_t > K$), then we have that

$$\begin{aligned} \lim_{h\to 0} CS(t,t+h) &= -\frac{1}{h}\log P(V_{t+h}>K|V_t>K) \\ &\approx -\frac{1}{h}(P(V_{t+h}>K|V_t>K)-1) \\ &= \frac{1}{h}P(V_{t+h}\leq K|V_t>K) \end{aligned} \tag{2.2.12}$$

Using Eq. (2.2.8), we can conclude that

$$\lim_{h\to 0} CS(t,t+h) = 0 \tag{2.2.13}$$

2.3 Actual versus Risk-Neutral Default Probabilities

The default probability given in Eq. (2.2.7) has been derived under historical measure, but it can be combined with an equilibrium model of underlying expected returns to produce estimates of expected returns. We can express an interesting relationship between the physical and risk-neutral world via the the capital asset pricing model (CAPM). The premium over the risk-free rate that an investor should require to invest in a risky asset depends on its holding period, volatility, and the current market price of risk. The CAPM framework states

$$\begin{aligned} \mu - r &= \beta(\mu_M - r) \\ &= \frac{cov(\mu, \mu_M)}{\sigma_M^2}(\mu_M - r) \\ &= \frac{cov(\mu, \mu_M)}{\sigma_M \sigma} \frac{\mu_M - r}{\sigma_M} \sigma \end{aligned} \tag{2.3.1}$$

where μ_M is the market return and β is the covariance of the asset's return with the market return relative to the variance of the market. If we denote by ρ the correlation between the return of the asset and the market return and

$$\lambda = \frac{\mu_M - r}{\sigma_M} \tag{2.3.2}$$

we can express the excess market return as:

$$\mu - r = \lambda \rho \sigma \tag{2.3.3}$$

The parameter λ is called the market price of risk. It measures the tradeoffs between risk and return that are made for securities depending on the asset values V_t. For an intuitive understanding of Eq. (2.3.3) notice that the variable σ can be interpreted as the quantity of risk present in the asset value. On the right-hand side, we are therefore multiplying the quantity of risk with its price, weighted by the correlation between asset return and market return (the larger such correlation, the larger the potential gain resulting from the investment in the risky asset V_t). The left-hand side is the expected return in excess of the risk-free interest rate that is required to compensate for this risk. In other words, the term $\lambda\rho\sigma$ measures the extent that the excess return required by investors on the firm's securities is affected by the dependence on the asset. If $\lambda\rho\sigma > 0$, investors require a higher expected return to compensate for this risk; if $\lambda\rho\sigma < 0$, the dependence of the security on the firm's asset value causes investors to require a lower return; if $\lambda\rho\sigma = 0$, then the investor does not require any extra compensation, which would imply that he is not taking any risk, thus its return should be the same as investing at the risk-free rate r. When this is the case (i.e., $\mu = r$, then we say that we are in a *risk-neutral* world, and this is the framework assumed by the modern asset pricing theory to determine the price of securities.

2.4 Reduced Form Models

Reduced form models go back to Artzner and Delbaen (1995), Jarrow and Turnbull (1995), and Duffie and Singleton (1999). Unlike structural models, they do not argue why a firm defaults, but model default as a Poisson-type event which occurs completely unexpectedly. The default time is defined as

$$\varsigma = \inf \left\{ t : \int_0^t \lambda_s ds \geq Exp(1) \right\} \tag{2.4.1}$$

where $Exp(1)$ denotes an exponentially distributed random variable with parameter 1. The process λ_s is a stochastic process called the *intensity* process. It is an ad-hoc combination of financial variables often fit to market spreads and is exogenously related to the the firms dynamics.

Eq. (2.4.1) implies immediately that the probability of non defaulting up to time T is given by

$$P(\varsigma > T) = e^{-\int_0^T \lambda_s ds} \tag{2.4.2}$$

which leads to the pricing formula for the bond

$$\begin{aligned} B(t,T) &:= e^{-r(T-t)} E[K\mathbf{1}_{\varsigma>T}] \\ &= Ke^{-r(T-t)} P\{\varsigma > T\} \\ &= Ke^{-r(T-t)} e^{-\int_t^T \lambda_s ds} \end{aligned} \tag{2.4.3}$$

If λ_s becomes constant, then Eq. (2.4.3) simplifies to

$$B(t,T) = e^{-(r+\lambda)(T-t)} \tag{2.4.4}$$

and thus the credit spread is given by the constant λ.

The nice feature of reduced form model is that as $t \to T$, the credit spreads do not approach zero, but stay positive; we have seen an example above for the case when the intensity is constant. Therefore, they remove the disturbing feature of zero-credit spreads in the short term. However, as evident from Eq. (2.4.1), there is no natural link of defaults to the underlying dynamics of the firm‘s cash flows and financial statements, thus those models are often criticized because they lose the micro-economic interpretation of the default time.

2.5 Incomplete Information Models

The incomplete information framework attempts to unify structural and reduced form models under a common perspective by taking the advantages of both methods. Underlying all credit models is the default process. In traditional structural models default can be anticipated, and as seen in Section 2.2, there is no short-term credit risk that would require compensation. In reduced form models, it is assumed that default cannot be anticipated, so there is short-term credit risk by assumption, see Section 2.4. The default process is parameterized through an intensity λ and model default probabilities and security prices are immediately implied by the exogenous intensity dynamics. Instead of focusing on the default intensity and making ad-hoc assumptions about its dynamics, *incomplete information* models specify the trend based on a model definition of default. There are mainly two approaches to introduce short-term uncertainties into structural models. The first possibility is to include jumps in the firm value, see Zhou (2001). In this situation, there is always a chance that the firm value jumps below the default barrier, and thus default cannot be anticipated.

However, there is also a chance that the firm diffuses to the barrier, as in traditional structural models, and in this case default can be anticipated. Thus, depending on the state of the world, there may or may not be short-term credit risk. There is another approach which guarantees that default cannot be anticipated and thus produces short-term credit risk. This approach is obtained after reconsidering the informational assumptions underlying the traditional structural credit risk models which assume that the information used to calibrate and run the model is observed perfectly. Such information includes the firm value process along with its parameters, and the default barrier. In the incomplete information framework, we assume that the information about these quantities is imperfect; this means that we are not sure either of the true value of the firm or of the condition of the firm that will trigger default and consequently default becomes a complete surprise. Several models with incomplete information have been proposed in the literature, some of which are discussed next. Cetin et al. (2004) and Guo et al. (2009) propose an approach in which the market is assumed to only partially observe, and possibly with a lag, relevant information concerning the state of the firm. Brody et al. (2007) and Brody et al. (2008) derive bond pricing formulas in an economic model where information about the actual cash flows of the debt obligation are obscured to market participants by a Gaussian noise process which vanishes when the time of each required cash flow is reached. Giesecke and Goldberg (2004) add incompleteness to structural models by assuming that the default barrier is a stochastic process, thus investors cannot deduce the distance to default from the firm's fundamentals as in Merton or Black-Cox models. Frey and Runggaldier (2007) consider a model in which the intensity is driven by unobserved state processes, and the calculation of measures of risk such as default probability leads to a nonlinear filtering problem. In all these cases, default becomes an inaccessible stopping time for the market, thus yielding a reduced form credit risk model. Duffie and Lando (2001) propose a model with endogenous default threshold, but in which the market only observes noisy or delayed accounting reports from which investors have to draw inference of the true asset value of the firm. This model creates a nonzero instantaneous hazard rate of default, thus implying a nonzero short- term credit spread. Even further, they prove that structural models with incomplete information are fully consistent with intensity models in the sense that the instantaneous hazard rate of default is the intensity of the default indicator process with respect to the information set observed in the market (noisy accounting reports), and not by the firm (true state). We next report a sketch of the main results obtained in Duffie and Lando (2001), being those highly related to the ones presented in this thesis, and we will provide a more detailed comparison in Chapter 3. Their assumption is that the firm's value is observed up to additive noise, namely $Y_t = V_t + U_t$ where U_t is a white noise sequence. Let

$$\varsigma = \inf\{t : V_t \leq K\} \tag{2.5.1}$$

and denote by $f(x)$ the density on the initial asset value of the firm. Then we have

$$\lim_{h \to 0} \frac{1}{h} \int_0^\infty P(\varsigma \leq h | V_0 = x) f(x) dx = \frac{1}{2}\sigma^2 f'(K) \tag{2.5.2}$$

as opposed to the case when the initial asset value of the firm is known with certainty, in which case, using the result in Eq. (2.2.8), we would obtain that $PD(0,h) = 0$.

Chapter 3

Credit Risk Modeling With Misreporting

3.1 Introduction

In a statement released in 2007, the US Treasury Secretary complained about the huge number of accounting data revisions in the US market. He claimed that some 1500 firms revise their balance sheet figures every year, and that means, he claimed, that there must be something wrong with the system. So, the debate on accounting transparency that arose at the beginning of the century is still open. Recent literature has focused on the impact of these events on the evaluation of corporate liabilities, namely equity and bonds. On theoretical grounds, Duffie and Lando (2001) was the first attempt to model the impact of accounting noise on the credit spread term structure. On empirical grounds, Yu (2005) proved that accounting noise is actually priced in the market: a risk premium is charged to the credit spreads of firms that adopt less transparency. Fraud events that took place both in the US (Worldcom, Tyco, Enron, etc.) and in Europe (Cirio, Marconi, Parmalat, etc.) in the first years of the century raised the issue of distinguishing between unbiased noise due to measurement errors and cases of deliberate fraud. Inspired by these cases Cherubini and Manera (2006) model the effect of deliberate misreporting on accounting statements through the introduction of a probability of fraud which the market updates whenever new information about balance sheet is issued. Brigo and Morini (2006) consider the effect of accounting reliability by modeling the ratio between the level of default barrier and the value of company assets as a random variable, where pessimistic scenarios, possibly corresponding to fraud in accounting, are associated with larger values of this ratio. None of the above studies models explicitly the dynamics of misreporting.

We introduce a credit risk framework which incorporates the misreporting event as an intrinsic feature, and is estimable using market and accounting data. We explicitly model the misreporting dynamics and also calibrate a simple version of our model to the data for the Parmalat company around its bankruptcy. The results indicate that the amount of misreporting was not negligible, and that by ignoring it, the model would have resulted in a large overestimation of the firm's volatility.

Misreporting may only arise if the market has incomplete knowledge of the manager's objective function, since then he may be better off with the option to misreport, see Fisher and Verrecchia (2000). If instead the market is assumed to have rational expectations and perfect knowledge of manager's objective, then it can back out perfectly misreporting in equilibrium. We work under the incomplete knowledge assumption since the exact nature of the manager's compensation, his time horizon, and his litigation risk and reputation costs associated with biased reporting are often unavailable to the market.

Although it is not easy to estimate the managerial risk of misreporting, recent studies have started addressing this issue. For example, Wang (2007) proposes a bivariate probit model to recover the probability of committing fraud from the probability of detected fraud. The rest of the chapter is organized as follows. Section 3.2 describes and applies the Wang model to Tyco, a major case of misreporting. Section 3.3 describes the components of the proposed model which deals with misreporting. Section 3.4 gives explicit formulas for bond and equity prices under a Merton framework. Since such formulas are not directly computable, we provide in Section 3.5 computable expressions for bond and equity prices. Section 3.6 provides formulas for the default measures under our credit risk framework.

3.2 Estimating Risk of Misreporting

The objective of this section is to explain that it is possible to estimate the risk of misreporting from financial and accounting data. Therefore, using appropriate calibration methodologies, it would be possible to use the credit risk model proposed in the following sections in industrial contexts to estimate the amount of misreporting and separate it from the effect of asset volatility. As an example, we present here the methodology proposed by Wang (2007) which recovers the amount of fraud from a set of financial and balance sheet indicators. More specifically, she considers a bivariate-probit model as follows. Let F_i^* denote firm i's potential to commit fraud, and D_i^* the firm's i potential of getting caught conditional on the event that the fraud has been committed. Then

$$\begin{aligned} F_i^* &= \mathbf{z}_{F,i}\beta_F + u_i \\ D_i^* &= \mathbf{z}_{D,i}\beta_D + v_i \end{aligned} \tag{3.2.1}$$

where $\mathbf{z}_{F,i}$ is a row vector with elements explaining the firm's i potential to commit fraud, and $\mathbf{z}_{D,i}$ contain elements explaining the firm's i potential to get caught, and u_i, v_i are Gaussian random variables with correlation coefficient ρ.

Since undetected fraud is (by definition) unobservable, they define the random variables

$$\begin{aligned} F_i &= \mathbf{1}_{F_i^*>0} \\ D_i &= \mathbf{1}_{D_i^*>0} \\ Q_i &= F_i D_i \end{aligned} \tag{3.2.2}$$

Here $Q_i = 1$ if the firm has committed fraud and has been detected, while $Q_i = 0$ if the firm has not committed fraud or has committed fraud but it has not been detected.

Using a sample of firms belonging to a heterogeneous number of sectors (see Wang, 2007), they recover the parameters β_F and β_D of their model by maximization of the following likelihood function

$$L(\beta_F, \beta_D) = \sum_{q_i=1} \log[P(Q_i = 1)] + \sum_{q_i=0} \log[P(Q_i = 0)] \tag{3.2.3}$$

where $P(Q_i = 1) = \phi(\mathbf{z}_{F,i}, \beta_F, \mathbf{z_{D,i}}, \beta_D, \rho)$ for a given function ϕ (see Wang, 2007).

We take the model coefficient estimates as computed by their procedure, and report their values in Table 3.1.

Predictors	β_F
Return on Asset (ROA)	1.72 (2.46)
Growth in External Financing (EFG)	3.64 (4.89)
Research and Development (R&D)	5.28 (3.05)
Investing cash flow (ICF)	1.42 (1.71)
Outsider ownership	1.21 (1.03)
(Outsider ownership)^2	1.45 (0.76)
Board size	-0.02 (-0.35)
Log asset value (Log-Ass)	0.09 (0.65)
Age	0.01 (1.78)
Technology	0.09 (0.21)
Service	-0.32 (-0.52)
Trade	0.88 (-1.28)

Table 3.1: Model predictors

The predictors Technology, Service, and Trade represents the percentile of the business involved in technology, service, and trade respectively. The ICF predictor is the amount of money spent in investing over the book value of assets. Outsider ownership is calculated as the ratio of the inside directors over the total number of directors.

We next report the time series of predictors calculated for the firm Tyco, a well-known case of misreporting in United States history. The data used to compute those predictors have been taken from the Edgar database on a three-month basis for the period ranging from January 2001 to December 2003. Such time frame includes the misreporting period which covers the years 2001 and 2002. The tech component of Tyco was 100%, with service and trade accounting for 0 %. The board consisted of 11 directors, 8 of which came from outside. Thus, for the period of interest the value for the board size predictor was 11, while the value for the outside ownership predictor was 73%.

The other predictor values are reported in Table 3.2

The above predictors are all expressed in percentile, except for the log-asset value which is expressed in million of dollars, and age which is expressed in years. We can now use Eq. (3.2.1) for each date, and calculate the probability of committing fraud for Tyco across time. This is reported in Figure 3.1, which shows a probability of misreporting (around 80%) during the time when misreporting did occur.

The discussion above further supports our argument that it is worth considering misreporting risk as an

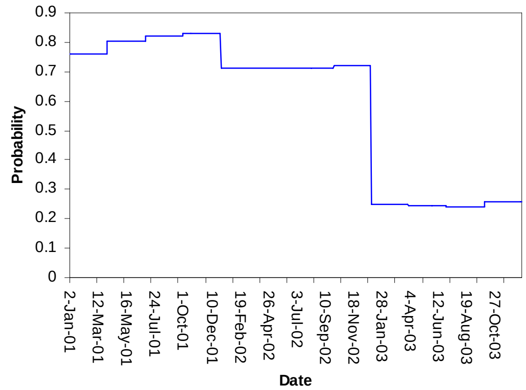

Figure 3.1: The probability of fraud as a function of time obtained for Tyco using the model 17 proposed in Wang (2007)

additional risk factor in credit models (besides volatility risk and accounting noise) since it would be possible to recover it from accounting data using appropriate models such as the one shown in this section.

3.3 Model Definition

We consider a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ with the following system of stochastic difference equations, a generalized version of a Hidden Markov Model:

$$V_k = e^{x_k} \tag{3.3.1}$$

$$x_k = x_{k-1} + (\mu(\theta_{k-1}) - 0.5\sigma(\theta_{k-1})^2)\Delta_k + \sigma(\theta_{k-1})v_k \tag{3.3.2}$$

$$\theta_k = \Gamma(\theta_{k-1}, x_k, \varrho, w_k) \tag{3.3.3}$$

$$z_k = x_k + h(\theta_{k-1}) + \nu(\theta_{k-1})u_k \tag{3.3.4}$$

Here, V_k describes the evolution of the asset value of the firm and is modeled as a discretized geometric Brownian motion, with x_k being the log-asset value process, and $x_0 = \mathcal{N}(\mu_0, \sigma_0^2)$, and drift μ and volatility

Dates	ROA	EFG	ICF	Log-Ass	Age
3/31/2001	5.2%	23.7%	12%	10.9	41
6/30/2001	5.2%	24.6%	8%	11.6	41.3
9/30/2001	5.2%	23.9%	0.01%	11.2	41.6
12/31/2001	5.2%	24.5%	1%	11.6	41.9
3/31/2002	-10%	3.7%	2.2%	11.7	42
6/30/2002	-10%	5.3%	4%	11.2	42.3
9/30/2002	-10%	5.5%	4.3%	11.1	42.6
12/31/2002	-10%	5.5%	1.6%	11.1	42.9
3/31/2003	-1.6%	2%	2.5%	11.1	43
6/30/2003	1.6%	2%	3.5%	11.1	43.3
9/30/2003	1.6%	1.9%	3.6%	11.1	43.6
12/31/2003	1.6%	1.9%	0.2%	11	43.9

Table 3.2: Values of predictors

σ both depending on the parameter θ. Moreover, z_k describes the released observation to the outsiders, with Δ_k denoting the time between consecutive observations, typically associated with release of balance sheet reports which often occurs on a quarterly basis. We assume that $\{v_k\}$, $\{w_k\}$, $\{u_k\}$ are independent sequences of i.i.d. Gaussian random variables with zero mean and unit variance.

Moreover, $\Theta = \{\theta^{(1)} = 1, \theta^{(2)} = 2, \ldots, \theta^{(m)} = m\}$ is a finite set of integer modes of cardinality m, Γ is a measurable mapping $\Theta \times \mathbb{R} \times \mathbb{R} \to \mathbb{M}$, and h is a measurable mapping $\Theta \to \mathbb{R}$ assumed to be time-invariant for notational simplicity only.

The interpretation is as follows. We can think of θ as a random variable which designates the *report model* used by the manager of the firm to release the observations to the investors. Such report model affects both the future evolution of the actual asset value, see Eq. (3.3.1), and the value released to the outsiders by the manager of the firm, see Eq. (3.3.4). More precisely, the released log-asset value z_k depends on the report model θ_{k-1} in place during the time interval $[t_{k-1}, t_k]$ through the function h which models the amount of misreporting associated with a given report model. Depending on whether $h(\theta_{k-1})$ is positive or negative, an overstatement or an understatement of the actual performance of the firm will occur when the report model θ_{k-1} is selected by the manager. The situation of no distortion occurring can be modeled by having $h(\theta_{k-1}) = 0$. The parameter $\nu(\theta_{k-1})$ captures the variance of accounting noise associated with the report model θ_{k-1}.

The set of report models is assumed to be finite. Eq. (3.3.3) models the choice of the current report model θ used by the firm's manager as dependent on the last report model, the state of the firm at the time immediately preceding the release of the observation and other factors affecting misreporting represented by the vector ϱ. Such factors can be the quality of corporate governance, the litigation and reputation costs associated with getting caught, or the manager bonus triggering threshold. As already mentioned in the introduction, outsiders do not always have a perfect knowledge of managerial objectives. We model this lack of information with a Gaussian random variable w_k, and assume that the market estimates the model θ used by the manager using the function Γ which depends on the optimal manager's choice of the report model. Such choice is the solution of an optimization problem, where the manager maximizes his expected utility function of misreporting, typically depending on the manager's stock ownership and equity compensation minus his disutility of getting caught. Such optimization problem depends on the true state x_k, the previous report model θ_{k-1} and w_k. The dependence on the previous report model is introduced in our model to denote the fact that the misreporting event also depends on the past managerial behavior.

3.4 Structural Model with Misreporting

3.4.1 The Pricing Framework

Assuming a Merton-type structural model, we propose a valuation framework in which bond and equity prices can be calculated as risk-neutral conditional expectations, i.e., assuming that the investor is neutral with respect to risk, as explained in Section 2.3. We assume a fixed maturity T of the debt, and that the default event can only occur at maturity, which happens if the actual asset value is below the nominal value of the debt, assumed to be constant. We define the random variable

$$\varsigma = \begin{cases} T & \text{if } V_T \leq K \\ \infty & \text{if } V_T > K \end{cases} \tag{3.4.1}$$

and denote by $\mathscr{F}_t^{z,\varsigma} = \mathscr{F}_t^z \vee \sigma(s \wedge \varsigma, s \leq t)$ the sigma algebra generated by the observations enlarged with the information generated by the default indicator random variable ς. Before proceeding with the analysis, we state a useful result from Jeanblanc and Rutkowski (2000).

Proposition 3.4.1 (Projection Formula). *Let A be a bounded, $\mathscr{F}_T^z$-measurable random variable. Then for every $t \leq T$:*

$$E[\mathbf{1}_{\varsigma>T}A|\mathscr{F}_t^{z,\varsigma}] = \mathbf{1}_{\varsigma>t}\frac{E[\mathbf{1}_{\varsigma>T}A|\mathscr{F}_t^z]}{P(\varsigma > t|\mathscr{F}_t^z)} \tag{3.4.2}$$

We apply the pricing methodology proposed in Coculescu et al. (2008) to our credit risk framework. Let us denote by P^* the pricing risk-neutral probability, i.e., the probability measure under which the investor is neutral with respect to risk. We define the risk-neutral estimate of the variable V_T as the $\mathscr{F}_T^z$ measurable random variable

$$\hat{V}_T = \frac{1}{Z_T}E^{P^*}[\mathbf{1}_{V(T)>K}V_T|\mathscr{F}_T^z] \tag{3.4.3}$$

where $Z_T = P^*(V_T > K|\mathscr{F}_T^z)$, and define a defaultable contingent claim as an integrable, $\mathscr{F}_T^z$ measurable random variable of the form:

$$d_T = \mathbf{1}_{V_T>K}f(\hat{V}_T) + \mathbf{1}_{V_T\leq K}g(V_T) \tag{3.4.4}$$

The main difference with the complete information models is that the defaultable claims are assumed to be evaluated using the estimate $\hat{V}_T$ when the firm is not in the default state and the true value is only observed at default. The price of a defaultable claim is then computed in Coculescu et al. (2008) as

$$d_t = e^{-r(T-t)}E^{P^*}[\mathbf{1}_{V_T>K}f(\hat{V}_T) + \mathbf{1}_{V(T)\leq K}g(V_T)|\mathscr{F}_t^{z,\varsigma}] \tag{3.4.5}$$

In our case, if the defaultable claim is a bond, $f(x) = K$, $g(x) = x$, with K being the nominal value of the debt. Therefore, we get that the time t-price of the bond is:

$$\begin{aligned} B(t,T) &= e^{-r(T-t)}E^{P^*}[K\mathbf{1}_{V_T>K} + \mathbf{1}_{V_T\leq K}V_T|\mathscr{F}_t^{z,\varsigma}] \\ &= e^{-r(T-t)}(K - E^{P^*}[(K-V_T)^+|\mathscr{F}_t^z]) \end{aligned} \tag{3.4.6}$$

If the defaultable claim is equity, $f(x) = (x-K)$ and $g(x) = 0$. Therefore, we obtain that the time t price of the equity is:

$$E(t,T) = e^{-r(T-t)}E^{P^*}[(\hat{V}_T - K)\mathbf{1}_{V_T>K}|\mathscr{F}_t^{z,\varsigma}] \tag{3.4.7}$$

We have:

$$\begin{aligned}\mathbf{1}_{V_T>K}(\hat{V}_T-K) &= \mathbf{1}_{V_T>K}\left(\frac{E^{P^*}[\mathbf{1}_{V_T>K}V_T|\mathscr{F}_T^z]}{P^*(V_T>K|\mathscr{F}_T^z)}-K\right) \\ &= E^{P^*}[\mathbf{1}_{V_T>K}V_T|\mathscr{F}_T^{z,\varsigma}]-\mathbf{1}_{V_T>K}K \end{aligned} \tag{3.4.8}$$

where the first equation is obtained using definition (3.4.3), while the second equation follows from the projection formula (3.4.2). Therefore,

$$\begin{aligned} E(t,T) &= e^{-r(T-t)}E^{P^*}[E^{P^*}[\mathbf{1}_{V_T>K}V_T|\mathscr{F}_T^{z,\varsigma}]|\mathscr{F}_t^z]-E^{P^*}[\mathbf{1}_{V_T>K}K|\mathscr{F}_t^z] \\ &= e^{-r(T-t)}E^{P^*}[(V_T-K)^+|\mathscr{F}_t^z] \end{aligned} \tag{3.4.9}$$

We next discuss the choice of the pricing measure P^*. Our market is incomplete since the measurement equation of our filtering model exhibits jumps of m possible different sizes, where m is the number of report models used by manager. Assuming that the released log-asset value z_k is the only traded asset, this means that the jump risk cannot be hedged away. In addition, we have a discrete-time model with continuously valued normally distributed noise. The specific measure P^* has to be inferred from market prices, either via modeling of the market price of risk or of the dynamics of the Radon-Nikodim derivative (see Runggaldier, 2004).

3.4.2 Exact Filtering

The calculation of the equity and bond prices requires the computation of a conditional expectation, which means that the time T filtering density $p_{T|t}(x)$ of our filtering model given by Eq. (3.3.2–3.3.4) needs to be evaluated. This in turn requires the calculation of the time t filtering density $p_{t|t}(x)$. Such calculation will be developed in Chapter 5, Section 5.2. Here, we report the final result which is given by

$$p_{t|t}(x)=\frac{\sum_{l=1}^m p_{t,t}^l(\mathbf{x})}{\int_{\mathbb{R}}\sum_{r=1}^m p_{t,t}^r(\mathbf{y})d\mathbf{y}} \tag{3.4.10}$$

where $p_{t,t}^l(\mathbf{x})$ is recursively defined as

$$p_{t,t}^l(\mathbf{x})=L_{t|t}^l(\mathbf{x})\sum_{r=1}^m\int_{\mathbb{R}}\lambda^{r,l}(y)p(x|y,l)p_{t-1|t-1}^r(\mathbf{y})d\mathbf{y} \tag{3.4.11}$$

3.4.3 Equity and Bond Prices

In order to simplify the exposition, the pricing formulas derived in this section assume a filtering model in which $\mu(i)=\mu$, $\sigma(i)=\sigma$, i.e., the expected return rate and the volatility of the asset is not affected by the report model used. We follow an argument similar to the one used in Merton jump diffusion model, i.e., replace μ with r, thus assuming that the investor becomes risk neutral if the expected short-term return yields the same as the risk-free rate, see also Section 2.3 for more details. We assume that the biases $h(i)$ and the accounting risks $\nu(i)$ are the same under the historical and risk neutral measure. This is because the bias in the released log-asset value depends on the vector including factors specific to the executive such as the reputation and litigation costs associated with misreporting. Such risk is "non-systematic" risk and cannot be diversified away, thus we assume the same form under both measures.

Let us denote $\tau = T - t$ and introduce functions $d_1(x)$ and $d_2(x)$ as

$$\begin{aligned} d_1(x) &= \frac{x - \log(K) + (r + 0.5\sigma^2)\tau}{\sigma\sqrt{\tau}} \\ d_2(x) &= d_1(x) - \sigma\sqrt{\tau} \end{aligned} \tag{3.4.12}$$

Moreover, notice that since the mode process $\{\theta_k\}$ take values in a finite set and Γ is a mapping into Θ, Eq. (3.3.3) induces state-dependent mode transition probabilities as follows:

$$\begin{aligned} \lambda^{i,j}(x) &:= P(\theta_k = j|\theta_{k-1} = i, x_k = x) \\ &= \int_{\mathbb{R}} P(w_k, \theta_k = j|\theta_{k-1} = i, x_k = x)dw_k \\ &= \int_{\mathbb{R}} P(\theta_k = j|\theta_{k-1} = i, w_k, x_k = x)f(w_k)dw_k \\ &= \int_{\mathbb{R}} \mathbf{1}_{\{j=\Gamma(i,x,w_k)\}} f(w_k)dw_k \end{aligned} \tag{3.4.13}$$

The price of the bond is given by the following proposition, proved in Appendix A.

Proposition 3.4.2. *The time t price of a bond maturing at T is given by:*

$$B(t,T) = Ke^{-r\tau} - E_{p_{t|t}}[Ke^{-r\tau}N(-d_2(Y)) - e^y N(-d_1(Y))] \tag{3.4.14}$$

where $p_{t|t}$ is the density computed as described in Chapter 3, Eq. (5.2.5), Y is a random variable with density $p_{t|t}$, and the functions d_1 and d_2 are defined in Eq. (3.4.12).

Similarly, the equity price $E(t,T)$ is given by

Proposition 3.4.3. *The price of the equity at time t is given by:*

$$E(t,T) = E_{p_{t|t}}[e^Y N(d_1(Y)) - Ke^{-r\tau}N(d_2(Y))] \tag{3.4.15}$$

3.5 Bond and Equity Price Computation

Although the filtering density (3.4.11) can be explicitly obtained, it is not amenable to an efficient implementation. First of all, the recursive expression (3.4.11) shows that an exponentially increasing number of terms have to interact to obtain the unnormalized density at time t_k. Additionally, it involves the evaluation of non-Gaussian integrals due to the appearance of the terms $\lambda^{j,l}$, and such integrals may in general be computationally expensive to evaluate. Furthermore, the normalization step required to obtain the posterior density in Eq. (3.4.10) involves an evaluation of m spatial integrations, thus increasing the computational burden even further. This makes it therefore practically impossible to compute exactly bond and equity prices in our credit risk framework, since they both require taking expectation with respect to the actual filtering density. In Chapter 5, Section 5.3, we describe a filtering scheme which is used to approximate the density (3.4.10). Such scheme approximates the true density with a Gaussian mixture. Let us indicate with

$$\{\hat{x}^j_{t|t}, \hat{\sigma}^j_{t|t}, \mu^j_t\} \tag{3.5.1}$$

the means, variances, and weights of the components of the such mixture density for the log-asset value at the initial time. Then approximate formulas for equity and bond prices are given in Section 3.5.1.

3.5.1 Approximate Bond and Equity Prices

We provide a directly implementable expression for the price of equity and debt in our proposed model. The formulas will be given as equalities, although they are to be considered as approximations due to the filtering density of the initial log-asset value being approximated by a Gaussian mixture in our nonlinear filtering model. We now state an obvious result as a lemma which will be extensively used hereafter:

Lemma 3.1. *Let Z be a random variable having a density of the form*

$$f_Z(z) = \sum_{i=1}^{n} w_i f_{Z_i}(z), \quad \sum_{i=1}^{n} w_i = 1 \tag{3.5.2}$$

where $f_{Z_i}(z)$ is the pdf of a measurable random variable Z_i. Then the expectation of a function f of Z may be expressed in terms of the expectations of the function f of the component random variables Z_i, each of them taken with respect to the probability density f_{Z_i}

$$E[f(Z)] = \sum_{i=1}^{n} w_i E[f(Z_i)] \tag{3.5.3}$$

For reading purposes, we denote the variance $(\hat{\sigma}^j_{t|t})^2$ of the j-th mixture component as $\hat{\sigma}^{2,j}_{t|t}$. It follows, from Eq. (3.3.2) and from the assumptions made in Section 3.4.3 that both drift and volatility are mode independent, that the conditional density of the log-asset value x_T at maturity has a Gaussian mixture density specified by

$$\hat{p}_{T|t}(x) = \sum_j \mu^j_t n(x; \hat{x}^j_{T|t}, \hat{\sigma}^j_{T|t}) \tag{3.5.4}$$

where $\hat{x}^j_{T|t} = \hat{x}^j_{t|t} + (r - 0.5\sigma^2)\tau$ and $\hat{\sigma}^j_{T|t} = \sqrt{\hat{\sigma}^{2,j}_{t|t} + \sigma^2\tau}$.

Using Lemma 3.1 we have:

$$E[e^{-r\tau}(K - V_T)^+ | \mathscr{F}^z_t] = e^{-r\tau} \sum_j \mu^j_t E[(K - V^j_{T|t})^+] \tag{3.5.5}$$

where $V^j_{T|t}$ is the exponential of a Gaussian random variable with mean $\hat{x}^j_{T|t}$ and standard deviation $\hat{\sigma}^j_{T|t}$ as defined above. Moreover, let

$$d_1^j = \frac{\hat{x}_{t|t}^j - \log(K) + (r + 0.5\sigma^2)\tau}{\sqrt{\hat{\sigma}_{t|t}^{2,j} + \sigma^2\tau}}, \qquad d_2^j = d_1^j - \left(\sqrt{\hat{\sigma}_{t|t}^{2,j} + \sigma^2\tau}\right) \tag{3.5.6}$$

Then the price of the bond is given by the following proposition, proved in Appendix B

Proposition 3.5.1. *The price of the bond at time t is given by:*

$$B(t,T) = Ke^{-r\tau} - \sum_j \mu_t^j (Ke^{-r\tau} N(-d_2^j) - e^{\hat{x}_{t|t}^j + 0.5\hat{\sigma}_{t|t}^{2,j}} N(-d_1^j)) \tag{3.5.7}$$

A similar calculation can be done for the equity price $E(t,T)$ leading to

Proposition 3.5.2. *The price of the equity at time t is given by:*

$$E(t,T) \quad = \quad \sum_j \mu_t^j (e^{\hat{x}_{t|t}^j + 0.5\hat{\sigma}_{t|t}^{2,j}} N(d_1^j) - Ke^{-r\tau} N(d_2^j))$$

3.6 Computation of Default Measures

Credit risk models have in common the goal of explaining two quantities, default probability and loss given default, whose product can then be used to compute the credit spread. In the next subsections we give closed form expressions for conditional default probability, recovery rate, and credit spreads under our framework, and assuming that the proposed filtering approximation scheme is used to approximate the density of the initial asset value.

3.6.1 Default Probability

The default may occur only at time T, and it occurs if $V_T \leq K$. We define the default probability as

$$PD(t,T) = \mathcal{P}(V_T \leq K | V_t > K, \mathscr{F}_t^z) \tag{3.6.1}$$

A calculation presented in Appendix D shows that the default probability in our model is given by a weighted sum of conditional default probabilities, where each weight w is proportional to the weight of the Gaussian density in the mixture and the distance to default of the asset value estimated using the mean of such density.

Proposition 3.6.1. *The probability of default at time T as seen at time t is given by:*

$$PD(t,T) = \sum_j w_t^j \mathcal{P}(Y_j \leq -d_2(\hat{x}_{t|t}^j) | X_j \leq dd_t^j) \tag{3.6.2}$$

where $w_t^j = \dfrac{\mu_t^j N(dd_t^j)}{\sum_{i=1}^m \mu_t^i N(dd_t^i)}$, $dd_t^j = \dfrac{\hat{x}_{t|t}^j - \log(K)}{\hat{\sigma}_{t|t}^j}$, *$d_2$ is defined in Eq. (3.4.12) and (X_j, Y_j) has a bivariate* *normal distribution function with positive correlation $\rho_{X_j Y_j}$:*

$$\mu_{X_j} = \mu_{Y_j} = 0, \sigma^2_{X_j} = 1,\ \sigma^2_{Y_j} = 1 + \frac{\hat{\sigma}^{2,j}_{t|t}}{\sigma^2 \tau},\ \rho_{X_j Y_j} = \frac{\hat{\sigma}^{j}_{t|t}}{\sqrt{\sigma^2 \tau + \hat{\sigma}^{2,j}_{t|t}}} \tag{3.6.3}$$

We can think of X_j and Y_j as two positively correlated random variables, where X_j is measurable using the information available by time t and reveals information about the distance to default of the current asset value estimated using the mean of the j-th Gaussian density in the approximation, while Y_j is measurable using the information available at maturity T and denotes a default event if it is smaller than the default threshold.

The default probability computed using our model reduces to the Merton default probability as the uncertainty around the true value x_t gets to zero, and the conditional pdf $p(x_t|\mathscr{F}^z_t)$ approaches a sum of delta functions $\sum_j \delta(x_t - \hat{x}^j_{t|t})$. Then we obtain

$$\begin{aligned} PD(t,T) &= \sum_j w^j_t \lim_{\hat{\sigma}^j_{t|t} \to 0} \mathcal{P}(Y_j \leq -d_2(\hat{x}^j_{t|t}) | X_j \leq dd^j_t) \\ &= \sum_j w^j_t \lim_{\hat{\sigma}^j_{t|t} \to 0} \mathcal{P}(Y_j \leq -d_2(\hat{x}^j_{t|t})) \\ &= \sum_j w^j_t \mathcal{P}(Y \leq -d_2(x_t)) \\ &= N(-d_2(x_t)) \end{aligned} \tag{3.6.4}$$

where the second line follows because Y_j becomes uncorrelated from X_j as $\hat{\sigma}^j_{t|t} \to 0$ since $\rho_{X_j Y_j} \to 0$. The last line of Eq. (3.6.4) corresponds to the probability of default in the standard Merton model.

3.6.2 Expected Recovery Rate

When default occurs, the recovery rate $RR(t,T)$ is given by the ratio of the asset value to the nominal value of the debt, V_T/K. However, this is true only if $V_T < K$, otherwise no default happens and no recovery can be observed. More formally, the expected recovery rate, $RR(t,T)$, in Merton model is defined as:

$$RR(t,T) = E\left[\frac{V_T}{K}\middle| V_T < K\right] \tag{3.6.5}$$

Altman et al. (2000) give an explicit expression of the recovery rate in terms of the ratio of two standard Gaussian probability functions, i.e.,

$$RR(t,T) = \frac{V_t}{K} e^{r\tau} \frac{N(-d_1(\log(V_t))}{N(-d_2(\log(V_t)))} \tag{3.6.6}$$

In our model, we have to condition on the set of received observations, thus leading to the following definition of recovery rate at default:

$$RR(t,T) = E\left[\frac{V_T}{K}\middle| V_t > K, V_T < K, \mathscr{F}_t^z\right] \tag{3.6.7}$$

The detailed calculation is presented in Appendix E. Here, only the final result is stated:

Proposition 3.6.2. *The recovery rate at time T as seen at time t is given by:*

$$RR(t,T) = e^{r\tau}\sum_j w_t^j \frac{\exp\{\hat{x}_{t|t}^j + 0.5\hat{\sigma}_{t|t}^{2,j}\}}{K}\frac{\mathcal{P}(\xi_j \le dd_t^j, \psi_j \le -d_1(\hat{x}_{t|t}^j)}{\mathcal{P}(X_j \le dd_t^j, Y_j \le -d_2(\hat{x}_{t|t}^j))} \tag{3.6.8}$$

where

$$w_t^j = \mu_t^j \frac{\mathcal{P}(X_j \le dd_t^j, Y_j \le -d_2(\hat{x}_{t|t}^j))}{\sum_{i=1}^m \mu_t^i \mathcal{P}(X_i \le dd_t^i, Y_i \le -d_2(\hat{x}_{t|t}^i))} \tag{3.6.9}$$

and (ξ_j, ψ_j) is a bivariate Gaussian with negative correlation coefficient $\rho_{\xi_j\psi_j}$:

$$\mu_{\xi_j} = -\hat{\sigma}_{t|t}^j,\ \mu_{\psi_j} = \frac{\hat{\sigma}_{t|t}^{2,j}}{\sigma\sqrt{\tau}}, \sigma_{\xi_j}^2 = 1,\ \sigma_{\psi_j}^2 = 1 + \frac{\hat{\sigma}_{t|t}^{2,j}}{\sigma^2\tau},\ \rho_{\xi_j\psi_j} = -\frac{\hat{\sigma}_{t|t}^j}{\sqrt{\sigma^2\tau + \hat{\sigma}_{t|t}^{2,j}}} \tag{3.6.10}$$

while (X_j, Y_j) is the bivariate defined in Eq. (3.6.3)

As already pointed out in Section 3.6.1, when the model consists only of Eq. (3.3.2), $\hat{\sigma}_{t|t} = 0$ and the conditional pdf $p(x_t|\mathscr{F}_t^z)$ becomes the delta function $\delta(x_t - \hat{x}_{t|t})$. Moreover, both (X, Y) and (ξ, ψ) become uncorrelated pairs since their correlation coefficient is zero. Therefore, the cumulative distribution function of X cancels the cumulative distribution function of ξ, and both Y and ψ converge in distribution to a standard Gaussian. Hence, we recover the expected default rate in the standard Merton model given by Eq. (3.6.6).

3.6.3 The Term Structure of Credit Spreads

The bond price $B(t,T)$ can be expressed as

$$B(t,T) = e^{-r\tau}E[K\mathbf{1}_{\{V_T>K\}} + V_T\mathbf{1}_{\{V_T\le K\}}|\mathscr{F}_t^z] \tag{3.6.11}$$

From the definition of credit spread

$$CS(t,T) = -\frac{1}{\tau}\log\left(\frac{B(t,T)}{Ke^{-r\tau}}\right) \tag{3.6.12}$$

it follows immediately using Eq. (3.6.11) that

$$CS(t,T) = -\frac{1}{\tau}\log(1 - PD(t,T)LGD(t,T)) \tag{3.6.13}$$

where $LGD(t,T) := 1 - RR(t,T)$ is the *loss given default*. Using the previously derived results, we obtain:

Proposition 3.6.3. *The credit spread at time T as seen at time t is given by:*

$$\begin{aligned} CS(t,T) &= -\frac{1}{\tau}\log\Bigg[1-\frac{\sum_j \mu_t^j \mathcal{P}(X_j \le dd_t^j, Y_j \le -d_2(\hat{x}_{t|t}^j))}{\sum_{i=1}^m \mu_t^i N(dd_t^i)} \\ &\quad +\frac{\sum_j \mu_t^j \exp\{r\tau + \hat{x}_{t|t}^j + 0.5\hat{\sigma}_{t|t}^{2,j}\}\mathcal{P}(\xi_j \le dd_t^j, \psi_j \le -d_1(\hat{x}_{t|t}^j))}{K\sum_{i=1}^m \mu_t^i N(dd_t^i)}\Bigg] \end{aligned} \tag{3.6.14}$$

where (X_j, Y_j) is the bivariate Gaussian defined in Eq. (3.6.3) and (ξ_j, ψ_j) is the bivariate Gaussian defined in Eq. (3.6.10).

Differently from Merton model, the credit spread for short maturities does not approach zero in our credit risk framework, but is given by

$$\lim_{\tau\to 0} CS(t,\tau) = \frac{\sigma^2}{2}\frac{\sum_j \mu_t^j n_t^j}{\sum_i \mu_t^i N(dd_t^i)} \tag{3.6.15}$$

where

$$n_t^j = \frac{1}{\sqrt{2\pi\hat{\sigma}_{t|t}^{2,j}}}\exp\left\{-\frac{(\log(K)-\hat{x}_{t|t}^j)^2}{2\hat{\sigma}_{t|t}^{2,j}}\right\} \tag{3.6.16}$$

Eq. (3.6.15) can be easily derived by taking the limits of Eq. (3.6.14) as $\tau \to 0$. Eq. (3.6.15) may be rewritten as:

$$\lim_{\tau\to 0} CS(t,\tau) = \frac{\sigma^2}{2}\sum_j w_t^j \frac{n_t^j}{N(dd_t^j)} \tag{3.6.17}$$

where

$$w_t^j = \frac{\mu_t^j N(dd_t^j)}{\sum_i \mu_t^i N(dd_t^i)} \tag{3.6.18}$$

We next want to evidence the similarity and differences between the short-term credit spread of our model and the one resulting from Duffie and Lando (2001). As it can be checked by comparing our formula in Eq. (3.6.17) and their formula in Eq. (2.5.2), the only difference lies in the term multiplying $\frac{\sigma^2}{2}$. In their framework such term is $f'(K)$, which can intuitively interpreted as follows. If the asset value is close to the default threshold, and the density is very steep at that threshold, then the asset value is more likely to slide below the default threshold in the short term. However, they do not put any economic structure on their density. On the contrary, our short-term credit spread is determined by the different reporting models which can be potentially used across time by managers. This is highlighted in Eq. (3.6.17) which has the functional form of a weighted sum of instantaneous hazard rates on the default threshold, with each term $\frac{n_t^j}{N(dd_t^j)}$ being the ratio of the survival density conditioned on the j-th Gaussian density being the actual density of the log-asset value and the corresponding survival probability. If at the current default threshold K the firm did not default, then the probability that default would occur when the default threshold is increased by an amount ΔK is given by $\Delta K\frac{n_t^j}{N(dd_t^j)}$, when $\hat{x}_{t|t}^j$ is believed to be the initial log-asset value mean and $\hat{\sigma}_{t|t}^{2,j}$ the initial log-asset value variance.

In this context, each term $\frac{n_t^j}{N(dd_t^j)}$ can be thought as a quantitative measure of the likelihood that the asset value falls below K in the short term if the estimate $\hat{x}^{j}_{t|t}$ is believed to be the correct log-asset value estimate, and such estimate corresponds to a particular reporting path followed by the manager.

Chapter 4

Empirical Analysis of misreporting: Model Calibration and Statistical Analysis

4.1 Introduction

In this chapter we present an empirical analysis of misreporting. Section 4.2 presents a novel calibration scheme for a simple version of the general model presented in Chapter 3 and applies it to the Parmalat case, a well-known case of deliberate misreporting, around the time of its bankruptcy. Many calibration approaches in literature for structural models have been proposed in the literature, see Leland (2004), Huang and Huang (2003), and Eom et al. (2003) for some relevant references. However, they would not be applicable to our case since they assume complete information on the firm's asset value. Our approach can be thought of as a generalization of Duan approach, which is briefly reported in Section 4.2.1. Section 4.3 uses CDS quotes for a sample of U.S. companies involved in misreporting and bootstraps the default probabilities. The findings show an inversion of the CDS term structure curve at around the time when misreporting was discovered.

4.2 Statistical Calibration

4.2.1 Estimation Procedure for Merton Model

We briefly report the calibration procedure for the Merton model proposed in Duan (1994) and Duan (2000). Since the asset value in the Merton model is assumed to evolve according to Eq. (2.2.1), we have that the conditional density of V_i given $V_{i-1} = v_{i-1}$ is

$$\phi(v_i|v_{i-1}) = \frac{1}{\sqrt{2\pi}\sigma\sqrt{t_i - t_{i-1}}} \exp\left(-\frac{\log v_i - \log v_{i-1} - (\mu - 0.5\sigma^2)(t_i - t_{i-1})^2}{2\sigma^2(t_i - t_{i-1})}\right)\frac{1}{v_i} \tag{4.2.1}$$

The asset value being unobserved, it is implied from the equity price using the fact that in the Merton model equity is a call option on the asset value as evidenced in Eq. (2.2.6). Letting

$$v_i(e_i) = BS_{call}^{-1}(i, e_i, K, T_i, \sigma) \tag{4.2.2}$$

where e_i denotes the equity price at time t_i and K the amount of debt due at maturity T_i. This allows us to write the likelihood function in terms of the equity price as

$$L(e_1, e_2, \ldots, e_n|e_0) = \prod_{i=1}^{n} \phi(v_i(e_i)|v_{i-1}(e_{i-1})) \frac{1}{\frac{\partial BS_{call}(t, v_i(e_i), K, T_i, \sigma)}{\partial v}} \tag{4.2.3}$$

and therefore the maximum likelihood estimator can be applied using available equity data.

4.2.2 Estimation Procedure for the Proposed Model

We consider a simple case of our model in Eq. (3.3.1–3.3.4) where at any time, the manager can either report the log-asset value correctly or bias it by a fixed amount h. We further assume constant drift μ and volatility σ in Eq. (3.3.2) and constant variance ν in the measurement eq. (3.3.4). Since we are assuming the same distortion throughout the period, we do not have the model selection eq. (3.3.3) under this framework. Let us denote by $\Delta_k = t_k - t_{k-1}$ the time between consecutive observations. The filter density is then Gaussian with mean $\hat{x}_{k|k}$ and variance $\sigma^2_{k|k}$ given by the Kalman filter as

$$(\hat{x}_{k|k}, \sigma^2_{k|k}) = \mathbf{KF}(\hat{x}_{k|k-1}, \hat{\sigma}^2_{k|k-1}, z_k, h, \nu) \tag{4.2.4}$$

where

$$\hat{x}_{k|k-1} = \hat{x}_{k-1|k-1} + (\mu - 0.5\sigma^2)\Delta_k \tag{4.2.5}$$

$$\hat{\sigma}^2_{k|k-1} = \hat{\sigma}^2_{k-1|k-1} + \sigma^2\Delta_k \tag{4.2.6}$$

$$G = \frac{\hat{\sigma}^2_{k|k-1}}{\hat{\sigma}^2_{k|k-1} + \nu^2} \tag{4.2.7}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + G \cdot (z_k - \hat{x}_{k|k-1}) \tag{4.2.8}$$

$$\hat{\sigma}^2_{k|k} = (1 - G)\hat{\sigma}^2_{k|k-1} \tag{4.2.9}$$

Moreover, it is easily seen that under these assumptions the price of equity $E(t, T)$ in Eq. (3.4.15) reduces to

$$E(t, T) = e^{\hat{x}_{t|t} + 0.5\hat{\sigma}^2_{t|t}} N(d_1(\hat{x}_{t|t})) - Ke^{-r\tau} N(d_2(\hat{x}_{t|t})) \tag{4.2.10}$$

We next describe a maximum likelihood estimator procedure which recovers the system parameters from a time series of market prices of equity. Differently from Duan (1994, 2000), who implies the exact asset values from equity prices, we imply the reported values from equity prices. Let us first assume that a sample of released observations $z_1, \ldots, z_n$ was observed on the market at discrete dates $\{t_1, t_2, \ldots, t_n\}$. Eq. (3.3.4) implies that the log-likelihood function is given by:

$$LL(z_1, z_2, \ldots, z_n) = -\frac{n}{2}\log(2\pi) - \sum_{i=1}^{n}\log(\hat{\nu}_{i|i-1}) - \sum_{i=1}^{n}\frac{(z_i - \hat{z}_{i|i-1})^2}{2\hat{\nu}_{i|i-1}^2} \tag{4.2.11}$$

where $\hat{z}_{i|i-1}$ and $\hat{\nu}_{i|i-1}^2$ are given by

$$\hat{z}_{i|i-1} = \hat{x}_{i|i-1} + h \qquad \hat{\nu}_{i|i-1}^2 = \hat{\sigma}_{i|i-1}^2 + \nu^2 \tag{4.2.12}$$

However, when we do not directly observe z_i, we need to imply it from equity data and this is done using the approach described next. Eq. (4.2.10) shows that the equity price at time t_i for a maturity T_i can be expressed as

$$\begin{aligned} E(t_i, T_i) &= f(\hat{x}_{i|i}, \hat{\sigma}_{i|i}^2, \mu, \sigma, h, \nu, r) \\ &= g(z_i, \hat{x}_{i|i-1}, \hat{\sigma}_{i|i-1}^2, \mu, \sigma, h, \nu, r) \end{aligned} \tag{4.2.13}$$

where f and g are two deterministic functions. The latter equation follows from the updating step of the Kalman filter given in Eq. (4.2.8) and Eq. (4.2.9).

Expressing the equity at time t_i in terms of z_i is particularly convenient because it would allow implying z_i through inversion of g, in case g is invertible. This turns out to be the case, since the first-order derivative of g with respect to z_i

$$\frac{\partial g(y_i, \hat{x}_{i|i-1}, \hat{\sigma}_{i|i-1}^2, \mu, \sigma, h, \nu)}{\partial z_i} = e^{\hat{x}_{i|i} + 0.5\hat{\sigma}_{i|i}^2} N(d_3^i) G \tag{4.2.14}$$

with d_3^i defined by

$$d_3^i = \frac{\log(\hat{x}_{i|i}/K_i) + (r + 0.5\sigma^2)\tau_i - \hat{\sigma}_{i|i}^2}{\sqrt{\hat{\sigma}_{i|i}^2 + \sigma^2\tau_i}} \tag{4.2.15}$$

is positive. Here τ_i is the difference $T_i - t_i$ and K_i represents the amount of outstanding debt at time t_i. Therefore, we can recursively imply the observations z_i from the market prices of equities $e_1, \ldots, e_i$, where e_j denotes the equity price at time t_j, using the three-step procedure described next:

(1) Predict the mean and variance of the true log-asset value using the time propagation formulas of the Kalman filter given in Eq. (4.2.5) and Eq. (4.2.6).

(2) Imply z_i from g, i.e., numerically compute

$$z_i = g^{-1}(e_i; \hat{x}_{i|i-1}, \hat{\sigma}_{i|i-1}^2, \mu, \sigma, h, \nu, r) \tag{4.2.16}$$

where the notation $g^{-1}(e_i; \hat{x}_{i|i-1}, \hat{\sigma}_{i|i-1}^2, \mu, \sigma, h, \nu, r)$ indicates that in the inversion we are considering g as a function only of the variable z_i and keeping the remaining parameters fixed.

(3) Update the mean and variance of the true log-asset value using the correction formulas of the Kalman filter given in Eq. (4.2.8) and Eq. (4.2.9).

Let $p_i = (\hat{x}_{i|i-1}, \hat{\sigma}^2_{i|i-1}, \mu, \sigma, h, \nu)$. Then we can rewrite the log-likelihood function as

$$\begin{aligned} LL(e_1, e_2, \ldots, e_n) \quad = \quad & -\frac{n}{2}\log(2\pi) - n\log(\hat{\nu}_{i|i-1}) - \sum_{i=1}^{n} \frac{(g^{-1}(e_i; p_i) - \hat{y}_{i|i-1})^2}{2\hat{\nu}^2_{i|i-1}} - \\ & \sum_{i=1}^{n} \log(G \cdot N(d_3^i)) - \sum_{i=1}^{n} (\hat{x}_{i|i} + 0.5\hat{\sigma}^2_{i|i}) \end{aligned} \tag{4.2.17}$$

where $\hat{z}_{i|i-1}$ is computed using Eq. (4.2.12) in step (1) of the above procedure. The value $z_i = g^{-1}(e_i; p_i)$ is implied from step (2), while $\hat{x}_{i|i}$ and $\hat{\sigma}^2_{i|i}$ are computed in step (3) along with d_3^i.

4.2.3 Application to the Parmalat Case

We apply the above methodology to estimate h, μ, σ, ν^2 for the case of Parmalat, an Italian food firm which experienced a crisis during the years 2002–2003 and resulted in the largest bankruptcy in European history. We chose Parmalat since this has been already investigated in the credit risk literature, see (see Brigo and Morini, 2006, Cherubini and Manera, 2006), and therefore it allows us to make a comparison of our results with theirs.

During those two years Parmalat was repeatedly announcing issuance of bonds despite its balance sheet statements reporting huge amount of available cash liquidity that was not used. On December 8, 2003, it was suddenly discovered that its claimed liquidity of four billion Euros did not exist, and that eight million Euros in bonds of investors' money had evaporated as well. An illustration of Parmalat financial distress is reported in Figure 4.1 using CDS, equity and debt data.

In Figure 4.1 the default risk is estimated as the difference between the quasi-debt, i.e., the value of debt discounted with the risk-free rate, and the fair value of debt, computed discounting debt with the risk-free rate plus the five-year CDS spread. The bottom plot of Figure 4.1 shows that the crisis was announced by a decline in the stock price combined with a simultaneous increase of the default risk. Moreover, the top plot shows a steep increase of the credit spread curves at around the time of default.

We collected the daily stock data from January 1, 2002 to December 1, 2003 and computed the daily value of equity multiplying the number of outstanding shares by the stock price. At each time t_i, we assume that the outstanding debt K_i has five-year maturity and proxy it with the long-term debt recovered from the balance sheet statements. We use the five-year treasury yield as the discount factor. We run the calibration procedure using our method and compare the parameter estimates with the ones obtained by running the maximum likelihood estimation procedure by Duan (1994, 2000) on the standard Merton model. The estimates reported in Table 4.1 show that the hidden parameter h plays an important role in the specification of the model. If it were omitted, as it happens for the Merton model, then the reduced value of the firm due to accounting misreporting would simply be explained by an increase of the asset volatility which rises to 17%. Our estimate of asset volatility of 11% matches closely with the estimates of asset volatility found by different calibration procedures, (see Brigo and Morini, 2006, Cherubini and Manera, 2006). Our estimate 0.2052 of the distortion factor recalls the findings of Cherubini and Manera (2006), who perform a historical analysis of risk-neutral probabilities of fraud via a modification of Merton model and

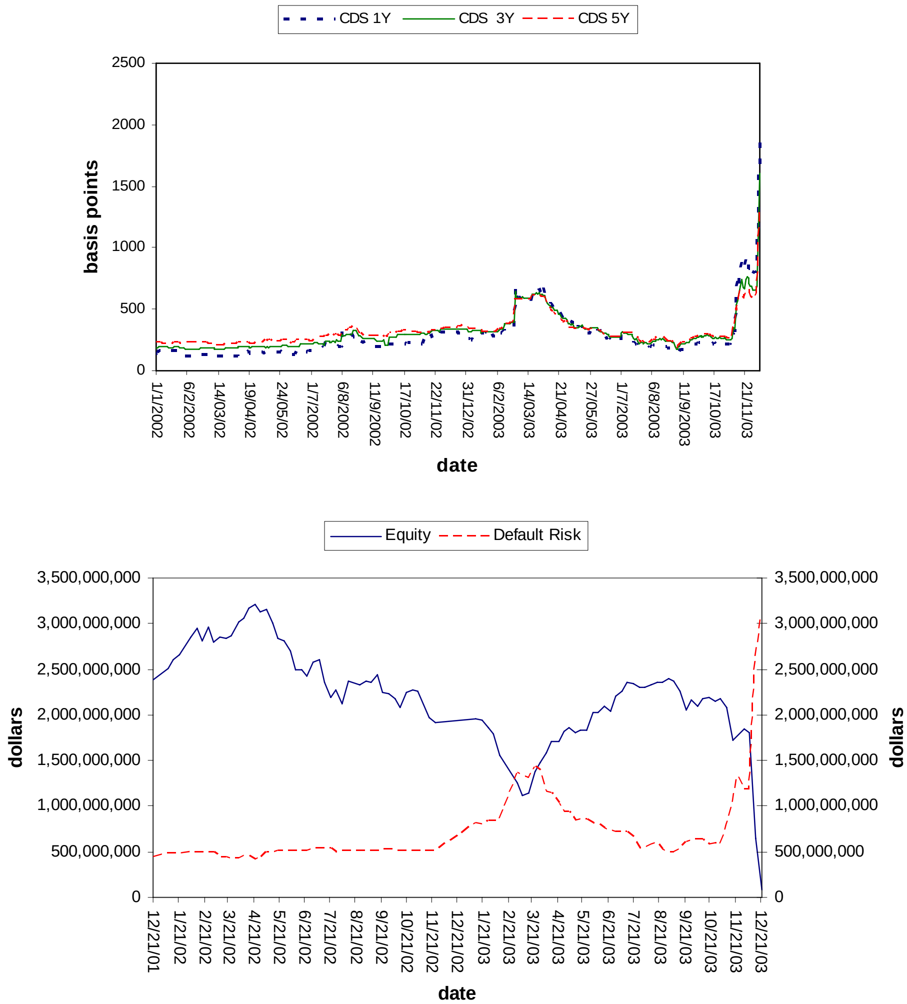

Figure 4.1: Parmalat crisis

	MLE Estimates	
Parameters	Proposed Model	Merton Model
μ	-1.3% (0.0055)	-0.8% (0.054)
σ	11.2% (0.0022)%	17% (0.007)
h	0.2052 (0.0024)	
ν	0.002 (0.0002)%	

Table 4.1: Parameter estimates

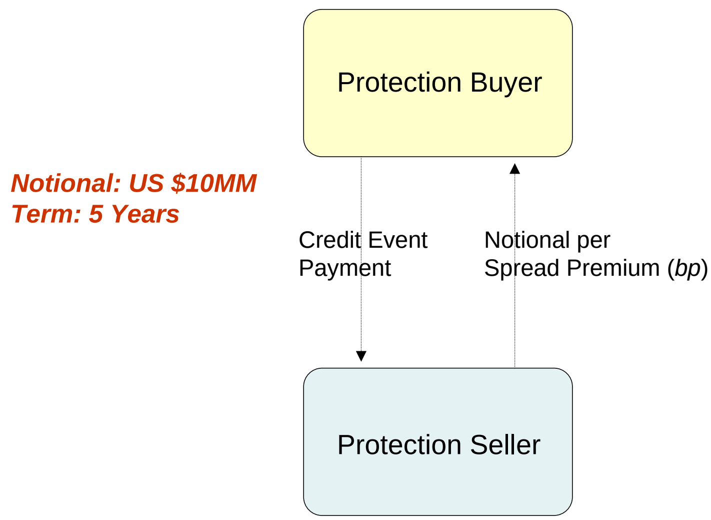

Figure 4.2: Schematic illustration of a CDS contract

extract an implied probability of fraud of about 0.2. Although there are differences in the assumptions and analysis between our work and theirs, making it hard to compare, using the simplest version of our model we find qualitatively similar results to the ones they obtain with the full power of their model.

4.3 Empirical Analysis of Misreporting

We present a statistical analysis of the term structure of the credit default swap curve (CDS) as well as of the default probabilities bootstrapped from CDS data. Section 4.3.1 briefly recalls the definition of a CDS instrument. Section 4.3.2 explains the bootstrapped procedure used in implying the default probabilities. Section 4.3.3 applies the above procedure to three known cases of occurred misreporting and draws inferences about the information content of those financial instruments in misreporting situations.

4.3.1 CDS Contracts

The credit default swap is the most liquid credit derivative instrument. A CDS is an agreement between two parties to exchange the credit risk of an issuer called reference entity. The buyer of the CDS pays a periodic fee to the seller until the maturity of the swap contract or the occurrence of the default event (whichever comes earlier). The seller of the CDS faces the obligation, until maturity of the swap, to pay the losses incurred by the protection buyer as a result of the credit event. Such credit event can be bankruptcy, failing to pay outstanding debt obligations, etc. A schematic illustration of a CDS contract is reported in Figure 4.2.

When a credit event occurs, the protection seller has to make up for losses on general exposures to the name under the contract. There are several settlement agreements between the buyer and seller on how to deal with the occurred credit event. The most popular is *physical delivery*. Under this agreement the protection buyer is allowed to deliver whatever debt obligation issued by the name and typically the natural choice is to deliver the cheapest one.

We next describe the payments of buyer and seller. For simplicity, we assume that payments occur at the end of the each period. The payment from the protection seller at the end of any coupon period (time t_i) can be represented as

$$1_{t_{i-1} \leq \varsigma \leq t_i} LGD \tag{4.3.1}$$

where $1_{t_{i-1}}$ is the indicator function taking value 1 if default occurs between time t_{i-1} and time t_i, while LGD is the loss given default paid by the protection seller to the protection buyer at the occurrence of the default event. At time t_i the protection buyer will pay instead

$$s 1_{\varsigma > t_{i-1}} \tag{4.3.2}$$

i.e, the premium s conditional on the reference entity having survived up to time t_{i-1}. It is assumed that if the name defaults in a period the contract ends in that period.

We next explain how to evaluate the payments of buyer and seller. Denote $Q(t) = P(\varsigma > t)$ the survival probability of the issuer up to time t under the risk-neutral measure. The present value of the payment done at time t_i by the protection seller is obtained by taking the expectation of the risky cash flow in Eq. (4.3.1). Using simple survival analysis, it is easily seen that this is given by

$$e^{-r(t-t_i)}[Q(t_{i-1}) - Q(t_i)]LGD \tag{4.3.3}$$

where $[Q(t_{i-1}) - Q(t_i)]$ is the probability of a default event occurring between t_{i-1} and t_i, in which case the protection seller will have to pay the promised LGD. By a similar argument, the present value of the payment done by the protection buyer at time t_i is

$$e^{-r(t-t_i)} Q(t_{i-1}) s \tag{4.3.4}$$

where $Q(t_{i-1})$ is the probability that no default event has occurred before t_{i-1}, in which case the protection buyer will have to pay the premium s at time t_i. The running premium s is the break-even value, i.e., the value which makes the present value of the protection buyer cash-flow equal to the present value of the protection seller cash-flow. For a n-year CDS contract, which thus consists of n payment periods, the spread s, as seen at time t, will have to satisfy the following equation

$$\sum_{i=1}^{n} e^{-r(t_i-t)} Q(t_{i-1}) s = \sum_{i=1}^{n} e^{-r(t_i-t)} [Q(t_{i-1}) - Q(t_i)] LGD \tag{4.3.5}$$

4.3.2 CDS-Implied Default Probabilities

The CDS instruments allow to investigate the term structure of the default probability for the case of obligors on which CDS contracts are quoted for several maturities. In this case, it is possible to bootstrap the term structure of the default probability in a very similar manner to that by which we usually back out the zero-coupon factor curve from swap rates. More precisely, assume that a CDS spread s_1 is quoted for protection over a one year horizon. For the sake of simplicity, assume that the spread is paid in one instance at the end of the year, and it is paid even if the name has defaulted. Moreover, assume that before time t_1 no default has occurred yet. Using Eq. (4.3.5) with $n = 1$, we obtain

$$s_1 = [1 - Q(t_1)]LGD \tag{4.3.6}$$

from which we can immediately imply

$$Q(t_1) = 1 - \frac{s_1}{LGD} \tag{4.3.7}$$

The value for $Q(t_1)$ can then be plugged into Eq. (4.3.5) applied to a 2 year CDS contract, thus obtaining

$$\sum_{i=1}^{2} e^{-r(t_i - t)} Q(t_{i-1}) s_2 = \sum_{i=1}^{2} e^{-r(t_i - t)} [Q(t_{i-1}) - Q(t_i)]LGD \tag{4.3.8}$$

from which we can back out $Q(t_2)$, since we have $Q(t_1)$ and s_2 which is the 2 year CDS quote observed on the market. We can continue iterating this process until extracting the survival probabilities for all maturities.

4.3.3 Signals of Misreporting Inferred from CDS

We report the results of an empirical investigation which evidences the impact of misreporting and financial distress on the term structure of the CDS curve and the resulting bootstrapped default probabilities. We found that the misreporting event is typically announced by an inversion of the term structure of the CDS spread curve. We next report the results of this analysis on three major cases of misreporting in United States history, namely on Enron, WorldCom, and Tyco, and make some comments about them. We start presenting the results for the Tyco case. On March 13, 2001, Tyco announced a 9.2 billion cash and stock deal to take over the CIT group, a commercial finance company. This CIT deal led to sudden discovery of management misconduct. Tyco shares dropped sharply at the end of January 2002, as a proxy report was filed with the Securities and Exchange Commission disclosing that the Tyco CFO received a 10 million fee on conclusion of the CIT Group deal, and that another $10 million went to a charity where he was a director.

The crisis at the end of January 2002 is evidenced very clearly from Figures 4.3 and 4.4. The shape of the term structure curve was inverted, so that the 3 year CDS contract was quoting on average about 100 basis points (bps) less than the 1 year contract, and the 5 year contract was quoting about 190 bps lower. The inverted shape persisted across all of year 2002 and was only reversed in late 2003. An inverted structure of the CDS spread is indeed typical of periods of particular financial distress. This means that the market prices a very high probability of default in the first year, while the probability is lower for the years further in the future. The rational behind is that if the company survives this initial period, then it will be in better shape and less likely to (conditionally) default over subsequent periods. All this is confirmed from data of the cumulative default probability which are displayed in Figure 4.4. The 1 year default probability was about 1.39% on average in the last quarter of 2001; 5 year default probability was about 5 times larger, i.e.,

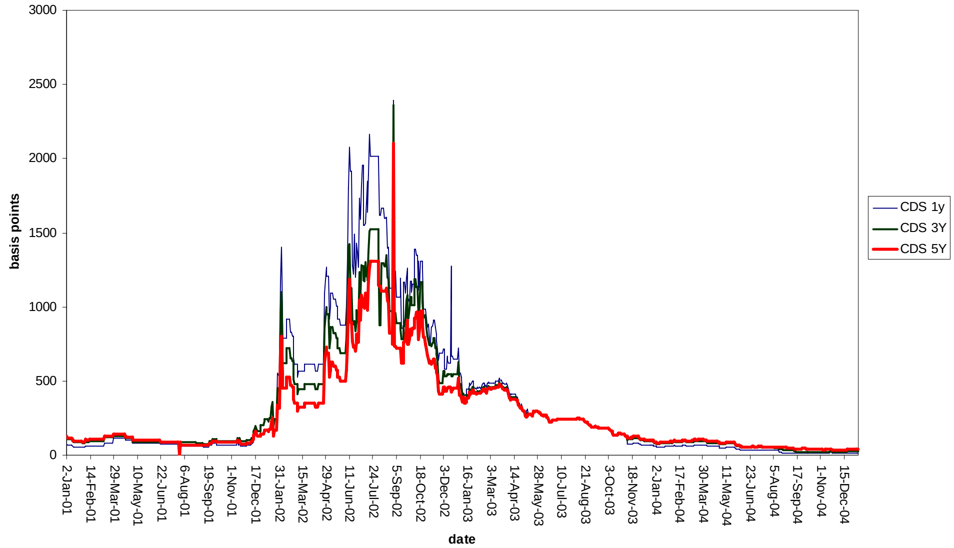

Figure 4.3: 1 year, 3 year, and 5 year CDS of Tyco

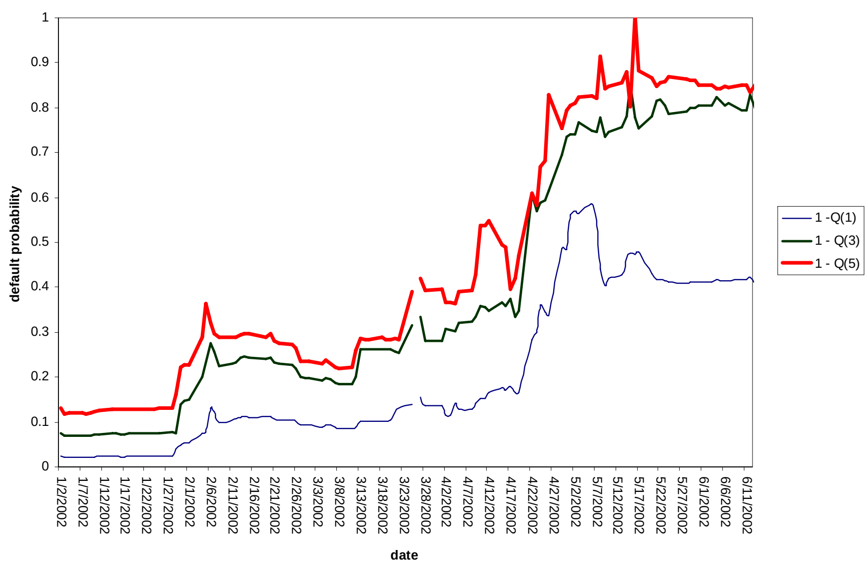

Figure 4.4: 1 year, 3 year, and 5 year CDS implied default probability of Tyco

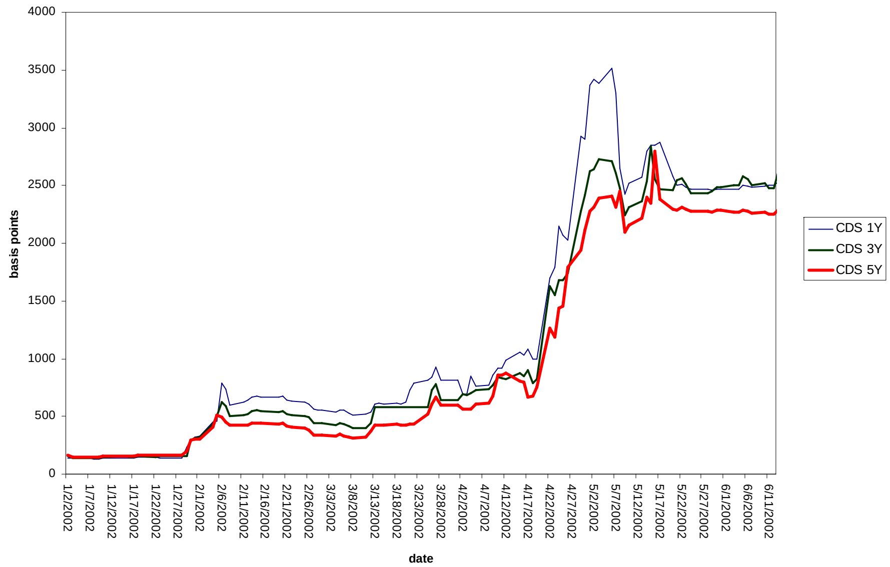

Figure 4.5: 1 year, 3 year, and 5 year CDS of WorldCom

7.51%. This means that the market was attributing good financial health to the firm and was not suspecting fraud. In the first quarter of 2004, the 1 year default probability climbed to 9.56%, while the 5 year default probability grew up to 22.90%. Since the 3 year default probability was 21%, thus very close to the 5 year default probability, the market was implying a probability of default between years 2005 and 2007 of about 1.90%. This is close to the 1 year default probability in the pre-crisis situation, meaning that the market was believing that either the company would have defaulted three years from now, or not defaulted at all.

We next report the same figures for WorldCom, the biggest corporate bankruptcy in the history of the United States. WorldCom was a giant leading the telecommunication industry at the end of last century and was handling 50% of the US internet traffic and 50% of e-mail worldwide. The fraud began when it was discovered that the former chief financial officer at WorldCom had directed employees to falsify balance sheets to hide more than $3.8 billion of expenses. The fraud was discovered in March 2002, when the internal audit found accounting irregularities consisting of $400 million that had been set aside to boost WorldCom's income. After asking for documents supporting capital expenditures and discovering that no supporting accounting standards existed, the Internal audit explained irregularities to the Audit Committee on June 20, 2002, and on July 21, 2002, WorldCom filed for bankruptcy.

Figure 4.5 reports the CDS term structure in the year of crisis. The CDS spreads climbed up about at the same time as the Tyco crisis. At the end of the January, the spreads increased substantially, and the term structure of the spread became inverted. Two months later, the 1 year CDS quoted 608 basis points on average, and the 3 and 5 year maturities respectively quoted 147 and 194 basis points less than that. Again, this corresponds to a massive increase of the default probability in the first year horizon that does not seem to spread out to the following years. In January, the average 1 year default probability was 2.30%, the 3 year default probability was 6.30%, and the 5 year default probability was 12.60%. In the following two months the crisis caused an increase of the 1 year default probability to more than 10%, while the 3

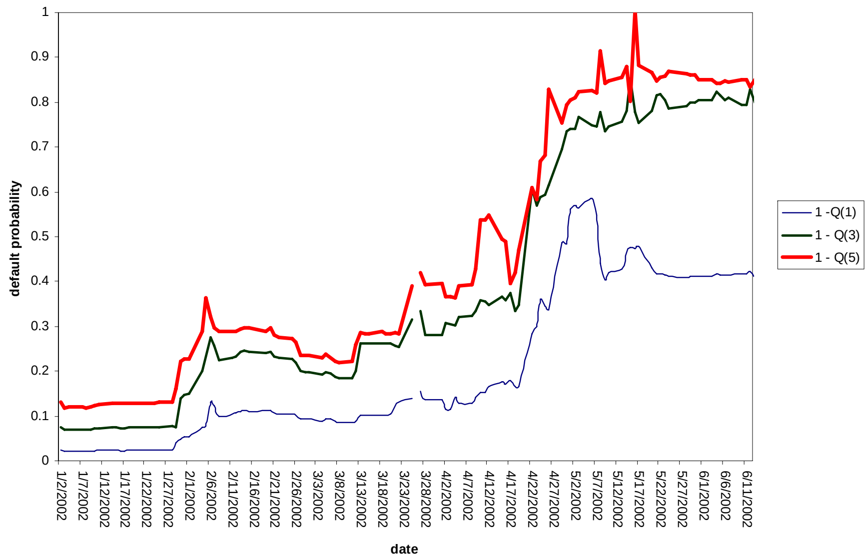

Figure 4.6: 1 year, 3 year, and 5 year CDS implied default probability of WorldCom

and 5 year default probability climbed to 22.66% and 27.95%, respectively. From this, it can be implied that the implied probability of defaulting between third and fifth year was about 5.30%, even less than in the pre-crisis period. Again the market was believing that, if had the company had defaulted, it would have done so in the next three years. This is confirmed from facts since WorldCom was only part of stock market history three years later.

We conclude this empirical investigation with the well-known Enron case, a large player in the natural gas and energy supply. Enron was raising off balance sheet debt by resorting to almost a thousand special purpose entities funded by banks and outside investors. This practice allowed Enron to exploit a loophole in the US Generally Accepted Accounting Standard (US GAAP) to circumvent the principle of consolidation. Under the GAAP, the balance sheet of a special purpose entity needs not be consolidated with the originator if two conditions hold: outside investors must supply at least 3% of the funds, and the originator cannot control the disposition of the assets under control of the entity. This technique enables increasing leverage in the assets of Enron, without reporting this stock of debt in the company accounting data. This leverage effect boosted growth of the company when the fundamentals of its business were good, but by the same token accelerated the crisis when the market turned against Enron, which happened between 2000 and 2001.

Signals of crisis began on October 17, 2001, when Enron announced a downward revision of its third quarter results by more than 1 billion dollars, mainly due to investment losses. This triggered a series of events leading to Moody's downgrading of Enron credit to junk and the subsequent filing of Enron for Chapter 11 which occurred on December 2, 2001.

Even in this case the crisis was announced by the inversion of the CDS spread term structure, which took place in mid-October 2001 when the first signals of crisis started appearing. It was the first time when inversion of the CDS term structure took place, and the one year CDS climbed higher than 500 basis points. The 1 year default probability remained between 1.6% and 2.4%, the 3 year default probability

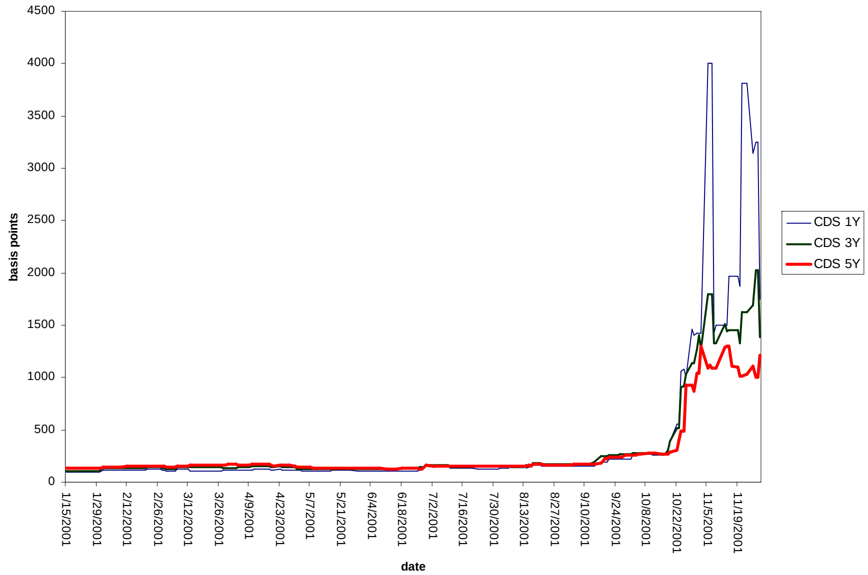

Figure 4.7: 1 year, 3 year, and 5 year CDS of Enron

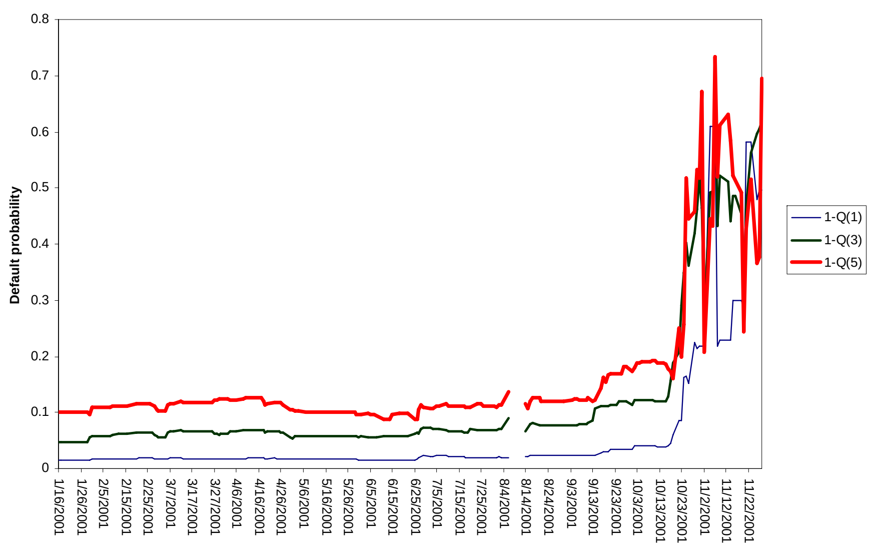

Figure 4.8: 1 year, 3 year, and 5 year CDS implied default probability of Enron

stayed between 6% and 8%, and the 5 year default probability between 10% and 12%. This means that the probability of defaulting between three and five year from that time was roughly 4%. On the day of the crisis, October 22nd, the one year default probability climbed to 8.47%, the 3 year default probability jumped to 20.8% and the 5 year one jumped to 24.94%, thus without any increase on the probability of defaulting between 3 and 5 years from now. Again, the market was believing that if the company had defaulted, it would have done so in the next three years. This was actually the case, since Enron did not manage to survive until the end of the year.

Hence, this empirical investigation showed that cases of misreporting have similarities. They all manifest a steepness of the credit spread curve, with short maturity spreads getting larger than long maturity spread. They all exhibit an increase of the probability of defaulting in the short term, leaving the probability of defaulting in the next years roughly the same as in the pre-crisis situation.

Chapter 5

Stochastic Filtering For Jump Systems

This chapter is devoted to the presentation of an approximate filtering scheme used to compute the filtering density needed for pricing securities of our credit model in Chapter 3.

We recall, from the previous chapter, that the system presented in Eq. (3.3.1–3.3.4) is a jump linear system with state dependent transition probabilities. Even if an analytical expression for the optimal filtering density may be provided (see Section 5.2), it turns out to be computationally prohibitive to evaluate the density on a given input. To this purpose, we develop an approximation method, which allows evaluating the conditional density of a given state in polynomial time. Although the filtering problem associated with our credit risk model given in Eq. (3.3) is one dimensional, we study the filtering problem in the multi-dimensional case, namely allowing both the state and the observation to be n-dimensional vectors. When both the state and the observation are scalar, the developed procedure would return the filtering density used to price bonds in Eq. (3.4.14) and equity in Eq. (3.4.15).

The output of the proposed scheme is an approximation of the true unnormalized posterior density with a mixture of Gaussian densities. Such densities are selected as the solution of a convex programming problem aiming at finding the sparsest possible expansion in terms of Gaussian densities, which is as close as possible to the actual unnormalized density. Our proposed scheme is validated on both theoretical and empirical grounds. From a theoretical point of view, we provide an upper bound on the total variation distance between the true filter density and the approximation density returned from our scheme. Such upper bound may be on-line computed using the system parameters, the received observations, and the constraints imposed on the convex optimization problem, without any recourse to Monte-Carlo simulations. Therefore, the developed bound would allow the user efficiently control the propagation of the approximation error introduced by the method over time. On empirical grounds, we compare the proposed method to the IMM and GPB2 estimators of Bar-Shalom et al. (2002), two widely used heuristic procedures used to deal with Markovian jump systems. We show that the estimation quality is superior at the expenses of an increase in computational burden, which tends to be a bottleneck for high-dimensional state space systems. This happens because the proposed method is based on a gridding of the state space and has complexity growing exponentially with the size of the state space. However, this is not a major concern for us since the state space model associated with our credit risk framework consists of one-dimensional state and observation equations.

The rest of the chapter is organized as follows. Section 5.1 introduces the general state and observation model along with the state transition process. Section 5.2 develops the exact recursive Bayesian filter. Section 5.3 describes the approximation approach and develops a bound on the quality of the density estimator. Section 5.4 presents numerical experiments to empirically validate the proposed procedure.

5.1 Setup and Problem Formulation

The dynamics of the state are modeled by the jump linear switching system

$$\mathbf{x}_k = \mathbf{A}(\theta_{k-1})\mathbf{x}_{k-1} + \mathbf{Q}(\theta_{k-1})\mathbf{w}_k \tag{5.1.1}$$

where $\mathbf{x}_k$ is the n dimensional state of the system at time t_k, and θ_k is the mode switching process. We assume that θ_k is a discrete time finite state process taking values in $\{1,\ldots,d\}$. For any $1 \leq r \leq d$, $\mathbf{A}(r)$ is a square n dimensional matrix governing the dynamics associated to mode r. The noise process $\mathbf{w}_k$ is a n' dimensional Gaussian vector consisting of n' independent Gaussian random variables with zero mean and unit covariance matrix, while $\mathbf{Q}(r)$ is a mode dependent $n \times n'$ positive semi-definite matrix. We assume that the pair $(\mathbf{x}_k, \theta_k)$ is a homogeneous Markov chain with transition matrix P^{Δ} defined as

$$\begin{aligned} P^{\Delta}(\mathbf{x}, j, \mathbf{y}, i) &:= P(\mathbf{x}_k = \mathbf{x}, \theta_k = j | \mathbf{x}_{k-1} = \mathbf{y}, \theta_{k-1} = i) \\ &= P(\theta_k = j | \theta_{k-1} = i, \mathbf{x}_k = \mathbf{x}) \cdot \\ &\quad \cdot P(\mathbf{x}_k = \mathbf{x} | \mathbf{x}_{k-1} = \mathbf{y}, \theta_{k-1} = i) \end{aligned} \tag{5.1.2}$$

and prior distribution on the chain

$$\begin{aligned} \pi(\mathbf{u}, j) &:= P(\mathbf{x}_0 = \mathbf{u}, \theta_0 = j) \\ &= P(\mathbf{x}_0 = \mathbf{u}) P(\theta_0 = j | \mathbf{x}_0 = \mathbf{u}) \end{aligned} \tag{5.1.3}$$

where the initial density $P(\mathbf{x}_0 = \mathbf{u})$ is assumed to be Gaussian with mean μ_0 and covariance matrix $\mathbf{\Sigma}_0$. It is clear from Eq. (5.1.2) and Eq. (5.1.3) that the mode switching process θ_k is statistically dependent on the state $\mathbf{x}_k$. In the special case that it is not, then the mode switching process θ_k becomes a Markov chain and we recover the mode dependent state transition probabilities as in classical Markov jump linear systems. Such systems have been extensively studied and techniques for computing the filtering density have been provided, such as the IMM estimator introduced by Blom and Bar-Shalom (1988), see also Li and Jilkov, 2005, for an excellent survey. Other approaches include particle filtering methods, see (see Doucet et al., 2001).

We next discuss the structure of the observation model. The measurement $\mathbf{z}_k$ produced at time t_k is described by the following equation

$$\mathbf{z}_k = \mathbf{H}(\theta_{k-1})\mathbf{x}_k + \mathbf{R}(\theta_{k-1})\mathbf{v_k} \tag{5.1.4}$$

The observation matrix $\mathbf{H}$ is a mode dependent $m \times n$ matrix. The measurement noise $\mathbf{v}_k$ is a m' dimensional vector consisting of m' independent Gaussian random variables with zero mean and unit covariance matrix, while $\mathbf{R}$ is a mode dependent $m \times m'$ matrix. For notation simplicity, both the design matrix $\mathbf{H}$ and the transition matrix $\mathbf{A}$ are assumed to be time invariant.

Our goal is to develop an efficient algorithm to recover the unnormalized posterior density

$$p^l_{k,k}(\mathbf{x}) := P(\mathbf{x}_k = \mathbf{x}, \theta_{k-1} = l, \mathscr{F}^z_k), l = 1, \ldots, d \tag{5.1.5}$$

where $\mathscr{F}^z_k = \{\mathbf{z}_1, \ldots, \mathbf{z}_k\}$ denotes the set of observations received up to and including sampling time t_k. Eq. (5.1.5) then allows obtaining the joint state-mode posterior hybrid density, and the posterior density of the state through normalization.

5.2 Exact Filter Derivation

We use a recursive expression for the unnormalized posterior density, instead of working directly with its normalized counterpart, following the approach presented in Blom and Bloem (2007), Theorem 1.

Such density is obtained through an interaction of d Bayesian filters, with each filter being an unnormalized posterior density

$$p^l_{k,k}(\mathbf{x}) := P(\mathbf{x}_k = \mathbf{x}, \theta_{k-1} = l, \mathscr{F}^z_k),\ \ l = 1, \ldots, d \tag{5.2.1}$$

The unnormalized prediction density is defined as

$$p^l_{k,k-1}(\mathbf{x}) := P(\mathbf{x}_k = \mathbf{x}, \theta_{k-1} = l, \mathscr{F}^z_{k-1}) \tag{5.2.2}$$

and can be developed as

$$\begin{aligned} &\int_{\mathbb{R}^n} \sum_{r=1}^{d} P(\mathbf{x}_k = \mathbf{x}, \mathbf{x}_{k-1} = \mathbf{y}, \theta_{k-1} = l, \theta_{k-2} = r, \mathscr{F}^z_{k-1}) d\mathbf{y} \\ &\quad = \quad \sum_{r=1}^{d} \int_{\mathbb{R}^n} \lambda^{r,l}(\mathbf{y}) p(\mathbf{x}|\mathbf{y}, l) p^r_{k-1,k-1}(\mathbf{y}) d\mathbf{y} \end{aligned} \tag{5.2.3}$$

where $p(\mathbf{x}|\mathbf{y}, l)$ is a Gaussian density with mean $\mathbf{A}(l)\mathbf{y}$ and covariance $\mathbf{Q}(l)\mathbf{Q}(l)'$. The above decomposition steps follow from straightforward application of Bayes rule. The unnormalized joint state-mode posterior density may then be obtained as

$$\begin{aligned} p^l_{k,k}(\mathbf{x}) &= L^l_{k|k}(\mathbf{x}) p^l_{k,k-1}(\mathbf{x}) \qquad k \geq 2 \\ p^l_{1,1}(\mathbf{x}) &= L^l_{1|1}(\mathbf{x}) \int_{\mathbb{R}^n} p(\mathbf{x}|\mathbf{u}, l)\pi(\mathbf{u}, l) p(\mathbf{u}) d\mathbf{u} \end{aligned} \tag{5.2.4}$$

where the correction term $L^l_{k|k}(\mathbf{x})$ is a Gaussian density with mean $\mathbf{H}(l)\mathbf{x}$ and covariance $\mathbf{R}(l)\mathbf{R}(l)'$ due to the structure of the observation equation (5.1.4).

From the above derivations, we can see that the filters interact with each other according to fundamental Bayesian rules, leading to the recursive expression given in Eq. (5.2.4) for the unnormalized density. It is important to notice that such recursion is linear. The linearity of the recursion allows provision of an upper bound for the error committed by our method at any sampling time t_k.

The posterior density may then be obtained from the interacting bayesian filters through normalization as

$$p_{k|k}(\mathbf{x}) = \frac{\sum_{l=1}^d p^l_{k,k}(\mathbf{x})}{\int_{\mathbb{R}^n} \sum_{l=1}^d p^l_{k,k}(\mathbf{x}) d\mathbf{x}} \tag{5.2.5}$$

5.3 The Filter Approximation Scheme

The filter described in Section 5.2 cannot be implemented efficiently. First of all, the recursive expression (5.2.4) shows that an exponentially increasing number of filters have to interact to obtain the unnormalized density at time t_k. Additionally, Eq. (5.2.3) involves the evaluation of non-Gaussian integrals due to the appearance of the terms $\lambda^{j,l}(\mathbf{y})$ in $p^l_{k,k-1}(\mathbf{x})$, and such integrals are computationally expensive to evaluate. Furthermore, the normalization step required to obtain the posterior density in Eq. (5.2.5) involves an evaluation of d spatial integrations, thus increasing the computational burden even further. In order to deal with these computational issues, we propose an approach which at every step k approximates the unnormalized density using a restricted set of Gaussian densities, selected from a prescribed base set. The number of densities and the weight of each density are recovered as the solution of a convex *second-order cone* programming problem.

5.3.1 The Approximation Method

Let $p(\mathbf{x})$ be the density which we wish to approximate. Let I and J be two set of indices. We construct a *base* set of multivariate Gaussian densities

$$B = \{n_{i,j}(\mathbf{x})\}_{i \in I, j \in J} \tag{5.3.1}$$

where $n_{i,j}(\mathbf{x})$ stands for the multivariate Gaussian density with mean $\boldsymbol{\mu}_i$ and covariance $\boldsymbol{\Sigma}_j$. Moreover, we require

$$\begin{aligned} \boldsymbol{\mu}_{i_1} &\neq \boldsymbol{\mu}_{i_2}, \qquad \forall i_1 \neq i_2 \\ \boldsymbol{\Sigma}_{j_1} &\neq \boldsymbol{\Sigma}_{j_2} \qquad \forall j_1 \neq j_2 \end{aligned} \tag{5.3.2}$$

meaning that the means and the covariances of the Gaussian densities in B are all different. We choose a *training* set

$$\mathfrak{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_q) \tag{5.3.3}$$

of size q containing vectors in $\mathbb{R}^n$. Let us define the matrix

$$\boldsymbol{\Phi}(\mathfrak{X}) = \begin{pmatrix} \mathbf{n}_1(\mathbf{x}_1) & \mathbf{n}_2(\mathbf{x}_1) & \dots & \mathbf{n}_{|I|}(\mathbf{x}_1) \\ \mathbf{n}_1(\mathbf{x}_2) & \mathbf{n}_2(\mathbf{x}_2) & \dots & \mathbf{n}_{|I|}(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{n}_1(\mathbf{x}_q) & \mathbf{n}_2(\mathbf{x}_q) & \dots & \mathbf{n}_{|I|}(\mathbf{x}_q) \end{pmatrix} \tag{5.3.4}$$

where $\mathbf{n}_i(\mathbf{x}_l) = (n_{i,1}(\mathbf{x}_l), n_{i,2}(\mathbf{x}_l), \dots, n_{i,|J|}(\mathbf{x}_l))$, i.e., a row vector whose j-th entry is the multivariate Gaussian density in B with mean μ_i and covariance $\boldsymbol{\Sigma}_j$ evaluated at $\mathbf{x}_l$.

Moreover, we assume that $q < |I| \times |J|$, i.e., the size of the training set, is strictly smaller than the cardinality of the base set B. Let $\mathbf{p} = (p(\mathbf{x}_1), \dots, p(\mathbf{x}_q))'$. The linear system

$$\mathbf{p} = \boldsymbol{\Phi}\mathbf{z} \tag{5.3.5}$$

is solvable and overdetermined being $q < |I| \times |J|$. Although we could solve the system and then approximate the density $p(\mathbf{x})$ with $\phi\mathbf{z}(\mathbf{x})$, we notice that such approach would require propagating a number of Gaussian densities equal to the size q of the training set, and therefore it would scale linearly with the size of the training set, making a real time implementation computationally intensive. Our goal is to approximate $\mathbf{p}$ using a short linear combination of Gaussian densities and at the same time minimize the introduced approximation error. Therefore, we look for the sparsest representation of $\mathbf{p}(\mathbf{x})$ in the following sense:

$$\min ||\boldsymbol{v}||_0 \quad \text{subject to } ||\mathbf{p} - \boldsymbol{\Phi}\boldsymbol{v}||_2 \leq \iota \tag{P1}$$

(P1) is a mathematical programming problem with decision variables $\boldsymbol{v}$, and $||\boldsymbol{v}||_0$ denotes the number of nonzero entries of the vector $\boldsymbol{v}$, i.e.,

$$||\boldsymbol{v}||_0 = |\{(i,j) : v_{(i,j)} \neq 0\}| \tag{5.3.6}$$

where $v_{(i,j)}$ is the entry of the $|I| \times |J|$ dimensional vector $\boldsymbol{v}$ multiplying the Gaussian density $n_{i,j}$. If $\boldsymbol{v}^*$ is the solution of (P1), we would approximate $\mathbf{p}(\mathbf{x})$ with $\phi\boldsymbol{v}^*(\mathbf{x})$. However, this is of little practical use, since (P1) is non-convex and easy to solve only in the special case when the system $\mathbf{p} = \boldsymbol{\Phi}\boldsymbol{v}$ admits a unique solution (see Donoho and Elad, 2003). In all other cases, its solution usually requires an intractable combinatorial search. In order to overcome this computational burden, basis pursuit (see Chen et al., 2001) has become the most widely used numerical shortcut to approximate the solution to (P1). Such method replaces the l_0 norm with its closest convex penalty function, the l_1 norm, and solves

$$\min ||\boldsymbol{\varpi}||_1 \quad \text{subject to } ||\mathbf{p} - \boldsymbol{\Phi}\boldsymbol{\varpi}||_2 \leq \epsilon \tag{P2}$$

with decision variable $\boldsymbol{\varpi}$. The problems (P1) and (P2) differ only in the choice of the objective function, with the latter using an l_1 norm as a proxy for the sparsity count. However, unlike (P1), (P2) is a convex second-order cone programming problem. This can be seen by rewriting the problem (P2) in the following form

$$\min \mathbf{1^T}\tilde{\mathbf{u}} \quad \text{subject to} - \tilde{\mathbf{u}} \leq \boldsymbol{\varpi} \leq \tilde{\mathbf{u}} \quad ||\mathbf{p} - \boldsymbol{\Phi}\boldsymbol{\varpi}||_\mathbf{2} \leq \epsilon \tag{P3}$$

where we added the slack optimization variable $\tilde{\mathbf{u}}$. In the above formulation $\mathbf{1}$ is a vector of ones and the vector inequality $\boldsymbol{\varpi} \leq \tilde{\mathbf{u}}$ means component-wise, i.e., $w_i \leq \tilde{u}_i$ for all i. Being a convex programming problem, (P2) can be solved efficiently using standard optimization algorithms. Therefore, we will approximate the actual density $\mathbf{p}$ as

$$\mathbf{p}(\mathbf{x}) \approx \boldsymbol{\Phi}\boldsymbol{\varpi}(\mathbf{x}) \tag{5.3.7}$$

5.3.2 Sparsity of the approximation

In this section, we discuss the relation between the approximation errors ι and ϵ resulting from the solution of problems (P1) and (P2) and the sparsity of the corresponding solutions. Let $\Lambda \subseteq I \times J$ denote a generic subset. We define the inner product between two vectors as their dot product, i.e.,

$$< \mathbf{x}, \mathbf{y} >= \sum_i x_i y_i \tag{5.3.8}$$

and denote by

$$n_{i,j}(\mathfrak{X}) = (n_{i,j}(\mathbf{x}_1), n_{i,j}(\mathbf{x}_2), \ldots, n_{i,j}(\mathbf{x}_q)) \tag{5.3.9}$$

the vector obtained evaluating the Gaussian density $n_{i,j}$ at all points of the training set $\mathfrak{X}$. We define the *cumulative coherence* as

$$c(s) = \max_{|\Lambda|=s} \max_{(i,j)\notin\Lambda} \sum_{(u,v)\in\Lambda} |< n_{i,j}(\mathfrak{X}), n_{u,v}(\mathfrak{X}) >| \tag{5.3.10}$$

which measures how much a collection of fixed densities resembles a fixed, distinct density, with the convention that $c(0) = 0$. The maximum correlation of the density vector $\mathbf{p}$ with the set of base densities B is defined as

$$\text{maxcor}(\mathbf{p}) = \max_{(i,j)\in I\times J} \frac{|< \mathbf{p}, n_{i,j}(\mathfrak{X}) >|}{||\mathbf{p}||_2} \tag{5.3.11}$$

We now report the following result from citetTropprelax, as:

Lemma 5.1. *Let $\mathbf{a_{opt}} = \boldsymbol{\Phi}\boldsymbol{\upsilon}^*$ where $\boldsymbol{\upsilon}^*$ is an optimal solution to the mathematical programming problem (P1) with tolerance ι. If $||\boldsymbol{\upsilon}^*||_0 = s$, and $c(s) \leq \frac{1}{3}$, and we select*

$$\epsilon \geq \iota\sqrt{1 + 6s \ \text{maxcor}(\mathbf{p} - \mathbf{a_{opt}})^\mathbf{2}} \tag{5.3.12}$$

then the solution $\boldsymbol{\varpi}^*$ *to (P2) with tolerance* ϵ *is at least as sparse as the solution* $\boldsymbol{\upsilon}^*$*, i.e.,*

$$||\boldsymbol{\varpi}^*||_0 \leq ||\boldsymbol{\upsilon}^*||_0 \tag{5.3.13}$$

In words, the above lemma states that if our base set B has a sufficiently high degree of independence, namely, $c(s) \leq \frac{1}{3}$, then we do not lose sparsity by using the l_1 norm instead of the l_0 norm in the programming problem, although this happens at the expenses of an increase in the approximation error ϵ which will depend upon the sparsity of the solution of (P1) and the correlation of the residual of $(\mathbf{p} - \mathbf{a_{opt}})$ with the base set B.

5.3.3 The Filter Approximation

We describe how to construct an approximation to the actual unnormalized density $p^l_{k,k}(\mathbf{x})$ at time t_k from an existing set of approximations to the unnormalized density $\{p^r_{k-1,k-1}(\mathbf{y})\}_{r=1,\ldots,d}$ available from time t_{k-1}. For any $r = 1, \ldots, d$, let us denote by $\hat{p}^r_{k-1,k-1}(\mathbf{y})$ such an approximation which would have been computed in the previous step of the approximation procedure. Let

$$\Omega^l_k = \{(i,j) \in I \times J : [\boldsymbol{\varpi}^l_k]_{(i,j)} \neq 0\} \tag{5.3.14}$$

where $\boldsymbol{\varpi}^l_k$ is the solution of (P2) obtained when approximating $p^l_{k,k}(\mathbf{x})$, i.e.,

$$\hat{p}^r_{k-1,k-1}(\mathbf{y}) \approx \sum_{(i,j)\in\Omega^r_{k-1}} [\boldsymbol{\varpi}^r_{k-1}]_{(i,j)} n_{i,j}(\mathbf{y}) \tag{5.3.15}$$

In Eq. (5.3.15), we want to highlight that we are only propagating from time t_{k-1} to time t_k the Gaussian densities in B with nonzero weight. Our goal is to approximate the actual filter density $p^l_{k,k}(\mathbf{x})$ at time t_k given by Eq. (5.2.4). This is done in two separate steps. The output of the first step is an approximate density $\check{p}^l_{k,k}(\mathbf{x})$ computed as follows:

$$\begin{aligned} \check{p}^l_{k,k}(\mathbf{x}) &:= L^l_{k|k}(\mathbf{x})\sum_{r=1}^{d}\int_{\mathbb{R}^n}\lambda^{r,l}(\mathbf{y})p(\mathbf{x}|\mathbf{y},l)\hat{p}^r_{k-1,k-1}(y)d\mathbf{y} \\ &\approx L^l_{k|k}(\mathbf{x})\sum_{r=1}^{d}\int_{\mathbb{R}^n}\lambda^{r,l}(\mathbf{y})p(\mathbf{x}|\mathbf{y},l)p^r_{k-1,k-1}(\mathbf{y})d\mathbf{y} \\ &= p^l_{k,k}(\mathbf{x}) \end{aligned} \tag{5.3.16}$$

where the approximation step consists in replacing each term $p^r_{k-1,k-1}(\mathbf{y})$ with its previously computed approximation $\hat{p}^r_{k-1,k-1}(\mathbf{y})$. Plugging Eq. (5.3.15) into Eq. (5.3.17), we obtain

$$\begin{aligned} \check{p}^l_{k,k}(\mathbf{x}) = L^l_{k|k}(\mathbf{x})\sum_{r=1}^{d}\sum_{(i,j)\in\Omega^r_{k-1}}[\boldsymbol{\varpi}^r_{k-1}]_{(i,j)} \\ \cdot\int_{\mathbb{R}^n}\lambda^{r,l}(\mathbf{y})p(\mathbf{x}|\mathbf{y},l)n_{i,j}(\mathbf{y})d\mathbf{y} \end{aligned} \tag{5.3.17}$$

Eq. (5.3.17) shows that we have obtained an approximation density $\check{p}^l_{k,k}(\mathbf{x})$ for $p^l_{k,k}(\mathbf{x})$ consisting of $\cdot \sum_{r=1}^d |\Omega^r_{k-1}|$ components. Moreover, $\check{p}^l_{k,k}(\mathbf{x})$ is no longer a mixture of Gaussians due to the appearance of the state-dependent mode transition term $\lambda^{r,l}(\mathbf{y})$. To this purpose, we approximate $\check{p}^l_{k,k}(\mathbf{x})$ further before propagating it to the next step sampling time t_{k+1}. This is the second step of our methodology, which is described next. We first compute the mean and standard deviation of the unnormalized density $\check{p}^l_{k,k}(\mathbf{x})$. These can be efficiently computed by rewriting the integral as a sum of expectations, each one taken with respect to two independent random variables, one being a Gaussian and another being a Gaussian mixture. For computing the mean, we proceed as follows:

$$
\begin{aligned}
\check{\mathbf{x}}^l_{k,k} &= \int_{\mathbb{R}^n} \check{p}^l_{k,k}(\mathbf{x})\mathbf{x}d\mathbf{x} \\
&= \sum_{r=1}^{d} ||\boldsymbol{\varpi}^r_{k-1}]||_1 \mathbb{E}_{\mathbf{X},\mathbf{Y}_r}[\mathbf{x}L^l_{k|k}(\mathbf{x})\ \lambda^{r,l}(\mathbf{y})] \qquad (5.3.18)
\end{aligned}
$$

where $\mathbf{X}$ and $\mathbf{Y}_r$ are dependent vectors from which samples are drawn as follows. We first sample $\mathbf{y}$ from $\mathbf{Y}_r$ according to the Gaussian mixture distribution with density given by

$$
\mathbf{Y}_r \sim \sum_{(i,j)\in\Omega^r_{k-1}} \frac{[\boldsymbol{\varpi}^r_{k-1}]_{(i,j)}}{||\boldsymbol{\varpi}^r_{k-1}||_1} n_{i,j}(\mathbf{y}) \qquad (5.3.19)
$$

We next sample $\mathbf{x}$ from a Gaussian distribution with density given by

$$
\mathbf{X} \sim n(\mathbf{x}; \mathbf{A}(\mathbf{l})\mathbf{y}, \mathbf{Q}(\mathbf{l})) \qquad (5.3.20)
$$

Monte-Carlo methods can be applied to efficiently approximate the above expectation. A similar argument can be followed to approximate the covariance matrix of $\check{p}^l_{k,k}(\mathbf{x})$, let us call it $\check{\boldsymbol{\Sigma}}^l_{k,k}$. We then choose an n-dimensional ellipsoid $G \in \mathbb{R}^n$ defined as

$$
(\mathbf{x} - \check{\mathbf{x}}^l_{k,k})'(\check{\boldsymbol{\Sigma}}^l_{k,k})^{-1}(\mathbf{x} - \check{\mathbf{x}}^l_{k,k}) \qquad (5.3.21)
$$

Both the training points in $\boldsymbol{\mathfrak{X}}$ and the means of the Gaussian densities in B are selected to be on an equally spaced grid covering the ellipsoid G. The covariances of the densities in B are defined by specifying the standard deviation of each state space component σ_i, $i \in 1, \ldots n$, and the correlation factors between them, $\rho_{i,j}$, $i, j \in 1, \ldots, n$. For each covariance matrix, $\sigma_i = \sigma$, i.e., the standard deviation is assumed to be the same for each component. The σ values decay exponentially from $\sigma_{\max}$ to $\sigma_{\min}$ to capture the behavior of the Gaussian density which exhibits exponential tails. The correlation factors $\rho_{i,j}$ are equally spaced in the interval $(-1, 1)$. We next evaluate the density $\check{p}^l_{k,k}(\mathbf{x})$ on the points in $\boldsymbol{\mathfrak{X}}$, thus obtaining a vector

$$
\check{\boldsymbol{p}}^l_{k,k} = (\check{p}^l_{k,k}(\mathbf{x}_1), \ldots, \check{p}^l_{k,k}(\mathbf{x}_q)) \qquad (5.3.22)
$$

Let ϵ^l_k be the smallest component of the vector $\check{\boldsymbol{p}}^l_{k,k}$. Then we solve the optimization problem (P2) as follows:

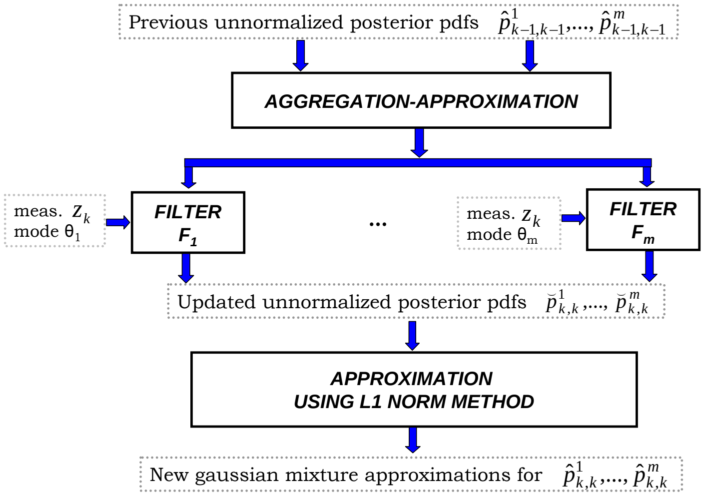

Figure 5.1: One cycle of the estimator

$$\min ||\boldsymbol{\varpi}^l_k||_1 \quad \text{subject to } ||\breve{\boldsymbol{p}}^l_{k,k} - \boldsymbol{\Phi}\boldsymbol{\varpi}^l_k||_2 \leq \epsilon^l_k \tag{P4}$$

If $\boldsymbol{\varpi}^{l,*}_k$ denotes the optimal solution to (P4), the approximation density propagated to step $k+1$ is

$$\hat{p}^l_{k,k}(\mathbf{x}) = \sum_{(i,j)\in\Omega^l_k} [\boldsymbol{\varpi}^{l,*}_k]_{(i,j)} n_{i,j}(\mathbf{x}) \tag{5.3.23}$$

We summarize the whole procedure with a block diagram of the estimator in Figure 5.1. It is clear from Figure 5.1 that the mode conditional densities are approximated right after the mixing step.

5.3.4 Distance between Approximate and Optimal Filter

We provide an upper bound on the total variation distance between the actual density $p^l_{k,k}(\mathbf{x})$ and the approximate density $\hat{p}^l_{k,k}(\mathbf{x})$ obtained using our method. Let us denote

$$\tilde{p}^l_{k,k}(\mathbf{x}) := p^l_{k,k}(\mathbf{x}) - \hat{p}^l_{k,k}(\mathbf{x}) \tag{5.3.24}$$

The total variation distance between $p^l_{k,k}(\mathbf{x})$ and $\hat{p}^l_{k,k}(\mathbf{x})$ is defined as

$$||\tilde{p}^l_{k,k}||_1 := \int_{\mathbb{R}^n} |\tilde{p}^l_{k,k}(\mathbf{x})| d\mathbf{x} \tag{5.3.25}$$

Let ρ be the distance between two vectors in the training set $\mathfrak{X}$. If those vectors are chosen using the methodology described in Section 5.3.3, it is reasonable to assume that

$$p^l_{k,k}(\mathbf{y}) \approx 0, \mathbf{y} \notin G \tag{5.3.26}$$

In words, $p^l_{k,k}(\mathbf{x})$ concentrates most of its probability mass inside $\mathfrak{X}$.

Since our approximation scheme finds the sparsest possible approximation of the density with a guaranteed error bound ϵ, the mass of the approximate density $\hat{p}_{k,k}^{l}(\mathbf{x})$ recovered by (P2) would also be negligible on the region $\mathbb{R}^n - G$. This is because the selection of any Gaussian component whose mean lies outside the grid would increase the l_1 norm of the optimal weight vector without reducing the approximation error any further since the training set does not contain any vector outside the grid. Therefore, the recovered solution would not be optimal. This justifies the following approximation for the l_1 norm

$$||\tilde{p}_{k,k}^{l}||_1 \approx \int_G |\tilde{p}_{k,k}^{l}(\mathbf{x})| d\mathbf{x} \tag{5.3.27}$$

We next state the final result under the form of two theorems, one for the density approximation at initial time, and one for the approximation at a generic sampling time. The detailed proofs are reported in Appendix G and H, respectively.

Define

$$C^{(l)} = \sup_{\mathbf{x}\in\mathbb{R}^n} \frac{||\nabla L_{k|k}^{l}(\mathbf{x})||_\infty}{(2\pi)^{\frac{n}{2}} \det(\mathbf{Q}(l))^{\frac{1}{2}}} + \sup_{\mathbf{x}\in\mathbb{R}^n} \frac{||\nabla p(\mathbf{x}|\mathbf{y},l)||_\infty}{(2\pi)^{\frac{m}{2}} \det(\mathbf{R}(l))^{\frac{1}{2}}} \tag{5.3.28}$$

where

$$\begin{aligned} \nabla L_{k|k}^{l}(\mathbf{x}) &= -\frac{e^{-\frac{1}{2}(\mathbf{z}_k-\mathbf{H}(l)\mathbf{x})'\mathbf{R}(l)^{-1}(\mathbf{z}_k-\mathbf{H}(l)\mathbf{x})}}{(2\pi)^{\frac{m}{2}}\det(\mathbf{R}(l))^{\frac{1}{2}}} \\ &\quad \cdot \mathbf{R}(l)^{-1}(\mathbf{z}_k - \mathbf{H}(l)\mathbf{x}) \\ \nabla p(\mathbf{x}|\mathbf{y},l) &= -\frac{e^{-\frac{1}{2}(\mathbf{x}-\mathbf{A}(l)\mathbf{y})'\mathbf{Q}(l)^{-1}(\mathbf{x}-\mathbf{A}(l)\mathbf{y})}}{(2\pi)^{\frac{n}{2}}\det(\mathbf{Q}(l))^{\frac{1}{2}}} \\ &\quad \cdot \mathbf{Q}(l)^{-1}(\mathbf{x} - \mathbf{A}(l)\mathbf{y}) \end{aligned} \tag{5.3.29}$$

Notice that $L_{k|k}^{l}(\mathbf{x})$ is independent of the time step k due to the time invariant assumption on the design matrix $\mathbf{H}$ made in Section 5.1. Let

$$C_k = \sup_{(j,h)\in\Omega_k^l} \sup_{\mathbf{x}\in\mathbb{R}^n} ||\nabla n_{j,h}(\mathbf{x})||_\infty \tag{5.3.30}$$

and recall that $\boldsymbol{\varpi}_k^{l,*}$ is the optimal solution recovered when using (P4) to approximate $p_{k,k}^{l}(\mathbf{x})$.

Theorem 5.1. *The total variation distance between the actual density and the approximation density at t_1 is*

$$||\tilde{p}_{1,1}^{l}||_1 \leq q(e_1^{l,2} + e_1^{l,3} + e_1^{l,4}) \qquad \forall l = 1,\ldots,d \tag{5.3.31}$$

where

$$\begin{aligned} e_1^{l,2} &:= \tfrac{1}{2}C^{(l)} n\rho^{n+1} E_{n(\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0)}[\lambda^l] \\ e_1^{l,3} &:= \epsilon_1^l \rho^n \\ e_1^{l,4} &:= C_1 n\rho^{n+1} \cdot ||\boldsymbol{\varpi}_1^{l,*}||_1 \end{aligned} \tag{5.3.32}$$

Let us denote by $\mathbf{S}(l) = \mathbf{H}\mathbf{Q}(l)\mathbf{H}^T + \mathbf{R}(l)$ the mode conditioned measurement prediction covariance matrix.

Theorem 5.2. *Suppose*

$$||\tilde{p}^{r}_{k-1,k-1}||_1 \leq \delta_r \qquad \forall r = 1, \ldots d \tag{5.3.33}$$

Then, the total variation distance at time t_k is

$$||\tilde{p}^{l}_{k,k}||_1 \leq e^{l,1}_k + q(e^{l,2}_k + e^{l,3}_k + e^{l,4}_k) \tag{5.3.34}$$

for any $l = 1, \ldots, d$, where

$$\begin{aligned}
e^{l,1}_k &:= \frac{1}{(2\pi)^{\frac{n}{2}} \det(\mathbf{S}_l)^{\frac{1}{2}}} \sum_{r=1}^{d} \sup_{\mathbf{y} \in \mathbb{R}^n} \lambda^{r,l}(\mathbf{y}) \delta_r \\
e^{l,2}_k &:= \frac{1}{2} C^{(l)} n \rho^{n+1} \sum_{r=1}^{d} \sum_{(i,j) \in \Omega^r_k} [\boldsymbol{\varpi}^{r,*}_{k-1}]_{(i,j)} E_{n_{i,j}}[\lambda^{r,l}] \\
e^{l,3}_k &:= \epsilon^l_k \rho^n \\
e^{l,4}_k &:= C_k n \rho^{n+1} \cdot ||\boldsymbol{\varpi}^{l,*}_k||_1
\end{aligned} \tag{5.3.35}$$

In case when the transition matrix of the chain P^{Δ} is state independent, then $\lambda^{r,l}(\mathbf{y})$ becomes independent on $\mathbf{y}$ and the terms $e^{l,1}_k$ and $e^{l,2}_k$ given in Theorem 5.2 take the following form

$$\begin{aligned}
e^{l,1}_k &= \frac{1}{(2\pi)^{\frac{n}{2}} \det(\mathbf{S}_l)^{\frac{1}{2}}} \sum_{r=1}^{d} \lambda^{r,l} \delta_r \\
e^{l,2}_k &= \frac{1}{2} C^{(l)} n \rho^{n+1} \sum_{r=1}^{d} \lambda^{r,l} ||\boldsymbol{\varpi}^{r,*}_{k-1}||_1
\end{aligned} \tag{5.3.36}$$

5.3.5 Computational Requirements

In this section we compare the computational complexity of our methodology with the one of the IMM estimator (see Blom and Bar-Shalom, 1988), a widely used heuristic to compute the filter density of a jump linear system. The IMM consists of two major steps:

- Interaction step. The initial condition for the filter matched to a certain mode is obtained by mixing the state estimates of all filters at the previous time under the assumption that this particular mode is in effect at the current time. This stage has complexity $O(d^2)$.

- Filtering step. The prediction and update steps are performed for each mode using the newly received measurement along with the initial condition matched to the same mode and computed earlier. This is done using a Kalman filter and thus has complexity $O(n^3)$, with n being the dimension of the state space.

Therefore, the overall complexity of the method is $O(d^2 + n^3)$. Our approach has higher computational complexity than the IMM approach due to the amount of time needed to set up and solve the second order cone programming problem (P2). Before solving the optimization problem, it is required to compute $\check{\boldsymbol{p}}^{l}_{k,k}$

for any mode l and this has complexity $O(d^2|I||J|q)$. This is because $\check{p}^l_{k,k}(\mathbf{x})$ in Eq. (5.3.17) may be written as

$$\check{p}^l_{k,k}(\mathbf{x}) = \sum_{r=1}^{d} \sum_{(i,j)\in\Omega^l_k} [\boldsymbol{\varpi}^r_{k-1}]_{(i,j)}\ \mathbb{E}_{n_{i,j}}[L^l_{k|k}(\mathbf{x})\lambda^{r,l}(\mathbf{Y}_{i,j})p(\mathbf{x}|\mathbf{Y}_{i,j},l)] \tag{5.3.37}$$

where $\mathbf{Y}_{i,j}$ is a Gaussian random vector with mean $\boldsymbol{\mu}_i$ and covariance $\boldsymbol{\Sigma}_j$. For each $\mathbf{x} \in \boldsymbol{\mathfrak{X}}$ in the training set, it is required to compute $d|I||J|$ Gaussian integrals which can be done in constant time using Hermite polynomials. This has to be repeated for each mode, thus the total complexity is $O(d^2q|\boldsymbol{\mathfrak{X}}||I||J|)$. Then the optimization problem needs to be solved for each mode, and this can be done in time $O((|I||J|)^2q)$ using the algorithm proposed by Lobo et al. (1998).

Thus the complexity of the approach grows quadratically with the size of the base set and the number of modes, and linearly in the size of the training set. This makes the approach slower than the IMM estimator for high dimensional systems where a larger base set and training set would be required to have good approximation quality. However, the approach would be very efficient for low-dimensional systems.

5.3.6 Quantitative Evaluation of the TVD

We perform a quantitative evaluation of the total variation distance upper bound derived in Section 5.3.4. This is done by means of a simple experiment considering a one-dimensional bimodal state space system defined as

$$\begin{aligned} x_k &= x_{k-1} + Q(l)w_k \\ z_k &= x_k + Rv_k \end{aligned} \tag{5.3.38}$$

where w_k is a white noise sequence with variance depending on the mode in place at time k, while the measurement noise v_k is assumed to be mode-independent. The two-mode dependent process noises are zero-mean and only differ in the choice of the standard deviation which is set to $Q(1) = 0.7\frac{m}{s^2}$ for mode 1, and to $Q(2) = 4\frac{m}{s^2}$ for mode 2. The standard deviation of the measurement noise is $R = 2\frac{m}{s^2}$. The duration of each sampling period is 1.5 seconds.

The mode transitions are assumed to be state independent with transition matrix given by

$$\lambda^{i,j} = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix} \tag{5.3.39}$$

and initial mode distribution given by the vector $[0.95, 0.05]$. The true trajectory is simulated for 14 sampling times using a single process noise sequence having standard deviation $0.5\frac{m}{s^2}$ and zero mean for all times except for the sampling times from $k = 3$ to $k = 7$ where

$$\begin{array}{ccc} E[w_3] = 10\frac{m}{s^2} & E[w_4] = 9\frac{m}{s^2} & E[w_5] = 8\frac{m}{s^2} \\ E[w_6] = 7\frac{m}{s^2} & E[w_7] = 6\frac{m}{s^2} & \end{array}$$

The target is assumed to start at $x = 0$. The approximation accuracy ϵ in the programming problem (P2) is set to 10^{-5}. The set of μs used in the base set of Gaussian densities consists of equally spaced values ranging

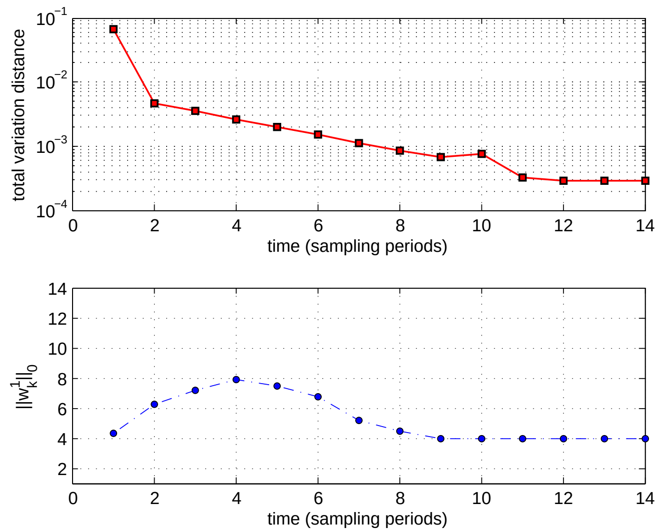

Figure 5.2: The top graph reports the total variation distance upper bound for the density $p_{k,k}^1(\boldsymbol{x}),\ k = 1, \ldots, 14$. The bottom graph reports the number of nonzero components selecting from the programming problem (P2) in the approximation $\hat{p}_{k,k}^1(x),\ k = 1, \ldots, 14$.

from $\mu_{\min} = -80\ m$ to $\mu_{\max} = 80\ m$, with distance $1m$ between consecutive μs. The standard deviations of the base densities are chosen to be exponentially decaying, with the largest being $\sigma_{\max} = 1$, and $\sigma_j = \frac{1}{1.3^j}$, for $j = 1, \ldots, 10$. This results in a base set of 1771 Gaussian densities. The training set of points x is chosen to be equally spaced from $-120\ m$ to $120\ m$, with distance $1\ m$ from each other.

We perform a number of one-hundred Monte-carlo runs. The scale of values taken by the actual unnormalized density is on the order of $10^{-2} - 10^{-3}$.

Figures 5.2 and 5.3 illustrate the behavior of the total variation bound for the unnormalized density $p_{k,k}^1(x)$ and $p_{k,k}^2(x)$, and also report the number of nonzero weights selected in the programming problem (P2). The total variation distance upper bound plotted in the figures is recursively on-line computed for each sampling time using formula (5.3.34). As evident from Eq. (5.3.34), the evaluation of the bound requires knowing the weights contributing to the density approximations $\hat{p}_{k,k}^1(x)$ and $\hat{p}_{k,k}^2(x)$ at each time k. Those weights are recovered as the solution of the two convex programming problems $P2$, one for each mode. We solve them using CVX, an efficient convex optimization solver (see Grant and Boyd, 2008) It appears from the graphs that, after an initial setup period, the total variation distance between the actual and the approximation densities becomes on the order of $10^{-3} - 10^{-4}$, a quite satisfactory result if we consider that the order of magnitude of the density is about 10^{-3}. Moreover, the number of needed Gaussian components in the approximation stabilizes at four, after an initial setup period where the number of approximation densities reaches a peak of 7.91 (for the first mode) and 12.25 (for the second mode). This means that few components are needed in each approximation step. Moreover, it appears that the approximation error introduced by the method is kept low and has an overall tendency to decrease over time.

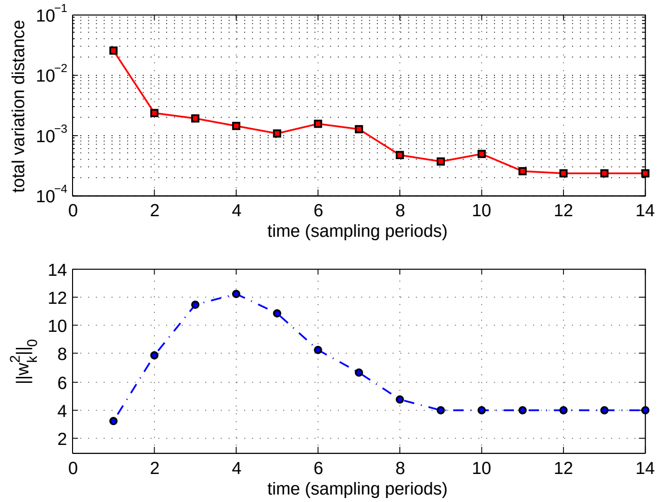

Figure 5.3: Total variation distance upper bound for the density $p^2_{k,k}(\boldsymbol{x}), k = 1, \ldots, 14$. The bottom graph reports the number of nonzero components selecting from the programming problem (P2) in the approximation $\hat{p}^2_{k,k}(x),\ k = 1, \ldots, 14$.

5.4 A Target Tracking Example

We evaluate the prediction accuracy and the computational power of our scheme on a target tracking problem. The considered scenario consists of an aircraft flying in the (x^1, x^2) plane at an initial velocity $[0, 120\frac{m}{s}]$, along the vertical direction. The aircraft is supposed to execute a 90° right turn (with turn rate of $\frac{3^\circ}{s}$) when the trajectory change point (TCP_P) is reached, see Figure 5.4 for an illustration. The sensor sampling period is 10 seconds. After the turn, the aircraft continues straight along the horizontal direction. However, in practice it can turn earlier at a difference trajectory change point (TCP_A) and we assume that the distance d_{PA} is Gaussian with zero mean and standard deviation $\sqrt{10}$ meters.

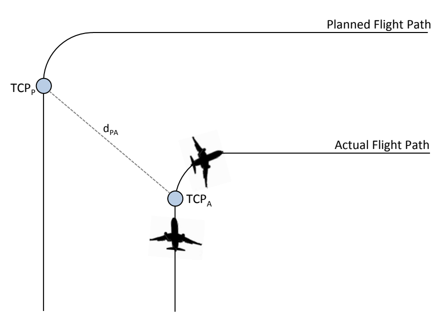

Figure 5.4: Actual versus expected trajectory of the aircraft

This scenario is challenging because the deviation from non-manoeuvering motion is very large, while the duration of the manoeuver is short. The second-order system for this motion, with position-only measurements is:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathrm{diag}(\mathbf{A},\mathbf{A})\mathbf{x}_k + \mathrm{diag}(\mathbf{\Gamma},\mathbf{\Gamma})\mathbf{w}_k \\ \mathbf{z}_k &= \mathrm{diag}(\mathbf{H},\mathbf{H})\mathbf{x}_k + \mathbf{v}_k \end{aligned} \tag{5.4.1}$$

where

$$\mathbf{x}_k = [x_k^1 \ \dot{x}_k^1 \ x_k^2 \ \dot{x}_k^2]' \qquad \mathbf{z}_k = [z_k^1 \ z_k^2]'$$

$$\mathbf{A} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \qquad \mathbf{\Gamma} = \begin{bmatrix} \frac{1}{2}T^2 \end{bmatrix} \qquad \mathbf{H} = [1 \ 0]$$

The transition probability from the straight mode (mode "1") to the maneuvering mode (mode "2"), here denoted as $\lambda^{1,2}(\mathbf{x})$, is defined as

$$\lambda^{1,2}(\mathbf{x}) \quad := \quad \lambda^{1,2}(x,y) = \alpha_{1,2} + \beta_{1,2} n([x,y],[0,\mathrm{TCP_P}],\mathbf{M}) \tag{5.4.2}$$

From Eq. (5.4.2), we see that the probability of turning is only dependent on the (x, y) aircraft position and is independent of the other state components. For this experiment, we set $\alpha_{1,2} = 0.03$, $\beta_{1,2} = 16\pi$, and $\mathbf{M} = 10m \cdot \mathbf{I}_2$, where $\mathbf{I}_2$ denotes the 2×2 identity matrix. Here $\alpha_{1,2}$ is modeling the probability that the aircraft turns at a different point than TCP, while $\beta_{1,2}$ models the probability that the aircraft turns at the expected TCP. The two process noise sequences assumed in the model are zero-mean and only differ in the choice of the process noise covariance $\boldsymbol{Q}$, which is set to a two-dimensional diagonal matrix with equal entries. For the straight mode (also referred to as mode 1) such entries are set to $(0.3\frac{m}{s^2})^2$, and for the maneuvering mode (also referred to as mode 2) such entries are set to $(6\frac{m}{s^2})^2$. The covariance matrix of the measurement noise is two dimensional with diagonal entries set to $(100m)^2$ and is the same for both modes. For the true dynamics, the process noise covariance is assumed to be a two-dimensional diagonal matrix having both entries equal to $(0.4\frac{m}{s^2})^2$. The maneuver is obtained by changing the mean of the process noise $\mathbb{E}[\mathbf{w_k}]$ from time $t_k = 7$ to $t_k = 9$ to $\mathbb{E}[\mathbf{w}_7] = \mathbb{E}[\mathbf{w}_9] = (6, -1.61)'m/s^2$ and $\mathbb{E}[\mathbf{w}_8] = (4.39, -4.39)'m/s^2$. At all other times $\mathbb{E}[\mathbf{w}_k] = 0$. The base set B consisted of about 700 Gaussian densities and the training set $\boldsymbol{\mathfrak{X}}$ of about 50 points for each time step.

We compare our method with the IMM and GPB2 estimator and the reported results consist of an average across one-hundred Monte-Carlo runs. The transition probability function used within IMM and GPB2 is

$$\lambda(\mathbb{E}[\boldsymbol{x}_k|\mathscr{F}_{k-1}^z]) \tag{5.4.3}$$

i.e., the function λ in Eq. (5.4.2) evaluated at the state predicted at time k on the basis of the observations $\mathscr{F}_{k-1}^z$. The results in Figures 5.5–5.6 show that our method outperforms both IMM and GPB2 in terms of estimation accuracy. Our method has generally a better estimation performance than IMM and GPB2, with GPB2 sometimes coming closer to our method, especially for the estimation of velocity components. This

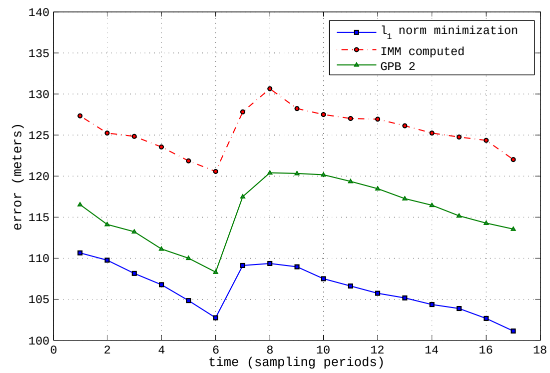

Figure 5.5: Coordinate-combined position estimation error

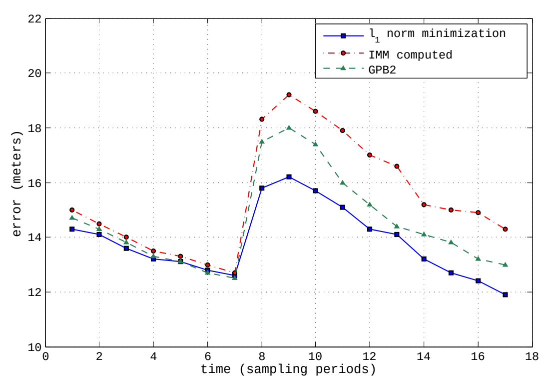

Figure 5.6: Coordinate-combined velocity estimation error

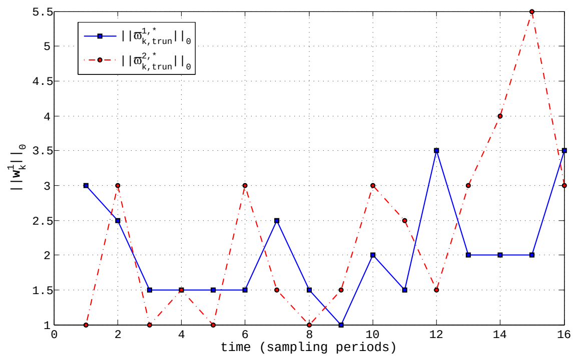

Figure 5.7: The number of nonzero Gaussian components used in the density approximations $\hat{p}^1_{k,k}(x)$ and $\hat{p}^2_{k,k}(x)$

may be attributed to the fact that our method always propagates more than two densities at each sampling step, whereas the IMM always propagates only two per time step. Our method also exhibits a slightly better tracking performance during the manouevering stage which may be explained by the fact that is considering actual state-dependent transitions (using Eq. (5.4.2)), whereas IMM and GPB2 consider estimated state dependent transitions (using Eq. (5.4.3)).

With respect to the computational time, our method is slower. For each Monte Carlo run, we took the average time needed to compute the state estimate over all sampling times. We then averaged out those times across all runs. Times are measured on a 2.20 GHz processor with 4 Gb of RAM. The times required by IMM, GPB2 and the proposed method are, respectively, 5.0 msec, 10 msec, and 150 msec, showing that our method is slower than IMM by a factor of 30 and slower than GPB2 by a factor of 15.

Chapter 6

Bilateral Counterparty Risk Valuation

6.1 Introduction

It has been reported by Canabarro and Duffie (2004) that the volume of outstanding over-the-counter derivatives has grown exponentially over the past 15 years. More specifically, market surveys conducted by the International Swaps and Derivatives Association (ISDA) show notional amounts of outstanding interest rate and currency swaps reaching 866 billion US dollars in 1987, 17.7 trillion US dollars in 1995, and 99.8 trillion US dollars in 2002; an astonishing compounded growth rate of 37.2% per year. Over-the-counter contracts are traded (and privately negotiated) directly between the two parties, without going through an exchange or other intermediary. Such contracts include, for example, credit default swaps, forward rate agreements, and interest rate swaps see (see Brigo and Mercurio, 2006). Since they are not traded on an exchange, there is no central clearinghouse that insures against the default risk of both parties, therefore both parties of the contracts are exposed to *counterparty risk*, namely the risk that one party defaults on its payment obligation.

Correct modeling and valuation of counterparty risk has become even more relevant in recent years as some protection sellers such as mono-line insurers or investment banks have witnessed increasing default probabilities or even default events, the case of Lehman Brothers being a clear example. Moreover, regulatory issues related to the IAS 39 framework encourage the inclusion of counterparty risk into the contract valuation.

In this chapter we introduce a general arbitrage-free valuation framework for bilateral counterparty default risk and assume absence of guarantees such as collateral. We call the two parties of the contract the *investor* and her *counterparty*. By 'bilateral' we intend to point out that the default of the investor is included into the framework, contrary to earlier works. This brings about symmetry, so that the price of the position including counterparty risk to the investor is exactly the opposite of the price of the position to the counterparty. This is clearly not the case if each of the two parties computes the present value assuming itself to be default-free and allowing for default of the other party only. This asymmetry would not matter in situations where financial investors had high credit quality and counterparties rather low one. Indeed, in such a case both parties would consider the investor as default-free and the counterparty as defaultable, so that inclusion of the investor default would be pointless, given that it happens in almost no scenario. However, recent events show that it is no longer realistic to take the credit quality of the financial institution for granted and to be highly superior to that of a general counterparty, no matter how prestigious or important the financial institution.

This chapter generalizes to the bilateral symmetric case the asymmetric unilateral setting proposed in Brigo and Chourdakis (2008), Brigo and El-Bachir (2008), Brigo and Masetti (2005), and Brigo and Pallavicini (2007), where only the counterparty is subject to default risk, while the investor is assumed to be default free. We provide a general formula that gives the bilateral risk credit valuation adjustment (BR-CVA) for portfolios exchanged between a default risky investor and a default risky counterparty. Such formula shows that the adjustment to the investor is the difference between two discounted options terms, a discounted call option in scenarios of early default time of the counterparty minus a discounted put option in scenarios of early default times of the investor, both options being on the residual net present value of the portfolio at the relevant default times and having zero strike. The BR-CVA seen from the point of view of the counterparty is exactly the opposite. We allow for correlation between default of the investor, default of the counterparty and underlying portfolio risk factors, and for volatilities and dynamics in the credit spreads and in the underlying portfolio, all arbitrage free.

We then specialize our analysis to Credit Default Swaps (CDS) as underlying portfolio, generalizing the work of Brigo and Chourdakis (2008) who deal with unilateral and asymmetric counterparty risk for these contracts. Featuring a CDS as underlying, a third default time enters the picture, namely the default time for the reference credit of the CDS. We therefore assume that all three entities are subject to default risk and that the default events of investor, counterparty and reference credit are correlated. We then propose a numerical methodology to evaluate the resulting BR-CVA formula. We investigate the impact of both credit spread volatility and default correlation on the credit valuation adjustment. Most of previous approaches on CDS counterparty risk only focus on unilateral counterparty risk. One exception is Duffie and Huang (1996) who considers the bilateral setting under a much simpler static model of counterparty exposure, which is calculated on average across a market. Moreover, most of the existing studies ignore the effect of volatility on the adjustment and mainly focus on default correlation. Hull and White (2000) study the counterparty risk problem by resorting to barrier correlated models. Collin-Dufresne et al. (2004) model default intensities as deterministic constants using default indicators of other names as feeds. The need for explicitly modeling credit spread volatility is even more pronounced in the case of CDS if the underlying reference contract is itself a CDS, as the credit valuation adjustment would involve CDS options and it is very undesirable to model options without volatility in the underlying asset.

We analyze the impact of credit spreads volatility and default correlation through some numerical examples and find significant sensitivity to credit spreads volatility, especially under scenarios of large default correlation.

The rest of the chapter is organized as follows. Section 6.2 gives the general BR-CVA formula and refers to the appendix for the mathematical proof. Section 6.3 defines the underlying framework needed for applying the above methodology to credit default swaps, including the stochastic intensity model and the trivariate copula function for correlating defaults. Section 6.4 develops a numerical method to implement the formula in the case when the underlying contract is a CDS and gives the pseudo code of the algorithm used to run numerical simulations. Section 6.5 present numerical experiments on different default correlation scenarios and changes in credit spread volatility. Section 6.6 presents a specific application for computing the mark-to-market value of a CDS contract agreed between British Airways, Lehman Brothers, and Royal Shell.

6.2 Arbitrage-Free Valuation of Bilateral Counterparty Risk

We refer to the two names involved in the financial contract and subject to default risk as

$$\begin{aligned} \text{investor} &\quad\rightarrow\quad \text{name "0"} \\ \text{counterparty} &\quad\rightarrow\quad \text{name "2"} \end{aligned}$$

In cases where the portfolio exchanged by the two parties is also a default sensitive instrument, we introduce a third name referring to the reference credit of that portfolio

$$\text{reference credit} \ \rightarrow \text{name "1"}$$

If the portfolio is not default sensitive then name "1" can be removed.

We denote by τ_0, (τ_1), and τ_2, respectively the default times of the investor, (reference credit), and counterparty. We place ourselves in a probability space $(\Omega, \mathcal{G}, \mathcal{G}_t, \mathbb{Q})$. The filtration $\mathcal{G}_t$ models the flow of information of the whole market, including credit and $\mathbb{Q}$ is the risk neutral measure. This space is endowed also with a right-continuous and complete sub-filtration $\mathcal{F}_t$ representing all the observable market quantities but the default event, thus $\mathcal{F}_t \subseteq \mathcal{G}_t := \mathcal{F}_t \vee \mathcal{H}_t$. Here, $\mathcal{H}_t = \sigma(\{\tau_0 \leq u\}(\vee\{\tau_1 \leq u\}) \vee \{\tau_2 \leq u\} : u \leq t)$ is the right-continuous filtration generated by the default events, either of the investor or of his counterparty (and of the reference credit if the underlying portfolio is credit sensitive).

Let us call T the final maturity of the payoff which we need to evaluate and let us define the stopping time

$$\tau = \min\{\tau_0, \tau_2\} \tag{6.2.1}$$

If $\tau > T$, there is neither default of the investor, nor of his counterparty during the life of the contract and they both can fulfill the agreements of the contract. On the contrary, if $\tau \leq T$ then either the investor or her counterparty (or both) defaults. At τ, the Net Present Value (NPV) of the residual payoff until maturity is computed. We then distinguish two cases:

- $\tau = \tau_2$. If the NPV is negative (respectively positive) for the investor (defaulted counterparty), it is completely paid (received) by the investor (defaulted counterparty) itself. If the NPV is positive (negative) for the investor (counterparty), only a recovery fraction $\mathrm{R}_{\mathrm{EC},2}$ of the NPV is exchanged.

- $\tau = \tau_0$. If the NPV is positive (respectively negative) for the defaulted investor (counterparty), it is completely received (paid) by the defaulted investor (counterparty) itself. If the NPV is negative (positive) for the defaulted investor (counterparty), only a recovery fraction $\text{R}_{\text{EC},0}$ of the NPV is exchanged.

Let us define the following (mutually exclusive and exhaustive) events ordering the default times

$$\begin{aligned}
A &= \{\tau_0 \le \tau_2 \le T\} \qquad E = \{T \le \tau_0 \le \tau_2\} \\
B &= \{\tau_0 \le T \le \tau_2\} \qquad F = \{T \le \tau_2 \le \tau_0\} \\
C &= \{\tau_2 \le \tau_0 \le T\} \\
D &= \{\tau_2 \le T \le \tau_0\}
\end{aligned} \tag{6.2.2}$$

Let us call $\Pi^D(t,T)$ the discounted payoff of a generic defaultable claim at t and $\text{C}_{\text{ASHFLOWS}}(u,s)$ the net cash flows of the claim without default between time u and time s, discounted back at u, all payoffs seen from the point of view of the investor. Then, we have $\text{NPV}(\tau_i) = \mathbb{E}_{\tau_i}\{\text{C}_{\text{ASHFLOWS}}(\tau_i,T)\}$, $i = 0,2$. Let us denote by $D(t,T)$ the price of a zero coupon bond with maturity T. We have

$$\begin{aligned}
\Pi^D(t,T) = \; & \mathbf{1}_{E\cup F}\text{C}_{\text{ASHFLOWS}}(t,T) \\
& +\mathbf{1}_{C\cup D}\left[\text{C}_{\text{ASHFLOWS}}(t,\tau_2) + D(t,\tau_2)\left(\text{R}_{\text{EC},2}\,(\text{NPV}(\tau_2))^+ - (-\text{NPV}(\tau_2))^+\right)\right] \\
& +\mathbf{1}_{A\cup B}\left[\text{C}_{\text{ASHFLOWS}}(t,\tau_0) + D(t,\tau_0)\left((\text{NPV}(\tau_0))^+ - \text{R}_{\text{EC},0}\,(-\text{NPV}(\tau_0))^+\right)\right]
\end{aligned} \tag{6.2.3}$$

This last expression is the general price of the payoff under bilateral counterparty default risk. Indeed, if there is no early default, this expression reduces to risk neutral valuation of the payoff (first term on the right-hand side). In case of early default of the counterparty, the payments due before default occurs are received (second term), and then if the residual net present value is positive only the recovery value of the counterparty $\text{R}_{\text{EC},2}$ is received (third term), whereas if it is negative it is paid in full by the investor (fourth term). In case of early default of the investor, the payments due before default occurs are received (fifth term), and then if the residual net present value is positive it is paid in full by the counterparty to the investor (sixth term), whereas if it is negative only the recovery value of the investor $\text{R}_{\text{EC},0}$ is paid to the counterparty (seventh term).

Let us denote by $\Pi(t,T)$ the discounted payoff for an equivalent claim with a default-free counterparty, i.e., $\Pi(t,T) = \text{C}_{\text{ASHFLOWS}}(t,T)$. We then have the following

Proposition 6.2.1. (General bilateral counterparty risk pricing formula). *At valuation time t, and conditional on the event $\{\tau > t\}$, the price of the payoff under bilateral counterparty risk is*

$$\begin{aligned}
\mathbb{E}_t\left\{\Pi^D(t,T)\right\} = \; & \mathbb{E}_t\left\{\Pi(t,T)\right\} \\
& +\text{LGD}_0 \cdot \mathbb{E}_t\left\{\mathbf{1}_{A\cup B}\cdot D(t,\tau_0)\cdot[-\text{NPV}(\tau_0)]^+\right\} \\
& -\text{LGD}_2 \cdot \mathbb{E}_t\left\{\mathbf{1}_{C\cup D}\cdot D(t,\tau_2)\cdot[\text{NPV}(\tau_2)]^+\right\}
\end{aligned} \tag{6.2.4}$$

where $LGD = 1 - R_{EC}$ is the Loss Given Default *and the recovery fraction R_{EC} is assumed to be deterministic. It is clear that the value of a defaultable claim is the value of the corresponding default-free claim plus a long position in a put option (with zero strike) on the residual NPV giving nonzero contribution only in scenarios where the investor is the earliest to default (and does so before final maturity) plus a short position in a call option (with zero strike) on the residual NPV giving nonzero contribution in scenarios where the counterparty is the earliest to default (and does so before final maturity),*

Proposition 6.2.1 is stated in Brigo (2008) without a proof. Here, we provide a mathematical proof in Appendix J. The adjustment is called bilateral counterparty risk credit valuation adjustment (BR-CVA) and it may be either positive or negative depending on whether the counterparty is more or less likely to default than the investor and on the volatilities and correlation. The mathematical expression is given by

$$\begin{aligned} \text{BR-CVA}(t, T, \text{LGD}_{0,1,2}) \quad &= \quad \text{LGD}_2 \cdot \mathbb{E}_t \left\{ \mathbf{1}_{C \cup D} \cdot D(t, \tau_2) \cdot [\text{NPV}(\tau_2)]^+ \right\} \\ &\quad -\text{LGD}_0 \cdot \mathbb{E}_t \left\{ \mathbf{1}_{A \cup B} \cdot D(t, \tau_0) \cdot [-\text{NPV}(\tau_0)]^+ \right\} \end{aligned} \tag{6.2.5}$$

where the right-hand side in Eq. (6.2.5) depends on T through the events A, B, C, D, and $\text{LGD}_{012} = (\text{LGD}_0, \text{LGD}_1, \text{LGD}_2)$ is shorthand notation to denote the dependence on the loss given defaults of the three names.

Remark 6.2.1. (Symmetry versus Asymmetry). *With respect to earlier results on counterparty risk valuation, Eq. (6.2.5) has the great advantage of being symmetric. This is to say that if "2" were to compute counterparty risk of her position towards "1", she would find exactly $-\text{BR-CVA}(t, T, \text{LGD}_{0,1,2})$. However, if each party computed the adjustment by assuming itself to be default-free and considering only the default of the other party, then the adjustment calculated by "0" would be*

$$\text{LGD}_2 \cdot \mathbb{E}_t \left\{ \mathbf{1}_{\tau_2 < T} \cdot D(t, \tau_2) \cdot [\text{NPV}(\tau_2)]^+ \right\}$$

whereas the adjustment calculated by "2" would be

$$\text{LGD}_0 \cdot \mathbb{E}_t \left\{ \mathbf{1}_{\tau_0 < T} \cdot D(t, \tau_0) \cdot [-\text{NPV}(\tau_0)]^+ \right\}$$

and they would not be each one the opposite of the other. This means that only in the first case the two parties agree on the value of the counterparty risk adjustment.

Remark 6.2.2. (Change in sign). *Earlier results on counterparty risk valuation, concerned with a default-free investor and asymmetric, would find an adjustment to be subtracted that is always positive. However, in our symmetric case even if the initial adjustment is positive due to*

$$\text{LGD}_2 \cdot \mathbb{E}_t \left\{ \mathbf{1}_{C \cup D} \cdot D(t, \tau_2) \cdot [\text{NPV}(\tau_2)]^+ \right\} > \text{LGD}_0 \cdot \mathbb{E}_t \left\{ \mathbf{1}_{A \cup B} \cdot D(t, \tau_0) \cdot [-\text{NPV}(\tau_0)]^+ \right\}$$

the situation may change in time, to the point that the two terms may cancel or that the adjustment may *change sign as the credit quality of "0" deteriorates and that of "2" improves, so that the inequality changes direction.*

6.3 Application to Credit Default Swaps

In this section we use the formula developed in Section 6.2 to evaluate the BR-CVA in credit default swap contracts (CDS). Section 6.3.1 recalls the general formula for CDS evaluation. Section 6.3.2 introduces the copula models used to correlate the default events. Section 6.3.3 recalls the CIR model used for the stochastic intensity of the three names. Section 6.3.4 applies the general BR-CVA formula to calculate the adjustment for CDS contracts.

6.3.1 CDS Payoff

We assume deterministic interest rates, which leads to independence between τ_1 and $D(0,t)$. Our results hold also true for the case of stochastic rates independent of default times. The receiver CDS valuation, for a CDS selling protection LGD_1 at time 0 for default of the reference entity between times T_a and T_b in exchange of a periodic premium rate S_1 is given by

$$\begin{aligned}\mathrm{CDS}_{a,b}(0,S_1,\mathrm{LGD}_1) &= S_1\Big[-\int_{T_a}^{T_b} D(0,t)(t-T_{\gamma(t)-1})d\mathbb{Q}(\tau_1>t)\\ &\quad+\sum_{i=a+1}^{b}\alpha_i D(0,T_i)\mathbb{Q}(\tau_1>T_i)\Big]\\ &\quad+\mathrm{LGD}_1\Big[\int_{T_a}^{T_b} D(0,t)d\mathbb{Q}(\tau_1>t)\Big] \qquad (6.3.1)\end{aligned}$$

where $\gamma(t)$ is the first payment period T_j following time t. Let us denote by

$$\mathrm{NPV}(T_j,T_b) := \mathrm{CDS}_{a,b}(T_j,S,\mathrm{LGD}_1) \qquad (6.3.2)$$

the residual NPV of a receiver CDS between T_a and T_b evaluated at time T_j, with $T_a < T_j < T_b$. Eq. (6.3.2) can be written similarly to Eq. (6.3.1), except that now evaluation occurs at time T_j and has to be conditioned on the information set available to the market at T_j. This leads to

$$\begin{aligned}\mathrm{CDS}_{a,b}(T_j,S_1,\mathrm{LGD}_1) &= \mathbf{1}_{\tau_1>T_j}\overline{\mathrm{CDS}}_{a,b}(T_j,S_1,LGD_1)\\ &:= \mathbf{1}_{\tau_1>T_j}\Big\{S_1\Big[-\int_{\max\{T_a,T_j\}}^{T_b} D(T_j,t)(t-T_{\gamma(t)-1})d\mathbb{Q}(\tau_1>t|\mathcal{G}_{T_j})\\ &\quad+\sum_{i=\max\{a,j\}+1}^{b}\alpha_i D(T_j,T_i)\mathbb{Q}(\tau_1>T_i|\mathcal{G}_{T_j})\Big]\\ &\quad+\mathrm{LGD}_1\Big[\int_{\max\{T_a,T_j\}}^{T_b} D(T_j,t)d\mathbb{Q}(\tau_1>t|\mathcal{G}_{T_j})\Big]\Big\} \qquad (6.3.3)\end{aligned}$$

6.3.2 Default Correlation

We consider a reduced form model that is stochastic in the default intensity for the investor, counterparty and CDS reference credit. The default correlation between the three names is defined through a dependence structure on the exponential random variables characterizing the default times of the three names. Such dependence structure is modeled using a trivariate copula function. A brief overview of copula functions is provided in Appendix I. For additional information about the subject, the reader is referred to Nelsen (2006) and Cherubini et al. (2004).

Let us denote by $\lambda_i(t)$ and $\Lambda_i(t) = \int_0^t \lambda_i(s)ds$, respectively, the default intensity and cumulated intensity of name i evaluated at time t. We recall that $i = 0$ refers to the investor, $i = 1$ refers to the reference credit and $i = 2$ to the counterparty. We assume λ_i to be independent of λ_j for $i \neq j$, and assume each of them to be strictly positive almost everywhere, thus implying that Λ_i is invertible. We stress the fact that independence of λs across names does not mean that the default event of one name does not change the default probability or intensity of other names, as discussed in Brigo and Chourdakis (2008), Section 4. We place ourselves in a Cox process setting, where

$$\tau_i = \Lambda_i^{-1}(\xi_i), \; i = 0, 1, 2 \tag{6.3.4}$$

with ξ_0, ξ_1, and ξ_2 being standard (unit-mean) exponential random variables whose associated uniforms

$$U_i = 1 - \exp\{-\xi_i\} \tag{6.3.5}$$

are correlated through a Gaussian trivariate copula function

$$C_{\boldsymbol{R}}(u_0, u_1, u_2) = \mathbb{Q}(U_0 < u_0, U_1 < u_1, U_2 < u_2) \tag{6.3.6}$$

with $\boldsymbol{R} = [r_{i,j}]_{i,j=0,1,2}$ being the correlation matrix parameterizing the Gaussian copula. Notice that a trivariate Gaussian copula implies bivariate Gaussian marginal copulas. We show it for the case of the bivariate copula connecting the reference credit and the counterparty, but the same argument applies to the two other bivariate Gaussian copulas. Let (X_1, X_2, X_3) be a standard Gaussian vector with correlation matrix $\boldsymbol{R}$ and let $\Phi_{\boldsymbol{R}}$ be the distribution function of a multivariate Gaussian random variable with correlation matrix $\boldsymbol{R}$. For any pair of indices $i \neq j$, $0 \leq i, j \leq 2$, we denote by $\boldsymbol{R}_{i,j}$ the $2 \cdot 2$ submatrix formed by the intersection of row i and row j with column i and column j. We then have

Lemma 6.1. *A trivariate Gaussian copula with correlation matrix $\boldsymbol{R}$ induces marginal bivariate Gaussian copulas.*

Proof. We prove that the bivariate Gaussian copula $C_{1,2}(u_1, u_2)$ connecting the reference credit and the counterparty is obtained from the trivariate Gaussian copula by taking the submatrix $\boldsymbol{R}$ formed by the intersection of rows 1 and 2 with columns 1 and 2. We have

$$\begin{aligned}
\mathbb{Q}(U_1 < u_1, U_2 < u_2) &= \mathbb{Q}(U_1 < u_1, U_2 < u_2, U_3 < 1) \\
&= C_{\boldsymbol{R}}(u_1, u_2, 1) \\
&= \lim_{b \to 1^-} \Phi_{\boldsymbol{R}}(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \Phi^{-1}(b)) \\
&= \lim_{b \to 1^-} \mathbb{Q}(X_1 < \Phi^{-1}(u_1), X_2 < \Phi^{-1}(u_2), X_3 < \Phi^{-1}(b)) \\
&= \mathbb{Q}(X_1 < \Phi^{-1}(u_1), X_2 < \Phi^{-1}(u_2)) \cdot \\
&\quad \lim_{b \to 1^-} \mathbb{Q}(X_3 < \Phi^{-1}(b) | X_1 < \Phi^{-1}(u_1), X_2 < \Phi^{-1}(u_2)) \\
&= \Phi_{\boldsymbol{R}_{1,2}}(\Phi^{-1}(u_1), \Phi^{-1}(u_2)) \\
&:= C_{1,2}(u_1, u_2) \qquad (6.3.7)
\end{aligned}$$

By a similar calculation, the bivariate copula connecting the investor and the reference credit is given by $C_{0,1}$ with associated correlation matrix $\boldsymbol{R}_{0,1}$ and the bivariate copula connecting the investor and the counterparty is given by $C_{0,2}$ with associated correlation matrix $\boldsymbol{R}_{0,2}$. □

6.3.3 CIR Stochastic Intensity Model

We use the CIR process (see Cox et al., 1985) to model the default intensity of the investor, reference credit and counterparty. We do not introduce correlation between the three intensity processes, but model default correlation through a trivariate copula function on the exponential triggers of the default times. This is because typically spread correlation has a much lower impact on dependence of default times than default correlation. The correlation structure underlying the trivariate copula function may be estimated by the implied correlation in the quoted indices tranches markets such as i-Traxx and CDX or calculated from the asset correlation of the names. We assume the following stochastic intensity model Brigo and Alfonsi (2005), Brigo and El-Bachir (2008) for the three names

$$\lambda_j(t) = y_j(t) + \psi_j(t; \boldsymbol{\beta}_j), \; t \geq 0, j = 0, 1, 2 \qquad (6.3.8)$$

where ψ is a deterministic function, depending on the parameter vector $\boldsymbol{\beta}$ (which includes y_0), that is integrable on closed intervals. We assume each y_j to be a Cox Ingersoll Ross (CIR) process (see Brigo and Mercurio, 2006) given by

$$dy_j(t) = \kappa_j(\mu_j - y_j(t))dt + \nu_j \sqrt{y_j(t)} dZ_j(t) + J_{M_t,j} dM_j(t), j = 0, 1, 2 \qquad (6.3.9)$$

where Js are i.i.d. positive jump sizes that are exponentially distributed with mean χ_j, and M_j are Poisson processes with intensity m_j measuring the arrival of jumps in the intensity λ_j. The parameter vectors are $\boldsymbol{\beta}_j = (\kappa_j, \mu_j, \nu_j, y_j(0), \chi_j, m_j)$ with each vector component being a positive deterministic constant. We relax the condition of inaccessibility of the origin $2\kappa_j\mu_j > \nu_j^2$ so that we do not limit the CDS implied volatility generated by the model. We assume the Z_j to be standard Brownian motion processes under the risk-neutral measure. We define the following integrated quantities which will be extensively used in the remainder of the paper

$$\Lambda_j(t) = \int_0^t \lambda_j(s)ds, \quad Y_j(t) = \int_0^t y_j(s)ds, \quad \Psi_j(t;\boldsymbol{\beta}_j) = \int_0^t \psi_j(s;\boldsymbol{\beta}_j)ds \tag{6.3.10}$$

Here, we focus on intensities without jumps, i.e., $m_j = 0$. Brigo and El-Bachir (2008) consider in detail the tractable model with jumps, and this extension will be applied in future work.

6.3.4 Bilateral Risk Credit Valuation Adjustment for Receiver CDS

We next proceed to the valuation of the BR-CVA adjustment for the case of CDS payoff given by Eq. (6.3.1). We state the result as a proposition.

Proposition 6.3.1. *The* BR-CVA *at time t for a receiver CDS contract (protection seller) running from time T_a to time T_b with premium S is given by*

$$\begin{aligned}\text{BR-CVA-CDS}_{a,b}(t,S,\text{LGD}_{0,1,2}) \;=\;& \text{LGD}_2 \cdot \mathbb{E}_t\left\{\mathbf{1}_{C\cup D}\cdot D(t,\tau_2)\cdot\left[\mathbf{1}_{\tau_1>\tau_2}\overline{CDS}_{a,b}(\tau_2,S,LGD_1)\right]^+\right\}\\ & -\text{LGD}_0\cdot\mathbb{E}_t\left\{\mathbf{1}_{A\cup B}\cdot D(t,\tau_0)\cdot\left[-\mathbf{1}_{\tau_1>\tau_0}\overline{CDS}_{a,b}(\tau_0,S,LGD_1)\right]^+\right\}\end{aligned} \tag{6.3.11}$$

Proof. We have

$$\begin{aligned}\mathbb{E}_t\left\{\mathbf{1}_{C\cup D}\cdot D(t,\tau_2)\cdot[\text{NPV}(\tau_2)]^+\right\} &= \mathbb{E}_t\left\{\mathbf{1}_{C\cup D}\cdot D(t,\tau_2)\cdot[\text{CDS}_{a,b}(\tau_2,S,\text{LGD}_1)]^+\right\}\\ &= \mathbb{E}_t\left\{\mathbf{1}_{C\cup D}\cdot D(t,\tau_2)\cdot\left[\mathbf{1}_{\tau_1>\tau_2}\overline{\text{CDS}}_{a,b}(\tau_2,S,\text{LGD}_1)\right]^+\right\}\end{aligned} \tag{6.3.12}$$

where the first equality in Eq. (6.3.12) follows by definition, while the last equality follows from Eq. (6.3.3). Similarly, we have

$$\begin{aligned}\mathbb{E}_t\left\{\mathbf{1}_{A\cup B}\cdot D(t,\tau_0)\cdot[-\text{NPV}(\tau_0)]^+\right\} &= \mathbb{E}_t\left\{\mathbf{1}_{A\cup B}\cdot D(t,\tau_0)\cdot[-\text{CDS}_{a,b}(\tau_0,S,\text{LGD}_1)]^+\right\}\\ &= \mathbb{E}_t\left\{\mathbf{1}_{A\cup B}\cdot D(t,\tau_0)\cdot\left[-\mathbf{1}_{\tau_1>\tau_0}\overline{\text{CDS}}_{a,b}(\tau_0,S,\text{LGD}_1)\right]^+\right\}\end{aligned} \tag{6.3.13}$$

The proof follows using the expression of BR-CVA which is given by Eq. (6.2.5). □

To summarize, in order to compute the counterparty risk adjustment, we determine the value of the CDS contract on the reference credit "1", at the point in time τ_2 at which the counterparty "2" defaults. The reference name "1" has survived this point and there is a bivariate copula $C_{1,2}$ given by Eq. (6.3.7) which connects the default times of the reference credit and the counterparty "2". Similarly, in order to compute the investor risk adjustment, we determine the value of the CDS contract on the reference credit "1" at the point in time τ_0 at which the counterparty "0" defaults. The reference name "1" has survived this point and there is a bivariate copula $C_{0,1}$ given by the analogous of Eq. (6.3.7) which connects the default times of the reference credit and the investor "0". It is clear from Eq. (6.3.3) that the only terms we need to know in order to compute (6.3.12) and (6.3.13) are

$$\mathbf{1}_{\tau_1>\tau_2}\mathbb{Q}(\tau_1 > t|\mathcal{G}_{\tau_2}) \tag{6.3.14}$$

and

$$\mathbf{1}_{\tau_1>\tau_0}\mathbb{Q}(\tau_1 > t|\mathcal{G}_{\tau_0}) \tag{6.3.15}$$

In the next section, we generalize the numerical method proposed in Brigo and Chourdakis (2008) to calculate quantities (6.3.14) and (6.3.15) for the bilateral counterparty case.

6.4 Monte-Carlo Evaluation of the BR-CVA Adjustment

We propose a numerical method based on Monte-Carlo simulations to calculate the BR-CVA for the case of CDS contracts. Section 6.4.1 specifies the simulation method used to generate the sample paths of the CIR process. Section 6.4.2 gives a method to calculate (6.3.14), while Section 6.4.3 gives the complete numerical algorithm for calculating the BR-CVA in (6.2.4).

6.4.1 Simulation of CIR Process

We use the well-known fact (see Brigo and Mercurio, 2006) that the distribution of $y(t)$ given $y(u)$, for some $u < t$ is, up to a scale factor, a noncentral chi-square distribution. More precisely, the transition law of $y(t)$ given $y(u)$ can be expressed as

$$y(t) = \frac{\nu^2(1-e^{-\kappa(t-u)})}{4\kappa}\chi'_d\left(\frac{4\kappa e^{-\kappa(t-u)}}{\nu^2(1-e^{-\kappa(t-u)})}y(u)\right) \tag{6.4.1}$$

where

$$d = \frac{4\kappa\mu}{\nu^2} \tag{6.4.2}$$

and $\chi'_u(v)$ denotes a noncentral chi-square random variable with u degrees of freedom and non-centrality parameter v. In this way, if we know $y(0)$, we can simulate the process $y(t)$ exactly on a discrete time grid by sampling from the noncentral chi-square distribution.

6.4.2 Calculation of Survival Probability

We state the result in the form of a proposition. Such result will be used in Section 6.4.3 to develop a numerical algorithm for computing the BR-CVA.

Proposition 6.4.1.

$$\mathbf{1}_{\tau_1>\tau_2}\mathbb{Q}(\tau_1 > t|\mathcal{G}_{\tau_2}) = \mathbf{1}_{\bar{A}} + \mathbf{1}_{t<\tau_2}\mathbf{1}_{\tau_1>\tau_2}\int_{\overline{U}_1}^{1} F_{\Lambda_1(t)}(-\log(1-u_1))dC_{1|2}(u_1;U_2)$$

where

$$\begin{aligned}
\bar{A} &= \{t < \tau_2 < \tau_1\} \\
C_{1|2} &= \frac{f(u,v) - f(\overline{U}_1, v)}{1 - f(\overline{U}_1, v)} \\
\overline{U}_1 &:= 1 - \exp\{-Y_1(\tau_2) - \Psi_1(\tau_2; \boldsymbol{\beta}_1)\} \\
f(u,v) &= \Phi\left(\frac{\Phi^{-1}(u) - r_{1,2}\Phi^{-1}(v)}{\sqrt{1 - r_{1,2}^2}}\right)
\end{aligned} \tag{6.4.3}$$

and $F_{\Lambda_1(t)}$ denotes the cumulative distribution function of the cumulative (shifted) intensity of the CIR process which can be retrieved inverting the characteristic function of the integrated CIR process. The proof of the proposition is reported in Appendix K.

6.4.3 The Numerical BR-CVA Adjustment Algorithm

We give the pseudo code of the numerical algorithm used to calculate the BR-CVA for the CDS payer and receiver. In the pseudo code below, the variable α_i represents the time elapsing between payment period t_{i-1} and t_i measured in years, the variable Δ represents the fineness of the grid used to evaluate the integral of the survival probability in Eq. (6.4.3) and the variable δ represents the fineness of the time grid used to evaluate the integral in Eq. (6.3.3). The variable $x_{\max}$ represents the maximum x value for which the cumulative distribution function is implied from the characteristic function. The expression CDF(x_k) corresponds to a subroutine call which calculates the cumulative distribution function at x_k. This can be done by inversion of the characteristic function of the integrated CIR process using fractional Fourier transform methods such as in Chourdakis (2005). The inputs to the main procedure Calculate Adjustment are the number N of Monte Carlo runs and the market quote S_1 of the 5 year CDS spread of the reference entity.

6.5 Numerical Results

We consider an investor (name "0") trading a five-years CDS contract on a reference name (name "1") with a counterparty (name "2"). Both the investor and the counterparty are subject to default risk. We experiment on different levels of credit risk and credit risk volatility of the three names, which are specified by the parameters of the CIR processes in Table 6.1. We recall that the survival probabilities associated with a CIR intensity process are given by

$$\begin{aligned}
\mathbb{Q}(\tau_i > t) &:= \mathbb{E}[e^{-Y_i(t)}] \\
&= P^{CIR}(0, t, \boldsymbol{\beta}_i)
\end{aligned} \tag{6.5.1}$$

Algorithm 1 $[\overline{\text{BR-CVA\_R}}, \overline{\text{BR-CVA\_P}}]$ = CALCULATE ADJUSTMENT (N, S)

```
for i = 1 : N do
    Generate τ_0, τ_1, and τ_2 using Eq. (6.3.5) and Eq. (6.3.6).
    if τ_2 < τ_0 and τ_2 < T_b then
        if τ_1 > τ_2 then
            [BR-CVA_R_2, BR-CVA_P_2] = CDSAdjust(T_{γ(τ_2)}, S, LGD_1, 2)
            CUM_BR-CVA_R = CUM_BR-CVA_R + LGD_2· BR-CVA_R_2
            CUM_BR-CVA_P = CUM_BR-CVA_P + LGD_2· BR-CVA_P_2
        end if
    end if
    if τ_0 < τ_2 and τ_0 < T_b then
        if τ_1 > τ_0 then
            [BR-CVA_R_0, BR-CVA_P_0] = CDSAdjust(T_{γ(τ_0)}, S, LGD_1, 0)
            CUM_BR-CVA_R = CUM_BR-CVA_R - LGD_0· BR-CVA_R_0
            CUM_BR-CVA_P = CUM_BR-CVA_P - LGD_0· BR-CVA_P_0
        end if
    end if
end for
BR-CVA_R (overline) = CUM_BR-CVA_R / N
BR-CVA_P (overline) = CUM_BR-CVA_P / N
```

Algorithm 2 $[CDS_R, CDS_P]$ = CDSADJUST(T_j, S, LGD$_1$, $index$)

```
Term_1 = Term_2 = Term_3 = 0
t_start = max(T_a, T_j)
Q_prev = ComputeProb(t_start, T_j, index)
for t = t_start + δ : δ : T_b do
    Q_curr = ComputeProb(t, T_j, index)
    Term_1 = Term_1 + D(T_j, t − δ)(t − δ − T_{γ(t−δ)−1})(Q_curr − Q_prev)
    Term_3 = Term_3 + D(T_j, t − δ)(Q_curr − Q_prev)
    Q_prev = Q_curr
end for
for t_i = t_start + α_i : α_i : T_b do
    Q_curr = ComputeProb(t_i, T_j, index)
    Term_2 = Term_2 + α_i · D(T_j, t_i) · Q_curr
end for
CDS_val = S · (Term_2 − Term_1) + LGD_1 · Term_3
if index == 2 then
    CDS_R = D(t, T_j) · max(CDS_val, 0)
    CDS_P = D(t, T_j) · max(−CDS_val, 0)
end if
if index == 0 then
    CDS_R = D(t, T_j) · max(−CDS_val, 0)
    CDS_P = D(t, T_j) · max(CDS_val, 0)
end if
```

Credit Risk Levels	$y(0)$	κ	μ	Credit Risk volatilities	ν
low	0.0001	0.9	0.001	low	0.01
middle	0.01	0.80	0.02	middle	0.2
high	0.03	0.50	0.05	high	0.5

Table 6.1: The credit risk levels and credit risk volatilities parameterizing the CIR processes

where $P^{CIR}(0, t, \boldsymbol{\beta}_i)$ is the price at time 0 of a zero coupon bond maturing at time t under a stochastic interest rate dynamics given by the CIR process with $\boldsymbol{\beta}_i = (y_i(0), \kappa_i, \mu_i, \nu_i)$ being the vector of CIR parameters,

Algorithm 3 $Q_i = \textsc{ComputeProb}(t, T_j, index)$

$U_{index} = 1 - \exp\{-Y_{index}(T_j) - \Psi_{index}(T_j; \boldsymbol{\beta}_{index})\}$
$\bar{U}_1 = 1 - \exp\{-Y_1(T_j) - \Psi_1(T_j; \boldsymbol{\beta}_1)\}$
for $x_k = 0 : \Delta : x_{\max}$ **do**
  $p_k = \mathrm{CDF}(x_k)$
  $u_k = 1 - \exp\{-x_k - \psi_1(t)\}$
  Compute $f_k = C_{1|2}(u_k; U_{index})$ using Eq. (K.0.5) and Eq. (K.0.6)
end for
$Q_i = \sum_{(u_k, p_k, f_k): u_k > \bar{U}_1} p_k (f_{k+1} - f_k)$

Maturity	Low Risk	Middle Risk	High risk
1y	0	92	234
2y	0	104	244
3y	0	112	248
4y	1	117	250
5y	1	120	251
6y	1	122	252
7y	1	124	253
8y	1	125	253
9y	1	126	254
10y	1	127	254

Table 6.2: Break-even spreads in basis points generated using the parameters of the CIR processes in Table 6.1. The first column is generated using low credit risk and credit risk volatility. The second column is generated using middle credit risk and credit risk volatility. The third column is generated using high credit risk and credit risk volatility.

$i = 0, 1, 2$. Such price may be computed in close form (see Brigo and Mercurio, 2006) and is given by

$$P^{CIR}(0, t, \boldsymbol{\beta}_i) = A(t, T) e^{-B(t,T) r(t)} \tag{6.5.2}$$

where

$$\begin{aligned} A(t,T) &= \left[\frac{2 h_i e^{(h_i + \kappa_i)(T-t)/2}}{2h_i + (\kappa_i + h_i)\left(e^{h_i (T-t)/2} - 1\right)} \right]^{2\kappa_i \mu_i} \\ B(t,T) &= \frac{2\left(e^{h_i(T-t)} - 1\right)}{2h_i + (\kappa_i + h_i)\left(e^{h_i(T-t)/2} - 1\right)} \\ h_i &= \sqrt{\kappa_i^2 + 2\nu_i^2} \end{aligned} \tag{6.5.3}$$

and r_t denotes the risk-free interest rate at time t. We report in Table 6.2 the break-even spreads zeroing (6.3.3) in S, with survival probabilities given by Eq. (6.5.1) and CIR parameters $\boldsymbol{\beta}_{low}$, $\boldsymbol{\beta}_{middle}$, $\boldsymbol{\beta}_{high}$ obtained from Table 6.1.

The evaluation time t and the starting time T_a of the CDS contract are both set to zero. The end time T_b of the contract is set to five years. It is assumed that payments are exchanged every three months. The loss given defaults of the low, middle, and high risk entity are, respectively, set to $\mathrm{LGD}_{low} = 0.6$, $\mathrm{LGD}_{middle} = 0.65$, $\mathrm{LGD}_{high} = 0.7$. We assume that the spreads in Table 6.2 are the spreads quoted in the markets for the three names under consideration. We recover the integrated shift $\Psi(t;\boldsymbol{\beta})$ which makes the model survival probabilities consistent with the market survival probabilities coming from Table 6.2 whenever we change the CIR parameters. In mathematical terms, for any $i = 0, 1, 2$, we impose that

$$\begin{aligned} \mathbb{Q}(\tau_i > t)_{model} &:= \mathbb{E}[e^{-\Lambda_i(t)}] \\ &= \mathbb{Q}(\tau_i > t)_{market} \end{aligned} \tag{6.5.4}$$

The market survival probability for name "i" is bootstrapped from the market CDS quotes reported in Table 6.2. Such bootstrap procedure is performed assuming a piecewise linear hazard rate function. From the definition of the integrated process $\Psi_j(t;\boldsymbol{\beta}_j)$ given in Eq. (6.3.10), we can restate Eq. (6.5.4) as

$$\begin{aligned} \Psi_i(t;\boldsymbol{\beta}_i) &= \log\left(\frac{\mathbb{E}[e^{-Y_i(t)}]}{\mathbb{Q}(\tau_i > t)_{market}}\right) \\ &= \log\left(\frac{P^{CIR}(0,t,\boldsymbol{\beta}_i)}{\mathbb{Q}(\tau_i > t)_{market}}\right) \end{aligned} \tag{6.5.5}$$

where the last equality in Eq. (6.5.5) follows from Eq. (6.5.1). We first study a case where the quoted market spreads of the investor has a low credit risk profile (β_{low}), the reference entity has high credit risk profile (β_{high}), while the counterparty has middle credit risk (β_{middle}). We vary the correlation between reference credit and counterparty as well as the credit risk volatility ν_1 of the reference credit. Since the focus is mostly on credit spreads volatility, we calculate the implied CDS volatility produced by the choice of parameters $\boldsymbol{\beta}_1 = (y_0(1), \mu_1, \kappa_1, \nu_1)$ for a hypothetical CDS option, maturing in one year and in case of exercise entering into a four year CDS contract on the underlying reference credit "1". The objective of the experiments is to measure the impact of correlation and credit spreads volatility on the BR-CVA. The triple (x, y, z) represents the correlation of the trivariate Gaussian copula, with x denoting the correlation between the investor and reference credit, y denoting the correlation between the investor and the counterparty and z denoting the correlation between the reference credit and the counterparty. The values $\overline{\text{BR-CVA\_P}}$ and $\overline{\text{BR-CVA\_R}}$ are respectively the Monte Carlo estimates of the CDS payer and receiver counterparty risk adjustments. The theoretical formula for payer risk adjustment is given by

$$\begin{aligned} &\mathrm{LGD}_2 \cdot \mathbb{E}_{\mathrm{t}}\left\{\mathbf{1}_{\mathrm{C}\cup\mathrm{D}} \cdot \mathrm{D}(\mathrm{t},\tau_2) \cdot [-\mathrm{NPV}(\tau_2)]^+\right\} \\ &-\mathrm{LGD}_0 \cdot \mathbb{E}_{\mathrm{t}}\left\{\mathbf{1}_{\mathrm{A}\cup\mathrm{B}} \cdot \mathrm{D}(\mathrm{t},\tau_0) \cdot [\mathrm{NPV}(\tau_0)]^+\right\} \end{aligned} \tag{6.5.6}$$

which follows from adapting the formula in Eq. (6.2.4), given for the receiver counterparty adjustment, to the case of the payer counterparty adjustment. Tables 6.3 and 6.4 report the results obtained.

The scenarios considered in Table 6.3 and 6.4 assume an investor with an extremely low credit risk profile and thus are similar to those considered in Brigo and Chourdakis (2008), where the investor is assumed to be default-free. Therefore, it is not surprising that we find similar results. Similarly to them, we can see that as the correlation between counterparty and reference credit increases, the BR-CVA decreases, except for large correlation values where the BR-CVA goes down. This is the case because if the reference credit

(r_{01}, r_{02}, r_{12})	**Vol. parameter** ν_1 **CDS Implied vol**	**0.01** **1.5%**	**0.10** **15%**	**0.20** **28%**	**0.30** **37%**	**0.40** **42%**	**0.50** **42%**
(0, 0, -0.99)	$\overline{\text{BR-CVA\_P}}$	0.0(0.0)	0.0(0.0)	-0.0(0.0)	-0.0(0.0)	-0.0(0.0)	-0.0(0.0)
	$\overline{\text{BR-CVA\_R}}$	28.6(1.3)	28.8(1.3)	28.7(1.3)	28.8(1.3)	28.7(1.3)	28.7(1.3)
(0, 0, -0.90)	$\overline{\text{BR-CVA\_P}}$	0.0(0.0)	0.0(0.0)	-0.0(0.0)	-0.0(0.0)	0.0(0.0)	0.1(0.1)
	$\overline{\text{BR-CVA\_R}}$	28.7(1.3)	28.8(1.3)	28.7(1.3)	28.8(1.3)	28.8(1.3)	28.6(1.3)
(0, 0, -0.60)	$\overline{\text{BR-CVA\_P}}$	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.2(0.1)	0.2(0.1)	0.7(0.3)
	$\overline{\text{BR-CVA\_R}}$	26.5(1.3)	26.5(1.3)	26.0(1.3)	25.3(1.2)	24.8(1.2)	24.6(1.2)
(0, 0, -0.20)	$\overline{\text{BR-CVA\_P}}$	0.0(0.0)	0.4(0.1)	1.2(0.2)	2.4(0.4)	2.5(0.4)	2.2(0.4)
	$\overline{\text{BR-CVA\_R}}$	9.5(0.6)	9.8(0.6)	10.8(0.6)	11.4(0.7)	12.0(0.7)	12.2(0.7)
(0, 0, 0)	$\overline{\text{BR-CVA\_P}}$	5.7(0.3)	6.0(0.4)	6.2(0.5)	7.3(0.7)	7.1(0.8)	5.1(0.6)
	$\overline{\text{BR-CVA\_R}}$	-0.0(0.0)	0.4(0.1)	1.8(0.2)	3.3(0.2)	4.3(0.3)	5.0(0.3)
(0, 0, 0.20)	$\overline{\text{BR-CVA\_P}}$	26.0(1.4)	25.5(1.4)	22.6(1.4)	19.8(1.4)	17.0(1.4)	14.5(1.3)
	$\overline{\text{BR-CVA\_R}}$	-0.0(0.0)	0.0(0.0)	0.2(0.0)	0.4(0.0)	0.5(0.1)	0.6(0.1)
(0, 0, 0.60)	$\overline{\text{BR-CVA\_P}}$	76.3(4.6)	76.7(4.6)	66.6(4.1)	56.0(3.6)	48.9(3.2)	45.1(2.9)
	$\overline{\text{BR-CVA\_R}}$	-0.0(0.0)	-0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
(0, 0, 0.90)	$\overline{\text{BR-CVA\_P}}$	75.6(6.0)	76.5(6.0)	77.3(5.7)	65.9(5.0)	64.6(4.6)	62.6(4.4)
	$\overline{\text{BR-CVA\_R}}$	-0.0(0.0)	-0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
(0, 0, 0.99)	$\overline{\text{BR-CVA\_P}}$	0.0(0.0)	12.5(2.6)	40.6(4.3)	40.8(4.1)	54.9(4.5)	62.0(4.7)
	$\overline{\text{BR-CVA\_R}}$	-0.0(0.0)	-0.0(0.0)	0.0(0.0)	0.1(0.0)	0.1(0.0)	0.0(0.0)

Table 6.3: BR-CVA in basis points for the case when $\nu_2 = 0.01$ and $\nu_0 = 0.01$; numbers within round brackets represent the Monte-Carlo standard error. The CDS contract on the reference credit has a five-years maturity.

(r_{01}, r_{02}, r_{12})	**Vol. parameter** ν_1 **CDS Implied vol**	**0.01** **1.5%**	**0.10** **15%**	**0.20** **28%**	**0.30** **37%**	**0.40** **42%**	**0.50** **42%**
(0, 0, -0.99)	$\overline{\text{BR-CVA\_P}}$	0.0(0.0)	0.0(0.0)	-0.0(0.0)	-0.0(0.0)	-0.0(0.0)	0.1(0.1)
	$\overline{\text{BR-CVA\_R}}$	29.0(1.3)	28.7(1.3)	28.1(1.3)	28.8(1.3)	28.3(1.3)	29.1(1.3)
(0, 0, -0.90)	$\overline{\text{BR-CVA\_P}}$	0.0(0.0)	0.0(0.0)	-0.0(0.0)	-0.0(0.0)	0.1(0.1)	0.0(0.0)
	$\overline{\text{BR-CVA\_R}}$	29.8(1.4)	28.0(1.3)	27.7(1.3)	28.7(1.3)	27.6(1.3)	28.4(1.3)
(0, 0, -0.60)	$\overline{\text{BR-CVA\_P}}$	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.3(0.1)	0.4(0.2)	0.5(0.2)
	$\overline{\text{BR-CVA\_R}}$	27.1(1.3)	24.8(1.2)	25.8(1.2)	24.6(1.2)	23.7(1.2)	24.1(1.2)
(0, 0, -0.20)	$\overline{\text{BR-CVA\_P}}$	0.0(0.0)	0.5(0.1)	1.7(0.3)	2.0(0.3)	3.1(0.5)	2.7(0.4)
	$\overline{\text{BR-CVA\_R}}$	9.3(0.6)	9.5(0.6)	10.9(0.6)	11.1(0.6)	11.3(0.7)	12.0(0.7)
(0, 0, 0)	$\overline{\text{BR-CVA\_P}}$	6.3(0.3)	5.3(0.4)	6.7(0.6)	6.7(0.7)	6.6(0.8)	6.0(0.7)
	$\overline{\text{BR-CVA\_R}}$	-0.0(0.0)	0.3(0.0)	2.2(0.2)	3.4(0.2)	4.1(0.3)	4.9(0.3)
(0, 0, 0.20)	$\overline{\text{BR-CVA\_P}}$	25.8(1.4)	23.5(1.4)	21.3(1.3)	19.2(1.4)	16.8(1.4)	14.1(1.2)
	$\overline{\text{BR-CVA\_R}}$	-0.0(0.0)	-0.0(0.0)	0.3(0.0)	0.5(0.1)	0.7(0.1)	0.7(0.1)
(0, 0, 0.60)	$\overline{\text{BR-CVA\_P}}$	71.9(4.4)	70.2(4.4)	65.2(4.1)	55.1(3.6)	46.4(3.1)	42.4(2.9)
	$\overline{\text{BR-CVA\_R}}$	-0.0(0.0)	-0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)	0.0(0.0)
(0, 0, 0.90)	$\overline{\text{BR-CVA\_P}}$	77.3(5.9)	82.6(6.2)	77.9(5.8)	66.4(5.0)	63.6(4.5)	61.9(4.3)
	$\overline{\text{BR-CVA\_R}}$	-0.0(0.0)	-0.0(0.0)	0.0(0.0)	0.1(0.0)	0.0(0.0)	0.0(0.0)
(0, 0, 0.99)	$\overline{\text{BR-CVA\_P}}$	43.0(4.8)	49.8(5.1)	56.1(5.2)	50.2(4.5)	57.7(4.5)	62.3(4.7)
	$\overline{\text{BR-CVA\_R}}$	0.0(0.0)	0.1(0.0)	0.2(0.1)	0.3(0.1)	0.3(0.1)	0.3(0.1)

Table 6.4: BR-CVA in basis points for the case when $\nu_2 = 0.2$ and $\nu_0 = 0.01$; numbers within round brackets represent the Monte-Carlo standard error. The CDS contract on the reference credit has a five-years maturity.

(r_{01}, r_{02}, r_{12})	**Base Scenario**	**Risky Counterparty**	**Risky Investor**	**Risky Ref**	**Safe Ref**
(0, 0, 0)	6.2(0.4)	3.6(0.2)	-0.8(0.1)	-0.0(0.0)	5.5(0.4)
(0, 0, 0.1)	15.1(0.9)	11.9(0.5)	-0.7(0.1)	0.0(0.0)	13.8(0.8)
(0, 0, 0.3)	36.6(2.1)	35.1(1.4)	-0.5(0.2)	0.2(0.0)	35.0(2.0)
(0, 0, 0.6)	72.7(4.4)	88.2(3.8)	-0.8(0.0)	0.7(0.1)	69.7(4.4)
(0, 0, 0.9)	75.2(5.9)	194.2(8.2)	-0.8(0.0)	1.4(0.4)	71.0(5.8)
(0, 0, 0.99)	25.3(3.7)	281.2(11.5)	-0.8(0.0)	1.9(0.7)	19.6(3.3)
(0, 0.1, 0)	5.9(0.4)	3.5(0.2)	-0.8(0.1)	-0.0(0.0)	5.0(0.4)
(0, 0.3, 0)	6.2(0.4)	3.5(0.2)	-0.8(0.0)	-0.0(0.0)	5.1(0.4)
(0, 0.6, 0)	5.9(0.4)	3.2(0.2)	-0.8(0.0)	-0.0(0.0)	4.4(0.3)
(0, 0.9, 0)	5.2(0.3)	3.4(0.2)	-0.8(0.0)	-0.0(0.0)	4.0(0.3)
(0, 0.99, 0)	5.1(0.3)	3.2(0.2)	-0.8(0.0)	-0.0(0.0)	3.5(0.3)
(0.1, 0, 0)	6.1(0.4)	3.6(0.2)	-0.1(0.0)	0.0(0.0)	5.8(0.4)
(0.3, 0, 0)	6.2(0.4)	3.6(0.2)	-0.0(0.0)	0.0(0.0)	5.8(0.4)
(0.6, 0, 0)	6.0(0.4)	3.6(0.2)	0.0(0.0)	0.0(0.0)	5.8(0.4)
(0.9, 0, 0)	6.0(0.4)	3.6(0.2)	-0.0(0.0)	-0.1(0.0)	6.0(0.4)
(0.99, 0, 0)	6.2(0.4)	3.6(0.2)	-1.7(0.1)	-0.1(0.0)	6.1(0.4)

Table 6.5: BR-CVA under five different riskiness scenarios. The CIR volatilities are set to $\nu_0 = \nu_1 = \nu_2 = 0.1$. The correlation triple has only one nonzero entry

and the counterparty are 99% correlated, then the exponential triggers ξ_1 and ξ_2 are almost identical; if we consider the scenario where $\nu_1 = \nu_2 = 0.01$, then the intensity of both processes are almost deterministic, with $\lambda_1 > \lambda_2$ having the name "1" high credit risk and name "2" middle credit risk. Thus, to a first approximation $\tau_1 = \frac{\xi_1}{\lambda_1}$ is higher than $\tau_2 = \frac{\xi_2}{\lambda_2}$, and consequently the reference credit always defaults before the counterparty, resulting in no adjustment. Table 6.4 has qualitatively a similar behavior although the larger value of the counterparty volatility $\nu_2 = 0.2$ tends to smooth out the negligible BR-CVA obtained in Table 6.3 for small values of reference credit volatility and high default correlation between counterparty and reference credit.

Tables 6.5, 6.6 and 6.7 report the BR-CVA adjustments for the Payer CDS under a set of five different riskiness scenarios. The assignments of credit risks in Table 6.1 to the three names determine the scenario in place. We have:

- Scenario 1 (*Base Scenario*). The investor has low credit risk, the reference entity has high credit risk, the counterparty has middle credit risk
- Scenario 2 (*Risky Counterparty*). The investor has low credit risk, the reference entity has middle credit risk and the counterparty has high credit risk.
- Scenario 3 (*Risky Investor*). The counterparty has low credit risk, the reference entity has middle credit risk, while the investor has high credit risk.
- Scenario 4 (*Risky Reference entity*). Both investor and counteparty have middle credit risk, while the reference entity has high credit risk.
- Scenario 5 (*Safe Reference entity*). Both investor and counterparty have high credit risk, while the reference entity has low credit risk.

(r_{01}, r_{02}, r_{12})	Base Scenario	Risky Counterparty	Risky Investor	Risky Ref	Safe Ref
(0, 0.5, 0.5)	54.8(3.5)	62.7(2.8)	-0.6(0.2)	0.5(0.0)	47.8(3.3)
(0, 0.2, 0.9)	72.9(5.8)	183.0(8.0)	-0.8(0.0)	1.3(0.4)	64.5(5.6)
(0, 0.9, 0.2)	21.0(1.3)	20.3(0.8)	-0.8(0.0)	0.1(0.0)	15.2(1.2)
(0.5, 0, 0.5)	64.3(3.7)	67.9(2.9)	-0.0(0.0)	0.6(0.0)	61.7(3.7)
(0.2, 0, 0.9)	81.8(6.2)	193.4(8.2)	-0.1(0.0)	1.5(0.4)	79.8(6.1)
(0.9, 0, 0.2)	27.0(1.5)	23.4(0.9)	0.0(0.1)	0.1(0.0)	26.3(1.4)
(0.5, 0.5, 0)	6.6(0.4)	3.4(0.2)	-0.0(0.0)	0.0(0.0)	6.2(0.4)
(0.2, 0.9, 0)	5.5(0.3)	3.5(0.2)	-0.1(0.0)	0.0(0.0)	4.3(0.3)
(0.9, 0.2, 0)	7.0(0.4)	3.5(0.2)	-0.0(0.0)	-0.0(0.0)	6.1(0.4)

Table 6.6: BR-CVA under five different riskiness scenarios. The CIR volatilities are set to $\nu_0 = \nu_1 = \nu_2 = 0.1$. The correlation triple has two nonzero entries

(r_{01}, r_{02}, r_{12})	Base Scenario	Risky Counterparty	Risky Investor	Risky Ref	Safe Ref
(0.8, 0.5, 0.2)	25.9(1.4)	23.3(0.9)	0.2(0.1)	0.1(0.0)	25.1(1.4)
(0.8, 0.2, 0.5)	60.9(3.6)	65.9(2.9)	0.3(0.2)	0.5(0.0)	60.1(3.6)
(0.5, 0.8, 0.2)	24.1(1.4)	22.5(0.9)	0.0(0.1)	0.1(0.0)	19.9(1.3)
(0.5, 0.2, 0.8)	89.0(6.0)	142.3(6.2)	-0.0(0.0)	1.1(0.2)	84.0(5.8)
(0.2, 0.5, 0.8)	75.1(5.5)	133.1(6.0)	-0.1(0.0)	0.9(0.2)	62.2(5.1)
(0.2, 0.8, 0.5)	51.7(3.4)	59.3(2.7)	-0.1(0.0)	0.4(0.0)	38.5(3.0)
(0.2, 0.2, 0.2)	25.2(1.4)	22.5(0.9)	0.1(0.1)	0.1(0.0)	22.8(1.3)
(0.5, 0.5, 0.5)	58.2(3.6)	64.4(2.8)	0.4(0.3)	0.5(0.0)	53.7(3.5)
(0.8, 0.8, 0.8)	79.5(5.7)	132.0(6.0)	-0.0(0.0)	0.8(0.2)	69.8(5.3)

Table 6.7: BR-CVA under five different riskiness scenarios. The CIR volatilities are set to $\nu_0 = \nu_1 = \nu_2 = 0.1$. The correlation triple has all nonzero entries

Tables 6.5, 6.6, and 6.7 show the important role played by default correlation. For example, if one looks at the second column of those tables, one notices that as the correlation between counterparty and reference credit gets larger, the BR-CVA increases due to the copula contagion coming from the default of the reference credit and to the large riskiness level of the counterparty. If the only correlation is between reference credit and counterparty, then the adjustment is the largest, while when correlation starts manifesting also between all other pairs (see second column of Tables 6.6 and 6.7), then the adjustment tends to decrease, since the default contagion induced by investor tends to reduce the number of times when the risky counterparty is the earliest to default.

An interesting pattern emerges from the first two columns and five rows of Table 6.5. Although the counterparty is riskier in Scenario 2, we have that the copula contagion effect starts kicking in as the correlation becomes larger than 60% and makes the BR-CVA adjustment in Scenario 2 significantly larger than the corresponding adjustment in Scenario 1. Unlikely from Tables 6.3 and 6.4, we see that low credit risk volatility values have the opposite effect and amplify the adjustment; this is because for 99% correlation between reference credit and counterparty, we now have that the counterparty always defaulting earlier is riskier.

Consistently with the results in Table 6.3 and 6.4, no BR-CVA takes place in scenario *Risky Reference Entity* due to the fact that the reference entity has the highest credit risk profile and being the credit risk volatilities of the three names relatively low (0.1), the reference entity always tends to be the first to default, thus resulting in no adjustment taking place.

6.6 Application to a Market Scenario

We apply the methodology to calculate the mark-to-market price of a five-year CDS contract between British Airways (counterparty) and Lehman Brothers (investor) on the default of Royal Dutch Shell (reference credit). We consider two CDS contracts. In the first contract Lehman Brothers buys 5-year protection on Shell from British Airways on January 5, 2006. In the second contract, Lehman Brothers sells 5-year protection on Shell to British Airways on January 5, 2006. In both contracts, British Airways computes the mark-to-market value of the contract on May 1, 2008. We consider different correlation scenarios among the three names. The CDS quotes of the three names on those dates are reported in Tables 6.8 and 6.9.

Maturity	Royal Dutch Shell	Lehman Brothers	British Airways
1y	4	6.8	10
2y	5.8	10.2	23.2
3y	7.8	14.4	50.6
4y	10.1	18.7	80.2
5y	11.7	23.2	110
6y	15.8	27.3.3	129.5
7y	19.4	30.5	142.8
8y	20.5	33.7	153.6
9y	21	36.5	162.1
10y	21.4	38.6	168.8

Table 6.8: Market spread quotes in basis points for Royal Dutch Shell, Lehman Brothers, and British Airways on January 5, 2006

Maturity	Royal Dutch Shell	Lehman Brothers	British Airways
1y	24	203	151
2y	24.6	188.5	230
3y	26.4	166.75	275
4y	28.5	152.25	305
5y	30	145	335
6y	32.1	136.3	342
7y	33.6	130	347
8y	35.1	125.8	350.6
9y	36.3	122.6	353.3
10y	37.2	120	355.5

Table 6.9: Market spread quotes in basis points for Royal Dutch Shell, Lehman Brothers and British Airways on May 1, 2008

The procedure used for the mark-to-market value valuation is detailed next:

(a) We take the CDS quotes of British Airways, Lehman Brothers, and Royal Dutch Shell on January 5, 2006, and calibrate the parameters of the CIR processes associated to the three names assuming zero shift and inaccessibility of the origin. The results obtained from the calibration are reported in Table 6.10.

(b) We calculate the value of the five year risk-adjusted CDS contract starting at T_a = January 5, 2006,

Credit Risk Levels	$y(0)$	κ	μ	Credit Risk volatilities	ν
Lehman Brothers (name "0")	0.0001	0.036	0.0432	low	0.0553
Royal Dutch Shell (name "1")	0.0001	0.0394	0.0219	middle	0.0192
British Airways (name "2")	0.00002	0.0266	0.2582	high	0.0003

Table 6.10: The CIR parameters of Lehman Brothers, Royal Dutch Shell, and British Airways calibrated to the market quotes of CDS on January 5, 2006

and ending five years later at T_b = January 5, 2011, as

$$CDS^D_{a,b}(T_a, S_1, \mathrm{LGD}_{0,1,2}) = CDS_{a,b}(T_a, S_1, \mathrm{LGD}_1) - \text{BR-CVA-CDS}_{a,b}(T_a, S_1, \mathrm{LGD}_{0,1,2}) \tag{6.6.1}$$

where $S_1 = 120$ bps is the five-year spread quote of Royal Dutch Shell at time T_a, $\mathrm{CDS}_{a,b}(T_a, S_1, \mathrm{LGD}_1)$ is the value of the equivalent CDS contract which does not account for counterparty risk given by Eq. (6.3.1), and the loss given default of the three names are taken from a market provider and equal to 0.6.

(c) Let T_c = May 1, 2008, be the time at which British Airways calculates the mark to market value of the CDS contract. We keep the CIR parameters of British Airways, and Royal Dutch Shell at the same values calibrated in (a). We vary the volatility of the CIR process associated to Lehman Brothers, while keeping the other parameters fixed. We take the market CDS quotes of Lehman Brothers, British Airways and Royal Dutch Shell and recompute the shift process as

$$\begin{aligned}\Psi_i(t;\boldsymbol{\beta}_i) &= \log\left(\frac{\mathbb{E}[e^{-Y_i(t)}\ ]}{\mathbb{Q}(\tau_i > t)_{market}}\right)\\ &= \log\left(\frac{P^{CIR}(0,t,\boldsymbol{\beta}_i)}{\mathbb{Q}(\tau_i > t)_{market}}\right)\end{aligned} \tag{6.6.2}$$

for any $T_c < t < T_d$, where T_d = May 1, 2013. Here the market survival probabilities $\mathbb{Q}(\tau_i > t)_{market}$, are stripped from the CDS quotes of the three names at the date, May, 1, 2008, using a piece-wise linear hazard rate function. We compute the value $\mathrm{CDS}^D_{c,d}(T_c, S_1, \mathrm{LGD}_1)$ of a risk- adjusted CDS contract starting at T_c and maturing at T_d, where the five year running spread premium, as well as the loss given defaults of the three parties, are the same as in T_a. We have

$$\mathrm{CDS}^D_{c,d}(T_c, S_1, \mathrm{LGD}_{0,1,2}) = \mathrm{CDS}_{c,d}(T_c, S_1, \mathrm{LGD}_1) - \text{BR-CVA-CDS}_{c,d}(T_c, S_1, \mathrm{LGD}_{0,1,2}) \tag{6.6.3}$$

where $\mathrm{CDS}_{c,d}(T_c, S_1, \mathrm{LGD}_1)$ is the value of the equivalent CDS contract which does not account for counterparty risk, and $\text{BR-CVA-CDS}_{c,d}(T_c, S_1, \mathrm{LGD}_1)$ is the adjustment for the period $[T_c, T_d]$ calculated at time T_c.

(d) We calculate the mark-to-market value of the CDS contract as follows:

$$MTM_{a,c}(S_1, \mathrm{LGD}_{0,1,2}) = CDS^D_{c,d}(T_c, S_1, \mathrm{LGD}_{0,1,2}) - \frac{CDS^D_{a,b}(T_a, S_1, \mathrm{LGD}_{0,1,2})}{D(T_a, T_c)} \tag{6.6.4}$$

Table 6.11 reports the MTM value of the CDS contract between British Airways and Lehman Brothers on default of Royal Dutch Shell under a number of correlation scenarios. The CDS contract agreed on January 5, 2006, is marked to market on May, 1, 2008 by British Airways using the four-step procedure described above. We check the effect of the increasing riskiness of Lehman Brothers by varying the volatility of the CIR process associated to Lehman. Table 6.11 reveals large sensitivity of the BR-CVA to default correlation

(r_{01}, r_{02}, r_{12})	**Vol ν_{Leh}** **CDS Implied vol**	**0.01** **1.5%**	**0.10** **15%**	**0.20** **28%**	**0.30** **37%**	**0.40** **42%**	**0.50** **42%**
(-0.3, -0.1, 0.7)	(LEH Pay, BAB Rec)	-56.0(5.6)	-55.7(5.7)	-55.6(5.6)	-55.6(5.6)	-55.8(5.6)	-55.6(5.7)
	(BAB Pay, LEH Rec)	84.2(0.0)	84.2(0.0)	84.2(0.0)	84.2(0.0)	84.2(0.0)	84.2(0.0)
(-0.1, -0.3, 0.8)	(LEH Pay, BAB Rec)	-41.6(9.1)	-41.7(9.1)	-41.7(9.1)	-41.6(9.1)	-41.7(9.1)	-41.6(9.1)
	(BAB Pay, LEH Rec)	83.8(0.1)	83.8(0.1)	83.8(0.1)	83.7(0.1)	83.8(0.1)	83.8(0.1)
(0, 0, -0.5)	(LEH Pay, BAB Rec)	-84.2(0.0)	-84.2(0.0)	-84.2(0.0)	-84.2(0.0)	-84.2(0.0)	-84.2(0.0)
	(BAB Pay, LEH Rec)	86.3(0.5)	86.3(0.5)	86.2(0.5)	86.2(0.5)	86.2(0.5)	86.2(0.5)
(0, 0, 0.6)	(LEH Pay, BAB Rec)	-62.2(3.9)	-62.4(3.9)	-62.4(3.9)	-62.3(3.9)	-62.6(3.9)	-62.3(3.9)
	(BAB Pay, LEH Rec)	82.8(0.2)	82.8(0.2)	82.7(0.2)	82.7(0.2)	82.7(0.2)	82.7(0.2)
(0, 0, 0.9)	(LEH Pay, BAB Rec)	-32.4(10.9)	-31.7(10.9)	-31.4(10.9)	-31.6(10.9)	-32.8(10.9)	-32.5(10.9)
	(BAB Pay, LEH Rec)	82.9(0.2)	82.8(0.2)	82.8(0.2)	82.8(0.2)	82.8(0.2)	82.8(0.2)
(0, 0.7, 0)	(LEH Pay, BAB Rec)	-81.9(0.3)	-82.0(0.3)	-82.0(0.3)	-82.0(0.3)	-82.0(0.3)	-82.0(0.3)
	(BAB Pay, LEH Rec)	83.1(0.2)	83.1(0.2)	83.0(0.2)	83.0(0.2)	82.9(0.2)	83.0(0.2)
(0, 0.7, 0.7)	(LEH Pay, BAB Rec)	-53.9(6.8)	-54.0(6.8)	-53.4(6.8)	-53.7(6.8)	-53.6(6.8)	-54.1(6.8)
	(BAB Pay, LEH Rec)	83.0(0.2)	83.0(0.2)	83.0(0.2)	82.9(0.2)	83.0(0.2)	83.0(0.2)
(0.1, 0.3, 0.8)	(LEH Pay, BAB Rec)	-54.6(6.4)	-53.6(6.6)	-54.5(6.4)	-53.6(6.6)	-53.8(6.6)	-53.7(6.6)
	(BAB Pay, LEH Rec)	81.5(0.5)	81.5(0.5)	81.4(0.5)	81.4(0.5)	81.5(0.5)	81.4(0.5)
(0.25, 0.25, 0.9)	(LEH Pay, BAB Rec)	-39.5(10.8)	-39.0(10.9)	-38.6(10.9)	-38.8(10.9)	-40.3(10.8)	-39.7(10.8)
	(BAB Pay, LEH Rec)	79.0(1.0)	79.0(1.0)	78.9(1.0)	78.9(1.0)	78.9(1.0)	78.9(1.0)
(0.3, 0.1, 0.7)	(LEH Pay, BAB Rec)	-58.7(4.6)	-59.1(4.6)	-58.9(4.6)	-58.9(4.6)	-59.3(4.6)	-58.8(4.6)
	(BAB Pay, LEH Rec)	77.9(1.2)	77.9(1.2)	77.8(1.2)	77.9(1.2)	77.9(1.2)	77.8(1.2)
(0.5, 0.5, 0.5)	(LEH Pay, BAB Rec)	-68.7(2.7)	-68.8(2.7)	-68.7(2.7)	-68.8(2.7)	-69.0(2.7)	-68.9(2.7)
	(BAB Pay, LEH Rec)	74.0(2.3)	74.0(2.3)	73.9(2.3)	74.1(2.3)	73.9(2.3)	73.9(2.3)
(0.7, 0, 0)	(LEH Pay, BAB Rec)	-81.3(0.3)	-81.3(0.3)	-81.3(0.3)	-81.4(0.3)	-81.2(0.3)	-81.3(0.3)
	(BAB Pay, LEH Rec)	63.9(4.4)	64.0(4.4)	63.8(4.4)	64.2(4.4)	64.0(4.4)	63.9(4.4)
(0.7, 0, 0.7)	(LEH Pay, BAB Rec)	-49.6(6.5)	-49.6(6.5)	-49.6(6.5)	-49.8(6.5)	-50.0(6.5)	-49.8(6.5)
	(BAB Pay, LEH Rec)	63.4(4.4)	63.5(4.4)	63.3(4.4)	63.8(4.4)	63.4(4.4)	63.4(4.4)
(0.9, 0.9, 0.9)	(LEH Pay, BAB Rec)	-44.6(9.1)	-44.8(9.0)	-44.5(9.1)	-44.7(9.1)	-44.4(9.1)	-44.7(9.1)
	(BAB Pay, LEH Rec)	53.5(7.1)	53.4(7.1)	53.6(7.1)	53.8(7.1)	54.1(7.0)	53.8(7.1)

Table 6.11: Value of the CDS contract between British Airways and Lehman Brothers on default of Royal Dutch Shell agreed on January 5, 2006, and marked to market by British Airways on May 1, 2008. The pairs (LEH Pay, BAB Rec) and (BAB Pay, LEH Rec) denote respectively the mark-to-market value when British Airways is the CDS receiver and CDS payer. The mark-to-market value of the CDS contract without risk adjustment when British Airways is respectively payer (receiver) is 84.2(-84.2) bps, due to the widening of the CDS spread curve of Royal Dutch Shell.

and very small sensitivity to credit risk volatility of Lehman.

When Lehman is negatively correlated or uncorrelated with Royal Dutch Shell and British Airways, which are instead positively correlated, then the likelihood of Lehman being the first to default tends to decrease. Moreover, the value of the CDS contract at the default time is very likely to be negative because of the copula contagion effect of Royal Dutch Shell on British Airways and of the increased five-year spread quotes of Shell. Therefore, when Lehman is the CDS receiver, the above considerations and Eq. (6.2.4) show that the Lehman BR-CVA becomes negligible and consequently the mark-to-market value of the contract equals its non-risk-adjusted counterpart. Conversely, when Lehman is the payer, then the default of British Airways is associated with a positive CDS contract value for Lehman at the default time. Eq. (6.2.4) thus implies that the value of the CDS contract for the investor Lehman gets reduced and consequently the mark-to-market value of the CDS contract for British Airways gets increased.

We next invert the role of Lehman Brothers and Royal Dutch Shell, hence Royal Dutch Shell becomes the investor and Lehman Brothers the reference entity. This is a natural case to examine in that the CVA entails an option on the underlying asset, i.e., an option on the Lehman CDS now. Since a CDS option depends strongly on the spread volatility, we are now able to examine this aspect directly.

We consider the following two CDS contracts. In the first contract Royal Dutch Shell buys 5 year protection on Lehman from British Airways on January 5, 2006. In the second contract, Royal Dutch Shell sells 5 year protection on Lehman to British Airways on January 5, 2006. As in the earlier case, British Airways computes the mark-to-market value of the two CDS contracts on May 1, 2008.

We consider again the effect of Lehman's volatility on the mark-to-market value of the CDS contract for British Airways on May 1, 2008, under the same correlation scenarios as above. The results are reported in Table 6.12.

(r_{01}, r_{02}, r_{12})	Vol ν_{Leh} **CDS Implied vol**	**0.01** **1.5%**	**0.10** **15%**	**0.20** **28%**	**0.30** **37%**	**0.40** **42%**	**0.50** **42%**
(-0.3, -0.1, 0.7)	(RDSPLC Pay, BAB Rec)	-448.5(11.9)	-433.6(13.2)	-422.6(14.1)	-403.7(15.7)	-387.1(16.6)	-364.1(17.8)
	(BAB Pay, RDSPLC Rec)	527.0(0.7)	526.3(1.0)	525.2(1.2)	524.9(1.3)	523.5(1.6)	522.0(2.1)
(-0.1, -0.3, 0.8)	(RDSPLC Pay, BAB Rec)	-440.3(12.6)	-427.1(13.9)	-417.3(14.7)	-393.2(16.7)	-381.2(17.2)	-352.7(18.9)
	(BAB Pay, RDSPLC Rec)	528.6(0.2)	528.3(0.3)	527.6(0.5)	527.4(0.5)	526.5(0.7)	525.4(1.0)
(0, 0, -0.5)	(RDSPLC Pay, BAB Rec)	-502.0(3.4)	-495.7(4.1)	-486.7(4.8)	-473.2(5.9)	-460.7(6.8)	-440.7(8.5)
	(BAB Pay, RDSPLC Rec)	526.2(1.2)	525.4(1.7)	525.2(1.7)	524.5(1.9)	523.5(2.3)	522.6(2.6)
(0, 0, 0.6)	(RDSPLC Pay, BAB Rec)	-529.2(0.0)	-529.0(0.1)	-528.1(0.4)	-527.0(1.0)	-524.3(0.8)	-513.6(2.0)
	(BAB Pay, RDSPLC Rec)	534.2(0.5)	533.2(0.6)	531.8(0.7)	530.2(0.6)	528.1(0.8)	526.8(1.2)
(0, 0, 0.9)	(RDSPLC Pay, BAB Rec)	-523.6(0.5)	-519.8(1.0)	-514.9(1.4)	-510.1(1.6)	-500.4(2.5)	-489.7(3.1)
	(BAB Pay, RDSPLC Rec)	529.0(0.1)	528.8(0.1)	528.7(0.2)	528.4(0.3)	527.9(0.5)	527.4(0.7)
(0, 0.7, 0)	(RDSPLC Pay, BAB Rec)	-485.6(6.2)	-476.5(7.0)	-466.1(8.0)	-450.8(9.5)	-434.9(10.7)	-417.7(11.7)
	(BAB Pay, RDSPLC Rec)	528.9(0.1)	528.6(0.3)	528.4(0.3)	528.2(0.4)	527.7(0.6)	527.1(0.8)
(0, 0.7, 0.7)	(RDSPLC Pay, BAB Rec)	-523.6(0.5)	-520.1(0.8)	-514.6(1.4)	-508.7(1.8)	-499.9(2.3)	-488.5(3.0)
	(BAB Pay, RDSPLC Rec)	521.3(3.0)	519.7(3.7)	518.8(3.8)	518.1(4.1)	516.6(4.6)	515.3(5.0)
(0.1, 0.3, 0.8)	(RDSPLC Pay, BAB Rec)	-459.7(9.2)	-447.5(10.3)	-436.1(11.5)	-415.1(13.2)	-399.3(14.1)	-374.0(15.8)
	(BAB Pay, RDSPLC Rec)	529.1(0.0)	528.8(0.1)	528.3(0.3)	528.1(0.3)	527.4(0.4)	526.2(0.8)
(0.25, 0.25, 0.9)	(RDSPLC Pay, BAB Rec)	-473.5(8.1)	-463.7(9.2)	-453.3(10.0)	-436.6(11.6)	-421.2(12.6)	-398.2(14.1)
	(BAB Pay, RDSPLC Rec)	528.1(0.3)	527.6(0.5)	526.9(0.7)	526.6(0.7)	525.6(1.0)	524.4(1.4)
(0.3, 0.1, 0.7)	(RDSPLC Pay, BAB Rec)	-481.6(6.3)	-472.6(7.2)	-461.5(8.2)	-443.4(10.2)	-430.2(10.6)	-408.2(12.2)
	(BAB Pay, RDSPLC Rec)	529.2(0.0)	529.2(0.0)	529.1(0.1)	529.1(0.0)	528.8(0.1)	527.8(0.4)
(0.5, 0.5, 0.5)	(RDSPLC Pay, BAB Rec)	-481.3(6.4)	-471.3(7.3)	-459.7(8.3)	-445.3(9.7)	-429.6(10.7)	-407.2(12.2)
	(BAB Pay, RDSPLC Rec)	526.5(0.8)	525.7(1.2)	524.5(1.5)	524.1(1.5)	522.6(1.9)	521.0(2.3)
(0.7, 0, 0)	(RDSPLC Pay, BAB Rec)	-484.2(6.1)	-474.1(6.9)	-463.6(8.1)	-449.9(9.1)	-429.0(11.3)	-415.8(11.6)
	(BAB Pay, RDSPLC Rec)	520.8(3.0)	519.0(3.8)	518.0(3.9)	517.0(4.2)	515.2(4.8)	513.6(5.2)
(0.7, 0, 0.7)	(RDSPLC Pay, BAB Rec)	-434.9(12.5)	-418.9(14.0)	-403.7(15.1)	-384.4(16.5)	-362.5(17.7)	-338.6(19.3)
	(BAB Pay, RDSPLC Rec)	528.0(1.2)	528.0(1.2)	528.0(1.2)	528.0(1.2)	528.0(1.2)	528.0(1.2)
(0.9, 0.9, 0.9)	(RDSPLC Pay, BAB Rec)	-495.6(4.0)	-487.8(4.7)	-479.0(5.4)	-458.5(7.7)	-448.5(7.6)	-430.4(8.9)
	(BAB Pay, RDSPLC Rec)	529.2(0.0)	529.2(0.0)	529.1(0.0)	529.2(0.0)	529.2(0.0)	528.0(0.4)

Table 6.12: Value of the CDS contract between British Airways and Royal Dutch Shell on default of Lehman Brothers agreed on January 5, 2006, and marked to market by British Airways on May 1, 2008. The pairs (RDSPLC Pay, BAB Rec) and (BAB Pay, RDSPLC Rec) denote respectively the mark-to-market value when British Airways is the CDS receiver and CDS payer. The mark-to-market value of the CDS contract without risk adjustment when British Airways is respectively payer (receiver) is 529(-529) bps, due to the widening of the CDS spread curve of Lehman Brothers.

Unlike the previous case, we observe significant sensitivity of the BR-CVA to both correlation and credit spread volatility of the reference entity Lehman when BA is the CDS receiver. The widening of the CDS curve of Lehman makes the value of the CDS contract negative for the receiver BA. However, the highest default rate of BA, given by the widening of its CDS curve, translates into a negative adjustment for the investor Royal Shell (whose CDS contract value is positive) as seen from Eq. (6.2.4), and consequently into an increasing mark-to-market value of the CDS contract for BA.

As the riskness of Lehman increases (larger CIR Lehman volatilities), we can see from Table 6.12 that the value of the CDS contract for British Airways increases. This is a consequence of our formula in Eq. (6.2.4) which shows that the adjustment involves a short position in a call option on the residual net value of the Lehman CDS portfolio evaluated at the time when BA defaults and a long position in a default put option on the residual net value of the Lehman CDS portfolio evaluated at the time when Shell defaults. Given that options increase with volatility and that the event of BA defaulting is more likely than the event of Shell defaulting because of the wider BA CDS term structure, we have that the short position in the call option is the driving factor of the adjustment, and it lowers the value of the risk-adjusted CDS contract for Shell as the underlying Lehman CIR volatility increases.

6.7 Conclusions

In this chapter we have provided a general arbitrage-free framework for calculating the bilateral counterparty credit valuation adjustment (BR-CVA) for payoffs exchanged between two parties, an investor and her counterparty. First, we have provided a general formula for computing the BR-CVA of financial contracts which are traded over-the-counter. Then, we have specialized our analysis to the case where also the underlying portfolio is sensitive to a third credit event, and in particular to the case where the underlying portfolio is a credit default swap on a third entity. We have then developed a Monte-Carlo numerical scheme to evaluate the formula and thus compute the BR-CVA in the specific case of credit default swap contracts. The Monte-Carlo scheme combines Fourier transform methods with filtration switching formulas and provides an efficient algorithm for numerically evaluating the adjustment. We have provided a case study in Section 6.5, and experimented with different levels of credit risk and credit risk volatilities of the three names as well as with different scenarios of default correlation. The results obtained confirm that the adjustment is sensitive to both default correlation and credit spreads volatility, having richly structured patterns that cannot be captured by rough multipliers. This points out that attempting to adapting the capital adequacy methodology (Basel II) to evaluating wrong-way risk by means of rough multipliers is not feasible. Our analysis confirms that also in the bilateral-symmetric case, wrong-way risk —namely the supplementary risk that one undergoes when the correlation assumes the worst possible value— has a structured pattern that cannot be captured by simple multipliers applied to the zero correlation case.

Chapter 7

Conclusions

7.1 Concluding Remarks

This thesis extends the credit risk literature along two main directions

- Credit Risk Modeling: It proposes a new credit risk framework for modeling deliberate distortion of the reported firm's asset values. The distortion is applied by insiders of the firm such as managers, and is not directly observable by the outside market which can only infer it from available data. The amount of distortion depends on the current performance of the firm and it also has implications on the future management of the firm.

- Counterparty risk valuation: It provides a general framework for calculating the bilateral counterparty credit valuation adjustment (BR-CVA) for payoffs exchanged between two parties, an investor and her counterparty, both bearing risk of default.

After an introduction to credit risk modeling and basic preliminaries presented in Chapters 1 and 2, Chapter 3 presents the theory underlying the proposed credit model. It describes its building blocks and derives explicit and computable expressions for bond and equity prices, conditional default probability, recovery rate, and credit spreads.

The estimate of the log-asset value of the firm in the proposed model is a crucial quantity for determining the price of corporate securities. The mathematical framework of our credit model, namely a jump linear system with state-dependent transition probabilities, leads to solving a nonlinear filtering problem evaluating the conditional density of the asset value. The problem is computationally untractable since the conditional log-asset value filter density would require an exponentially increasing number of terms. In order to deal with the exploding computational complexity, we have proposed in Chapter 5 a novel methodology which approximates the actual log-asset value density with a Gaussian mixture. Such methodology selects the smallest possible number of Gaussian components which approximate the true density within a desired level of accuracy through the solution of a second-order cone programming problem.

Chapter 4 shows the power of the proposed model in dealing with real situations of misreporting. It first introduces a novel estimation procedure for a simplified version of our model and then applies it to the Parmalat case. Such estimation procedure implies the market observations of the reported log-asset value, instead of the actual log-asset value which is only known to insiders, from equity prices. The obtained results are in line with the existing findings in the literature, and (a) confirm the importance of modeling accounting distortion in credit risk models, and (b) evidence that its omission may result in exaggeratedly large estimates of asset volatility.

Chapter 6 addresses the problem of bilateral risk evaluation in over-the-counter contracts. It develops a general arbitrage-free framework and specializes the analysis to the case where also the underlying portfolio is sensitive to a third credit event, and in particular to the case where the underlying portfolio is a credit default swap on a third entity. It develops a Monte-Carlo numerical scheme to evaluate the formula and computes the BR-CVA in the specific case of credit default swap contracts. It proposes case studies to measure the impact of both default correlation and credit spreads volatility on the adjustment, and shows that they both have a significant impact on the adjustment evaluation, thus adding to the existing literature, which only measures the impact of default correlation on the adjustment.

7.2 Future Work

There are several directions which we can follow to extend and improve the results of this thesis. We next outline some of them.

- Model endogenization. It would be interesting to define an optimal level of debt K^* and an equity maximizing bankruptcy trigger $B < K*$ at which shareholders's incentives to continue operating the firm's assets vanish.

- Dynamic consequences of accounting transparency. We could look at the dynamic implications of misreporting on the future evolution of asset value. This would require formulating misreporting as an optimal controlled diffusion problem with partial information. The bias to the asset value added by the managers would be optimally chosen in such a way as to maximize their utility function at some time horizon. Although some work on analytical methods for controlled diffusion process with partial information has been done, see Bjork et al. (2008) and Haussmann and Sass (2004), those techniques would not be able to handle our credit risk framework. There would be a need to develop efficient numerical methods combining statistical Bayesian filtering with utility maximization methods.Additional challenges would be posed by the calibration of those models. Currently, all calibrations use a constant and exogenous value for asset volatility and reporting bias, thus a new calibration methodology would be needed to handle the new scenario.

- Hedging of credit risk. The problem of credit risk hedging is very relevant and important nowadays due to the high numbers of credit rating downgrading or even default events being experienced. The challenge would be the construction of an appropriate model in which a family of options written on credit- sensitive instruments can be valued and hedged. Some efforts, although at a very abstract level start appearing Bielecki et al. (2008).

Bibliography

E. Altman, A. Resti, and A. Sironi. Analyzing and explaining default recovery rates. *Technical Report: The International Swaps and Derivatives Association*, 2000. 25

P. Artzner and F. Delbaen. Default risk insurance and incomplete markets. *Mathematical Finance*, 5: 187–195, 1995. 11

Y. Bar-Shalom, X.R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, 2002. 42

T. Bielecki, M. Jeanblanc, and M. Rutkowski. Hedging of credit default swaptions in a hazard process model. *Working paper, available at* `http://www.defaultrisk.com/pp_crdrv169.htm`, December 2008. 82

T. Bjork, M. Davis, and C. Landen. Optimal investment under partial information. *Preprint*, 2008. 82

F. Black and J. Cox. Valuing corporate securities: some effects of bond indenture provisions. *Journal of Finance*, 31:351–367, 1976. 6

H. Blom and Y. Bar-Shalom. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Transaction on Automatic Control*, 33:780–783, 1988. 43, 52

H. Blom and E. Bloem. Exact Bayesian and particle filtering of stochastic hybrid systems. *IEEE Transactions on Aerospace and Electronic Systems*, 43:55–70, 2007. 44

D. Brigo. Counterparty risk valuation with stochastic dynamical models: Impact of volatilities and correlations. In *Proceedings of 5th World Business Strategies Fixed Income Conference*, September 2008. 63

D. Brigo and A. Alfonsi. Credit default swap calibration and derivatives pricing with the ssrd stochastic intensity model. *Finance and Stochastics*, 9:29–42, 2005. 66

D. Brigo and A. Capponi. Bilateral counterparty risk with application to cds. *Submitted to Risk Magazine.* (Submission Date: January 26, 2009). 4

D. Brigo and K. Chourdakis. Counterparty risk for credit default swaps: Impact of spread volatility and default correlation. *Working paper, Fitch solutions*, 2008. 60, 65, 68, 72

D. Brigo and N. El-Bachir. An exact formula for default swaptions' pricing in the ssrjd stochastic intensity model. *Mathematical Finance*, 2008. 60, 66, 67

D. Brigo and M. Masetti. *Counterparty Credit Risk Modeling: Risk Management, Pricing and Regulation*, chapter 10, Risk Neutral Pricing of Counterparty Risk. Risk Books, London, 2005. 60

D. Brigo and F. Mercurio. *Interest Rate Models: Theory and Practise —with Smile, Inflation and Credit*. Springer Verlag, 2nd edition, 2006. 59, 66, 68, 71

D. Brigo and M. Morini. Cds calibration with tractable structural models under uncertain credit quality. *Risk Magazine*, 19, 2006. 15, 32

D. Brigo and A. Pallavicini. *Numerical Methods for Finance*, chapter Counterparty Risk under Correlation between Default and Interest Rates. Chapman Hall, 2007. 60

D. Brody, L. Hughston, and A. Macrina. Beyond hazard rates: a new approach to credit risk modelling. In *Advances in Mathematical Finance, volume in honour of D. Madan*, pages 231–257. Birkhuser and Springer, 2007. 13

D. Brody, L. Hughston, and A. Macrina. Information-based asset pricing. *International Journal of Theoretical and Applied Finance*, 11:107–142, 2008. 13

E. Canabarro and D. Duffie. Credit risk for otc derivatives portfolios: Exposure and valuation. *ALM of Financial Institutions, Institutional Investor Books ed. Leo Tilman*, 2004. 59

A. Capponi. A convex optimization approach to filtering in jump systems with state dependent transition probabilities. *Forthcoming in Automatica Elsevier*, 2009a. 4

A. Capponi. A calibration method for structural models of credit risk with reporting bias. In *Proceedings of the IEEE Conference on Computational Intelligence in Financial Engineering*, 2009b. 4

A. Capponi and J. Cvitanić. Credit risk modelling with misreporting and incomplete information. *International Journal of Theoretical and Applied Finance*, 12:83–112, 2009. 3

U. Cetin, R. Jarrow, P. Protter, and Y. Yildirim. Modeling credit risk with partial information. *Annals of Applied Probability*, 14:1167–1178, 2004. 13

S. Chen, D. Donoho, and M. Saunders. Atomic decomposition by basis pursuit. *SIAM Review*, 43:129–159, 2001. 46

U. Cherubini and M. Manera. Hunting the living dead: A peso problem in corporate liabilities data. *ICFAI Journal of Financial Risk Management*, 3:44–56, 2006. 15, 32

U. Cherubini, E. Luciano, and W. Vecchiato. *Copula Methods in Finance*. Wiley Finance, 2004. 65

K. Chourdakis. Option pricing using the fractional fft. *Journal of Computational Finance*, 8:1–18, 2005. 69

D. Coculescu, H. Geman, and M. Jeanblanc. Valuation of default sensitive claims under imperfect information. *Finance and Stochastics*, 12:195–218, 2008. 20

P. Collin-Dufresne, R. Goldstein, and J. Hugonnier. A general formula for pricing defaultable securities. *Econometrica*, 72:1377–1407, 2004. 6, 60

J. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–408, 1985. 66

D. Donoho and M. Elad. Maximal sparsity representation via l_1 minimization. *Proceedings of National Academy of Sciences*, 100:2197–2202, 2003. 46

A. Doucet, N. Gordon, and V. Krishnamurthy. Particle filters for state estimation of jump Markov linear systems. *IEEE Transactions on Signal Processing*, 49:613–624, 2001. 43

J. Duan. Correction: Maximum likelihood estimation using price data of the derivative contract. *Mathematical Finance*, 10:461–462, 2000. 29

J. Duan. Maximum likelihood estimation using price data of the derivative contract. *Mathematical Finance*, 4:155–167, 1994. 29

D. Duffie and M. Huang. Swap rates and credit quality. *Journal of Finance*, 51:921–950, 1996. 60

D. Duffie and D. Lando. Term structure of credit spreads with incomplete accounting information. *Econometrica*, 63:633–664, 2001. 13, 15, 27

D. Duffie and K. Singleton. Modeling term structures of defaultable bonds. *Review of Financial Studies*, 12: 687–720, 1999. 11

Y. Eom, J. Helwege, and J. Huang. Structural models of corporate bond pricing: An empirical analysis. *Review of Financial Studies*, 17:499–544, 2003. 29

P. Fisher and R. Verrecchia. Reporting bias. *The Accounting Review*, 75:229–245, 2000. 15

R. Frey and W. Runggaldier. Credit risk and incomplete information: a nonlinear-filtering approach. *Preprint*, 2007. 13

R. Geske. The valuation of corporate liabilities as compound options. *Journal of Financial and Quantitative Analysis*, 12:541–552, 1977. 6

K. Giesecke and L. R. Goldberg. Forecasting default in the face of uncertainty. *Journal of Derivatives*, 12: 11–25, 2004. 13

M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer, 2008. 54

X. Guo, R. A. Jarrow, and Y. Zeng. Credit risk models with incomplete information. *Forthcoming in Mathematics of Operations Research*, 2009. 13

U. Haussmann and J. Sass. Optimizing the terminal wealth under partial information: The drift process as a continuous time markov chain. *Finance and Stochastics*, 8:553–577, 2004. 82

J. Huang and M. Huang. How much of the corporate-treasury yield spread is due to credit risk? *Working paper, Pennsylvania State and Stanford Universities*, 2003. 29

J. Hull and A. White. Valuing credit default swaps: Modeling default correlations. *Working paper, University of Toronto*, 2000. 60

R. Jarrow and S. Turnbull. Pricing derivatives on financial securities subject to credit risk. *Journal of Finance*, 50:53–86, 1995. 11

M. Jeanblanc and M. Rutkowski. *Mathematical Finance: Theory and Practise*, volume 1, chapter Modeling Default Risk: An Overview, pages 71–269. *Higher Education Press*, Beijing, 2000. 20

I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus.* (Graduate Texts in Mathematics). Springer-Verlag, New York, 1995. 9

D. Lando. *Credit Risk Modeling: Theory and Applications.* Princeton University Press, 2004. 5

H. E. Leland. Prediction of default probabilities in structural models of debt. *Journal of Investment Management*, 2:5–20, 2004. 29

H. E. Leland and K. Toft. Optimal capital structure, endogeneous bankruptcy, and the term structure of credit spreads. *Journal of Finance*, 51:987–1019, 1996. 6

X. Rong Li and V. P. Jilkov. A survey of maneuvering target tracking: Multiple-model methods. *IEEE Transactions on Aerospace and Electronic Systems*, 41:1255–1321, 2005. 43

M. Lobo, L. Vanderberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, 284:193–228, 1998. 53

F. Longstaff and E. Schwartz. A simple approach for valuing risky fixed and floating rate debt. *Journal of Finance*, 50:789–819, 1995. 6

R. C. Merton. On the pricing of corporate debt: the risk structure of interest rates. *Journal of Finance*, 29: 449–470, 1974. 6

R. Nelsen. *An Introduction to Copulas.* Springer Verlag, 2006. 65

W. J. Runggaldier. *Mathematics of Finance (G.Yin and Q.Zhang eds.)*, volume 351, chapter Estimation via stochastic filtering in financial market models, pages 309–318. Contemporary Mathematics, America Mathematical Society, 2004. 21

P. Schonbucher. *Credit derivatives pricing models: Models, Pricing, Implementation.* Wiley Finance, 2003. 109

T. Wang. Real investment and corporate securities fraud. Working paper, University of Minnesota, 2007. 16, 17, 18

F. Yu. Accounting transparency and the term structure of credit spreads. *Journal of Financial Economics*, 75:53–84, 2005. 15

C. Zhou. The term structure of credit spreads with jump risk. *Journal of Banking and Finance*, 25:2015–2040, 2001. 12

Appendix A

Chapter 3: Bond and Equity Prices

We recall from Section 3.4 that the time t price of a bond with maturity T is given by

$$B(t,T) = e^{-r\tau}(K - E^{P^*}[(K - V_T)^+|\mathscr{F}_t^z]) \tag{A.0.1}$$

Following the assumptions made in Section 3.4.3 the probability measure P^* is obtained replacing μ with r and assuming that the biases $h(i)$ and the accounting risks $\nu(i)$ are the same under the historical and risk-neutral measure. We then have

$$\begin{aligned} E[e^{-r\tau}(K - V_T)^+|\mathscr{F}_t^z] &= e^{-r\tau}\int_{\mathbb{R}}\int_{\mathbb{R}} P(x_T = x|x_t = y)p_{t|t}(y)(K - e^x)^+ dydx \\ &= e^{-r\tau}\int_{\mathbb{R}} p_{t|t}(y)\int_{\mathbb{R}} P(x_T = x|x_t = y)(K - e^x)^+ dxdy \end{aligned} \tag{A.0.2}$$

where we have used the assumption made in Section 3.4.3 that the volatility risk $\sigma(l)$ is mode independent. The inner integral can be explicitly computed since $P(x_T = x|x_t = y)$ is a Gaussian density with mean $\mu_y = y + (r - 0.5\sigma^2)\tau$ and variance $\sigma^2\tau$, thus

$$\int_{\mathbb{R}} P(x_T = x|x_t = y)(K - e^x)^+ dx = E_{n_{\mu_y,\sigma^2\tau}}[(K - e^X)^+] \tag{A.0.3}$$

where $n_{\mu_y,\sigma^2\tau}$ denotes the Gaussian density with mean μ_y and variance $\sigma^2\tau$ and X is a random variable with density $n_{\mu_y,\sigma^2\tau}$.

We have

$$\begin{aligned} E_{n_{\mu_y,\sigma^2\tau}}[(K - e^X)^+] &= E_{n_{\mu_y,\sigma^2\tau}}[K\mathbf{1}_{\{K\geq e^X\}}] - E_{n_{\mu_y,\sigma^2\tau}}[e^X\mathbf{1}_{\{K\geq e^X\}}] \\ &:= I_1 - I_2 \end{aligned} \tag{A.0.4}$$

It is easy to see that I_1 is given by:

$$I_1 = KN(-d_2(y))$$

where $d_2(y) = \dfrac{\mu_y - \log(K)}{\sigma\sqrt{\tau}}$. We compute I_2 as:

$$\begin{aligned} I_2 &= \int_{-\infty}^{-d_2(y)} \exp\{\mu_y + \sigma\sqrt{\tau}z\} n(z) dz \\ &= e^{y+r\tau} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d_2(y)} \exp\{-0.5(z - \sigma\sqrt{\tau})^2\} dz \\ &= e^{y+r\tau} N(-d_1(y)) \end{aligned} \tag{A.0.5}$$

where $d_1(y) = d_2(y) + \sigma\sqrt{\tau}$. Therefore,

$$E_{n(x;\mu_y,\sigma^2\tau)}[(K - e^x)^+] = KN(-d_2(y)) - e^{y+r\tau} N(-d_1(y)) \tag{A.0.6}$$

Thus, we have that

$$E[e^{-r\tau}(K - V_T)^+ | \mathscr{F}_t^z] = E_{p_{t|t}}[Ke^{-r\tau} N(-d_2(Y)) - e^Y N(-d_1(Y))] \tag{A.0.7}$$

where Y is a random variable with density $p_{t|t}$ and the price of the bond is then given by

$$B(t,T) = Ke^{-r\tau} - E_{p_{t|t}}[Ke^{-r\tau} N(-d_2(Y)) - e^y N(-d_1(Y))] \tag{A.0.8}$$

Appendix B

Chapter 3: Approximate Bond and Equity Prices

We recall from Section 3.5.1 that the price of the bond at time t is given by

$$B(t,T) = e^{-r\tau}\left(K - \sum_j \mu_t^j E[(K - V_{T|t}^j)^+]\right) \tag{B.0.1}$$

where $V_{T|t}^j$ is the exponential of a Gaussian random variable with mean $\hat{x}_{T|t}^j = \hat{x}_{t|t}^j + (r - 0.5\sigma^2)\tau$ and standard deviation $\hat{\sigma}_{T|t}^j = \sqrt{\hat{\sigma}_{t|t}^{2,j} + \sigma^2\tau}$. We have

$$\begin{aligned} E[(K - V_{T|t}^j)^+] &= E[K\mathbf{1}_{\{K \geq V_{T|t}^j\}}] - E[V_{T|t}^j\mathbf{1}_{\{K \geq V_{T|t}^j\}}] \\ &:= I_1^j - I_2^j \end{aligned} \tag{B.0.2}$$

It is easy to see that I_1^j is given by:

$$\begin{aligned} I_1^j &= KP\left(\mathcal{N}(0,1) \leq \frac{\log(K) - \hat{x}_{t|t}^j - (r - 0.5\sigma^2)\tau}{\sqrt{\hat{\sigma}_{t|t}^{2,j} + \sigma^2\tau}}\right) \\ &= KN(-d_2^j) \end{aligned} \tag{B.0.3}$$

where $d_2^j = \dfrac{\hat{x}_{t|t}^j - \log(K) + (r - 0.5\sigma^2)\tau}{\sqrt{\hat{\sigma}_{t|t}^{2,j} + \sigma^2\tau}}$. The expression for I_2^j is given by:

$$\begin{aligned} I_2^j &= \int_{-\infty}^{-d_2^j} \exp\left\{\hat{x}_{t|t}^j + (r - 0.5\sigma^2)\tau + \sqrt{\hat{\sigma}_{t|t}^{2,j} + \sigma^2\tau}\, z\right\} n(z)dz \\ &= \exp\{\hat{x}_{t|t}^j + 0.5\hat{\sigma}_{t|t}^{2,j} + r\tau\}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-d_2^j} \exp\left\{-0.5\left(z - \sqrt{\hat{\sigma}_{T|t}^{2,j} + \sigma^2\tau}\right)^2\right\} dz \\ &= \exp\{\hat{x}_{t|t}^j + 0.5\hat{\sigma}_{t|t}^{2,j} + r\tau\}N(-d_1^j) \end{aligned} \tag{B.0.4}$$

where $d_1^j = d_2^j + \sqrt{\hat{\sigma}_{t|t}^{2,j} + \sigma^2\tau}$. Thus

$$E[e^{-r\tau}(K - V_{T|t}^j)^+] = Ke^{-r\tau}N(-d_2^j) - \exp\{\hat{x}_{t|t}^j + 0.5\hat{\sigma}_{t|t}^{2,j}\}N(-d_1^j) \tag{B.0.5}$$

Therefore, plugging Eq. (B.0.5) into Eq. (B.0.1), we have that the bond price is given by

$$B(t,T) = Ke^{-r\tau} - \sum_j \mu_t^j (Ke^{-r\tau} N(-d_2^j) - \exp\{\hat{x}_{t|t}^j + 0.5\hat{\sigma}_{t|t}^{2,j}\} N(-d_1^j)) \tag{B.0.6}$$

Appendix C

Chapter 3: Bivariate Integrals

It is possible to establish a relation between integrals of the form

$$\int_{-\infty}^{x_L} e^{ax} N(bx+c)n(x)dx \tag{C.0.1}$$

for some constants a, b, c and the bivariate normal distribution function. This is done as follows:

$$\begin{aligned}
&\int_{-\infty}^{x_L} e^{ax} N(bx+c)n(x)dx \\
&\qquad = \int_{-\infty}^{x_L}\int_{-\infty}^{c} e^{ax} n(bx+y)n(x)dydx \\
&\qquad = \frac{e^{0.5a^2}}{2\pi}\int_{-\infty}^{x_L}\int_{-\infty}^{c} e^{-\frac{(1+b^2)(x-a)^2}{2}-\frac{(y+ab)^2}{2}-b(y+ab)(x-a)}dydx
\end{aligned} \tag{C.0.2}$$

This can be related to the bivariate normal distribution as follows:

$$B(a,b,c) := \int_{-\infty}^{x_L} e^{ax} N(bx+c)n(x)dx = e^{0.5a^2}\mathcal{P}(X \le x_L, Y \le c) \tag{C.0.3}$$

where (X, Y) has a bivariate normal distribution with

$$\mu_X = a,\ \mu_Y = -ab,\ \sigma_x^2 = 1,\ \sigma_j^2 = 1+b^2,\ \rho = -\frac{b}{\sqrt{1+b^2}} \tag{C.0.4}$$

Appendix D

Chapter 3: Probability of Default

Set $l = r - 0.5\sigma^2$. Then, we have

$$\begin{aligned}
PD(t,T) &= \mathcal{P}(V_T \leq K | V_t \geq K, \mathscr{F}_t^z) \\
&= \frac{\mathcal{P}(V_T \leq K, V_t \geq K | \mathscr{F}_t^z)}{\mathcal{P}(V_t \geq K | \mathscr{F}_t^z)} \\
&= \frac{\mathcal{P}(X_T \leq \log(K), X_t \geq \log(K) | \mathscr{F}_t^z)}{\mathcal{P}(X_t \geq \log(K) | \mathscr{F}_t^z)} \\
&= \frac{\mathcal{P}(l\tau + \sigma\mathcal{N}(0,\tau) \leq -X_t + \log(K), X_t > \log(K) | \mathscr{F}_t^z)}{\mathcal{P}(X_t \geq \log(K) | \mathscr{F}_t^z)} \qquad \text{(D.0.1)}
\end{aligned}$$

where the last equality follows because we can decompose $X_T = X_t + \mathcal{N}(l\tau, \sigma^2\tau)$ as it can be easily checked. Denote $dd_t^j = \dfrac{\hat{x}_{t|t}^j - \log(K)}{\hat{\sigma}_{t|t}^j}$. Developing the probability (D.0.1) using the associated density functions, we obtain that the denominator is given by:

$$\mathcal{P}(X_t \geq \log(K) | \mathscr{F}_t^z) = 1 - \sum_{j=1}^{m} \mu_t^j N(-dd_t^j) = \sum_{j=1}^{m} \mu_t^j N(dd_t^j) \qquad \text{(D.0.2)}$$

while the numerator $\mathcal{P}(l\tau + \sigma\mathcal{N}(0,\tau) \leq -X_t + \log(K), X_t > \log(K) | \mathscr{F}_t^z)$ can be computed as follows:

$$\begin{aligned}
&= \int_{\log(K)}^{\infty} \left(\sum_{j=1}^{m} \mu_t^j n(x; \hat{x}_{t|t}^j, \hat{\sigma}_{t|t}^{2,j}) \right) \left(\int_{-\infty}^{-x+\log(K)} n(y; l\tau, \sigma^2\tau) dy \right) dx \\
&= \sum_{j=1}^{m} \mu_t^j \int_{\log(K)}^{\infty} n(x; \hat{x}_{t|t}^j, \hat{\sigma}_{t|t}^{2,j}) N\left(\frac{-x + \log(K) - l\tau}{\sigma\sqrt{\tau}} \right) dx \\
&= \sum_{j=1}^{m} \mu_t^j \int_{\log(K)}^{\infty} \hat{\sigma}_{t|t}^j n\left(\frac{x - \hat{x}_{t|t}^j}{\hat{\sigma}_{t|t}^j} \right) N\left(\frac{-x + \log(K) - l\tau}{\sigma\sqrt{\tau}} \right) dx \\
&= \sum_{j=1}^{m} \mu_t^j \int_{-\infty}^{dd_t^j} n(y) N\left(\frac{y\hat{\sigma}_{t|t}^j - \hat{x}_{t|t}^j + \log(K) - l\tau}{\sigma\sqrt{\tau}} \right) dy \qquad \text{(D.0.3)}
\end{aligned}$$

where the last line follows from a change of variable $y = \dfrac{\hat{x}^j_{t|t} - x}{\hat{\sigma}^j_{t|t}}$. Using the result in Appendix C, we can conclude that the numerator is a convex combination of bivariate cumulative distribution functions, i.e.,

$$\mathcal{P}(V_T \leq K, V_t \geq K|\mathscr{F}^z_t) = \sum_{j=1}^{m} \mu^j_t \mathcal{P}(X_j \leq dd^j_t, Y_j \leq -d_2(\hat{x}^j_{t|t})) \tag{D.0.4}$$

where $dd^j_t = \dfrac{\hat{x}^j_{t|t} - \log(K)}{\hat{\sigma}^j_{t|t}}$, d_2 is the function defined in Eq. (3.4.12), and (X_j, Y_j) has a bivariate normal distribution function with positive correlation $\rho_{X_jY_j}$:

$$\mu_{X_j} = \mu_{Y_j} = 0, \sigma^2_{X_j} = 1,\ \sigma^2_{Y_j} = 1 + \frac{\hat{\sigma}^{2,j}_{t|t}}{\sigma^2\tau},\ \rho_{X_jY_j} = \frac{\hat{\sigma}^j_{t|t}}{\sqrt{\sigma^2\tau + \hat{\sigma}^{2,j}_{t|t}}} \tag{D.0.5}$$

Defining $w^j_t = \dfrac{\mu^j_t N(dd^j_t)}{\sum_{i=1}^m \mu^i_t N(dd^i_t)}$, we obtain

$$PD(t,T) = \sum_j w^j_t \mathcal{P}(Y_j \leq -d_2(\hat{x}^j_{t|t})|X_j \leq dd^j_t) \tag{D.0.6}$$

Appendix E

Chapter 3: Recovery Rate

Set $l = (r - 0.5\sigma^2)$, $Z_\tau = l\tau + \sigma\mathcal{N}(0,\tau)$. We have

$$\begin{aligned} RR(t,T) &:= \frac{1}{K}E[V_T | V_T < K, V_t > K, \mathscr{F}_t^z] \\ &= \frac{1}{K}E[e^{X_t+Z_\tau} | X_t > \log(K), Z_\tau < \log(K) - X_t, \mathscr{F}_t^z] \\ &= \frac{\int_{\log(K)}^{\infty} e^x \int_{-\infty}^{\log(K)-x} e^{l\tau+\sigma\sqrt{\tau}y} n(y)dy \sum_{j=1}^m \mu_t^j n(x; \hat{x}_{t|t}^j \hat{\sigma}_{t|t}^{2,j})dx}{\mathcal{P}(X_t + Z_\tau < \log(K), X_t > \log(K) | \mathscr{F}_t^z)} \end{aligned} \tag{E.0.1}$$

The denominator can be decomposed as follows:

$$\mathcal{P}(X_t + Z_\tau < \log(K), X_t > \log(K) | \mathscr{F}_t^z) = PD(t,T)P(X_t > \log(K) | \mathscr{F}_t^z) \tag{E.0.2}$$

where both of the quantities above have been computed in Appendix D. The numerator in the last line of Eq. (E.0.1) can be written as:

$$\begin{aligned} &= e^{r\tau} \sum_{j=1}^m \mu_t^j \int_{\log(K)}^{\infty} e^x n(x; \hat{x}_{t|t}^j, \hat{\sigma}_{t|t}^{2,j}) N\left(\frac{\log(K) - x - l\tau - \sigma^2\tau}{\sigma\sqrt{\tau}}\right)dx \qquad \text{(E.0.3)} \\ &= e^{r\tau} \sum_{j=1}^m \mu_t^j e^{\hat{x}_{t|t}^j} \int_{-\infty}^{dd_t^j} e^{-y\hat{\sigma}_{t|t}^j} n(y) N\left(\frac{-\hat{x}_{t|t}^j + y\hat{\sigma}_{t|t}^j + \log(K) - l\tau - \sigma^2\tau}{\sigma\sqrt{\tau}}\right)dy \end{aligned}$$

where $dd_t^j = \dfrac{\hat{x}_{t|t}^j - \log K}{\hat{\sigma}_{t|t}^j}$. The last line of the derivation above follows from the change of variable $y = \dfrac{\hat{x}_{t|t}^j - x}{\hat{\sigma}_{t|t}^j}$. Using the result in Appendix C, we obtain that the numerator is again a convex combination of bivariate Gaussian cumulative distribution functions given by

$$e^{r\tau} \sum_{j=1}^m \mu_t^j e^{\hat{x}_{t|t}^j + 0.5\hat{\sigma}_{t|t}^{2,j}} \mathcal{P}(\xi_j \le dd_t^j, \psi_j \le -d_1(\hat{x}_{t|t}^j)) \tag{E.0.4}$$

where d_1 is the function defined in Eq. (3.4.12), and (ξ_j, ψ_j) is a bivariate Gaussian with negative correlation coefficient $\rho_{\xi_j\psi_j}$:

$$\mu_{\xi_j} = -\hat{\sigma}^{j}_{t|t},\ \mu_{\psi_j} = \frac{\hat{\sigma}^{2,j}_{t|t}}{\sigma\sqrt{\tau}}, \sigma^2_{\xi_j} = 1,\ \sigma^2_{\psi_j} = 1 + \frac{\hat{\sigma}^{2,j}_{t|t}}{\sigma^2\tau},\ \rho_{\xi_j\psi_j} = -\frac{\hat{\sigma}^{j}_{t|t}}{\sqrt{\sigma^2\tau + \hat{\sigma}^{2,j}_{t|t}}} \tag{E.0.5}$$

Defining $w^j_t = \dfrac{\mu^j_t \mathcal{P}(X_j \le dd^j_t, Y_j \le -d_2(\hat{x}^j_{t|t}))}{\sum_{i=1}^m \mu^i_t \mathcal{P}(X_i \le dd^i_t, Y_i \le -d_2(\hat{x}^i_{t|t}))}$, we obtain

$$RR(t,T) = e^{r\tau} \sum_{j=1}^{m} w^j_t \frac{e^{\hat{x}^j_{t|t} + \frac{\hat{\sigma}^{2,j}_{t|t}}{2}}}{K} \frac{\mathcal{P}(\xi_j \le dd^j_t, \psi_j \le -d_1(\hat{x}^j_{t|t}))}{\mathcal{P}(X_j \le dd^j_t, Y_j \le -d_2(\hat{x}^j_{t|t}))} \tag{E.0.6}$$

Appendix F

Chapter 5: Facts about Lipschitz Functions

The l_p norm of an n dimensional vector $\mathbf{x}$ is defined as follows:

$$||\mathbf{x}||_p = \begin{cases} \left[\sum_i |x_i|^p\right]^{\frac{1}{p}} & 1 \leq p < \infty \\ \max_{i=1,\ldots n} |x_i| & p = \infty \end{cases}$$

Lemma F.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$. $\forall \mathbf{x}_1, \mathbf{x}_2$ it holds that*

$$|f(\mathbf{x}_1) - f(\mathbf{x}_2)| \leq C_p ||\mathbf{x}_1 - \mathbf{x}_2||_q \tag{F.0.1}$$

where $C_p = \sup_{\mathbf{x}} ||\nabla f(\mathbf{x})||_p$, $\nabla f(\mathbf{x}) = (\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n})$ and $\frac{1}{p} + \frac{1}{q} = 1$, $1 \leq p, q \leq \infty$.

We next state an obvious result as a lemma. Such result states that the product of two bounded Lipschitz functions is a bounded Lipschitz function.

Lemma F.2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$. Let $C_{f,p}, M_f, C_{g,p}, M_g$ constants such that for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$, we have*

$$\begin{aligned} |f(\mathbf{x}_1) - f(\mathbf{x}_2)| &\leq C_{f,p} ||\mathbf{x}_1 - \mathbf{x}_2||_p \\ |g(\mathbf{x}_1) - g(\mathbf{x}_2)| &\leq C_{g,p} ||\mathbf{x}_1 - \mathbf{x}_2||_p \end{aligned} \tag{F.0.2}$$

and

$$\begin{aligned} |f(\mathbf{x})| &\leq M_f \\ |g(\mathbf{x})| &\leq M_g \end{aligned} \tag{F.0.3}$$

Then,

$$|f(\mathbf{x}_1)g(\mathbf{x}_1) - f(\mathbf{x}_2)g(\mathbf{x}_2)| \leq (M_f C_{g,p} + M_g C_{f,p}) ||\mathbf{x}_1 - \mathbf{x}_2||_p \tag{F.0.4}$$

Appendix G

Chapter 5: Total Variation distance at initial time

Using approximation (5.3.27), we can write

$$||p^l_{1,1}(\mathbf{x}) - \hat{p}^l_{1,1}(\mathbf{x})||_1 \approx \int_G |p^l_{1,1}(\mathbf{x}) - \hat{p}^l_{1,1}(\mathbf{x})| d\mathbf{x} \tag{G.0.1}$$

Since the vectors in the training set are equally spaced in G, we can partition G into n-dimensional cubes and denote by G_i the n-dimensional cube in the partition having $\mathbf{x}_i$ as the bottom left vertex of the cube.

Using the triangular inequality, we have

$$\begin{aligned}
&\int_G |p^l_{1,1}(\mathbf{x}) - \hat{p}^l_{1,1}(\mathbf{x})| d\mathbf{x} \\
&\quad = \sum_{i=1}^{q} \int_{G_i} |p^l_{1,1}(\mathbf{x}) - \hat{p}^l_{1,1}(\mathbf{x})| d\mathbf{x} \\
&\quad \leq \sum_{i=1}^{q} \int_{G_i} |p^l_{1,1}(\mathbf{x}) - p^l_{1,1}(\mathbf{x}_i)| d\mathbf{x} \\
&\qquad + \int_{G_i} |p^l_{1,1}(\mathbf{x}_i) - \hat{p}^l_{1,1}(\mathbf{x}_i)| d\mathbf{x} \\
&\qquad + \int_{G_i} |\hat{p}^l_{1,1}(\mathbf{x}_i) - \hat{p}^l_{1,1}(\mathbf{x})| d\mathbf{x}
\end{aligned} \tag{G.0.2}$$

Next, we analyze separately the three terms in Eq. (G.0.2). The first term is upper bounded as follows

$$\begin{aligned}
&\int_{G_i} |p^l_{1,1}(\mathbf{x}) - p^l_{1,1}(\mathbf{x}_i)| d\mathbf{x} \\
&\quad = \int_{G_i} \left| \int_{\mathbb{R}^n} P(\mathbf{x}_0 = \mathbf{y}) \lambda^l(\mathbf{y}) \right. \\
&\qquad \left. \cdot (L^l_{1|1}(\mathbf{x}) p(\mathbf{x}|\mathbf{y}, l) - L^l_{1|1}(\mathbf{x}_i) p(\mathbf{x}_i|\mathbf{y}, l) d\mathbf{y}) \right| d\mathbf{x} \\
&\quad \leq \int_{\mathbb{R}^n} P(\mathbf{x}_0 = \mathbf{y}) \lambda^l(\mathbf{y}) \\
&\qquad \cdot \int_{G_i} |L^l_{1|1}(\mathbf{x}) p(\mathbf{x}|\mathbf{y}, l) - L^l_{1|1}(\mathbf{x}_i) p(\mathbf{x}_i|\mathbf{y}, l)| d\mathbf{x} d\mathbf{y} \\
&\quad := e^{l,2}_1
\end{aligned} \tag{G.0.3}$$

where the last step in Eq. (G.0.3) follows from application of Fubini's theorem which allows exchange of the order of the integrations. Both $L^l_{1|1}(\mathbf{x})$ and $p(\mathbf{x}|\mathbf{y}, l)$ are multivariate Gaussian densities. We consider both of them as functions of $\mathbf{x}$, thus treating $\mathbf{y}$ as a constant in $p(\mathbf{x}|\mathbf{y}, l)$. We have that the l_∞ norms of their gradients are both finite. Notice that although $p(\mathbf{x}|\mathbf{y}, l)$ is dependent on y, being its mean a function of $\mathbf{y}$, the l_∞ norm of its gradient is independent of $\mathbf{y}$ because the mean of a Gaussian density does not affect the l_∞ norm of the gradient.

Moreover,

$$L^l_{1|1}(\mathbf{x}) \leq \frac{1}{[2\pi]^{\frac{m}{2}} \det[\mathbf{R}(l)]^{\frac{1}{2}}}$$

$$p(\mathbf{x}|\mathbf{y}, l) \leq \frac{1}{[2\pi]^{\frac{n}{2}} \det[\mathbf{Q}(l)]^{\frac{1}{2}}} \tag{G.0.4}$$

We can then apply Lemma F.1 setting $p = \infty$ and $q = 1$ along with Lemma F.2 to obtain

$$\begin{aligned}
&\int_{G_i} |p^l_{1,1}(\mathbf{x}) - p^l_{1,1}(\mathbf{x}_i)| d\mathbf{x} \\
&\quad \leq \quad C^{(l)} \int_{\mathbb{R}^n} P(x_0 = \mathbf{y}) \lambda^l(\mathbf{y}) \left(\int_{G_i} ||\mathbf{x} - \mathbf{x}_i||_1 d\mathbf{x} \right) d\mathbf{y} \\
&\quad \leq \quad \frac{1}{2} C^{(l)} n \rho^{n+1} E_{n(\mu_0, \mathbf{\Sigma}_0)}[\lambda^l]
\end{aligned} \tag{G.0.5}$$

where $C^{(l)}$ is computed as indicated in Lemma F.2 and given by

$$C^{(l)} = \sup_{\mathbf{x} \in \mathbb{R}^n} \frac{||\nabla L^l_{1|1}(\mathbf{x})||_\infty}{(2\pi)^{\frac{n}{2}} \det(\mathbf{Q}(l))^{\frac{1}{2}}} + \sup_{\mathbf{x} \in \mathbb{R}^n} \frac{||\nabla p(\mathbf{x}|\mathbf{y}, l)||_\infty}{(2\pi)^{\frac{m}{2}} \det(\mathbf{R}(l))^{\frac{1}{2}}} \tag{G.0.6}$$

Here, we recall that ρ denotes the side length of each cube in the partition of G.

The second term in Eq. (G.0.2) can be upper bounded as

$$\int_{G_i} |p^l_{1,1}(\mathbf{x}_i) - \hat{p}^l_{1,1}(\mathbf{x}_i)| d\mathbf{x} \leq \epsilon \rho^n := e^{l,3}_1 \tag{G.0.7}$$

which follows from the l_2 norm of the error in the programming problem (P2) being bounded by ϵ.

We now provide an upper bound for the third term in Eq. (G.0.2).

$$\begin{aligned}
&\int_{G_i} |\hat{p}^l_{1,1}(\mathbf{x}_i) - \hat{p}^l_{1,1}(\mathbf{x})| d\mathbf{x} \\
&\quad = \quad \int_{G_i} \left| \sum_{(j,h) \in \Omega^l_1} [\varpi^l_1]_{(j,h)} [n_{j,h}(\mathbf{x}_i) - n_{j,h}(\mathbf{x})] \right| d\mathbf{x} \\
&\quad \leq \quad \int_{G_i} \sum_{(j,h) \in \Omega^l_1} [\varpi^l_1]_{(j,h)} |n_{j,h}(\mathbf{x}_i) - n_{j,h}(\mathbf{x})| d\mathbf{x} \\
&\quad \leq \quad \int_{G_i} \sum_{(j,h) \in \Omega^l_1} [\varpi^l_1]_{(j,h)} C_{j,h} ||\mathbf{x} - \mathbf{x}_i||_1 d\mathbf{x} \\
&\quad \leq \quad \frac{1}{2} C n \rho^{n+1} \cdot ||\varpi^l_1||_1 := e^{l,4}_1
\end{aligned} \tag{G.0.8}$$

where $C_{j,h} = \sup_{\mathbf{x}\in\mathbb{R}^n} ||\nabla n_{j,h}(\mathbf{x})||_\infty$ and $C_1 = \max_{(j,h)\in\Omega_1^l} C_{j,h}$. The inequalities above follow from application of the Lipschitz property of the multivariate Gaussian density, after using the approximation (5.3.27) introduced in Section 5.3.4.

Combining Eq. (G.0.5), (G.0.7), and (G.0.8), we can conclude that

$$||\tilde{p}^l_{1,1}||_1 \leq q(e_1^{l,2} + e_1^{l,3} + e_1^{l,4}) \tag{G.0.9}$$

Appendix H

Chapter 5: Total Variation Distance at Time k

It holds that

$$||p^l_{k,k} - \hat{p}^l_{k,k}||_1 \leq ||p^l_{k,k} - \check{p}^l_{k,k}||_1 + ||\check{p}^l_{k,k} - \hat{p}^l_{k,k}||_1 \tag{H.0.1}$$

We start providing an upper bound for the first term

$$||p^l_{k,k} - \check{p}^l_{k,k}||_1 = \int_{\mathbb{R}^n} |p^l_{k,k}(\mathbf{x}) - \check{p}^l_{k,k}(\mathbf{x})| d\mathbf{x} \tag{H.0.2}$$

We have

$$\begin{aligned}
&\int_{\mathbb{R}^n} |p^l_{k,k}(\mathbf{x}) - \check{p}^l_{k,k}(\mathbf{x})| d\mathbf{x} \\
&= \int_{\mathbb{R}^n} L^l_{k|k}(\mathbf{x}) \left| \sum_{r=1}^{d} \int_{\mathbb{R}^n} \lambda^{r,l}(\mathbf{y}) p(\mathbf{x}|\mathbf{y}, l) \tilde{p}^r_{k-1,k-1}(\mathbf{y}) d\mathbf{y} \right| d\mathbf{x} \\
&\leq \int_{\mathbb{R}^n} L^l_{k|k}(\mathbf{x}) \sum_{r=1}^{d} \int_{\mathbb{R}^n} \lambda^{r,l}(\mathbf{y}) p(\mathbf{x}|\mathbf{y}, l) |\tilde{p}^r_{k-1,k-1}(\mathbf{y})| d\mathbf{y} d\mathbf{x} \\
&= \sum_{r=1}^{d} \int_{\mathbb{R}^n} \lambda^{r,l}(\mathbf{y}) |\tilde{p}^r_{k-1,k-1}(\mathbf{y})| \left(\int_{\mathbb{R}^n} L^l_{k|k}(\mathbf{x}) p(\mathbf{x}|\mathbf{y}, l) d\mathbf{x} \right) d\mathbf{y} \\
&:= e^{l,1}_k
\end{aligned} \tag{H.0.3}$$

The inside integral evaluates to a Gaussian density

$$\int_{\mathbb{R}^n} L^l_{k|k}(\mathbf{x}) p(\mathbf{x}|\mathbf{y}, l) d\mathbf{x} = n(\mathbf{z}_k, \mathbf{m}_{l,y}, \mathbf{S}_l) \tag{H.0.4}$$

where

$$\begin{aligned}
\mathbf{m}_{l,\mathbf{y}} &= \mathbf{H}(l)\mathbf{A}(l)\mathbf{y} \\
\mathbf{S}_l &= \mathbf{H}\mathbf{Q}(l)\mathbf{H}^T + \mathbf{R}(l)
\end{aligned} \tag{H.0.5}$$

Therefore, we can rewrite Eq. (H.0.3) as

$$
\begin{aligned}
&\sum_{r=1}^{d}\int_{\mathbb{R}^n}\lambda^{r,l}(\mathbf{y})|\tilde{p}^r_{k-1,k-1}(\mathbf{y})|n(\mathbf{z}_k,\mathbf{m}_{l,y},\mathbf{S}_l)d\mathbf{y}\\
\leq\;& \frac{1}{(2\pi)^{\frac{n}{2}}\det(\mathbf{S}_l)^{\frac{1}{2}}}\sum_{r=1}^{d}\sup_{\mathbf{y}\in\mathbb{R}^n}\lambda^{r,l}(\mathbf{y})\int_{\mathbb{R}^n}|\tilde{p}^r_{k-1,k-1}(\mathbf{y})|d\mathbf{y}\\
\leq\;& \frac{1}{(2\pi)^{\frac{n}{2}}\det(\mathbf{S}_l)^{\frac{1}{2}}}\sum_{r=1}^{d}\sup_{\mathbf{y}\in\mathbb{R}^n}\lambda^{r,l}(\mathbf{y})\delta_r
\end{aligned}
\tag{H.0.6}
$$

where the first step follows form the monotonicity property of an integral after bounding both $\lambda^{r,l}(\mathbf{y})$ and $n(\mathbf{z}_k,\mathbf{m}_{l,y},\mathbf{S}_l)$ from above, while the last step follows from the assumption of the theorem.

Next, we provide an upper bound for the second term in the inequality (H.0.1) which is approximated according to Eq. (5.3.27) as

$$
||\check{p}^l_{k,k}-\hat{p}^l_{k,k}||_1 \approx \int_G |\check{p}^l_{k,k}(\mathbf{x})-\hat{p}^l_{k,k}(\mathbf{x})|d\mathbf{x} \tag{H.0.7}
$$

We first give an upper bound for the integral in Eq. (H.0.7) using the triangular inequality

$$
\begin{aligned}
&\int_G |\check{p}^l_{k,k}(\mathbf{x})-\hat{p}^l_{k,k}(\mathbf{x})|d\mathbf{x}\\
&\quad= \sum_{i=1}^{q}\int_{G_i}|\check{p}^l_{k,k}(\mathbf{x})-\hat{p}^l_{k,k}(\mathbf{x})|d\mathbf{x}\\
&\quad\leq \sum_{i=1}^{q}\Bigg(\int_{G_i}|\check{p}^l_{k,k}(\mathbf{x})-\check{p}^l_{k,k}(\mathbf{x}_i)|d\mathbf{x}\\
&\qquad\quad+\int_{G_i}|\check{p}^l_{k,k}(\mathbf{x}_i)-\hat{p}^l_{k,k}(\mathbf{x}_i)|d\mathbf{x}\\
&\qquad\quad+\int_{G_i}|\hat{p}^l_{k,k}(\mathbf{x}_i)-\hat{p}^l_{k,k}(\mathbf{x})|d\mathbf{x}\Bigg)
\end{aligned}
\tag{H.0.8}
$$

and then analyze separately the three terms in Eq. (H.0.8). The first term in Eq. (H.0.8) is upper bounded as follows

$$
\begin{aligned}
&\int_{G_i}|\check{p}^l_{k,k}(\mathbf{x})-\check{p}^l_{k,k}(\mathbf{x}_i)|d\mathbf{x}\\
=\;& \int_{G_i}\Bigg|\sum_{r=1}^{d}\int_{\mathbb{R}^n}\hat{p}^r_{k-1,k-1}(\mathbf{y})\lambda^{r,l}(\boldsymbol{y})\\
&\cdot\Big(L^l_{k|k}(\mathbf{x})p(\mathbf{x}|\mathbf{y},l)-L^l_{k|k}(\mathbf{x}_i)p(\mathbf{x}_i|\mathbf{y},l)d\mathbf{y}\Big)\Bigg|d\mathbf{x}\\
\leq\;& \sum_{r=1}^{d}\int_{\mathbb{R}^n}\hat{p}^r_{k-1,k-1}(\mathbf{y})\lambda^{r,l}(\mathbf{y})\\
&\cdot\Big(\int_{G_i}|L^l_{k|k}(\mathbf{x})p(\mathbf{x}|\mathbf{y},l)-L^l_{k|k}(\mathbf{x}_i)p(\mathbf{x}_i|\mathbf{y},l)|d\mathbf{x}\Big)d\mathbf{y}\\
:=\;& e^{l,2}_k
\end{aligned}
\tag{H.0.9}
$$

where the last inequality in (H.0.9) follows from bringing the absolute value inside the integral and then applying Fubini's theorem to exchange the order of integration. Applying Lemma F.1 setting $p = \infty$ and $q = 1$ along with Lemma F.2, we obtain

$$\begin{aligned}
&\int_{G_i} |\breve{p}^l_{k,k}(\mathbf{x}) - \breve{p}^l_{k,k}(\mathbf{x}_i)|d\mathbf{x} \\
&\leq \quad C^{(l)} \sum_{r=1}^{d} \int_{\mathbb{R}^n} \hat{p}^r_{k-1,k-1}(\mathbf{y})\lambda^{r,l}(\mathbf{y})(\int_{G_i} ||\mathbf{x} - \mathbf{x}_i||_1 d\mathbf{y})d\mathbf{x} \\
&= \quad C^{(l)}\rho^n ||\mathbf{x}_{i+1} - \mathbf{x}_i||_1 \sum_{r=1}^{d} \int_{\mathbb{R}^n} \hat{p}^r_{k-1,k-1}(\mathbf{y})\lambda^{r,l}(\mathbf{y}) \\
&= \quad \frac{1}{2} C^{(l)} n\rho^{n+1} \sum_{r=1}^{d} \sum_{(i,j)\in\Omega^r_k} [\varpi^r_{k-1}]_{(i,j)} E_{n_{i,j}}[\lambda^{r,l}]
\end{aligned} \tag{H.0.10}$$

where $C^{(l)}$ is the same Lipschitz constant given in Eq. (G.0.6), the transition matrix $\mathbf{A}$, the design matrix $\mathbf{H}$, and the mode dependent noise covariance matrices all being time invariant. We recall that ρ denotes the side length of each cube in the grid G. The last equality in Eq. (H.0.10) just follows from the definition of $\hat{p}^r_{k-1,k-1}(\mathbf{y})$ which is a weighted sum of Gaussian densities.

The second term in Eq. (H.0.8) is upper bounded as

$$\int_{G_i} |\breve{p}^l_{k,k}(\mathbf{x}_i) - \hat{p}^l_{k,k}(\mathbf{x}_i)|d\mathbf{x} \leq \epsilon\rho^n := e^{l,3}_k \tag{H.0.11}$$

which follows from the fact that the l_2 norm of the error in the programming problem (P2) is bounded from above by ϵ.

Using the exact same argument as in Appendix G, we can upper bound the third term in Eq. (H.0.8) as

$$\begin{aligned}
&\int_{G_i} |\hat{p}^l_{1,1}(\mathbf{x}_i) - \hat{p}^l_{1,1}(\mathbf{x})|d\mathbf{x} \\
&\qquad \leq \quad C_k n\rho^{n+1} \cdot ||\varpi^l_k||_1 := e^{l,4}_k
\end{aligned} \tag{H.0.12}$$

where $C_k = \max_{(j,h)\in\Omega^l_k} C_{j,h}$ and $C_{j,h} = \sup_{\mathbf{x}\in\mathbb{R}^n} ||\nabla n_{j,h}(\mathbf{x})||_\infty$.

Combining Eq. (H.0.6), (H.0.10), (H.0.11), and (H.0.12), we can conclude that

$$||\tilde{p}^l_{k,k}|| \leq e^{l,1}_k + q(e^{l,2}_k + e^{l,3}_k + e^{l,4}_k) \tag{H.0.13}$$

Appendix I

Chapter 6: Brief Overview of Copula Functions

Copula functions represent a general and flexible tool to measure association among probabilities. Association is a more general concept than correlation, encompassing both linear and nonlinear relationships among variables. While correlation measures linear relationships between variables, copula functions are invariant to any linear or nonlinear transformation of the variables. The basic idea is based on the *principle of probability integral transformation*. Consider variables X and Y with marginal distributions F_X and F_Y. The principle states that transformations $u \equiv F_X(X)$ and $v \equiv F_Y(Y)$ have uniform distribution in $[0, 1]$. Marginal distributions can be inverted, and if they are continuous such inverse is unique. Consider now the joint distribution $H(X, Y)$. This can be written as

$$H(X, Y) = H(F_X^{-1}(u), F_Y^{-1}(v)) \equiv C(u, v) \tag{I.0.1}$$

where $C(u, v)$ is called the copula function representing association between X and Y. So, every joint distribution can be written as a function taking the marginal distributions as arguments. On the contrary, one can prove that if $C(u, v)$ satisfy appropriate requirements, then plugging univariate distributions into it generates a joint distribution. The requirements for $C(u, v)$ are summarized in what is known as Sklar's theorem, and are reported below.

Definition I.0.1. *A copula function $C(u, v)$ has domain in the two-dimensional unit hypercube and range in the unit interval and satisfies the following requirements:*

1. ***Groundedness**: $C(u, 0) = C(0, v) = 0$*
2. ***Identity of marginals**: $C(u, 1) = v$, $C(1, v) = u$*
3. ***2-increasing**: $C(u_1, v_1) - C(u_1, v_2) - C(u_2, v_1) + C(u_2, v_2) \geq 0$, with $u_1 > u_2$,$v_1 > v_2$.*

Appendix J

Chapter 6: Proof of the General Counterparty Risk Pricing Formula

Proof. We have that

$$\begin{aligned}\Pi(t,T) &= \textsc{Cashflows}(t,T) \\ &= \mathbf{1}_{A\cup B}\textsc{Cashflows}(t,T) + \mathbf{1}_{C\cup D}\textsc{Cashflows}(t,T) + \mathbf{1}_{E\cup F}\textsc{Cashflows}(t,T) \qquad \text{(J.0.1)}\end{aligned}$$

since the events in Eq. (6.2.2) form a complete set. From the linearity of the expectation, we can rewrite the right-hand side of Eq. (6.2.4) as

$$\mathbb{E}_t\{\Pi(t,T) + \mathrm{LGD}_0 \cdot \mathbf{1}_{A\cup B} \cdot D(t,\tau_0) \cdot [-\mathrm{NPV}(\tau_0)]^+ - \mathrm{LGD}_2 \cdot \mathbf{1}_{C\cup D} \cdot D(t,\tau_2) \cdot [\mathrm{NPV}(\tau_2)]^+\} \qquad \text{(J.0.2)}$$

We can then rewrite the formula in Eq. (J.0.2) using Eq. J.0.1 as

$$\begin{aligned}&= E_t[\mathbf{1}_{A\cup B}\textsc{Cashflows}(t,T) + (1-\textsc{Rec}_{,0})\mathbf{1}_{A\cup B}D(t,\tau_0)[-NPV(\tau_0)]^+ \\ &\quad +\mathbf{1}_{C\cup D}\textsc{Cashflows}(t,T) + (\textsc{Rec}_{,2}-1)\mathbf{1}_{C\cup D}D(t,\tau_2)[NPV(\tau_2)]^+ \\ &\quad +\mathbf{1}_{E\cup F}\textsc{Cashflows}(t,T)] \\ &= E_t[\mathbf{1}_{A\cup B}\textsc{Cashflows}(t,T) + (1-\textsc{Rec}_{,0})\mathbf{1}_{A\cup B}D(t,\tau_0)[-NPV(\tau_0)]^+] \\ &\quad +E_t[\mathbf{1}_{C\cup D}\textsc{Cashflows}(t,T) + (\textsc{Rec}_{,2}-1)\mathbf{1}_{C\cup D}D(t,\tau_2)[NPV(\tau_2)]^+] \\ &\quad +E_t[\mathbf{1}_{E\cup F}\textsc{Cashflows}(t,T)] \qquad \text{(J.0.3)}\end{aligned}$$

We next develop each of the three expectations in the equality of Eq. (J.0.3).

The expression inside the first expectation can be rewritten as

$$\begin{aligned}&\mathbf{1}_{A\cup B}\textsc{Cashflows}(t,T) + (1-\textsc{Rec}_{,0})\mathbf{1}_{A\cup B}D(t,\tau_0)[-NPV(\tau_0)]^+ \\ =\ &\mathbf{1}_{A\cup B}\textsc{Cashflows}(t,T) + \mathbf{1}_{A\cup B}D(t,\tau_0)[-NPV(\tau_0)]^+ - \textsc{Rec}_{,0}\mathbf{1}_{A\cup B}D(t,\tau_0)[-NPV(\tau_0)]^+\end{aligned}$$

(J.0.4)

Conditional on the information at τ_0, the expectation of the expression in Eq. (J.0.4) is equal to

$$\begin{aligned}
& \mathbb{E}_{\tau_0}\left[\mathbf{1}_{A\cup B}\textsc{Cashflows}(t,T) + \mathbf{1}_{A\cup B}D(t,\tau_0)\left(-NPV(\tau_0)\right)^+ - \textsc{Rec}_{,0}\mathbf{1}_{A\cup B}D(t,\tau_0)[-NPV(\tau_0)]^+\right] \\
= \; & \mathbb{E}_{\tau_0}[\mathbf{1}_{A\cup B}[\textsc{Cashflows}(t,\tau_0) + D(t,\tau_0)\textsc{Cashflows}(\tau_0,T) + D(t,\tau_0)\left(-\mathbb{E}_{\tau_0}\left[\textsc{Cashflows}(\tau_0,T)\right]\right)^+ \\
& -\textsc{Rec}_{,0}D(t,\tau_0)[-NPV(\tau_0)]^+]] \\
= \; & \mathbf{1}_{A\cup B}[\textsc{Cashflows}(t,\tau_0) + D(t,\tau_0)\mathbb{E}_{\tau_0}\left[\textsc{Cashflows}(\tau_0,T)\right] + D(t,\tau_0)\left(-\mathbb{E}_{\tau_0}\left[\textsc{Cashflows}(\tau_0,T)\right]\right)^+ \\
& -\textsc{Rec}_{,0}D(t,\tau_0)[-NPV(\tau_0)]^+] \\
= \; & \mathbf{1}_{A\cup B}[\textsc{Cashflows}(t,\tau_0) + D(t,\tau_0)\left(\mathbb{E}_{\tau_0}\left[\textsc{Cashflows}(\tau_0,T)\right]\right)^+ - \textsc{Rec}_{,0}D(t,\tau_0)[-NPV(\tau_0)]^+] \\
= \; & \mathbf{1}_{A\cup B}[\textsc{Cashflows}(t,\tau_0) + D(t,\tau_0)\left(NPV(\tau_0)\right)^+ - \textsc{Rec}_{,0}D(t,\tau_0)[-NPV(\tau_0)]^+]
\end{aligned}$$

where the first equality in Eq. (J.0.5) follows because

$$\mathbf{1}_{A\cup B}\textsc{Cashflows}(t,T) = \mathbf{1}_{A\cup B}[\textsc{Cashflows}(t,\tau_0) + D(t,\tau_0)\textsc{Cashflows}(\tau_0,T)] \tag{J.0.5}$$

with the default time τ_0 is always smaller than T under the event $A \cup B$. Conditioning the obtain result on the information available at t, and using the fact that $E_t[E_{\tau_0}[.]] = E_t[.]$ due to $t < \tau_0$, we obtain that the first term in Eq. (J.0.3) is given by

$$\mathbb{E}_t\left[\mathbf{1}_{A\cup B}\left[\textsc{Cashflows}(t,\tau_0) + D(t,\tau_0)(NPV(\tau_0))^+ - \textsc{Rec}_{,0}D(t,\tau_0)(-NPV(\tau_0))^+\right]\right] \tag{J.0.6}$$

which coincides with the expectation of the third term in Eq. (6.2.3).

We next repeat a similar argument for the second expectation in Eq. (J.0.3). We have

$$\begin{aligned}
& \mathbf{1}_{C\cup D}\textsc{Cashflows}(t,T) + (\textsc{Rec}_{,2} - 1)\mathbf{1}_{C\cup D}D(t,\tau_2)[NPV(\tau_2)]^+ \\
= \; & \mathbf{1}_{C\cup D}\textsc{Cashflows}(t,T) - \mathbf{1}_{C\cup D}D(t,\tau_2)[NPV(\tau_2)]^+ + \textsc{Rec}_{,2}\mathbf{1}_{C\cup D}D(t,\tau_2)[NPV(\tau_2)]^+
\end{aligned} \tag{J.0.7}$$

Conditional on the information available at time τ_2, we have

$$\begin{aligned}
& \mathbb{E}_{\tau_2}\left[\mathbf{1}_{C\cup D}\textsc{Cashflows}(t,T) - \mathbf{1}_{C\cup D}D(t,\tau_2)\left(NPV(\tau_2)\right)^+ \textsc{Rec}_{,2}\mathbf{1}_{C\cup D}D(t,\tau_2)[NPV(\tau_2)]^+\right] \\
= \; & \mathbb{E}_{\tau_2}[\mathbf{1}_{C\cup D}[\textsc{Cashflows}(t,\tau_2) + D(t,\tau_2)\textsc{Cashflows}(\tau_2,T) - D(t,\tau_2)\left(\mathbb{E}_{\tau_2}\left[\textsc{Cashflows}(\tau_2,T)\right]\right)^+ \\
& +\textsc{Rec}_{,2}D(t,\tau_2)[NPV(\tau_2)]^+]] \\
= \; & \mathbf{1}_{C\cup D}[\textsc{Cashflows}(t,\tau_2) + D(t,\tau_2)\mathbb{E}_{\tau_2}\left[\textsc{Cashflows}(\tau_2,T)\right] - D(t,\tau_2)\left(\mathbb{E}_{\tau_2}\left[\textsc{Cashflows}(\tau_2,T)\right]\right)^+ \\
& +\textsc{Rec}_{,2}D(t,\tau_2)[NPV(\tau_2)]^+] \\
= \; & \mathbf{1}_{C\cup D}\left[\textsc{Cashflows}(t,\tau_2) - D(t,\tau_2)\left(\mathbb{E}_{\tau_2}\left[-\textsc{Cashflows}(\tau_2,T)\right]\right)^+ + \textsc{Rec}_{,2}D(t,\tau_2)[NPV(\tau_2)]^+\right] \\
= \; & \mathbf{1}_{C\cup D}\left[\textsc{Cashflows}(t,\tau_2) - D(t,\tau_2)(-NPV(\tau_2))^+ + \textsc{Rec}_{,2}D(t,\tau_2)[NPV(\tau_2)]^+\right]
\end{aligned} \tag{J.0.8}$$

where the first equality follows because

$$\mathbf{1}_{C\cup D}\textsc{Cashflows}(t,T) = \mathbf{1}_{C\cup D}[\textsc{Cashflows}(t,\tau_2) + D(t,\tau_2)\textsc{Cashflows}(\tau_2,T)] \tag{J.0.9}$$

with the default time τ_2 always smaller than T under the event $C \cup D$. Conditioning the obtain result on the information available at $t < \tau_2$, we obtain that the second term in Eq. (J.0.3) is given by

$$\mathbb{E}_t\left[\mathbf{1}_{C\cup D}\left[\textsc{Cashflows}(t,\tau_2) + D(t,\tau_2)\textsc{Rec}_{,2}(NPV(\tau_2))^+ - D(t,\tau_2)(-NPV(\tau_2))^+\right]\right] \tag{J.0.10}$$

which coincides exactly with the expectation of the second term in Eq. (6.2.3).

The third expectation in Eq. (J.0.3) coincides with the first term in Eq. (6.2.3), therefore their expectations ought to be the same. Since we have proven that the expectation of each term in Eq. (6.2.3) equals the expectation of the corresponding term in Eq. (J.0.3), the desired result is obtained. □

Appendix K

Chapter 6: Proof of the Survival Probability Formula

Proof. We have

$$\begin{aligned}\mathbf{1}_{\tau_1>\tau_2}\mathbb{Q}(\tau_1>t|\mathcal{G}_{\tau_2}) &= \mathbf{1}_{t<\tau_2<\tau_1}+\mathbf{1}_{\tau_2<t}\mathbf{1}_{\tau_1\geq\tau_2}\mathbb{E}[\mathbb{Q}(\Lambda_1(t)<\xi_1|\mathcal{G}_{\tau_2},\xi_1)|\mathcal{G}_{\tau_2},\{\tau_1>\tau_2\}]\\ &= \mathbf{1}_{\bar{A}}+\mathbf{1}_{\tau_2<t}\mathbf{1}_{\tau_1\geq\tau_2}\mathbb{E}[F_{\Lambda_1(t)-\Lambda_1(\tau_2)}(\xi_1-\Lambda_1(\tau_2))|\mathcal{G}_{\tau_2},\{\tau_1>\tau_2\}] \qquad \text{(K.0.1)}\end{aligned}$$

The last step, following from the fact that the $\Lambda_1(t)<\xi_1$ is the same as $\Lambda_1(t)-\Lambda_1(\tau_2)<\xi_1-\Lambda_1(\tau_2)$ and the right-hand side $\xi_1-\Lambda_1(\tau_2)$, becomes known once we condition on ξ_1 and $\mathcal{G}_{\tau_2}$. Here by $F_{\Lambda_1(t)-\Lambda_1(\tau_2)}$, we indicate the cumulative distribution function of the integrated (shifted) CIR process $\Lambda_1(t)-\Lambda_1(\tau_2)$. Since $\xi_1=-\log(1-U_1)$ and $\tau_1=\Lambda_1^{-1}(\xi_1)$, we can rewrite Eq. (K.0.1) as

$$\begin{aligned}&\mathbf{1}_{\bar{A}}+\mathbf{1}_{\tau_2<t}\mathbf{1}_{\tau_1\geq\tau_2}\mathbb{E}[F_{\Lambda_1(t)-\Lambda_1(\tau_2)}(-\log(1-U_1)-\Lambda_1(\tau_2))|\mathcal{G}_{\tau_2},\{\xi_1>\Lambda_1(\tau_2)\}]=\\ &= \mathbf{1}_{\bar{A}}+\mathbf{1}_{\tau_2<t}\mathbf{1}_{\tau_1\geq\tau_2}\mathbb{E}[F_{\Lambda_1(t)-\Lambda_1(\tau_2)}(-\log(1-U_1)-\Lambda_1(\tau_2))|\mathcal{G}_{\tau_2},\{-\log(1-U_1)>\Lambda_1(\tau_2)\}]\\ &= \mathbf{1}_{\bar{A}}+\mathbf{1}_{\tau_2<t}\mathbf{1}_{\tau_1\geq\tau_2}\mathbb{E}[F_{\Lambda_1(t)-\Lambda_1(\tau_2)}(-\log(1-U_1)-\Lambda_1(\tau_2))|\mathcal{G}_{\tau_2},U_1>1-\exp(-\Lambda_1(\tau_2)\})]\end{aligned} \qquad \text{(K.0.2)}$$

Let us denote $\bar{U}_1=1-\exp(-\Lambda_1(\tau_2))$. Then we can rewrite Eq. (K.0.2) as

$$\begin{aligned}&\mathbf{1}_{\bar{A}}+\mathbf{1}_{\tau_2<t}\mathbf{1}_{\tau_1\geq\tau_2}\int_0^1 F_{\Lambda_1(t)-\Lambda_1(\tau_2)}(-\log(1-U_1)-\Lambda_1(\tau_2))d\mathbb{Q}(U_1<u_1|\mathcal{G}_{\tau_2},U_1>\bar{U}_1)]=\\ &= \mathbf{1}_{\bar{A}}+\mathbf{1}_{\tau_2<t}\mathbf{1}_{\tau_1\geq\tau_2}\int_0^1 F_{\Lambda_1(t)}(-\log(1-U_1))d\mathbb{Q}(U_1<u_1|\mathcal{G}_{\tau_2},U_1>\bar{U}_1)]\\ &= \mathbf{1}_{\bar{A}}+\mathbf{1}_{\tau_2<t}\mathbf{1}_{\tau_1\geq\tau_2}\int_{\bar{U}_1}^1 F_{\Lambda_1(t)}(-\log(1-U_1))d\mathbb{Q}(U_1<u_1|\mathcal{G}_{\tau_2},U_1>\bar{U}_1)] \qquad \text{(K.0.3)}\end{aligned}$$

We next need to provide a methodology to compute the conditional distribution

$$C_{1|2}(u_1;U_2):=\mathbb{Q}(U_1<u_1|\mathcal{G}_{\tau_2},\{\tau_1>\tau_2\}) \qquad \text{(K.0.4)}$$

Eq. (K.0.4) may be rewritten as

$$
\begin{aligned}
C_{1|2}(u_1; U_2) &= \mathbb{Q}(U_1 < u_1 | U_2, U_1 > \overline{U}_1) \\
&= \frac{\mathbb{Q}(U_1 < u_1 | U_2) - \mathbb{Q}(U_1 < \overline{U}_1 | U_2)}{1 - \mathbb{Q}(U_1 < \overline{U}_1 | U_2)} \\
&= \frac{\frac{\partial C_{1,2}(u_1, U_2)}{\partial u_2} - \frac{\partial C_{1,2}(\overline{U}_1, U_2)}{\partial u_2}}{1 - \frac{\partial C_{1,2}(\overline{U}_1, U_2)}{\partial u_2}}
\end{aligned} \tag{K.0.5}
$$

Notice that U_2 is explicitly known once we condition on $\mathcal{G}_{\tau_2}$ and is given by $1 - \exp(\Lambda_2(\tau_2))$. The expressions in Eq. (K.0.5) may be computed in closed form for the Gaussian copula and are given Schonbucher (2003) by

$$
\frac{\partial C(u, v)}{\partial u} = \Phi \left(\frac{x - r_{1,2} y}{\sqrt{1 - r_{1,2}^2}} \right) \tag{K.0.6}
$$

where $x = \Phi^{-1}(u)$ and $y = \Phi^{-1}(v)$, and $r_{1,2}$ is the correlation between the reference credit and the counterparty. □

A similar methodology can be repeated to compute Eq. (6.3.15) and is not reproduced here.