Université catholique de Louvain

Louvain School of Management
LIDAM & Louvain Finance

Credit Risk Modelling: beyond standard intensity models

Cheikh Mbaye

Thesis submitted in partial fulfilment
of the requirements of the degree of
Docteur en sciences économiques et de gestion

Composition of the jury:
Prof. Frédéric Vrins (UCLouvain), Adviser
Prof. Damiano Brigo (Imperial College London)
Prof. Monique Jeanblanc (Université d'Évry)
Prof. Olivier Le Courtois (EMLyon)
Prof. Wim Schoutens (KULeuven)
Prof. Pierre Devolder (UCLouvain), President

Louvain-la-Neuve, December 2019

To my parents

Abstract

The most extensively studied form of credit risk is the default risk which is the risk that an obligor does not honour his payment obligations. Therefore, the main tool in the area of credit risk modelling is a sound specification of the random time of default. In this thesis, we investigate different ways to model the default time of a given reference entity. We do this by going beyond the existing standard reduced-form models mostly characterized by their weakness to reproduce market credit spreads in terms of volatility featured and perfect calibration. We then address these problems as follows. First, we allow a two-sided jump to the default intensity process which gives an appealing solution but solves partially the calibration problem because it doesn't fully preserve the range of the initial intensity process when forcing a perfect fit to market data. Secondly, we adopt a time change approach allowing for exact calibration to market survival probability curves without affecting the range of the original intensity process. Third, we consider a more general setup where we still preserve a reduced-form setting but under which the default time follows a structural approach with partial information flow. The last essay considers a direct modelling of the conditional survival probability associated to a default time. This approach is a case of "no-immersion" which is more convenient for application to credit derivatives which are characterized by strong wrong-way risk and high credit spreads.

Acknowledgments

It is a very pleasant task to thank, at the end of this work, those without whom it could not have been realized.

I especially thank my PhD supervisor, Prof. Frédéric Vrins, for agreeing to mentor me and guide me throughout my work. For the attention he has kindly guaranteed me, his availability and valuable assistance till the realisation of this thesis. I thank him introducing me to this passionate topic of research.

I would also like to thank Prof. Wim Schoutens and Prof. Olivier Le Courtois that kindly accepted to be part of my panel and dedicated time and energy to go through this thesis. I am very grateful to their useful comments and encouragements during the assessment phases of the thesis.

My special thank goes to Prof. Abass Sagna for having introduced me to the field of Mathematical Finance by supervising my first master thesis in this area. He has provided a valuable support in this thesis by giving me the opportunity for having a research visiting at ENSIIE. I have the privilege of co-authoring a paper with him and his advices helped me expand my research topic and enrich the thesis. I thank Prof. Damien Simon in the same way.

I warmly thank Prof. Monique Jeanblanc for hosting us in the Mathematics department at the Universté d'Évry. I also thank her for stimulating discussions, mainly, at Évry and occasionally when I met her in many conferences in Mathematical Finance. His numerous and valuable comments improve a lot the quality of this thesis. I am very happy that she agreed to be part of the jury.

I thank Prof. Gilles Pagès for his help and encouragements and for earlier insightful discussions about a chapter of this thesis. In the same way, I am grateful to Prof. Donatien Hainaut for having discussed a chapter of the thesis and his useful comments.

I am grateful to Thomas Hvolby for nice discussions in the area of credit risk and for introducing me the Rcpp coding language.

Beside Prof. Frédéric Vrins, Prof. Wim Schoutens, Prof. Olivier Le Courtois and Prof. Monique Jeanblanc, I would like to thank the other jury members of the thesis: Prof. Pierre Devolder and Prof. Damiano Brigo for giving the honour to be part of my committee.

I am specially thankful to all CORE, ISBA and LFIN members and my closest colleagues of LFIN for having made the workplace a wonderful environment and for challenging discussions during the thesis. I take the opportunity to thank my office-mates Younes El Hichou and Nathan Lassance for passing stimulating moments with them and sharing research distress.

My family members and loved ones have been the most ardent supporters, and have also been the main victims of the time I have devoted to the realization of this thesis. I thank you all.

This research was funded by the National Bank of Belgium and UCLouvain via an FSR grant.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Financial mathematics are nowadays tools for both industries and sovereigns, providing judicious and efficient means to assess risks that institutions can take. This is possible because of the success of sophisticated quantitative methodologies in helping professionals manage financial risk. Credit risk modelling is now a key area in financial mathematics and is a well-known risk affecting securities. Therefore, it is important to design reliable default models. Despite the contributions coming from others areas such as economics, our topic will be relevant in solving some of the central questions faced by the risk management of today that became central in light of the 2008 financial crisis. Before the crisis, the big financial institutions were assumed to be free of default risk but, after the consequences of the crisis leading to the collapse of Lehman Brothers, the market needs of modelling default risk have been increasing. In particular, the post-crisis regulation included the modelling of the capital requirement of firms and added some credit adjustment to the valuation under credit risk. One of the main problem faced for modelling such products is the correlation between assets and default known as wrong-way risk which will be developed further and its modelling framework is a central point in this thesis. In addition, the international accounting rules (IFRS[1]) force banks to account for credit derivatives such as default swaps in the trading book and assess their fair value. In summary, prudential regulations impose strict constraints on market players in managing their risks and allocating their own funds. Thus, the assessment of credit risk is a central problem for financial institutions and investors in the defaultable market who must analyze the individual risk of their clients and the overall risk of their loans. This especially leads, in one side, to more complex structured products that need to be carefully accounted in the modelling

[1]International Financial Reporting Standards.

process.

Although, the banking regulation developed a set of measures that aimed at providing more transparency in the risk management of banks, it will have a significant weight on the economy. Indeed, these severe capital rules could force banks to charge more the trading book due to default risk arising from defaultable products that will become more expensive. To fill this gap, regulators recommended a simplification of the global derivatives market that could effectively lead to modelling mismatch and liquidity issues in the class of derivative products subject to default.

However, in financial markets and more specifically in investment banks internal models, the complexity is maintained and further supported by the academics word-class expert in the field of quantitative finance who recommend to consider an advanced approach, allowing institutions to implement the corresponding models the best way they can.

> *"Trying [...] to impose a simplistic and potentially inadequate methodology top-down to all institutions would result in undesirable consequences that are easy to imagine."* The Dilemma of Regulators and Basel III. In Brigo et al., 2013.

Following the same directions, the main purpose of this thesis is to enhance standard intensity-based default models (known as Cox models) by addressing its current major limitations, namely the small "implied credit spread volatility" generated in such a setup and their inability to perfectly fit CDS market data. Finding alternative frameworks leading to larger implied volatilities and exact calibration to CDS quotes is fundamental in order to correctly price credit swap options and credit value adjustment under wrong-way risk.

In particular, an important step is to model the investor's information flow. In credit risk, and particularly in the reduced-form approach a convenient arrangement is to use the theory of *enlargement of filtration*. Throughout the thesis, the investor's information will be modelled through a specific use of progressive enlargement of filtrations: the usual reference filtration which contains the information available in the market is extended to a more general setup by adding the information coming from the default time. Indeed, instead of working on the reference filtration as usual, we will consider a full filtration which contain both the information available in the default-free market and the one coming from to a default event. However, within the approach based on the progressive enlargement of filtration, *immersion property* (which states that the martingales in the reference filtration remain martingales in the full filtration) is fundamental. Indeed, the absence of arbitrage in finance is closely linked to the property of martingale satisfied by the assets. In a Cox framework, the immersion property is satisfied and there is no arbitrage opportunities in this model. But, if immersion is not satisfied, an arbitrage-free model is not guaranteed. In such a case, knowing the properties of the default time is fundamental in order to deal with arbitrages. Both examples of default models including the ones satisfying the immersion property and the ones beyond the immersion setting will be studied in this thesis.

In order to reach the main objectives of the thesis, we have explored four directions.

The first one dealt with the construction of a default intensity model which can capture jumps in credit spreads and prices of credit-sensitive securities and at the same time preserves the correlation between assets and default intensities. Many existing papers are criticized due to their inability to capture possible jumps in credit spreads with possibility to jump back down. But empirical literature shows that the actual credit spreads curves are sometimes downward-sloping. A time-changed intensity model featuring two-sided jump thereby maintaining the initial dependence structure between assets and intensities is discussed in Chapter 3.

The second route is devoted to the so-called calibration problem which is a standard step that needs to be handled in Finance. It consists of turning a model such that it best-fits market quotes at a given time. Brigo et al. already introduced a deterministic shift version in order to solve this problem. Unfortunately, this model is very limited when focusing on positive processes as in a credit risk application. In Chapter 4, we discussed the limitations of using the deterministic shift approach to deal with the modelling of default intensities and proposed an alternative version allowing a perfect fit to any valid CDS market curve.

In the third direction, we investigated the case where investors have no access to the state variable triggering the default but observe a correlated price process, while keeping the structural economic explanation of the default event. This phenomenon is commonly referred to *partial information* and is faced by investors in most practical situations regarding the firm's assets. A modelling approach in such a setup properly implemented in an enlargement of filtration setting is presented in Chapter 5.

The fourth and final direction of the thesis proposed a modelling framework that completely gets out of Cox models. In a technical language, it is a setup where immersion property does not hold. This approach is an extremely interesting and promising research area typically in credit risk, but triggers very challenging mathematical problems. Hence the main objective of this part of the thesis (Chapter 6) is to complete the theoretical study justifying the absence of arbitrage opportunities out of the immersion setup. A central point is indeed to guarantee that the resulting model can be arbitrage-free.

1.1 Structure of the thesis

The structure of the thesis is organized as follows.
Chapter 2 gives the review of the standard models in credit risk which can be divided in two main categories, namely, structural models and reduced-form models. We define a set of their extensions commonly used in the current literature as benchmark models due to their tractability. Furthermore, we recall two classes of credit derivatives that will be used to test the performances of our main results compared to the benchmark models: the credit valuation adjustment (CVA) and the credit default swap (CDS).

Chapter 3 proposes a default model that accounts for both tractability and flexibility. In such models, the survival probability has to be known in closed form (for calibration purposes), the model should be able to fit any valid credit default swap (CDS) curve, should lead to large volatilities (in line with CDS options) and finally should be able to feature significant wrong-way risk (WWR) impact in CVA. In this case, the Cox–Ingersoll–Ross (CIR) model combined with independent positive jumps and deterministic shift (JCIR++) is a very good candidate. In practice however, there is a strong limit on the model parameters that can be chosen, and thus on the resulting WWR impact. This is because only non-negative shifts are allowed for consistency reasons, whereas the upwards jumps of the JCIR++ need to be compensated by a downward shift. To limit this problem, we consider a variation of Cox's model using *jump diffusion* time-changed techniques. It features two-side jumps, thereby pushing upwards the implied volatility on one side and, solving partially the negativity issues on the other side since the resulting intensity process can jump downward. Alternatively, in a multivariate setup like CVA, time-changing the intensity partly kills the potential correlation with the exposure process and destroys WWR impact. It can also introduce a forward looking effect that can lead to arbitrage opportunities. In the same chapter, we use the time change techniques in a way that the above issues are avoided. In addition, we show that the resulting process allows to introduce a large WWR effect compared to the JCIR++ model. Furthermore, we propose a methodology to reduce the computational cost of the resulting Monte Carlo framework by using an adaptive control variate procedure.

Chapter 4 addresses the so-called *calibration problem* which consists of fitting in a tractable way a given model to a specified term structure like, e.g., yield or default probability curves. Time-homogeneous jump-diffusions like Vasicek or Cox-Ingersoll-Ross (possibly coupled with compounded Poisson jumps, JCIR), are tractable processes but have limited flexibility; they fail to replicate actual market curves. The deterministic shift extension of the latter (Hull-White or JCIR++) is a simple but yet efficient solution that is widely used by both academics and practitioners. However, the shift approach is often not appropriate when positivity is required, which is a common constraint when dealing with credit spreads or default intensities. In this chapter, we tackle this problem by adopting a *deterministic* time change approach. By doing so, this chapter completely solves the negativity issues of the shift approach that were partly fixed in Chapter 3. In addition, on the top of providing an elegant solution to the calibration problem under positivity constraint, we show that the model can feature additional interesting properties in terms of implied volatilities. In the context of valuation adjustment under wrong-way risk and credit default swaption, this model offers an appealing alternative to the shift in such cases.

The problem of how to estimate credit risk by assigning default event to a company's assets is the subject of Chapter 5. In that case, it is usually assumed that the firm's value is observable at any time. However, this is a very big assumption since the company's assets is not necessarily a tradable liquid asset on a financial market in continuous time. We use a structural model where a noisy version of the firm value is observed to compute the conditional default probability for a general diffusion, given the noisy observations and the information on the default indicator. This will give us the opportunity to link the structural and reduced-form approaches as the resulting default time is a totally inaccessible stopping time. Furthermore, we propose a numerical procedure based on the recursive quantization method to approximate the conditional default probability for general diffusion of the unobserved process. Finally, we perform numerical tests to evaluate the performance of the method and suggest an application to the pricing of CDS options.

Chapter 6 studies a case of "no-immersion" resulting from a direct modelling of the

Azéma supermatingale without clearly specifying a definition of the default time by using the concept of *conic martingales*. Conic martingales refer to Brownian martingales evolving between bounds. Among other potential applications, they have been suggested for the sake of modelling conditional survival probabilities under partial information, as usual in reduced-form models. Yet, conic martingale default models have a special feature; in contrast to the class of Cox models, they fail to satisfy the immersion property. Hence, it is not clear whether this setup is arbitrage-free or not. In this chapter, we study the relevance of conic martingales-driven default models for practical applications. We first introduce an arbitrage-free conic martingale, the Φ-martingale, by showing that it fits in the class of Dynamized Gaussian copula model of Crépey et al., thereby providing an explicit construction scheme for the default time. In particular, the Φ-martingale features interesting properties inherent on its construction easing the practical implementation. Eventually, we apply this model to CVA pricing under wrong-way risk and CDS options giving us the opportunity to explore the volatility and covariance effects generated in such a beyond immersion setup compared to a Cox setup.

1.2 Published materials

The thesis is composed of four papers split in four different chapters and each paper is either published or submitted.

Chapter 3 is published in Mbaye and Vrins, 2018. A first version was published in the Proceedings of the ICCF 2017.

Chapter 4 is available as a preprint in Mbaye and Vrins, 2019 and has been submitted for publication. An earlier version of the paper has been presented at the 10^{th} World Congress of the Bachelier Finance Society, July 2018, in Dublin and also at the 11^{th} European Summer School in Mathematical Finance, August 2018, in Paris. A recent version has been presented at the Quantitative Finance and Risk Analysis (QFRA) conference, June 2019, in Kos (Greece).

Chapter 5 is a joint work with Abass Sagna[2] & Frédéric Vrins and is submitted.

[2]Fédération de Mathématiques d'Evry, Laboratoire Analyse et Probabilités, 23 Boulevard de France, 91037 Evry, & ENSIIE, e-mail: `abass.sagna@ensiie.fr`. This research benefited from the support of the " Chaire Marchés en Mutation", Fédération Bancaire Française.

Chapter 6 is a joint work with Frédéric Vrins and is submitted.

Chapter 2

Standard credit risk models

In this chapter, we first present the standard (classic) approaches to modelling credit risk. There are two main approaches. The first one known as the *firm-value* or *structural* default model is essentially based on Black and Scholes, 1973 formulas for option pricing. The work by Merton, 1974 is one of the earliest of these models and it builds the basis of credit risk models. There exists other structural models extending and thus eliminating some weaknesses of the Merton model such as *first passage time* models (see for example Black and Cox, 1976; Geske, 1977; Longstaff and Schwartz, 1995). The second approach is known under the name of *reduced-form* or *intensity-based* credit model. It focuses directly on modelling the default probability by means of the intensity process and is based on the work by Duffie and Singleton, 1999. Extended version of this model and details illustrations can be found in Duffie and Singleton, 2003; Bielecki and Rutkowski, 2002 or more recently in Lando, 2004; Brigo and Mercurio, 2006; Jeanblanc et al., 2009; Schoutens and Cariboni, 2009.

After recalling the details of the modelling framework, we give two examples of credit derivatives that we will use to analyse the performances of the chosen credit risk models by their risk-neutral valuation. The pricing methodology will be discussed and a general pricing formula of each chosen asset class will be given for all models. The fact that a pricing formula admits a closed-form or not is strongly linked to the level of tractability of the chosen credit model.

Throughout the thesis, we consider a fixed time horizon T^* and a probability space $(\Omega, \mathcal{G}, \mathbb{Q})$. Our financial market can feature default-free entities and, to ease the exposition, a single credit-risky entity, which default time is modelled by the random time τ. Hence, we deal with two classes of financial instruments:

those which future cashflows (hence prices) are not impacted by the default of the risky reference entity (called *default-free assets* in the sequel), and those who are (*defaultable assets*). To deal with those products, we consider several flows of information, modelled as filtrations satisfying the usual conditions. They will be formally specified for each model below, but they can be intuitively introduced as follows. The full market information is noted $\mathbb{G} = (\mathcal{G}_t, t \in [0,T^*])$ and all risk factors and price processes considered in this thesis are $\mathbb{G}$-adapted. Then, we define the filtration specific to the default event, i.e., the natural filtration of the default indicator $\mathbb{D} = (\mathcal{D}_t, t \in [0,T^*])$, $\mathcal{D}_t = \sigma(\mathbb{1}_{\{\tau\leq u\}}, u \leq t, t \in [0,T^*])$. Eventually, $\mathbb{F} = (\mathcal{F}_t, t \in [0,T^*])$ is a sub-filtration of $\mathbb{G}$ and is referred to as the assets filtration. Notice that $\mathbb{F}$ should not be considered as the information conveyed by the risk factors driving the default-free assets only. Indeed, some processes impacting the default likelihood could be $\mathbb{F}$-adapted, too. Based on the two approaches used to model the default time in credit risk, the default time τ is either a stopping time in the assets filtration $\mathbb{F}$, or is a stopping time in the full filtration $\mathbb{G}$. We assume in the sequel that $\mathcal{G} = \mathcal{G}_{T^*}$ and $\tau > 0$. The default time τ is said to be a *predictable stopping time* if there exists an announcing increasing sequence of stopping times $(\tau_n)_{n\geq 1}$, such that $\lim_{n\to\infty} = \tau$, $\mathbb{Q}$-a.s. A default time which avoids all predictable stopped time is said to be *totally inaccessible*, that is, a stopping time τ such that $\mathbb{Q}(\tau = \bar{\tau} < \infty) = 0$, for any predictable stopping time $\bar{\tau}$. The probability $\mathbb{Q}$ stands for an equivalent martingale probability measure, so that every payoff discounted at the $\mathbb{F}$-adapted risk-free rate r is a $(\mathbb{Q},\mathbb{G})$-martingale. Notice that the discounted prices of default-free assets (which future cashflows do not depend on τ) are $(\mathbb{Q},\mathbb{F})$-martingales as well.

We assume that the bank account numeraire is driven by

$$d\beta_t = r_t\beta_t dt, \quad \beta_0 = 1 \ . \tag{2.1}$$

Thus, the price at time t of a risk-free zero coupon bond maturing at T is

$$P_t(T) := \mathbb{E}\left[e^{-\int_t^T r_s ds} \middle| \mathcal{F}_t \right] \tag{2.2}$$

Eventually, we assume that the market provides us with the survival curve G under the pricing measure, which represents the current $\mathbb{Q}$-distribution that the reference entity survives up to some point in time, i.e.,

$$G(t) = \mathbb{Q}(\tau > t|\mathcal{G}_0) = \mathbb{Q}(\tau > t) \ , \tag{2.3}$$

where we assume that $\mathcal{G}_0$ is trivial. In practice, the G curve is obtained by a bootstrapping procedure, that is, by considering the market prices of defaultable instruments like credit-risky bonds or credit default swaps (CDS), and reverse-engineering the risk-neutral valuation formula. Obviously, the G function must start from 1, remain positive and be decreasing. As standard in the literature and in line with the market practice, we assume that G is differentiable Markit, 2004. Therefore, the market-implied survival probability curve observed at time 0 can be parametrized as

$$G(t) = e^{-\int_0^t h(s)ds} , \tag{2.4}$$

for some positive function h called *hazard rate*[1].

The purpose of a dynamic default model is to generate a set of probability curves at some future time $t > 0$, that is, to model the probability that a default event occurs after a given time $T \in [t, T^*]$ given the information available at time $t \leq T$. Mathematically speaking, the model aims at providing

$$Q_t(T) := \mathbb{Q}(\tau > T | \mathcal{G}_t) . \tag{2.5}$$

These curves are needed for pricing (e.g., options on CDS or credit valuation adjustment) or risk-management purposes. We refer to Cesari et al., 2009 for a couple of examples.

2.1 Structural approach

This section provides a short review of the literature on the structural models and the derivation of the corresponding survival probabilities.

In structural models, default is associated with the value of the firm's financial assets. Usually, the dynamics of the assets' value is given and the default event is defined as a boundary condition on this process, typically the debt book value of the company. Let's denote by X_t the value of the firm at time t, for any $t \leq T^*$ and L its debt value at maturity T.

2.1.1 Merton model

The Merton model is the pioneer of the structural models and relies on the option pricing methodology. Note that in this model, default can only happen at maturity

[1] If G is not differentiable, the instantaneous rate of default h does not exit. This is not compatible with standard market assumptions (for instance, it may lead to assign a positive default probability at some specific times) and thus will not be considered here.

T. By definition, the default time is equal to the maturity if the firm value is below the debt level L and is infinity (that means no default) otherwise. Hence, the default time is given by

$$\tau = T\mathbb{1}_{\{X_T<L\}} + \infty\mathbb{1}_{\{X_T\geq L\}}$$

and the assets value of the firm X is assumed to be described by a diffusion stochastic process following a geometric Brownian motion:

$$dX_t = rX_t dt + \sigma X_t dW_t, \quad X_0 > 0 \tag{2.6}$$

where r is the risk free rate (assumed here to be constant), σ is the volatility parameter and W is a Brownian motion.

The Merton model assumes that $\mathbb{F}$ is the filtration generated by the Brownian motion W and $\mathbb{G}$ coincides with $\mathbb{F}$.

The value of the firm X is also given, according to the general accounting equation, by

$$X_t = D(t,T) + E_t$$

where $D(t,T)$ is the value at time t of a zero coupon bond with maturity T and E_t represents the equity price at time t. At maturity T, the outstanding debt, the firm value and the debt at maturity are related by

$$D(t,T) = \min(L, X_T) = L - (L - X_T)^+$$

so that

$$E_T = (X_T - L)^+ \ .$$

Therefore, the equity is the price of an option written on the asset of the firm with strike price equal to the face value L of liabilities: the face value of a zero coupon bond. From Black-Scholes formula, the price of such an option is given by

$$E_t = X_t\Phi(d_1) - Le^{-r(T-t)}\Phi(d_2)$$

where

$$d_1 = \frac{\log(X_t/L) + \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = d_1 - \sigma\sqrt{T-t} \tag{2.7}$$

and

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{u^2}{2}}du \tag{2.8}$$

is the cumulative standard normal distribution.

In this case, the survival probability (2.5) under the Merton model is then equal to

$$Q_t(T) = \mathbb{Q}(X_T \geq L|\mathcal{G}_t) = \Phi(d_2)\ . \tag{2.9}$$

An important limitation of the Merton model is the assumption that a firm can only default at maturity. This problem is resolved by the first-passage time models such as the Black and Cox model that we introduce in the following section.

2.1.2 Black and Cox model

As recalled above, Merton's model does not allow for a premature default. Black and Cox, 1976 extend Merton's model to the case where safety covenants provide the firm's bondholders with the right to force the firm into bankruptcy and obtain the ownership of the assets. They postulate that as soon as the firm's asset cross a lower threshold, the bondholders take over the firm. Although Black and Cox considered a time-dependent barrier, we assume a simplified model with a constant barrier $\mathbf{a}$. Therefore, the model assumes that the firm defaults when its value falls below a pre-specified level, meaning that the default time is defined as

$$\tau := \inf\{u \geq 0 : X_u \leq \mathbf{a}\}\,,\quad 0 < \mathbf{a} < X_0 \tag{2.10}$$

where $\inf \emptyset := +\infty$, as usual.

In the Black and Cox model, the default time τ is a stopping time in the assets filtration $\mathbb{F}$ and $\mathbb{G}$ coincides with $\mathbb{F}$.

In this case, the survival probability (2.5) in this first-passage time model is given by

$$Q_t(T) = \Phi(h_1(X_t, T-t)) - \left(\frac{\mathbf{a}}{X_t}\right)^{\frac{2r}{\sigma^2}-1} \Phi(h_2(X_t, T-t)) \tag{2.11}$$

where h_1 and h_2 are given by

$$\begin{aligned} h_1(x,u) &= \frac{1}{\sigma\sqrt{u}}\left(\log\left(\frac{x}{\mathbf{a}}\right) + \left(r - \frac{1}{2}\sigma^2\right)u\right), \\ h_2(x,u) &= \frac{1}{\sigma\sqrt{u}}\left(\log\left(\frac{\mathbf{a}}{x}\right) + \left(r - \frac{1}{2}\sigma^2\right)u\right)\ . \end{aligned} \tag{2.12}$$

Notice that the model can be extended in a more general setting by considering a time-dependent or a stochastic random barrier but this makes harder the task of finding the law of the corresponding first passage time models. Furthermore, the analytical complexity of the model is increased when considering a stochastic interest rates. This theoretical complexity makes difficult to obtain closed form expressions of the firm's capital structure components or for the survival probability, leading to the use of numerical procedures.

Most structural models and specially the Black and Cox model suffer from the drawback that they fail to reproduce spread consistent with empirical observations. They reproduce lower spreads that decrease to zero for short maturities. This limitation comes from the fact that the valuation takes place in the filtration of the firm's assets, where the default event is a predictable stopping time. Being predictable, short-term default risk is not priced by the models. The predictability issue were addressed by, among others, Brigo and Morini, 2006 and Brigo and Tarenghi, 2005 who included random default barriers and time dependent volatilities. The same problem were also fixed, for example, in Schonbucher, 1996; Zhou, 2001 where the authors introduce unpredictable jumps into the dynamics of the firm's value process. Chen and Kou, 2009 introduced a structural default model by incorporating a two-sided jump process in the firm dynamics leading to a variety of shapes for the implied volatility of equity options with a decent calibration performance to CDS spreads. It follows that, in these models, the default time is no longer a predictable stopping time. Recently, Scherer and Huttner, 2016 applied the default model introduced by Chen and Kou, 2009 to the valuation of CDS options and related optionalities such as extension risk using an efficient Monte Carlo algorithm based on Brownian bridges. Ballotta and Fusai, 2015 presented a multivariate version of a structural default model with jumps using Lévy processes in order to evaluate the bilateral credit value adjustment in the presence of wrong-way risk and the bilateral debt value adjustment for equity contracts. In a more recent paper, Ballotta et al., 2019 proposed an integrated pricing framework for credit value adjustment for equity and commodity products by taking into account the inclusion of risk mitigations clauses such as netting, collateral and initial margin provisions.

To summarize, in all these structural models, the default is explained by economic variables related the firm's asset process and the default event is defined in terms of a boundary condition on this process. Nevertheless, the default should be announced by exogenous state variables of the economy and it is unclear that a firm's asset value is such a variable. Furthermore, in practice, the firm's asset value is not a tradable security. In order to circumvent this issue, it is possible to rely on a modified structural model with partial information (which is the subject of Chapter 5, see also Duffie and Lando, 2001 or Coculescu et al., 2008) or simply consider a reduced-form approach that we now introduce.

2.2 Reduced-form approach

The reduced-form approach, usually referred to as the *intensity* approach is a second family of credit risk models beside the structural models. Reduced-form models focus directly on modelling the default probability by means of the intensity process. The basic idea behind these models is that there is a possibility of default for an obligor at any time, in contrast with the Merton model. In such a setup, the filtration $\mathbb{G}$ is obtained by progressively enlarging $\mathbb{F}$ with $\mathbb{D}$: $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{D}_t$ for any $t \geq 0$. We shall write

$$\mathbb{G} = \mathbb{F} \vee \mathbb{D} \, , \tag{2.13}$$

where $\mathbb{F}$ is given. The important thing is that, given $\mathcal{F}_t$, it should not be possible to determine whether the default event took already place or not.
Hence τ is a $\mathbb{D}$- and a $\mathbb{G}$-stopping time, but not necessarily an $\mathbb{F}$-stopping time.
Furthermore, we assume that the default admits an $\mathbb{F}$-adapted intensity λ_t, for any $t \leq T^\star$ [2].
Special attention is paid here to the $\mathcal{H}$-*hypothesis*, which postulates the *immersion property* of $\mathbb{F}$ in $\mathbb{G}$.

Definition 1 (Immersion property)**.** *The filtration* $\mathbb{F}$ *is said to be immersed in* $\mathbb{G}$ *if any* $(\mathbb{Q}, \mathbb{F})$*-martingale is a* $(\mathbb{Q}, \mathbb{G})$*-martingale.*

This property is very fundamental since the arbitrage pricing theory is strongly linked by the property of martingale satisfied by the assets. Because if the $\mathbb{F}$-market is arbitrage-free, the no-arbitrage condition is guaranteed in the $\mathbb{G}$-market model if immersion holds[3]. For example, the Cox model that we will introduce in Section 2.2.2 is a setup where the immersion property is satisfied (see for example Jeanblanc and Le Cam, 2009b). A “no-immersion” case will be provided in Chapter 6.

[2] Suppose that there exists Λ, $\mathbb{F}$-predictable such that $\mathbb{1}_{\{\tau \leq t\}} - \Lambda_{t \wedge \tau}$ is a $\mathbb{G}$-martingale. If Λ is absolutely continuous (i.e, $\Lambda_t = \int_0^t \lambda_s ds$), the intensity is λ.

[3] It can be shown that the immersion property is necessarily satisfied under the assumption that the $\mathbb{F}$-market model is complete and arbitrage-free, and the extended $\mathbb{G}$-market model is arbitrage-free (for details, see Blanchet-Scalliet and Jeanblanc, 2004 or Jeanblanc and Le Cam, 2009a)

For more details about the immersion property and the theory of enlargement of filtrations, we refer the reader to Elliott et al., 2000; Aksamit et al., 2014; Aksamit and Jeanblanc, 2017.
Before defining the Cox model, we recall some properties of the survival probabilities in a reduced-form setup.

2.2.1 Conditional survival probabilities

Although prices are given by considering the full market information that is, by computing $\mathbb{G}$-conditional expectations, the considered setup allows us to work in the sub-filtration $\mathbb{F}$ thanks to the *Key lemma*. This fundamental theorem allows one to write the $\mathcal{G}_s$-conditional expectation of $X\mathbb{1}_{\{\tau>t\}}$ as the $\mathcal{F}_s$-conditional expectation of $Xe^{-\int_s^t \lambda_u du}$, rescaled by $\mathbb{1}_{\{\tau>s\}}$, for every $\mathbb{Q}$-integrable and $\mathcal{F}_t$-measurable random variable X, $s \leq t$. It is originally due to Dellacherie and Meyer, 1980, although its use in financial applications have been put forward by Bielecki and Rutkowski, 2002 (see, e.g., Bielecki et al., 2011 for numerous examples in credit risk and Brigo and Vrins, 2018 for a specific application in counterparty credit risk). More explicitly, the $\mathcal{G}_t$-conditional probability can be written as a ratio of $\mathcal{F}_t$-probabilities, scaled by a survival indicator Bielecki et al., 2011, Lemma 3.2.1.

$$\mathbb{Q}(\tau > T|\mathcal{G}_t) = \mathbb{1}_{\{\tau>t\}} \frac{\mathbb{Q}(\tau > T|\mathcal{F}_t)}{\mathbb{Q}(\tau > t|\mathcal{F}_t)} \,. \tag{2.14}$$

We denote by $S_t = \mathbb{Q}(\tau > t|\mathcal{F}_t)$ the $\mathbb{F}$-*survival process* of τ, also known as the *Azéma supermartingale* in the probability literature. The following lemma holds.

Lemma 1. *Let Z be an $\mathbb{F}$-predictable process such that the random variable $Z_\tau \mathbb{1}_{\{\tau\leq T\}}$ is $\mathbb{Q}$-integrable. Then we have, for every $t \leq T$,*

$$\mathbb{1}_{\{\tau>t\}}\mathbb{E}(Z_\tau \mathbb{1}_{\{\tau\leq T\}}|\mathcal{G}_t) = -\mathbb{1}_{\{\tau>t\}} S_t^{-1} \mathbb{E}\left(\int_t^T Z_u dS_u \mid \mathcal{F}_t \right) \tag{2.15}$$

Proof. See Proposition 5.1.1. in Bielecki and Rutkowski, 2002. □

Introducing the following notation for the $\mathcal{F}_t$-conditional survival probability curve

$$S_t(T) := \mathbb{Q}(\tau > T|\mathcal{F}_t) = \mathbb{E}\left[\mathbb{1}_{\{\tau>T\}}|\mathcal{F}_t\right] \,, \tag{2.16}$$

one gets that the $\mathcal{G}_t$-risk-neutral probability of the event $\{\tau > T\}$, $T \geq t$, can be written as

$$Q_t(T) := \mathbb{1}_{\{\tau>t\}} \frac{S_t(T)}{S_t(t)} \,. \tag{2.17}$$

We set

$$\overline{Q}_t(T) := \frac{S_t(T)}{S_t(t)} \ . \tag{2.18}$$

Interestingly, for every $T \in [0, T^*]$, $(S_t(T)\ ,\ t \in [0, T])$ is a $(\mathbb{Q}, \mathbb{F})$-martingale valued in $[0, 1]$. By contrast,

$$S_t := S_t(t) = \mathbb{E}\left[\mathbb{1}_{\{\tau > t\}} | \mathcal{F}_t\right] \tag{2.19}$$

is also valued in $[0, 1]$, but is a $(\mathbb{Q}, \mathbb{F})$-supermartingale. Indeed, from the tower law, we have for $s \geq t$,

$$\mathbb{E}[S_s|\mathcal{F}_t] = \mathbb{E}\left[[\mathbb{1}_{\{\tau>s\}}|\mathcal{F}_s]|\mathcal{F}_t\right] = \mathbb{E}[\mathbb{1}_{\{\tau>s\}}|\mathcal{F}_t] = \mathbb{Q}(\tau > s|\mathcal{F}_t) \leq \mathbb{Q}(\tau > t|\mathcal{F}_t) = S_t \ ,$$

since $\{\omega \in \Omega : \tau(\omega) > s\} \subseteq \{\omega \in \Omega : \tau(\omega) > t\}$. Clearly, $S_0 = S_0(0) = 1$ from (2.19) because $\tau > 0$ and, from (2.17), $Q_0(T) = S_0(T) = G(T)$. Notice that from the tower law again, the expectation of the survival process at time T is nothing but the probability that the reference entity survives up to T, as seen from time $t = 0$:

$$\mathbb{E}[S_T] = \mathbb{E}\left[\mathbb{E}[\mathbb{1}_{\{\tau>T\}}|\mathcal{F}_T]\right] = \mathbb{E}\left[\mathbb{1}_{\{\tau>T\}}\right] = \mathbb{Q}(\tau > T) = G(T) \ .$$

One can consider this expression as a kind of constraint (or calibration) that the default model must satisfy at time 0. In the general case, note that not all models generate a curve $\mathbb{E}[S_.]$ compatible with the form of $G(\cdot)$ given in (2.3). It is however the case for all models such that τ is a continuous random variable satisfying $\mathbb{Q}(\tau \leq T^*) < 1$. In this case, τ admits a density, and if G admits a derivative α, $\mathbb{E}[S_t] = \mathbb{Q}(\tau > t) = 1 - \int_0^t \alpha(s)ds > 0$ for all $t \in [0, T^*]$. This expectation can be written as in (2.4) provided that $h(t) := \frac{\alpha(t)}{1-\int_0^t \alpha(s)ds}$ (see for example Chapter 4 in Bielecki and Rutkowski, 2002).

Clearly, the process S is a bounded $\mathbb{G}$-supermartingale. A fundamental result from stochastic calculus stipulates that any càdlàg survival process S admits a unique Doob-Meyer decomposition (Dellacherie and Meyer, 1975; Bielecki et al., 2011)

$$S_t = A_t + M_t \ , \tag{2.20}$$

where M is a $(\mathbb{Q}, \mathbb{F})$-martingale and A is an $\mathbb{F}$-predictable decreasing process, satisfying $M_0 = 0$ and $A_0 = 1$. If S is continuous, then so are A and M provided that the $\mathbb{F}$-martingales are continuous or τ avoids $\mathbb{F}$-stopping times[4]. In addition if A is absolutely continuous with respect to the Lebesgue measure, $dA_t = -\mu_t dt$ and $dM_t = \sigma_t dB_t$ where μ is a positive process, μ, σ are $\mathbb{F}$-adapted and B a $(\mathbb{Q}, \mathbb{F})$-Brownian motion (see Bielecki et al., 2008). Hence, whenever $S_t > 0$ for every $t \in]0, T^*]$, then the dynamics of the survival process can be written as

$$dS_t = -\lambda_t S_t dt + \sigma_t dB_t, \quad S_0 = 1. \tag{2.21}$$

In this case, the compensated process

$$\mathbb{1}_{\{\tau \leq t\}} - \int_0^{t \wedge \tau} \lambda_s ds = D_t - \int_0^{t \wedge \tau} \lambda_s ds$$

follows a $\mathbb{G}$-martingale and $\lambda_t := \mu_t / S_t$ is an $\mathbb{F}$-progressively measurable non-negative process called the *default intensity* with respect to $\mathbb{F}$.

Remark 1. *Note that the above results can be extended to the case where S is not assumed to be continuous, under the assumption that τ avoids the $\mathbb{F}$-stopping times and that A is a continuous process. The continuity of A is required to obtain a continuous compensator of D, which means, to work with a totally inaccessible $\mathbb{G}$-stopping time τ. The hypothesis on continuity of S is also not needed under the assumption that A is absolutely continuous with respect to the Lebesgue measure (see Bielecki et al., 2008 for more details or Chapter 6 in Bielecki and Rutkowski, 2002).*

Remark 2. *It is common to expect the survival process S to be decreasing. This is a feature that is indeed met in the* usual *default models. Surprisingly or not, this is just a special case: it is clear from* (2.21) *that S is decreasing if and only if $M \equiv 0$, i.e., $\sigma \equiv 0$. This behavior, admittedly weird at first sight, will be discussed later. At this stage, observe that $S_t(T)$ in* (2.16) *is decreasing with respect to T for $T \in [t, T^*]$ but is a $(\mathbb{Q}, \mathbb{F})$-martingale. In particular, it does* not *decrease with t. This behavior simply results from the fact that we are computing probabilities under partial information and, in such circumstances, one is allowed to change her mind about past events, at least as long as those events remain* unobserved *(i.e., as long as they are not measurable).*

[4]The default time τ avoids $\mathbb{F}$-stopping times if and only if $\mathbb{Q}(\tau = \xi) = 0$ for every $\mathbb{F}$-stopping time ξ.

2.2.2 The Cox construction

The *Cox setup* is the most popular approach for dynamic intensity-based (reduced form) modelling. It is originally due to Lando, 1998 and Duffie and Singleton, 1999. We refer to Brigo and Mercurio, 2006 for extensive applications in credit risk. In contrast with the firm-value (from which default occurs when the firm's asset breaches a default barrier, a.k.a. Merton or structural models, Merton, 1974) the default event is triggered by the first jump of a Poisson process with stochastic intensity λ. Equivalently, τ can be modelled as the first passage of $\int_0^\cdot \lambda_s ds$ above a random threshold $\mathcal{E}$:

$$\tau := \inf\{t \geq 0 : \Lambda_t \geq \mathcal{E}\}, \quad \Lambda_t := \int_0^t \lambda_s ds \; . \tag{2.22}$$

In this model, λ is a non-negative, $\mathbb{F}$-adapted process and $\mathcal{E}$ is a random variable with unit exponential distribution, independent from $\mathcal{F}_{T^*}$. Hence, one can choose as $\mathbb{F}$ the natural filtration of λ (possibly enlarged with the factors impacting the default-free assets), $\mathbb{D}$ collapses to the filtration generated by the pair $(\lambda, \mathcal{E})$.

Because λ is positive $\mathbb{Q}$-a.s., Λ is increasing, hence, the survival process S defined in (2.19) reduces to

$$S_t = \mathbb{Q}(\tau > t|\mathcal{F}_t) = \mathbb{Q}(\Lambda_t \leq \mathcal{E}|\mathcal{F}_t) = e^{-\Lambda_t} = e^{-\int_0^t \lambda_s ds} \; . \tag{2.23}$$

The dynamics of the survival process are given by

$$dS_t = -\lambda_t S_t dt \; , \tag{2.24}$$

showing that λ in (2.22) actually corresponds to the default intensity introduced in (2.21). Moreover, the Cox setup is a case where immersion holds which corresponds to a very special Doob-Meyer decomposition: the survival process S is decreasing and has no martingale part ($M \equiv 0$).

It is also easy to compute the conditional survival probabilities under both filtrations. For instance, the $\mathcal{F}_t$-conditional survival probability that $\tau > T$ is obviously a $(\mathbb{Q}, \mathbb{F})$-martingale on $[0, T]$, and reads as

$$S_t(T) = \mathbb{Q}(\tau > T|\mathcal{F}_t) = \mathbb{Q}(\Lambda_T \leq \mathcal{E}|\mathcal{F}_t) = e^{-\Lambda_t}\mathbb{E}\left[e^{-\int_t^T \lambda_s ds}\middle|\mathcal{F}_t\right] \; . \tag{2.25}$$

The corresponding $\mathcal{G}_t$-conditional survival probability is, from (2.17), given by

$$Q_t(T) = \mathbb{Q}(\tau > T|\mathcal{G}_t) = \mathbb{1}_{\{\tau>t\}}e^{\Lambda_t}\mathbb{E}\left[e^{-\Lambda_T}\middle|\mathcal{F}_t\right] = \mathbb{1}_{\{\tau>t\}}\mathbb{E}\left[e^{-\int_t^T \lambda_s ds}\middle|\mathcal{F}_t\right] \; . \tag{2.26}$$

2.3 Shifted homogeneous affine models

In this section, we recall the definitions of some tractable models such as affine term structure models and the deterministic shift approach. We expose the definitions in a general case of interest rates and default intensities term structure, since the models remain valid in both cases.

2.3.1 Affine processes and affine jump-diffusions

Affine term structure models (ATSM) models are widely used in finance because they offer an appealing modelling framework : they are scarce, and empirical evidences suggest that they depict relatively well the market dynamics. In their simplest version, they feature a single factor, but empirical evidences suggest that they better perform when several factors are considered. Three factors seem to be a good compromise Dai and Singleton, 2000. The problems discussed in this thesis rely on the specific structure of discount curves that are generated by ATSM models, which is independent of the number of factors involved. Hence, we restrict ourselves to consider single-factor models. Affine models are characterized as follows Filipovic, 2005.

Definition 2 (Affine process). *An affine process is any process $\mathbb{F}$-adapted y satisfying for $s \leq t$*

$$P_s^y(t) := \mathbb{E}\left[e^{-\int_s^t y_u du} \middle| \mathcal{F}_s \right] = e^{A_s^y(t) - B_s^y(t) y_s} \tag{2.27}$$

where A_s^y, B_s^y are differentiable functions satisfying $A_s^y(s) = B_s^y(s) = 0$.

Provided that the A, B functions are known, the analytical form (2.27) facilitates in a tremendous way the perfect fit problem when the considered $P_s^y(t)$ function takes the form of the conditional expectation in (2.27), as illustrated on the risk-free and risky discounting applications. This explains why such models are so popular in term-structure modelling.

It is known (see, e.g., Brigo and Mercurio, 2006 and Duffie and Kan, 1996) that every diffusion with affine drift and diffusion coefficients, regular enough so that a solution exists, is an affine process. Similarly, every jump-diffusion with such types of drift and variance coefficients and independent compound Poisson jumps (i.e., exponentially-distributed jumps arriving according to a Poisson process) is also affine.

Definition 3 (Affine jump-diffusions, AJD). *A stochastic process y is called an affine jump-diffusion if its dynamics take the form*

$$dy_t = (a(t) + b(t)y_t)dt + \sqrt{c(t) + d(t)y_t}dW_t + dJ_t \tag{2.28}$$

with W an $\mathbb{F}$-Brownian motion and J a compound Poisson process independent from W, defined according to $J_t := \sum_{i=1}^{N_t} \zeta_i$ where J is an $\mathbb{F}$-adapted Poisson process with instantaneous jump rate $\omega(t) \geq 0$ and ζ_i's are i.i.d. exponentially distributed random variables with mean $\alpha \geq 0$. In the special case where the parameters $(a, b, c, d, \alpha, \omega)$ are constant, y is said time-homogeneous, or simply homogeneous, or HAJD.

As explained above, affine models are specifically relevant in our context when A, B are known in closed form. This is the case for HAJD. Three important homogeneous cases are the Ornstein-Uhlenbeck, the square-root diffusion and the square-root jump-diffusion. The first one, widely known as the Vasicek (VAS) model, corresponds to the special case where $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\varrho, -\kappa, \eta^2, 0, 0, 0)$ and $y_0 \in \mathbb{R}$. The second model is the Cox-Ingersoll-Ross (CIR), and is associated to

$$(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\varrho, -\kappa, 0, \delta^2, 0, 0),$$

with $y_0, \varrho > 0$. Eventually, the JCIR is an extension of the CIR, associated with parameters $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\varrho, -\kappa, 0, \delta^2, \alpha, \omega)$. The speed of mean-reversion κ is assumed to be positive in all models. In contrast to VAS which is a Gaussian model, the CIR and JCIR models are non-negative.

2.3.2 A deterministic shift extension

The starting point is to notice that the limited capacities of homogeneous models result from their rigid parametric form. Therefore, an interesting route is to consider a family of models x defined as time-dependent transform of a base HAJD model y in such a way that the model's tractability is not affected.

In this section, we recall the general deterministic shift extension approach. The latter has been introduced in the seminal paper Brigo and Merccurio, 2001 in order, precisely, to perfectly fit any given term structure of interest rates or hazard rates, typically a market curve $P^{market}(\cdot)$. In this model, $x := x^\varphi$ is defined as a HAJD (y) that is shifted in a time-dependent way using a deterministic function φ:

$$x_t^\varphi := y_t + \varphi(t) \ . \tag{2.29}$$

Interestingly, $P^{\varphi}(t) := P_0^{x^{\varphi}}(t)$ where x^{φ} remains affine (although no longer homogeneous) is hence analytically tractable in terms of calibration since

$$P^{x^{\varphi}}(t) = e^{A^{x^{\varphi}}(t) - B^{x^{\varphi}}(t) x_0^{\varphi}} \tag{2.30}$$

with

$$\begin{aligned} A^{x^{\varphi}}(t) &= A^{y}(t) - \int_0^t \varphi(u) du + B^{y}(t)\varphi(0) \ , \\ B^{x^{\varphi}}(t) &= B^{y}(t) \ . \end{aligned}$$

In this case, the function φ is uniquely determined by solving the calibration equation

$$P^{\varphi}(t) = P^{market}(t)$$

yielding

$$\varphi(t) = -\frac{d}{dt} \ln \frac{P^{market}(t)}{P^{y}(t)} \ . \tag{2.31}$$

2.4 Credit derivatives instrument

A credit derivative is a contract whose underlying asset is a debt instrument. It is a flow exchange agreement between two counterparties that does not directly involve the issuer of the underling debt. It reduces or increases exposure to bonds issued by sovereign or corporate entities and is generally used (i) to express positive or negative expectations about the evolution of the credit quality of a single issuer or a portfolio of issuers, regardless of the exposure that the investor may have (ii) to hedge the credit risk of a portfolio of bonds or loans.

Credit derivatives are currently the key instruments in the trading of credit risk and in the risk management of financial and corporate institutions.

In this section, we derive the pricing methodology of two classes of credit derivatives, mainly, the credit default swaps and the counterparty credit value adjustment.

2.4.1 CDS and CDSO

A credit default swap (CDS) is a financial instrument used by two parties – called the protection buyer and the protection seller – to transfer to the protection seller the financial loss that the protection buyer would suffer if a particular default event happened to a third party called the *reference entity*. Typically, we set τ as the default time of the latter. In a default swap contracted at time t, started at time T_a with maturity T_b, the protection buyer pays a coupon (or spread) k at a set of payment dates $T_{a+1}, \dots, T_b$ as long as the reference entity does not default. The protection seller agrees to make a single payment LGD to the protection buyer if the default occurs between T_a and T_b. When applicable, the protection buyer makes a final payment corresponding to the spread accrued since the last payment date before default. For more details about the mechanics of this product, we refer to Brigo and Alfonsi, 2005 and Brigo and El-Bachir, 2010.[5]

It is well-admitted that "pure credit instruments" like CDS or CDS option (CDSO) are quite insensitive to the stochasticity of the interest rates in realistic conditions. This has been discussed explicitly for the CIR base model in Brigo and Alfonsi, 2005 and Brigo and Cousot, 2006. Hence, we consider a deterministic short rate process in the following CDS formulas, which is stressed by the notation $r_u = r(u)$. In this case, one simply gets $P_t(T) = e^{-\int_t^T r(u)du} = \beta_t/\beta_T$, where β is given in (2.1).

In the sequel, we deal with the pricing of a CDSO. Because CDSO is an option on CDS, we start by recalling the no-arbitrage pricing equation of a CDS. We note t the valuation time and work on the set $\tau > t$ as it is pointless to price a CDS (or a CDSO) post-default. From the perspective of the protection buyer, the time-t value of a 1 dollar notional CDS $CDS_t(a, b, k)$ starting at time T_a with maturity T_b, $t \leq T_a < T_b$, a spread k and (known) loss given default $LGD = (1-R)$ is given by the difference of the conditional risk-neutral expectations of the protection and premium cashflows (see Brigo and Alfonsi, 2005):

$$\begin{aligned} CDS_t(a,b,k) &= \mathbb{E}\left[(1-R)\mathbb{1}_{\{T_a \leq \tau \leq T_b\}} P_t(\tau) | \mathcal{G}_t\right] \\ &\quad -k\,\mathbb{E}\left[\sum_{i=a+1}^{b}\left(\mathbb{1}_{\{\tau \geq T_i\}}\alpha_i P_t(T_i) + \mathbb{1}_{\{T_{i-1} \leq \tau < T_i\}}\alpha_i \frac{\tau - T_{i-1}}{T_i - T_{i-1}} P_t(\tau)\right)\middle| \mathcal{G}_t\right] \end{aligned}$$

with α_i the day count fraction between dates T_{i-1} and T_i which, in a standard CDS, is around 0.25 (quarterly payment dates), R is the recovery rate assumed to be constant[6] and $P_t(\cdot)$ is given in (2.2). In an enlargement of filtration setup, this

[5] For more details about the actual market conventions, we refer the interested reader to Markit, 2004 and Markit, March 13, 2009.

[6] We assume here the recovery rate R to be constant for a sake of simplicity since we focus more on the credit risk impact implied by the definition of the random time of default. Nevertheless, R is the value at time τ of the cheapest bond that is deliverable under the terms of the contract and thus may be time-dependent or even stochastic, since the delivery option is explicitly recognized.

expression can be developed explicitly thanks to Lemma 1:

$$CDS_t(a,b,k) = \mathbb{1}_{\{\tau>t\}} \left(-(1-R) \int_{T_a}^{T_b} P_t(u) \partial_u \overline{Q}_t(u) du - k \; C_t(a,b) \right) \; , \qquad (2.32)$$

where $C_t(a,b)$ is the risky duration, i.e., the time-t value of the CDS premia paid during the life of the contract when the spread is 1:

$$C_t(a,b) := \sum_{i=a+1}^{b} \alpha_i P_t(T_i) \overline{Q}_t(T_i) - \int_{T_{i-1}}^{T_i} \frac{u - T_{i-1}}{T_i - T_{i-1}} \alpha_i P_t(u) \partial_u \overline{Q}_t(u) du$$

provided that $\overline{Q}$, given in (2.18) is differentiable.

The spread which, at time t, sets the forward start CDS at 0, called *par spread*, is given by:

$$\mathbb{1}_{\{\tau>t\}} s_t(a,b) := \mathbb{1}_{\{\tau>t\}} \frac{-(1-R) \int_{T_a}^{T_b} P_t(u) \partial_u \overline{Q}_t(u) du}{C_t(a,b)} \; . \qquad (2.33)$$

The no-arbitrage price of a call option on such a contract at time $t = 0$ becomes

$$\begin{aligned} PSO(a,b,k) &= \mathbb{E}\left[\beta_{T_a}^{-1} (CDS_{T_a}(a,b,k))^+ \right] \qquad (2.34) \\ &= \beta_{T_a}^{-1} \mathbb{E}\left[S_{T_a} \left((1-R) - \sum_{i=a+1}^{b} \int_{T_{i-1}}^{T_i} g_i(u) P_{T_a}(u) \overline{Q}_{T_a}(u) du \right)^+ \right] \; , \end{aligned}$$

where $g_i(u) := (1-R)(r(u) + \delta_{T_b}(u)) + k \frac{\alpha_i}{T_i - T_{i-1}} (1 - (u - T_{i-1}) r(u))$, with $\delta_s(u)$ the Dirac delta function centered at s, β is given (2.1) and S in (2.19). We refer to Brigo and El-Bachir, 2010 for a detailed proof of the CDS options formula (2.34).

Such kind of options have little liquidity. Models are then often compared in terms of their capabilities to generate large “implied volatilities”. Indeed, empirical evidences show that this is a typical feature of CDS option quotes, when disclosed. Therefore, we compare the models in terms of their “Black volatilities”: the volatility that one needs to plug in a “Black-Scholes” type of model to reproduce the model prices.

In this context, the Black-Scholes model works as follows. We start by noting that the forward start CDS can be written in terms of the difference between the fair and the agreed premium cashflows. Indeed, the former corresponds to the protection leg. Inserting (2.33) in (2.32) yields

$$CDS_t(a,b,k) = \mathbb{1}_{\{\tau>t\}} (s_t(a,b) - k) \; C_t(a,b) \; . \qquad (2.35)$$

Equation (2.35) suggests to use $C(a, b)$ as a numeraire in order to price CDS options. Unfortunately, a direct application of this result is not possible since $C(a, b)$ can be zero and thus cannot be defined as a numeraire which has to be strictly positive. In order to solve this problem, an important contribution is given in Schonbucher, 2003 using survival measures when working in the filtration $\mathbb{F}$. However, the model allows for Black and Scholes formulas but differs from standard market models, since the survival measures are not equivalent to the pricing measure $\mathbb{Q}$. Fortunately, a standard CDS market model under which the CDS spreads are consistent with simple market payoffs is proposed in Brigo and Morini, 2005 when working under the filtration $\mathbb{G}$. The authors introduced a subfiltration structure allowing all measures to be equivalent to the pricing measure and the chosen numeraire $C_t(a, b)$ is strictly positive provided that the Azéma supermartingale satisfies $S_t > 0$. Thus, the par spread $s_t(a, b)$ in (2.33) is well defined and positive.

Therefore, using the results based on Brigo and Morini, 2005, the price of the Payer default swaption becomes

$$PSO(a,b,k) = \mathbb{E}\left[(s_{T_a}(a,b)-k)^+ \; C_{T_a}(a,b)\beta_{T_a}^{-1}\right] = C_0(a,b)\mathbb{E}^{(a,b)}\left[(s_{T_a}(a,b)-k)^+\right] \tag{2.36}$$

where $\mathbb{E}^{(a,b)}$ stands for the expectation under the equivalent measure $\mathbb{Q}^{(a,b)}$, associated with the numeraire $C(a, b)$. Interestingly, it is clear from (2.33) that the par spread $s(a, b)$ is a $\mathbb{Q}^{(a,b)}$-martingale on $[0, T_a]$. Hence, the Black-Scholes model for CDSO naturally postulates $\mathbb{Q}^{(a,b)}$-martingale dynamics for the par spread:

$$ds_t(a,b) = \bar{\sigma} s_t(a,b) dW_t^s, \quad t \leq T_a$$

where W^s is a $\mathbb{Q}^{(a,b)}$-Brownian motion. Eventually, the expectation in (2.36) is given by the standard Black-Scholes formula by setting $r = 0$. Hence, the Black-Scholes price of the PSO is given by

$$PSO^{Black}(a,b,k,\bar{\sigma}) = C_0(a,b)\left[s_0(a,b)\Phi(d_3) - k\Phi(d_4)\right]$$

where

$$d_3 = \frac{\ln \frac{s_0(a,b)}{k} + \frac{1}{2}\bar{\sigma}^2 T_a}{\bar{\sigma}\sqrt{T_a}}, \qquad d_4 = d_3 - \bar{\sigma}\sqrt{T_a}$$

and Φ is given in (2.8).

The Black volatility associated to a model price $PSO^{model}(a, b, k)$ is thus the volatility $\bar{\sigma}$ satisfying $PSO^{model}(a, b, k) = PSO^{Black}(a, b, k, \bar{\sigma})$.

Note that all models need to be calibrated to the market curve G in (2.4), and thus have to satisfy the calibration equation

$$Q_0(t) = G(t) \ .$$

In this case, all models agree at inception to the par spread

$$s_0(a,b) = \frac{(1-R)\int_{T_a}^{T_b} P_0(u)h(u)G(u)du}{C_0(a,b)} \ .$$

where h is given in (2.4).

2.4.2 Credit Value Adjustment

A major concern of the post-crisis regulation is the modelling of the capital requirement of firms tacking into account some credit adjustments to the valuation under credit risk. Counterparty credit risk is defined as the risk that the counterparty of an over-the-counter (OTC) deal will default before the maturity of the contract. The latter can be seen as an option given to the counterparty, and can be priced in a risk-neutral setup by adjusting the OTC derivative, leading to CVA. The latter is nothing but the expected losses due to the missed payments associated to the OTC portfolio. In a risk-neutral specification and assuming $\tau > 0$, the value at time t of the CVA is expressed as:

$$\mathrm{CVA}_t = \beta_t \mathbb{E}\left[(1-R)\frac{V_\tau^+}{\beta_\tau}\mathbb{1}_{\tau<T}\middle|\mathcal{G}_t\right] \tag{2.37}$$

for a $\mathcal{G}_\tau$-measurable exposure V, recovery rate R and bank account β given in (2.1).

The time-t CVA formula (2.37) can be expressed in term of $\mathcal{F}_t$ conditional expectations with an integral involving the Azéma supermartingale S_t given in (2.19). Using Lemma 1, the time-0 CVA simply yields

$$\mathrm{CVA} = -\mathbb{E}\left[(1-R)\int_0^T \frac{V_u^+}{\beta_u} dS_u\right] \tag{2.38}$$

In the independent case, CVA depends only on the discounted expected positive exposure (EPE): $\mathbb{E}\left[\frac{V_t^+}{\beta_t}\right]$ since all models are calibrated to the same market survival probability curve G. In this case, the independent CVA formula is given by:

$$\mathrm{CVA}^\perp = -(1-R)\int_0^T \mathbb{E}\left[\frac{V_u^+}{\beta_u}\right] dG(u) = (1-R)\int_0^T \mathbb{E}\left[\frac{V_u^+}{\beta_u}\right] h(u)G(u)du \ . \tag{2.39}$$

Generally speaking however the latter expression does not hold, and CVA depends on the joint dynamics of the exposure and credit worthiness. We cannot get such a simpler formula as in (2.39) and we have to take in account the dependency between credit and exposure. Wrong-way risk (WWR) is the additional risk related to this dependency.

Chapter 3

A subordinated CIR intensity model with application to wrong-way risk CVA

3.1 Introduction

Since the 2008 crisis, regulators suggest financial institutions pay specific attention to counterparty credit risk (CCR) when valuing over-the-counter (OTC) deals. In this context, CCR refers to the possibility that the counterparty of the transaction can default before the maturity of the contract. The CCR can be accounted for either by setting up a strong collateralisation agreement, or by charging a Credit Value Adjustment (CVA) to absorb the corresponding expected losses. One of the main challenge when pricing such adjustments is to account for the potential dependency of the exposure with counterparty's credit quality, a phenomenon commonly referred to as wrong-way risk (WWR).

In that respect, a popular framework fitting in the class of reduced-form models is to consider the default intensity to be governed by the CIR++ (Brigo and Alfonsi, 2005) or the JCIR++ (Brigo and El-Bachir, 2010) process. In essence, the intensity is modeled as a CIR or a jump diffusion CIR (JCIR) dynamics shifted in a deterministic way so as to fit a given CDS term structure. However, this model suffers from an important restriction: the resulting intensity process (including the shift) needs to be positive. Because for tractability reasons the jumps in the JCIR++ model are upwards only, increasing the jump activity (e.g., to increase the implied spread volatility) under the constraint of keeping the survival probability curve unchanged introduces a shift function that tends to be more and more negative. Because negative shifts should be ruled out for consistency reasons, this puts limits on the CIR or JCIR parameters that can be chosen. This will further limit the WWR impact in CVA applications. One way to limit the appearance of shift functions with negative values would be to allow for both upwards and downwards jumps, without affecting the tractability of the model.

In this chapter, we consider the approach of Mendoza-Arriaga and Linetsky, 2014 to model the default intensity as a time-changed CIR in a CVA context. This poses several problems that need to be addressed. First, the time-changed process will destroy the potential correlation between the intensity and the exposure increments. As a consequence, this would lead to a weak WWR impact. Second, the time-changed approach may also introduce arbitrage opportunities via a forward looking effect. Eventually, even if some techniques exist to deal with WWR in a semi-analytical way (see, e.g., Vrins, 2017; Brigo and Vrins, 2018), one generally has to rely on Monte Carlo simulations. Standard Monte Carlo methods are known to be computationally intensive. In addition, because of the stochastic clock, the time-changed model is very time consuming due to the fact that the simulation is done in a random grid.

Our contribution in this chapter is multiple. First, we propose a way to use the time-changed model of Mendoza-Arriaga and Linetsky in the context of CVA avoiding both the correlation destruction and the appearance of arbitrage opportunities. This is achieved by reconstructing the exposure process in a "synchronous" way with the intensity, while preserving the original exposure's dynamics. Second, we show via numerical experiments that the corresponding model is indeed able to generate larger WWR CVA figures compared to JCIR++ without facing the inconsistency issue resulting from a negative shift. Eventually, we propose a variance reduction technique based on the *adaptive control variate* to reduce the computational cost.

The chapter is organized as follows. In Section 3.2, we recall some definitions related to the benchmark models considered in this chapter. In Section 3.3, we introduce the subordinated model combined with a reconstruction of the exposure process avoiding possible arbitrage opportunities resulting from the time change. Section 3.4 reviews the CVA computation for the benchmark models and the new subordinated model. In Section 3.5, we present the numerical experiments including the comparison of the diffusion models and the time-changed model in term of WWR effects and the control variate technique. The last section contains some concluding remarks and perspectives.

3.2 Diffusion default intensity models

Before defining the subordinated model, we recall the definitions of the chosen standard intensity benchmark models. In this study, we consider the square-root jump diffusion default intensity models and their extended shifted version SSRD Brigo and Alfonsi, 2005 and SSRJD Brigo and El-Bachir, 2010 (see Section 2.3 for more details).

3.2.1 The diffusion intensity market model

We consider a reduced-form model (see Section 2.2) where $\mathbb{F}$ is the filtration generated by the stochastic market variables (interest rates, default intensities, prices, etc), except default events. In this setup, the components of the vector $\mathbf{W} = (W^\beta, W^V, W^\perp)$ are part of risk drivers. In particular, W^β governs the dynamics of the risk-free rate r, hence that of the bank account numéraire (2.1).

Under risk-neutral measure $\mathbb{Q}$, all prices of tradable assets divided by β are $\mathbb{F}$-martingales between two cash-flow dates. The second Brownian motion W^V drives the dynamics of the portfolio price process

$$dV_t = b(V_t)dt + \sigma(V_t)dW_t^V, \quad V_0 > 0. \tag{3.1}$$

We assume the coefficients b, σ to be regular enough to guarantee that a unique strong solution to this SDE exists. Finally, we model the default time τ of our counterparty as in the Cox setup (see Section 2.2.2) where the default intensity λ is driven by a Brownian motion correlated to W^V, $W^\lambda := \rho W^V + \sqrt{1-\rho^2}W^\perp$, $W^\perp \perp W^V$, $\rho \in [-1, 1]$, but the threshold $\mathcal{E}$ is independent from $\mathbb{F}$. Since the random time τ is constructed through a Cox process, $\mathcal{H}$-Hypothesis in Definition 1 holds between the filtrations $\mathbb{F}$ and $\mathbb{G}$ and the model is arbitrage free.

3.2.2 The CIR++ and JCIR++ intensity models

A convenient way to define the intensity process λ is to set $\lambda_t = y_t$ where y follows a CIR SDE, i.e., takes the form (2.28) with $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\varrho, -\kappa, 0, \delta^2, 0, 0)$:

$$dy_t = \kappa(\varrho - y_t)dt + \delta\sqrt{y_t}dW_t^\lambda, \quad y_0 > 0. \tag{3.2}$$

By doing so, the intensity process becomes a mean-reverting square-root process with speed of mean reversion κ, long-term mean ϱ and volatility δ, usually chosen to satisfy the Feller constraint $2\kappa\beta > \delta^2$.

In order to describe the appearance of positive jumps in the default intensity process, we consider the JCIR model defined as in (2.28) with $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\varrho, -\kappa, 0, \delta^2, \alpha, \omega)$:

$$dy_t = \kappa(\varrho - y_t)dt + \delta\sqrt{y_t}dW_t^\lambda + dJ_t \tag{3.3}$$

where J is $\mathbb{F}$-adapted and independent of $\mathbf{W}$.

Adding non-negative jumps independent from W^λ in the SDE (3.2) increases the volatility of the intensity process. These two choices belong to the class of Affine models: the time-t survival probability (2.26) has the simple form (according to Definition 2)

$$Q_t(T) = \mathbb{1}_{\{\tau>t\}} e^{A_t^y(T) - B_t^y(T) y_t} \tag{3.4}$$

Shifting the process y in a time-dependent way does not affect the above relationship as long as the shift is deterministic but provides full flexibility in terms of calibration capabilities. Therefore, one typically consider $\lambda_t = y_t + \varphi(t)$ where y is a CIR or JCIR process and φ is chosen such that the model and market survival probability curves coincide at inception: $\mathbb{E}[S_t^\varphi] := \mathbb{E}[e^{-\int_0^t (y_s + \varphi(s))ds}] = G(t)$, where S^φ is the Azéma supermartingale associated to the deterministic shift model. The corresponding models are known as CIR++ and JCIR++, depending on whether y features jumps or not. The main advantage of adding the shift is that we can fit exactly any term structure of hazard rates and derive analytical formulas both for bonds and European options. In particular, the CIR++ and JCIR++ models remain affine, we just need to replace $A_t^y(T)$ by $A_t^y(T)e^{-\int_0^t \varphi(s)ds}$ in (3.4) (see Section 2.3.2 for more details). In this case, following Equation (2.31), the shift φ, at any time t, is given by (for both CIR++ and JCIR++ models)

$$\varphi(t) = -\frac{d}{dt} \ln \frac{G(t)}{Q_0(t)}. \tag{3.5}$$

The weakness of this approach is that we can guarantee the positivity of intensities only through restrictions on model parameters such that $\varphi \geq 0$. Indeed, y can take values arbitrarily close to zero, so that the condition $\lambda \geq 0$ $\mathbb{Q}$-a.s. is equivalent to saying $\varphi \geq 0$. However in general, φ has to correct the function $Q_0(\cdot)$ so as it sticks, thanks to the shift, to the target function $G(\cdot)$. Given a set of model parameters for y, nothing prevents φ to become negative, in general. But should it take negative values, the resulting model fails to be a Cox-type, and the $\mathcal{H}$-hypothesis does not hold anymore. This problem is of high importance in practice. Indeed, the CIR model has low volatility to fit CDS curves, and increasing the volatility just breaks

the Feller condition. It is possible to increase the volatility without breaking the Feller constraint using JCIR. However, the affine form of the JCIR model requires the jumps to be independent from the diffusion part, so that positivity allows for upwards jumps only. This of course tends to increase the mean of λ, and hence to decrease $Q_0(\cdot)$ (the shift can go quickly negative leading to negative intensities which is inconsistent to the Cox model). Reciprocally, fitting a given target curve $G(\cdot)$ with non-negative shift only puts constraints on the jump sizes/rates, hence on the attainable volatility.
However, one cannot increase the activity of J without bounds. By doing so indeed, the calibration constraint $Q_0(\cdot) = G(\cdot)$ drives the implied shift function φ downwards. As φ cannot take negative values, there is a strong limit on the jump rates and/or sizes that one can use while preserving the consistency of the model. In the next section, we propose an alternative model that is less subject to suffer from the above problems.

3.3 Time-changed intensity model

As explained earlier, the fact that the jumps in the JCIR model rapidly push φ downwards results from the fact that the jumps are positive only. This would not be the case if the jumps could go in both directions. Yet, it is not enough just to use symmetric jumps in JCIR++: this would break the positiveness of y in (3.3) if J is independent from W^λ.
One possibility consists in modelling λ as a time-changed version of a standard intensity process like the CIR process using the time change idea of Mendoza-Arriaga and Linetsky, 2014. On that respect, we define the time-changed CIR (TC-CIR) model by subordinating the CIR process y in (3.2) with a jump-process

$$\theta_t := t + J'_t \text{ [1]} \tag{3.6}$$

where J' is a compound Poisson process independent of J and $\mathbf{W}$ but with jump arrival rate ω' and exponential jump size α' defined with a Poisson process N'. That is, we define a new process y^θ by $y^\theta_t = y_{\theta_t}$ where θ is the stochastic clock defined above. If the stochastic clock features jumps, the resulting time-changed process would still be positive, and would feature jumps in both directions. This would provide a mean to increase the volatility of λ avoiding the implied shift of the

[1] It is possible that θ is a subordinator with drift $a > 0$ (i.e., $\theta_t = at + J'_t$) but we focus here on the case $a = 1$.

time-changed model to become negative too quickly as the jumps activity increases. As y^θ is no longer affine, we need to apply the procedure developed by Mendoza-Arriaga and Linetsky, 2014 to get a closed-form for the survival probability. This approach is a time-changed CIR default intensity by means of subordination in the sense of Bochner (see Mendoza-Arriaga and Linetsky, 2014, for more details). Based on a Cox model, it is analytically tractable by means of eigenfunction expansions of relevant semigroups, yielding closed-form pricing of defaultable zero coupon bonds.

3.3.1 The time-changed market model

Consider a reduced-form setting with the corresponding time change probability space $(\Omega, \mathcal{G}_T^\theta, \mathbb{G}^\theta, \mathbb{Q})$ with $\mathcal{G}_t^\theta = \mathcal{G}_{\theta_t}$ and $\mathbb{G}^\theta = (\mathcal{G}_{\theta_t})_{t\geq 0}$. To introduce the time-changed defaultable market, we consider the default time as defined in (2.22) in order to determine the corresponding intensity of the time-changed model. Let's define the corresponding indicator process of D defined in Chapter 2 by $D_t^\theta := \mathbb{1}_{\{\tau\leq\theta_t\}}$, $t \geq 0$. To introduce the time-changed filtration, we need first to define an inverse subordinator process $(L_t := \inf\{s \geq 0 : \theta_s > t\}, t \geq 0)$. Let $\mathbb{L} = (\mathcal{L}_t)_{t\geq 0}$ be its completed natural filtration and $\mathbb{H} = (\mathcal{H}_t)_{t\geq 0}$ the enlarged filtration with $\mathcal{H}_t = \mathcal{G}_t \vee \mathcal{L}_t$. We then define our time-changed filtration $\mathbb{H}^\theta = (\mathcal{H}_t^\theta)_{t\geq 0}$ by $\mathcal{H}_t^\theta = \mathcal{H}_{\theta_t}$. Hence, the time-changed bivariate process $(y_t^\theta, D_t^\theta)_{t\geq 0}$ is $\mathbb{H}^\theta$-adapted and càdlàg and is an $\mathbb{H}^\theta$-semimartingale (see Mendoza-Arriaga and Linetsky 2014 for details). In this setup, from the Doob-Meyer decomposition of D^θ, our time-changed intensity is (see Theorem 3.3 (iii) in Mendoza-Arriaga and Linetsky 2014) given by $\lambda_t^{\mathbb{H}^\theta} = (1 - D_t^\theta)\lambda_t^\theta$,

$$\lambda_t^\theta = k^\theta(y_t^\theta) \quad \text{with} \quad k^\theta(x) = x + \int_{(0,\infty)} \left(1 - e^{A_0^y(s) - B_0^y(s)x}\right) \nu(ds) \tag{3.7}$$

where $\nu(ds) = \omega'\alpha' e^{-\alpha' s} ds$ and A_t^y, B_t^y are the same as in (3.4). Hence, if k^θ is a function from $\mathbb{R}^+$ to $\mathbb{R}^+$ and y is an intensity process, λ^θ defines a new intensity process and can be used to define a new default time using the Cox framework used above:

$$\tau^\theta := \inf\left\{t \geq 0 : \int_0^t \lambda_s^\theta ds \geq \mathcal{E}\right\} \tag{3.8}$$

and we have that $\{\tau^\theta \leq t\} \equiv \{\tau \leq \theta_t\}$ $\mathbb{Q}$- a.s., with

$$D_t^\theta = \mathbb{1}_{\{\tau^\theta \leq t\}}. \tag{3.9}$$

In this setup, the new time-t survival probability to survive up to time T is

$$Q_t^\theta(T) := \mathbb{Q}\left(\tau^\theta > T | \mathcal{H}_t^\theta\right) = \mathbb{1}_{\{\tau^\theta > t\}} \frac{\mathbb{E}[S_T^\theta | \mathcal{F}_t^\theta]}{S_t^\theta} = \mathbb{1}_{\{\tau^\theta > t\}} \frac{S_t^\theta(T)}{S_t^\theta(t)} \tag{3.10}$$

where $S_t^\theta = \mathbb{Q}(\tau^\theta > t \,|\, \mathcal{F}_t^\theta) = e^{-\int_0^t \lambda_s^\theta ds}$ is the Azéma supermartingale and $S_t^\theta(T) = \mathbb{Q}(\tau^\theta > T \,|\, \mathcal{F}_t^\theta)$ the risk-neutral survival probability in the time-changed model. In a multivariate setup in general and in the specific case of CVA application in particular, the time-changed approach presents a problem. Indeed, λ^θ is an intensity, which typically can be correlated with other processes (e.g., V and β). If we correlate these Brownian motions, two problems arise. First, because of the time change, the correlation between the intensity λ^θ and (V, β) is partially destroyed. Indeed, V_t and β_t depend on W^V and W^β on $[0, t]$ whereas the intensity λ^θ depends on W^λ on $[0, \theta_t]$ with $\theta_t \geq t$. This methodology thus destroys the dependence between the processes λ and (V, β) as the intensity or the size of jumps of θ increases. Another problem, probably even more important, is related to arbitrage opportunities. The knowledge of λ^θ at t contains information on W^λ up to θ_t. If $\rho \neq 0$, this introduces a forward looking effect on V. It is therefore important to work with a model in which all processes remain synchronized, but without changing the law of $V_.$ and $\beta_.$ as originally specified. To do this, it is enough to rebuild new Brownian motions $(\widetilde{W}^V, \widetilde{W}^\beta)$ so that the increments of (V, β) remain synchronized with those of λ^θ.

Lemma 2. *Let W be a Brownian motion and t_i be the time of the i^{th} jump of a Poisson process N'_t. Then the process*

$$\widetilde{W}_t := \sum_{i=0}^{N'_t - 1} \int_{t_i \wedge t}^{t_{i+1}^- \wedge t} dW_{\theta_s} + (W_{\theta_t} - W_{\theta_{N'_t}}) \tag{3.11}$$

is an $\mathbb{F}^\theta$-Brownian motion and behaves exactly as W sampled on the time grid.

Proof. $\widetilde{W}$ is a continuous $\mathbb{F}^\theta$-local martingale with $\widetilde{W}_0 = 0$ and for every $t \geq 0$, $\langle \widetilde{W} \rangle_t = t$. By Lévy's characterization theorem, the process $(\widetilde{W}_t)_{t \geq 0}$ is an $\mathbb{F}^\theta$-Brownian motion. □

The dynamics of V can thus be equivalently described in terms of $\widetilde{W}^V$ setting $W = W^V$ on $\mathbb{F}^\theta$ as:

$$d\widetilde{V}_t = b(\widetilde{V}_t)dt + \sigma(\widetilde{V}_t)d\widetilde{W}_t^V, \quad \widetilde{V}_0 = V_0. \tag{3.12}$$

Applying the same procedure of Lemma 2 on W^{β}, the corresponding copy of W^{β} on the time change grid is $\widetilde{W}^{B}$. The latter will govern the new dynamics of the risk-free rate $\tilde{r}$ leading to that of the bank account numéraire given by

$$d\widetilde{\beta}_t = \tilde{r}_t \widetilde{\beta}_t dt, \quad \widetilde{\beta}_0 = 1. \tag{3.13}$$

Remark 3. It is worth stressing the fact that keeping W^V as the driver of the exposure process V (instead of $\widetilde{W}^V$) does not completely destroy the WWR effect resulting from the correlation ρ between W^V and W^{λ}. The reason is that even if the *instantaneous correlation between the infinitesimal increments* of λ^{θ} and V are mutually independent as from the first jump of θ, the correlation between λ^{θ}_t and V_t is non-zero for any $t > 0$ whenever $\rho \neq 0$. This is because these processes depend on the integrals of increments of Brownian motions on some time intervals. Between two jumps of the clock in particular, the Brownian increments driving the change in the intensity process λ^{θ} can be independent from the Brownian increments driving the exposure V on a same time period, but the increments of the first Brownian motion on a given time interval can be dependent on the second Brownian motion on *another* interval. This explains that two processes λ^{θ} and V can be dependent on each other even if the instantaneous correlation of their increments vanishes because of the time change. By contrast, the correlation between λ^{θ}_t and V_t is lower than that of λ^{θ}_t and $\widetilde{V}_t$: by synchronizing the changes of the Brownian motion driving V with the one driving λ^{θ}, one maximizes the attainable correlation. For instance, let W, B be two Brownian motions with instantaneous correlation ρ, i.e., $d\langle W, B\rangle_t = \rho dt$ and define B^{δ} as $B^{\delta}_t = B_{t+\delta}$ where $\delta > 0$. The correlation between the increments of W and B on the interval $[s,t]$ is ρ whereas that between W and B^{δ} is $\rho(t-(s+\delta))^+/(t-s)$ whose absolute value is no greater than $|\rho|$.

Theorem 1. *Discounted payoffs driven by $\widetilde{V}$ and $\widetilde{\beta}$ are $\mathbb{H}^{\theta}$-martingales under $\mathbb{Q}$.*

Proof. We know that the discounted payoffs driven by V and β are $(\mathbb{Q}, \mathbb{F})$-martingales, hence they are $(\mathbb{Q}, \mathbb{G})$-martingales due to the $\mathcal{H}$-Hypothesis. By construction, $\widetilde{V}$ and $\widetilde{\beta}$ have the same $\mathbb{F}^{\theta}$ dynamics as V and β under $\mathbb{F}$ (see Lemma 2). Hence, the discounted payoffs driven by $\widetilde{V}$ and $\widetilde{\beta}$ are $(\mathbb{Q}, \mathbb{F}^{\theta})$-martingales. Because τ^{θ} is modelled with a Cox process, immersion holds and they remain martingales when progressively enlarging $\mathbb{F}^{\theta}$ with τ^{θ}. This shows that they are $(\mathbb{Q}, \mathbb{G}^{\theta})$-martingales, hence $(\mathbb{Q}, \mathbb{H}^{\theta})$-martingales (since $\mathcal{L}_{\theta_t} = t$). $\square$

As explained earlier, time-changing the intensity can introduce a forward-looking effect and thus arbitrage opportunities. Indeed, prices of non-dividend paying asset discounted at the risk-free rate would not be a martingale under $\mathbb{Q}$ in the natural filtration generated by $(W^V, W^{\lambda^\theta})$ when $\rho \neq 0$. This is formalized in the next lemma proven in the Appendix.

Lemma 3. *Suppose that V is a martingale with respect to $\mathbb{F}^{W^V}$, the natural filtration of W^V. It is also a martingale with respect to $\mathbb{F}^{W^V} \vee \mathbb{F}^{W^\lambda}$. We note W^{λ^θ} the time-changed version of W^λ. For instance suppose that for some $s \in [0,t]$, we have $s < \theta_s < t$. Then, V may not be a $\mathbb{F}^{W^V} \vee \mathbb{F}^{W^{\lambda^\theta}}$-martingale.*

3.3.2 The time-changed CIR++ intensity model

To define the shifted time-changed CIR (TC-CIR++) model, we need to have a closed form of the survival probability of the TC-CIR model. For that, we only need to compute the Laplace transform of our time change process θ in order to apply the idea devised in Mendoza-Arriaga and Linetsky, 2014.

Laplace transform of a Lévy subordinator: A *subordinator* (see Applebaum, 2009) is a one-dimensional Lévy process that is non-decreasing (a.s.). Such processes can be thought of as a random model of time evolution, since if $(\mathcal{T}_t)_{t\geq 0}$ is a subordinator we have

$$\mathcal{T}_t \geq 0 \quad \text{a.s. for each } t > 0,$$

and

$$\mathcal{T}_{t_1} \leq \mathcal{T}_{t_2} \quad \text{a.s. whenever } t_1 \leq t_2.$$

By the Lévy-Khintchine formula, the Laplace transform of the Lévy subordinator $\mathcal{T}$ is

$$\mathbb{E}[e^{-u\mathcal{T}_t}] = \int_{[0,\infty)} e^{-us}\pi_t(ds) = e^{-t\phi(u)}$$
$$\text{with } \phi(u) = \gamma u + \int_{(0,\infty)} (1 - e^{-us})\nu(ds)$$

where $\pi_t(ds)$ is the transition kernel, ψ is a Lévy exponent, $\gamma \geq 0$ is the drift and $\nu(ds)$ is the Lévy measure of the subordinator satisfying the integrability condition $\int_{(0,\infty)}(s \wedge 1)\nu(ds) < \infty$.

Our time change jump-process θ is a Lévy subordinator and its Laplace transform can be found easily using the Lévy-Khintchine formula. Hence, for any $u \in \mathbb{R}$, the Laplace transform of θ reads

$$\mathbb{E}[e^{-u\theta_t}] = \mathbb{E}[e^{-u(t+J'_t)}] = e^{-ut}\mathbb{E}[e^{-uJ'_t}] \,. \tag{3.14}$$

From the exponential formula, a Lévy-Khintchine formula representation of the compound Poisson process yields

$$\mathbb{E}[e^{-uJ'_t}] = e^{-t\psi(u)} \tag{3.15}$$

where ψ is given by

$$\psi(u) = \int_{\mathbb{R}`\{0\}} (1 - e^{-us})\omega'\alpha' e^{-\alpha' s} ds \mathbb{1}_{\{s>0\}} = \int_{(0,\infty)} \omega'\alpha'(1 - e^{-us})e^{-\alpha' s} ds. \tag{3.16}$$

It follows that θ satisfies the Lévy-Khintchine formula with $\gamma = 1$ and $\nu(ds) = \omega'\alpha' e^{-\alpha' s}ds$, and thus

$$\mathbb{E}[e^{-u\theta_t}] = e^{-t\phi(u)} \quad , \quad \phi(u) = u\left(\frac{u + \alpha' + \omega'}{u + \alpha'}\right). \tag{3.17}$$

Knowing ϕ, the time change survival probability takes the closed form

$$Q^\theta_t(T) = \mathbb{1}_{\{\tau^\theta > t\}} \sum_{n=1}^{\infty} e^{-\phi(\lambda_n)(T-t)} f_n(0)\varphi_n(y^\theta_t) \tag{3.18}$$

where λ_n, f_n and φ_n are given in Mendoza-Arriaga and Linetsky, 2014.
The TC-CIR++ model is obtained by defining the time-changed intensity process as $\lambda^\theta_t = k^\theta(y^\theta_t) + \varphi^\theta(t)$ and finding φ^θ such that $\mathbb{E}[S^\theta_t] = G(t)$. Hence

$$\varphi^\theta(t) = -\frac{d}{dt} \ln \frac{G(t)}{Q^\theta_0(t)}. \tag{3.19}$$

3.4 CVA in a reduced-form setup

The reduced-form approach relies on a change of filtrations. In this section, we derive the CVA formulas in both cases where the default intensity is given by the square-root diffusions or the time-changed model. Let R be the recovery rate of the counterparty (assumed to be constant) and the exposure processes V and $\widetilde{V}$ respectively given by (3.1) and (3.12). We refer to Section 2.4.2 for more details.

3.4.1 CVA formula in the diffusion model

From the $\mathcal{H}$-Hypothesis (due to the Cox setup), payoffs driven by V and β are $\mathbb{G}$-martingales. In addition, since V/β is $\mathbb{Q}$-integrable and $\mathbb{F}$-predictable, assuming deterministic recovery rate R, the time-0 CVA expression is given by (2.38)

Discretizing the integral in (2.38) with a numerical scheme, the Monte-Carlo estimation of time-0 CVA becomes

$$\widehat{\mathrm{CVA}}_0 := -(1-R)\frac{1}{m}\sum_{i=1}^{m}\sum_{k=1}^{n}\frac{V_{t_k}^{+,(i)}}{\beta_{t_k}^{(i)}}\Delta S_{t_k}^{(i)}, \quad n = \frac{T}{\Delta} \tag{3.20}$$

where $\Delta S_{t_k}^{(i)} = S_{t_k}^{(i)} - S_{t_{k-1}}^{(i)}$. The right-hand side results from Monte Carlo approximation, by taking the sample mean of m time-integrals discretized in n intervals of length Δ.

In the specific case where τ is independent from the discounted exposure (i.e., $\rho = 0$), the *independent* CVA formula is given by (2.39).

In other words, CVA only depends *separately* on the expected discounted exposure $\mathbb{E}\left[\frac{V_u^+}{\beta_u}\right]$ and the prevailing risk-neutral survival probability curve $G(.)$. We refer to Brigo and Vrins, 2018 for more details.

3.4.2 CVA formula in the time-changed model

As $\widetilde{V}/\widetilde{\beta}$ is $\mathbb{Q}$-integrable and $\mathbb{F}^\theta$-predictable, assuming $\tau^\theta > 0$ and using Theorem 1, in a similar way as before, but using $(\widetilde{V}, \widetilde{\beta})$ instead of (V, β) (having the same dynamics under $\mathbb{F}^\theta$ as (V, β) under $\mathbb{F}$), but using the intensity of τ^θ with Azéma supermartingale S^θ, Lemma 1 yields

$$\begin{aligned}
\mathrm{CVA}_t &= \mathbb{1}_{\{\tau^\theta > t\}}\widetilde{\beta}_t\mathbb{E}\left[(1-R)\frac{\widetilde{V}_{\tau^\theta}^+}{\widetilde{\beta}_{\tau^\theta}}\mathbb{1}_{\{\tau^\theta \le T\}}\middle|\mathcal{H}_t^\theta\right] \\
&= \mathbb{1}_{\{\tau^\theta > t\}}\widetilde{\beta}_t\mathbb{E}\left[(1-R)\frac{\widetilde{V}_{\tau^\theta}^+}{\widetilde{\beta}_{\tau^\theta}}\mathbb{1}_{\{\tau^\theta \le T\}}\middle|(\tau^\theta \le t) \vee \mathcal{F}_t^\theta\right] \\
&= -\mathbb{1}_{\{\tau^\theta > t\}}\frac{\widetilde{\beta}_t}{S_t^\theta}\mathbb{E}\left[(1-R)\int_t^T\frac{\widetilde{V}_u^+}{\widetilde{\beta}_u}dS_u^\theta\middle|\mathcal{F}_t^\theta\right] .
\end{aligned} \tag{3.21}$$

Therefore, the time-0 CVA in the time-changed model can be approximated using m paths of Monte Carlo simulations as

$$\widehat{\mathrm{CVA}}_0 := -(1-R)\frac{1}{m}\sum_{i=1}^{m}\sum_{k=1}^{n}\frac{\widetilde{V}_{t_k}^{+,(i)}}{\widetilde{\beta}_{t_k}}\Delta S_{t_k}^{\theta,(i)}, \quad n = \frac{T}{\Delta} . \tag{3.22}$$

In the specific case where τ^θ is independent from the discounted exposure, we obtain the *independent* CVA formula in the time-changed model

$$\mathrm{CVA}_0^\perp = -(1-R)\int_0^T\mathbb{E}\left[\frac{\widetilde{V}_u^+}{\widetilde{\beta}_u}\right]d\,\mathbb{E}[S_u^\theta] . \tag{3.23}$$

Observe that by construction of φ and φ^θ, $\mathbb{E}[S_t^\varphi] = \mathbb{E}[S_t^\theta] = G(t)$. Hence, $\mathrm{CVA}_0^\perp$ agrees with both models under the calibration constraint. This means

$$\mathrm{CVA}_0^\perp = -(1-R)\int_0^T \mathbb{E}\left[\frac{V_u^+}{\beta_u}\right] dG(u)) = -(1-R)\int_0^T \mathbb{E}\left[\frac{\widetilde{V}_u^+}{\widetilde{\beta}_u}\right] dG(u). \quad (3.24)$$

3.5 Numerical experiments

In this section, we start by defining the simulation procedure of the bivariate process $(y^\theta, \widetilde{V})$. The CVA is computed using standard Monte Carlo simulation and the performance of the shifted time-changed model in terms of WWR is compared to the CIR++ and JCIR++ stochastic intensity models. For the sake of simplicity, the recovery rate R and the interest rate r are assumed to be constant and set to zero (i.e., $R = 0$ and $\beta_t = \widetilde{\beta}_t = 1$) to put the focus on the treatment of the credit-exposure dependency.

3.5.1 Simulation procedure of $(y^\theta, \widetilde{V})$

Let's denote by $\mathcal{T} = \{0, \Delta, 2\Delta, \ldots, T\}$ the time-t grid and $\mathcal{T}^\theta = \{0, \theta_\Delta, \theta_{2\Delta}, \ldots, \theta_T\}$ the grid at time θ_t. To refine the grid $\mathcal{T}^\theta$, we define grids $\mathcal{T}_i^\theta, i = 1, \ldots, n$, as

$$\mathcal{T}_i^\theta := \bigcup_{j=1}^{n_i - 1} \left\{\theta_{t_{i-1}} + j\frac{(\theta_{t_i} - \theta_{t_{i-1}})}{n_i}\right\},\ n_i = \left\lceil \frac{(\theta_{t_i} - \theta_{t_{i-1}})}{\Delta} \right\rceil,\ \mathcal{T}_{fine}^\theta := \mathcal{T}^\theta \bigcup \left\{\bigcup_{i=1}^{n} \mathcal{T}_i^\theta\right\}$$

The simulation grid $\mathcal{T}_{fine}^\theta$ contains $\mathcal{T}^\theta$ and is completed in such a way that (after sorting), the step between two consecutive points is no greater than the chosen time step δ to keep control on the discretization error independently of the jump sizes of θ.

To simulate y^θ, we simulate the CIR process y in (3.2) on $\mathcal{T}_{fine}^\theta$ using Diop's scheme (Euler approximation with positive-parts adjustment to ensure non-negativity), see Diop, 2003 and Alfonsi, 2005:

$$y_{(i+1)\Delta} = y_{i\Delta} + \kappa(\varrho - y_{i\Delta}^+)\Delta + \delta\sqrt{y_{i\Delta}^+}(W_{(i+1)\Delta}^\lambda - W_{i\Delta}^\lambda),\ i = 0, \ldots, n-1. \quad (3.25)$$

y^θ is obtained by extracting in $\mathcal{T}_{fine}^\theta$ the corresponding values of y on the grid $\mathcal{T}^\theta$. To obtain $\widetilde{V}$, we simulate W^V on $\mathcal{T}_{fine}^\theta$, extract its corresponding values on $\mathcal{T}^\theta$ and use respectively (3.11) and (3.12).

3.5.2 Numerical results

In this section, we compare the performance of the three models studied above in terms of WWR impact for a simple forward-type Gaussian exposure and Brownian swap bridge exposure. We fix the CIR parameters $(\kappa, \varrho, \delta, y_0)$ as in Brigo and Vrins, 2018 and search for the jump parameters (ω, α) of JCIR (affecting directly the intensity) and (ω', α') of TC-CIR (affecting the stochastic clocks and only indirectly the jumps in the intensity) such that $\psi \geq 0$ and the WWR impact is maximum. The two considered sets of model parameters are given in Table 3.1.

Set	y_0	κ	ϱ	δ	ω	α	ω'	α'
1	0.03	0.02	0.161	0.08	0.07	0.08	0.60	0.512
2	0.04	0.072	0.05	0.045	0.07	0.05	0.40	0.49

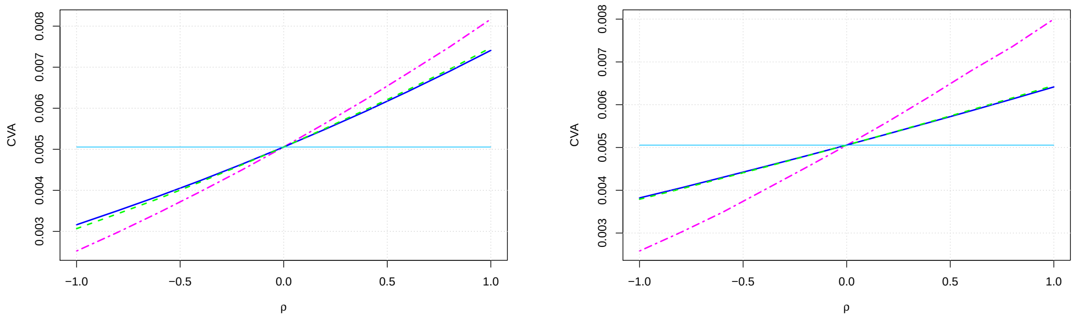
Table 3.1: Parameter sets considered for the numerical experiments.

Brownian exposure

We set the coefficients of the exposure dynamics in Equation (3.1) to $b(V_t) \equiv 0$, $\sigma(V_t) \equiv \sigma$ and $V_0 = 0$ which leads to

$$dV_t = \sigma dW_t^V. \tag{3.26}$$

This example illustrates a three-year forward contract or total return swap. In Figure 3.1, we plot the CVA as a function of the correlation ρ for the three considered models. We notice that due to the shift constraint, the JCIR model has slightly more WWR impact than the CIR model (because of the jumps) while the TC-CIR model has a larger WWR impact.



(a) Parameter Set 1 (b) Parameter Set 2

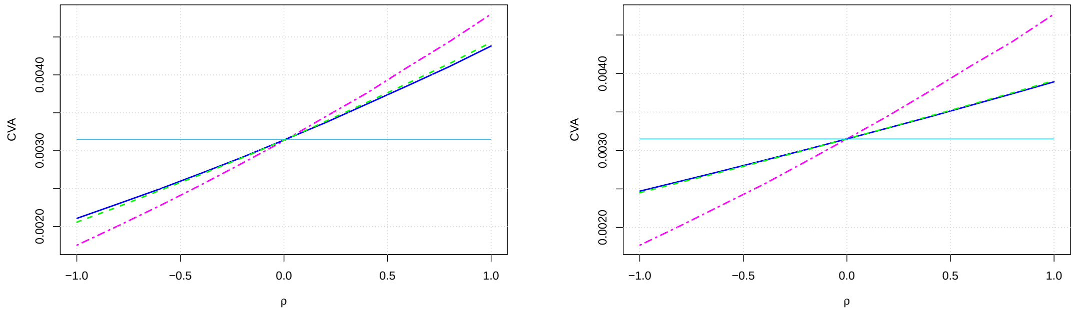
Figure 3.1: CVA figures for CIR++ (blue), JCIR++ (dotted green), TC-CIR++ (dotted magenta) and the analytical CVA with zero-correlation (cyan) using Monte Carlo method with $m = 10^6$, $\delta = 0.01$. Profiles: 3Y Gaussian exposure, $\sigma = 8\%$. Hazard rate is $h(t) = 5\%$.

Brownian bridge (with drift) exposure

The drifted Brownian bridge is obtained by choosing in (3.1) $b(V_t) \equiv \gamma(T-t) - \frac{V_t}{T-t}$, $\sigma(V_t) \equiv \sigma$ and $V_0 = 0$, so that

$$dV_t = \left[\gamma(T-t) - \frac{V_t}{T-t}\right] dt + \sigma dW_t^V \tag{3.27}$$

where γ stands for the future expected moneyness of an interest rate swap implied by the forward curve and σ controls the exposure volatility. We get similar results in term of WWR impact as in the previous example.



(a) Parameter Set 1 (b) Parameter Set 2

Figure 3.2: CVA figures for CIR++ (blue), JCIR++ (dotted green), TC-CIR++ (dotted magenta) and the analytical CVA with zero-correlation (cyan) using Monte Carlo method with $m = 10^6$, $\delta = 0.01$. Profiles: 3Y Swap exposure, $(\sigma, \gamma) = (8\%, 0.1\%)$. Hazard rate is $h(t) = 5\%$.

In the numerical examples, we consider the dynamics of the exposure process to be governed by a Brownian motion or Brownian bridge. In practice however, shifted geometric Brownian motion might be worth considering as a more realistic model to mimic real portfolios, using moment matching techniques (see Brigo et al., 2013, Chapter 4).

3.5.3 Adaptive control variate Monte Carlo estimator

The standard Monte Carlo method applied to the TC-CIR++ model is time consuming. The reason is that the time-step needs to be kept relatively small (to limit the discretization errors) whereas the simulation horizon is governed by θ_T which can be much larger than T. For example, the CPU time of the CVA computation using CIR++ intensity takes 74.5 seconds with 10^6 paths while the one with a TC-CIR++ intensity takes about 4 hours with the same number of paths[2]. In order to reduce the computational cost, we propose to adopt a variance reduction technique called *adaptive control variate*. In this section, we briefly recall the idea of control variate, describe its adaptive implementation and then transpose it to our CVA application.

Our purpose is to find an estimator of $\mathbb{E}[Y]$ from m independent, identically distributed observations of Y. The unbiased sample-mean estimator of $\mathbb{E}[Y]$ is

[2] The code was executed with a DELL CORE i5 vPro and the coding language is Rcpp.

$\hat{Y} := \frac{1}{m}\sum_{k=1}^{m} Y_k$. The idea of control variate consists of finding an alternative unbiased estimator $\tilde{Y}$ with lower variance compared to $\hat{Y}$ by using a control variate Z with known expectation. Consider a generic pair (Y, Z) of random variables with independent, identically distributed copies (Y_k, Z_k), $k \in \{1, 2, \dots, m\}$ and define

$$Y_k^\mu := Y_k - \mu\Xi_k \quad , \quad \Xi_k := Z_k - \mathbb{E}[Z] \; . \tag{3.28}$$

The sample-mean estimator of Y_k^μ is an alternative unbiased estimator of $\mathbb{E}[Y]$, and is given by

$$\hat{Y}^\mu = \frac{1}{m}\sum_{k=1}^{m} Y_k^\mu = \frac{1}{m}\sum_{k=1}^{m}(Y_k - \mu\Xi_k) \; . \tag{3.29}$$

Its variance is equal to that of $\hat{Y}$ for $\mu = 0$ but is minimum for $\mu = \mu^*$ where

$$\mu^* := \frac{\mathrm{Cov}(Y, \Xi)}{\mathrm{Var}(\Xi)} = \frac{\mathbb{E}[Y\Xi]}{\mathbb{E}[\Xi^2]} \; . \tag{3.30}$$

Because $\mathrm{Var}(\hat{Y}^{\mu^*}) = (1 - \mathrm{Corr}^2(Y, \Xi))\mathrm{Var}(\hat{Y})$, this approach is interesting when choosing the control variate Ξ highly correlated with Y.
In practice however, the optimal constant μ^* needs to be itself estimated. The *adaptive* control variate uses a different value for μ for every index $k \in \{1, 2, \dots, m\}$:

$$V_k := \frac{1}{k}\sum_{i=1}^{k}\Xi_i^2, \quad C_k := \frac{1}{k}\sum_{i=1}^{k} Y_i\Xi_i \quad \text{and} \quad \mu_k := \frac{C_k}{V_k} \; , \tag{3.31}$$

where $\mu_0 := 0$ and μ_{k-1} is the best estimator of μ^* at step k (see Pagès 2018 for more details). Eventually, the *adaptive control variate* estimator of $\mathbb{E}[Y]$ is given by

$$\tilde{Y}^\mu := \frac{1}{m}\sum_{k=1}^{m} Y_k^{\mu_{k-1}} = \hat{Y} - \frac{1}{m}\sum_{k=1}^{m}\mu_{k-1}Z_k + \frac{\mathbb{E}[Z]}{m}\sum_{k=1}^{m}\mu_{k-1} \; . \tag{3.32}$$

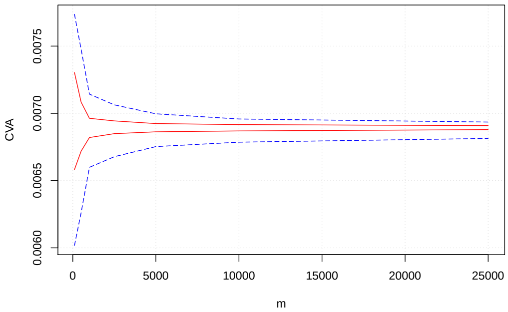
In our CVA application, we are interested in estimating CVA_0 which is nothing but $\mathbb{E}[Y]$ with $Y = -\int_0^T V_u^+ dS_u$ in the CIR and JCIR models (2.38) or $Y = -\int_0^T \widetilde{V}_u^+ dS_u^\theta$ in the TC-CIR model (3.21). We take as control variable $Z = -\int_0^T V_u^+ dS_u^\perp$ or $Z = -\int_0^T \widetilde{V}_u^+ dS_u^{\theta,\perp}$, respectively, where $S^\perp$ (resp. $S^{\theta,\perp}$) is the survival process S (resp. S^θ) associated to the intensity λ (resp. λ^θ) simulated using $W^\perp$ instead of W^λ.[3] In light of the above development, this choice is appealing because Z is correlated with Y (via the exposure process as well as the

[3] Alternatively, if the trajectories of V and S are stored, it is enough to shuffle those of V (or S), so as to combine, in the CVA computation, the i-th exposure's sample path with the $\pi(i) \neq i$ survival process' sample path.

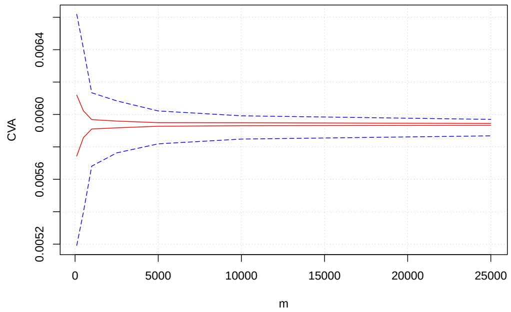
$W^\perp$ component of the survival process) whereas the expectation of Z is known in closed form and corresponds to $\mathrm{CVA}_0^\perp$ given in (3.24). In the adaptive procedure, the m pairs (Y_k, Z_k) are given by integrals (Y, Z) over the corresponding scenario. They are independent and identically distributed copies of Y and Z (up to discretization error resulting from the integral computation and simulation scheme). Eventually, our *adaptive* CVA estimator for the CIR and JCIR models reads

$$\widetilde{\mathrm{CVA}}_0 = \widehat{\mathrm{CVA}}_0 - \frac{1}{m}\sum_{k=1}^{m}\sum_{j=1}^{n}\mu_{k-1}V_{t_j}^{+,(k)}\Delta S_{t_j}^{\perp,(k)} + \frac{\mathrm{CVA}_0^\perp}{m}\sum_{k=1}^{m}\mu_{k-1} \tag{3.33}$$

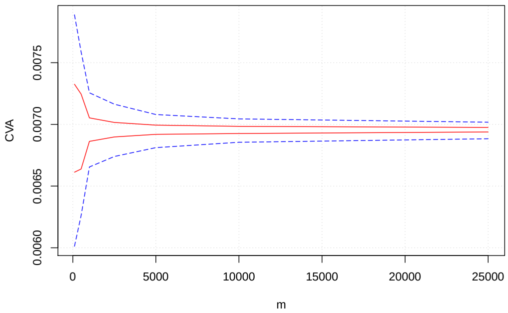
with $\widehat{\mathrm{CVA}}_0$ given in (3.20) and (3.22), respectively. The TC-CIR estimator takes a similar form provided that one replaces (V, S) by $(\widetilde{V}, S^{\theta,\perp})$.
In figure 3.3, we compare the confidence interval at level 95% of the CVA computation resulting from standard Monte Carlo and adaptive control variate. We apply this technique for the three considered models (CIR, JCIR and TC-CIR) using parameter Set 1 with a Gaussian exposure. Similar results are detailed in the Appendix when using the same Gaussian exposure profile but with the parameter Set 2.



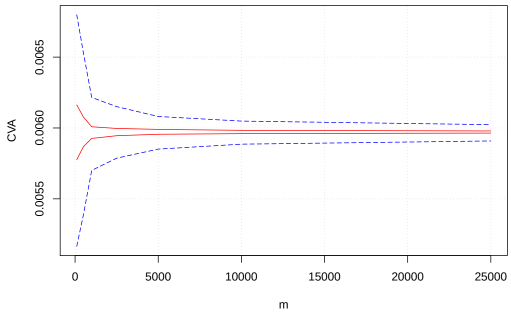
(a) CIR, $\rho = 0.8$



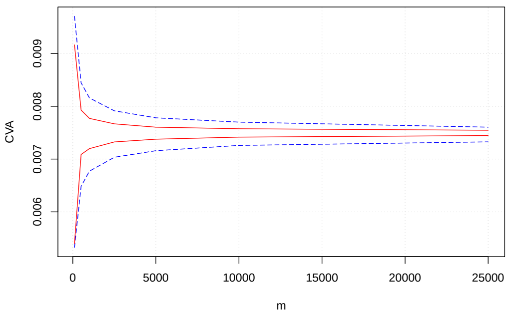
(b) CIR, $\rho = 0.4$



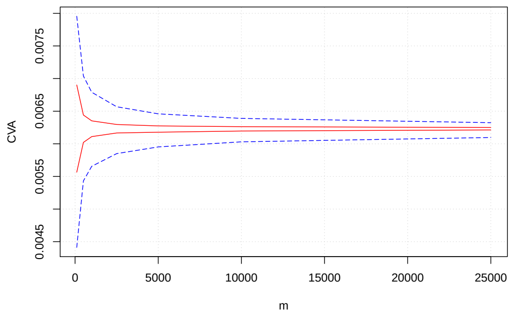
(c) JCIR, $\rho = 0.8$



(d) JCIR, $\rho = 0.4$



(e) TC-CIR, $\rho = 0.8$



(f) TC-CIR, $\rho = 0.4$

Figure 3.3: Confidence interval at level 95% of CVA figures for adaptive control variate (red) and standard Monte Carlo (dotted blue) using parameter Set 1. Profiles: 3Y Gaussian exposure, $\sigma = 8\%$. Hazard rate is $h(t) = 5\%$.

We observe clearly that the variance reduction technique adopted allows to reduce significantly the computational cost of the standard Monte Carlo as a solution to CVA computation in presence of WWR. Since Y and the chosen control Z are more correlated as $|\rho|$ decreases, we observe that the variance is reduced again when $|\rho|$ is decreasing and the convergence of the *adaptive* estimator is faster in this case. For example with $\rho = 0.4$, if we fix a level of precision at 1 bp and using a confidence level of 95%, the TC-CIR++ without control variate requires $m = 32,175$ paths to achieve the desired level of precision while TC-CIR++ with control variate requires only $m = 851$ paths. The corresponding computational time is respectively 8 minutes 30 seconds for the TC-CIR++ without control variate and 13 seconds for TC-CIR++ with the control variate estimator. We then reduce the computational cost for about 40 times by using control variate techniques.

3.5.4 Comparison between TC-CIR++ and Basel III formula

We already show that the time-changed model makes CVA more sensitive to WWR risk compared to the diffusion models. In order to characterize the magnitude of sensitivity to WWR, we provide a comparison of our CVA in the TC-CIR++ model (CIR++ and JCIR++ as well) with the one obtained by the Basel III formula. In the case without WWR, the *independent* CVA Basel III formula is given by (see Hull and White 2012, Section 3 for more details):

$$\mathrm{CVA}^{\mathrm{basel}} = (1-R)\sum_{i=1}^{n} q_i \nu_i \tag{3.34}$$

where q_i is the probability of a default for the counterparty unwind between times t_{i-1} and t_i (with $t_i (0 \le i \le n)$, $t_0 = 0$, $t_n = T$ and $t_0 < t_1 < t_2 < \ldots < t_n$), $\nu_i = \nu\left(t_i^\star\right)$ with $t_i^\star = 0.5(t_{i-1} + t_i)$ is the value of a derivative that has the payoff equal to the dealer's net exposure (i.e., dealer's expected exposure) to the counterparty at time $t_i^\star$ and R is the recovery rate (assumed to be constant as above). If s_i is the counterparty's credit spread for a maturity of t_i, we have that

$$q_i = e^{-h(t_{i-1})t_{i-1}} - e^{-h(t_i)t_i}, \quad h(t_i) = \frac{s_i}{1-R} \tag{3.35}$$

where $h(t_i)$ is the hazard rate at time t_i and assuming zero interest rate, the ν_i's are given by the risk-neutral expected exposure $\mathbb{E}[V^+_{t_i^\star}]$.

In our numerical setting, we clearly verify that the *independent* CVAs arising from Basel III formula and the TC-CIR++ model coincide. To deal with WWR, the regulators use the α multiplier to increase the exposure in the version of the model without WWR (3.34). The Basel II rules set α equal to 1.4 or allow banks to use their one models, with the floor for alpha of 1.2. Indeed, it is clear that the Basel III parameter α cannot be considered as a way to reproduce market-credit correlation but can capture the dependence or market-credit covariance. This means that the Basel III formula can only capture the dependence between portfolio and credit but disregards market and credit volatilities. Therefore, dynamic models are, from this perspective, a major improvement of the Basel type formula. Table 3.2 and Table 3.3 show the values of α which correspond to the CIR++, JCIR++ and TC-CIR++ models using respectively Gaussian and swap type exposures with the two parameter sets defined in Table 3.1. It is clear that when $\rho = 0$, $\alpha \approx 1$. In

	ρ	-1.0	- 0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1.0
A	Set 1	0.62	0.69	0.76	0.84	0.92	1.00	1.08	1.17	1.27	1.36	1.46
	Set 2	0.76	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.21	1.26
B	Set 1	0.61	0.68	0.75	0.83	0.91	1.00	1.09	1.18	1.28	1.37	1.48
	Set 2	0.75	0.79	0.84	0.89	0.94	1.00	1.05	1.11	1.16	1.22	1.27
C	Set 1	0.50	0.59	0.68	0.78	0.89	1.00	1.11	1.23	1.36	1.48	1.62
	Set 2	0.51	0.60	0.69	0.79	0.90	1.00	1.10	1.22	1.34	1.45	1.58

Table 3.2: The values of α generated by the CIR++ model (A), JCIR++ (B) and TC-CIR++ (C) using Gaussian exposure with different correlations and considering parameter Set 1 and Set 2.

	ρ	-1.0	- 0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8	1.0
A	Set 1	0.67	0.73	0.79	0.86	0.92	1.00	1.07	1.15	1.22	1.30	1.39
	Set 2	0.78	0.82	0.87	0.91	0.95	1.00	1.04	1.09	1.14	1.18	1.23
B	Set 1	0.65	0.72	0.78	0.85	0.92	1.00	1.07	1.15	1.23	1.32	1.41
	Set 2	0.78	0.82	0.86	0.91	0.95	1.00	1.04	1.09	1.14	1.19	1.24
C	Set 1	0.55	0.63	0.72	0.80	0.90	1.00	1.09	1.19	1.30	1.41	1.52
	Set 2	0.56	0.64	0.73	0.81	0.91	1.00	1.09	1.19	1.30	1.40	1.51

Table 3.3: The values of α generated by the CIR++ model (A), JCIR++ (B) and TC-CIR++ (C) using swap exposure with different correlations and considering parameter Set 1 and Set 2.

particular, the time-changed model can reproduce the Basel III case of $\alpha = 1.4$

and can feature larger α's (which corresponds to larger WWR effect) compared to the one prescribed in the Basel approach. To sum up, the TC-CIR++ model is able to generate higher α's compared to the others, without leading to unrealistic values compared to Basel requirements.

3.6 Conclusions

Among the reduced-form intensity models, affine models such as CIR++ process received much attention. The latter consists of a time-homogeneous mean-reverting square-root diffusion shifted in a deterministic way so as to fit a given probability term-structure. In order to increase the attainable volatilities, one can add jumps to the CIR++ dynamics. If the jumps are independent and positive, one obtains the so-called JCIR++ model, which remains affine. The problem however is that the model-implied survival probability curve decreases when increasing the activity of the jumps because they are one-sided. The calibration of the model curve to the market curve being achieved via the shift function, the latter decreases when increasing jumps activity. Consequently, in order to avoid facing "negative intensities", one is limited in the activity of the jumps that can be used. For instance, this specificity limits the attainable values for value-at-risk on CDS, CDS options or wrong-way risk CVA.
An alternative intensity model that allows for two-side jumps is the time-changed idea of Mendoza-Arriaga & Linetsky. Because jumps can be both positive and negative, the model-implied survival probability curve is expected to decrease less rapidly when increasing the jumps' activity compared to the JCIR++. Hence, the problem of facing negative shift function (i.e., "negative intensities") is expected to be less severe, so that larger "volatility's effect" could be generated. This motivates the use of the above model for CVA purposes. Yet, using the time-changed intensity approach in a multivariate framework requires specific precautions. Without specific adjustments indeed, the time-changed technique partly destroys the potential correlation between intensity and exposure which impacts negatively the attainable WWR effect. More importantly, it features forward-looking effects that can generate arbitrage opportunities. In this chapter, we have shown how the time-changed model can be used in a consistent and efficient way by reconstructing the exposure dynamics in a "synchronous" way such that the above problems can be avoided. The computational issue inherent to the time-changed technique is tackled by proposing a variance reduction technique based on adaptive control variate. Eventually, numerical simulations show that under calibration constraint

to a given term structure, the time-changed model can give larger WWR effects compared to the CIR++ and JCIR++ models.
In this work, we have restricted ourselves to unilateral CVA, i.e., disregard the bank's own credit risk. In a bilateral context, two defaults have to be considered. Technically, the extension of the model to multiple counterparties can be addressed by using correlated TC-CIR++ intensities. Dealing with multiple defaults can trigger some computational issues in this approach. Indeed, it does not seem realistic to rely on a single clock applicable to all intensity processes. This in turns means that the synchronization procedure of the Brownian motions has to take place between any two consecutive jumps, independently of the corresponding clock. This can become quickly very time consuming and computationally intensive, and further stresses the need for variance reduction techniques, such as the one proposed in the chapter. Fortunately, CVA is typically calculated at a netting set level, i.e., independently for each counterparty, and then sum up to obtain the CVA of the global portfolio. Hence, in practice, at most one (in the unilateral) or two (in the bilateral) defaults have to be managed at a time.
Analyzing the model's ability to generate higher CDS spread's volatility is another important question from both risk-management and pricing perspectives, which is left for future work.

Appendices

In this section, we give the proof of Lemma 3 and the variance reduction figures using the parameter Set 2 in Table 3.1.

3.A Proof of Lemma 3

Proof. Let $V_t = V_s\, e^{-\frac{\sigma^2}{2}(t-s)+\sigma(W_t^V - W_s^V)}$. Clearly, because V is adapted to $\mathbb{F}^{W^V}$ and $\mathbb{F}^{W^V} \vee \mathbb{F}^{W^\lambda} = \mathbb{F}^{W^V} \vee \mathbb{F}^{W^\perp}$,

$$\mathbb{E}\left[V_t \,|\, \mathcal{F}_s^{W^V} \vee \mathcal{F}_s^{W^\lambda}\right] = \mathbb{E}\left[V_t \,|\, \mathcal{F}_s^{W^V}\right] = V_s\, e^{-\frac{\sigma^2}{2}(t-s)}\mathbb{E}\left[e^{\sigma(W_t^V - W_s^V)}\right] = V_s \tag{3.36}$$

However, the increments of W^V after $\theta_s > s$ are independent both from $\mathcal{F}_s^{W^V}$ and from $\mathcal{F}_{\theta_s}^{W^\lambda} := \mathcal{F}_s^{W^{\lambda^\theta}}$. Hence, using tower law,

$$\begin{aligned}
\mathbb{E}\left[V_t \,|\, \mathcal{F}_s^{W^V} \vee \mathcal{F}_s^{W^{\lambda^\theta}}\right] &= V_s\, e^{-\frac{\sigma^2}{2}(t-s)}\mathbb{E}\left[e^{\sigma(W_t^V - W_s^V)}|\mathcal{F}_{\theta_s}^{W^\lambda}\right] \\
&= V_s\, e^{-\frac{\sigma^2}{2}(t-s)} e^{\frac{\sigma^2}{2}(t-\theta_s)}\mathbb{E}\left[e^{\sigma(W_{\theta_s}^V - W_s^V)}|\mathcal{F}_{\theta_s}^{W^\lambda}\right] \qquad (3.37)\\
&= V_s\, e^{-\frac{\sigma^2}{2}(\theta_s-s)}\mathbb{E}\left[e^{\sigma(W_{\theta_s}^V - W_s^V)}|\mathcal{F}_{\theta_s}^{W^\lambda}\right] .
\end{aligned}$$

Assuming in the sequel that $\rho \neq 0$,

$$\mathbb{E}\left[e^{\frac{\sigma}{\rho}[(W_{\theta_s}^\lambda - W_s^\lambda)-\sqrt{1-\rho^2}(W_{\theta_s}^\perp - W_s^\perp)]}|\mathcal{F}_{\theta_s}^{W^\lambda}\right] = e^{\frac{\sigma}{\rho}(W_{\theta_s}^\lambda - W_s^\lambda)}\mathbb{E}\left[e^{\frac{-\sigma\sqrt{1-\rho^2}}{\rho}(W_{\theta_s}^\perp - W_s^\perp)}|\mathcal{F}_{\theta_s}^{W^\lambda}\right] \tag{3.38}$$

Observe that $W_{\theta_s}^\perp - W_s^\perp$ is not independent from $\mathcal{F}_{\theta_s}^{W^\lambda}$ as this information set gives us the value of $W_{\theta_s}^\lambda - W_s^\lambda$:

$$W_{\theta_s}^\lambda - W_s^\lambda = \rho\left(W_{\theta_s}^V - W_s^V\right) + \sqrt{1-\rho^2}\left(W_{\theta_s}^\perp - W_s^\perp\right)$$

The computation of the above conditional expectation amounts to evaluate the moment generating function (MGF) $\varphi\left(-\frac{\sigma\sqrt{1-\rho^2}}{\rho}\right)$ associated to the Normal variable $W_{\theta_s}^\perp - W_s^\perp$ for which one knows the value of its weighted sum with another independent variable. More explicitly, we are looking for the MGF of $X = \sqrt{1-\rho^2}\sqrt{\theta_s - s}Z_1$ such that $X + Y = W_{\theta_s}^\lambda - W_s^\lambda$ with $Y = \rho\sqrt{\theta_s - s}Z_2$, Z_1, Z_2 independent, identically distributed standard Normal. It can be shown (see, e.g., Vrins 2018) that, given $X + \omega_2 Z_2 = c$, $X = \omega_1 Z_1 \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma})$ with $\tilde{\sigma}^2 = \left(\omega_1^{-2} + \omega_2^{-2}\right)^{-1}$ and $\tilde{\mu} = c\tilde{\sigma}^2/\omega_2^2$. Using the values of ω_1, ω_2 and $c = W_{\theta_s}^\lambda - W_s^\lambda$,

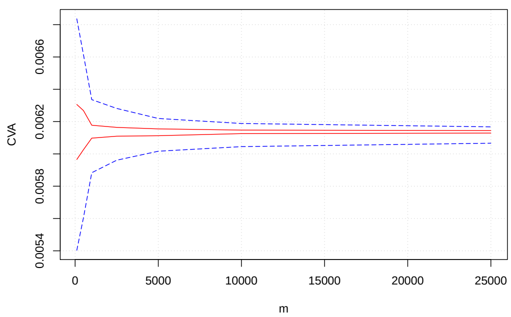
$$\varphi(t) = e^{(1-\rho^2)(W_{\theta_s}^\lambda - W_s^\lambda)t + \rho^2(1-\rho^2)(\theta_s - s)\frac{t^2}{2}} \tag{3.39}$$

Eventually,

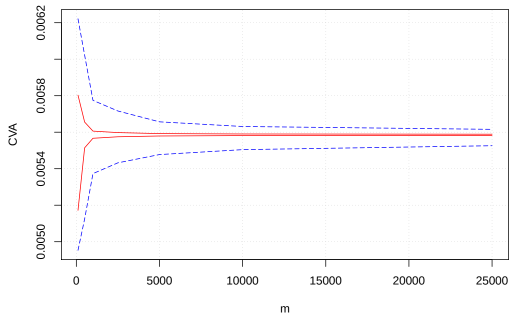
$$\begin{aligned}
\mathbb{E}\left[V_t \,|\, \mathcal{F}_s^{W^V} \vee \mathcal{F}_s^{W^{\lambda^\theta}}\right] &= V_s\, e^{-\frac{\sigma^2}{2}(\theta_s - s)} e^{\frac{\sigma}{\rho}(W^\lambda_{\theta_s} - W^\lambda_s)} e^{(1-\rho^2)(W^\lambda_{\theta_s} - W^\lambda_s)\left(-\frac{\sigma\sqrt{1-\rho^2}}{\rho}\right)} \\
&\quad \times e^{\rho^2(1-\rho^2)(\theta_s - s)\left(-\frac{\sigma\sqrt{1-\rho^2}}{\rho\sqrt{2}}\right)^2} \\
&= V_s\, e^{\frac{\sigma}{\rho}\left[1-(1-\rho^2)^{\frac{3}{2}}\right](W^\lambda_{\theta_s} - W^\lambda_s) + \frac{\sigma^2}{2}\left[(1-\rho^2)^2 - 1\right](\theta_s - s)} \\
&\neq V_s\,.
\end{aligned} \tag{3.40}$$

□

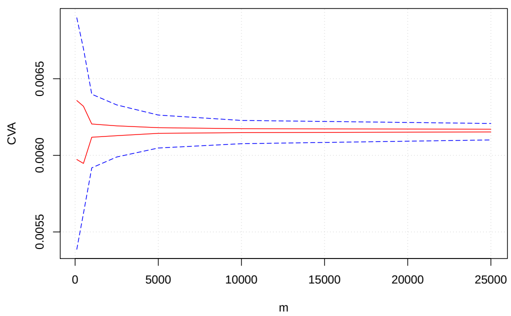
3.B Variance reduction figures using Set 2



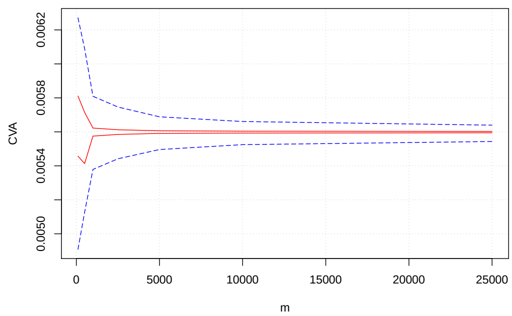
(a) CIR, $\rho = 0.8$



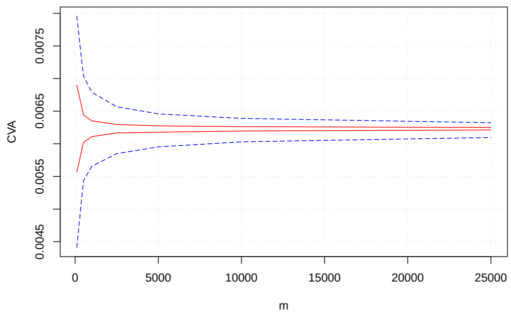
(b) CIR, $\rho = 0.4$



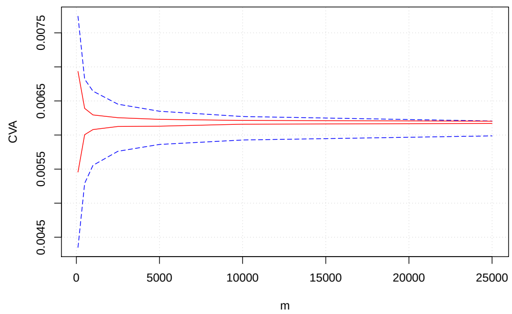
(c) JCIR, $\rho = 0.8$



(d) JCIR, $\rho = 0.4$



(e) TC-CIR, $\rho = 0.8$



(f) TC-CIR, $\rho = 0.4$

Figure 3.4: Confidence interval at level 95% of CVA figures for adaptive control variate (red) and standard Monte Carlo (dotted blue) using parameter Set 2. Profiles: 3Y Gaussian exposure, $\sigma = 8\%$. Hazard rate is $h(t) = 5\%$.

Chapter 4

Affine term-structure models: a time approach with perfect fit to market curves

4.1 Introduction

Model calibration is a standard problem in many areas of finance (Brigo and Mercurio, 2006; Joshi, 2003; Veronesi, 2010). It consists of tuning a model such that it "best fits" market quotes at a given time. As an example, financial markets provide a set of prices associated with liquid instruments, that openly trade on the market. Alongside with risk management (hedging), the main purpose of a model here is to act as an "interpolation/extrapolation" tool, i.e., to obtain the value of products at a given time t for which the market does not disclose prices in a transparent way. This could happen because either the product to be priced is "exotic" (i.e., is too "special", it does not quote openly on a platform, only on a bilateral basis) or because its cashflow schedule is not in line with the products that currently trade openly at t (a situation that commonly happens since products that were "standard" at inception, may have time-to-expiry or moneyness levels that are no longer "standard" afterwards).

Mathematically, model calibration is nothing but an optimization problem. Starting from a set of prices quoted on the market (called "market prices") for a set of specific financial products (called "calibration instruments"), model calibration consists of computing the model parameters such that the prices generated by the model (called "model prices") best fit to the market prices, according to some error function. Model calibration is crucial in finance; it is strongly related to arbitrage opportunities. In practice, only models that are able to reproduce the market prices of "simple instruments" (either in a perfect way, or at least up to the bid-ask spread) are trustworthy enough when it comes to pricing other instruments. For instance, one can price exotic derivatives (like barrier options) using stochastic volatility model like Heston in a semi-analytical way Carr and Madan, 1999; Heston, 1993. The parameters of the Heston model will be obtained by "calibration" to a volatility surface, i.e., to a set of liquid ("plain vanilla") options, like European calls and puts of various strikes and maturities. The justification behind this is that in a no-arbitrage, complete market setup, the price of an option can be obtained by computing the cost of setting up a self-financing hedging strategy. This cost depends on the prevailing prices of the hedging instruments. If the model fails to correctly price the latter, there is no chance it can correctly price the option.

In this work, we focus on financial calibration problems arising in other asset classes: interest-rates and credit Brigo and Mercurio, 2006; Duffie and Singleton, 2003. When specifying an interest rate model to price a derivative on, say, the Libor 3M index, one needs to make sure that the model generates a discount curve that is in line with that extracted from market quotes of simpler Libor 3M-indexed products. In this case, the set of calibration instruments could be forward rate agreements (FRA), interest rate swaps (IRS), as well as vanilla cap/floors or swaptions. Similarly, adjusting the value of derivatives for counterparty risk (a problem known as credit valuation adjustment, or CVA) generally involves a stochastic model to represent the default of the counterparty with whom the trade is executed. The default probability of the counterparty can be extracted from a set of calibration instruments, which prices are driven by the default likelihood of the counterparty, e.g., corporate bonds of credit default swaps (CDS). In this context, the default model must be "calibrated" in such a way that the default probability curve generated by the stochastic model agrees with that implied from the prices of the corresponding instruments (see, e.g., Gregory, 2010 and Stein and Pong, 2011 for a general overview of CVA and Brigo et al., 2014 for a discussion of bilateral CVA in presence of collateralization agreements).

The calibration constraint rises practical issues. Indeed, the models that are actually used in the industry must have a tractability that is compatible with real-time pricing but, as explained above, must be flexible enough to match the information conveyed by the calibration instruments. Affine term structure models (ATSM) have been extensively used in fixed income modeling because of their analytical tractability. See, e.g., Duffie et al., 2003 and Duffie and Kan, 1996 for an excellent review and a mathematical analysis of this class of processes. In practice, homogeneous affine jump diffusion (HAJD) models are extremely popular. The Vasicek (Ornstein-Uhlenbeck) model Vasicek, 1977 is a short-rate model being widely used in both industry and academia. It is a time-homogeneous affine diffusion model that postulates Gaussian dynamics. If negative rates need to be ruled out, positive dynamics like the CIR (Cox-Ingersoll-Ross, also known as square-root diffusion, SRD) Cox et al., 1985 can be preferred, possibly with independent compounded Poisson jumps (JCIR or SRJD). However, it is in general impossible with either models (even in a multi-factor setup) to achieve a perfect fit: the flexibility of HAJD is limited, they are in general unable to generate a given discount curve. The same problem arises when dealing with credit derivatives: it is generally impossible to make sure that the default intensity process, modeled with HAJD dynamics, will generate a default probability curve that is in line with the corresponding curve, exogenously given by the market via the calibration instruments.

Several routes can be followed to deal with this issue. The first one consists of disregarding this lack of flexibility. Nevertheless, working with a model that fails to yield a perfect fit to the market is often unacceptable in practice. Indeed, as explained above, the models are used to value derivatives positions, and a mismatch with the market can introduce a tremendous bias in the valuation of the book of companies or financial institutions. Another possibility is to significantly increase the complexity of the models. This is often to be avoided in practice, for computational, identification or over-fitting issues. A trade-off consists of extending the "simple models" in such a way that they can fit to the market. In fact, several authors show that a great flexibility can be obtained by shifting HAJD models in a deterministic way. Dybvig, 1997 for instance shows that the term structure of interest rates can be reproduced by adding a deterministic shift to Ho and Lee, 1986 or Vasicek processes. Later, Brigo and Merccurio, 2001 extended this idea to a broader class of models, thereby providing a simple but very clever solution to the calibration problem. Instead of considering a HAJD, one could simply adjust it with a deterministic function φ: the resulting process will have the required flexibility. When shifted this way, the Vasicek, CIR and JCIR models respectively correspond to the Hull-White, the CIR++ or the JCIR++ (a.k.a. SSRJD) models Brigo and Mercurio, 2006. This trick is actually very powerful: it solves the calibration problem at no cost, since the model's dynamics remain affine. Moreover, the shift function is known analytically, as a function of the parameters of the underlying HAJD and the market curve to be fitted by the model.

Yet, this approach suffers from an important limitation. Because of the shift, there is no reason that the range of the shifted process agrees with that of the underlying HAJD process. For instance, shifting a positive process with a deterministic function may result in a process that could take on negative values. It all depends on the mismatch between the information conveyed by the calibration instruments on the one hand, and the parameters of the underlying HAJD on the other hand. In general, there is no reason to believe that the implied shift function will preserve the range of the HAJD model. This is problematic in many cases, and in credit risk modeling in particular: negative default intensities, for instance, make no sense. To circumvent this issue, one could think of adding a non-negativity constraint on the shift in the calibration step. But, as we will show, this drastically restricts the parameters of the underlying HAJD, hence the randomness embedded in the model. This explains why this solution is often not considered by practitioners: the shift approach (without positivity constraint) remains the standard approach, even if positivity is required, theoretically speaking. It seems that in absence of a valid alternative, one actually prefers to rely on a model providing a perfect fit, even though the latter suffers from theoretical inconsistencies.

In this chapter we introduce an alternative to the deterministic shift. Using an equally simple – but intrinsically different – technique, we adjust a HAJD so as to allow for a perfect fit to a given market curve, without affecting the model's tractability, but also without introducing the aforementioned inconsistencies. More specifically, instead of shifting a HAJD, we time-change it. Time change techniques were first studied in 1965 Dambis, 1965; Dubins and Schwartz, 1965. The first application to finance dates back from the early 2000. Geman et al. used Lévy processes and interpreted the new time scale as the business time, in contrast with the calendar time Geman et al., 2001. This was then applied to stochastic volatility models Carr et al., 2003. Thanks to subordinated Lévy models, the authors introduced the leverage effect, as well as a long-term skew. Many other financial applications of time change techniques can be found in the review Swishchuk, 2016. More recently, Mendoza-Arriaga and Linetsky used stochastic time change processes to introduce two-side jumps in positive processes. The analytical tractability of the resulting model is preserved to some extend. This model has been recently applied to counterparty credit risk Mbaye and Vrins, 2018. In this work, we exploit the time change idea in yet another way, to solve a completely different problem. Our purpose is to time-change HAJDs so as to obtain models with the desired calibration flexibility, without affecting tractability and preserving the range of the original process. The intuition is that by slowing down or speeding up the time of the latent HAJD, at the appropriate rate, one would obtain a model that could fit most discount curves, and actually every default probability curve. Moreover, the time change function is easily found using simple numerical methods (namely, inversion of easy functions or ordinary differential equation). Eventually, our time-changed HAJD is proven to feature larger implied volatilities compared to the corresponding valid (i.e., non-negative) shifted HAJD. To illustrate the power of our approach, we provide two applications taken from credit: pricing of CDS options and computation of derivatives pricing accounting for counterparty risk under exposure-credit dependence (wrong-way risk, WWR). In either cases, all the considered default models perfectly fit the risk-neutral default probability curve extracted from market quotes associated to the CDS of the reference entity. The obtained results illustrate the nice feature of large implied volatilities : they are able to generate larger option prices compared to the shift approach calibrated on a same probability curve under non-negativity constraint.

Eventually, observe that although we focus on examples featuring reduced-form models when pricing of credit-sensitive instruments, our approach is of potential use for other models, in many areas of finance and insurance. Ongoing work suggests that it can be applied to other default models, including the firm-value (structural) models Merton, 1974, but also to linear-rational (polynomial) models Filipovic et al., 2016. Other models could be considered as well, like Jeanblanc and Vrins, 2018 or Crépey et al., 2013.In terms of applications, the proposed method can be used in life insurance, to calibrate mortality rate to mortality tables. Our time-changed process could also be used to model prepayment rates in mortgage-backed securities (MBS). These products naturally exhibit a negative convexity due to a the negative relationship between interest and prepayment rates: householders tend to refinance their loans when interest rates drop. This calls for a stochastic prepayment (i.e., positive) rate, that will be negatively correlated with interest rates, and which parameters could be calibrated so as to agree with the averaged values given in the PSA measure, the indicator attached to MBS securities that characterizes the prepayment speed in MBS Veronesi, 2010. Eventually, the proposed method could be applied in many other applications, including the modeling of performance degradation of devices or materials through time, which average outstanding performances evolution are given according to some quality standards.

The chapter is organized as follows. In Section 4.2 the calibration problem is introduced and two specific cases (cashflow discounting and probability curves) are discussed. We then recall in Section 2.3 how a shifted version of time-homogeneous affine jump diffusions can fit every discount curve. We pay specific attention to the case where the resulting process needs to meet a positivity constraint. We then introduce in Section 4.3 our alternative model, specifically devoted to this case, focusing on the most common HAJD, namely the Vasicek and JCIR (generalizing the CIR) models. Eventually, we compare in Section 4.4 our model's performance to that of the shift approach on three different pricing problems taken from credit risk: CDS curve calibration, pricing of CDS options and pricing of credit valuation adjustment under wrong-way risk.

4.2 The calibration problem

Consider a given time-s *market curve* $P_s^{market}(t)$, $t \geq s$. The calibration problem consists of finding, for a given model, the (set of) parameter(s) $\Xi = \Xi^\star$ such that the corresponding *model curve* $P_s^{model}(t) = P_s^{model}(t;\Xi)$ "best fits" the market curve, according to some criterion. Mathematically speaking, this is an optimization problem that consists of finding a set of parameters that minimizes an error function between model and market values,

$$\Xi^\star := \arg\min_{\Xi} \left\| P_s^{model}(\cdot;\Xi) - P_s^{market}(\cdot) \right\| , \tag{4.1}$$

where $\|f(\cdot) - g(\cdot)\|$ represents a divergence measure between two functions f, g. In practice, one often computes the mean-square error (MSE) between f and g on a set of maturities $\mathcal{T} := \{T_1, \ldots, T_n\}$:

$$\|f(\cdot) - g(\cdot)\| := \frac{1}{n}\sum_{i=1}^{n} \left(f(T_i) - g(T_i)\right)^2 . \tag{4.2}$$

A model with parameter Ξ is said to *perfectly fit* the market up to horizon T whenever $P_s^{model}(t;\Xi) = P_s^{market}(t)$ for all $s \leq t \leq T$ or using a shorthand notation, $P_s^{model} \equiv P_s^{market}$.

Remark 4. *It is important to notice that the calibration problem consists of adjusting the parameters of a model based on a current set of prices. Therefore, it differs from the usual estimation problem, which refers to the calibration of parameters using historical time-series, using, e.g., maximum likelihood or generalized method of moments. Here, one starts by choosing a parametric model that is somehow* *supported by empirical evidences (like mean-reverting square-root), and then estimates the parameters not from history, but using a market-implied (risk-neutral) setting, i.e., directly from current market prices.*

4.2.1 Setup

A model can be either static or dynamic. For instance, the Nelsen-Siegel model Nelson and Siegel, 1987 postulates a parametric form for the yield curve, but is a static model: the resulting curve does not correspond to the yield curve generated by the dynamics of a stochastic model. We focus on dynamic models in the sequel.

We consider the Cox setup of Section 2.2.2 where $\mathbb{G}$ corresponds to the filtration generated by the stochastic market variables (risk factors, prices, interest rates, default intensities, default event, etc), and all process are $\mathbb{F}$-adapted except those featuring the default time τ. In this case, the immersion property holds between $\mathbb{F}$ and $\mathbb{G}$.

In the sequel, we shall focus on a specific class of P^{model} and P^{market} functions: we assume they are *discount curves*, a set of functions that we now define.

Definition 4 (Discount curve)**.** *A time-s discount curve is any differentiable function of the form*

$$P_s : [s, \infty) \to \mathbb{R}_0^+, \quad t \mapsto P_s(t)$$

satisfying $P_s(s) = 1$.

In the specific $s = 0$ case, a *time-0 discount curve* $P_0(t)$ is simply called a *discount curve* and is noted $P(t)$, assuming implicitly that $t \geq 0$. Any time-s discount curve admits an exponential-integral from, as we now show.

Lemma 4. *Every time-s discount curve P_s admits a representation in terms of time-s instantaneous forward rate curve f_s:*

$$P_s(t) = e^{-\int_s^t f_s(u)du}, \quad t \geq s.$$

If moreover $P_s(t)$ is strictly decreasing on (s, ∞), then $f_s(t)$ is strictly positive for all $t > s$.

Proof. Since $P_s(t) > 0$ for all $t \geq s$ and $P_s(t)$ is differentiable with respect to t on (s, ∞), then one can define the time-s instantaneous forward rate function as $f_s(t) := \frac{-1}{P_s(t)} \frac{d}{dt} P_s(t) = -\frac{d}{dt} \ln P_s(t)$ for all $t > s$; the value $f_s(s)$ is not identified but can be defined by, e.g., the limit as $t \downarrow s$. Moreover, if P_s is strictly decreasing on (s, ∞) then $f_s(t) > 0$ for all $t > s$. $\square$

Discount curves are of paramount importance in finance. As suggested by the name, $P_s(t)$ allows one to compute the time-s value of a cashflow paid at time $t \geq s$, both in a credit risk-free and credit risky setup. As a special case of the second framework, they encompass survival probability curves, defined as one minus cumulative distribution functions. This is elaborated in the next two subsections.

Discounting in a default-free market

In this application, $P_s(t)$ stands for the time-s price of a risk-free zero-coupon bond (ZCB) with maturity t and face value 1, denominated in a given currency. In particular, $P_s^{market}(t)$ and $P_s^{model}(t;\Xi)$ respectively give the market and the model prices of that instrument.

Indeed, in a no-arbitrage setup, the price of a financial instrument paying a single cashflow (payoff) at a given maturity is given by the risk-neutral conditional expectation of the payoff, discounted at the risk-free rate from the payment date (maturity) back to the valuation date. Adopting a short-rate model, $P_s^{model}(t)$ corresponds to the $\mathbb{Q}$-expectation of the stochastic discount factor, the negative exponential of the risk-free short-rate process r, integrated from the valuation time s up to the payment time t, conditional upon the information prevailing at the pricing time Brigo and Mercurio, 2006:

$$P_s^{model}(t) = \mathbb{E}\left[1\beta_s/\beta_t \middle| \mathcal{G}_s\right] = \mathbb{E}\left[e^{-\int_s^t r_u du}\middle| \mathcal{G}_s\right] =: P_s^r(t) \ .$$

where β is given in (2.1). In this context, we aim at finding a model x to depict the risk-free short rate dynamics r that would be tractable enough, and provide a perfect fit to any yield curve, the curve that gives the set of prices of ZCBs with increasing maturities.

Discounting in a defaultable market

Adopting the same framework as before, the time-s price of a zero-coupon bond paying one unit of currency at time $t \geq s$ *contingent on the fact that the issuer doesn't default prior to the payment date* is given by a similar expression as before. It suffices to replace the risk-free payoff 1 by the risky one, namely $\mathbb{1}_{\{\tau>t\}}$, where the random variable τ represents the default time of the issuer and $\mathbb{1}_A$ is the indicator function defined as 1 if A is true and zero otherwise. Mathematically,

$$P_s^{model}(t) = \mathbb{E}\left[\mathbb{1}_{\{\tau>t\}}\beta_s/\beta_t\middle| \mathcal{G}_s\right] =: \bar{P}_s^r(t) \ .$$

In such a context, $\bar{P}_s^r$ corresponds to a risky discounting, where the term *risk* is referring to the possibility for the issuer not to meet her financial obligations.

To proceed, we need to model the default event. To that end, we consider a Cox model where the default time τ is given by (2.22).

The function $\bar{P}_s^r(t)$ can be proven to be a time-s discount curve in many cases. To show this, we first notice that one can replace $\mathcal{G}_s$ by $\mathcal{F}_s$ in the expression providing the time-s price of the risk-free ZCB since immersion holds:

$$P_s^r(t) = \mathbb{E}\left[\beta_s/\beta_t \middle| \mathcal{G}_s\right] = \mathbb{E}\left[\beta_s/\beta_t \middle| \mathcal{F}_s \vee \mathcal{D}_s\right] = \mathbb{E}\left[\beta_s/\beta_t \middle| \mathcal{F}_s\right] \ .$$

Applying the Key lemma (see (2.14)) to the risky ZCB formula above yields

$$\bar{P}_s^r(t) = \mathbb{1}_{\{\tau>s\}}\,\mathbb{E}\left[e^{-\int_s^t \lambda_u du}\beta_s/\beta_t \middle| \mathcal{F}_s\right] = \mathbb{1}_{\{\tau>s\}}\,\mathbb{E}\left[e^{-\int_s^t (\lambda_u+r_u)du} \middle| \mathcal{F}_s\right] =: \mathbb{1}_{\{\tau>s\}} P_s^{\lambda+r}(t).$$

Eventually, in the special case where $r \equiv 0$, $\beta \equiv 1$ so that $\bar{P}_s^r(t)$ collapses to

$$\bar{P}_s^r(t) = \mathbb{E}\left[\mathbb{1}_{\{\tau>t\}}|\mathcal{G}_s\right] = \mathbb{Q}\left(\tau > t|\mathcal{G}_s\right) = \mathbb{1}_{\{\tau>s\}} P_s^{\lambda}(t) \ .$$

Hence, on the event $\{\tau > s\}$, $\bar{P}_s^r(t) = P_s^{\lambda}(t)$ agrees with the survival probability function associated with τ, conditional upon $\mathcal{G}_s$.

In this specific context, we are interested in a model x to depict the dynamics of the intensity process λ that would be tractable enough, and provide a perfect fit to any valid survival probability curve extracted from the prices of defaultable instruments like corporate bonds or credit default swaps (CDS).

Remark 5. *Notice that in contrast to rates, that can – and some of them currently do – take negative value, non-negativity is a formal requirement when x represents an intensity process λ. A default model featuring "negative intensities" is theoretically flawed, and is problematic. Indeed, modelling the event $\{\tau > t\}$ as $\{\Lambda_t < \mathcal{E}\}$ yields a survival indicator process $\mathbb{1}_{\{\tau>t\}}$ that might jump both up and down, i.e., the reference entity could be "brought back to life". One could of course think of replacing the default event using a first passage time, thereby revisiting the default time definition as $\tau := \inf\{t \geq 0 : \Lambda_t \geq \mathcal{E}\}$. However, one looses the analytical tractability for the survival probability since in this case, $\mathbb{Q}(\tau > t)$ does no longer agree with $\mathbb{Q}(\Lambda_t < \mathcal{E}) = \mathbb{E}\left[e^{-\int_0^t \lambda_u du}\right] = P_0^{\lambda}(t)$. The $x \geq 0$ constraint is also a natural requirement when it represents a credit spread.*

4.2.2 The perfect fit problems

Equation (4.1) suggests that the calibration problem consists of finding the parameters of a given model to minimize the discrepancies between market and model curves up to a time horizon T. However, it is clear that the choice of the model class will also have a substantial impact. Indeed, depending on the model chosen, the minimum of the error function could be large, small or even zero, in which case the perfect fit is obtained.
Inspired by the financial problems mentioned in Section 4.2.1, we consider the following problems directly related to the perfect fit constraint (up to a given time horizon T, that is implicit in the sequel). The first one does not impose any constraint on the process x to consider.

Problem 1. *Find a tractable process x satisfying*

$$P_s^x(t) := \mathbb{E}\left[e^{-\int_s^t x_u du}\middle|\mathcal{F}_s\right] = P_s(t)$$

for $t \in [0,T]$ and every given discount curve P_s.

Depending on the application at hand, one may need to impose additional constraints on x. As suggested by the risky discounting example, non-negativity is a crucial one. This leads us to consider a second (constrained) problem.

Problem 2. *Find a tractable positive process x (i.e., such that $\mathbb{Q}(x_t \geq 0) = 1$ and $\mathbb{Q}(x_t > 0) > 0$ for all $t \in [0,T]$) satisfying*

$$P_s^x(t) := \mathbb{E}\left[e^{-\int_s^t x_u du}\middle|\mathcal{F}_s\right] = P_s(t)$$

for $t \in [0,T]$ and every strictly decreasing discount curve P_s.

In either problems, *tractability* refers to the fact that model calibration (4.1) – that features an optimization over the parameter space – is not too cumbersome, computationally. Solving this optimization problem typically requires many iterations, hence numerous evaluations of the objective function. This suggests that a highly desirable feature of the model is to admit a closed form expression for P^{model} or, at least, that the latter can be computed without having to rely on time-consuming numerical methods like, e.g., Monte Carlo simulations.

4.2.3 Homogeneous affine models and shift extension

In order to solve these two problems, we consider what is probably the most tractable family of models, namely *affine processes* and, more specifically *time-homogeneous affine processes*. Indeed, for a one-factor affine model $y := (y_t)_{t\in[0,T]}$, many expressions are available analytically, as well as for its integrated version $Y := (Y_t)_{t\in[0,T]}$, $Y_t := \int_0^t y_u du$. In particular, $P_s^y(t) := \mathbb{E}\left[e^{-\int_s^t y_u du}\middle|\mathcal{F}_s\right]$ is merely the conditional moment generating function of $Y_t - Y_s$, $t \geq s$.

We refer to Section 2.3.1 for more details where $P_s^y(t)$ is given in (2.27) and has a closed-form:

$$P_s^y(t) = e^{A_s^y(t;\Xi)-B_s^y(t;\Xi)y_s} =: P_s^y(t;\Xi) \tag{4.3}$$

where Ξ is the (set of) parameter(s) governing y and A_s^y, B_s^y are differentiable functions satisfying $A_s^y(s;\Xi) = B_s^y(s;\Xi) = 0$.
When the initial value y_0 is part of the parameters, we note the parameter set Ξ_0.

The analytical form (4.3) facilitates in a tremendous way calibration procedures such as (4.1) when the considered P^{model} function takes the form of the conditional expectation in (2.27).
For such processes, the function P_s^y is thus well-defined for every s, is positive, and satisfies $P_s^y(s) = 1$. It is therefore a time-s discount curve in the sense of Definition 4 since it is obviously differentiable on (s,∞). For instance, P_s^r and P_s^λ in the above two examples are time-s discount curves whenever r and λ are affine processes, respectively.

We recall (and drive) some the properties of these processes in Appendix 4.A for further references.

Observe that the sum of two affine processes x, y is, generally speaking, not an affine process. Hence, it is not clear whether the risky discounting curve $P_s^{\lambda+r}$ is a time-s discount curve, even in the simple case where both r, λ are affine processes. Obviously, this holds true in the independence case since $P_s^{x+y} = P_s^x P_s^y$, and the product of two time-s discount curves is itself a time-s discount curve.

In the sequel, we consider a specific pricing time, say $s = 0$ without loss of generality, and drop the observation time subscript for conciseness. The next lemma provides sufficient conditions on y for P^y to be a discount curve in the general case.

Lemma 5. *Let T be a fixed time horizon. Then, P^y is a discount curve whenever y is positive and* $\sup_{t\in[0,T]} y_t$ *is integrable.*

Proof. Let us fix $t \leq T$. Hence,

$$\left|\frac{d}{dt}e^{-\int_0^t y_u du}\right| = |y_t e^{-\int_0^t y_u du}| \leq y_t \leq \sup_{t\in[0,T]} y_t$$

for all $t \in [0,T]$. Noting that $\sup_{t\in[0,T]} y_t$ is integrable, one can use Lemma 13 with $\Psi(t,w) \leftarrow e^{-\int_0^t y_u(w) du}$ and $Z(\omega) \leftarrow S_T^y := \sup_{t\in[0,T]} y_t$, justifying the swap between the derivative and expectation operators:

$$\frac{d}{dt}P^y(t) = \frac{d}{dt}\mathbb{E}\left[e^{-\int_0^t y_u du}\right] = \mathbb{E}\left[\frac{d}{dt}e^{-\int_0^t y_u du}\right] = -\mathbb{E}\left[y_t e^{-\int_0^t y_u du}\right] ,$$

where the right-hand side is bounded by the expectation of Z, which is integrable. $\square$

Remark 6. *The derivative of P^{x+y} is proven to exist similarly whenever x and y satisfy all the assumptions of Lemma 5.*

In order for this assumption (integrability of the sup) to be useful, it needs to be "checkable'. We would like to give easier conditions (e.g., from the SDE of y) that would guarantee that $S_T^y := \sup_{t\in[0,T]} |y_t|$ satisfies $\mathbb{E}[S_T^y] < \infty$.

Lemma 6. *Let W be a Brownian motion, J a compound Poisson process with constant jump intensity ω and the jump sizes are exponentially distributed with mean α, and y, a positive process solving*

$$dy_t = \mu(t,y_t)dt + \sigma(t,y_t)dW_t + dJ_t$$

where y_0 is positive, $\mathbb{E}[\int_0^T |\mu(t,y_t)|dt] < \infty$ and $\mathbb{E}[\int_0^T \sigma^2(t,y_t)dt] < \infty$. Then,

$$\mathbb{E}[|S_T^y|] = \mathbb{E}[S_T^y] < \infty \quad \textit{where } S_T^y := \sup_{t\in[0,T]} |y_s| .$$

Proof. See Section 4.B. $\square$

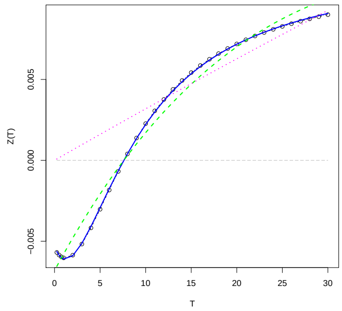
One can check that P^{x+y} is a discount curve when x, y are HAJD and positive, possibly driven by correlated Brownian motions. Indeed, they satisfy the assumptions of Lemma 6.

HAJD models like VAS, CIR and JCIR seem appropriate to solve problems 1 and 2. Unfortunately, they do not allow for a perfect fit to a given discount curve P, except in very special cases. Indeed, it is not possible in general, for such type of processes x, to find Ξ (or Ξ_0) such that $P^{model} := P^x(\cdot;\Xi) \equiv P$, even up to a finite horizon T.

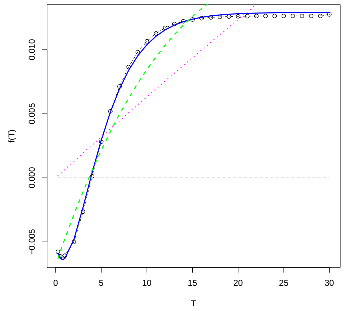
Example 1. *We illustrated the lack of flexibility of the VAS and CIR model to fit interest rates curves of AAA-rated euro area central government bonds retrieved from the European Central Bank website on January 29th, 2019. The market provides the yield curve (Z^{market}) – which can easily be converted into a ZCB price curve (P^{market}) – as well as the instantaneous forward curve (f^{market}). One can see from Fig. 4.1 that whatever the type of curve that is fitted ($Z = Z^{market}$, $f = f^{market}$ or $P = P^{market}$), none of the VAS and CIR models succeed to replicate the market quotes.*[1] *For the sake of comparison, we also plot the best Nelsen-Siegel (NS) curves Nelson and Siegel, 1987. Although it does not perfectly fit market curves either, it provides a much better approximation. VAS, CIR and NS are estimated using a MSE criterion similar to* (4.2) *with f, g denoting model and market curves, and $\mathcal{T}$ represents the maturities for which a zero-rate is available.*[2]

[1]The authors are grateful to Linqi Wang for providing the exhibits of Fig. 4.1.

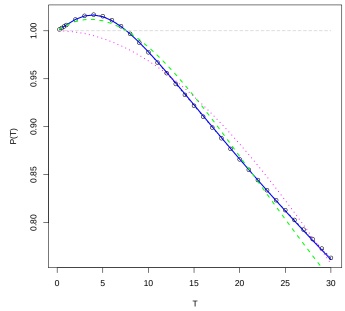
[2]Obviously, a non-equally weighted approach would be preferred in the case where the fit needs to be better on some maturities.



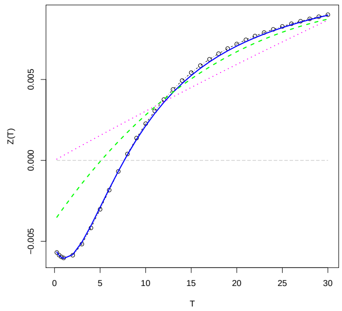
(a) Z curves (fit on Z^{market})



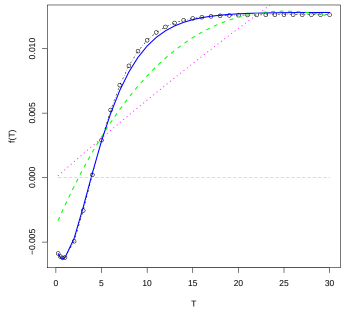
(b) f curves (fit on Z^{market})



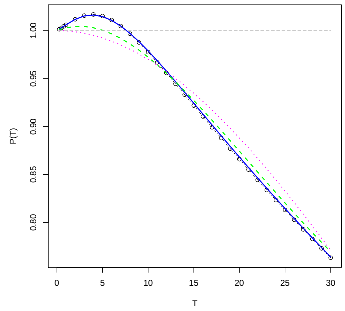
(c) P curves (fit on Z^{market})



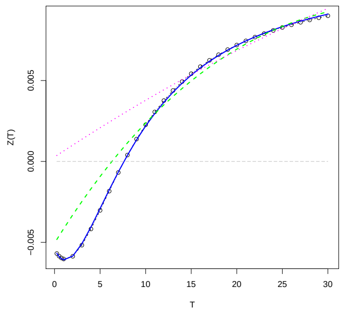
(d) Z curves (fit on f^{market})



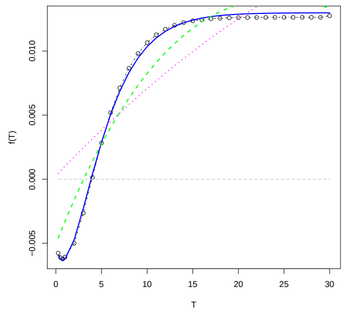
(e) f curves (fit on f^{market})



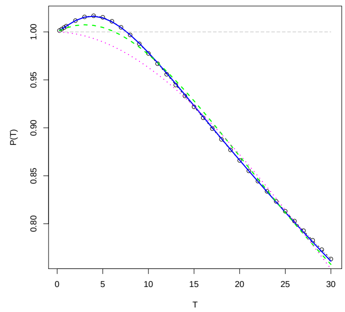
(f) P curves (fit on f^{market})



(g) Z curves (fit on P^{market})



(h) f curves (fit on P^{market})



(i) P curves (fit on P^{market})

Figure 4.1: Yield curves (Z, left), instantaneous forward curves (f, center) and ZCB price curves (P, right) generated by the Vasicek (green, dashed), CIR (magenta, dotted) and the Nelsen-Siegel (solid, blue) models when they are fitted on Z^{market} curve (top), f^{market} curve (middle) and P^{market} curve (bottom) shown in black (circles). The yield curve Z^{market} downloaded from the European Central Bank website and corresponds to AAA-rated euro area central government bonds. Although Nelsen-Siegel proves to be relatively satisfactory compared to the curves generated by the affine diffusions models, none of them provide a perfectly fit with the market curve, whatever the calibration target chosen: Z^{market}, f^{market} or P^{market}.

The above example shows that these HAJDs lack flexibility to replicate market curves. This can be addressed by performing a simple extension, introduced in the next section.

Another class of models that address the calibration issues such as Problem 1 is the general deterministic approach. We refers to Section 2.3.2 for more details. In this case, the model $x := x^\varphi$ is defined as in (2.29). The resulting process x^φ remains affine and is thus analytically tractable (see (2.30)).

Clearly, the dynamics of x^φ are easily obtained from that of y. Indeed, assuming

$$dy_t = \mu(t, y_t)dt + \sigma(t, y_t)dW_t + dJ_t \;, \tag{4.4}$$

the dynamics of x^φ read, when φ is differentiable, as

$$dx_t^\varphi = dy_t + \varphi'(t)dt = (\mu(t, x_t^\varphi - \varphi(t)) + \varphi'(t))dt + \sigma(t, x_t^\varphi - \varphi(t))dW_t + dJ_t,\; x_0^\varphi = y_0 + \varphi(0).$$

It can be shown that in the particular case where y is a HAJD, then x^φ remains an AJD, even though no longer homogeneous, unless $\varphi(t)$ is constant. For instance, if the dynamics of y obey (2.28), then x^φ is governed by the same type of dynamics since

$$dx_t^\varphi = \left(a^\varphi(t) + b(t)x_t^\varphi\right)dt + \sqrt{c^\varphi(t) + d(t)x_t^\varphi}\,dW_t + dJ_t \;. \tag{4.5}$$

where $a^\varphi(t) := a(t) + \varphi'(t) - b(t)\varphi(t)$ and $c^\varphi(t) := c(t) - d(t)\varphi(t)$. It is well known that the process $J_t - \omega\alpha t, t \geq 0$ is an $\mathbb{F}$-martingale. As already noticed in Brigo and Merccurio, 2001, whatever the base model y, the parameter Ξ and the discount curve P^{market}, there always exists a shift function $\varphi(t) = \varphi^\star(t; \Xi)$ that provides a perfect fit between the x^φ-model and the market. This is summarized in the next lemma.

Remark 7. *The shift approach may look suspicious: adding a deterministic function to a stochastic process is arguably a somewhat artificial way to fix the model's limitations in terms of calibration. As clear from* (4.5)*, the shifted model simply amounts to consider an inhomogeneous model. For instance, the Vasicek model* $(a(t), b(t), c(t), d(t)) = (0, -\kappa, \eta, 0)$ *shifted with* $\varphi(t) \leftarrow \int_0^t \varrho(s)e^{-\kappa(t-s)}ds$ *yields a HAJD with* $(a(t), b(t), c(t), d(t)) = (\kappa\varrho(t), -\kappa, \eta, 0)$*, which is known as the Hull-White (HW) model Hull and White, 1990. Moreover, the later is itself a particular case of the Heath-Jarrow-Morton (HJM) model Heath et al., 1992 which consists of modeling the entire instantaneous forward curve* $f_s(t)$ *with* $df_s(t) = \mu(s, t)dt + \eta e^{-\kappa(t-s)}dW_s$ *where the drift* $\mu(s, t)$ *is given by no-arbitrage, and* *provided that the initial discount curve and the long-term mean obey the relationship* $\varrho(t) = \frac{d}{dt} f_0(t) + \kappa f_0(t) + \frac{\eta^2}{2\kappa}(1 - e^{-2\kappa t})$*. Therefore, any instantaneous forward curve* f^{market} *(hence discount curve* P^{market}*) can be fitted with either models provided that one takes* $f_0(t) \leftarrow f^{market}(t)$ *as initial curve (HJM), the corresponding long-term mean* $\varrho(t)$ *(HW), or the associated shift* $\varphi(t)$ *(shifted Vasicek). These models became very popular among practitioners, essentially because of their ability to replicate market curves, i.e., to solve Problem 1.*

Lemma 7. *The* x*-model defined according to* (2.29) *where* y *is a HAJD solves Problem 1 provided that*

$$\varphi(t) \leftarrow \varphi^\star(t;\Xi) := \frac{d}{dt} \ln \frac{P^y(t;\Xi)}{P^{market}(t)} = f^{market}(t) - f^y(t;\Xi) \ . \tag{4.6}$$

where f^{market} *and* f^y *are the instantaneous forward rate functions associated with* P^{market} *and* P^y*, respectively.*

Proof. Indeed, because y is a HAJD, P^y is a discount curve and from Lemma 4, it admits a representation in terms of forward rates f^y. By assumption, same holds true for P^{market}. Eventually,

$$P^{x^\varphi}(t;\Xi) = \mathbb{E}\left[e^{-\int_0^t x_u du}\right] = e^{-\int_0^t \varphi^\star(u;\Xi)du}\,\mathbb{E}\left[e^{-\int_0^t y_u du}\right] = e^{-\int_0^t f^{market}(u)du} = P^{market}(t).$$

The model is tractable since $f^y(t;\Xi) = -\frac{d}{dt} \ln P^y(t;\Xi)$ can be computed in closed form. $\square$

It is worth noting that, for a given model y, the perfect fit can be attained for every parameters Ξ. This suggests that the calibration problem (4.1) is ill-posed. Indeed, the choice of Ξ is completely arbitrary since the error between P^{x^φ} and P^{market} can be set to zero for any Ξ, provided that one chooses $\varphi(t) \leftarrow \varphi^\star(t;\Xi)$. In particular, one could take the null process for y (i.e., $y_t = 0$ and Ξ such that $dy_t = 0$) and $\varphi(t) = f^{market}(t)$. This trivial choice rends x^φ deterministic, which is most likely not the desired result. A common practice to circumvent this indeterminacy is thus either (i) to extend the set of calibration instruments, incorporating products that are sensitive to volatility (like interest-rate or credit options in the above asset classes), or (ii) to require the y-model to fit the market "as best as possible" (to get $\Xi^\star$) and then take $\varphi(t;\Xi^\star)$ as shift function:

$$\begin{aligned} &P^{model}(t) := P^{x^\varphi}(t;\Xi^\star) \quad \text{where} \\ &\Xi^\star := \arg\min_{\Xi} \|P^y(\cdot;\Xi) - P^{market}(\cdot)\| \ , \quad \varphi(t) \leftarrow \varphi^\star(t) := \varphi^\star(t;\Xi^\star) \ . \end{aligned} \tag{4.7}$$

This approach is particularly relevant when no or little "volatility-sensitive" instruments are quoted on the market. The role of the shift is thus merely to compensate the remaining discrepancies between the market curve P^{market} and $P^y(\cdot;\Xi^\star)$, the one generated by the "best" parametric model y. Adding a shift to the VAS, CIR or JCIR models yield the Hull-White, CIR++ or JCIR++, respectively Brigo and Mercurio, 2006.[3]

Remark 8. *On the top of the appealing affine structure, the shifted model is highly tractable because many statistical properties of the process are available in closed form. Indeed, as recalled in Section 4.A, the k-th moment* $m^y(k,t) := \mathbb{E}[y_t^k]$ *and the moment generating function (MGF)* $\psi^y(u,t) := \mathbb{E}[e^{uy_t}]$ *of a time-homogeneous affine model* y *are known analytically, as well as those of their time-integrals* $Y_t := \int_0^t y_u du$, $m^Y(k,t)$ *and* $\psi^Y(u,t)$. *Due to the simple shift structure, the corresponding expressions for* x^φ, *the shifted model, are readily available. For instance, the k-th moment of* x_t^φ *and* $X_t^\varphi := \int_0^t x_s^\varphi ds$ *are given by Newton's binomial formula applied to* $(y_t + \varphi(t))^k$ *and the MGFs simply collapse to* $\psi^{x^\varphi}(u,t) = e^{u\varphi(t)}\psi^y(u,t)$ *and* $\psi^{X^\varphi}(u,t) = e^{u\int_0^t \varphi(s)ds}\psi^Y(u,t)$.

4.2.4 Dealing with the positivity constraint

As discussed above, the deterministic shift extension nicely solves Problem 1. In order to solve Problem 2 however, one first considers a non-negative base process y. Yet, there is no reason that the shifted process x^φ would remain non-negative. For instance, taking CIR dynamics for y, x^φ is non-negative on $[s,t]$ if and only if $\min_{u\in[s,t]} \varphi(u) \geq 0$. From (4.6), the shift function depends both on the y model (and its parameters Ξ) and on the market curve.

Remark 9. *Observe that the optimization problem* (4.7) *is contradictory with non-negative shift functions. Indeed, by construction of* $\Xi^\star$, $P^y(\cdot;\Xi^\star)$ *passes through* P^{market}. *Consequently, the shift* $\varphi(t) \leftarrow \varphi^\star(t;\Xi^\star)$ *will lead to a perfect fit, but will correct for both negative and positive errors. In other words,* φ *will change of sign. Therefore, this strategy does not provide a valid solution to Problem 2. This will be illustrated on a real example in Section 4.4.1.*

In order to satisfy the non-negativity constraint mentioned in Problem 1, one needs to force the non-negativity constraint on the shift at the optimal parameters. The

[3]Notice that the Hull-White model is a Vasicek model where the long-term mean parameter is replaced by a deterministic function of time.

shift function under positivity constraint is referred to with the notation $\varphi^{\star,+}(t)$ to stress the difference with the unconstrained counterpart, $\varphi^{\star}(t)$.

Lemma 8. *Let y be a HAJD that is non-negative on $[0, T]$ with parameters $\Xi^{\star}$ given by*

$$\Xi^{\star,+} := \arg\min_{\Xi} \|P^y(\cdot;\Xi) - P^{market}(\cdot)\| \text{ subject to } f^y(t;\Xi) \leq f^{market}(t),\ \forall\, 0 \leq t \leq T. \tag{4.8}$$

Then, the x^{φ}-model (2.29) with $\varphi(t) \leftarrow \varphi^{\star,+}(t) := \varphi^{\star}(t;\Xi^{\star,+})$ solves Problem 2.

Proof. The condition on the instantaneous forward rates ensures that shift function $\varphi^{\star,+}$ will be non-negative on $[0, T]$; this is obvious from (4.6). Hence, since y is assumed to be non-negative, so is the shifted process x^{φ}. Moreover, taking $\varphi(s) \leftarrow \varphi^{\star}(s;\Xi)$ yields a perfect fit for every Ξ, by construction. □

Notice that there always exists a set of parameters Ξ such that the constraint is met. Indeed, all parameters Ξ associated to the deterministic case $y \equiv 0$ yield $f^y(\cdot;\Xi) \equiv 0$. Clearly, the constraint is met since $f^{market}(t)$ is strictly positive given that P^{market} is strictly decreasing, by assumption. The shift is simply given by the market forward rate $\varphi(t) \leftarrow \varphi^{\star,+}(t) = f^{market}(t)$. However, the trivial process parameter is likely not to be the desired one.

In order to deal with Problem 2, we need to consider a non-negative base model y. Given that we focus on HAJDs, we consider the CIR and JCIR models. To make the distinction between the two shifted models, we call S-(J)CIR the (J)CIR process shifted with $\varphi(t) \leftarrow \varphi^{\star}(t) = \varphi^{\star}(t;\Xi^{\star})$ (i.e., without positivity constraint, and parameter $\Xi^{\star}$ given by (4.7)) and PS-(J)CIR the (J)CIR process shifted with $\varphi(t) \leftarrow \varphi^{\star,+}(t) = \varphi^{\star}(t;\Xi^{\star,+})$ (i.e., under positivity constraint, and parameter $\Xi^{\star}$ given by (4.8)). Although the PS-(J)CIR allows both for a perfect fit and the non-negativity constraint, one may argue that it is not as tractable as the (J)CIR. Indeed, the optimization problem (4.8) is more difficult than (4.7) due to the constraint on the instantaneous forwards, even if some sufficient conditions on the parameters can be found. Second, and probably more importantly, this constraint is binding, in the sense that it often deeply impacts the optimal parameter $\Xi^{\star,+}$. Even it is unlikely that the optimal solution corresponds to the deterministic case, it often yields dynamics associated to rates that feature “little randomness”. This will be illustrated in Section 4.4, first by comparing the variance of the integrated S-CIR and PS-CIR processes, as well as the impact when dealing with financial applications. These two points are discussed in Brigo and Mercurio, 2006, sec.

3.9.3, p.107-109. To circumvent this issue in an interest rate framework, the authors suggest to relax the strict positivity constraint. By working in a setup where positivity is expected but not guaranteed, they obtain a process that yields much more realistic results in terms of implied volatility levels. This is perfectly fine in such a context as positivity of rates might be desirable (in some cases), but zero is by no means a strict lower bound (neither theoretically nor practically). Yet, this is much more problematic when it comes to model such things as default intensities, because this kind of applications requires both strict positivity and, typically, large volatility. Increasing the variance of the CIR++ process without breaking Feller's constraint[4] can be achieved by incorporating compounded Poisson jumps (JCIR++) but, unfortunately, increasing the jump activity while maintaining the calibration to a given market curve f^{market} is difficult under the positivity constraint. Indeed, the minimum of the implied shift function is driven down when increasing the jump activity because the difference $f^{\mathrm{JCIR}}(t)-f^{\mathrm{CIR}}(t)$ is non-negative and increases with ω, α for $\alpha, \omega > 0$ (see Section 4.A.3). This observation combined with (4.6) leads to a lower shift function φ for the JCIR++ than for the corresponding CIR++. For this reason, there is a need for an alternative to CIR++ and JCIR++ that would combine (i) tractability, (ii) the prefect fit feature, (iii) the large implied volatility and (iv) positivity.

4.3 The deterministic time-changed extension

In order to circumvent the drawbacks of the deterministic shift extension with regards to Problem 2, we propose a different approach. In the same spirit as the shift, we aim at finding a model x by adjusting a time-homogeneous affine model y, that would benefit from a set of desirable properties.

4.3.1 Model setting

The x-model is obtained by time-changing a HAJD y using a specific (but deterministic) *clock* $\Theta(t)$ that may differ from the calendar clock. A clock is a time change function that can differ from identity, but having specific properties.

[4] Increasing the volatility of the CIR++ process by increasing the diffusion paramter δ just breaks the Feller's condition ($2\kappa\varrho \geq \delta^2$) and leads to an intensity process that almost surely equals to zero at a given time interval.

Definition 5 (Clock)**.** *A clock is an application*

$$\Theta : \mathbb{R}^+ \to \mathbb{R}^+,\ t \mapsto \Theta(t)$$

that is a grounded, increasing and differentiable. In other words, a clock is any function Θ *of the form*

$$\Theta(t) := \int_0^t \theta(u)du \quad where \quad \theta(u) > 0\ , \quad \forall\, u \geq 0\ .$$

Clearly, $\Theta(t) = t$ is the calendar clock, and any function of the form $\Theta(t) = kt$, $k > 0$, is again a clock, corresponding to a constant rescaling of the calendar time.

Similarly to (2.29), we define our model as $x = x^\theta$, obtained from the following transform of the base process y:

$$x_t^\theta := \theta(t) y_{\Theta(t)}\ . \tag{4.9}$$

The dynamics of x^θ are given by Ito's product rule. Defining the process $y^\theta := \left(y_{\Theta(t)}\right)_{t\in[0,T]}$, one gets

$$dx_t^\theta = y_t^\theta d\theta(t) + \theta(t) dy_t^\theta\ , \tag{4.10}$$

where the dynamics of y^θ are given in the below lemma.

Lemma 9. *Let* Θ *be a clock and consider a base model* y *with dynamics* (4.4)*. Then, the dynamics of* y^θ *take the form*

$$dy_t^\theta = \mu\left(\Theta(t), y_t^\theta\right)\theta(t)dt + \sigma\left(\Theta(t), y_t^\theta\right)\sqrt{\theta(t)}dB_t + dJ_t^\theta, \quad y_0^\theta = y_0 \tag{4.11}$$

where B *is an* $\mathbb{F}^\theta$*-Brownian motion,* $\mathbb{F}^\theta := (\mathcal{F}_{\Theta(t)})_{t\in[0,T]}$ *and* J^θ *an inhomogeneous compounded Poisson process with jump size mean* α *and time-*t *intensity* $\omega\theta(t)$*.*

Proof. By definition, we have

$$y_t^\theta := y_{\Theta(t)} = y_0 + \int_0^{\Theta(t)} \mu(u, y_u)du + \int_0^{\Theta(t)} \sigma(u, y_u)dW_u + \int_0^{\Theta(t)} dJ_u\ .$$

Hence,

$$\int_0^{\Theta(t)} \mu(u, y_u)du = \int_0^t \mu\left(\Theta(u), y_{\Theta(u)}\right)\theta(u)du = \int_0^t \mu\left(\Theta(u), y_u^\theta\right)\theta(u)du\ ,$$

and

$$\int_0^{\Theta(t)} \sigma(u, y_u)dW_u = \int_0^t \sigma\left(\Theta(u), y_{\Theta(u)}\right)dW_{\Theta(u)} = \int_0^t \sigma\left(\Theta(u), y_u^\theta\right)\sqrt{\theta(s)}dB_u\ .$$

Indeed, Θ is a clock, hence $\theta > 0$ and the process $W^\theta := (W_{\Theta(t)})_{t\in[0,T]}$ is a local martingale with quadratic variation $\langle W^\theta, W^\theta\rangle_t = \Theta(t)$. From Jeanblanc et al., 2009, the process $B := (B_t)_{t\in[0,T]}$ defined as

$$B_t := \int_0^t \frac{1}{\sqrt{\theta(u)}} dW_{\Theta(u)} \tag{4.12}$$

is then an $\mathbb{F}^\theta$-Brownian motion. Differentiating y_t^θ leads to (4.11). With regards to the compounded Poisson process, notice that $dJ_t = \zeta_{N_t} dN_t$ and $dJ_t^\theta = dJ_{\Theta(t)} = \zeta_{N_{\Theta(t)}} dN_{\Theta(t)}$. The process N^θ defined as $N_t^\theta := N_{\Theta(t)}$ is a Poisson process with instantaneous intensity $\omega\theta(t)$. Hence, the dynamics of J^θ are given by $J_0^\theta = 0$ and $\zeta_{N_t^\theta} dN_t^\theta$, so that J^θ is a compounded Poisson process with jump size mean α and time-t instantaneous rate of jumps arrival, $\omega\theta(t)$. $\square$

This model looks appealing for several reasons. First, just as the shift extension, it is a deterministic adjustment of a base model and is hence expected to be tractable when the latter is, say, an HAJD. Second, because x_t^θ is a positive rescaling of the process y sampled at time $\Theta(t)$, the range of x^θ is linked to that of y. In particular, if the range of y is $\mathbb{R}$, as for the Vasicek model, then so is the range of x^θ. However, if y is non-negative as in the (J)CIR case, then so is x^θ. Hence, this solves the drawback of the shift approach related to Problem 2. Eventually, the time-dependent feature of the clock rate θ is expected to provide additional flexibility in the calibration properties of x^θ with respect to that of the homogeneous model y. Two questions remain open in this respect. First, we need to clarify the circumstances under which the model provides a perfect fit. Second, in the case where the perfect fit can be achieved, we need to provide an efficient procedure to compute the resulting "optimal clock", $\Theta^\star$. The price to pay is that, in contrast with the shift extension, the time-changed model is not fully flexible. Indeed, starting with a given model y, the x^θ model can only generate specific shapes for discount curves. We are thus more dependent on the initial choice of the base model y. Fortunately, it turns out that a perfect fit is achievable for a wide set of market curves, including all decreasing discount curves, considered in Problem 2. This is clearly the most important case since (i) it corresponds to the case where the shift approach fails to provide a convincing solution and (ii) it is probably the most common case in practice, since it encompasses the class of discount curves with non-negative rates (or, more generally, with non-negative instantaneous forward rates), as well as the set of all continuous survival probability curves. Moreover, even if the mathematical expression of the clock $\Theta^\star$ is not available in closed form, its numerical computation turns out to be easy. This leads us to the first fundamental result of the chapter.[5]

Theorem 2. *Let P^{market} be a discount curve and y a model such that P^y is a discount curve. Define the x^θ-model as in* (4.9)*. Then, $P^{x^\theta} \equiv P^{market}$ provided that $\Theta \leftarrow \Theta^\star$ where $\Theta^\star$ satisfies the first-order ODE*

$$\theta^\star(t) := \frac{d}{dt}\Theta^\star(t) = \frac{f^{market}(t)}{f^y(\Theta^\star(t))}\ , \tag{4.13}$$

with f^{market}, f^y the corresponding instantaneous forward curves. Moreover, if P^{market} and P^y are strictly decreasing, the solution to (4.13) *exists, is a clock, and is given by*

$$\Theta^\star(t) := Q^y\left(P^{market}(t)\right)\ , \tag{4.14}$$

where Q^y is the inverse of the base-model discount curve, P^y.

Proof. See Section 4.C. □

Observe that the optimal clock $\Theta^\star$ actually depends from the y-model parameters Ξ. Just like for the shift, we actually have $\Theta^\star(t) = \Theta^\star(t;\Xi)$. Although other frameworks are possible, we set $\Xi = \Xi^\star$ as in (4.7). Similar to the function φ in the shift approach, the purpose of the clock Θ is then to absorb the remaining errors between $P^y(\cdot;\Xi^\star)$ and P^{market}.

4.3.2 Time-changed homogeneous affine diffusions

As in the shift extension, a time-changed model x^θ enjoys a similar tractability level to that of the base model y. Indeed the k-th moment is $m^{x^\theta}(k,t) = \theta(t)^k m^y(k,\Theta(t))$ and moment generating function is $\psi^{x^\theta}(u,t) = e^{u\theta(t)}\psi^y(u,\Theta(t))$, whereas those of X_t^θ coincide with those of $Y_{\Theta(t)}$. Hence, a tractable model x^θ can be obtained by considering HAJD processes as base model y. We illustrate our method by analyzing two calibration problems that can be solved by considering the Vasicek and the JCIR processes.
It is clear from Lemma 9 that in the particular case where y is a HAJD, then x^θ is a scaled version of an inhomogeneous affine jump diffusion (AJD), unless $\theta(t)$ is a

[5]When no confusion is possible, the explicit reference to the model parameters Ξ is avoided to ease the notations.

positive constant, in which case it remains a HAJD. To see this, suppose that the dynamics of y obey (2.28). From Lemma 9, y^θ is governed by

$$dy_t^\theta = \left(a(\Theta(t)) + b(\Theta(t))y_t^\theta\right)\theta(t)dt + \sqrt{\left(c(\Theta(t)) + d(\Theta(t))y_t^\theta\right)\theta(t)}dB_t + dJ_t^\theta \ . \tag{4.15}$$

Interestingly, y^θ is still an AJD. In the sequel, we focus on the special case where the base model y is a HAJD, i.e., takes the form (2.28) with constant parameters $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\varrho, -\kappa, \eta^2, \delta^2, \alpha, \omega)$. To simplify the notation, we specify the model parameters using the vector $\Xi = (\kappa, \varrho, \eta, \delta, \alpha, \omega)$.

Time-changed Vasicek

Our time change approach can be easily used to solve Problem 1 in the most common case where the forward curve f^{market} is arbitrary (monotonic, humped, etc) provided that it is positive. As there is no constraint on the range of the process x^θ, let us postulate Vasicek dynamics for the base process with parameters $\Xi = (\kappa, \varrho, \eta, 0, 0, 0)$:

$$dy_t = \kappa(\varrho - y_t)dt + \eta dW_t, \quad y_0 \in \mathbb{R} \ .$$

The forward curve associated to this model is given by $f^y(t) = f^{\mathrm{VAS}}(t) := f_0^{\mathrm{VAS}}(t)$ in (4.19):

$$f^{\mathrm{VAS}}(t) = (1 - \mathrm{e}^{-\kappa t})\frac{\kappa^2\varrho - \eta^2/2}{\kappa^2} + \frac{\eta^2}{2\kappa^2}\,\mathrm{e}^{-\kappa t}(1 - \mathrm{e}^{-\kappa t}) + y_0\,\mathrm{e}^{-\kappa t} \ . \tag{4.16}$$

It can thus be used to select an appropriate Vasicek model. The next corollary provides guidelines to generate decreasing discount curve, associated with the most common case of positive instantaneous forwards.

Corollary 1. *Let P^{market} be a strictly decreasing market curve. Then, for every Vasicek model with parameters satisfying $y_0 \geq 0$ and $2\kappa^2\varrho > \eta^2$, there exists a clock $\Theta^\star$ such that $P^{x^\theta} \equiv P^{market}$.*

Proof. Because y is a Vasicek process, P^y is a discount curve. Moreover, the conditions $y_0 \geq 0$ and $2\kappa^2\varrho > \eta^2$ guarantee that the forward curve (4.16) is strictly positive, hence P^y is strictly decreasing. From Theorem 2, the clock $\Theta^\star$ exists and is given by (4.13) with f^y given in (4.16). □

Notice that the dynamics of the time-changed Vasicek model x^θ_t are given by (4.10) with

$$dy^\theta_t = \kappa(\varrho - y^\theta_t)\theta(t)dt + \eta\sqrt{\theta(t)}dB_t, \quad y^\theta_0 = y_0 \ ,$$

showing that y^θ remains a Gaussian process. Fitting perfectly a strictly decreasing discount curve (without further constraints on the process) is a special case of Problem 1, that can also be solved using the shift approach (2.29) by taking $x \leftarrow x^\varphi$ where y is a Vasicek with arbitrary parameters Ξ and $\varphi(t) \leftarrow \varphi^\star(t;\Xi) = f^{market}(t) - f^{\mathrm{VAS}}(t)$. The main interest of the time-changed approach is actually when considering Problem 2.

Time-changed (J)CIR

The following result is the second main contribution of the chapter. It shows that the time change approach $x \leftarrow x^\theta$ provides a solution to Problem 2.

Corollary 2. *Let y be an almost-surely positive HAJD with parameters Ξ. Then, the model x^θ defined in (4.9) with $\Theta \leftarrow \Theta^\star(t;\Xi)$ solves Problem 2.*

Proof. Because y is a HAJD, P^y is a discount curve and is tractable analytically. Moreover, the latter is strictly decreasing since y is almost-surely positive. We conclude the proof by relying on Theorem 2. □

Let us now consider the JCIR model, i.e., the HAJD with $\Xi = (\kappa, \varrho, 0, \delta, \alpha, \omega)$. The CIR is recovered as a special case by choosing (α, ω) such that $\alpha\omega = 0$.

Then,

$$dy_t = \kappa(\varrho - y_t)dt + \delta\sqrt{y_t}dW_t + dJ_t, \quad y_0 > 0 \ , \tag{4.17}$$

where κ, ϱ, δ are strictly positive constants and ω, α are non-negative. The optimal clock $\Theta^\star$ leading to the perfect fit to a given strictly decreasing curve P^{market} is given by (4.13) where the forward curve associated to this model is given by $f^y(t) = f^{\mathrm{JCIR}}(t) := f^{\mathrm{JCIR}}_0(t)$ in (4.23) :

$$\begin{aligned} f^{\mathrm{JCIR}}(t) \;&=\; \frac{2\kappa\varrho(e^{t\gamma}-1)}{2\gamma+(\kappa+\gamma)(e^{t\gamma}-1)} + y_0\frac{4\gamma^2 e^{t\gamma}}{[2\gamma+(\kappa+\gamma)(e^{t\gamma}-1)]^2} \\ &+ \frac{2\omega\alpha(e^{t\gamma}-1)}{2\gamma+(\kappa+\gamma+2\alpha)(e^{t\gamma}-1)} \ , \end{aligned}$$

where $\gamma := \sqrt{\kappa^2 + 2\delta^2}$. The dynamics of the time-changed process $x_t^\theta = \theta(t) y_t^\theta$ are given by (4.10) with

$$dy_t^\theta = \kappa(\varrho - y_t^\theta)\theta(t)dt + \delta\sqrt{\theta(t) y_t^\theta} dB_t + dJ_t^\theta, \quad y_0^\theta = y_0 \, .$$

where B is an $\mathbb{F}^\theta$-Brownian motion and J^θ is an inhomogeneous compound Poisson process with jump size mean α and time-t instantaneous rate of arrival $\omega\theta(t)$.
The time change technique applied to a JCIR (TC-JCIR) therefore solves Problem 2. In particular, in contrast to the PS-JCIR (that focuses on parameters such that φ is positive), the positivity constraint on x^θ is automatically satisfied for every (strictly decreasing) market curve and every Ξ (such that y is not trivially equal to 0). However, we have shown that it is possible to ensure positivity by considering the PS-JCIR, $x^{\varphi,+}$. Working with $\Xi^{\star,+}$ instead of $\Xi^\star$ can make the job, but at the expenses of having a process $x^{\varphi,+}$ that is, to a large extend, deterministic (i.e., $x_t^{\varphi,+}$ varies in a small neighborhood around $f^{market}(t)$). Consequently, TC-JCIR model are expected to feature a higher volatility compared to the corresponding PS-JCIR, at least up to some time horizon. This is summarized in the next theorem, which is the third main result of the chapter.

Theorem 3. *Let P^{market} be a strictly decreasing discount curve and y be a JCIR++ process with parameter Ξ such that the perfect fit JCIR++ model $x_t^{\varphi^\star}$ is positive. Then, the ODE (4.13) with $f^y(t) = f^{\mathrm{JCIR}}(t;\Xi)$ given by (4.18) admits a solution that satisfies $\Theta^\star(t) = \Theta^\star(t;\Xi) \geq t$. Moreover, the variance of the corresponding perfect fit TC-JCIR model $x_t^{\theta^\star}$ satisfies:*

1) $\mathbb{V}\left[X_t^{\theta^\star}\right] \geq \mathbb{V}\left[X_t^{\varphi^\star}\right]$, $\forall\, t \geq 0$,

2) $\mathbb{V}\left[x_t^{\theta^\star}\right] \geq \mathbb{V}\left[x_t^{\varphi^\star}\right]$ *if one of the following holds:*

   i) $y_0 = \varrho + \frac{\omega\alpha}{\kappa}$,

   ii) f^{market} *constant and* $y_0 \leq \varrho + \frac{\omega\alpha}{\kappa}$,

   iii) $y_0 > \varrho + \frac{\omega\alpha}{\kappa}$ *and* $t < \Theta^{\star -1}(t_1)$,

   iv) $(\kappa\varrho + \omega\alpha)/\gamma < y_0 < \varrho + \frac{\omega\alpha}{\kappa}$ *and* $t > \Theta^{\star -1}(t_2)$

where

$$t_1 := \frac{1}{\kappa} \ln\left(1 + \frac{y_0 + 2\omega\alpha^2/\gamma^2}{y_0 - \varrho - \omega\alpha/\kappa}\right) \quad \text{and} \quad t_2 := \frac{1}{\gamma} \ln \frac{(\gamma-\kappa)(\kappa\varrho + y_0\gamma + \omega\alpha) - 2\omega\alpha^2}{(\kappa+\gamma)(y_0\gamma - \kappa\varrho - \omega\alpha) - 2\omega\alpha^2}.$$

Proof. See Section 4.D. □

To sum up, the TC-JCIR model (including the TC-CIR) provides an elegant solution to Problem 2: the process $x^{\theta^\star}$ is non-negative (in contrast with the S-JCIR $x^{\varphi^\star}$), is almost as tractable as the simple JCIR diffusion (in contrast with the PS-JCIR $x^{\varphi^{\star,+}}$), provides a perfect fit to every strictly decreasing discount curve (as both JCIR++ models) and features, to some extend, a larger variance (compared to the PS-JCIR $x^{\varphi^{\star,+}}$). In particular, it is observed, empirically, that the variance of the integral of the TC-JCIR remains similar to that of the unconstrained (i.e., flawed, but high-volatility) S-JCIR model $x^{\varphi^\star}$. Therefore, when a positivity constraint is required, the TC-JCIR avoids the drawbacks of the JCIR++ models. The only price to pay is that the clock is not available in closed form, but requires a (simple) numerical inversion. The properties of the model, namely the perfect fit and high-variance features, are illustrated on various applications taken from credit risk modeling.

4.4 Application to Credit Risk Modelling

We consider a reduced-form default model as in Section 4.2.1 by using a CIR base model y (i.e., (4.17) with $J \equiv 0$). The default intensity λ is modelled either as a CIR++ ($\lambda \leftarrow \lambda_t^\varphi := y_t + \varphi(t)$) or using the TC-CIR ($\lambda \leftarrow \lambda_t^\theta := \theta(t)y_{\Theta(t)}$). Observe that depending on the pair (P^{market}, Ξ), the CIR++ process can feature negative values. This will be the case when taking $\Xi \leftarrow \Xi^\star$ given using the MSE approach (4.7), unless there is an explicit constraint as in (4.8), leading to take $\Xi \leftarrow \Xi^{\star,+}$. Bear in mind that when λ represents an intensity process, the S-CIR model (λ^φ) is actually flawed as there is a non-zero probability to observe negative intensities, and $P^{\lambda^\varphi}(t)$ cannot be interpreted as a survival probability associated to a Cox model. Yet, we give the results of the model as a benchmark since, as explained in the introduction, it is a very standard approach.

We compare the CIR++ (S-CIR and PS-CIR) to the TC-CIR on several aspects related to a real case example where the reference entity is Ford Inc. We also discuss the TC-JCIR case when relevant. We first analyze the perfect fit feature of both types of models, as well as the non-negativity property of λ. We then compare the variance of the integrated processes Λ. We then analyse their behaviors in two different applications, namely the pricing of various credit default swaptions (a.k.a. CDS options, or CDSO) with Ford as reference entity, or on the credit valuation adjustment (CVA) of prototypical FRA and IRS exposures where Ford is the trade counterparty.

We consider a deterministic short rate process[6], leading to $P^r_s(t) = \beta_s/\beta_t = e^{-\int_s^t r(u)du}$.

In the sequel, we first illustrate the perfect fit feature of S-CIR, PS-CIR and TC-CIR when the default model is calibrated on the survival probability curve of Ford Inc. We then use the model to price CDSO and compute CVA figures.

4.4.1 Perfect fit of CDS term-structure

We consider the CDS term-structure of Ford Inc, and show that considering a set of parameter Ξ, there exist φ and Θ that yield a perfect fit. In the sequel, we drop the star superscript on the shift and clock functions. Hence, $\Xi^\star$ corresponds to the CIR parameter optimized without constraint to a given P^{market} curve, and φ and Θ refer to the corresponding optimal shift and clock functions. The corresponding parameters found under a non-negativity constraint are noted $\Xi^{\star+}, \varphi^+$ and Θ^+, respectively.

The CDS term-structure consists of a set of par spreads associated with CDS of various maturities. The time-t par spread $s_t(T_i)$ of a CDS contract of maturity T_i is defined as the contract spread k that sets the value of the CDS contract to 0 at time t. The par spreads have been taken from Bloomberg on November 12, 2018 and are shown on the table below.

Maturity (years)	1	3	5	7	10
Spread (bps)	18.3	136.6	191.9	267.6	280.6

Table 4.1: CDS spread term structure of Ford Inc. on November 12, 2018. Source: Bloomberg.

In this context, the market curve P^{market} to be fitted is the risk-neutral survival probability curve $G(\cdot)$, defined in (2.4), associated with the default time τ of a given reference entity (here, Ford Inc.). It can be extracted from CDS quotes by inverting the no-arbitrage pricing formulae of the corresponding financial instruments. In practice, one only has a couple of calibration equations, say n, given by the number of market quotes (here, $n = 5$). It is therefore not possible to estimate the full (i.e., infinite-dimensional) market curve G without further assumptions. It

[6]Given that the interest rates have little impact on the figure and that our main objective is to discuss the impact of the default model, we considered zero risk-free rate in the numerical applications below.

is common market practice to consider the CDS model from the International Swap and Derivative Association (ISDA) – a.k.a the JP Morgan model – Markit, 2004, that provides a slightly simplified version of the actual no-arbitrage pricing formula applying to CDSs. In this approach, the curve G is parametrized via a positive hazard rate function h, playing a similar role as the instantaneous forward rate f^{market}, where h is itself parametrized by n constants $h_1, h_2, \ldots, h_n$ bootstrapped from the spreads $s_1, s_2, \ldots, s_n$ associated with the maturities $T_1, T_2, \ldots, T_n$. Let us focus on the horizon $T = T_n$. It is market practice to assume that h is piecewise constant between the maturities, i.e., to postulate the parametric form:

$$h(t) = \sum_{i=1}^{n} \mathbb{1}_{\{T_{i-1} \le t < T_i\}} h_{i-1} \ ,$$

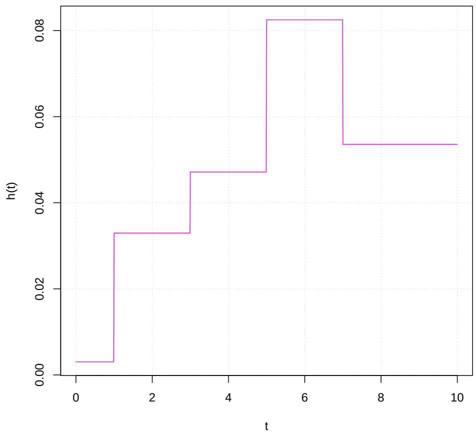
where $T_0 := 0$, $h_0 := \frac{s_1}{1-R}$ with $LGD := 1 - R$, $R = 40\%$ the assumed recovery rate of the firm and h_i's are positive constants. Even if less standard, another specifications like, e.g., a piecewise linear parametrization could be preferred:

$$h(t) = \sum_{i=1}^{n} \mathbb{1}_{\{T_{i-1} \le t < T_i\}} \left[\frac{h_i - h_{i-1}}{T_i - T_{i-1}} (t - T_{i-1}) + h_{i-1} \right] \ .$$

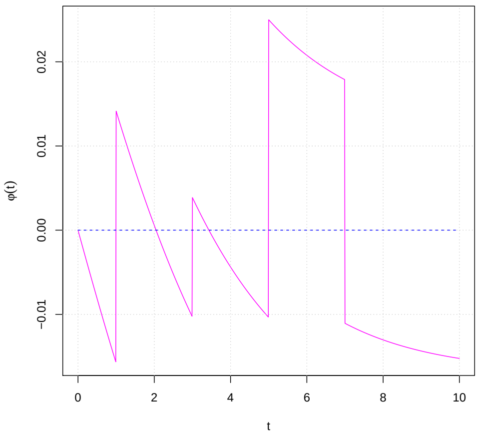
These two different specifications of the hazard rate function are considered on panels (a) of Figure 4.2 and 4.3, respectively. These frameworks yield similar (yet, slightly different) market curves $G(t)$ (green curves on panels (d)). For each of them, we start by computing the "best" base CIR model y. In line with market practice, we take $\Xi \leftarrow \Xi^\star$ using (4.1) with $P^{model} \leftarrow P^y$ considering (4.2) as error function and $\mathcal{T}$ the set of available liquid CDS maturities available[7]. In each case, we consider the two adjusted intensity models associated to the optimal shift (φ, given by (4.7)) or optimal clock (Θ, given by (4.14)). The latter are shown on panels (b) and (c), respectively. The model curves P^{λ^φ} (S-CIR) and P^{λ^θ} (TC-CIR) are shown in magenta on panels (d); they agree with each other, and collapse to $G(t)$ due to the perfect fit. Notice that the parametrization of the hazard rate function has little importance: the survival probability curves G, P^y and P^λ are very similar in either cases. Similarly, the clock functions Θ look very similar in both panels (c) of Fig. 4.2 and 4.3. However, it can be seen by comparing panels

[7]Notice that in all the thesis, we do not address the the stability issue of the calibrated parameters over time. The raison is that the theoretical models used to perform the calibration are the CIR (and JCIR, see Chapter 6) which are very standard models and belong to the class of affine models that are widely studied in the credit and interest rates term-structure literature (see for example Duffie and Kan, 1996, Duffie and Gârleanu, 2001, Brigo and Mercurio, 2006 or Brigo et al., 2013).

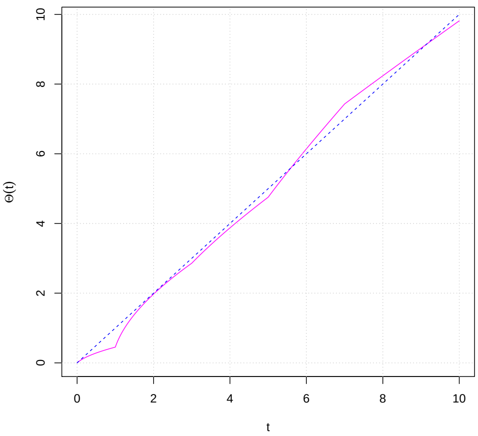
(b) that the shift functions φ are drastically different depending on the assumption regarding the parametric form of h.



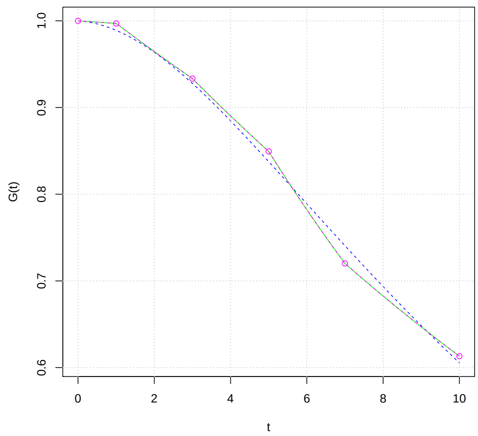
(a) Piecewise constant hazard rate (h)



(b) Optimal shift ($\varphi = \varphi^\star$)

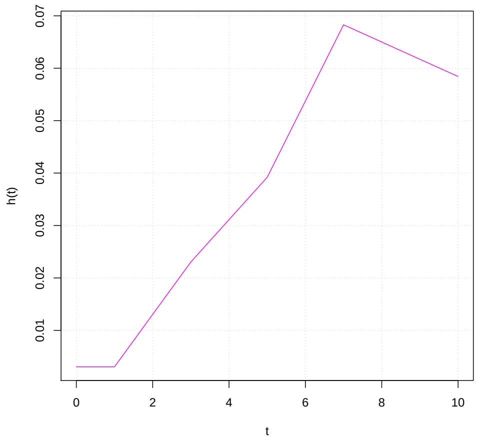


(c) Optimal clock ($\Theta = \Theta^\star$)

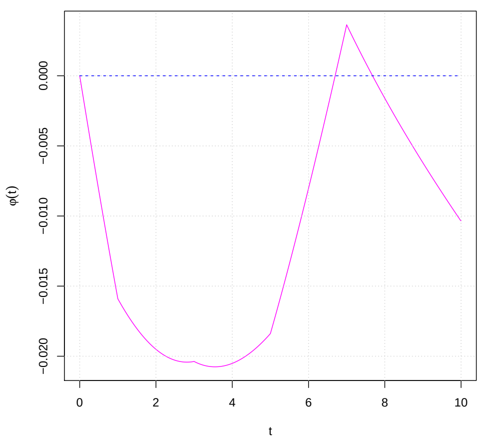


(d) Survival probability ($P^y, G = P^{\lambda^\varphi} = P^{\lambda^\theta}$)

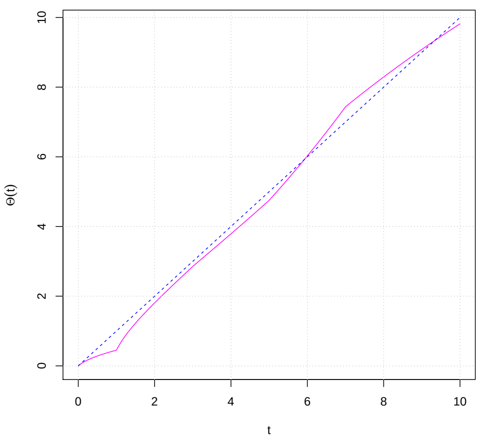
Figure 4.2: Fitting Ford Inc. CDS term-structure with adjusted CIR models. The survival probability curve $G(t)$ is parametrized with a piecewise constant hazard rate function $h(t)$ extracted from market prices taken from Bloomberg on November 12, 2018, panel (a). The base model y is a CIR with parameters $\Xi = \Xi^\star$ where $\Xi^\star = (\kappa, \varrho, \eta, \delta, \alpha, \omega) = (0.0555, 0.3018, 0, 0.2939, 0, 0)$ is obtained from (4.7) and $y_0 = h_0$. The shift function $\varphi(t) \leftarrow \varphi^\star(t, \Xi^\star)$ is shown in panel (b). Panel (c) gives the clock $\Theta(t) \leftarrow \Theta^\star(t; \Xi^\star)$. Eventually, panel (d) yields the survival probability curves given by the market ($G(t)$, green), or associated to $\mathbb{Q}(\tau(\lambda) > t)$ for various intensity models λ : the best base model $\lambda \leftarrow y$ (leading to $\mathbb{Q}(\tau(y) > t) = P^y(t, \Xi^\star)$, dashed blue), $\lambda \leftarrow \lambda^\varphi$ (S-CIR) and $\lambda \leftarrow \lambda^\theta$ (TC-CIR model). By construction of φ and Θ, the last two curves coincide (magenta) and agree with $G(t)$.



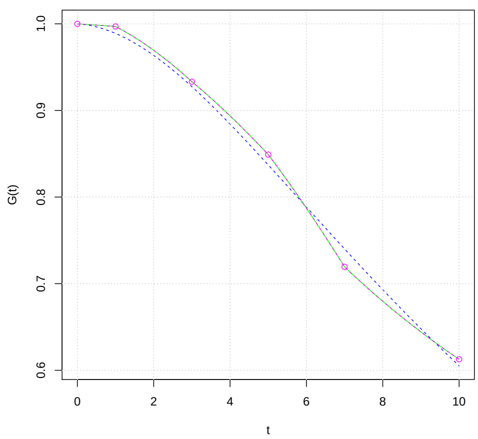
(a) Piecewise linear hazard rate (h)



(b) Optimal shift ($\varphi = \varphi^{\star}$)



(c) Optimal clock ($\Theta = \Theta^{\star}$)



(d) Survival probability ($P^{y}, G = P^{\lambda^{\varphi}} = P^{\lambda^{\theta}}$)

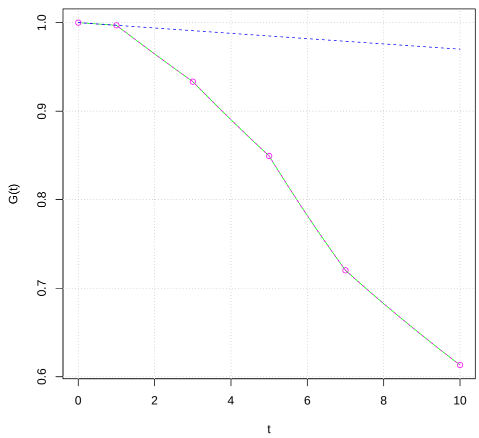
Figure 4.3: Fitting Ford Inc. CDS term-structure with adjusted CIR. The survival probability curve $G(t)$ is parametrized with a piecewise linear hazard rate function $h(t)$ extracted from market prices taken from Bloomberg on Novermber 12, 2018, panel (a). The base model y is a CIR with parameters $\Xi = \Xi^{\star}$ where $\Xi^{\star} = (\kappa, \varrho, \eta, \delta, \alpha, \omega) = (0.0620, 0.2729, 0, 0.2926, 0, 0)$ is obtained from eq. (4.7) and $y_0 = h_0$. The shift function $\varphi(t) \leftarrow \varphi^{\star}(t, \Xi^{\star})$ is shown in panel (b). Panel (c) gives the clock $\Theta(t) \leftarrow \Theta^{\star}(t; \Xi^{\star})$. Eventually, panel (d) yields the survival probability curves given by the market ($G(t)$, green), or associated to $\mathbb{Q}(\tau(\lambda) > t)$ for various intensity models λ : the best base model $\lambda \leftarrow y$ (leading to $\mathbb{Q}(\tau(y) > t) = P^{y}(t, \Xi^{\star})$, dashed blue), $\lambda \leftarrow \lambda^{\varphi}$ (S-CIR) and $\lambda \leftarrow \lambda^{\theta}$ (TC-CIR model). By construction of φ and Θ, the last two curves coincide (magenta) and agree with $G(t)$.

The parameters used in the numerical examples in the rest of the chapter are given in Table 4.2.

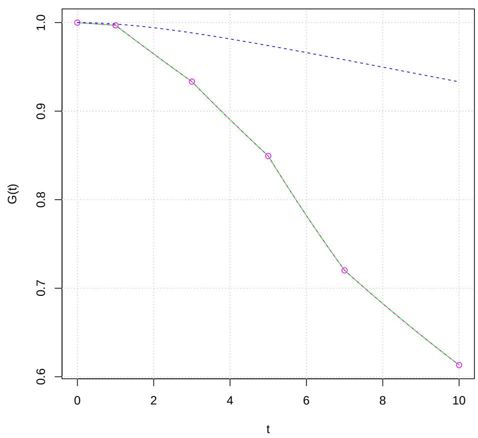
Ξ	κ	ϱ	δ	y_0
$\Xi^\star$	0.0555	0.3018	0.2939	h_0
$\Xi^{\star,+}$	0.2118	0.0030	0.0006	h_0
$\Xi_0^\star$	0.0624	0.2975	0.3343	0.0000
$\Xi_0^{\star,+}$	$3.8252.10^{-01}$	$9.6881.10^{-03}$	$1.5195.10^{-01}$	$3.2093.10^{-10}$

Table 4.2: Calibration parameters using Ford piecewise constant hazard rate. Parameters $\Xi^\star$ and $\Xi^{\star,+}$ correspond to the parameters of the CIR model y with and without positivity constraint, where y_0 is set exogenously to the first level of the piecewise hazard rate function, $h_0 = 0.0030$. The other parameters, $\Xi_0^\star$ and $\Xi_0^{\star,+}$, correspond to the similar cases but where y_0 is a parameter that enters the optimization procedure. In all cases, we have taken $\alpha = \omega = 0$.

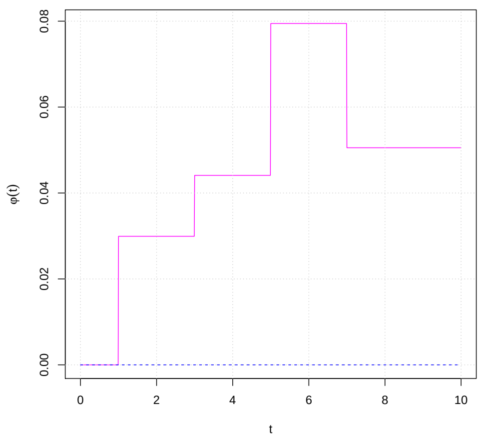
Notice that in both Figure 4.2 and 4.3, the shift function φ can take negative values. This means that the shift approach, S-CIR, yields negative default intensities λ^φ and, calibrated that way, is flawed. In particular, we cannot interpret λ^φ as a default intensity associated to a Cox process. This contrasts with the TC-CIR approach since λ^θ is a positive process if so is y. To fix this issue in a CIR++ framework, one needs to rely on PS-CIR. We note the corresponding processes y^+ and $\lambda^{\varphi,+}$. As illustrated on Figure 4.4 with our Ford example, this procedure is very restrictive: it leads to a curve P^y that is decreasing at a very low rate. In particular, the shape of $P^{\lambda^{\varphi,+}}$ essentially results from the shift, not from the base model y. This is problematic: it basically amounts to say that $h \approx \varphi$, i.e., that the PS-CIR process $\lambda^{\varphi,+}$ is essentially deterministic. This will put strong limitations on the resulting default model, and will be further discussed in the remaining subsections.



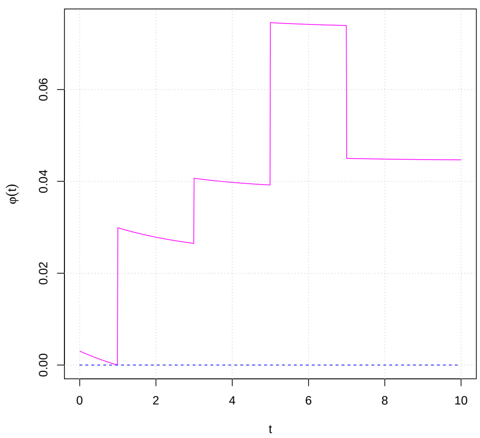
(a) Survival probability $(P^{y^+}, G = P^{\lambda^{\varphi,+}})$



(b) Survival probability $(P^{y^+}, G = P^{\lambda^{\varphi,+}})$



(c) Shift function $(\varphi = \varphi^{\star,+})$



(d) Shift function $(\varphi = \varphi^{\star,+})$

Figure 4.4: Fitting Ford Inc. CDS term-structure using $\lambda^{\varphi^\star,+}$ (PS-CIR). Panels (a) and (c) correspond to $y_0^+ = h_0$ whereas panels (b) and (d) correspond to the case where y_0^+ is one of the optimized parameters. The survival probability curve $G(t)$ is parametrized with a piecewise constant hazard rate function $h(t)$ extracted from market prices taken from Bloomberg on November 12 2018. The parameters $\Xi^{\star,+}$ are computed under the constraint $\varphi(t) \leftarrow \varphi^\star(t; \Xi^{\star,+}) \geq 0$. The base model y^+ is a CIR with parameters $\Xi = \Xi^{\star,+}$ (left) and $\Xi = \Xi_0^{\star,+}$ (right).

4.4.2 Variance analysis

Interesting observations can be made regarding the variance of the various integrated processes. As shown in the next two sections, they will have important consequences when considering financial applications, where Λ plays a central role in governing volatility and covariance effects.
First, observe that the integrated CIR process with optimal parameter $\Xi^\star$ is expected to feature a larger variance compared to the integrated CIR with parameter $\Xi^{\star,+}$. Because of the shift constraint, the discount curve P^y in the latter case rends to be much flatter than in the former case, i.e., one expects to have, in general

$$P^y(t;\Xi^{\star,+}) \geq P^y(t;\Xi^\star) \ .$$

This can be observed from panels (a) and (b) of Figure 4.4. When working with $\Xi^{\star,+}$, a substantial part of the shape of $P^{market} = G$ comes from the deterministic shift. This amounts to limit the randomness of the process. Not surprisingly, this will impact the variance of the integrated process Y. Indeed, because the discount curve of the CIR process with parameter $\Xi^{\star,+}$ generally dominates that of the CIR process with parameter $\Xi^\star$, one intuitively expects the variance of the CIR with parameter $\Xi^\star$ to be larger than that of the CIR with parameter $\Xi^{\star,+}$, due to the zero lower bound. In other words, even if it seems difficult to provide a formal proof, one expects intuitively the following to hold, in general:

$$v^{\Lambda^\varphi}(t) = v^Y(t;\Xi^\star) \geq v^Y(t;\Xi^{\star,+}) = v^{\Lambda^{\varphi,+}}(t) \ .$$

This is indeed the case on Figure 4.5: v^{Λ^φ} (dotted blue) dominates $v^{\Lambda^{\varphi,+}}$ (solid blue).

Second, observe that for a given base process y, the variance of the integrated TC-CIR is always larger than that of the integrated PS-CIR. Indeed, when working under the positivity constraint (i.e., when y is driven by $\Xi^{\star,+}$), we necessarily have $\Theta^+(t) := \Theta(t;\Xi^{\star,+}) \geq t$, in agreement with Theorem 3. Because for any parameter, the variance of Y is an increasing function of time (Theorem 3 again) we have, for $\Xi \leftarrow \Xi^{\star,+}$ in particular,

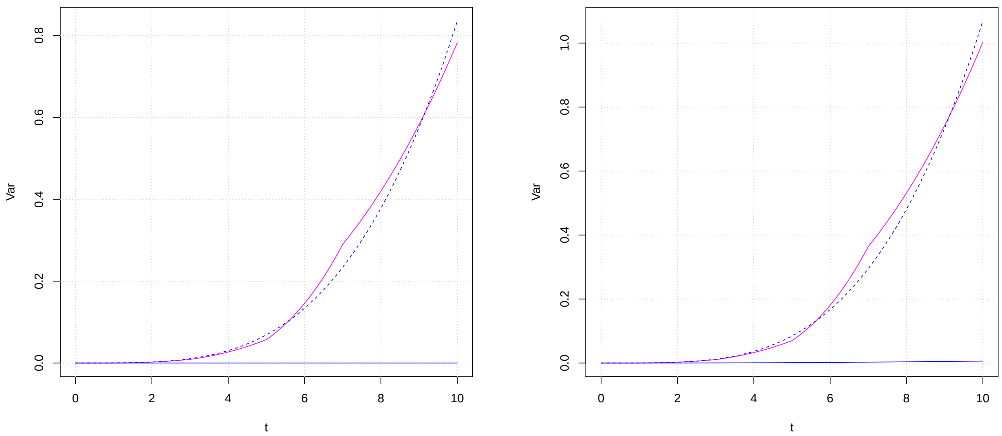
$$v^{\Lambda^{\theta,+}}(t) = v^Y(\Theta^+(t);\Xi^{\star,+}) \geq v^Y(t;\Xi^{\star,+}) = v^{\Lambda^{\varphi,+}}(t) \ .$$

Third, we observe from Figure 4.5 that, in this example at least, the variance of the TC-CIR using $\Xi^\star$ is comparable to the variance of the S-CIR:

$$v^{\Lambda^\theta}(t) \approx v^Y(t;\Xi^\star) = v^{\Lambda^\varphi}(t) \ .$$

The fact that the variance of the S-CIR is expected to be close to that of the corresponding TC-CIR model can be understood intuitively as follows. As explained above the parameter $\Xi \leftarrow \Xi^\star$ computed using (4.7) leads the HAJD y to best fits the market curve, and the clock is used to absorb the remaining discrepancies. Therefore, one expects the clock not to deviate much from the actual time, i.e, $\theta(t) \approx 1$ and the two processes to behave similarly. In particular, the parameters of y^θ are those of y scaled by $\theta(t)$, and $x_t^\theta = \theta(t) y_t^\theta \approx y_t^\theta$, at least when the fit between P^{market} and the base HAJD model P^y is not too poor; see (4.15).

To sum up, we observe that when dealing with CIR++ under a positivity constraint, one has to choose between a valid (but low-volatility) PS-CIR process $\lambda^{\varphi,+}$, or a flawed (by high-volatility) S-CIR one λ^φ. By contrast, the TC-CIR model λ^θ is always valid (Corollary 2), always feature a variance that is larger than the PS-CIR counterpart (Theorem 3), and its variance is actually comparable to the large levels generated by the S-CIR. The TC-CIR thus proves to be a solid challenger to CIR++ models. In particular, its features are specifically interesting when dealing with actual credit risk applications, as we now point out based on two case studies.



(a) $y_0 = h_0$ (b) y_0 optimized

Figure 4.5: Variances of the integrated versions of $\lambda^{\varphi,+}$ (PS-CIR $\Xi = \Xi^{\star,+}$ (left) and $\Xi = \Xi_0^{\star,+}$ (right), solid blue), λ^φ (S-CIR $\Xi = \Xi^\star$ (left) and $\Xi = \Xi_0^\star$ (right), dashed blue), and λ^θ (TC-CIR with $\Xi = \Xi^\star$ (left) and $\Xi = \Xi_0^\star$ (right), magenta).

4.4.3 Pricing CDS options

In this section, we test the models performances in term of the pricing of CDS options. For more details about the pricing methodology of CDS contracts, we refer to Section 2.4.1. Considering the modelling framework of the current chapter, a call option on a CDS contract, $PSO(a,b,k)$ starting at T_a with maturity T_b and spread k, is given by (2.34) when setting

$$\overline{Q}_t(T) := P^{\lambda}_t(T), \quad P_t(T) := P^{r}_t(T), \quad \text{and} \quad S_t = e^{-\Lambda_t} \ .$$

Replacing the base intensity model (λ) by its shifted ($\lambda^{\varphi}, \lambda^{\varphi,+}$) or time-changed ($\lambda^{\theta}$) versions leads to model prices noted $PSO^{\varphi}(a,b,k)$, $PSO^{\varphi,+}(a,b,k)$ and $PSO^{\theta}(a,b,k)$, respectively. Interestingly, these models are equally tractable as they feature similar expressions that can be written in terms of the base process λ or its time integral, Λ. For instance, dropping the Ξ for short,

$$\Lambda^{\varphi}_{T_a} = \Lambda_{T_a} + \int_0^{T_a} \varphi(u)du \ , \quad \Lambda^{\theta}_{T_a} = \Lambda_{\Theta(T_a)} \ ,$$

and

$$\begin{aligned} P^{\lambda^{\varphi}}_{T_a}(u) &= P^{\lambda}_{T_a}(u) e^{\int_{T_a}^{u}\varphi(s)ds} = e^{A_{T_a}(u) - B_{T_a}(u)\lambda_{T_a} + \int_{T_a}^{u}\varphi(s)ds} \ , \\ P^{\lambda^{\theta}}_{T_a}(u) &= P^{\lambda}_{\Theta(T_a)}(u) = e^{A_{\Theta(T_a)}(u) - B_{\Theta(T_a)}(u)\lambda_{\Theta(T_a)}} \ , \end{aligned}$$

Recall that these expressions have a closed form when λ is a (J)CIR process.In the sequel, we compare the models in term of their CDS implied volatilities (see Section 2.4.1).

In all cases, the intensity process λ is calibrated to the market, i.e., $P^{\lambda}(t) = G(t)$. In other words, choosing, e.g., a CIR process for the base intensity process λ combined with the correct shift with (φ^+) or without (φ) positivity constraint, or eventually using the correct clock rate θ, all three models yield the same survival probability curve ($P^{\lambda^{\varphi}}(t) = P^{\lambda^{\varphi,+}}(t) = P^{\lambda^{\theta}}(t) = G(t)$). Hence, all these models agree on the par spread (2.37).

We compare the S-CIR, the PS-CIR and the TC-CIR. The base HAJD process y in TC-CIR is taken to be the same as that of the S-CIR. One can see from Table 4.3 that the S-CIR features large implied volatilities. Recall however that it allows for negative intensities, hence is not appropriate. The PS-CIR model are not capable of generating large volatility levels, in line with the previous discussion. The TC-CIR fits in between: it rules out negative intensities, while maintaining substantial volatility levels.

One might be concerned by the fact that the implied volatilities of the TC-CIR remain relatively small. This can be addressed in two ways. First, one can play with the parameter Ξ. However, the Feller constraint is often required to hold, which sets limits on the process' volatility. Another approach consists of considering a JCIR model as HAJD. Indeed, JCIR is often considered when large volatilities are required. However, as explained in Section 4.2.4, increasing the volatility by boosting the jump activity while maintaining the calibration to a given market curve G reinforces the positivity issue. Fortunately, we do not have this problem in the TC-JCIR. One can drastically increase the jump activity without impacting the positivity of the TC-JCIR. As a consequence, the TC-JCIR seems very much appropriate when one needs a positive but yet high-volatility process. This is illustrated on Table 4.4 using the same jump parameters as those given in Brigo and Mercurio, 2006. We keep the same parameter Ξ as before for the diffusion part, and play with the jump rate (ω) and jump size (α) in the compound Poisson process J. In every case, the clock Θ is chosen such that the model perfectly fits Ford's survival probability curve. Interestingly, the (positive) TC-JCIR model can feature much larger implied volatility levels than the PS-CIR. The results for the S-JCIR are also larger than the PS-CIR, but they are not shown because the negative intensity problem is magnified.

T_a	T_b	CIR++		TC-CIR
		λ^{φ}	$\lambda^{\varphi,+}$	λ^{θ}
1	3	67.12%	1.03%	43.68%
1	5	45.10%	1.25%	26.92%
1	7	27.72%	1.64%	16.86%
1	10	21.52%	1.65%	12.86%
3	5	61.30%	0.85%	57.16%
3	7	34.88%	1.15%	36.33%
3	10	27.65%	1.08%	27.17%
5	7	34.81%	1.16%	42.60%
5	10	30.96%	0.93%	31.19%
7	10	45.37%	0.63%	38.93%

T_a	T_b	CIR++		TC-CIR
		λ^{φ}	$\lambda^{\varphi,+}$	λ^{θ}
1	3	65.21%	9.48%	44.00%
1	5	42.35%	5.88%	26.56%
1	7	25.30%	4.87%	16.30%
1	10	19.37%	2.94%	12.27%
3	5	63.87%	8.02%	58.67%
3	7	35.20%	4.50%	36.10%
3	10	27.02%	3.44%	26.62%
5	7	35.77%	4.59%	43.32%
5	10	30.35%	3.70%	30.61%
7	10	45.05%	5.16%	39.40%

Table 4.3: Black volatilities for at-the-money ($k = s_0(a,b)$) CDS options implied by CIR++ models (S-CIR and PS-CIR) and the TC-CIR model with $y_0 = h_0$ (left) and y_0 optimized (right) using Monte Carlo simulation (500K paths with time step 0.01). In all the considered cases, the CIR++ model without shift is not valid since $\inf\{\lambda_t^{\varphi},\ t \in [0,T_a]\} < 0$. Among the two valid intensity models ($\lambda_t^{\varphi,+}$ and λ_t^{θ}), the latter exhibits a much higher implied volatility.

T_a	T_b	TC-JCIR (ω,α)		
		$(0,0)$	$(0.1,0.1)$	$(0.15,0.15)$
1	3	43.68%	79.04%	100.17%
1	5	26.92%	48.07%	69.69%
1	7	16.86%	30.43%	44.00%
1	10	12.86%	23.33%	33.75%
3	5	57.16%	65.50%	82.60%
3	7	36.33%	42.85%	53.17%
3	10	27.17%	33.17%	41.36%
5	7	42.60%	49.13%	60.11%
5	10	31.19%	37.34%	45.78%
7	10	38.93%	40.59%	46.02%

Table 4.4: Black volatilities for at-the-money ($k = s_0(a,b)$) CDS options implied by the TC-JCIR model (jump arrival rate ω and jump size α) using Monte Carlo simulation (10^6 paths with time step 0.01) and parameter set $\Xi = \Xi^{\star}$ but for various jump parameters (α,ω).

4.4.4 Wrong-way risk impact in credit valuation adjustments

All the models considered in this chapter follow a Cox setup and thus are arbitrage-free since the immersion property is satisfied. Therefore, according to (2.38), the current ($t = 0$) value of the CVA is expressed as:

$$\mathrm{CVA} = \mathbb{E}\left[(1-R)\int_0^T \frac{V_u^+}{\beta_u}\lambda_u e^{-\Lambda_u} du\right] . \tag{4.18}$$

where V stands for the exposure process.

The CVA of the shifted and the time-changed models, CVA^φ and CVA^θ, correspond to above expression, replacing (λ, Λ) by $(\lambda^\varphi, \Lambda^\varphi)$ and $(\lambda^\theta, \Lambda^\theta)$, respectively. The purpose of this section is to illustrate the order of magnitude of CVA figures that can be obtained with either models. In particular, we do not aim at representing a specific exposure. Instead, we simplify the analysis by considering two prototypical dynamics:

$$\begin{aligned} dV_t &= \nu dW_t^V \ , \\ dV_t &= \left(\gamma(T-t) - \frac{V_t}{T-t}\right) dt + \nu dW_t^V \ . \end{aligned}$$

where W^V is an $\mathbb{F}$-Brownian motion. The first SDE is that of a martingale, and can depict the evolution of the discounted price of a forward contract prior to its cashflow date. The second SDE corresponds to a Brownian bridge with drift, and mimics the dynamics of the discounted price of an asset paying continuous dividends. These two models have been previously used in Vrins, 2017 and Brigo and Vrins, 2018 to describe, in a schematic way, exposures of FRA and IRS. Calibration to actual exposures give indicative value for the parameters.

In general, there is no reason to assume that the Brownian motion driving the default intensity (W) would be independent of the Brownian motion driving the exposure (W^V): it depends on the problem at hand. Usually, we consider the general case of WWR effect, obtained by introducing a correlation between the Brownian drivers, i.e., we assume that $dW_t dW_t^V = \rho dt$[8]. In particular, for the TC-CIR, we apply the synchronization procedure devised in Chapter 3 in order to preserve the correlation after time-changing the intensity process. In the special

[8] The market credit correlation ρ can be estimated using basic statistical techniques (see, for example, Fleck and Schmidt, 2005 or Rosen and Saunders, 2012) or historical data (see Cepedes et al., 2010).

case where the default time of the counterparty is independent from the discounted exposure (i.e., $\rho = 0$, that is no wrong-way risk) one can easily deduce from (4.18) the *independent* CVA formula

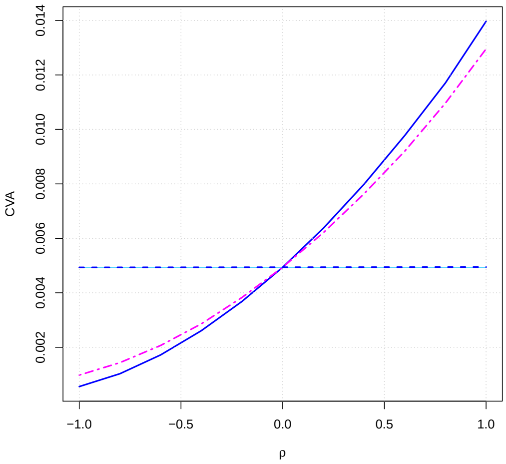
$$\mathrm{CVA}^{\perp} = -(1-R)\int_0^T \mathbb{E}\left[V_u^+\right] d\,\mathbb{E}\left[e^{-\Lambda_u}\right] = (1-R)\int_0^T f^{\lambda}(u)\,\mathbb{E}\left[V_u^+\right] P^{\lambda}(u)du\,.$$

Recall that whatever the chosen model, it is assumed to be calibrated to the survival probability curve G, extracted from CDS prices. This leads to $P^{\lambda}(u) = G(t)$, and to the optimal shift and clock functions, namely φ or φ^+ in the S-CIR and PS-CIR cases, and Θ for the TC-CIR. In this case, $\mathrm{CVA}^{\perp}$ does not depend on the default model:

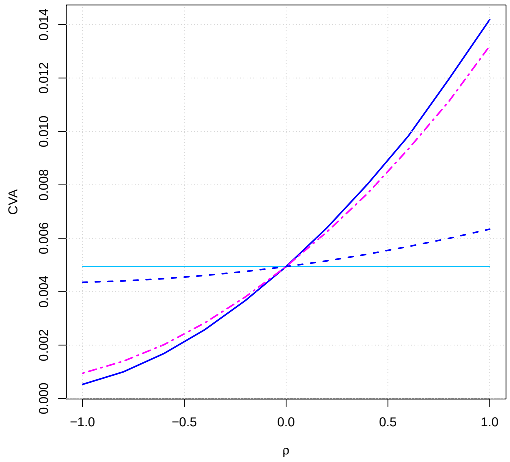
$$\mathrm{CVA}^{\perp} = -(1-R)\int_0^T \mathbb{E}\left[V_u^+\right] dG(u) = (1-R)\int_0^T h(u)\mathbb{E}\left[V_u^+\right] G(u)du\,.$$

However, the independent case $\rho = 0$ is unrealistic, and may lead to severe over or underestimations of CVA Kim and Leung, 2016; Brigo and Vrins, 2018; Breton and Marzouk, 2018. Under WWR, CVA becomes model-dependent. Figure 4.6 shows the evolution of CVA with respect to ρ for three different models: λ^{φ} (CIR++ without constraint, solid blue), $\lambda^{\varphi,+}$ (CIR++ with constraint, dashed blue) and λ^{θ} (TC-CIR, dashed magenta), all calibrated to Ford's survival probability curve G as before. Under no-WWR, the CVA is equal to the independent CVA (cyan): it is flat, model-free and can be computed using a simple integration. Under WWR, the CVAs are computed using Monte Carlo simulations (100K paths, time step of 0.01) and adaptive control variate [9]. The TC-CIR and S-CIR models exhibit the largest WWR effects and seem therefore appropriate to deal with high WWR applications. Recall that only TC-CIR is valid here as S-CIR gives room to negative intensities. The PS-CIR however is almost flat, equal to the independent CVA. This can be understood from the fact that WWR is essentially a covariance effect between V and $e^{-\Lambda}$. Hence, the models featuring large variance for Λ exhibit larger WWR effects at any (non-zero) fixed correlation level ρ. Eventually, TC-CIR provides an appealing trade-off: on the one hand, as the PS-CIR, it rules out the negative intensity problem inherent to the S-CIR model. But on the other hand, it preserves, to some extend, the variance of the S-CIR model, and therefore exhibits a much larger variance compared to PS-CIR.

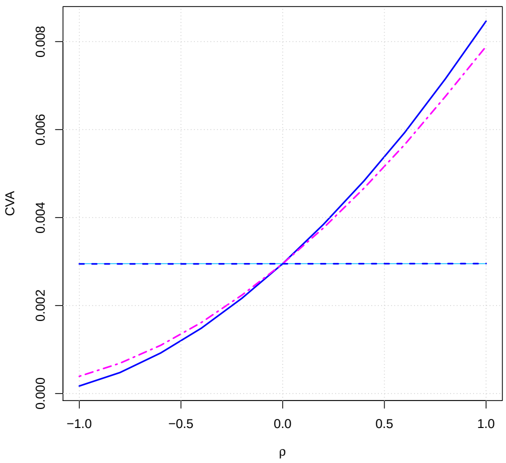
[9]See Chapter 3 for the implementation of the adaptive control variate applied on CVA computation.



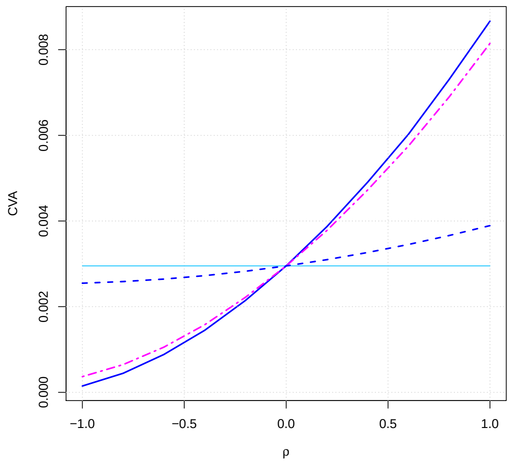
(a) $dV_t = \nu dW_t^V$, $y_0 = h_0$



(b) $dV_t = \nu dW_t^V$, y_0 optimized



(c) $dV_t = \left(\gamma(T-t) - \frac{V_t}{T-t}\right) dt + \nu dW_t^V$, $y_0 = h_0$



(d) $dV_t = \left(\gamma(T-t) - \frac{V_t}{T-t}\right) dt + \nu dW_t^V$, y_0 optimized

Figure 4.6: Impact of the exposure-credit correlation ρ on CVA levels for prototypical 5Y Forward (top) and Swap exposure (bottom) with $y_0 = h_0$ (left) or y_0 optimized (right), $\nu = 8\%$ and $\gamma = 0.1\%$. The curves correspond to different intensity models: $\lambda^{\varphi,+}$ (PS-CIR $\Xi = \Xi^{\star,+}$ (left) and $\Xi = \Xi_0^{\star,+}$ (right), dotted blue), λ^{φ} (S-CIR $\Xi = \Xi^{\star}$ (left) and $\Xi = \Xi_0^{\star}$ (right), solid blue), and λ^{θ} (TC-CIR $\Xi = \Xi^{\star}$ (left) and $\Xi = \Xi_0^{\star}$ (right), dotted magenta). The case without wrong-way risk corresponds to the flat (cyan) line.

4.5 Conclusions

The calibration problem consists of finding the parameters of a model x so as to perfectly fit a given market curve. The perfect fit is an important feature in a pricing context, that is connected to no-arbitrage opportunities and corrects valuation of trading positions. This calls for two important features: the model x must be (i) flexible enough (to be able to generate various shapes) and (ii) tractable enough (to facilitate the parameters' optimization procedure). Time-homogeneous affine models like Vasicek, CIR or JCIR are very good candidates in this respect, and are widely used in interest rates and credit risk modelling. However, as such, they only feature a couple of constants and hence lack calibration flexibility. The deterministic shift extension offers an appealing solution. It consists of starting with a tractable base model y, that is shifted in a deterministic way with a function φ. The resulting process $x_t = y_t + \varphi(t)$ becomes fully flexible. Indeed, any discount or survival probability curve can be generated by such a model. Moreover, it has a tractability level that is very similar to that of y because φ is deterministic. Eventually, for every market curve, the shift $\varphi^\star$ that leads to the perfect fit is known in closed form, as a function of the y parameters and the market curve. However, this method is less appealing when the model x needs to fulfill some range constraints. Among those, non-negativity is of primary importance when modelling interest rates (depending on the type of economy at hand), credit spread or default intensities. In the deterministic shift approach indeed, starting with a non-negative base process y is not enough to guarantee that so will be x, without additional constraint on φ. Furthermore, this constraint becomes more and more severe when increasing the process volatility, due to the zero lower bound.

It seems obvious to rule out models allowing for "negative volatilities". However, surprisingly, the same does not seem to apply when it comes to "negative intensities". Yet, both are equally flawed. We believe the reason is twofold: first, negative intensities do not directly generate numerical problems (in contrast with volatilities that often appear in square-roots), so that the issue is less "obvious", second, there is a lack of a sound alternative. The positivity constraint can be dealt with by including a non-negativity constraint on φ. However, this again raises two problems. First the parameter optimization problem becomes more difficult and second, the resulting process x then features a much lower variance than without the constraint, which contradicts empirical evidences. Therefore, one often prefers to disregard the "negative intensities" issue, giving the priority to stochasticity and perfect fit.

In this chapter, we develop such an alternative. It simply consists of time-changing a positive homogeneous affine jump-diffusion. The model remains tractable, positive, the optimal clock is found by simple inversion and features larger implied volatility compared to the shift approach. Moreover, the perfect fit is achievable for a broad class of discount curves, including all decreasing discount curves. In this case, the corresponding time change function Θ is available using a simple inversion. This method thus proves to be a competitive challenger to the shift approach, at least under the (very common) positivity constraint, and when large volatility levels are needed.

Appendices

4.A Properties of some Homogeneous Affine Jump-Diffusions

Let y be a $\mathbb{F}$-adapted jump-diffusion introduced in Definition 3 and $Y_t := \int_0^t y_u du$ its integrated version. We denote $v^x(t) = v^x(t;\Xi) := \mathbb{V}[x_t]$ the variance of a stochastic process x parametrized by Ξ at time t. Without explicit mention, all the results below that are given without proofs can be found in, e.g., Brigo and Mercurio, 2006. New results are given in lemmas for further reference.

4.A.1 Vasicek model

The Vasicek model corresponds to the special HAJD case $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\varrho, -\kappa, \eta^2, 0, 0, 0)$. The A^y, B^y functions in (2.27) are given by:

$$\begin{aligned} A_s^y(t;\Xi) &= \left(\varrho - \frac{\eta^2}{2\kappa^2}\right)(B_s^y(t;\Xi) + s - t) - \frac{\eta^2}{4\kappa} B_s^y(t;\Xi)^2 \ , \\ B_s^y(t;\Xi) &= \frac{1}{\kappa}\left(1 - \mathrm{e}^{-\kappa(t-s)}\right) \ . \end{aligned}$$

The forward curve associated to this model is proven to be

$$f_s^{\mathrm{VAS}}(t) := (1-\mathrm{e}^{-\kappa(t-s)})\frac{\kappa^2\varrho - \eta^2/2}{\kappa^2} + \frac{\eta^2}{2\kappa^2}\,\mathrm{e}^{-\kappa(t-s)}(1-\mathrm{e}^{-\kappa(t-s)}) + y_t\,\mathrm{e}^{-\kappa(t-s)} \ . \tag{4.19}$$

Moreover, both y and Y are Normally distributed at all times, with

$$\begin{aligned} \mathbb{E}[y_t] &= y_0^{-\kappa t} + \varrho\kappa B_0^y(t;\Xi) \ , \\ \mathbb{E}[Y_t] &= y_0 B_0^y(t;\Xi) + \varrho\left(t - B_0^y(t;\Xi)\right) \ , \\ v^y(t) &= \frac{\eta^2}{2\kappa}\left(1 - e^{-2\kappa t}\right) \ , \\ v^Y(t) &= \frac{\eta^2}{\kappa^2}\left[t + \frac{1 - e^{-2\kappa t}}{2\kappa} - 2B_0^y(t;\Xi)\right] \ . \end{aligned}$$

Lemma 10. *Let y be a Vasicek process and Y its time integral. The functions $v^y(t)$ and $v^Y(t)$ are increasing with respect to t.*

Proof. It is obvious for $v^y(t)$, and for $v^Y(t)$, a few manipulations lead to

$$\frac{d}{dt}v^Y(t) = \left(\frac{\eta}{\kappa}(1 - e^{-\kappa t})\right)^2 \geq 0 \ .$$

$\square$

4.A.2 CIR model

The CIR model corresponds to the special HAJD case $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\varrho, -\kappa, 0, \delta^2, 0, 0)$. The A^y, B^y functions in eq. (2.27) are given by:

$$
\begin{aligned}
A_s^y(t;\Xi) &= \frac{2\kappa\varrho}{\delta^2} \ln \frac{2\gamma \exp\{(\kappa+\gamma)(t-s)/2)\}}{2\gamma + (\kappa+\gamma)(\exp\{(t-s)\gamma\}-1)}, \\
B_s^y(t;\Xi) &= \frac{2(\exp\{(t-s)\gamma\}-1)}{2\gamma + (\kappa+\gamma)(\exp\{(t-s)\gamma\}-1)} .
\end{aligned}
$$

where $\gamma := \sqrt{\kappa^2 + 2\delta^2}$. The forward curve associated to this model is given by Brigo and Mercurio, 2006

$$
f_s^{\mathrm{CIR}}(t) := \frac{2\kappa\varrho(e^{(t-s)\gamma}-1)}{2\gamma + (\kappa+\gamma)(e^{(t-s)\gamma}-1)} + y_t \frac{4\gamma^2 e^{(t-s)\gamma}}{[2\gamma + (\kappa+\gamma)(e^{(t-s)\gamma}-1)]^2} . \tag{4.20}
$$

Important characteristics of the CIR processes can be computed explicitly, (see, e.g., Dufresne, 2001). For instance, y is distributed as a non-central chi-squared. The two first order moments of y and Y are respectively given by

$$
\begin{aligned}
\mathbb{E}[y_t] &= y_0 e^{-\kappa t} + \varrho\left(1 - e^{-\kappa t}\right) , \\
\mathbb{E}[y_t^2] &= y_0^2 e^{-2\kappa t} + \left(\frac{\delta^2}{2\kappa} + \varrho\right)\left[2y_0\left(e^{-\kappa t} - e^{-2\kappa t}\right) + \varrho\left(1 - e^{-\kappa t}\right)^2\right] , \\
\mathbb{E}[Y_t] &= t\varrho + \frac{(y_0 - \varrho)}{\kappa}(1 - e^{-\kappa t}) , \\
\mathbb{E}[Y_t^2] &= \left(\frac{y_0 - \varrho}{\kappa}\right)^2 + \delta^2\left(\frac{y_0}{\kappa^3} - \frac{5\varrho}{2\kappa^3}\right) + t\varrho\left(\frac{2y_0}{\kappa} - \frac{2\varrho}{\kappa} + \frac{\delta^2}{\kappa^2}\right) + t^2\varrho^2 \\
&\quad + e^{-\kappa t}\left[-2\left(\frac{y_0 - \varrho}{\kappa}\right)^2 + \frac{2\varrho\delta^2}{\kappa^3} + t\left(-\frac{2y_0\varrho}{\kappa} + \frac{2\varrho^2}{\kappa} - \frac{2y_0\delta^2}{\kappa^2} + \frac{2\varrho\delta^2}{\kappa^2}\right)\right] \\
&\quad + e^{-2\kappa t}\left[\left(\frac{y_0 - \varrho}{\kappa}\right)^2 - \frac{y_0\delta^2}{\kappa^3} + \frac{\varrho\delta^2}{2\kappa^3}\right] .
\end{aligned}
$$

In contrast with the Vasicek model, the variance of the CIR is not always increasing monotonously with time; it depends on the parameters. However, the variance of the integrated CIR is increasing. These properties are proven in the next lemma, and will be central in the proof of Theorem 3.

Lemma 11. *Let y be a CIR process and Y its time integral. Then,*

$$
\begin{aligned}
v^y(t) &= \frac{\delta^2}{\kappa}\left[y_0(e^{-\kappa t} - e^{-2\kappa t}) + \frac{\varrho}{2\kappa}(1-e^{-\kappa t})^2\right] , && (4.21)\\
v^Y(t) &= \frac{\delta^2}{\kappa^3}\left[\left(2\varrho - 2\kappa t(y_0-\varrho) - (y_0 - \varrho/2)\, e^{-\kappa t}\right) e^{-\kappa t}\right] \\
&\quad + \frac{\delta^2}{\kappa^3}\left[t\kappa\varrho + (y_0 - 5\varrho/2)\right] . && (4.22)
\end{aligned}
$$

The function $v^y(t)$ is increasing if $\varrho \geq y_0$. Otherwise, it is first increasing up to a time $t^\star$, and then decreasing on $(t^\star, \infty)$. By contrast, $v^Y(t)$ is always increasing.

Proof. The computation of the variances is trivial from the first two moments recalled above. The derivative of the variance of the CIR is given by

$$
\begin{aligned}
\frac{d}{dt}v^y(t) &= y_0\delta^2(-e^{-\kappa t} + 2e^{-2\kappa t}) + \varrho\delta^2 e^{-\kappa t}(1-e^{-\kappa t}) \\
&= \delta^2 e^{-\kappa t}(\varrho - y_0) - \delta^2 e^{-2\kappa t}(\varrho - 2y_0) .
\end{aligned}
$$

This expression has a root on the positive half-line at

$$
t^\star = \frac{1}{\kappa}\ln\left(1 + \frac{y_0}{y_0 - \varrho}\right)
$$

only if $y_0 > \varrho$. Otherwise, $v^y(t)$ is always increasing in t. The derivative of the variance of the integrated CIR with respect of time can be written, after some manipulations, as

$$
\frac{d}{dt}v^Y(t) = \frac{\delta^2}{\kappa^2}\left[2y_0 e^{-\kappa t}(e^{-\kappa t} - (1-\kappa t)) + \varrho(1 - (2\kappa t + e^{-\kappa t})e^{-\kappa t})\right] .
$$

Because $e^{-x} \geq 1 - x$ for all $x \geq 0$ and y_0, κ are positive constants, the first term is positive for all $t \geq 0$. On the other hand, $\varrho > 0$, and it is enough to check that $1 - g(\kappa t) \geq 0$ for all $t \geq 0$, with $g(x) := (2x + e^{-x})e^{-x}$. Clearly, $g(0) = 1$ and $g'(x) = 2e^{-x}(1 - x - e^{-x}) \leq 0$ for all $x \geq 0$. Hence, $1 - g(\kappa t) \geq 0$ for all $t \geq 0$. □

4.A.3 JCIR model

The characteristics of the JCIR can be obtained by adjusting those of the corresponding CIR, i.e., with same initial value and diffusion parameters. We note z the former and y the latter, and similarly for their integrated versions (Z and Y, respectively). Hence, if the parameter set for the CIR (y, Y) is $\Xi_0 = (\kappa, \varrho, \delta, 0, 0, y_0)$, the parameter set of the corresponding JCIR (z, Z) is $\Xi_0 = (\kappa, \varrho, \delta, \alpha, \omega, z_0)$ with $z_0 = y_0$ and $\alpha, \omega \geq 0$. The functions associated to the discount curve are given by

$$\begin{aligned} A^z_s(t;\Xi) &= A^y_s(t;\Xi) + \frac{\alpha\omega}{\delta^2/2 - \kappa\alpha - \alpha^2} \ln \frac{2\gamma e^{\frac{\gamma+\kappa+2\alpha}{2}(t-s)}}{2\gamma + (\kappa+\gamma+2\alpha)(e^{(t-s)\gamma} - 1)} , \\ B^z_s(t;\Xi) &= B^y_s(t;\Xi) . \end{aligned}$$

From the above functions, it is easy to see that the forward curve associated to this model reads as

$$f^{\mathrm{JCIR}}_s(t) := f^{\mathrm{CIR}}_s(t) + \frac{2\omega\alpha(e^{(t-s)\gamma} - 1)}{2\gamma + (\kappa+\gamma+2\alpha)(e^{(t-s)\gamma} - 1)} , \tag{4.23}$$

where $f^{\mathrm{CIR}}_s(t)$ is given in (4.20). For every valid parameters, $f^{\mathrm{JCIR}}_s(t) \geq f^{\mathrm{CIR}}_s(t)$ for all $t \geq s$. Regarding the moments, we have the following result.

Lemma 12. *Let y (resp. Y) be a CIR (resp. integrated CIR) and z (resp. Z) be a JCIR (resp. integrated JCIR) with same initial value, same diffusion parameters but with jumps governed by (ω, α). Then,*

$$\begin{aligned} \mathbb{E}[z_t] &= \mathbb{E}[y_t] + \frac{\omega\alpha}{\kappa}(1 - e^{-\kappa t}) , \\ \mathbb{E}[Z_t] &= \mathbb{E}[Y_t] + \frac{\omega\alpha}{\kappa^2}\left(\kappa t - (1 - e^{-\kappa t})\right) . \\ v^z(t) &= v^y(t) + \omega\alpha\left[\frac{\delta^2}{2}\left(\frac{1 - e^{-\kappa t}}{\kappa}\right)^2 + \alpha\frac{1 - e^{-2\kappa t}}{\kappa}\right] , \\ v^Z(t) &= v^Y(t) + \frac{\alpha\omega}{\kappa^3}\left[\frac{1 - e^{-\kappa t}}{\kappa}\left(\xi(3 - e^{-\kappa t}) - 4\delta^2\right) + 2\delta^2 t e^{-\kappa t} + t\left(2\alpha\kappa + \delta^2\right)\right] , \end{aligned}$$

where $\xi := \delta^2/2 - \alpha\kappa$. The function $v^z(t)$ is increasing with respect to t unless $y_0 > \varrho + \omega\alpha/\kappa$, in which case it is first increasing up to a time t_1, and then decreasing on (t_1, ∞). Moreover, $v^z(t) \geq v^y(t)$, $v^Z(t) \geq v^Y(t)$ and $v^Z(t)$ is always increasing.

Proof. Applying Ito's lemma we can solve the JCIR SDE (4.17) by

$$z_t = z_0 e^{-\kappa t} + \varrho(1 - e^{-\kappa t}) + \delta\int_0^t e^{-\kappa(t-s)}\sqrt{z_s}dW_s + \int_0^t e^{-\kappa(t-s)}dJ_s , \tag{4.24}$$

and find the SDE governing the integrated JCIR process

$$Z_t = t\varrho + \frac{(z_0 - \varrho)}{\kappa}(1 - e^{-\kappa t}) + \delta\int_0^t\int_0^s e^{-\kappa(s-u)}\sqrt{z_u}dW_u ds + \int_0^t\int_0^s e^{-\kappa(s-u)}dJ_u ds . \tag{4.25}$$

From (4.24), we can write

$$
\begin{aligned}
\mathbb{E}[z_t] &= z_0 e^{-\kappa t} + \varrho(1-e^{-\kappa t}) + \mathbb{E}\left[\int_0^t e^{-\kappa(t-s)} dJ_s\right] \\
&= \mathbb{E}[y_t] + \omega\alpha \int_0^t e^{-\kappa(t-s)} ds \\
&= \mathbb{E}[y_t] + \frac{\omega\alpha}{\kappa}(1-e^{-\kappa t}) \; .
\end{aligned}
$$

Using Ito isometry,

$$
\begin{aligned}
v^z(t) &= \mathbb{E}[(z_t - \mathbb{E}[z_t])^2] \\
&= \mathbb{E}\left[\left(\delta \int_0^t e^{-\kappa(t-s)} \sqrt{z_s} dW_s + \int_0^t e^{-\kappa(t-s)} dJ_s - \omega\alpha \int_0^t e^{-\kappa(t-s)} ds\right)^2\right] \\
&= \mathbb{E}\left[\left(\delta \int_0^t e^{-\kappa(t-s)} \sqrt{z_s} dW_s\right)^2\right] + \mathbb{E}\left[\left(\int_0^t e^{-\kappa(t-s)} (dJ_s - \omega\alpha ds)\right)^2\right] \\
&= v^y(t) + \frac{\delta^2 \omega\alpha}{\kappa} \int_0^t \mathrm{e}^{-2\kappa(t-s)} (1-e^{-\kappa s}) ds + 2\omega\alpha^2 \int_0^t e^{-2\kappa(t-s)} ds \\
&= v^y(t) + \omega\alpha \left[\frac{\delta^2}{2}\left(\frac{1-e^{-\kappa t}}{\kappa}\right)^2 + \alpha \frac{1-e^{-2\kappa t}}{\kappa}\right] \; .
\end{aligned}
$$

Using a similar procedure applied to (4.25) combined with Fubini's theorem, one can derive the expectation and variance of the integrated JCIR :

$$
\begin{aligned}
\mathbb{E}[Z_t] &= t\varrho + \frac{(z_0 - \varrho)}{\kappa}(1-e^{-\kappa t}) + \mathbb{E}\left[\int_0^t \int_0^s e^{-\kappa(s-u)} dJ_u ds\right] \\
&= \mathbb{E}[Y_t] + \mathbb{E}\left[\int_0^t \int_s^t e^{-\kappa u} du e^{\kappa s} dJ_s\right] \\
&= \mathbb{E}[Y_t] + \frac{\omega\alpha}{\kappa} \int_0^t (1-e^{-\kappa(t-s)}) ds \\
&= \mathbb{E}[Y_t] + \frac{\omega\alpha}{\kappa^2}\left(\kappa t - (1-e^{-\kappa t})\right)
\end{aligned}
$$

and

$$\begin{aligned}
v^Z(t) &= \mathbb{E}[(Z_t - \mathbb{E}[Z_t])^2] \\
&= \mathbb{E}\left[\left(\delta \int_0^t \int_0^s e^{-\kappa(s-u)}\sqrt{z_u}dW_u ds + \int_0^t \int_0^s e^{-\kappa(s-u)}dJ_u ds - \mathbb{E}\left[\int_0^t \int_0^s e^{-\kappa(s-u)}dJ_u ds\right]\right)^2\right] \\
&= \mathbb{E}\left[\left(\frac{\delta}{\kappa}\int_0^t (1-e^{-\kappa(t-s)})\sqrt{z_s}dW_s + \frac{1}{\kappa}\int_0^t (1-e^{-\kappa(t-s)})dJ_s - \frac{\omega\alpha}{\kappa}\int_0^t (1-e^{-\kappa(t-s)})ds\right)^2\right] \\
&= \mathbb{E}\left[\left(\frac{\delta}{\kappa}\int_0^t (1-e^{-\kappa(t-s)})\sqrt{z_s}dW_s\right)^2\right] + \mathbb{E}\left[\left(\frac{1}{\kappa}\int_0^t (1-e^{-\kappa(t-s)})(dJ_s - \omega\alpha ds)\right)^2\right] \\
&= v^Y(t) + \frac{\delta^2\omega\alpha}{\kappa^2}\int_0^t (1-e^{-\kappa(t-s)})^2(1-e^{-\kappa s})ds + \frac{2\omega\alpha^2}{\kappa^2}\int_0^t (1-e^{-\kappa(t-s)})^2 ds \\
&= v^Y(t) + \frac{\alpha\omega}{\kappa^3}\left[\frac{1-e^{-\kappa t}}{\kappa}\left((\delta^2/2 - \alpha\kappa)(3-e^{-\kappa t}) - 4\delta^2\right) + 2\delta^2 t e^{-\kappa t} + t\left(2\alpha\kappa + \delta^2\right)\right] .
\end{aligned}$$

Notice that the above results can be obtained using another procedure, namely by deriving once or twice the characteristic function $\Psi_t(u,v) = \mathbb{E}[e^{uz_t + qZ_t}]$ of (z_t, Z_t) which can be recovered from equation (B.9) in the draft version of Duffie and Gârleanu's paper, that is available for download on the authors' webpage and published in Duffie and Gârleanu, 2001. This procedure is however much heavier. The derivative of the variance of the JCIR is given by

$$\begin{aligned}
\frac{d}{dt}v^z(t) &= y_0\delta^2(-e^{-\kappa t} + 2e^{-2\kappa t}) + \varrho\delta^2 e^{-\kappa t}(1-e^{-\kappa t}) + \frac{\delta^2\omega\alpha}{\kappa}e^{-\kappa t}(1-e^{-\kappa t}) \\
&\quad + 2\omega\alpha^2 e^{-2\kappa t} \\
&= \delta^2 e^{-\kappa t}(\varrho - y_0 + \omega\alpha/\kappa) - \delta^2 e^{-2\kappa t}(\varrho - 2y_0 + \omega\alpha/\kappa - 2\omega\alpha^2/\delta^2) .
\end{aligned}$$

This expression has a root on the positive half-line at

$$t_1 := \frac{1}{\kappa}\ln\left(1 + \frac{y_0 + 2\omega\alpha^2/\delta^2}{y_0 - \varrho - \omega\alpha/\kappa}\right) \tag{4.26}$$

only if $y_0 > \varrho + \omega\alpha/\kappa$. Otherwise, $v^y(t)$ is always increasing in t.
It seems intuitive that the variance of the JCIR cannot be smaller than that of the CIR, and similarly for the integrated versions. Because of the mean reverting effect however, this needs to be confirmed. It is obvious that $v^z(t) - v^y(t) \geq 0$. The term associated to $v^Z(t) - v^Y(t)$ starts at zero (since obviously $v^Z(0) = v^Y(0) = 0$). This difference is increasing:

$$\frac{d}{dt}(v^Z(t) - v^Y(t)) = \frac{\delta^2\omega\alpha}{\kappa^3}\left(1 - (2\kappa t + e^{-\kappa t})e^{-\kappa t}\right) + \frac{2\omega\alpha^2}{\kappa^2}\left(1 - e^{-\kappa t}\right)^2 \geq 0 .$$

Indeed, the second term is obviously positive and the first term takes the form $\frac{\delta^2\omega\alpha}{\kappa^3}(1-g(\kappa t))$ where the function $g(x) = (2x+e^{-x})e^{-x}$ is shown to be bounded by 1 for $x \geq 0$ in the proof of Lemma 11. This shows that $v^Z(t) \geq v^Y(t)$. Because both $v^Y(t)$ (from Lemma 11) and $v^Z(t) - v^Y(t)$ are increasing; $v^Z(t)$ is itself increasing. □

4.B Proof of Lemma 6

We start with the lemma giving sufficient conditions to swap the expectation and derivative operators which can be found in, e.g., Pagès, 2018.

Lemma 13. *Let I be a nontrivial interval of $\mathbb{R}$, $\mathcal{B}(I)$ the Borel set of I and $\Psi : I \times \Omega \to \mathbb{R}$, $(x, \omega) \mapsto \Psi(x, \omega)$ be a $\mathcal{B}(I) \otimes \mathcal{G}$-measurable function. If the function Ψ satisfies:*

(i) For every $x \in I$, the random variable $\Psi(x,\, \omega) \in L^1$,

(ii) $\Psi_x(x, \omega) := \frac{\partial \Psi(x,\omega)}{\partial x}$ exists for all $x \in I$ a.s.,

(iii) There exists $Z \in L^1$ such that for every $x \in I$,

$$|\Psi_x(x, \omega)| \leq Z(\omega) \text{ a.s. } .$$

Then the function $\psi(x) := \mathbb{E}[\Psi(x,\, \omega)]$ is defined and differentiable at every $x \in I$ with derivative

$$\frac{d\psi(x)}{dx} = \mathbb{E}\left[\Psi_x(x, \omega)\right] \; .$$

We now proceed with the proof of Lemma 6. The solution of the SDE is

$$y_t = \underbrace{y_0 + \int_0^t \mu(s, y_s)ds}_{A_t} + \underbrace{\int_0^t \sigma(s, y_s)dW_s}_{M_t} + J_t \;\Rightarrow\; |y_t| \leq |A_t| + |M_t| + J_t \; ,$$

showing that

$$S_T^y \leq \underbrace{\sup_{t\in[0,T]} |A_t|}_{S_T^A} + \underbrace{\sup_{t\in[0,T]} |M_t|}_{S_T^M} + \underbrace{\sup_{t\in[0,T]} J_t}_{S_T^J} \; .$$

We show in the sequel that S_T^A, S_T^M and S_T^J are integrable. This would conclude the proof since it would lead to

$$\mathbb{E}[|S_T^y|] = \mathbb{E}[S_T^y] \leq \mathbb{E}[S_T^A] + \mathbb{E}[S_T^M] + \mathbb{E}[S_T^J] < \infty \; .$$

Suppose that $\mathbb{E}[\int_0^T |\mu(s,y_s)|ds] < \infty$. Then,

$$S_T^A = \sup_{t\in[0,T]} |y_0 + \int_0^t \mu(s,y_s)ds| \le y_0 + \sup_{t\in[0,T]} \int_0^t |\mu(s,y_s)|ds \le y_0 + \int_0^T |\mu(s,y_s)|ds \ .$$

showing that $\mathbb{E}[|S_T^A|] = \mathbb{E}[S_T^A] \le y_0 + \mathbb{E}[\int_0^T |\mu(s,y_s)|ds] < \infty$.

On the other hand, M is a martingale, so that $|M|$ is a submartingale:

$$\mathbb{E}[|M_t|\ |\mathcal{F}_s] \ge |\,\mathbb{E}[M_t|\mathcal{F}_s]| = |M_s| \ .$$

We can then apply Doob's inequality,

$$\mathbb{E}[S_T^M] = \mathbb{E}[\sup_{t\in[0,T]} |M_t|] \le \frac{e}{e-1}\left(1 + \mathbb{E}[|\ |M_T| \log^+ |M_T|\ |]\right) \ .$$

Using $-\frac{1}{e} \le x\log x \le x^2$ for $x \ge 0$, $|x\log^+ x| = x\log^+ x \le |x\log x| \le \max(e^{-1}, x^2)$:

$$\begin{aligned} \mathbb{E}[|M_T|\log^+|M_T|] &\le \mathbb{E}[\max(e^{-1}, M_T^2)] \\ &= \mathbb{E}[e^{-1}1_{\{M_T^2\le e^{-1}\}} + M_T^2 1_{\{M_T^2 > e^{-1}\}}] \\ &\le e^{-1} + \mathbb{E}[M_T^2] \ . \end{aligned}$$

Hence,

$$\mathbb{E}[S_T^M] \le \frac{e}{e-1}\left(1 + e^{-1} + \mathbb{E}[M_T^2]\right) \ .$$

Using Ito isometry, $\mathbb{E}[M_T^2] = \mathbb{E}\left[\int_0^T \sigma^2(t,y_t)dt\right]$ which is bounded, by assumption.

Similarly, one can prove that $\mathbb{E}[S_T^J]$ is finite by applying the Doob's inequality to the martingale $(J_t - \omega\alpha t), t \le T$. Indeed,

$$J_t = |J_t - \omega\alpha t + \omega\alpha t| \le |J_t - \omega\alpha t| + \omega\alpha t, \quad \forall\, t \le T \ ,$$

which implies that

$$\begin{aligned} \mathbb{E}[S_T^J] &\le \mathbb{E}[\sup_{t\in[0,T]} |J_t - \omega\alpha t|] + \omega\alpha T \\ &\le \frac{e}{e-1}\left(1 + e^{-1} + \mathbb{E}[(J_T - \omega\alpha T)^2]\right) + \omega\alpha T \\ &= \frac{e}{e-1}\left(1 + e^{-1} + 2\omega\alpha^2 T\right) + \omega\alpha T < \infty \ . \end{aligned}$$

This concludes the proof.

4.C Proof of Theorem 2

Observe first that for every θ and every t, one gets

$$\int_0^t x_u^\theta du = \int_0^t \theta(u) y_{\theta(u)} du = \int_0^{\Theta(t)} y_u du \ .$$

Hence, the expectation of their negative exponentials agree as well:

$$P^{x^\theta}(t) = P^y(\Theta(t)) \ .$$

The specific clock rate $\theta^\star$ given by the calibration equation thus satisfies, for all t,

$$P^{market}(t) = P^{x^{\theta^\star}}(t) = P^y(\Theta^\star(t)) \ . \tag{4.27}$$

Turning this equality in terms of instantaneous forward rates yields

$$\int_0^t f^{market}(u)du = \int_0^{\Theta^\star(t)} f^y(u)du \ .$$

Eq. (4.13) is just the differential form of the latter.

It is not clear, in general, to determine when this ODE admits a solution. However, a simple case is when P^y is a strictly decreasing discount curve. In this case indeed, P^y admits an inverse on the positive half line, noted Q^y. Apply Q^y to (4.27) yields $\Theta^\star = Q^y(P^{market}(t))$. Furthermore, the inverse of a decreasing function is decreasing, and the combination of two decreasing functions is itself increasing. Hence, if P^{market} is decreasing, $\Theta^\star(t)$ is continuous and strictly increasing. Moreover, $\Theta^\star(0) = Q^y\left(P^{market}(0)\right) = Q^y(1) = 0$. Hence, $\Theta^\star$ exists, and is a clock.

4.D Proof of Theorem 3

It is known from Corollary 2 that for any (non-trivial) (J)CIR process y with parameter Ξ, there exits a clock $\Theta^\star(t) = \Theta^\star(t;\Xi)$ that yields a perfect fit between the curves P^x generated y the TC-JCIR $x_t^{\theta^\star} := \theta^\star(t) y_{\Theta^\star(t)}$. Same holds true for the JCIR++, $x_t^{\varphi^\star}$. This means:

$$P^{x^{\varphi^\star}}(t;\Xi) = P^{market}(t) = P^{x^{\theta^\star}}(t;\Xi) \ ,$$

or equivalently,

$$e^{-\int_0^t \varphi^\star(u)du} P^y(t;\Xi) = P^{market}(t) = P^y(\Theta^\star(t);\Xi) \ .$$

Because y is a JCIR, it can be arbitrarilly close to 0 at any time t, hence the calibration constraint amounts to force $\varphi^\star(t) \geq 0$ (or equivalently, $f^{\mathrm{JCIR}}(t) \leq f^{\mathrm{market}}(t)$) $\forall\, t \geq 0$. This implies that $P^y(\Theta^\star(t); \Xi) \leq P^y(t; \Xi)$. Because $P^y(.; \Xi)$ is a decreasing function, the last inequality is equivalent to $\Theta^\star(t) \geq t$.
To prove 1), we start from the increasingness of $v^Y(t)$ (Lemma 12). Hence, $v^Y(\Theta^\star(t)) \geq v^Y(t)$ since $\Theta^\star(t) \geq t$.

From (4.23), we have, after some computations,

$$\begin{aligned}
\frac{d}{dt} f^{\mathrm{JCIR}}(t) &= \frac{4\gamma^2 e^{t\gamma}\left[\kappa\varrho(\gamma-\kappa+(\kappa+\gamma)e^{t\gamma}) + y_0\gamma(\gamma-\kappa-(\kappa+\gamma)e^{t\gamma})\right]}{[2\gamma+(\kappa+\gamma)(e^{t\gamma}-1)]^3} \\
&\quad + \frac{\omega\alpha(2\gamma+(\kappa+\gamma)(e^{t\gamma}-1))}{[2\gamma+(\kappa+\gamma)(e^{t\gamma}-1)]^3} \\
&= \frac{4\gamma^2 e^{t\gamma}\left[(\gamma-\kappa)(\kappa\varrho+y_0\gamma+\omega\alpha)-2\omega\alpha^2\right]}{[2\gamma+(\kappa+\gamma)(e^{t\gamma}-1)]^3} \\
&\quad + \frac{4\gamma^2 e^{2t\gamma}\left[(\gamma+\kappa)(\kappa\varrho-y_0\gamma+\omega\alpha)+2\omega\alpha^2\right]}{[2\gamma+(\kappa+\gamma)(e^{t\gamma}-1)]^3}\,.
\end{aligned}$$

From this expression, one can check that f^{JCIR} is strictly increasing if $y_0 < \varrho + \omega\alpha/\kappa$ and $y_0\gamma \leq \kappa\varrho + \omega\alpha$. It is strictly decreasing if $y_0 \geq \varrho + \omega\alpha/\kappa$. Otherwise, i.e., if $y_0 < \varrho + \omega\alpha/\kappa$ and $y_0\gamma > \kappa\varrho + \omega\alpha$, the derivative has a root at

$$t_2 := \frac{1}{\gamma} \ln \frac{(\gamma-\kappa)(\kappa\varrho+y_0\gamma+\omega\alpha)-2\omega\alpha^2}{(\kappa+\gamma)(y_0\gamma-\kappa\varrho-\omega\alpha)-2\omega\alpha^2}\,,$$

i.e., f^{JCIR} is first increasing, then decreasing.
The constraint $\varphi^\star(t) \geq 0$ for all t, simply means that $f^{market}(t) \geq f^{\mathrm{JCIR}}(t)$ and so $\theta^\star(t) \geq \frac{f^{\mathrm{JCIR}}(t)}{f^{\mathrm{JCIR}}(\Theta^\star(t))}$. Observe that the condition $y_0 = \varrho + \omega\alpha/\kappa$ corresponds to the case where $v^y(t)$ is increasing and $f^{\mathrm{JCIR}}(t)$ is decreasing, hence (i) holds.
If f^{market} is constant, we have that $f^{market}(t) \geq f^{\mathrm{JCIR}}(t)$ which implies that $f^{market}(t) \geq f^{\mathrm{JCIR}}(\Theta^\star(t))$. Clearly, if $f^{market}(t)$ is constant or f^{JCIR} is decreasing, then $\theta^\star(t) \geq 1$. And if $v^y(t)$ is increasing, $v^y(\Theta^\star(t)) \geq v^y(t)$ since $\Theta^\star(t) \geq t$. From the fact that $\mathbb{V}[x_t^{\theta^\star}] := \theta^\star(t)^2 v^y(\Theta^\star(t))$ and the variation of $v^y(t)$ (Lemma 12), (ii), (iii) and (iv) follow.
In particular, taking $\omega\alpha = 0$, we recover the CIR case which corresponds to the TC-CIR model.

Chapter 5

Conditional survival probabilities under partial information: a recursive quantization approach with applications

5.1 Introduction

In the recent decades, the market for credit derivatives has grown and the focus on credit risk has increased among academics and practitioners. This has been reinforced by the 2008 post-crisis regulation obliging financial institutions to perform a variety of regulatory capital requirements including credit risk management. Consequently, it has been generated a demand for successfully estimating the credit risk of a firm such as defaultable corporate securities. In that respect, the structural approach in credit risk is a very important direction of research. It is originated to the seminal work of Merton, 1974 and uses the dynamics of structural variables of a firm, such as asset and debt to model the time of default under the assumption that default can only happen at maturity. Then, debt and credit spreads can be simply derived using the classical option pricing theory. However, Merton's model does not allow for premature default in the sense that the default may only occurs at the maturity of the claim. The first passage time models were then introduced, relaxing this restrictive and unrealistic feature. Among them, we have the well known Black and Cox, 1976 model which adds a time-dependent barrier and further takes into account the so-called safety covenants allowing the bondholders with the right to force the firm to bankruptcy if it is doing poorly. Despite the model's positive feature, the Black and Cox model has few parameters and is not easily calibrated to structural data such as CDS quotes along different maturities. To that end, extensions of the same models called AT1P and SBTV were introduced in Brigo and Tarenghi, 2004 and Brigo and Tarenghi, 2005 allowing exact calibration to credit spreads using efficient closed-form formulas for default probabilities. Hence the market value of the firm is a key input of these structural models.

In practice however, it is difficult for investors in secondary markets to assess the value of the firm's assets, a phenomenon commonly referred to as incomplete information. In this case, modeling the firm value in a Black-Scholes framework is problematic since the model assumes that the firm's underlying asset is traded. Moreover, in such a framework, the default time is a predictable default time leading to low credit spreads for short maturities. For these reasons, Duffie and Lando, 2001 propose a model where the investors have only partial information on the firm value and observe at discrete time intervals a noisy accounting report. In this model, the default time is totally inaccessible in the market filtration and admits an intensity that is proportional to the derivative of the conditional density of the unobserved firm value at the default barrier. As a result, the corresponding short term spreads are always higher compared to the complete information short term spreads. Furthermore, Duffie and Lando provide an interesting link between structural and reduced-form models. Alternatively, some extensions of this model based on noisy information in continuous time can be found among others in Coculescu et al., 2008 or recently in Frey et al., 2019. In particular, in Coculescu et al., 2008, the conditional default probabilities are computed under the assumption that the fundamental process which triggers the default is a continuous and invertible function of a Gaussian martingale.

In this chapter, we both generalize and improve results derived in earlier studies. The models presented in Callegaro and Sagna, 2013 and Profeta and Sagna, 2014 can deal with arbitrary firm-value diffusions, but are heavy. Moreover, the considered information flow is only made of a noisy version of the firm-value. This is not realistic as in practice, investors can obviously observe the default state of the firm. This larger filtration is considered in Coculescu et al., 2008, but under the restrictive assumption that the firm-value process is a continuous and invertible function of a Gaussian martingale. In this work, we consider the same information set as Coculescu et al., 2008 but relax the restriction regarding the firm-value dynamics. To deal with this general case, we propose a numerical scheme based on fast quantization recently introduced in Pagès and Sagna, 2015. This technique is faster compared to Callegaro and Sagna, 2013 and Profeta and Sagna, 2014 as there is no need to rely on Monte-Carlo simulations to compute the conditional survival probabilities. A detailed analysis of the error induced by the approximation is provided. Eventually, we illustrated our method on the pricing of CDS option credit derivatives.

The reminder of the chapter is organized as follows. In Section 5.2, we introduce the model. Different information flows and corresponding survival probabilities will be discussed. Section 5.3 presents the estimations of the survival probabilities using the recursive quantization method and the stochastic filtering theory. We then give a brief introduction to the recursive quantization method. Section 6.3 is devoted to the results of the numerical experiments.

5.2 The model

Let us assume the probability space introduced in Chapter 2, modelling the uncertainty of our economy. We consider a structural default model, and represent the default time of a reference entity as the first passage time of firm value process below a default barrier. More precisely, the default time τ takes the form (2.10), where the stochastic process X represents the actual value of a firm and $\mathbf{a} > 0$ stands for the default barrier.

We consider a partial information model where the true value of the firm X is not observable and we only observe Y, which is correlated with X. We suppose a general diffusion framework where the dynamics of X and Y are governed by the following stochastic differential equations (SDEs) :

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, & X_0 = x_0, \\ dY_t = h(t, Y_t, X_t)dt + \nu(t, Y_t)dW_t + \delta(t, Y_t)d\widetilde{W}_t, & Y_0 = y_0, \end{cases} \tag{5.1}$$

where $(W, \widetilde{W})$ is a standard two-dimensional Brownian motion under the probability measure $\mathbb{Q}$ and $0 \le t \le T$, T is a finite time horizon. We suppose that the functions $b, \sigma, \nu, \delta : [0, +\infty) \times \mathbb{R} \to \mathbb{R}$ are Lipschitz in x uniformly in t. Moreover we assume that h is locally bounded and Lipschitz in (y, x), uniformly in t and that $\nu(t, y) \ge 0$ and $\sigma(t, y) \ge 0$ for every $(t, y) \in [0, +\infty) \times \mathbb{R}$. These conditions ensure that the above SDEs admit a unique strong solution.

5.2.1 Information flows

One of the major critiques of the Black and Cox models is that in practice, the firm value is not observable. It is therefore not realistic to consider that the information available to the investor is $\mathbb{F}^X := (\mathcal{F}_t^X)_{t\geq 0}$, $\mathcal{F}_t^X := \sigma(X_u, 0 \leq u \leq t)$. A more realistic framework has been proposed in Callegaro and Sagna, 2013 where the investor information is given by the natural filtration $\mathbb{F}^Y$ of a noisy version Y of the process X. However, one might argue that this way of modelling the information is not realistic either since given $\mathcal{F}_t^Y$, the investor is unable to know whether the reference entity defaulted or not by time t. In other words, the default indicator process $D_t := \mathbb{1}_{\{t\geq\tau\}}, t \geq 0$, is not adapted to $\mathbb{F}^Y$.

In this chapter, we address this point by considering a more realistic information flow, defined as the progressive enlargement of $\mathbb{F}^Y$ with $\mathbb{D}$, the natural filtration of the default indicator process.

In other words, we have two investors' information flows. On the one hand, the information available to the common investor $\mathbb{G}$, is defined as the progressive enlargement of the natural filtration of the default indicator process with that of the noisy firm-value:

$$\mathcal{G}_t := \mathcal{F}_t^Y \vee \mathcal{D}_t, \quad t \geq 0$$

In this setup, the following relationships hold:

$$\mathbb{D} \subsetneq \mathbb{F}^X = \mathbb{F}^W \qquad \text{and} \qquad \mathbb{F}^Y \subsetneq \mathbb{G} .$$

where $\mathbb{F}^W$ is the natural filtration of the Brownian motion driving X.

On the other hand, the natural filtration of the actual (i.e., noise-free) firm-value process, $\mathbb{F}^X$, could be seen as the information available to insiders. Economically, X would represent the value of the firm, which is unobservable to the common investors, while Y might be the market price of an asset issued by the firm, accessible to all market participants, and $\mathbf{a}$ would stand for the solvency capital requirement imposed by regulators.

Remark 10. *The results in this chapter can be straightforwardly extended in the case when the default barrier* $\mathbf{a}$ *is a piecewise constant function of time* $\mathbf{a} : [0,\infty) \to [0,\infty)$*, with* $0 < \mathbf{a}(0) < x_0$*. In all the other situations, more technical results have to be applied and we are not going to focus on these aspects. Nevertheless, let us mention that in, e.g., Pötzelberger and Wang, 2001 crossing probabilities for the Brownian motion are obtained in the case when the (double) barrier is a piecewise* *linear function on $[0,T]$ and approximations for crossing probabilities are obtained for general nonlinear bounds.*

5.2.2 Survival probability

We are here interested in computing the survival probability of the firm up to time t conditional upon $\mathcal{G}_s$, $s \le t \le T$:

$$\mathbb{Q}\left(\tau > t|\mathcal{G}_s\right) = \mathbb{E}\left(\mathbb{1}_{\{\tau>t\}}\middle|\mathcal{G}_s\right) \tag{5.2}$$

Note that this probability collapse to zero whenever $\{\tau \le s\}$.
Using the Markov property of X and the fact that the two Brownian motions W and $\widetilde{W}$ are independent, we have the following lemma (see Profeta and Sagna, 2014).

Lemma 14.

$$\mathbb{Q}\left(\tau > t|\mathcal{F}_s^Y\right) = \mathbb{E}\left[\mathbb{1}_{\{\tau>s\}}F(s,t,X_s)|\mathcal{F}_s^Y\right]. \tag{5.3}$$

where, for every $x \in \mathbb{R}$,

$$F(s,t,x) := \mathbb{Q}\left(\inf_{s\le u\le t} X_u > \mathbf{a}\middle|X_s = x\right) \tag{5.4}$$

is the conditional survival probability under full information.

Applying the chain rule of the conditional expectation, we show the following result.

Proposition 1. *We have, for $s \le t$,*

$$\mathbb{Q}\left(\tau > t\middle|\mathcal{G}_s\right) = \mathbb{1}_{\{\tau>s\}}\frac{\mathbb{E}\left[\mathbb{1}_{\{\tau>s\}}F(s,t,X_s)|\mathcal{F}_s^Y\right]}{\mathbb{Q}\left(\tau > s\middle|\mathcal{F}_s^Y\right)} \tag{5.5}$$

Furthermore, it holds on the set $\{\tau > s\}$ that

$$\mathbb{Q}\Big(\tau > t|\mathcal{F}_s^Y\Big) \le \mathbb{Q}\Big(\tau > t|\mathcal{G}_s\Big). \tag{5.6}$$

Proof. Using the *Key lemma* (see (2.14)), we have

$$\mathbb{Q}\left(\tau > t|\mathcal{G}_s\right) = \mathbb{1}_{\{\tau>s\}}\frac{\mathbb{Q}\left(\tau > t|\mathcal{F}_s^Y\right)}{\mathbb{Q}\left(\tau > s|\mathcal{F}_s^Y\right)}. \tag{5.7}$$

The proof is completed by using Lemme 14. The particular case follows from (5.7) since on the event $\{\tau > s\}$,

$$\mathbb{Q}\Big(\tau > t|\mathcal{F}_s^Y\Big) = \mathbb{Q}(\tau > t|\mathcal{G}_s)\,\mathbb{Q}\Big(\tau > s|\mathcal{F}_s^Y\Big) \leq \mathbb{Q}(\tau > t|\mathcal{G}_s)$$

or, equivalently, $\mathbb{Q}\Big(\tau \leq t|\mathcal{G}_s\Big) \leq \mathbb{Q}\Big(\tau \leq t|\mathcal{F}_s^Y\Big)$.

□

Remark 11. *The inequality* (5.6)*, confirmed by the numerical experiments in Section 6.3, means that the less we have information on the state of the system* ($\mathcal{F}_s^Y \subset \mathcal{F}_s^Y \vee \mathcal{D}_s = \mathcal{G}_s$)*, the higher the default probability. This also shows the difference with Profeta and Sagna, 2014 where the quantity of interest is just* $\mathbb{Q}\Big(\tau > t|\mathcal{F}_s^Y\Big)$.

5.2.3 The problem

We are interested in computing the survival probability of the reference entity up to time t conditional upon the investor's information up to time s, $\mathbb{Q}(\tau \geq t|\mathcal{G}_s)$, $0 \leq s < t$. In order to even more comply with real market practice, we further consider that we can only access to, say, *discrete time* observations of Y up to time s. To that end, let us start by fixing a time discretization grid over $[0, t]$:

$$0 = t_0 < \cdots < t_m = s < t_{m+1} < \cdots < t_n = t.$$

Our aim is to approximate the right hand side of (5.5) by recursive quantization. In some specific models (those for which (5.1) admits an explicit solution (X, Y), like in the Black-Scholes framework), we will consider the discrete trajectories $(X_{t_k}, Y_{t_k})_{k=0,\ldots,n}$. In more general models, we need to make a discrete time approximation of the quantity of interest. To this end, we suppose that we have access to a trajectory of Y sampled at m times: $(\bar{Y}_{t_0}, \ldots, \bar{Y}_{t_m})$, with $t_0 = 0$ and $t_m = s$ (which in practice will be approximated from the paths of the Euler scheme associated to the stochastic process Y) and will estimate (5.2) by

$$\frac{\mathbb{Q}\left(\tau > t\middle|\mathcal{F}_s^{\bar{Y}}\right)}{\mathbb{Q}\left(\tau > s\middle|\mathcal{F}_s^{\bar{Y}}\right)},$$

on the event $\{\tau > s\}$, where $\mathcal{F}_s^{\bar{Y}} = \sigma(\bar{Y}_{t_k},\ t_k \leq s) = \sigma(\bar{Y}_{t_0}, \ldots, \bar{Y}_{t_m})$.

5.2.4 Discrete time approximation

We denote by $\bar{X}$ the continuous Euler scheme associated to the process X in Equation (5.1), namely:

$$\bar{X}_s = \bar{X}_{\underline{s}} + b(\underline{s}, \bar{X}_{\underline{s}})(s - \underline{s}) + \sigma(\underline{s}, \bar{X}_{\underline{s}})(W_s - W_{\underline{s}}), \quad \bar{X}_0 = x_0,$$

with $\underline{s} = t_k$ if $s \in [t_k, t_{k+1})$, for $k = 0, \ldots, n$. Based on the Euler scheme, we introduce the discretized version of our state-observation processes $(\bar{X}, \bar{Y})$

$$\begin{cases} \bar{X}_{t_{k+1}} = \bar{X}_{t_k} + b(t_k, \bar{X}_{t_k})\Delta_k + \sigma(t_k, \bar{X}_{t_k})(W_{t_{k+1}} - W_{t_k}) \\ \bar{Y}_{t_{k+1}} = \bar{Y}_{t_k} + h(t_k, \bar{Y}_{t_k}, \bar{X}_{t_k})\Delta_k + \nu(t_k, \bar{Y}_{t_k})(W_{t_{k+1}} - W_{t_k}) + \delta(t_k, \bar{Y}_{t_k})(\widetilde{W}_{t_{k+1}} - \widetilde{W}_{t_k}) \end{cases} \tag{5.8}$$

where $k \in \{0, \ldots, n-1\}$ for the signal process and $k \in \{0, \ldots, m-1\}$ for the observation process and where $\Delta_k := t_{k+1} - t_k$.
Supposing that we have access to a discrete trajectory of Y, $(\bar{Y}_{t_0}, \ldots, \bar{Y}_{t_m})$, our first goal is to approximate (recall that $t_m = s$)

$$\frac{\mathbb{Q}\Big(\tau > t |\, \mathcal{F}_s^Y\Big)}{\mathbb{Q}\Big(\tau > s |\, \mathcal{F}_s^Y\Big)} \quad \text{by} \quad \frac{\mathbb{Q}\Big(\tau_{\bar{X}} > t | \mathcal{F}_s^{\bar{Y}}\Big)}{\mathbb{Q}\Big(\tau_{\bar{X}} > s | \mathcal{F}_s^{\bar{Y}}\Big)}, \tag{5.9}$$

where (recall Equation (2.10))

$$\tau_{\bar{X}} := \inf\{u \geq 0, \bar{X}_u \leq \mathbf{a}\}.$$

Using the Brownian Bridge method and the Markov property of $(\bar{X}_{t_k}, \bar{Y}_{t_k})_k$, we show that the quantity (5.9) can be written in a closed formula.

Theorem 4. *We have:*

$$\frac{\mathbb{Q}\Big(\tau_{\bar{X}} > t | \mathcal{F}_s^{\bar{Y}}\Big)}{\mathbb{Q}\Big(\tau_{\bar{X}} > s | \mathcal{F}_s^{\bar{Y}}\Big)} = \Psi(\bar{Y}_{t_0}, \ldots, \bar{Y}_{t_m}), \tag{5.10}$$

where $s = t_m$, $t = t_n$ and for $y = (y_0, \ldots, y_m) \in \mathbb{R}^{m+1}$,

$$\Psi(y) = \frac{\mathbb{E}\Big[\bar{F}(t_m, t_n, \bar{X}_{t_m}) K_{\mathbf{a}}^m L_y^m\Big]}{\mathbb{E}[K_{\mathbf{a}}^m L_y^m]}, \tag{5.11}$$

with

$$K_{\mathbf{a}}^m = \prod_{k=0}^{m-1} G_{\Delta_k \sigma_k^2}^{\bar{X}_{t_k}, \bar{X}_{t_{k+1}}}(\mathbf{a}), \qquad L_y^m = \prod_{k=0}^{m-1} g_k(\bar{X}_{t_k}, y_k; \bar{X}_{t_{k+1}}, y_{k+1})$$

and where for every $x \in \mathbb{R}$,

$$\bar{F}(t_m, t_n, x) = \mathbb{E}\Big[\prod_{k=m}^{n-1} G_{\Delta_k \sigma_k^2}^{\bar{X}_{t_k}, \bar{X}_{t_{k+1}}}(\mathbf{a}) \Big| \bar{X}_{t_m} = x\Big]. \tag{5.12}$$

The function g_k is defined by

$$\begin{aligned} g_k(x_k, y_k; x_{k+1}, y_{k+1}) &= \frac{\mathbb{Q}\Big((\bar{X}_{t_{k+1}}, \bar{Y}_{t_{k+1}}) = (x_{k+1}, y_{k+1}) | (\bar{X}_{t_k}, \bar{Y}_{t_k}) = (x_k, y_k)\Big)}{\mathbb{Q}\Big(\bar{X}_{t_{k+1}} = x_{k+1} | \bar{X}_{t_k} = x_k\Big)} \\ &= \frac{1}{(2\pi\Delta_k)^{1/2}\delta_k} \exp\left(-\frac{\nu_k^2}{2\delta_k^2 \Delta_k}\Big(\frac{x_{k+1} - m_k^1}{\sigma_k} - \frac{y_{k+1} - m_k^2}{\nu_k}\Big)^2\right) \end{aligned} \tag{5.13}$$

with $m_k^1 := x_k + b_k \Delta_k$, $m_k^2 := y_k + h_k \Delta_k$, $\sigma_k = \sigma(t_k, x_k)$, $b_k = b(t_k, x_k)$, $\delta_k = \delta(t_k, y_k)$, $h_k = h(t_k, y_k, x_k)$ and $\nu_k = \nu(t_k, y_k)$. Finally,

$$\begin{aligned} G_{\Delta_k \sigma_k^2}^{x_k, x_{k+1}}(\mathbf{a}) &= \mathbb{Q}\Big(\inf_{u \in [t_k, t_{k+1}]} \bar{X}_u \geq \mathbf{a} | \bar{X}_{t_k} = x_k\Big) \\ &= \left(1 - \exp\left(-\frac{2(x_k - \mathbf{a})(x_{k+1} - \mathbf{a})}{\Delta_k \sigma_k^2}\right)\right) \mathbb{1}_{\{x_k \geq \mathbf{a};\, x_{k+1} \geq \mathbf{a}\}}. \end{aligned} \tag{5.14}$$

Proof. Following Theorem 2.5. in Profeta and Sagna, 2014, we have,

$$\mathbb{Q}\Big(\tau_{\bar{X}} > t | \mathcal{F}_s^{\bar{Y}}\Big) = \frac{\mathbb{E}\Big[\bar{F}(t_m, t_n, \bar{X}_{t_m}) K_{\mathbf{a}}^m L_y^m\Big]}{\mathbb{E}[L_y^m]}$$

and

$$\mathbb{Q}\Big(\tau_{\bar{X}} > s | \mathcal{F}_s^{\bar{Y}}\Big) = \frac{\mathbb{E}\Big[K_{\mathbf{a}}^m L_y^m\Big]}{\mathbb{E}[L_y^m]}.$$

The result follows. □

The question of interest is now to know how to estimate efficiently $\Psi(y)$ for $y = (\bar{Y}_{t_0}, \ldots, \bar{Y}_{t_m})$. Owing to the form of the random vector $K_{\mathbf{a}}^m$, we may put it together with L_y^m to be reduced to similar formula as the filter estimate in a standard nonlinear filtering problem. In other word we may write for $y := (y_0, \ldots, y_m)$,

$$\Psi(y) = \frac{\mathbb{E}\Big[\bar{F}(t_m, t_n, \bar{X}_{t_m})\, L_{y,\mathbf{a}}^m\Big]}{\mathbb{E}[L_{y,\mathbf{a}}^m]},$$

where

$$L_{y,\mathbf{a}}^m = \prod_{k=0}^{m-1} g_k(\bar{X}_{t_k}, y_k; \bar{X}_{t_{k+1}}, y_{k+1}) \times G_{\Delta_k \sigma_k^2}^{\bar{X}_{t_k}, \bar{X}_{t_{k+1}}}(\mathbf{a}).$$

5.3 Approximation by recursive quantization

Notice that defining the operator $\pi_{y,m}$, for every bounded measurable function f, by

$$\pi_{y,m} f := \mathbb{E}\Big[f(\bar{X}_{t_m}) L^m_{y,\mathbf{a}}\Big],$$

we have

$$\Psi(y) = \frac{\pi_{y,m}\bar{F}(t_m, t_n, \cdot)}{\pi_{y,m}\mathbf{1}} =: \Pi_{y,m}\bar{F}(t_m, t_n, \cdot), \tag{5.15}$$

where $\mathbf{1}(x) = 1$, for every real x. Then it is enough to tell how to compute the numerator

$$\pi_{y,m}\bar{F}(t_m, t_n, \cdot) = \mathbb{E}\Big[\bar{F}(t_m, t_n, \bar{X}_{t_m}) \prod_{k=0}^{m-1} g^{\mathbf{a}}_k(\bar{X}_{t_k}, y_k; \bar{X}_{t_{k+1}}, y_{k+1})\Big]$$

where

$$g^{\mathbf{a}}_k(\bar{X}_{t_k}, y_k; \bar{X}_{t_{k+1}}, y_{k+1}) = g_k(\bar{X}_{t_k}, y_k; \bar{X}_{t_{k+1}}, y_{k+1}) \times G^{\bar{X}_{t_k}, \bar{X}_{t_{k+1}}}_{\Delta_k \sigma_k^2}(\mathbf{a}).$$

At this stage, several methods involving Monte Carlo simulations as the particle method can be used to approximate $\Pi_{y,m}$. Optimal quantization is an alternative and some times as a substitute to the Monte Carlo method to approximate such a quantity (we refer to Sellami, 2008 for a comparison of particle like methods and optimal quantization methods).
To use the optimal quantization methods we have to quantize the marginals of the process $(\bar{X}_{t_k})_k$, means, to represent every marginal $\bar{X}_{t_k}$, $k = 0, \ldots, n$, by a discrete random variable $\widehat{X}^{\Gamma_k}_{t_k}$ (we will simply denote it $\widehat{X}_{t_k}$ when there is no ambiguity) taking N_k values $\Gamma_k = \{x^k_1, \ldots, x^k_{N_k}\}$. As we will see later, we have also need in our context to compute the transition probabilities $\hat{p}^{ij}_k = \mathbb{Q}(\widehat{X}_{t_k} = x^k_j | \widehat{X}_{t_{k-1}} = x^{k-1}_i)$, for $i = 1, \ldots, N_{k-1}$; $j = 1, \ldots, N_k$. To this end, we may use stochastic algorithms to get the optimal grids and the associated transition probabilities (see, e.g., Pagès and Pham, 2005; Sellami, 2008). This method works well but may be very time consuming. The so-called marginal functional quantization method (see Callegaro and Sagna, 2013; Profeta and Sagna, 2014; Sagna, 2012) is used as an alternative to the previous method. It consists to construct the marginal quantizations by considering the ordinary differential equation (ODE) resulting to the substitution of the Brownian motion appearing in the dynamics of X in (5.1) by a quadratic quantization of the Brownian motion (see Luschgy and Pagès, 2006). This procedure performs the marginal quantizations quite instantaneous and works well enough from the numerical point of view even if the rate of convergence

(which has not been computed yet from the theoretical point of view) seems to be poor. As an alternative to the two previous methods, we propose the recursive marginal quantization (also called fast quantization) method introduced in Pagès and Sagna, 2015. It consists of quantizing the process $(\bar{X}_{t_k})_{k=0,\dots,n}$, based on a recursive method involving the conditional distributions $\bar{X}_{t_{k+1}}|\bar{X}_{t_k}$, $k = 0, \dots, n-1$. For the problem of interest, this last method is more performing than the previous ones due to its computation speed and to its robustness.

On the other hand, the function $\bar{F}$ has been estimated by Monte Carlo method in Callegaro and Sagna, 2013; Profeta and Sagna, 2014. For competitiveness reasons of the recursive quantization w.r.t. the previously raised methods, we propose here to approximate both quantities Π and $\bar{F}$ by the recursive quantization method.

5.3.1 Approximation of $\Pi_{y,m}$ by recursive quantization

Given that the denominator in the right hand side of (5.15) has a similar form as the numerator, we will only show how to compute the numerator. We remark that $\pi_{y,m}$ can be computed from the following recursive formula:

$$\pi_{y,k} = \pi_{y,k-1} H_{y,k}, \qquad k = 1, \dots, m, \tag{5.16}$$

where, for every $k = 1, \dots, m$, and for every bounded and measurable function f, the transition kernel $H_{y,k}$ is defined by

$$H_{y,k} f(z) = \mathbb{E}\left[f(\bar{X}_{t_k}) g^{\mathbf{a}}_{k-1}(\bar{X}_{t_{k-1}}, y_{k-1}; \bar{X}_{t_k}, y_k) | \bar{X}_{t_{k-1}} = z \right]$$

with

$$H_{y,0} f := \mathbb{E}[f(\bar{X}_0)].$$

In fact, for any bounded Borel function f we have

$$\begin{aligned}
\pi_{y,k} f &= \mathbb{E}\Big[f(\bar{X}_{t_k}) \prod_{\ell=0}^{k-1} g^{\mathbf{a}}_{\ell}(\bar{X}_{t_\ell}, y_\ell; \bar{X}_{\ell+1}, y_{\ell+1}) \Big] \\
&= \mathbb{E}\Big[\mathbb{E}\Big(f(\bar{X}_{t_k}) \prod_{\ell=0}^{k-1} g^{\mathbf{a}}_{\ell}(\bar{X}_{t_\ell}, y_\ell; \bar{X}_{\ell+1}, y_{\ell+1}) \Big| \mathcal{F}^{\bar{X}}_{k-1} \Big) \Big].
\end{aligned}$$

Since $\bar{X}$ is a Markov process we deduce that

$$\begin{aligned}\pi_{y,k}f &= \mathbb{E}\Big[\mathbb{E}\Big(f(\bar{X}_{t_k})g^{\mathbf{a}}_{k-1}(\bar{X}_{t_{k-1}},y_{k-1};\bar{X}_k,y_k)\Big|\mathcal{F}^{\bar{X}}_{k-1}\Big)\prod_{\ell=0}^{k-2} g^{\mathbf{a}}_{\ell}(\bar{X}_{t_\ell},y_\ell;\bar{X}_{\ell+1},y_{\ell+1})\Big]\\ &= \mathbb{E}\Big[H_{y,k}f(\bar{X}_{t_{k-1}})\prod_{\ell=0}^{k-2} g^{\mathbf{a}}_{\ell}(\bar{X}_{t_\ell},y_\ell;\bar{X}_{\ell+1},y_{\ell+1})\Big]\\ &= \pi_{y,k-1}H_{y,k}f.\end{aligned}$$

Then, when we have access to the quantization of the marginals of the process $\bar{X}$, the functional $\pi_{y,k}$ can be approximated recursively by optimal quantization as $\hat{\pi}_{y,k} = \hat{\pi}_{y,k-1}\widehat{H}_{y,k}$ where for every $k \geq 1$, $\widehat{H}_{y,k}$ is a matrix $N_k \times N_{k-1}$ which components $\widehat{H}^{i,j}_{y,k}$ read

$$\widehat{H}^{ij}_{y,k} = g^{\mathbf{a}}_{k-1}(x^i_{k-1},y_{k-1};x^j_k,y_k)\,\hat{p}^{ij}_k\,\delta_{x^j_k}$$

where

$$p^{ij}_k = \mathbb{Q}(\widehat{X}_{t_k} = x^j_k|\widehat{X}_{t_{k-1}} = x^i_{k-1})$$

and $(\widehat{X}_{t_k})_k$ is the quantization of the process $(\bar{X}_t)_{t\geq 0}$ over the time steps $t_k, k = 1,\dots,m$: on the grids $\Gamma_k = \{x^1_k,\dots,x^{N_k}_k\}$, of sizes N_k.
As a consequence, the quantity of interest $\Pi_{y,m}\bar{F}(t_m,t_n,\cdot)$ is estimated by

$$\widehat{\Pi}_{y,m}\bar{F}(t_m,t_n,\cdot) = \sum_{i=1}^{N_m}\widehat{\Pi}^i_{y,m}\bar{F}(t_m,t_n,x^i_m). \tag{5.17}$$

where

$$\widehat{\Pi}^i_{y,m} := \frac{\widehat{\pi}^i_{y,m}}{\sum_{j=1}^{N_m}\widehat{\pi}^i_{y,m}}, \quad i = 1,\dots,N_m$$

and where $\widehat{\pi}_{y,m}$ is the estimation (by optimal quantization) of $\pi_{y,m}$ defined recursively by

$$\begin{cases}\widehat{\pi}_{y,0} = \widehat{H}_{y,0}\\ \widehat{\pi}_{y,k} = \widehat{\pi}_{y,k-1}\widehat{H}_{y,k} := \Big[\sum_{i=1}^{N_{k-1}}\widehat{H}^{i,j}_{y,k}\widehat{\pi}^i_{y,k-1}\Big]_{j=1,\dots,N_k}, \quad k = 1,\dots,m\end{cases} \tag{5.18}$$

with

$$\widehat{H}^{ij}_{y,k} = g^{\mathbf{a}}_{k-1}(x^i_{k-1},y_{k-1};x^j_k,y_k)\,\hat{p}^{ij}_k\,\delta_{x^j_k}. \tag{5.19}$$

Our aim is now to use the (marginal) recursive quantization method to estimate the $F(t_m,t_n,x^i_m)'$s.

5.3.2 Approximation of $\bar{F}(t_m, t_n, \cdot)$ by recursive quantization

Recall that for every x,

$$\bar{F}(t_m, t_n, x) = \mathbb{E}\Big(\prod_{k=m}^{n-1} G_{\Delta_k \sigma_k^2}^{\bar{X}_{t_k}, \bar{X}_{t_{k+1}}}(\mathbf{a}) \Big| \bar{X}_{t_m} = x\Big).$$

As previously, we remark that if we define the functional $\pi_{n,m}$ by

$$\Big(\pi_{n,m} f\Big)(x) = \mathbb{E}\Big(f(\bar{X}_{t_n}) \prod_{k=m}^{n-1} G_{\Delta_k \sigma_k^2}^{\bar{X}_{t_k}, \bar{X}_{t_{k+1}}}(\mathbf{a}) \Big| \bar{X}_{t_m} = x\Big),$$

for every bounded and measurable function f, then $F(t_m, t_n, x)$ reads

$$\bar{F}(t_m, t_n, x) = \Big(\pi_{n,m} \mathbf{1}\Big)(x).$$

Now, for every bounded and measurable function f, defining as previously the transition kernel H_k as,

$$\Big(H_k f\Big)(z) = \mathbb{E}\left(f(\bar{X}_{t_k}) G_{\Delta_{k-1} \sigma_{k-1}^2}^{\bar{X}_{t_{k-1}}, \bar{X}_{t_k}}(\mathbf{a}) | \bar{X}_{t_{k-1}} = z\right),$$

for every $k = m+1, \ldots, n$ and setting

$$H_m f = \mathbb{E}\left[f(\bar{X}_{t_m})\right] \tag{5.20}$$

yields for every $k = m+1, \ldots, n$,

$$\begin{aligned}
(\pi_{k,m} f)(x) &= \mathbb{E}\Big(\mathbb{E}\Big(f(\bar{X}_{t_k}) \prod_{i=m}^{k-1} G_{\Delta_k \sigma_k^2}^{\bar{X}_{t_k}, \bar{X}_{t_{k+1}}}(\mathbf{a}) \Big| (\bar{X}_{t_\ell})_{\ell=m,\ldots,k-1}\Big) \Big| \bar{X}_{t_m} = x\Big) \\
&= \mathbb{E}\Big(\mathbb{E}\Big(f(\bar{X}_{t_k}) G_{\Delta_{k-1} \sigma_{k-1}^2}^{\bar{X}_{t_{k-1}}, \bar{X}_{t_k}}(\mathbf{a}) | (\bar{X}_{t_\ell})_{\ell=m,\ldots,k-1}\Big) \prod_{\ell=m}^{k-2} G_{\Delta_\ell \sigma_\ell^2}^{\bar{X}_{t_\ell}, \bar{X}_{t_{\ell+1}}}(\mathbf{a}) | \bar{X}_{t_m} = x\Big) \\
&= \mathbb{E}\Big(\mathbb{E}\Big(f(\bar{X}_{t_k}) G_{\Delta_{k-1} \sigma_{k-1}^2}^{\bar{X}_{t_{k-1}}, \bar{X}_{t_k}}(\mathbf{a}) | \bar{X}_{t_{k-1}}\Big) \prod_{\ell=m}^{k-2} G_{\Delta_\ell \sigma_\ell^2}^{\bar{X}_{t_\ell}, \bar{X}_{t_{\ell+1}}}(\mathbf{a}) | \bar{X}_{t_m} = x\Big) \\
&= (\pi_{k-1,m} H_k f)(x).
\end{aligned}$$

Consequently, if one has access to the recursive quantizations $(\widehat{X}_{t_k})_{k=m,\ldots,n}$ and the transition probabilities $\{\hat{p}_k^{ij}, k = m+1, \ldots, n\}$ of the process $(\bar{X}_{t_k})_{k=m,\ldots,n}$, the quantity $F(t_m, t_n, x)$ will be estimated by

$$\widehat{F}(t_m, t_n, x) = \sum_{j=1}^{N_n} \widehat{\pi}_{n,m}\, \delta_{\{x_m^j = x\}}, \tag{5.21}$$

where the $\widehat{\pi}_{n,m}$'s are defined from the following recursive formula

$$\begin{cases} \widehat{\pi}_{m,m} = \widehat{H}_m \\ \widehat{\pi}_{k,m} = \widehat{\pi}_{k-1,m}\widehat{H}_k := \Big[\sum_{i=1}^{N_{k-1}} \widehat{H}_k^{i,j}\widehat{\pi}_{k-1,m}\Big]_{j=1,\ldots,N_k}, k = m+1,\ldots,n \end{cases} \tag{5.22}$$

with

$$\widehat{H}_k^{ij} = G_{\Delta_{k-1}\sigma_{k-1}^2}^{x_{k-1}^i, x_k^j}(\mathbf{a})\, \hat{p}_k^{ij}\, \delta_{x_k^j}, \quad i = 1,\ldots,N_{k-1}; j = 1,\ldots,N_k.$$

5.3.3 Approximation of $\Pi_{y,m}\bar{F}(t_m, t_n, \cdot)$ by recursive quantization

Combining equations (5.17) and (5.21), the conditional survival probability $\Pi_{y,m}F(t_m, t_n, \cdot)$ will be estimated (for a fixed trajectory $(y_0, \ldots, y_m)$ of the observation process $(Y_{t_0}, \ldots, Y_{t_m})$) by

$$\widehat{\Pi}_{y,m}\widehat{F}(t_m, t_n, \cdot) = \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} \widehat{\Pi}_{y,m}^i \widehat{\pi}_{n,m}\, \delta_{x_m^j}. \tag{5.23}$$

A rigorous analysis of the error resulting from the approximation of $\Pi_{y,m}\bar{F}(t_m, t_n, \cdot) = \Big[\Pi_{y,m}\Big(\pi_{n,m}\mathbf{1}\Big)\Big](\cdot)$ by $\Big[\widehat{\Pi}_{y,m}\Big(\widehat{\pi}_{n,m}\mathbf{1}\Big)\Big](\cdot)$ is detailed in Sagna et al., 2019 and is out of the scope of this thesis. We refer to the former paper for a detailed proof of the error approximation.

Remark 12. *In Section 6.3, the formula* (5.23) *will be compared to the one of interest in Profeta and Sagna, 2014:* $\mathbb{Q}\Big(\tau_{\bar{X}} > t|\mathcal{F}_s^{\bar{Y}}\Big)$*, which reads (following the previous notations)*

$$\mathbb{Q}\Big(\tau_{\bar{X}} > t|\,(\bar{Y}_{t_0}, \ldots, \bar{Y}_{t_m}) = y\Big) = \frac{\mathbb{E}\Big[\bar{F}(t_m, t_n, \bar{X}_{t_m})L_{y,\mathbf{a}}^m\Big]}{\mathbb{E}[L_y^m]}. \tag{5.24}$$

The conditional probability has been approximated in Profeta and Sagna, 2014 via an hybrid Monte Carlo - optimal quantization method. It may be approximated following the procedure we propose using only optimal quantization method as

$$\widehat{\varpi}_{y,m}\widehat{F}(t_m, t_n, \cdot) = \sum_{i=1}^{N_m}\sum_{j=1}^{N_n} \widehat{\varpi}_{y,m}^i \widehat{\pi}_{n,m}\, \delta_{x_m^j} \tag{5.25}$$

where the $\widehat{\varpi}_{y,m}^i$*'s are obtained from* (5.18) *by replacing the function* $g_k^{\mathbf{a}}$ *by* g_k *of equation* (5.13)*.*

Let us say now how to quantize the signal process X from the recursive quantization method.

5.3.4 The recursive quantization method

The optimal vector quantization was originally developed since the 1950's to optimally discretize a continuous (stationary) signal in view of its transmission in the field of signal processing (see Gersho and Gray, 1988). It was introduced by Pagès, 1998 as a quadrature formula for numerical integration, and by Bally et al., 2001 and Bally and Pagès, 2003 for conditional expectation approximations in order to price multi-asset American style options. We refer to Graf and Luschgy, 2000 for a more detailed introduction to optimal vector quantization theory, and to Pagès, 2015 for an introduction more oriented to applications in numerical probability. For a more complete exposition regarding the applications in numerical integration with respect to the distribution of a random vector taking values in $\mathbb{R}^d$, we refer to Chapter 5 in Pagès, 2018.

Let X be a given $\mathbb{R}^d$-valued random defined on $(\Omega, \mathcal{F}, \mathbb{Q})$ with distribution $\mathbb{Q}_X$. The $L^r(\mathbb{Q}_X)$-optimal quantization problem of size N for X (or for the distribution $\mathbb{P}_X$) aims to approximate X by a Borel function of X taking at most N values. If $X \in L^r(\mathbb{Q})$ and defining $\|X\|_r := (\mathbb{E}|X|^r)^{1/r}$ where $|\cdot|$ denotes an arbitrary norm on $\mathbb{R}^d$, this turns out to solve the following optimization problem (see, e.g., Graf and Luschgy, 2000):

$$\begin{aligned} e_{N,r}(X) &= \inf\{\|X - \widehat{X}^{\Gamma}\|_r, \Gamma \subset \mathbb{R}^d,\ \mathrm{card}(\Gamma) \le N\} \\ &= \inf_{\substack{\Gamma \subset \mathbb{R}^d \\ \mathrm{card}(\Gamma) \le N}} \Big(\int_{\mathbb{R}^d} d(x, \Gamma)^r d\mathbb{Q}_X(x) \Big)^{1/r}, \end{aligned} \tag{5.26}$$

where $\widehat{X}^{\Gamma}$, the quantization of X on the subset $\Gamma = \{x_1, \ldots, x_N\} \subset \mathbb{R}^d$ (called a codebook, an N-quantizer or a grid) is defined by

$$\widehat{X}^{\Gamma} = \mathrm{Proj}_{\Gamma}(X) := \sum_{i=1}^{N} x_i \mathbb{1}_{\{X \in C_i(\Gamma)\}}$$

and where $(C_i(\Gamma))_{i=1,\ldots,N}$ is a Borel partition (Voronoi partition) of $\mathbb{R}^d$ satisfying for every $i \in \{1, \ldots, N\}$,

$$C_i(\Gamma) \subset \{x \in \mathbb{R}^d : |x - x_i| = \min_{j=1,\ldots,N} |x - x_j|\}.$$

Keep in mind that for every $N \ge 1$, the infimum in (5.26) is reached at one grid at least. Any N-quantizer realizing this infimum is called an L^r-optimal N-quantizer. Moreover, if $\mathrm{card}(\mathrm{supp}(\mathbb{Q}_X)) \ge N$ then the optimal N-quantizer is of size N (see

Graf and Luschgy, 2000 or Pagès, 1998). On the other hand, the quantization error, $e_{N,r}(X)$, decreases to zero at an $N^{-1/d}$-rate as the grid size N goes to infinity. This convergence rate (known as Zador Theorem) has been investigated in Bucklew and Wise, 1982 and Zador, 1982 for absolutely continuous probability measures under the quadratic norm on $\mathbb{R}^d$. A detailed study of the convergence rate under an arbitrary norm on $\mathbb{R}^d$ and for both absolutely continuous and singular measures may be found in Graf and Luschgy, 2000.

The recursive quantization of the Euler scheme of an $\mathbb{R}^d$-valued diffusion process has been introduced in Pagès and Sagna, 2015. The method allows to speak of fast online quantization and consists on a sequence of quantizations $(\widehat{X}^{\Gamma_k}_{t_k})_{k=0,\dots,n}$ of the Euler scheme $(\bar{X}_{t_k})_{k=0,\dots,n}$ defined recursively as

$$\begin{aligned}\widetilde{X}_0 &= \bar{X}_0,\\ \widehat{X}^{\Gamma_k}_{t_k} &= \mathrm{Proj}_{\Gamma_k}(\widetilde{X}_{t_k}) \quad \text{and} \quad \widetilde{X}_{t_{k+1}} = \mathcal{E}_k(\widehat{X}^{\Gamma_k}_{t_k}, Z_{k+1}),\ k = 0,\dots,n-1, \end{aligned} \tag{5.27}$$

where $(Z_k)_{k=1,\dots,n}$ is an i.i.d. sequence of $\mathcal{N}(0; I_q)$-distributed random vectors, independent from $\bar{X}_0$ and

$$\mathcal{E}_k(y,z) = y + \Delta b(t_k,y) + \sqrt{\Delta}\sigma(t_k,y)z, \quad y \in \mathbb{R}^d,\ z \in \mathbb{R}^q,\ k = 0,\dots,n-1.$$

The sequence of quantizers satisfies for every $k \in \{0,\dots,n\}$,

$$\Gamma_k \in \operatorname{argmin}\{\widetilde{D}_k(\Gamma),\ \Gamma \subset \mathbb{R}^d,\ \mathrm{card}(\Gamma) \le N_k\},$$

where for every grid $\Gamma \subset \mathbb{R}^d$, $\widetilde{D}_{k+1}(\Gamma) := \mathbb{E}\Big[\mathrm{dist}(\widetilde{X}_{t_{k+1}},\Gamma)^2\Big]$.

This recursive quantization method raises some problems among which the computation of the quadratic error bound $\|\bar{X}_{t_k} - \widehat{X}^{\Gamma_k}_{t_k}\|_2 := \Big(\mathbb{E}|\bar{X}_{t_k} - \widehat{X}^{\Gamma_k}_{t_k}|_2\Big)^{1/2}$, for every $k = 0,\dots,n$. It has been shown in Pagès and Sagna, 2015; Pagès and Sagna, 2018 that for any sequences of (quadratic) optimal quantizers Γ_k for $\widetilde{X}^{\Gamma_k}_{t_k}$, for every $k = 0,\dots,n-1$, the quantization error $\|\bar{X}_{t_k} - \widehat{X}^{\Gamma_k}_{t_k}\|_2$ is bounded by the cumulative quantization errors $\|\widetilde{X}_{t_i} - \widehat{X}^{\Gamma_i}_{t_i}\|_2$, for $i = 0,\dots,k$. This result is obtained under the following assumptions and is stated in Proposition 2 below:

1. L^2*-Lipschitz assumption.* The mappings $x \mapsto \mathcal{E}_k(x, Z_{k+1})$ from $\mathbb{R}^d$ to $L^2(\Omega,\mathcal{A},\mathbb{Q})$, $k = 1:n$ are Lipschitz continuous, i.e.,

$$(\mathrm{Lip}) \quad \equiv \quad \forall\, x,\, x' \in \mathbb{R}^d, \quad \Big\|\mathcal{E}_k(x,Z_{k+1}) - \mathcal{E}_k(x',Z_{k+1})\Big\|_2 \le [\mathcal{E}_k]_{\mathrm{Lip}}|x-x'|,\ k = 1,\dots,n.$$

2. L^p*-linear growth assumption.* Let $p \in (2,3]$.

$$(\mathrm{SL})_p \ \equiv\ \forall\, k \in \{1,\dots,n\},\ \forall\, x \in \mathbb{R}^d, \quad \mathbb{E}|\mathcal{E}_k(x,Z_{k+1})|^p \le \alpha_{p,k} + \beta_{p,k}|x|^p.$$

Proposition 2. *Let $\widehat{X} = (\widehat{X}_{t_k})_{k=0:n}$ be defined by* (5.27) *and suppose that all the grids Γ_k are quadratic optimal. Assume that both assumptions* (Lip) *and* $(\mathrm{SL})_p$ *(for some $p \in (2,3]$) hold and that $X_0 \in L^p(\mathbb{Q})$. Then,*

$$\left\|\bar{X}_{t_k} - \widehat{X}_{t_k}\right\|_2 \leq C_{d,p} \sum_{i=0}^{k} [\mathcal{E}_{i+1:k}]_{\mathrm{Lip}} \left[\sum_{\ell=0}^{i} \alpha_{p,\ell} \beta_{p,\ell+1:i}\right]^{\frac{1}{p}} N_i^{-\frac{1}{d}} \tag{5.28}$$

where $C_{d,p} > 0$ and $\alpha_{p,0} = \mathbb{E}\,|X_0|^p = \|X_0\|_p^p$, $\beta_{p,\ell:i} = \prod_{m=\ell}^{i} \beta_{p,m}$ (with $\prod_{\emptyset} = 1$) and

$$[\mathcal{E}_{i:k}]_{\mathrm{Lip}} := \prod_{\ell=i}^{k} [\mathcal{E}_\ell]_{\mathrm{Lip}},\ 1 \leq \ell \leq k \leq n \qquad \textit{and} \qquad [\mathcal{E}_{k+1:k}]_{\mathrm{Lip}} = 1.$$

The associated probability weights and transition probabilities are computed from explicit formulas we recall in the following result.

Proposition 3. *Let Γ_{k+1} be a quadratic optimal quantizer for the marginal random variable $\widetilde{X}_{t_{k+1}}$. Suppose that the quadratic optimal quantizer Γ_k for $\widetilde{X}_{t_k}$ is already computed and that we have access to its associated weights $\mathbb{Q}(\widetilde{X}_{t_k} \in C_i(\Gamma_k))$, $i = 1, \ldots, N_k$. The transition probability $\hat{p}_k^{ij} = \mathbb{Q}(\widetilde{X}_{t_{k+1}} \in C_j(\Gamma_{k+1}) | \widetilde{X}_{t_k} \in C_i(\Gamma_k)) = \mathbb{Q}(\widehat{X}_{t_{k+1}} = x_{k+1}^j | \widehat{X}_{t_k} = x_k^i)$ is given by*

$$\hat{p}_k^{ij} = \Phi\Big(x_{k+1,j+}(x_i^k)\Big) - \Phi\Big(x_{k+1,j-}(x_i^k)\Big), \tag{5.29}$$

where $\Phi(\cdot)$ is given in (2.8),

$$x_{k+1,j-}(x) := \frac{x_{k+1}^{j-1/2} - m_k(x)}{v_k(x)} \qquad \textit{and} \qquad x_{k+1,j+}(x) := \frac{x_{k+1}^{j+1/2} - m_k(x)}{v_k(x)},$$

with $m_k(x) = x + \Delta b(t_k, x)$, $v_k(x) = \sqrt{\Delta}\sigma(t_k, x)$ and, for $k = 0, \ldots, n-1$ and for $j = 1, \ldots, N_{k+1}$,

$$x_{k+1}^{j-1/2} = \frac{x_{k+1}^j + x_{k+1}^{j-1}}{2},\ x_{k+1}^{j+1/2} = \frac{x_{k+1}^j + x_{k+1}^{j+1}}{2}, \quad \textit{with } x_{k+1}^{1/2} = -\infty, x_{k+1}^{N_{k+1}+1/2} = +\infty.$$

Once we have access to the marginal quantizations and to its associated transition probabilities, the right hand side of (5.23) can be computed explicitly.

5.4 Numerical results

In the remaining part of the chapter, we analyze the behavior of our approach from various standpoints by relying on a Black-Scholes type of model. More specifically, the dynamics of the signal process X and observation process Y take the form

$$\begin{cases} dX_t = X_t(rdt + \sigma dW_t), & X_0 = x_0 \ , \\ dY_t = Y_t(rdt + \sigma dW_t + \delta d\widetilde{W}_t), & Y_0 = y_0 \ , \end{cases} \tag{5.30}$$

meaning that

$$\frac{dY_t}{Y_t} = \frac{dX_t}{X_t} + \delta d\widetilde{W}_t \ .$$

In other words, the yield associated with the observation process Y is a noisy version of the yield of the signal process X.

We start by analyzing the model's behavior in terms of conditional survival probabilities in Section 5.4.1. Then, in Section 5.4.2, the impact of the noise parameter δ on the implied volatilities of CDS options is discussed.

5.4.1 Comparison of conditional survival probabilities

First, we compare the function $F(s,t,x) = \mathbb{P}(\inf_{s \leq u \leq t} X_u > \mathbf{a} | X_s = x)$ to its quantized version $\widehat{F}(s,t,x)$ defined in (5.21). This is possible because in the Black-Scholes setup (5.30), the exact expression of $F(s,t,x)$ is known analytically:

$$F(s,t,x) = \Phi(h_1(x, t-s)) - \left(\frac{\mathbf{a}}{x}\right)^{2\sigma^{-2}(r-\sigma^2/2)} \Phi(h_2(x, t-s)) \ , \tag{5.31}$$

where

$$\begin{aligned} h_1(x,u) &= \frac{1}{\sigma\sqrt{u}} \left(\log\left(\frac{x}{\mathbf{a}}\right) + \left(r - \frac{1}{2}\sigma^2\right) u \right) , \\ h_2(x,u) &= \frac{1}{\sigma\sqrt{u}} \left(\log\left(\frac{\mathbf{a}}{x}\right) + \left(r - \frac{1}{2}\sigma^2\right) u \right) \ . \end{aligned}$$

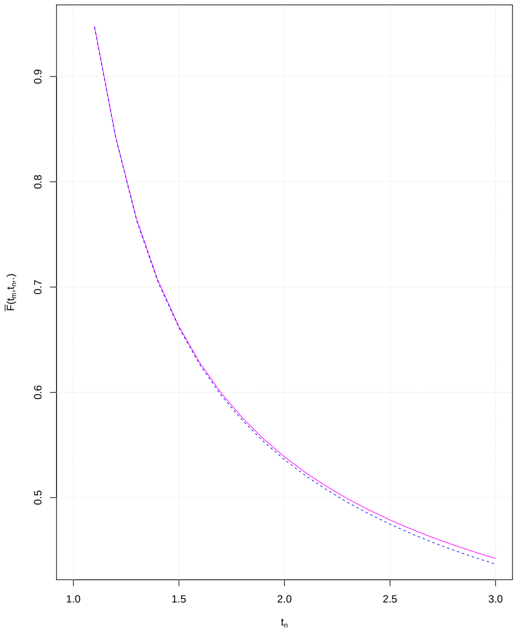
Second, we compare the conditional default probability $\mathbb{P}(\tau_{\bar{X}} \leq t | \mathcal{F}_s^{\bar{Y}} \vee \mathcal{F}_s^{\bar{H}})$ to $\mathbb{P}(\tau_{\bar{X}} \leq \ t \mid \mathcal{F}_s^{\bar{Y}})$. The first conditional default probability is estimated by (5.23). The second one is given by equation (5.25).

The purpose here is to understand what is the specific impact of the information regarding the default state, i.e., $\mathcal{F}_s^{\bar{H}}$ on the conditional survival probability. More explicitly, we compare the conditional probability of $\tau_{\bar{X}} \leq t$ given $\mathcal{F}_s^{\bar{Y}}$ or $\mathcal{F}_s^{\bar{Y}} \vee \mathcal{F}_s^{\bar{H}}$.

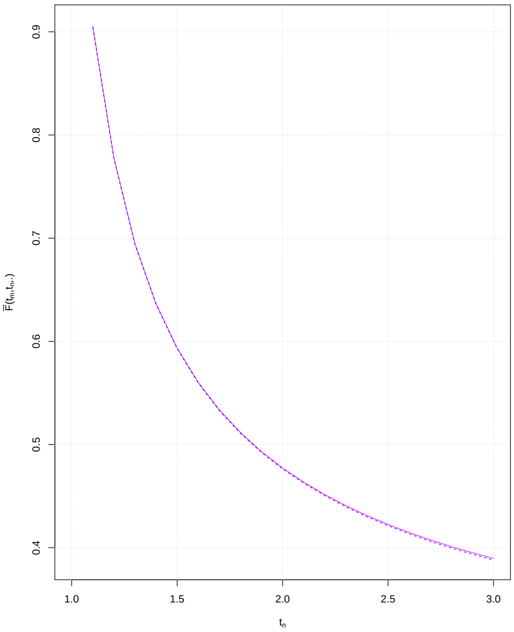
Remark 13. *Notice that the working example* (5.30) *can lead to explicit solutions. In such a simple case, one could perfectly consider a formula similar to* (5.11)*, but relying on the exact solutions on the SDEs of X and Y; the only change will come* *from the function* g_k *which involves the conditional density of* $(X_{t_{k+1}}, Y_{t_{k+1}})$ *given* (X_{t_k}, Y_{t_k})*. However, recall that our purpose here is to provide a general framework, in which the signal and the observation processes can be arbitrary diffusions, with no explicit solutions. Therefore, we do not make use of the specific explicit solutions* (X, Y) *but, instead, restrict ourselves to work with their discretized versions, that is, with* $(\bar{X}, \bar{Y})$*.*

Impact of the quantization

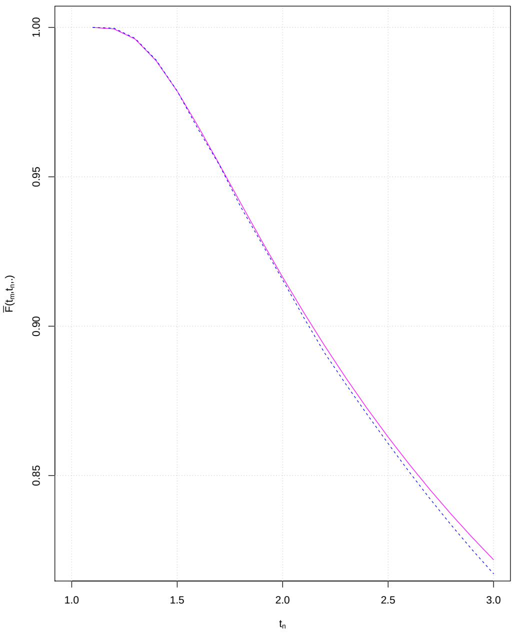
To compare the functions $F(s, t, \cdot)$ and $\widehat{F}(s, t, \cdot)$, we choose the following set of parameters (like those of Profeta and Sagna, 2014): $r = 0.03$, $\sigma = 0.09$, $\delta = 0.5$, $x_0 = y_0 = 86.3$ and $\mathbf{a} = 76$. Figure 5.1 shows the convergence of the quantized function $\widehat{F}(t_m, t_n, x)$ towards the exact one, $F(t_m, t_n, x)$. We set $t_m = 1$, $t_n \in [1.1, 3]$ and x to be one particular point on the grid $\{x_m^i, i = 1, \ldots, N_m\}$ (see (5.21)). Having fixed t_m, $F(t_m, t_n, x)$ depends on both t_n and x. Therefore, to show the convergence, we take a fixed x and plot both $F(t_m, t_n, x)$ and $\widehat{F}(t_m, t_n, x)$ with respect to t_n. The number of discretization points m is set to 50 and the convergence is achieved by increasing the number of quantization points N_n.



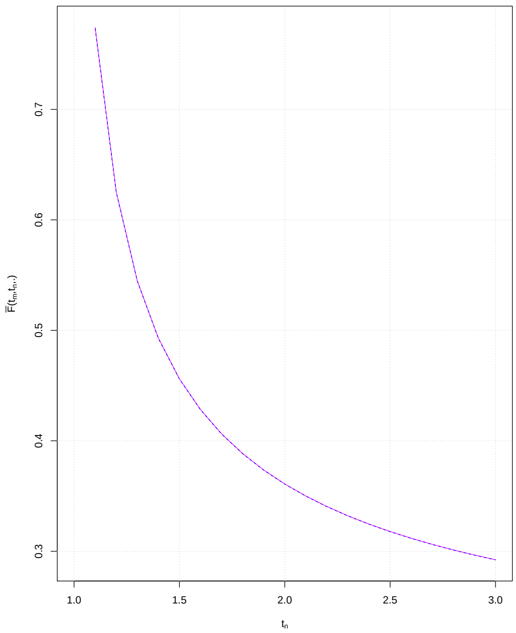
(a) $N_n = 50$



(b) $N_n = 100$



(c) $N_n = 50$

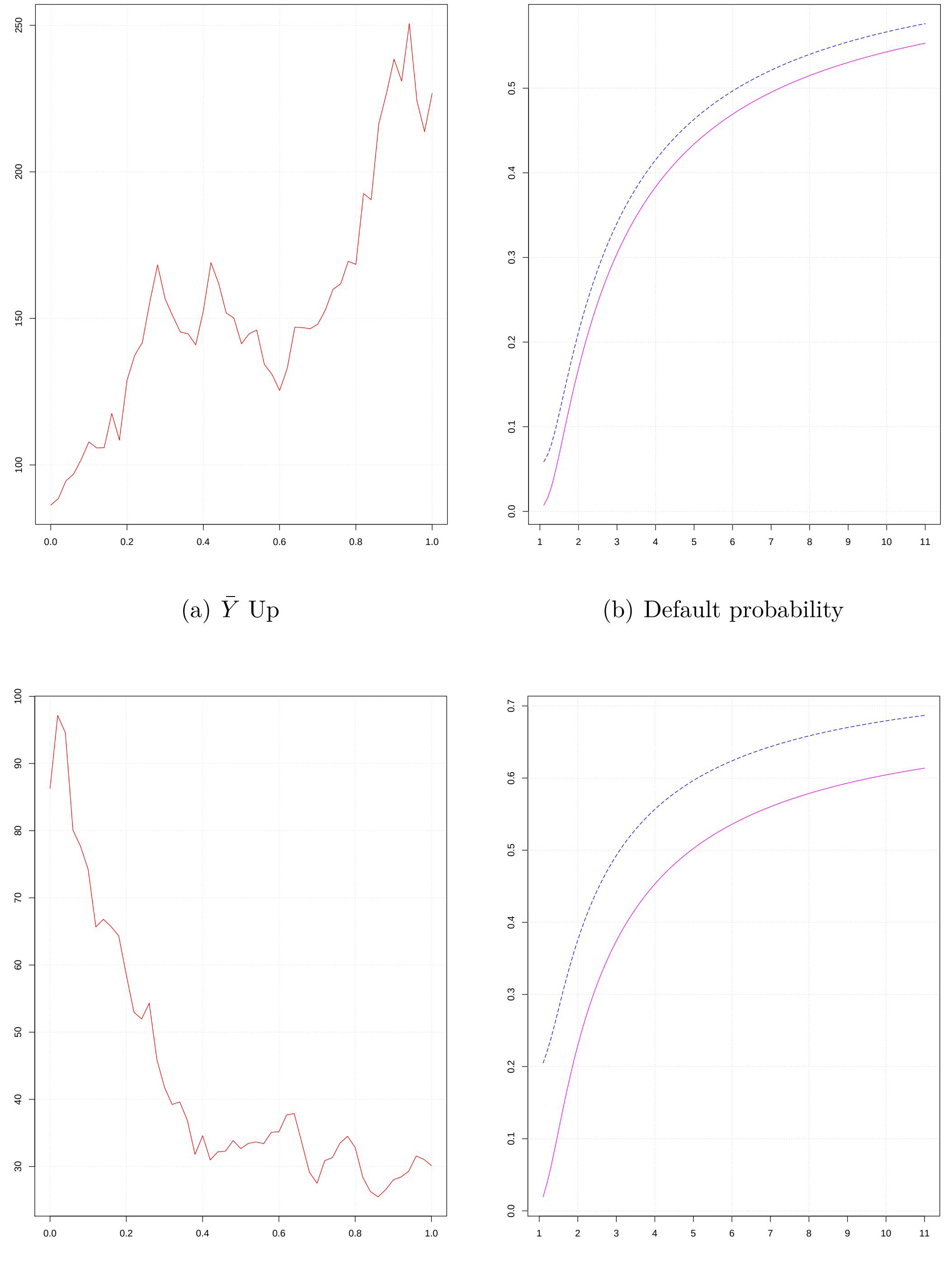


(d) $N_n = 400$

Figure 5.1: Convergence of $\bar{F}(t_m, t_n, \cdot)$ estimated by $\widehat{F}(t_m, t_n, \cdot)$ (dotted blue) toward $F(t_m, t_n, \cdot)$ (solid magenta).

Impact of the default information

In order to check the statements in Remark 11, we proceed with the numerical comparison between the conditional default probabilities $\mathbb{P}(\tau_{\bar{X}} \leq t | \mathcal{F}_s^{\bar{Y}} \vee \mathcal{F}_s^{\bar{H}})$ and $\mathbb{P}(\tau_{\bar{X}} \leq t | \mathcal{F}_s^{\bar{Y}})$, respectively estimated by (5.23) and (5.25). Setting $s = t_m = 1$, $t = t_n$ and considering the same parameter set as in the previous figure but with $t_n \in [1.1, 11]$, Figure 5.2 depicts the trajectories of the observation process $\bar{Y}$ from 0 to t_m and the associated conditional default probabilities as a function of t_n. First, we notice that inequality (5.6) holds since, for the two trajectories of $\bar{Y}$ represented in Fig. 5.2(a) and Fig. 5.2(c), the curve $\mathbb{P}(\tau_{\bar{X}} \leq t | \mathcal{F}_s^{\bar{Y}})$ (dotted blue) dominates the curve $\mathbb{P}(\tau_{\bar{X}} \leq t | \mathcal{F}_s^{\bar{Y}} \vee \mathcal{F}_s^{\bar{H}})$ (solid magenta) in both Fig. 5.2(b) and Fig. 5.2(d). Second, it is clear by comparing these last two figures that the gap between these two curves is larger for a downward trajectory of $\bar{Y}$ than for an upward trend. This can be understood from the fact that the firm is less exposed to default risk in the first case than in the second, and that the actual information about the default state of a firm is less relevant for a very creditworthy firm than for company with a higher default likelihood.



(a) $\bar{Y}$ Up

(b) Default probability

(c) $\bar{Y}$ Down

(d) Default probability

Figure 5.2: Trajectories of the observation process $\bar{Y}$ (solid red), Y up: panel (a), $\bar{Y}$ Down: panel (c) and the associated conditional default probabilities functions $\mathbb{Q}(\tau_{\bar{X}} \leq t | \mathcal{F}_s^{\bar{Y}} \vee \mathcal{F}_s^{\bar{H}})$ (solid magenta) and $\mathbb{Q}(\tau_{\bar{X}} \leq t | \mathcal{F}_s^{\bar{Y}})$ (dotted blue): panels (b) and (d), with $t_m = 1$ and $t_n \in [1.1, 11]$ and number of quantization points $N_n = N_m = 30$.

5.4.2 Application to CDS option pricing

In this section, we derive a CDS option formula via the quantization method before analyzing the CDS options prices and implied volatilities generated by the model. This will allow to give a full pricing formula of credit swaps derivatives in a firm value approach using partial information theory and optimal quantization. In addition, the fact that we add the default filtration in the model indicating whether default has already taken place or not is very important in this case as it is pointless to price a default swap post-default. For more details about CDS pricing methodology and CDS implied volatility analysis, we refer to Section 2.4.1.

The time-0 no-arbitrage price of a call option on a CDS contract starting at time T_a, with maturity T_b and strike (or spread) k is given by (2.34).
The random terms inside the expectation (2.34) mainly the survival processes S_{T_a} and $\overline{Q}_{T_a}(\cdot)$ are ready to be computed using optimal quantization. To do so, using equations (5.23) and (5.25), one only needs to set

$$\overline{Q}_{T_a}(u) = \widehat{\Pi}_{y,a}\widehat{F}(T_a, u, \cdot) \quad \text{and} \quad S_{T_a} = \widehat{\varpi}_{y,a}\mathbf{1} \; . \tag{5.32}$$

Hence the randomness in the expectation (2.34) is only from the observation process Y simulated from time 0 to T_a. This means that we should not need a lot of paths when estimating the expectation (2.34) using Monte Carlo simulation after computing the above mentioned survival processes using optimal quantization. This in turn motivates to fully estimate (2.34) using a hybrid Monte Carlo-optimal quantization procedure.

We now assess the numerical results based on the model's applications to the pricing of CDS option. The model's parameter set is the same as before except here we take $\sigma = 5\%$ and δ is varying. Table 6.1 shows the estimated values of a European payer CDS option and the corresponding Black's volatilities with different strikes and different values of δ. First, we observe that both CDS option prices and the implied volatilities are increasing with with the volatility of the noise δ. This can be explained by the fact that, the higher δ, the noisier the observations are and the higher the default probability. Since δ measures the degree of transparency of the firm, this will have a positive impact on the prices of the CDS option, hence on the corresponding implied volatilities. In contrast, while the option prices are always decreasing with the strike, this is not the case with the implied volatilities except for $\delta = 2\%$ and $\delta = 3\%$. In the case where $\delta = 1\%$, the implied volatility is increasing with respect to the strike. Hence with the help of the parameter δ, one can observe different levels of skewness.

k (bps)	Payer			Implied vol (%)		
	$\delta = 1\%$	$\delta = 2\%$	$\delta = 3\%$	$\delta = 1\%$	$\delta = 2\%$	$\delta = 3\%$
52.9	0.004655	0.007214	0.009276	69.44	133.84	196.80
66.2	0.003739	0.006077	0.008032	73.36	124.16	173.70
79.4	0.003107	0.005298	0.007081	76.85	121.08	161.57

Table 5.1: CDS options and corresponding Black volatilities (with spread $s_0(a,b) =$ 66.20 pbs, $T_a = 1$ and $T_b = 3$) implied by the structural model using Monte Carlo simulation $(1.5 \cdot 10^5)$ paths and for various volatility parameter δ and different strikes (80%, 100%, 120%)$s_0(a,b)$ and σ =5%.

Notice that in this example, we focus more on the numerical performances of the model and do not address the calibration problem. Hence, we use the model implied term structure given by the time-0 model survival probability curve as a CDS term structure. Calibration issues of the model to real market data will be investigated in a future work.

5.5 Conclusions

In this chapter, a new structural model for credit risk has been proposed, generalizing earlier works. Our model deals with an incomplete information, where the default state and a noisy observation of the firm valued are accessible to the investor. It is therefore an extension of Coculescu et al., 2008, as the firm-value triggering the default is no longer restricted to be a continuous and invertible function of a Gaussian martingale, but can be any diffusion.

This more general framework benefits however from a limited analytical tractability. Therefore, we propose a numerical method that relies on nonlinear filtering theory associated with recursive quantization. Compared to earlier works such as Callegaro and Sagna, 2013 or Profeta and Sagna, 2014, our numerical procedure is based on the fast quantization method recently introduced in Pagès and Sagna, 2015, which avoids the use of Monte Carlo simulations. A rigorous analysis of the global error induced to the estimation of the survival processes is performed. We analyze the shapes of the default probabilities which are characterized by a path-dependent feature keeping the memory of all the path of the observed process. Eventually we quantify the impact of the volatility of the noise impacting the firm-value process on the pricing of CDS options and the corresponding implied volatilities using a hybrid Monte Carlo-optimal quantization method.

In future research, we will first investigate the calibration issues of the model which can be tackled by either using observed prices or CDS quotes. In this case, our model can be easily extended to other works dealing with exact calibration to survival probabilities such as including a specific time-dependent barrier (Brigo and Tarenghi, 2004) or using time change techniques (see Chapter 4). Another possible research area is to deal with the price of general default sensitive securities. While we have derived a full quantization scheme to estimate the conditional default probabilities, this was not the case in the pricing of CDS option which required additional Monte Carlo simulations in order to be estimated. To derive a full quantization scheme for the pricing of defaultable claims, a possible route is to exploit the functional quantization method.

Chapter 6

An arbitrage-free conic martingale model with application to credit risk

Term-structure models that is, models that allow to generate a set of curves at future times, starting from a given curve at time 0 are particularly popular in interest rates to model discount or forward curves. In this context, the interest rates model is chosen so as to yield a deterministic discount curve at time 0, $P_0(T)$, say, and various discount curves at any time $t > 0$, $P_t(T)$, $T \geq t$. Note that the curve $P_0(T)$ is deterministic, but those associated to future times, $P_t(T)$, depend on the evolution of the underlying stochastic model up to t. This can be achieved by relying on short-rate models (Vasicek, Hull-White, CIR, etc) or instantaneous forward rates (Heath-Jarrow-Morton (HJM) or market models). Term-structure models are equally useful in credit risk applications, to model the dependency of credit spreads to the maturity. More generally, term-structure models are required to model the default (or survival) probability curve prevailing at time t, $Q_t(T)$. Just like interest rates models, we start by assuming a given curve at time 0 (here, a survival probability curve $Q_0(T) = G(T)$), and let the stochastic model generate future curves, noted $Q_t(T)$, $T \geq t$.

Given the well-known equivalence between short-rate interest rates models and intensity-based default models, the machinery developed in the interest rates literature can be recycled in credit risk applications, possibly with some restrictions. Indeed, while negative rates can be tolerated (or even desired), such a thing like a "negative intensity" makes no sense.

In standard reduced-form models, credit risk is handled by modelling the default time as the first jump of a stochastic process. The jump likelihood is controlled by the intensity process. The later is most of the time stochastic, and in any case is restricted to be positive or, at least, non-negative. This specific setup corresponds to the *Cox framework*. This specific class of reduced-form models is popular due to its tractability and the fact that it satisfies the *immersion property*. The later guarantees the absence of arbitrage opportunities. Following Jeulin, 1980; Elliott et al., 2000; Jeanblanc and Rutkowski, 2000, a reduced-form model for defaultable claims can be constructed by considering a full filtration obtained by progressively enlarging the default-free market filtration with the default time. The full filtration is usually considered as the relevant filtration in credit risk models: it represents the information available on the market, to be used for pricing and hedging defaultable claims. When working in such a setup, the most fundamental object attached to the random default time is certainly the conditional survival process, known as the *Azéma supermartingale*. Another fundamental behaviour ensuring the no-arbitrage condition in the enlarged filtration is the above-mentioned immersion property. For more details about the immersion property and the enlargement of filtration theory we refer the reader to, among others, Jeanblanc and Le Cam, 2009a and Jeanblanc and Song, 2011. Alternatively, when considering such a framework, a set of problems concerns the specification of the dynamics of the default intensities. In order to ease the calibration of the model, one naturally choose a specification allowing for an easy solution of the pricing problem. Several routes are possible, but some of them run into other problems. One could think about postulating Gaussian dynamics, but this could mean negative default intensities in a large number of classes. Yet, from a practical perspective, it is common to consider the homogeneous affine term structure models with positive dynamics such as CIR (see Duffie and Singleton, 1999) or JCIR (Cox-Ingersoll-Ross with independent compound Poisson jumps, see Brigo and El-Bachir, 2010) to deal with the intensity process. Although very popular, these models present some drawbacks. The classical time-homogeneous affine models such as CIR do not have enough flexibility when it comes to perfectly fit a given market curve. To circumvent this issue, these models were extended by starting with a non-negative time-homogeneous affine (i.e., very tractable) model and adding a deterministic shift, leading to the well known CIR++ (SSRD) or JCIR++ (SSRJD) introduced in Brigo and Mercurio, 2006 and further studied in Brigo and Alfonsi, 2005 and Brigo and El-Bachir, 2010 for specific applications in credit derivatives. The shift approach is appealing since it solves the perfect fit problem and preserves the affine property of the dynamics. However, when shifting the intensity process in a deterministic way, there is no guarantee that the resulting process remains positive when forcing the fit to a given market curve. This problem can be handled by including a non-negativity constraint on the shift function when optimizing the parameters of the time-homogeneous intensity process. More recently, the calibration problem has been solved using a different deterministic adjustment: the shift function is replaced by a time change. These two modifications of the shift extension are referred to as the positive-shift (PS-(J)CIR) and the time-changed (TC-(J)CIR) in Chapter 4. We refer to Duffie and Kan, 1996; Duffie and Singleton, 2003 for general results within the affine term structure models and to Chapter 4 regarding the perfect fit problem.

Interestingly, all these models fit in the class of Cox models. Indeed, in all these cases, the associated Azéma supermartingale is decreasing; its Doob-Meyer decomposition exhibits no martingale component. This shows that such models actually correspond to a very special case. Yet, the above property is interesting: a vanishing martingale part in the Doob-Meyer decomposition of the Azéma supermartingale proves the associated models to be arbitrage-free. Although a few models featuring a martingale component in their Azéma supermartingale have been discussed in the literature Bielecki et al., 2008; Jeanblanc and Vrins, 2018; Crépey and Song, 2014, little work has been done to actually make "non-Cox models" workable.

In this chapter, we deal with a class of default models, *conic martingales* or the *martingale approach*, recently introduced in Vrins, 2014 and Vrins and Jeanblanc, 2015 and further developed in Jeanblanc and Vrins, 2018. Conic martingales offer a modelling framework that completely gets out of Cox models. It consists of a direct modeling of the Azéma supermartingale and is a setup where immersion property does not hold. As explained above, it is therefore not clear whether this model is free of arbitrage opportunities. A central point in this chapter is indeed to give an answer to this question but in a more practical perspective. While very promising, conic martingales trigger important mathematical challenges and deserve in depth technical analysis when dealing with arbitrages opportunities in a no-immersion setup. To fill this gap, we rely on recent results introduced by Crépey and Song, 2014 to show that a special case of conic martingales, the Φ-martingale, belongs to another class of default models that are arbitrage free. In particular, we pay attention to the fact that the Φ-martingale possesses interesting analytical properties which rends it quite suitable for credit risk applications.

In Section 6.1, we recall some standard Cox models (like JCIR++, TC-JCIR and HJM) and briefly review the martingale approach with a particular focus on the Φ-martingale case. For both models, we provide the default time definition. Section 6.2 introduces the study of the no-arbitrage property under immersion and beyond immersion. The arbitrage free property of the Φ-martingale model is then established. In Section 6.3, we compare the performances of the Φ-martingale model with the JCIR++ and TC-JCIR models to the pricing of two classes of credit derivatives: credit valuation (CVA) adjustment under the presence of wrong-way risk and credit default swap (CDS) option, before concluding in section 6.4.

6.1 Intensity-based default models

We consider the general reduced-form setup of Section 2.2 where the default time τ admits an intensity λ. In this case, we are interested to give an overview of some default models that are characterized by the definitions of the couple (τ, λ). We consider three different models and introduce a new one. The first model is the well-known deterministic shift extension to the Cox-Ingersoll-Ross model with compound Poisson jumps (JCIR++ or SSRJD) extensively studied in Brigo and Mercurio, 2006. It will serve as comparison for the other –more recent– models further considered. The second model is a time-changed version of the JCIR, called TC-JCIR, introduced in Chapter 4. It is indeed an interesting alternative to the JCIR++ model. The third model is a defaultable Heath-Jarrow-Morton (HJM) model Bielecki et al., 2011. As shown below, all these approaches fit in the same class of default models, namely *Cox models*. Eventually, the last model, which is of interest here, is based on conic martingales originally introduced in Vrins, 2014 and further studied in Jeanblanc and Vrins, 2018. For each model, we derive the $\mathbb{F}$ and $\mathbb{G}$-conditional survival probability curves (i.e., $S_{\cdot}(T)$ and $Q_{\cdot}(T)$), as well as the Azéma supermartingale ($S_{\cdot}$) respectively defined by (2.16), (2.17) and (2.19).

6.1.1 Basic reduced-form models: the Cox setup

In the Cox setup, the default time is given by (2.22). For more details regarding the properties of this setup, we refer to Section 2.2.2.

In the special case where λ is an affine process, the above conditional expression takes the usual exponential-affine form:

$$S_t(T) = e^{-\Lambda_t} P_t^{\lambda}(T) \ , \quad Q_t(T) = \mathbb{1}_{\{\tau > t\}} P_t^{\lambda}(T) \tag{6.1}$$

where $P^x_t(T)$ is given by (2.27) for $0 \leq t \leq T \leq T^*$. To ease the notation, we set $P^x(t) := P^x_0(t)$.

We give below three examples of Cox models.

JCIR++

The JCIR++ model (see Section 2.3 postulates the following dynamics for the intensity process:

$$\lambda^\varphi_t = x_t + \varphi(t) \tag{6.2}$$

where φ is a deterministic function and x is a time-homogeneous JCIR model

$$dx_t = \kappa(\varrho - x_t)dt + \delta\sqrt{x_t}dB_t + dJ_t, \quad x_0 \geq 0, \tag{6.3}$$

i.e., follows (2.28) with $(a(t), b(t), c(t), d(t), \alpha, \omega(t)) = (\kappa\varrho, -\kappa, 0, \delta^2, \alpha, \omega)$ and J is independent of B.

In this model, τ is defined as in (2.22) but with intensity $\lambda \leftarrow \lambda^\varphi$. Therefore, $\mathbb{F}$ is chosen to be the natural filtration of x (possibly enlarged with the factors impacting the default-free assets), while $\mathbb{G}$ is given by (2.13).

As recalled in Chapter 3, the purpose of the shift function φ is to guarantee a perfect fit between the survival probability $\mathbb{Q}(\tau > t)$ implied by the model with the curve $G(t)$ in (2.4) extracted from market data. Mathematically, it is determined such that $\mathbb{E}[S_t] = G(t)$ for all $t \in [0, T^*]$. This identifies the prevailing shift function for a given set $(x_0, \kappa, \varrho, \delta)$:

$$\mathbb{E}[S_t] = P^{\lambda^\varphi}(t) = e^{-\int_0^t \varphi(s)ds} P^x(t) = G(t) \quad \Rightarrow \quad \varphi(t) = -\frac{d}{dt} \ln \frac{G(t)}{P^x(t)} \,. \tag{6.4}$$

The conditional survival probabilities that $\tau > t$ become

$$S_t(T) = e^{-\int_0^T \varphi(s)ds} e^{-\int_0^t x_s ds} P^x_t(T) = \frac{G(T)}{e^{\int_0^t x_s ds}} \frac{P^x_t(T)}{P^x(T)} \,,$$

and

$$Q_t(T) = \mathbb{1}_{\{\tau > t\}} \frac{G(T)}{G(t)} \frac{P^x(t)}{P^x(T)} P^x_t(T) \,.$$

This model has the advantage of being able to perfectly fit any (continuous) survival probability curve G implied from the market without affecting analytical tractability (in terms of prices of zero-coupon bonds and European options). Moreover, thanks to the jump process J, it can generate large implied volatilities without breaking Feller's constraint, i.e., such that the origin is not accessible for λ. Unfortunately, it suffers from an important drawback: we cannot guarantee the positiveness of the intensity process λ^{φ} (hence, of λ since x can be arbitrarily close to 0) without ad-hoc constraints when computing the parameters $(x_0, \kappa, \varrho, \delta)$. This drawback becomes more and more serious when increasing the activity of the jump process J since only positive jumps are allowed for tractability reasons. Therefore, increasing the jump activity under the constraint that $\mathbb{E}[S_t] = G(t)$ for a given curve G requires to lower the shift, possibly to the negative territory. We refer to Chapter 4 for a detailed analysis of the negativity intensity issue of the JCIR++ model.

TC-JCIR

The TC-JCIR model is an alternative to the JCIR++ aiming to solve the negative intensity issue without losing neither the analytical tractability nor the calibration flexibility of the JCIR++. The model flexibility is achieved by time-changing the non-negative x-model in a deterministic way. Therefore, although similar in principle with the deterministic shift extension of time-homogeneous models introduced above, the positiveness of the intensity process is guaranteed by construction. More specifically, the intensity is modeled as

$$\lambda_t^{\theta} = \theta(t) x_t^{\theta} \ , \quad x_t^{\theta} := x_{\Theta(t)} \ ,$$

where x follows JCIR dynamics, $\Theta(t)$ is a time change function called a *clock* and $\theta(t) := \Theta'(t)$ is the *clock rate*. In this model, τ is defined as in (2.22) but with $\lambda \leftarrow \lambda^{\theta}$. Therefore, $\mathbb{F}^{\theta} := (\mathcal{F}_{\Theta(t)})_{t\in[0,T]}$ is chosen to be the natural filtration of x^{θ} (possibly enlarged with the factors impacting the default-free assets), while $\mathbb{D}$ and $\mathbb{G}$ are given by their corresponding time change filtrations.

The clock plays a similar role as the shift in the JCIR++ model: it is chosen such that

$$\mathbb{E}[S_t] = P^{\lambda^{\theta}}(t) = P^x(\Theta(t)) = G(t) \quad \Rightarrow \quad \Theta(t) = Q^x(G(t)) \ , \tag{6.5}$$

where Q^x is the inverse of P^x. We refer to Chapter 4 for more details about this model.

The conditional survival probabilities that $\tau > t$ are given by

$$S_t(T) = \mathbb{E}\left[e^{-\Lambda_T^{\theta}} \middle| \mathcal{F}_t \right] = \mathbb{E}\left[e^{-\int_0^{\Theta(T)} x_s ds} \middle| \mathcal{F}_t \right] = \frac{P^x_{Q^x(G(t))}(Q^x(G(T)), x_{Q^x(G(t))})}{\exp\{\int_0^{Q^x(G(t))} x_s ds\}} \ ,$$

and

$$Q_t(T) = \mathbb{1}_{\{\tau>t\}} P^x_{Q^x(G(t))}(Q^x(G(T)), x_{Q^x(G(t))}) .$$

It can be shown that the clock solving equation (6.5) takes the form $\Theta(t) := \int_0^t \theta(s)ds$ where θ is non-negative, leading to a valid time change function. Hence, $\mathbb{Q}(\lambda_t \geq 0) = 1$ for all t, solving the negative intensity issue. Interestingly, the process x^θ remains affine if so is x, although not necessarily time-homogeneous affine, obviously. In the sequel we take consider JCIR dynamics for x, i.e., the same as for the JCIR++, for the sake of comparison.

HJM intensities

The general expression of $Q_t(T)$ in (2.26) suggests that the reduced-form models (and JCIR++ and TC-JCIR in particular) can be considered as *short intensity* models, by analogy with *short rate* models in the interest rates literature. Indeed, the conditional expectation agrees with the time-t no-arbitrage price of a default-free zero-coupon bond price with maturity T provided that λ stands for the short risk-free rate. It is possible to revisit these models *à la Heath-Jarrow-Merton*, by modelling directly the term structure of the future default intensities, i.e., by modelling the hazard rate *curve* at once.

In Schonbucher, 1988, Schonbucher models the default-free and defaultable instantaneous forward rate curves with a same Brownian motion. The instantaneous forward curve associated with the risk-free rate is noted $f_t(u)$. The time-t no-arbitrage price of a default-free zero-coupon bond price with maturity T becomes $e^{-\int_t^T f_t(u)du}$ as the process f_t is $\mathcal{F}_t$-measurable. This expression can take a similar form to the short-rate expression $P_t(T)$ given in (2.2) provided that we set $r_t := f_t(t)$ with initial condition $f_0(T) = -\frac{d}{du}P_0(u)\Big|_{u=T}$. In this setup, only the diffusion coefficients of $f.(T)$ need to be specified; the drift is given by a no-arbitrage argument. A similar term structure model is assumed for the instantaneous forward curves associated with the *defaultable* instruments, $\bar{f}_t(T)$. It turns out that defining the *credit spread* process $\lambda_t(T) := \bar{f}_t(T) - f_t(T)$ $(0 \leq t \leq T\ ,\ \ T \in [0, T^*])$, the process $\lambda_t := \lambda_t(t)$ is strictly positive, $\mathbb{Q}$-a.s. In fact, the latter can be interpreted as the default intensity, in the sense that the default time can be defined as in (2.22).

A slightly different point of view is considered in Chiarelli et al., 2011 where the author start from a similar setup but model $f_t(T)$ and $\lambda_t(T)$ with two correlated Brownian motion to obtain the dynamics of $\bar{f}_t(T)$ satisfying $\bar{f}_t := \bar{f}_t(t) = f_t(t) +$

$\lambda_t(t) = r_t + \lambda_t$. Here again, $\lambda_t(T)$ can be interpreted as the $\mathcal{F}_t$-measurable hazard rate curve prevailing at time t. In either HJM frameworks, the drift of $\bar{f}.(T)$ (hence that of $\lambda.(T)$) is given by no-arbitrage, and it holds that

$$Q_t(T) = \mathbb{1}_{\{\tau>t\}}\mathbb{E}\left[e^{-\int_t^T \lambda_s ds}\middle| \mathcal{F}_t\right] = \mathbb{1}_{\{\tau>t\}}e^{-\int_t^T \lambda_t(s)ds} \;.$$

Compared to the JCIR++ model, HJM models are appealing for several reasons. First, the calibration equation $\mathbb{E}[S_t] = G(t)$ is automatically satisfied by imposing the initial condition $\lambda_0(T) = h(T)$:

$$\mathbb{Q}(\tau > T) = \mathbb{Q}(\tau > T|\mathcal{G}_0) = Q_0(T) = e^{-\int_0^T h(s)ds} = G(T) \;.$$

The problem of negative intensities in the JCIR++ model can thus be ruled out by choosing appropriate dynamics for $\lambda.(T)$, as the calibration equation is handled by the initial condition. Second, one can directly model the shape of the volatility of instantaneous forward intensities via the diffusion coefficients. It is worth pointing out that although this can be interested for the sake of dealing with risky rates (like Libor rates, as in Fanelli, 2016), it is probably less relevant for pure credit applications due to the scarcity of quotes on the credit options' market. Moreover, as pointed out in Fanelli, 2016, this model is not easy to deal with in practice. Anyway, as shown above, it still fit in the same class of Cox models (just like JCIR++ and TC-JCIR). Instead, we consider a similar – but different – alternative, called *martingale approach*.

6.1.2 Extended reduced-form models: Conic martingale

In contrast with the models introduced above, the martingale approach is a framework that does not fit in the Cox setup. Instead of defining τ directly as in (2.22) for Cox models or via an intensity, this approach consists of modelling the $\mathbb{F}$-conditional survival probability curves (2.16) directly using diffusion martingales,

$$dS_t(T) = \sigma(t, S_t(T))dB_t \;, \quad t \leq T \;. \tag{6.6}$$

The requirement $\mathbb{E}[S_t] = G(t)$ is automatically satisfied by choosing the initial condition $S_0(T) = G(T)$.

Remark 14. *Because we are modelling the $\mathbb{F}$-conditional probabilities with the help of the Brownian motion B, the latter must be $\mathbb{F}$-adapted. Hence, $\mathbb{F}$ can be chosen as the natural filtration of B (possibly enlarged with the factors impacting* *the default-free assets). Notice that we do not provide an explicit construction scheme for τ. Therefore, at this stage, we cannot specify explicitly the filtrations $\mathbb{D}$ and $\mathbb{G}$. This is a major drawback of the martingale models. This point will be addressed later in the chapter in the particular case of Φ-martingales.*

To the best of our knowledge, the martingale approach was first considered in Cesari et al., 2009 in the context of counterparty risk on credit derivatives. They postulate a diffusion coefficient of the form $\sigma(t, T)$, Cesari et al., 2009, section 3.3.1. Obviously, because of the lack of state-dependency, this setup leads to Gaussian dynamics,

$$S_t(T) = S_0(T) + \int_0^t \sigma(s,T)dB_s \ \sim \mathcal{N}\left(S_0(T), \int_0^t \sigma^2(s,T)ds\right) .$$

Just like the JCIR++, TC-JCIR and HJM models introduced above, the model can be perfectly calibrated to the market: imposing the initial condition $S_0(T) = G(T)$ leads to $\mathbb{E}[S_t] = G(t)$. Notice that (6.6) is a *family* of SDEs. For each $T \in [0, T^*]$, the process $S_t(T)$, $0 \leq t \leq T$ represents the evolution of the $\mathcal{F}_t$-conditional probability that $\tau > T$. It is obviously not a binary process since τ is *not* an $\mathbb{F}$-stopping time.[1] Nevertheless, each of those processes must belong to the interval $[0, 1]$. This is clearly violated by Gaussian dynamics. Specifying the shape of the diffusion coefficient such that $S_.(T) \in [0, 1]$ and, at the same time, get a tractable model is not easy. To circumvent this problem, we can directly model $S_.(T)$ with conic martingales, i.e., with martingales evolving within a specific range. Conic martingales have been introduced in Vrins, 2014 and further studied in Jeanblanc and Vrins, 2018. Credit risk has been mentioned as a potential application but without being further developed.

Following Jeanblanc and Vrins, 2018, the $\mathbb{F}$-conditional probability curves are modeled in one go. To make sure that $S_t(T) \in [0, 1]$, we start from a family of latent processes, and map them into a function with appropriate image:

$$S_t(T) := F(Z_{t,T}) , \tag{6.7}$$

where $F : \mathbb{R} \longrightarrow [0, 1]$ is a $\mathcal{C}^2$ invertible function and $Z_{t,T}$, $0 \leq t \leq T, T \leq T^*$, a family of diffusions driven by the same Brownian motion B[2],

$$dZ_{t,T} = a(t, Z_{t,T})dt + \eta(t, Z_{t,T})dB_t \ , \quad Z_{0,T} := F^{-1}(G(T)) . \tag{6.8}$$

[1]Observe from (2.25) that this feature is not a specificity of the conic martingale approach. It results from the definition of $S_t(T)$, and is obviously shared by Cox models.

[2]Note that more general dynamics of the process $Z_{t,T}$ could be considered here, e.g., models featuring infinite activity features (Levy). Nevertheless, using diffusions, we have a more complete market, which is better for hedging.

This is similar in spirit to one-factor HJM models in the sense that we directly model probability curves with the help of a single Brownian motion. Moreover, the drift in (6.8) is uniquely determined by the martingale property of $S_t(T)$:

$$a(t,z) = \frac{\eta^2(t,z)}{2}\psi(z)\ , \quad \psi(z) := -F''(z)/F'(z)\ . \tag{6.9}$$

This is a simple consequence of Itô's lemma (see Jeanblanc and Vrins, 2018 for more details).

However, there is a fundamental difference with HJM models: the martingale approach does not belong to the class of Cox models. Indeed, in contrast with intensity models where S is decreasing (2.24), the decomposition of S in the martingale approach *does* feature a non-zero martingale part. In the conic martingale case for instance,[3]

$$dS_t = -\frac{F'\left(F^{-1}(S_t)\right)}{F'\left(F^{-1}(G(t))\right)}h(t)G(t)dt + F'\left(F^{-1}(S_t)\right)\eta\left(t, F^{-1}(S_t)\right)dB_t\ .$$

In the sequel, we assume that the diffusion coefficient of $Z_{t,T}$ is a bounded function of time, i.e., $\eta(t,z) = \eta(t)$ where $0 < \eta^2(t) < \infty$ on $[0, T^*]$. Moreover, we assume that the score function ψ of the mapping F is Lipschitz continuous. Then, for each $u \leq T^*$, $0 \leq t \leq u$, the SDE

$$dZ_{t,u} = \frac{\eta^2(t)}{2}\psi(Z_{t,u})dt + \eta(t)dB_t \tag{6.10}$$

has a strong, pathwise unique solution on $[0, T^*]$ (see Kloeden and Platen, 1999). Since F is a bijection, it is invertible, and the diffusion coefficient of $S_t(T)$ in (6.6) takes the form $\sigma(t,z) = \eta(t)F'\left(F^{-1}(z)\right)$.

The Φ-martingale default model

A special case consists of considering $F = \Phi$, the cumulative distribution of the standard normal random variable. This is a very particular case where $Z_{\cdot,T}$ are Gaussian processes. Indeed, the score function ψ collapses to the identity. Therefore, the drift in (6.8) is linear and the diffusion coefficient is a bounded, implying that the SDE admits a unique strong solution, and leading to a tractable model. In this model, the process $S_\cdot(T)$ corresponds to Φ-martingale Jeanblanc and Vrins, 2018.

[3] Notice that it is not enough to replace T by t to get the dynamics of $Z_{t,t}$, as it corresponds to the dynamics of $Z_{t,T}$ for a fixed T. The dynamics of S are obtained by applying Itô-Wentzell's lemma to $F(Z_{t,t})$ with $F(Z_{0,t}) = G(t)$.

Definition 6 (The Φ-martingale default model)**.** *The Φ-martingale model postulates that $S_{\cdot}(T)$ is a Gaussian process mapped to the standard normal cumulative distribution function, i.e.,*

$$S_t(T) = \Phi(Z_{t,T}) \tag{6.11}$$

with

$$Z_{t,T} = \Phi^{-1}(G(T))e^{\int_0^t \frac{\eta^2(s)}{2}ds} + \int_0^t \eta(s)e^{\int_s^t \frac{\eta^2(u)}{2}du}dB_s \tag{6.12}$$

$$\sim \mathcal{N}\left(\Phi^{-1}(G(T))e^{\int_0^t \frac{\eta^2(s)}{2}ds},\ e^{\int_0^t \eta^2(s)ds} - 1\right) . \tag{6.13}$$

This model is of particular interest for several reasons. First, when $\eta \equiv 1$, the process (6.11) can be seen as the analog of the Brownian motion (martingale valued on $\mathbb{R}$) and its Doléans-Dade exponential (martingale valued in $\mathbb{R}^+$) but for the $[0, 1]$ range; see Jeanblanc and Vrins, 2018 for a discussion. This feature has been noticed independently by Carr, who called the corresponding process *Bounded Brownian motion*, Carr, 2017. Second, the solution $S_t(T) \in [0, 1]$ is known in closed form for all $0 \le t \le T$, $T \in [0, T^*]$, and the probability distribution of $S_t(T)$ is known analytically from (6.13). Third, this model automatically meets the calibration equation, by construction. Indeed, for every $X \sim \mathcal{N}(\mu, \sigma^2)$, it holds

$$\mathbb{E}[\Phi(X)] = \Phi\left(\frac{\mu}{\sqrt{1+\sigma^2}}\right) .$$

Hence,

$$\mathbb{E}[S_t] = \mathbb{E}[\Phi(Z_{t,t})] = \Phi\left(\frac{\Phi^{-1}(G(t))e^{\int_0^t \frac{\eta^2(s)}{2}ds}}{\sqrt{e^{\int_0^t \eta^2(s)ds}}}\right) = G(t) .$$

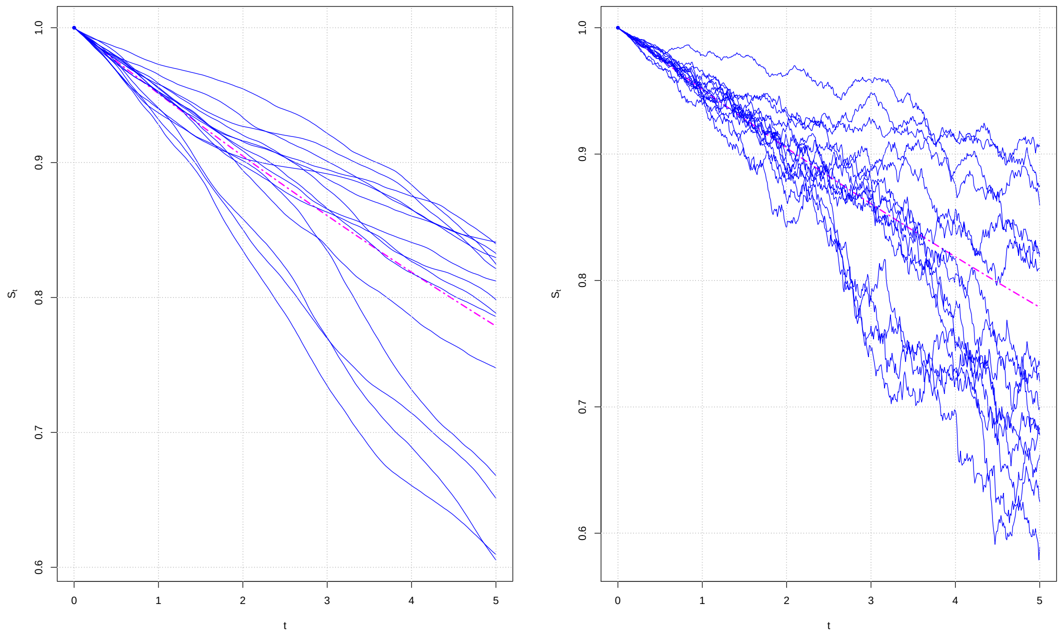
The only "free" parameter is thus the time-dependent volatility function η, controlling the randomness of the survival probabilities. As explained above, sparsity is often considered as an asset in credit derivatives. By using the properties of Φ, it is easy to show that the diffusion coefficient associated with the Φ-martingale in (6.6) is $\sigma(t, z) = \eta(t)\phi(\Phi^{-1}(z))$.
The associated Azéma supermartingale S_t is given by the following Itô process

$$S_t = 1 + \int_0^t e^{\int_0^s \frac{\eta^2(u)}{2}du} \frac{\phi(\Phi^{-1}(S_s))}{\phi(\Phi^{-1}(G(s))} dG(s) + \int_0^t \eta(s)\phi(\Phi^{-1}(S_s))dB_s . \tag{6.14}$$

Differentiating (6.14) shows that it is a supermartingale satisfying (2.21) with

$$\lambda_t S_t = e^{\int_0^t \frac{\eta^2(u)}{2}du} \frac{\phi(\Phi^{-1}(S_t))}{\phi(\Phi^{-1}(G(t))} h(t)G(t) \quad \text{and} \quad \sigma_t = \eta(t)\phi(\Phi^{-1}(S_t)) . \tag{6.15}$$

In this approach, the martingale part of the decomposition (2.21) is non-zero (so S_t is not decreasing in contrast with the intensity approach) which gives the power to the conic martingale approach to be more convenient when modelling in credit risk compared to the Cox setup. Figure 6.1 gives a comparison of the sample paths of S_t between a Cox model using CIR intensity and the Φ-martingale model.



(a) Cox with CIR intensity (b) Φ-martingale

Figure 6.1: Sample paths of S_t (solid blue) and $\mathbb{E}[S_t] = G(t)$ (dotted magenta): $\eta = 0.15$ and CIR parameters $\kappa = 0.35$, $\beta = 0.045$, $\delta = 0.15$, $\lambda_0 = 0.035$ and $h(t) = 5\%$.

Default time definition in martingale models

The conic martingale setup provides an appealing way to model future survival probability curves with the correct range and therefore, is an interesting alternative to Cesari et al., 2009. However, the standard approach in default modelling is to first define τ, e.g., as a first-passage time, and then compute the survival probabilities of interest, as in (2.22) for Cox. At this stage however, it is not clear how one can construct a default time τ associated with given dynamics for the Azéma supermartinagle in the case of conic martingale models. This is not a secondary question. As recalled in the introduction, it is important to associate a definition for τ in order to deal with potential arbitrages issues in the enlargement of filtration setup.

A natural question to ask is whether one could still use the intensity process λ to define τ as in (2.22) when the martingale part of the Doob-Meyer decomposition of the Azéma supermartingale in (2.21) does not vanish. After all, Λ is still increasing. It turns out that, generally speaking, this is not correct: defining τ as in (2.22) leads to a random time which distribution is not compatible with the survival process S, as we now show.

Lemma 15. *Consider a model whose Azéma supermartingale S takes the Doob-Meyer decomposition* (2.21)*. and let us note τ the default time associated with this model. Now, define $\tilde{\tau}$ as in* (2.22) *where λ is the intensity process in* (2.21)*. Then, $\tau \not\sim \tilde{\tau}$, in general.*

Proof. It is clear that $\tilde{\tau}$ is the default time in a Cox setup which survival process solves $d\tilde{S}_t = -\lambda_t \tilde{S}_t dt$ with $\tilde{S}_0 = 1$, i.e., $\tilde{S}_t = \exp\{-\int_0^t \lambda_s ds\}$. Because the same intensity process λ enters both S and $\tilde{S}$, the solution to (2.21) can be written as the multiplicative form $S_t = \tilde{S}_t \tilde{M}_t$ where $\tilde{M}_t := \exp\{-\int_0^t \frac{\sigma_s^2}{2S_s^2} ds + \int_0^t \frac{\sigma_s}{S_s} dB_s\}$ is a martingale with unit expectation. Indeed, $S_0 = \tilde{S}_0 M_0 = 1$ and from Itô's product rule,

$$dS_t = (-\lambda_t \tilde{S}_t dt)\tilde{M}_t + \tilde{S}_t \left(\frac{\tilde{M}_t \sigma_t}{S_t} dB_t \right) = -\lambda_t S_t dt + \sigma_t dB_t \ .$$

From the Tower law, we have $\mathbb{Q}(\tilde{\tau} > t) = \mathbb{E}[\tilde{S}_t]$ and $\mathbb{Q}(\tau > t) = \mathbb{E}[S_t]$, where

$$\mathbb{E}[S_t] = \mathbb{E}\left[\tilde{S}_t \tilde{M}_t\right] = \mathbb{C}ov\left(\tilde{S}_t, \tilde{M}_t\right) + \mathbb{E}\left[\tilde{S}_t\right] \ .$$

Clearly, σ depends on S (in a non-linear way as dS must vanish when $S \downarrow 0$ or $S \uparrow 1$), hence on λ. Therefore, $\tilde{S}$ and $\tilde{M}$ both depend on λ, and there is no reason for their covariance to vanish, in general. □

Remark 15. *In the limit where $\eta \equiv 0$, so is σ, each process $S_.(T)$ becomes a trivial martingale (i.e., a constant equal to $G(T)$), and the model becomes deterministic, $S_t = G(t)$. Hence, the survival process solves*

$$dS_t = -\lambda(t) S_t dt \ ,$$

where the deterministic intensity function satisfies $\lambda(t) = h(t)$. This becomes similar to a Cox model with a deterministic intensity λ given by the hazard rate *function h. One can then define the default time as in* (2.22) *with* $\lambda_t \leftarrow h(t)$*. Similarly, when the intensity process is deterministic, so is* $\tilde{S}$ *and* $\mathbb{C}\mathrm{ov}\left(\tilde{S}_t, \tilde{M}_t\right) = 0$*. Therefore,* $\mathbb{Q}(\tau > t) = \mathbb{E}[S_t] = \mathbb{E}[\tilde{S}_t] = \mathbb{Q}(\tilde{\tau} > t)$ *and both* $\tau, \tilde{\tau}$ *have the same survival function given by* $G(t)$*, showing that in the case of a deterministic intensity, one can define* τ *as in* (2.22)*. But these two examples are special cases. In general, we cannot define* τ *using* (2.22) *in models whose Azéma supermartingale is non-decreasing.*

6.2 Immersion and arbitrages

In this chapter, we consider several information flows, characterized by the knowledge (or not) of the default indicator. This does not trigger any problem in Cox processes, where an explicit construction scheme is available for τ. In this case indeed, the various filtrations (namely, $\mathbb{F}$, $\mathbb{D}$ and $\mathbb{G} = \mathbb{F} \vee \mathbb{D}$) are well identified. For instance, it is therefore relatively easy to check that the knowledge of the default indicator does not provide a superior information to $\mathbb{F}$ when it comes to pricing default-free assets. This is however much more difficult to verify when there is no explicit definition for τ. In this case indeed, $\mathbb{D}$ and hence $\mathbb{G}$ are not explicitly identified. We refer to Aksamit et al., 2014 for more details and explicit examples of classical arbitrages using the knowledge of τ. To avoid these issues, a basic reduced-form approach under the standard Cox model has already been proposed in order to deal with counterparty risk modelling Duffie and Singleton, 1999; Brigo et al., 2013; Crépey, 2015. Our purpose is how can we extend this approach beyond such a basic Cox setup using conic martingale models without facing arbitrage opportunities.

6.2.1 Models without martingale part (Cox models)

Default models that belong to the class of Cox models are known to be arbitrage-free. Indeed, it can be shown that there is no arbitrage opportunity provided that every $\mathbb{F}$-martingale remains a $\mathbb{G}$-martingale (see Jeanblanc and Le Cam, 2009a). This condition, first introduced under the name of $\mathcal{H}$ hypothesis in Brémaud and Yor, 1978, is commonly referred to as the *immersion property.* Cox models provide a very convenient modeling environment in this respect, as they are proven to always satisfy the immersion property. Indeed, it is known (see, e.g., Bielecki et al., 2011, Remark 3.2.1. (iii)) that the immersion property is equivalent to

$$S_t = S_t(t) = \mathbb{Q}(\tau > t|\mathcal{F}_t) = \mathbb{Q}(\tau > t|\mathcal{F}_{T^*}) = S_{T^*}(t), \text{for any}\, t,\, T^* > t\ . \tag{6.16}$$

Cox models satisfy the above condition (hence the immersion property):

$$S_{T^*}(t) = \mathbb{Q}(\tau > t|\mathcal{F}_{T^*}) = \mathbb{Q}(\Lambda_t \leq \mathcal{E}|\mathcal{F}_{T^*}) = \mathbb{Q}(\Lambda_t \leq \mathcal{E}|\mathcal{F}_t) = e^{-\Lambda_t} = S_t$$

for every $t \in [0, T^*]$, where we have used that $\mathcal{E}$ is independent from $\mathcal{F}_{T^*}$ and Λ is $\mathbb{F}$-adapted.

6.2.2 Models with martingale part (Conic martingale models)

It is easy to see that the condition (6.16) implies that the martingale part in the Doob-Meyer decomposition of S must vanish. Indeed, $S_t(T)$ is decreasing in T for all t. In particular, $S_{T^*}(T)$ is decreasing in T, too from (6.16). As a consequence, immersion cannot hold if S features a non-trivial martingale part. Therefore, existence of potential arbitrage opportunities in such models requires more attention compared to Cox models. Interestingly, it has been shown in Crépey and Song, 2014 that a suitable redued-form can be applied beyond the immersion setup under some conditions satisfied by the dynamic Gaussian copula (DGC) credit model. A total valuation adjustment (TVA) price process of a general defaultable security based on the dynamic Gaussian copula model have been proposed. In this section, we show that the Φ-martingale default model rules out arbitrage opportunities in the sense that it is a particular case of a DGC when an additional condition on the diffusion parameter η is satisfied. To do this, we first recall how to define a corresponding default time to the Φ-martingale model.

Lemma 15 calls for a procedure to construct τ in the martingale setup. We show below that this is easy for the Φ-martingale approach, as it fits in the class of Dynamized Gaussian copula models, for which a closed-form expression of the default time has been given; see Crépey et al., 2013.

Definition 7 (Dynamized Gaussian Copula model)**.** *The dynamized Gaussian copula (DGC) model is a default model where the default time is defined as*

$$\tau = \ell^{-1}\left(\int_0^\infty f(s)dB_s\right) \tag{6.17}$$

where B is an $\mathbb{F}$-adapted Brownian motion, f is a square integrable function with unit L^2-norm and $\ell : \mathbb{R}_+ \to \mathbb{R}$ is a differentiable increasing function from satisfying $\lim_{u\to 0} \ell(u) = -\infty$ and $\lim_{u\to\infty} \ell(u) = +\infty$.

Proposition 4. *The default time associated with the conic martingale model* (6.7) *can be defined as* (6.17) *if and only if $F = \Phi$ and $\int_0^\infty \eta^2(u)du = +\infty$.*

Proof. Let us start by computing $S_t(T)$ in the DGC model. Suppose that τ is given by (6.17) and define $\varsigma^2(t) := \int_t^\infty f^2(s)ds$ and $m_t := \int_0^t f(s)dB_s$. Then,

$$\{\tau > t\} = \left\{ \int_t^\infty f(s)dB_s > \ell(t) - m_t \right\} . \tag{6.18}$$

The Itô integral is distributed as a zero-mean normal variable with variance $\varsigma(t)$, so that the $\mathbb{F}$-conditional survival process of τ collapses to

$$S_t(T) = \mathbb{Q}(\tau > T|\mathcal{F}_t) = \Phi\left(\frac{m_t - \ell(T)}{\varsigma(t)}\right) . \tag{6.19}$$

Let us now set $F = \Phi$ and show that this is the form taken by the $\mathcal{F}_t$-conditional survival probability of the event $\tau > T$ in the Φ-martingale model provided that

$$\ell(u) = -\Phi^{-1}(G(u)) \quad \text{and} \quad f(s) = \eta(s)e^{-\int_0^s \frac{\eta^2(u)}{2}du}. \tag{6.20}$$

Using these notations, we can write the solution in (6.13) as

$$Z_{t,T} = \frac{Z_{0,T} + \int_0^t \eta(s)e^{-\int_0^s \frac{\eta^2(u)}{2}du}dB_s}{e^{-\int_0^t \frac{\eta^2(s)}{2}ds}} = \frac{\Phi^{-1}(G(T)) + m_t}{\varsigma(t)} \tag{6.21}$$

Indeed, remark that f in (6.20) is of unit L^2-norm (as in Definition 7) if

$$e^{-\int_0^\infty \eta^2(u)du} = 0$$

or equivalently

$$\int_0^\infty \eta^2(u)du = +\infty.$$

So that,

$$\varsigma^2(t) = \int_t^\infty f^2(s)ds = \int_t^\infty \eta^2(s)e^{-\int_0^s \eta^2(u)du} = e^{-\int_0^t \eta^2(s)ds}.$$

From (6.11),

$$S_t(T) = \Phi\left(\frac{\Phi^{-1}(G(T)) + m_t}{\varsigma(t)}\right) \tag{6.22}$$

which agrees with (6.19).

Let us now show that $S_t(T)$ associated with the conic martingale default model with $F \neq \Phi$ cannot be written as (6.19). It is easy to show that if ψ is regular enough for (6.10) to admit a unique strong solution,

$$S_t(T) = F\left(\frac{m_t - l_t(T)}{\varsigma_t}\right) ,$$

where

$$l_t(T) = -F^{-1}(G(T)) + \int_0^t \left(\psi(Z_{s,T}) - Z_{s,T}\right)\frac{\eta^2(s)}{2} e^{-\int_0^s \frac{\eta^2(u)}{2} du} ds. \tag{6.23}$$

This agrees with (6.19) if and only if $F = \Phi$, leading to $\psi(x) = x$ and $l_t(T) = -F^{-1}(G(T)) = l(T)$. $\square$

The next corollary shows that the Φ-martingale model is an arbitrage free default model in the sense of Crépey and Song, 2014 that allows for automatic calibration to CDS market quotes insured by a specific function ℓ in (6.20) under a specific condition on the diffusion process η.

Corollary 3. *The Φ-martingale model is a case of "non immersion" arbitrage free default model if $\int_0^\infty \eta^2(u)du = +\infty$.*

Proof. The result follows with a direct application of Proposition 4 and Theorem 6.2 in Crépey and Song, 2014. $\square$

Notice that on the top of being arbitrage-free, the Φ-martingale default model features some attractive additional properties in a practical perspective. First the distribution of $S_t(T)$ is known in closed form. Second, since in our case, we first specify the dynamics of $Z_{t,T}$, this intuitively implies the choices of f and ℓ in (6.20), where, in particular, ℓ is given by the calibration constraint, $\ell(u) = Z_{0,u} := \Phi^{-1}(S_{0,u}) = \Phi^{-1}(G(u))$. Third, another contribution comes from the fact that one has an exact scheme for S and then for λS via (6.15) easing the numerical computation of respectively (2.34) and (2.38) since $S_t(T)$ can be rewritten as a function of S_t (see Appendix 6.A for more details).

6.3 Numerical experiments

In this section, we provide numerical examples by considering the Φ-martingale default model in (6) with constant diffusion coefficient $\eta(t) = \eta$ (in this case, the

condition $\int_0^\infty \eta^2(u)du = +\infty$ is satisfied). We choose as benchmark the JCIR++ model devised in section 6.1.1 which is a very standard approach when it comes to deal with high credit spread (see for example Brigo and El-Bachir, 2010). The only drawback with this model is that when forcing perfect fit to market data, the implied shift function φ can become negative leading to negative intensities as shown in Chapter 3 and Chapter 4. To avoid this issues, one solution consists of imposing a positive shift constraint leading to the PS-JCIR model, but this restricts the level of volatility featured by the model (see Chapter 4 for more details). Another solution is to consider the time change approach introduced in Chapter 4 leading to the TC-JCIR model as discussed in section 6.1.1.
The performances of the Φ-martingale model will be compared to respectively the PS-JCIR and the TC-CIR (i.e., TC-JCIR without jumps) models using real market data when the default counterparty is Ford. We then consider two different applications in credit risk namely the pricing of credit value adjustment (CVA) in the presence of wrong-way risk (WWR) effects and credit default swap options (CDSO) by considering Ford as reference entity. The considered Ford's CDS spreads are given in Table 4.1 of Chapter 4.

As we can see with Table 4.1, the counterparty's term structure is not fully known at each point in time but only provides data at some maturities. In this context, we need further assumptions in order to construct the market curve G associated to the default time τ of the reference entity. To do so, we assume piecewise constant hazard rates bootstrapped form the CDS spread associated to each maturity of Table 4.1. This is a common market practice procedure known as the JP Morgan model Markit, 2004. Once the market curve G is fully determined, the next step is to calibrate the models parameters to the obtained market curve. While the Φ-martingale model provide an automatic calibration, the PS-JCIR and the TC-CIR models's parameters need to calibrated to the market curve. This is done using an optimization procedure searching the models parameters that minimize the discrepancies between model and market risk-neutral survival probability curves (see Chapter 4). Notice that for the spacial case of the PS-JCIR, the optimization problem includes an additional constraint (i.e., $\varphi \geq 0$) ensuring non-negativity of the intensity process governed by the JCIR++ model. The parameters obtained after calibration are

$$\kappa = 0.0624, \quad \varrho = 0.2975, \quad \delta = 0.3343, \quad x_0 = 0.0000$$

for the TC-CIR model, and

$$\kappa = 0.4382, \quad \varrho = 0.0086, \quad \delta = 0.0396, \quad x_0 = 0.1051$$

for the PS-JCIR model with jumps parameters

$$\omega = 9.5619 \cdot 10^{-10}, \quad \alpha = 3.1508 \cdot 10^{-10}. \tag{6.24}$$

Observe that the jumps parameters of the PS-JCIR model in (6.24) are very close to zero that would mean that the PS-JCIR is not able to reproduce significantly much more volatility compared to the corresponding PS-CIR model. Numerical examples about the PS-CIR model showing its limitations to reproduce high WWR effects and CDSO implied volatilities are provided in Chapter 4. In what follows, we will focus on the numerical comparison of the three considered models (Φ-martingale, PS-JCIR and TC-CIR) in term of their ability to feature both WWR and CDSO implied volatilities.

6.3.1 Application to wrong-way risk CVA

As shown in section 6.2, all the considered default models (Φ-martingale, PS-JCIR and TC-CIR) are free of arbitrage opportunities and a general risk-neutral valuation formula of the time-0 CVA on $[0, \tau \wedge T]$ (neglecting margin effects) is given by Corollary 3 and (2.38):

$$\mathrm{CVA} = \mathbb{E}\left[(1-R)\int_0^T \frac{V_u^+}{\beta_u}\lambda_u S_u du\right] \tag{6.25}$$

where $\lambda_t S_t$ is given by (6.15) for the Φ-martingale model, $\lambda_t^\varphi e^{-\int_0^t \lambda_u^\varphi du}$ for the PS-JCIR model and $\lambda_t^\theta e^{-\int_0^t \lambda_u^\theta du}$ for the TC-CIR model. The bank account $\beta_t = e^{\int_0^t r_s ds}$ where r is the risk-free rate that we assume to be driven by the following Vasicek dynamics:

$$dr_t = \gamma(\theta - r_t)dt + \sigma dW_t, \quad r_0 \in \mathbb{R}$$

where γ, θ, σ are positive constants and W is an $\mathbb{F}$-Brownian motion.

In order to compute the CVA, one therefore needs a dynamic and tractable model for the exposure process V. Among the OTC derivatives, Interest Rate Swaps (IRS) are very popular and sensible to default. This motivates to deal with counterparty risk in IRS portfolio in our CVA example. An IRS contract implies two counterparties: B (most often a bank), the payer, exchanges a fixed rate K for a floating rate F with a defaultable counterparty C (here, Ford), called the receiver, at say quarterly payments dates $\mathcal{T}_{a+1}, \ldots, \mathcal{T}_b$. The contract starts at $\mathcal{T}_a$ and ends at $T = \mathcal{T}_b$. However, if C defaults before the maturity of the contract (at the default time τ), B looses part of the contract and will only receive the recovered part of the exposure . CVA is the expected losses (the non-recovered part) do to the default of C. This IRS contract can be used to manage interest rates risk. The discounted payoff at time t with 1 notional from B's side can be expressed as:

- if $t \leq \mathcal{T}_a$,
$$V_t = \sum_{i=a+1}^{b} \Delta_i P_t(\mathcal{T}_i)(F_t(\mathcal{T}_{i-1}, \mathcal{T}_i) - K)$$

- if $\mathcal{T}_{j-1} < t \leq \mathcal{T}_j$,
$$V_t = (F_{\mathcal{T}_{j-1}}(\mathcal{T}_{j-1}, \mathcal{T}_j) - K)\Delta_j P_t(\mathcal{T}_j) + \sum_{i=j+1}^{b} \Delta_i P_t(\mathcal{T}_i)(F_t(\mathcal{T}_{i-1}, \mathcal{T}_i) - K)$$

where
$$F_t(\mathcal{T}_{i-1}, \mathcal{T}_i) = \frac{P_t(\mathcal{T}_{i-1}) - P_t(\mathcal{T}_i)}{P_t(\mathcal{T}_i)\Delta_i}$$
is the forward rate prevailing at t between $\mathcal{T}_{i-1}$ and $\mathcal{T}_i$ and $P_t(T) = \mathbb{E}[\beta_t/\beta_T|\mathcal{F}_t]$ also given in (2.2) is the time-t risk free zero-coupon bond price with maturity T. In general, V and τ are not independent and in this case we have to take into account the dependency between default probability and exposure. This dependency is known as WWR and is experienced here by correlating the Brownian motions B and W meaning that $dB_t dW_t = \rho dt$. In the independent case (no-WWR or $\rho = 0$), CVA depends only on the discounted expected positive exposure (EPE): $\mathbb{E}\left[\frac{V_t^+}{\beta_t}\right]$ since all models are calibrated to the same market survival probability curve G. In this case, the independent CVA formula is given by (2.39).

Under No-WWR the CVA is partially model-dependent because the EPE in the right hand side of (2.39) is not known in closed-form and is computed using Monte Carlo simulation. Figure 6.2 depicts the EPE profile with respect to time. As we can observe it is in line with the fact that the value of a swap contract is null at inception and the maturity time. In addition, the shape of this figure is not surprising as it when payments start to occur at $\mathcal{T}_{a+1}$, there is less money at stake and the potential risk reduces instantly after each cash flow because the interest rate risk is proportional to the number of remaining payments. Between payments, the potential losses is increasing but this increase is less apparent as between payments, the next floating rate is fixed.

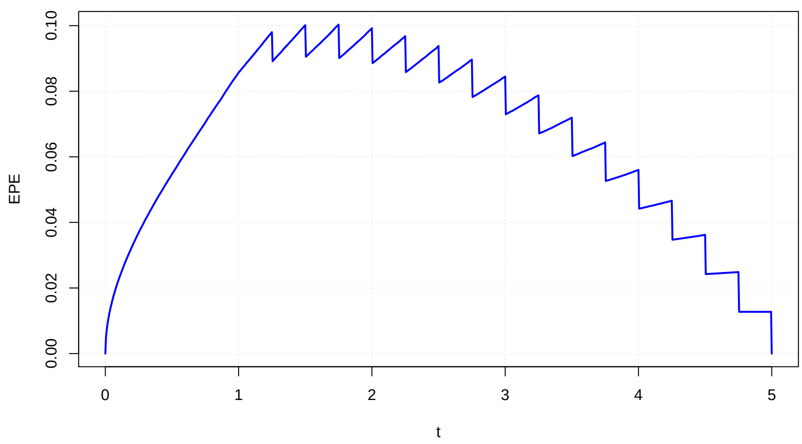


Figure 6.2: Discounted expected positive exposure (EPE) computed by Monte Carlo simulation using $5 \cdot 10^5$ paths and time step of 0.1%. IRS with Vasicek parameters $\gamma = 0.4$, $\theta = 0.026$, $\sigma = 0.14$ and $r_0 = 0.0165$ fitted to a flat interest rate market curve with constant yield of 3%. We used $\mathcal{T}_a = 1$, $\mathcal{T}_b = 5$, fixed rate $K = F_0$ and quarterly payment dates $\Delta_i = 0.25$.

Under WWR, the CVA is fully model-dependent and potentially increases with respect to the correlation ρ. In Figure 6.3, we plot the CVA in function of ρ for the three considered models: PS-JCIR (solid green), TC-CIR (dotted blue) and Φ-martingale (dashed magenta with $\eta \in \{0.10, 0.15, 0.20, 0.50, 0.75\}$), calibrated to the same survival probability curve G given by Ford's CDS term structure. Thus due to the calibration constraint, all models agree with the special case of no-WWR ($\rho = 0$): the independent CVA lined up in cyan. Further, we observe that CVA generally increases with the correlation parameter ρ for all the considered models meaning that all the models are able to reproduce the WWR effect. However, if we evaluate the models performances in term of their capability to feature high WWR, one can notice that the PS-JCIR is less competitive. Do to the positivity constraint, it features the lowest WWR impact which is comparable to the one featured by the Φ-martingale with $\eta = 10\%$. In particular, when increasing the volatility parameter η, the WWR produced by the Φ-martingale model increases as well and is comparable to the one generated by the TC-CIR model when $\eta = 75\%$. In summary, the Φ-martingale model allows for automatic calibration to CDS quotes and can reproduce a range of WWR effects (even high) with the help of one parameter, η which is a nice feature here as there is no quotes for CDS options.

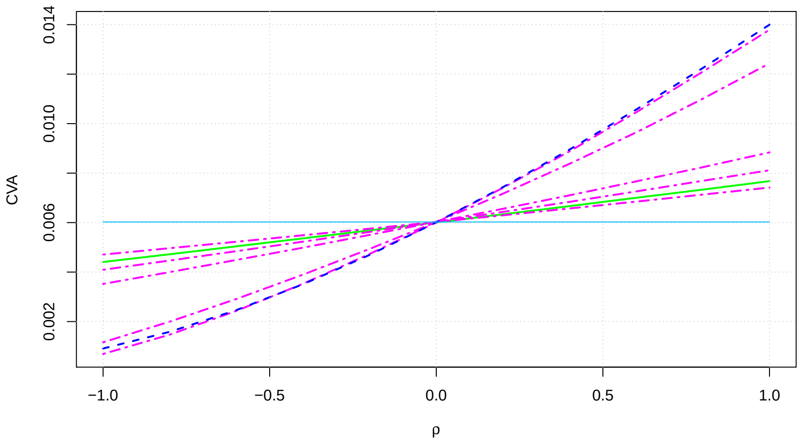


Figure 6.3: CVA figures as a function of the correlation ρ for PS-JCIR (solid green), TC-CIR (dotted blue), Φ-martingale (dotted magenta, $\eta \in \{0.10, 0.15, 0.20, 0.50, 0.75\}$) and the CVA with zero-correlation (solid cyan) using Monte Carlo method with 10^6 paths, time step 0.01. Profiles: 5Y IRS exposure detailed in 6.2.

6.3.2 Volatility surface of CDS option

We now proceed to the assessment of the volatility of CDS spreads in term of CDS option implied volatilities. A CDS (call) option with maturity T_a on a single name gives the invertor the right to enter at T_a a *payer* CDS on a single name with contractual spread k and termination time $T_b > T_a$. The corresponding no-arbitrage price is given at time $t = 0$ by (2.34) with

$$\overline{Q}_t(T) := \mathbb{1}_{\{\tau > t\}} Q_t(T) .$$

Clearly, formula (2.34) can be implemented using the three considered models (Φ-martingale, PS-JCIR and TC-CIR) by considering the corresponding conditional survival process of each model.
A CDS option is typically quoted on the market in term of its Black implied volatility $\bar{\sigma}$ which is based on the assumption that the credit spread follows a geometric Brownian motion (See Section 2.4.1).

Given the forward spread and risky annuities, we can compute the implied volatilities for payers written on the same underlying CDS for the three models (Φ-martingale, PS-JCIR and TC-CIR) using at-the-money payer (Table 6.1) or with different strikes (Table 6.2). We have assumed zero interest rates in the numerical applications since we are focusing on credit impact. As expected, Table 6.1 shows that the PS-JCIR model generates small implied volatilities compared the other models. In addition, it is not difficult to notice that the TC-CIR model features much more implied volatilities compared to the PS-JCIR but the level of volatility is relatively small. It could be possible to increase level of volatility implied by the TC-CIR by playing with the model's parameters but Feller's constraint is required and puts limits to the volatility magnitude that can be achieved. In contrast, the Φ-martingale model gives the freedom to increase the volatility level without facing any constraint and this in turn allows to increases the CDS option implied volatility without bounds. This is very important when dealing with a counterparty of poor credit quality characterized by a high degree of risk as it was the case of many firms in the recent global crisis, a property that still lot of existing models fail to capture. Another interesting feature is to notice that the Φ-martingale is very parsimonious which is an advantage in credit characterized by lack of information on which to calibrate except the initial curve G.

T_a	T_b	PS-JCIR (%)	TC-CIR (%)	Φ-martingale (%)			
				$\eta = 10\%$	$\eta = 15\%$	$\eta = 20\%$	$\eta = 50\%$
1	3	19.55	44.00	20.24	30.11	39.91	99.46
1	5	11.90	26.56	17.31	25.64	34.00	83.86
1	7	7.33	16.30	14.82	21.79	28.74	70.59
1	10	5.72	12.27	13.64	20.06	26.40	64.52
3	5	7.79	58.67	15.06	22.40	29.84	76.67
3	7	4.48	36.10	13.24	19.60	26.03	66.34
3	10	3.47	26.62	12.36	18.32	24.36	61.59
5	7	3.11	43.32	12.40	18.39	24.55	65.34
5	10	2.49	30.61	11.58	17.24	23.02	61.29
7	10	2.47	39.40	10.24	15.33	20.61	58.56

Table 6.1: Black volatilities for at-the-money ($k = s_0(a,b)$) payer CDS options implied by the PS-JCIR, TC-CIR and the Φ-martingale models using Monte Carlo simulation (2.10^6 paths with time step 0.01) for various volatility parameter η.

Table 6.2 shows how the CDS option price evolves when increasing the strike k. We observe first that the payer CDS option decreases when the strike prices increases which is evident since a payer with higher strike price is worthless. Alternatively, the Φ-martingale model seems to be more sensitive with respect to the strike and can give different levels of price even with lower or higher strike with the help of the parameter η. In contrast, the parameters of the other models are fixed due to the calibration constraint (PS-JCIR and TC-CIR) or the positivity constraint

(PS-JCIR).

k (bps)	PS-JCIR	TC-CIR	Φ-martingale	
			$\eta = 15\%$	$\eta = 20\%$
200	148.96	152.07	176.12	198.56
220	85.64	116.54	126.99	153.89
240	39.55	92.34	88.06	117.27
260	14.08	74.18	58.67	87.68
280	3.83	60.16	37.84	64.50
300	0.79	49.26	23.52	46.95

Table 6.2: European payer (bps) with maturity $T_a = 1$ year to enter into a single-name CDS (Ford Inc.) with $T_b = 5$ year maturity with different strikes implied by the PS-JCIR, TC-CIR and Φ-martingale models using Monte Carlo method with 10^6 paths and time step 0.01. The forward spread is $s_0(a, b) = 238$ bps and the volatility parameter of the Φ-martingale model is $\eta = 0.15$.

6.4 Conclusions

The most popular class of credit risk models is undoubtedly the set of Cox models with stochastic default intensities governed by positive dynamics such as CIR or JCIR. If the intensity dynamics are simple enough (like time-homoegeneous square-root diffusions), these models allow for closed form solutions for the prices of defaultable bonds or even options, and are arbitrage-free by construction since the immersion property is satisfied. However, this simplicity comes at the price of drawbacks that are manifest in actual credit risk applications. The first one is that, given the few number of parameters at hand, such models are not flexible enough to allow a good fit to the prices prescribed by the market. To circumvent the calibration issues, the deterministic shift extension, leading in particular to CIR++ and JCIR++ models, is a very good alternative but is often problematic in practice as it features negative intensities. Adding a non-negativity constraint is a simple fix, but which often drastically limits the model's volatility. A time-change version seems to help in this respect, as it allows to get a perfect fit while generating volatilities being somewhat larger. Second, a modeling framework that satisfies the immersion property is quite a specific configuration, corresponding to a decreasing Azéma supermartingale.
In this chapter, we have developed a defaultable term structure model that does not fit in the class of Cox models. The conic martingale approach seems a promising alternative to the standard stochastic default intensity model in this respect. It

fills the gap of negative intensities since all survival probabilities are bounded and belong to $[0,1]$ by construction. Moreover, although it goes beyond immersion, we have shown that a particular case of conic martingales models, the Φ-martingale approach, is a special case of the dynamic Gaussian copula (DGC) introduced by Crépey and Song, 2014, and therefore is arbitrage-free. Furthermore, this allows us to identify an explicit construction scheme for the default time. Eventually, the model is sparse, and can exhibit a large volatility impact, which is interesting for option and counterparty credit risk pricing. These interesting features have been illustrated in two examples: the valuation adjustment under counterparty credit (CVA) risk and the pricing of CDS option. For all these reasons, the Φ-martingale model seems to be an appealing trade-off between theory and practice and provides an interesting tool for credit risk practitioners such as CVA traders and risk managers.

Dealing with arbitrage for a general conic martingales model seems to be less evident. Nevertheless, additional conditions on the score function ψ show that conic martingales belong to the density models of El Karoui et al., 2010 which are examples of beyond immersion models satisfying the $\mathcal{H}'$-hypothesis (i.e., the martingales on the default-free filtration are semimartingales in the full filtration). Future research will investigate the special case of the positive density hypothesis, a sufficient condition ensuring the no-arbitrage property in this class of models.

Appendices

6.A Exact scheme of S via Z

In this section, we show that is possible to simulate exactly the survival process S with the Φ-martingale model in contrast with the other models considered in this chapter that require a full truncation scheme such as Euler to be simulated. Let's consider (6.12) by setting the diffusion coefficient $\eta(t) = \eta$ to be constant to simplify the exposition. One gets

$$Z_t := Z_{t,t} = \Phi^{-1}(G(t))e^{\frac{\eta^2}{2}t} + \eta \int_0^t e^{\frac{\eta^2}{2}(t-s)} dB_s \, . \tag{6.26}$$

Using (6.26) and considering a time step Δ, we can express the exact distribution of $Z_{t+\Delta}$ in term of Z_t:

$$\begin{aligned}
Z_{t+\Delta} &= e^{\frac{\eta^2}{2}(t+\Delta)} \left((\Phi^{-1}(G(t+\Delta)) + \eta \int_0^t e^{-\frac{\eta^2}{2}s} dB_s + \eta \int_t^{t+\Delta} e^{-\frac{\eta^2}{2}s} dB_s \right) \\
&= e^{\frac{\eta^2}{2}(t+\Delta)} \left(\Phi^{-1}(G(t+\Delta)) - \Phi^{-1}(G(t)) + Z_t e^{-\frac{\eta^2}{2}t} + \eta \int_t^{t+\Delta} e^{-\frac{\eta^2}{2}s} dB_s \right) \\
&\sim Z_t e^{\frac{\eta^2}{2}\Delta} + \left[\Phi^{-1}(G(t+dt)) - \Phi^{-1}(G(t)) \right] e^{\frac{\eta^2}{2}(t+\Delta)} + \sqrt{e^{\eta^2\Delta} - 1}\, Y \, .
\end{aligned}$$

where $Y \sim \mathcal{N}(0,1)$ independent from Z_t.
From the latter, we can deduce the exact simulation of S by setting $S_t = \Phi(Z_t)$. In addition, the survival process $S_t(T)$ the can also be derived from S_t. Indeed, using (6.26) again,

$$\begin{aligned}
Z_{t,T} &= \Phi^{-1}(G(T))e^{\frac{\eta^2}{2}t} + \eta \int_0^t e^{\frac{\eta^2}{2}(t-s)} dB_s \\
&= Z_t + \left[\Phi^{-1}(G(T)) - \Phi^{-1}(G(t)) \right] e^{\frac{\eta^2}{2}t}
\end{aligned} \tag{6.27}$$

so that

$$S_t(T) = \Phi\left(\Phi^{-1}(S_t) + \left[\Phi^{-1}(G(T)) - \Phi^{-1}(G(t)) \right] e^{\frac{\eta^2}{2}t} \right)$$

and finally

$$\overline{Q}_t(T) = \frac{\Phi\left(\Phi^{-1}(S_t) + \left[\Phi^{-1}(G(T)) - \Phi^{-1}(G(t)) \right] e^{\frac{\eta^2}{2}t} \right)}{S_t} \, .$$

Chapter 7

Conclusion

Credit risk is at the heart of the bank's activity, whose primary role is to finance a variety of clients while taking controlled risks. The apprehension of credit risk has profoundly changed in recent years characterized by financial crises that have shown, in particular, the importance of considering the default risk. We addressed credit risk modelling across multiple angles.

First, we provided a methodological contribution to deal with the CVA computation in Chapter 3 by introducing a new intensity model relying on the time change approach of Mendoza-Arriaga and Linetsky. As the chosen stochastic clock features jumps, the time-changed intensity process can include negative and positive jumps, while remaining positive. By doing this, we resolved the tension between the necessity to fit the CDS curve, the constraints on the jump process (standards jump diffusions are constrained to feature negative jumps when positive dynamics are required), and the attainable volatility of survival probabilities. This allows greater flexibility in the modelling of volatility while sticking to the calibration constraint. In order to allow a perfect calibration, the model were implemented by adding a deterministic shift function to the time-changed intensity dynamics. The wrong-way risk were generated by correlating the Brownians governing the exposure and the default intensity. However, time-changing the intensity process introduces a time lag between intensity and exposure, thereby destroying the correlation between the two processes. It may also introduce a forward looking effect leading to potential arbitrage opportunities. We solved this issue by properly synchronising the Brownian motions driving the two processes. We then derived the expression for the CVA and showed, by means of Monte Carlo simulations, that our model has a greater wrong-way risk impact. Eventually, to deal with the computational issues, we further present a methodological improvement of the Monte

Carlo simulation that consists in using an adaptive control variate. However, even appealing, this model partially solve the perfect calibration problem since it does not preserve the range of the intensity process because of the shift function.

In order to solve this problem, that is, to perfectly fit a given term structure without facing the negativity issue of the intensity process, an alternative solution is proposed in Chapter 4. The solution consisted of time-changing, in a deterministic and continuous way, a time-homogeneous diffusion model. Hence, by speeding up or slowing down the underlying deterministic clock, one makes sure that one can exactly fit a range of zero-coupon bond curves (including all the decreasing survival probability curves). The positiveness of the resulting time-changed intensity process is guaranteed by construction. In the context of credit risk modelling, this allowed us to address, on the one hand, the perfect fit problem of homogeneous non-negative affine term structure models and negativity issues of the the deterministic shift approach, and on the other hand, to improve the volatility process generated by such affine models. The model's performances were tested using real market data based on Ford Inc. CDS term structure. When fitting Ford's survival probability curve, the time change model provided a perfect fit with an increasing clock function, while the shift approach yielded negative values in this case. A variance analysis has shown that the volatility featured in the two models are comparable. Adding a positive constraint to the shift function drastically limits the volatility level of the resulting shifted intensity process which is almost deterministic and close to zero. In addition, the numerical experiments revealed that the time change model outperformed the shift approach in terms of high CDS implied volatilities and wrong-way risk impact.

Third, to account for an economic explanation of the default event, we introduced a structural model in Chapter 5. In this context, the default time is modelled as the first passage time of the firm's assets below a pre-specified barrier. In practice however, the firm's value is not observable and investors face incomplete information regarding the firm's assets. We avoided this problem by modelling the default time trough a structural model under partial information. In our setting, the investors have no access to the exact value of the firm triggering the default but instead, have a noisy estimation. This was accounted for by observing another process correlated to the actual firm's value. Furthermore, we considered the fact that investors can also observe the default event as in reduced-form models. This setup is a hybrid type of model unifying the structural and reduced-form approaches. Similar approaches have already been studied in the credit risk literature, but they assumed simpler forms of the firm process in order to compute the conditional survival probabilities. We proposed a numerical procedure based on recursive quantization method and stochastic filtering to estimate the conditional survival probabilities for a general diffusion of the firm value process. A numerical illustration of the model showed a path dependent feature of default probabilities that keep the memory of all the path of the observed process. This path-dependent feature is very important as it relates past information in the current prices, and is not observed in a Markov setting as in standard structural models. An example of application based on CDS option pricing were provided.

Finally, we considered a new class of default models in a beyond immersion setting that are arbitrage free in Chapter 6: the Φ-martingale. This class of models belongs to a wider class of default models called *conic martingale*. The main idea introduced in Jeanblanc and Vrins, 2018 is to provide a direct modelling of the conditional survival probability up to some maturity as a martingale evolving in $[0, 1]$ and then deduce the dynamics of the Azéma supermartingale. This setup is more convenient to deal with credit risk modelling compared to a Cox setup. This is because in a Cox model, the Azéma supermartingale has no martingale part while in a beyond immersion setup, this martingale part is preserved, thus providing a noisier source of default. In a Cox setup, the no-arbitrage property is guaranteed by the immersion property while in a conic martingale model, we need additional tools in order to prove the no-arbitrage condition since this model is a case of "no-immersion". By relying on existing results in the current credit risk literature, we showed that the Φ-martingale default model is a case of "no-immersion" that is arbitrage free. Furthermore, this model is proven to feature nice statistical properties and benefits from automatic calibration. It was applied to the pricing of CDS option and CVA computation. As a result, the model is parsimonious (because can provide different levels of prices with only one parameter) and featured higher implied volatilities and wrong-way impact compared to the chosen benchmark models (that belong to the Cox setup).

This current work can be extended in several routes. Among them, we mention three as possible future research areas. A first promising area is for example to extend the CVA formulas to a more general setting by considering a total valuation adjustment commonly referred to as the "XVA". Second, our default models can be applied for modelling the interbank risk via a proper estimation of the basis spread due to discrepancies between interbank rates. The modelling of the prepayment rates for a portfolio of mortgages might also be an interesting research perspective.

Finally, the Φ-martingale belongs to the class of the density models of El Karoui et al., 2010 which is a more general setup compared to the intensity process. The intensity process allows for a knowledge of the conditional distribution of the default only "before the default". This lack of information is crucial while working in a multi-default setting. It will be important to compute the prices of a portfolio derivative on the disjoint sets before the first default, after the first default and before the second default and so on. The density approach is suitable in this after-default study and explains why the intensity approach is inadequate for this case. Therefore, another possible research area is to study the multiple defaults case when using the Φ-martingale as a default model.

Bibliography

Aksamit, A. and M. Jeanblanc (2017). *Enlargement of Filtrations with Finance in View*. Springer.

Aksamit, A., T. CHOULLI, J. DENG, and M. Jeanblanc (2014). “Arbitrages in a Progressive Enlargement Setting”. *Arbitrage, Credit and Informational Risks.* Ed. by Jeanblanc Hillairet and Jiao. Peking University Series in Mathematics, pp. 53–86.

Alfonsi, A. (2005). *On the discretization schemes for the CIR (and other Bessel squared) processes.* Tech. rep. CERMICS (Université Marne-la-Vallée).

Applebaum, D. (2009). *Lévy processes and stochastic calculus, second edition.* Cambridge University Press.

Ballotta, L. and G. Fusai (2015). “Counterparty credit risk in a multivariate structural model with jumps”. *Finance, Presses universitaires de Grenoble* 36.1, pp. 39–79.

Ballotta, L., G. Fusai, and D. Marazzina (2019). “Integrated structural approach to Credit Value Adjustment”. *European Journal of Operational Research, Elsevier* 272.3, pp. 1143–1157.

Bally, V. and G. Pagès (2003). “Error analysis of the quantization algorithm for obstacle problems”. *Stoch. Process. Appl.* 106.1, pp. 1–40.

Bally, V., G. Pagès, and J. Printems (2001). “A stochastic quantization method for non-linear problems”. *Monte Carlo Methods Appl.* 7.1, pp. 21–34.

Bielecki, T. and M. Rutkowski (2002). *Credit risk : modeling, valuation and hedging.* Springer finance. Springer.

Bielecki, T., M. Jeanblanc, and M. Rutkowski (2008). “Pricing and trading credit default swaps in a hazard process model”. *The Annals of Applied Probability* 18.6, pp. 2495–2529.

– (2011). *Credit Risk Modeling.* Tech. rep. Center for the Study of Finance and Insurance, Osaka University, Osaka (Japan).

Black, F. and J. C. Cox (1976). “Valuing corporate securities: Some effects of bonds indenture provisions”. *Journal of Finance* 31, pp. 351–367.

Black, F. and M. Scholes (1973). “The pricing of Options and Corporate Liabilities”. *Journal of Political Economy* 81, pp. 637–659.

Blanchet-Scalliet, C. and M. Jeanblanc (2004). “Hazard rate for credit risk and hedging defaultable contingent claims”. *Finance and Stochastics* 8, pp. 145–159.

Brémaud, P. and M. Yor (1978). “Changes of filtrations and probability measures”. *Z. Wahrsh. Verw. Gebiete. MR0511775* 45, pp. 269–295.

Breton, M. and O. Marzouk (2018). “Evaluation of counterparty risk for derivatives with early-exercise features”. *Journal of Economic Dynamics and Control* 88, pp. 1–20.

Brigo, D. and A. Alfonsi (2005). “Credit Default Swaps Calibration and Option Pricing with the SSRD Stochastic Intensity and Interest Rate Model”. *Finance and Stochastics* 9, pp. 29–42.

Brigo, D. and L. Cousot (2006). “A Comparison Between the SSRD model and the Market Model for CDS Options Pricing”. *International Journal of Theoretical and Applied Finance* 9.3.

Brigo, D. and N. El-Bachir (2010). “An Exact Formula For Default Swaptions Pricing in the SSRJD Stochastic Intensity Model”. *Mathematical Finance* 20.3, pp. 365–382.

Brigo, D. and F. Merccurio (2001). “A Deterministic-Shift Extension to Analytically-Tractable and Time-Homogeneous Short-Rate Models”. *Finance and Stochastics* 5, pp. 369–388.

Brigo, D. and F. Mercurio (2006). *Interest Rate Models - Theory and Practice.* Springer.

Brigo, D. and M. Morini (2005). “CDS Market Formulas and Models”. *Working paoer, Banca IMI.*

– (2006). “Structural credit calibration”. *Risk*, pp. 78–83.

Brigo, D. and M. Tarenghi (2004). “Credit default swap calibration and equity swap valuation under counterpaty risk with tractable structural model”. *In Proceedings of FEA 2004 conference at MIT.*

– (2005). “Credit default swap calibration and counterparty risk valuation with a scenario-based first passage model”. *In Proceedings of the Counterparty Credit Risk 2005 C.R.E.D.I.T. conference, Venice, 22-3 Sptember* 1.

Brigo, D. and F. Vrins (2018). “Disentangling Wrong-Way Risk: Pricing CVA via change of measures and drift adjustment”. *European Journal of Operational Research* 269, pp. 254–264.

Brigo, D., M. Morini, and A. Pallavicini (2013). *Counterparty Credit Risk, Collateral and Funding*. Wiley.

Brigo, D., A. Capponi, and A. Pallavicini (2014). "Arbitrage-free bilateral counterparty risk valuation under collateralization and application to Credit Default Swaps". *Mathematical Finance* 24.1, pp. 125–146.

Bucklew, J. A. and G. L. Wise (1982). "Multidimensional asymptotic quantization theory with r-th power distribution measures". *IEEE Trans. Inform. Theory* 28, pp. 239–247.

Callegaro, G. and A. Sagna (2013). "An application to credit risk of optimal quantization methods for nonlinear filtering". *The Journal of Computational Finance* 16.4, pp. 123–156.

Carr, P. (2017). "Bounded Brownian Motion". *Risks* 5.4.

Carr, P. and D. Madan (1999). "Option valuation using the fast Fourier transform". *Journal of Computational Finance* 2.4, pp. 61–73.

Carr, P., H. Geman, D. Madan, and M. Yor (2003). "Stochastic volatility for Lévy processes". *Mathematical Finance* 13, 345–382.

Cepedes, J., J. de Juan Herrero, D. Rosen, and Saunders D. (2010). "Effective Modeling of Wrong Way Risk, Counterparty Credit Risk Capital and Alpha in Basel II". *Journal of Risk Model Validation* 4.1, pp. 71–98.

Cesari, G., J. Aquilina, N. Charpillon, Z. Filipovic, G. Lee, and I. Manda (2009). *Modelling, pricing and hedging counterparty credit exposure - A Technical Guide*. Springer.

Chen, N. and V. G. Kou (2009). "Credit Spreads, Optimal Capital Structure, and Implied Volatility with Endogenous Default and Jump Risk". *Mathematical Finance* 19.3, pp. 343–378.

Chiarelli, C., V. Fanelli, and S. Musti (2011). "Modelling the evolution of credit spreads using the Cox process within the HJM framework: A CDS option pricing model". *European Journal of Operational Research* 208.2, pp. 95–108.

Coculescu, D., H. Geman, and M. Jeanblanc (2008). "Valuation of Default Sensitive Claims under Imperfect Information". *Finance and Stochastics* 12.2, pp. 195–218.

Cox, JC., J.E. Ingersoll, and S.A. Ross (1985). "A Theory of the Term Structure of Interest Rates". *Econometrica* 53, pp. 385–407.

Crépey, S. (2015). "Bilateral Counterparty Risk under Funding Constraints (Part I & II)". *Mathematical Finance* 25.1, pp. 1–50.

Crépey, S. and S. Song (2014). *Counterparty Risk and Funding : Immersion and Beyond*. Tech. rep. https://hal.archives-ouvertes.fr/hal-00989062v2. <hal-00989062v2>.

Crépey, S., M. Jeanblanc, and D. Wu (2013). “Informationally dynamized Gaussian copula”. *International Journal of Theoretical and Applied Finance (IJTAF), World Scientofic Publishing Co. Pte. Ltd.* 16.02, pp. 1–29.

Dai, Q. and K. Singleton (2000). “Specification Analysis of Affine Term Structure Models”. *Journal of Finance* 55.5, pp. 1943–1978.

Dambis, K.E. (1965). “On the decomposition of continuous submartingales”. *Theory of Probability and its Applications* 10, pp. 4091–4100.

Dellacherie, C. and P. A. Meyer (1975). *Probabilités et Potentiel - Espaces Mesurables.* Hermann.

– (1980). *Probabilités et Potentiel - Théorie des martingales.* Hermann.

Diop, A. (2003). “Sur la discrétisation et le comportement à petit bruit de l’EDS mutlidimensionelles dont les coefficients sont à dérivées singulières”. PhD thesis. INRIA.

Dubins, L. and G. Schwartz (1965). “On continuous martingales”. *Proceedings of the National Academy of Sciences* 53, pp. 913–916.

Duffie, D. and N. Gârleanu (2001). “Risk and Valuation of Collateralized Debt Obligations”. *Financial Analysts Journal* 57.1, pp. 41–59.

Duffie, D. and R. Kan (1996). “A Yield-Factor Model of Interest Rates”. *Mathematical Finance* 6, pp. 379–406.

Duffie, D. and D. Lando (2001). “Term structure of credit spreads with incomplete accounting information”. *Econometrica* 69, pp. 633–664.

Duffie, D. and K. Singleton (1999). “Modeling term structures of defaultable bonds”. *Review of Financial Studies* 12, pp. 687–720.

– (2003). “Credit risk: pricing, measurement, and management”. *Princeton University Press.*

Duffie, D., D. Filipovic, and W. Schachermayer (2003). “Affine Processes and Applications in Finance”. *Annals of Applied Probability* 13.3, pp. 984–1056.

Dufresne, D. (2001). *The Integrated Square-Root Process.* Tech. rep. Research Paper 90 Center for Actuarial Studies Department of Economics University of Melbourne.

Dybvig, P.H. (1997). “Bond and Bond Option Pricing Based on the Current Term Structure”. *Mathematics of Derivatives Securities (pp. 271-293).* Ed. by M. Dempster and S. Pliska. Cambridge University Press.

El Karoui, N., M. Jeanblanc, and Y. Jiao (2010). “What happens after a default: The conditional density approach”. *Stochastic Processes and their Applications* 120, pp. 1011–1032.

Elliott, R.J., M. Jeanblanc, and M. Yor (2000). “On models of default risk”. *Math. Finance* 10, pp. 179–196.

Fanelli, V. (2016). “A defaultable HJM modelling of the Libor rate for pricing Basis Swaps after the credit crunch”. *European Journal of Operational Research* 249.1, pp. 238–244.

Filipovic, D. (2005). “Time-inhomogeneous affines processes”. *stochastic processes and their application*, pp. 639–659.

Filipovic, D., M. Larsson, and A.B. Trolle (2016). “Linear-Rational Term Structure Models”. *The Journal of Finance*, n/a–n/a.

Fleck, M. and A. Schmidt (2005). “Analysis of Basel II treatment of counterparty credit risk”. *Counterparty Credit Risk Modelling, Pykhtin, M. (ed). Risk Books, London*, pp. 229–261.

Frey, R., L. Rosler, and D. Lu (2019). “Corporate security prices in structural credit risk models with incomplete information”. *Mathematical Finance* 29, pp. 84–116.

Geman, H., D. Madan, and M. Yor (2001). “Time changes for Lévy processes”. *Mathematical Finance* 11, pp. 79–96.

Gersho, A. and R. M. Gray (1988). “Special issue on quantization”. *I-II (A. Gersho and R. M. Gray eds.) IEEE Trans. Inf. Theory* 28, pp. 139–148.

Geske, R. (1977). “The valuation of corporate liabilities as coumpound options”. *Journal of Financial and Quantitative Analysis* 12.14, pp. 541–552.

Graf, S. and H. Luschgy (2000). *Foundations of Quantization for Probability Distributions.* Lect. Notes in Math. 1730. Springer, Berlin.

Gregory, J. (2010). *Counterparty Credit Risk*. Wiley Finance.

Heath, D., R. Jarrow, and A. Morton (1992). “Bond Pricing and the Term Structure of Interest Rates”. *Econometrica* 60, pp. 77–105.

Heston, S. (1993). “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”. *The Review of Financial Studies* 6.2, pp. 327–343.

Ho, T. and S. B. Lee (1986). “Term Structure Movements and Pricing Interest Rate Contingent Claims”. *The Journal of Finance* 41.5, pp. 1011–1029.

Hull, J. and A. White (1990). “Pricing Interest Rate Derivative Securities”. *The Review of Financial Studies* 3, pp. 573–592.

– (2012). “CVA and Wrong-Way Risk”. *Financial Analysts Journal* 68.5, pp. 58–69.

Jeanblanc, M. and Y. Le Cam (2009a). “Immersion Property and Credit Risk Modelling”. *Optimality and Risk - Modern Trends in Mathematical Finance.* Ed. by F. Delbaen. Springer, pp. 99–132.

Jeanblanc, M. and Y. Le Cam (2009b). “Progressive enlargement of filtration with initial times”. *Stochastic Processes and their Applic.* 119, pp. 2523–2543.

Jeanblanc, M. and M. Rutkowski (2000). “Modeling default risk: Mathematical tools”. *Fixed Income and Credit risk modeling and Management, New York University, Stern School of business, Statistics and Operations Research Department, Workshop.*

Jeanblanc, M. and S. Song (2011). “Random times with given survival probability and their F-martingale decomposition formula”. *Stochastic Processes and their Applications* 121.6, pp. 1389–1410.

Jeanblanc, M. and F. Vrins (2018). “Conic Martingales from Stochastic Integrals”. *Mathematical Finance* 28.2, pp. 516–535.

Jeanblanc, M., M. Yor, and M. Chesney (2009). *Mathematical Methods for Financial Markets*. Berlin: Springer Verlag.

Jeulin, T. (1980). “Semi-martingales et grossissements d’une filtration”. *Lecture Note in Mathematics, Springer* 833.

Joshi, M. (2003). *The Concepts and Practice of Mathematical Finance*. Cambridge University Press.

Kim, J. and T. Leung (2016). “Pricing derivatives with counterparty risk and collateralization: A fixed point approach”. *European Journal of Operational Research* 249.2, pp. 525–539.

Kloeden, P. E. and E. Platen (1999). *Numerical Solution of Stochastic Differential Equations*. Applications of Mathematics (stochastic modelling and applied probability). Springer.

Lando, D. (1998). “On Cox Processes and Credit-Risky Securities”. *Review of Derivatives Research* 2, pp. 99–120.

– (2004). *Credit Risk Modeling: Theory and Applications.* Pinceton Series in Finance. Pinceton University Press.

Longstaff, F. A. and E. S. Schwartz (1995). “A simple approach for valuing risyk fixed and floating rate debt”. *The Journal of Finance* 50, pp. 789–819.

Luschgy, H. and G. Pagès (2006). “Functional quantization of a class of Brownian diffusions: A constructive approach”. *Stochastic Processes & Their Applications* 116, pp. 310–336.

Markit (2004). *ISDA CDS Standard Model*. Tech. rep. http://www.cdsmodel.com/cdsmodel/. Markit.

– (March 13, 2009). *The CDS Big-Bang: understanding the changes to the Global CDS Contract and North American conventions*. Tech. rep. http://www.markit.com/cds/announcements/resource/cds_big_bang.pdf/. Markit.

Mbaye, C. and F. Vrins (2018). “A Subordinated CIR Intensity Model with Application to Wrong-Way Risk CVA”. *International Journal of Theoretical and Applied Finance (IJTAF), World Scientofic Publishing Co. Pte. Ltd.* 21.07, pp. 1–22.

– (2019). *Affine term structure models: a time-changed approach with perfect fit to market curves.* Tech. rep. https://arxiv.org/abs/1903.04211. UCLouvain, Louvain Finance Center.

Mendoza-Arriaga, R. and V. Linetsky (2014). “Time-changed CIR default intensities with two-sided mean-reverting jumps”. *Annals of Applied Probability* 24.2, pp. 811–856.

Merton, R. (1974). “On the Pricing of Corporate Debt: the Risk Structure of Interest Rates”. *Journal of Finance* 29.2, pp. 449–470.

Nelson, C. and A. Siegel (1987). “Parsimonious Modelling of Yield Curves”. *Journal of Business* 60.

Pagès, G. (1998). “A space vector quantization method for numerical integration”. *J. Computational and Applied Mathematics* 89, pp. 1–38.

– (2015). “Introduction to optimal quantization for numerics”. *ESMAIM Proc. Surv.* 48, pp. 28–79.

– (2018). *Numerical Probability: An introduction with Applications to Finance.* Springer, Cham.

Pagès, G. and H. Pham (2005). “Optimal quantization methods for nonlinear filtering with discrete time observations”. *Bernoulli* 11.5, pp. 893–932.

Pagès, G. and A. Sagna (2015). “Recursive marginal quantization of the Euler scheme of a diffusion process”. *Applied Mathematical Finance.* 22.15, pp. 463–498.

– (2018). “A general weak and strong error analysis of the recursive quantization with an application to jump diffusions”. *preprint.*

Pötzelberger, K. and L. Wang (2001). “Boundary crossing probability for Brownian motion”. *Journal of applied probability* 38, pp. 152–164.

Profeta, C. and A. Sagna (2014). “Conditional hitting time estimation in a nonlinear filtering model by the Brownian bridge method”. *Stochastics: An International Journal of Probability and Stochastic Processes.*

Rosen, D. and D. Saunders (2012). “CVA the Wrong Way”. *Journal of Risk Management in Financial Institutions* 5, pp. 252–272.

Sagna, A. (2012). “Pricing of barrier options by marginal functional quantization method”. *Monte Carlo Methods and Applications* 17(4), pp. 371–398.

Sagna, A., C. Mbaye, and F. Vrins (2019). *Conditional survival probabilities under partial information: a recursive quantization approach with applications*. Tech. rep. https://arxiv.org/abs/1909.01970.

Scherer, M. and A. Huttner (2016). “A note on the valuation of CDS options and extension risk in a structural model with jumps”. *International Journal of Financial Ingineering* 03.02.

Schonbucher, P. (1988). “Term structure modeling of defaultable bonds”. *Review of Derivatives Research* 2, pp. 161–192.

– (1996). *The term structure of defaultable bond prices*. Tech. rep. Discussion paper, No. B-384, Bonn University.

– (2003). *A Note on Survival Measures and the Pricing of Options on Credit Default Swaps*. Tech. rep. Bonn University, Statistics Department.

Schoutens, W. and J. Cariboni (2009). *Lévy Processes in Credit Risk*. A John Wiley and Sons Ltd.

Sellami, A. (2008). “Comparative survey on nonlinear filtering methods: the quantization and the particle filtering approaches”. *Journal of Statistical Computation and Simulation* 78(2), pp. 93–113.

Stein, H. and K. Pong (2011). “Counterparty Valuation Adjustments”. *Credit Risk Frontiers: Subprime crisis, Pricing and Hedging, CVA, MBS, Ratings and Liquidity*. Ed. by Brigo Bielecki and Patras. Wiley/Bloomberg Press. Chap. 15.

Swishchuk, A. (2016). *Change of Time Methods in Quantitative Finance*. SpringerBriefs in Mathematics. Springer.

Vasicek, O. (1977). “An Equilibrium Characterization of the Term Structure”. *Journal of Financial Economics* 5, pp. 177–188.

Veronesi, P. (2010). *Fixed Income Analysis: Valuation, Risk, and Risk Management*. Wiley.

Vrins, F. (2014). “Conic Martingales”. *8th World Congress of the Bachelier Finance Society*. Brussels (Belgium).

– (2017). “Wrong-Way Risk CVA Models with Analytical EPE Profiles under Gaussian Exposure Dynamics”. *International Journal of Theoretical and Applied Finance* 20.7, p. 1750045.

– (2018). “Sampling the multivariate normal distribution under a weighted sum constraint”. *Risks* 6.3.

Vrins, F. and M. Jeanblanc (2015). “The Phi-martingale”. *International Colloquium of Actuarial Association*. Oslo (Norway).

Zador, P. (1982). “Asymptotic quantization error of continuous signals and the quantization dimension”. *IEEE Trans. Inform. Theory* 28, pp. 139–149.

Zhou, C. (2001). "The term structure of credit spreads with jump risk". *Journal of Banking & Finance* 25.11, pp. 2015–2040.