Filtering and Incomplete Information in Credit Risk

Rüdiger Frey[1] and Thorsten Schmidt[2]

March 1, 2010

Abstract

This chapter studies structural and reduced-form credit risk models under incomplete information using techniques from stochastic filtering. We start with a brief introduction to stochastic filtering. Next we cover the pricing of corporate securities (debt and equity) in structural models under partial information. Furthermore we study the construction of a dynamic reduced-form credit risk model via the innovations approach to nonlinear filtering, and we discuss pricing, calibration and hedging in that context. The paper closes with a number of numerical case studies related to model calibration and the pricing of credit index options.

Key words: Credit risk, credit derivatives, filtering, incomplete information, structural models, intensity-based models.

Contents

[1]Department of mathematics, Universität Leipzig, mail: ruediger.frey@math.uni-leipzig.de

[2]TU Chemnitz, mail: thorsten.schmidt@mathematik.tu-chemnitz.de

1 Introduction

The recent turmoil in credit markets highlights the need for a sound methodology for managing books of credit derivatives. In particular, it is commonly agreed that the pricing and the risk management of relatively liquid products such as corporate bonds, CDSs, or index-based derivatives (e.g. synthetic CDO tranches) should be based on *dynamic* credit risk models. The development of such models is however a challenging problem: a successful model needs to capture the dynamic evolution of credit spreads and the dependence structure of defaults in a realistic way, while being at the same time tractable and parsimonious.

In this chapter we show that incomplete information and filtering techniques are a very useful tool in this regard. We begin with a short introduction to stochastic filtering in Section 2. In Section 3 we give an overview of the application of filtering to credit risk models. Next we discuss in detail structural models with incomplete information on the asset value. In particular, we explain that the pricing of many corporate securities naturally leads to a nonlinear filtering problem; this problem is then solved by a Markov chain approximation. This part of the exposition (Section 4 and 5) is based on the seminal paper by Duffie and Lando [17] and on our own work [27].

Sections 6 and 7 are devoted to filtering in reduced-form models. We present in detail the construction of a dynamic credit portfolio model via the innovations approach to nonlinear filtering. Moreover, pricing of credit derivatives, model calibration, hedging and various aspects of the numerical implementation of the model are considered. A number of numerical case studies in Section 7 illustrate practical aspects of the model. In particular, we discuss the performance of calibration strategies and present new results on the pricing of credit index options. This part of the chapter largely follows our own paper [28].

A survey on nonlinear filtering in interest-rate and credit risk models with a focus default-free term structure models can be found in [25]. Further related literature is discussed in the body of the paper.

2 A Short Introduction to Stochastic Filtering

Factor models are frequently employed in financial mathematics, since they lead to fairly parsimonious models. Stochastic filtering comes into play when these factors are observed only indirectly, possibly because they are hidden in additional noise. We present a small introduction to filtering which is inspired by [14]; for a detailed exposition we refer to [2]. We start with a small example.

Example 2.1. Consider a normally distributed random variable $X \sim \mathcal{N}(0, \sigma^2)$ (the so-called signal). Assume that X cannot be observed directly but with additional noise; that is an analyst of the system observes the sequence $Y = (Y_1, Y_2, \dots, Y_n)$, where

$$Y_i = X + W_i, \tag{1}$$

and $W_1, W_2, \dots, W_n$ are i.i.d. with $W_i \sim \mathcal{N}(0, s^2)$, independent of X. The natural estimate of X given the observation Y is the conditional expectation $\mathbb{E}(X|Y_1, \dots, Y_n)$. As X and Y are jointly normal, X can be decomposed in the following way:

$$X = a_1 Y_1 + \dots + a_n Y_n + \xi \tag{2}$$

where ξ is normally distributed and independent of Y. The coefficients a_i can be computed as follows. Consider for simplicity the case $n = 2$. Then

$$\mathrm{Cov}(X, Y_1) = \mathrm{Cov}(X, X + W_1) = \mathrm{Var}(X) = \sigma^2 = \mathrm{Cov}(X, Y_2).$$

On the other hand

$$\begin{aligned} \mathrm{Cov}(X, Y_1) &= \mathrm{Cov}(a_1 Y_1 + a_2 Y_2 + \xi, Y_1) = a_1 \mathrm{Var}(Y_1) + a_2 \mathrm{Var}(X) \\ &= a_1(\sigma^2 + s^2) + a_2 \sigma^2 \end{aligned}$$

and similarly $\mathrm{Cov}(X, Y_2) = a_1 \sigma^2 + a_2(\sigma^2 + s^2)$. This gives two linear equations for a_1 and a_2 and we obtain $a_1 = a_2 = \sigma^2/(2\sigma^2 + s^2)$. Hence, the estimate of X turns out to be

$$\mathbb{E}(X \mid Y_1, Y_2) = a_1 Y_1 + a_2 Y_2 = \frac{\sigma^2}{2\sigma^2 + s^2}\big(Y_1 + Y_2\big).$$

More generally, the estimate of X for arbitrary n is given by

$$\mathbb{E}(X \mid Y_1, \dots, Y_n) = \frac{\sigma^2}{n\sigma^2 + s^2} \sum_{i=1}^{n} Y_i. \tag{3}$$

The conditional variance $\mathrm{Var}(X \mid Y_1, \dots, Y_n)$ is given by the variance of ξ in (2) and computes to $(\sigma^2 s^2)/(n\sigma^2 + s^2)$.

Stochastic filtering. Generalizing this example, stochastic filtering is concerned with the following problem: consider a set of time points $\mathcal{T}$; in the discrete setting typically $\mathcal{T} = \{1, 2, \dots\}$ or in continuous-time $\mathcal{T} = [0, \infty)$. The unobservable variable of interest X is called *signal* or *state process*. It is a stochastic process $X = (X_t)_{t \in \mathcal{T}}$. The observation is given by the *observation process* $Y = (Y_t)_{t \in \mathcal{T}}$, and we denote by $\mathcal{F}_t^Y = \sigma(Y_s : s \le t)$ the information generated by the observation until time t.

In filtering one wants to estimate X based on the observation of Y. A major goal is to describe the conditional distribution of X_t given the $\mathcal{F}_t^Y$. The conditional distribution can be computed if one knows

$$\mathbb{E}\big(\phi(X_t)|\mathcal{F}_t^Y\big) \tag{4}$$

for a reasonably large class of functions ϕ. For computational reasons it is important to obtain this expression in a recursive way. In the sequel we describe several standard filtering problems and the corresponding recursive algorithms for evaluating the conditional expectation (4) in each of the models.

2.1 The Kalman-Bucy filter in discrete time

The Kalman-Bucy-filter is the simplest practically relevant case where the filtering problem has an explicit solution. The setup is an extension of Example 2.1. As before X stands for the unobserved factor process, while Y represents the observation process. Consider the following model on $\mathcal{T} = \{1, 2, \dots\}$:

$$\begin{aligned} X_t &= a(t)X_{t-1} + b(t)W_t \\ Y_t &= c(t)X_t + d(t)V_t \end{aligned} \tag{5}$$

where $W = (W_1, W_2, \dots)$ and $V = (V_1, V_2, \dots)$ are sequences of independent, standard normally distributed random variables and where the distribution of X_0 is given. This setup has the following interpretation: the factor process X evolves through a stochastic difference equation (note that (1) is the special case with $a(t) = 1$ and $b(t)$=0). The observation Y contains $c(t)X_t$ plus additive noise. For simplicity, we assume that a, b, c and d are deterministic, real-valued functions[3]. It iss well-known that in this case the conditional distribution of X_t given $\mathcal{F}_t^Y$ is normal, so that it suffices to determine the mean and the variance of this distribution.

The Kalman-Bucy-filter is a recursive procedure for computing the conditional mean and variance of X. It works in two steps: Assume that until time t

$$X_{t|t} := \mathbb{E}\big(X_t|\mathcal{F}_t^Y\big)$$

has been computed. A first observation is that from t to $t+1$ the process X evolves according to (5). Taking this into account, one computes the *prediction*

$$X_{t+1|t} := \mathbb{E}\big(X_{t+1}|\mathcal{F}_t^Y\big) = a(t)X_{t|t}.$$

The next step incorporates the new observation at $t + 1$ given by Y_{t+1}. A part of Y_{t+1}, namely $X_{t+1|t}$, can be predicted on the basis of the information available at time t, so that the *innovation* (the part of Y_{t+1} which actually carries new information) is given by $Y_{t+1} - X_{t+1|t}$. Only the innovation therefore matters for the updating to $X_{t+1|t+1}$. It can be shown that $X_{t+1|t+1}$ is given by a recursive updating rule of the form

$$X_{t+1|t+1} = X_{t+1|t} + L_t(Y_{t+1} - X_{t+1|t}) \tag{6}$$

where

$$L_t = \frac{c(t)P_{t+1|t}}{c(t)^2 P_{t+1|t} + d(t)^2},$$

$$P_{t+1|t} = a(t)^2 P_{t|t} + b(t)^2, \text{ and } P_{t+1|t+1} = P_{t+1|t}(1 - L_t c(t)).$$

Here P is the conditional variance, that is $P_{s|t} = \mathbb{E}\big((X_s - X_{s|t})^2|\mathcal{F}_t^Y\big)$. Note that P is a deterministic function of time and independent of the particular realization of the observation process Y.

The Kalman-Bucy has also been applied to Gaussian models which are nonlinear by linearizing the nonlinear coefficient functions around the current estimate of X_t. This procedure is called *extended Kalman filter*. Kalman filtering is frequently employed in the empirical analysis of swap and credit spreads, see for instance [21].

[3]The Kalman-Bucy filter can be generalised to the multi-dimensional case and $a, \dots, d$ may be adapted with respect to the filtration $\mathcal{F}_t^Y = \sigma(Y_s : s \le t)$.

2.2 Filtering in continuous time

The standard continuous-time filtering problem is of the form

$$\begin{aligned} dX_t &= a(t, X_t)dt + b(t, X_t)dW_t \\ dY_t &= c(t, X_t)dt + dV_t. \end{aligned} \tag{7}$$

for independent Brownian motions W and V. The model (7) can be viewed as continuous-time version of a discrete-time filtering setup such as (5), as is illustrated in the following example.

Example 2.2. Translating (5) to more general time points $t_k := k\Delta$, we consider the observation $\tilde{c}(t_k)X_{t_k} + \varepsilon_k$ for $\varepsilon_k \sim N(0, \Delta)$ and $\tilde{c}(\cdot) = \Delta c(\cdot)$. In continuous-time one considers instead the *cumulative observation process*

$$Y_t := \sum_{t_k \le t} \Big(\tilde{c}(t_k)X_{t_k} + \varepsilon_k \Big).$$

Then one has for Δ small:

$$Y_t \approx \int_0^t c(s)X_s ds + V_t. \tag{8}$$

The generalisation of this equation leads to (7).

In the innovations approach to nonlinear filtering the conditional distribution of X_t given $\mathcal{F}_t^Y$ in the model (7) is characterized by a stochastic differential equation (SDE) as follows. First, denote by $\mathscr{L}$ the generator of the Markovian diffusion X:

$$\mathscr{L}\phi(t, x) = \phi_t a(t, x) + \phi_x(t, x) + \frac{1}{2}\sum_{i,j=1}^n v_{ij}(t, x)\phi_{x_i x_j}(t, x)$$

for any function $\phi(t, x) \in C^{1,2}$, where we set $v(t, x) := b(t, x)b(t, x)^\top$. Note that by the Itô formula $\phi(t, X_t) - \int_0^t \mathscr{L}\phi(t, X_s)ds$ is a (local) martingale so that locally $\mathscr{L}\phi(t, X_t)dt$ gives the expected change of the process $\phi(t, X_t)$. Denote for a generic function $f(t, x)$

$$\widehat{f}_t := \mathbb{E}\big(f(t, X_t)|\mathcal{F}_t^Y\big).$$

The innovations approach leads to the following SDE, called *Kushner-Stratonovich equation*,

$$d\widehat{\phi}_t = \widehat{(\mathscr{L}\phi)}_t dt + \big(\widehat{c\phi}_t - \widehat{c}_t\widehat{\phi}_t\big) \cdot \big(dY_t - \widehat{c}_t\, dt\big). \tag{9}$$

This equation is driven by the *innovation*

$$dY_t - \widehat{c}_t\, dt = dY_t - \mathbb{E}\big(dY_t|\mathcal{F}_t^Y\big).$$

As in the case of the Kalman-Bucy filter, the filter equation (9) contains two parts: $\widehat{\mathscr{L}\phi}_t$ represents the expected change of $\phi(X_t)$, and the second term gives the update with respect to the new information which we called innovation.

Equation (9) is in general an infinite-dimensional equation: in order to determine $\widehat{\phi}$ one needs $\widehat{c\phi}$; this in turn requires $\widehat{c^2\phi}$ and so on. A substantial part of the modern filtering literature is concerned with finding finite-dimensional approximations to this equation which can be implemented on a computer; see for instance [9] or Part II of [2].

Equation (9) remains true if the diffusion X is replaced by a general Markov process, just the generator $\mathscr{L}$ and the class of functions ϕ need to be adjusted in a proper way. For instance, in the case of a finite-state Markov chain the generator is given by the matrix of transition intensities. The corresponding filter formulas can be found in Section 2.4.

2.3 Observations as a jump process

Alternatively, the observations could be given by a doubly-stochastic Poisson process with intensity depending on the factor process X. For a concrete example, suppose that N is a standard Poisson process with intensity one, that X is a diffusion and that the observation Y is a time-changed Poisson process. Formally,

$$\begin{aligned} dX_t &= a(t, X_t)dt + b(t, X_t)dW_t \\ Y_t &= N_{\Lambda_t} \quad \text{for } \Lambda_t = \int_0^t \lambda(X_s)ds\,. \end{aligned} \tag{10}$$

In applications to credit risk the jump-process Y typically represents default events in a given reference portfolio. The Kushner-Stratonovich equation for the model (10) takes the form

$$d\widehat{\phi}_t = \widehat{\mathscr{L}\phi}_t dt + \frac{1}{\widehat{\lambda}_t}\Big(\widehat{\lambda\phi}_t - \widehat{\lambda}_t\widehat{\phi}_t\Big) \cdot \big(dY_t - \widehat{\lambda}_t\, dt\big); \tag{11}$$

see for instance [6] for a detailed derivation.

2.4 The case of Markov chains

If X is a finite-state Markov chain the Kushner-Stratonovich equation (9) reduces to a finite-dimensional SDE system. Assume w.l.o.g. that X has values $\{1, \dots, K\}$ and denote the transition intensities of X by $(q(i,j))_{1 \le i,j \le K}$. The conditional distribution of X_t given $\mathcal{F}_t^Y$ is given by the probabilities

$$\pi_t^k := \mathbb{P}(X_t = k|\mathcal{F}_t^Y).$$

From (9) one obtains, letting $\phi(x) = 1_{\{x=k\}}$, the dynamics of the conditional distribution π:

$$d\pi_t^k = \sum_{i=1}^N \pi_t^i q(i,k)dt + \pi_t^k\Big(c(t,k) - \sum_{i=1}^N \pi_t^i c(t,i)\Big) \cdot \big(dY_t - \sum_{i=1}^N \pi_t^i c(t,i)dt\big). \tag{12}$$

An illustration of the filter is given in Figure 1. Similar formulas can be given if Y follows a jump process. For further details on filtering in the case of finite-state Markov chains consider [19]. We will use the Kushner-Stratonovich equation for finite-state Markov chains in the construction of a reduced-form credit risk model in Section 6.

Markov-chain approximations can be used as computational tool for computing the filter for more general state variables; see for instance [9] for details. There are other situations where the Kushner-Stratonovich equation admits a finite-dimensional solution. The most prominent example is the continuous-time Kalman-Bucy filter; see for instance [2].

3 Credit Risk Models under Incomplete Information

In this section we explain how incomplete information (the fact that some state variables are not fully observed by investors) frequently arises in credit risk models, so that filtering techniques naturally come into play.

We start with some notation. Consider a portfolio of m firms. The default time of firm i is denoted by the random variable $\tau_i > 0$. Let $D_{t,i} = 1_{\{\tau_i \le t\}}$ denote the current default state of firm i. $D_{t,i}$ is zero if the company did not default until t and jumps to one at the default

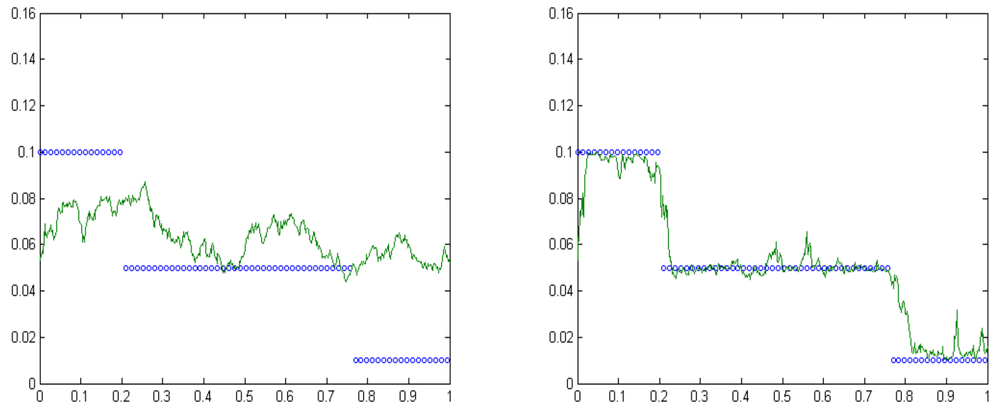

Figure 1: A simulated trajectory of the unobservable Markov chain X (circles) and the filter estimate $\widehat{X} = \mathbb{E}(X_t \mid \mathcal{F}_t^Y)$. The left figure has high observation noise σ_V while the right one has low observation noise. High noise translates to a low precision in the filter estimate and vice versa.

time. The current default state of the portfolio is $D_t = (D_{t,1}, \ldots, D_{t,m})$ and the default history up to time t is given by $\mathcal{F}_t^D := \sigma(D_s \colon s \leq t)$. The corresponding filtration is denoted by $\mathbb{F}^D$. Throughout we work on a filtered probability space $(\Omega, \mathcal{G}, \mathbb{G}, Q)$ and all stochastic processes considered will be $\mathbb{G}$-adapted. Typically Q will be the risk-neutral measure used for pricing. Moreover, we assume for simplicity that default-free interest rates are deterministic and equal to $r > 0$.

Existing dynamic credit risk models can be grouped into two classes: *structural* and *reduced-form models*. Structural models originated from Black and Scholes [4], Merton [38], and Black and Cox [3]. Important contributions to the literature on reduced-form models are [32], [33] [18] and [5] among others. Further details on credit risk models can be found in [37], [44], [45] or [34]. In structural as well as in reduced-form models it makes sense to assume that investors have imperfect information on some of the state variables of the models. A rich literature on credit risk models under incomplete information deals with this aspect.

Structural models. Here one starts by modeling the asset value A of the firm under consideration. Given some, possibly random, default barrier $K = (K_t)_{t \geq 0}$, default happens at the first time when A crosses the the barrier K, i.e.

$$\tau = \inf\{t \geq 0 \colon A_t \leq K_t\} \,. \tag{13}$$

The default barrier is often interpreted as the value of the liabilities of the firm; then default happens at the first time that the asset value of a firm is too low to cover its liabilities. If A is a diffusion, then the default time τ is a predictable stopping time with respect to the global filtration $\mathbb{G}$ to which A and K are adapted. It is well-documented that this fact leads to very low values for short-term credit spreads, contradicting most of the available empirical evidence.

The natural state variable in a structural model are thus the asset value A of the firm and, if liabilities are stochastic, the liability-level K. It is difficult for investors to assess the value of these variables. There are many reasons for this: accounting reports offer only noisy information on the asset value; market- and book values can differ as intangible assets are difficult to value; part of the liabilities are usually bank loans whose precise terms are unknown to the public, and many more. Starting with the seminal work of Duffie and Lando [17], a growing literature therefore studies models where investors have only noisy information about A and/or K; the conditional distribution of the state variables given investor information is then computed by Bayesian updating or filtering arguments. Examples of this line of research include [17], [39], [10], [43] and [27]; we discuss the works [17] and [27] in detail below. Interestingly, it turns out that the distinction between structural and reduced-form models is in fact a distinction between full and partial observability of asset values and liabilities (see [31]): in the models mentioned above the default time is predictable with respect to the global filtration $\mathbb{G}$ but becomes totally inaccessible with respect to the investor filtration $\mathbb{F}$. As a consequence, the default time admits an intensity in the investor filtration and the short-term credit spreads achieve realistic levels, as is explained in Section 4 below.

Reduced-form models. In this model class one models directly the law of the default time τ. Typically, τ is modeled as a totally inaccessible stopping time with respect to the global filtration $\mathbb{G}$, and it is assumed that τ admits a $\mathbb{G}$-intensity λ. This intensity is termed *risk-neutral default intensity* (recall that we work under the risk-neutral measure Q). Formally, $\lambda = (\lambda_t)_{t\geq 0}$ is a $\mathbb{G}$-predictable process such that

$$1_{\{\tau \leq t\}} - \int_0^{t\wedge\tau} \lambda_s \, ds \tag{14}$$

is a $\mathbb{G}$-martingale. Dependence between defaults is typically generated by assuming that the default intensities depend on a common factor process X. Denote by $\mathbb{F}^X := (\mathcal{F}_t^X)_{t\geq 0} = (\sigma(X_s : s \leq t))_{t\geq 0}$ the filtration generated by the factor X. The simplest construction is that of *conditionally independent, doubly stochastic default times*: here it is assumed that given the realization of the factor process the default times τ_i are conditionally independent with intensities $\lambda_i(X_t)$, i.e.

$$Q(\tau_1 > t_1, \ldots, \tau_m > t_m \mid \mathcal{F}_\infty^X) = \prod_{i=1}^m \exp\Big(- \int_0^{t_i} \lambda_i(X_s)\, ds \Big). \tag{15}$$

In applications X is usually treated as a latent process whose current value must be inferred from observable quantities such as prices or the default history. A theoretically consistent way for doing this is to determine via filtering the conditional distribution of X_t given investor information $\mathbb{F}$. Models of this type include the contributions by [46], [11], [16], [28] and [26]. The last two papers are discussed in Section 6.

Remark 3.1. We will see below that the introduction of incomplete information generates *information-driven default contagion* (both in structural and in reduced-form models): the news that some obligor has defaulted causes an update in the conditional distribution of the unobserved factor and hence to a jump in the $\mathbb{F}$-default intensity of the surviving firms. This in turn leads to interesting dynamics of credit spreads, compare the discussions on pages 10 and 16.

4 Structural Models I: Duffie and Lando [17]

In this section we discuss the influential paper by Duffie and Lando [17]. As previously, let Q be the risk-neutral measure and $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ be the filtration which represents full information. The asset value A is assumed to follow a geometric Brownian motion with drift μ, volatility σ and initial value A_0, such that

$$A_t = A_0 \exp\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right), \tag{16}$$

where W is a Browninan motion. The default barrier K and the inital value A_0 are taken to be constant and therefore the default time is $\tau := \inf\{t \geq 0\colon A_t \leq K\}$. It is assumed that A is not directly observable. Rather, investors observe default, and they receive *noisy accounting reports* at deterministic times $t_1, t_2, \cdots$. This is modelled by assuming they observe the random variables

$$Y_i := \ln A_{t_i} + U_i$$

at t_i, where $U_1, U_2, \ldots$ is a sequence of independent, normally distributed random variables, independent of A. Formally, with $D_t := 1_{\{\tau \leq t\}}$, the information available to investors is given by

$$\mathcal{F}_t := \sigma(D_s : s \leq t) \vee \sigma(\{Y_i : t_i \leq t\}).$$

We now study the computation of survival probabilities, default intensities and credit spreads. We start with the situation under full information. By the Markov property of A one has, for $T \geq t$,

$$\begin{aligned} Q(\tau > T \mid \mathcal{G}_t) &= 1_{\{\tau > t\}} Q\left(\inf_{s \in (t,T)} A_s > K \mid \mathcal{G}_t\right) \\ &= 1_{\{\tau > t\}} Q\left(\inf_{s \in (t,T)} A_s > K \mid A_t\right) \\ &=: 1_{\{\tau > t\}} F_\tau(t, T, A_t). \end{aligned}$$

Note that the mapping $T \mapsto F_\tau(t, T, v)$ gives the risk-neutral survival probabilities of the firm under full information at time t, given that $A_t = v$. This probability is easily computed using standard results on the first passage time of Brownian motions with drift. Using iterated conditional expectations one obtains the survival probability in the investor filtration:

$$\begin{aligned} Q(\tau > T \mid \mathcal{F}_t) &= \mathbb{E}\Big(Q(\tau > T \mid \mathcal{G}_t) \mid \mathcal{F}_t\Big) \\ &= 1_{\{\tau > t\}} \int_{\log K}^{\infty} F_\tau(t, T, v)\, \pi_{A_t|\mathcal{F}_t}(dv); \end{aligned}$$

here $\pi_{A_t|\mathcal{F}_t}$ denotes the conditional distribution of A_t given $\mathcal{F}_t$. In [17] this distribution is computed in an elementary way, involving Bayes' formula and properties of first passage time of Brownian motion. Section 5.1 shows how filtering techniques can be used in this context.

Next we turn to the *default intensity* in the model with incomplete information. It can be shown that under some regularity conditions the default intensity is given by

$$\lambda_t = \lim_{h \downarrow 0} \frac{1}{h} Q\big(t < \tau \leq t + h \mid \mathcal{F}_t\big), \tag{17}$$

provided this limit exists for all $t \geq 0$ almost surely (see [1] for details). Duffie and Lando show that such a λ_t exists and is given on $\{\tau > t\}$ by

$$\lambda_t = \frac{\sigma^2 K}{2} \frac{\partial}{\partial v} f_{A_t|\mathcal{F}_t}(K)\,, \tag{18}$$

where $f_{A_t|\mathcal{F}_t}$ denotes the Lebesgue-density of the conditional distribution of A_t given $\mathcal{F}_t$.

Bonds and credit spreads. A defaultable zero-coupon bond with zero recovery pays one unit of account at maturity T if no default happened and zero otherwise. Hence, in this setup its price $p(t,T)$ at time t equals

$$1_{\{\tau>t\}} e^{-r(T-t)} Q\big(\tau > T \mid \mathcal{F}_t\big) = 1_{\{\tau>t\}} e^{-r(T-t)} \int_{\log K}^{\infty} F_\tau(t,T,v)\, \pi_{A_t|\mathcal{F}_t}(dv).$$

Therefore zero-coupon bond prices can be expressed as an average with respect to the filter distribution. The credit spread $c(t,T)$ of the bond gives the yield over the risk-free short-rate. Formally it is given by

$$c(t,T) = -\frac{1}{T-t}\big(\log p(t,T) - \log p_0(t,T)\big), \tag{19}$$

where $p_0(t,T)$ denotes the price at time t of the risk-free zero-coupon bond with maturity T. Hence we get on $\{\tau > t\}$ that $c(t,T) = \frac{-1}{T-t} \log Q(\tau > T \mid \mathcal{F}_t)$. In particular, we obtain

$$\lim_{T\to t} c(t,T) = -\lim_{T\to t} \Big(\frac{\partial}{\partial T} Q(\tau > T \mid \mathcal{F}_t)\Big) = \lambda_t,$$

where the second equation follows from (17). This shows that the introduction of incomplete information typically leads to non-vanishing short-term credit spreads.

Other debt-related securities such as credit default swaps (CDS) can be priced in a straightforward fashion once the conditional survival function given investor information is at hand.

5 Structural models II: Frey & Schmidt [27]

In [27], the Duffie-Lando model is extended in two directions: first, nonlinear filtering techniques based on Markov-chain approximations are employed in order to determine the conditional distribution of the asset value given the investor-information; second, the paper introduces dividend payments and discusses the pricing of the firm's equity under incomplete information. We begin with a discussion of the filtering part.

5.1 The filtering part

The model. Here we present a slightly simplified version of the model discussed in [27]. Similarly as in the Duffie-Lando model the asset value A is given by the geometric Brownian motion (16), so that the log-asset value $X_t := \log A_t$ satisfies $X_t = X_0 + \big(\mu - \frac{1}{2}\sigma^2\big)\, t + \sigma W_t$. The default time τ is

$$\tau := \inf\{t \geq 0 \colon A_t \leq K\} = \inf\{t \geq 0 \colon X_t \leq \log K\}\,.$$

Investors observe the default state of the firm; moreover, they receive information related to the state of the company such as information (news) given by analysts, articles in newspapers, etc. It is assumed that this information is discrete, corresponding for instance to buy/hold/sell recommendations or rating information. Formally, news events on the company are issued at time points t_n^I, $n \geq 1$; the news obtained at t_n^I is denoted by I_n, which takes values in the discrete state space $\{\ell_1, \dots, \ell_{M^I}\}$. The conditional distribution of I_n given the asset value of the company at t_n^I is denoted by $\nu_I(I_n|x)$ where

$$\nu_I(\ell_j|x) := Q(I_n = \ell_j | X_{t_n^I} = x).$$

Summarizing, the information of investors at time t is given by

$$\mathcal{F}_t := \mathcal{F}_t^D \vee \sigma\big(I_n \colon t_n^I \leq t\big). \tag{20}$$

Filtering. In order to determine the conditional distribution $\pi_{X_t|\mathcal{F}_t}$ with minimal technical difficulties, the log-asset value process X is approximated by a finite-state discrete-time Markov chain as follows: define for a $\Delta > 0$ the grid

$$\{t_k^\Delta = k\Delta : \, k \in \mathbb{N}\}.$$

Let $(X_k^\Delta)_{k\in\mathbb{N}}$ be a discrete-time finite-state Markov chain with state space $\{m_1^\Delta, \dots, m_{M^\Delta}^\Delta\}$ and transition probabilities p_{ij}^Δ. Define the induced process X^Δ by $X_t^\Delta = X_k^\Delta$ for $t \in [t_k^\Delta, t_{k+1}^\Delta)$. In [27] it is assumed that the chain $(X_k^\Delta)_{k\in\mathbb{N}}$ is close to the continuous log-asset-value process X in the sense that X^Δ converges in distribution to X as $\Delta \to 0$; it is shown that this implies that the conditional distribution $\pi_{X_t^\Delta|\mathcal{F}_t}$ converges weakly to $\pi_{X_t|\mathcal{F}_t}$ as $\Delta \to 0$. The approximating Markov chain can be chosen to be trinomial with transition probabilities determined by matching the first and second moment of the transition probabilities of X^Δ with those of X; see [27] for details.

In the sequel we keep Δ fixed and therefore mostly omit it from the notation. The conditional distribution $\pi_{X_{t_k}^\Delta|\mathcal{F}_{t_k}}$ is summarized by the probability vector $\pi(k) = (\pi^1(k), \dots, \pi^M(k))$ with

$$\pi^j(k) := Q\big(X_k = m_j \mid \mathcal{F}_{t_k}\big).$$

The initial filter distribution $\pi(0)$ can be inferred from the initial distribution of X_0 which is a model primitive. There is a simple explicit recursive updating rule for the probability vector $\pi(k)$ as we show next. It is convenient to formulate the updating rule in terms of *unnormalized probabilities* $\sigma(k) \propto \pi(k)$ ($\propto$ standing for proportional to); the vector $\pi(k)$ can then be obtained by normalization:

$$\pi^j(k) = \frac{\sigma^j(k)}{\sum_{i=1}^M \sigma^i(k)}.$$

Proposition 5.1. *For $k \geq 1$ and $t_k < \tau$ denote by $N_k := \{n \in \mathbb{N} : t_{k-1} < t_n^I \leq t_k\}$ the set of indices of news arrivals in the period $(t_{k-1}, t_k]$, and recall that p_{ij} are the transition probabilities of the approximating Markov chain X. Then for $j = 1, \dots, M$ we have that*

$$\sigma^j(k) = 1_{\{m_j > \log K\}} \sum_{i=1}^M \Big(p_{ij}\, \sigma^i(k-1) \prod_{n\in N_k} \nu_I\big(I_n|m_i\big)\Big). \tag{21}$$

Here we use the convention that the product over a empty set is equal to one.

Formula (21) explains how to obtain $\sigma(k)$ from $\sigma(k-1)$ and the new observation received over $(t_{k-1}, t_k]$. The derivation of this formula is quite instructive. First note that given the new information arriving in $(t_{k-1}, t_k]$, (21) forms a linear and in particular a positively homogeneous mapping Γ such that $\sigma(k) = \Gamma\sigma(k-1)$. Hence it is enough to show that $\pi(k) \propto \Gamma\pi(k-1)$. In order to compute $\pi(k)$ from $\pi(k-1)$ and the new observation we proceed in two steps. In Step 1 we compute (up to proportionality) an auxiliary vector of probabilities $\tilde{\pi}(k-1)$ with

$$\tilde{\pi}^i(k-1) = Q\big(X_{k-1} = m_i \mid \mathcal{F}_k^-\big),\ 1 \le i \le M, \tag{22}$$

where $\mathcal{F}_k^- := \mathcal{F}_{t_{k-1}} \vee \sigma\big(\{I_n : n \in N_k\}\big)$. In filtering terminology this is a smoothing step as the conditional distribution of X_{k-1} is updated using the new information arriving in $(t_{k-1}, t_k]$. In Step 2 we determine (again up to proportionality) $\pi(k)$ from the auxiliary probability vector $\tilde{\pi}(k-1)$ using the dynamics of (X_k) and the additional information that $\tau > t_k$. We begin with Step 2. Since $\{\tau > t_k\} = \{\tau > t_{k-1}\} \cap \{X_k > \log K\}$, we get

$$\begin{aligned} Q\left(X_k = m_j \mid \mathcal{F}_{t_k}\right) &\propto Q\left(X_k = m_j, X_k > \log K \mid \mathcal{F}_k^-\right) \\ &= \sum_{i=1}^{M} Q\left(X_k = m_j, X_k > \log K, X_{k-1} = m_i \mid \mathcal{F}_k^-\right) \\ &= 1_{\{m_j > \log K\}} \sum_{i=1}^{M} p_{ij}\, \tilde{\pi}^i(k-1)\,. \end{aligned} \tag{23}$$

Next we turn to the smoothing step. Note that given $X_{k-1} = m_i$, the likelihood of the news observed over $(t_{k-1}, t_k]$ equals $\prod_{n \in N_k} \nu_I\big(I_n | m_i\big)$, and we obtain

$$\tilde{\pi}^i(k-1) \propto \pi^i(k-1) \cdot \prod_{n \in N_k} \nu_I\big(I_n | m_i\big).$$

Combining this with equation (23) gives (21).

5.2 Pricing the firm's equity

Next we discuss the pricing of the firm's equity or shares. This is of interest for at least two reasons: on the theoretical side this analysis sheds some light on the relation between share price and asset value in the presence of incomplete information; on the practical side this is a prerequisite for the pricing of certain hybrid corporate securities such as equity default swaps.

In [27] the pre-default value of the firm's equity S is defined as expected value of the future discounted dividends up to the default of the firm. Simplifying slightly the original setup of [27], we assume that dividends are paid at dividend dates t_n^d; the dividend paid at t_n^d is given by the random variable

$$d_n = \delta_n A_{t_n^d}, \quad \text{for } \delta_n \in [0,1) \text{ iid, independent of } A \text{ with mean } \bar{\delta}\ .$$

Formally we thus have

$$S_t = \mathbb{E}\Big(\sum_{t < t_n^d < \tau} e^{-r(t_n^d - t)} \delta_n A_{t_n^d} \mid \mathcal{F}_t \Big). \tag{24}$$

Denote by $\mathbb{E}\big(\sum_{t<t_n^d<\tau} e^{-r(t_n^d-t)}\delta_n A_{t_n^d} \mid \mathcal{G}_t\big)$ the equity value under full information. Since A is a Markov process, on $\{\tau > t\}$ the latter is given by some function $S(t, A_t)$ of time and the current equity value. Using the tower property of conditional expectations we thus get

$$S_t = \mathbb{E}\bigg(\mathbb{E}\Big(\sum_{t<t_n^d<\tau} e^{-r(t_n^d-t)}\delta_n A_{t_n^d} \mid \mathcal{G}_t\Big) \mid \mathcal{F}_t\bigg) = 1_{\{\tau>t\}}\mathbb{E}\big(S(t, A_t) \mid \mathcal{F}_t\big),$$

and the conditional distribution of the right can be computed using the approximation of the filter distribution given in Proposition 5.1.[4]

Example 5.2 (A closed-form solution for the equity value under full information)**.** A slight modification of the setup leads to a closed form expression for the function $S(\cdot)$. For this we assume that the dividend dates are the jump times of a Poisson process with intensity λ^d, corresponding to the average number of dividend dates per year. With frequent dividend payments, such as quarterly or semi-annually, the equity value obtained under the assumption of Poissonian dividend dates is a good approximation of its counterpart for fixed dividend dates. The advantage of this assumption is that the pre-default equity value becomes independent of calender time t. Proposition 2.4 of [27] states that for $\mu < r$ the full-information value of the firm's equity equals $1_{\{\tau>t\}}S(A_t)$ with

$$S(v) = \frac{\lambda^d\bar{\delta}}{r-\mu}\Big[v - \Big(\frac{v}{K}\Big)^{\alpha^*} K\Big]\,. \tag{25}$$

Here $\alpha^* < 0$ is the unique negative root of the quadratic equation

$$\alpha\mu + \frac{1}{2}\sigma^2\alpha(\alpha-1) = r.$$

Note that S is concave in v and approaches the line $v \mapsto v \cdot \frac{\lambda^d\bar{\delta}}{(r-\mu)}$ as v tends to infinity. This line corresponds to the value of the firm's equity for $K = 0$ and therefore $\tau = \infty$. The qualitative behaviour of S is illustrated in Figure 2.

5.3 Further applications

We briefly discuss further results obtained in [27].

Estimating the asset values from equity values. The filter estimate of the previous section corresponds to a fundamental valuation approach: one tries to assess the value of the firm's assets from economic information such as news. When the stock of the firm is liquidly traded, one could alternatively compute a market implied estimator of the asset value by inverting some pricing formula that relates asset- and equity value. The KMV-methodology is a typical example where this approach is used, see [13]. Formally, given the current equity value S^* observed in the market and a valuation formula under full information of the form $S_t = S(t, A_t)$, S strictly increasing in v, the corresponding *equity-implied estimator* EE_t is given by the solution of the equation $S(t, \mathrm{EE}_t) = S^*$.

In [27] it is shown that this estimator performs well if the conditional variance of A_t given the investor information $\mathcal{F}_t$ is small, that is if the the observations received by investors carry

[4] Strictly speaking, observed dividends contain information about A as well and should therefore be included in the filtering result. This can be done analogously as in Proposition 5.1; for details we refer to [27].

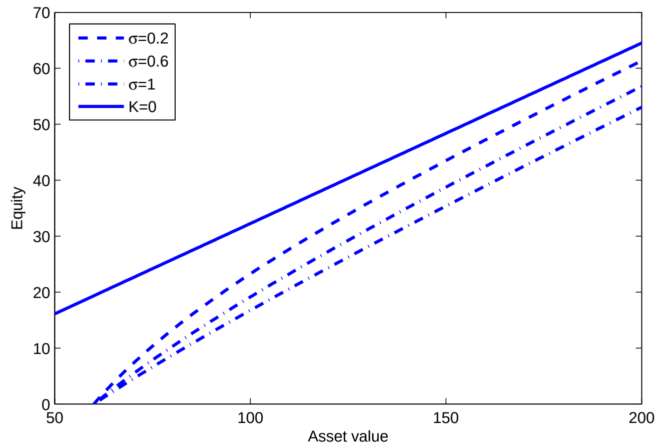

Figure 2: Value of the firm's equity as function of the asset value according to (25) for different σ and with $K = 60$. The straight line is the equity value for $K = 0$.

a lot of information about the "true value" of A_t. On the other hand the estimator performs poorly if the conditional variance of A_t is comparatively large. Moreover, it is possible to study the bias of the equity estimator via Jensen's inequality. If the function $v \mapsto S(t, v)$ is concave as in Example 5.2 one obtains that

$$\mathrm{EE}_t \leq \mathbb{E}(A_t|\mathcal{F}_t);$$

if $v \mapsto S(t, v)$ is convex, the inequality is reversed.

Equity value and default intensity. It can be shown that relation (18) for the default intensity under incomplete information in the Duffie-Lando model carries over to the setup of [27]. Given this result, in [27] the relation between equity value and default intensity λ_t is studied. It turns out that for fixed firm characteristics μ and σ a hyperbolic relation of the form $\lambda_t = h(S_t) := \frac{\alpha}{S^\rho}$ - as it is imposed in certain hybrid models such as [36] - describes the relation between stock price and default intensity well. If these characteristics vary however, the relation between default intensity and equity value breaks down completely.

Pricing of hybrid securities. As mentioned before, the model could in principle be used for the pricing and hedging of hybrid securities such as equity default swaps or convertible bonds. For this one needs to give a description of the stock price dynamics in the investor filtration, using a suitable variant of the Kushner-Stratonovich equations (9) and (11); see for instance [24]. In the next section we will show how a similar approach can be carried out in the context of reduced-form models.

6 Constructing Reduced-form Credit Risk Models via Nonlinear Filtering

Now we turn our attention to reduced form models under incomplete information. We discuss in detail our own model proposed in [28]. A key idea in that paper is to use the innovations approach to nonlinear filtering in order to derive the price dynamics of credit derivatives. We will show that this leads to a fairly tractable model with rich dynamics of credit spreads and a lot of flexibility for calibration.

6.1 The Setup

We consider a portfolio that contains credit derivatives on m firms. As before the default state of the portfolio is described by the process $D = (D_{t,1}, \dots, D_{t,m})_{t \geq 0}$ with $D_{t,i} = 1_{\{\tau_i \leq t\}}$, $\mathbb{G}$ represents the global filtration to which all processes are adapted, and we work directly under the risk neutral pricing measure Q. The model is driven by some factor process X, modelled as a finite-state Markov chain with state space $S^X := \{1, \dots, K\}$. The default time τ_i has $\mathbb{G}$-default intensity $\lambda_i(X_t)$ where $\lambda_1, \dots, \lambda_m$ are given functions from S^X to $(0, \infty)$. Then, as in Equation (14), $D_{t,i} - \int_0^{\tau_i \wedge t} \lambda_i(X_s)\, ds$ is a martingale w.r.t. the full-information filtration $\mathbb{G}$. Moreover, we assume that the τ_i are conditionally independent given X (compare (15)). In this setup the process (X, D) is jointly Markov.

However, we assume that X is unobservable and that prices of traded securities are given as conditional expectation with respect to a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ which is called *market information*. The filtration $\mathbb{F}$ is generated by a process Y giving observations of X in additive noise and by the default history of the firms under consideration, that is $\mathbb{F} = \mathbb{F}^Y \vee \mathbb{F}^D$. Y is of the form

$$Y_t = \int_0^t a(X_s) ds + V_t \tag{26}$$

with a Brownian motion V, independent of X and D. As intended, X is not $\mathbb{F}$-adapted. The process Y models the information contained in security prices; it is not directly linked to observable economic quantities. We come back to this point when we discuss calibration strategies for the model in Section 6.4 below. Throughout the rest of the paper we denote by $\widehat{U}_t := \mathbb{E}(U_t | \mathcal{F}_t)$ the *projection* of a generic process U on the market filtration $\mathbb{F}$.

Example 6.1. (*A one-factor model*) In the numerical part we will consider a one-factor model where X represents the global state of the economy. For this we model the default intensities under full information as *increasing* functions $\lambda_i : \{1, \dots, K\} \to (0, \infty)$. Note that this implies that 1 represents the best state (lowest default intensities) and that K corresponds to the worst state; moreover, the default intensities are comonotonic. In the special case of a homogeneous model the default intensities of all firms are identical, $\lambda_i(\cdot) \equiv \lambda(\cdot)$. In that situation one could assume that $a(\cdot) = c \ln \lambda(\cdot)$. Here the constant $c \geq 0$ models the information-content of Y: for $c = 0$, Y carries no information, whereas for c large the state X_t can be observed with high precision.

Denote by $(q(i,k))_{1 \leq i,k \leq K}$ the generator matrix of X so that $q(i,k)$, $i \neq k$, gives the intensity of a transition from state i to state k. In this paper we consider two possible choices for this matrix. First, let the factor process be constant, $X_t \equiv X$ for all t. In that case $q(i,k) \equiv 0$, and filtering reduces to Bayesian analysis. A model of this type is known as *frailty model*, see also [46]. Below we will see that the frailty model can be viewed as a dynamic version of the implied copula model of [30]. Second, we consider the case where X has *next neighbour dynamics*, that is the chain jumps from X_t only to the neighbouring points $X_t \pm 1$ (with the obvious modifications for $X_t = 0$ and $X_t = K$).

Nonlinear filtering problems. In this setup the computation of important economic quantities leads to nonlinear filtering problems in a natural way. Consider first the pricing of credit derivatives. The payoff H of a typical credit derivative depends on the default-state of the portfolio at the maturity date T; in mathematical terms H is an $\mathcal{F}_T^D$-measurable random variable. Examples include defaultable zero-coupon bonds, CDSs or CDOs; see Subsection 6.3 for details. In line with risk-neutral pricing we define the price of the claim by the conditional expectation of the discounted payoff under the risk-neutral measure:

$$\widehat{H}_t := \mathbb{E}\big(e^{-r(T-t)}H \mid \mathcal{F}_t\big);$$

note that this definition involves the market filtration $\mathcal{F}_t$. As (X, Y) is Markovian it follows that for typical payoffs $\mathbb{E}\big(e^{-r(T-t)}H \mid \mathcal{G}_t\big)$ is a function of t, X_t and D_t which we denote by $h(t, X_t, D_t)$. By the tower property of the conditional expectation we obtain

$$\widehat{H}_t = \mathbb{E}\Big(\mathbb{E}\big(e^{-r(T-t)}H \mid \mathcal{G}_t\big) \mid \mathcal{F}_t\Big) = \mathbb{E}\big(h(t, X_t, D_t) \mid \mathcal{F}_t\big). \tag{27}$$

Since D_t is observable, in order to compute $\widehat{H}_t$ we need to determine the conditional distribution of X_t given $\mathcal{F}_t$, i.e. we have to solve a nonlinear filtering problem. This problem is studied in the next Section.

The intensity of τ_i with respect to a smaller information set $\mathbb{F}$ with $\mathbb{F}^D \subset \mathbb{F} \subset \mathbb{G}$ is given by projecting the $\mathbb{G}$-default intensity on the smaller filtration $\mathbb{F}$ (see Chapter II of [6]). Hence in our setup the $\mathbb{F}$-default intensity of firm i is given by

$$\hat{\lambda}_{t,i} := \mathbb{E}\left(\lambda_i(X_t) \mid \mathcal{F}_t\right), \quad t \le \tau_i \tag{28}$$

i.e. the computation of default intensities in the market filtration leads to a nonlinear filtering problem as well.

Model performance. We are convinced that this model has a number of attractive features. First, actual computations are done mostly in the context of the hypothetical model where X is fully observable. Since the latter has a simple Markovian structure, computations become relatively straightforward. Second, the fact that prices of traded securities are constructed by projection on the market filtration $\mathbb{F}$ leads to rich credit-spread dynamics: the proposed approach accommodates *spread risk* (random fluctuations of credit spreads between defaults) and *default contagion*; see for instance Figure 3 below. Finally, the approach gives great flexibility in terms of calibration methodologies, as is discussed in detail in Section 6.4.

6.2 Filtering and factor representation of market prices

Since X is a finite state Markov chain, the conditional distribution of X_t given $\mathcal{F}_t$ is described by the vector $\pi_t = (\pi_t^1, \dots, \pi_t^K)^\top$ with $\pi_t^k := Q(X_t = k|\mathcal{F}_t)$. In particular, we have for a generic function $a \colon \{1, \dots, K\} \to \mathbb{R}$ the relation

$$\widehat{a}_t := \widehat{a(X_t)} = \sum_{k=1}^{K} \pi_t^k a(k).$$

Proposition 6.2 shows that π is the solution of a K-dimensional SDE system. This system is driven by the $\mathbb{F}$-Brownian motion m^Y given by

$$m_t^Y := Y_t - \int_0^t \widehat{a}_s \, ds \tag{29}$$

and by the compensated default indicator process $M = (M_{t,1}, \dots, M_{t,m})'_{t \geq 0}$ with

$$M_{t,j} := D_{t,j} - \int_0^t (1 - D_{s-,j}) \, (\widehat{\lambda_j})_s ds. \tag{30}$$

Recall that $(q(i,k))_{1 \leq i,k, \leq K}$ is the generator matrix of X. In [28] the following result is established.

Proposition 6.2. *The vector $\pi_t = (\pi_t^1, \dots, \pi_t^K)'$ solves the SDE-system*

$$d\pi_t^k = \sum_{i=1}^K q(i,k) \pi_t^i dt + (\gamma^k(\pi_{t-}))^\top \, dM_t + (\alpha^k(\pi_t))^\top \, dm_t^Y \,, \tag{31}$$

with coefficients given by

$$\gamma_j^k(\pi_t) = \pi_t^k \Big(\frac{\lambda_j(k)}{\sum_{i \in S^X} \lambda_j(i) \pi_t^i} - 1 \Big), \quad 1 \leq j \leq m, \tag{32}$$

$$\alpha^k(\pi_t) = \pi_t^k \Big(a(k) - \sum_{i \in S^X} \pi_t^i a(i) \Big) . \tag{33}$$

Note that the diffusion part of (31) and in particular the function α^k from (33) has the same form as in equation (12); the form of γ^k from (32) is closely related to the Kushner-Stratonovich equation for point process observation (11).

Proposition 6.2 permits us to give an explicit expression for *contagion effects* induced by incomplete information. More precisely, consider two firms $i \neq j$. Then it follows from (32) that the jump in the default intensity of firm i at the default time τ_j of firm j is given by

$$\widehat{\lambda}_{\tau_j,i} - \widehat{\lambda}_{\tau_j-,i} = \sum_{k=1}^K \lambda_i(k) \cdot \pi_{\tau_j-}^k \Big(\frac{\lambda_j(k)}{\sum_{l=1}^K \lambda_j(l) \pi_{\tau_j-}^l} - 1 \Big) = \frac{\mathrm{cov}^{\pi_{\tau_j-}} \left(\lambda_i, \lambda_j \right)}{\mathbb{E}^{\pi_{\tau_j-}} (\lambda_j)} \,. \tag{34}$$

Here cov^π as well as $\mathbb{E}^\pi$ denote the covariance (expectation) w.r.t. the probability measure π on S^X, and π_{τ_j-} gives the conditional distribution of X immediately prior to the default event. According to (34), default contagion increases with increasing correlation of the random variables $\lambda_i(\cdot)$ and $\lambda_j(\cdot)$ under π_{τ_j-}, which is perfectly in line with economic intuition.

In [28] it is shown that the process $(D_t, \pi_t)_{t\geq}$ is a Markov process in the market filtration $\mathbb{F}$; the generator $\mathscr{L}$ of this process is an integro-differential operator. Hence the prices of credit derivatives can be expressed in terms of D_t and π_t, as is discussed in detail in the next section. The process (D, π) will therefore be called the *market state* process. The following algorithm can be used to generate a trajectory of the market state process:

Algorithm 6.3. (i) Generate a trajectory of X using a standard algorithms for the simulation of Markov chains.

(ii) Generate for the trajectory of X constructed in Step (i) a trajectory of the default indicator D and the noisy information Y. For the simulation of D one can use known methods for simulating conditionally independent, doubly stochastic random times as given in Section 9.6 of [37].

(iii) Solve (numerically) for the given trajectory of D and Y the SDE-system (31), e.g. via Euler approximation.

Once a trajectory of the market state process (D, π) is at hand, the price path of a credit derivative can be simulated using the relation $\widehat{H}_t = \sum_{k=1}^K h(t,k,D_t)\pi_t^k$, $h(\cdot)$ the full-information value of the claim as in (27).

6.3 Pricing

In this section we discuss the pricing of credit derivatives in more detail. Basically all credit derivatives common in practice fall in one of the following two classes:

- *Options on the default state:* this class comprises derivatives with a cash-flow stream that depends on the default history of the underlying portfolio so that it is $\mathbb{F}^D$-adapted; examples are corporate bonds, CDSs and CDOs.
- *Options on traded assets:* this class contains derivatives whose payoff depends on the future market value of traded credit products. Examples include options on corporate bonds or options on CDS indices and CDO tranches.

The pricing methodology for these product classes differs, so that they are discussed separately.

Options on the default state. Let the $\mathbb{F}^D$-adapted process $(H_t)_{0\le t\le T}$ be the cumulative cash-flow stream associated with the claim. Then by risk-neutral pricing its ex-dividend price at time t is defined to be

$$\widehat{H}_t = \mathbb{E}\Big(\int_t^T e^{-r(s-t)}\, dH_s \mid \mathcal{F}_t \Big). \tag{35}$$

Denote by $h(t,k,d) := \mathbb{E}\Big(\int_t^T e^{-r(s-t)}\, dH_s \mid (X_t, D_t) = (k,d) \Big)$ the full-information price of the claim. Similarly as in (27), double conditioning on the full-information filtration $\mathbb{G}$ leads to the relation

$$\widehat{H}_t = \sum_{k=1}^K \pi_t^k h(t,k,D_t). \tag{36}$$

Note that $\widehat{H}_t$ depends only on the current market state (D_t, π_t) and on the function $h(\cdot)$ that gives the hypothetical value under full information; the precise form of the the function $a(\cdot)$ from (26) and thus of the dynamics of π is irrelevant. The dynamics of π do however matter in the computation of hedging strategies; see [28] for details.

Example 6.4. We discuss zero bonds and CDS.

- Consider a zero bond on firm i with maturity T and zero recovery. Here $H_t \equiv 0$ for $t < T$ and $H_T = 1_{\{\tau_i > T\}}$. By standard results on bond pricing with doubly stochastic default times (see for instance [33]) the full-information value is given by

$$h_i(t,k,d) = 1_{\{d_i=0\}}\mathbb{E}\Big(e^{-\int_t^T r+\lambda_i(X_s)ds} | X_t = k \Big); \tag{37}$$

The price of the bond at time t is then given by $\widehat{H}_{t,i} = \sum_{k=1}^K \pi_t^k h_i(t,k,D_t)$.

- Next consider a CDS on name i. Denote by $t_1 < \cdots < t_N = T$ the premium payment dates and by x the spread of the contract. The cumulative cash-flow stream of the premium leg is then given by $H_t^{\text{prem}} = x \sum_{t_n \le t} 1_{\{\tau_i > t_n\}}$, whereas the cumulative cash-flow stream of the default leg equals $H_t^{\text{def}} = \delta \int_0^t dD_{s,i}$, $\delta \in (0,1)$ the loss given default of the firm. It is well-known that the full-information value of the premium leg at time t is equal to $x 1_{\{\tau_i > t\}} V_i^{\text{prem}}(t,k)$ with

$$V_i^{\text{prem}}(t,k) = \sum_{t_n > t} \mathbb{E}\Big(e^{-\int_t^{t_n} r + \lambda_i(X_s) ds} | X_t = k \Big) ; \tag{38}$$

the full-information value of the default leg equals $\delta 1_{\{\tau_i > t\}} V_i^{\text{def}}(t,k)$ with

$$V_i^{\text{def}}(t,k) = \mathbb{E}\Big(\int_t^T \lambda_i(X_s) e^{-\int_t^s r + \lambda_i(X_u) du} \, ds | X_t = k \Big). \tag{39}$$

Given the spread x the value at time t is thus given by $\sum_{k=1}^K \pi_t^k \big(x\, V_i^{\text{prem}}(t,k) - V_i^{\text{def}}(t,k) \big)$, and the fair spread at time t is

$$x_t^* := \frac{\delta \sum_{k=1}^K \pi_t^k V_i^{\text{def}}(t,k)}{\sum_{k=1}^K \pi_t^k V_i^{\text{prem}}(t,k)}.$$

Analogous arguments are used in the pricing of CDS indices and CDOs.

Remark 6.5 (Computation of full-information value $h(\cdot)$). For bonds and CDSs the computation of h amounts to computing (37) and (38), respectively. There is an easy solution to this task involving the exponential of the generator matrix of X, see [20] and [29]. In the case of CDOs, a solution of this problem via Laplace transforms can be found in [15]. Alternatively, a two stage method that employs the conditional independence of defaults given $\mathcal{F}_\infty^X$ can be used. For this one first generates a trajectory of X via Monte Carlo. Given this trajectory, the loss distribution can then be evaluated using one of the known methods for computing the distribution of the sum of independent (but not identically distributed) Bernoulli variates. Finally, the loss distribution is computed by averaging over the sampled trajectories of X. An extensive numerical case study comparing the different approaches is given in [47].

Options on traded assets. Assume now that N basic options on the default state are traded on the market, and denote their ex-dividend price at time t by $\widehat{p}_{\tilde{t},1}, \dots, \widehat{p}_{\tilde{t},N}$. Then the payoff of an option on traded assets is of the form $\tilde{g}(D_{\tilde{T}}, \widehat{p}_{\tilde{T},1}, \dots, \widehat{p}_{\tilde{T},N})$, to be paid at maturity $\tilde{T} \le T$ (T is the maturity of the underlying products). From (36) the payoff of the option can be written in the form $g\,(D_{\tilde{T}}, \pi_{\tilde{T}})$, where g is implicitly defined. Since the market state (D, π) is an $\mathbb{F}$-Markov process, the price of the option at time $t < \tilde{T}$ is given by a function of time and the current market state,

$$\mathbb{E}\left(e^{-r(\tilde{T}-t)} g(D_{\tilde{T}}, \pi_{\tilde{T}}) | \mathcal{F}_t \right) = g(t, D_t, \pi_t). \tag{40}$$

By standard results from Markov process theory the function g is a solution of the backward equation

$$\partial_t g(\cdot) + \mathscr{L} g(\cdot) = 0,$$

$\mathscr{L}$ the generator of (D,π). However, the market state is usually a high-dimensional process so that the practical computation of $g(\cdot)$ has to be based on Monte Carlo methods, using Algorithm 6.3. Note that for an option on the default state the function $g(\cdot)$ does typically depend on the entire generator $\mathscr{L}$ of (D,π) and hence on the form of $a(\cdot)$.

Example 6.6 (options on a CDS index)**.** Index options are a typical example for an option on a traded asset. Denote by $\tilde{T} < T$ the maturity of the contract and of the underlying CDS index. Upon exercise the owner of the option holds a protection-buyer position on the underlying index with a pre-specified spread $\bar{x}$ (the exercise spread of the option); moreover, he obtains the cumulative portfolio loss up to time $\tilde{T}$ given by

$$L_{\tilde{T}} = \sum_{i=1}^{m} \delta 1_{\{\tau_i \le \tilde{T}\}} \,.$$

Denote by $V^{\mathrm{def}}(t, X_t, D_t)$ and $V^{\mathrm{prem}}(t, X_t, D_t)$ the full-information value of the default and the premium payment leg of the CDS index. In our setup the value of the option at maturity $\tilde{T}$ is then given by the following function of the market state at $\tilde{T}$:

$$g(D_{\tilde{T}}, \pi_{\tilde{T}}) = \Big(L_{\tilde{T}} + \sum_{k \le K} \pi^k_{\tilde{T}} \big(V^{\mathrm{def}}(\tilde{T}, k, D_{\tilde{T}}) - \bar{x} V^{\mathrm{prem}}(\tilde{T}, k, D_{\tilde{T}}) \big) \Big)^+ \,. \tag{41}$$

Numerical examples are given in Subsection 7.3 below.

6.4 Calibration

As we have just seen, the price of the credit derivatives common in practice is given by a function of the current market state (D,π). Here a major issue arises: we view the process Y generating the market filtration $\mathbb{F}$ as some kind of abstract information. Then the process π is not directly observable for investors. On the other hand, pricing formulas need to be evaluated using only publicly available information. An key point for the application of the model is therefore to determine π_t from prices of traded securities observed at time t, that is model calibration. We discuss two approaches, standard calibration based on linear or convex optimization and a calibration approach via filtering proposed in [26].

Standard calibration. Standard calibration means that we determine π_t by minimizing some distance between market prices and model prices at time t. This is facilitated substantially by the observation that the set of all probability vectors consistent with the price information at a given point in time t can be described in terms of a set of linear inequalities.

Example 6.7. We discuss zero coupon bonds and CDSs as representative examples:

- Consider a zero coupon bond on firm i and suppose that at t we observe bid and ask quotes $\underline{p} \le \overline{p}$ for the bond. In order to be consistent with this information, a solution π of the calibration problem at t needs to satisfy the linear inequalities

$$\underline{p} \le \sum_{k=1}^{K} p_i(t,k) \pi^k \le \overline{p} \,.$$

- Consider a CDS contract on firm i and suppose that at time t we observe bid and ask spreads $\underline{x} \le \overline{x}$ for the contract. Then π must satisfy the following two inequalities:

$$\sum_{k=1}^{K} \pi^k \big(\underline{x} V_i^{\mathrm{prem}}(t,k) - \delta V_i^{\mathrm{def}}(t,k) \big) \le 0\,,$$
$$\sum_{k=1}^{K} \pi^k \big(\overline{x} V_i^{\mathrm{prem}}(t,k) - \delta V_i^{\mathrm{def}}(t,k) \big) \ge 0\,.$$

  Moreover, π needs to satisfy the obvious linear constraints $\pi^k \ge 0$ for all k and $\sum_{k=1}^{K} \pi^k = 1$.

Standard linear programming techniques can be used to detect if the system of linear inequalities corresponding to the available market quotes is nonempty and to determine a solution π. In case that there is more than one probability vector π consistent with the given price information at time t, a unique solution π^* of the calibration problem can be determined by a suitable *regularization procedure.* For instance one could choose π^* by minimizing the relative entropy to the uniform distribution. This leads to the convex optimization problem

$$\pi^* = \operatorname{argmin} \Big\{ \sum_{k=1}^{K} \pi^k \ln \pi^k : \pi \text{ is consistent with the price information in } t \Big\}.$$

Calibration via filtering. Alternatively π_t can be estimated from historical price data by nonlinear filtering. By (36), the price of a traded credit product in the market filtration is given by a function $g(t, D_t, \pi_t)$ of time and of the current market state. Assume that investors observe this price with a small amount of noise. The noise represents observation errors such as bid-ask spreads and transmission errors as well as errors in the model specification. As explained in Subsection 2.2, the noisy price observation can be modelled by the process U with dynamics

$$dU_t = g(t, D_t, \pi_t)dt + w d\tilde{V}_t, \tag{42}$$

where $\tilde{V}$ is a Brownian motion independent of all other processes, and where the constant $w > 0$ models the error variance in the price observation. In this context estimating π_t amounts to finding the mean of conditional distribution of π_t given the information available to investors at time t, where the latter is described by the σ-field $\mathcal{F}_t^I := \mathcal{F}_t^D \vee \mathcal{F}_t^U$. Recall that π_t solves the SDE (31). In order to determine the conditional distribution of π_t given $\mathcal{F}_t^I$ one therefore has to solve a second nonlinear filtering problem with signal process $(\pi_t)_{t\ge 0}$ and observations given by the default state D and the noisy price information U. From a filtering viewpoint this is a challenging problem with usually high-dimensional signal π, observations of mixed type (diffusion and marked point processes) and with common jumps of observation D and signal π. This problem is studied in detail in [26]; in particular, that paper proposes a numerical solution via particle filtering. Numerical results are presented in the next section.

Calibration via filtering is appealing conceptually: new price information at t is used to update the a priori distribution of π_t given past price information up to time $t-1$, say, but this a-priori distribution (and hence the history of prices) is not disregarded altogether. In that sense the method provides an interpolation between "historical estimation" of model parameters and standard calibration procedures.

Remark 6.8. Of course, in order to use the model one needs to determine also the parameters of $a(\cdot)$ and - in case of next neighbour dynamics - parameters of the generator matrix of X. An approach to this problem via time-series methods can be found in [29].

6.5 Hedging

Hedging is a key issue in the management of portfolios of credit derivatives. The standard practice adopted on credit markets is to use sensitivity-based hedging strategies computed by ad hoc rules within the static base-correlation framework; see for instance [40]. Clearly, it is desirable to work with hedging strategies which are based on a methodologically sound approach instead. Using the previous results on the dynamics of credit derivatives it is possible to derive model-based dynamic hedging strategies. A detailed derivation of these strategies can be found in the original paper [28]. A key issue in the computation of hedging strategies is the fact the market is typically incomplete (that is most claims cannot be replicated perfectly), as the price of the traded credit derivatives follows a jump-diffusion process. In order to deal with this problem the concept of risk minimization as introduced by [22] is therefore used. Risk-minimization is well-suited for the hedging of credit derivatives, as the ensuing hedging strategies are relatively easy to compute and as it suffices to know the risk-neutral dynamics of credit derivative prices.

The dynamic hedging of credit derivatives is also studied in [23], [35] or [12], albeit in a different setup.

7 Numerical case studies

In order to illustrate the application of the model to practical problems we present a number of small numerical case studies on model-dynamics, calibration and on the pricing of credit index options. We concentrate on homogeneous models throughout, while the inhomogeneous situation is covered in [28].

7.1 Dynamics of credit spreads and of π

As remarked earlier, the fact that in our model the prices of traded securities are given by the conditional expectation with respect to the market filtration leads to rich credit-spread dynamics with random fluctuations of credit spreads between defaults and default contagion. This is illustrated in Figure 3 by a simulated credit-spread trajectory. The fluctuation of credit spreads between defaults as well contagion effects at default times (e.g. around $t = 600$) can be spotted clearly. The right graph gives the corresponding trajectory of the solution π of the Kushner-Stratonovich equation (31). State probabilities fluctuate in response to the fluctuations of D; moreover, there are shifts in the distribution π at default events. Both graphs have been created for the case where X is a Markov chain with next-neighbour dynamics.

7.2 Calibration

We discuss calibration for the frailty model where $X_t \equiv X$ and hence the generator matrix of X is identically zero, see also Example 6.1. In the frailty model default times are independent, exponentially distributed random variables given $X = k$, and dependence is created by mixing

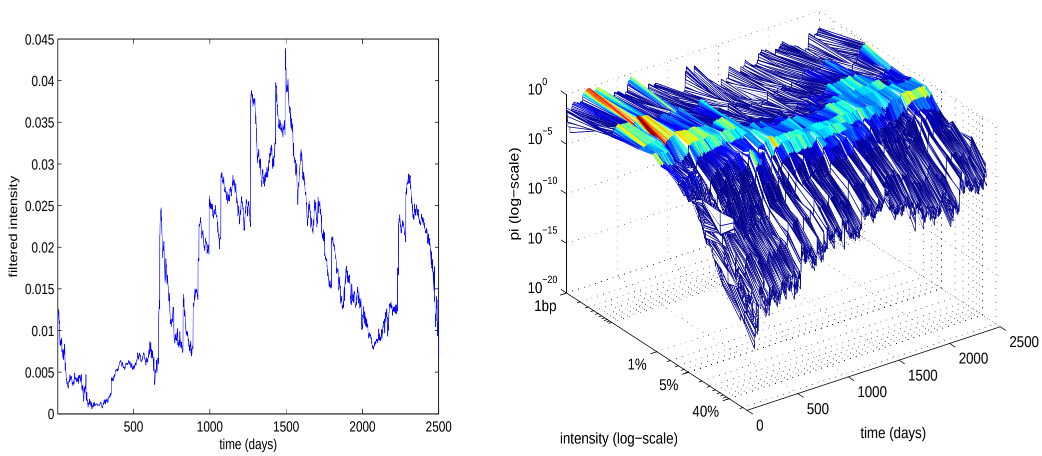

Figure 3: A simulated path of credit spreads under zero recovery (left) and the corresponding trajectory of the solution π of the Kushner-Stratonovich equation (right); time is measured in days. Note that on the right graph logarithmic scaling is being used.

over the states of X. A static model of this form (no dynamics of π) has been proposed by [30] under the label *implied copula model*; see also [42]. Since prices of CDS-indices and CDO tranches are independent of the dynamics of π (recall the discussion surrounding (36) above), for these products pricing and standard calibration in the dynamic frailty model and in the static implied copula models coincide. However, our framework permits the pricing of tranche- and index options and the derivation of model-based hedging strategies. Both issues cannot be addressed in the static implied copula models.

Since in the frailty model default times are independent given the current value of X, computing the full-information value of traded securities is particularly easy. On the other hand, the long-run dynamics of credit spreads implied by the frailty model are quite unrealistic, as the filter learns the "true value" of X over time. However, since prices of CDS-indices and CDO tranches are independent of the dynamics of π this is not a problem for the calibration of the model to index data and tranche data or for the pricing of bespoke tranches. The frailty model is however not well-suited for the pricing of options on traded assets if the maturity $\tilde{T}$ of the option is large.

Standard calibration to itraxx spreads. We begin with an example for a calibration of the model to observed tranche and index spreads of the itraxx. We consider a homogeneous model with $|S^X| = 9$; the values of the one-year default intensity are given in Table 1 below. The model was calibrated to tranche and index spread data from 2004, 2006, 2008 and 2009. The data from 2004 and 2006 are typical for tranche and index spreads before the credit crisis; the data from 2008 and 2009 on the other hand represent the state of the market during the crisis. In order to determine a solution π^* of the calibration problem we use the methodology described in Section 6.4, with very satisfactory results. The resulting values for π are given in Table 1. We clearly see that with the emergence of the credit crisis the calibration procedure puts more mass on states where the default intensity is high; in particular, the extreme state

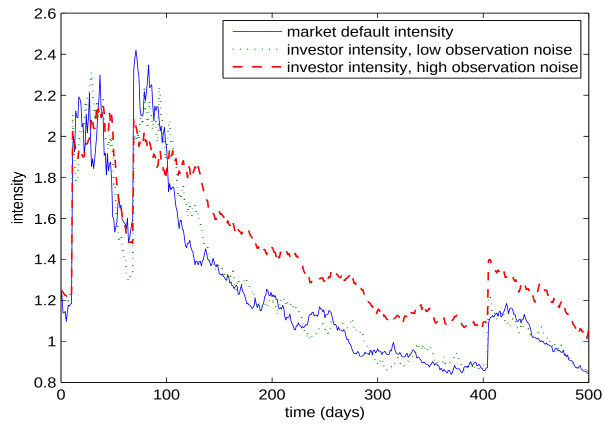

Figure 4: A trajectory of the market default intensity $\widehat{\lambda}_t$ and of the investor estimate $E(\widehat{\lambda}_t \mid \mathcal{F}_t^I)$ for different observation noise.

where $\lambda = 70\%$ gets a probability of around 3%. This reflects the increased awareness of future defaults and the increasing risk aversion in the market after the advent of the crisis. The fact that the the model-implied probability of "Armageddon-scenarios" increases as the credit crisis unfolds can also be observed in other model types; see for instance [8].

λ (in %)	0.01	0.3	0.6	1.2	2.5	4.0	8.0	20	70
π^*, data from 2004	12.6	22.9	42.0	17.6	2.5	1.45	0.54	0.13	0.03
π^*, data from 2006	22.2	29.9	39.0	7.6	1.2	0.16	0.03	0.03	0.05
π^*, data from 2008	1.1	7.9	57.6	10.8	11.7	4.9	1.26	1.79	2.60
π^*, data from 2009	0.0	13.6	6.35	42.2	22.3	12.5	0.0	0.00	3.06

Table 1: Results of the calibration to itraxx spread data (index and tranches) for different data sets from several years; the components of π^* are given in percentage points.

Calibration via filtering. Next we illustrate the filter approach to model calibration with numerical results from [26]. The quantity to be estimated via filtering is the default intensity in the market filtration $\widehat{\lambda}_t$ which can be viewed as approximation for the short-term credit spread. Numerical results are given in Figure 4, where the filter estimate $\mathbb{E}(\widehat{\lambda}_t \mid \mathcal{F}_t^I)$ is given for a high and a low value of the observation noise w. Note that for low observation noise the estimator $\mathbb{E}(\widehat{\lambda}_t \mid \mathcal{F}_t^I)$ tracks $\widehat{\lambda}_t$ quite well. Further details are given in [26].

7.3 Pricing of credit index options

Options on a CDS index introduced in Example 6.6 are a typical example for an option on traded asset. In practice this contract is usually priced by a fairly ad-hoc procedure: it is assumed that the so-called loss adjusted spread (the sum of the value of the default payments over the time period $(\tilde{T}, T]$ and of the front-end protection $L_{\tilde{T}}$, divided by the value of

c	0.5	1	2	5
moneyness $\bar{x}/x_0 = 0.8$	1.53	1.56	1.62	1.83
moneyness $\bar{x}/x_0 = 1$	1.75	1.75	1.76	1.93
moneyness $\bar{x}/x_0 = 1.2$	1.95	1.95	1.95	2.04

Table 2: Implied volatilities for a option to buy protection on the CDS index with Implied volatilities are computed via the Pedersen (2003) [41] approach.

the premium payments over $(\tilde{T}, T])$ is lognormally distributed under a suitable martingale measure so that the value of the option can be computed via the Black formula. Prices are then quoted in terms of implied volatilities; see [41] and [7] for further details. Beyond convenience there is no justification for the lognormality assumption in the literature. In particular, it is unclear if a dynamic model for the evolution of spreads and credit losses can be constructed that supports the lognormality assumption and the use of the Black formula, and there is no empirical justification for this assumption either.

The filter-model discussed here on the other hand offers the possibility to price this product in the context of a consistent model for the joint evolution of defaults and credit spreads. In our numerical experiments we worked in the following setup: we used the same frailty model as in the calibration to itraxx data; the function $a(\cdot)$ from (26) was given by $a(k) = c \ln \lambda(k)$ for varying values of c; the value π_0 at the starting day of the contract was the 2009-value from Table 1, i.e. the model was calibrated to tranche spreads and index spreads on that date; the time to maturity $\tilde{T}$ of the option was taken equal to three months[5]. Prices were computed using Monte Carlo simulation.

Table 2 presents our pricing results for varying values of c (varying local spread volatility) and varying moneyness $\bar{x}/x_0$ ($\bar{x}$ the exercise spread of the option as given in (41) and x_0 the index spread at inception of the option). We can see the following

- The model generates *volatility skews:* options with high moneyness (out of the money options) tend to have higher implied volatilities than in the money options. This appears reasonable: out of the money potions provide protection against the adverse scenario of rising spreads and/or many losses during the runtime of the option. Such a protection tends to be more expensive than the protection against benign scenarios. The obvious analogy is the skew for equity options, where implied volatilities for out-of-the-money put options (which offer protection against the adverse scenario of falling markets) are higher than implied volatilities for out-of-the-money calls.

- Increasing the value of c tends to lead to higher implied volatilities. Nonetheless it shows that for options on traded assets the choice of the function $a(\cdot)$ does indeed have an impact on the price of the option. Given market quotes for credit index options this observation could of course be used to calibrate parameters of $a(\cdot)$.

We also used next neighbour dynamics for X to price the option. This led to a slightly smoother distribution of the credit spread at $\tilde{T}$, but the impact on option prices and implied volatilities was found to be very small. Finally we looked at the distribution of the loss-adjusted spread in our model. Recall that in the literature it is frequently assumed that

[5]Short maturities of 3–6 months are the market standard for index options, longer-term contracts are hardly traded as the composition of the underlying index changes every 6 months.

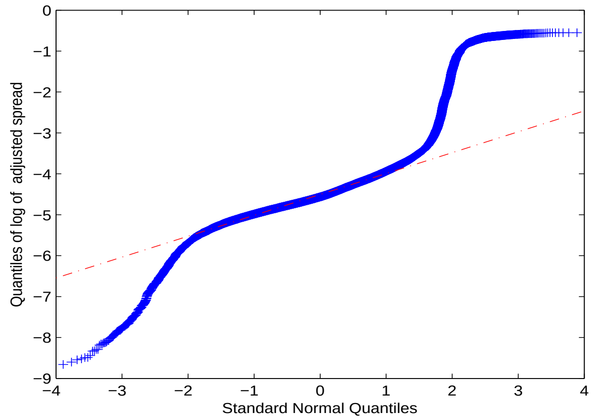

Figure 5: Quantile-quantile plot of logarithmic loss-adjusted spread against the normal distribution. The S-shaped form of the plot clearly points to heavy tails.

this spread is log-normally distributed. In Figure 5 we therefore give a quantile-quantile plot of logarithmic loss-adjusted spreads in our model against the normal distribution. The S-shaped form of the plot clearly points to heavy tails.

Unfortunately, market quotes for index options are relatively scarce so that we could not test our pricing results empirically. However, our findings clearly caution against the thoughtless use of the Black formula and of market models in credit index markets, despite of the obvious success of this methodology in the default-free interest world.

References

[1] T. Aven. A theorem for determining the compensator of a counting process. *Scandandinavian Journal of Statistics*, 12(1):69–72, 1985.

[2] A. Bain and D. Crisan. *Fundamentals of Stochastic Filtering*. Springer, New York, 2009.

[3] F. Black and J. C. Cox. Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance*, 31:351–367, 1976.

[4] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[5] C. Blanchet-Scalliet and M. Jeanblanc. Hazard rate for credit risk and hedging defaultable contingent claims. *Finance and Stochastics*, 8:145–159, 2004.

[6] P. Brémaud. *Point Processes and Queues*. Springer Verlag. Berlin Heidelberg New York, 1981.

[7] D. Brigo and M. Morini. Arbitrage-free pricing of credit index options. the no-armageddon-pricing measure and the role of correlation after the subprime crisis. preprint, available on www.defaultrisk.com, 2007.

[8] D. Brigo, A. Pallavicini, and R. Torresetti. Credit Models and the Crisis, or: How I learned to stop worrying and love the CDO. working paper, Imperial College London, 2009.

[9] A. Budhiraja, L. Chen, and C. Lee. A survey of nonlinear methods for nonlinear filtering problems. *Physica D*, 230:27–36, 2007.

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

[11] P. Collin-Dufresne, R. Goldstein, and J. Helwege. Is credit event risk priced? Modeling contagion via the updating of beliefs. Preprint, Carnegie Mellon University, 2003.

[12] R. Cont and Y. H. Kan. Dynamic hedging of portfolio credit derivatives. *Working paper*, 2008.

[13] P. Crosbie and J. R. Bohn. *Modeling Default Risk*. KMV Corporation [http://www.kmv.kom/insight/index.html], 1997-2001.

[14] M. H. A. Davis and S. I. Marcus. An introduction to nonlinear filtering. In M. Hazewinkel and J. C. Willems, editors, *Stochastic Systems: The Mathematics of Filtering and Identifications and Applications*, pages 53–75. Reidel Publishing Company, 1981.

[15] G. di Graziano and L. C. G. Rogers. A new approach to the modeling and pricing of correlation credit derivatives. *Int. Jour. Theor. Appl. Fin.*, 64:2089–2123, 2009.

[16] D. Duffie, A. Eckner, G. Horel, and L. Saita. Frailty correlated default. *Journal of Finance*, 64, 2089–2123, 2009.

[17] D. Duffie and D. Lando. Term structures of credit spreads with incomplete accounting information. *Econometrica*, 69:633–664, 2001.

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

[19] R. Elliott. New finite-dimensional filters and smoothers for noisily observed markov chains. *IEEE Trans. Info. theory*, IT-39:265–271, 1993.

[20] R. J. Elliott and R. S. Mamon. A complete yield curve description of a Markov interest rate model. *International Journal of Theoretical and Applied Finance*, 6:317 – 326, 2003.

[21] P. Feldhütter and D. Lando. Decomposing swap spreads. *Journal of Financial Economics*, 88:375 – 405, 2008.

[22] H. Föllmer and D. Sondermann. Hedging of non-redundant contingent-claims. In W. Hildenbrand and A. Mas-Colell, editors, *Contributions to Mathematical Economics*, pages 147–160. North Holland, 1986.

[23] R. Frey and J. Backhaus. Dynamic hedging of synthetic CDO-tranches with spread- and contagion risk. To appear in Journal of Economic Dynamics and Control, 2009.

[24] R. Frey and D. Lu. Pricing and hedging of hybrid credit products: a filtering approach. Preprint, Universität Leipzig, 2010, in preparation.

[25] R. Frey and W. Runggaldier. Nonlinear filtering in models for interest-rate and credit risk. To appear in *Handbook of Nonlinear Filtering* by Oxford University Press, 2008.

[26] R. Frey and W. Runggaldier. Pricing credit derivatives under incomplete information: a nonlinear filtering approach. to appear in *Finance and Stochastics*, 2008.

[27] R. Frey and T. Schmidt. Pricing corporate securities under noisy asset information. *Mathematical Finance*, 19(3):403 – 421, 2009.

[28] R. Frey and T. Schmidt. Pricing and hedging of credit derivatives via the innovations approach to nonlinear filtering. Preprint. Universität Leipzig, 2010.

[29] A. Herbertsson and R. Frey. Pricing and hedging index-CDS options in a nonlinear filtering model. *Working paper*, 2010 (in preparation).

[30] J. Hull and A. White. The implied copula model. *The Journal of Derivatives*, 2006.

[31] R. Jarrow and P. Protter. Structural versus reduced-form models: a new information based perspective. *Journal of Investment management*, 2:1–10, 2004.

[32] R. Jarrow and S. Turnbull. Pricing options on financial securities subject to default risk. *Journal of Finance*, 5:53–86, 1995.

[33] D. Lando. On Cox processes and credit risky securities. *Review of Derivatives Research*, 2:99–120, 1998.

[34] D. Lando. *Credit Risk Modeling: Theory and Applications.* Princeton University Press. Princeton, New Jersey, 2004.

[35] J. Laurent, A. Cousin, and J. Fermanian. Hedging default risk of CDOs in Markovian contagion models. workinp paper, ISFA Actuarial School, Université de Lyon, 2007.

[36] V. Linetsky. Pricing equity derivatives subject to bankruptcy. *Mathematical Finance*, 16(2):255–282, 2006.

[37] A. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, 2005.

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

[39] H. Nakagawa. A filtering model on default risk. *J. Math. Sci. Univ. Tokyo*, 8:107–142, 2001.

[40] M. Neugebauer. Understanding and hedging risks in synthetic CDO tranches. Fitch Special Report, August 2006.

[41] C. Pedersen. Valuation of portfolio credit default swaptions. working paper, Lehman Brothers Quantitative Credit Research, 2003.

[42] D. Rosen and D. Saunders. Valuing CDOs of bespoke portfolios with implied multi-factor models. *The Journal of Credit Risk*, 2009.

[43] T. Schmidt and A. Novikov. A structural model with unobserved default boundary. *Applied Mathematical Finance*, 15(2):183 – 203, 2008.

[44] T. Schmidt and W. Stute. Credit risk – a survey. *Contemporary Mathematics*, 336: 75–115, 2004.

[45] P. Schönbucher. *Credit Derivates Pricing Models*. John Wiley & Sons. New York, 2003.

[46] P. Schönbucher. Information-driven default contagion. Preprint, Department of Mathematics, ETH Zürich, 2004.

[47] R. A. Wendler. Über die Bewertung von Kreditderivaten unter unvollständiger Information. *Blätter der DGVFM*, 30:379 – 394, 2009.