Nonlinear Filtering in Models for Interest-Rate and Credit Risk

Rüdiger Frey[1] and Wolfgang Runggaldier[2]

June 23, 2009[3]

Abstract

We consider filtering problems that arise in Markovian factor models for the term structure of interest rates and for credit risk. Investors are supposed to have only incomplete information about the factors and so their current state has to be inferred/filtered from observable financial quantities. Our main goal is the pricing of derivative instruments in the interest rate and credit risk contexts, but also other applications are discussed.

Keywords: Stochastic filtering, Incomplete information in finance, Term structure of interest rates, Credit risk, Derivatives, Filtering and parameter estimation.

1 Introduction

Modern financial mathematics is mainly concerned with the pricing and hedging of derivative securities, with portfolio optimization, and with risk management and the statistical analysis of financial data. All these activities are based on mathematical models for the dynamics of the underlying economic quantities such as security prices. These models need to capture the complicated nonlinear dynamics of real asset prices while being at the same time parsimonious and numerically tractable. Factor models have proven to be a useful tool for meeting these conflicting objectives, since the quantities of interest can be expressed in terms of relatively few factors. Moreover, with Markovian factor processes, Markov-process techniques can be fruitfully employed. In most financial applications of factor models investors have only incomplete information about the state of the factor process, essentially for the following reasons: first, some factors are associated with economic quantities which are hard to observe precisely such as instantaneous interest rates, volatilities, or the asset value of a firm; second, abstract factors without direct economic interpretation are often included in the specification of a model in order to increase its flexibility. When applying the model, the current state of the factors therefore needs to be inferred from observable quantities such as historical price data. Filtering is an elegant and theoretically consistent way for doing this, which is why filtering techniques are increasingly being used in all areas of financial mathematics.

In the present paper we concentrate on the application of filtering techniques in the context of incomplete-information-models for interest-rate and credit risk that are of the type of jump-diffusion models. Our main concern is the pricing of derivatives via martingale methods; hedging and parameter estimation are touched upon occasionally.[4] This focus is motivated by trends in the current literature and by our own research interests over the last few years. We remark

[1]Department of Mathematics, University of Leipzig, D-04009 Leipzig, Germany, frey@math.uni-leipzig.de.

[2]Department of Mathematics, University of Padua, Via Trieste 63, 35121 Padova, Italy, runggal@math.unipd.it.

[3]This paper is submitted to the *Handbook of Nonlinear Filtering* (D. Crisan and B. Rozovski, eds.), to be published by Oxford University Press. The authors thank Carl Chiarella and Abdel Gabih for useful comments.

[4]For a discussion of portfolio optimization under incomplete information and the ensuing nonlinear filtering problems we refer to [54].

at this point that nonlinear filtering has been applied very successfully to pricing, hedging, and parameter-estimation problems in marked-point-process models driven by an unobservable volatility factor; see for instance [34], [33], [31], [60], [38], [20] or [19].

The outline of the paper is as follows: In Section 2 we give a brief introduction to arbitrage-free models for the term-structure of interest rates with a particular emphasis on factor models. This sets the scene for our discussion of term-structure models under incomplete information in Section 3. Here we start with a general result which shows that arbitrage-free prices with respect to the sub-filtration representing the information actually available to investors can be computed by projection. In the remainder of Section 3 this principle is applied within specific factor models for the term structure of interest rates and this leads to a number of interesting filtering problems. Sections 4, 5 and 6 are devoted to an analysis of nonlinear filtering in dynamic credit risk models: in Section 4 we give an overview of key modeling approaches and explain how and where incomplete information enters; in Section 5 we discuss nonlinear filtering problems in the context of firm-value models with noisily observed asset value; Section 6 deals with reduced-form models. Section 7 summarizes the paper. Rather than aiming at a complete description of available results, we will concentrate on a few illustrative models, many of them coming from our own activity in the field. We assume throughout that the reader is familiar with standard nonlinear filtering theory; a comprehensive modern account can be be found in the recent monograph [2].

Throughout the paper we denote by $(\mathcal{G}_t)$ the global or full-information filtration, so that all processes introduced will be $(\mathcal{G}_t)$ adapted; the information actually available to investors is represented by the sub-filtration $(\mathcal{F}_t)$. Moreover, we generally adopt bold-face notation for vectors and vector-valued stochastic processes.

2 The term structure of interest rates. Full information.

In this section we give a brief introduction to models for the term structure of interest rates; details and further information can for instance be found in [6].

Bonds and interest rates. A zero-coupon bond or T-bond is a contract guaranteeing a unit amount at a given future date T without intermediate payments; the price of such a contract at a time $t \leq T$ is denoted by $p(t,T)$. The collection of bond prices $p(t,T)$, $T \geq t$, completely describes the term structure of interest rates, or, equivalently, the time-value of money at a given point in time t. Various notions of interest rates can be defined from the family $p(t,T)$, $T \geq t$. An important example is the *simple compounded interest rate* for the future time period $[T,S]$ and contracted at $t < T$, denoted by $L(t;T,S)$. This rate is given by

$$L(t;T,S) = \frac{p(t,T) - p(t,S)}{(S-T)\,p(t,S)} = \frac{1}{S-T}\left[\frac{p(t,T)}{p(t,S)} - 1\right] . \tag{1}$$

Assuming that, as a function of T, $p(t,T)$ is sufficiently regular, letting $S \downarrow T$, one obtains the instantaneous *forward rate*

$$f(t,T) = \lim_{S \downarrow T} L(t;T,S) = -\frac{\partial}{\partial T} \log p(t,T) \tag{2}$$

By its definition, $f(t,T)$ represents the rate, evaluated at $t < T$, for an instantaneous borrowing at T. From (2), using the fact that $p(T,T) = 1$, we also get the inverse relationship

$$p(t,T) = \exp\left(-\int_t^T f(t,u)du\right), \tag{3}$$

so that there is a one-to-one relationship between the family of bond prices $p(t,T)$, $T \geq t$ and the family of forward rates $f(t,T)$, $T \geq t$. The (instantaneous) *short rate* is finally defined by $r_t := f(t,t)$.

Martingale pricing. As mentioned in the introduction, our main concern in this paper is the pricing of derivatives via martingale methods. This methodology is based on a widely used economic principle, namely the notion of absence of arbitrage. This principle basically states that, in equilibrium, the prices of the assets on a given market have to be such that by investing in this market it is not possible to make a sure profit without risk. According to the so-called *first fundamental theorem of asset pricing* the mathematical counterpart of this principle is the existence of an equivalent martingale measure. This is a measure Q^N, equivalent to the physical/real-world measure P, so that the prices of all the assets in a given market expressed in units of a given reference asset (*numeraire*) N with price $N_t > 0$ are Q^N martingales with respect to a given generic filtration $(\mathcal{H}_t)$ representing the information available to investors in the model. Formally, the price of a non-dividend-paying traded asset $(S_t)_{t\geq 0}$ thus satisfies for all $t \leq T$

$$\frac{S_t}{N_t} = E^{Q^N}\left(\frac{S_T}{N_T} \mid \mathcal{H}_t\right) \tag{4}$$

Under the popular *martingale modeling approach* relation (4) is used for constructing the price-dynamics of the traded assets as follows: suppose that the value of a security[5] at some given future date T is given by a known $\mathcal{H}_T$-measurable random variable Π_T. A prime case in point is a T bond where $\Pi_T \equiv 1$. Given a numeraire N, a candidate martingale measure Q^N and a filtration $(\mathcal{H}_t)$, the price Π_t of this security at $t \leq T$ is then defined to be

$$\Pi_t = N_t\, E^{Q^N}\left(\frac{\Pi_T}{N_T} \mid \mathcal{H}_t\right). \tag{5}$$

Model parameters are determined by the requirement that the model-implied price Π_t from (5) should coincide with the price observed on the market; this goes under the label *calibration to market data.* In a second step the price of non-traded derivatives is defined by the analogous expression to (5). In this way it is automatically ensured that the resulting model is arbitrage-free and that derivatives are priced consistently with the prices of traded assets. A frequently used numeraire is the so-called money market account (locally risk free asset) that is the asset with value $B_t = B_0 \exp\left(\int_0^t r_s ds\right)$, r the short rate. The martingale measure corresponding to B as numeraire is commonly denoted by Q.

Note that the real-world measure P does not enter in this approach, and in fact it is common practice to set up a pricing model for derivatives without specifying the real-world dynamics of security prices. A conceptual problem may however arise at this point, since the measure Q^N need not be unique and since different martingale measures can lead to different prices for non-traded

[5]For simplicity we tacitly assume that the security does not generate any intermediate cash flows such as dividend- or interest payments.

derivatives. This problem is closely related to the so-called *completeness* of the market (see for instance Chapter 8, 10 and 14 of [6]). Economic criteria for choosing one of these measures do usually invoke the physical measure P and martingale modeling is no longer sufficient. In practical applications of derivative pricing models this issue is largely neglected and the chosen martingale measure is kept fixed, a praxis which is also adopted in the present paper.

Heath-Jarrow-Morton approach. We proceed now to derive dynamic models for the term structure that do not allow for arbitrage opportunities. A recent such modeling approach is the so-called *Heath-Jarrow-Morton (HJM) approach* [43]. Under this approach one models directly the dynamics of the forward rates and derives from there the dynamics of bond prices and related quantities. Here we restrict ourselves to Wiener driven models and assume that the forward rate dynamics are of the form

$$df(t,T) = \alpha(t,T)dt + \boldsymbol{\sigma}(t,T)d\mathbf{W}_t, \tag{6}$$

where $\mathbf{W}_t$ is a d-dimensional Wiener process on a given filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t), Q)$ and $\alpha(\cdot,T)$, $\boldsymbol{\sigma}(\cdot,T)$ are adapted processes with values in $\mathbb{R}$ and $\mathbb{R}^d$ respectively. Note that (6) may be interpreted as a system of infinite stochastic differential equations, one for each T.

The fact of having in principle infinitely many assets, given by the bonds of the various maturities T, implies that with a model as in (6) one might introduce arbitrage into the market. A simple way to preclude arbitrage opportunities is to specify the dynamics of the processes $f(\cdot,T)$ in such a way that the given measure Q is a martingale measure. It is well-known that (modulo some integrability conditions) Q is a martingale measure if and only the if the so-called *HJM-drift condition* is satisfied, that is the following relation between the drift α and the volatility $\boldsymbol{\sigma}$ has to hold:

$$\alpha(t,T) = \boldsymbol{\sigma}(t,T)\int_t^T \boldsymbol{\sigma}'(t,u)du\,; \tag{7}$$

see e.g. [6] for details. Rewriting (6) in integral form, namely

$$f(t,T) = f^*(0,T) + \int_0^t \alpha(s,T)ds + \int_0^t \boldsymbol{\sigma}(s,T)d\mathbf{W}_s\,, \tag{8}$$

one sees that, in the HJM setup, the inputs for a model defined under a martingale measure Q are: i) the volatility structure $\boldsymbol{\sigma}(t,T)$; ii) the initially observed forward rate curve $f^*(0,T)$. The structure of the model is thus specified by specifying $\boldsymbol{\sigma}(t,T)$.

Factor models. In the given setup the models are a-priori infinite-dimensional and one may ask whether, by a judicious choice of the volatility structure $\boldsymbol{\sigma}(t,T)$ in the HJM framework (6), they may become equivalent to a model driven by a finite-dimensional factor process. The question has a positive answer and a general account on this issue may be found in [5]. For the filter application below we recall here a specific case from [14]. Take $d = 1$ and let

$$\sigma(t,T) = g(r_t)\,e^{-\lambda(T-t)} \quad \text{with} \quad g(r) = \sigma_0\,|r|^\delta\,, \tag{9}$$

where r_t is the short rate and $\sigma_0, \delta, \lambda$ are parameters to be determined from market prices. It can be shown, see [14], that in this case the entire term structure can be expressed as driven by two forward rate processes $f(\cdot,T_1), f(\cdot,T_2)$ with maturities T_1, T_2 that may be chosen arbitrarily. One has in fact

$$p(t,T) = \exp\left\{-\bar{\alpha}_0(t,T) - \bar{\alpha}_1(t,T)f(t,T_1) - \bar{\alpha}_2(t,T)f(t,T_2)\right\} \tag{10}$$

for suitable functions $\bar{\alpha}_i : [0,T] \to \mathbb{R}$; $i = 0,1,2$. Choosing $T_1 = t$ and $T_2 = \tau > t$ arbitrary but fixed so that $f(t,T_1) = r_t, f(t,T_2) = f(t,\tau)$, one obtains (see always [14]) a Markovian system for $\mathbf{X}_t := (r_t, f(t,\tau))$ of the form

$$\begin{cases} dr_t &= \big(\beta_0(t) + \beta_1(t) r_t + \beta_2(t) f(t,\tau)\big)\, dt + g(r_t) dW_t \\ \\ df(t,\tau) &= \big(\gamma_0(t) + \gamma_1(t) r_t + \gamma_2(t) f(t,\tau)\big)\, dt + g(r_t) e^{-\lambda(\tau - t)} dW_t \end{cases} \tag{11}$$

for suitable time functions $\beta_i(t), \gamma_i(t)$, $i = 0,1,2$. The two-dimensional Markovian factor process $(r_t, f(t,\tau))$ drives now the entire term structure in the sense that

$$p(t,T) = \exp\{-\alpha_0(t,T) - \alpha_1(t,T) r_t - \alpha_2(t,T) f(t,\tau)\} \tag{12}$$

for suitable functions $\alpha_i(t,T)$ that correspond to the $\bar{\alpha}_i(t,T)$ in (10) for $T_1 = t, T_2 = \tau$.

An alternative way for constructing factor models is to specify a finite-dimensional Markovian factor process $\mathbf{X}$ and to represent the term structure in the form $p(t,T) = F^T(t,\mathbf{X}_t)$ for a suitable family of functions $F^T(t,\boldsymbol{x})$, $T \geq t$. In this way one ensures a-priori that the whole term structure evolves on a finite-dimensional manifold. A special case are the classical short-rate models where $\mathbf{X}$ is identified with the short rate r itself (r is then modeled as a Markov process), so that bond prices take the form $p(t,T) = F^T(t,r_t)$.

In order to exclude the possibility of arbitrage one has to impose appropriate conditions on the family $F^T(t,\boldsymbol{x})$, $T \geq t$. One way to proceed is to apply Itô's formula to $F^T(t,\mathbf{X}_t)$ and to derive dynamics for $p(t,T)$ and, via (2), the corresponding dynamics of $f(t,T)$. On the forward-rate dynamics one imposes the HJM drift condition, which leads to a PDE for $F^T(t,\boldsymbol{x})$, usually called *term structure equation*. One context where this PDE becomes relatively easily solvable by means of ordinary differential equations are the so-called *affine term structure models*. In the next example we present a special case; we shall come back to this example in our discussion of term structure models under incomplete information in Section 3.2 below.

Example 2.1 (Linear-Gaussian factor models)**.** On $(\Omega, \mathcal{G}, (\mathcal{G}_t), Q)$ consider an N-dimensional factor process $\mathbf{X}$ satisfying the linear-Gaussian dynamics

$$d\mathbf{X}_t = F\,\mathbf{X}_t dt + D\, d\mathbf{W}_t \tag{13}$$

with $\mathbf{W}$ an M-dimensional ($M \geq N$) $(Q, (\mathcal{G}_t))$-Wiener process and with F and D parametric matrices such that D has full rank. It can be shown that in this case the term structure is exponentially affine in $\mathbf{X}_t$, i.e.

$$p(t,T) = \exp\{A(t,T) - \mathbf{B}(t,T)\mathbf{X}_t\} \tag{14}$$

for deterministic functions $A(\cdot,T)\colon [0,T] \to \mathbb{R}$ and $\mathbf{B}(\cdot,T)\colon [0,T] \to \mathbb{R}^N$. It follows from the HJM drift condition that $A(\cdot,T)$ and $\mathbf{B}(\cdot,T)$ have to satisfy the following system of ODEs

$$\begin{cases} \frac{\partial}{\partial t}\mathbf{B}(t,T) + \mathbf{B}(t,T)\, F + \boldsymbol{b}(t) = 0 \\ \\ \frac{\partial}{\partial t} A(t,T) + \frac{1}{2}\mathbf{B}(t,T) D D' \mathbf{B}'(t,T) - a(t) = 0, \end{cases} \tag{15}$$

with terminal condition $A(T,T) = \mathbf{B}(T,T) = 0$. Here $\boldsymbol{b}(t)$ is a parametric function which has to be calibrated together with the matrices F and D; $a(t)$ is defined via $a(t) = f^*(0,t) +$

$\frac{1}{2}\int_0^t \beta_T(s,t)ds$ where $\beta(t,T) = \mathbf{B}(t,T)DD'\mathbf{B}'(t,T)$ and where $f^*(0,t)$ are the initially observed forward rates. For further use note that the log-prices are of the form

$$Y_t^T := \log\, p(t,T) = A(t,T) - \mathbf{B}(t,T)\mathbf{X}_t, \tag{16}$$

so that log-prices are affine functions of the factors. From $r_t = f(t,t)$ and $f(t,T) = -\frac{\partial}{\partial T}\log p(t,T)$ (see (2)) one immediately has that the short rate r_t can be expressed as a linear combination of the factors as-well. Applying Itô's formula one then obtains the following dynamics type

$$\begin{aligned} dr_t &= \left(\alpha_t^0 + \boldsymbol{\beta}_t^0\mathbf{X}_t\right)dt + \boldsymbol{\sigma}_t^0 d\mathbf{W}_t \\ dY_t^T &= \left(\alpha_t^T + \boldsymbol{\beta}_t^T\mathbf{X}_t\right)dt + \boldsymbol{\sigma}_t^T d\mathbf{W}_t \end{aligned} \tag{17}$$

for suitable coefficients. For a general discussion about affine term-structure models we refer to [24] or [6].

3 The term structure of interest rates. Incomplete information.

3.1 Pricing under incomplete information and nonlinear filtering

If the factor process $\mathbf{X}$ is observable, or equivalently, if we work under the global filtration $(\mathcal{G}_t)$, bond prices can be obtained in the form $p(t,T;\mathbf{X}_t) = F^T(t,\mathbf{X}_t)$. Moreover, in most cases of interest the function F^T can be computed explicitly. The picture changes if we assume that the information available to investors corresponds to a sub-filtration $\mathcal{F}_t \subset \mathcal{G}_t$ such that $\mathbf{X}$ is not $(\mathcal{F}_t)$-adapted. In the following lemma we show how to pass under the martingale pricing approach from the full information prices $p(t,T;\mathbf{X}_t)$ to arbitrage-free prices in the investor filtration $(\mathcal{F}_t)$; the latter will be denoted by $\hat{p}(t,T)$.

Lemma 3.1. *Let N be a given numeraire that is adapted to the investor filtration $(\mathcal{F}_t)$ and choose a corresponding martingale measure Q^N. Denote by $p(t,T;\mathbf{X}_t) = N_t E^{Q^N}\left(1/N_T \mid \mathcal{G}_t\right)$ arbitrage-free bond prices under full information, and by $\hat{p}(t,T) := N_t E^{Q^N}\left(1/N_T \mid \mathcal{F}_t\right)$ the corresponding arbitrage-free prices with respect to the investor filtration $(\mathcal{F}_t)$. Then one has that*

$$\hat{p}(t,T) = E^{Q^N}\left(p(t,T;\mathbf{X}_t) \mid \mathcal{F}_t\right). \tag{18}$$

In particular, if the savings account B is $(\mathcal{F}_t)$-adapted, we obtain $\hat{p}(t,T) = E^Q(p(t,T;\mathbf{X}_t) \mid \mathcal{F}_t)$.

Proof: By the very definition of a martingale, the fact that the bond prices at maturity T are equal to 1 and the assumption that $N_t \in \mathcal{F}_t$, for the first statement we have that

$$\hat{p}(t,T) = N_t\, E^{Q^N}\left(\frac{1}{N_T} \mid \mathcal{F}_t\right) = E^{Q^N}\left(N_t E^{Q^N}\left(\frac{1}{N_T} \mid \mathcal{G}_t\right) \mid \mathcal{F}_t\right) = E^{Q^N}\left(p(t,T;\mathbf{X}_t) \mid \mathcal{F}_t\right).$$

The second statement is then immediate.

Comments. The result can be extended to general $\mathcal{F}_T$-measurable claims and to credit-risky securities in an obvious way. The lemma shows that in order to obtain arbitrage-free prices in the investor filtration, one has to compute the conditional expectation in (18), which amounts to solving a filtering problem. In abstract terms the solution of this filtering problem is given by the *optional projection* of the process $(p(t,T;\mathbf{X}_t))_{t\le T}$ on $(\mathcal{F}_t)$; the latter is usually denoted by $\widehat{p(t,T;\mathbf{X}_t)}$, which motivates the notation $\hat{p}(t,T)$. Note moreover, that the conditional expectation in (18) has to be computed with respect to the chosen martingale measure, so that martingale pricing leads to filtering problems under the martingale measure Q^N (rather than the physical measure P).[6] Suppose finally that for a certain maturity $\bar{T}$ the price of the $\bar{T}$-bond is assumed to be observable (in mathematical terms, $(\mathcal{F}_t)$-adapted), and moreover equal to the model-value $(p(t,\bar{T};\mathbf{X}_t))_{t\le\bar{T}}$. In that case we obviously have $\hat{p}(t,\bar{T}) = p(t,\bar{T};\mathbf{X}_t)$ so that the filtered model is "automatically" calibrated to the observed bond price; we will encounter a specific example of this in the next subsection.

In the rest of this section we describe some specific models. Rather than aiming at a complete description of available results, we shall concentrate on a few illustrative examples that come mostly from our own activities in this field.

3.2 Filtering in affine factor models

In this subsection we discuss the application of the pricing principle from Lemma 3.1 in the context of the linear-Gaussian factor model of Example 2.1; our description is based on the analysis of [41] and [40]. We consider two different scenarios with incomplete information about the factor process $\mathbf{X}$. In both cases investors observe (possibly with noise) a finite number of yields $y(t,T_i) = -\frac{1}{T_i-t}\log p(t,T_i)$, $i=1,\cdots,n$, or equivalently the logarithmic bond prices $Y_t^i = \log p(t,T_i)$, and in addition the short rate. In the first scenario the observations of the yields and of the short rate are given by perturbed versions of the theoretical model values. In the second case it is assumed that model values can be observed exactly; however, the factor process $\mathbf{X}$ will be high-dimensional so that its current value $\mathbf{X}_t$ cannot be inferred from the observed model values. Both scenarios lead to a linear filtering problem; we shall also mention an extension to nonlinear filtering.

1. Filtering with observations given by perturbed model values. Recall the dynamics of the factor process $\mathbf{X}$, of the short-rate r and of the logarithmic bond-prices Y^i from (17). Here we assume that perturbed versions $\tilde{r}$ and $\tilde{Y}^i$ are observable; these perturbed versions are generated by adding independent Wiener-type observation noises v_t^i, $i=0,\cdots,n$ to the original processes. The investor filtration is thus given by

$$\mathcal{F}_t = \sigma\left(\tilde{r}_s, \tilde{Y}_s^i;\ s\le t,\ i=1,\cdots,n\right), \tag{19}$$

where state process $\mathbf{X}$ and observations $\tilde{r}, \tilde{Y}^1,\ldots,\tilde{Y}^n$ have the following dynamics (for $t < \min\{T_i\colon 1\le i\le n\}$)[7]

$$\begin{cases} d\mathbf{X}_t &= F\mathbf{X}_t dt + D\,d\mathbf{W}_t \\ d\tilde{r}_t &= \left(\alpha_t^0 + \boldsymbol{\beta}_t^0\mathbf{X}_t\right)dt + \boldsymbol{\sigma}_t^0 d\mathbf{W}_t + dv_t^0 \\ d\tilde{Y}_t^i &= \left(\alpha_t^i + \boldsymbol{\beta}_t^i\mathbf{X}_t\right)dt + \boldsymbol{\sigma}_t^i d\mathbf{W}_t + (T_i - t)dv_t^i\,;\quad i=1,\cdots,n\,; \end{cases} \tag{20}$$

[6]Filtering problems with respect to the physical measure will be discussed in Section 3.4 and 3.5 below.

[7]By adjusting appropriately the filter so that the log-prices of already matured bonds are not anymore taken into account, we may let t go also beyond $\min\{T_i\colon 1\le i\le n\}$.

the time-dependent volatility of the additional noise reflects the fact that bond-price volatility converges to zero as time approaches the maturity date of the bond. Since for the given model we are in the affine term structure context of (14), for the prices $\hat{p}(t,T)$ we have that

$$\hat{p}(t,T) = E(p(t,T;\, \mathbf{X}_t) \mid \mathcal{F}_t) = \exp(A(t,T))\, E\big(\exp(-\mathbf{B}(t,T)\mathbf{X}_t) \mid \mathcal{F}_t\big), \tag{21}$$

where the last term corresponds to the conditional moment generating function of $\mathbf{X}_t$ given $\mathcal{F}_t$. Since the filtering model in (20) is linear-Gaussian, the filter distribution is Gaussian as well so that, denoting its conditional mean and covariance by $\boldsymbol{m}_t$ and Σ_t respectively, from (21) one obtains

$$\hat{p}(t,T) = \exp\left\{ A(t,T) - \mathbf{B}(t,T)\boldsymbol{m}_t + \frac{1}{2}\mathbf{B}(t,T)\Sigma_t\mathbf{B}'(t,T) \right\} \tag{22}$$

For the given model the pricing under incomplete information can thus be accomplished by solving the system of ODEs in (15) and the Kalman filter corresponding to (20).

Taking a financial point of view this simple model is not completely satisfactory for the following two reasons: first, recall from Lemma 3.1 that formula (21) is justified if B is $(\mathcal{F}_t)$-adapted. In that case the short rate is strictly speaking $(\mathcal{F}_t)$-adapted as-well (since $r_t = \frac{d}{dt} \ln B_t$), contradicting (19). However, a very small amount of observation noise for B (which, from a practical point-of-view would still permit the use of Lemma 3.1) leads to a substantial observation noise for the short rate $r_t = \frac{d}{dt} \ln B_t$, so that the assumption that the short-rate cannot be observed perfectly can be defended. Second, there is also the problem that, for maturities T_i corresponding to liquid bonds, $\hat{p}(t,T_i)$ does in general not coincide with the observed values for these maturities (recall the third point in the comments directly after Lemma 3.1). In the next paragraph we discuss a variant of the model that overcomes these issues.

2. Filtering with exact observations of the theoretical prices. We assume now that the dimension N of the factor process $\mathbf{X}$ is strictly larger than the number of traded bonds with observable prices. This occurs for instance in the case when maturity-specific idiosyncratic factors are being added (see the situations considered in [41], [40]). In this case there is no need to add exogenous noise terms to justify a filtering setup, and the observation filtration is given by

$$\mathcal{F}_t = \sigma\left(r_s, Y_s^i\, ;\, s \le t,\ i = 1, \cdots, n\right)\,, \tag{23}$$

where, in line with (16) and (17),

$$r_t = a(t) + \boldsymbol{b}(t)\,\mathbf{X}_t \ \text{ and } \ Y_t^i = A(t,T_i) - \mathbf{B}(t,T_i)\,\mathbf{X}_t\,, \quad i = 1, \cdots, n\,. \tag{24}$$

While still linear, this is a degenerate filtering problem. Adapting a procedure from [30] we shall now reduce it to a non-degenerate problem via a change of coordinates. Recall that the observations $\mathbf{Y}_t := [r_t, Y_t^1, \cdots, Y_t^n]$ are *affine* functions of $\mathbf{X}_t$,

$$\mathbf{Y}_t = \boldsymbol{\mu}_t + M_t\,\mathbf{X}_t \tag{25}$$

for an appropriate $(n+1)$-vector $\boldsymbol{\mu}_t$ and some $(n+1, N)$-matrix M_t with $N > n+1$. Moreover, our assumptions on the linear Gaussian factor model in Example 2.1 ensure that M_t has full rank.

Introduce now some $(N-n-1,N)$ matrix L_t such that the $(N\times N)$-matrix $\begin{pmatrix} L_t \\ M_t \end{pmatrix}$ is invertible; this is always possible as M_t was assumed to have full rank. Define the $N-n-1$-dimensional process

$$\bar{\mathbf{X}}_t := L_t \mathbf{X}_t \tag{26}$$

and note that for appropriate matrices Φ_t and Ψ_t one has

$$\mathbf{X}_t = \begin{pmatrix} L_t \\ M_t \end{pmatrix}^{-1} \begin{pmatrix} \bar{\mathbf{X}}_t \\ \mathbf{Y}_t - \boldsymbol{\mu}_t \end{pmatrix} =: \Phi_t \bar{\mathbf{X}}_t + \Psi_t \left(\mathbf{Y}_t - \boldsymbol{\mu}_t\right) . \tag{27}$$

Using the linearity of the dynamics of $\mathbf{X}$ we can now derive a closed-form linear-Gaussian system for the pair $(\bar{\mathbf{X}}, \mathbf{Y})$. In fact, from (26), (13) and (27) it then follows

$$\begin{aligned} d\bar{\mathbf{X}}_t &= \dot{L}_t \mathbf{X}_t dt + L_t d\mathbf{X}_t = \left(\dot{L}_t + L_t F\right) \mathbf{X}_t dt + L_t D\, d\mathbf{W}_t \\ &= \left(\dot{L}_t + L_t F\right) \Phi_t \bar{\mathbf{X}}_t + \left(\dot{L}_t + L_t F\right) \Psi_t \mathbf{Y}_t - \left(\dot{L}_t + L_t F\right) \Psi_t \boldsymbol{\mu}_t + L_t D\, d\mathbf{W}_t \\ &=: \alpha_t\, \bar{\mathbf{X}}_t + \beta_t\, \mathbf{Y}_t + \gamma_t + \delta_t\, d\mathbf{W}_t\,, \end{aligned} \tag{28}$$

where $\alpha_t, \beta_t, \gamma_t, \delta_t$ are implicitly defined. Analogously, from (25), (13) and (27)

$$\begin{aligned} d\mathbf{Y}_t &= \dot{\boldsymbol{\mu}}_t dt + \dot{M}_t \mathbf{X}_t dt + M_t d\mathbf{X}_t = \left[\dot{\boldsymbol{\mu}}_t + \left(\dot{M}_t + M_t F\right) \mathbf{X}_t\right] dt + M_t D\, d\mathbf{W}_t \\ &= \left[\dot{\boldsymbol{\mu}}_t - \left(\dot{M}_t + M_t F\right) \Psi_t \boldsymbol{\mu}_t\right] dt + \left(\dot{M}_t + M_t F\right) \Phi_t \bar{\mathbf{X}}_t dt + \left(\dot{M}_t + M_t F\right) \Psi_t \mathbf{Y}_t dt + M_t D\, d\mathbf{W}_t \\ &=: \phi_t\, \bar{\mathbf{X}}_t + \psi_t\, \mathbf{Y}_t + \rho_t + \sigma_t\, d\mathbf{W}_t\,, \end{aligned} \tag{29}$$

where, again, $\phi_t, \psi_t, \rho_t, \sigma_t$ are implicitly defined. We can now formulate a non-degenerate filtering problem for the unobserved state variable process $\bar{\mathbf{X}}_t$ with observations $\mathbf{Y}_t$ as follows

$$\begin{cases} d\bar{\mathbf{X}}_t &= \alpha_t\, \bar{\mathbf{X}}_t + \beta_t\, \mathbf{Y}_t + \gamma_t + \delta_t\, d\mathbf{W}_t \\ d\mathbf{Y}_t &= \phi_t\, \bar{\mathbf{X}}_t + \psi_t\, \mathbf{Y}_t + \rho_t + \sigma_t\, d\mathbf{W}_t \end{cases} \tag{30}$$

This system is of the linear, conditionally Gaussian type and it leads thus to a Gaussian conditional (filter) distribution that we denote by $\pi_{\bar{\mathbf{X}}_t|\mathcal{F}_t} = \mathcal{N}\left(\bar{\mathbf{X}}_t; \bar{\boldsymbol{m}}_t, \bar{P}_t\right)$, and where the mean $\bar{\boldsymbol{m}}_t$ and covariance $\bar{P}_t$ can be computed via the Kalman filter. We then have from Lemma 3.1 that

$$\begin{aligned} \hat{p}(t,T) &= E^Q\left(p(t,T;\mathbf{X}_t) \mid \mathcal{F}_t\right) = E^Q\left(p\left(t,T;\left(\Phi_t \bar{\mathbf{X}}_t + \Psi_t \left(\mathbf{Y}_t - \boldsymbol{\mu}_t\right)\right)\right) \mid \mathcal{F}_t\right) \\ &= \int p\big(t,T;\left(\Phi_t \bar{\boldsymbol{x}} + \Psi_t \left(\mathbf{Y}_t - \boldsymbol{\mu}_t\right)\right)\big)\, \pi_{\bar{\mathbf{X}}_t|\mathcal{F}_t}(d\bar{\boldsymbol{x}}). \end{aligned} \tag{31}$$

Nonlinear extensions. The setup in the example of this subsection can be generalized in various ways as is indicated by the following two dual setups. For the first setup one keeps the linear-Gaussian dynamics (13) for the factors, but instead of (14) one considers an exponentially quadratic term structure model of the form

$$p(t,T) = \exp\left[A(t,T) - \mathbf{B}(t,T)\mathbf{X}_t - \mathbf{X}_t' C(t,T)\mathbf{X}_t\right] \tag{32}$$

Notice that, for a linear-Gaussian factor model as in (13), more general exponentially polynomial term structure models lead to arbitrage for a degree larger that two (see [29]) so that (32) represents the most general nonlinear generalization of (14) that does not lead to arbitrage. For the second setup one keeps the exponentially affine structure (14) but considers instead of (13) a scalar square root process of the form

$$dX_t = F(X_t - b_t)\,dt + \sqrt{X_t}\,D\,dW_t\,. \tag{33}$$

With these nonlinear extensions of the model the filtering problem with perturbed observations of the state becomes nonlinear; it seems that a finite-dimensional filter does not exist. The second (degenerate) filtering problem is even more challenging, since the solution-approach described above does not extend to the nonlinear case.

3.3 Constructing term structure models via nonlinear filtering

In [48] the innovations approach to nonlinear filtering is used in order to construct a factor model for bond prices; here we sketch a simplified version of the approach. The author studies a model where the short rate dynamics under a martingale measure are of the form

$$dr_t = a(t, r_t, X_t)\,dt + b\,dW_t \tag{34}$$

with X a scalar finite-state Markov chain with state space $\{1, \dots, K\}$. In [48] the process X is assumed to be unobservable; the investor filtration is given by $\mathcal{F}_t = \sigma\,(r_s\,;\, s \le t)$, so that only the short rate is observable. As before, the bond-pricing problem is approached via a two-step procedure: first one determines the bond prices under full observation. Given the Markovianity of the pair (r, X), these prices are of the form $p(t,T) = F^T(t; r_t, X_t)$ with the function $F^T(\cdot)$ such that the resulting prices do not allow for the possibility of arbitrage. According to Lemma 3.1, bond prices under incomplete information are then given by

$$\hat{p}(t,T) = E\left(F^T(t; r_t, X_t) \mid \mathcal{F}_t\right) =: \pi_t F^T\,. \tag{35}$$

Instead of first determining the filter distribution $\pi_{X_t|\mathcal{F}_t}$, in [48] the author determines directly the dynamics of the filtered value $\pi_t F^T$ of the bond prices. Using Itô's formula, she obtains first the semimartingale representation of the full-information bond price $F^T(t; r_t, X_t)$. From there, following the innovations approach to nonlinear filtering, she then obtains directly the dynamics of the filtered bond prices $\pi_t F^T$ decomposed into a finite variation part and a term driven by the innovations process

$$\bar{W}_t = \frac{1}{b}\Big(r_t - \int_0^t \pi_s\big(a(s, r_s, X_s)\big)\,ds\Big)\,,\ \text{where } \pi_s(a(s, r_s, X_s)) := \int a(s, r_s, x)\pi_{X_s|\mathcal{F}_s}(dx).$$

Let $p_t^k = Q(X_t = k \mid \mathcal{F}_t)$, $1 \le k \le K$. Since $\hat{p}(t,T) = \pi_t F^T = \sum_{k=1}^K p_t^k F^T(t, r_t; k)$, the ensuing term structure model has a natural factor structure with factor given by $\boldsymbol{p}_t := (p_t^1, \dots, p_t^K)$. The dynamics of the factor vector $\boldsymbol{p}$ (which summarizes the conditional distribution $\pi_{X_t|\mathcal{F}_t}$) can be computed via the Wonham filter (see [59] or [27]).

The idea of using nonlinear filtering for the construction of a term structure model is undoubtedly very elegant; a similar approach in the context of credit risk models is discussed in the third example of Section 6.3 below. However, from a financial point of view the assumption that investors observe only the short rate is somewhat problematic: bonds with certain prominent maturities are usually liquidly traded, so that one would like to calibrate the model also to bond-price information.

3.4 Filtering of the market price of risk

Filtering in mathematical finance can be performed also for econometric- and risk-management applications where it is usually most appropriate to study the filtering problem under the physical measure. If in that case the observations include prices that are expressed as expectations under a martingale measure, one ends up with a situation where one has to work simultaneously with the physical measure and with a martingale measure. The obvious thing is then to express everything under the same measure. Since the martingale measure serves mainly the purpose of guaranteeing absence of arbitrage, it is most natural to express everything under the physical measure.

As an example we start from the SDE-system (11) for the short rate and some instantaneous forward rate, defined under a martingale measure Q so that absence of arbitrage is guaranteed. To transform the system into an equivalent one under the physical measure $P \sim Q$, we introduce the integrable and adapted *market price of risk* process ψ_t that allows one to pass from Q to P in the sense that, using now the symbol W_t^Q to specify a Wiener process under Q, the process $W_t := W_t^Q - \int_0^t \psi_s ds$ is a Wiener process under P (*Girsanov measure transformation*). Using a mean-reverting diffusion model for the evolution of ψ_t under P, the system (11) extends then to the following system defined under the physical measure P

$$\begin{cases} dr_t &= \left[\beta_0(t) + \beta_1(t) r_t + \beta_2(t) f(t,\tau) + g(r_t)\psi_t\right] dt + g(r_t) dW_t \\ df(t,\tau) &= \left[\gamma_0(t) + \gamma_1(t) r_t + \gamma_2(t) f(t,\tau) + g(r_t) e^{-\lambda(\tau-t)}\psi_t\right] dt + g(r_t) e^{-\lambda(\tau-t)} dW_t \\ d\psi_t &= \kappa(\bar{\psi} - \psi_t) dt + b \mid \psi_t \mid^{\gamma} dW_t \end{cases} \tag{36}$$

with the totality of the parameters given by the vector $(\sigma_0, \delta, \lambda, \kappa, \bar{\psi}, b, \gamma)$. A filter application in this context can be found in [15]. There the unobserved state vector is $\mathbf{X}_t = [r_t, f(t,\tau), \psi_t]$, while the observations are noisy observations of a finite number of given forward rates. For further aspects in this context see [56] and [55]. Notice finally that by filtering the market price of risk, this quantity (and hence also the corresponding martingale measure) continuously adapts to the current market situation.

3.5 Parameter Estimation in Term-Structure Models

Market models are mostly specified as families of models that depend on certain parameters. The parameters are usually identified by matching as best as possible the theoretical model prices with the actually observed market prices. This goes under the name of *calibration to the market.* Calibration leads to a form of point estimation that may however lead to unstable estimates and without indication of their accuracy. In a filtering context one may instead consider a dynamic parameter estimation as part of the filtering problem and such a dynamic estimation enhances the possibility for the model to continuously adapt to the current market situation.

Two major approaches to this effect may be considered: i) *combined filtering and parameter estimation*; ii) *EM (expectation maximization)* combined with filtering. In the approach via combined filtering and parameter estimation one considers an extended state $(\mathbf{X}_t, \theta)$ where θ denotes the vector of parameters that are now considered as random variables according to the Bayesian point of view and one determines recursively the joint conditional (filter) distribution $\pi_{(\mathbf{X}_t,\theta)|\mathcal{F}_t}$. An example for this approach is presented next.

Combined filtering and parameter estimation with interest-rate observations ([4]). As explained in Section 2, in the context of HJM-models with a volatility structure as in (9), forward rates and bond prices follow a factor model with factor process $\mathbf{X}$ given by two instantaneous rates. Instantaneous (continuously compounded) forward rates are a mathematical abstraction and cannot be directly observed on the market (at most proxies are observable). Simple (discretely compounded) rates such as the LIBOR rates on the other hand are regularly quoted on interest markets so that they can be considered observable. The latter are related to the bond prices via (1), which in turn are related to $\mathbf{X}_t$ via $p(t,T;\mathbf{X}_t) = F^T(t,\mathbf{X}_t)$. Since $\mathbf{X}_t = [r_t, f(t,\tau)]$ satisfies the diffusion model (11), by stochastic differentiation one can then derive stochastic dynamics for the LIBOR rates. In this context in [4] a model is studied where the observation filtration $(\mathcal{F}_t)$ is generated by noisy observations of LIBOR rates. More precisely, by adding an independent observation noise to the LIBOR rates the authors in [4] obtain a non-degenerate nonlinear filtering problem to estimate $\mathbf{X}_t = [r_t, f(t,\tau)]$ and the parameters $(\sigma_0, \delta, \lambda)$ of the volatility function $\sigma(t,T)$ in (9), i.e. to estimate the theoretical instantaneous rates, on the basis of the observations of the LIBOR rates.

Parameter estimation via the EM algorithm. The EM algorithm is based on the following: let a given family of models, parameterized by θ, induce a family of probability measures P^θ that are assumed to be absolutely continuous with respect to a given reference measure P^0. Putting

$$Q(\theta,\theta') := E_{\theta'}\left\{\log \frac{dP^\theta}{dP^{\theta'}} \mid \mathcal{F}_t\right\} \tag{37}$$

the algorithm iterates through the following two steps:

i) compute $Q(\theta,\theta')$ for θ' given, θ arbitrary (*expectation step*)

ii) determine $\theta^* = argmax_\theta\, Q(\theta,\theta')$ and return to i) with $\theta' = \theta^*$ (*maximization step*).

The algorithm stops as soon as the maximizing values in two successive iterations are sufficiently close.

Since the EM algorithm is based on an absolutely continuous change of measure, the parameters entering the coefficient of the observation noise cannot be estimated via EM and have to be estimated by other methods, e.g. on the basis of the empirical quadratic variation. The other parameters can in principle be estimated via EM and the maximization step leads to solving the system of equations obtained by putting $\frac{\partial Q(\theta,\theta')}{\partial\theta} = 0$. The resulting system involves various conditional expectations that can be computed on the basis of the filtering results (see e.g.[28]): in continuous time, if the state and observation noises are independent, filtering alone suffices; if they are not independent, also smoothing is required.

There exist other approaches as well, in particular in a discrete time setup. One of them is based on the maximization of the innovations likelihood, which is in fact of the type of maximum likelihood estimation. The parameter estimation approaches are mentioned here only in the context of term structure models; they can however be easily carried over also to credit risk models (see e.g. [30]).

4 Nonlinear filtering in credit risk models

In this section we give a brief introduction to dynamic credit risk models and explain how incomplete information and nonlinear filtering enter in credit risk modeling; a detailed discussion of specific models is given in Sections 5 and 6 below.

4.1 Dynamic Credit Risk Models and Credit Derivatives

Dynamic credit risk models are concerned with the modeling of the default times in a given portfolio of firms. In our discussion of credit risk models we use the following notation: the firms under consideration are indexed by $i \in \{1, \dots, m\}$; the random time $\tau_i > 0$ denotes the default time of firm i; the current default state of firm i is described by the default indicator process $Y_{t,i} = \mathbf{1}_{\{\tau_i \le t\}}$, jumping from zero to one at $t = \tau_i$; the current default state of the portfolio is described by $\mathbf{Y}_t = (Y_{t,1}, \dots, Y_{t,m})$; the *default history* up to time t is given by $\mathcal{F}_t^{\mathbf{Y}} := \sigma(\mathbf{Y}_s \colon s \le t)$. Since our focus is primarily on pricing problems, we model the dynamics of the objects of interest directly under some risk-neutral measure Q. Throughout we therefore work on a filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t), Q)$; as in the interest-rate part, $(\mathcal{G}_t)$ represents the full-information filtration, so that all stochastic processes considered will be $(\mathcal{G}_t)$-adapted. Moreover, in line with most of the credit risk literature, we assume in this part of the paper that default-free interest rates are deterministic and equal to $r > 0$.

A large part of the credit risk literature is concerned with the pricing of *credit derivatives*. These are securities whose payoff is linked to default events in a given reference portfolio. In abstract terms the payoff of a credit derivative is thus given by some $\mathcal{F}_T^{\mathbf{Y}}$-measurable random variable H. Important examples include defaultable zero-coupon bonds and default payments. The payoff of a defaultable zero coupon bond issued by firm i with maturity T and zero recovery is given by $H = \mathbf{1}_{\{\tau_i > T\}} = 1 - Y_{T,i}$; the price at $t < T$ of this bond will be denoted by $p_i(t,T)$. A default payment of size δ on firm i with maturity T has a payoff of size δ directly at τ_i, provided $\tau_i \le T$. By combining zero coupon bonds and default payments other important products such as Credit Default Swaps (CDSs) or corporate bonds with recovery payments can be constructed; see Section 9.4 of [51] for further details. An important quantity in this context is the *credit spread* of firm i, denoted $c_i(t,T)$, $t \le T \wedge \tau_i$. This quantity measures the difference in the continuously compounded yields of a defaultable zero coupon bond issued by firm i and of the corresponding default-free zero coupon bond (denoted here by $p_0(t,T)$), and reflects thus the market's assessment of the likelihood of the default of firm i. Formally, $c_i(t,T)$ is given by

$$c_i(t,T) := -\frac{1}{T-t} \left(\log p_i(t,T) - \log p_0(t,T) \right) . \tag{38}$$

Existing dynamic credit risk models can be grouped into two classes: *structural* and *reduced-form models*. Structural models originated from Black and Scholes [8], Merton [52], and Black and Cox [7]. Important contributions to the literature on reduced-form models are [45], [49] [26] and [9]; a more complete list of references can be found in the textbooks [3], [50], [51], among others.

In structural models one starts by modeling the asset values $V_i = (V_{t,i})_{t \ge 0}$ of the firms under consideration; usually V_i is modeled as a diffusion process. Given some default barrier $K_i = (K_{t,i})_{t \ge 0}$, the default time τ_i is then defined to be the first passage time of V_i at the barrier

K_i, i.e.

$$\tau_i = \inf\{t \geq 0\colon V_{t,i} \leq K_{t,i}\}\,. \qquad (39)$$

The default barrier is often interpreted as the value of the liabilities of the firm; with this interpretation (39) states that, in line with economic intuition, default happens at the first time that the asset value of a firm is too low to cover its liabilities. Note that the default time τ_i defined in (39) is a predictable stopping time with respect to the global filtration $(\mathcal{G}_t)$ to which V_i and K_i are adapted. It is well-documented that the fact that τ_i is $(\mathcal{G}_t)$-predictable leads to very low values for short-term credit spreads (in particular $\lim_{h\to 0} c_i(t, t+h) = 0$), contradicting most of the available empirical evidence.

In reduced-form models on the other hand, the precise mechanism leading to default is left unspecified; rather one models directly the law of the default times τ_i or of the associated default indicator process Y_i. Typically τ_i is modeled as a totally inaccessible stopping time with respect to the global filtration $(\mathcal{G}_t)$, admitting a $(Q, (\mathcal{G}_t))$-intensity λ_i (termed *risk-neutral default intensity*). Formally, $\lambda_i = (\lambda_{t,i})_{t\geq 0}$ is a $(\mathcal{G}_t)$-predictable process such that

$$Y_{t,i} - \int_0^{t\wedge\tau_i} \lambda_{s,i} ds \quad \text{is a} \quad (Q, (\mathcal{G}_t))\text{-martingale.} \qquad (40)$$

In reduced-form models dependence between defaults is often generated by assuming that the default intensities do depend on a common factor process $\mathbf{X}_t \in \mathbb{R}^d$, i.e. $\lambda_{t,i} = \lambda_i(\mathbf{X}_t)$ for suitable functions $\lambda_i\colon \mathbb{R}^d \to (0, \infty)$. The simplest construction is that of *conditionally independent, doubly stochastic default times.* Here it is assumed that given $\mathcal{F}^{\mathbf{X}}_\infty$, the τ_i are conditionally independent with

$$P(\tau_i > t \mid \mathcal{F}^{\mathbf{X}}_\infty) = \exp(-\int_0^t \lambda_i(\mathbf{X}_s) ds), \quad t > 0;$$

see for instance Section 9.6 of [51] for details.

4.2 Incomplete Information

In both modeling paradigms it makes sense to assume that investors have imperfect information on some of the state variables of the models; this has given rise to a rich literature on credit risk models under incomplete information.

In a structural model the natural state variables are given by the asset value V_i or the log-asset value X_i and - with stochastic liabilities - by the liability-levels K_i of the firms under consideration. It is difficult for investors in secondary markets to precisely assess the values of these state variables for a number of reasons: accounting reports might be noisy; market- and book-values can differ as intangible assets such as R&D (research and development)-results or client-relationships are difficult to value; part of the liabilities are usually bank loans whose precise terms are unknown to the public; and many more. Hence, starting with the seminal work of Duffie and Lando [25], a growing literature studies models where investors have only noisy information about V_i and/or K_i; the conditional distribution of the state variables given investor information $\mathcal{F}_t$ is then computed by Bayesian updating or filtering arguments. Examples of this line of research include [25], [53]; [39], [16] and [36]; some of these papers are discussed in more detail in Section 5 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 e.g. Jarrow and Protter [44]): in the models mentioned above the default time

τ_i that is predictable with respect to the global filtration $(\mathcal{G}_t)$ becomes totally inaccessible with respect to the investor filtration $(\mathcal{F}_t)$ and moreover admits an intensity. This leads furthermore to a realistic behavior of short-term credit spreads, as is explained in Section 5.

In typical reduced-form models default intensities are assumed to depend on some Markovian factor process $\mathbf{X}$ which here becomes the natural state variable process. In applications $\mathbf{X}$ is usually not identified with observable quantities but treated as a latent process whose current value must be inferred from observables such as prices or the default history. A theoretically consistent way for doing this is to determine - via Bayesian updating or filtering arguments - $\pi_{\mathbf{X}_t|\mathcal{F}_t}$, the conditional distribution of $\mathbf{X}_t$ given investor information $\mathcal{F}_t$. Reduced-form credit risk models with incomplete information include the contributions by [57], [17] and [22] as well as our own work [35] and [37]. The structure of the models [57], [17] and [22] is relatively similar: default intensities are driven by an unobservable factor $\mathbf{X}$; the default times are conditionally independent, doubly stochastic random times; the investor information $(\mathcal{F}_t)$ is given by the default history of the portfolio, augmented by economic covariates. In [57], and [17] the unobservable factors are modeled by a static random vector $\mathbf{X}$ which is termed *frailty*; the conditional distribution $\pi_{\mathbf{X}|\mathcal{F}_t}$ is determined via Bayesian updating. In [22] the unobservable (scalar) factor X is modeled as an Ornstein-Uhlenbeck process. This latter paper has an empirical focus: dynamic Bayesian methodology is used in order to estimate the model-parameters from historical default data; moreover, filtering is used in order to determine the conditional mean of X_t, given the history of defaults and covariates. This analysis provides strong evidence for the assertion that an unobservable stochastic process driving default intensities (a so-called dynamic frailty) is needed on top of observable covariates in order to explain the clustering of defaults in historical data, a finding which strongly supports the use of filtering methodology in credit risk models.

Our own work [35] on filtering in reduced-form credit risk models extends these contributions in a number of ways, at least from a methodological viewpoint. To begin with, we consider a more general investor filtration that contains noisily observed prices on top of the default history of the portfolio. Moreover, the problem of finding the conditional distribution of $\pi_{\mathbf{X}_t|\mathcal{F}_t}$ is studied in a general jump-diffusion model for $\mathbf{X}$ and default indicator $\mathbf{Y}$ that includes most reduced-form credit risk models from the literature and in particular the analysis of [57], [17] and [22] as special case. Our discussion of reduced-form models with unobservable state variables in Section 6 is therefore based mainly on [35] and the companion paper [37].

Introducing incomplete information into credit portfolio models has interesting implications for the dynamics of credit derivative prices and credit spreads (both in structural and in reduced-form models), since the successive updating of the conditional distribution $\pi_{\mathbf{X}|\mathcal{F}_t}$ in reaction to incoming default observations generates so-called *information-driven default contagion*: the news that some obligor has defaulted leads to an update in $\pi_{\mathbf{X}_t|\mathcal{F}_t}(d\boldsymbol{x})$ and hence to a jump in the $(\mathcal{F}_t)$-default intensity of the surviving firms, as will be explained in more detail below. In the context of reduced-form models this was first pointed out by [57] and [17], whereas default contagion in structural models is studied among others in [39]; empirical evidence for contagious effects is provided for instance in [17]. Note that a similar phenomenon did not occur in our discussion of interest-rate market models, essentially because there the investor information $(\mathcal{F}_t)$ was generated by continuous (Wiener-driven) processes.

5 Filtering in Structural Models

In this section we discuss structural credit risk models under incomplete information and some of the ensuing nonlinear filtering problems.

5.1 The model of Duffie and Lando [25]

The setup. Recall that we work on a filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t), Q)$, Q the risk-neutral measure and $(\mathcal{G}_t)$ the full-information filtration. Throughout this section we focus on models for the default of a single firm, so that the index i giving the identity of the firm can be omitted. We assume that the asset value V follows a geometric Brownian motion on this filtered probability space with drift μ, volatility σ and initial value V_0. Consider then as (scalar) state variable

$$X_t := \log V_t = X_0 + \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t, \tag{41}$$

W a Browninan motion on $(\Omega, \mathcal{G}, (\mathcal{G}_t), Q)$. In [25] the default barrier K is taken constant. The *default time* is thus given by the stopping time

$$\tau := \inf\{t \geq 0 \colon V_t < K\} = \inf\{t \geq 0 \colon X_t < \log K\}. \tag{42}$$

It is assumed that V is not directly observable. Rather, investors observe default; moreover, they receive "noisy accounting reports" at deterministic times $t_1, t_2, \cdots$, that is they observe random variables $Z_i = X_{t_i} + U_i$ where $(U_i)_{i \in \mathbb{N}}$ is a sequence of independent, normally distributed random variables, independent of X (or V). Formally, with $Y_t := \mathbf{1}_{\{\tau \leq t\}}$, the investor filtration is

$$\mathcal{F}_t := \mathcal{F}_t^Y \vee \sigma(\{Z_i \colon t_i \leq t\}). \tag{43}$$

The default barrier K and the initial asset value X_0 are supposed to be known.

Survival probabilities, default intensity and credit spreads. By the Markov property of V (or X) one has, for $T \geq t$,

$$\begin{aligned} Q(\tau > T \mid \mathcal{G}_t) &= \mathbf{1}_{\{\tau > t\}} Q\left(\inf_{s \in (t,T)} V_s > K \mid \mathcal{G}_t\right) = \mathbf{1}_{\{\tau > t\}} Q\left(\inf_{s \in (t,T)} V_s > K \mid V_t\right) \\ &=: \mathbf{1}_{\{\tau > t\}} \bar{F}_\tau(t, T, V_t). \end{aligned}$$

Note that for $T \geq t$ the mapping $T \mapsto \bar{F}_\tau(t, T, v)$ gives the (risk-neutral) survival probabilities of the firm under full information as of time t, given that $V_t = v$; $\bar{F}_\tau$ is easily computed using standard results on the first passage time of Brownian motion with drift. Using iterated conditional expectations one gets for the survival probability in the investor filtration

$$Q(\tau > T \mid \mathcal{F}_t) = E(Q(\tau > T \mid \mathcal{G}_t) \mid \mathcal{F}_t) = \mathbf{1}_{\{\tau > t\}} \int_{\log K}^{\infty} \bar{F}_\tau(t, T, e^x)\, \pi_{X_t|\mathcal{F}_t}(dx). \tag{44}$$

Next turn to the $(Q, \mathcal{F}_t)-$*default intensity* λ_t of τ. It can be shown that under some regularity conditions one has

$$\lambda_t = \lim_{h \downarrow 0} \frac{1}{h} Q\{t < \tau \leq t + h \mid \mathcal{F}_t\}, \tag{45}$$

provided this limit exists for all $t \geq 0$ almost surely (see [10], [1] for details). Duffie and Lando in [25] now show that such a λ_t exists and is given by

$$\lambda_t = \frac{1}{2}\sigma^2 \frac{\partial}{\partial x}\pi(X_t \in dx \mid \mathcal{F}_t)_{|x=\log K}, \quad \tau \geq t,$$

where $\pi(X_t \in dx \mid \mathcal{F}_t)$ denotes the Lebesgue-density of the filter distribution $\pi_{X_t|\mathcal{F}_t}(dx)$. (The fact that the derivative of the conditional density exists at $x = \log K$ is part of their result.)

Finally we discuss bond prices and credit spreads in the Duffie-Lando model under incomplete information. We get for the price of a defaultable zero-coupon bond with zero recovery, denoted by $p_1(t,T)$,

$$p_1(t,T) = \mathbf{1}_{\{\tau>t\}} e^{-r(T-t)} Q\{\tau > T \mid \mathcal{F}_t\} = \mathbf{1}_{\{\tau>t\}} e^{-r(T-t)} \int_{\log K}^{\infty} \bar{F}_\tau(t,T,e^x)\, \pi_{X_t|\mathcal{F}_t}(dx) \tag{46}$$

i.e. zero-coupon bond prices can be expressed as an average with respect to the filter distribution $\pi_{X_t|\mathcal{F}_t}(dx)$. The price of a default payment is also easily computed once the survival probability in the investor filtration is at hand. These pricing results are of course special cases of the general pricing principle from Lemma 3.1 in Section 3.

The credit spread $c(t,T)$ introduced in (38) satisfies on $\{\tau > t\}$ the following relation (since r is assumed deterministic one has $p_0(t,T) = e^{-r(T-t)}$)

$$c(t,T) = \frac{-1}{T-t} \log Q(\tau > T \mid \mathcal{F}_t)\,. \tag{47}$$

In particular, we get

$$\lim_{T \downarrow t} c(t,T) = -\frac{\partial}{\partial T} \log Q\{\tau > T \mid \mathcal{F}_t\}_{|T=t} = \lambda_t$$

where the second equality follows from (45). This shows that the introduction of incomplete information typically leads to non-vanishing short-term credit spreads.

Computing the filter distribution. We have seen that in order to determine risk sensitive financial quantities such as defaultable bond prices or credit spreads, one needs to determine the conditional distribution (*filter distribution*) $\pi_{X_t|\mathcal{F}_t}(dx)$. In [25] this problem is tackled in an elementary way, involving Bayes' formula and properties of first passage time of Brownian motion. We do not discuss the details here; in the next subsection we show how proper filtering arguments can be used in order to determine (approximately) $\pi_{X_t|\mathcal{F}_t}(dx)$.

5.2 The model of Frey & Schmidt [36]

In [36], the basic Duffie-Lando model is extended in essentially two directions. On the financial side the paper introduces dividend payments and discusses the pricing of the firm's equity under incomplete information. On the mathematical side nonlinear filtering techniques and Markov-chain approximations are employed in order to determine the conditional distribution of the log-asset value X_t given the investor-information $\mathcal{F}_t$.

Here we concentrate on the filtering part. The setup of the model under full information is as in Subsection 5.1: the log-asset value X is given by the arithmetic Brownian motion (41) and, in line with (42), the default time τ is the first passage time of X at the barrier $\log K$.[8] Investors

[8] Below we present a slightly simplified version of the model discussed in [36].

observe the default state of the firm; moreover, they receive pieces of economic information (news) related to the state of the company such as information 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 (for simplicity) deterministic 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 $\{\ell_1, \ldots, \ell_{M^I}\}$. The conditional distribution of I_n given $\mathcal{G}_{t_n^I}$ is denoted by

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

Summarizing, the information of secondary market investors at time t is given by the σ-field

$$\mathcal{F}_t := \mathcal{F}_t^Y \vee \sigma\Big(\{I_n \colon t_n^I \leq t\}\Big). \tag{48}$$

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 X^Δ as follows: define for a given time discretization $\Delta > 0$ 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 $\Xi^\Delta = \{m_1^\Delta, \ldots, m_{M^\Delta}^\Delta\}$ and transition probabilities p_{ij}^Δ, $1 \leq i, j \leq M^\Delta$, and define the induced process X^Δ by $X_t^\Delta = X_k^\Delta$ for $t \in [t_k^\Delta, t_{k+1}^\Delta)$. In [36] 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 $\Delta \to 0$.

In the sequel we keep Δ fixed and omit it from our notation. Obviously the conditional distribution $\pi_{X_{t_k}^\Delta|\mathcal{F}_{t_k}}$ is summarized by the probability vector $\boldsymbol{\pi}(k) = (\pi_1(k), \ldots, \pi_{M^\Delta}(k))$ with $\pi_j(k) := Q\big(X_k = m_j \mid \mathcal{F}_{t_k}\big)$. It is possible to give explicit recursive updating rules for the probability vector $\boldsymbol{\pi}(k)$. In fact, due to the discrete nature of the problem, this is fairly easy, as is illustrated by the simple proof of Proposition 5.1 below. It will be convenient to formulate the updating rule in terms of "unnormalized probabilities" $\boldsymbol{\sigma}(k) \propto \boldsymbol{\pi}(k)$ ($\propto$ standing for proportional to); the vector $\boldsymbol{\pi}(k)$ can then be obtained by normalization.

The initial filter distribution $\boldsymbol{\pi}(0)$ can be inferred from the (known) initial distribution of X_0. For $k \geq 1$ we have the following updating rule.

Proposition 5.1. *For $k \geq 1$ and $t_k < \tau$, denote by $N_k^I := \{n \in \mathbb{N} : t_{k-1} < t_n^I \leq t_k\}$ the set of indices of news arrivals in the period $(t_{k-1}, t_k]$. Then, with the convention that $\prod_\emptyset$ (the product over the empty set) equals 1,*

$$\sigma_j(k) = \mathbf{1}_{\{m_j > \log K\}} \sum_{i=1}^{M^\Delta} \Big\{ p_{ij}\, \sigma_i(k-1) \prod_{n \in N_k^I} \nu_I\big(I_n | m_i\big) \Big\}, \quad j = 1, \ldots, M^\Delta. \tag{49}$$

Proof. Given the new information arriving in $(t_{k-1}, t_k]$, the updating rule (49) forms a linear and in particular a positively homogeneous mapping Γ such that $\boldsymbol{\sigma}(k) = \Gamma\boldsymbol{\sigma}(k-1)$. Hence it is enough to show that $\boldsymbol{\pi}(k) \propto \Gamma\boldsymbol{\pi}(k-1)$. In order to compute $\boldsymbol{\pi}(k)$ from $\boldsymbol{\pi}(k-1)$ and the new information in $(t_{k-1}, t_k]$ we proceed in two steps. In Step 1 we compute (up to proportionality) an auxiliary vector of probabilities $\tilde{\boldsymbol{\pi}}(k-1)$ with

$$\tilde{\pi}_i(k-1) = Q\big(X_{k-1} = m_i \mid \mathcal{F}_k^-\big),\ 1 \leq i \leq M^\Delta, \tag{50}$$

where $\mathcal{F}_k^- := \mathcal{F}_{t_{k-1}} \vee \sigma\big(\{I_n \colon n \in N_k^I\}\big)$. 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) $\boldsymbol{\pi}(k)$ from the auxiliary probability vector $\tilde{\boldsymbol{\pi}}(k-1)$ 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

$$
\begin{aligned}
Q\left(X_k = m_j \mid \mathcal{F}_{t_k}\right) &\propto Q\left(X_k = m_j, X_k > \log K \mid \mathcal{F}_k^-\right) \\
&= \sum_{i=1}^{M^\Delta} Q\left(X_k = m_j, X_k > \log K, X_{k-1} = m_i \mid \mathcal{F}_k^-\right) \\
&= \mathbf{1}_{\{m_j > \log K\}} \sum_{i=1}^{M^\Delta} p_{ij}\, \tilde{\pi}_i(k-1)\,. \qquad (51)
\end{aligned}
$$

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 $\prod_{n \in N_k^I} \nu_I\big(I_n | m_i\big)$, and we obtain

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

Combining this with equation (51) gives the result. □

5.3 Further related work

There is a rich literature on structural credit risk models under incomplete information; here we briefly discuss some contributions which cannot be treated in detail for reasons of space.

The filtering model of Nakagawa [53]. In [53] the author considers a slightly generalized version of the basic Duffie-Lando model (see Subsection 5.1). The main difference is that in [53] the investor filtration is given by $\mathcal{F}_t = \mathcal{F}_t^Y \vee \mathcal{F}_t^Z$, where the process Z has dynamics $dZ_t = a(X_t)dt + d\beta_t$ for a Brownian motion β independent of W (observations of the state in additive Gaussian noise). The principal goal is to determine the form of the default intensity $(\lambda_t)_{t \geq 0}$ with respect to the investor filtration $(\mathcal{F}_t)$. Note that this is a non-standard filtering problem, as the default time τ does not admit an intensity with respect to the global filtration $(\mathcal{G}_t)$. In order to deal with this problem the author applies an equivalent change of measure, so that under the new measure $\tilde{Q}$ the process Z is independent of X and Y. The $(\tilde{Q}, (\mathcal{F}_t))$-intensity of τ can be computed explicitly using results from [47] or [25]. The $(Q, (\mathcal{F}_t))$-default intensity can then be computed via a suitable Girsanov-theorem for point processes once an explicit martingale representation of the density martingale $L_t = E^{\tilde{Q}}\big(dQ/d\tilde{Q} \mid \mathcal{F}_t\big)$ is at hand. In order to compute this representation the author projects the density martingale $\tilde{L}_t = E^{\tilde{Q}}\big(dQ/d\tilde{Q} \mid \mathcal{G}_t\big)$ (which is easily computed via the usual Girsanov theorem) on the filtration $(\mathcal{F}_t)$ using arguments from the innovations approach to nonlinear filtering.

Further work. In [39] a structural portfolio model is considered. In contrast to the papers discussed so-far, in the model of [39] the asset value is observable, whereas the liabilities are subject to random shocks which cannot be observed. Bayesian updating is used in order to compute the conditional distribution of the liabilities given the investor information. The author points out that in case that liability shocks are correlated across firms, the model leads to information-based default contagion.

The authors in [16] study a Duffie-Lando-type model with $\mathcal{F}_t = \mathcal{F}_t^Y \vee \mathcal{F}_t^Z$; in their setup the process Z solves an SDE driven by a Brownian motion β which is correlated with the Brownian motion W driving the asset-value process. In this paper bond prices are computed via the hazard-function approach to reduced-form credit risk models (see [9]). Finally, in [13] the authors study models where only the sign of the firm's cash flow is available. Both papers are mathematically interesting. However, filtering arguments play only a minor role, so that we do not enter into a deeper discussion.

6 Filtering in Reduced Form Models

This section is concerned with the pricing of credit derivatives in reduced-form portfolio credit risk models under incomplete information; the presentation is largely based on our own papers [35] and [37].

6.1 Pricing credit derivatives and nonlinear filtering

As in the previous section we work on a filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t), Q)$ where Q represents the risk neutral pricing measure and $(\mathcal{G}_t)$ the full-information filtration. Recall that the default state of the portfolio under consideration is summarized by the default indicator process $\mathbf{Y} = (Y_{t,1}, \dots, Y_{t,m})_{t\geq 0}$ with $Y_{t,i} = \mathbf{1}_{\{\tau_i \leq t\}}$. We assume that there is some d-dimensional process $\mathbf{X}$ (the state process) and functions $\lambda_i \colon \mathbb{R}^d \to (0, \infty)$ such that $\lambda_i(\mathbf{X}_t)$ is the $(Q, (\mathcal{G}_t))$-default intensity of firm i, or, equivalently, that $Y_{t,i} - \int_0^{\tau_i \wedge t} \lambda_i(\mathbf{X}_s)\, ds$ is a $(Q, (\mathcal{G}_t))$-martingale. Moreover, we assume that the pair of processes $(\mathbf{X}_t, \mathbf{Y}_t)_{t\geq 0}$ is jointly Markov. For concreteness one may think of a model with conditionally independent, doubly stochastic default times where default intensities are functions of some Markovian factor process $\mathbf{X}$ (see also the first example in Subsection 6.3 below). As in the previous sections we denote by $(\mathcal{F}_t)$ the information actually observable to investors. Following [35] we assume that $(\mathcal{F}_t)$ contains the default history $\mathcal{F}_t^{\mathbf{Y}} = \sigma(\mathbf{Y}_s \colon s \leq t)$ and observations of functions of $\mathbf{X}$ in additive Gaussian noise. Formally, $\mathcal{F}_t := \mathcal{F}_t^{\mathbf{Y}} \vee \mathcal{F}_t^{\mathbf{Z}}$ where the l-dimensional process $\mathbf{Z}$ is given by

$$d\mathbf{Z}_t = \boldsymbol{a}_t(\mathbf{X}_t, \mathbf{Y}_t)dt + d\boldsymbol{\beta}_t \ , \tag{52}$$

with $\boldsymbol{\beta}_t = (\beta_{t,1}, \dots, \beta_{t,l})$ an l-dimensional Brownian motion on $(\Omega, \mathcal{G}, (\mathcal{G}_t), Q)$, independent of $\mathbf{X}$ and $\mathbf{Y}$. In order to avoid technical difficulties, $\boldsymbol{a}_t(\cdot)$ is assumed to be bounded. We shall see below that $\mathbf{Z}$ can be interpreted as theoretical prices for traded credit derivatives, observed in additive noise. Note that $\mathbf{X}$ is not $(\mathcal{F}_t)$-adapted, due to the independence of $\mathbf{X}$ and $\boldsymbol{\beta}$.

Pricing credit derivatives. Recall that a credit derivative is a security with $\mathcal{F}_T^{\mathbf{Y}}$-measurable payoff H; specific examples include defaultable zero-coupon bonds or CDSs as introduced in Section 4.1. In accordance with the general pricing principle from Lemma 3.1 we define the *theoretical* or *full-information price* of a credit derivative by $\tilde{H}_t := E\big(e^{-r(T-t)}H \mid \mathcal{G}_t\big)$, $t \leq T$, where the constant $r > 0$ denotes the default-free short rate. By the assumed Markovianity of the pair $(\mathbf{X}, \mathbf{Y})$, for typical credit derivatives the process $\tilde{H}_t$ is of the form $\tilde{H}_t = \boldsymbol{a}_t(\mathbf{X}_t, \mathbf{Y}_t)$ for some functions $\boldsymbol{a}_t : \mathbb{R}^d \times \{0,1\}^m \to \mathbb{R}$ with $t \in [0, T]$. This is obvious, if the payoff is of the form

$H = h(\mathbf{Y}_T)$ as in the case of a defaultable zero-coupon bond; it holds true for most other credit derivatives such as the default-payment introduced in Section 4.1 as-well. For non-traded credit derivatives we define the *investor price* by $H_t := E\big(e^{-r(T-t)}H \mid \mathcal{F}_t\big)$. We get from Lemma 3.1 (or by a direct application of iterated conditional expectations) that

$$H_t = E(\boldsymbol{a}_t(\mathbf{X}_t, \mathbf{Y}_t) \mid \mathcal{F}_t)\,. \tag{53}$$

Since $\mathbf{Y}_t$ is observable, in order to compute H_t we thus need to determine the conditional distribution $\pi_{\mathbf{X}_t|\mathcal{F}_t}$, which amounts to solving a nonlinear filtering problem.

Now we come back to the economic interpretation of $\mathbf{Z}$. Assume that investors have noisy information about the theoretical price $\boldsymbol{a}_t(\mathbf{X}_t, \mathbf{Y}_t)$ of l traded credit derivatives. In a discrete-time framework it is natural to assume that the observed market quotes are of the form $\boldsymbol{z}_{t_k} = \boldsymbol{a}_{t_k}(\mathbf{X}_{t_k}, \mathbf{Y}_{t_k}) + \boldsymbol{\epsilon}_k$ for time points $t_k = k\Delta$ and an iid sequence $(\boldsymbol{\epsilon}_k)_k$ of independent noise-variables $\epsilon_1, \ldots, \epsilon_l$ with mean zero and finite variance. The noise variables model transmission and observation errors as well as temporary deviations of market quotes from theoretical prices. In continuous time one considers instead the *cumulative observation process* $\mathbf{Z}_t^\Delta := \Delta \sum_{t_k \le t} \boldsymbol{z}_{t_k}$. Then we have for Δ small, using Donsker's invariance principle,

$$\mathbf{Z}_t^\Delta = \sum_{t_k \le t} \boldsymbol{a}_{t_k}(\mathbf{X}_{t_k}, \mathbf{Y}_{t_k})\Delta + \Delta \sum_{t_k \le t} \boldsymbol{\epsilon}_k \approx \int_0^t \boldsymbol{a}_s(\mathbf{X}_s, \mathbf{Y}_s)ds + \boldsymbol{\beta}_t. \tag{54}$$

Remark 6.1 (Default intensities and default contagion)**.** It is well-known that default-intensities with respect to sub-filtrations can be computed by projection (see for instance Chapter II of [10]). Hence the risk-neutral $(\mathcal{F}_t)$-default intensity of firm j is given by the left-continuous version of

$$\hat{\lambda}_{t,j} := E\left(\lambda_j(\mathbf{X}_t) \mid \mathcal{F}_t\right) = \int_{\mathbb{R}^d} \lambda_j(\boldsymbol{x})\, \pi_{\mathbf{X}_t|\mathcal{F}_t}(d\boldsymbol{x})\,, \quad t \le \tau_j, \tag{55}$$

so that in order to compute this quantity we again need the conditional distribution $\pi_{\mathbf{X}_t|\mathcal{F}_t}(d\boldsymbol{x})$. Relation (55) illustrates nicely the notion of information-based default contagion that was already mentioned in Section 4.2: new default information such as the news that obligor $i \neq j$ has defaulted leads to an update in the conditional distribution $\pi_{\mathbf{X}_t|\mathcal{F}_t}(d\boldsymbol{x})$ and hence to a jump in the $(\mathcal{F}_t)$-default intensity of firm j. Note that this leads to a downward jump in the model value (and hence an increase in the credit spread) of a zero-coupon bond issued by some non-defaulted firm.

6.2 A general jump-diffusion model

Following [35] we next introduce a jump-diffusion model for the joint dynamics of $\mathbf{X}$ and $\mathbf{Y}$. This model is fairly general and includes most reduced-form models from the literature as special cases; specific examples are discussed in the next subsection. We assume that the factor process $\mathbf{X} = (X_{t,1}, \ldots, X_{t,d})_{t \ge 0}$ and the default indicator process $\mathbf{Y}$ solve the following SDE on $(\Omega, \mathcal{G}, (\mathcal{G}_t), Q)$

$$\mathbf{X}_t = \mathbf{X}_0 + \int_0^t \boldsymbol{b}(\mathbf{X}_{s-})ds + \int_0^t \sigma(\mathbf{X}_{s-})d\mathbf{W}_s + \int_0^t \int_E K^{\mathbf{X}}(\mathbf{X}_{s-}, u)\mathcal{N}(ds, du), \tag{56}$$

$$Y_{t,j} = Y_{0,j} + \int_0^t \int_E (1 - Y_{s-,j})K_j^{\mathbf{Y}}(\mathbf{X}_{s-}, u)\mathcal{N}(ds, du),\ 1 \le j \le m. \tag{57}$$

Here $\mathbf{W}$ is a standard k-dimensional Brownian motion; drift $\boldsymbol{b} = (b_1, \ldots, b_d)$ and dispersion matrix $\sigma = (\sigma_{i,l})$, $1 \le i \le d, 1 \le l \le k$ are functions from $S^{\mathbf{X}}$ to $\mathbb{R}^d$ and $\mathbb{R}^{d \times k}$ respectively, $S^{\mathbf{X}} \subset \mathbb{R}^d$ is the state space of $\mathbf{X}$; $\mathcal{N}(ds, du)$ denotes a $(Q, (\mathcal{G}_t))$-standard Poisson random measure on $\mathbb{R}_+ \times E$, E some Euclidean space, with compensator measure $F_{\mathcal{N}}(du)ds$; $\mathbf{W}$ and $\mathcal{N}$ are independent; $\mathbf{X}_0$ is a random vector taking values in $S^{\mathbf{X}} \subset \mathbb{R}^d$; $\mathbf{Y}_0$ is a given element of $\{0,1\}^m$. Moreover, $K_j^{\mathbf{Y}}(\boldsymbol{x}, u) \in \{0,1\}$ for all $\boldsymbol{x}, u$ and all $1 \le j \le m$, so that the solution of (57) is in fact of the form $Y_{t,j} = \mathbf{1}_{\{\tau_j \le t\}}$. Define the sets

$$D_i^{\mathbf{X}}(\boldsymbol{x}) := \{u \in E : K_i^{\mathbf{X}}(\boldsymbol{x}, u) \neq 0\}, \quad 1 \le i \le d, \tag{58}$$

$$D_j^{\mathbf{Y}}(\boldsymbol{x}) := \{u \in E : K_j^{\mathbf{Y}}(\boldsymbol{x}, u) \neq 0\}, \quad 1 \le j \le m. \tag{59}$$

By definition of $D_i^{\mathbf{X}}$, the process $\Delta X_{t,i} \neq 0$ if and only of $\mathcal{N}(\{t\} \times D_i^{\mathbf{X}}) > 0$; similarly, for a non-defaulted firm j we have $\tau_j = t$ if and only if $\mathcal{N}(\{t\} \times D_j^{\mathbf{Y}}) > 0$.

In addition to several regularity conditions ensuring the existence and uniqueness of a solution, in [35] it is assumed that for all $1 \le j_1 < j_2 \le m$ and all $\boldsymbol{x} \in S^{\mathbf{X}}$ one has $F_{\mathcal{N}}(D_{j_1}^{\mathbf{Y}}(\boldsymbol{x}) \cap D_{j_2}^{\mathbf{Y}}(\boldsymbol{x})) = 0$. This assumption ensures that for $j_1 \neq j_2$ the processes Y_{j_1} and Y_{j_2} have no common jumps so that there are no joint defaults. Note however, that the model (56), (57) allows for common jumps of $\mathbf{X}$ and $\mathbf{Y}$. More precisely, there is a strictly positive probability that the factor process $\mathbf{X}$ jumps at τ_j, if

$$F_{\mathcal{N}}\big(D_j^{\mathbf{Y}}(\mathbf{X}_{\tau_j -}) \cap D_i^{\mathbf{X}}(\mathbf{X}_{\tau_j -})\big) > 0 \text{ for some } 1 \le i \le d. \tag{60}$$

Note moreover, that by definition of the compensator of a Poisson random measure,

$$Y_{t,j} - \int_0^t (1 - Y_{s-,j}) F_{\mathcal{N}}(D_j^{\mathbf{Y}}(\mathbf{X}_{s-}))\, ds, \quad t \ge 0,$$

is a $(\mathcal{G}_t)$-martingale, so that $\lambda_j(\mathbf{X}_{t-}) := F_{\mathcal{N}}(D_j^{\mathbf{Y}}(\mathbf{X}_{t-}))$ is the $(\mathcal{G}_t)$-default intensity of firm j.

6.3 Examples

Next we present a number of specific examples which show that a great variety of models are covered by the system (56), (57).

Conditionally independent defaults. Consider a model with conditionally independent, doubly-stochastic default times and assume that $\mathbf{X}$ follows a jump diffusion model of the form

$$d\mathbf{X}_t = \boldsymbol{b}(\mathbf{X}_t)dt + \sigma(\mathbf{X}_t)d\mathbf{W}_t + d\mathbf{J}_t, \tag{61}$$

where $\mathbf{J}$ is an $\mathbb{R}^d$-valued compound Poisson process with compensator measure $F_{\mathbf{J}}(d\boldsymbol{x})ds$. A popular model of this form is the affine jump-diffusion model of [23]. Such a model can be included in the framework (56), (57) as follows. Take $E = \mathbb{R}^d \times \mathbb{R}$, $F_{\mathcal{N}} = F_{\mathbf{J}} \times \nu$, ν Lebesgue-measure on $\mathbb{R}$, and put

$$K_j^{\mathbf{Y}}(\boldsymbol{x}, \boldsymbol{u}) = \mathbf{1}_{\{\left[\sum_{i=1}^{j-1} \lambda_i(\boldsymbol{x}), \sum_{i=1}^{j} \lambda_i(\boldsymbol{x})\right]\}}(u_{d+1}), \quad 1 \le j \le m, \text{ and} \tag{62}$$

$$K_i^{\mathbf{X}}(\boldsymbol{x}, \boldsymbol{u}) = u_i \mathbf{1}_{\{[-1,0)\}}(u_{d+1}), \quad 1 \le i \le d. \tag{63}$$

Note that $K^{\mathbf{X}}$ and $K^{\mathbf{Y}}$ have been chosen so that $F_{\mathcal{N}}(D_i^{\mathbf{X}}(\boldsymbol{x}) \cap D_j^{\mathbf{Y}}(\boldsymbol{x})) = 0$ for all $1 \le i \le d$, all $1 \le j \le m$, and all $\boldsymbol{x}$ in $S^{\mathbf{X}}$.

A Markov-chain model with infectious defaults. Next we consider models where the state process jumps in reaction to default events. A simple example is provided by the following generalization of the *infectious-defaults model* of [21]. Here X is taken as scalar and modeled as a finite-state Markov chain with state space $S^X = \{1, \dots, K\} \subset \mathbb{R}$; the default intensity of firm j is given by $\lambda_j(X_t)$ for increasing functions $\lambda_j : S^X \to \mathbb{R}^+$. At a default time τ_n, X jumps upward by one unit with probability p_{ξ_n} (which may depend on the identity ξ_n of the nth defaulting firm), and remains constant with probability $1 - p_{\xi_n}$ (unless, of course, if $X_{\tau_n -} = K$, where X remains constant). If the system is in an "ignited state", that is if $X_t \geq 2$, X_t jumps to $X_t - 1$ with intensity $\gamma(X_t)$; these downward jumps occur independently of the default history. An upward jump of X at a default can be viewed as manifestation of counterparty risk and/or default contagion, as the default intensities of the remaining firms are increased. This leads to a downward jump in the model value (and hence an increase in the credit spread) of a zero-coupon bond issued by some non-defaulted firm. This model can be embedded in the framework (56), (57) by a proper choice of $F_{\mathcal{N}}$, $K^{\mathbf{X}}$ and $K^{\mathbf{Y}}$; see [35] for details.

The information-based model of Frey, Schmidt and Gabih [37]. This model is of interest in our context, since the dynamics of model values and of the state variable process are themselves derived by filtering arguments. In [37] the framework of this section is slightly extended and three different layers of information are considered: full information, so-called market information $(\mathcal{G}_t)$, and finally information of secondary-market investors $(\mathcal{F}_t)$. Here the full-information setup is an additional layer of information used for the construction of the model; we denote the corresponding filtration by $(\tilde{\mathcal{G}}_t)$. It is assumed that under full information the default times are conditionally independent doubly stochastic random times, and that $(\tilde{\mathcal{G}}_t)$-default intensities are driven by a finite-state Markov chain Ψ with state space $\{1, \dots, K\}$ and generator matrix Q^Ψ. The filtration $\mathcal{G}_t \subset \tilde{\mathcal{G}}_t$ represents the information used by the market in determining the theoretical equilibrium prices of traded credit derivatives. In accordance with the definition of $\tilde{H}_t$ in Section 6.1, the theoretical price of the traded credit derivatives is defined by $\tilde{H}_{t,j} = E\big(\exp(-r(T-t))H_j \mid \mathcal{G}_t\big)$, $1 \leq j \leq l$. Define the vector $\boldsymbol{p}_t$ of conditional-probabilities by

$$\boldsymbol{p}_t := (p_t^1, \dots, p_t^K) \text{ where } p_t^k := Q(\Psi_t = k \mid \mathcal{G}_t),\ 1 \leq k \leq K\,. \tag{64}$$

The process $\boldsymbol{p} = (\boldsymbol{p}_t)_{t \geq 0}$ is a natural state variable process for the model in the market filtration; $\boldsymbol{p}$ thus plays the role of the process $\mathbf{X}$ introduced in (56). The reasons are the following: first, denoting the $(\tilde{\mathcal{G}}_t)$-default intensities by $\nu_i(\Psi_t)$, the $(\mathcal{G}_t)$-default intensities are given by $\lambda_i(\boldsymbol{p}_t) := \sum_{k=1}^K p_t^k \nu_i(k)$. Moreover, note that by the $(\tilde{\mathcal{G}}_t)$-Markovianity of Ψ and $\mathbf{Y}$, the conditional expectation $E\big(\exp(-r(T-t))H_j \mid \tilde{\mathcal{G}}_t\big)$ is given by some function $\tilde{a}_{t,j}(\Psi_t, \mathbf{Y}_t)$, at least for $H_j = h_j(\mathbf{Y}_T)$. By iterated conditional expectations theoretical prices can therefore be expressed as functions of t, $\boldsymbol{p}_t$ and $\mathbf{Y}_t$ as well:

$$\tilde{H}_{t,j} = E(\tilde{a}_{t,j}(\Psi_t, \mathbf{Y}_t) \mid \mathcal{G}_t) = \sum_{k=1}^K p_t^k\, \tilde{a}_{t,j}(k, \mathbf{Y}_t) =: a_{t,j}(\boldsymbol{p}_t, \mathbf{Y}_t)\,,\ 1 \leq j \leq l. \tag{65}$$

In particular, theoretical prices have a linear factor structure with factor process $\boldsymbol{p}$.

In [37] it is assumed that $\mathcal{G}_t = \mathcal{F}_t^Y \vee \mathcal{F}_t^U$, where the process U models in abstract form the information aggregated in the equilibrium prices of traded securities. Mathematically, U is given by $U_t = \int_0^t \mu(\Psi_s)ds + B_t$, B a standard $(\tilde{\mathcal{G}}_t)$-Brownian motion independent of Ψ and Y. In a spirit similar to that of the Landen-model discussed in Section 3.3, the innovations approach to nonlinear filtering is used in order to determine the dynamics of the process $\boldsymbol{p}$, and with it the dynamics of theoretical prices. We briefly sketch the main steps. Consider the $(\mathcal{G}_t)$-adapted processes

$$M_{t,i} := Y_{t,i} - \int_0^{t\wedge\tau_i} \lambda_i(\boldsymbol{p}_s)\, ds\,,\ 1 \le i \le m, \text{ and } W_t := U_t - \int_0^t \mu(\boldsymbol{p}_s)\, ds\,,$$

where $\mu(\boldsymbol{p}_t) = \sum_{k=1}^K p_t^k \mu(k)$. The processes $\mathbf{M} = (M_{t,1}, \ldots, M_{t,d})_{t\ge 0}$ and W are the *innovations processes*. In particular, it is well-known that $\mathbf{M}$ is a $(\mathcal{G}_t)$-martingale and that W is $(\mathcal{G}_t)$-Brownian motion; moreover, as shown in [37], every $(\mathcal{G}_t)$ martingale can be represented as stochastic integral wrt $\mathbf{M}$ and W.

For a generic process Γ denote by $\widehat{\Gamma}$ the optional projection of Γ with respect to the market filtration $(\mathcal{G}_t)$. Consider now a generic $(\tilde{\mathcal{G}}_t)$-semimartingale with canonical decomposition of the form

$$J_t = J_0 + \int_0^t A_s ds + M_t^J\,;$$

here A is some adapted right-continuous process, and M^J is an $(\tilde{G}_t)$-martingale with $[M^J, B] = 0$ and $[M^J, Y_i] = 0$ for all i. It is then shown in [37], that the optional projection $\widehat{J}_t$ has the representation

$$\widehat{J}_t = \widehat{J}_0 + \int_0^t \widehat{A}_s ds + \int_0^t \boldsymbol{\gamma}_s^\top d\mathbf{M}_s + \int_0^t \alpha_s dW_s, \tag{66}$$

where, with $\mu_t := \mu(\Psi_t)$ and $\nu_{t,j} := \nu_j(\Psi_t)$, the integrands $\boldsymbol{\gamma}$ and α are given by

$$\alpha_t = (\widehat{J\mu})_t - \widehat{J}_t\widehat{\mu}_t, \text{ and } \gamma_{t,j} = \frac{1}{(\widehat{\nu}_j)_{t-}}\Big((\widehat{J\nu_j})_{t-} - (\widehat{J})_{t-}(\widehat{\nu}_j)_{t-}\Big) \quad j = 1, \ldots, m. \tag{67}$$

The proof of this result is based on standard arguments from the innovations approach to nonlinear filtering. Consider now the semimartingales $J_{t,k} = \mathbf{1}_{\{\Psi_t = k\}}$, $1 \le k \le K$ and note that $\boldsymbol{p}_t = \widehat{\mathbf{J}}_t$. Since the semimartingale decomposition of J_k is given by $J_{t,k} = J_{0,k} + \int_0^t Q_{\Psi_s,k}^{\Psi}\, ds + M_t^{J_k}$, by applying (66) to J_k one therefore obtains the following K-dimensional SDE for the process $\boldsymbol{p}$:

$$dp_t^k = \sum_{i=1}^K Q_{i,k}^{\Psi}\, p_t^i dt + \sum_{i=1}^m \boldsymbol{\gamma}_i^k(\boldsymbol{p}_{t-})\, dM_{t,i} + \delta^k(\boldsymbol{p}_{t-})\, dW_t,\ 1 \le k \le K, \tag{68}$$

with coefficients given by the functions

$$\gamma_j^k(\boldsymbol{p}) = p^k\Big(\frac{\nu_j(k)}{\sum_{n=1}^K \nu_j(n) p^n} - 1\Big)\,,\ \delta^k(\boldsymbol{p}) = p^k\Big(\mu(k) - \sum_{n=1}^K p^n \mu(n)\Big)\,; \tag{69}$$

see again[37] for details.

Similarly as in (52), in [37] the information set of secondary market investors is given by $\mathcal{F}_t = \mathcal{F}_t^{\mathbf{Y}} \vee \mathcal{F}_t^{\mathbf{Z}}$ for $\mathbf{Z}_t = \int_0^t \boldsymbol{a}_s(\boldsymbol{p}_s, \mathbf{Y}_s) ds + \boldsymbol{\beta}_t$, $\boldsymbol{\beta}$ a Brownian motion independent of $\mathbf{Y}$ and U. By an argument analogous to that leading to (53), the computation of prices for secondary-market investors leads to the problem of finding $\pi_{\boldsymbol{p}_t|\mathcal{F}_t}$ so that the solution $\boldsymbol{p}$ of the filtering problem with respect to the market information $(\mathcal{G}_t)$ becomes the state variable of the filtering problem with respect to the investor filtration $(\mathcal{F}_t)$. The latter filtering problem is covered by our setup, as the processes $\boldsymbol{p}$ and $\mathbf{Y}$ follow an SDE-system of the form (56) and (57). Note in particular that at a default time the probability vector $\boldsymbol{p}$ is updated according to (68), so that there are common jumps in the state variable $\boldsymbol{p}$ and the observation $\mathbf{Y}$. In [37] the authors discuss also the hedging of credit derivatives from the viewpoint of secondary-market investors. For this they rely on the concept of *risk-minimization with restricted information* as introduced in [58]. It is shown that the solution of the hedging problem again leads to the problem of finding $\pi_{\boldsymbol{p}_t|\mathcal{F}_t}$.

While complicated at first sight, the model of [37] has a number of attractive features. To begin with, by (65), the main numerical task is the evaluation of the functions $\tilde{a}_j(t, k, \boldsymbol{y})$; as these functions are computed in the simple full-information setup, computations become relatively easy even for non-homogeneous models. Moreover, the model generates a rich set of price dynamics with randomly fluctuating credit spreads and default contagion, and it has a natural factor structure.

6.4 Filter equations

In the setup of Subsection 6.2, the determination of $\pi_{\mathbf{X}_t|\mathcal{F}_t}(d\boldsymbol{x})$ becomes an interesting nonlinear filtering problem with observations of mixed type (generated by marked point processes and diffusions) and with common jumps in the observation $\mathbf{Y}$ and the state process $\mathbf{X}$. This is problem is non-standard and merits a discussion in the context of the present survey.

Filtering problems with common jumps of the unobserved state process and of the observations have previously been discussed in the literature. Initial results can be found in [42]; the papers [46] and [12] are concerned with scalar observations described by a pure jump process. The recent paper [19] on the other hand treats the filtering problem for a very general marked point process model but without common jumps of the state- and the observation process. All these papers are based on the innovations approach to nonlinear filtering. Assuming that investor information is equal to the default history $(\mathcal{F}_t^{\mathbf{Y}})$, [32] gives a simple filter algorithm for the affine jump-diffusion model of [23] with conditionally independent doubly stochastic default times.

In line with [35], in the present paper we follow an alternative route which is based on ideas from the reference probability approach and takes into account the particular structure of the given general model. In this way we obtain new general recursive filter equations and, as a byproduct, an explicit expression for the joint likelihood of state process and observations.[9] In the case that $(\mathbf{X}, \mathbf{Y})$ is a finite-state continuous-time Markov chain our filter equations give rise to a finite-dimensional filter. We describe now in more detail some of the results from [35] and present in particular the filter equations.

Preliminaries. Since the approach in [35] is based on ideas from the reference probability approach, we first mention some preliminaries to this effect. For ease of notation and without loss of generality in what follows we shall simply denote by $\boldsymbol{a}_t(\mathbf{X}_t)$ the process with the components $a_{t,1}(\mathbf{X}_t, \mathbf{Y}_t), \cdots, a_{t,l}(\mathbf{X}_t, \mathbf{Y}_t)$ from the drift of $\mathbf{Z}$ as introduced in (52). It will be convenient to define the processes $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ on a product space $(\Omega, \mathcal{G}, (\mathcal{G}_t), R^0)$ so that $\mathbf{Z}$ is independent of $\mathbf{X}$ and $\mathbf{Y}$. Denote by $(\Omega_2, \mathcal{G}_2, (\mathcal{G}_t^2), P^{0,l})$ the l-dimensional Wiener space with coordinate process $\mathbf{Z}$, i.e. $\mathbf{Z}_t(\omega_2) = \omega_2(t)$. Given some probability space $(\Omega_1, \mathcal{G}_1, (\mathcal{G}_t^1), P)$ supporting a solution $(\mathbf{X}, \mathbf{Y})$

[9] With common jumps of $\mathbf{X}$ and $\mathbf{Y}$ such an expression cannot be derived from standard filtering results in a straightforward way, essentially because $\mathbf{X}$ and $\mathbf{Y}$ cannot be made independent by a change of measure.

of the SDE-system (56), (57), let $\Omega := \Omega_1 \times \Omega_2$, $\mathcal{G} = \mathcal{G}_1 \otimes \mathcal{G}_2$, $\mathcal{G}_t = \mathcal{G}_t^1 \otimes \mathcal{G}_t^2$, $R^0 := P \otimes P^{0,l}$, and put for $\omega = (\omega_1, \omega_2) \in \Omega$

$$\mathbf{X}_t(\omega) := \mathbf{X}_t(\omega_1),\ \mathbf{Y}_t(\omega) := \mathbf{Y}_t(\omega_1),\ \text{and } \mathbf{Z}_t(\omega) := \mathbf{Z}_t(\omega_2).$$

Note that this implies that under R^0, $\mathbf{Z}$ is l-dimensional Brownian motion, independent of $\mathbf{X}$ and $\mathbf{Y}$. Introduce then a Girsanov-type measure transformation of the form $\frac{dR}{dR^0}_{|\mathcal{F}_t} = L_t$ with

$$L_t = L_t(\omega_1, \omega_2) = \exp\Big\{ \int_0^t \big(\boldsymbol{a}_s(\mathbf{X}_s(\omega_1))\big)' \, d\mathbf{Z}_s(\omega_2) - \frac{1}{2} \int_0^t \|\boldsymbol{a}_s(\mathbf{X}_s(\omega_1))\|^2 \, ds \Big\} \tag{70}$$

and note that L is indeed a R^0-martingale since $\boldsymbol{a}_t(\cdot)$ was assumed to be bounded. Under R, the process $\mathbf{Z}$ has the original dynamics (52), while the law of $\mathbf{X}$ and $\mathbf{Y}$ remains unchanged.

In the sequel we discuss how to obtain the filter distribution $\pi_{\mathbf{X}_t|\mathcal{F}_t}$ in *weak form*, that is we want to compute for a generic bounded and continuous function $h\colon \mathbb{R}^d \to \mathbb{R}$ the conditional expectation

$$\pi_t h := E(h(\mathbf{X}_t) \mid \mathcal{F}_t) = \int_{\mathbb{R}^d} h(\boldsymbol{x}) \pi_{\mathbf{X}_t|\mathcal{F}_t}(d\boldsymbol{x}) \,. \tag{71}$$

The well-known Kallianpur-Striebel formula gives

$$\pi_t h = \frac{E^{R^0}(h(\mathbf{X}_t) L_t \mid \mathcal{F}_t)}{E^{R^0}(L_t \mid \mathcal{F}_t)} \,, \tag{72}$$

so that, to compute $\pi_t h$, it suffices to compute the numerator on the right-hand side in (72). Recalling that $\mathcal{F}_t = \mathcal{F}_t^{\mathbf{Y}} \vee \mathcal{F}_t^{\mathbf{Z}}$, we reduce next the conditioning on $\mathcal{F}_t$ to a conditioning on $\mathcal{F}_t^{\mathbf{Y}}$. Using the Fubini theorem and the product-structure of $(\Omega, \mathcal{F}, (\mathcal{F}_t), R^0)$ we get with $L_t = L_t(\omega_1, \omega_2)$ as introduced in (70)

$$E^{R^0}(h(\mathbf{X}_t) L_t \mid \mathcal{F}_t^{\mathbf{Y}} \vee \mathcal{F}_t^{\mathbf{Z}})(\omega) = E^P(h(\mathbf{X}_t) L_t(\cdot, \omega_2) \mid \mathcal{F}_t^{\mathbf{Y}})(\omega_1) \,. \tag{73}$$

In order to compute $\pi_t h$ we thus have to evaluate the conditional expectation on the right hand side of (73). Note that this involves only the first component $(\Omega_1, \mathcal{G}_1, (\mathcal{G}_t^1), P)$ of the underlying probability space and hence only the joint law of $\mathbf{X}$ and $\mathbf{Y}$; expectations with respect to that law will be simply denoted by E (instead of E^P).

To derive the filter between default times we have to modify the kernel $K^{\mathbf{X}}(\cdot)$ in the dynamics (56) for $\mathbf{X}$ because, conditional on not having jumps in $\mathbf{Y}$, certain jumps in $\mathbf{X}$ cannot occur. More precisely, we shall consider the following kernel, which here is slightly generalized wrt the corresponding definition in Section 6.2 by making it explicitly dependent also on $\boldsymbol{y}$, namely

$$\bar{K}^{\mathbf{X}}(\boldsymbol{x}, \boldsymbol{y}, u) := \begin{cases} 0, & \text{if } \ u \in \bar{D}^{\mathbf{Y}}(\boldsymbol{x}, \boldsymbol{y}) := \bigcup_{\{j\,:\,\boldsymbol{y}_j = 0\}} D_j^{\mathbf{Y}}(\boldsymbol{x}, \boldsymbol{y}) \,, \\ K^{\mathbf{X}}(\boldsymbol{x}, u) & \text{else.} \end{cases} \tag{74}$$

We shall denote by $\bar{\mathbf{X}}_t$ the process corresponding to $\bar{K}^{\mathbf{X}}(\cdot)$ and by $\bar{\mathbf{Y}}_t$ the process in (57) obtained when replacing $\mathbf{X}_t$ by $\bar{\mathbf{X}}_t$ there. The law of $(\bar{\mathbf{X}}_t, \bar{\mathbf{Y}}_t)$ with initial condition $(\bar{\mathbf{X}}_0 = \boldsymbol{x}, \bar{\mathbf{Y}}_0 = \boldsymbol{y})$ will be denoted by $\bar{P}_{(\boldsymbol{x},\boldsymbol{y})}$ and the corresponding expectation by $\bar{E}_{(\boldsymbol{x},\boldsymbol{y})}$.

Given these preliminaries the filtering results now take the form of a recursion over the successive default times T_n, $1 \le n \le m$. Since in the generic interval $[T_{n-1}, T_n)$ only the new price information $(\mathbf{Z}_s)_{s \ge T_{n-1}}$ matters, in the sequel we use the lighter notation

$$\mathbf{Z}_s^n := \mathbf{Z}_{s+T_{n-1}} \quad \text{and} \quad \boldsymbol{a}_s^n(\cdot) := \boldsymbol{a}_{s+T_{n-1}}(\cdot) \,. \tag{75}$$

Filtering between defaults. The main result here is the following (see [35])

Theorem 6.2. *Given two successive default times T_{n-1}, T_n, we have for $t \in [T_{n-1}, T_n)$*

$$\pi_t h \propto \int_{\mathbb{R}^d} \pi_{T_{n-1}}(d\boldsymbol{x})\, \bar{E}_{(\boldsymbol{x},\mathbf{Y}_{T_{n-1}})}\Big(h\left(\bar{\mathbf{X}}_{t-T_{n-1}}\right) L^n_{t-T_{n-1}} \exp\Big\{ -\int_0^{t-T_{n-1}} \bar{\lambda}(\bar{X}_s, \mathbf{Y}_{T_{n-1}}) ds \Big\}\Big) \tag{76}$$

where $\bar{E}_{(\boldsymbol{x},\boldsymbol{y})}$ is the expectation as introduced above, $\pi_{T_{n-1}}(d\boldsymbol{x})$ is the filter distribution at $t = T_{n-1}$ and the process $L^n = (L^n_t)_{t \geq 0}$ is defined by

$$L^n_t = \exp\Big\{ \int_0^t (\boldsymbol{a}^n_s)'(\bar{X}_s)\, d\mathbf{Z}^n_s - \frac{1}{2}\int_0^t \left\| \boldsymbol{a}^n_s(\bar{X}_s) \right\|^2 ds \Big\}.$$

Filtering at a default time. By (72) and (73) at a generic default time one has

$$\pi_{T_n} h \propto E(h(\mathbf{X}_{T_n}) L_{T_n} \mid \mathcal{F}^{\mathbf{Y}}_{T_n}).$$

Notice now that, due to the possibility of common jumps between $\mathbf{X}$ and $\mathbf{Y}$, the expressions $E(h(\mathbf{X}_{T_n})L_{T_n} \mid \mathcal{F}^{\mathbf{Y}}_{T_n})$ and $E(h(\mathbf{X}_{T_n^-})L_{T_n} \mid \mathcal{F}^{\mathbf{Y}}_{T_n})$ do not necessarily coincide. We shall therefore proceed along two steps. In Step 1 we show that one can obtain the conditional expectation $E(h(\mathbf{X}_{T_n})L_{T_n} \mid \mathcal{F}^{\mathbf{Y}}_{T_n})$ once one is able to compute $E(g(\mathbf{X}_{T_n^-})L_{T_n} \mid \mathcal{F}^{\mathbf{Y}}_{T_n})$ for a generic function $g(\cdot)$. In this step we use the joint distribution of the jumps $\Delta\mathbf{X}_{T_n}$ and $\Delta\mathbf{Y}_{T_n}$ and hence the particular structure of the given model. In Step 2 we then compute the latter of those two quantities via Bayesian updating.

Step 1 (Reduction to the filter distribution of $\mathbf{X}_{T_n^-}$). Here one can show (see [35])

Proposition 6.3. *We have the relation*

$$E(h(\mathbf{X}_{T_n})L_{T_n} \mid \mathcal{F}^{\mathbf{Y}}_{T_n}) = E(g(\mathbf{X}_{T_n^-}, \xi_n)L_{T_n} \mid \mathcal{F}^{\mathbf{Y}}_{T_n}),$$

where ξ_n is the identity of the firm defaulting at T_n, and where the function g is given by

$$g(\boldsymbol{x}, j) = \begin{cases} F_{\mathcal{N}}\big(D^{\mathbf{Y}}_j(\boldsymbol{x})\big)^{-1} \displaystyle\int_{D^{\mathbf{Y}}_j(\boldsymbol{x})} h\big(\boldsymbol{x} + K^{\mathbf{X}}(\boldsymbol{x}, u)\big) F_{\mathcal{N}}(du), & \text{if } F_{\mathcal{N}}\big(D^{\mathbf{Y}}_j(\boldsymbol{x})\big) > 0, \\ h(\boldsymbol{x}), & \text{else.} \end{cases} \tag{77}$$

Step 2 (Updating of the conditional distribution of $\mathbf{X}_{T_n^-}$). Here we have

Theorem 6.4. *Given the information that a default has actually occurred at $t = T_n$ and given the identity ξ_n of the defaulting firm, for a generic function $g : \mathbb{R}^d \to \mathbb{R}$ we have*

$$E\big(g(\mathbf{X}_{T_n^-})L_{T_n} \mid \mathcal{F}^{\mathbf{Y}}_{T_n}\big) \propto \int_{\mathbb{R}^d} \pi_{T_{n-1}}(d\boldsymbol{x})\, \bar{E}_{(\boldsymbol{x},\mathbf{Y}_{T_{n-1}})}\Big(g(\bar{\mathbf{X}}_{T_n - T_{n-1}})$$
$$\times L^n_{T_n - T_{n-1}} \cdot \lambda_{\xi_n}\left(\bar{\mathbf{X}}_{T_n-T_{n-1}}, \mathbf{Y}_{T_{n-1}}\right) \exp\Big\{ -\int_0^{T_n - T_{n-1}} \bar{\lambda}(\bar{\mathbf{X}}_s, \mathbf{Y}_{T_{n-1}})\, ds \Big\}.$$

Filter equations for finite-state Markov chains We show here how the general filter equations specialize for the case when $(\mathbf{X}, \mathbf{Y})$ form a finite state Markov chain thereby also showing how these equations allow the computations to be performed explicitly.

We assume w.l.o.g. that the state space of $(\mathbf{X}, \mathbf{Y})$ is $\{1, \ldots, K\} \times \{0,1\}^m$ so that X can be considered as scalar. Denote the transition intensities of $(X, \mathbf{Y})$ by $q(k, \boldsymbol{y}; \tilde{k}, \tilde{\boldsymbol{y}})$. In line with our general framework we restrict the transition intensities so that default is an absorbing state and so that there are no simultaneous defaults. Hence, denoting the current state by $(k, \boldsymbol{y})$, there are three possible transitions of $(X, \mathbf{Y})$. First there may be a transition from $(k, \boldsymbol{y})$ to $(h, \boldsymbol{y})$, $h \neq k$; this transition occurs with intensity $\bar{q}^{\boldsymbol{y}}_{k,h} := q(k, \boldsymbol{y}; h, \boldsymbol{y})$. Second, there may be a 'contagious default', i.e. for $i \in \{1, \ldots, m\}$ with $y_i = 0$ and $h \neq k$ there may be a transition from $(k, \boldsymbol{y})$ to $(h, \boldsymbol{y}^i)$, where $\boldsymbol{y}^i$ is obtained from $\boldsymbol{y}$ by flipping the ith coordinate. Third we may have a 'pure default', i.e. a transition from $(k, \boldsymbol{y})$ to $(k, \boldsymbol{y}^i)$. In particular, the default intensity of a non-defaulted firm i is equal to $\lambda_i(k, \boldsymbol{y}) = \sum_{h=1}^K q(k, \boldsymbol{y}; h, \boldsymbol{y}^i)$. This Markov chain model can be included in the general framework of (56), (57) by a specific choice of K^X and $K^{\mathbf{Y}}$; see [35] for details.

In this finite-state Markov case the filter distribution can be summarized by the K-dimensional process $\boldsymbol{\pi}_t = (\pi_t^1, \ldots, \pi_t^K)$ with $\pi_t^i := P(X_t = i \mid \mathcal{F}_t)$. Obviously, it suffices to compute an unnormalized version of $\boldsymbol{\pi}_t$. Put for $h \in \{1, \ldots, K\}$

$$\sigma_t^h[n, \boldsymbol{y}] := \sum_{i=1}^K \bar{E}_{(i,\boldsymbol{y})} \left(1_{\{\bar{X}_t = h\}} L_t^n \exp\left\{ - \int_0^t \bar{\lambda}(\bar{X}_s, \boldsymbol{y})\, ds \right\} \right) \pi_{T_{n-1}}(\{i\}), \tag{78}$$

so that $\sigma_t[n, \boldsymbol{y}]g = \sum_{h=1}^K \sigma_t^h[n, \boldsymbol{y}]g(h)$. We have the following Zakai-equation for $\boldsymbol{\sigma}_t$ (see [35])

Proposition 6.5. *Between default times the process $\boldsymbol{\sigma}_t = \boldsymbol{\sigma}_t[n, \boldsymbol{y}]$ solves the SDE*

$$d\sigma_t^i = \Big(\sum_{k=1}^K \bar{q}^{\boldsymbol{y}}_{k,i} \sigma_t^k - \bar{\lambda}(i, \boldsymbol{y}) \sigma_t^i \Big)\, dt + \sigma_t^i\, (\boldsymbol{a}_t^n)'(i)\, d\mathbf{Z}_t^n, \quad 1 \leq i \leq K, \tag{79}$$

with initial condition $\sigma_0^i = \pi_{T_{n-1}}(\{i\})$.

At a default time T_n the filter distribution is updated as follows. Compute first

$$P\left(X_{T_{n^-}} = i \mid \mathcal{F}_{T_n}\right) := \frac{\lambda_{\xi_n}(i, \mathbf{Y}_{T_{n-1}}) \sigma^i_{T_n - T_{n-1}}[n, \mathbf{Y}_{T_{n-1}}]}{\sum_{k=1}^K \lambda_{\xi_n}(k, \mathbf{Y}_{T_{n-1}}) \sigma^k_{T_n - T_{n-1}}[n, \mathbf{Y}_{T_{n-1}}]}$$

(ξ_n the identity of the defaulting firm) and then

$$\begin{aligned} \pi^i_{T_n} :=& \pi_{T_n}(\{i\}) = \sum_{h \neq i} P\left(X_{T_n^-} = h \mid \mathcal{F}_{T_n}\right) \frac{q\big(h, \mathbf{Y}_{T_{n-1}}; i, \mathbf{Y}_{T_n}\big)}{\sum_{j=1}^K q\big(h, \mathbf{Y}_{T_{n-1}}; j, \mathbf{Y}_{T_n}\big)} \\ &+ P\left(X_{T_n^-} = i \mid \mathcal{F}_{T_n}\right) \frac{q\big(i, \mathbf{Y}_{T_{n-1}}; i, \mathbf{Y}_{T_n}\big)}{\sum_{j=1}^K q\big(i, \mathbf{Y}_{T_{n-1}}; j, \mathbf{Y}_{T_n}\big)}, \quad 1 \leq i \leq K. \end{aligned} \tag{80}$$

The computability of these expressions hinges upon the solvability of the SDE in (79); various considerations to this effect are given in [35].

Further results. In [35] a number of additional results can be found. To begin with, a novel filter-approximation result is established; this justifies the use of the Markov-chain-filter as a computational tool for general state variable processes. Moreover, it is shown how to adapt particle filters such as the algorithm of [18] to models with joint jumps of $\mathbf{X}$ and $\mathbf{Y}$. This is important from a computational point of view: while Markov chain approximations are an effective tool for models where $\mathbf{X}$ is 2–3 dimensional, computations become prohibitively expensive in higher dimensions. Suitable particle filters on the other hand are a viable numerical scheme for moderate dimensions of the state process; see for instance [11] for an elaboration of this point in the context of standard nonlinear filtering problems.

7 Summary

We have discussed stochastic filtering problems that arise in the context of incomplete-information models for the term structure of interest rates and credit risk. The models considered were Markovian factor models. At the level of investors in secondary markets, the precise values of these factors are difficult to assess for a number of reasons, and so they have to be treated as latent factors to be filtered on the basis of the actual market information.

The main objective has been the pricing of derivative instruments. Other problems, including parameter estimation/calibration, have been touched upon only briefly. We have typically followed a two-step procedure: in the first step we determined the quantities of interest under full information as functions of the factors; in the second step we then derived their values under the actual market information by projecting the full information values on the subfiltration representing the market information. This is where filtering came in. In order to rule out the possibility of arbitrage, prices are expressed as expectations under a martingale/pricing measure. For pricing problems the filtering problems were therefore formulated directly under a martingale measure; for other purposes the real world/physical measure was found to be more appropriate.

References

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

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

[3] T. Bielecki and M. Rutkowski. *Credit Risk: Modeling, Valuation, and Hedging.* Springer, Berlin 2002.

[4] R. Bhar, C. Chiarella, H. Hung, and W. Runggaldier. The volatility of the instantaneous spot interest rate implied by arbitrage pricing - a dynamic Bayesian approach. *Automatica*, 42:1381–1393, 2005.

[5] T. Björk and L. Svensson. On the existence of finite dimensional realizations for nonlinear forward rate models. *Mathematical Finance*, 11:205–243, 2001.

[6] T. Björk. *Arbitrage Theory in Continuous Time.* Oxford University Press, Oxford, 2nd edition, 2004.

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

[8] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Polit. Economy*, 81(3):637–654, 1973.

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

[10] P. Brémaud. *Point Processes and Queues: Martingale Dynamics*. Springer, New York, 1981.

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

[12] C. Ceci and A. Gerardi. A model for high frequency data under partial information: a filtering approach. *International Journal of Theoretical and Applied Finance (IJTAF)*, 9:555-576, 2006.

[13] U. Cetin, R. Jarrow, P. Protter, and Y. Yildirim. Modeling credit risk witth partial information. *Annals of Applied Probability*, 14(3):1167–1178, 2004.

[14] C. Chiarella and O.K. Kwon. Forward rate dependent Markovian transformations of the Heath-Jarrow-Morton term structure model. *Finance and Stochastics*, 5:237–257, 2001.

[15] C. Chiarella, S. Pasquali, and W. Runggaldier. On filtering in Markovian term structure models. *Advances in Applied Probability*, 33:794–809, 2001.

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

[17] 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.

[18] D. Crisan and T. Lyons. A particle approximation of the solution of the Kushner-Stratonovich equation. *Probability Theory and Related Fields*, 115:549–578, 1999.

[19] J. Cvitanic, R. Liptser, and B. Rozovski. A filtering approach to tracking volatility from prices observed at random times. *The Annals of Applied Probability*, 16:1633–1652, 2006.

[20] J. Cvitanic, B. Rozovski, and Il. Zalyapin. Numerical estimation of volatility values from discretely observed diffusion data. *Journal of Computational Finance*, 9:1–36, 2006.

[21] M. Davis and V. Lo. Infectious defaults. *Quant. Finance*, 1:382–387, 2001.

[22] D. Duffie, A. Eckner, G. Horel, and L. Saita. Frailty correlated default. Preprint, Graduate School of Business, Stanford University 2008. Forthcoming in *Journal of Finance*.

[23] D. Duffie and N. Garleanu. Risk and valuation of collateralized debt obligations. *Financial Analysts Journal*, 57(1):41–59, 2001.

[24] D. Duffie and R. Kan. A yield factor model of interest rates. *Math. Finance*, 6, 1996.

[25] D. Duffie and D. Lando. Term structure of credit risk with incomplete accounting observations. *Econometrica*, 69:633–664, 2001.

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

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

[28] R. J. Elliott and V. Krishnamurthy. Exact finte-dimensional filters for maximum likelihood parameter estimation of continuous-time linear Gaussian systems. *SIAM Journal on Control and Optimization*, 35:1908–1923, 1997.

[29] D. Filipovic. Separable term structures and the maximal degree problem. *Mathematical Finance*, 12:341–349, 2002.

[30] C. Fontana. Affine multi-factor credit risk models under incomplete information: Filtering and parameter estimation. Thesis, University of Padova, 2007.

[31] R. Frey. Risk-minimization with incomplete information in a model for high frequency data. *Math. Finance*, 10(2):215–225, 2000.

[32] R. Frey, C. Prosdocimi, and W. Runggaldier. Affine credit risk models under incomplete information. In J. Akahori, S. Ogawa, and S. Watanabe, editors, *Stochastic Processes and Application to Mathematical Finance*, pages 97–113, Singapore, 2007. World Scientific.

[33] R. Frey and W. Runggaldier. Risk-minimizing hedging strategies under restricted information: the case of stochastic volatility models observed only at discrete random times. *Math. Methods Operations Res.*, 50(3):339–350, 1999.

[34] R. Frey and W. Runggaldier. A nonlinear filtering approach to volatility estimation with a view towards high frequency data. *International Journal of Theoretical and Applied Finance*, 4:199–210, 2001.

[35] R. Frey and W. Runggaldier. Pricing credit derivatives under incomplete information: a nonlinear filtering approach. preprint, Universität Leipzig, 2008. Available from www.math.uni-leipzig.de/˜ frey/publications-frey.html.

[36] R. Frey and T. Schmidt. Pricing corporate securities under noisy asset information. Preprint, Universität Leipzig, forthcoming in *Mathematical Finance*, 2007.

[37] R. Frey, T. Schmidt, and A. Gabih. Pricing and hedging of credit derivatives via nonlinear filtering. Preprint, Universität Leipzig, 2007. Available from www.math.uni-leipzig.de/˜ frey/publications-frey.html.

[38] A. Gerardi and P. Tardelli. Filtering on a Partially Observed Ultra-High-Frequency Data Model. *Acta Applicandae Mathematicae*, 91:193–205, 2006.

[39] K. Giesecke and L. R. Goldberg. Sequential defaults and incomplete information. *Journal of Risk*, 7:1–26, 2004.

[40] A. Gombani, S. Jaschke, and W. Runggaldier. A filtered no arbitrage model for term structures with noisy data. *Stochastic Processes and Applications*, 115:381–400, 2005.

[41] A. Gombani and W. Runggaldier. A filtering approach to pricing in multifactor term structure models. *International Journal of Theoretical and Applied Finance*, 4:303–320, 2001.

[42] B. I. Grigelionis. On stochastic equations for nonlinear filtering problem of stochastic processes. *Lietuvos Matematikos Rinkinys*, 12:37–51, 1972.

[43] D. Heath, R. Jarrow, and A. Morton, Bond pricing and the term structure of interest rates : a new methodology for contingent claims valuation. *Econometrica* 60:77–105, 1992.

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

[45] R. A. Jarrow and S. M. Turnbull. Pricing derivatives on financial securities subject to credit risk. *J. Finance*, L(1):53–85, 1995.

[46] W. Kliemann, G. Koch, and F. Marchetti. On the unnormalized solution of the filtering problem with counting process observations. *IEEE*, IT-36:1415–1425, 1990.

[47] S. Kusuoka. A remark on default risk models. *Adv. Math. Econ.*, 1:69–81, 1999.

[48] C. Landen. Bond pricing in a hidden markov model of the short rate. *Finance and Stochastics*, 4:371–389, 2001.

[49] D. Lando. Cox processes and credit risky securities. *Rev. Derivatives Res.*, 2:99–120, 1998.

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

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

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

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

[54] H. Pham. *This Volume.*

[55] E. Platen and W. J. Runggaldier. A benchmark approach to filtering in finance. *Asia-Pacific Financial Markets*, 11:79–105, 2005.

[56] W. J. Runggaldier. Estimation via stochastic filtering in financial market models. In : G. Yin, and Q. Zhang, editors, *Mathematics of Finance. Contemporary Mathematics, Vol. 351*, pages 309–318. American Mathematical Society, Providence R.I., 2004.

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

[58] M. Schweizer. Risk minimizing hedging strategies under restricted information. *Math. Finance*, 4:327–342, 1994.

[59] W. M. Wonham. Some applications of stochastic differential equations to optimal non-linear filtering,. *SIAM Journal on Control and Optimization*, 2:347–369, 1965.

[60] Y. Zeng. A partially observed model for micromovement of asset prices with Bayes estimation via filtering. *Mathematical Finance*, 13:411–444, 2003.