Pricing Credit Derivatives under Incomplete Information: a Nonlinear-Filtering Approach

Rüdiger Frey · Wolfgang Runggaldier

Version from January 23, 2009

Abstract This paper considers a general reduced form pricing model for credit derivatives where default intensities are driven by some factor process **X**. The process **X** is not directly observable for investors in secondary markets; rather, their information set consists of the default history and of noisy price observation for traded credit products. In this context the pricing of credit derivatives leads to a challenging nonlinear filtering problem. We provide recursive updating rules for the filter, derive a finite dimensional filter for the case where **X** follows a finite state Markov chain and propose a novel particle filtering algorithm. A numerical case study illustrates the properties of the proposed algorithms.

Keywords Credit derivatives, nonlinear filtering, marked point processes
AMS classification : 91B28, 93E11, 60G55
JEL classification: G13, C11

1 Introduction

This paper is concerned with the pricing of credit derivatives in reduced form portfolio credit risk models under incomplete information. We consider models where the default intensities of the firms in a given portfolio are driven by some Markov process **X**. We assume that **X** is not directly observable for investors trading in secondary markets; rather, their information set, denoted $(\mathcal{F}_t^I)$, is restricted to some special type of publicly available information, namely historical default- and price information. In this incomplete information context

Rüdiger Frey: Department of Mathematics, University of Leipzig, 04009 Leipzig, Germany. E-mail: ruediger.frey@math.uni-leipzig.de
Wolfgang Runggaldier: Department of Mathematics, University of Padua, Via Trieste 63, 35121 Padova, Italy. E-mail: runggal@math.unipd.it
Financial support from the European Science Foundation via the programme Advanced Mathematical Methods for Finance is gratefully acknowledged.

the pricing of credit derivatives leads to a two step procedure: in the first step one computes theoretical prices, (termed *model values*), with respect to a large filtration $(\mathcal{F}_t)$ such that $\mathbf{X}$ is $(\mathcal{F}_t)$-adapted; here Markov process techniques can be fruitfully employed. In the second step the price of non traded credit derivatives (from the viewpoint of secondary market investors) is computed by projecting model values onto the investor information $(\mathcal{F}_t^I)$. This projection is essentially a nonlinear filtering problem; one has to determine $\pi_{\mathbf{X}_t|\mathcal{F}_t^I}(d\boldsymbol{x})$, the conditional distribution of $\mathbf{X}_t$ given $\mathcal{F}_t^I$.

Credit risk models with incomplete information have been considered previously in the literature. Kusuoka (1999), Duffie & Lando (2001), Giesecke (2004), Jarrow & Protter (2004), Giesecke (2004), Coculescu, Geman, & Jeanblanc (2008) and Frey & Schmidt (2006) are concerned with structural models where the value of assets and/or liabilities is not directly observable. Reduced form credit risk models with incomplete information such as our paper have been considered by Schönbucher (2004), Collin-Dufresne, Goldstein & Helwege (2003) and Duffie, Eckner, Horel & Saita (2006). The structure of the latter three models is relatively similar: default intensities are driven by an unobservable factor (process) $\mathbf{X}$; given information about $\mathbf{X}$, the default times are conditionally independent, doubly stochastic random times; finally, the investor information $(\mathcal{F}_t^I)$ is given by the default history of the portfolio, augmented by economic covariates. Schönbucher (2004), and Collin-Dufresne et al. (2003) model the unobservable factor by a static random vector $\mathbf{X}$, called *frailty*; the conditional distribution $\pi_{\mathbf{X}|\mathcal{F}_t^I}$ is determined via Bayesian updating. Both papers point out that the successive updating of $\pi_{\mathbf{X}|\mathcal{F}_t^I}$ in reaction to incoming default observations generates *information-driven default contagion*: the news that some obligor has defaulted leads to an update in $\pi_{\mathbf{X}|\mathcal{F}_t^I}(d\boldsymbol{x})$ and hence typically to a jump in the $(\mathcal{F}_t^I)$-default intensity of the surviving firms. Duffie et al. (2006) model the unobservable factor $\mathbf{X}$ by an Ornstein-Uhlenbeck process. Their paper contains interesting empirical results; in particular, the analysis provides strong support 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.

Our paper differs from these contributions in two directions: First, in order to better capture the complicated price dynamics of traded credit products, we work in a general jump-diffusion model for the joint dynamics of the state process $\mathbf{X}$ and the default indicator process $\mathbf{Y}$ (the jump process associated with the default times). Our framework covers most reduced form models from the literature and includes in particular models where $\mathbf{X}$ may jump in reaction to a default event. Joint jumps of $\mathbf{X}$ and $\mathbf{Y}$ arise naturally in a variety of credit risk models, and several examples are given in Section 2.3. Joint jumps can in particular be used for the modeling of so-called *physical default contagion*, that is the fact that the default of a major corporation has at least temporarily an adverse impact on the survival probability of certain remaining firms. As shown in a recent empirical analysis of Azizpour & Giesecke (2008), physical default contagion is relevant even after the inclusion of frailty effects into the model. Our model is thus accounting for both potential sources for jumps in the prices of credit derivatives at defaults, namely physical and information-based contagion. Second, we use a different information set: in our setup the investor information $(\mathcal{F}_t^I)$ contains theoretical prices of traded credit derivatives observed in additive noise in addition to the default history of the firms under consideration. This is important, as market quotes for traded credit products are a crucial piece of information in any pricing model for credit derivatives.

In order to determine the conditional distribution $\pi_{\mathbf{X}_t|\mathcal{F}_t^I}(d\boldsymbol{x})$ in our setup we have to solve a challenging nonlinear filtering problem with mixed observations of marked point processes and diffusion type and with common jumps of point process observation $\mathbf{Y}$ and state process $\mathbf{X}$.Filtering problems with common jumps of the unobserved state process and of the observations have previously been discussed in the literature. First results can be found in Grigelionis (1972); the papers Kliemann, Koch & Marchetti (1990) and Ceci & Gerardi (2006) are concerned with scalar observations described by a pure jump process. The recent paper Cvitanic, Liptser & Rozovski (2006) 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 follow the innovations approach to nonlinear filtering.

In the present paper we take an alternative route which is based on ideas from the reference probability approach. In this way we obtain new general recursive filter equations. In case that $(\mathbf{X}, \mathbf{Y})$ is a finite state continuous time Markov chain our filter equations give rise to a finite dimensional filter. This is important for two reasons: on one hand many credit risk models that have recently appeared in the literature are of this form; on the other hand Markov chain models can be used as computational tools in models with general state variable processes. We establish a novel convergence result which justifies the use of Markov chain approximations in our setup. Moreover, using our filter formulas we are able to adapt - to our knowledge for the first time - particle filters such as the algorithm of Crisan & Lyons (1999) to models with joint jumps of $\mathbf{X}$ and $\mathbf{Y}$. Suitable particle filters are a viable numerical scheme for higher dimensions of the state spaces where Markov-chain approximations fail (Budhiraja, Chen & Lee 2007). We carry out numerical experiments illustrating the performance of both filtering algorithms.

The paper is organized as follows. In Section 2 we introduce our setup and provide various examples; moreover, we discuss the pricing of credit derivatives under incomplete information. The ensuing filtering problem is then studied in Sections 3 to 5.

2 An Information-based Approach to Credit-Derivatives Pricing

2.1 The Model

We work on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$; all stochastic processes considered will be $(\mathcal{F}_t)$-adapted. Since our focus is on the pricing of credit derivatives via martingale methods, P is interpreted as risk-neutral pricing measure. Throughout we consider a fixed credit portfolio consisting of a set of m firms. Our model is of the bottom-up type, that is we model the stochastic evolution of the default-state of the individual firms in the portfolio. The $(\mathcal{F}_t)$-stopping time τ_i denotes the default time of firm i and the current default state of the portfolio is summarized by the default indicator process $\mathbf{Y} = (Y_{t,1}, \dots, Y_{t,m})_{t \geq 0}$ with $Y_{t,i} = 1_{\{\tau_i \leq t\}}$. We assume that the factor process $\mathbf{X} = (X_{t,1}, \dots, X_{t,d})_{t \geq 0}$ and the default indicator process $\mathbf{Y}$ solve the following SDE system

$$\mathbf{X}_t = \mathbf{X}_0 + \int_0^t b(\mathbf{X}_{s-}, \mathbf{Y}_{s-}) ds + \int_0^t \sigma(\mathbf{X}_{s-}, \mathbf{Y}_{s-}) d\mathbf{W}_s + \int_0^t \int_E K^{\mathbf{X}}(\mathbf{X}_{s-}, \mathbf{Y}_{s-}, u) \mathcal{N}(ds, du), \tag{2.1}$$

$$Y_{t,j} = Y_{0,j} + \int_0^t \int_E (1 - Y_{s-,j}) K_j^{\mathbf{Y}}(\mathbf{X}_{s-}, \mathbf{Y}_{s-}, u) \mathcal{N}(ds, du), \ 1 \leq j \leq m. \tag{2.2}$$

Here $\mathbf{W}$ is a standard k-dimensional Brownian motion; the drift $b = (b_1, \dots, b_d)$ and the dispersion matrix $\sigma = (\sigma_{i,\ell})$, $1 \leq i \leq d, 1 \leq \ell \leq k$ are functions from $S^{\mathbf{X}} \times \{0,1\}^m$ 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 $(P, (\mathcal{F}_t))$-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 $\{0,1\}^m$. We assume that $K_j^{\mathbf{Y}}(\boldsymbol{x}, \boldsymbol{y}, u) \in \{0,1\}$ for all $\boldsymbol{x}, \boldsymbol{y}, u$ and all $1 \leq j \leq m$, so that the solution of (2.2) is in fact of the form $Y_{t,j} = 1_{\{\tau_j \leq t\}}$. By $\Sigma_{\mathbf{X}}(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{d \times d}$ we denote the matrix $\sigma(\boldsymbol{x}, \boldsymbol{y})\sigma'(\boldsymbol{x}, \boldsymbol{y})$. Define

$$D_i^{\mathbf{X}}(\boldsymbol{x}, \boldsymbol{y}) := \{u \in E : K_i^{\mathbf{X}}(\boldsymbol{x}, \boldsymbol{y}, u) \neq 0\}, \quad 1 \leq i \leq d, \tag{2.3}$$

$$D_j^{\mathbf{Y}}(\boldsymbol{x}, \boldsymbol{y}) := \{u \in E : K_j^{\mathbf{Y}}(\boldsymbol{x}, \boldsymbol{y}, u) \neq 0\}, \quad 1 \leq j \leq m. \tag{2.4}$$

We make the following assumptions.

A1. For every pair $(\boldsymbol{x}, \boldsymbol{y}) \in S^{\mathbf{X}} \times \{0,1\}^m$ the SDE system (2.1), (2.2) has a global solution with $\mathbf{X}_0 = \boldsymbol{x}, \mathbf{Y}_0 = \boldsymbol{y}$; moreover, pathwise uniqueness holds. Sufficient conditions for A1 are discussed in Remark 2.1 below.

A2. For all $1 \le i \le d$, $1 \le j \le m$ and all $T \ge 0$ we have

$$E\Big(\int_0^T F_{\mathcal{N}}(D_i^{\mathbf{X}}(\mathbf{X}_s, \mathbf{Y}_s))\,ds\Big) + E\Big(\int_0^T F_{\mathcal{N}}(D_j^{\mathbf{Y}}(\mathbf{X}_s, \mathbf{Y}_s))ds\Big) < \infty.$$

A3. For all $1 \le j_1 < j_2 \le m$ and all $(\boldsymbol{x}, \boldsymbol{y}) \in S^{\mathbf{X}} \times \{0,1\}^m$ we have

$$F_{\mathcal{N}}(D_{j_1}^{\mathbf{Y}}(\boldsymbol{x}, \boldsymbol{y}) \cap D_{j_2}^{\mathbf{Y}}(\boldsymbol{x}, \boldsymbol{y})) = 0\,.$$

Assumption A2 ensures that the expected number of jumps of $\mathbf{X}$ on every time interval $[0,T]$ is finite and that $\tau_i > 0$ for all firms i such that $Y_{0,i} = 0$. Assumption A3 ensures that for $j_1 \ne j_2$ 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 (2.1), (2.2) 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

$$F_{\mathcal{N}}\big(D_j^{\mathbf{Y}}(\mathbf{X}_{\tau_j-}, \mathbf{Y}_{\tau_j-}) \cap D_i^{\mathbf{X}}(\mathbf{X}_{\tau_j-}, \mathbf{Y}_{\tau_j-})\big) > 0 \text{ for some } 1 \le i \le d\,. \quad (2.5)$$

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

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

is an $(\mathcal{F}_t)$-martingale, so that $\lambda_j(\mathbf{X}_{t-}, \mathbf{Y}_{t-}) := F_{\mathcal{N}}(D_j^{\mathbf{Y}}(\mathbf{X}_{t-}, \mathbf{Y}_{t-}))$ is the $(\mathcal{F}_t)$-default intensity of firm j. In Subsection 2.3 below we show how various reduced-form credit risk models can be constructed as solutions of the SDE system (2.1), (2.2).

Remark 2.1 (Sufficient conditions for A1) There are various types of conditions ensuring strong existence and uniqueness for the SDE-system (2.1), (2.2). In Theorem 2.2 of Kliemann et al. (1990) strong existence and uniqueness is proved under growth conditions on $b(\boldsymbol{x}, \boldsymbol{y})$, $\Sigma_X(\boldsymbol{x}, \boldsymbol{y})$, $F_{\mathcal{N}}(D_i^{\mathbf{X}}(\boldsymbol{x}, \boldsymbol{y}))$, $F_{\mathcal{N}}(D_j^{\mathbf{Y}}(\boldsymbol{x}, \boldsymbol{y}))$ and under the additional assumption that for every fixed $\boldsymbol{y}$ the SDE $d\mathbf{X}_t = b(\mathbf{X}_t, \boldsymbol{y})ds + \sigma(\mathbf{X}_t, \boldsymbol{y})d\mathbf{W}_s$ has a unique weak solution which is moreover a Feller process. Alternatively, one can impose growth and Lipschitz conditions on the data of the problem; see for instance Appendix 1, Section 4 of Ceci & Gerardi (2006).

Some notation. Typically we take $\mathbf{Y}_0 = \mathbf{0}$. In that case Assumption A2 permits us to introduce the *ordered default times* $0 = T_0 < T_1 < \cdots < T_m < T_{m+1} := \infty$ and the counting process $N_t := \max\{n \le m \colon T_n \le t\}$; for $n \ge 1$ the rv ξ_n denotes the identity of the firm defaulting at T_n. The sequence $(T_n, \xi_n)_{1 \le n \le m}$ gives a representation of $\mathbf{Y}$ as marked point process. The σ-field $\mathcal{H}_t = \sigma(\mathbf{Y}_s : s \le t) = \sigma\big(\{(T_n, \xi_n) : n = 1, \ldots, N_t\}\big)$ is the internal filtration of $\mathbf{Y}$ or, in economic terms, the default history of the portfolio at time t. Note that any $\mathcal{H}_t$-measurable function $\boldsymbol{a}_t(\cdot)\colon \Omega \times S^{\mathbf{X}} \to \mathbb{R}^\ell$ is of the form

$$\boldsymbol{a}_t(\omega;\boldsymbol{x}) = \sum_{n=0}^{m} 1_{\{T_n(\omega)\le t<T_{n+1}(\omega)\}}\, \boldsymbol{a}_n\big(t,\boldsymbol{x};\, T_j(\omega),\xi_j(\omega) : 1\le j\le n\big), \quad (2.6)$$

for functions $\boldsymbol{a}_n\colon [0,\infty)\times S^{\mathbf{X}} \times \big([0,\infty)\times\{1,\dots,m\}\big)^n \to \mathbb{R}^\ell$. For further use we finally define the functions

$$\bar\lambda_j(\boldsymbol{x},\boldsymbol{y}) = \sum_{i=1}^{j}(1-y_i)\lambda_i(\boldsymbol{x},\boldsymbol{y}),\ 0\le j\le m,\ \text{and}\ \bar\lambda(\boldsymbol{x},\boldsymbol{y}) := \bar\lambda_m(\boldsymbol{x},\boldsymbol{y}). \quad (2.7)$$

2.2 Credit Derivatives and Incomplete Information

Credit derivatives are securities whose payoff at maturity T depends on the default history of some underlying reference portfolio; in abstract terms their payoff is hence given by some $\mathcal{H}_T$-measurable random variable H. Examples include corporate bonds, credit default swaps or collaterized debt obligations (CDOs). We study the pricing of these securities using the popular martingale modeling approach; as mentioned before, P represents the martingale measure used for pricing. For simplicity we assume that the default-free short rate $r_t > 0$ is deterministic. Recall that the σ-field $\mathcal{F}_t^I \subset \mathcal{F}_t$ represents the information available to secondary market investors at time t; a formal description of $(\mathcal{F}_t^I)$ is given in Assumption A4 below. We introduce two notions for the the value/price of a credit derivative with maturity $T>t$ and payoff H. The *model-value* $\tilde H_t$ is defined to be

$$\tilde H_t := E\Big(e^{-\int_t^T r_s ds} H \mid \mathcal{F}_t\Big), \quad t\le T\,. \quad (2.8)$$

In case that H is traded we view the model value $\tilde H$ as theoretical price of the claim; in our setup actual market quotes may however deviate temporarily from theoretical prices (see the discussion following Assumption A4 below). For non traded credit derivatives we define the *secondary market price* by

$$H_t := E\Big(e^{-\int_t^T r_s ds} H \mid \mathcal{F}_t^I\Big). \quad (2.9)$$

Note that in our context H_t is the correct notion of the secondary market price for a non traded credit derivative since this quantity is defined with respect to the information set actually available to investors.

By the Markovianity of the pair $(\mathbf{X},\mathbf{Y})$, the model value $\tilde H_t$ is of the form $\tilde H_t = \boldsymbol{a}_t(\mathbf{X}_t)$ for some $\mathcal{H}_t$-measurable random function $\boldsymbol{a}_t : \Omega\times S^{\mathbf{X}} \to \mathbb{R}$ as in (2.6). Now we get by iterated conditional expectations

$$H_t = E\Big(E\big(e^{-\int_t^T r_s ds} H \mid \mathcal{F}_t\big) \mid \mathcal{F}_t^I\Big) = E\big(\boldsymbol{a}_t(\mathbf{X}_t) \mid \mathcal{F}_t^I\big)\,. \quad (2.10)$$

In order to compute the secondary-market price H_t from the right hand side of (2.10) we thus need to determine in weak form the conditional distribution $\pi_{\mathbf{X}_t|\mathcal{F}_t^I}(d\boldsymbol{x})$. This is a typical nonlinear filtering problem which is studied in detail in Sections 3, 4 and 5.[1]

Remark 2.2 It is well-known that intensities with respect to subfiltrations can be computed by projection. Hence the $(\mathcal{F}_t^I)$-default intensity of firm j is given by the left-continuous version of

$$\hat{\lambda}_{t,j} = E\left(\lambda_j(\mathbf{X}_t, \mathbf{Y}_t) \mid \mathcal{F}_t^I\right) = \int_{\mathbb{R}^d} \lambda_j(\boldsymbol{x}, \mathbf{Y}_t)\, \pi_{\mathbf{X}_t|\mathcal{F}_t^I}(d\boldsymbol{x})\,, \quad t \le \tau_j. \tag{2.11}$$

In particular, new default information such as the news that obligor $i \neq j$ has defaulted leads to an update in the conditional distribution $\pi_{\mathbf{X}_t|\mathcal{F}_t^I}(d\boldsymbol{x})$ and hence to a jump in the $(\mathcal{F}_t^I)$-default intensity of firm j.

The investor filtration. We assume that investors observe the default history of the portfolio under consideration. Moreover, we assume that they have noisy information about the model value $\boldsymbol{a}_t(\mathbf{X}_t) = (a_{t,1}(\mathbf{X}_t), \dots, a_{t,\ell}(\mathbf{X}_t))$ of ℓ traded credit derivatives. As explained below, in continuous time the appropriate way to model this is to assume that investors observe the vector $\boldsymbol{a}_t(\mathbf{X}_t)$ in additive Gaussian noise. We therefore make the assumption

A4. $\mathcal{F}_t^I = \mathcal{H}_t \vee \mathcal{F}_t^{\mathbf{Z}}$, $t \ge 0$, where the ℓ-dimensional process $\mathbf{Z}$ solves the SDE

$$d\mathbf{Z}_t = \boldsymbol{a}_t(\mathbf{X}_t)dt + v d\boldsymbol{\beta}_t\,. \tag{2.12}$$

Here $\boldsymbol{\beta}$ is an ℓ-dimensional standard Brownian motion on $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, independent of $\mathbf{X}$ and $\mathbf{Y}$; v denotes an *invertible* $\ell \times \ell$ matrix of constants; $\boldsymbol{a}_t(\cdot)\colon \Omega \times S^{\mathbf{X}} \to \mathbb{R}^\ell$ is an $\mathcal{H}_t$-measurable random function as in (2.6); moreover, the functions $\boldsymbol{a}_n$ in (2.6) are continuous and bounded.

Now we turn to the financial interpretation of $\mathbf{Z}$. Suppose that secondary market investors observe market quotes for the traded credit derivatives such as (logarithmic) prices or spreads at discrete points in time $t_k = k\Delta$, and that these quotes are of the form $\boldsymbol{z}_{t_k} = \boldsymbol{a}_{t_k}(\mathbf{X}_{t_k}) + \boldsymbol{\epsilon}_k$ for an iid sequence of $\mathbb{R}^\ell$-valued noise variables $(\boldsymbol{\epsilon}_k)_k$, independent of $\mathbf{X}$, with $E(\boldsymbol{\epsilon}_1) = \mathbf{0}$ and $\operatorname{cov}(\boldsymbol{\epsilon}_1) = \tilde{\Sigma}_{\boldsymbol{\epsilon}}$. Assume that $\tilde{\Sigma}_{\boldsymbol{\epsilon}}$ is positive definite and choose an invertible root $\tilde{v}$. Define the scaled *cumulative observation process* $\mathbf{Z}_t^\Delta := \Delta \sum_{t_k \le t} \boldsymbol{z}_{t_k}$ and let $v = \sqrt{\Delta}\tilde{v}$. Then we have for Δ small, using Donsker's invariance principle,

$$\mathbf{Z}_t^\Delta = \sum_{t_k \le t} \boldsymbol{a}_{t_k}(\mathbf{X}_{t_k})\Delta + \Delta \sum_{t_k \le t} \boldsymbol{\epsilon}_k \approx \int_0^t \boldsymbol{a}_s(\mathbf{X}_s)ds + v\boldsymbol{\beta}_t. \tag{2.13}$$

[1] In the important case of credit default swaps and CDOs (2.10) applies separately to the premium payment leg and the default payment leg of the transaction; the fair secondary market spread is then computed by equating the secondary market price of both legs of the transaction.

We have $\Sigma_{\mathbf{Z}} := vv' = \Delta\tilde{\Sigma}_{\boldsymbol{\epsilon}}$, so that the instantaneous covariance matrix of $\mathbf{Z}$ in (2.13) is proportional to the covariance matrix $\tilde{\Sigma}_{\boldsymbol{\epsilon}}$ of the noise variables and inversely proportional to the observation frequency $1/\Delta$. Note that with this interpretation of $\mathbf{Z}$ the information contained in observable market quotes is taken into account when pricing non-traded claims according to (2.9). The noise - $(\boldsymbol{\epsilon}_k)_k$ respectively $v\boldsymbol{\beta}$ - represents classical observation errors such as bid-ask spreads, transmission errors or non-simultaneous quotes as well as (spurious) deviations of market quotes from theoretical prices. In applications the error covariance matrix v has to be chosen by the modeler by balancing flexibility (the fact that with low error variances $\pi_{\mathbf{X}_t|\mathcal{F}_t^I}$ reacts swiftly to new price information) against stability of the filtering algorithm which is being used.

Remark 2.3 As pointed out to us by a referee, traders might base their decisions also on general economic conditions represented by observable covariates. This case is not covered by Assumption A4. However, observable covariates can easily be incorporated into our framework. The idea of interpreting observed prices of derivatives as noisy observations of functions of some unobserved factor processes is pursued also in Gombani, Jaschke & Runggaldier (2005).

2.3 Examples

The following examples show that a great variety of models are covered by our framework.

1. We begin with the standard models with conditionally independent, doubly stochastic default times. In these models it is assumed that $\mathbf{X}$ follows a jump diffusion model of the form

$$d\mathbf{X}_t = b(\mathbf{X}_t)dt + \sigma(\mathbf{X}_t)d\mathbf{W}_t + d\mathbf{J}_t, \tag{2.14}$$

$\mathbf{J}$ an $\mathbb{R}^d$-valued compound Poisson process with compensator measure $F_{\mathbf{J}}(d\boldsymbol{x})ds$; moreover, given $\mathcal{F}_\infty^X$, the default times are conditionally independent with hazard rate $\lambda_i(\mathbf{X}_t)$. A typical representative of this model class is the popular affine jump-diffusion model of Duffie & Garleanu (2001). A possible choice for $\mathcal{N}$, $K^{\mathbf{X}}$ and $K^{\mathbf{Y}}$ is 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

$$K_j^{\mathbf{Y}}(\boldsymbol{x}, \boldsymbol{u}) = 1_{\left[\sum_{i=1}^{j-1}\lambda_i(\boldsymbol{x}), \sum_{i=1}^{j}\lambda_i(\boldsymbol{x})\right]}(u_{d+1}), \quad 1 \le j \le m, \text{ and} \tag{2.15}$$

$$K_i^{\mathbf{X}}(\boldsymbol{x}, \boldsymbol{u}) = u_i 1_{[-1,0)}(u_{d+1}), \quad 1 \le i \le d. \tag{2.16}$$

Note that by the choice of $K^{\mathbf{X}}$ and $K^{\mathbf{Y}}$, $F_{\mathcal{N}}(D_i^{\mathbf{X}}(\boldsymbol{x}) \cap D_j^{\mathbf{Y}}(\boldsymbol{x})) = 0$ for all $1 \le i \le d$, $1 \le j \le m$ and all $\boldsymbol{x}$ in $S^{\mathbf{X}}$. A filter algorithm for the special case where $(\mathcal{F}_t^I) = (\mathcal{H}_t)$ (only default information) and where the coefficients in (2.14) are affine functions of $\mathbf{X}_t$ is proposed in Frey, Prosdocimi & Runggaldier (2007).

2. Next we discuss a version of the *infectious-defaults model* of Davis & Lo (2001); here the state process does jump in reaction to default events. Assume that X is modeled as a finite state Markov chain with state space $S^X = \{1,\dots,K\}$ and that 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 T_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 $X_{T_n-} = K$, where X remains constant). If the system is in an "ignited state", i.e. 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 physical default contagion, as the default intensities of the remaining firms are increased. This leads to a downward jump in the model value (and hence to an increase in the credit spread) of a zero-coupon bond issued by non-defaulted firms. In Section 4.1 we show how this model can be embedded into the general framework (2.1), (2.2) by a proper choice of $F_{\mathcal{N}}$, $K^{\mathbf{X}}$ and $K^{\mathbf{Y}}$.

3. Finally we turn to the model of Frey, & Schmidt (2009) (FS-model in the sequel). This model is of interest in our context, since the state variable process is itself the solution of a filtering problem, so that the common jumps of state variable and default indicator are generated by information effects. In the FS-model it is assumed that the default times are conditionally independent doubly stochastic random times on some filtered probability space $(\Omega, \mathcal{F}, (\tilde{\mathcal{F}}_t), P)$ and that the $(\tilde{\mathcal{F}}_t)$-default intensities are driven by an unobservable finite-state Markov chain Ψ with state space $\{1,\dots,K\}$ and generator matrix Q^{Ψ} modelling the state of the economy. Frey, & Schmidt (2009) consider a market where ℓ credit derivatives with payoff $H_1, \dots H_\ell$ are traded. The model value of these contracts is defined as conditional expectations with respect to the information set $\mathcal{F}_t = \mathcal{F}_t^Y \vee \mathcal{F}_t^U$, termed market filtration, i.e. we have $\tilde{H}_{t,j} = E\big(\exp(-\int_t^T r_s\,ds)H_j \mid \mathcal{F}_t\big)$. The process U is given by $U_t = \int_0^t \alpha(\Psi_s)ds + B_t$, B a standard $(\tilde{\mathcal{F}}_t)$-Brownian motion independent of Ψ and Y; it models in abstract form the information about the state of the economy Ψ contained in the prices of traded credit derivatives. Define the conditional probabilities

$$p_t^k := P(\Psi_t = k \mid \mathcal{F}_t),\ 1 \leq k \leq K, \ \text{ and let } \boldsymbol{p}_t := (p_t^1, \dots, p_t^K). \tag{2.17}$$

The process $\boldsymbol{p} = (\boldsymbol{p}_t)_{t\geq 0}$ is a natural state variable process for the model in the market filtration; it thus plays the role of the process $\mathbf{X}$ from (2.1). The reasons are the following: first, denoting the $(\tilde{\mathcal{F}}_t)$-default intensities by $\nu_i(\Psi_t)$, the $(\mathcal{F}_t)$-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 $(\tilde{\mathcal{F}}_t)$-Markovianity of Ψ and $\mathbf{Y}$, the conditional expectation $E\big(\exp(-\int_0^{T-t} r_s ds)H_j \mid \tilde{\mathcal{F}}_t\big)$ is given by some $\mathcal{H}_t$-measurable function $\tilde{a}_{t,j}(\Psi_t)$. By iterated conditional expectations theoretical prices can therefore be expressed as $\mathcal{H}_t$-measurable functions of $\boldsymbol{p}_t$ as well:

$$\tilde{H}_{t,j} = E(\tilde{a}_{t,j}(\Psi_t) \mid \mathcal{F}_t) = \sum_{k=1}^{K} p_t^k \tilde{a}_{t,j}(k) =: a_{t,j}(\boldsymbol{p}_t)\,,\ 1 \le j \le \ell. \tag{2.18}$$

Using the innovations approach to nonlinear filtering, in FS the following K-dimensional SDE for the process $\boldsymbol{p} = (p_t^1, \dots, p_t^K)_{t\ge 0}$ is derived:

$$dp_t^k = \sum_{i=1}^{K} Q_{i,k}^{\Psi}\, p_t^i dt + \sum_{i=1}^{m} \boldsymbol{\gamma}_i^k(\boldsymbol{p}_{t-})\, d\big(Y_{t,i} - \lambda_i(\boldsymbol{p}_t)dt\big) + \delta^k(\boldsymbol{p}_{t-})\, dW_t\,. \tag{2.19}$$

Here $W_t = U_t - \int_0^t \sum_{k=1}^K \alpha(k) p_s^k\, ds$ is $(\mathcal{F}_t)$-Brownian motion; the coefficients in (2.19) are given by the functions

$$\gamma_j^k(\boldsymbol{p}) = p^k\Big(\frac{\nu_j(k)}{\sum_{n=1}^K \nu_j(n) p^n} - 1\Big)\,,\ \delta^k(\boldsymbol{p}) = p^k\Big(\alpha(k) - \sum_{n=1}^{K} p^n \alpha(n)\Big)\,. \tag{2.20}$$

In the FS-model the process U is unobservable for secondary market investors; these investors have to back out the conditional distribution of probability vector $\boldsymbol{p}_t$ given historical default- and price information. Assuming that their information set $(\mathcal{F}_t^I)$ is as in Assumption A4 of the present paper, this leads to a nonlinear filtering problem with state variable process $\boldsymbol{p}$ and observations $\mathbf{Y}$ and $\mathbf{Z}$ with $\mathbf{Z}_t = \int_0^t \boldsymbol{a}_s(\boldsymbol{p}_s)ds + v\boldsymbol{\beta}_t$. Thus the solution of the filtering problem with respect to $(\mathcal{F}_t)$ becomes the state variable $\boldsymbol{p} = \mathbf{X}$ of the filtering problem with respect to the investor filtration $(\mathcal{F}_t^I)$. The latter filtering problem is covered by our setup, as the processes $\boldsymbol{p}$ and $\mathbf{Y}$ follow an SDE-system of the form (2.1) (2.2). Note in particular that at a default time the probability vector $\boldsymbol{p}$ is updated according to (2.19), so that there are common jumps of state variable $\boldsymbol{p}$ and observation $\mathbf{Y}$.

The FS-model has a number of attractive features: first, by (2.18) the main numerical task is the evaluation of the functions $\tilde{a}_{t,j}(\psi)$; as this function is computed in a simple setup with conditionally independent defaults, computations become relatively easy. Moreover, the model generates a rich set of price dynamics with randomly fluctuating credit spreads and default contagion. Results of numerical filter experiments for the FS-model are reported in Section 4.3.

3 Filter Equations

The remainder of this paper is devoted to the analysis of the the following filtering problem which is the crucial step in the computation of secondary market prices: given a bounded function $f : \mathbb{R}^d \to \mathbb{R}$ develop a recursive approach to computing $\pi_t f := E(f(\mathbf{X}_t) \mid \mathcal{F}_t^I)$.

We begin with a brief overview of our analysis. As a first step, in Subsection 3.1 we use a well known change of measure argument from the reference probability approach in order to reduce the filtering problem to the case where $(\mathcal{F}_t^I)$ consists only of the default history $(\mathcal{H}_t)$. In Subsection 3.2 we study the dynamics of $\mathbf{X}$ between default events. This is a non standard step which is necessary, given the common jumps of $\mathbf{X}$ and $\mathbf{Y}$. Equipped with these results we can derive general filter formulas in Subsections 3.3 and 3.4. Section 4 is devoted to computational aspects: we derive a finite dimensional filter for the case when $(\mathbf{X}, \mathbf{Y})$ follows a finite state Markov chain, adapt existing particle-filtering algorithms to our more general setup and present some numerical experiments. In Section 5 we finally discuss the the filter convergence for finite-state Markov approximations.

Notation. In the sequel we use the following pieces of notation. By $\Sigma_{\mathbf{Z}}^{-1} = (vv')^{-1}$ we denote the inverse of the instantaneous covariance matrix of $\mathbf{Z}$; for any vector $\boldsymbol{a} \in \mathbb{R}^\ell$ we define $\|\boldsymbol{a}\|^2_{\Sigma_{\mathbf{Z}}^{-1}} := \boldsymbol{a}'\Sigma_{\mathbf{Z}}^{-1}\boldsymbol{a}$.

3.1 Measure transformation and reduction to $(\mathcal{H}_t)$

It will be convenient to model the processes $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ on a product space $(\Omega, \mathcal{F}, (\mathcal{F}_t), R^0)$ so that $\mathbf{Z}$ is independent of $\mathbf{X}$ and $\mathbf{Y}$ and to revert to the original model dynamics via an equivalent change of measure. For this we denote by $(\Omega_2, \mathcal{F}_2, (\mathcal{F}_t^2), P^{0,\ell})$ the ℓ-dimensional Wiener space with coordinate process $\boldsymbol{\beta}^0$, i.e. $\boldsymbol{\beta}^0_t(\omega_2) = \omega_2(t)$. Given some probability space $(\Omega_1, \mathcal{F}_1, (\mathcal{F}_t^1), P)$ supporting a solution $(\mathbf{X}, \mathbf{Y})$ of the SDE-system (2.1), (2.2), let $\Omega := \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$, $\mathcal{F}_t = \mathcal{F}_t^1 \otimes \mathcal{F}_t^2$, $R^0 := P \times P^{0,\ell}$, and put for $\omega = (\omega_1, \omega_2) \in \Omega$

$$\mathbf{X}_t(\omega) := \mathbf{X}_t(\omega_1),\ \mathbf{Y}_t(\omega) := \mathbf{Y}_t(\omega_1), \text{ and } \boldsymbol{\beta}^0_t(\omega) := \boldsymbol{\beta}^0_t(\omega_2).$$

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

$$\begin{aligned} L_t(\omega_1, \omega_2) &= \exp\Big\{ \int_0^t \big(v^{-1}\boldsymbol{a}_s(\mathbf{X}_s)(\omega_1)\big)' \, d\boldsymbol{\beta}^0_s(\omega_2) - \frac{1}{2}\int_0^t \|\boldsymbol{a}_s(\mathbf{X}_s)(\omega_2)\|^2_{\Sigma_{\mathbf{Z}}^{-1}} \, ds \Big\} \\ &= \exp\Big\{ \int_0^t \boldsymbol{a}'_s(\mathbf{X}_s)(\omega_1)\Sigma_{\mathbf{Z}}^{-1} d\mathbf{Z}_s(\omega_2) - \frac{1}{2}\int_0^t \|\boldsymbol{a}_s(\mathbf{X}_s)(\omega_1)\|^2_{\Sigma_{\mathbf{Z}}^{-1}} \, ds \Big\}, \end{aligned} \tag{3.1}$$

and note that the process L is indeed a R^0-martingale of mean one as $\boldsymbol{a}_t(\cdot)$ is bounded by A4. Using the Girsanov theorem for Brownian motion we therefore obtain that, under R, the process $\mathbf{Z}$ has the original dynamics (2.12); moreover, since $\boldsymbol{\beta}^0$ is orthogonal to both $\mathbf{W}$ and the martingale that results from compensating the counting measure $\mathcal{N}$, the above measure transformation induces no changes in the law of $\mathbf{X}$ and $\mathbf{Y}$. Hence under R the triple of processes $(\mathbf{X},\mathbf{Y},\mathbf{Z})$ has indeed the correct joint law. Notice finally that by (3.1), L can be expressed in terms of the observation $\mathbf{Z}$. By the following Bayes formula, known as Kallianpur-Striebel formula (Kallianpur & Striebel 1968), we then have

$$\pi_t f := E^R(f(\mathbf{X}_t) \mid \mathcal{F}_t^I) = \frac{E^{R^0}(f(\mathbf{X}_t)L_t \mid \mathcal{F}_t^I)}{E^{R^0}(L_t \mid \mathcal{F}_t^I)}\,, \tag{3.2}$$

so that, to compute $\pi_t f$, it suffices to compute the numerator on the right-hand side in (3.2).

Recall that $\mathcal{F}_t^I = \mathcal{H}_t \vee \mathcal{F}_t^{\mathbf{Z}}$. Next we reduce the conditioning on $\mathcal{F}_t^I$ to a conditioning on $\mathcal{H}_t$. Since $\mathbf{Z}_t = v\boldsymbol{\beta}_t^0$ with v invertible, we have $\mathcal{F}_t^{\mathbf{Z}} = \mathcal{F}_t^{\boldsymbol{\beta}^0}$, i.e. given $\mathcal{F}_t^{\mathbf{Z}}$, the process $\omega_2(s)$, $s \le t$ is "known." Using the Fubini theorem and the product structure of $(\Omega,\mathcal{F},(\mathcal{F}_t),R^0)$ we therefore get

$$E^{R^0}(f(\mathbf{X}_t)L_t \mid \mathcal{H}_t \vee \mathcal{F}_t^{\mathbf{Z}})(\omega) = E^P(f(\mathbf{X}_t)L_t(\cdot,\omega_2) \mid \mathcal{H}_t)(\omega_1)\,. \tag{3.3}$$

In order to compute $\pi_t f$ we thus have to evaluate the conditional expectation on the right hand side of (3.3). Note that this involves only the first component $(\Omega_1,\mathcal{F}_1,(\mathcal{F}_t^1),P)$ of the underlying probability space and hence only the joint law of $\mathbf{X}$ and $\mathbf{Y}$. In orderto ease the notation expectations with respect to that law will be simply denoted by E (instead of E^P); moreover the arguments of L_t will usually be omitted.

3.2 Dynamics of $\mathbf{X}$ between default times

Next we discuss the dynamics of $\mathbf{X}$ for $t \in [T_{n-1},T_n)$ i.e. between default times. This is a prerequisite for the filtering equations in the next subsections and for the derivation of approximation results in Section 5. For this purpose we define the new kernel

$$\bar{K}^{\mathbf{X}}(\boldsymbol{x},\boldsymbol{y},u) := \begin{cases} 0, & \text{if} \quad u \in \bar{D}^{\mathbf{Y}}(\boldsymbol{x},\boldsymbol{y}) := \bigcup_{\{j\,:\,\boldsymbol{y}_j=0\}} D_j^{\mathbf{Y}}(\boldsymbol{x},\boldsymbol{y})\,, \\ K^{\mathbf{X}}(\boldsymbol{x},\boldsymbol{y},u) & \text{else.} \end{cases} \tag{3.4}$$

We will see shortly that the kernel $\bar{K}^{\mathbf{X}}$ governs the jumps of $\mathbf{X}$ between default times. For instance, in case of the extended Davis Lo model, $K^X(k,\boldsymbol{y},u) = -1_{[-\gamma(k),0]}(u)$, $k = 1,\cdots,K$, reflecting the fact that between defaults X can only jump downwards.

Consider, for $t > T_{n-1}$, the SDE system

$$\widetilde{\mathbf{X}}_t = \mathbf{X}_{T_{n-1}} + \int_{T_{n-1}}^{t} b(\widetilde{\mathbf{X}}_s, \mathbf{Y}_{T_{n-1}})ds + \int_{T_{n-1}}^{t} \sigma(\widetilde{\mathbf{X}}_s, \mathbf{Y}_{T_{n-1}})d\mathbf{W}_s$$

$$+ \int_{T_{n-1}}^{t} \int_{E} \bar{K}^{\mathbf{X}}(\widetilde{\mathbf{X}}_{s-}, \mathbf{Y}_{T_{n-1}}, u)\mathcal{N}(ds, du), \tag{3.5}$$

$$\widetilde{Y}_{t,j} = \mathbf{Y}_{T_{n-1},j} + \int_{T_{n-1}}^{t} \int_{E} (1 - \widetilde{Y}_{s-,j})K_j^{\mathbf{Y}}(\widetilde{\mathbf{X}}_{s-}, \widetilde{\mathbf{Y}}_{T_{n-1}}, u)\mathcal{N}(ds, du), \quad 1 \le j \le m.$$

Comparing the system (3.5) with the original model dynamics (2.1), (2.2), we see that $K^{\mathbf{X}}$ has been replaced with $\bar{K}^{\mathbf{X}}$; moreover, in the coefficients of the equation the "initial value" $\mathbf{Y}_{T_{n-1}}$ replaces $\mathbf{Y}_{t-}$. Since, for $T_{n-1} \le t < T_n$, there are no atoms of $\mathcal{N}(ds, du)$ in $\{t\} \times \bar{D}^{\mathbf{Y}}(\mathbf{X}_{t-}, \mathbf{Y}_{T_{n-1}})$, by definition of $\bar{K}^{\mathbf{X}}$ in (3.4) we have for $T_{n-1} \le t < T_n$ the equality

$$\int_{T_{n-1}}^{t} \int_{E} K^{\mathbf{X}}(\mathbf{X}_{s-}, \mathbf{Y}_{s-}, u)\mathcal{N}(ds, du) = \int_{T_{n-1}}^{t} \int_{E} \bar{K}^{\mathbf{X}}(\mathbf{X}_{s-}, \mathbf{Y}_{T_{n-1}}, u)\mathcal{N}(ds, du)\,.$$

Strong uniqueness of (3.5) implied by Assumption A1 therefore yields that a.s.

$$\mathbf{X}_t = \widetilde{\mathbf{X}}_t,\ T_{n-1} \le t < T_n,\ \text{and } \mathbf{Y}_t = \widetilde{\mathbf{Y}}_t,\ T_{n-1} \le t \le T_n\ (T_n \text{ included})\,. \tag{3.6}$$

Denote by $\bar{P}_{(\boldsymbol{x},\boldsymbol{y})}$ the law of the solution $(\bar{\mathbf{X}}, \bar{\mathbf{Y}})$ to the SDE system (3.5) starting at $t = 0$ in the point $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^d \times \{0,1\}^m$. Now the law of the processes $\big(\widetilde{\mathbf{X}}_{t+T_{n-1}}, \widetilde{\mathbf{Y}}_{t+T_{n-1}}\big)_{t\ge 0}$ is obviously equal to $\bar{P}_{(\mathbf{X}_{T_{n-1}}, \mathbf{Y}_{T_{n-1}})}$ and so $\bar{P}_{(\mathbf{X}_{T_{n-1}}, \mathbf{Y}_{T_{n-1}})}$ also governs the evolution of the original process $(\mathbf{X}_{t-}, \mathbf{Y}_t)$ for $T_{n-1} < t \le T_n$.

Define the stopping time $\bar{T}_1 := \inf\{t \ge 0 \,:\, \Delta\bar{\mathbf{Y}}_t \ne 0\}$ and denote by $\bar{\xi}_1 \in \{1, \cdots, m\}$ the identity of the first jump firm in the "bar-model". Note that $\bar{T}_1$ is a standard doubly stochastic random time. Hence we have

$$\bar{P}_{(\boldsymbol{x},\boldsymbol{y})}\Big(\bar{T}_1 > t \mid \mathcal{F}_\infty^{\bar{\mathbf{X}}}\Big) = \exp\Big\{ -\int_0^t \bar{\lambda}(\bar{\mathbf{X}}_s, \boldsymbol{y})ds\Big\} \text{ and} \tag{3.7}$$

$$h_{\bar{T}_1,\bar{\xi}_1|\mathcal{F}_\infty^{\bar{\mathbf{X}}}}(t, i \mid \bar{\mathbf{X}}) = \lambda_i(\bar{\mathbf{X}}_t, \boldsymbol{y}) \exp\Big\{ -\int_0^t \bar{\lambda}(\bar{\mathbf{X}}_s, \boldsymbol{y})ds\Big\}\,, \tag{3.8}$$

where $h_{\bar{T}_1,\bar{\xi}_1|\mathcal{F}_\infty^{\bar{\mathbf{X}}}}$ is the conditional density of $\bar{T}_1, \bar{\xi}_1$ under $\bar{P}_{(\boldsymbol{x},\boldsymbol{y})}$, given $\mathcal{F}_\infty^{\bar{\mathbf{X}}}$ (see e.g. Section 9.6.2 of McNeil, Frey & Embrechts (2005)). Properties (3.7) and (3.8) will be essential for the derivation of the filter equations.

3.3 Filtering between default times

Overview. Our filter formulas take the form of a recursion over the ordered default times $0 = T_0 < T_1 < \cdots < T_m$. Denote the distribution of $\mathbf{X}_0$ given $\mathcal{F}_0^I$ (the initial filter distribution) by $\pi_0(d\boldsymbol{x})$ and for $1 \le n \le m$ by $\pi_{T_n}(d\boldsymbol{x})$ the filter distribution at $t = T_n$. In Theorem 3.1 we consider a time point $t \in [T_{n-1}, T_n)$ and show how $\pi_t f$ can be derived from $\pi_{T_{n-1}}(d\boldsymbol{x})$ and from the trajectory $(\mathbf{Z}_s)_{T_{n-1} \le s \le t}$, representing the new price information received over $[T_{n-1}, t]$. In Subsection 3.4 below we explain how to compute $\pi_{T_n} f$ from $\pi_{T_{n-1}}(d\boldsymbol{x})$, the new price information received over $[T_{n-1}, T_n]$ and the new default information (T_n, ξ_n). Since only the new price information $(\mathbf{Z}_s)_{s \ge T_{n-1}}$ matters for these considerations, in the sequel we use the lighter notation

$$\mathbf{Z}_s^n := \mathbf{Z}_{s+T_{n-1}} \quad \text{and} \quad \boldsymbol{a}_s^n(\cdot) := \boldsymbol{a}_{s+T_{n-1}}(\cdot)\,. \tag{3.9}$$

Theorem 3.1 *Consider a time point $t \in [T_{n-1}, T_n)$ for two successive default times T_{n-1} and T_n. Then $\pi_t f$ is proportional to*

$$\int_{\mathbb{R}^d} \pi_{T_{n-1}}(d\boldsymbol{x})\, \bar{E}_{(\boldsymbol{x}, \mathbf{Y}_{T_{n-1}})}\Big(f\left(\bar{\mathbf{X}}_{t-T_{n-1}}\right) L_{t-T_{n-1}}^n \exp\Big\{ - \int_0^{t-T_{n-1}} \bar{\lambda}(\bar{\mathbf{X}}_s, \mathbf{Y}_{T_{n-1}}) ds \Big\}\Big), \tag{3.10}$$

where the process $L^n = (L_u^n)_{u \ge 0}$ is defined by

$$L_u^n = \exp\Big\{ \int_0^u (\boldsymbol{a}_s^n)'(\bar{\mathbf{X}}_s) \Sigma_{\mathbf{Z}}^{-1}\, d\mathbf{Z}_s^n - \frac{1}{2} \int_0^u \big\| \boldsymbol{a}_s^n(\bar{\mathbf{X}}_s) \big\|_{\Sigma_{\mathbf{Z}}^{-1}}^2\, ds \Big\}\,. \tag{3.11}$$

The proportionality factor is given by (3.10) evaluated at $f(\boldsymbol{x}) \equiv 1$.

Proof Recall that by (3.2) and (3.3), $\pi_t f \propto E\big(f(\mathbf{X}_t) L_t \mid \mathcal{H}_t\big)$. Denote by $\mathcal{F}_t^{T_n} = \sigma\big(1_{\{T_n \le s\}} \colon s \le t\big)$ the filtration generated by the indicator of the random time T_n. Now note that for $T_{n-1} \le t < T_n$, $\mathcal{H}_t = \mathcal{H}_{T_{n-1}} \vee \mathcal{F}_t^{T_n}$. By the so called Dellacherie-formula (see for instance Lemma 3.1 in Elliott, Jeanblanc & Yor (2000)) we get for any integrable, $\mathcal{F}_\infty$-measurable random variable U that

$$E\big(U 1_{\{T_n > t\}} \mid \mathcal{H}_t\big) = 1_{\{T_n > t\}} \frac{E\big(U 1_{\{T_n > t\}} \mid \mathcal{H}_{T_{n-1}}\big)}{P\big(T_n > t \mid \mathcal{H}_{T_{n-1}}\big)}\,. \tag{3.12}$$

With $U = f(\mathbf{X}_t) L_t$ we therefore obtain for $t \in [T_{n-1}, T_n)$

$$E(f(\mathbf{X}_t) L_t \mid \mathcal{H}_t) \propto E(f(\mathbf{X}_t) L_t 1_{\{T_n > t\}} \mid \mathcal{H}_{T_{n-1}})\,. \tag{3.13}$$

By double conditioning we get

$$\begin{aligned} &E\big(f(\mathbf{X}_t) L_t 1_{\{T_n > t\}} \mid \mathcal{H}_{T_{n-1}}\big) \\ &\quad = E\Big(L_{T_{n-1}} E\Big(f(\mathbf{X}_t) \frac{L_t}{L_{T_{n-1}}}\, 1_{\{T_n > t\}} \mid \mathcal{F}_{T_{n-1}} \Big) \mid \mathcal{H}_{T_{n-1}} \Big)\,. \end{aligned} \tag{3.14}$$

Recall the definition of the process L^n from (3.11). By the Markov property of $(\mathbf{X}, \mathbf{Y})$ and the equality in law discussed in the previous subsection below equation (3.6), the inner conditional expectation in (3.14) equals

$$\bar{E}_{(\mathbf{X}_{T_{n-1}}, \mathbf{Y}_{T_{n-1}})}\Big(f(\bar{\mathbf{X}}_{t-T_{n-1}}) L^n_{t-T_{n-1}} \, 1_{\{\bar{T}_1 > t-T_{n-1}\}}\Big)\,. \tag{3.15}$$

Using the survival function of $\bar{T}_1$ as given in (3.7) and double conditioning on $\mathcal{F}^{\bar{\mathbf{X}}}_\infty$, (3.15) is equal to

$$\bar{E}_{(\mathbf{X}_{T_{n-1}}, \mathbf{Y}_{T_{n-1}})}\left(f(\bar{\mathbf{X}}_{t-T_{n-1}})\, L^n_{t-T_{n-1}} \exp\Big\{-\int_0^{t-T_{n-1}} \bar{\lambda}(\bar{\mathbf{X}}_s, \mathbf{Y}_{T_{n-1}})\, ds\Big\}\right).$$

Using the Kallianpur Striebel formula (3.2) and relation (3.3), expression (3.14) is thus proportional to

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

proving the theorem.

3.4 Filtering at a default time T_n

Again by (3.2) and (3.3), at a generic default time T_n we have

$$\pi_{T_n} f \propto E(f(\mathbf{X}_{T_n}) L_{T_n} \mid \mathcal{H}_{T_n})\,.$$

Notice now that, due to the possibility of common jumps between $\mathbf{X}$ and $\mathbf{Y}$, the expressions $E(f(\mathbf{X}_{T_n}) L_{T_n} \mid \mathcal{H}_{T_n})$ and $E(f(\mathbf{X}_{T_n-}) L_{T_n} \mid \mathcal{H}_{T_n})$ do not necessarily coincide. We shall therefore proceed along two steps. In Step 1 - which is specific to the case of common jumps of $\mathbf{X}$ and $\mathbf{Y}$ - we show that one can obtain the conditional expectation $E(f(\mathbf{X}_{T_n}) L_{T_n} \mid \mathcal{H}_{T_n})$ once one is able to compute $E(g(\mathbf{X}_{T_n-}) L_{T_n} \mid \mathcal{H}_{T_n})$ 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 $\mathbf{X}_{T_n-}$)

Proposition 3.2 *We have the relation*

$$E(f(\mathbf{X}_{T_n}) L_{T_n} \mid \mathcal{H}_{T_n}) = E(g(\mathbf{X}_{T_n-}; \mathbf{Y}_{T_{n-1}}, \xi_n) L_{T_n} \mid \mathcal{H}_{T_n}).$$

Here the function g is given by

$$g(\boldsymbol{x}; \boldsymbol{y}, j) = \begin{cases} \int_{D^{\mathbf{Y}}_j(\boldsymbol{x}, \boldsymbol{y})} f\big(\boldsymbol{x} + K^{\mathbf{X}}(\boldsymbol{x}, \boldsymbol{y}, u)\big)\, \nu_j(du) & ,\ \textit{for } F_{\mathcal{N}}\big(D^{\mathbf{Y}}_j(\boldsymbol{x}, \boldsymbol{y})\big) > 0 \\ f(x) & ,\ \textit{else}; \end{cases} \tag{3.16}$$

the measure $\nu_j(d\boldsymbol{x})$ *is defined by* $\nu_j(A) = F_{\mathcal{N}}\big(D_j^{\boldsymbol{y}}(\boldsymbol{x},\boldsymbol{y})\big)^{-1} F_{\mathcal{N}}(A \cap D_j^{\boldsymbol{y}}(\boldsymbol{x},\boldsymbol{y})$, *provided that the denominator is strictly positive.*

Remark 3.3 Note that without common jumps of $\mathbf{X}$ and $\mathbf{Y}$, i.e. in case that $F_{\mathcal{N}}\big(D_j^{\mathbf{Y}}(\boldsymbol{x},\boldsymbol{y}) \cap D_i^{\mathbf{X}}(\boldsymbol{x},\boldsymbol{y})\big) = 0$ for all $i, j, \boldsymbol{x}, \boldsymbol{y}$, we have that $K^{\mathbf{X}}(\boldsymbol{x},\boldsymbol{y},u)$ is zero on $D_j^{\mathbf{Y}}(\boldsymbol{x},\boldsymbol{y})$ $F_{\mathcal{N}}-$a.s., so that

$$\int\limits_{D_j^{\mathbf{Y}}(\boldsymbol{x},\boldsymbol{y})} f\big(\boldsymbol{x} + K^{\mathbf{X}}(\boldsymbol{x},\boldsymbol{y},u)\big)\nu_j(du) = f(\boldsymbol{x})\,.$$

Consequently $g(\boldsymbol{x},\boldsymbol{y},j) \equiv f(\boldsymbol{x})$ in that case, even for $F_{\mathcal{N}}\big(D_j^{\mathbf{Y}}(\boldsymbol{x},\boldsymbol{y})\big) > 0$, and so Step 1 becomes superfluous.

Proof By (2.1), (2.2) we have for $t \in (T_{n-1}, T_n]$ and for j with $Y_{T_{n-1},j} = 0$ the equality

$$\big\{(T_n, \xi_n) = (t, j)\big\} = \big\{\mathcal{N}\left(\{t\} \times D_j^{\mathbf{Y}}(\mathbf{X}_{t-}, \mathbf{Y}_{T_{n-1}})\right) = 1\big\}\ ,$$

or, in words, $(T_n, \xi_n) = (t, j)$ if and only if $\mathcal{N}(ds, du)$ has an atom in the set $\{t\} \times D_j^{\mathbf{Y}}(\mathbf{X}_{t-}, \mathbf{Y}_{T_{n-1}})$. Hence

$$\mathcal{H}_{T_n} \subset \widetilde{\mathcal{F}}_{T_n-} := \mathcal{F}_{T_n-} \vee \sigma\Big(1_{\big\{\mathcal{N}\big(\{T_n\}\times D_{\xi_n}^{\mathbf{Y}}(\mathbf{X}_{T_n-}, \mathbf{Y}_{T_{n-1}})\big)=1\big\}}\Big)\,. \tag{3.17}$$

Since moreover $\mathbf{X}_{T_n} = \mathbf{X}_{T_n-} + \int_E K^{\mathbf{X}}(\mathbf{X}_{T_n-}, \mathbf{Y}_{T_{n-1}}, u)\mathcal{N}(\{T_n\}, du)$, we get by double conditioning, using (3.17),

$$E(f(\mathbf{X}_{T_n})L_{T_n} \mid \mathcal{H}_{T_n}) = E\Bigg(E\Bigg(L_{T_n} \\ \times \int\limits_E f\big(\mathbf{X}_{T_n-} + K^{\mathbf{X}}(\mathbf{X}_{T_n-}, \mathbf{Y}_{T_{n-1}}, u)\big)\,\mathcal{N}(\{T_n\}, du) \mid \widetilde{\mathcal{F}}_{T_n-}\Bigg) \mid \mathcal{H}_{T_n}\Bigg).$$

Now note that L_{T_n} and $\mathbf{X}_{T_n-}$ are $\mathcal{F}_{T_n-}$ measurable whereas $\mathcal{N}(\{T_n\}, du)$ is independent of $\mathcal{F}_{T_n-}$ with compensator measure $F_{\mathcal{N}}(du)$. Moreover, given $\mathcal{F}_{T_n-}$, conditioning on $\widetilde{\mathcal{F}}_{T_n-}$ is equivalent to conditioning on the fact that $\mathcal{N}(\{T_n\}, du)$ has an atom in the set $\{T_n\} \times D_{\xi_n}^{\mathbf{Y}}\big(\mathbf{X}_{T_n-}, \mathbf{Y}_{T_{n-1}}\big)$. Hence the inner conditional expectation is equal to $g(\mathbf{X}_{T_n-}; Y_{T_{n-1}}, \xi_n)L_{T_n}$, and the result follows.

Step 2 (Updating of the conditional distribution of $\mathbf{X}_{T_n-}$)

Theorem 3.4 *Recall the definition of the process L^n from* (3.11)*. 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*

$$E\big(g(\mathbf{X}_{T_n-})L_{T_n} \mid \mathcal{H}_{T_n}\big) \propto \int_{\mathbb{R}^d} \pi_{T_{n-1}}(d\boldsymbol{x})\, \bar{E}_{(\boldsymbol{x},\mathbf{Y}_{T_{n-1}})}\Big(g(\bar{\mathbf{X}}_{T_n-T_{n-1}}) \times L^n_{T_n-T_{n-1}}\, \lambda_{\xi_n}\big(\bar{\mathbf{X}}_{T_n-T_{n-1}}, \mathbf{Y}_{T_{n-1}}\big) \exp\Big\{-\int_0^{T_n-T_{n-1}} \bar{\lambda}(\bar{\mathbf{X}}_s, \mathbf{Y}_{T_{n-1}})\, ds\Big\}\Big). \tag{3.18}$$

Proof The proof goes in two steps: first we reduce the claim to a statement with respect to $\bar{P}_{(\boldsymbol{x},\mathbf{Y}_{T_{n-1}})}$; secondly we apply Bayesian updating. Since $\mathcal{H}_{T_n} = \mathcal{H}_{T_{n-1}} \vee \sigma(T_n, \xi_n)$ we get

$$\begin{aligned} &E\big(g(\mathbf{X}_{T_n-})L_{T_n} \mid \mathcal{H}_{T_n}\big) \\ &\quad = E\Big(L_{T_{n-1}} E\Big(g(\mathbf{X}_{T_n-})\tfrac{L_{T_n}}{L_{T_{n-1}}} \mid \mathcal{F}_{T_{n-1}}, T_n, \xi_n\Big) \mid \mathcal{H}_{T_{n-1}}, T_n, \xi_n\Big)\,. \end{aligned} \tag{3.19}$$

We now concentrate on the inner expectation. Using (3.6) we get that

$$\begin{aligned} &E\Big(g(\mathbf{X}_{T_n-})\tfrac{L_{T_n}}{L_{T_{n-1}}} \mid \mathcal{F}_{T_{n-1}}, T_n = t, \xi_n = i\Big) = E\Big(g(\widetilde{\mathbf{X}}_{T_n-}) \\ &\times \exp\Big\{\int_{T_{n-1}}^{T_n} \boldsymbol{a}'_s(\widetilde{\mathbf{X}}_s)\Sigma_{\mathbf{Z}}^{-1}\, d\mathbf{Z}_s - \frac{1}{2}\int_{T_{n-1}}^{T_n} \Big\|\boldsymbol{a}_s(\widetilde{\mathbf{X}}_s)\Big\|^2_{\Sigma_{\mathbf{Z}}^{-1}}\, ds\Big\} \Big| \mathcal{F}_{T_{n-1}}, T_n = t, \xi_n = i\Big). \end{aligned}$$

Now note that, due to the equality in law as discussed after relation (3.6), given $\mathcal{F}_{T_{n-1}}$, the joint law of $\big((\mathbf{X}_s)_{T_{n-1}\le s<T_n}, T_n - T_{n-1}, \xi_n\big)$ equals the law of $\big((\bar{\mathbf{X}}_s)_{0\le s<\bar{T}_1}, \bar{T}_1, \bar{\xi}_1\big)$ under $\bar{P}_{(\mathbf{X}_{T_{n-1}},\mathbf{Y}_{T_{n-1}})}$. Moreover, $\Delta\bar{\mathbf{X}}_{\bar{T}_1} = 0$ a.s. Hence the last term equals

$$\begin{aligned} &\bar{E}_{(\mathbf{X}_{T_{n-1}},\mathbf{Y}_{T_{n-1}})}\Big(g(\bar{\mathbf{X}}_{\bar{T}_1}) \exp\Big\{\int_0^{\bar{T}_1} (\boldsymbol{a}^n_s)'(\bar{\mathbf{X}}_s)\Sigma_{\mathbf{Z}}^{-1}\, d\mathbf{Z}^n_s \\ &\quad - \frac{1}{2}\int_0^{\bar{T}_1} \big\|\boldsymbol{a}^n_s(\bar{\mathbf{X}}_s)\big\|^2_{\Sigma_{\mathbf{Z}}^{-1}}\, ds\Big\} \Big| \bar{T}_1 = t - T_{n-1}, \bar{\xi}_1 = i\Big)\,. \end{aligned} \tag{3.20}$$

We are now in a position to do the Bayesian updating. Recall from (3.8) the form of the conditional density $h_{\bar{T}_1,\bar{\xi}_1|\mathcal{F}^{\bar{\mathbf{X}}}_\infty}$. By double conditioning on $\mathcal{F}^{\bar{\mathbf{X}}}_\infty \vee \sigma\big(\bar{T}_1, \bar{\xi}_1\big)$ and the Bayes formula, (3.20) is easily seen to be proportional to

$$\begin{aligned} &\bar{E}_{(\mathbf{X}_{T_{n-1}},\mathbf{Y}_{T_{n-1}})}\Big(g(\bar{\mathbf{X}}_{t-T_{n-1}})\, L^n_{t-T_{n-1}} \\ &\quad \times \lambda_i(\bar{\mathbf{X}}_{t-T_{n-1}}, \mathbf{Y}_{T_{n-1}}) \exp\Big\{-\int_0^{t-T_{n-1}} \bar{\lambda}(\bar{\mathbf{X}}_s, \mathbf{Y}_{T_{n-1}})ds\Big\}\Big)\,. \end{aligned}$$

Combining this with (3.19), thereby using (3.2) and (3.3), gives the result.

4 Filter Computation

In this section we discuss two approaches for turning the filter equations from the previous section into a computable filtering algorithm. In Section 4.1 we derive a finite dimensional filtering algorithm (Algorithm 4.3) for the case where the pair process $(\mathbf{X}, \mathbf{Y})$ follows a finite state Markov chain. Models of this type are frequently being used in portfolio credit risk modelling (with observable $\mathbf{X}$); examples include the infectious defaults model discussed in Section 2.3 or the Markov-chain models of Arnsdorf & Halperin (2007) and of Frey & Backhaus (2007). Moreover, the results for the finite state Markov case can be used to construct a filter approximation for general jump-diffusion models, as will be shown in Section 5 below. While Markov chain approximations are very useful for lower dimensional state processes, computations become prohibitively expensive as soon as the dimension of $\mathbf{X}$ becomes moderately large. In Subsection 4.2 we therefore explain how the filter equations from Section 3 can be used to construct a particle filtering algorithm for the jump-diffusion model (2.1) (2.2). The results of numerical experiments are reported in Subsection 4.3

4.1 Filter equations for finite-state Markov chains

A general Markov chain model. Assume that the pair process $(\mathbf{X}, \mathbf{Y})$ follows a finite-state Markov chain. W.l.o.g. we assume that the state space of $(\mathbf{X}, \mathbf{Y})$ is given by the set $\{1, \ldots, K\} \times \{0,1\}^m$; in particular, X is now scalar. We denote the transition intensities of (X, 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 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 $\lambda_i(k, \boldsymbol{y}) = \sum_{h=1}^{K} q(k, \boldsymbol{y}; h, \boldsymbol{y}^i)$.

In order to apply the results from the previous section we have to model the dynamics of $(X, \mathbf{Y})$ in a way that fits into the general framework (2.1), (2.2). Since there are three types of transitions we take $E = \mathbb{R} \times \{1, 2, 3\}$ and, with $u = (u_1, u_2)$, $F_{\mathcal{N}}(du) = \nu(du_1) \otimes \sum_{j=1}^{3} \delta_j(u_2)$ with $\nu(\cdot)$ the Lebesgue and

$\delta_j(\cdot)$ the Dirac measures. We put

$$K^X(k,\boldsymbol{y},u) := \sum_{h\neq k}(h-k)1_{\big(\sum_{\{\ell<h:\ \ell\neq k\}}\bar q^y_{k,\ell}\,,\,\sum_{\{\ell<h+1:\ell\neq k\}}\bar q^y_{k,\ell}\big]}(u_1)\,1_{\{1\}}(u_2) \tag{4.1}$$
$$+\sum_{\{i:y_i=0\}}\sum_{h\neq k}(h-k)1_{\big(\sum_{\{\ell<h:\ell\neq k\}}q(k,\boldsymbol{y};\ell,\boldsymbol{y}^i)\,,\,\sum_{\{\ell<h+1:\ \ell\neq k\}}q(k,\boldsymbol{y};\ell,\boldsymbol{y}^i)\big]}(u_1)\,1_{\{2\}}(u_2)$$

and for $1\le i\le m$

$$K_i^{\mathbf{Y}}(k,\boldsymbol{y},u) := 1_{\{y_i=0\}}\Big(1_{\big(0\,,\,\sum_{\{\ell\neq k\}}q(k,\boldsymbol{y};\ell,\boldsymbol{y}^i)\big]}(u_1)\,1_{\{2\}}(u_2)$$
$$+\,1_{\big(0\,,\,q(k,\boldsymbol{y},k,\boldsymbol{y}^i\big]}(u_1)1_{\{2\}}(u_3)\Big) \tag{4.2}$$

It follows that the the kernel $\bar K^{\mathbf{X}}(\cdot)$ defined in (3.4) is given by the right hand side of (4.1), as can also be seen from the fact that for $t\in[T_n,T_{n+1})$ the factor process X follows a finite state Markov chain with transition intensities $\bar q_{k,h}^{Y_{T_n}}$, $k\neq h$; we denote the corresponding generator matrix by $\bar Q^{\boldsymbol{y}}$.

The filter equations. In the finite-state Markov case the filter distribution can be summarized by the K-dimensional process $\boldsymbol{\pi}_t=(\pi_t^1,\dots,\pi_t^K)$ with $\pi_t^i := P(X_t=i\mid\mathcal{F}_t^I)$. Obviously, it suffices to compute an un-normalized version of $\boldsymbol{\pi}_t$. The key step in applying the filtering results of Section 3, in particular Theorems 3.1 and 3.4, is thus the evaluation of expressions of the form

$$\sigma_t g[n,\boldsymbol{y}] := \sum_{i=1}^K \pi_{T_{n-1}}(\{i\})\,\bar E_{(i,\boldsymbol{y})}\Big(g\left(\bar X_t\right)L_t^n\exp\Big\{-\int_0^t\bar\lambda(\bar X_s,\boldsymbol{y})\,ds\Big\}\Big) \tag{4.3}$$

for generic $g\colon\{1,\dots,K\}\to\mathbb{R}$, $\boldsymbol{y}\in\{0,1\}$. Put for $h\in\{1,\dots,K\}$

$$\sigma_t^h[n,\boldsymbol{y}] := \sum_{i=1}^K \bar\pi_{T_{n-1}}(\{i\})\,E_{(i,\boldsymbol{y})}\Big(1_{\{\bar{\mathbf{X}}_t=h\}}L_t^n\exp\Big\{-\int_0^t\bar\lambda(\bar X_s,\boldsymbol{y})\,ds\Big\}\Big), \tag{4.4}$$

so that $\sigma_t[n,\boldsymbol{y}]g=\sum_{h=1}^K\sigma_t^h[n,\boldsymbol{y}]g(h)$. In the next proposition we derive a Zakai-type SDE for the vector process $\boldsymbol{\sigma}_t[n,\boldsymbol{y}]=(\sigma_t^1[n,\boldsymbol{y}],\dots,\sigma_t^K[n,\boldsymbol{y}])$ that, as follows immediately from the previous development, represents a vector of unnormalized conditional probabilities.

Proposition 4.1 *The process $\boldsymbol{\sigma}_t=\boldsymbol{\sigma}_t[n,\boldsymbol{y}]$ solves the SDE*

$$d\sigma_t^i=\Big(\sum_{k=1}^K\bar q^{\boldsymbol{y}}_{k,i}\sigma_t^k-\bar\lambda(i,\boldsymbol{y})\sigma_t^i\Big)\,dt+\sigma_t^i\,(\boldsymbol{a}_t^n)'(i)\Sigma_{\mathbf{Z}}^{-1}\,d\mathbf{Z}_t^n\,,\quad 1\le i\le K\,, \tag{4.5}$$

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

Proof A similar reasoning as in Section 3.1 yields $\sigma_t^h \propto \bar{R}_\pi\left(\bar{\mathbf{X}}_t = h \mid \mathcal{F}_t^{\mathbf{Z}^n}\right)$ where under $\bar{R}_\pi$, $\bar{X}$ is a Markov chain with generator matrix $\bar{Q}^{\boldsymbol{y}}$, initial distribution π, and where $\mathbf{Z}_t^n$ solves the SDE $d\mathbf{Z}_t^n = \boldsymbol{a}_t^n(\bar{X}_t)dt + vd\boldsymbol{\beta}_t^n$, with $\boldsymbol{\beta}_t^n = \left(\boldsymbol{\beta}_{t+T_{n-1}}\right)_{t\geq 0}$.

The statement can now be derived from general results in Corollary 3.9 of Elliott (1993). For this purpose we need to define an appropriate process H that we choose as $H_t := \exp\left\{-\int_0^t \bar{\lambda}(X_s, \boldsymbol{y})ds\right\}$, so that $dH_t = -\bar{\lambda}(X_t, \boldsymbol{y})H_t dt$. With this choice of H the coefficients in (3.1) of Elliott (1993) become $\beta \equiv \delta \equiv 0$ and $\alpha_t = -\bar{\lambda}(X_t, \boldsymbol{y})H_t$. Equation (4.5) now corresponds to relation (3.18) in Elliott (1993).

The filter distribution at a default time T_n. Here we have the following result.

Corollary 4.2 *Using the convention* $0/0 = 0$, *we have for* $1 \leq i \leq K$

$$\begin{aligned}\pi_{T_n}^i := \pi_{T_n}(\{i\}) = \sum_{h\neq i} P\left(X_{T_n-} = h \mid \mathcal{F}_{T_n}^I\right) \frac{q\big(h, \mathbf{Y}_{T_{n-1}}; i, \mathbf{Y}_{T_n}\big)}{\sum_{j=1}^K q\big(h, \mathbf{Y}_{T_{n-1}}; j, \mathbf{Y}_{T_n}\big)} \\ + P\left(X_{T_n-} = i \mid \mathcal{F}_{T_n}^I\right) \frac{q\big(i, \mathbf{Y}_{T_{n-1}}; i, \mathbf{Y}_{T_n}\big)}{\sum_{j=1}^K q\big(i, \mathbf{Y}_{T_{n-1}}; j, \mathbf{Y}_{T_n}\big)}.\end{aligned} \tag{4.6}$$

Proof The result can be established by applying Proposition 3.2 to the kernels K^X and $K_i^{\mathbf{Y}}$ introduced previously. Alternatively, an analogous reasoning as in the proof of Proposition 3.2 can be used to show that the expressions

$$\frac{q\big(h, \mathbf{Y}_{T_{n-1}}; i, \mathbf{Y}_{T_n}\big)}{\sum_{j=1}^K q\big(h, \mathbf{Y}_{T_{n-1}}; j, \mathbf{Y}_{T_n}\big)} \quad \text{and} \quad \frac{q\big(i, \mathbf{Y}_{T_{n-1}}; i, \mathbf{Y}_{T_n}\big)}{\sum_{j=1}^K q\big(i, \mathbf{Y}_{T_{n-1}}; j, \mathbf{Y}_{T_n}\big)}$$

give the conditional probability that X jumps from state h to state i at T_n respectively that X stays in state i at T_n, given the observed transition of $\mathbf{Y}$.

We summarize our filtering results for the finite-state Markov case in the following algorithm. Recall that $\left(\boldsymbol{\sigma}_s[n, \boldsymbol{y}]\right)_{s\geq 0}$ denotes the solution of (4.5) with initial value $\boldsymbol{\sigma}_0^i[n, \boldsymbol{y}] = \pi_{T_{n-1}}(\{i\})$, $\mathbf{Z}_s^n = \mathbf{Z}_{s+T_{n-1}}$ and $\boldsymbol{a}_s^n(\cdot) = \boldsymbol{a}_{s+T_{n-1}}(\cdot)$.

Algorithm 4.3 (Filtering for a finite-state Markov chain.)

1. Set $n = 1$, $T_0 = 0$, $\mathbf{Y}_0 = 0$, and denote the initial filter distribution by π_0.
2. Compute $\sigma_t^i[n, \mathbf{Y}_{T_{n-1}}]$ for $t \geq T_{n-1}$ according to (4.5) (for the actual computation see the next paragraph).
3. For $t \in [T_{n-1}, T_n)$, set (see Theorem 3.1) $\pi_t^i := \dfrac{\sigma_{t-T_{n-1}}^i[n, \mathbf{Y}_{T_{n-1}}]}{\sum_{k=1}^K \sigma_{t-T_{n-1}}^k[n, \mathbf{Y}_{T_{n-1}}]}$

4. For $t = T_n$ compute first (see Theorem 3.4)

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

and determine then $\pi_{T_n} = \left[\pi^1_{T_n}, \ldots, \pi^K_{T_n}\right]$ according to (4.6). Replace n by $n+1$ and return to Step 2.

Note that all quantities appearing in Algorithm 4.3 can be expressed in term of the transition intensities of the Markov chain $(X, \mathbf{Y})$.

Solving the Zakai-equation (4.5). In order to apply Algorithm 4.3, we need to solve the SDE (4.5). Here two cases can be distinguished. If $(\mathcal{F}^I_t) = (\mathcal{H}_t)$ or, equivalently, $\boldsymbol{a}^n_t(\cdot) \equiv 0$ equation (4.5) reduces to the ODE-system

$$\frac{d}{dt}\sigma^i_t = \sum_{k=1}^K \bar{q}^{\boldsymbol{y}}_{k,i}\sigma^k_t - \bar{\lambda}(i, \boldsymbol{y})\sigma^i_t\,, \quad 1 \le i \le K. \tag{4.7}$$

In vector notation the solution of this equation is given by the matrix exponential $\boldsymbol{\sigma}_t[n, \boldsymbol{y}] = \boldsymbol{\sigma}_0[n, \boldsymbol{y}] \exp\left\{t\left((\bar{Q}^{\boldsymbol{y}})' - \operatorname{diag}(\bar{\lambda}(1, \boldsymbol{y}) \cdots, \bar{\lambda}(K, \boldsymbol{y}))\right)\right\}$. This matrix exponential can be computed by diagonalizing the matrix $\bar{Q}^{\boldsymbol{y}}$. Alternatively, one can apply numerical schemes for ODEs to equation (4.7).

If $\boldsymbol{a}^n_t(\cdot) \neq 0$, (4.5) is a stochastic differential equation. Numerical methods for solving this equation are based on time discretization, e.g. according to the Euler-Maruyama scheme. This is the natural approach if only discrete time observations of $\mathbf{Z}$ are available; for a discussion of technical details we refer to Clark (1978). An alternative approach, again due to Clark (1978), is to reduce the stochastic differential equation (4.5) to a deterministic one via a well chosen factorization. To this effect notice that a straightforward application of Itô's formula allows to show the following

Lemma 4.4 *We have* $\sigma^i_t := A_{t,i} \cdot r_i(t, \mathbf{A})$ *where*

$$A_{t,i} := \exp\left\{\int_0^t (\boldsymbol{a}^n_s)'(i)\Sigma^{-1}_{\mathbf{Z}}\, d\mathbf{Z}^n{}_s - \frac{1}{2}\int_0^t \|\boldsymbol{a}^n_s(i)\|^2_{\Sigma^{-1}_{\mathbf{Z}}}\, ds\right\},\ i = 1, \ldots, K \tag{4.8}$$

and where $r_i(\cdot, \mathbf{A})$ *solves the ODE-system*

$$\frac{d}{dt}r_i(t, \mathbf{A}) = A^{-1}_{t,i}\sum_{k=1}^K A_{t,k}\left(\bar{q}^{\boldsymbol{y}}_{k,i} - \bar{\lambda}(i, \boldsymbol{y})\delta_{ik}\right)r_k(t, \mathbf{A})\,, \ \ \boldsymbol{r}(0, \mathbf{A}) = \boldsymbol{\sigma}_0\,. \tag{4.9}$$

In order to use this result one has to determine a trajectory of $\mathbf{A} = (A_{t,1}, \cdots, A_{t,K})_{|t \ge 0}$ given the observed trajectory of the process $\mathbf{Z}^n_t$. This can be accomplished by using stochastic partial integration to allow for a pathwise evaluation of the stochastic integral in (4.8); we refer to Davis (1978) for details and for a deeper discussion of pathwise non-linear filtering.

In particle filtering the conditional distribution $\pi_{\mathbf{X}_t|\mathcal{F}_t^I}$ is approximated by the occupation measure $\tilde{\pi}_t$ of a branching particle system with particles in the state space $S^{\mathbf{X}}$ of $\mathbf{X}$. This branching system is constructed by a recursion over discrete time steps $t_k = k\Delta$, $k = 0, 1, \ldots$. The measure $\tilde{\pi}_{t_{k+1}}$ is constructed from $\tilde{\pi}_{t_k}$ in a two-stage procedure. In the *prediction step* one generates for each particle $\boldsymbol{x}_k^i$ in the system at time t_k a trajectory of the SDE (3.5) of length Δ with initial value $\boldsymbol{x}_k^i$. In the *updating step*, the new particle system is constructed by letting each particle branch into a random number of offsprings; the mean number of offsprings is defined in accordance with Theorems 3.1 and 3.4 (see Step 3 and 4 of Algorithm 4.5 below). Moreover, at a default time the particles are shifted by a random amount according to the conditional jump-size distribution of $\mathbf{X}$ at a default time as given in Proposition 3.2.

Let $\beta(t_k)$ denote the number of particles at time t_k, and consider (in accordance with (3.5)) for $(\boldsymbol{x}, \boldsymbol{y}) \in S^{\mathbf{X}} \times \{0,1\}^m$ the SDE

$$\bar{\mathbf{X}}_t = \boldsymbol{x} + \int_0^t b(\bar{\mathbf{X}}_s, \boldsymbol{y}) ds + \int_0^t \sigma(\bar{\mathbf{X}}_s, \boldsymbol{y}) d\mathbf{W}_s + \int_0^t \int_E \bar{K}^{\mathbf{X}}(\bar{\mathbf{X}}_{s-}, \boldsymbol{y}, u) \mathcal{N}(ds, du), \tag{4.10}$$

The evolution of the particle system can then be described as follows:

Algorithm 4.5 (Particle filtering.)

1. The initial state $\tilde{\pi}_0$ is given by the occupation measure of $\beta(0)$ particles of mass $1/\beta(0)$, i.e. $\tilde{\pi}_0 = \beta(0)^{-1} \sum_{i=1}^{\beta(0)} \delta_{\boldsymbol{x}^i(0)}$; here $\{\boldsymbol{x}^1(0), \ldots, \boldsymbol{x}^{\beta(0)}(0)\}$ represent independent draws from the initial distribution π_0.
2. (Prediction step) Given the particles $\{\boldsymbol{x}_k^1, \ldots, \boldsymbol{x}_k^{\beta(t_k)}\}$ in the system at time t_k, generate for $i = 1, \ldots, \beta(t_k)$ independent trajectories $\bar{\mathbf{X}}^i = (\bar{\mathbf{X}}^i(s))_{0 \le s \le \Delta}$ of the SDE (4.10) with starting value $\boldsymbol{x} = \boldsymbol{x}_k^i$ and $\boldsymbol{y} = \mathbf{Y}_{t_k}$.
3. (Updating step/no defaults) Assume that there is no default in $(t_k, t_{k+1}]$. Given the new noisy price observation $(\mathbf{Z}_t)_{t_k \le s \le t_{k+1}}$, define (in accordance with Theorem 3.1) for each trajectory $\bar{\mathbf{X}}^i$, $1 \le i \le \beta(t_k)$, the weights

$$\begin{aligned} L^i := \exp\Big\{ &- \int_0^\Delta \bar{\lambda}(\bar{\mathbf{X}}_s^i, \mathbf{Y}_{t_k})\, ds \\ &+ \int_0^\Delta \boldsymbol{a}'_{t_k+s}(\bar{\mathbf{X}}_s^i) \Sigma_{\mathbf{Z}}^{-1}\, d\mathbf{Z}_{t_k+s} - \frac{1}{2} \int_0^\Delta \left\| \boldsymbol{a}_{t_k+s}(\bar{\mathbf{X}}_s^i) \right\|^2_{\Sigma_{\mathbf{Z}}^{-1}}\, ds \Big\} \end{aligned} \tag{4.11}$$

   In a numerical implementation the stochastic integral in (4.11) may be computed by Euler approximation. Define

$$\mu^i := \frac{\beta(t_k) L^i}{\sum_{j=1}^{\beta(t_k)} L^j},$$

and denote by $[\mu^i]$ the integer part of μ^i. At t_{k+1} each particle $\boldsymbol{x}_k^i$ in the system at t_k produces independently a random number $m(i)$ of offsprings with mean number of offsprings equal to μ^i; in order to minimize the variance of $m(i)$ it is assumed that $m(i)$ has support $\{[\mu^i], [\mu^i]+1\}$. Note that together with the requirement $E(m(i)) = \mu_i$ this determines the distribution of $m(i)$. The positions of the $m(i)$ offsprings of particle i are given by $\bar{\mathbf{X}}^i(\Delta)$ (the endpoint of the trajectory with initial value $\boldsymbol{x}_k^i$). We set $\beta(t_{k+1}) := \sum_{i=1}^{\beta(t_k)} m(i)$ and denote the new particles at time t_{k+1} by $\{\boldsymbol{x}_{k+1}^1, \dots, \boldsymbol{x}_{k+1}^{\beta(t_{k+1})}\}$. The approximation to the filter distribution at time t_{k+1} is then given by

$$\tilde{\pi}_{t_{k+1}} = \beta(t_{k+1})^{-1} \sum_{i=1}^{\beta(t_{k+1})} \delta_{\boldsymbol{x}_{k+1}^i} . \tag{4.12}$$

4. (Updating step at a default time) If there is a default event in $(t_k, t_{k+1}]$, we use Theorem 3.4 and proceed as follows[2]. Denote by $\xi \in \{1, \dots, m\}$ the identity of the defaulting firm and put $\tilde{L}^i := \widehat{\lambda}_\xi\big(\bar{\mathbf{X}}^i(\Delta)\big)L^i$ with L^i as in (4.11). The number of offsprings $m(i)$ is determined by the same mechanism as in Step 3 but with $\tilde{L}^1, \dots, \tilde{L}^{\beta(t_k)}$ instead of $L^1, \dots, L^{\beta(t_k)}$. In accordance with Proposition 3.2 the position of the offsprings of particle i is given by $\bar{\mathbf{X}}^i(\Delta) + K^X(\bar{\mathbf{X}}^i(\Delta), \mathbf{Y}_{t_k}, U)$, where $U \sim \nu_\xi(du)$ with

$$\nu_\xi(A) := \frac{F_{\mathcal{N}}\big(A \cap D_\xi^{\mathbf{Y}}(\bar{\mathbf{X}}^i(\Delta), \mathbf{Y}_{t_k})\big)}{F_{\mathcal{N}}\big(D_\xi^{\mathbf{Y}}(\bar{\mathbf{X}}^i(\Delta), \mathbf{Y}_{t_k})\big)} .$$

The measure $\tilde{\pi}_{t_{k+1}}$ is then again given by (4.12).

This algorithm has a number of advantages, in particular for high-dimensional problems. First, particles with small weights (corresponding to a-posteriori unlikely trajectories of $\mathbf{X}$) have a low probability of being carried forward, so that the particles concentrate mostly in the more probable regions of the state space. Moreover, the computational effort increases only linearly in the dimensionality of the state process. Obviously, particle filtering algorithms based on resampling instead of branching (see e.g. Budhiraja et al. (2007)) can be adapted to our setting in an analogous manner.

4.3 Numerical experiments

Next we present results from numerical experiments illustrating the performance of the Markov-chain filter (Algorithm 4.3) and of the proposed particle filtering algorithm. We work in a scalar version of the affine jump-diffusion model of Duffie & Garleanu (2001), termed (scalar) DG-model, and in a special version of the Frey, & Schmidt (2009)-model, termed FS-model. The scalar

[2] Since the interval length is short we can neglect the possibility of more than one default per time step. Moreover, we may assume that the default happens exactly at $t = t_{k+1}$.

DG-model has a low-dimensional state space and is therefore a useful test case where Markov-chain filter and particle filter can be compared. The FS-model is more demanding: in this model the state space of the signal process is of higher dimension (we take $K = 4$ for the simulation study), and there are common jumps of state space and default indicator process. For this model we we relied exclusively on particle filtering.

For simplicity we assume that all firms have identical default intensities. In the DG-model the signal is given by the default intensity $\lambda_t = \lambda X_t$, $\lambda > 0$, of the firms under consideration; in the FS-model the signal corresponds to the solution $\boldsymbol{p} = (p_t^1, \dots, p_t^4)_{t \geq 0}$ of the SDE (2.19) representing the filter-distribution of the market. The continuous observation process Z is given by $dZ_t = \lambda X_t dt + v d\beta_t$ (for the DG model) and by $dZ_t = \sum_{k=1}^4 \lambda(k) p_t^k dt + v d\beta_t$ (for the FS-model) respectively; in both cases the parameter v models the size of the observation noise.[3] Note that the drift term in the observation represents the $(\mathcal{F}_t)$-default intensity of the firms under consideration; this term can be interpreted as credit spread of a short term bond or CDS-contract (assuming for simplicity a zero recovery rate). All simulations were carried out with $m = 100$ firms.

Results. Numerical results are displayed in Figures 4.1 and 4.2 below. Inspection of these graphs points to the following observations.

- In both models, for low observation noise (low σ_ϵ), the filtered default intensity is very close to the market default intensity (see left panel of Figure 4.1 respectively the bottom right panel of Figure 4.2). Interestingly, for low observation noise[4] the filtered probabilities $(E(\boldsymbol{p}_t \mid \mathcal{F}_t^I))$ are quite close to the market probability vector $\boldsymbol{p}_t$, as is revealed by the two left panels of Figure 4.2; this is quite remarkable, since the signal is a four-dimensional process whereas Z (the continuous part of the observation process) is one dimensional. For high σ_ϵ the filter performance is of course somewhat worse but still quite good. Overall we found that the information contained in Z has a stronger impact on the precision of the filter than the default history; this is not surprising since for realistic parameter values defaults are rare events.
- In the DG-model we can compare the Markov-chain filter and a particle filtering algorithm. Both algorithms give roughly similar results even for a coarse discretisation of the state space of X. While the Markov chain filter is significantly faster, for low observation noise the numerical integration of the Zakai equation tends to produce small numerical instabilities as can be seen from the spikes on the left plot of Figure 4.1.

[3] Recall from Subsection 2.2 that in order to model discrete observations occurring on a fine time grid $t_k = k\Delta$ one should take $v = \sigma_\epsilon \sqrt{\Delta}$, where σ_ϵ represents the standard deviation of the discrete observation noise.

[4] The value $\sigma_\epsilon = 0.5\%$ or $\sigma_\epsilon = 0.2\%$ may seem very small at first sight, but one should keep in mind that the default intensities are of the order of 1 or 2 %

Note finally that Figure 4.2 gives a nice numerical illustration of the phenomenon of information based default contagion: at a default the conditional probability of being in State 1 (where default intensities are assumed to be low) is decreased, whereas the conditional probability of being in States 3 and 4 (where default intensities are high) jumps upwards.

5 Filter Approximations

A viable approach to solve filtering problems for a general state variable $\mathbf{X}$ is to consider approximations of $\mathbf{X}$ by a sequence $\mathbf{X}^m$ of finite state Markov chains as a computational tool. In this section we provide the necessary theoretical basis for this and show the convergence of the filter computed for a finite state approximation $\mathbf{X}^m$ to the filter corresponding to the original state variable process $\mathbf{X}$. The proof of this result relies heavily on our general filter formulas.

An alternative representation of the filter. We begin with an alternative expression for the filter; while more abstract than the results from Section 4, this expression is well suited for deriving approximation results. Consider a pair of processes $(\mathbf{X}, \mathbf{Y})$ solving the original SDE system (2.1) (2.2); in particular, the dynamics of $\mathbf{X}$ are given in terms of the drift vector $b(\boldsymbol{x}, \boldsymbol{y})$, the dispersion matrix $\sigma(\boldsymbol{x}, \boldsymbol{y})$ or equivalently the matrix $\Sigma_{\mathbf{X}}(\boldsymbol{x}, \boldsymbol{y}) = \sigma(\boldsymbol{x}, \boldsymbol{y})\sigma(\boldsymbol{x}, \boldsymbol{y})'$ and the kernel $K^{\mathbf{X}}(\boldsymbol{x}, \boldsymbol{y}, u)$. Fix some $t > 0$. Put as before $N_t := \max\{n \leq m : T_n \leq t\}$ and recall that the sequence (T_n, ξ_n), $n = 1, \ldots, N_t$ or equivalently the process $(\mathbf{Y}_s)_{s \leq t}$ represents the default history up to time t. In the sequel we will always work with respect to given and fixed default observations $(\hat{T}_n, \hat{\xi}_n)$, $n = 0, \ldots, \hat{N}_t$ respectively $(\hat{\mathbf{Y}}_s)_{s \leq t}$; the "hat-notation" is meant to indicate that the default observations are fixed and can hence be considered deterministic.

Recall now the definition (3.4) of the kernel $\bar{K}^{\mathbf{X}}$. Denote by $(\Omega^d, \mathcal{F}^d, (\mathcal{F}_t^d))$ the Skorokhod space $D^d([0, \infty))$ with its canonical filtration and denote the coordinate process on Ω^d by $\bar{\mathbf{X}}$. For reasons that will become apparent in the sequel, we define on $(\Omega^d, \mathcal{F}^d, (\mathcal{F}_t^d))$ a predictable vector process by

$$B_t = \int_0^t b(\bar{\mathbf{X}}_s, \hat{\mathbf{Y}}_s) ds + \int_0^t \int_E \bar{K}^{\mathbf{X}}(\bar{\mathbf{X}}_{s-}, \hat{\mathbf{Y}}_{s-}, u) F_{\mathcal{N}}(du)\, ds \tag{5.1}$$

$$+ \sum_{n=1}^{\hat{N}_t} F_{\mathcal{N}}\big(D^{\mathbf{Y}}_{\hat{\xi}_n}(\bar{\mathbf{X}}_{\hat{T}_n -}, \hat{\mathbf{Y}}_{\hat{T}_{n-1}})\big)^{-1} \int_{D^{\mathbf{Y}}_{\hat{\xi}_n}(\bar{\mathbf{X}}_{\hat{T}_n -}, \hat{\mathbf{Y}}_{\hat{T}_{n-1}})} \bar{K}^{\mathbf{X}}(\bar{\mathbf{X}}_{\hat{T}_n -}, \hat{\mathbf{Y}}_{\hat{T}_{n-1}}, u)\, F_{\mathcal{N}}(du);$$

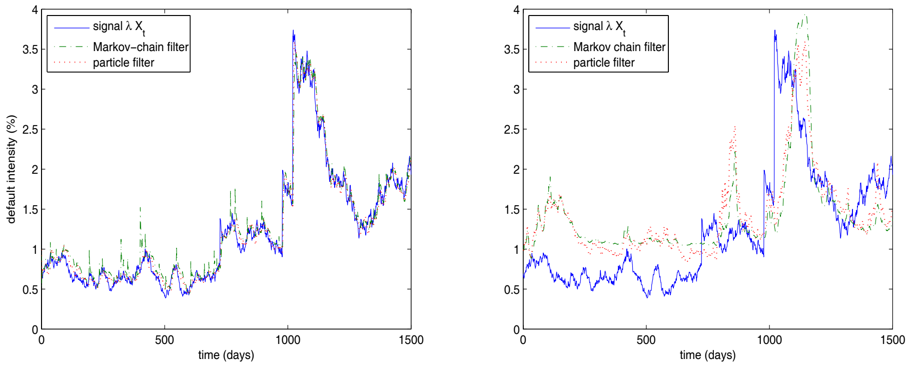

Fig. 4.1 Filtering results for the DG-model. Left: results for low observation noise ($v = 1/8000$ respectively $\sigma_\epsilon = 0.2\%$); right: results for high observation noise ($v = 1/320$ respectively $\sigma_\epsilon = 5\%$).

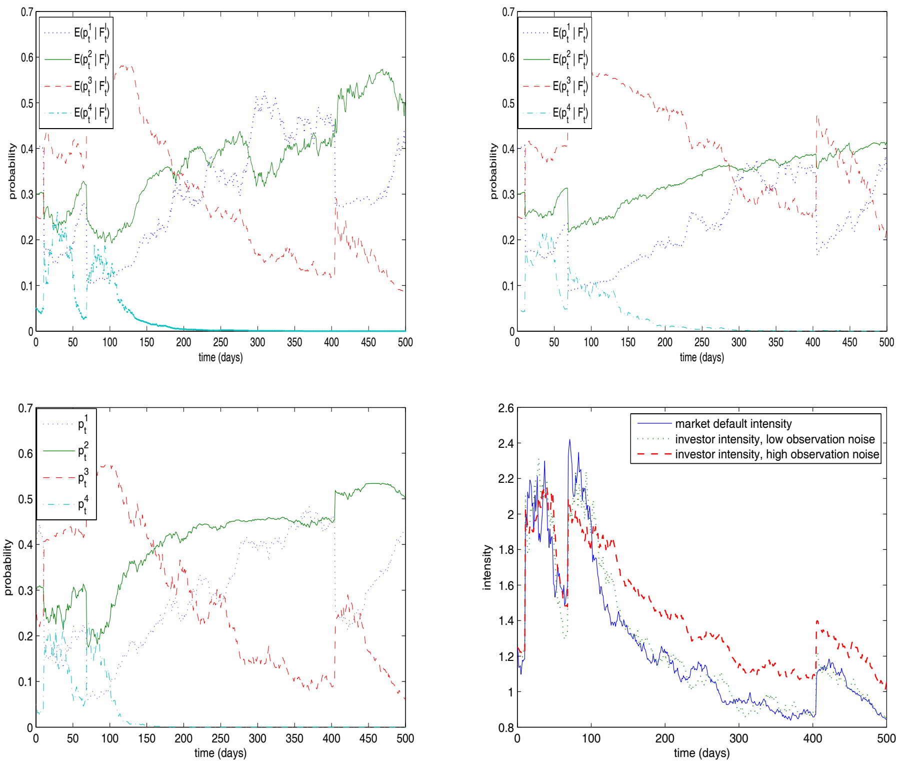

Fig. 4.2 Filtering results for the FS-model. Top left: trajectories of $E(p_t^k \mid \mathcal{F}_t^I)$ for low observation noise ($v = 1/3200$ resp. $\sigma_\epsilon = 0.5\%$); top right: trajectories of $E(p_t^k \mid \mathcal{F}_t^I)$ for high observation noise ($v = 1/400$ resp. $\sigma_\epsilon = 4\%$); bottom left: trajectories of signal $\boldsymbol{p}_t = (p_t^1, \ldots, p_t^4)$; bottom right: market default intensity $\sum_{k=1}^4 \lambda(k) p_t^k$ compared with investor default intensity $\sum_{k=1}^4 \lambda(k) E(p_t^k \mid \mathcal{F}_t^I)$ for high and low observation noise.

a predictable $\mathbb{R}^{d\times d}$-valued process $\widetilde{C}_t^{ij}$, $1 \le i, j \le d$, by

$$\widetilde{C}_t^{ij} = \int_0^t \Sigma_{\mathbf{X}}(\bar{\mathbf{X}}_s, \hat{\mathbf{Y}}_s)ds + \int_0^t \int_E \bar{K}_i^{\mathbf{X}}(\bar{\mathbf{X}}_{s-}, \hat{\mathbf{Y}}_{s-}, u)\bar{K}_j^{\mathbf{X}}(\bar{\mathbf{X}}_{s-}, \hat{\mathbf{Y}}_{s-}, u)F_{\mathcal{N}}(du)\, ds \tag{5.2}$$
$$+ \sum_{n=1}^{\hat{N}_t} F_{\mathcal{N}}\big(D_{\hat{\xi}_n}^{\mathbf{Y}}(\bar{\mathbf{X}}_{\hat{T}_n-}, \hat{\mathbf{Y}}_{\hat{T}_{n-1}})\big)^{-1} \int_{D_{\hat{\xi}_n}^{\mathbf{Y}}(\bar{\mathbf{X}}_{\hat{T}_n-}, \hat{\mathbf{Y}}_{\hat{T}_{n-1}})} \bar{K}_i^{\mathbf{X}}(\bar{\mathbf{X}}_{s-}, \hat{\mathbf{Y}}_{s-}, u)\bar{K}_j^{\mathbf{X}}(\bar{\mathbf{X}}_{s-}, \hat{\mathbf{Y}}_{s-}, u)F_{\mathcal{N}}(du);$$

and finally a predictable random measure ν on $[0,\infty)\times\mathbb{R}^d$ given for bounded and measurable $\gamma : \mathbb{R}^d \to \mathbb{R}$ by

$$\int_0^t \gamma(\boldsymbol{x})\nu(ds, d\boldsymbol{x}) = \int_0^t \int_E \gamma\left(\bar{K}^{\mathbf{X}}(\bar{\mathbf{X}}_{\hat{T}_n-}, \hat{\mathbf{Y}}_{\hat{T}_{n-1}}, u)\right) F_{\mathcal{N}}(du)ds \tag{5.3}$$
$$+ \sum_{n=1}^{\hat{N}_t} F_{\mathcal{N}}\big(D_{\hat{\xi}_n}^{\mathbf{Y}}(\bar{\mathbf{X}}_{\hat{T}_n-}, \hat{\mathbf{Y}}_{\hat{T}_{n-1}})\big)^{-1} \int_{D_{\hat{\xi}_n}^{\mathbf{Y}}(\bar{\mathbf{X}}_{\hat{T}_n-}, \hat{\mathbf{Y}}_{\hat{T}_{n-1}})} \gamma\left(\bar{K}^{\mathbf{X}}(\bar{\mathbf{X}}_{\hat{T}_n-}, \hat{\mathbf{Y}}_{\hat{T}_{n-1}}, u)\right) F_{\mathcal{N}}(du)\,.$$

The triple $(B, \widetilde{C}, \nu)$ has the typical form of *modified semimartingale characteristics* (Definition II.2.16. in Jacod & Shiryaev (2003)). We assume that

A5. The martingale problem associated with $(B, \widetilde{C}, \nu)$ and initial law π_0 is well posed, i.e. there is a unique probability measure $\bar{R}$ on Ω^d such that $\bar{X}$ is a semimartingale with modified characteristics $(B, \widetilde{C}, \nu)$ and initial law π_0.

Furthermore, denote as in Section 3.1 by $(\Omega_2, \mathcal{F}_2, P^{0,\ell})$ the ℓ-dimensional Wiener space with coordinate process $\boldsymbol{\beta}^0$ and let $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{R}^0)$ be the product space

$$(\bar{\Omega}, \bar{\mathcal{F}}, \bar{R}^0) = (\Omega^d \times \Omega_2, \mathcal{F}^d \otimes \mathcal{F}_2, \bar{R} \times P^{0,\ell})\,. \tag{5.4}$$

According to the results from Section 3 (see, in particular, (3.2), (3.3) as well as Theorems 3.1, 3.4 and Proposition 3.2) we obtain that, for a given time t, for a given default history $\{(\hat{T}_n, \hat{\xi}_n) \colon n \le \hat{N}_t\}$, and for a bounded and continuous function $f : S^{\mathbf{X}} \to \mathbb{R}$, the filter $\pi_t f$ can be expressed in the form

$$\pi_t f(\omega_2) \propto E^{\bar{R}}\left(f(\bar{\mathbf{X}}_t)L_t^1 L_t^2(\cdot, \omega_2)\right)\,, \tag{5.5}$$

where

$$L_t^1 = \prod_{n=1}^{\hat{N}_t} \left\{\lambda_{\hat{\xi}_n}(\bar{\mathbf{X}}_{\hat{T}_n-}, \hat{\mathbf{Y}}_{\hat{T}_{n-1}}) \exp\Big(-\int_{\hat{T}_{n-1}}^{\hat{T}_n} \bar{\lambda}(\bar{\mathbf{X}}_s, \hat{\mathbf{Y}}_{\hat{T}_{n-1}})ds\Big)\right\} \exp\Big(-\int_{\hat{T}_{N_t}}^t \bar{\lambda}(\bar{\mathbf{X}}_s, \hat{\mathbf{Y}}_{\hat{T}_{N_t}})ds\Big), \tag{5.6}$$

and where, by analogy with (3.1),

$$L_t^2(\cdot, \omega_2) = \exp\left\{\int_0^t \boldsymbol{a}_s'(\bar{\mathbf{X}}_s)\Sigma_{\mathbf{Z}}^{-1}d\mathbf{Z}_s(\omega_2) - \tfrac{1}{2}\int_0^t \left\|\boldsymbol{a}_s(\bar{\mathbf{X}}_s)\right\|^2_{\Sigma_{\mathbf{Z}}^{-1}}\, ds\right\}, \tag{5.7}$$
$$\mathbf{Z}_t(\omega_2) = v\boldsymbol{\beta}_t^0(\omega_2) = v\,\omega_2(t).$$

Approximating filter sequence. Consider a sequence of processes $(\mathbf{X}^m, \mathbf{Y}^m)_{m\in\mathbb{N}}$, solving the SDE system

$$\mathbf{X}_t^m = \mathbf{X}_0^m + \int_0^t \int_E K^{\mathbf{X},m}(\mathbf{X}_{s-}^m, \mathbf{Y}_{s-}^m, u)\mathcal{N}(ds, du) \tag{5.8}$$

$$Y_{t,j}^m = Y_{0,j} + \int_0^t \int_E (1 - Y_{s-,j}^m) K_j^{\mathbf{Y}}(\mathbf{X}_{s-}^m, \mathbf{Y}_{s-}^m, u)\mathcal{N}(ds, du). \tag{5.9}$$

Note that $K_j^{\mathbf{Y}}$ is independent of m. In applications $K^{\mathbf{X},m}(\cdot)$ will be chosen so that $(\mathbf{X}^m, \mathbf{Y}^m)$ is a finite state Markov chain as e.g. in Section 4.1.

Given the default observation $(\hat{T}_n, \hat{\xi}_n)$, $n = 1, \ldots, \hat{N}_t$ respectively $(\hat{\mathbf{Y}}_s)_{s\le t}$, introduce the modified semimartingale characteristics $(B^m, \widetilde{C}^m, \nu^m)$ given by (5.1), (5.2) and (5.3) with $b = \Sigma = 0$ and $\bar{K}^{\mathbf{X},m}$ instead of $\bar{K}^{\mathbf{X}}$. Choose a sequence of initial distributions π_0^m on the state space of $\mathbf{X}^m$ such that π_0^m converges weakly to π_0 and denote by the measure $\bar{R}^m$ on $(\Omega^d, \mathcal{F}^d, (\mathcal{F}_t^d))$ the solution of the martingale problem associated with the characteristics $(B^m, \widetilde{C}^m, \nu^m)$ and the initial distribution π_0^m (we assume that this martingale problem is well-posed for every m). The filter $\pi_t^m f$ in the approximating model can be expressed in the form

$$\pi_t^m f(\omega_2) \propto E^{\bar{R}^m}\left(f(\bar{\mathbf{X}}_t) L_t^1 L_t^2(\cdot, \omega_2)\right) . \tag{5.10}$$

Below we shall give conditions on $(B^m, \widetilde{C}^m, \nu^m)$ so that the sequence of measures $\bar{R}^m$ converges weakly to $\bar{R}$ (denoted $\bar{R}^m \Rightarrow \bar{R}$) as $m \to \infty$. Assuming for a moment such a convergence, here we show first the ensuing convergence of the filters. We need the following additional assumption.

A6. The default intensities $\lambda_j(\boldsymbol{x}, \boldsymbol{y})$ are bounded and continuous in $\boldsymbol{x}$.

Theorem 5.1 *Fix some $t > 0$ and a default history $\{(\hat{T}_n, \hat{\xi}_n)\colon n = 1, \ldots, \hat{N}_t\}$. Suppose that Assumptions A1 to A6 hold and that $\bar{R}^m \Rightarrow \bar{R}$. Then $\pi_t^m f(\omega_2)$ converges $P^{0,\ell}$-stochastically to $\pi_t f(\omega_2)$, i.e. $P^{0,\ell} - \lim_{m\to\infty} \pi_t^m f = \pi_t f$.*

Remark 5.2 1.) Although the convergence is in a weaker form than a.s. convergence, it still implies that, for m sufficiently large, the probability that $\pi_t^m f$ differs from $\pi_t f$ by a given amount can be made arbitrarily small.
2.) Note that Theorem 5.1 is a filter approximation result for a model where signal and observation cannot be made independent via a measure transformation. This sets the result apart from filter approximation results as in Zeng (2003) which are based on general results by T. Kurtz and E. Goggins concerning weak convergence of conditional expectations. By the same token, Zeng (2003) obtains only weak convergence of the approximating filters, while here we obtain convergence in probability.

Proof (of Theorem 5.1) Denote by $\delta_d(\boldsymbol{x},\boldsymbol{y})$ the Prokhorov metric on $D^d([0,\infty))$ (see Jacod & Shiryaev (2003), Chapter VI, (1.26)). By Skorokhod embedding, the weak convergence $\bar{\mathbb{R}}^m \Rightarrow \bar{R}$ implies that there is some probability space - denoted again by $(\Omega^d, \mathcal{F}^d, \bar{R})$ for simplicity - and processes $\bar{\mathbf{X}}^m$ and $\bar{\mathbf{X}}$ with laws $\bar{R}^m$ and $\bar{R}$ respectively such that $\lim_{m\to\infty}\delta_d(\bar{\mathbf{X}}^m,\bar{\mathbf{X}}) = 0$, $\bar{R}$ a.s. and hence also $\bar{R}^0 = \bar{R}\times P^{0,\ell}$ a.s. Now we have the following two Lemmas, whose proof is relegated to Appendix A.

Lemma 5.3 *Consider bounded and continuous functions $f(\cdot)$ and $\lambda(\cdot)$ We get for processes $\bar{\mathbf{X}}^m, \bar{\mathbf{X}}$ as above that*

$$\lim_{m\to\infty} f(\bar{\mathbf{X}}^m_t) = f(\bar{\mathbf{X}}_t) \quad \bar{R}^0 - a.s. \tag{5.11}$$

$$\lim_{m\to\infty} \lambda(\bar{\mathbf{X}}^m_{\hat{T}_n-}) = \lambda(\bar{\mathbf{X}}_{\hat{T}_n-}), \quad \bar{R}^0 - a.s., \quad n = 1,\ldots,\hat{N}_t \tag{5.12}$$

$$\lim_{m\to\infty} \int_{t_1}^{t_2} \bar{\lambda}(\bar{\mathbf{X}}^m_s, \hat{\mathbf{Y}}_s)ds = \int_{t_1}^{t_2} \bar{\lambda}(\bar{\mathbf{X}}_s, \hat{\mathbf{Y}}_s)ds\,; \quad t_1 < t_2 < t;\ \bar{R}^0 - a.s. \tag{5.13}$$

Lemma 5.4 *Let $L^{2,m}_t$ denote the process L^2_t defined in (5.7), but with $\bar{\mathbf{X}}^m$ replacing $\bar{\mathbf{X}}$ there. Then, for processes $\bar{\mathbf{X}}^m, \bar{\mathbf{X}}$ as above, one has* $\bar{R}^0 - \lim_{m\to\infty} L^{2,m}_t = L^2_t$.

Now we return to the proof of Theorem 5.1. From the boundedness of f and $L^{1,m}$ (see A6), the definition of $L^{1,m}$ according to (5.6) (with $\bar{\mathbf{X}}^m$ instead of $\mathbf{X}$) and the fact that $E^{\bar{R}^0}\left((L^{2,m})^2\right) \le C < \infty$ (due to the boundedness of $\boldsymbol{a}_t(\cdot)$) we obtain uniform integrability for the sequence $(f(\bar{\mathbf{X}}^m_t)\, L^{1,m}_t L^{2,m}_t)_{m\in\mathbb{N}}$. From Lemma 5.3 as well as from Lemma 5.4 we then obtain that

$$f(\bar{\mathbf{X}}^m_t)\, L^{1,m}_t L^{2,m}_t \;\to\; f(\bar{\mathbf{X}}_t)\, L^1_t L^2_t \quad \text{in} \quad \mathcal{L}^1(\bar{\Omega}, \bar{\mathcal{F}}, \bar{R}^0)\,,\; m\to\infty\,. \tag{5.14}$$

Using the product-form of $\bar{R}^0$, this $\mathcal{L}^1$-convergence can be written more explicitly as

$$\lim_{m\to\infty} \int_{\Omega_2} E^{\bar{R}}\left(\left|f(\bar{\mathbf{X}}^m_t)\, L^{1,m}_t L^{2,m}_t(\cdot,\omega_2) - f(\bar{\mathbf{X}}_t)\, L^1_t L^2_t(\cdot,\omega_2)\right|\right) P^{0,\ell}(d\omega_2) = 0.$$

It follows that the inner expectation converges to zero in $\mathcal{L}^1(\Omega_2, \mathcal{F}_2, P^{0,\ell})$ and therefore also $P^{0,\ell}$-stochastically, which proves the theorem.

Weak convergence. In the remaining part of this subsection, under a couple of additional assumptions on the model, we give conditions on the modified characteristics of $\bar{\mathbf{X}}^m$ for which we obtain weak convergence of $\bar{\mathbf{X}}^m$ to $\bar{\mathbf{X}}$. For this purpose we base ourselves on Theorems 3.35 and 2.11 in chapter IX of Jacod & Shiryaev (2003). The additional assumptions on the model are:

A7.1 *(a strengthening of A2).* There exists $A \subset E$ with A compact such that $D^{\mathbf{X}}_i(\boldsymbol{x},\boldsymbol{y}) \subset A$ for all $\boldsymbol{x},\boldsymbol{y}$ and all $i = 1,\cdots,d$.

A7.2 There exists a positive constant H such that for all $\boldsymbol{x}, \boldsymbol{y}$, all $i = 1, \cdots, d$ and all $u \in E$,

$$\sum_{i=1}^{d} \mid b_i(\boldsymbol{x}, \boldsymbol{y}, u) \mid \leq H \, ; \quad \sum_{i=1}^{d} \mid \bar{K}_i^{\mathbf{X}}(\boldsymbol{x}, \boldsymbol{y}, u) \mid \leq H \, ; \quad \sum_{i=1}^{d} \mid \Sigma_{\mathbf{X}}^{ii}(\boldsymbol{x}, \boldsymbol{y}) \mid \leq H \, .$$

Following Jacod & Shiryaev (2003) we introduce the following class of test functions: $\mathcal{C}_1 := \{\gamma : \mathbb{R}^d \to \mathbb{R} \colon \gamma \text{ continuous}, |\gamma(\boldsymbol{x})| < 1, \gamma(0) = 0\}$. Furthermore, for generic d, $\delta_d(\boldsymbol{x}, \boldsymbol{y})$ denotes the Prokhorov metric on the d-dimensional Skorohod space.

Proposition 5.5 *Let Assumption A1 to A7 hold and suppose that for the characteristics of the approximating Markov chains $\bar{\mathbf{X}}^m$ one has*

$$\begin{cases} \delta_d(B^m, B \circ \bar{\mathbf{X}}) & \xrightarrow{\bar{R}^m} 0\,, \\ \delta_{d^2}(\tilde{C}^m, \tilde{C} \circ \bar{\mathbf{X}}) & \xrightarrow{\bar{R}^m} 0\,, \\ \delta_1(\gamma * \nu^m, (\gamma * \nu) \circ \bar{\mathbf{X}}) & \xrightarrow{\bar{R}^m} 0 \text{ for all } \gamma \in \mathcal{C}_1\,. \end{cases} \tag{5.15}$$

where $\bar{R}^m$ is the sequence of measures making the coordinate process a semimartingale with modified characteristics $(B^m, \tilde{C}^m, \nu^m)$ and initial distribution π_0^m with $\pi_0^m \Longrightarrow \pi_0$. Then, for a given default sequence $(\hat{T}_n, \hat{\xi}_n)$, we have the weak convergence $\bar{R}^m \Longrightarrow \bar{R}$ for $m \to \infty$.

The conditions (5.15) can be taken as guidelines when choosing the approximating sequence of finite state Markov chains. While given in a somewhat abstract form here, they assume a more specific form for a given problem at hand. The proof of the proposition is given in Appendix A.

A Additional proofs for Section 5

Proof of Lemma 5.3. Note that by the form of the characteristics of $\bar{\mathbf{X}}$, for $t \notin J_t := \{\hat{T}_n : n = 1, \ldots, \hat{N}_t\}$ we have $\Delta\bar{\mathbf{X}}_t = 0$, $\bar{R}^0-$a.s. Given the $\bar{R}^0-$a.s. convergence of $\bar{\mathbf{X}}^m$ to $\bar{\mathbf{X}}$ in the Prokhorov metric, (5.11) follows from Proposition VI.2.1 (b.5) in Jacod & Shiryaev (2003). On the other hand, if $\nu(\{\hat{T}_n\} \times \mathbb{R}^d) > 0$ then also $\nu^m(\{\hat{T}_n\} \times \mathbb{R}^d) > 0$ for m sufficiently large, so that Proposition VI.2.1 (b.6) in Jacod & Shiryaev (2003) implies that for $\hat{T}_n \leq t$ we have $\bar{R}^0-$a.s.

$$\lim_{m\to\infty} \bar{\mathbf{X}}^m_{\hat{T}_n-} = \bar{\mathbf{X}}_{\hat{T}_n-} \text{ and } \lim_{m\to\infty} \bar{\mathbf{X}}^m_{\hat{T}_n} = \bar{\mathbf{X}}_{\hat{T}_n}. \tag{A.1}$$

Relation (A.1) then implies (5.12) and, for $t \in J_t$, also (5.11). Relation (5.13) is obvious from the definition of the Skorokhod topology.

Proof of Lemma 5.4. We first consider the stochastic integral terms. Since $\mathbf{Z}_t = v\boldsymbol{\beta}_t^0$, we get

$$E^{\bar{R}^0}\Big(\Big(\int_0^t \boldsymbol{a}_s'(\bar{\mathbf{X}}_s^m)\Sigma_{\mathbf{Z}}^{-1}d\mathbf{Z}_s - \int_0^t \boldsymbol{a}_s'(\bar{\mathbf{X}}_s)\Sigma_{\mathbf{Z}}^{-1}d\mathbf{Z}_s\Big)^2\Big) = E^{\bar{R}^0}\Big(\int_0^t \big\|(\boldsymbol{a}_s(\bar{\mathbf{X}}_s^m) - \boldsymbol{a}_s(\bar{\mathbf{X}}_s))\big\|^2_{\Sigma_{\mathbf{Z}}^{-1}}\,ds\Big) \;\to\; 0\,, \quad m\to\infty\,, \tag{A.2}$$

where the convergence follows from the assumption of continuity and boundedness of the function $\boldsymbol{a}_t(\cdot)$ (see Assumption A4) as well as from bounded convergence. Next, always by the Assumption A4 as regards $\boldsymbol{a}_t(\cdot)$ as well as by the triangle inequality we also have,

$$\int_0^t \Big(\big\|\boldsymbol{a}_s(\bar{\mathbf{X}}_s^m)\big\|^2_{\Sigma_{\mathbf{Z}}^{-1}} - \big\|\boldsymbol{a}_s(\bar{\mathbf{X}}_s)\big\|^2_{\Sigma_{\mathbf{Z}}^{-1}}\Big)\,ds \le 2\,\|\boldsymbol{a}\| \int_0^t \big\|(\boldsymbol{a}_s(\bar{\mathbf{X}}_s^m) - \boldsymbol{a}_s(\bar{\mathbf{X}}_s))\big\|_{\Sigma_{\mathbf{Z}}^{-1}}\,ds\,,$$

and, as $m\to\infty$, the right hand side converges to 0 R^0-a.s. This convergence and relations (A.2) imply that for $m\to\infty$,

$$\int_0^t \boldsymbol{a}_s'(\bar{\mathbf{X}}_s^m)\Sigma_{\mathbf{Z}}^{-1}d\mathbf{Z}_s - \frac{1}{2}\int_0^t \|\boldsymbol{a}_s(\bar{\mathbf{X}}_s^m)\|\,2_{\Sigma_{\mathbf{Z}}^{-1}}\,ds \overset{\bar{R}^0}{\longrightarrow} \int_0^t \boldsymbol{a}_s'(\bar{\mathbf{X}}_s)\Sigma_{\mathbf{Z}}^{-1}d\mathbf{Z}_s - \frac{1}{2}\int_0^t \|a_s(\bar{\mathbf{X}}_s)\|^2_{\Sigma_{\mathbf{Z}}^{-1}}\,ds$$

and therefore also $\bar{R}^0 - \lim_{m\to\infty} L_t^{2,m} = L_t^2$ for arbitrary fixed t.

Proof of Proposition 5.5. The proof is based on the following two lemmas.

Lemma A.1 *Under the assumptions of Proposition 5.5 the sequence of measures $\bar{R}^m$, $m\in\mathbb{N}$, is tight.*

Proof The proof is based on that for Theorem 3.35 in Chapter IX of Jacod & Shiryaev (2003). We show that, under our assumptions, Conditions i),ii),v) and vi) of that theorem are satisfied.

Condition i) *(Strong majorization hypothesis II)* can, on the basis of Assumption A7.1 and A7.2, be seen to be satisfied if in the definition of the strong majorization condition (Definition 3.11 in Chapter IX of Jacod & Shiryaev (2003)) one takes as deterministic increasing càdlàg functions the following:

$$F_t = [2H + H(1+H)F_{\mathcal{N}}(A)]\,t + H(H+1)\,\hat{N}_t \tag{A.3}$$

with A and H from Assumptions A7.1 and A7.2 respectively; for $\gamma\in\mathcal{C}_1$ we take

$$F_t^\gamma = F_{\mathcal{N}}(A)\,t + \hat{N}_t\,. \tag{A.4}$$

Recall that we work with a given observed sequence $(\hat{T}_n,\hat{\xi}_n)$, so that the functions F_t and F_t^γ are deterministic functions of time. The remaining conditions are immediate: Condition ii) *(Condition on the big jumps)* is automatically satisfied under the given assumptions; Condition v) holds by assumption on the initial conditions; Condition vi) corresponds to (5.15).

The statement of the lemma follows now from the first part of the proof of Theorem 3.35 in Chapter IX of Jacod & Shiryaev (2003), which in turn is based on Theorem 3.20 in the same chapter.

Lemma A.2 *Under the hypotheses of Proposition 5.5 every weakly converging sequence* $\bar{\mathbf{X}}^m$ *has as limit a semimartingale process with modified characteristics* $(B, \tilde{C}, \nu)$

Proof Here we rely on Theorem 2.11 in Chapter IX of Jacod & Shiryaev (2003). Condition i) in that Theorem is satisfied since the assumption made by requiring (5.15) is stronger than this condition. Condition ii) *(majorization condition)* is here a rather immediate consequence of assumptions A7.1 andA7.2.

There remains Condition iii), namely the *(Continuity condition)* for the modified characteristics $(B, \tilde{C}, \nu)$ in the Skorokhod topology. This condition has to hold a.s. with respect to the limit measure, in our case $\bar{R}$. Next recall that each of the characteristics is composed of terms expressed as an integral with respect to time and one term given by a sum over $\hat{T}_n$. The time integrals are automatically continuous. The term given by the sum over $\hat{T}_n$ can on the other hand be treated by analogy to Lemma 5.3 noticing that we have $\lambda_j(\mathbf{X}_{t-}, \hat{\mathbf{Y}}_{t-}) = F_{\mathcal{N}}(D_j^{\mathbf{Y}}(\mathbf{X}_{t-}, \hat{\mathbf{Y}}_{t-}))$ and for $\lambda_j(\cdot, \boldsymbol{y})$ we have the continuity assumption A6. Having now all conditions of Theorem 2.11 in Chapter IX of Jacod & Shiryaev (2003) satisfied, the statement here follows from that same Theorem.

We can now conclude with the proof of Proposition 5.5: By the tightness result of Lemma A.1 we have that every sequence $\bar{R}^m$ has a weakly converging subsequence. By Lemma A.2 each of these converging sequences has a weak limit that corresponds to the same modified characteristics $(B, \tilde{C}, \nu)$, namely those of the original process $\bar{\mathbf{X}}$. By Assumption A5 the weak limit is unique and this gives us the statement of the proposition.

References

Arnsdorf, M. & Halperin, I. (2007), 'BSLP: Markovian bivariate spread-loss model for portfolio credit derivatives', working paper, JP Morgan.

Azizpour, S. & Giesecke, K. (2008), 'Self-exciting corporate defaults', preprint, Stanford University.

Budhiraja, A., Chen, L. & Lee, C. (2007), 'A survey of nonlinear methods for nonlinear filtering problems', *Physica D* **230**, 27–36.

Ceci, C. & Gerardi, A. (2006), 'A model for high frequency data under partial information: a filtering approach', *International Journal of Theoretical and Applied Finance (IJTAF)* pp. 1–22.

Clark, J. (1978), The design of robust approximations to the stochastic differential equations of nonlinear filtering, *in* J. Skwirzynski, ed., 'Communication Systems and Random Process Theory', Sijthoff and Noordhoff, pp. 721–734.

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

Collin-Dufresne, P., Goldstein, R. & Helwege, J. (2003), 'Is credit event risk priced? modeling contagion via the updating of beliefs', Preprint, Carnegie Mellon University.

Crisan, D. & Lyons, T. (1999), 'A particle approximation of the solution of the Kushner-Stratonovich equation', *Probability Theory and Related Fields* **115**, 549–578.

Cvitanic, J., Liptser, R. & Rozovski, B. (2006), 'A filtering approach to tracking volatility from prices observed at random times', *The Annals of Applied Probability* **16**, 1633–1652.

Davis, M. (1978), Pathwise non-linear filtering, *in* M. Hazewinkel & J. Willems, eds, 'Stochastic Systems: The Mathematics of Filtering and Identification and Applications', D. Reidel Publishing Company, pp. 505–528.

Davis, M. & Lo, V. (2001), 'Infectious defaults', *Quant. Finance* **1**, 382–387.

Duffie, D., Eckner, A., Horel, G. & Saita, L. (2006), 'Frailty correlated default', forthcoming in *Journal of Finance*.

Duffie, D. & Garleanu, N. (2001), 'Risk and valuation of collateralized debt obligations', *Financial Analyst's Journal* **57**(1), 41–59.

Duffie, D. & Lando, D. (2001), 'Term structure of credit risk with incomplete accounting observations', *Econometrica* **69**, 633–664.

Elliott, R. (1993), 'New finite-dimensional filters and smoothers for noisily observed markov chains', *IEEE Trans. Info. theory* **IT-39**, 265–271.

Elliott, R., Jeanblanc, M. & Yor, M. (2000), 'On models of default risk', *Mathematical Finance* **10**, 179–195.

Frey, R. & Backhaus, J. (2007), 'Dynamic hedging of synthetic CDO-tranches with spread- and contagion risk', preprint, department of mathematics, Universität Leipzig. available from `http://www.math.uni-leipzig.de/ frey/publications-frey.html`.

Frey, R., Prosdocimi, C. & Runggaldier, W. (2007), Affine credit risk models under incomplete information, *in* J. Akahori, S. Ogawa & S. Watanabe, eds, 'Stochastic Processes and Application to Mathematical Finance', World Scientific, Singapore, pp. 97–113.

Frey, R. & Schmidt, T. (2006), 'Pricing corporate securities under noisy asset information', preprint, Universität Leipzig, forthcoming in *Mathematical Finance*.

Frey, R. & Schmidt, T. (2009), 'Credit Innovation: Pricing and hedging of credit derivatives via the innovations approach to nonlinear filtering', preprint, Universität Leipzig.

Giesecke, K. and Goldberg, L. (2004), 'Sequential defaults and incomplete information', *Journal of Risk* **7**, 1–26.

Gombani, A., Jaschke, S. & Runggaldier, W. (2005), 'A filtered no arbitrage model for term structures with noisy data', *Stochastic Processes and Applications* **115**, 381–400.

Grigelionis, B. (1972), 'On stochastic equations for nonlinear filtering problem of stochastic processes', *Lietuvos Matematikos Rinkinys* **12**, 37–51.

Jacod, J. & Shiryaev, A. (2003), *Limit Theorems for Stochastic Processes*, 2nd edn, Springer Verlag, Berlin.

Jarrow, R. & Protter, P. (2004), 'Structural versus reduced-form models: a new information based perspective', *Journal of Investment management* **2**, 1–10.

Kallianpur, G. & Striebel, C. (1968), 'Estimation of stochastic systems: arbitrary systems process with additive white noise observation errors', *Annals of Mathematical Statistics* **39**, 785–801.

Kliemann, W., Koch, G. & Marchetti, F. (1990), 'On the unnormalized solution of the filtering problem with counting process observations', *IEEE* **IT-36**, 1415–1425.

Kusuoka, S. (1999), 'A remark on default risk models', *Adv. Math. Econ.* **1**, 69–81.

McNeil, A., Frey, R. & Embrechts, P. (2005), *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University Press, Princeton, New Jersey.

Schönbucher, P. (2004), 'Information-driven default contagion', Preprint, Department of Mathematics, ETH Zürich.

Zeng, Y. (2003), 'A partially observed model for micromovement of asset prices with Bayes estimation via filtering', *Mathematical Finance* **13**, 411–444.