Probability in the Engineering and Informational Sciences, **25**, 2011, 393–418.
doi:10.1017/S0269964811000076

EXPLICIT COMPUTATIONS FOR A FILTERING PROBLEM WITH POINT PROCESS OBSERVATIONS WITH APPLICATIONS TO CREDIT RISK

VINCENT LEIJDEKKER
Product Analysis
ABN AMRO Bank N.V.
Gustav Mahlerlaan 10
1082 PP Amsterdam
The Netherlands
E-mail: vincent.leijdekker@gmail.com

PETER SPREIJ
Korteweg-de Vries Institute for Mathematics
Universiteit van Amsterdam
PO Box 94248
1090GE Amsterdam
The Netherlands
E-mail: spreij@uva.nl

We consider the filtering problem for a doubly stochastic Poisson or Cox process, where the intensity follows the Cox–Ingersoll–Ross model. In this article we assume that the Brownian motion, which drives the intensity, is not observed. Using filtering theory for point process observations, we first derive the dynamics for the intensity and its moment-generating function, given the observations of the Cox process. A transformation of the dynamics of the conditional moment-generating function allows us to solve in *closed form* the filtering problem, between the jumps of the Cox process as well as at the jumps, which constitutes the main contribution of the article. Assuming that the initial distribution of the intensity is of the Gamma type, we obtain an *explicit* solution to the filtering problem for all $t > 0$. We conclude the article with the observation that the resulting conditional moment-generating function at time t, after N_t jumps, corresponds to a mixture of $N_t + 1$ Gamma distributions. Currently, the model that we analyze has become popular in credit risk modeling, where one uses the intensity-based approach for the modeling of default times of one or more companies. In this approach, the default times are defined as the jump times of a Cox

 0269-9648/11 $25.00

process. In such a model, one only has access to observations of the Cox process, and thus filtering comes in as a natural technique in credit risk modeling.

1. INTRODUCTION

The main results of this article are explicit, closed-form expressions for the solution of a filtering problem with counting process observations, where the unobserved intensity process is a solution to a square root stochastic differential equation. As a matter of fact, the explicit solution we provide is split into a part that concerns the update of the filter at jump times and a part that solves the problem between jump times. This, of course, reflects the usual strategy for filtering problems with counting observations. The evolution of the filter between jump times is commonly expressed by a partial differential equation (PDE) for the conditional moment-generating function. In general, an explicit solution to this PDE is impossible to get, if only were it for the fact that it can be viewed as an infinite-dimensional problem, reflecting that the filter itself is, in general, infinite dimensional. However, for the specific choice of the state process that we have made, we are able to provide explicit solutions. The choice for our specification of the state process is made upon two considerations. First, it is known that for a conditional Poisson process where the intensity is a random variable having a Gamma distribution, the filtered intensity at a time t also has a Gamma distribution, with parameters depending on t and the value N_t of the counting process. However, in the case that we analyze, the random intensity is also evolving in time; it solves a stochastic differential equation of the Cox–Ingersoll–Ross (CIR) type, which admits a stationary solution for which all marginal distributions also belong to the Gamma family. By choosing the initial distribution of the intensity properly, we are able to come up with explicit expressions for the conditional moment-generating function, and we also show that the filtered intensities have distributions that are mixtures of Gamma distributions.

Another reason to study the chosen state process is that it comes up naturally in a simple model for credit risk, which has become a major field of interest in financial mathematics. Indeed, in [23] the author considered this model for the intensity. Further, Duffie [9] considered more general affine models for credit risk. The filtering problem in this setup has previously been studied in [12], where the focus was more on the update for the filter on jump times, whereas we also treat the evolution between jump times in great detail. The filtering problem as such has already been mentioned in [3], where the state process was assumed to follow an Ornstein–Uhlenbeck process and the intensity of the counting process was taken to be the square of the state process, which is easily shown to be a CIR process. Although in [3] attention has been paid to the evolution of the filter between jump times, an explicit formula for the solution of the resulting PDE has not been given. We obtain this part of the solution analytically by providing a closed-form solution to a PDE. Furthermore, we follow a different approach to obtain the recursive solution at jump times as compared to [12]. By

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

combining these solutions, we obtain a solution for all $t > 0$. It is further observed that the resulting conditional moment-generating function at time t corresponds to a mixture of $N_t + 1$ Gamma distributions according to some discrete distribution.

Let us give some background for credit risk modeling and explain why filtering is a natural tool in this field of research. The main goal in credit risk is the modeling of the default time of a company or default times of several companies. The default times are often modeled using the so-called intensity-based approach, as opposed to the firm value approach. Here, the default time of a company is modeled as the first jump time of a Cox process, of which the intensity is driven by some stochastic process (e.g., Brownian motion) or, in case of more than one company, as consecutive jump times of this Cox process. This approach enables one to calculate survival probabilities and to price financial derivatives depending on the default of one or more companies, such as defaultable bonds and credit default swaps. We refrain from a further presentation of these issues, as it is beyond the objectives of this article and refer to [20] and [22] for detailed expositions. Overviews of the intensity-based modeling approach can be found in [10,15,18]. In this approach, it is a common assumption that the driving process can be observed; that is, the observed filtration is generated by the Cox process, which can be seen as the default counting process, and by the driving process.

In this article it is assumed that the driving process is *not* observed, and thus only a point process N_t is observed, which introduces a stochastic filtering problem for point processes. For instance, one may want to compute the conditional probability that default has not occurred yet, given the observations, the conditional survival probability. In particular, we assume the intensity to follow the CIR model, where the driving Brownian motion is not observed. See Example 3.1 for the computation of the conditional survival probability. For general results on nonlinear filtering in models for credit risk, we refer to [13,14].

Our results are obtained under the assumption that the CIR process follows a stochastic differential equation (SDE) with constant parameters. We briefly discuss what happens if we let the parameters also depend on time. Such a model is more attractive from a practical point, as it allows for more flexible modeling. In general, we will then lose the attractive feature of obtaining closed-form solutions. However, if one restricts the model by taking parameters that are piecewise constant functions, closed-form solutions still exist. In practice, these piecewise constant models have become popular in credit risk, as its flexibility does not destroy calibration procedures; see [19].

The article is organized as follows: In Section 2 the CIR model is discussed and some results for the case of full information are discussed. Next, in Section 3 the filtering problem is introduced and some background is given for filtering of point process observations. First, the filtering formulas from [5] are given, and the equations for the conditional intensity and conditional moment-generating function are derived. Then in the second part of Section 3, we introduce *filtering by the method of the probability of reference*, and the filtering equations are transformed using the ideas introduced in [3]. Section 4 deals with the filtering problem between the jump times of the point process, given the initial distribution of the intensity at jump times. In Section 5 the filtering problem is solved at jump times, and an explicit, recursive solution is

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

obtained, which combines the solutions between and at jumps. Further, the resulting conditional moment-generating function is analyzed and it is observed that this function agrees with the moment-generating function of a mixture of Gamma distributions. The section concludes with an illustration of the mixing probabilities. Finally, in Section 6 we discuss possible extensions of the problem under consideration, where the parameters of the SDE for the state process are allowed to be time-varying or where the state process is more dimensional.

2. MODEL AND BACKGROUND

The main goal of the article is to derive explicit closed-form expressions for a filtering problem with counting process observations. Filtering problems with such observations have been studied already some 30 years ago; see, for example, [4,24,25] and the later appearing book [5]. Recently, this kind of problem has gained renewed interest in the field of credit risk modeling (see also [12]), as we will outline below. One of the main goals in this field is the modeling of the default time of a company or the default times of several companies. Over the years, two approaches have become popular, the *structural approach* and the *intensity-based approach*. In the structural approach, the company value is modeled, for example as a (jump-)diffusion, and the company defaults when its value drops below a certain level. This approach is discussed in more detail in, for example, [2,11,15]. In the intensity-based approach, the default time is modeled as the first jump of a point process (e.g., a Poisson process or, more general, a Cox process, which is an inhomogeneous Poisson process conditional on the realization of its intensity). In the case that one considers more than one company, one can model the default times as consecutive jump times of the Cox process. In [10,15,18] this modeling approach is discussed in more detail, and [21] provides a detailed application. In this article we focus on the intensity-based approach, where the intensity λ_t of the Cox process, a nonnegative process, has an affine structure, similar to interest term structure models [8]. This means that the intensity process λ_t follows a SDE of the form

$$\mathrm{d}\lambda_t = (a + b\lambda_t)\,\mathrm{d}t + \sqrt{c + d\lambda_t}\,\mathrm{d}W_t, \tag{2.1}$$

for a Brownian motion W_t, with $d > 0$. In particular, the focus is on the CIR square root model with mean reversion [7] for the intensity, where the intensity λ_t satisfies

$$\mathrm{d}\lambda_t = -\alpha(\lambda_t - \mu_0)\,\mathrm{d}t + \beta\sqrt{\lambda_t}\,\mathrm{d}W_t. \tag{2.2}$$

In [17, Sect. 6.2.2.] one finds parameter restrictions for this model that guarantee positivity of λ_t. Naturally one should start with a positive initial value λ_0, and if $\alpha\mu_0 \geq \beta^2/2$, then λ_t remains strictly positive with probability 1. Note that using the transformation $X_t = \lambda_t + c/d$ and by a reparametrization, X_t satisfies the general SDE (2.1) and λ_t satisfies (2.2). This implies that the general form (2.1) and the CIR intensity (2.2) are in fact equivalent. Therefore, the CIR intensity will be considered in most of the remainder of this article.

A big advantage of the affine setup that we choose in this article, is that many relevant quantities in credit risk can be calculated explicitly. Using the formulas from [17, Sect. 6.2.2.] one can, for example, easily calculate the survival probability $\mathbb{P}(\tau > t|\mathcal{F}_s)$, with $t > s$ and $\mathcal{F}_t = \mathcal{F}_t^N \vee \mathcal{F}_t^Y$, where the former filtration is generated by the point process N_t and the latter is generated by some process Y_t driving the intensity process.

Example 2.1: Consider, on the filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$, a random time $\tau > 0$ as the first jump time of a Cox process N_t, whose intensity follows the CIR model (2.2). Further assume that $\mathcal{F}_t = \mathcal{F}_t^N \vee \mathcal{F}_t^W$, where $\mathcal{F}_t^W$ is the filtration generated by the Brownian motion that drives the intensity process. Then one can calculate the survival probability for $t > s$ as

$$\mathbb{P}(\tau > t|\mathcal{F}_s) = 1_{\{\tau > s\}} \mathbb{E}\left[e^{-\int_s^t \lambda_u \, \mathrm{d}u} \middle| \mathcal{F}_s^W \right], \tag{2.3}$$

which follows from formulas in [2, Chap. 6]. Because λ_t is a Markov process, one can condition on λ_s instead of $\mathcal{F}_s^W$. An application of Proposition 6.2.4. from [17] to (2.3) yields

$$\mathbb{P}(\tau > t|\mathcal{F}_s) = 1_{\{\tau > s\}} \exp\left(-\alpha\mu_0\phi(t-s) - \lambda_s\psi(t-s)\right), \tag{2.4}$$

where

$$\begin{aligned}
\phi(t) &= -\frac{2}{\beta^2}\log\left(\frac{2\gamma e^{t(\gamma+\alpha)/2}}{\gamma-\alpha+e^{t\gamma}(\gamma+\alpha)}\right),\\
\psi(t) &= \frac{2\left(e^{\gamma t}-1\right)}{\gamma-\alpha+e^{t\gamma}(\gamma+\alpha)},\\
\gamma &= \sqrt{\alpha^2+2\beta^2}.
\end{aligned}$$

Other relevant quantities, such as the price of a defaultable bond, can also be calculated analytically, under some restrictions on the interest rate (e.g., by posing that the interest rate evolves deterministically). In [12] some of these quantities are considered in more detail.

It is a common assumption, which is also followed above, that the filtration $\mathcal{F}_t$ is built up using two filtrations $\mathcal{F}_t^Y$ and $\mathcal{F}_t^N$, where the first filtration represents the information about the process driving the intensity and the second filtration contains information about past defaults. In this article it is assumed that the factor Y is not observed which results in a filtering problem of a point process.

In the following sections the problem is introduced formally and solved for the case where the intensity follows the CIR model.

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

3. THE FILTERING PROBLEM

In filtering theory, one deals with the problem of partial observations. Suppose that a process Z_t on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is adapted to the filtration $\mathcal{F}_t$. Furthermore, let the process Y_t be observed, where Y_t is measurable with respect to a smaller filtration $\mathcal{F}_t^Y \subsetneq \mathcal{F}_t$. One is then interested in conditional expectations of the form $\widehat{Z}_t = \mathbb{E}\left[Z_t|\mathcal{F}_t^Y\right]$, and one tries to find the dynamics of the process $\widehat{Z}_t$—for instance, by showing that it is the solution of a stochastic differential equation.

In this section the filtering problem is considered in the case that a point process is observed. First, some general theory about filtering with point process observations is discussed, and Example 2.1 is continued within the filtering setup. The calculation of the survival probability depends on the conditional moment-generating function, for which an SDE is derived. In the second part of this section, this equation is transformed in such a way that the filtering problem allows an explicit solution.

3.1. Filtering Using Point Process Observations

In the case of point process observations, the observed process Y_t is equal to the point process N_t, with $\mathcal{F}_t$-intensity λ_t. The process Z_t is assumed to follow the SDE

$$\mathrm{d}Z_t = a_t\,\mathrm{d}t + \mathrm{d}M_t \tag{3.1}$$

for an $\mathcal{F}_t$-progressive measurable a_t, with $\int_0^t |a_s|\,\mathrm{d}s < \infty$, and an $\mathcal{F}_t$-martingale M_t. The filtering problem is often cast as the calculation of the conditional expectation $\mathbb{E}[Z_t|\mathcal{F}_t^N] =: \widehat{Z}_t$. Using the filtering formulas from [5, Chap. IV] and under the conditions of its Theorem T2, a representation of the solution to this filtering problem can be found. In the case that the martingale M_t and the observed point process have no jumps in common, one has

$$\mathrm{d}\widehat{Z}_t = \widehat{a}_t\,\mathrm{d}t + \left(\frac{\widehat{Z\lambda}_{t-}}{\widehat{\lambda}_{t-}} - \widehat{Z}_{t-}\right)\left(\mathrm{d}N_t - \widehat{\lambda}_t\,\mathrm{d}t\right), \tag{3.2}$$

with $\widehat{a}_t := \mathbb{E}[a_t|\mathcal{F}_t^N]$ and $X_{t-} := \lim_{s\uparrow t} X_s$.

Example 3.1 (Example 2.1 continued): When one wants to calculate the survival probability given $\mathcal{F}_t^N$, one has $Z_t = 1_{\{\tau>t\}}$. Combining this with the survival probability in the case of full information, one can calculate the survival probability $\mathbb{P}(\tau > t|\mathcal{F}_s^N)$:

$$\begin{aligned}\mathbb{P}\left(\tau > t|\mathcal{F}_s^N\right) &= \mathbb{E}\left[\mathbb{P}\left(\tau > t|\mathcal{F}_s^N \vee \mathcal{F}_s^W\right)\middle|\,\mathcal{F}_s^N\right]\\ &= 1_{\{\tau>s\}}\exp\left(-\alpha\mu_0\phi(t-s)\right)\mathbb{E}\left[\exp(-\psi(t-s)\lambda_s)|\,\mathcal{F}_s^N\right],\end{aligned}$$

which can be calculated if an expression for the conditional moment-generating function $\widehat{f}(s,t) := \mathbb{E}\left[e^{s\lambda_t}\middle|\,\mathcal{F}_t^N\right]$ is available.

Example 3.1 illustrates that one can calculate the survival probability if the conditional moment-generating function $\widehat{f}(s,t)$ is known. As a first step in the determination of this function, the SDEs of $\widehat{\lambda}_t := \mathbb{E}\left[\lambda_t | \mathcal{F}_t^N\right]$ and $\widehat{f}(s,t)$ are determined. First, Itô's formula is used to obtain the SDE for $e^{s\lambda_t}$, where λ_t satisfies (2.2):

$$\mathrm{d}e^{s\lambda_t} = \left[\left(-\alpha s + \frac{1}{2}s^2\beta^2\right)\frac{\partial}{\partial s}e^{s\lambda_t} + s\alpha\mu_0 e^{s\lambda_t}\right]\mathrm{d}t + \beta\sqrt{\lambda_t}e^{s\lambda_t}\,\mathrm{d}W_t. \tag{3.3}$$

The filtered versions are obtained by applying (3.2). One obtains for $\widehat{\lambda}_t$ that

$$\mathrm{d}\widehat{\lambda}_t = -\alpha(\widehat{\lambda}_t - \mu_0)\,\mathrm{d}t + \left(\frac{\widehat{\lambda^2_{t-}}}{\widehat{\lambda}_{t-}} - \widehat{\lambda}_{t-}\right)\left(\mathrm{d}N_t - \widehat{\lambda}_t\,\mathrm{d}t\right), \tag{3.4}$$

and for $\widehat{f}(s,t)$, one finds

$$\begin{aligned}\mathrm{d}\widehat{f}(s,t) &= \left[\left(-\alpha s + \frac{1}{2}s^2\beta^2\right)\frac{\partial}{\partial s}\widehat{f}(s,t) + s\alpha\mu_0\widehat{f}(s,t)\right]\mathrm{d}t \\ &\quad + \left(\frac{(\partial/\partial s)\widehat{f}(s,t-)}{\widehat{\lambda}_{t-}} - \widehat{f}(s,t-)\right)\left(\mathrm{d}N_t - \widehat{\lambda}_t\,\mathrm{d}t\right).\end{aligned} \tag{3.5}$$

In general, filtering equations are very difficult, if possible at all, to solve explicitly, as the first equation involves terms with $\widehat{\lambda_t^2}$ and the second equation involves combinations of $\widehat{\lambda}_t$ and $\widehat{f}(s,t)$. In order to solve these equations, one should also have equations for $\widehat{\lambda_t^2}$, but this involves $\widehat{\lambda_t^3}$ and so on, assuming that they exist. So instead of trying to solve these equations directly, a different approach is considered in order to find an expression for $\widehat{f}(s,t)$.

3.2. Filtering by the Method of Probability of Reference

In order to solve the problem introduced above, the *filtering by the method of probability of reference* is considered; see [5, Chap. VI] or [3, Sect. 2]. In this approach, a second probability measure $\mathbb{P}_0$ and intensity process λ_t^0 are introduced, such that $N_t - \int_0^t \lambda_s^0\,\mathrm{d}s$ is a martingale with respect to $\mathcal{F}_t$ under $\mathbb{P}_0$. Corresponding to this change of measure, one has the likelihood ratio, or density process Λ, given by $\Lambda_t := \mathbb{E}_0[\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{P}_0}|\mathcal{F}_t]$ (where $\mathbb{E}_0$ denotes expectation under $\mathbb{P}_0$), which is a martingale under $\mathbb{P}_0$ by construction. Moreover, one has (see [5, Theoms. VI T2, VI T7, and III T9]).

$$\Lambda_t = 1 + \int_0^t \Lambda_{s-}\frac{\lambda_{s-} - \lambda_{s-}^0}{\lambda_{s-}^0}\left(\mathrm{d}N_s - \lambda_s^0\,\mathrm{d}s\right). \tag{3.6}$$

Conversely, one defines for given positive predictable processes λ_t and λ_t^0 the process

$$\Lambda_t = \exp\!\left(\int_0^t \log\frac{\lambda_s}{\lambda_s^0}\,\mathrm{d}N_s - \int_0^t(\lambda_s - \lambda_s^0)\,\mathrm{d}s\right)$$

is a martingale under $\mathbb{P}_0$ under the condition that $\mathbb{E}_0\Lambda_t = 1$ for all $t \geq 0$. This likelihood ratio turns out to be a useful tool to solve the filtering problem for $\widehat{f}(s,t)$.

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

It is known that under appropriate conditions (see, e.g., [5, Result VI R8] for the case $\lambda_t^0 \equiv 1$) that the filtered version of this likelihood ratio, $\widehat{\Lambda}_t := \mathbb{E}_0\left[\Lambda_t | \mathcal{F}_t^N\right]$, follows an equation similar to (3.6). One has

$$\widehat{\Lambda}_t = 1 + \int_0^t \widehat{\Lambda}_{s-} \frac{\widehat{\lambda}_{s-} - \widehat{\lambda}^0_{s-}}{\widehat{\lambda}^0_{s-}} \left(\mathrm{d}N_s - \widehat{\lambda}^0_s \,\mathrm{d}s\right).$$

To solve the filtering problem for $\widehat{f}(s,t)$, an auxiliary function $g(s,t)$ is introduced. It is defined by

$$g(s,t) := \widehat{f}(s,t)\widehat{\Lambda}_t \exp\left(-\int_0^t \widehat{\lambda}^0_u \,\mathrm{d}u\right). \tag{3.7}$$

The exponent is used in order to obtain a simpler SDE of $g(s,t)$. After a solution to this equation has been found, one can obtain $\widehat{f}(s,t)$ by

$$\widehat{f}(s,t) = \frac{g(s,t)}{g(0,t)}. \tag{3.8}$$

It is directly clear that the first and third component of $g(s,t)$ are positive, and from (3.10), it follows that also the second component is positive; thus, the division in (3.8) is well defined. The solution to the filtering problem is obtained as soon as an expression for $g(s,t)$ is found. In Proposition 3.2, an SDE is derived for $g(s,t)$ for the intensity following the CIR model.

Proposition 3.2: *Let $g(s,t)$ be given by (3.7), then one has, for $t \geq 0$,*

$$\mathrm{d}g(s,t) = \left[s\mu_0\alpha g(s,t) + \left(\frac{1}{2}s^2\beta^2 - s\alpha - 1\right)\frac{\partial}{\partial s} g(s,t)\right]\mathrm{d}t + \left[\left(\widehat{\lambda}^0_{t-}\right)^{-1}\frac{\partial}{\partial s} g(s,t-) - g(s,t-)\right]\mathrm{d}N_t. \tag{3.9}$$

Proof: As a first step in proving (3.9), one can rewrite the function $g(s,t)$. Denoting the consecutive jump times of N_t by T_n ($n \geq 0$, and putting $T_0 = 0$), an alternative expression for $\widehat{\Lambda}_t$ is given by

$$\widehat{\Lambda}_t = \prod_{T_n \leq t}\left(\frac{\widehat{\lambda}_{T_n-}}{\widehat{\lambda}^0_{T_n-}}\right)\exp\left(-\int_0^t \left(\widehat{\lambda}_u - \widehat{\lambda}^0_u\right)\mathrm{d}u\right), \tag{3.10}$$

which can be checked by a direct calculation. From this it is easy to see that

$$g(s,t) \overset{(3.7)}{=} \widehat{f}(s,t)\widehat{\Lambda}_t \exp\left(-\int_0^t \widehat{\lambda}^0_u \,\mathrm{d}u\right) = \widehat{f}(s,t)\prod_{T_n \leq t}\left(\frac{\widehat{\lambda}_{T_n-}}{\widehat{\lambda}^0_{T_n-}}\right)\exp\left(-\int_0^t \widehat{\lambda}_u \,\mathrm{d}u\right) =: \widehat{f}(s,t)\widehat{L}_t.$$

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

For $\widehat{L}_t$, one finds the SDE

$$\mathrm{d}\widehat{L}_t = \frac{\widehat{L}_{t-}\widehat{\lambda}_{t-}}{\widehat{\lambda}^0_{t-}}\left(\mathrm{d}N_t - \widehat{\lambda}^0_t\,\mathrm{d}t\right) - \widehat{L}_{t-}\,\mathrm{d}N_t.$$

The SDE in (3.9) follows from the product rule

$$\begin{aligned}\mathrm{d}g(s,t) =& \widehat{f}(s,t-)\,\mathrm{d}\widehat{L}_t + \widehat{L}_{t-}\,\mathrm{d}\widehat{f}(s,t) + \Delta\widehat{f}(s,t)\Delta\widehat{L}_t\\ =& \widehat{f}(s,t-)\left(\frac{\widehat{L}_{t-}\widehat{\lambda}_{t-}}{\widehat{\lambda}^0_{t-}}\left(\mathrm{d}N_t - \widehat{\lambda}^0_t\,\mathrm{d}t\right) - \widehat{L}_{t-}\,\mathrm{d}N_t\right)\\ &+ \widehat{L}_{t-}\Bigg(\left[\left(-\alpha s + \frac{1}{2}s^2\beta^2\right)\frac{\partial}{\partial s}\widehat{f}(s,t) + s\alpha\mu_0\widehat{f}(s,t)\right]\mathrm{d}t\\ &+ \left(\frac{(\partial/\partial s)\widehat{f}(s,t-)}{\widehat{\lambda}t-} - \widehat{f}(s,t-)\right)\left(\mathrm{d}N_t - \widehat{\lambda}_t\,\mathrm{d}t\right)\Bigg)\\ &+ \left(\frac{(\partial/\partial s)\widehat{f}(s,t-)}{\widehat{\lambda}_{t-}} - \widehat{f}(s,t-)\right)\left(\frac{\widehat{L}_{t-}\widehat{\lambda}_{t-}}{\widehat{\lambda}^0_{t-}} - \widehat{L}_{t-}\right)\mathrm{d}N_t.\end{aligned}$$

Collecting the terms before $\mathrm{d}t$ and $\mathrm{d}N_t$, one obtains

$$\begin{aligned}\mathrm{d}g(s,t) =& \Bigg(-\widehat{\lambda}_t\widehat{f}(s,t)\widehat{L}_t + \left(-\alpha s + \frac{1}{2}s^2\beta^2\right)\frac{\partial}{\partial s}\widehat{f}(s,t)\widehat{L}_t\\ &+ s\alpha\mu_0\widehat{f}(s,t)\widehat{L}_t - \frac{\partial}{\partial s}\widehat{f}(s,t)\widehat{L}_t + \widehat{f}(s,t)\widehat{L}_t\widehat{\lambda}_t\Bigg)\mathrm{d}t\\ &+ \Bigg(\frac{\widehat{f}(s,t-)\widehat{L}_{t-}\widehat{\lambda}_{t-}}{\widehat{\lambda}^0_{t-}} - \widehat{f}(s,t-)\widehat{L}_{t-}\\ &+ \frac{\widehat{L}_{t-}(\partial/\partial s)\widehat{f}(s,t-)}{\widehat{\lambda}_{t-}} - \widehat{L}_{t-}\widehat{f}(s,t-) + \frac{(\partial/\partial s)\widehat{f}(s,t-)\widehat{L}_{t-}}{\widehat{\lambda}^0_{t-}}\\ &- \frac{\widehat{f}(s,t-)\widehat{L}_{t-}\widehat{\lambda}_{t-}}{\widehat{\lambda}^0_{t-}} - \frac{(\partial/\partial s)\widehat{f}(s,t-)\widehat{L}_{t-}}{\widehat{\lambda}_{t-}} + \widehat{f}(s,t-)\widehat{L}_{t-}\Bigg)\mathrm{d}N_t.\end{aligned}$$

The result follows by simplifying the last equation. ■

The right-hand side of (3.9) depends only on $g(s,t)$ and its partial derivative with respect to s. In the next section this equation is solved between jumps, and in Section 5 the equation is solved at jump times of the process N_t.

4. FILTERING BETWEEN JUMPS

In the previous sections the filtering problem for point processes has been defined in general terms, and the problem has further been considered for an intensity following

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

the CIR model. To solve the filtering problem, one has to solve (3.9). This equation can be split up into a PDE between jumps of the process N_t and an equation at jumps. In this section the equation between jumps is solved for a general initial condition at time $T > 0$. Later, T will be considered as a jump time of N_t. Note that an initial condition for $g(s,t)$ is given as

$$g(s,T) = \widehat{f}(s,T)\widehat{\Lambda}_T \exp\left(-\int_0^T \widehat{\lambda}_u^0 \, \mathrm{d}u\right).$$

For $T = 0$, it follows that

$$g(s,0) = \widehat{f}(s,0) = \mathbb{E}\left[e^{s\lambda_0} \middle| \mathcal{F}_0^N\right] = \mathbb{E}\left[e^{s\lambda_0}\right],$$

which is the moment-generating function of the intensity at time $t = 0$, since $\mathcal{F}_0^N = \{\emptyset, \Omega\}$.

Before the solution to (3.9) is found, the specific case is considered, in which all of the parameters in the CIR model are set to zero. Albeit a simple example, the analysis of it sheds some light on the approach that will be followed for the general case.

Example 4.1: Consider the CIR model in which all the parameters are set to zero. This results in a constant intensity and, thus, $\mathrm{d}\lambda_t = 0$. The filter equations (3.4) and (3.5) reduce to

$$\mathrm{d}\widehat{\lambda}_t = \left(\frac{\widehat{\lambda^2_{t-}}}{\widehat{\lambda}_{t-}} - \widehat{\lambda}_{t-}\right)\left(\mathrm{d}N_t - \widehat{\lambda}_t \, \mathrm{d}t\right),$$

$$\mathrm{d}\widehat{f}_t = \left(\frac{\widehat{\lambda_{t-}f_{t-}}}{\widehat{\lambda}_{t-}} - \widehat{f}_{t-}\right)\left(\mathrm{d}N_t - \widehat{\lambda}_t \, \mathrm{d}t\right).$$

The PDE for $g(s,t)$ between jumps reduces to

$$\frac{\partial}{\partial t} g(s,t) = -\frac{\partial}{\partial s} g(s,t).$$

With an initial condition $g(s,T) = w(s)$, one easily finds that the solution to this equation is

$$g(s,t) = w(s - t + T).$$

In fact, what happens in this example is nothing else than computing the a posteriori mean and a posteriori moment-generating function of λ given the observations N_s, $s \le t$. We have thus considered a classical Bayesian problem.

In the next section this example is considered once more, where the filter at jump times is considered. We proceed with the case of an intensity following the CIR model.

PROPOSITION 4.2: *Let λ_t follow the CIR model (2.2) and let $g(s,t)$ be given by (3.7), with an initial condition at time T, $g(s,T) = w(s)$. Then, for $T \leq t < T_n$, with T_n the first jump time of N_t after T, $g(s,t)$, solves the PDE*

$$\frac{\partial}{\partial t} g(s,t) = s\mu_0\alpha g(s,t) + \frac{1}{2\rho}(\rho s - \alpha + \tau)(\rho s - \alpha - \tau)\frac{\partial}{\partial s} g(s,t), \tag{4.1}$$

where $\rho := \beta^2$ and $\tau := \sqrt{\alpha^2 + 2\beta^2}$. The unique solution to this equation is given by

$$g(s,t) = e^{\theta(\alpha-\tau)(t-T)} \left(\frac{2\tau}{\rho s(e^{-\tau(t-T)} - 1) + (\tau - \alpha)e^{-\tau(t-T)} + \tau + \alpha} \right)^{2\theta}$$
$$\times\, w\left(\frac{s\left((\alpha+\tau)e^{-\tau(t-T)} + \tau - \alpha\right) + 2e^{-\tau(t-T)} - 2}{\rho s(e^{-\tau(t-T)} - 1) + (\tau - \alpha)e^{-\tau(t-T)} + \tau + \alpha} \right), \tag{4.2}$$

where $\theta := \mu_0\alpha/\rho$.

PROOF: The PDE (4.2) for $g(s,t)$ follows directly from Proposition 3.2, as the jump part of this equation can be discarded.

To obtain a solution to this equation, a candidate solution is derived by making a number of transformations of the independent variables, until a simple PDE is found, which can be solved explicitly using known techniques. This candidate solution can then be checked to be the solution by calculating its partial derivatives and inserting these into (4.1).

The first transformation is given by

$$(s,t) \longrightarrow \left(\frac{\rho s - \alpha + \tau}{\rho s - \alpha - \tau}, t \right) =: (u,t). \tag{4.3}$$

Instead of $g(s,t)$, one writes $f_1(u,s)$, in terms of the new variable u. Using this transformation and the PDE for $g(s,t)$, one can derive a PDE for $f_1(u,t)$, by expressing s in terms of u and expressing the partial derivatives of $g(s,t)$ as partial derivatives of $f_1(u,t)$. The resulting PDE for $f_1(u,t)$ is

$$\frac{\partial}{\partial t} f_1(u,t) = \mu_0\alpha \left(\frac{\alpha}{\rho} + \frac{\tau(u+1)}{\rho(u+1)} \right) f_1(u,t) - \tau u \frac{\partial}{\partial u} f_1(u,t).$$

The second transformation that is used is given by

$$(u,t) \longrightarrow \left(\frac{\log(u)}{\tau}, t \right) =: (v,t),$$

where, for the time being, u is tacitly understood to be positive. Instead of the function $f_1(s,t)$, one considers the function $f_2(v,t) := f_1(u,t)$, in terms of the new variable v.

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

This transformation results in a PDE for $f_2(v,t)$:

$$\frac{\partial}{\partial t} f_2(v,t) = \mu_0\alpha\left(\frac{\alpha}{\rho} + \frac{\tau(e^{\tau v}+1)}{\rho(e^{\tau v}-1)}\right) f_2(v,t) - \frac{\partial}{\partial v} f_2(v,t).$$

The final transformation is given by

$$f_3(v,t) := \log(f_2(v,t)),$$

which results in the PDE for $f_3(v,t)$:

$$\frac{\partial}{\partial t} f_3(v,t) + \frac{\partial}{\partial v} f_3(v,t) = \mu_0\alpha\left(\frac{\alpha}{\rho} + \frac{\tau(e^{\tau v}+1)}{\rho(e^{\tau v}-1)}\right). \quad \textbf{(4.4)}$$

This equation can be solved using the method of characteristics, which is explained in [6, Chaps. 1 and 8], for example. Using this technique, the PDE is transformed in an ordinary differential equation (ODE) by introducing new variables $\xi(v,t)$ and $\zeta(v,t)$. The former is used to replace both v and t, and the latter is used to parameterize the initial curve. To be able to solve the PDE, an initial condition is required for $f_3(v,t)$. By applying all of the previous transformations to the initial condition $g(s,T) = w(s)$, with $t \geq T$, one obtains the initial condition

$$f_3(v,T) = \log\left(w\left(\frac{e^{\tau v}(\tau+\alpha)+\tau-\alpha}{\rho\,(e^{\tau v}-1)}\right)\right) =: G(v).$$

Next, one has to solve the differential equations

$$\frac{\partial}{\partial \xi} t(\xi,\zeta) = 1, \qquad \frac{\partial}{\partial \xi} v(\xi,\zeta) = 1,$$

with the initial conditions $t(0,\zeta) = T$ and $v(0,\zeta) = \zeta$. The unique solution to these equations is trivially given by

$$t(\xi,\zeta) = \xi + T, \qquad v(\xi,\zeta) = \xi + \zeta,$$

respectively. Inverting these expressions yields

$$\xi(v,t) = t - T, \qquad \zeta(v,t) = v - t + T,$$

respectivley. Using these transformations, the PDE (4.4) can be transformed into the ODE

$$\begin{aligned}\frac{\partial}{\partial \xi} f_3(\xi,\zeta) &= \mu_0\alpha\left(\frac{\alpha}{\rho} + \frac{\tau(e^{\tau(\xi+\zeta)}+1)}{\rho(e^{\tau(\xi+\zeta)}-1)}\right)\\ &= \frac{\mu_0\alpha(\alpha+\tau)}{\rho} + \frac{2\tau\mu_0\alpha}{\rho(e^{\tau(\xi+\zeta)}-1)}\\ &= \theta(\alpha+\tau) + \frac{2\tau\theta}{e^{\tau(\xi+\zeta)}-1},\end{aligned} \quad \textbf{(4.5)}$$

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

where $\theta = \mu_0\alpha/\rho$. This ODE can be solved for the given initial condition $f_3(v, T) = G(v)$. To derive the solution, one can start with a candidate solution

$$f_3(\xi,\zeta) = C_1 \log\left(e^{\tau(\xi+\zeta)} - 1\right) + C_2\xi + C_3.$$

For $\xi = 0$, one has $f_3(0,\zeta) = C_1 \log(e^{\tau\zeta} - 1) + C_3$ and f_3 has partial derivative with respect to ξ:

$$\frac{\partial}{\partial\xi} f_3(\xi,\zeta) = \tau C_1 + \frac{C_1\tau}{e^{\tau(\xi+\zeta)} - 1} + C_2.$$

Using the initial condition $f_3(0,\zeta) = G(\zeta)$, together with the ODE (4.5), one can find the values of $C_1,\ C_2$, and C_3:

$$\begin{aligned} C_1 &= 2\theta, \\ C_2 &= \theta(\alpha - \tau), \\ C_3 &= G(\zeta) - 2\theta \log(e^{\tau\zeta} - 1). \end{aligned}$$

This leads to the unique solution

$$f_3(\xi,\zeta) = \theta(\alpha-\tau)\xi + 2\theta\log(e^{\tau(\xi+\zeta)} - 1) + G(\zeta) - 2\theta\log(e^{\tau\zeta} - 1). \tag{4.6}$$

The proof of the uniqueness of this solution is postponed to the end of this proof.

Replacing ξ by $t - T$ and ζ by $v - t + T$ results in

$$\begin{aligned} f_3(v,t) = \theta(\alpha-\tau)(t-T) + 2\theta\log\left(\frac{e^{\tau v} - 1}{e^{\tau(v-t+T)} - 1}\right) \\ + \log\left(w\left(\frac{e^{\tau(v-t+T)}(\tau+\alpha) + \tau - \alpha}{\rho\left(e^{\tau(v-t+T)} - 1\right)}\right)\right). \end{aligned} \tag{4.7}$$

Next, one obtains a candidate solution for $g(s,t)$, by reversing all of the transformations. This gives

$$\begin{aligned} f_2(s,t) &= e^{\theta(\alpha-\tau)(t-T)}\left(\frac{e^{\tau v} - 1}{e^{\tau(v-t+T)} - 1}\right)^{2\theta} w\left(\frac{e^{\tau(v-t+T)}(\tau+\alpha) + \tau - \alpha}{\rho\left(e^{\tau(v-t+T)} - 1\right)}\right), \\ f_1(s,t) &= e^{\theta(\alpha-\tau)(t-T)}\left(\frac{u - 1}{ue^{-\tau(t-T)} - 1}\right)^{2\theta} w\left(\frac{ue^{-\tau(t-T)}(\tau+\alpha) + \tau - \alpha}{\rho\left(ue^{-\tau(t-T)} - 1\right)}\right). \end{aligned}$$

By performing the last substitution, (4.3), an expression for $g(s,t)$ is obtained. One has

$$g(s,t) = e^{\theta(\alpha-\tau)(t-T)}\left(\left(\frac{\rho s-\alpha+\tau}{\rho s-\alpha-\tau}-1\right)\left(\frac{\rho s-\alpha+\tau}{\rho s-\alpha-\tau}e^{-\tau(t-T)}-1\right)^{-1}\right)^{2\theta}$$
$$\times\, w\left(\left(\frac{\rho s-\alpha+\tau}{\rho s-\alpha-\tau}e^{-\tau(t-T)}(\tau+\alpha)+\tau-\alpha\right)\right.$$
$$\left.\times\left(\rho\left(\frac{\rho s-\alpha+\tau}{\rho s-\alpha-\tau}e^{-\tau(t-T)}-1\right)\right)^{-1}\right)$$
$$= e^{\theta(\alpha-\tau)(t-T)}\left(\frac{2\tau}{\rho s(e^{-\tau(t-T)}-1)+(\tau-\alpha)e^{-\tau(t-T)}+\tau+\alpha}\right)^{2\theta}$$
$$\times\, w\left(\frac{s\left((\alpha+\tau)e^{-\tau(t-T)}+\tau-\alpha\right)+2e^{-\tau(t-T)}-2}{\rho s(e^{-\tau(t-T)}-1)+(\tau-\alpha)e^{-\tau(t-T)}+\tau+\alpha}\right),$$

where it was used that $(\alpha+\tau)(\tau-\alpha) = 2\rho$. By inserting this candidate into (4.1), one can check that it indeed is the solution.

The last thing to proof is the uniqueness of the solution to (4.1). As all of the transformations are clearly one-to-one, the uniqueness of this solution should follow from the uniqueness of the solution to (4.5). It is easy to see that the solution to this equation is unique, as the difference of two possible solutions, with the same initial condition, has a zero derivative, which implies that the two solutions are in fact equal. ■

The result of Proposition 4.2 tells us that one can calculate $g(s,t)$, for $T \le t < T_n$, where T_n is the first jump time of N_t after T. In order to completely solve the filtering problem, one further has to solve (3.9) at jump times. This is the topic of the next section, in which a recursive solution will be obtained for the case in which λ_0 has a Gamma distribution.

5. FILTERING AT JUMP TIMES AND A GENERAL SOLUTION

In the previous section the filtering problem has been solved between jumps, for an arbitrary initial condition $w(s)$ for $g(s,t)$, at time $T > 0$. In this section the filtering problem is solved at jump times, first for Example 4.1 and after that for the case where the intensity follows the CIR model.

Example 5.1 (Example 4.1 continued): At jumps, one obtains from (3.9),

$$\Delta g(s,t) = \left(\frac{\partial/\partial s\, g(s,t-)}{\widehat{\lambda}^0_{t-}} - g(s,t-)\right)\Delta N_t.$$

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

From this identity it easily follows that at a jump time $T > 0$,

$$g(s, T) = \left(\widehat{\lambda}^0_{T-}\right)^{-1} \frac{\partial}{\partial s} g(s, T-). \tag{5.1}$$

Combining the results between jumps and at jumps, one can obtain the solution to

$$\mathrm{d}g(s,t) = -\frac{\partial}{\partial s} g(s,t)\,\mathrm{d}t + \left(\frac{(\partial/\partial s) g(s,t-)}{\widehat{\lambda}^0_{t-}} - g(s,t-)\right)\mathrm{d}N_t.$$

At each jump time T_n, one has to take the derivative of the function $g(s,t)$, and divide by $\lambda^0_{T_n-}$; the resulting function can then be used as initial condition for the interval $[T_n, T_{n+1})$. Using an initial condition $g(s,0) = w(s)$, one obtains the solution

$$g(s,t) = w^{(N_t)}(s-t) \prod_{n=1}^{N_t} \left(\widehat{\lambda}^0_{T_n-}\right)^{-1},$$

where $w^{(n)}(s)$ denotes the nth derivative of $w(s)$. The conditional moment-generating function is found from (3.8) and is given by

$$\widehat{f}(s,t) = \frac{g(s,t)}{g(0,t)} = \frac{w^{(N_t)}(s-t)}{w^{(N_t)}(-t)}.$$

If one assumes that $\lambda_0 \sim \Gamma(\alpha, \beta)$, one has

$$f(s,0) = \widehat{f}(s,0) = \left(\frac{\beta}{\beta - s}\right)^{\alpha}, \qquad \widehat{f}(s,t) = \left(\frac{\beta + t}{\beta + t - s}\right)^{\alpha + N_t}. \tag{5.2}$$

From this follows that at time $t > 0$, λ_t given $\mathcal{F}^N_t$ is distributed according to $\Gamma(\alpha + N_t, \beta + t)$. Further, $\widehat{\lambda_t}$ can easily be derived by a differentiation with respect to s:

$$\widehat{\lambda}_t = \frac{\partial}{\partial s}\widehat{f}(s,t)\Big|_{s=0} = \frac{\alpha + N_t}{\beta + t}.$$

We recognize here, in a slightly more general form (we condition on $\mathcal{F}^N_t$ and not just on N_t), a classical result on conjugate distributions from Bayesian analysis: If a random variable has a Poisson distribution with random parameter λ that a priori follows a Gamma distribution, the posterior distribution of λ also belongs to the Gamma family.

The solution in this example was easy to find, which could be expected, because λ_t is constant over time in this case. The general CIR model for the intensity is more complicated, but in the remainder of this section this problem is solved. At jumps, one has the same equation as in Example 5.1, which is already solved in (5.1). In Theorem 5.2, solution for $g(s,t)$ for the CIR model is given. Before this theorem is stated, some notation is introduced. Let $x, y \in \mathbb{R}$, $t \geq 0$ and put

$$A(x,t,y) := x\left((\tau-\alpha)e^{-\tau t}+\tau+\alpha\right)+2y\left(1-e^{-\tau t}\right), \tag{5.3}$$

$$B(s,t) := \rho s\left(e^{-\tau t}-1\right)+(\tau-\alpha)e^{-\tau t}+\tau+\alpha, \tag{5.4}$$

$$C(x,t,y) := y\left((\alpha+\tau)e^{-\tau t}+\tau-\alpha\right)+\rho x\left(1-e^{-\tau t}\right). \tag{5.5}$$

This notation allows us to write the general solution between jumps, (4.2), as

$$g(s,t) = e^{\theta(\alpha-\tau)(t-T)}\left(\frac{2\tau}{B(s,t-T)}\right)^{2\theta} w\left(\frac{C(-2/\rho, t-T, s)}{B(s,t-T)}\right). \tag{5.6}$$

Next, let $T_1, T_2, \ldots$ denote the jump times and let $T_0 = 0$. Then introduce the following notation:

$$\mathscr{A}(t,T_0) := A(\phi,t,1) \quad \text{for } 0 \leq t < T_1, \tag{5.7}$$

$$\mathscr{A}(t,T_n) := A(\mathscr{A}(T_n,T_{n-1}), t-T_n, \mathscr{C}(T_n,T_{n-1})) \quad \text{for } T_n \leq t < T_{n+1}, \tag{5.8}$$

$$\mathscr{C}(t,T_0) := C(\phi,t,1) \quad \text{for } 0 \leq t < T_1, \tag{5.9}$$

$$\mathscr{C}(t,T_n) := C(\mathscr{A}(T_n,T_{n-1}), t-T_n, \mathscr{C}(T_n,T_{n-1})) \quad \text{for } T_n \leq t < T_{n+1}. \tag{5.10}$$

With this notation, the main result of this article can be stated. A recursive solution to the filtering problem is obtained for the case where λ_0 has a Gamma distribution.

THEOREM 5.2: *Let $\lambda_0 \sim \Gamma(2\theta,\phi)$, for $\phi > 0$ and $\theta = \mu_0\alpha/\rho > 0$. Then one has*

$$\widehat{f_0}(s) = g(s,0) = \left(\frac{\phi}{\phi - s}\right)^{2\theta},$$

which is the moment-generating function of the $\Gamma(2\theta,\phi)$ distribution. With the notation introduced in (5.3)–(5.5) and (5.7)–(5.10), one further has, for $T_n \leq t < T_{n+1}$,

$$g(s,t) = K(t)p_n(s,t)\left(\frac{1}{\mathscr{A}(t,T_n) - s\mathscr{C}(t,T_n)}\right)^{2\theta+n}, \tag{5.11}$$

where $p_0(s,t) \equiv 1$, and for $n \geq 1$, $p_n(s,t)$ is a polynomial of degree n in s, which satisfies the recursion

$$\begin{aligned} p_n(s,t) = B^n(s,t-T_n)\Bigg[&p_{n-1}\left(\frac{C(-2/\rho,t-T_n,s)}{B(s,t-T_n)},T_n\right)(2\theta+n-1)\mathscr{C}(T_n,T_{n-1}) \\ &+ \partial_1\left(p_{n-1}\left(\frac{C(-2/\rho,t-T_n,s)}{B(s,t-T_n)},T_n\right)\right) \\ &\times\left(\mathscr{A}(T_n,T_{n-1}) - \frac{C(-2/\rho,t-T_n,s)}{B(s,t-T_n)}\mathscr{C}(T_n,T_{n-1})\right)\Bigg], \end{aligned} \tag{5.12}$$

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

where ∂_1 *denotes the derivative with respect to the first argument of* p_n *and*

$$K(t) = e^{\theta(\alpha-\tau)t}(2\tau\phi)^{2\theta} \prod_{m\geq 1,\ T_m\leq t} \left(\frac{(2\tau)^{2\theta}}{\widehat{\lambda}^0_{T_m-}} \right). \qquad \textbf{(5.13)}$$

In the proof of this theorem, the following lemma is used.

LEMMA 5.3: *With the notation from (*5.3*)–(*5.5*) and (*5.7*)–(*5.10*), the following relations hold for* $n \geq 1$ *and* $x, y \in \mathbb{R}$*:*

(i) $\mathscr{A}(T_n, T_n) = 2\tau\mathscr{A}(T_n, T_{n-1})$*;*
(ii) $\mathscr{C}(T_n, T_n) = 2\tau\mathscr{C}(T_n, T_{n-1})$*;*
(iii) $xB(s,t) - yC(-2/\rho, t, s) = A(x,t,y) - sC(x,t,y)$*.*

PROOF:

(i) From (5.8) and (5.3) it follows that

$$\begin{aligned} \mathscr{A}(T_n, T_n) &= A(\mathscr{A}(T_n, T_{n-1}), 0, \mathscr{C}(T_n, T_{n-1})) \\ &= \mathscr{A}(T_n, T_{n-1}) \left((\tau-\alpha)e^0 + \tau + \alpha\right) + \mathscr{C}(T_n, T_{n-1}) \left(1 - e^0\right) \\ &= 2\tau\mathscr{A}(T_n, T_{n-1}). \end{aligned}$$

(ii) This follows along the same lines as in (i), using (5.10) and (5.5).

(iii) Using (5.3)–(5.5) one finds

$$\begin{aligned} xB(s,t) - yC\left(-\frac{2}{\rho}, t, s\right) &= x\left(\rho s\left(e^{-\tau t} - 1\right) + (\tau-\alpha)e^{-\tau t} + \tau + \alpha\right) \\ &\quad - y\left(s\left((\alpha+\tau)e^{-\tau t} + \tau - \alpha\right) + 2\left(1 - e^{-\tau t}\right)\right) \\ &= x\left((\tau-\alpha)e^{-\tau t} + \tau + \alpha\right) + 2y\left(1 - e^{-\tau t}\right) \\ &\quad - s\left(y\left((\alpha+\tau)e^{-\tau t} + \tau - \alpha\right) + x\rho\left(1 - e^{-\tau t}\right)\right) \\ &= A(x,t,y) - sC(x,t,y). \end{aligned}$$

■

Now, Theorem 5.2 can be proved.

PROOF OF THEOREM 5.2: For each n it has to be shown that (5.11) holds at T_n and between T_n and T_{n+1}. First, this is shown for $n = 0$. Then the induction step is proved for $n \geq 1$.

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

$n = 0$: For $t = T_0 = 0$, one has, by assumption,

$$g(s,0) = \left(\frac{\phi}{\phi - s}\right)^{2\theta}.$$

From (5.11) one finds

$$\begin{aligned} g(s,0) &= K(0)p_0(s,0)\left(\frac{1}{\mathscr{A}(0,0) - s\mathscr{C}(0,0)}\right)^{2\theta} \\ &= e^0(2\tau\phi)^{2\theta}\left(\frac{1}{A(\phi,0,1) - sC(\phi,0,1)}\right)^{2\theta} \\ &= \left(\frac{2\tau\phi}{2\tau\phi - 2\tau s}\right)^{2\theta} = \left(\frac{\phi}{\phi - s}\right)^{2\theta}. \end{aligned}$$

Next, the interval up to the first jump time, $0 < t < T_1$, is considered. From (5.6) and the expression for $w(s) = g(s,0)$, one finds

$$\begin{aligned} g(s,t) &= e^{\theta(\alpha-\tau)t}\left(\frac{2\tau}{B(s,t)}\right)^{2\theta}\left(\frac{\phi}{\phi - C(-2/\rho,t,s)/B(s,t)}\right)^{2\theta} \\ &= e^{\theta(\alpha-\tau)t}(2\tau\phi)^{2\theta}\left(\frac{1}{B(s,t)\phi - C(-2/\rho,t,s)}\right)^{2\theta} \\ &= K(t)p_0(s,t)\left(\frac{1}{\mathscr{A}(t,0) - s\mathscr{C}(t,0)}\right)^{2\theta}, \end{aligned}$$

which is the same expression as in (5.11) for $n = 0$. The final step in the above derivation follows from Lemma 5.3 (iii), with $x = \phi$ and $y = 1$, together with the definition of $K(t)$ in (5.13).

$n \geq 1$: Now, it remains to prove the induction step. Therefore, one can assume that (5.11) holds for $n-1$. It then remains to show that the equation holds for n at T_n and between T_n and T_{n+1}. First, the jump is considered. Thus, one has to calculate the derivative of $g(s,t)$ with respect to s and take the left limit in $t = T_n$; further, the derivative is divided by $\widehat{\lambda}^0_{T_n-}$. By (5.1) one has

$$\begin{aligned} g(s,T_n) &= \left(\widehat{\lambda}^0_{T_n-}\right)^{-1}\frac{\partial}{\partial s}g(s,T_n-) \\ &= \left(\widehat{\lambda}^0_{T_n-}\right)^{-1}\frac{\partial}{\partial s}\Bigg(K(T_n-)p_{n-1}(s,T_n-) \\ &\qquad \times\left(\frac{1}{\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1})}\right)^{2\theta+n-1}\Bigg). \end{aligned} \quad \textbf{(5.14)}$$

Calculating the derivative with respect to s leads to

$$g(s,t) = \left(\widehat{\lambda}^0_{T_n-}\right)^{-1} K(T_n-) \times \left[p_{n-1}(s,T_n)(2\theta+n-1)\mathscr{C}(T_n,T_{n-1}) + \frac{\partial}{\partial s} p_{n-1}(s,T_n) \times (\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1}))\right] \times \left(\frac{1}{\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1})}\right)^{2\theta+n}. \quad \textbf{(5.15)}$$

From Lemma 5.3 parts (i) and (ii) follows that for the denominator in (5.15), one has

$$\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1}) = (2\tau)^{-1}\left(\mathscr{A}(T_n,T_n) - s\mathscr{C}(T_n,T_n)\right).$$

Hence, (5.15) can be written as

$$g(s,T_n) = \left(\widehat{\lambda}^0_{T_n-}\right)^{-1} K(T_n-)(2\tau)^{2\theta}(2\tau)^n \times \left[p_{n-1}(s,T_n)(2\theta+n-1)\mathscr{C}(T_n,T_{n-1}) + \frac{\partial}{\partial s} p_{n-1}(s,T_n) \times (\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1}))\right] \times \left(\frac{1}{\mathscr{A}(T_n,T_n) - s\mathscr{C}(T_n,T_n)}\right)^{2\theta+n}. \quad \textbf{(5.16)}$$

From (5.13), it is easy to see that $K(T_n) = K(T_n-)\left(\widehat{\lambda}^0_{T_n-}\right)^{-1}(2\tau)^{2\theta}$, and further, one has $2\tau = B(s,0) = B(s,T_n - T_n)$. From this follows that (5.16) can be written as

$$g(s,T_n) = K(T_n)B^n(s,T_n - T_n) \times \left[p_{n-1}(s,T_n)(2\theta+n-1)\mathscr{C}(T_n,T_{n-1}) + \frac{\partial}{\partial s} p_{n-1}(s,T_n) (\mathscr{A}(T_n,T_{n-1}) - s\mathscr{C}(T_n,T_{n-1}))\right] \times \left(\frac{1}{\mathscr{A}(T_n,T_n) - s\mathscr{C}(T_n,T_n)}\right)^{2\theta+n}.$$

This can be simplified further using the definition of $p_n(s,t)$ as given in (5.12), together with the identity $C(-2/\rho, 0, s) = \tau s$. This results in

$$g(s,T_n) = K(T_n)p_n(s,T_n)\left(\frac{1}{\mathscr{A}(T_n,T_n) - s\mathscr{C}(T_n,T_n)}\right)^{2\theta+n},$$

which is the required result at $t = T_n$. Finally, one has to check that (5.11) holds for $T_n < t < T_{n+1}$. For this one, can use the general solution (5.6) with initial condition $w(s) = g(s, T_n)$. One finds

$$g(s,t) = e^{\theta(\alpha-\tau)(t-T_n)} \left(\frac{2\tau}{B(s,t-T_n)} \right)^{2\theta} e^{\theta(\alpha-\tau)T_n} (2\tau\phi)^{2\theta} \times \prod_{m\geq 1,\ T_m \leq T_n} \left(\frac{(2\tau)^{2\theta}}{\widehat{\lambda}^0_{T_m-}} \right) p_n \left(\frac{C(-2/\rho, t-T_n, s)}{B(s,t-T_n)}, T_n \right) \times \left(\frac{1}{\mathscr{A}(T_n,T_n) - (C(-(2/\rho), t-T_n, s)/B(s,t-T_n))\mathscr{C}(T_n,T_n)} \right)^{2\theta+n}.$$

Simplifying this expression yields

$$g(s,t) = e^{\theta(\alpha-\tau)t} (2\tau\phi)^{2\theta} \left(\prod_{m=1}^{n} \left(\frac{(2\tau)^{2\theta}}{\widehat{\lambda}^0_{T_m-}} \right) \right) (2\tau)^{2\theta} \times B^n(s,t-T_n) p_n \left(\frac{C(-(2/\rho), t-T_n, s)}{B(s,t-T_n)}, T_n \right) \times \left(\frac{1}{2\tau B(s,t-T_n)\mathscr{A}(T_n,T_{n-1}) - 2\tau C(-(2/\rho), t-T_n, s)\mathscr{C}(T_n,T_{n-1})} \right)^{2\theta+n}.$$

An application of Lemma 5.3, with $x = \mathscr{A}(T_n, T_{n-1})$ and $y = \mathscr{C}(T_n, T_{n-1})$, and the definitions of $\mathscr{A}(t, T_n)$ and $\mathscr{C}(t, T_n)$ in (5.8) and (5.10) together with the definition of $K(t)$ results in

$$g(s,t) = K(t) \frac{1}{(2\tau)^n} B^n(s,t-T_n) p_n \left(\frac{C(-(2/\rho), t-T_n, s)}{B(s,t-T_n)}, T_n \right) \times \left(\frac{1}{\mathscr{A}(t,T_n) - s\mathscr{C}(t,T_n)} \right)^{2\theta+n}.$$

Next, with the definition of $p_n(s, T_n)$ from (5.12), evaluated in $t = T_n$, together with $C(x, 0, y) = 2\tau y$ and $B(s, 0) = 2\tau$, one rewrites this to

$$g(s,t) = K(t) \left(\frac{1}{\mathscr{A}(t,T_n) - s\mathscr{C}(t,T_n)} \right)^{2\theta+n} \frac{1}{(2\tau)^n} B^n(s,t-T_n) \times (2\tau)^n \left[p_{n-1} \left(\frac{C(-(2/\rho), t-T_n, s)}{B(s,t-T_n)}, T_n \right) (2\theta+n-1)\mathscr{C}(T_n,T_{n-1}) + \partial_1 \left(p_{n-1} \left(\frac{C(-(2/\rho), t-T_n, s)}{B(s,t-T_n)}, T_n \right) \right) \right.$$

$$\times\left(\mathscr{A}(T_n,T_{n-1})-\frac{C(-(2/\rho),t-T_n,s)}{B(s,t-T_n)}\mathscr{C}(T_n,T_{n-1})\right)\Bigg]$$

$$=K(t)p_n(s,t)\left(\frac{1}{\mathscr{A}(t,T_n)-s\mathscr{C}(t,T_n)}\right)^{2\theta+n}.$$

In the final step, the definition of $p_n(s,t)$ is used—this time evaluated in t—which concludes the proof of (5.11). From the definitions of $B(s,t)$ and $C(x,t,y)$, with $y=s$, which are both linear in s, follows that $p_n(s,t)$ is a polynomial of degree n in s. ■

This theorem provides a recursive solution to (3.9), for the case λ_0 is distributed according to a $\Gamma(2\theta,\phi)$ distribution. From (3.8) it already known that the conditional moment-generating function can easily be obtained from an expression for $g(s,t)$. Now that this has been found, the conditional moment-generating function $\widehat{f}(s,t)$ can be obtained easily.

Corollary 5.4: *Under the assumptions of Theorem* 5.2, *the conditional moment-generating function* $\widehat{f}(s,t)$, *for* $T_n\le t<T_{n+1}$, *can be expressed as*

$$\widehat{f}(s,t)=q_n(s,t)\left(\frac{Q(t,T_n)}{Q(t,T_n)-s}\right)^{2\theta+n}, \tag{5.17}$$

where

$$q_n(s,t)=\frac{p_n(s,t)}{p_n(0,t)}\quad and \quad Q(t,T_n)=\frac{\mathscr{A}(t,T_n)}{\mathscr{C}(t,T_n)}.$$

Here, $q_n(s,t)$ *is a polynomial of degree* n *in* s.

Proof: The result follows directly from (3.8), Theorem 5.2, and the definitions of q_n and Q. ■

With the derivation of the conditional moment-generating function, the filtering problem has been solved, and one is able to calculate conditional default probabilities using the results in Example 3.1. To conclude this section, it is observed that the conditional moment-generating function in (5.17) corresponds to a mixture of Gamma distributions.

Remark 5.5: Corollary 5.4 provides an expression for $\widehat{f}(s,t)$ that involves the polynomial $q_n(\cdot,t)$. Deriving an explicit expression for $q_n(s,t)=p_n(s,t)/p_n(0,t)$ for any $n\ge 0$ is quite complicated, but we can write

$$q_n(s,t)=\sum_{i=0}^{n}R_i^n(t)s^i,$$

where the coefficients $R_i^n(t)$ of the polynomial follow directly from the coefficients of the polynomial in s, $p_n(s,t)$, which, in turn, can be obtained using the recursion (5.12).

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

Next, one can consider $n+1$ independent random variables Γ_i, where $\Gamma_i \sim \Gamma(2\theta + n - i, Q(t, T_n))$, for $i = 0, 1, \dots, n$. Further, consider the discrete random variable M^n, independent of the Γ_i, which assumes the values $0, 1, \dots, n$, with probabilities $\pi_i^n(t)$, and define the random variable

$$X_t^n = \sum_{i=0}^{n} 1_{\{M=i\}} \Gamma_i.$$

The moment-generating function of X_t^n can easily be found, as Γ_i and M^n are independent, hence,

$$\begin{aligned} \mathbb{E}\left[e^{sX_t^n}\right] &= \sum_{i=0}^{n} \mathbb{E}\left[e^{s\Gamma_i} 1_{\{M=i\}}\right] \\ &= \sum_{i=0}^{n} \pi_i^n(t) \mathbb{E}\left[e^{s\Gamma_i}\right] = \sum_{i=0}^{n} \pi_i^n(t) \left(\frac{Q(t,T_n)}{Q(t,T_n) - s}\right)^{2\theta+n-i}. \end{aligned} \tag{5.18}$$

The goal is to show that by choosing the probabilities correctly, the moment-generating function of X_t^n equals the conditional moment-generation function $\widehat{f}(s,t)$. Therefore, (5.18) is first rewritten as

$$\mathbb{E}\left[e^{sX_t^n}\right] = \left(\frac{Q(t,T_n)}{Q(t,T_n) - s}\right)^{2\theta+n} \sum_{i=0}^{n} \pi_i^n(t) \left(\frac{Q(t,T_n) - s}{Q(t,T_n)}\right)^{i}.$$

To have that both moment-generating functions $\widehat{f}(s,t)$ and (5.18) are equal, it is required that

$$q_n(s,t) = \sum_{i=0}^{n} R_i^n(t) s^i = \sum_{i=0}^{n} \pi_i^n(t) \left(\frac{Q(t,T_n) - s}{Q(t,T_n)}\right)^{i}.$$

The right-hand side of this equation can be written as

$$\sum_{i=0}^{n} \pi_i^n(t) Q(t,T_n)^{-i} \sum_{j=0}^{i} \binom{i}{j} Q(t,T_n)^{i-j} s^j (-1)^j.$$

This equation can be turned into a polynomial in s by interchanging the summations, which leads to

$$\begin{aligned} &\sum_{j=0}^{n} \sum_{i=j}^{n} \binom{i}{j} \pi_i^n(t) Q(t,T_n)^{-j} s^j (-1)^j \\ &\quad = \sum_{j=0}^{n} s^j \left((-1)^j Q(t,T_n)^{-j} \sum_{i=j}^{n} \binom{i}{j} \pi_i^n(t) \right). \end{aligned}$$

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

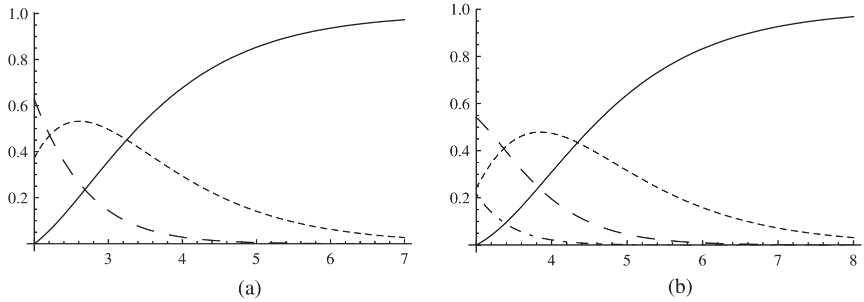

FIGURE 1. Graphs of the mixing probabilities after two jumps of the process N_t, (a), and after three jumps, (b). The values of the previous jump times, T_1 and T_2 in case (a), and T_1, T_2 and T_3 in case (b), are taken as $T_i = i$, such that one is able to calculate the $\pi_j^n(t)$. The model parameters are chosen to be $\alpha = 0.5$, $\beta = 0.5$, $\mu_0 = 0.4$, and $\phi = 4.0$.

The moment-generating functions are equal when

$$R_j^n(t) = (-1)^j Q(t, T_n)^{-j} \sum_{i=j}^{n} \binom{i}{j} \pi_i^n(t)$$

for $j = 0, 1, \ldots, n$. This can be solved iteratively, starting from $j = n$, which results in the probabilities

$$\pi_j^n(t) = (-1)^j R_j^n(t) Q(t, T_n)^j - \sum_{i=j+1}^{n} \pi_i^n(t) \binom{i}{j}. \tag{5.19}$$

It is not immediately clear from (5.19) that the $\pi_j^n(t)$ are all nonnegative and sum to 1. It turns out, however, that this is indeed the case for $T_n \leq t < T_{n+1}$, which means that the $\pi_j^n(t)$ can be interpreted as probabilities. It is, however, far from trivial to provide a general proof for all $n \geq 0$. We confine ourselves to illustrate this fact by some examples. In Figure 1, two graphs are given in which the probabilities are plotted.

6. REMARKS ON MODEL EXTENSIONS

We briefly discuss two ways of extending the model that we have considered in this article. First, we consider time varying parameters for the intensity. This is a more realistic assumption from a practical point of view. Next, we look at a simple multifactor specification of the intensity, where we see that the calculations for the one-dimensional case do not carry over to this multifactor case. In the situations in which explicit solutions cannot be obtained—which is the rule—one has to resort to numerical methods. One of the options is to approximate the state process by a finite-state Markov chain (or rather a sequence of them), for which the associated filtering

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

problem admits a finite-dimensional solution. This approach has been first developed for filtering with conditionally Gaussian observations, see [16]. For a recent financial application, see [14]. As an alternative, one could think of the use of particle filters, see [1] or [26] for a recent example in finance.

6.1. Time-Varying Parameters

In this subsection we briefly outline the consequences for our results when we replace the constant parameters in (2.2) with time-varying ones. Clearly, this introduces more flexibility of the model. So, we have $\alpha(t)$ instead of α, $\mu_0(t)$ instead of μ_0, and so for. Many results in Sections 2 and 3 remain valid upon substitution of the constants by their time-varying counterparts. In particular, (4.1) will change into

$$\frac{\partial}{\partial t}g(s,t) = s\mu_0(t)(\rho(t)s - \alpha(t) + \tau(t))(\rho(t)s - \alpha(t) - \tau(t))\frac{\partial}{\partial s}g(s,t). \qquad \textbf{(6.1)}$$

However, an explicit closed-form solution for $g(s,t)$ that we were able to give for the constant parameter case by (4.2) is, in general, impossible to obtain. The main reason for this is that transformation as given in (4.3) now introduces additional dependence on t and a simple PDE for $f_1(u,t)$ cannot be given. This complication carries over to similar ones for the functions $f_2(u,t)$ and $f_3(u,t)$.

If one uses piecewise constant functions for the parameters (as an approximation if needed), closed-form solutions are still possible, although they will be given by complex expressions. We briefly outline how to get these. Suppose that $0 < t_1, t_2, \ldots$ (with $t_i \to \infty$) denote the time instants where the parameters possibly change value. Consider a realization of the jump times $T_1, T_2, \ldots$. On each interval $[T_{n-1}, T_n)$ ($n \geq 1$), we relabel the t_i that fall in this interval by $\{t_1^n, \ldots, t_{k_n}^n\}$, which could be an empty set, in which case we can simply use (4.2) with the prevailing parameter values. Suppose now that this set is nonempty. On the subinterval $[T_{n-1}, t_1^n)$, we can compute the solution $g(s,t)$ to (6.1) again according to (4.2), eventually yielding $g(s, t_1^n-)$. Then we consider the PDE (6.1) on the interval $[t_1^n, t_2^n)$ with initial condition at t_1^n (instead of T) $w(s) = g(s, t_1^n-)$, and the values of the parameters on this interval. With the appropriate modifications, (4.2) can be used again. One then proceeds in this way until the final interval $[t_{k_n}^n, T_n)$ is reached, which eventually produces $g(s, T_n-)$.

We conclude by stating that more flexibility of the model by introducing time-varying, but piecewise constant parameter functions also leads to closed-form expressions, although they are more cumbersome to write down.

6.2. Multifactor Intensity

A second extension of the model that we have considered is to assume that the intensity is driven by more than one Brownian motion, or factor. To illustrate the difficulties that emerge in such an extension, we look at a very simple two-factor model for the intensity:

$$\lambda_t = \lambda_{1,t} + \lambda_{2,t},$$
$$\mathrm{d}\lambda_{i,t} = -\alpha_i \left(\lambda_{i,t} - \mu_i\right) \mathrm{d}t + \beta_i \sqrt{\lambda_{i,t}}\, \mathrm{d}W_{i,t}, \quad \text{for } i = 1, 2,$$

where W_1 and W_2 are independent Brownian motions and $\lambda_{1,t}$ and $\lambda_{2,t}$ both follow the CIR model with suitable parameter restrictions.

When we apply the filtering formulas (3.2) to λ_t we find

$$\mathrm{d}\widehat{\lambda_t} = \left(-\alpha_1 \left(\widehat{\lambda}_{1,t} - \mu_1\right) - \alpha_2 \left(\widehat{\lambda}_{2,t} - \mu_2\right)\right) \mathrm{d}t + \left(\frac{\widehat{\lambda^2_{t-}}}{\widehat{\lambda}_{t-}} - \widehat{\lambda}_{t-}\right) \left(\mathrm{d}N_t - \widehat{\lambda}_t\, \mathrm{d}t\right).$$

Just as in the one-dimensional case, this involves the term $\widehat{\lambda_t^2}$ and, thus, we again consider the conditional moment-generating function $\widehat{f}(s,t) = \mathbb{E}\left[e^{s\lambda_t} \,\middle|\, \mathcal{F}_t^N\right]$. Therefore, we have to determine the dynamics of $e^{s\lambda_t}$. An application of Itô's formula yields

$$\mathrm{d}e^{s\lambda_t} = \Bigg[\left(-\alpha_1 s + \frac{1}{2}s^2\beta_1^2\right)\lambda_{1,t}e^{s\lambda_t} + \left(-\alpha_2 s + \frac{1}{2}s^2\beta_2^2\right)\lambda_{2,t}e^{s\lambda_t}$$
$$+ s\left(\alpha_1\mu_1 + \alpha_2\mu_2\right)\Bigg]\mathrm{d}t + se^{s\lambda_t}\left(\beta_1\sqrt{\lambda_{1,t}}\,\mathrm{d}W_{1,t} + \beta_2\sqrt{\lambda_{2,t}}\,\mathrm{d}W_{2,t}\right). \quad \textbf{(6.2)}$$

Comparing the terms in the square brackets above with those in (3.3), we directly observe that we have lost an important feature. In the one-dimensional case, we could write the term $\lambda_t e^{s\lambda_t}$ as $\partial e^{s\lambda_t}/\partial s$, eventually resulting in the PDE that we could solve explicitly in Proposition 4.2. In the two-factor model, the derivative of $e^{s\lambda_t}$ with respect to s results in $\left(\lambda_{1,t} + \lambda_{2,t}\right) e^{s\lambda_t}$. The terms $\lambda_{i,t}e^{s\lambda_t}$ in (6.2) thus cannot be written as $\partial e^{s\lambda_t}/\partial s$. This shows that a solution, similar to that of Proposition 4.2, cannot be obtained.

Alternatively, one could consider the conditional moment-generating function of $(\lambda_{1,t}, \lambda_{2,t})$, given by $h(s_1, s_2, t) = \mathbb{E}\left[e^{s_1\lambda_{1,t}+s_2\lambda_{2,t}} \,\middle|\, \mathcal{F}_t^N\right]$, and derive its dynamics. As we introduce an additional variable, the eventual PDE will be of a higher dimension and thus more complex. Obtaining an explicit closed-form solution, if it exists, will be a substantially harder task and is beyond the scope of the present article. We conclude that it is far from straightforward to extend the explicit solution that we have obtained to models with more than one.

References

1. Bain, A. & Crisan, D. (2009). *Fundamentals of stochastic filtering*. New York: Springer.
2. Bielecki, T.R. & Rutkowski, M. (2002). *Credit risk: Modeling, valuation and hedging*. New York: Springer.
3. Boel, B.K. & Beneš, V.E. (1980). Recursive nonlinear estimation of a diffusion acting as the rate of an observed Poisson process. *IEEE Transactions on Information Theory* 26(5): 561–575.
4. Brémaud, P. & Jacod, J. (1977). Processus ponctuels et martingales: résultats récents sur la modélisation et le filtrage. *Advances in Applied Probability* 9(2): 362–416.

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press

5. Brémaud, P. (1981). *Point processes and queues*. New York: Springer-Verlag.
6. Chester, C.R. (1971). *Techniques in partial differential equations*. New York: McGraw-Hill.
7. Cox, J.C., Ingersoll, J.E., & Ross, S.A. (1985). A theory of the term structure of interest rates. *Econometrica* 53(2): 385–407.
8. Duffie, D. & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance* 6: 379–406.
9. Duffie, D. (2005). Credit risk modeling with affine processes. *Journal of Banking and Finance* 29: 2751–2802.
10. Elizalde, A. (2005). Credit risk models I: Default correlation in intensity models. Available from http://abelelizalde.com/pdf/survey1%20-%20intensity.pdf
11. Elizalde, A. (2005). Credit risk models II: Structural models. Available from http://abelelizalde.com/pdf/survey2%20-%20structural.pdf
12. Frey, R., Prosdocimi, C., & Runggaldier, W.J. (2007) Affine credit risk under incomplete information. In *Stochastic processes and applications to mathematical finance Proceedings of the 6th Ritsumeikan international symposium*, J. Akahori, S. Ogawa, S.Watanabe (eds.). Singapore: World Scientific, pp. 97–113.
13. Frey, R. & Runggaldier, W.J. (2010). Nonlinear filtering in models for interest-rate and credit risk. In D. Crisan, B. Rozovski (eds.) *Handbook on nonlinear filtering*. Oxford: Oxford University Press.
14. Frey, R. & Runggaldier, W.J. (2010). Pricing credit derivatives under incomplete information: a nonlinear-filtering approach. *Finance and Stochastics* 14: 495–526.
15. Giesecke, K. (2004). Credit risk modeling and valuation: An introduction. Working paper, Cornell University.
16. Kushner, H.J. & Dupuis, P. (2001). Numerical Methods for Stochastic Control Problems in Continuous Time. In *Stochastic Modelling and Applied Probability*. New York: Springer, vol. 24.
17. Lamberton, D. & Lapeyre, B. (1996). *Introduction to stochastic calculus applied to finance*. New York: Chapman & Hall.
18. Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives Research* 2: 99–120.
19. Luo, L.S.J. (2005). Bootstrapping default probability curves. *Journal of Credit Risk* 1: 169–179.
20. Rajan, A., McDermott, G., & Roy, R. (2007). *The structured credit handbook*. New York: Wiley.
21. Schönbucher, P.J. (2002). A tree implementation of a credit spread model for credit derivatives. *Journal of Computational Finance* 6(2): 1–38.
22. Schönbucher, P.J. (2003). *Credit derivatives pricing models*. New York: Wiley.
23. Schonbucher, P.J. (2003). *Information-driven default contagion*. Working paper, ETH Zürich.
24. Segall, A., Davis, M.H.A., & Kailath, T. (1975). Nonlinear filtering with counting observations. *IEEE Transactions on Information Theory*, 21: 143–149.
25. Van Schuppen, J.H. (1977). Filtering, prediction and smoothing for counting process observations, a martingale approach. *SIAM Journal of Applied*, 32(3): 552–570.
26. Xiong, J. & Zeng, Y. (2006). *Branching particle approximation to the filtering problem with counting process observations*, preprint.

https://doi.org/10.1017/S0269964811000076 Published online by Cambridge University Press