Modeling credit risk and credit derivatives

Leijdekker, V.J.G.

Citation for published version (APA):
A Filtering Problem with Applications to Credit Risk

In the first chapter of this thesis we have given a brief introduction into the modeling of credit risk and credit derivatives. In this chapter, that is based on [LS09], we present a model for the loss process of a credit risky portfolio. We model this loss process as a Cox process, where the default intensity is of the Cox-Ingersoll-Ross type. It is assumed that the Brownian motion driving the intensity process is not observed. With filtering using point process observations, we derive a recursive solution for the conditional moment generating function of the default counting process, given that we only observe the defaults.

2.1 Introduction

The main results of this chapter 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 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 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 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 [Sch03b] the author considers this model for the intensity. Further, [Duf05] considers more general affine models for credit risk. The filtering problem in this set up has previously been studied in [FPR07], 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 [BB80], 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 [BB80] 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 partial differential equation. Furthermore, we follow a different approach to obtain the recursive solution at jump times as compared to [FPR07]. By 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 chapter and refer to [RMR07] and [Sch03a] for detailed expositions. Overviews of the intensity-based modeling approach can be found in [Lan98], [Gie04] and [Eli05]. In this approach, it is a common assumption that the driving process can be observed, i.e. 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 chapter 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. In particular the intensity is assumed to follow the Cox-Ingersoll-Ross (CIR) model, where the driving Brownian motion is not observed.

Our results are obtained under the assumption that the CIR process follows a 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, since it allows for more flexible modeling. In general we will then loose the attractive feature of obtaining closed form solutions. But if one restricts the model by taking parameters which are piecewise constant functions, closed form solutions still exist. In practice these piecewise constant models have become popular in credit risk as its flexibility doesn’t destroy calibration procedures, see [Luo05].

The chapter is organized as follows: In Section 2.2 the Cox-Ingersoll-Ross model is discussed and some results for the case of full information are discussed. Next, in Section 2.3, the filtering problem is introduced and some background is given for filtering of point process observations. First, the filtering formulas from [Bré81] are given, and the equations for the conditional intensity and conditional moment generating function are derived. Then, in the second part of Section 2.3, we introduce filtering by the method of the probability of reference, and the filtering equations are transformed using the ideas introduced in [BB80]. Section 2.4 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 2.5, the filtering problem is solved at jump times, and an explicit, recursive solution is 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 2.6 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.

### 2.2 Model and Background

The main goal of the chapter 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 e.g. [SDK75], [Sch77] and [BJ77] and the later appearing book [Bré81]. Recently this kind of problem has gained renewed interest in the field of credit risk modeling, see also [FPR07], 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 e.g. [Gie04], [BR02] and [Eli05]. 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 case one considers more than one company, one can model the default times as consecutive jump times of the Cox Process. In [Lan98], [Gie04] and [Eli05] this modeling approach is discussed in more detail, and [Sch02] provides a detailed application. In this chapter we focus on the intensity-based approach, where the intensity $\lambda_t$ of the Cox process, a nonnegative process, has an affine structure, similar to interest term structure models [DK96]. This means that the intensity process $\lambda_t$ follows a stochastic differential equation (SDE) of the form:

$$d\lambda_t = (a + b\lambda_t)dt + \sqrt{c + d\lambda_t}dW_t,$$

(2.1)

for a Brownian Motion $W_t$, with $d > 0$. In particular, the focus is on the Cox-Ingersoll-Ross square root (CIR) model with mean reversion, [CIR85], for the intensity, where the intensity $\lambda_t$ satisfies

$$d\lambda_t = -\alpha(\lambda_t - \mu_0)dt + \beta\sqrt{\lambda_t}dW_t.$$

(2.2)

In [LL96, Section 6.2.2.] one finds parameter restrictions for this model which guarantee positivity of $\lambda_t$. Naturally one should start with a positive initial value $\lambda_0$, and if $\alpha\mu_0 \geq \beta^2/2$, then $\lambda_t$ remains strictly positive with probability one. Note that using the transformation $X_t = \lambda_t + c/d$ and by a reparametrization, $X_t$ satisfies the general SDE (2.1), and $\lambda_t$ 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 chapter.

A big advantage of the affine setup that we choose in this chapter, is that many relevant quantities in credit risk can be calculated explicitly. Using the formulas from [LL96, Section 6.2.2.] one can, for example, easily calculate the survival probability $\mathbb{P}(\tau > t | F_s)$, with $t > s$ and $F_t = F_t^N \vee F_t^W$, where the former filtration is generated by the point process $N_t$ and the latter by some process $Y_t$ driving the intensity process.

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

$$\mathbb{P}(\tau > t | F_s) = 1_{\{\tau > s\}} \mathbb{E} \left[ e^{-\int_s^t \lambda_u du} \big| F_s^W \right],$$

(2.3)

which follows from formulas in [BR02, Chapter 6]. Since $\lambda_t$ is a Markov process, one can condition on $\lambda_s$ instead of $F_s^W$. An application of Proposition 6.2.4. from [LL96]
to (2.3) yields
\[
\mathbb{P}(\tau > t|\mathcal{F}_s) = 1_{\{\tau > s\}} \exp \left( -\alpha \mu_0 \varphi(t-s) - \lambda s \psi(t-s) \right),
\]
where
\[
\varphi(t) = -\frac{2}{\beta^2} \log \left( \frac{2 \gamma e^{t(\gamma+\alpha)/2}}{\gamma - \alpha + e^{t\gamma(\gamma+\alpha)}} \right),
\]
\[
\psi(t) = \frac{2(e^{\gamma t} - 1)}{\gamma - \alpha + e^{t\gamma(\gamma+\alpha)}},
\]
\[
\gamma = \sqrt{\alpha^2 + 2\beta^2}.\]

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 [FPR07] 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 chapter 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.

### 2.3 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 \(\mathcal{F}_t^Y \subseteq \mathcal{F}_t\). One is then interested in conditional expectations of the form \(\hat{Z}_t = \mathbb{E}[Z_t|\mathcal{F}_t^Y]\), and one tries to find the dynamics of the process \(\hat{Z}_t\), 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 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.

### 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 jump times \(T_n\), and with \(\mathcal{F}_t\)-intensity \(\lambda_t\). The process \(Z_t\) is assumed
to follow the SDE
\[ dZ_t = a_t \, dt + dM_t, \tag{2.5} \]
for an \( \mathcal{F}_t \)-progressive measurable \( a_t \), with \( \int_0^t |a_s| \, ds < \infty \), and an \( \mathcal{F}_t \)-local martingale \( M_t \). The filtering problem is often cast as the calculation of the conditional expectation
\[ E[Z_t|\mathcal{F}_N_t] =: \hat{Z}_t. \]
Using the filtering formulas from [Bré81, Chapter IV], a representation of the solution to this filtering problem can be found. In case the (local) martingale \( M_t \) and the observed point process have no jumps in common, one has:
\[ d\hat{Z}_t = \hat{a}_t \, dt + \left( \frac{\hat{Z}_{\lambda t^-}}{\hat{\lambda}_t^-} - \hat{Z}_{t^-} \right) \left( dN_t - \hat{\lambda}_t d\lambda_t \right), \tag{2.6} \]
with \( \hat{a}_t := E[a_t|\mathcal{F}_N_t] \), and \( X_{t^-} := \lim_{s \uparrow t} X_s \).

**Example 2.2** (Example 2.1 continued). When one wants to calculate the survival probability given \( \mathcal{F}_N_t \), 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
\[ P(\tau > t|\mathcal{F}_N_s) = \frac{1}{\hat{\lambda}_t} \exp\left(-\alpha \mu_0 \psi(t-s)\right), \]
which can be calculated if an expression for the conditional moment generation function
\[ \hat{f}(s, t) := E[e^{s\lambda t}|\mathcal{F}_N_t] \]
is available.

The above example illustrates that one can calculate the survival probability if the conditional moment generating function \( \hat{f}(s, t) \) is known. As a first step in the determination of this function, the SDEs of \( \hat{\lambda}_t := E[\lambda_t|\mathcal{F}_N_t] \) and \( \hat{f}(s, t) \) are determined. First Itô’s formula is used to obtain the SDE for \( e^{s\lambda_t} \), where \( \lambda_t \) satisfies (2.2)
\[ de^{s\lambda_t} = \left[ \left( -\alpha s + \frac{1}{2} s^2 \beta^2 \right) \frac{\partial}{\partial s} e^{s\lambda_t} + s\alpha \mu_0 e^{s\lambda_t} \right] \, dt + \beta \sqrt{\lambda_t} e^{s\lambda_t} \, dW_t. \tag{2.7} \]
The filtered versions are obtained by applying formula (2.6). One obtains for \( \hat{\lambda}_t \)
\[ d\hat{\lambda}_t = -\alpha(\hat{\lambda}_t - \mu_0) \, dt + \left( \frac{\hat{\lambda}_t^2}{\hat{\lambda}_t^-} - \hat{\lambda}_t^- \right) \left( dN_t - \hat{\lambda}_t d\lambda_t \right), \tag{2.8} \]
and for \( \hat{f}(s, t) \) one finds
\[ d\hat{f}(s, t) = \left[ \left( -\alpha s + \frac{1}{2} s^2 \beta^2 \right) \frac{\partial}{\partial s} \hat{f}(s, t) + s\alpha \mu_0 \hat{f}(s, t) \right] \, dt + \left( \frac{\partial}{\partial s} \hat{f}(s, t^-) \right) \left( dN_t - \hat{\lambda}_t d\lambda_t \right). \tag{2.9} \]
In general, filtering equations are very difficult, if possible at all, to solve explicitly, since the first equation involves terms with $\hat{\lambda}_t^2$ and the second equation involves combinations of $\hat{\lambda}_t$ and $\hat{f}(s,t)$. In order to solve these equations one should also have equations for $\hat{\lambda}_t^2$, but this involves $\hat{\lambda}_t^3$ 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 $\hat{f}(s,t)$.

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 [Bré81, chapter VI] or [BB80, Section 2]. In this approach a second probability measure $\mathbb{P}_0$ and intensity process $\lambda_0^t$ are introduced, such that $N_t - \int_0^t \lambda_s^0 ds$ 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 $\Lambda$, given by

$$\Lambda_t := \mathbb{E} \left[ \frac{d\mathbb{P}}{d\mathbb{P}_0} \bigg| \mathcal{F}_t \right] = 1 + \int_0^t \Lambda_s - \frac{\lambda_s - \lambda_0^0}{\lambda_s - \lambda_0^0} \left( dN_s - \lambda_0^0 ds \right).$$

(2.10)

This likelihood ratio turns out to be a useful tool to solve the filtering problem for $\hat{f}(s,t)$. It is known, see e.g. [Bré81] for the case $\lambda_0^t \equiv 1$, that the filtered version of this likelihood ratio, $\hat{\Lambda}_t := \mathbb{E} \left[ \Lambda_t \bigg| \mathcal{F}_N^t \right]$ follows an equation similar to (2.10). One has

$$\hat{\Lambda}_t = 1 + \int_0^t \hat{\Lambda}_s - \frac{\hat{\lambda}_s - \hat{\lambda}_0^0}{\hat{\lambda}_s - \hat{\lambda}_0^0} \left( dN_s - \hat{\lambda}_0^0 ds \right)$$

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

$$g(s,t) := \hat{f}(s,t)\hat{\Lambda}_t \exp \left( - \int_0^t \hat{\lambda}_u^0 du \right).$$

(2.11)

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 $\hat{f}(s,t)$ by

$$\hat{f}(s,t) = \frac{g(s,t)}{g(0,t)}.$$

(2.12)

It is directly clear that the first and third component of $g(s,t)$ are positive, and from (2.14) follows that also the second component is positive, and thus the division in (2.12) is well defined. The solution to the filtering problem is obtained as soon as an expression for $g(s,t)$ is found. In Proposition 2.3 an SDE is derived for $g(s,t)$ for the intensity following the CIR model.
Proposition 2.3. Let \( g(s, t) \) be given by (2.11), then one has, for \( t \geq 0 \)
\[
\begin{align*}
\frac{dg(s, t)}{dt} &= \left[ s \mu_0 \alpha g(s, t) + \left( \frac{1}{2} s^2 \beta^2 - \alpha \right) \frac{\partial}{\partial s} g(s, t) \right] \, dt \\
&\quad + \left[ \left( \lambda^0_t \right)^{-1} \frac{\partial}{\partial s} g(s, t - ) - g(s, t - ) \right] \, dN_t.
\end{align*}
\] (2.13)

Proof. As a first step in proving (2.13), one can rewrite the function \( g(s, t) \). Denoting the jump times of \( N_t \) as \( T_n \), an alternative expression for \( \hat{\Lambda}_t \) is given by
\[
\hat{\Lambda}_t = \prod_{T_n \leq t} \left( \frac{\hat{\lambda}_{T_n -}}{\hat{\lambda}^0_{T_n -}} \right) \exp \left( - \int_0^t \left( \hat{\lambda}_u - \hat{\lambda}^0_u \right) \, du \right),
\] (2.14)
which can be checked by a direct calculation. From this it is easy to see that
\[
g(s, t) = \hat{f}(s, t) \hat{\Lambda}_t \exp \left( - \int_0^t \hat{\lambda}_u \, du \right)
= \hat{f}(s, t) \prod_{T_n \leq t} \left( \frac{\hat{\lambda}_{T_n -}}{\hat{\lambda}^0_{T_n -}} \right) \exp \left( - \int_0^t \hat{\lambda}_u \, du \right) =: \hat{f}(s, t) \hat{L}_t.
\]

For \( \hat{L}_t \) one finds the SDE
\[
d\hat{L}_t = \frac{\hat{L}_t - \hat{\lambda}_t - \hat{\lambda}^0_t}{\lambda^0_t} \left( dN_t - \hat{\lambda}^0_t \, dt \right) - \hat{L}_t - dN_t.
\]
The SDE in (2.13) follows from the product rule
\[
\begin{align*}
\frac{dg(s, t)}{dt} &= \hat{f}(s, t - ) d\hat{L}_t + \hat{L}_t - d\hat{f}(s, t) + \Delta \hat{f}(s, t) \Delta \hat{L}_t \\
&= \hat{f}(s, t - ) \left( \frac{\hat{L}_t - \hat{\lambda}_t - \hat{\lambda}^0_t}{\lambda^0_t} \left( dN_t - \hat{\lambda}^0_t \, dt \right) - \hat{L}_t - dN_t \right) \\
&\quad + \hat{L}_t - \left( \left( - \alpha s + \frac{1}{2} s^2 \beta^2 \right) \frac{\partial}{\partial s} \hat{f}(s, t) + s \alpha \mu_0 \hat{f}(s, t) \right) \, dt \\
&\quad + \left( \frac{\partial}{\partial s} \hat{f}(s, t - ) \frac{\lambda^0_t}{\lambda^0_t} - \hat{f}(s, t - ) \right) \left( dN_t - \hat{\lambda}_t \, dt \right) \\
&\quad + \left( \frac{\partial}{\partial s} \hat{f}(s, t - ) - \hat{f}(s, t - ) \right) \left( \frac{\hat{L}_t - \hat{\lambda}_t - \hat{\lambda}^0_t}{\lambda^0_t} - \hat{L}_t - \hat{\lambda}^0_t \right) \, dN_t.
\end{align*}
\]
Collecting the terms before \( dt \) and \( dN_t \), one obtains the equation

\[
dg(s,t) = \left( -\hat{\lambda}_t \hat{f}(s,t) \hat{L}_t + \left( -\alpha s + \frac{1}{2} s^2 \beta^2 \right) \frac{\partial}{\partial s} \hat{f}(s,t) \hat{L}_t \right) dt
\]

\[
+ s \alpha \mu_0 \hat{f}(s,t) \hat{L}_t - \frac{\partial}{\partial s} \hat{f}(s,t) \hat{L}_t + \hat{f}(s,t) \hat{L}_t \hat{\lambda}_t \right) dt
\]

\[
+ \left( \hat{f}(s,t-) \frac{\hat{L}_t - \hat{\lambda}_t}{\lambda^0_{t-}} - \hat{f}(s,t-) \hat{L}_t \right.
\]

\[
+ \frac{\hat{L}_t - \hat{\lambda}_t}{\lambda^0_{t-}} \left. - \hat{f}(s,t-) \frac{\hat{L}_t - \hat{\lambda}_t}{\lambda^0_{t-}} + \hat{f}(s,t-) \hat{L}_t \right)
\]

\[
\] \( dN_t. \)

The result follows by simplifying the last equation.

The right hand side of (2.13) 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 2.5 the equation is solved at jump times of the process \( N_t \).

### 2.4 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 the Cox-Ingersoll-Ross model. To solve the filtering problem, one has to solve equation (2.13). This equation can be split up into a partial differential equation 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 on \( T \) will be considered as a jump time of \( N_t \). Note that an initial condition for \( g(s,t) \) is given as

\[
g(s,T) = \hat{f}(s,T) \hat{\Lambda}_T \exp \left( - \int_0^T \hat{\lambda}_0^0 du \right)
\]

For \( T = 0 \) it follows that

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

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

Before the solution to (2.13) is found, the specific case is considered, in which all 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 2.4. Consider the CIR model in which all the parameters are set to zero. This results in a constant intensity, and thus \( d\lambda_t = 0 \). The filter equations (2.8) and (2.9) reduce to
\[
\begin{align*}
\dot{\lambda}_t &= \left( \frac{\lambda^2_t}{\lambda_t} - \lambda_t \right) (dN_t - \lambda_t dt) \\
\dot{f}_t &= \left( \frac{\lambda_t f_t}{\lambda_t} - f_t \right) (dN_t - \lambda_t dt).
\end{align*}
\]

The partial differential equation 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
\[
g(s,t) = w(s - t + T).
\]

\[\diamond\]

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 2.5. Let \( \lambda_t \) follow the Cox-Ingersoll-Ross model (2.2), and let \( g(s,t) \) be given by (2.11), 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 partial differential equation
\[
\frac{\partial}{\partial t} g(s,t) = s\mu_0 \alpha g(s,t) + \frac{1}{2\rho} (\rho s - \alpha + \tau)(\rho s - \alpha - \tau) \frac{\partial}{\partial s} g(s,t),
\]
where \( \rho := \beta^2 \) and \( \tau := \sqrt{\alpha^2 + 2\beta^2} \). The unique solution to this equation is given by
\[
g(s,t) = e^{\theta(\alpha - \tau)(t-T)} \left( \frac{2\tau}{\rho s(e^{-\tau(t-T)} - 1) + (\tau - \alpha)e^{-\tau(t-T)} + \tau + \alpha} \right)^{2\theta} \\
\times w \left( \frac{s((\alpha + \tau)e^{-\tau(t-T)} + \tau - \alpha) + 2e^{-\tau(t-T)} - 2}{\rho s(e^{-\tau(t-T)} - 1) + (\tau - \alpha)e^{-\tau(t-T)} + \tau + \alpha} \right),
\]
where \( \theta := \frac{\mu_0 \alpha}{\rho} \).

Proof. The partial differential equation (2.16) for \( g(s,t) \) follows directly from Proposition 2.3, since 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 (2.15).

The first transformation is given by

\[(s, t) \rightarrow \left(\frac{\rho s - \alpha + \tau}{\rho s - \alpha - \tau}, t\right) =: (u, t).\] (2.17)

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

\[(u, t) \rightarrow \left(\frac{\log(u)}{\tau}, t\right) =: (v, t),\]

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 \(v\). This transformation results in a partial differential equation 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 \(f_3(v, t)\):

\[\frac{\partial}{\partial t} f_3(v, t) + \frac{\partial}{\partial v} f_3(v, t) = \mu_0 \alpha \left(\frac{\alpha}{\rho} + \frac{\tau(e^{\tau v} + 1)}{\rho(e^{\tau v} - 1)}\right).\] (2.18)

This equation can be solved using the method of characteristics, which is explained in chapter 1 and 8 of [Che71], for example. Using this technique the partial differential equation is transformed in an ordinary differential equation 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 the previous transformations to the initial condition \(g(s, T) = w(s)\), with \(t \geq T\), one obtains the initial condition

\[f_3(v, T) = \log \left(\frac{e^{\tau v}(\tau + \alpha) + \tau - \alpha}{\rho(e^{\tau v} - 1)}\right) =: G(v).\]
Next one has to solve the differential equations

\[ \frac{\partial}{\partial \xi} t(\xi, \zeta) = 1, \quad \frac{\partial}{\partial \xi} v(\xi, \zeta) = 1, \]

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

\[ t(\xi, \zeta) = \xi + T, \quad v(\xi, \zeta) = \xi + \zeta. \]

Inverting these expressions, yields

\[ \xi(v, t) = t - T, \quad \zeta(v, t) = v - t + T. \]

Using these transformations, the partial differential equation (2.18) can be transformed into the ordinary differential equation (ODE)

\[ \frac{\partial}{\partial \xi} f_3(\xi, \zeta) = \mu_0 \alpha \left( \frac{\alpha}{\rho} + \frac{\tau(e^{\tau(\xi+\zeta)} + 1)}{\rho(e^{\tau(\xi+\zeta)} - 1)} \right) = \frac{\mu_0 \alpha(\alpha + \tau)}{\rho} + \frac{2\tau \mu_0 \alpha}{\rho(e^{\tau(\xi+\zeta)} - 1)} = \theta(\alpha + \tau) + \frac{2\tau \theta}{e^{\tau(\xi+\zeta)} - 1}, \tag{2.19} \]

where \( \theta = \frac{\mu_0 \alpha}{\rho} \). This ordinary differential equation 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 \( f_3(0, \zeta) = C_1 \log \left( e^{\tau \zeta} - 1 \right) + C_3 \), and \( f_3 \) has partial derivative with respect to \( \xi \):

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

Using the initial condition \( f_3(0, \zeta) = G(\zeta) \), together with the ODE (2.19), one can find the values of \( C_1, C_2 \) and \( C_3 \):

\[ C_1 = 2\theta, \]
\[ C_2 = \theta(\alpha - \tau), \]
\[ C_3 = G(\zeta) - 2\theta \log \left( e^{\tau \zeta} - 1 \right). \]

This leads to the unique solution

\[ f_3(\xi, \zeta) = \theta(\alpha - \tau) \xi + 2\theta \log \left( e^{\tau(\xi+\zeta)} - 1 \right) + G(\zeta) - 2\theta \log \left( e^{\tau \zeta} - 1 \right). \tag{2.20} \]

The proof of the uniqueness of this solution is postponed to the end of this proof.
Reversing substitution (2.17) results in the expression for \( g(s, t) \) obtained. One

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

where it was used that \((\alpha + \tau)(\tau - \alpha) = 2\rho\). By inserting this candidate into equation
\[
(2.15), \text{one can check that it indeed is the solution.}
\]

The last thing to prove is the uniqueness of the solution to equation (2.15). As all
the transformations are clearly one-to-one, the uniqueness of this solution should
follow from the uniqueness of the solution to equation (2.19). 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 zero derivative, which implies that the two solutions
are in fact equal.
2.5 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 2.4, and after that for the case where the intensity follows the CIR model.

Example 2.6 (Example 2.4 (continued)). At jumps one obtains from equation (2.13)

\[
\Delta g(s, t) = \left( \frac{\partial}{\partial s} g(s, t) - g(s, t) \right) \Delta N_t.
\]

From this identity it easily follows that at a jump time \( T > 0 \):

\[
g(s, T) = \left( \hat{\lambda}_T^0 \right)^{-1} \frac{\partial}{\partial s} g(s, T). \tag{2.22}
\]

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

\[
dg(s, t) = \frac{\partial}{\partial s} g(s, t) dt + \left( \frac{\partial}{\partial s} g(s, t) - g(s, t) \right) dN_t.
\]

At each jump time \( T_n \), one has to take the derivative of the function \( g(s, t) \), and divide by \( \lambda_{T_n}^0 \); 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

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

where \( w^{(n)}(s) \) denotes the \( n \)-th derivative of \( w(s) \). The conditional moment generating function is found from (2.12), and is given by

\[
\hat{f}(s, t) = \frac{g(s, t)}{g(0, t)} = \frac{w^{(N_t)}(s - t)}{w^{(N_t)}(-t)}.
\]

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

\[
f(s, 0) = \hat{f}(s, 0) = \left( \frac{\beta}{\beta - s} \right)^\alpha, \quad \hat{f}(s, t) = \left( \frac{\beta + t}{\beta + t - s} \right)^{\alpha + N_t}. \tag{2.23}
\]

From this follows that at time \( t > 0 \), \( \lambda_t \) given \( F_t^N \) is distributed as \( \Gamma(\alpha + N_t, \beta + t) \). Further \( \lambda_t \) can easily be derived by a differentiation with respect to \( s \):

\[
\hat{\lambda}_t = \left. \frac{\partial}{\partial s} \hat{f}(s, t) \right|_{s=0} = \frac{\alpha + N_t}{\beta + t}.
\]

\[\Diamond\]
The solution in this example was easy to find, which could be expected, since $\lambda_t$ is constant over time in this case. The general Cox-Ingersoll-Ross model for the intensity is more complicated, but in the remainder of this section, also this problem is solved. At jumps one has the same equation as in Example 2.6, which is already solved in (2.22). In Theorem 2.7 the 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

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

(2.24)

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

(2.25)

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

(2.26)

This notation allows us to write the general solution between jumps, (2.16), as

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

(2.27)

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

$$A(t, T_0) := A(\varphi, t, 1) \quad \text{for } 0 \leq t < T_1, \quad (2.28)$$

$$A(t, T_n) := A(A(T_n, T_{n-1}), t - T_n, C(T_n, T_{n-1})) \quad \text{for } T_n \leq t < T_{n+1}, \quad (2.29)$$

$$C(t, T_0) := C(\varphi, t, 1) \quad \text{for } 0 \leq t < T_1, \quad (2.30)$$

$$C(t, T_n) := C(A(T_n, T_{n-1}), t - T_n, C(T_n, T_{n-1})) \quad \text{for } T_n \leq t < T_{n+1}. \quad (2.31)$$

With this notation, the main result of this chapter can be stated. A recursive solution to the filtering problem is obtained, for the case where $\lambda_0$ has a Gamma distribution.

**Theorem 2.7.** Let $\lambda_0 \sim \Gamma(2\theta, \varphi)$, for $\varphi > 0$ and $\theta = \frac{\nu_0 \alpha}{\rho} > 0$. Then one has

$$\hat{f}_0(s) = g(s, 0) = \left( \frac{\varphi}{\varphi - s} \right)^{2\theta},$$

which is the moment generating function of the $\Gamma(2\theta, \varphi)$ distribution. With the notation introduced in (2.24)-(2.26) and (2.28)-(2.31) one further has, for $T_n \leq t < T_{n+1},$

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

(2.32)
where \( p_0(s, t) \equiv 1 \), and for \( n \geq 1 \), \( p_n(s, t) \) is a polynomial of degree \( n \) in \( s \), that satisfies the recursion,

\[
p_n(s, t) = B^n(s, t - T_n) \left[ p_{n-1} \left( \frac{C\left(\frac{-2}{\rho}, t - T_n, s\right)}{B(s, t - T_n)} \right), T_n \right) (2\theta + n - 1) \mathcal{C}(T_n, T_{n-1}) \\
+ \partial_1 \left( p_{n-1} \left( \frac{C\left(\frac{-2}{\rho}, t - T_n, s\right)}{B(s, t - T_n)} \right), T_n \right) \\
\times \left( \mathcal{A}(T_n, T_{n-1}) - \frac{C\left(\frac{-2}{\rho}, t - T_n, s\right)}{B(s, t - T_n)} \mathcal{C}(T_n, T_{n-1}) \right),
\]

(2.33)

where \( \partial_1 \) denotes the derivative with respect to the first argument of \( p_n \), and

\[
K(t) = e^{\theta(\alpha - \tau)t} (2\tau \varphi)^{2\theta} \prod_{m \geq 1, T_m \leq t} \left( \frac{(2\tau)^{2\theta}}{\lambda_{T_m}^0} \right).
\]

(2.34)

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

**Lemma 2.8.** With the notation from (2.24)-(2.26) and (2.28)-(2.31), the following relations hold for \( n \geq 1 \) and \( x, y \in \mathbb{R} \):

(i) \( \mathcal{A}(T_n, T_n) = 2\tau \mathcal{A}(T_n, T_{n-1}) \)

(ii) \( \mathcal{C}(T_n, T_n) = 2\tau \mathcal{C}(T_n, T_{n-1}) \)

(iii) \( xB(s, t) - yC\left(\frac{-2}{\rho}, t, s\right) = A(x, t, y) - sC(x, t, y) \).

**Proof.**

(i) From equations (2.29) and (2.24) follows that

\[
\mathcal{A}(T_n, T_n) = A(\mathcal{A}(T_n, T_{n-1}), 0, \mathcal{C}(T_n, T_{n-1})) \\
= \mathcal{A}(T_n, T_{n-1}) \left( (\tau - \alpha)e^0 + \tau + \alpha \right) + \mathcal{C}(T_n, T_{n-1}) \left( 1 - e^0 \right) \\
= 2\tau \mathcal{A}(T_n, T_{n-1}).
\]

(ii) This follows along the same lines as in (i), using equations (2.31) and (2.26).

(iii) Using Equations (2.25) and (2.26) one finds:

\[
xB(s, t) - yC\left(\frac{-2}{\rho}, t, s\right) = x \left( \rho s \left( e^{-\tau t} - 1 \right) + (\tau - \alpha)e^{-\tau t} + \tau + \alpha \right) \\
- y \left( s \left( (\alpha + \tau)e^{-\tau t} + \tau - \alpha \right) + 2 \left( 1 - e^{-\tau t} \right) \right).
\]
Rearranging terms, and using Equations (2.24) and (2.26), we find
\[
x B(s, t) - y C \left( -\frac{2}{\rho}, t, s \right) = x \left( (\tau - \alpha) e^{-\tau t} + \tau + \alpha \right) + 2y \left( 1 - e^{-\tau t} \right)
- s \left( y \left( (\alpha + \tau) e^{-\tau t} + \tau - \alpha \right) + x \rho \left( 1 - e^{-\tau t} \right) \right)
- s C(x, t, y),
\]
which establishes the result.

Now, Theorem 2.7 can be proved.

**Proof of Theorem 2.7.** For each \( n \) it has to be shown that (2.32) 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 \( n \geq 1 \).

\( n = 0 \): For \( t = T_0 = 0 \) one has by assumption:
\[
g(s, 0) = \left( \frac{\varphi}{\varphi - s} \right)^{2\theta}.
\]
From (2.32) one finds:
\[
g(s, 0) = K(0)p_0(s, 0) \left( \frac{1}{\varphi(0, 0) - sC(0, 0)} \right)^{2\theta}
= e^0 \left( 2\tau \varphi \right)^{2\theta} \left( \frac{1}{A(\varphi, 0, 1) - sC(\varphi, 0, 1)} \right)^{2\theta}
= \left( \frac{2\tau \varphi}{2\tau \varphi - 2\tau s} \right)^{2\theta} = \left( \frac{\varphi}{\varphi - s} \right)^{2\theta}.
\]

Next the interval up to the first jump time, \( 0 < t < T_1 \), is considered. From (2.27) and the expression for \( w(s) = g(s, 0) \), one finds:
\[
g(s, t) = e^{\theta(\alpha - \tau)t} \left( \frac{2\tau}{B(s, t)} \right)^{2\theta} \left( \frac{\varphi}{\varphi - C\left( \frac{2}{\rho}, t, s \right)} \right)^{2\theta}
= e^{\theta(\alpha - \tau)t} \left( 2\tau \varphi \right)^{2\theta} \left( \frac{1}{B(s, t) \varphi - C\left( \frac{2}{\rho}, t, s \right)} \right)^{2\theta}
= K(t)p_0(s, t) \left( \frac{1}{\varphi(t, 0) - sC(t, 0)} \right)^{2\theta},
\]
which is the same expression as in (2.32) for $n = 0$. The final step in the derivation above follows from Lemma 2.8 (iii), with $x = \varphi$ and $y = 1$, together with the definition of $K(t)$ in (2.34).

$n \geq 1$: Now it remains to prove the induction step. Therefore one can assume that equation (2.32) 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 $\hat{\lambda}_{T_n-}$. By (2.22) one has

$$g(s, T_n) = \left(\hat{\lambda}_{T_n-}^0\right)^{-1} \frac{\partial}{\partial s} g(s, T_n)$$

$$= \left(\hat{\lambda}_{T_n-}^0\right)^{-1} \frac{\partial}{\partial s} \left(K(T_n-) p_{n-1}(s, T_n)\right)$$

$$\times \left(\frac{1}{\mathcal{A}(T_n, T_{n-1}) - s\mathcal{E}(T_n, T_{n-1})}\right)^{2\theta + n - 1}.$$  \hspace{1cm} (2.35)

Calculating the derivative with respect to $s$, leads to

$$g(s, t) = \left(\hat{\lambda}_{T_n-}^0\right)^{-1} K(T_n-) \left[p_{n-1}(s, T_n)(2\theta + n - 1)\mathcal{E}(T_n, T_{n-1})ight]$$

$$+ \frac{\partial}{\partial s} p_{n-1}(s, T_n) \left(\mathcal{A}(T_n, T_{n-1}) - s\mathcal{E}(T_n, T_{n-1})\right)$$

$$\times \left(\frac{1}{\mathcal{A}(T_n, T_{n-1}) - s\mathcal{E}(T_n, T_{n-1})}\right)^{2\theta + n}. \hspace{1cm} (2.36)$$

From Lemma 2.8 (i) and (ii) follows that for the denominator in (2.36) one has

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

Hence (2.36) can be written as

$$g(s, T_n) = \left(\hat{\lambda}_{T_n-}^0\right)^{-1} K(T_n-)(2\tau)^{2\theta}(2\tau)^{n-1} \left[p_{n-1}(s, T_n)(2\theta + n - 1)\mathcal{E}(T_n, T_{n-1})\right]$$

$$+ \frac{\partial}{\partial s} p_{n-1}(s, T_n) \left(\mathcal{A}(T_n, T_{n-1}) - s\mathcal{E}(T_n, T_{n-1})\right)$$

$$\times \left(\frac{1}{\mathcal{A}(T_n, T_{n-1}) - s\mathcal{E}(T_n, T_{n-1})}\right)^{2\theta + n}.$$  \hspace{1cm} (2.37)
From (2.34) it is easy to see that \( K(T_n) = K(T_{n-}) \left( \hat{\lambda}_{T_{n-}}^0 \right)^{-1} (2\tau)^{2\theta} \), and further one has \( 2\tau = B(s,0) = B(s,T_n - T_n) \). From this follows that (2.37) can be written as

\[
g(s,T_n) = K(T_n)B^n(s,T_n - T_n) \left[ p_{n-1}(s,T_n)(2\theta + n - 1)\mathcal{C}(T_n,T_{n-1}) \right.
\]

\[
+ \frac{\partial}{\partial s} p_{n-1}(s,T_n) \left( \mathcal{A}(T_n,T_{n-1}) - s\mathcal{C}(T_n,T_{n-1}) \right) \]

\[
\times \left( \frac{1}{\mathcal{A}(T_n,T_n) - s\mathcal{C}(T_n,T_n)} \right)^{2\theta + n}.
\]

This can be simplified further using the definition of \( p_n(s,t) \) as given in (2.33), together with the identity \( C \left( \frac{-2}{\rho}, 0, s \right) = \tau s \). This results in

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

which is the required result at \( t = T_n \). Finally one has to check that (2.32) holds for \( T_n < t < T_{n+1} \). For this one can use the general solution (2.27) with initial condition \( w(s) = g(s,T_n) \). One finds

\[
g(s,t) = e^{\theta(\alpha - \tau)(t - T_n)} \left( \frac{2\tau}{B(s,t - T_n)} \right)^{2\theta} \left( \frac{2\tau}{\hat{\lambda}_{T_{m-}}^0} \right)^{2\theta} \left( \frac{2\tau}{B(s,t - T_n)} \right)^{2\theta}
\]

\[
\times \prod_{m \geq 1, \ T_m \leq T_n} \left( \frac{(2\tau)^{2\theta}}{\hat{\lambda}_{T_{m-}}^0} \right) p_n \left( \frac{C \left( \frac{-2}{\rho}, t - T_n, s \right)}{B(s,t - T_n)}, T_n \right)
\]

\[
\times \left( \frac{1}{\mathcal{A}(T_n,T_n) - s\mathcal{C}(T_n,T_n)} \right)^{2\theta + n}.
\]

Simplifying this expression yields:

\[
g(s,t) = e^{\theta(\alpha - \tau)(2\tau)^{2\theta}} \left( \frac{2\tau}{\hat{\lambda}_{T_{m-}}^0} \right)^{2\theta} \left( \frac{2\tau}{B(s,t - T_n)} \right)^{2\theta} \left( \frac{2\tau}{B(s,t - T_n)} \right)^{2\theta}
\]

\[
\times \left( \frac{1}{2\tau B(s,t - T_n)\mathcal{A}(T_n,T_{n-1}) - 2\tau C \left( \frac{-2}{\rho}, t - T_n, s \right)\mathcal{C}(T_n,T_{n-1})} \right)^{2\theta + n}.
\]
An application of Lemma 2.8, with $x = \mathcal{A}(T_n, T_{n-1})$ and $y = \mathcal{C}(T_n, T_{n-1})$, and the definitions of $\mathcal{A}(t, T_n)$ and $\mathcal{C}(t, T_n)$ in (2.29) and (2.31), together with the definition of $K(t)$ results in

$$g(s, t) = K(t) \frac{1}{(2\tau)^n} B^n(s, t - T_n) p_n \left( \frac{C \left( -\frac{2}{\rho}, t - T_n, s \right)}{B(s, t - T_n)}, T_n \right)$$

$$\times \left( \frac{1}{\mathcal{A}(t, T_n) - s \mathcal{C}(t, T_n)} \right)^{2\theta + n}.$$

Next, with the definition of $p_n(s, T_n)$ from (2.33), evaluated in $t = T_n$, together with $C(x, 0, y) = 2\tau y$ and $B(s, 0) = 2\tau$ one rewrites this to

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

$$\times (2\tau)^n \left[ p_{n-1} \left( \frac{C \left( -\frac{2}{\rho}, t - T_n, s \right)}{B(s, t - T_n)}, T_n \right) (2\theta + n - 1) \mathcal{C}(T_n, T_{n-1})$$

$$+ \partial_t \left( p_{n-1} \left( \frac{C \left( -\frac{2}{\rho}, t - T_n, s \right)}{B(s, t - T_n)}, T_n \right) \right)$$

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

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

In the final step the definition of $p_n(s, t)$ is used, this time evaluated in $t$, which concludes the proof of (2.32). From the definition of $B(s, t)$ and $C(x, t, y)$, with $y = s$, which are both linear in $s$, it follows that $p_n(s, t)$ is a polynomial of degree $n$ in $s$. □

This theorem provides a recursive solution to equation (2.13), in case $\lambda_0$ is distributed according to a $\Gamma(2\theta, \varphi)$ distribution. From (2.12) it already known that the conditional moment generating function can easily be obtained from an expression for $g(s, t)$. Now this has been found, the conditional moment generating function $\hat{f}(s, t)$ can be obtained easily.

**Corollary 2.9.** Under the assumptions of Theorem 2.7 the conditional moment generating function $\hat{f}(s, t)$, for $T_n \leq t < T_{n+1}$, can be expressed as:

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

(2.38)
where
\[ q_n(s, t) = \frac{p_n(s, t)}{p_n(0, t)} \text{ and } Q(t, T_n) = \frac{\mathcal{A}(t, T_n)}{\mathcal{C}(t, T_n)}. \]

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

**Proof.** The result follows directly from equation (2.12), Theorem 2.7 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 2.2. To conclude this section it is observed that the conditional moment generating function in (2.38) corresponds to a mixture of Gamma distributions.

**Remark 2.10.** Corollary 2.9 provides an expression for \( \hat{f}(s, t) \) that involves the polynomial \( q_n(\cdot, t) \). Deriving an explicit expression for \( q_n(s, t) = \frac{p_n(s, t)}{p_n(0, t)} \) for any \( n \geq 0 \) is quite complicated, but we can write
\[ q_n(s, t) = \sum_{i=0}^{n} R^n_i(t) s^i, \]
where the coefficients, \( R^n_i(t) \), 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 (2.33).

Next, one can consider \( n + 1 \) independent random variables \( \Gamma_i \), where \( \Gamma_i \sim \Gamma(2\theta + n - i, Q(t, T_n)) \), for \( i = 0, 1, \ldots, n \). Further, consider the discrete random variable \( M^n \), independent of the \( \Gamma_i \), which assumes the values 0, 1, \ldots, \( n \), with probabilities \( \pi^n_i(t) \), and define the random variable
\[ X^n_t = \sum_{i=0}^{n} 1_{\{M=i\}} \Gamma_i. \]

The moment generating function of \( X^n_t \) can easily be found, as \( \Gamma_i \) and \( M^n \) are independent, hence
\[ \mathbb{E}[e^{sX^n_t}] = \sum_{i=0}^{n} \mathbb{E}[e^{s \Gamma_i} 1_{\{M=i\}}] = \sum_{i=0}^{n} \pi^n_i(t) \mathbb{E}[e^{s \Gamma_i}] = \sum_{i=0}^{n} \pi^n_i(t) \left( \frac{Q(t, T_n)}{Q(t, T_n) - s} \right)^{2\theta + n - i}. \]

The goal is to show that, by choosing the probabilities correctly, the moment generating function of \( X^n_t \) equals the conditional moment generation function \( \hat{f}(s, t) \).
Therefore (2.39) is first rewritten as

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

To have that both moment generating functions $\hat{f}(s, t)$ and (2.39) are equal, it is required that

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

The right hand side of this equation can be written as

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

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

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

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

The moment generating functions are equal when

$$R^n_j(t) = (-1)^j Q(t, T_n)^{-j} \sum_{i=j}^{n} \binom{i}{j} \pi^n_i(t),$$

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

$$\pi^n_j(t) = (-1)^j R^n_j(t) Q(t, T_n)^j - \sum_{i=j+1}^{n} \pi^n_i(t) \binom{i}{j}.$$ 

(2.40)

It is not immediately clear from (2.40) that the $\pi^n_j(t)$ are all non-negative and sum to one. It turns out however that this is indeed the case for $T_n \leq t < T_{n+1}$, which means that the $\pi^n_j(t)$ 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 2.1, two graphs are given in which the probabilities are plotted.
2.6 Remarks on Model Extensions

We briefly discuss two ways of extending the model that we have considered in this chapter. 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 multi-factor specification of the intensity, where we see that the calculations for the one-dimensional case do not carry over to this multi-factor case.

Time Varying Parameters

In this section 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 $\alpha$, $\mu_0(t)$ instead of $\mu_0$ etc. Many results in Sections 2 and 3 remain valid upon substitution of the constants by their time varying counterparts. In particular Equation (2.15) will change into

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

But an explicit closed form solution for $g(s, t)$ that we were able to give for the constant parameter case by Equation (2.16) is in general impossible to obtain. The main reason for this is that transformation as given in (2.17) 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

Figure 2.1: 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 $\pi^n_j(t)$. The model parameters are chosen to be $\alpha = 0.5$, $\beta = 0.5$, $\mu_0 = 0.4$ and $\varphi = 4.0$. 

(a) $\pi^n_0$ (short dashed line), $\pi^n_1$ (long dashed line) and $\pi^n_2$ (solid line)  
(b) $\pi^n_0$ (dot dashed line), $\pi^n_1$ (short dashed line), $\pi^n_2$ (long dashed line) and $\pi^n_3$ (solid line)
\( t_i \to \infty \) 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 \( \{t_{1n}, \ldots, t_{kn}\} \), which could be an empty set, in which case we can simply use (2.16) with the prevailing parameter values. Suppose now that this set is non-empty. On the subinterval \([T_{n-1}, t_{1n})\), we can compute the solution \( g(s, t) \) to (2.41) again according to (2.16), eventually yielding \( g(s, t_{1n}) \). Then we consider the PDE (2.41) on the interval \([t_{1n}, t_{2n})\) with initial condition at \( t_{1n} \) (instead of \( T \)) \( w(s) = g(s, t_{1n}) \) and the values of the parameters on this interval. With the appropriate modifications, formula (2.16) can be used again. One then proceeds in this way until the final interval \([t_{kn}, T_n)\) 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.

**Multi-Factor 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,

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

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 (2.6) to \( \lambda_t \) we find

\[
\text{d}\hat{\lambda}_t = \left( -\alpha_1 (\hat{\lambda}_{1,t} - \mu_1) - \alpha_2 (\hat{\lambda}_{2,t} - \mu_2) \right) \text{d}t + \left( \frac{\lambda_{1,t}^2}{\lambda_{1,t} - \hat{\lambda}_t} - \frac{\lambda_{2,t}^2}{\lambda_{2,t} - \hat{\lambda}_t} \right) (\text{d}N_t - \hat{\lambda}_t \text{d}t).
\]

Just as in the one-dimensional case, this involves the term \( \frac{\lambda_{1,t}^2}{\lambda_{1,t} - \hat{\lambda}_t} \), and thus we again consider the conditional moment generating function \( \hat{f}(s, t) = \mathbb{E} [e^{s\lambda_t} | \mathcal{F}_t^N] \). Therefore we have to determine the dynamics of \( e^{s\lambda_t} \). An application of Itô’s formula yields

\[
\text{d}e^{s\lambda_t} = \left[ \left( -\alpha_1 s + \frac{1}{2} s^2 \beta_1^2 \right) \lambda_{1,t} e^{s\lambda_t} + \left( -\alpha_2 s + \frac{1}{2} s^2 \beta_2^2 \right) \lambda_{2,t} e^{s\lambda_t} \right] \text{d}t + s e^{s\lambda_t} \left( \beta_1 \sqrt{\lambda_{1,t}} \text{d}W_{1,t} + \beta_2 \sqrt{\lambda_{2,t}} \text{d}W_{2,t} \right).
\] (2.42)

Comparing the terms in the square brackets above with those in (2.7), 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 2.5. In the two factor model, the derivative of $e^{s\lambda_t}$ with respect to $s$, results in $(\lambda_{1,t} + \lambda_{2,t}) e^{s\lambda_t}$. The terms $\lambda_{i,t} e^{s\lambda_t}$ in (2.42) thus cannot be written as $\partial e^{s\lambda_t} / \partial s$. This shows that a solution, similar to that of Proposition 2.5, cannot be obtained.

Alternatively, one could consider the conditional moment generating function of the pair $(\lambda_{1,t}, \lambda_{2,t})$, given by $h(s_1, s_2, t) = \mathbb{E} \left[ e^{s_1\lambda_{1,t} + s_2\lambda_{2,t}} \mid \mathcal{F}_t \right]$, 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 chapter. We conclude that it is far from straightforward to extend the explicit solution that we have obtained to models with more than one-factor.