Dependent Defaults and Credit Migrations*

Tomasz R. Bielecki
Department of Mathematics, The Northeastern Illinois University, Chicago, USA
Marek Rutkowski
Faculty of Mathematics and Information Science, Warsaw University of Technology,
Warszawa, Poland

This version: 11 March 2003

Abstract

The paper deals with the modeling of mutually dependent default times of several credit names through the intensity-based approach. We extend to the case of multiple ratings some models results due to Kusuoka (1999) and Jarrow and Yu (2001). The issue of the arbitrage valuation of related basket credit derivatives is also briefly examined. We argue that our approach permits in some cases to reduce significantly the dimensionality of the valuation problem at hand.

1   Dependent Intensities of Default Times

In the case of mutually dependent defaults, it is natural to assume that the default probability of a certain entity increases as soon as a related firm defaults on its obligations. Within the so-called reduced-form approach to the modeling of credit risk, this kind of dependence is reflected in the jump of the default intensity of a given firm at the default time of another entity. This specific method of modeling mutually dependent default times was examined by, among others, Schmidt (1998), Kusuoka (1999), Jeanblanc and Rutkowski (2000, 2001), Kijima and Muromachi (2000), Schönbucher and Schubert (2000), and Jarrow and Yu (2001). We present few results, which generalize the valuation formulae for corporate bonds obtained in the recent paper by Jarrow and Yu (2001). We focus on the modeling of mutually dependent default times and credit migrations of several credit names through the intensity-based approach. Although most presented results remain valid for a finite number of reference entities, for the sake of expositional clarity, we shall first concentrate on the special case of two reference entities. Subsequently, we shall present results concerning the general set-up; that is, the case of several entities and multiple ratings.

We shall argue that for some contracts the calculations can be reduced to the familiar results concerning the case of conditionally independent random times through a judicious choice of an equivalent probability measure. Finally, it should be stressed that in our approach the migration process can also be seen as a standard Markov chain (or a conditional Markov chain). However, for large pools of obligors the dimension of the state space of the corresponding joint migration process

is excessively large. Thus, direct application of standard tools from the theory of Markov chains becomes problematic, as it is computationally infeasible.

1.1   Jarrow and Yu (2001) Approach

Kusuoka's (1999) counter-example, examined in some detail in Jeanblanc and Rutkowski (2000), shows that the modeling of the counterparty risk in default-risk sensitive contracts constitutes a rather delicate issue. Jarrow and Yu (2001) argue that some difficulties can be circumvented through a judicious choice of reference filtrations. To explain the ideas that underpin their approach, we start by assuming that there are n firms in the economy; they are also informally referred to as 'counterparties' in the sequel. Jarrow and Yu (2001) propose to make a distinction between the *primary firms* and the *secondary firms*. The former class encompasses these entities whose probabilities of default are influenced by macroeconomic conditions, but not by the credit risk of counterparties. The pricing of bonds issued by primary firms can be done through the standard intensity-based methodology; in particular, it is natural to introduce in this context the state-variables process Y, representing the macroeconomic factors. Thus, it suffices to focus on securities issued by secondary firms, i.e., firms for which the intensity of default depends on the status of other firms.

To circumvent the issue of the mutual dependence of martingale intensities, Jarrow and Yu (2001) postulate that the information structure is asymmetric. Specifically, in their assessment of default probabilities investors take into account the observed defaults of primary firms, but they deliberately choose to disregard eventual defaults of secondary firms. Such an assumption is supported by real-life financial arguments of two kinds.

First, a secondary firm may be seen as a financial institution that has a long or a short position in the debt of a primary firm (e.g., a large corporation), so that the likelihood of its default depends on the status of this corporation. It is natural to assume that the situation is not symmetric, and the default probability of a primary firm depends only on macroeconomic factors.

In the second plausible financial interpretation, a primary firm may be seen as a large corporation, and a secondary firm as one of many relatively small dependent manufacturers. For instance, a large firm can be a major supplier for several small manufacturers, or there may be a lot of small suppliers for a large corporation.

The following set of assumptions underpins Jarrow and Yu (2001) approach.

Assumptions (JY) Let $\mathcal{I} = \{1, \ldots, n\}$ represent the set of all firms, and let $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ be the reference filtration generated by the relevant 'macroeconomic factors'. It is postulated that:

- for any firm from the set $\{1, \ldots, p\}$ of primary firms, the 'default intensity' depends only on the reference filtration $\mathbb{F}$,

- the 'default intensity' of each firm from the set $\{p+1, \ldots, n\}$ of secondary firms may depend not only on the filtration $\mathbb{F}$, but also on the status (default or no-default) of the primary firms.

The construction of the collection of default times $\tau_1, \ldots, \tau_n$ with the desired properties runs as follows. In the first step, we assume that we are given a family of $\mathbb{F}$-predictable intensity processes $\lambda^1, \ldots, \lambda^p$, and we produce a collection $\tau_1, \ldots, \tau_p$ of $\mathbb{F}$-conditionally independent random times through the canonical method (see, e.g., Section 9.1.2 in Bielecki and Rutkowski (2002)). Specifically, we set

$$\tau_i = \inf \left\{ t \in \mathbb{R}_+ : \int_0^t \lambda_u^i \, du \geq \eta_i \right\}, \tag{1.1}$$

where η_i, $i = 1, \ldots, p$ are mutually independent, identically distributed, random variables with unit exponential law under the martingale probability $\mathbb{Q}^*$.

In the second step, we assume that the underlying probability space $(\Omega, \mathcal{G}, \mathbb{Q}^*)$ is large enough to accommodate a family η_i, $i = p + 1, \ldots, n$ of mutually independent random variables, with unit exponential law under $\mathbb{Q}^*$, and such that these random variables are independent not only of the filtration $\mathbb{F}$, but also of the already constructed in the previous step default times $\tau_1, \ldots, \tau_p$ of primary firms.

The default times τ_i, $i = p + 1, \ldots, n$ are also defined by means of equality (1.1). However, the 'intensity processes' λ^i, $i = p + 1, \ldots, n$ are now given by the following generic expression:

$$\lambda_t^i = \mu_t^i + \sum_{l=1}^{p} \nu_t^{i,l} \mathbb{1}_{\{\tau_l \leq t\}}, \tag{1.2}$$

where μ^i and $\nu^{i,l}$ are $\mathbb{F}$-predictable stochastic processes. If the default of the jth primary firm does not affect the likelihood of default of the ith secondary firm, we set $\nu^{i,j} \equiv 0$ in (1.2).

We introduce the jump process $H_t^i = \mathbb{1}_{\{\tau_i \leq t\}}$ for $t \in \mathbb{R}_+$, and we write $\mathbb{H}^i = (\mathcal{H}_t^i)_{t \geq 0}$ to denote the natural filtration of the process H^i. Let $\mathbb{G} = \mathbb{F} \vee \mathbb{H}^1 \vee \ldots \vee \mathbb{H}^n$ stand for the enlarged filtration, that is, the smallest filtration encompassing filtrations $\mathbb{H}^1, \ldots, \mathbb{H}^n$ and $\mathbb{F}$. This means that $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$, where $\mathcal{G}_t = \sigma(\mathcal{F}_t, \mathcal{H}_t^1, \ldots, \mathcal{H}_t^n)$. Similarly, let $\tilde{\mathbb{F}} = \mathbb{F} \vee \mathbb{H}^{p+1} \vee \ldots \vee \mathbb{H}^n$ be the filtration generated by the 'macroeconomic factors' and the observations of default times of secondary firms.

It is worthwhile to mention that:

- the default times $\tau_1, \ldots, \tau_p$ of primary firms, as given by (1.1), are no longer conditionally independent when we replace the reference filtration $\mathbb{F}$ by the larger filtration $\tilde{\mathbb{F}}$,

- for each primary firm, its default intensity with respect to the filtration $\tilde{\mathbb{F}}$ differs from the corresponding default intensity λ^i with respect to $\mathbb{F}$, in general.

The last observation indicates that the processes $\lambda^1, \ldots, \lambda^p$ do not represent the conditional probabilities of survival, unless we disregard the information flow generated by default processes of secondary firms. Put another way, a one-way dependence in default intensities is not possible. If the intensity of default of firm A jumps at the time of default of firm B, a similar effect will show up in the default intensity of firm B at the time of default of firm A. Of course, the concept of default intensity heavily depends on the choice of filtration.

1.1.1 Case of two firms

To clarify the last statement, we shall examine in detail a special case of the Jarrow and Yu model. We consider only two firms, A and B say, and we postulate that the first one represents a primary firm, while the second is a secondary firm. Let the $\mathbb{F}$-predictable process λ^1 represent the $\mathbb{F}$-intensity of default for firm A. The default time τ_1 is given by the standard formula

$$\tau_1 = \inf \left\{ t \in \mathbb{R}_+ : \int_0^t \lambda_u^1 \, du \geq \eta_1 \right\},$$

where η_1 is a random variable, independent of the filtration $\mathbb{F}$, and exponentially distributed under the martingale probability $\mathbb{Q}^*$. For the second firm, the 'intensity' of default λ^2 is assumed to satisfy

$$\lambda_t^2 = \mu_t^2 + \nu_t^{2,1} \mathbb{1}_{\{\tau_1 \leq t\}},$$

where μ^2 and $\nu^{2,1}$ are positive, $\mathbb{F}$-predictable processes. We set

$$\tau_2 = \inf \big\{ t \in \mathbb{R}_+ : \int_0^t \lambda_u^2 \, du \geq \eta_2 \big\},$$

where η_2 is a random variable with the unit exponential law under $\mathbb{Q}^*$, independent of $\mathbb{F}$, and such that η_1 and η_2 are mutually independent under $\mathbb{Q}^*$. From the construction above, it is apparent that the following properties are valid:

- the process λ^1 represents the intensity of τ_1 with respect to the filtration $\mathbb{F}$,

- the process λ^2 represents the intensity of τ_2 with respect to the filtration $\mathbb{F} \vee \mathbb{H}^1$,

- the process λ^1 does not represent the intensity of τ_1 with respect to the filtration $\mathbb{F} \vee \mathbb{H}^2$ (see, e.g., Kusuoka (1999)).

We shall now apply the present setup to the valuation of corporate bonds. To this end, we assume that we have already specified some arbitrage-free model of the default-free term structure. In particular, we are given a filtered probability space $(\tilde{\Omega}, \mathbb{F}, \mathbb{P}^*)$, where $\mathbb{P}^*$ is the spot martingale measure for the Treasury bonds market. As usual, we denote by $B(t,T)$ the price at time t of a unit zero-coupon Treasury bond which matures at time $T \geq t$.

To obtain closed-form representations for the values of corporate bonds, in addition, we postulate that $\lambda_t^1 = \lambda_1$ for some strictly positive constant λ_1, and λ^2 equals $\lambda_t^2 = \lambda_2 + (\alpha_2 - \lambda_2) \mathbb{1}_{\{\tau_1 \leq t\}}$, for some strictly positive constants λ_2 and α_2. Notice the size of the jump of the process λ^2 at the random time τ_1 may be either positive or negative, depending on the specific financial interpretation.

To construct the default times τ_1 and τ_2, we enlarge the probability space in a standard way, so that we end up with the enlarged probability space $(\Omega, \mathbb{G}, \mathbb{Q}^*)$, and the two mutually independent, exponentially distributed, random variables η_1, η_2 that are also independent of the filtration $\mathbb{F}$ under $\mathbb{Q}^*$. As usual, we shall write $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t^1 \vee \mathcal{H}_t^2$ for every $t \in \mathbb{R}_+$.

For any date $T > 0$, we shall also introduce the forward martingale measure $\mathbb{Q}_T$ associated with $\mathbb{Q}^*$. To this end, recall that under the spot martingale measure $\mathbb{P}^*$ (and thus also under $\mathbb{Q}^*$) the dynamics of the price process $B(t,T)$ are:

$$dB(t,T) = B(t,T)\big(r_t \, dt + b(t,T) \, dW_t^*\big),$$

where the volatility $b(\cdot, T)$ follows an $\mathbb{F}$-progressively measurable stochastic process, and W^* is a standard Brownian motion with respect to $\mathbb{F}$ under $\mathbb{P}^*$ (W^* is also a standard Brownian motion with respect to both $\mathbb{F}$ and $\mathbb{G}$ under $\mathbb{Q}^*$). The probability measure $\mathbb{Q}_T$ is given on $(\Omega, \mathcal{G}_T)$ through the Radon-Nikodým density

$$\frac{d\mathbb{Q}_T}{d\mathbb{Q}^*}\Big|_{\mathcal{G}_T} = \exp\Big(\int_0^T b(u,T) \, dW_u^* - \frac{1}{2} \int_0^T b^2(u,T) \, du \Big), \quad \mathbb{Q}^*\text{-a.s.}$$

In view of the assumed independence, it is clear that the random variables η_1, η_2 have identical probabilistic properties under the spot martingale measure $\mathbb{Q}^*$ and under the associated forward martingale measure $\mathbb{Q}_T$. For the sake of concreteness, we shall assume that the bonds issued by the firms A and B are subject to the fractional recovery of Treasury value scheme with constant recovery rates δ_1 and δ_2, respectively. For other types of recovery schemes, the foregoing results need to be modified in a suitable way.

Let us first report the results obtained by Jarrow and Yu (2001). They show that at any date $t \leq T$ the bond issued by the primary firm has the following value:

$$D_1(t,T) = B(t,T)\big(\delta_1 + (1 - \delta_1)e^{-\lambda_1(T-t)}\mathbb{1}_{\{\tau_1 > t\}}\big). \tag{1.3}$$

This valuation formula is rather obvious when it is postulated that

$$D_1(t,T) = B(t,T)\,\mathbb{E}_{\mathbb{Q}_T}\big(\mathbb{1}_{\{\tau_1 > T\}} + \delta_1\mathbb{1}_{\{\tau_1 \leq T\}}\,\big|\,\mathcal{F}_t \vee \mathcal{H}_t^1\big),$$

that is, when the status of the secondary firm is not observed. It is noteworthy that the right-hand side of (1.3) also yields the correct value for $D_1(t,T)$ if it is defined through the standard formula:

$$D_1(t,T) = B(t,T)\,\mathbb{E}_{\mathbb{Q}_T}\big(\mathbb{1}_{\{\tau_1 > T\}} + \delta_1\mathbb{1}_{\{\tau_1 \leq T\}}\,\big|\,\mathcal{G}_t\big).$$

which takes into account the full information about macroeconomic factors and the status of both firms. The intuitive difference between the last two expressions is clear: the former assumes a priori that the occurrence of default of the secondary firm is not relevant for the valuation of a bond issued by the primary firm, while the latter relies on the complete information available at time t.

The calculation of the value of the bond issued by the secondary firm is more involved. We adopt here the usual formula based on the full information – that is, we set

$$D_2(t,T) = B(t,T)\,\mathbb{E}_{\mathbb{Q}_T}\big(\mathbb{1}_{\{\tau_2 > T\}} + \delta_2\mathbb{1}_{\{\tau_2 \leq T\}}\,\big|\,\mathcal{G}_t\big).$$

Let us introduce some useful notation. For $\lambda_1 + \lambda_2 - \alpha_2 \neq 0$, we write

$$c_{\lambda_1,\lambda_2,\alpha_2}(u) = \frac{1}{\lambda_1 + \lambda_2 - \alpha_2}\Big(\lambda_1 e^{-\alpha_2 u} + (\lambda_2 - \alpha_2)e^{-(\lambda_1 + \lambda_2)u}\Big).$$

Otherwise, we set

$$c_{\lambda_1,\lambda_2,\alpha_2}(u) = \big(1 + \lambda_1 u\big)e^{-(\lambda_1 + \lambda_2)u}.$$

The following proposition is borrowed from Jarrow and Yu (2001). Later, we are going to establish a general result that covers the Jarrow and Yu result as a special case.

Proposition 1.1 *The value of a zero-coupon bond issued by the secondary firm equals, on the event $\{\tau_1 > t\}$, that is, prior to default of the primary firm,*

$$D_2(t,T) = B(t,T)\big(\delta_2 + (1 - \delta_2)c_{\lambda_1,\lambda_2,\alpha_2}(T - t)\mathbb{1}_{\{\tau_2 > t\}}\big),$$

and on the set $\{\tau_1 \leq t\}$, that is, after default of the primary firm,

$$D_2(t,T) = B(t,T)\big(\delta_2 + (1 - \delta_2)e^{-\alpha_2(T-t)}\mathbb{1}_{\{\tau_2 > t\}}\big).$$

Assume, for the sake of simplicity that $\delta_2 = 0$, i.e., the bond issued by the secondary firm is subject to the zero recovery rule. Then for every $t \leq T$ Proposition 1.1 yields (we denote hereafter $\lambda = \lambda_1 + \lambda_2$)

$$D_2(t,T) = \mathbb{1}_{\{\tau_1 > t, \tau_2 > t\}}\frac{1}{\lambda - \alpha_2}\Big(\lambda_1 e^{-\alpha_2(T-t)} + (\lambda_2 - \alpha_2)e^{-\lambda(T-t)}\Big) + \mathbb{1}_{\{\tau_1 \leq t < \tau_2\}}\,e^{-\alpha_2(T-t)}.$$

1.2   Extension of Jarrow and Yu (2001) Results

We shall now argue that the assumption that some entities represent 'primary firms', while other entities are 'secondary firms', is actually irrelevant, and thus it can be relaxed. For the sake of expositional simplicity, we maintain here the assumption that $n = 2$; i.e., we consider the case of two firms, and we place ourselves in Kusuoka's (1999) set-up with $T^* = \infty$ (see also Schmidt (1998)). For detailed exposition of this framework, we refer to Section 7.3 in Bielecki and Rutkowski (2002). Let us only mention here that the compensated jump processes, for $t \in [0, T]$,

$$H_t^1 - \int_0^{t \wedge \tau_1} \lambda_u^{*1}\, du = H_t^1 - \int_0^{t \wedge \tau_1} \left(\lambda_1 \mathbb{1}_{\{\tau_2 > u\}} + \alpha_1 \mathbb{1}_{\{\tau_2 \leq u\}} \right) du$$

and

$$H_t^2 - \int_0^{t \wedge \tau_2} \lambda_u^{*2}\, du = H_t^2 - \int_0^{t \wedge \tau_2} \left(\lambda_2 \mathbb{1}_{\{\tau_1 > u\}} + \alpha_2 \mathbb{1}_{\{\tau_1 \leq u\}} \right) du$$

follow martingales under $\mathbb{Q}^*$ with respect to the joint filtration $\mathbb{G} = \mathbb{F} \vee \mathbb{H}^1 \vee \mathbb{H}^2$. For the sake of concreteness and simplicity, the parameters $\lambda_1, \lambda_2, \alpha_1$ and α_2 are assumed to be strictly positive constants, so that the reference filtration $\mathbb{F}$ plays no important role in the foregoing calculations. We may thus assume that it is trivial. Let us recall that the process λ^{*1} (λ^{*2}, resp.) is referred to as the *martingale intensity* of τ_1 (τ_2, resp.) with respect to the filtration $\mathbb{F} \vee \mathbb{H}^2$ ($\mathbb{F} \vee \mathbb{H}^1$, resp.)

An alternative (more direct) construction of random times with required properties is given in Shaked and Shanthikumar (1987), also in the case of more than two random times. It is thus worth stressing that the considerations below remain valid in the extended Jarrow and Yu (2001) framework for any $n \geq 2$.

1.2.1   Case of zero recovery

To start with, we postulate that both corporate bonds we are going to analyze are subject to the zero-recovery scheme, and the interest rate r is constant, so that $B(t, T) = e^{-r(T-t)}$ for every $t \leq T$. Due to the last assumption we have $\mathbb{Q}_T = \mathbb{Q}^*$ and we may take the trivial filtration as the reference filtration $\mathbb{F}$. Since the situation is symmetric, it suffices to analyze one bond only, for instance, a bond issued by the first firm. By definition, the price of this bond equals

$$D_1(t, T) = B(t, T)\, \mathbb{Q}^*(\tau_1 > T \,|\, \mathcal{G}_t) = B(t, T)\, \mathbb{Q}^*(\tau_1 > T \,|\, \mathcal{H}_t^1 \vee \mathcal{H}_t^2).$$

For the sake of comparison, we shall also evaluate the random variable

$$\tilde{D}_1(t, T) := B(t, T)\, \mathbb{Q}^*(\tau_1 > T \,|\, \mathcal{H}_t^2),$$

which models the price of the bond given the observations of the default of the second firm (but not of the first firm), and the random variable

$$\hat{D}_1(t, T) := B(t, T)\, \mathbb{Q}^*(\tau_1 > T \,|\, \mathcal{H}_t^1),$$

which represents the value of the bond based on the observations of the default of the first firm only. To make the formulae slightly shorter, we shall assume in the next result that $r = 0$, so that $B(t, T) = 1$ for every $t \in [0, T]$. Recall that we denote $\lambda = \lambda_1 + \lambda_2$.

Proposition 1.2 *The price $D_1(t,T)$ equals*

$$D_1(t,T) = \mathbb{1}_{\{\tau_1>t,\tau_2>t\}} \frac{1}{\lambda-\alpha_1} \left(\lambda_2 e^{-\alpha_1(T-t)} + (\lambda_1-\alpha_1)e^{-\lambda(T-t)} \right) + \mathbb{1}_{\{\tau_2\le t<\tau_1\}} e^{-\alpha_1(T-t)}.$$

Processes $\tilde{D}_1(t,T)$ and $\hat{D}_1(t,T)$ satisfy

$$\tilde{D}_1(t,T) = \mathbb{1}_{\{\tau_2>t\}} \frac{\lambda-\alpha_2}{\lambda-\alpha_1} \frac{(\lambda_1-\alpha_1)e^{-\lambda(T-t)} + \lambda_2 e^{-\alpha_1(T-t)}}{\lambda_1 e^{(\lambda-\alpha_2)t} + \lambda_2 - \alpha_2} + \mathbb{1}_{\{\tau_2\le t\}} \frac{(\lambda-\alpha_2)\lambda_2 e^{-\alpha_1(T-\tau_2)}}{\lambda_1 \alpha_2 e^{(\lambda-\alpha_2)\tau_2} + \lambda(\lambda_2-\alpha_2)}$$

and

$$\hat{D}_1(t,T) = \mathbb{1}_{\{\tau_1>t\}} \frac{\lambda_2 e^{-\alpha_1 T} + (\lambda_1-\alpha_1)e^{-\lambda T}}{\lambda_2 e^{-\alpha_1 t} + (\lambda_1-\alpha_1)e^{-\lambda t}}.$$

Proof. For the detailed calculations of conditional expectations, we refer to Section 7.3 in Bielecki and Rutkowski (2002). □

Few pertinent comments on the last result are in order. First, in case of a non-zero interest rate r, it suffices to multiply the right-hand sides in the valuation formulae of the last proposition by the price $B(t,T)$. Second, notice that the two processes $D_1(t,T)$ and $\hat{D}_1(t,T)$ obviously vanish after the default time τ_1 of the first firm (the latter remark does not apply to $\tilde{D}_1(t,T)$, however).

First formula. Although derived in a different set-up, formula for $D_1(t,T)$ and an analogous formula for $D_2(t,T)$ clearly coincide with the Jarrow and Yu (2001) valuation result for the bond issued by the secondary firm. Furthermore, it is interesting to observe that the bond price $D_1(t,T)$ ($D_2(t,T)$, resp.) does not depend on the value of α_2 (α_1, resp.) This means, of course, that in calculations of $D_1(t,T)$ we may assume without loss of generality that the equality $\lambda_2 = \alpha_2$ holds. By the same token, when searching for the price $D_2(t,T)$, we may set $\alpha_1 = \lambda_1$.

The last observation is in fact crucial. It shows that when valuing a bond issued by a given firm, we may postulate that the default of this firm affects the default intensity of the other firm (or, of other firms if $n > 2$), but we may equally well assume that the occurrence of default of this firm has no impact on default intensity for the other firm. This is also clear at the intuitive level: since we search for the pre-default value of a corporate bond, the behavior of default probabilities of other firms after the default of the issuer of a bond in question is totally irrelevant.

To conclude, when looking for $D_1(t,T)$ we only need to distinguish between the two following cases: $\lambda_1 \ne \alpha_1$ and $\lambda_1 = \alpha_1$, and we may always set $\lambda_2 = \alpha_2$. In the former case, the value of $D_1(t,T)$ is given by the general formula of Proposition 1.2. In the latter, this formula reduces to the standard result:

$$D_1(t,T) = \mathbb{1}_{\{\tau_1>t\}} B(t,T) e^{-\lambda_1(T-t)} = \mathbb{1}_{\{\tau_1>t\}} e^{-(r+\lambda_1)(T-t)}.$$

The same arguments apply to $D_2(t,T)$, so that for $\lambda_2 = \alpha_2$ we have

$$D_2(t,T) = \mathbb{1}_{\{\tau_2>t\}} B(t,T) e^{-\lambda_2(T-t)} = \mathbb{1}_{\{\tau_2>t\}} e^{-(r+\lambda_2)(T-t)}.$$

Second formula. The second valuation formula of Proposition 1.2 hinges on the assumption that we observe the default time τ_2, but the default time of the first firm is not observed. The following special cases are of interest:

- for the special case when $\lambda_1 = \alpha_1$ and $\lambda_2 \ne \alpha_2$, the process $\tilde{D}_1(t,T)$ equals

$$\tilde{D}_1(t,T) = \mathbb{1}_{\{\tau_2>t\}} B(t,T) \frac{(\lambda-\alpha_2)e^{-\lambda_1(T-t)}}{\lambda_1 e^{(\lambda-\alpha_2)t} + \lambda_2 - \alpha_2} + \mathbb{1}_{\{\tau_2\le t\}} B(t,T) \frac{(\lambda-\alpha_2)\lambda_2 e^{-\lambda_1(T-\tau_2)}}{\lambda_1 \alpha_2 e^{(\lambda-\alpha_2)\tau_2} + \lambda(\lambda_2-\alpha_2)}.$$

- if we assume instead that $\lambda_1 \neq \alpha_1$ but $\lambda_2 = \alpha_2$, we obtain

$$\tilde{D}_1(t,T) = \mathbb{1}_{\{\tau_2>t\}}B(t,T)\frac{(\lambda_1-\alpha_1)e^{-\lambda(T-t)} + \lambda_2 e^{-\alpha_1(T-t)}}{(\lambda-\alpha_1)e^{\lambda_1 t}} + \mathbb{1}_{\{\tau_2\leq t\}}B(t,T)\,e^{-\lambda_1\tau_2-\alpha_1(T-\tau_2)}.$$

It is interesting to observe that in both cases considered above, the price process $D_1(t,T)$ admits a discontinuity at the default time τ_2 of the second firm. This is a natural consequence of the fact that in both cases the default times τ_1 and τ_2 are mutually dependent under $\mathbb{Q}^*$.

- finally, for $\lambda_1 = \alpha_1$ and $\lambda_2 = \alpha_2$, we get

$$\tilde{D}_1(t,T) = B(t,T)e^{-\lambda_1 T} = B(t,T)\,\mathbb{Q}^*\{\tau_1 > T\}.$$

This result is also clear, since in this very special case the default times τ_1 and τ_2 are mutually independent under the martingale measure $\mathbb{Q}^*$.

Third formula. Finally, the last formula in Proposition 1.2 implicitly assumes that we are able to observe only the default time of the first firm (but not of the second). It is apparent that this result is also independent of the value of the parameter α_2. Again, in case when $\lambda_1 = \alpha_1$, we conclude that for any $t \leq T$ we have:

$$\hat{D}_1(t,T) = \mathbb{1}_{\{\tau_1>t\}}B(t,T)e^{-\lambda_1(T-t)} = D_1(t,T).$$

1.2.2 Case of non-zero recovery rates

The valuation results in the case on non-zero recovery are not much different from the special case of zero recovery scheme. Indeed, the payoff $D_i(T,T)$ at maturity can be represented as follows:

$$D_i(T,T) = \mathbb{1}_{\{\tau_i>T\}} + \delta_i\mathbb{1}_{\{\tau_i\leq T\}} = \delta_i + (1-\delta_i)\mathbb{1}_{\{\tau_i>T\}}$$

and thus

$$D_i(t,T) = B(t,T)\big(\delta_i + (1-\delta_i)\mathbb{Q}^*\{\tau_i > T \,|\, \mathcal{H}_t^1 \vee \mathcal{H}_t^2\}\big).$$

Explicit formulae for $D_i(t,T)$, as well as for $\tilde{D}_i(t,T)$ and $\hat{D}_i(t,T)$, can be thus obtained directly from Proposition 1.2. It is easy to see that the expressions for the values of $D_1(t,T)$ and $D_2(t,T)$ will coincide, up to a suitable change of notation, with the formula of Proposition 1.1. Moreover, when $\lambda_1 = \alpha_1$, we obtain (cf. formula (1.3))

$$D_1(t,T) = B(t,T)\big(\delta_1 + (1-\delta_1)e^{-\lambda_1(T-t)}\mathbb{1}_{\{\tau_1>t\}}\big),$$

while for $\lambda_2 = \alpha_2$, we get

$$D_2(t,T) = B(t,T)\big(\delta_2 + (1-\delta_2)e^{-\lambda_2(T-t)}\mathbb{1}_{\{\tau_2>t\}}\big).$$

1.2.3 Interpretation of martingale intensities

We shall provide an intuitive probabilistic interpretation for the martingale intensities:

$$\lambda_t^{*1} = \lambda_1\mathbb{1}_{\{\tau_2>t\}} + \alpha_1\mathbb{1}_{\{\tau_2\leq t\}}, \quad \lambda_t^{*2} = \lambda_2\mathbb{1}_{\{\tau_1>t\}} + \alpha_2\mathbb{1}_{\{\tau_1\leq t\}}.$$

Recall that the construction of default times given in Kusuoka (1999) relies on an equivalent change of a probability measure. A different, but essentially equivalent, construction of τ_1 and τ_2 runs as

follows: we take the two mutually independent, identically distributed, random variables η_i, $i = 1, 2$ with unit exponential law under $\mathbb{Q}^*$, and we set (see Shaked and Shanthikumar (1987))

$$\tau_1 = \begin{cases} \lambda_1^{-1}\eta_1, & \text{if } \lambda_1^{-1}\eta_1 \leq \lambda_2^{-1}\eta_2, \\ \lambda_2^{-1}\eta_2 + \alpha_1^{-1}(\eta_1 - \lambda_1\lambda_2^{-1}\eta_2), & \text{if } \lambda_1^{-1}\eta_1 > \lambda_2^{-1}\eta_2, \end{cases}$$

$$\tau_2 = \begin{cases} \lambda_2^{-1}\eta_2, & \text{if } \lambda_2^{-1}\eta_2 \leq \lambda_1^{-1}\eta_1, \\ \lambda_1^{-1}\eta_1 + \alpha_2^{-1}(\eta_2 - \lambda_2\lambda_1^{-1}\eta_1), & \text{if } \lambda_2^{-1}\eta_2 > \lambda_1^{-1}\eta_1. \end{cases}$$

The following result shows that the jump of the martingale intensity has the desired financial interpretation. For the proof of Proposition 1.3, the interested reader may consult Shaked and Shanthikumar (1987).

Proposition 1.3 *For $i = 1, 2$, $j \neq i$ and every $t \in \mathbb{R}_+$ we have*

$$\lambda_i = \lim_{h \downarrow 0} h^{-1} \mathbb{Q}^*(t < \tau_i \leq t + h \,|\, \tau_1 > t, \, \tau_2 > t)$$

and

$$\alpha_i = \lim_{h \downarrow 0} h^{-1} \mathbb{Q}^*(t < \tau_i \leq t + h \,|\, \tau_i > t, \, \tau_j \leq t).$$

2 Dependent Intensities of Credit Migrations

For the sake of notational simplicity, we continue considering two firms only. In this section, we shall examine the situation where the current financial standing of the ith firm (of course, $i = 1, 2$) is reflected through an appropriate credit ranking, whose generic value, denoted as k_i, belongs to a finite set of credit grades $\mathcal{K}_i = \{1, \ldots, K_i\}$, where $K_i \geq 2$ for $i = 1, 2$.

Remarks. In the previous section, we have examined a special case of the present setting with $K_1 = K_2 = 2$. That is, we have assumed there that the financial standing of each of the two firms could be classified into one of the two categories: *pre-default* ($k_i = 1$) and *default* ($k_i = 2$).

As observed in practice, the credit ranking of a corporation varies over time. We refer to this possibility as to the *credit migrations*. Our goal is to model the credit migrations of the ith firm in terms of a certain stochastic process, denoted by C^i, taking values in the finite state space $\mathcal{K}_i$. The process C^i is called the *credit migration process* (or simply the *migration process*). We fix the initial credit ranking of each firm, that is, we set C_0^i =const. for $i = 1, 2$, and we assume that at each future date $t > 0$ the credit rankings C_t^i, $i = 1, 2$, can be observed exactly. The issues related to the modeling of credit migrations are examined in detail in Chapters 12 and 13 in Bielecki and Rutkowski (2002).

2.1 Basic Assumptions

For the sake of expositional simplicity, we shall first assume that the reference filtration $\mathbb{F}$ is trivial. Let $\mathbb{F}^i = \mathbb{F}^{C^i}$, $i = 1, 2$, denote the natural filtration of the process C^i. We assume that each filtration $\mathbb{F}^i$ satisfies the so-called "usual conditions", and we denote $\mathbb{G} = \mathbb{F}^1 \vee \mathbb{F}^2$. We shall conduct our study under specific Markovian assumptions under the real-world probability $\mathbb{Q}$ and under the martingale measure $\mathbb{Q}^*$.[1]

---

[1] We refer to Section 11.3 in Bielecki and Rutkowski (2002) for detailed definitions and a discussion of the preservation of the conditional Markov assumption under an equivalent change of a probability measure.

- Assumption (M) Standard $\mathbb{G}$-Markov condition for the process $C = (C^1, C^2)$ under $\mathbb{Q}^*$: for any $t \leq s$ and $k_i \in \mathcal{K}_i$

$$\mathbb{Q}^*(C_s^1 = k_1, \, C_s^2 = k_2 \,|\, \mathcal{G}_t) = \mathbb{Q}^*(C_s^1 = k_1, \, C_s^2 = k_2 \,|\, \sigma(C_t^1, C_t^2)). \qquad (2.4)$$

- Assumption (CM) $\mathbb{F}^j$-conditional $\mathbb{G}$-Markov condition for the process C^i, $i = 1, 2$, $j \neq i$ under $\mathbb{Q}^*$: for any $t \leq s$ and $k_i \in \mathcal{K}_i$

$$\mathbb{Q}^*(C_s^i = k_i \,|\, \mathcal{G}_t) = \mathbb{Q}^*(C_s^i = k_i \,|\, \sigma(C_t^i) \vee \mathcal{F}_t^j).$$

Remarks. (i) Observe that condition (2.4) obviously implies the following equality:

$$\mathbb{Q}^*(C_s^i = k_i \,|\, \mathcal{G}_t) = \mathbb{Q}^*(C_s^i = k_i \,|\, \sigma(C_t^1, C_t^2)).$$

(ii) In the present context, the standard $\mathbb{G}$-Markov property of the process C^i would read:

$$\mathbb{Q}^*(C_s^i = k_i \,|\, \mathcal{G}_t) = \mathbb{Q}^*(C_s^i = k_i \,|\, \sigma(C_t^i)).$$

The latter condition is manifestly not interesting for us, as it essentially says that there is no dependence between the credit migrations of the two firms.

In the general case of a non-trivial reference filtration $\mathbb{F}$ we denote $\mathbb{G} = \mathbb{F} \vee \mathbb{F}^1 \vee \mathbb{F}^2$ and we assume that $\tilde{\mathbb{F}}$ is some filtration such that $\mathbb{F} \subseteq \tilde{\mathbb{F}} \subseteq \mathbb{G}$. We adopt the following generic definition of the conditional Markov property.

Definition 2.1 We say that a process C^i follows an $\tilde{\mathbb{F}}$-*conditional* $\mathbb{G}$-*Markov chain under* $\mathbb{Q}^*$ if for any $t \leq s$ and $k_i \in \mathcal{K}_i$ we have

$$\mathbb{Q}^*(C_s^i = k_i \,|\, \mathcal{G}_t) = \mathbb{Q}^*(C_s^i = k_i \,|\, \sigma(C_t^i) \vee \tilde{\mathcal{F}}_t).$$

Let us denote $\mathbb{G}^j = \mathbb{F} \vee \mathbb{F}^j$. Then for $i = 1, 2$ and $j \neq i$ the $\mathbb{G}^j$-conditional $\mathbb{G}$-Markov property of C^i under $\mathbb{Q}^*$ reads:

$$\mathbb{Q}^*(C_s^i = k_i \,|\, \mathcal{G}_t) = \mathbb{Q}^*(C_s^i = k_i \,|\, \sigma(C_t^i) \vee \mathcal{G}_t^j) = \mathbb{Q}^*(C_s^i = k_i \,|\, \sigma(C_t^i) \vee \mathcal{F}_t \vee \mathcal{F}_t^j).$$

2.2   Jarrow and Yu Set-up with Credit Migrations

We shall first place ourselves within the Jarrow and Yu (2001) set-up, as presented in Section 1.1. We deal throughout with only two firms: a primary firm and a secondary firm. First, we shall re-examine from the Markovian perspective the special case of the Jarrow and Yu framework discussed in Subsection 1.1.1.

2.2.1   Case of two rating grades

Consider the processes C^1 and C^2 defined as $C_t^i = H_t^i + 1$ for every $t \in \mathbb{R}_+$ and $i = 1, 2$. Both processes take values in the state space $\mathcal{K} = \{1, 2\}$, and clearly $\mathbb{F}^i = \mathbb{H}^i$ for $i = 1, 2$. The process C^1 is an ordinary Markov chain (with respect to the filtrations $\mathbb{F}^1$ and $\mathbb{G} = \mathbb{F}^1 \vee \mathbb{F}^2$) with the following infinitesimal generator matrix:

$$\Lambda^1 = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ 0 & 0 \end{pmatrix}.$$

Furthermore, the migration process C^2 is a conditionally Markov chain with respect to the filtration $\mathbb{F}^1$. Specifically, for any $k = 1, 2$ and every $s, t \in \mathbb{R}_+$ we have

$$\mathbb{Q}^*(C_{t+s}^2 = k \,|\, \mathcal{F}_t^1 \vee \mathcal{F}_t^2) = \mathbb{Q}^*(C_{t+s}^2 = k \,|\, \sigma(C_t^2) \vee \mathcal{F}_t^1) \left(= \mathbb{Q}^*(C_{t+s} = k \,|\, \sigma(C_t^1) \vee \sigma(C_t^2)) \right).$$

From Section 1.1, we know that the process $\lambda_t^2 = \lambda_2 + (\alpha_2 - \lambda_2)H_t^1$ is the $(\mathbb{F}^1, \mathbb{F}^1 \vee \mathbb{F}^2)$-martingale intensity of the default time τ_2. Therefore, the conditional infinitesimal generator process[2] for C^2 is given as:

$$\Lambda_t^2 = \begin{pmatrix} -\lambda_t^2 & \lambda_t^2 \\ 0 & 0 \end{pmatrix}.$$

For every $j, k \in \mathcal{K}$ and any dates $0 \le t \le s$, we denote by $p_{jk}^2(t, s)$ the conditional probability $\mathbb{Q}^*(C_s^2 = k \,|\, \mathcal{F}_t^1 \vee \sigma(C_t^2))$ on the event $\{C_t^2 = j\}$, i.e., the $\mathcal{F}_t^1$-conditional probability of transition from the state j to the state k over the time interval $[t, s]$ by the credit migration process C^2. Finally, we write $\mathcal{P}^2(t, s) = [p_{jk}^2(t, s)]_{j,k \in \mathcal{K}}$ to denote the conditional transition matrix for C^2.

Let us define a matrix-valued random field $\mathbf{P}^2(t, s)$ in terms of the following (random) ODE:

$$\frac{d\mathbf{P}^2(t, s)}{ds} = \mathbf{P}^2(t, s)\Lambda_s^2, \quad s \ge t, \quad \mathbf{P}^2(t, t) = \mathrm{Id}_2 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

It is clear that

$$\mathbf{P}^2(t, s) = \begin{pmatrix} e^{-\int_t^s \lambda_u^2 \, du} & 1 - e^{-\int_t^s \lambda_u^2 \, du} \\ 0 & 1 \end{pmatrix},$$

and it can be easily shown that the conditional transition matrix for C^2 admits the following representation:

$$\mathcal{P}^2(t, s) = \mathbb{E}_{\mathbb{Q}^*}(\mathbf{P}^2(t, s) \,|\, \mathcal{F}_t^1).$$

In particular, as can be verified by direct calculations, on the set $\{C_t^1 = 1\} = \{\tau_1 > t\}$, we have

$$p_{11}^2(t, s) = \frac{1}{\lambda_1 + \lambda_2 - \alpha_2}\left(\lambda_1 e^{-\alpha_2(s-t)} + (\lambda_2 - \alpha_2)e^{-(\lambda_1 + \lambda_2)(s-t)} \right)$$

if $\lambda_1 + \lambda_2 - \alpha_2 \ne 0$; otherwise,

$$p_{11}^2(t, s) = \left(1 + \lambda_1 t\right)e^{-(\lambda_1 + \lambda_2)(s-t)}.$$

On the set $\{C_t^1 = 2\} = \{\tau_1 \le t\}$, we have

$$p_{11}^2(t, s) = e^{-\alpha_2(s-t)}.$$

Observing that $p_{11}^2(t, s) = \mathbb{Q}^*(\tau_2 > s \,|\, \mathcal{F}_t^1 \vee \mathcal{F}_t^2)$ on the event $\{\tau_2 > t\}$, we see that the above results agree with the results of Section 1.1 (cf. Proposition 1.1).

2.2.2 Case of three rating grades

We shall now generalize the previous example to the case of three rating grades for each firm, that is, we set $K_1 = K_2 = 3$. For the primary firm, it suffices to consider an ordinary Markov chain C^1

---

[2]We refer to Bielecki and Rutkowski (2002), Chapter 11, for the definition of the conditional infinitesimal generator process.

on the state space $\mathcal{K} = \{1, 2, 3\}$ and with the generator (under the martingale probability $\mathbb{Q}^*$)

$$\Lambda^1 = \begin{pmatrix} -\lambda_{12}^1 - \lambda_{13}^1 & \lambda_{12}^1 & \lambda_{13}^1 \\ \lambda_{21}^1 & -\lambda_{21}^1 - \lambda_{23}^1 & \lambda_{23}^1 \\ 0 & 0 & 0 \end{pmatrix},$$

where the constants λ_{jk}^1 are strictly positive for every $j = 1, 2$ and $k = 1, 2, 3$ such that $j \neq k$. Thus, the state $K_1 = 3$ is the (only) absorbing state for the migration process C^1 (it corresponds to the default event). In the next step, proceeding as in Chapter 10 of Bielecki and Rutkowski (2002), we can construct the migration process C^2 as an $\mathbb{F}^1$-conditionally Markov chain on $\mathcal{K}$ with the $\mathbb{F}^1$-conditional infinitesimal generator process under $\mathbb{Q}^*$ of the following form:

$$\Lambda_t^2 = \begin{pmatrix} -\lambda_{12}^2(t) - \lambda_{13}^2(t) & \lambda_{12}^2(t) & \lambda_{13}^2(t) \\ \lambda_{21}^2(t) & -\lambda_{21}^2(t) - \lambda_{23}^2(t) & \lambda_{23}^2(t) \\ 0 & 0 & 0 \end{pmatrix},$$

where $\lambda_{jk}^2(t) = \sum_{l=1}^3 \alpha_{jk}^l \mathbb{1}_{\{C_t^1 = l\}}$ and $\alpha_{jk}^l > 0$ for every $j = 1, 2$ and $k = 1, 2, 3$ such that $j \neq k$. Similarly as in the preceding example, we introduce a matrix-valued random field $\mathbf{P}^2(t, s)$ by means of the following (random) ODE:

$$\frac{d\mathbf{P}^2(t, s)}{ds} = \mathbf{P}^2(t, s)\Lambda_s^2, \quad s \geq t, \quad \mathbf{P}^2(t, t) = \mathrm{Id}_3 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Again, we define the $\mathbb{F}^1$-conditional transition matrix $\mathcal{P}^2(t, s) = [p_{jk}^2(t, s)]_{j,k \in \mathcal{K}}$, where for every $s \geq t$

$$p_{jk}^2(t, s) := \mathbb{Q}^*(C_s^2 = k \,|\, \mathcal{F}_t^1 \vee \sigma(C_t^2))$$

on the event $\{C_t^2 = j\}$. Consequently, $\mathcal{P}^2(t, s) = \mathbb{E}_{\mathbb{Q}^*}(\mathbf{P}^2(t, s) \,|\, \mathcal{F}_t^1)$.

2.3   General Model with Credit Migrations

Our next goal is to examine from the Markovian perspective Kusuoka's set-up introduced in Section 1.2. In other words, we relax the assumption that one of the two firms is primary and the other one secondary. We shall first deal with the case of only two rating grades and we shall assume that the reference filtration $\mathbb{F}$ is trivial. Subsequently, in Section 2.3.2, we shall present results covering the general case – that is, the case of several rating grades and a non-trivial reference filtration.

2.3.1   Case of two firms and two rating grades

Since the detailed construction of the model will be given in Section 2.3.2, we shall concentrate in this section on the derivation of relevant formulae under the equivalent probability measure $\mathbb{Q}^*$. More specifically, we assume from the outset that we are given the two random times, τ_1 and τ_2, such that the $(\mathbb{H}^j, \mathbb{H}^1 \vee \mathbb{H}^2)$-martingale intensity of τ_i under the probability measure $\mathbb{Q}^*$ equals:

$$\lambda_t^{*i} = \lambda_i(1 - H_t^j) + \alpha_i H_t^j = \lambda_i \mathbb{1}_{\{\tau_j > t\}} + \alpha_i \mathbb{1}_{\{\tau_j \leq t\}}$$

for $i, j = 1, 2$, $i \neq j$. This means that for each $i = 1, 2$ the compensated jump process

$$M_t^{*i} = H_t^i - \int_0^t \lambda_u^{*i}(1 - H_u^i) \, du$$

follows a martingale under $\mathbb{Q}^*$ with respect to the joint filtration $\mathbb{G} = \mathbb{H}^1 \vee \mathbb{H}^2$.

As in Section 2.2.1, we associate with the random times τ_1 and τ_2 the two migration processes C^1 and C^2 defined as $C_t^i = 1 + H_t^i$. It is possible to check that for each $i, j = 1, 2$, $i \neq j$ the process C^i is a $\mathbb{F}^j$-conditionally Markov chain. Moreover, its conditional infinitesimal generator matrix process equals:

$$\Lambda_t^{*i} = \begin{pmatrix} -\lambda_t^{*i} & \lambda_t^{*i} \\ 0 & 0 \end{pmatrix}.$$

From now on, we shall focus on the process C^1 (the calculations below are also valid for C^2, after obvious modifications of notation). We intend to evaluate the probability

$$\mathbb{Q}^*(\tau_1 > s \,|\, \sigma(C_t^1) \vee \mathcal{F}_t^2) = \mathbb{Q}^*(C_s^1 = 1 \,|\, \sigma(C_t^1) \vee \mathcal{F}_t^2)$$

for $0 \leq t \leq s$. Due to the $\mathbb{F}^2$-conditional Markov property of C^1 under $\mathbb{Q}^*$, we have

$$\mathbb{Q}^*(C_s^1 = 1 \,|\, \sigma(C_t^1) \vee \mathcal{F}_t^2) = \mathbb{Q}^*(C_s^1 = 1 \,|\, \mathcal{F}_t^1 \vee \mathcal{F}_t^2) = \mathbb{Q}^*(\tau_1 > s \,|\, \mathcal{H}_t^1 \vee \mathcal{H}_t^2).$$

The above probability is manifestly equal to 0 on the event $\{C_t^1 = 2\} = \{\tau_1 \leq t\}$. We thus only need to compute it on the event $\{C_t^1 = 1\} = \{\tau_1 > t\}$.

Let us recall that $H_t^i = \mathbb{1}_{\{\tau_i \leq t\}} = \mathbb{1}_{\{C_t^i = 2\}}$. In order to carry out the calculations, we shall introduce a new probability measure $\mathbb{Q}$. First, we define the two auxiliary processes κ^1 and κ^2 by setting:

$$\kappa_t^1 = H_{t-}^2 \left(\frac{\alpha_1}{\lambda_1} - 1 \right), \quad \kappa_t^2 = H_{t-}^1 \left(\frac{\alpha_2}{\lambda_2} - 1 \right),$$

and we introduce the processes ξ^i, $i = 1, 2$, given by the formula:

$$\xi_t^i = -\frac{\kappa_t^i}{1 + \kappa_t^i}.$$

Next, we define a strictly positive process $\tilde{\eta}$ that solves the following SDE:

$$\tilde{\eta}_t = 1 + \sum_{i=1}^{2} \int_{]0,t]} \tilde{\eta}_{u-} \xi_u^i \, dM_u^{*i}.$$

We fix $T > 0$, and we define on $(\Omega, \mathcal{G})$ the probability measure $\mathbb{Q}$ equivalent to $\mathbb{Q}^*$ by setting:

$$\frac{d\mathbb{Q}}{d\mathbb{Q}^*} = \tilde{\eta}_T, \quad \mathbb{Q}^*\text{-a.s.}$$

Notice that

$$\frac{d\mathbb{Q}^*}{d\mathbb{Q}} = \eta_T, \quad \mathbb{Q}\text{-a.s.}$$

where the process $\eta := \tilde{\eta}^{-1}$ satisfies

$$\eta_t = 1 + \sum_{i=1}^{2} \int_{]0,t]} \eta_{u-} \kappa_u^i \, dM_u^i,$$

and where in turn for each $i = 1, 2$ the process M^i is given as

$$M_t^i = M_t^{*i} + \int_0^t \lambda_i \kappa_u^i \mathbb{1}_{\{C_u^i = 2\}} \, du = \mathbb{1}_{\{C_t^i = 2\}} - \int_0^t \lambda_i \mathbb{1}_{\{C_u^i = 1\}} \, du.$$

Notice that M^i is a martingale under $\mathbb{Q}$ with respect to the filtration $\mathbb{F}^i$, as well as with respect to the filtration $\mathbb{G}$. For the future reference, it is useful to note that $\eta_t = \eta_t^1 \eta_t^2$ where the processes η^i, $i = 1, 2$, satisfy

$$\eta_t^i = 1 + \int_{]0,t]} \eta_{u-}^i \kappa_u^i \, dM_u^i,$$

or, more explicitly,

$$\eta_t^1 = \mathbb{1}_{\{\tau_1 \leq \tau_2\}} + \mathbb{1}_{\{\tau_2 < t \leq \tau_1\}} e^{-(\alpha_1 - \lambda_1)(t - \tau_2)} + \mathbb{1}_{\{\tau_2 < t < \tau_1\}} \frac{\alpha_1}{\lambda_1} e^{-(\alpha_1 - \lambda_1)(\tau_1 - \tau_2)} + \mathbb{1}_{\{t \leq \tau_2 < \tau_1\}}$$

and

$$\eta_t^2 = \mathbb{1}_{\{\tau_2 \leq \tau_1\}} + \mathbb{1}_{\{\tau_1 < t \leq \tau_2\}} e^{-(\alpha_2 - \lambda_2)(t - \tau_1)} + \mathbb{1}_{\{\tau_1 < t < \tau_2\}} \frac{\alpha_2}{\lambda_2} e^{-(\alpha_2 - \lambda_2)(\tau_2 - \tau_1)} + \mathbb{1}_{\{t \leq \tau_1 < \tau_2\}}.$$

Also, we have that $\tilde{\eta} = \tilde{\eta}^1 \tilde{\eta}^2$, where $\tilde{\eta}^i = (\eta^i)^{-1}$.

At this point we observe that the probability measures $\mathbb{Q}^*$ and $\mathbb{Q}$ correspond to the probability measures $\mathbb{P}^*$ and $\mathbb{P}$ studied in Section 7.3 of Bielecki and Rutkowski (2002). In particular, from Proposition 7.3.1 in Bielecki and Rutkowski (2002) it follows that

$$\mathbb{1}_{\{C_t^1 = 1\}} \mathbb{Q}^*(C_s^1 = 1 \,|\, \sigma(C_t^1) \vee \mathcal{F}_t^2) \neq \mathbb{1}_{\{C_t^1 = 1\}} \mathbb{E}_{\mathbb{Q}^*}\left(e^{-\int_t^s \lambda_u^{*1} \, du} \,\Big|\, \mathcal{F}_t^2\right). \tag{2.5}$$

On the other hand, there is an important representation for the probability $\mathbb{Q}^*(C_s^1 = 1 \,|\, \sigma(C_t^1) \vee \mathcal{F}_t^2)$, that we shall provide now. For this, we need to consider another equivalent probability measure on $(\Omega, \mathcal{G})$. Thus, we define on $(\Omega, \mathcal{G})$ a probability measure $\mathbb{Q}^1$ by

$$\frac{d\mathbb{Q}^1}{d\mathbb{Q}^*} = \tilde{\eta}_T^2, \quad \mathbb{Q}^*\text{-a.s.},$$

or, equivalently, by

$$\frac{d\mathbb{Q}^1}{d\mathbb{Q}} = \eta_T^1, \quad \mathbb{Q}\text{-a.s.}$$

The probability measure $\mathbb{Q}^1$ corresponds to the probability measure $\mathbb{P}_1^*$ given by formula (7.21) in Bielecki and Rutkowski (2002). Making use of Proposition 7.3.1 in Bielecki and Rutkowski (2002), we conclude that the following result is valid.

Lemma 2.1 *For any $s \geq t$ we have*

$$\mathbb{Q}^*(C_s^1 = 1 \,|\, \mathcal{G}_t) = \mathbb{1}_{\{C_t^1 = 1\}} \mathbb{E}_{\mathbb{Q}^1}\left(e^{-\int_t^s \lambda_u^{*1} \, du} \,\Big|\, \mathcal{F}_t^2\right) \tag{2.6}$$

and

$$\mathbb{Q}^*(C_s^1 = 2 \,|\, \mathcal{G}_t) = \mathbb{1}_{\{C_t^1 = 2\}} + \mathbb{1}_{\{C_t^1 = 1\}} \mathbb{E}_{\mathbb{Q}^1}\left(1 - e^{-\int_t^s \lambda_u^{*1} \, du} \,\Big|\, \mathcal{F}_t^2\right) \tag{2.7}$$

We shall now provide an alternative method of deriving equality (2.6), which will prove useful in what follows. First, we define the two matrix-valued processes by means of the following (random) ODEs:

$$d\mathbf{P}^1(t) = \mathbf{P}^1(t) \Lambda_t^{*1} \, dt, \quad \mathbf{P}^1(0) = \mathrm{Id}_2,$$

and

$$d\mathbf{Q}^1(t) = -\Lambda_t^{*1}\mathbf{Q}^1(t)\,dt, \quad \mathbf{Q}^1(0) = \mathrm{Id}_2.$$

Observe that $\mathbf{P}^1(t)\mathbf{Q}^1(t) = \mathbf{Q}^1(t)\mathbf{P}^1(t) = \mathrm{Id}_2$ for $t \geq 0$. In fact, it is rather obvious that

$$\mathbf{P}^1(t) = \begin{pmatrix} e^{-\int_0^t \lambda_u^{*1}\,du} & 1 - e^{-\int_0^t \lambda_u^{*1}\,du} \\ 0 & 1 \end{pmatrix}$$

and

$$\mathbf{Q}^1(t) = \begin{pmatrix} e^{\int_0^t \lambda_u^{*1}\,du} & 1 - e^{\int_0^t \lambda_u^{*1}\,du} \\ 0 & 1 \end{pmatrix}.$$

Next, we fix $s \leq T$ and we define other two matrix-valued processes, for $t \in [0, s]$,

$$\mathbf{Y}(t) = \mathbb{E}_{\mathbb{Q}^1}(\mathbf{P}^1(s)\,|\,\mathcal{F}_t^2) \quad \text{and} \quad \mathbf{U}(t) = \mathbf{H}(t)\mathbf{Q}^1(t),$$

where

$$\mathbf{H}(t) = \begin{pmatrix} (1 - H_t^1) & H_t^1 \\ (1 - H_t^1) & H_t^1 \end{pmatrix} = \begin{pmatrix} \mathbb{1}_{\{C_t^1 = 1\}} & \mathbb{1}_{\{C_t^1 = 2\}} \\ \mathbb{1}_{\{C_t^1 = 1\}} & \mathbb{1}_{\{C_t^1 = 2\}} \end{pmatrix}.$$

Lemma 2.2 *The process* $\mathbf{U}$ *is a* $\mathbb{G}$*-martingale under* $\mathbb{Q}^*$ *and the process* $\mathbf{U}^* = \eta^2\mathbf{U}$ *is a* $\mathbb{G}$*-martingale under* $\mathbb{Q}^1$.

Proof. For the first statement, it is enough to check that $\mathbf{U}_{11}(t)$, $t \in [0, s]$, follows a $\mathbb{G}$-martingale under $\mathbb{Q}^*$, and this holds because

$$d\mathbf{U}_{11}(t) = -e^{\int_0^t \lambda_u^{*1}\,du}\left(dH_t^1 - \lambda_t^{*1}\,d(t \wedge \tau_1)\right) = -\mathbf{Q}_{11}^1(t)\,dM_t^{*1},$$

where $\mathbf{Q}_{11}^1(t)$ is the respective element of the matrix $\mathbf{Q}^1(t)$. The second claim follows from the fact that $d\mathbb{Q}^* = \eta_T^2\,d\mathbb{Q}^1$. $\square$

We shall now check that the process $\mathbf{Y}$ is a $\mathbb{G}$-martingale under $\mathbb{Q}^1$ as well. First, we note that this process is an $\mathbb{F}^2$-martingale under $\mathbb{Q}^1$. Next, we demonstrate the following result, which shows that the hypothesis (M.1) of Bielecki and Rutkowski (2002) is satisfied under $\mathbb{Q}^1$ by the filtrations $\mathbb{F}^2$ and $\mathbb{G}$.

Lemma 2.3 *Any* $\mathbb{F}^2$*-martingale under* $\mathbb{Q}^1$ *is also a* $\mathbb{G}$*-martingale under* $\mathbb{Q}^1$.

Proof. It is known (see, e.g., Section 6.1.1 in Bielecki and Rutkowski (2002)) that in the present set-up the condition (M.1) is equivalent to the following condition (M.2): For any $t \geq 0$, any bounded, $\mathcal{F}_\infty^2$-measurable random variable ξ, and any bounded, $\mathcal{G}_t$-measurable random variable ψ we have

$$\mathbb{E}_{\mathbb{Q}^1}(\xi\psi\,|\,\mathcal{F}_t^2) = \mathbb{E}_{\mathbb{Q}^1}(\xi\,|\,\mathcal{F}_t^2)\mathbb{E}_{\mathbb{Q}^1}(\psi\,|\,\mathcal{F}_t^2). \tag{2.8}$$

It is enough to establish (2.8) for $\xi = \mathbb{1}_{\{\tau_2 > s\}}$ and $\psi = \mathbb{1}_{\{\tau_1 \leq u\}}$ where $0 \leq u \leq t \leq s$. We have

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}^1}(\mathbb{1}_{\{\tau_2 > s\}}\mathbb{1}_{\{\tau_1 \leq u\}}\,|\,\mathcal{F}_t^2) &= \mathbb{1}_{\{\tau_2 > t\}}\frac{\mathbb{E}_{\mathbb{Q}^1}(\mathbb{1}_{\{\tau_2 > s\}}\mathbb{1}_{\{\tau_1 \leq u\}})}{\mathbb{Q}^1(\tau_2 > s)} = \mathbb{1}_{\{\tau_2 > t\}}\frac{\mathbb{E}_{\mathbb{Q}}(\eta_s^1\mathbb{1}_{\{\tau_2 > s\}}\mathbb{1}_{\{\tau_1 \leq u\}})}{\mathbb{E}_{\mathbb{Q}}(\eta_s^1\mathbb{1}_{\{\tau_2 > s\}})} \\ &= \mathbb{1}_{\{\tau_2 > t\}}\,e^{\lambda_2(t - s)}(1 - e^{-\lambda_1 u}). \end{aligned}$$

On the other hand

$$\mathbb{Q}^1(\tau_2 > s \,|\, \mathcal{F}_t^2) = \mathbb{1}_{\{\tau_2 > t\}} \frac{\mathbb{Q}^1(\tau_2 > s)}{\mathbb{Q}^1(\tau_2 > t)} = \mathbb{1}_{\{\tau_2 > t\}} e^{\lambda_2(t-s)},$$

and

$$\mathbb{1}_{\{\tau_2 > t\}} \mathbb{Q}^1(\tau_1 \le u \,|\, \mathcal{F}_t^2) = \mathbb{1}_{\{\tau_2 > t\}} \frac{\mathbb{E}_{\mathbb{Q}^1}\big(\mathbb{1}_{\{\tau_2 > t\}} \mathbb{1}_{\{\tau_1 \le u\}}\big)}{\mathbb{Q}^1(\tau_2 > t)} = \mathbb{1}_{\{\tau_2 > t\}}(1 - e^{-\lambda_1 u}).$$

This completes the proof of the lemma.   □

Corollary 2.1 *The process* $\mathbf{Y}$ *is a* $\mathbb{G}$*-martingale under* $\mathbb{Q}^1$.

The following proposition generalizes Lemma 2.1.

Proposition 2.4 *For any* $t \le s \le T$, *we have*

$$\mathbb{E}_{\mathbb{Q}^*}(\mathbf{H}(s) \,|\, \mathcal{G}_t) = \mathbf{H}(t) \, \mathbb{E}_{\mathbb{Q}^1}(\mathbf{Q}^1(t)\mathbf{P}^1(s) \,|\, \mathcal{F}_t^2).$$

Moreover, for any $\mathbb{Q}^1$*-integrable,* $\mathcal{F}_s^2$*-measurable random variable* Y *we have*

$$\mathbb{E}_{\mathbb{Q}^*}(\mathbf{H}(s)Y \,|\, \mathcal{G}_t) = \mathbf{H}(t) \, \mathbb{E}_{\mathbb{Q}^1}(\mathbf{Q}^1(t)\mathbf{P}^1(s)Y \,|\, \mathcal{F}_t^2).$$

Proof. We know that the processes $\mathbf{Y}$ and $\mathbf{U}^*$ are $\mathbb{G}$-martingales under $\mathbb{Q}^1$. Since the jump time of the process of finite variation $\mathbf{U}^*$ does not coincide with the jump time of the process $\mathbf{Y}$, the Itô product rule shows that the process $\mathbf{U}^*\mathbf{Y}$ also follows a $\mathbb{G}$-martingale under $\mathbb{Q}^1$. Thus, for $t \le s$ we obtain

$$\mathbb{E}_{\mathbb{Q}^1}(\mathbf{U}^*(s)\mathbf{Y}(s) \,|\, \mathcal{G}_t) = \mathbf{U}^*(t)\mathbf{Y}(t) = \eta_t^2 \mathbf{H}(t) \, \mathbb{E}_{\mathbb{Q}^1}(\mathbf{Q}^1(t)\mathbf{P}^1(s) \,|\, \mathcal{F}_t^2)$$

since $\mathbf{Q}^1(t)$ is $\mathcal{F}_t^2$-measurable. On the other hand,

$$\mathbb{E}_{\mathbb{Q}^1}(\mathbf{U}^*(s)\mathbf{Y}(s) \,|\, \mathcal{G}_t) = \mathbb{E}_{\mathbb{Q}^1}(\eta_s^2 \mathbf{H}(s)\mathbf{Q}^1(s)\mathbf{P}^1(s) \,|\, \mathcal{G}_t) = \mathbb{E}_{\mathbb{Q}^1}(\eta_s^2 \mathbf{H}(s) \,|\, \mathcal{G}_t)$$

since $\mathbf{Q}^1(s)\mathbf{P}^1(s) = 1$. Combining the equalities above, we obtain (2.4). Notice that equality (2.4) implies (2.6), in particular. To establish equality (2.4), it suffices to consider the process $\mathbf{Y}(t) = \mathbb{E}_{\mathbb{Q}^1}(\mathbf{P}^1(s)Y \,|\, \mathcal{F}_t^2)$ so that $\mathbf{Y}(s) = \mathbf{P}^1(s)Y$.   □

To complete the example let us compute the probability under $\mathbb{Q}^*$ that the first default occurs after time t, that is: $\mathbb{Q}^*(C_t^i \ne 2, \, i = 1, 2)$. According to Proposition 7.2.2 (or Lemma 7.3.1) in Bielecki and Rutkowski (2002), we have that

$$\mathbb{Q}^*(C_t^i \ne 2, \, i = 1, 2) = \mathbb{Q}^*(\hat{\tau} > t) = \mathbb{Q}^1(\hat{\tau} > t) = \mathbb{Q}(\hat{\tau} > t) = e^{-t(\lambda_1 + \lambda_2)},$$

where $\hat{\tau} := \inf\{t \ge 0 \,:\, C_t^1 = 2 \text{ or } C_t^2 = 2\}$. This result is expected, as the mutual dependence of default times appears only after the first default occurs. In other words, τ_1 and τ_2 are not independent under $\mathbb{Q}^*$, but may be treated as mutually independent if we are interested in the law of the minimum $\tau_1 \wedge \tau_2$ under $\mathbb{Q}^*$, and the probability measure $\mathbb{Q}^*$ may be replaced by $\mathbb{Q}$ in the calculation of this law. Later on we shall generalize this useful property.

Remarks. We have already discussed an interpretation of the martingale intensity λ^{*1}. It is apparent from this discussion that it is indeed the "local" (or infinitesimal) conditional intensity of the default

time τ_1, relative to the probability measure $\mathbb{Q}^*$. It is an interesting question whether there exists an $\mathbb{F}^2$-progressively measurable process, say $\tilde{\lambda}^1$, so that

$$\mathbb{1}_{\{C_t^1=1\}}\mathbb{Q}^*(C_s^1 = 1 \mid \sigma(C_t^1) \vee \mathcal{F}_t^2) = \mathbb{1}_{\{C_t^1=1\}}\mathbb{E}_{\mathbb{Q}^*}\left(e^{-\int_t^s \tilde{\lambda}_u^1 \, du} \, \Big| \, \mathcal{F}_t^2\right).$$

If so, then we could term such a process the $\mathbb{F}^2$-intensity of default time τ_1, relative to the probability measure $\mathbb{Q}^*$. We can not verify whether such a process exists, though. However, the following holds,

$$\mathbb{1}_{\{C_t^1=1\}}\mathbb{Q}^*(C_s^1 = 1 \mid \sigma(C_t^1) \vee \mathcal{F}_t^2) = \mathbb{1}_{\{C_t^1=1\}}\mathbb{E}_{\mathbb{Q}^1}\left(e^{-\int_t^s \lambda_u^{*1} \, du} \, \Big| \, \mathcal{F}_t^2\right)$$

$$= \mathbb{1}_{\{C_t^1=1\}}\mathbb{E}_{\mathbb{Q}^*}\left(\frac{\eta_T^2}{\eta_t^2}e^{-\int_t^s \lambda_u^{*1} \, du} \, \Big| \, \mathcal{F}_t^2\right) = \mathbb{1}_{\{C_t^1=1\}}\mathbb{E}_{\mathbb{Q}^*}\left(e^{S_t^1 - S_s^1} \, \Big| \, \mathcal{F}_t^2\right)$$

where the explicit formula for S^1 is easy to derive.

2.3.2 Case of two firms and multiple rating grades

We postulate that $K_1, K_2 \geq 2$ so that C^i takes values in $\mathcal{K}_i = \{1, 2, \ldots, K_i\}$. We first consider the two $\mathbb{F}$-conditional Markov chains C^i, $i = 1, 2$, defined on the underlying probability space $(\Omega, \mathcal{G}, \mathbb{Q})$, with the corresponding infinitesimal generators:

$$\Lambda_t^i = \begin{pmatrix} \lambda_{11}^i(t) & \lambda_{12}^i(t) & . & \lambda_{1K_i}^i(t) \\ \lambda_{21}^i(t) & \lambda_{22}^i(t) & . & \lambda_{2K_i}^i(t) \\ . & . & . & . \\ \lambda_{K_i-1,1}^i(t) & \lambda_{K_i-1,2}^i(t) & . & \lambda_{K_i-1,K_i}^i(t) \\ 0 & 0 & . & 0 \end{pmatrix}$$

where λ_{km}^i are strictly positive $\mathbb{F}$-progressively measurable stochastic processes for $k = 1, 2, \ldots, K_i-1$ and $m = 1, 2, \ldots, K_i$, $m \neq k$, and $\lambda_{kk}^i = -\sum_{m=1,m\neq k}^{K_i} \lambda_{km}^i$ for $k = 1, 2, \ldots, K_i-1$. In particular, the state $k = K_i$ is the (only) absorbing state for each chain. For the sake of notational simplicity, we postulate that $C_0^1 = C_0^2 = 1$ (this assumption can be easily relaxed). By construction, the processes C^1 and C^2 are also conditionally independent under $\mathbb{Q}$ with respect to the reference filtration $\mathbb{F}$.

In addition, we assume that we are given the family of stochastic matrices for $i = 1, 2$ and $l = 2, \ldots, K_j$, $j \neq i$,

$$\Lambda_t^{i|l} = \begin{pmatrix} \lambda_{11}^{i|l}(t) & \lambda_{12}^{i|l}(t) & . & \lambda_{1K_i}^{i|l}(t) \\ \lambda_{21}^{i|l}(t) & \lambda_{22}^{i|l}(t) & . & \lambda_{2K_i}^{i|l}(t) \\ . & . & . & . \\ \lambda_{K_i-1,1}^{i|l}(t) & \lambda_{K_i-1,2}^{i|l}(t) & . & \lambda_{K_i-1,K_i}^{i|l}(t) \\ 0 & 0 & . & 0 \end{pmatrix}$$

where $\lambda_{km}^{i|l}$ are strictly positive $\mathbb{F}$-progressively measurable stochastic processes for $k = 1, 2, \ldots, K_i-1$ and $m = 1, 2, \ldots, K_i$, $m \neq k$, and $\lambda_{kk}^{i|l} = -\sum_{m=1,m\neq k}^{K_i} \lambda_{km}^{i|l}$ for $k = 1, 2, \ldots, K_i - 1$. At the intuitive level, at any date $t > 0$, the entry $\lambda_{km}^{i|l}(t) > 0$ represents the intensity of transition from k to m over the time interval $[t, t + dt]$ by the process C^i conditional on the event $\{C_t^j = l\}$, where $j \neq i$ and

$l \in \mathcal{K}_j$. Notice that under the present convention concerning the initial condition, it is natural to set: $\Lambda^{i|1} = \Lambda^i$ for $i = 1, 2$. Of course, we need first to construct the process (C^1, C^2) with the above described properties. To this end, we shall introduce a probability measure $\mathbb{Q}^*$ equivalent to $\mathbb{Q}$ on $(\Omega, \mathcal{G})$. For $i = 1, 2$ we define the processes κ^i_{km}, $k = 1, 2, \ldots, K_i - 1$, $m = 1, 2, \ldots, K_i$, $m \neq k$, as follows:

$$\kappa^i_{km}(t) = \sum_{l=2}^{K_j} H^j_l(t-) \left(\frac{\lambda^{i|l}_{km}(t)}{\lambda^i_{km}(t)} - 1 \right), \quad j \neq i,$$

where we write $H^i_k(t) = \mathbb{1}_{\{C^i_t = k\}}$ for $i = 1, 2$ and $k \in \mathcal{K}_i$. We also define, for $i = 1, 2$ and any two states $k \neq m$, the transition counting process $H^i_{km}(t) = \sum_{0 < u \leq t} H^i_k(u-) H^i_m(u)$. Let us recall that for $i = 1, 2$ the process M^i_{km} given by the expression:

$$M^i_{km}(t) = H^i_{km}(t) - \int_0^t \lambda^i_{km}(u) H^i_k(u) \, du, \quad m \neq k$$

is known to follow a $\mathbb{G}^i$-martingale under $\mathbb{Q}$, and thus also a $\mathbb{G}$-martingale under $\mathbb{Q}$, where $\mathbb{G}^i = \mathbb{F} \vee \mathbb{F}^i$ and $\mathbb{G} = \mathbb{F} \vee \mathbb{F}^1 \vee \mathbb{F}^2$.

We can now define a strictly positive martingale η by means of the following SDE:

$$\eta_t = 1 + \sum_{i=1}^{2} \int_{]0,t]} \sum_{k=1}^{K_i-1} \sum_{m=1,m\neq k}^{K_i} \eta_{u-} \kappa^i_{km}(u) \, dM^i_{km}(u).$$

For a fixed, but otherwise arbitrary, $T > 0$ the probability measure $\mathbb{Q}^*$ is given through the formula:

$$\frac{d\mathbb{Q}^*}{d\mathbb{Q}} = \eta_T, \quad \mathbb{Q}\text{-a.s.}$$

Recall that we have denoted $\mathbb{G}^j = \mathbb{F} \vee \mathbb{F}^j$ and $\lambda^{i|1}_{km}(t) = \lambda^i_{km}(t)$.

Proposition 2.5 *For any $i, j = 1, 2$, $i \neq j$ the process C^i is a $\mathbb{G}^j$-conditional $\mathbb{G}$-Markov chain under $\mathbb{Q}^*$ and its transition $\mathbb{G}^j$-intensities under $\mathbb{Q}^*$ are:*

$$\lambda^{*i}_{km}(t) = \left(1 + \kappa^i_{km}(t) \right) \lambda^i_{km}(t) = \sum_{l=1}^{K_j} \mathbb{1}_{\{C^j_t = l\}} \lambda^{i|l}_{km}(t), \quad k \neq m.$$

Proof. The conditional Markov property follows from the combination of: (i) the fact that the density η only depends on the filtration $\mathbb{F}$ and the process $C = (C^1, C^2)$, (ii) the abstract Bayes formula, and (iii) the fact that the process C is a Markov chain with a given $\mathbb{F}$-intensity matrix under $\mathbb{Q}$. The processes λ^{*i}_{km} are the corresponding $\mathbb{G}^j$-intensities in view of the fact that the processes M^{*i}_{km} defined as

$$M^{*i}_{km}(t) = H^i_{km}(t) - \int_0^t \lambda^{*i}_{km}(u) H^i_k(u) \, du, \quad k \neq m,$$

are $\mathbb{G}$-martingales under $\mathbb{Q}^*$. $\square$

Similarly as in the previous section, we shall now introduce some auxiliary probability measures. For this, we define two densities η^i, $i = 1, 2$ by means of the following SDE:

$$\eta^i_t = 1 + \int_{]0,t]} \sum_{k=1}^{K_i-1} \sum_{m=1,m\neq k}^{K_i} \eta^i_{u-} \kappa^i_{km}(u) \, dM^i_{km}(u).$$

As before, we note that $\eta = \eta^1 \eta^2$. This easily follows from the integration by parts formula for the purely discontinuous martingales, combined with the fact that jumps of the processes C^1 and C^2 coincide with zero $\mathbb{Q}$ probability. Next, we also set $\tilde{\eta} := \eta^{-1}$ and $\tilde{\eta}^i := \left(\eta^i\right)^{-1}$ for $i = 1, 2$, so that $\tilde{\eta} = \tilde{\eta}^1 \tilde{\eta}^2$. We are now in the position to define a probability measure $\mathbb{Q}^1$ on $(\Omega, \mathcal{G})$ by setting:

$$\frac{d\mathbb{Q}^1}{d\mathbb{Q}^*} = \tilde{\eta}_T^2, \quad \mathbb{Q}^*\text{-a.s.}, \quad \frac{d\mathbb{Q}^*}{d\mathbb{Q}^1} = \eta_T^2, \quad \mathbb{Q}^1\text{-a.s.}$$

Equivalently,

$$\frac{d\mathbb{Q}^1}{d\mathbb{Q}} = \eta_T^1, \quad \mathbb{Q}\text{-a.s.}, \quad \frac{d\mathbb{Q}}{d\mathbb{Q}^1} = \tilde{\eta}_T^1, \quad \mathbb{Q}^1\text{-a.s.}$$

We shall now proceed similarly as in Subsection 2.2.2. That is, we define two matrix-valued processes by means of the following (random) ODEs:

$$d\mathbf{P}^1(t) = \mathbf{P}^1(t)\Lambda_t^{*1}\, dt, \quad \mathbf{P}^1(0) = \mathrm{Id}_{\mathrm{K}_1},$$

$$d\mathbf{Q}^1(t) = -\Lambda_t^{*1}\mathbf{Q}^1(t)\, dt, \quad \mathbf{Q}^1(0) = \mathrm{Id}_{\mathrm{K}_1},$$

where $\Lambda_t^{*1} = [\lambda_{km}^{*1}(t)]_{k,m\in\mathcal{K}_1}$. Observe that $\mathbf{P}^1(t)\mathbf{Q}^1(t) = \mathrm{Id}_{\mathrm{K}_1}$ for $t \geq 0$. Indeed, we have

$$d(\mathbf{P}^1(t)\mathbf{Q}^1(t)) = \mathbf{P}^1(t)\, d\mathbf{Q}^1(t) + \mathbf{Q}^1(t)\, d\mathbf{P}^1(t) = -\mathbf{P}^1(t)\Lambda_t^{*1}\mathbf{Q}^1(t)\, dt + \mathbf{P}^1(t)\Lambda_t^{*1}\mathbf{Q}^1(t)\, dt = \mathbf{0}$$

and obviously $\mathbf{P}^1(0)\mathbf{Q}^1(0) = \mathrm{Id}_{\mathrm{K}_1}$.

Let us fix $s \leq T$. We define the two matrix-valued processes $\mathbf{Y}$ and $\mathbf{U}$ by setting, for $t \in [0, s]$,

$$\mathbf{Y}(t) = \mathbb{E}_{\mathbb{Q}^1}(\mathbf{P}^1(s)Y \,|\, \mathcal{G}_t^2), \quad \mathbf{U}(t) = \mathbf{H}(t)\mathbf{Q}^1(t),$$

where Y is a $\mathbb{Q}^1$-integrable, $\mathcal{G}_s^2$-measurable random variable, and $\mathbf{H}(t)$ is the $K_1 \times K_1$ matrix:

$$\mathbf{H}(t) = \begin{pmatrix} H_1^1(t) & H_2^1(t) & . & H_{K_1}^1(t) \\ H_1^1(t) & H_2^1(t) & . & H_{K_1}^1(t) \\ . & . & . & . \\ H_1^1(t) & H_2^1(t) & . & H_{K_1}^1(t) \end{pmatrix}.$$

Lemma 2.4 *The process $\mathbf{U}$ is a $\mathbb{G}$-martingale under $\mathbb{Q}^*$ and the process $\tilde{\mathbf{U}} = \eta^2 \mathbf{U}$ is a $\mathbb{G}$-martingale under $\mathbb{Q}^1$.*

Proof. We first observe that for $t \leq s$

$$d\mathbf{U}_{kk'}(t) = \sum_{l=1}^{K_1} \mathbf{Q}_{lk'}^1(t)\, dH_l^1(t) - \sum_{m=1}^{K_1-1}\sum_{l=1}^{K_1} H_m^1(t)\lambda_{ml}^{*1}(t)\mathbf{Q}_{lk'}^1(t)\, dt$$

$$= \sum_{l=1}^{K_1} \mathbf{Q}_{lk'}^1(t)\Big(dH_l^1(t) - \sum_{m=1}^{K_1-1} H_m^1(t)\lambda_{ml}^{*1}(t)\, dt\Big).$$

Let us fix $l = 1$, and let us analyze the expression:

$$dH_1^1(t) - \sum_{m=1}^{K_1-1} H_m^1(t)\lambda_{m1}^{*1}(t)\, dt.$$

First, we observe that

$$dH_1^1(t) = \sum_{m=2}^{K_1-1} dH_{m1}^1(t) - \sum_{m=2}^{K_1} dH_{1m}^1(t).$$

Thus, noting that $\lambda_{11}^{*1}(t) = -\sum_{m=2}^{K_1} \lambda_{1m}^{*1}(t)$, we obtain

$$dH_1^1(t) - \sum_{m=1}^{K_1-1} H_m^1(t)\lambda_{m1}^{*1}(t)\,dt = \sum_{m=2}^{K_1-1} dM_{m1}^{*1}(t) - \sum_{m=2}^{K_1-1} dM_{1m}^{*1}(t).$$

The above argument carries over for $l = 2, \ldots, K_1 - 1$, and thus

$$dH_l^1(t) - \sum_{m=1}^{K_1-1} H_m^l(t)\lambda_{ml}^{*1}(t)\,dt = \sum_{m=1,m\neq l}^{K_1-1} dM_{ml}^{*1}(t) - \sum_{m=1,m\neq l}^{K_1} dM_{lm}^{*1}(t).$$

Finally, for $l = K_1$ we have

$$dH_{K_1}^1(t) - \sum_{m=1}^{K_1-1} H_m^1(t)\lambda_{mK_1}^{*1}(t)\,dt = dM_{K_1}^{*1}(t),$$

where

$$M_{K_1}^{*1}(t) := H_{K_1}^1(t) - \sum_{m=1}^{K_1-1} H_m^1(t)\lambda_{mK_1}^{*1}(t)$$

is a $\mathbb{G}$-martingale under $\mathbb{Q}^*$. The above observations prove the first statement. The second claim is now obvious, since we have $d\mathbb{Q}^* = \eta_T^2\,d\mathbb{Q}^1$.   □

The next natural step would be to demonstrate that the process $\mathbf{Y}$ is a $\mathbb{G}$-martingale under $\mathbb{Q}^1$ (obviously $\mathbf{Y}$ is a $\mathbb{G}^2$-martingale under $\mathbb{Q}^1$). To this end, it would be sufficient to prove the following conjecture: *any $\mathbb{G}^2$-martingale under $\mathbb{Q}^1$ is also a $\mathbb{G}$-martingale under $\mathbb{Q}^1$*. Due to the somewhat complicated mutual dependence of transition intensities it does not seem plausible that this conjecture holds true, however. For this reason, we take a slightly different (indeed, less general) approach. We introduce the stopped migration process $\tilde{C}_t^1 = C_t^1 \mathbb{1}_{\{\tau_1 > t\}} + C_{\tau_1}^1 \mathbb{1}_{\{\tau_1 \leq t\}}$ where $\tau_1 :=$ $\inf\{t \in \mathbb{R}_+ : C_t^1 \neq C_0^1\}$. Let $\tilde{\mathbb{F}}^1$ be the filtration generated by the process $\tilde{C}^1$ and let $\tilde{\mathbb{G}}^1 = \mathbb{F} \vee \tilde{\mathbb{F}}^1 \vee \mathbb{F}^2$ be the reduced filtration. Then we have the following counterpart of Lemma 2.3.

Lemma 2.5 *Any $\mathbb{G}^2$-martingale under $\mathbb{Q}^1$ is also a $\tilde{\mathbb{G}}^1$-martingale under $\mathbb{Q}^1$.*

Proof. Along the similar lines as the proof of Lemma 2.3.   □

As usual, we consider $t \leq s \leq T$. The auxiliary process $\tilde{Y}$ is given by (recall that $C_0^1 = 1$)

$$\tilde{Y}_t = \mathbb{E}_{\mathbb{Q}^1}\left(Y e^{-\int_0^t \lambda_{11}^{*1}(u)\,du} \,\Big|\, \mathcal{G}_t^2\right)$$

where Y is a $\mathbb{Q}^1$-integrable, $\mathcal{G}_s^2$-measurable random variable. In view of Lemma 2.5, it is clear that the following corollary is valid.

Corollary 2.2 *The process $\tilde{Y}$ is a $\tilde{\mathbb{G}}^1$-martingale under $\mathbb{Q}^1$.*

Furthermore, we set

$$\tilde{U}_t = \mathbb{1}_{\{\tilde{C}_t^1=1\}} \, e^{\int_0^t \lambda_{11}^{*1}(u)\,du} = \mathbb{1}_{\{\tau_1>t\}} \, e^{\int_0^t \lambda_{11}^{*1}(u)\,du}.$$

Lemma 2.6 *The process $\tilde{U}$ is a $\tilde{\mathbb{G}}^1$-martingale under $\mathbb{Q}^*$ and the product $U^* = \eta^2 \tilde{U}$ is a $\tilde{\mathbb{G}}^1$-martingale under $\mathbb{Q}^1$.*

Proposition 2.6 *For any $t \le s \le T$, and any $\mathbb{Q}^1$-integrable, $\mathcal{G}_s^2$-measurable random variable Y we have*

$$\mathbb{E}_{\mathbb{Q}^*}(\mathbb{1}_{\{\tilde{C}_s^1=1\}} Y \,|\, \mathcal{G}_t) = \mathbb{E}_{\mathbb{Q}^*}(\mathbb{1}_{\{\tilde{C}_s^1=1\}} Y \,|\, \tilde{\mathcal{G}}_t^1) = \mathbb{1}_{\{\tilde{C}_t^1=1\}} \, \mathbb{E}_{\mathbb{Q}^1}\left(Y e^{-\int_t^s \lambda_{11}^{*1}(u)\,du} \,\Big|\, \mathcal{G}_t^2 \right).$$

Proof. We shall mimic the proof of Proposition 2.4. We know that the processes $\tilde{Y}$ and U^* are $\tilde{\mathbb{G}}$-martingales under $\mathbb{Q}^1$. Since the jump time of the process of finite variation U^* does not coincide with the jump time of the process $\tilde{Y}$, the Itô product rule shows that the process $U^*\tilde{Y}$ also follows a $\tilde{\mathbb{G}}$-martingale under $\mathbb{Q}^1$. Thus, for $t \le s$ we obtain

$$\mathbb{E}_{\mathbb{Q}^1}(U_s^* \tilde{Y}_s \,|\, \tilde{\mathcal{G}}_t^1) = U_t^* \tilde{Y}_t = \eta_t^2 \mathbb{1}_{\{\tilde{C}_t^1=1\}} \, \mathbb{E}_{\mathbb{Q}^1}\left(Y e^{-\int_t^s \lambda_{11}^{*1}(u)\,du} \,\Big|\, \tilde{\mathcal{G}}_t^2 \right).$$

On the other hand,

$$\mathbb{E}_{\mathbb{Q}^1}(U_s^* \tilde{Y}_s \,|\, \tilde{\mathcal{G}}_t^1) = \mathbb{E}_{\mathbb{Q}^1}(\eta_s^2 \mathbb{1}_{\{\tilde{C}_t^1=1\}} Y \,|\, \tilde{\mathcal{G}}_t^1).$$

This ends the proof.   □

Corollary 2.3 *For any $s \ge t$ we have*

$$\mathbb{Q}^*(\tau_1 > s \,|\, \mathcal{G}_t) = \mathbb{Q}^*(\tau_1 > s \,|\, \tilde{\mathcal{G}}_t^1) = \mathbb{1}_{\{\tau_1>t\}} \, \mathbb{E}_{\mathbb{Q}^1}\left(e^{-\int_t^s \lambda_{11}^{*1}(u)\,du} \,\Big|\, \mathcal{G}_t^2 \right).$$

and

$$\mathbb{Q}^*(\tau_1 \le s \,|\, \mathcal{G}_t) = \mathbb{Q}^*(\tau_1 \le s \,|\, \tilde{\mathcal{G}}_t^1) = \mathbb{1}_{\{\tau_1 \le t\}} - \mathbb{1}_{\{\tau_1>t\}} \, \mathbb{E}_{\mathbb{Q}^1}\left(1 - e^{-\int_t^s \lambda_{11}^{*1}(u)\,du} \,\Big|\, \mathcal{G}_t^2 \right).$$

2.3.3 Case of several firms and multiple rating grades

In the case of n firms and several rating grades ($K_i \ge 2$ for the ith firm) the notation becomes rather heavy, and thus it is rather difficult to present a general, but in the same time reasonably transparent, result. Therefore, we shall concentrate on a specific issue of a risk-neutral valuation of a first-to-default swap. Unfortunately, since we assume here that the number of reference entities $n \ge 3$ this case is not directly covered by our previous results.

In the case of n reference entities, we consider the credit migration process $C = (C^1, \ldots, C^n)$. The construction of this process is based on the assumption that under the probability measure $\mathbb{Q}$ the following properties hold:

- each process C^i is an $\mathbb{F}$-conditionally $\mathbb{G}$-Markov chain with values in $\mathcal{K}_i = \{1, \ldots, K_i\}$ and the generator matrix Λ^i, where $\mathbb{F}$ is the reference filtration and $\mathbb{G} = \mathbb{F} \vee \mathbb{F}^1 \vee \ldots \vee \mathbb{F}^n$,

- processes C^i are $\mathbb{F}$-conditionally independent under $\mathbb{Q}$.

In addition, we are given a family of stochastic matrices $\Lambda^{i|l_1,\dots,l_{i-1},l_{i+1},\dots,l_n}$ for $i = 1,\dots,n$ and $l_j \in \mathcal{K}_j$ for any $j \neq i$. As before, to simplify the notation we adopt the convention that $(C_0^1,\dots,C_0^n) = (1,\dots,1)$. Notice that under this convention we have: $\lambda_{km}^{i|1,\dots,1}(t) = \lambda_{km}^i(t)$.

We shall use our standard notation for the indicator process $H_l^j(t) = \mathbb{1}_{\{C_t^j = l\}}$ and for the transition counting process $H_{km}^i(t) = \sum_{0 < u \leq t} H_k^i(u-)H_m^i(u)$.

In the second step, we introduce a probability measure $\mathbb{Q}^*$ equivalent to $\mathbb{Q}$ on $(\Omega, \mathcal{G})$. To this end, we observe that for each $i = 1,\dots,n$ the process

$$M_{km}^i(t) = H_{km}^i(t) - \int_0^t \lambda_{km}^i(u)H_k^i(u)\,du, \quad m \neq k$$

follows a $\mathbb{G}$-martingale under $\mathbb{Q}$. We set

$$\frac{d\mathbb{Q}^*}{d\mathbb{Q}} = \eta_T = \eta_T^1 \dots \eta_T^n, \quad \mathbb{Q}\text{-a.s.}$$

where for each i the process η^i solves the SDE:

$$\eta_t^i = 1 + \int_{]0,t]} \sum_{k=1}^{K_i-1} \sum_{m=1,m\neq k}^{K_i} \eta_{u-}^i \kappa_{km}^i(u)\,dM_{km}^i(u).$$

Furthermore,

$$\kappa_{km}^i(t) = \sum_{l_1=2}^{K_1} \dots \sum_{l_n=2}^{K_n} H_{l_1}^1(t-) \dots H_{l_n}^n(t-)\left(\frac{\lambda_{km}^{i|l_1,\dots,l_{i-1},l_{i+1},\dots,l_n}(t)}{\lambda_{km}^i(t)} - 1\right),$$

where the summation is over all state spaces except for $\mathcal{K}_i$ (i.e., the sum $\sum_{l_i=2}^{K_i}$ is not present). Let $\mathbb{G}^{(i)} = \mathbb{F} \vee \mathbb{F}^1 \vee \dots \vee \mathbb{F}^{i-1} \vee \mathbb{F}^{i+1} \vee \dots \vee \mathbb{F}^n$. The following result is a straightforward generalization Proposition 2.5 and thus its proof its omitted.

Proposition 2.7 *For any $i = 1, 2, \dots, n$ the process C^i is a $\mathbb{G}^{(i)}$-conditional $\mathbb{G}$-Markov chain under $\mathbb{Q}^*$ and its transition $\mathbb{G}^{(i)}$-intensities under $\mathbb{Q}^*$ are:*

$$\lambda_{km}^{*i}(t) = \left(1 + \kappa_{km}^i(t)\right)\lambda_{km}^i(t) = \sum_{l_1=2}^{K_1} \dots \sum_{l_n=2}^{K_n} H_{l_1}^1(t) \dots H_{l_n}^n(t)\lambda_{km}^{i|l_1,\dots,l_{i-1},l_{i+1},\dots,l_n}(t).$$

From now on we fix $p < n$ and we introduce the probability measure $\hat{\mathbb{Q}}$ by setting:

$$\frac{d\hat{\mathbb{Q}}}{d\mathbb{Q}} = \hat{\eta}_T = \hat{\eta}_T^1 \dots \hat{\eta}_T^n, \quad \mathbb{Q}\text{-a.s.,} \tag{2.9}$$

where

$$\hat{\eta}_t^i = 1 + \int_{]0,t]} \sum_{k=1}^{K_i-1} \sum_{m=1,m\neq k}^{K_i} \hat{\eta}_{u-}^i \hat{\kappa}_{km}^i(u)\,dM_{km}^i(u)$$

and (as before, the sum $\sum_{l_i=2}^{K_i}$ is not present)

$$\hat{\kappa}_{km}^i(t) = \sum_{l_1=2}^{K_1} \dots \sum_{l_p=2}^{K_p} H_{l_1}^1(t-) \dots H_{l_p}^p(t-)\left(\frac{\lambda_{km}^{i|l_1,\dots,l_p,1\dots,1}(t)}{\lambda_{km}^i(t)} - 1\right).$$

In words, under the auxiliary probability measure $\hat{\mathbb{Q}}$ the entities with indices $1, \ldots, p$ can be seen as 'primary firms' and the entities with indices $p+1, \ldots, n$ as 'secondary firms' in the sense of Jarrow and Yu (2001). The following result is a corollary to Proposition 2.7.

Corollary 2.4 *For any $i = 1, 2, \ldots, n$ the process C^i is a $\mathbb{G}^{(i)}$-conditional $\mathbb{G}$-Markov chain under $\hat{\mathbb{Q}}$ and its transition $\mathbb{G}^{(i)}$-intensities under $\hat{\mathbb{Q}}$ are:*

$$\hat{\lambda}^i_{km}(t) = \left(1 + \hat{\kappa}^i_{km}(t)\right)\lambda^i_{km}(t) = \sum_{l_1=2}^{K_1} \ldots \sum_{l_p=2}^{K_p} H^1_{l_1}(t) \ldots H^p_{l_p}(t)\lambda^{i|l_1,\ldots,l_p,1\ldots,1}_{km}(t).$$

Let $\hat{\mathbb{G}} = \mathbb{F} \vee \mathbb{F}^1 \vee \ldots \mathbb{F}^p$. Due to the specific choice of $\hat{\mathbb{Q}}$ we have also the following result that will prove useful in the valuation of a first-to-default swap. For each $i = p+1, \ldots, n$ we introduce the stopped process: $\tilde{C}^i_t = C^i_t \mathbb{1}_{\{\tau_i > t\}} + C^i_{\tilde{\tau}} \mathbb{1}_{\{\tau_i \leq t\}}$, where $\tau_i = \inf\{t \in \mathbb{R}_+ : C^i_t \neq C^i_0\}$.

Lemma 2.7 *The following statements are valid.*
(i) *Migration processes $C^{p+1}, \ldots, C^n$ follow $\hat{\mathbb{G}}$-conditionally independent, $\hat{\mathbb{G}}$-conditional $\mathbb{G}$-Markov chains under $\hat{\mathbb{Q}}$.*
(ii) *The random times $\tau_{p+1}, \ldots, \tau_n$ are $\hat{\mathbb{G}}$-conditionally independent under $\hat{\mathbb{Q}}$ with the $\hat{\mathbb{G}}$-hazard processes $\hat{\Gamma}^i_t = \sum_{j=2}^{K_i} \int_0^t \hat{\lambda}^i_{1j}(u)\,du$ for $i = p+1, \ldots, n$.*
(iii) *The joint ($\mathbb{F}$-conditional) laws of processes $C^1, \ldots, C^p$ and random times $\tau_{p+1}, \ldots, \tau_n$ under $\hat{\mathbb{Q}}$ and under $\mathbb{Q}^*$ are identical.*
(iv) *The random times $\tau_{p+1}, \ldots, \tau_n$ are $\hat{\mathbb{G}}$-conditionally independent under under $\mathbb{Q}^*$ with the $\hat{\mathbb{G}}$-hazard processes $\hat{\Gamma}^i_t$.*

Proof. The first two statements are direct consequences of the construction of the process $(C^1, \ldots, C^n)$ and the definitions of probability measures $\hat{\mathbb{Q}}$ and $\mathbb{Q}^*$. Part (iii) follows from the rather obvious fact that all probabilistic properties of the process $(C^1, \ldots, C^p, \tilde{C}^{p+1}, \ldots, \tilde{C}^n)$ are the same under $\hat{\mathbb{Q}}$ and $\mathbb{Q}^*$. The last statement is a consequence of (ii) and (iii). $\square$

Let $\tilde{\mathbb{F}}^i$ stand for the filtration generated by $\tilde{C}^i$ and let $\tilde{\mathbb{G}} = \mathbb{F} \vee \mathbb{F}^1 \vee \ldots \mathbb{F}^p \vee \tilde{\mathbb{F}}^{p+1} \vee \ldots \vee \tilde{\mathbb{F}}^n$.

Lemma 2.8 *Any $\hat{\mathbb{G}}$-martingale under $\hat{\mathbb{Q}}$ is also a $\tilde{\mathbb{G}}$-martingale under $\hat{\mathbb{Q}}$.*

Proof. Along the similar lines as the proof of Lemma 2.5. $\square$

The next result is a counterpart of Proposition 2.6. Recall that by convention $(C^1_0, \ldots, C^n_0) = (1, \ldots, 1)$.

Proposition 2.8 *Let $\hat{\tau} = \tau_{p+1} \wedge \ldots \wedge \tau_n$. For any $t \leq s \leq T$, and any $\hat{\mathbb{Q}}$-integrable, $\hat{\mathcal{G}}_s$-measurable random variable Y we have*

$$\mathbb{E}_{\mathbb{Q}^*}(\mathbb{1}_{\{\hat{\tau}>s\}}Y \mid \mathcal{G}_t) = \mathbb{E}_{\mathbb{Q}^*}(\mathbb{1}_{\{\hat{\tau}>s\}}Y \mid \tilde{\mathcal{G}}_t) = \mathbb{1}_{\{\hat{\tau}>t\}}\,\mathbb{E}_{\hat{\mathbb{Q}}}\left(Ye^{-\sum_{i=p+1}^n \int_t^s \hat{\lambda}^i_{11}(u)\,du} \,\middle|\, \hat{\mathcal{G}}_t\right).$$

Proof. In view of Lemma 2.8, to prove the proposition we can make use of the similar arguments as in the proof of Proposition 2.6. $\square$

In particular, we have

$$\mathbb{Q}^*(\hat{\tau} > s \,|\, \mathcal{G}_t) = \mathbb{Q}^*(\hat{\tau} > s \,|\, \tilde{\mathcal{G}}_t) = \mathbb{1}_{\{\hat{\tau} > t\}} \, \mathbb{E}_{\hat{\mathbb{Q}}}\left(e^{-\sum_{i=p+1}^{n} \int_t^s \hat{\lambda}_{11}^i(u)\,du} \,\Big|\, \hat{\mathcal{G}}_t\right).$$

and

$$\mathbb{Q}^*(\hat{\tau} \leq s \,|\, \mathcal{G}_t) = \mathbb{Q}^*(\hat{\tau} \leq s \,|\, \tilde{\mathcal{G}}_t) = \mathbb{1}_{\{\hat{\tau} \leq t\}} + \mathbb{1}_{\{\hat{\tau} > t\}} \, \mathbb{E}_{\hat{\mathbb{Q}}}\left(1 - e^{-\sum_{i=p+1}^{n} \int_t^s \hat{\lambda}_{11}^i(u)\,du} \,\Big|\, \hat{\mathcal{G}}_t\right).$$

2.3.4   First-to-jump swap

In order to illustrate our method, we shall now examine an example of a credit derivative. By analogy with the *first-to-default swap*, we call this contract the *first-to-jump swap*. Since the *premium leg* of the contract can be treated in a similar way, we shall focus on the *jump leg* only.

Without loss of generality, we assume that the payoff occurs at the first jump of the credit rating of any entity from the class $\{p+1, \ldots, n\}$. In addition, if $\hat{\tau} = \tau_j \leq T$ then the payoff at time $\hat{\tau}$ equals $Z_{\hat{\tau}}^j$, where Z^j, $j = p+1, \ldots, n$ are given $\hat{\mathbb{G}}$-predictable stochastic processes. Formally, the swap corresponds to the contingent claim (recall that we consider here only the jump leg of the contract)

$$Y = \sum_{j=p+1}^{n} Z_{\hat{\tau}}^j \mathbb{1}_{\{\hat{\tau} = \tau_j \leq T\}}$$

which settles at time $\hat{\tau}$. By definition, the ex-dividend price at time $t \leq T$ of the contract described above equals

$$\pi_t(Y) = B_t \, \mathbb{E}_{\mathbb{Q}^*}\left(B_{\hat{\tau}}^{-1} \mathbb{1}_{\{\hat{\tau} > t\}} Y \,\big|\, \mathcal{G}_t\right) = B_t \, \mathbb{E}_{\mathbb{Q}^*}\left(B_{\hat{\tau}}^{-1} \mathbb{1}_{\{t < \hat{\tau} \leq T\}} \sum_{j=p+1}^{n} Z_{\hat{\tau}}^j \mathbb{1}_{\{\hat{\tau} = \tau_j\}} \,\Big|\, \mathcal{G}_t\right).$$

Let us recapitulate the basic steps in the proposed method of risk-neutral valuation of this contract:

- Introduce a judiciously chosen probability measure $\hat{\mathbb{Q}}$ equivalent to the martingale measure $\mathbb{Q}^*$ on $(\Omega, \mathcal{G}_T)$ (cf. formula (2.9)),

- Check that $\tau_{p+1}, \ldots, \tau_n$ are $\hat{\mathbb{G}}$-conditionally independent under $\hat{\mathbb{Q}}$ and under $\mathbb{Q}^*$ with the same $\hat{\mathbb{G}}$-hazard processes (cf. Lemma 2.7),

- Use the standard method of valuing the first-to-default swap through conditional expectations under $\mathbb{Q}^*$ (cf. Kijima and Muromachi (2000) or Chapter 9 in Bielecki and Rutkowski (2002)).

2.4   Conclusions

To summarize, due to the complexity of the model of mutually dependent credit migration processes, it is advisable to study particular contracts on a case-by-case basis, rather than to attempt to establish a general approach. For instance, in the case of the first-to-change contract, the calculations can be reduced to the familiar results concerning the case of conditionally independent random times through a judicious choice of an equivalent probability measure. Finally, it should be acknowledged

that the migration process C can also be studied as a standard Markov chain with respect to its natural filtration (in the case of a trivial reference filtration) or an $\mathbb{F}$-conditional Markov chain (when the reference filtration $\mathbb{F}$ is non-trivial), so that the machinery developed for these classes of processes can be directly employed. Due to the large dimension of the state space for C (the number of states for C equals $K_1 K_2 \dots K_n$) such a direct approach does not seem to be efficient, however.

References

[1] Bielecki, T.R., Rutkowski, M. (1999) Defaultable term structure: Conditionally Markov approach. Working paper.

[2] Bielecki, T.R., Rutkowski, M. (2000) Multiple ratings model of defaultable term structure. *Math. Finance* 10, 125–139.

[3] Bielecki, T.R., Rutkowski, M. (2002) *Credit Risk: Modeling, Valuation and Hedging.* Springer-Verlag, Berlin Heidelberg New York.

[4] Jeanblanc, M., Rutkowski, M. (2000) Modelling of default risk: Mathematical tools. Working paper, Université d'Évry and Warsaw University of Technology.

[5] Jeanblanc, M., Rutkowski, M. (2001) Default risk and hazard process. *Mathematical Finance – Bachelier Congress 2000.* Springer-Verlag, Berlin Heidelberg New York.

[6] Jarrow, R.A., Yu, F. (2001) Counterparty risk and the pricing of defaultable securities. *J. Finance* 56, 1765-1799.

[7] Kijima, M., Muromachi, Y. (2000) Credit events and the valuation credit derivatives of basket type. *Rev. Derivatives Res.* 4, 55–79.

[8] Kusuoka, S. (1999) A remark on default risk models. *Adv. Math. Econom.* 1, 69–82.

[9] Shaked, M., Shanthikumar, J.G. (1987) The multivariate hazard construction. *Stochastic Process. Appl.* 24, 241–258.

[10] Schmidt, W.M. (1998) Modelling correlated defaults. Working paper.

[11] Schönbucher, P., Schubert, D. (2000) Copula-dependent default risk in intensity models. Working paper, Bonn University.