Quantitative Finance

ISSN: 1469-7688 (Print) 1469-7696 (Online) Journal homepage: http://www.tandfonline.com/loi/rquf20

Interacting default intensity with a hidden Markov process

Feng-Hui Yu, Wai-Ki Ching, Jia-Wen Gu & Tak-Kuen Siu

To cite this article: Feng-Hui Yu, Wai-Ki Ching, Jia-Wen Gu & Tak-Kuen Siu (2017) Interacting default intensity with a hidden Markov process, Quantitative Finance, 17:5, 781-794, DOI: 10.1080/14697688.2016.1237036

To link to this article: http://dx.doi.org/10.1080/14697688.2016.1237036

Published online: 07 Nov 2016.

Quantitative Finance, 2017
Vol. 17, No. 5, 781–794, http://dx.doi.org/10.1080/14697688.2016.1237036

Routledge
Taylor & Francis Group

Interacting default intensity with a hidden Markov process

FENG-HUI YU†, WAI-KI CHING*†, JIA-WEN GU† and TAK-KUEN SIU‡

†Advanced Modeling and Applied Computing Laboratory, Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong
‡Faculty of Business and Economics, Department of Applied Finance and Actuarial Studies, Macquarie University, Sydney, NSW 2109, Australia

(*Received 3 April 2016; accepted 8 September 2016; published online 7 November 2016*)

In this paper we consider a reduced-form intensity-based credit risk model with a hidden Markov state process. A filtering method is proposed for extracting the underlying state given the observation processes. The method can be applied to a wide range of problems. Based on this model, we derive the joint distribution of multiple default times without imposing stringent assumptions on the form of default intensities. Closed-form formulas for the distribution of default times are obtained which are then applied to solve a number of practical problems such as hedging and pricing credit derivatives. The method and numerical algorithms presented can be applicable to various forms of default intensities.

Keywords: Reduced-form intensity model; Default risk; Credit derivatives; Hidden Markov model (HMM)

1. Introduction

Modeling credit risk has long been a critical issue in credit risk management. Attention has been given to it especially since the global financial crisis in 2008. Credit risk modeling has a lot of applications, for example, pricing and hedging credit derivatives, as well as the management of credit portfolios. Models adopted in the finance industry can be classified into two major categories: structural firm value models and reduced-form intensity-based models. The first class of models was pioneered by Black and Scholes (1973) and Merton (1974). The key idea of the structural firm's value model is to model the default of a firm by using its asset value, where the asset value is governed by a geometric Brownian motion. When the asset value falls below a certain prescribed level, the default of the firm is triggered. The second kind of model was pioneered by Jarrow and Turnbull (1995) and Madan and Unal (1998). The main idea of reduced-form intensity-based models is to consider the defaults as exogenous processes and describe their occurrences with Poisson processes and their variants.

Interacting intensity-based default models are widely adopted to model portfolio credit risk and defaults. Since we focus on contagion models in this paper as in, for example, Giesecke (2008), we divide intensity-based credit risk models into top-down models and bottom-up models. The top-down models focus on modeling the default times at the portfolio level without reference to the intensities of individual entities. Based on this, one can also recover the individual entity's intensity with some method like random thinning, etc. Some works related to this class of models include Davis and Lo (2001), Brigo *et al*. (2006), Longstaff and Rajan (2008), Giesecke *et al*. (2011) and Cont and Minca (2013), etc. While the bottom-up model focuses on modeling the default intensities of individual reference entities and their aggregation to form a portfolio default intensity. Some works related to this class of models include Duffie and Garleanu (2001), Jarrow and Yu (2001), Schönbucher and Schubert (2001), Giesecke and Goldberg (2004), Duffie *et al*. (2006) and Yu (2007), etc. The differences between these two classes of models are the form of individual entity's default intensities and the way the portfolio aggregation is formed. In this paper we shall focus on a bottom-up model.

Based on the model developed by Lando (1998), Yu (2007) further extended the model and applied it to multiple defaults and their correlation. In addition, Yu adopted the total hazard construction method proposed by Norros (1986) and Shaked and Shanthikumar (1987) to simulate the distribution of default times which have interacting intensities. Zheng and Jiang (2009) then adopted this method and derived closed-form formulas for the multiple default distributions under their contagion model. Gu *et al*. (2013) introduced a recursive method to calculate the distribution of ordered default times, and Gu *et al*. (2014) further proposed a hidden Markov reduced-form model with a specific form of default intensities.

*Corresponding author. Email: wching@hku.hk

In this paper we develop a generalized reduced-form intensity-based credit model with a hidden Markov process. The model is applicable to a wide class of default intensities with various forms of dependent constructions. For the hidden Markov process, we also discuss a flexible method to extract the hidden state process given the observations processes without constraints on the dynamic structure of HMM, which may hopefully have applications in diverse fields. Then using the total hazard construction method by Yu (2007), we derive closed-form formulas for the joint default distribution. When the intensities are homogeneous and symmetric, analytic algorithm for the calculation of the distribution of ordered default times is also provided. The explicit formulas may enhance the computational efficiency in applications, for instance, pricing of credit derivatives. We remark that the results in Gu *et al.* (2014) are a special case of the method discussed here. In addition, we extend the total hazard construction method to the cases with hidden process to simulate the joint distribution of default times.

For practical applications of hidden Markov models (HMMs), an important step is filtering the hidden Markov processes. There are different approaches for filtering HMMs in the literature. One prominent and theoretically sound approach is based on a reference probability approach in, for example Elliott *et al.* (2008). The key idea of the reference probability approach is to start with a reference probability measure under which the observed dynamics become simpler and do not depend on the hidden state processes. Then filtering is done in a 'reference probability world' and unnormalized filters instead of normalized filters for the hidden state processes are derived. One advantage of the unnormalized filters is that they usually satisfy linear stochastic differential equations, which are called the Zakai equations in the literature. Normalized filters may not enjoy this nice property. Indeed, the use of the reference approach for filtering HMMs in the context of credit risk analysis has been explored in the literature. For example, Frey and Runggaldier (2010, 2011) focused on mathematical analysis of nonlinear filtering using a reference probability approach. They derived recursive filtering equations in a continuous-time model. Frey and Schmidt (2011) presented an analysis of nonlinear filtering, but they adopted the innovations approach and derived some filtering equations. Elliott and Siu (2013) also used the reference probability approach and derived filtering equations for continuous-time finite-state HMMs under their models. Elliott *et al.* (2014) applied the reference probability approach to derive filters for hidden Markov chains in the context of a double HMM-based Z-scores model. In Gu *et al.* (2014), another filtering method using moment generating functions was used in the context of credit risk modeling. An advantage of the approach used in Gu *et al.* (2014) is that explicit formulas for filtering and joint default distributions can be obtained under some parametric forms of default intensities. Here, our aim is to further explore the filtering approach in Gu *et al.* (2014) in our current modeling setup and derive explicit formulas. We shall illustrate the practical implementation of this approach using some numerical examples. It may be also interesting to explore the use of the reference probability approach for filtering hidden Markov processes in our current modeling setup. This may represent a potentially interesting topic for future research. In a recent paper by Elliott and Shen (2015a), an intensity-based credit risk model based on a self-exciting intensity which incorporated both frailty and default contagion was considered but without numerical examples. This paper may be related to our current paper. However, the structure of the intensity process considered in the paper by Elliott and Shen (2015a) may not be the same as ours. Furthermore, they used the reference probability approach and derived recursive filtering equations for the hidden state processes which are different from ours. Elliott and Shen (2016) also discuss a credit risk model with latent contagion and frailty for pricing, and hedging. In another recent paper by Elliott and Shen (2015b), an optimal capital structure for a corporation was considered in a regime-switching extension of the Black–Cox model. The modeling structure in Elliott and Shen (2015b) is different from ours. The former may belong to the class of structural firm values models while ours belongs to the class of reduced-form credit risk model.

The rest of this paper is structured as follows. Section 2 gives a snapshot of the interacting intensity-based default model with hidden Markov process. Section 3 presents the method for extracting the hidden state process from the observation processes. Section 4 derives the closed-form expression for the joint default distribution based on the total hazard construction method, and also gives an analytic formula for the distribution of ordered default times. Further, the extended total hazard construction method under a hidden Markov process to obtain the joint distribution of default times is also presented. Section 5 provides numerical methods for some situations in section 3, which may be used in both sections 3 and 4, and error analysis is also discussed. Section 6 illustrates an application of the proposed method in pricing credit derivatives. Finally, section 7 concludes the paper.

2. Model setup

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space where P is a risk-neutral probability measure, which is assumed to exist. Suppose there are K interacting entities, and we let $N_i(t) := 1_{\{\tau_i \le t\}}$, where τ_i is a stopping time, representing the default time of credit name i, for each $i = 1, 2, \ldots, K$. Suppose we have an underlying state process $(X_t)_{t\ge 0}$ describing the dynamics of the economic condition. Let $\mathcal{F}_t^X := \sigma(X_s, 0 \le s \le t) \vee \mathcal{N}$ where $\mathcal{N}$ represents all the null subsets of Ω in $\mathcal{F}$ and $\mathcal{C}_1 \vee \mathcal{C}_2$ is the minimal σ-algebra containing both the σ-algebras $\mathcal{C}_1$ and $\mathcal{C}_2$. We also let $\mathcal{H}_t := \sigma(X_t) \vee \mathcal{F}_t^N$ where

$$\mathcal{F}_t^N = \mathcal{F}_t^1 \vee \mathcal{F}_t^2 \vee \cdots \vee \mathcal{F}_t^K \quad \text{and} \quad \mathcal{F}_t^i := \sigma(1_{\{\tau_i \le s\}}, 0 \le s \le t) \vee \mathcal{N}.$$

We assume that for each $i = 1, 2, \ldots, K$, $N_i(t)$ possesses a nonnegative, $\{\mathcal{H}_t\}_{t\ge 0}$-adapted, intensity process λ_i satisfying

$$E\left(\int_0^t \lambda_i(s)\mathrm{d}s\right) < \infty, \quad t \ge 0, \tag{1}$$

such that the compensated process

$$M_i(t) := N_i(t) - \int_0^{t\wedge\tau_i} \lambda_i(s)\mathrm{d}s\ , \quad t \ge 0, \tag{2}$$

is an $(\{\mathcal{H}_t\}_{t\ge 0}, P)$-martingale. Note that after the default time τ_i, $N_i(t)$ will stay at the value one, so there is no need to compensate for $N_i(t)$ after time τ_i, see, for example, Elliott *et al.* (2000).

For all the market participants, we assume that they cannot observe the underlying process $(X_t)_{t\geq 0}$ directly. Instead, they observe the process $(Y_t)_{t\geq 0}$, revealing the delayed and noisy information of $(X_t)_{t\geq 0}$, and also observe the default process $(N_t^i)_{t\geq 0}$. Hence, the common information set available to the market participants at time t is $\mathcal{F}_t := \mathcal{F}_t^Y \vee \mathcal{F}_t^N$ where $\mathcal{F}_t^Y := \sigma(Y_s, 0 \leq s \leq t) \vee \mathcal{N}$. We further assume that $(X_t)_{t\geq 0}$ is an 'exogenous' process to $(N_t^i)_{t\geq 0}, i = 1, 2, \ldots, K$, i.e. For any t, the σ-fields $\mathcal{F}_\infty^X$ and $\mathcal{F}_t^N$ are conditionally independent given $\mathcal{F}_t^X$ and $P(\tau_i \neq \tau_j) = 1, i \neq j$.

To simplify our discussion, throughout the paper, we suppose that $(X_t)_{t\geq 0}$ is a two-state Markov chain taking a value in $\{x_0, x_1\}$. We assume the transition rates of the chain for '$x_0 \to x_1$' and '$x_1 \to x_0$' are θ_0 and θ_1, respectively. The observable process $(Y_t)_{t\geq 0}$ is again a two-state Markov chain taking a value in $\{y_0, y_1\}$, with transition rates depending on X_t, i.e. $\eta_0(X_t)(y_0 \to y_1)$ and $\eta_1(X_t)(y_1 \to y_0)$, where η_0 and η_1 are real-valued functions. At time 0, we suppose that X_0 is in state x_0 and Y_0 is in state y_0. We remark that the methods introduce later can still be applicable when the Markov chains X and Y have more than two states i.e. finite many states though more complicated notations may involve.

3. Extraction of hidden state process with observable processes

3.1. Some preliminaries

Let $\bar{T}_{i,k,j}(s_0, \Delta s)$ be the union of subintervals of time of the chain X in state x_k in the time interval $[s_0, s_0 + \Delta s]$ given the chain starts from $X_{s_0} = x_i$ and ends at $X_{s_0+\Delta s} = x_j$. For $i, j = 0, 1$, we let

$$\bar{T}_{i,j}(s_0, t) = (\bar{T}_{i,0,j}(s_0, t), \bar{T}_{i,1,j}(s_0, t))^T \quad \text{and} \quad \mathbf{u}(\bar{t}) = (u_0(\bar{t}), u_1(\bar{t}))^T$$

where $\bar{t} \in [s_0, s_0 + t]$. Note that $\bar{T}_{i,1,j}(s_0, t) = [s_0, s_0 + t]\backslash \bar{T}_{i,0,j}(s_0, t)$. Since jumps in chain Y and defaults are Poisson processes, using the concept of moment generating function, we define

$$\bar{\Psi}_{ij}(s_0, \mathbf{u}, t) = E\left[\exp\left\{\int_{\bar{T}_{i,0,j}(s_0,t)} u_0(\bar{t})\mathrm{d}\bar{t} + \int_{\bar{T}_{i,1,j}(s_0,t)} u_1(\bar{t})\mathrm{d}\bar{t}\right\}\right].$$

Note that $\mathbf{u}(\bar{t})$ is an arbitrary integrable function. This means, in this case, we can adopt this moment generating function. For instance, $\mathbf{u}(\bar{t})$ can be the transition rates of jumps in chain Y or the default rates which are the default intensities accumulated by all the entities by time $\bar{t}$ before default.

PROPOSITION 1 *Let* $\bar{\Phi}_{ij}(s_0, \mathbf{u}, t) = P_{ij}(t)\bar{\Psi}_{ij}(s_0, \mathbf{u}, t)$, *where* $P_{ij}(t)$ *is the probability that a process in state* x_i *will be in state* x_j *after a time of* t, *and* $i, j = 0, 1$. *Then*

$$\begin{cases} \bar{\Phi}_{ij}(s_0, \mathbf{u}, t) = \theta_i \displaystyle\int_0^t \exp\left(\int_{s_0}^{s_0+t-s} (u_i(\bar{t}) - \theta_i)\mathrm{d}\bar{t}\right) \bar{\Phi}_{jj}(s_0 + t - s, \mathbf{u}, s)\mathrm{d}s \\ \bar{\Phi}_{ii}(s_0, \mathbf{u}, t) = \theta_i \displaystyle\int_0^t \exp\left(\int_{s_0}^{s_0+t-s} (u_i(\bar{t}) - \theta_i)\mathrm{d}\bar{t}\right) \bar{\Phi}_{ji}(s_0 + t - s, \mathbf{u}, s)\mathrm{d}s + \exp\left(\int_{s_0}^{s_0+t} (u_i(\bar{t}) - \theta_i)\mathrm{d}\bar{t}\right) \end{cases} \tag{3}$$

where $i, j = 0, 1$.

Proof.

$$\begin{aligned} &\bar{\Psi}_{ij}(s_0, \mathbf{u}, t) \\ &= E\left[\exp\left(\int_{\bar{T}_{i,0,j}(s_0,t)} u_0(\bar{t})\mathrm{d}\bar{t} + \int_{\bar{T}_{i,1,j}(s_0,t)} u_1(\bar{t})\mathrm{d}\bar{t}\right)\right] \\ &= \frac{\theta_i}{P_{ij}(t)} \int_0^t e^{-\theta_i s} \cdot e^{\int_{s_0}^{s_0+s} u_i(\bar{t})\mathrm{d}\bar{t}} P_{jj}(t-s) \\ &\quad \times E\left[\exp\left(\int_{\bar{T}_{j,0,j}(s_0+s,t-s)} u_0(\bar{t})\mathrm{d}\bar{t} + \int_{\bar{T}_{j,1,j}(s_0+s,t-s)} u_1(\bar{t})\mathrm{d}\bar{t}\right)\right]\mathrm{d}s \\ &= \frac{\theta_i}{P_{ij}(t)} \int_0^t \exp\left(\int_{s_0}^{s_0+s} (u_i(\bar{t}) - \theta_i)\mathrm{d}\bar{t}\right) \\ &\quad \times P_{jj}(t-s)\bar{\Psi}_{jj}(s_0 + s, \mathbf{u}, t - s)\mathrm{d}s \\ &= \frac{\theta_i}{P_{ij}(t)} \int_0^t \exp\left(\int_{s_0}^{s_0+t-s} (u_i(\bar{t}) - \theta_i)\mathrm{d}\bar{t}\right) \\ &\quad \times P_{jj}(s)\bar{\Psi}_{jj}(s_0 + t - s, \mathbf{u}, s)\mathrm{d}s. \end{aligned}$$

We also have

$$\bar{\Psi}_{ii}(s_0, \mathbf{u}, t) = \frac{\theta_i}{P_{ii}(t)} \int_0^t e^{\int_{s_0}^{s_0+t-s} (u_i(\bar{t})-\theta_i)\mathrm{d}\bar{t}}$$

To specify the form of the intensities, we give the following notations. Suppose that at time t, N_t^D defaults have already occurred at $t_1, t_2, \ldots, t_{N_t^D}$ such that

$$0 = t_0 < t_1 < \cdots < t_{N_t^D} \leq t.$$

Then we denote $T_{N_t^D} = (t_1, \ldots, t_{N_t^D})$ the ordered N_t^D default times and $I_{N_t^D} = (j_1, \ldots, j_{N_t^D})$ the corresponding N_t^D defaulters, and the mth $(1 \leq m \leq K)$ defaulted obligor is j_m. We assume that $i > N_t^D$ and $t < \tau^i$, where τ^i is the obligor i's default time. Each process λ_i $(i = 1, \ldots, K)$, is $\{\mathcal{H}_t\}_{t\geq 0}$-predictable, that is to say $\lambda_i(t)$ is known given information about the chain X and all the default processes prior to time t. Then the intensity of τ^i may be written as $\lambda_t^i = \lambda_i(t|I_{N_t^D}, T_{N_t^D}, X_t)$ where X_t is the state of chain X at time t. Note that $(I_{N_t^D}, T_{N_t^D}, X_t) \in \mathcal{H}_t$.

Since the path of X is unobservable, while the path of Y and $N_i, (i = 1, \ldots, K)$ are observable, we can use the relationship between X, Y and N_i, $i = 1, \ldots, K$. To find the probability law of X. We apply the recursive method proposed in Gu *et al.* (2014) to calculate the conditional probability $P(X_t = x_i|\mathcal{F}_t)$, $(i = 0, 1, t \geq 0)$. Before discussing the method, we need to find the expressions for all the unknown items in the recursive formulas. In the process of finding the expressions, we also present moment generating function method to achieve our goal.

$$\times P_{ji}(s)\bar{\Psi}_{ji}(s_0+t-s,\mathbf{u},s)\mathrm{d}s + \frac{e^{\int_{s_0}^{s_0+t}(u_i(\bar{t})-\theta_i)\mathrm{d}\bar{t}}}{P_{ii}(t)}.$$

Replace $\bar{\Psi}_{ij}(s_0,\mathbf{u},t)$ by $\frac{\bar{\Phi}_{ij}(s_0,\mathbf{u},t)}{P_{ij}(t)}$, we can then get the system of equations in the proposition. □

We observe that when the expression of $\mathbf{u}(\bar{t})$ satisfies some 'good' property, equation (3) in the above proposition has a unique solution. The property is that $\mathbf{u}(\bar{t})$ does not have any direct relationship with time $\bar{t}$ even though it may have implied relationship with $\bar{t}$. This means $\mathbf{u}(\bar{t})$ can be written as $\mathbf{u}$. Then, not only the problem of solving equation (3) can be simplified, but some related definitions can also be simplified as well. Similar as before, let $T_{i,k,j}(\Delta s)$ be the occupation time of the chain X in state x_k in the time interval $[s, s+\Delta s]$ given the chain starting from $X_s = x_i$ and ending at $X_{s+\Delta s} = x_j$. For each $i, j = 0, 1$, we let

$$T_{i,j}(t) = (T_{i,0,j}(t), T_{i,1,j}(t))^T \quad \text{and} \quad \mathbf{u} = (u_0, u_1)^T \in \mathbb{R}^2.$$

The moment generating function of $T_{i,j}(t)$ is given by

$$\Psi_{ij}(\mathbf{u},t) = E(\exp\{\mathbf{u}^T T_{i,j}(t)\}).$$

Apply the same method to $\Psi_{ij}(\mathbf{u},t)$ as we have done to $\bar{\Psi}_{ij}(\mathbf{u},t)$, and let

$$\Phi_{ij}(\mathbf{u},t) = \Psi_{ij}(\mathbf{u},t)\cdot P_{ij}(t).$$

We can also get the equivalent equation (3) for $\Phi_{ij}(\mathbf{u},t)$, *i.e.*, replacing $\bar{\Phi}_{ij}(\mathbf{u},t)$ by $\Phi_{ij}(\mathbf{u},t)$, $(u_i(\bar{t})-\theta_i)$ by $(u_i-\theta_i)$ in equation (3). Then to solve the equivalent equation, it suffices to solve a linear system of O.D.E.s (c.f. Gu *et al.* 2014):

$$\frac{\partial\Phi(\mathbf{u},t)}{\partial t} = A\Phi(\mathbf{u},t),$$

where

$$\Phi(\mathbf{u},t) = \begin{bmatrix} \Phi_{11}(\mathbf{u},t) & \Phi_{12}(\mathbf{u},t) \\ \Phi_{21}(\mathbf{u},t) & \Phi_{22}(\mathbf{u},t) \end{bmatrix}$$
$$\text{and} \quad A = \begin{bmatrix} u_0-\theta_0 & \theta_0 \\ \theta_1 & u_1-\theta_1 \end{bmatrix}.$$

This linear system of O.D.E.s is known as the fundamental matrix equation in the literature. Then it is well-known that the equation has a unique solution which is called the fundamental matrix solution with the initial condition $\Phi_{ij}(\mathbf{u},0) = 1$, $(i, j = 0, 1)$ as

$$\Phi(\mathbf{u},t) = e^{At}\mathbf{1}\cdot\mathbf{1}^T$$

where $\mathbf{1}$ is the two-dimensional column vector with all entries being equal to 1. Hence we can get the solution for $\Psi_{ij}(\mathbf{u},t)$ by

$$\Psi_{ij}(\mathbf{u},t) = \frac{\Phi_{ij}(\mathbf{u},t)}{P_{ij}(t)}.$$

In practice, when the expressions of $u_i(\bar{t})$, $(i = 0, 1)$ are given, we can substitute them into the above equation (3), then intuitively we can check whether it has a solution. Note that the expressions of $u_i(\bar{t})$, $(i = 0, 1)$ determine whether the system is solvable. If it is solvable, then we can obtain the solution $\bar{\Phi}_{ij}(s_0,\mathbf{u},t)$, $(i, j = 0, 1)$. Note that the results in Gu *et al.* (2014) can be regarded as a special case that has a unique solution.

3.2. *Recursive formulas for extracting hidden process*

For $\widetilde{\omega}_t \in \mathcal{F}_t$, we can express $\widetilde{\omega}_t$ in a more clear way as follows:

$$\widetilde{\omega}_t = (N_t^Y, N_t^D, S_{N_t^Y}, I_{N_t^D}, T_{N_t^D})$$

where

- $S_{N_t^Y} = (s_1, s_2, \ldots, s_{N_t^Y})$,
- $I_{N_t^D} = (j_1, j_2, \ldots, j_{N_t^D})$,
- $T_{N_t^D} = (t_1, t_2, \ldots, t_{N_t^D})$,
- N_t^Y counts the number of jumps in chain Y by time t,
- N_t^D counts the number of defaults by time t,
- $(s_1, s_2, \ldots, s_{N_t^Y})$ is the collection of ordered jump times of the chain Y by time t, i.e. $0 < s_1 < \cdots < s_{N_t^Y} \le t$,
- $(t_1, t_2, \ldots, t_{N_t^D})$ is the collection of ordered default times by time t, i.e. $0 < t_1 < \cdots < t_{N_t^Y} \le t$,
- $(j_1, j_2, \ldots, j_{N_t^D})$ is the collection of ordered corresponding name of defaulters by time t, *i.e.*, name j_i defaults at time t_i.

Here $\widetilde{\omega}_t$ can be interpreted as the state of the stochastic dynamical system at time t. Given the information up to time t, i.e. $\mathcal{F}_t$, we divide the time period $[0, t]$ into $(N_t^Y + N_t^D)$ sub-periods, $[0, h_1], (h_1, h_2], \ldots, (h_{N_t^Y+N_t^D-1}, h_{N_t^Y+N_t^D}]$. In each of them, exactly one default or one jump in Y is observed. When there is no default or jump occurred by time t, the calculation of $P(X_t = x_i \mid \mathcal{F}_t)$ can be simplified and we shall introduce it later.

Define $\bar{I}_{N_t^D} = (1, 2, \ldots, K)\backslash I_{N_t^D}$. Suppose that s and $s+\Delta s$ are two endpoints of one sub-period. The following characterizes the computational method for $P(X_t = x_i \mid \mathcal{F}_t)$. For $\widetilde{\omega} \in \{t_k = s + \bar{t}_k \in (s, s+\Delta s]\}$,

$$\begin{aligned} &P(X_s = x_i \mid \mathcal{F}_{s+\Delta s}) \\ &= P(X_s = x_i \mid \mathcal{F}_s, t_k = s+\bar{t}_k, j_k = \beta) \\ &= \frac{P(X_s = x_i \mid \mathcal{F}_s)\cdot\left(\sum_{l=0,1} f_{t_k}^{i,l}(s+\bar{t}_k;\beta,s,\Delta s)\right)}{\sum_{j=0,1} P(X_s = x_j \mid \mathcal{F}_s)\cdot\left(\sum_{l=0,1} f_{t_k}^{j,l}(s+\bar{t}_k;\beta,s,\Delta s)\right)} \end{aligned} \tag{4}$$

and

$$\begin{aligned} &P(X_{s+\Delta s} = x_i \mid \mathcal{F}_{s+\Delta s}) \\ &= \sum_{j=0,1} P(X_s = x_j \mid \mathcal{F}_{s+\Delta s}) P(X_{s+\Delta s} = x_i \mid \mathcal{F}_{s+\Delta s}, X_s = x_j) \\ &= \sum_{j=0,1} P(X_s = x_j \mid \mathcal{F}_{s+\Delta s}) \frac{f_{t_k}^{j,i}(s+\bar{t}_k;\beta,s,\Delta s)}{\sum_{l=0,1} f_{t_k}^{j,l}(s+\bar{t}_k;\beta,s,\Delta s)} \end{aligned} \tag{5}$$

where

$$\begin{aligned} f_{t_k}^{j,i}(t;\beta,s,\Delta s)\mathrm{d}t = P(t_k \in \mathrm{d}t,\ & j_k = \beta, \\ & X_{s+\Delta s} = x_i \mid X_s = x_j, N_s^D, N_s^Y, I_{N_s^D}). \end{aligned}$$

Similarly, we have for $\widetilde{\omega} \in \{s_k = s + \bar{s}_k \in (s, s+\Delta s]\}$,

$$\begin{aligned} &P(X_s = x_i \mid \mathcal{F}_{s+\Delta s}) \\ &= \frac{P(X_s = x_i \mid \mathcal{F}_s)\left(\sum_{l=0,1} f_{s_k}^{i,l}(s+\bar{s}_k;s,\Delta s)\right)}{\sum_{j=0,1} P(X_s = x_j \mid \mathcal{F}_s)\left(\sum_{l=0,1} f_{s_k}^{j,l}(s+\bar{s}_k;s,\Delta s)\right)} \end{aligned} \tag{6}$$

and

$$P(X_{s+\Delta s}=x_i \mid \mathcal{F}_{s+\Delta s}) = \sum_{j=0,1} P(X_s=x_j \mid \mathcal{F}_{s+\Delta s}) \times \frac{f_{s_k}^{j,i}(s+\bar{s}_k; s, \Delta s)}{\left(\sum_{l=0,1} f_{s_k}^{j,l}(s+\bar{s}_k); s, \Delta s\right)} \quad (7)$$

where

$$f_{s_k}^{j,i}(t; s, \Delta s)\mathrm{d}t = P(s_k \in \mathrm{d}t, X_{s+\Delta s}=x_i \mid X_s=x_j, N_s^D, N_s^Y, I_{N_s^D}).$$

Combining equations (4)–(7), we obtain a recursive method for computing $P(X_t=x_i \mid \mathcal{F}_t)$ in terms of $f_{t_k}^{j,i}(s+\bar{t}_k; \beta, s, \Delta s)$ and $f_{s_k}^{j,i}(s+\bar{s}_k; s, \Delta s)$. That is to say, with the fact that $P(X_0 = x_0|\mathcal{F}_0) = 1$ and $P(X_0 = x_1|\mathcal{F}_0) = 0$, we can apply them to equation (4) or equation (6) according to $\mathcal{F}_t$, and then to get $P(X_0 = x_i|\mathcal{F}_{\Delta s})$ which are unknown in the calculation of $P(X_{\Delta s} = x_i|\mathcal{F}_{\Delta s})$ in equation (5) or equation (7). The equation to calculate $P(X_{\Delta s} = x_i|\mathcal{F}_{\Delta s})$ should be chosen according to $\mathcal{F}_t$ as well. By repeating this recursion procedure, we can obtain the desired conditional probabilities. We remark that the number of recursion procedures is equal to $N_t^Y + N_t^D$, which is the sum of jumps in the observable chain Y and defaults observed by time t. It does not seem that in practice the sum will go too large. Consequently, it is envisaged that the number of recursion procedures for filtering may not go too large.

To get the expressions for the desired $f_{t_k}^{j,i}(s+\bar{t}_k; \beta, s, \Delta s)$ and $f_{s_k}^{j,i}(s+\bar{s}_k; s, \Delta s)$, we need to use the method introduced in section 3.1. Replace $\mathbf{u}$ by $-(\eta_i(x_0), \eta_i(x_1))$, $i = 0, 1$ and we know that there exists unique solutions for Ψ_{ij}, $i, j = 0, 1$. Replace $\mathbf{u}(\bar{t})$ by $-(\lambda_i(x_0), \lambda_i(x_1))$, $i = 1, \ldots, K$ in equation (3), we then could check whether it is solvable or not. If it is solvable and has an analytical solution, then from the definition of $f_{t_k}^{j,i}(s+\bar{t}_k; \beta, s, \Delta s)$ and $f_{s_k}^{j,i}(s+\bar{s}_k; s, \Delta s)$, we get

$$\begin{aligned} &f_{s_k}^{j,i}(s+\bar{s}_k; s, \Delta s) \\ &= \sum_{l=0,1} P_{jl}(\bar{s}_k)P_{li}(\Delta s-\bar{s}_k)\eta_{C(N_s^Y)}(x_l) \\ &\quad\times \Psi_{jl}\left(-(\eta_{C(N_s^Y)}(x_0), \eta_{C(N_s^Y)}(x_1))^T, \bar{s}_k\right) \\ &\quad\times \Psi_{li}\left(-(\eta_{C(N_s^Y+1)}(x_0), \eta_{C(N_s^Y+1)}(x_1))^T, \Delta s-\bar{s}_k\right) \\ &\quad\times \bar{\Psi}_{jl}\Bigg(s, -\sum_{i\in\bar{I}_{N_s^D}} (\lambda_i(\bar{t}|I_{N_s^D}, T_{N_s^D}, x_0), \\ &\qquad\qquad \times \lambda_i(\bar{t}|I_{N_s^D}, T_{N_s^D}, x_1))^T, \bar{s}_k\Bigg) \\ &\quad\times \bar{\Psi}_{li}\Bigg(s+\bar{s}_k, -\sum_{i\in\bar{I}_{N_s^D}} (\lambda_i(\bar{t}|I_{N_s^D}, T_{N_s^D}, x_0), \\ &\qquad\qquad \times \lambda_i(\bar{t}|I_{N_s^D}, T_{N_s^D}, x_1))^T, \Delta s-\bar{s}_k\Bigg), \end{aligned}$$

$$\begin{aligned} &f_{t_k}^{j,i}(s+\bar{t}_k; \beta, s, \Delta s) \\ &= \sum_{l=0,1} P_{jl}(\bar{t}_k)P_{li}(\Delta s-\bar{t}_k)\lambda_\beta(s+\bar{t}_k|I_{N_s^D}, T_{N_s^D}, X_s=x_l) \\ &\quad\times \Psi_{jl}(-(\eta_{C(N_s^Y)}(x_0), \eta_{C(N_s^Y)}(x_1))^T, \bar{t}_k) \\ &\quad\times \Psi_{li}\left(-(\eta_{C(N_s^Y)}(x_0), \eta_{C(N_s^Y)}(x_1))^T, \Delta s-\bar{t}_k\right) \\ &\quad\times \bar{\Psi}_{jl}\Bigg(s, -\sum_{i\in\bar{I}_{N_s^D}} (\lambda_i(\bar{t}|I_{N_s^D}, T_{N_s^D}, x_0), \\ &\qquad\qquad \times \lambda_i(\bar{t}|I_{N_s^D}, T_{N_s^D}, x_1))^T, \bar{t}_k\Bigg) \\ &\quad\times \bar{\Psi}_{li}\Bigg(s+\bar{t}_k, -\sum_{i\in\bar{I}^*_{N_s^D}} (\lambda_i(\bar{t}|I^*_{N_s^D}, T^*_{N_s^D}, x_0), \\ &\qquad\qquad \times \lambda_i(\bar{t}|I^*_{N_s^D}, T^*_{N_s^D}, x_1))^T, \Delta s-\bar{t}_k\Bigg) \end{aligned}$$

where $I^*_{N_s^D} = I_{N_s^D}\bigcup\{\beta\}$, $T^*_{N_s^D} = T_{N_s^D}\bigcup\{t_\beta\}$ and

$$C(x) = \begin{cases} 1, & x+Y_0 \equiv 0 \pmod 2 \\ 0, & x+Y_0 \equiv 1 \pmod 2). \end{cases}$$

If up to time t, no jump or default has been observed, then we have the following: for $\widetilde{\omega} \in$ {no jump or default observed in $[0, t]$},

$$P(X_t=x_i \mid \mathcal{F}_t) = \frac{P(X_t=x_i, \text{no jump or default in}[0,t])}{\sum_{j=0,1} P(X_t=x_j, \text{no jump or default in}[0,t])}$$

where

$$\begin{aligned} &P(X_t=x_j, \text{no jump or default in}[0,t]) \\ &= P(X_t=x_j)\Psi_{0j}\left(-(\eta_{C(0)}(x_0), \eta_{C(0)}(x_1))^T, t\right) \times \bar{\Psi}_{0j}(0, \\ &\quad -\textstyle\sum_{i\in I}(\lambda_i(\bar{t}|I_{N_0^D}, T_{N_0^D}, x_0), \lambda_i(\bar{t}|I_{N_0^D}, T_{N_0^D}, x_1))^T, t). \end{aligned}$$

Note that if the jump intensities of chain Y, η_i $(i = 0, 1)$, are not as simple as in our assumptions and they are also related with time directly, i.e. $\eta_i(\bar{t})$, all the algorithms introduced above are still applicable and we just need to replace $\Psi_{ij}(-(\eta_{C(0)}(x_0), \eta_{C(0)}(x_1))^T, t)$ by $\bar{\Psi}_{ij}(-(\eta_{C(0)}(x_0), \eta_{C(0)}(x_1))^T, t)$, $i, j = 0, 1$. This replacement holds only when equation (3) given $\mathbf{u}(\bar{t}) = -(\eta_0(\bar{t}), \eta_1(\bar{t}))$ has an analytical solution.

Even if equation (3) does not admit an analytical solution given $u_i(\bar{t})$, $(i = 0, 1)$, we can still provide a numerical method in section 5. Now we know how to get $P(X_t = x_i|\mathcal{F}_t)$.

4. Default distributions

We derive the default distributions in this section. Besides deriving closed-form expressions for the default distributions, extended total hazard construction method for HMM to derive the joint default distribution is also presented.

4.1. *Closed-form expressions for default distributions*

In this subsection, we compute the conditional joint distribution of default times

$$P(\tau^1 > t^1, \tau^2 > t^2, \ldots, \tau^K > t^K \mid \mathcal{F}_t)$$

and the distribution of ordered default times

$$P(\tau^k > s \mid \mathcal{F}_t), \quad k = 1, 2, \ldots, K.$$

Notice that when $t = 0$, we don't have any information, the above two conditional probabilities become unconditional probabilities. As for the first probability, due to the Markov property of X_t and the structure of $\lambda_i(t)$, we have

$$\begin{aligned} &P(\tau^1 > t^1, \tau^2 > t^2, \ldots, \tau^K > t^K \mid \mathcal{F}_t) \\ &= \sum_{i=0,1} P(\tau^1 > t^1, \tau^2 > t^2, \ldots, \tau^K > t^K \mid \mathcal{F}_t^N, X_t = x_i) \\ &\quad \times P(X_t = x_i \mid \mathcal{F}_t). \end{aligned}$$

Since we know how to calculate $P(X_t = x_i \mid \mathcal{F}_t)$, we only need to compute the conditional joint probability $P(\tau^1 > t^1, \tau^2 > t^2, \ldots, \tau^K > t^K \mid \mathcal{F}_t^N, X_t = x_i)$.

Assume we first enter the market immediately after the N_t^Dth default of the K obligors at time t, to simplify the notations, we denote $m = N_t^D$, that means we have already known the information $T_m = (t_1, \ldots, t_m)$, $I_m = (j_1, \ldots, j_m)$ and $\mathcal{F}_t^Y$ by time t. Then we can get the following equation:

$$\begin{aligned} &P(\tau^1 > t^1, \tau^2 > t^2, \ldots, \tau^K > t^K \mid \mathcal{F}_t^N, X_t = x_i) \\ &\quad = P(\tau^{j_{m+1}} > t^{j_{m+1}}, \ldots, \tau^{j_K} > t^{j_K} \mid \tau^{j_1} \\ &\quad = t_1, \ldots, \tau^{j_m} = t_m, X_t = x_i). \end{aligned}$$

Furthermore, we also know the relationship that

$$\begin{aligned} f(t^{j_{m+1}}, \ldots, t^{j_K} \mid \mathcal{F}_t) &= (-1)^{K-m} \frac{\mathrm{d}^{K-m}}{\mathrm{d}t^{j_{m+1}} \ldots \mathrm{d}t^K} \\ &\quad \times P(\tau^1 > t^1, \tau^2 > t^2, \ldots, \tau^K > t^K \mid \mathcal{F}_t^N, X_t = x_i) \end{aligned}$$

where $f(t^{j_{m+1}}, \ldots, t^{j_K} \mid \mathcal{F}_t)$ is the conditional joint density function. Therefore, to obtain the desired conditional probability, it suffices to find its conditional joint density function.

Here we employ the approach introduced by Yu (2007) (called the total hazard construction method) to derive the conditional density function.

PROPOSITION 2 *The expression of the density function is in the form of expectation*

$$\begin{aligned} &f(t^{j_{m+1}}, \ldots, t^{j_K} \mid \mathcal{F}_t) \\ &= E\left[\sum_{l=m+1}^{K} \sum_{i \in \bar{I}_l} \lambda_i(t^{j_l} | I_l, T_l, X_{t^{j_l}}) \cdot \right. \\ &\quad \left. \times \exp\left(-\sum_{l=m+1}^{K}\left(\sum_{i\in\bar{I}_l}\int_{t_l}^{t^{j_l}} \lambda_i(u|I_l, T_l, X_u)\mathrm{d}u\right)\right)\right]. \end{aligned}$$

Proof. Without loss of generality, we assume that $t^{j_{m+1}} < \cdots < t^{j_K}$. In this case, $\tau^{m+1} - \tau^m$ would be the first default time we observed after entering the market. By using the total hazard construction method pioneered by Yu (2007) with the information already known, we draw a collection of independent standard exponential random variables: $(E_{j_{m+1}}, \ldots, E_{j_K})$. Then we know

$$\tau^{m+1} - \tau^m = \min_{i \in \bar{I}_m} \Lambda_i^{-1}(E_i) = \min_{i \in \bar{I}_m} \inf\{s \geq 0 : \Lambda_i(s) \geq E_i\}$$

which implies that

$$\begin{aligned} &P\left(\tau^{m+1} - \tau^m > t \mid \mathcal{F}_{\tau^m}\right) \\ &\quad = P\left(\min_{i \in \bar{I}_m} \inf\{s \geq 0 : \Lambda_i(s) \geq E_i\} > t\right). \end{aligned}$$

Suppose the information $\mathcal{F}_\infty^X$ is known, then

$$\begin{aligned} &P(\tau^{m+1} - \tau^m > t \mid \mathcal{F}_{\tau^m}) \\ &\quad = \prod_{i \in \bar{I}_m} P\left(E_i > \int_{t_m}^{t_m+t} \lambda_i(u|I_m, T_m, X_u)\mathrm{d}u\right) \\ &\quad = \prod_{i \in \bar{I}_m} \exp\left(-\int_{t_m}^{t_m+t} \lambda_i(u|I_m, T_m, X_u)\mathrm{d}u\right) \\ &\quad = \exp\left(-\sum_{i \in \bar{I}_m}\int_{t_m}^{t_m+t} \lambda_i(u|I_m, T_m, X_u)\mathrm{d}u\right). \end{aligned}$$

If we assume that $\tau^m < t < \tau^{m+1}$ and $t^i > \tau^i, i = 1, \ldots, m$, and let $\lambda^{m+1}(t)$ denote the $(m+1)$th default rate at time t, then

$$\lambda^{m+1}(t) = \sum_{i \in \bar{I}_m} \int_{t_m}^{t} \lambda_i(u|I_m, T_m, X_u)\mathrm{d}u.$$

Since

$$\begin{aligned} &P(\tau^{m+1} > t \mid \mathcal{F}_{\tau^m}, X_{s(t_m<s<\infty)}) \\ &\quad = e^{-\sum_{i\in\bar{I}_m}\int_{t_m}^{t} \lambda_i(u|I_m,T_m,X_u)\mathrm{d}u} = e^{-\lambda^{m+1}(t)} \end{aligned}$$

we have

$$\begin{aligned} &P(\tau^{j_{m+1}} > t^{j_{m+1}}, \ldots, \tau^{j_K} > t^{j_K} \mid \mathcal{F}_t, X_{s(t_m<s<\infty)}) \\ &= \prod_{l=m+1}^{K} P(\tau^{j_l} > t^{j_l} \mid \mathcal{F}_t, X_{s(t_m<s<\infty)}) \\ &= \prod_{l=m+1}^{K} e^{-\lambda^l(t^{j_l})} \\ &= \prod_{l=m+1}^{K} \exp\left(-\sum_{i\in\bar{I}_{l-1}}\int_{t_{l-1}}^{t^{j_l}} \lambda_i(u|I_{l-1}, T_{l-1}, X_u)\mathrm{d}u\right) \\ &= \exp\left(-\sum_{l=m+1}^{K}\left(\sum_{i\in\bar{I}_{l-1}}\int_{t_{l-1}}^{t^{j_l}} \lambda_i(u|I_{l-1}, T_{l-1}, X_u)\mathrm{d}u\right)\right) \end{aligned}$$

and therefore

$$\begin{aligned} &f(t^{j_{m+1}}, \ldots, t^{j_K} \mid \mathcal{F}_t, X_{s(t_m<s<\infty)}) \\ &= (-1)^{K-m}\frac{\mathrm{d}^{K-m}}{\mathrm{d}t^{j_{m+1}}\ldots\mathrm{d}t^K} P(\tau^{j_{m+1}} > t^{j_{m+1}}, \ldots, \tau^{j_K} \\ &\quad > t^{j_K} \mid \mathcal{F}_t, X_{s(t_m<s<\infty)}) \\ &= (-1)^{K-m}\frac{\mathrm{d}^{K-m}}{\mathrm{d}t^{j_{m+1}}\ldots\mathrm{d}t^K} \\ &\quad \times \exp\left(-\sum_{l=m+1}^{K}\left(\sum_{i\in\bar{I}_{l-1}}\int_{t_{l-1}}^{t^{j_l}} \lambda_i(u|I_{l-1}, T_{l-1}, X_u)\mathrm{d}u\right)\right) \\ &\quad\quad \times |_{t_{l-1}=t^{j_{l-1}}} \\ &= \prod_{l=m+1}^{K} \sum_{i\in\bar{I}_{l-1}} \lambda_i(t^{j_l}|I_{l-1}, T_{l-1}, X_{t^{j_l}}) \cdot \end{aligned}$$

$$\times \exp\left(-\sum_{l=m+1}^{K}\left(\sum_{i\in\bar{I}_{l-1}}\int_{t^{j_{l-1}}}^{t^{j_l}}\lambda_i(u|I_{l-1},T_{l-1},X_u)\mathrm{d}u\right)\right)$$

and

$$\begin{aligned}
&f(t^{j_{m+1}},\ldots,t^{j_K}\mid\mathcal{F}_t)\\
&= E[f(t^{j_{m+1}},\ldots,t^{j_K}\mid\mathcal{F}_t,X_{s(t_m<s<\infty)})]\\
&= E\Bigg[\prod_{l=m+1}^{K}\sum_{i\in\bar{I}_{l-1}}\lambda_i(t^{j_l}|I_{l-1},T_{l-1},X_{t^{j_l}})\cdot\\
&\quad\times\exp\left(-\sum_{l=m+1}^{K}\left(\sum_{i\in\bar{I}_{l-1}}\int_{t^{j_{l-1}}}^{t^{j_l}}\lambda_i(u|I_{l-1},T_{l-1},X_u)\mathrm{d}u\right)\right)\Bigg]\\
&= E\Bigg[\prod_{l=m+1}^{K}\sum_{i\in\bar{I}_{l-1}}\lambda_i(t^{j_l}|I_{l-1},T_{l-1},X_{t^{j_l}})\cdot\\
&\quad\times\exp\left(-\sum_{l=m+1}^{K}\left(\int_{t^{j_{l-1}}}^{t^{j_l}}\sum_{i\in\bar{I}_{l-1}}\lambda_i(u|I_{l-1},T_{l-1},X_u)\mathrm{d}u\right)\right)\Bigg].
\end{aligned}$$

□

If equation (3) in the previous section given $\mathbf{u}(\bar{t}) = -(\lambda_i(x_0),\lambda_i(x_1))$, $i = 1,\ldots,K$, has unique solutions, then we further have the following result.

PROPOSITION 3 *The explicit formula for calculating the desired density function is given by*

$$\begin{aligned}
&f(t^{j_{m+1}},\ldots,t^{j_K}\mid\mathcal{F}_t)\\
&\qquad=(-1)^{K-m}\cdot\sum_{l_{m+1}=0,1}\sum_{l_{m+2}=0,1}\cdots\sum_{l_K=0,1}\\
&\cdot\frac{\mathrm{d}(\bar{\Psi}_{il_{m+1}}(t^{j_m},-\sum_{i\in\bar{I}_m}(\lambda_i(\bar{t}|I_m,T_m,x_0),\lambda_i(\bar{t}|I_m,T_m,x_1))^T,t^{j_{m+1}}-t^{j_m}))}{\mathrm{d}t^{j_{m+1}}}\\
&\cdot\frac{\mathrm{d}(\bar{\Psi}_{l_{m+1}l_{m+2}}(t^{j_{m+1}},-\sum_{i\in\bar{I}_{m+1}}(\lambda_i(\bar{t}|I_{m+1},T_{m+1},x_0),\lambda_i(\bar{t}|I_{m+1},T_{m+1},x_1))^T,t^{j_{m+2}}-t^{j_{m+1}}))}{\mathrm{d}t^{j_{m+2}}}\\
&\cdot\;\cdots\cdot\frac{\mathrm{d}(\bar{\Psi}_{l_{K-1}l_K}(t^{j_{K-1}},-\sum_{i\in\bar{I}_{K-1}}(\lambda_i(\bar{t}|I_{K-1},T_{K-1},x_0),\lambda_i(\bar{t}|I_{K-1},T_{K-1},x_1))^T,t^{j_K}-t^{j_{K-1}}))}{\mathrm{d}t^{j_K}}
\end{aligned}$$

where $\bar{\Psi}_{ij}, i, j = 0, 1$ *are the moment generating function defined in section 3.*

Proof. We note that

$$\begin{aligned}
&E\Bigg[\prod_{l=m+1}^{K}\sum_{i\in\bar{I}_{l-1}}\lambda_i(t^{j_l}|I_{l-1},T_{l-1},X_{t^{j_l}})\\
&\cdot e^{-\sum_{l=m+1}^{K}(\int_{t^{j_{l-1}}}^{t^{j_l}}\sum_{i\in\bar{I}_{l-1}}\lambda_i(u|I_{l-1},T_{l-1},X_u)\mathrm{d}u)}\Bigg]\\
&=\sum_{l_{m+1}=0,1}\sum_{l_{m+2}=0,1}\cdots\sum_{l_K=0,1}\\
&E\Bigg[\sum_{i\in\bar{I}_m}\lambda_i(t^{j_{m+1}}|I_m,T_m,X_{t^{j_{m+1}}})\\
&\cdot e^{\int_{t^{j_m}}^{t^{j_{m+1}}}\sum_{i\in\bar{I}_m}\lambda_i(u|I_m,T_m,X_u)\mathrm{d}u}|X_{t^{j_m}}=i,X_{t^{j_{m+1}}}=l_{m+1}\Bigg]\\
&\cdot E\Bigg[\sum_{i\in\bar{I}_{m+1}}\lambda_i(t^{j_{m+2}}|I_{m+1},T_{m+1},X_{t^{j_{m+2}}})\\
&\cdot e^{\int_{t^{j_{m+1}}}^{t^{j_{m+2}}}\sum_{i\in\bar{I}_{m+1}}\lambda_i(u|I_{m+1},T_{m+1},X_u)\mathrm{d}u}|X_{t^{j_{m+1}}}\\
&\qquad=l_{m+1},X_{t^{j_{m+2}}}=l_{m+2}\Bigg]\\
&\cdots\cdot E\Bigg[\sum_{i\in\bar{I}_{K-1}}\lambda_i(t^{j_K}|I_{K-1},T_{K-1},X_{t^{j_K}})\\
&\cdot e^{\int_{t^{j_{K-1}}}^{t^{j_K}}\sum_{i\in\bar{I}_{K-1}}\lambda_i(u|I_{K-1},T_{K-1},X_u)\mathrm{d}u}|X_{t^{j_{K-1}}}\\
&\qquad=l_{K-1},X_{t^{j_K}}=l_K\Bigg]\\
&=(-1)^{K-m}\cdot\sum_{l_{m+1}=0,1}\sum_{l_{m+2}=0,1}\cdots\sum_{l_K=0,1}\\
&\times\frac{\mathrm{d}\left(E\left[e^{\int_{t^{j_m}}^{t^{j_{m+1}}}\sum_{i\in\bar{I}_m}\lambda_i(u|I_m,T_m,X_u)\mathrm{d}u}|X_{t^{j_m}}=i,X_{t^{j_{m+1}}}=l_{m+1}\right]\right)}{\mathrm{d}t^{j_{m+1}}}\\
&\cdot\frac{\mathrm{d}\left(E\left[e^{\int_{t^{j_{m+1}}}^{t^{j_{m+2}}}\sum_{i\in\bar{I}_{m+1}}\lambda_i(u|I_{m+1},T_{m+1},X_u)\mathrm{d}u}|X_{t^{j_{m+1}}}=l_{m+1},X_{t^{j_{m+2}}}=l_{m+2}\right]\right)}{\mathrm{d}t^{j_{m+2}}}\\
&\cdots\cdot\frac{\mathrm{d}\left(E\left[e^{\int_{t^{j_{K-1}}}^{t^{j_K}}\sum_{i\in\bar{I}_{K-1}}\lambda_i(u|I_{K-1},T_{K-1},X_u)\mathrm{d}u}|X_{t^{j_{K-1}}}=l_{K-1},X_{t^{j_K}}=l_K\right]\right)}{\mathrm{d}t^{j_K}}\\
&=(-1)^{K-m}\cdot\sum_{l_{m+1}=0,1}\sum_{l_{m+2}=0,1}\cdots\sum_{l_K=0,1}\\
&\times\frac{\mathrm{d}(\bar{\Psi}_{il_{m+1}}(t^{j_m},-\sum_{i\in\bar{I}_m}(\lambda_i(\bar{t}|I_m,T_m,x_0),\lambda_i(\bar{t}|I_m,T_m,x_1))^T,t^{j_{m+1}}-t^{j_m}))}{\mathrm{d}t^{j_{m+1}}}\cdot\\
&\times\frac{\mathrm{d}(\bar{\Psi}_{l_{m+1}l_{m+2}}(t^{j_{m+1}},-\sum_{i\in\bar{I}_{m+1}}(\lambda_i(\bar{t}|I_{m+1},T_{m+1},x_0),\lambda_i(\bar{t}|I_{m+1},T_{m+1},x_1))^T,t^{j_{m+2}}-t^{j_{m+1}}))}{\mathrm{d}t^{j_{m+2}}}\\
&\cdots\cdots\\
&\times\frac{\mathrm{d}(\bar{\Psi}_{l_{K-1}l_K}(t^{j_{K-1}},-\sum_{i\in\bar{I}_{K-1}}(\lambda_i(\bar{t}|I_{K-1},T_{K-1},x_0),\lambda_i(\bar{t}|I_{K-1},T_{K-1},x_1))^T,t^{j_K}-t^{j_{K-1}}))}{\mathrm{d}t^{j_K}}
\end{aligned}$$

□

Similarly if equations related to $\bar{\Psi}_{ij}$ do not have analytical solutions, then we can use the same approximation method which will be discussed in the next section to approximate $\bar{\Psi}_{ij}$ with Ψ_{ij}. Thus one can obtain an explicit approximation expression for the density function $f(t^{j_{m+1}},\ldots,t^{j_K}\mid\mathcal{F}_t)$. When the expressions of the default intensities are homogeneous and symmetric,

$$\begin{aligned}
&P(\tau^{j_{m+1}}<\cdots<\tau^{j_k}<s<\tau^{j_{k+1}}<\cdots<\tau^{j_K}\mid\mathcal{F}_t)\\
=&\int_{t_m}^{t}\int_{t^{j_{m+1}}}^{t}\cdots\int_{t^{j_{k-1}}}^{t}\int_{t}^{\infty}\int_{t^{j_{k+1}}}^{\infty}\\
&\cdots\int_{t^{j_{K-1}}}^{\infty}f(t^{j_{m+1}},t^{j_{m+2}},\ldots,t^{j_K}\mid\mathcal{F}_s)\mathrm{d}t^{j_K}\ldots\mathrm{d}t^{t_{j_{m+1}}}.
\end{aligned}$$

Because they are homogeneous and symmetric,

$$\begin{aligned}
P(\tau^{j_k}\le s<\tau^{j_{k+1}}\mid\mathcal{F}_t)&=(K-m)!P(\tau^{j_{m+1}}<\cdots<\tau^{j_k}\\
&\quad<s<\tau^{j_{k+1}}<\cdots<\tau^{j_K}\mid\mathcal{F}_t).
\end{aligned}$$

Furthermore, we have

$$P(\tau^{j_k}>s\mid\mathcal{F}_t)=\sum_{i=m}^{k-1}P(\tau^{j_i}\le s<\tau^{j_{i+1}}\mid\mathcal{F}_t).$$

4.2. Extended total hazard construction method for HMM

We further extend the total hazard construction method to make it applicable to various forms of default intensities modulated by a hidden Markov process, then to gain the joint default distribution.

The total hazard accumulated by obligor i by time t, denoted by $\psi_i(t|I_{N_t^D}, T_{N_t^D}, X_t)$, can be defined as follows:

$$\psi_i(t|I_{N_t^D}, T_{N_t^D}, X_t) = \sum_{l=0}^{N_t^D-1} \Lambda_i(t_{l+1} - t_l|I_l, T_l, X_{t_{l+1}}) + \Lambda_i(t - t_{N_t^D}|I_{N_t^D}, T_{N_t^D}, X_t) \quad (8)$$

where

$$\Lambda_i(s|I_l, T_l, X_{t_l+s}) = \int_{t_l}^{t_l+s} \lambda_i(\mu|I_l, T_l, X_\mu)\mathrm{d}\mu \quad (9)$$

is the total hazard accumulated by obligor i in the time interval $[t_l, t_l + s]$. Note that the default processes are independent unit exponential random variables. And we define the inverse function

$$\Lambda_i^{-1}(x|I_k, T_k, N_\infty^Y, S_{N_\infty^Y}) = \inf\{s : \Lambda_i(s|I_k, T_k, X_{t_k+s}) \geq x\}, x \geq 0 \quad (10)$$

where $(N_\infty^Y, S_{N_\infty^Y}) \in \mathcal{F}_\infty^Y$ is the entire historical path of Y, N_∞^Y is the entire number of jump in the chain Y and $S_{N_\infty^Y}$ is the collection of corresponding ordered jump times.

The total hazard can be constructed by the following recursive procedure:

Step 1 Generate a complete sample path of Y, and denote it as $(N_\infty^Y, S_{N_\infty^Y}) \in \mathcal{F}_\infty^Y$.
Generate a collection of i.i.d. unit exponential random variables $(E_1, \ldots, E_K)$.

Step 2 Let $j_1 = \arg\min\{\Lambda_i^{-1}(E_i) : i = 1, \ldots, K\}$ and define $\hat{\tau}^{j_1} = \Lambda_{j_1}^{-1}(E_{j_1})$.
Note that $T_1 = (t_1), t_1 = \hat{\tau}^{j_1}, I_1 = (j_1)$.

Step 3 (i)Assume that $(\hat{\tau}^{j_1}, \ldots, \hat{\tau}^{j_{m-1}})$ and the simulated path of $X_s(0 \leq s < \hat{\tau}^{j_{m-1}})$ are already obtained as $T_{m-1} = (t_1, \ldots, t_{m-1}), t_l = \hat{\tau}^{j_l}, l = 1, \ldots, m-1$ and $I_{m-1} = (j_1, \ldots, j_{m-1})$, where $m \geq 2$. By using the conditional probability of $P(X_s = x_i|\tilde{\mathcal{F}}_s)$, $i = 0, 1, x_0 = 0, x_1 = 1, s \geq \hat{\tau}^{j_{m-1}}$ and $\tilde{\mathcal{F}}_s = \mathcal{F}_s^Y \vee T_{m-1} \vee I_{m-1}$, we can generate a sequence of random numbers of X_s under this conditional probability.
We can then obtain the simulated path of $X_s, s \geq \hat{\tau}^{j_{m-1}}$ which will be useful in the calculation of $\Lambda_i^{-1}(x|I_{m-1}, T_{m-1}, N_\infty^Y, S_{N_\infty^Y})$. (ii) Note that $\bar{I}_{m-1} = (1, 2, \ldots, K)\backslash I_{m-1}$.
Therefore, with the information of T_{m-1}, I_{m-1} and the path of $X_s(0 \leq s < \hat{\tau}^{j_{m-1}}) \cup X_s(s \geq \hat{\tau}^{j_{m-1}})$, i.e. the path of X. We let $j_m = \arg\min\{\Lambda_i^{-1}(E_i - \psi_i(t_{m-1}|I_{m-1}, T_{m-1}, X_{t_{m-1}})|I_{m-1}, T_{m-1}, N_\infty^Y, S_{N_\infty^Y}) : i \in \bar{I}_{m-1}\}$ where $\psi_i(t_{m-1}|I_{m-1}, T_{m-1}, X_{t_{m-1}})$ is the total hazard accumulated by Name i under the condition of defaults and information of chain X by the $(m-1)$th default time, i.e. t_{m-1}.
Then we let $\hat{\tau}^{j_m} = t_{m-1} + \Lambda_{j_m}^{-1}(E_{j_m} - \psi_{j_m}(t_{m-1}| I_{m-1}, T_{m-1}, X_{t_{m-1}})|I_{m-1}, T_{m-1}, N_\infty^Y, S_{N_\infty^Y})$ and reserve the simulated path of $X_s, \hat{\tau}^{j_{m-1}} \leq s < \hat{\tau}^{j_m}$ at this step.
Thus, with the simulated path, we can get the simulated path of $X_s, 0 \leq s < \hat{\tau}^{j_m}$.

Step 4 If $m = K$, then stop. Otherwise, increase m by 1 and go to **Step 3**.

From the recursive procedure, we can obtain the distribution of $\hat{\tau}$. According to Shaked and Shanthikumar (1987) and Yu (2007), the distribution of $\hat{\tau}$ obtained from the above recursive processes is equal to the distribution of the original default time τ. This gives the following results.

PROPOSITION 4 *Let τ be the default time with the intensities*

$$\lambda_t^i = \lambda_i(t|I_{N_t^D}, T_{N_t^D}, X_t), \quad i = 1, \ldots, K$$

and the related jump processes satisfying the assumptions mentioned in section 2. Construct $\hat{\tau}$ according to Steps 1–4 with the intensity equal to

$$\lambda_i(t|I_{N_t^D}, T_{N_t^D}, X_t), \quad i = 1, 2, \ldots, K.$$

Let $\mathcal{F}'_t$ be the minimal filtration containing $\mathcal{F}_t^Y$ and the information of the default processes related to $\hat{\tau}$ by time t, and P' be the distribution of $(Y, \hat{\tau})$. Then every element in $\hat{\tau}^i$ has $(P', \mathcal{F}'_t)$-intensity of the form:

$$\lambda_i(t|I_{N_t^D}, T_{N_t^D}, X_t), \quad i = 1, 2, \ldots, K.$$

Therefore, we can generate τ by just generating $\hat{\tau}$.

5. Numerical approximation method

In this section, we consider an outstanding problem in Section 3. If equation (3) does not admit an analytical solution given $u_i(\bar{t})$, $(i = 0, 1)$ then we shall try to use another method to approximate the conditional probability $P(X_t = x_i|\mathcal{F}_t)$.

We can consider approximating $\bar{\Psi}_{ij}(s_0, \mathbf{u}, t)$ directly. As we mentioned before, it is because of the default intensities λ_i, $(i = 1, \ldots, K)$ which give equation (3) with

$$\mathbf{u}(\bar{t}) = -(\lambda_i(x_0), \lambda_i(x_1)), \quad i = 1, \ldots, K$$

does not have an analytical solution, and hence we cannot obtain closed-form expressions for $\bar{\Psi}_{ij}(s_0, \mathbf{u}, t)$. Thus, we need to approximate the moment generating function $\bar{\Psi}_{ij}(s_0, \mathbf{u}, t)$ when the default intensities are applied. If the error of $\bar{\Psi}_{ij}(s_0, \mathbf{u}, t)$ is less than any arbitrary ϵ then according to the expression of $f_{\bar{s}_k}^{j,i}(s + \bar{s}_k; s, \Delta s)$ and $f_{\bar{t}_k}^{j,i}(s + \bar{t}_k; \beta, s, \Delta s)$ given below, we know that their relative errors can be controlled. Furthermore, from the recursive method for $P(X_t = x_i|\mathcal{F}_t)$ presented in section 3, the error of this conditional probability may be controlled.

In the following, we are going to illustrate how the approximation works. When the length of the time interval length is small enough, without loss of generality, we can approximately assign t in the default intensities $\lambda_i(t)$ to be the left value of the concerned time interval, i.e. $t = s_0$ when the time interval is $[s_0, s_0 + \Delta\bar{s}]$. Then we can still apply the moment generating function given $\mathbf{u}(\bar{t}) = \mathbf{u}(s_0) = \bar{\mathbf{u}}$, and we know the corresponding equation (3) has a unique solution. But we need to ensure that by using this method, the error of $\bar{\Psi}_{ij}(s_0, \mathbf{u}, \Delta\bar{s})$

can be controlled such that it can be less than any arbitrarily given ϵ.

PROPOSITION 5 *The error control* $\Delta\Psi_{ij}(s_0, \mathbf{u}, \Delta\bar{s}) < \epsilon < 1$, *where* ϵ *is arbitrary, can be achieved by requiring* $\Delta\bar{s}$ *to satisfy*

$$\Delta\bar{s} < \frac{-\ln(1-\epsilon)}{K \cdot \lambda_{\max}(s_0)}$$

where

$$\lambda_{\max}(s_0) = \max_{i=1,\ldots,K}\{\lambda_i(s), s \in [0, s_0]\}$$

and

$$\Delta\Psi_{ij}(s_0, \mathbf{u}, \Delta\bar{s}) = |\Psi_{ij}(\bar{\mathbf{u}}, \Delta\bar{s}) - \bar{\Psi}_{ij}(s_0, \mathbf{u}, \Delta\bar{s})|$$

and

$$\bar{\mathbf{u}}(t) = \mathbf{u}(\tilde{s}_{k-1}) \quad \text{for} \quad t \in (\tilde{s}_{k-1}, \tilde{s}_k]$$

and

$$[0, s_0] = [\tilde{s}_0, \tilde{s}_1]\bigcup(\tilde{s}_1, \tilde{s}_2]\bigcup\cdots\bigcup(\tilde{s}_{n-1}, \tilde{s}_n].$$

Proof. Note that there are K entities, so when the default intensity is applied, i.e.

$$\mathbf{u}(\bar{t}) = -(\lambda_i(x_0), \lambda_i(x_1)) \quad \text{or}$$
$$\mathbf{u} = -(\lambda_i(x_0), \lambda_i(x_1)), \quad i = 1, \ldots, K,$$

we notice the relationships that

$$E[e^{-K\cdot\lambda_{\max}(s_0)\cdot\Delta\bar{s}}] \le \Psi_{ij}(\bar{\mathbf{u}}, \Delta\bar{s}) \le E[e^{K\cdot 0\cdot\Delta\bar{s}}]$$

and

$$E[e^{-K\cdot\lambda_{\max}(s_0)\cdot\Delta\bar{s}}] \le \bar{\Psi}_{ij}(s_0, \mathbf{u}, \Delta\bar{s}) \le E[e^{K\cdot 0\cdot\Delta\bar{s}}].$$

Since all $\lambda_i, i = 1, \ldots, K$ are nonnegative, we have the following relationship:

$$\Delta\Psi_{ij}(s_0, \mathbf{u}, \Delta\bar{s}) \le E[e^{K\cdot 0\cdot\Delta\bar{s}} - e^{-K\cdot\lambda_{\max}(s_0)\cdot\Delta\bar{s}}] < \epsilon$$

if and only if

$$e^{K\cdot\lambda_{\max}(s_0)\cdot\Delta\bar{s}} < \frac{1}{1-\epsilon}$$

if and only if

$$\Delta\bar{s} < \frac{-\ln(1-\epsilon)}{K \cdot \lambda_{\max}(s_0)}.$$

□

We can simply let $\Delta\bar{s} = \frac{-\ln(1-\epsilon)}{K\cdot\lambda_{\max}(s_0)}$, it is enough to ensure the error of $\bar{\Psi}_{ij}(s_0, \mathbf{u}, \Delta\bar{s})$ controllable. Here we are in the position to approximate $f_{\bar{s}_k}^{j,i}(s+\bar{s}_k; s, \Delta s)$. First, we partition the time interval $[s, s+\bar{s}_k]$ evenly with step size equal to $\Delta\bar{s} = \frac{-\ln(1-\epsilon)}{K\cdot\lambda_{\max}(s+\Delta s)}$, and denote $M_1(s, \Delta\bar{s}) = \left[\frac{\bar{s}_k}{\Delta\bar{s}}\right]$. That is to say,

$$[s, s+\bar{s}_k] = [s, s+\Delta\bar{s}]\bigcup[s+\Delta\bar{s}, s+2\Delta\bar{s}]$$
$$\times\bigcup\cdots\bigcup[s+M_1(s, \Delta\bar{s})\cdot\Delta\bar{s}, \bar{s}_k]$$

Moreover, we do the same thing for the remaining time interval: $[s+\bar{s}_k, \Delta s]$ and denote $M_2(s, \Delta\bar{s}) = \left[\frac{\Delta s - \bar{s}_k}{\Delta\bar{s}}\right]$. We denote $M_1 = M_1(s, \Delta\bar{s})$ and $M_2 = M_2(s, \Delta\bar{s})$. Now the explicit approximation formula is given as follows:

$$f_{\bar{s}_k}^{j,i}(s+\bar{s}_k; s, \Delta s) = \sum_{l=0,1}\sum_{l_1=0,1}\cdots\sum_{l_{M_1}=0,1}\sum_{\bar{l}_1=0,1}$$
$$\cdots\sum_{\bar{l}_{M_2}=0,1} P_{jl}(\bar{s}_k)P_{li}(\Delta s - \bar{s}_k)\eta_{C(N_s^Y)}(x_l)$$
$$\times\Psi_{jl}\left(-(\eta_{C(N_s^Y)}(x_0), \eta_{C(N_s^Y)}(x_1))^T, \bar{s}_k\right)$$
$$\times\Psi_{li}\left(-(\eta_{C(N_s^Y+1)}(x_0), \eta_{C(N_s^Y+1)}(x_1))^T, \Delta s - \bar{s}_k\right)$$
$$\times\Psi_{jl_1}\Bigg(-\sum_{i\in\bar{I}_{N_s^D}}(\lambda_i(s|I_{N_s^D}, T_{N_s^D}, x_0),$$
$$\times\lambda_i(s|I_{N_s^D}, T_{N_s^D}, x_1))^T, \Delta\bar{s}\Bigg)$$
$$\times\Psi_{l_1l_2}\Bigg(-\sum_{i\in\bar{I}_{N_s^D}}(\lambda_i(s+\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_0),$$
$$\times\lambda_i(s+\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_1))^T, \Delta\bar{s}\Bigg)$$
$$\times\cdots$$
$$\times\Psi_{l_{M_1}l}\Bigg(-\sum_{i\in\bar{I}_{N_s^D}}(\lambda_i(s+M_1\cdot\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_0),$$
$$\times\lambda_i(s+M_1\cdot\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_1))^T, \bar{s}_k - s - M_1\cdot\Delta\bar{s}\Bigg)$$
$$\times\Psi_{l\bar{l}_1}\Bigg(-\sum_{i\in\bar{I}_{N_s^D}}(\lambda_i(\bar{s}_k|I_{N_s^D}, T_{N_s^D}, x_0),$$
$$\times\lambda_i(\bar{s}_k|I_{N_s^D}, T_{N_s^D}, x_1))^T, \Delta\bar{s}\Bigg)$$
$$\times\Psi_{\bar{l}_1\bar{l}_2}\Bigg(-\sum_{i\in\bar{I}_{N_s^D}}(\lambda_i(\bar{s}_k+\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_0),$$
$$\times\lambda_i(\bar{s}_k+\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_1))^T, \Delta\bar{s}\Bigg)$$
$$\times\cdots$$
$$\times\Psi_{\bar{l}_{M_2}i}\Bigg(-\sum_{i\in\bar{I}_{N_s^D}}(\lambda_i(\bar{s}_k+M_2\cdot\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_0),$$
$$\times\lambda_i(\bar{s}_k+M_2\cdot\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_1))^T,$$
$$\times\Delta s - \bar{s}_k - M_2\cdot\Delta\bar{s}\Bigg).$$

Similarly, we can get

$$f_{\bar{t}_k}^{j,i}(s+\bar{t}_k; \beta, s, \Delta s) = \sum_{l=0,1}\sum_{l_1=0,1}\cdots\sum_{l_{\bar{M}_1}=0,1}\sum_{\bar{l}_1=0,1}$$
$$\cdots\sum_{\bar{l}_{\bar{M}_2}=0,1} P_{jl}(\bar{t}_k)P_{li}(\Delta s - \bar{t}_k)\lambda_\beta(s+\bar{t}_k|I_{N_s^D}, T_{N_s^D}, x_l)$$
$$\times\Psi_{jl}\left(-(\eta_{C(N_s^Y)}(x_0), \eta_{C(N_s^Y)}(x_1))^T, \bar{t}_k\right)$$
$$\times\Psi_{li}\left(-(\eta_{C(N_s^Y)}(x_0), \eta_{C(N_s^Y)}(x_1))^T, \Delta s - \bar{t}_k\right)$$
$$\times\Psi_{jl_1}\Bigg(-\sum_{i\in\bar{I}_{N_s^D}}(\lambda_i(s|I_{N_s^D}, T_{N_s^D}, x_0),$$
$$\times\lambda_i(s|I_{N_s^D}, T_{N_s^D}, x_1))^T, \Delta\bar{s}\Bigg)$$
$$\times\Psi_{l_1l_2}\Bigg(-\sum_{i\in\bar{I}_{N_s^D}}(\lambda_i(s+\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_0),$$

$$
\begin{aligned}
&\times \lambda_i(s+\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_1))^T, \Delta\bar{s}\Big)\\
&\times \cdots\\
&\times \Psi_{l_{\bar{M}_1}l}\Big(-\sum_{i\in\bar{I}_{N_s^D}}(\lambda_i(s+\bar{M}_1\cdot\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_0),\\
&\qquad \times \lambda_i(s+\bar{M}_1\cdot\Delta\bar{s}|I_{N_s^D}, T_{N_s^D}, x_1))^T,\\
&\qquad \times \bar{t}_k - s - \bar{M}_1\cdot\Delta\bar{s}\Big)\\
&\times \Psi_{l\bar{l}_1}\Big(-\sum_{i\in\bar{I}^*_{N_s^D}}(\lambda_i(\bar{t}_k|I^*_{N_s^D}, T^*_{N_s^D}, x_0),\\
&\qquad \times \lambda_i(\bar{t}_k|I^*_{N_s^D}, T^*_{N_s^D}, x_1))^T, \Delta\bar{s}\Big)\\
&\times \Psi_{\bar{l}_1\bar{l}_2}\Big(-\sum_{i\in\bar{I}^*_{N_s^D}}(\lambda_i(\bar{t}_k+\Delta\bar{s}|I^*_{N_s^D}, T^*_{N_s^D}, x_0),\\
&\qquad \times \lambda_i(\bar{t}_k+\Delta\bar{s}|I^*_{N_s^D}, T^*_{N_s^D}, x_1))^T, \Delta\bar{s}\Big)\\
&\times \cdots\\
&\times \Psi_{\bar{l}_{\bar{M}_2}i}\Big(-\sum_{i\in\bar{I}^*_{N_s^D}}(\lambda_i(\bar{s}_k+\bar{M}_2\cdot\Delta\bar{s}|I^*_{N_s^D}, T^*_{N_s^D}, x_0),\\
&\qquad \times \lambda_i(\bar{s}_k+\bar{M}_2\cdot\Delta\bar{s}|I^*_{N_s^D}, T^*_{N_s^D}, x_1))^T,\\
&\qquad \times \Delta s - \bar{t}_k - \bar{M}_2\Delta\bar{s}\Big)
\end{aligned}
$$

where $\bar{M}_1 = \left[\frac{\bar{t}_k}{\Delta\bar{s}}\right]$ and $\bar{M}_2 = \left[\frac{\Delta s - \bar{t}_k}{\Delta\bar{s}}\right]$.

$$
\begin{aligned}
&P(X_t = x_j, \text{no jump or default in}[0,t])\\
&\quad = \sum_{l_1=0,1}\sum_{l_2=0,1}\cdots\sum_{l_M=0,1} P(X_t = x_j)\\
&\qquad \times \Psi_{0j}\left(-(\eta_{C(0)}(x_0), \eta_{C(0)}(x_1))^T, t\right)\\
&\qquad \times \Psi_{0l_1}(-\sum_{i\in I}(\lambda_i(0|I_{N_0^D}, T_{N_0^D}, x_0),\\
&\qquad\quad \times \lambda_i(0|I_{N_0^D}, T_{N_0^D}, x_1))^T, \Delta\bar{s})\\
&\qquad \times \Psi_{l_1l_2}(-\sum_{i\in I}(\lambda_i(\Delta\bar{s}|I_{N_0^D}, T_{N_0^D}, x_0),\\
&\qquad\quad \lambda_i(\Delta\bar{s}|I_{N_0^D}, T_{N_0^D}, x_1))^T, \Delta\bar{s})\\
&\quad \times\cdots\\
&\qquad \times \Psi_{l_Mj}(-\sum_{i\in I}(\lambda_i(M\cdot\Delta\bar{s}|I_{N_0^D}, T_{N_0^D}, x_0),\\
&\qquad\quad \lambda_i(M\cdot\Delta\bar{s}|I_{N_0^D}, T_{N_0^D}, x_1))^T, t - M\cdot\Delta\bar{s})
\end{aligned}
$$

where $M = \left[\frac{t}{\Delta\bar{s}}\right]$.

Now we know how to ensure $\Delta\Psi_{ij}(s_0, \mathbf{u}, \Delta\bar{s}) < \epsilon$, and have the formulas for calculating $f^{j,i}_{\bar{s}_k}(s+\bar{s}_k; s, \Delta s)$, $f^{j,i}_{\bar{t}_k}(s+\bar{t}_k; \beta, s, \Delta s)$ and $P(X_t = x_j, \text{no jump or default in}[0,t])$. We then discuss how to choose ϵ such that the relative error of them can be controlled as small as we wish, i.e. ζ. Taking $f^{j,i}_{\bar{s}_k}(s+\bar{s}_k; s, \Delta s)$ as an example in the following discussion, the results related to the others are similar.

PROPOSITION 6 *To ensure the relative error of $f^{j,i}_{\bar{s}_k}(s+\bar{s}_k; s, \Delta s)$, i.e.*

$$\frac{|\bar{f}^{j,i}_{\bar{s}_k}(s+\bar{s}_k; s, \Delta s) - f^{j,i}_{\bar{s}_k}(s+\bar{s}_k; s, \Delta s)|}{f^{j,i}_{\bar{s}_k}(s+\bar{s}_k; s, \Delta s)}$$

where $f^{j,i}_{\bar{s}_k}(s+\bar{s}_k; s, \Delta s)$ denotes the real value, $\bar{f}^{j,i}_{\bar{s}_k}(s+\bar{s}_k; s, \Delta s)$ denotes the value calculated according to the approximation formula, be less than any arbitrary real number ζ, we can require the error of $\bar{\Psi}_{ij}(s+\Delta s, \mathbf{u}, \Delta\bar{s})$, i.e. ϵ, where $\Delta\bar{s} = \frac{-\ln(1-\epsilon)}{K\cdot\lambda_{\max}(s+\Delta s)}$, to satisfy the following conditions:

$$
\begin{cases}
2^{\frac{\bar{s}_k\cdot K\cdot\lambda_{\max}(s+\Delta s)}{-\ln(1-\epsilon)}}\left[\left(1+\epsilon\cdot e^{\frac{-\ln(1-\epsilon)}{K}}\right)^{\frac{\bar{s}_k\cdot K\cdot\lambda_{\max}(s+\Delta s)}{-\ln(1-\epsilon)}+1} - 1\right] < \frac{\zeta}{2}\\
2^{\frac{(\Delta s-\bar{s}_k)\cdot K\cdot\lambda_{\max}(s+\Delta s)}{-\ln(1-\epsilon)}}\left[\left(1+\epsilon\cdot e^{\frac{-\ln(1-\epsilon)}{K}}\right)^{\frac{(\Delta s-\bar{s}_k)\cdot K\cdot\lambda_{\max}(s+\Delta s)}{-\ln(1-\epsilon)}+1} - 1\right] < \frac{\zeta}{2}.
\end{cases}
$$

Proof. Notice that when $s_1 < s_2$, the following relationship

$$\frac{-\ln(1-\epsilon)}{K\cdot\lambda_{\max}(s_2)} \le \frac{-\ln(1-\epsilon)}{K\cdot\lambda_{\max}(s_1)}$$

would always be valid. That is to say, when we choose the numerical time step size $\Delta\bar{s}$ to ensure the error of $\bar{\Psi}_{ij}(s+\Delta s, \mathbf{u}, \Delta\bar{s})$ be less than ϵ, this step size would also ensure the error of $\bar{\Psi}_{ij}(s_0, \mathbf{u}, \Delta\bar{s})$ where $s_0 \in [0, s+\Delta s]$ be less than ϵ as well. Because $P(X_t = x_j)$ and $\bar{\Psi}_{ij}(s_0, \mathbf{u}, \Delta\bar{s})$ are always less than 1, from the expressions for calculating $f^{j,i}_{\bar{s}_k}(s+\bar{s}_k; s, \Delta s)$ above, to make sure that the error be less than ζ, we have the following relationships

$$\sum_{l_1=0,1}\cdots\sum_{l_{M_1}=0,1} \times \left(\frac{(\bar{\Psi}_{jl_1}+\epsilon)\cdot(\bar{\Psi}_{l_1l_2}+\epsilon)\cdots(\bar{\Psi}_{l_{M_1}l}+\epsilon) - \bar{\Psi}_{jl_1}\cdot\bar{\Psi}_{l_1l_2}\cdots\bar{\Psi}_{l_{M_1}l}}{\bar{\Psi}_{jl_1}\cdot\bar{\Psi}_{l_1l_2}\cdots\bar{\Psi}_{l_{M_1}l}}\right) < \frac{\zeta}{2}$$

which implies

$$\sum_{l_1=0,1}\cdots\sum_{l_{M_1}=0,1} \times \left(\left(1+\frac{\epsilon}{\bar{\Psi}_{jl_1}}\right)\cdot\left(1+\frac{\epsilon}{\bar{\Psi}_{l_1l_2}}\right)\cdots\left(1+\frac{\epsilon}{\bar{\Psi}_{l_{M_1}l}}\right) - 1\right) < \frac{\zeta}{2}$$

and

$$\sum_{\bar{l}_1=0,1}\cdots\sum_{\bar{l}_{M_2}=0,1} \times \left(\frac{(\bar{\Psi}_{l\bar{l}_1}+\epsilon)\cdot(\bar{\Psi}_{\bar{l}_1\bar{l}_2}+\epsilon)\cdots(\bar{\Psi}_{\bar{l}_{M_2}i}+\epsilon) - \bar{\Psi}_{l\bar{l}_1}\cdot\bar{\Psi}_{\bar{l}_1\bar{l}_2}\cdots\bar{\Psi}_{\bar{l}_{M_2}i}}{\bar{\Psi}_{l\bar{l}_1}\cdot\bar{\Psi}_{\bar{l}_1\bar{l}_2}\cdots\bar{\Psi}_{\bar{l}_{M_2}i}}\right) < \frac{\zeta}{2}$$

which implies

$$\sum_{\bar{l}_1=0,1}\cdots\sum_{\bar{l}_{M_2}=0,1} \times \left(\left(1+\frac{\epsilon}{\bar{\Psi}_{l\bar{l}_1}}\right)\cdot\left(1+\frac{\epsilon}{\bar{\Psi}_{\bar{l}_1\bar{l}_2}}\right)\cdots\left(1+\frac{\epsilon}{\bar{\Psi}_{\bar{l}_{M_2}i}}\right) - 1\right) < \frac{\zeta}{2}$$

where $l = 0, 1$. Notice that $\bar{\Psi}_{ij}(s_0, \mathbf{u}, \Delta\bar{s})$, $s_0 \in [s, s+\Delta s]$ in the above would always be greater than $e^{-\lambda_{\max}(s+\Delta s)\cdot\Delta\bar{s}}$ which is equal to $e^{\frac{\ln(1-\epsilon)}{K}}$. Thus we can replace each $\bar{\Psi}_{ij}(s_0, \mathbf{u}, \Delta\bar{s})$ in the above with $e^{\frac{\ln(1-\epsilon)}{K}}$ to find ϵ according to ζ. Also, note that

$$M_1 \le \frac{\bar{s}_k\cdot K\cdot\lambda_{\max}(s+\Delta s)}{-\ln(1-\epsilon)} \quad \text{and}$$

$$M_2 \le \frac{(\Delta s-\bar{s}_k)\cdot K\cdot\lambda_{\max}(s+\Delta s)}{-\ln(1-\epsilon)},$$

then the above equations can be rewritten as follows:

$$\begin{cases} 2^{M_1}\left[\left(1+\frac{\epsilon}{e^{\frac{\ln(1-\epsilon)}{K}}}\right)^{M_1+1}-1\right] \\ \leq 2^{\frac{\bar{s}_k\cdot K\cdot\lambda_{\max}(s+\Delta s)}{-\ln(1-\epsilon)}}\left[\left(1+\frac{\epsilon}{e^{\frac{\ln(1-\epsilon)}{K}}}\right)^{\frac{\bar{s}_k\cdot K\cdot\lambda_{\max}(s+\Delta s)}{-\ln(1-\epsilon)}+1}-1\right]<\frac{\zeta}{2} \\ 2^{M_2}\left[\left(1+\frac{\epsilon}{e^{\frac{\ln(1-\epsilon)}{K}}}\right)^{M_2+1}-1\right]\leq 2^{\frac{(\Delta s-\bar{s}_k)\cdot K\cdot\lambda_{\max}(s+\Delta s)}{-\ln(1-\epsilon)}} \\ \left[\left(1+\frac{\epsilon}{e^{\frac{\ln(1-\epsilon)}{K}}}\right)^{\frac{(\Delta s-\bar{s}_k)\cdot K\cdot\lambda_{\max}(s+\Delta s)}{-\ln(1-\epsilon)}+1}-1\right]<\frac{\zeta}{2} \end{cases}$$

□

All the conditions related to the relative errors of $f_{s_k}^{j,i}(s+\bar{s}_k; s, \Delta s)$, $f_{t_k}^{j,i}(s+\bar{t}_k; \beta, s, \Delta s)$ and $P(X_t = x_j$, no jump or default in $[0, t])$, similar to the above proposition should be satisfied to find a suitable ϵ. Therefore, the relative errors are controlled and the error of $P(X_t = x_i|\mathcal{F}_t)$ can also be controlled. We remark that suppose the expiry time is denoted as T_{expiry}, then all $\lambda_{\max}(s_0)$, $s_0 \in [0, T_{expiry}]$ in propositions 5 and 6 could simply be replaced by $\lambda_{\max} = \lambda_{\max}(T_{\text{expiry}})$.

6. Numerical experiments

In this section, we present numerical experiments to demonstrate our proposed methods. For the configuration of the parameters value in the hidden Markov chain X_t, we let the transition rates be $\theta_0 = 0.1$ and $\theta_1 = 0.1$, the initial state $x_0 = 0$. For the observable chain Y_t, we set the transition rates

$$\eta_0(x) = \begin{cases} 0.1, & x = x_0 \\ 0.2, & x = x_1 \end{cases}$$

and

$$\eta_1(x) = \begin{cases} 0.2, & x = x_0 \\ 0.1, & x = x_1. \end{cases}$$

and the initial state is $y_0 = 0$ as we assumed. The risk-free interest rate r is assumed to be 5%.

6.1. Numerical example 1

We consider the pricing of Credit Default Swaps (CDS). Assume that the buyer of the CDS agrees to pay premiums to the seller continuously over time at a fixed rate until the expiration time of the CDS contract. If the reference asset defaults prior to the expiry, then the seller will pay $1 to the buyer. Denote the seller as entity A, buyer as entity B and the reference asset of the CDS as entity C. Denote τ^A, τ^B, τ^C the default times and $\lambda_A, \lambda_B, \lambda_C$ the default intensities of entities A, B and C, respectively. Here the default intensities of these homogeneous three entities are assumed in the following form:

$$\lambda_i(t) = a + b^n(X(t))\cdot X(t) + c\cdot\left(\sum_{j\neq i} 1_{\{\tau^j\leq t\}}\right), \quad i, j = A, B, C, n \in \mathbb{N}^+$$

where a and c are constants, and $b^n(X(t))$ is assumed to take the following form:

$$b^n(X(t)) = \begin{cases} b_0, & X(t) = x_0 \\ b_1, & X(t) = x_1 \\ \vdots, & \vdots \\ b_n, & X(t) = x_n \end{cases},$$

for some distinct real constants $b_0, b_1, \ldots, b_n$.

Here $X(t)$ is the hidden state process and $\sum_{j\neq i} 1_{\{\tau^j\leq t\}}$ are the observable processes. Recall the assumption made in section 2 that $(X_t)_{t\geq 0}$ is an 'exogenous' process to $(N_t^i)_{t\geq 0}$, $i = 1, 2, \ldots, K$. Consequently, the two stochastic variables $X(t)$ and $\sum_{j\neq i} 1_{\{\tau^j\leq t\}}$ in the expression of default intensities are independent. For the coefficients in the default intensities, one may interpret them as the contributions of the factors to the total default correlation. For example, the parameter a may be interpreted as the impact size of a deterministic factor to the default intensities. If $a = 0$, there is no deterministic factor contributed to the default intensities; while $b_0, b_1, \ldots, b_n$ may be interpreted as different impact sizes for the default intensities under different economic states; and c is the impact size of every default process to the default intensities. If $c = 0$, defaults of entities are independent; otherwise, the default of one entity will increase the default intensities of other entities by size c. As we assumed before for simplicity, both $X(t)$ and $Y(t)$ are two-state Markov chains, and $X(t) \in \{0, 1\}$ and $Y(t) \in \{0, 1\}$, respectively. We further denote $b = b_1$, so the default intensities could be written as the following form:

$$\lambda_i(t) = a + b\cdot X(t) + c\cdot\left(\sum_{j\neq i} 1_{\{\tau^j\leq t\}}\right), \quad i, j = A, B, C.$$

Let y be the fixed premium rate, and suppose the issue time of the swap contract is 0, the expiry time is T, and we are at time s, then the present value of the premium payment from the buyer should be

$$E\left[\int_0^T e^{-rs} y 1_{\{s<\tau^A, s<\tau^B, s<\tau^C\}}\mathrm{d}s\right].$$

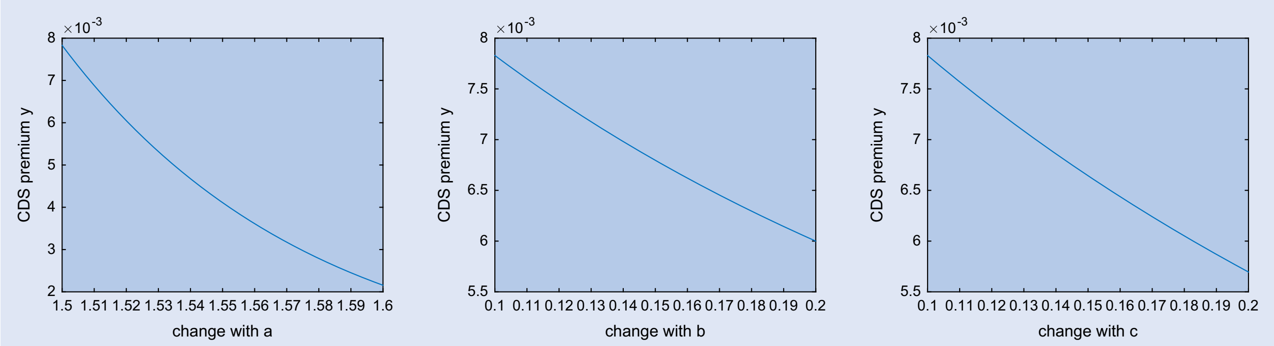
This means if any one of the three entities defaults, the buyer of the CDS contract would stop paying the premium. Similarly, the present the value of the seller should be

$$E\left[e^{-rT} 1_{\{T<\tau^A, T<\tau^B, \tau^C\leq T\}}\mathrm{d}s\right].$$

According to these two expressions, one can obtain the premium of the CDS in the following form:

$$y = \frac{E\left(e^{-rT} 1_{\{T<\tau^A, T<\tau^B, \tau^C\leq T\}}\mathrm{d}s\right)}{E\left(\int_0^T e^{-rs} 1_{\{s<\tau^A, s<\tau^B, s<\tau^C\}}\mathrm{d}s\right)}.$$

From the above formula, we know that to calculate y, we need to compute the joint density function $f(s < \tau^A, s < \tau^B, s < \tau^C)$ and the joint probability $P(T < \tau^A, T < \tau^B, \tau^C \leq T)$. Notice that $f(s < \tau^A, s < \tau^B, s < \tau^C)$ is actually equal to $f(\tau^1 > s)$ where τ^1 denotes the time of the first default out of the 3 entities, and $P(T < \tau^A, T < \tau^B, \tau^C \leq T) = P(\tau^1 \leq T < \tau^2)$. Here τ^1 has the same meaning as before, τ^2 denotes the time of the second default in the reference portfolio. Then we can apply the methods introduced in the previous sections to calculate the fixed premium rate y. The base setting



(a) premium y change with respect to a (b) premium y change with respect to b (c) premium y change with respect to c

Figure 1. Change of premium y with coefficients.

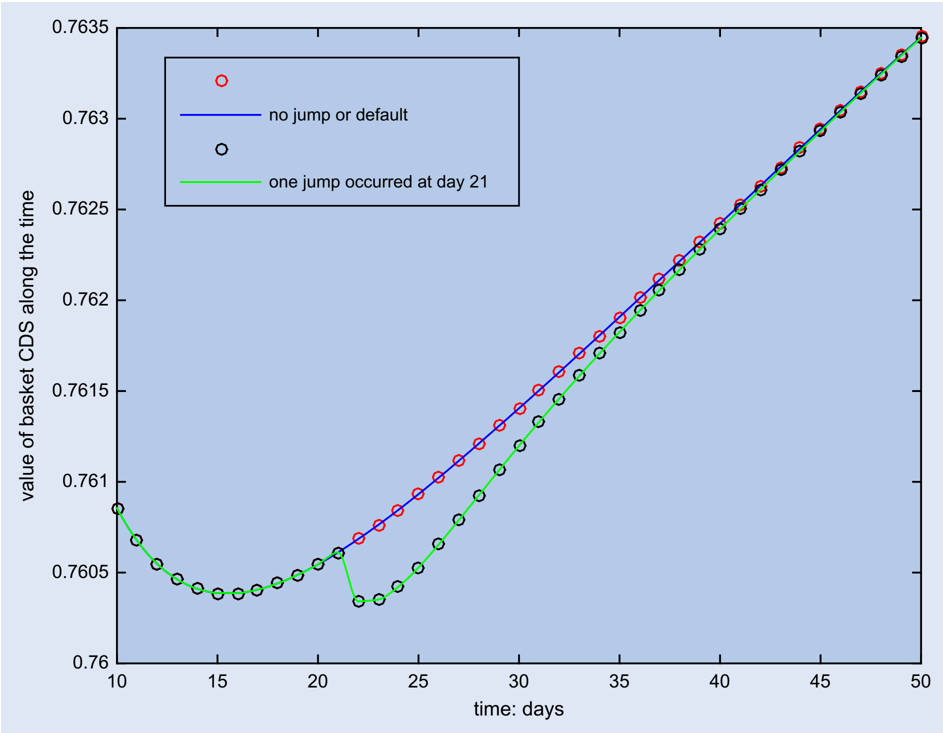


Figure 2. Change of CDS's value from day to day with the first default intensities.

of parameters are given in the following. For the contagion factors, we let $a = 1, b = 0.1, c = 0.1$. The expiry T is 5 years, and the initial time is 0. We change the coefficients a, b and c in the expressions of default intensities separately, and each time we keep the remaining coefficients unchanged to investigate the change in the CDS premium rate y.

From Figure 1(a)–(c), we find that the value of CDS premium rate y decreases as the coefficients a, b, and c increase.

6.2. *Numerical example 2*

We then consider a kth-to-default basket CDS contact. Assume that our portfolio contains $K = 10$ homogeneous entities, if k entities out of this portfolio default prior to the expiry time, then \$1 will be paid. For simplicity, this payment only occurs at the expiry time, but the payment of premium occurs at the initial time. Similar to the previous experiment, the entity i's default intensity is given by

$$\lambda_i(t) = a + b \cdot X(t) + c \cdot \left(\sum_{j \neq i} 1_{\{\tau^j \leq t\}} \right), \quad i, j = 1, 2, \ldots, K.$$

The value of this kth-to-default basket CDS at time t can be written as

$$V_k(t) = \exp\{-r(T - t)\} P(\tau^k \leq T \mid \mathcal{F}_t)$$

where τ^k denotes the kth-to-default time. For the state of chain X, x_0 and x_1 represent the 'good (bull market)' and 'bad (bear market)' economic state, respectively. While States y_0 and y_1 of chain Y represent the delayed information of 'bad' economic state and 'good' economic state, respectively. Here we also

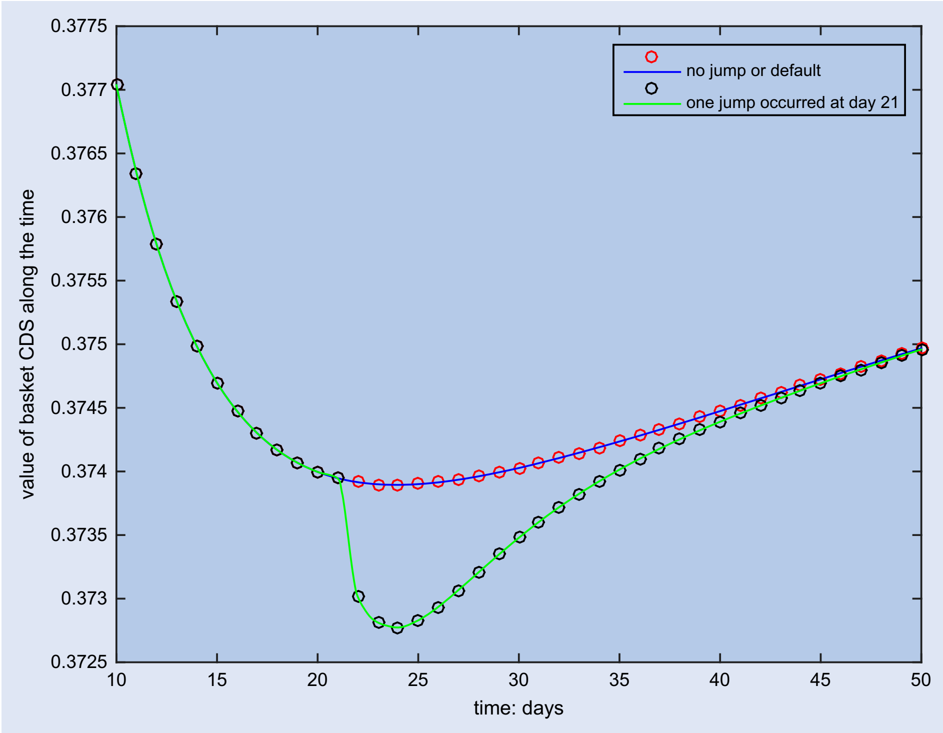


Figure 3. Change of CDS's value from day to day with the second default intensities.

assume that the total number of entities in the portfolio is $K = 10$. The calculation of $P(\tau^k \leq T \mid \mathcal{F}_s)$ can be obtained from $1 - P(\tau^{j_k} > T \mid \mathcal{F}_s)$ where

$$P(\tau^{j_k} > t \mid \mathcal{F}_s) = \sum_{i=m}^{k-1} P(\tau^{j_i} \leq t < \tau^{j_{i+1}} \mid \mathcal{F}_s).$$

The calculation of the probability $P(\tau^{j_i} \leq t < \tau^{j_{i+1}} \mid \mathcal{F}_s)$ is similar to the calculation of $P(\tau^1 \leq t < \tau^2)$ in Experiment 1. Without loss of generality, for simplicity, we consider the 1st-to-default basket CDS as $k = 1$. We further assume that the initial time is 0, and that we are at time $t = 10$ days now, and that the expiry time is $T = 100$ days. In the following experiments, we consider two scenarios. In Scenario 1, there is no jump in chain Y and default observed by the expiry time. In Scenario 2, there is one jump in chain Y between day 21 and day 22 but no default observed by expiration. According to the assumptions presented in section 2, we know that the initial state of chain X is $x_0 = 0$ and the initial state of chain Y is $y_0 = 0$. In addition, let the coefficients in default intensities be $a = 0.001$, $b = 0.001$, and $c = 0.001$. Then one can see the change of basket CDS values from day to day, and here we only provide the values from day $t = 10$ to day $t = 50$ as an example.

From Figure 2 we can see that as time goes by, the general tendency of basket CDS's value is increasing. When there is one jump in chain Y from state y_0 to state y_1, the value will drop suddenly. It is because at the beginning, the information of chain Y reflected a 'bad' economic condition, when it changed to state y_1 which representing a 'good' economic state, intuitively, the probability of defaults will drop suddenly, and the value of basket CDS will therefore drop suddenly as well.

As we mentioned before, our model and methods can be applicable to various forms of default intensities. Therefore, we further consider another form of default intensities which decay exponentially with time. The expression is as follows:

$$\lambda_i(t) = \left(a + c \cdot \sum_{j \neq i} 1_{\{\tau^j \leq t\}} \right) e^{-t} + b^n(X(t)) \cdot X(t),$$
$$i, j = 1, 2, \ldots, K, n \in \mathbb{N}^+.$$

Similar to Experiment 1, denote $b = b_1$, under our assumption, we have

$$\lambda_i(t) = \left(a + c \cdot \sum_{j \neq i} 1_{\{\tau^j \leq t\}} \right) e^{-t} + b \cdot X(t),$$
$$i, j = 1, 2, \ldots, K.$$

Two stochastic variables $X(t)$ and $\sum_{j \neq i} 1_{\{\tau^j \leq t\}}$ are independent under the assumption made in section 2. In the numerical experiments, all parameters in this default intensity keep the same as the previous case, then we can also calculate the value of basket CDS and observe it from day to day. Note that the impact size of a deterministic factor to the default intensities is $a \cdot e^{-t}$ which declines exponentially with time t, and the impact size of default processes also declines exponentially with time which becomes $c \cdot e^{-t}$.

From Figure 3, we notice that the overall value based on this form of default intensity is smaller than the previous one. The phenomenon can be explained as follows. As the default intensity exponentially decreased with time, the default probability will become smaller accordingly and therefore the value of basket CDS. For the same reason and similar explanations like before, the value will also jump down suddenly when one jump in observable chain Y from state y_0 to y_1 occurred.

7. Concluding remarks

In this paper we present a reduced-form intensity-based credit risk model with a hidden Markov process modeling the evolution of economic conditions over time. We also discuss a method to extract the underlying hidden state process from observable processes: the default processes and the stochastic process which reflects the delayed and noisy information about the hidden state process. The method may have a wide range of applications. Based on this, we develop a closed-form expression to obtain the joint default distribution with the hidden state process. After deriving this general formula, for the homogeneous contagion portfolio, we also give analytical formulas for the distribution of ordered default times. Further, we extend the total hazard construction method to get the joint distribution of default times for HMMs. We remark that the methods discussed may be applicable to various forms of default intensities. Algorithms for practical implementation of the methods are presented and their uses for pricing credit derivatives are illustrated. In the numerical experiments, we consider valuations for the CDSs premium rates of the regular and basket type with different expressions of default intensities which cover an exponential decay and a stochastic intensity process. We also study the sensitivities of premium rates with respect to changes in the underlying parameters in the regular CDS as an example.

Acknowledgements

The authors would like to thank the two anonymous referees and the editor for their helpful comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research work was supported by Research Grants Council of Hong Kong [grant number 17301214]; HKU CERG Grants and HKU Strategic Research Theme in Information and Computing.

References

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

Brigo, D., Pallavicini, A. and Torresetti, R., Calibration of CDO tranches with the dynamical generalized-Poisson loss model. Working Paper, Banca IMI, 2006. Available online at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=900549

Cont, R. and Minca, A., Recovering portfolio default intensities implied by CDO quotes. *Math. Finance.*, 2013, **23**(1), 94–121, doi:10.1111/j.1467-9965.2011.00491.x.

Davis, M. and Lo, V., Modeling default correlation in bond portfolios. In *Mastering Risk Volume 2: Applications*, edited by C. Alexander, pp. 141–151, 2001 (Prentice Hall: Upper Saddle River, NJ).

Duffie, D. and Garleanu, N., Risk and valuation of collateralized debt obligations. *Financ. Anal. J.*, 2001, **57**(1), 41–59.

Duffie, D., Saita, L. and Wang, K., Multi-period corporate default prediction with stochastic covariates. *J. Financ. Econ.*, 2006, **83**(3), 635–665.

Elliott, R.J., Aggoun, L. and Moore, J.B., *Hidden Markov Models: Estimation and Control*, Vol. 29, 2008 (Springer-Verlag New York: New York).

Elliott, R.J., Jeanblanc, M. and Yor, M., On models of default risk. *Math. Finance*, 2000, **10**(2), 179–195.

Elliott, R.J. and Shen, J., Credit risk and contagion via self-exciting default intensity. *Ann. Finance*, 2015a, **11**(3), 319–344.

Elliott, R.J. and Shen, J., Dynamic optimal capital structure with regime switching. *Ann. Finance*, 2015b, **11**(2), 199–220.

Elliott, R.J. and Shen, J., Credit risk with latent contagion and frailty: Default probabilities, pricing, and hedging. *Rev. Derivatives Res.* 2016, forthcoming.

Elliott, R.J., Siu, T.K. and Fung, E.S., A double HMM approach to altman Z-scores and credit ratings. *Expert Syst. Appl.*, 2014, **41**(4–2), 1553–1560.

Elliott, R.J. and Siu, T.K., An HMM intensity-based credit risk model and Filtering. In *State-Space Models: Applications Economics and Finance*, edited by Y. Zeng and S. Wu, Vol. 1, pp. 169–184, 2013 (Springer New York: New York).

Frey, R. and Runggaldier, W., Pricing credit derivatives under incomplete information: A nonlinear-filtering approach. *Finance Stoch.*, 2010, **14**(4), 495–526.

Frey, R. and Runggaldier, W., Nonlinear filtering in models for interest rate and credit risk, Chapter 32. In *Handbook of Nonlinear Filtering*, edited by D. Crisan and B. Rozovski, 2011 (Oxford University Press: Oxford).

Frey, R. and Schmidt, T., Filtering and incomplete information in credit risk, Chapter 7. In *Credit Risk Frontiers: Subprime Crisis, Pricing and Hedging, CVA, MBS, Ratings, and Liquidity*, edited by D. Brigo, T. Bielecki and F. Patras, 2011 (Wiley: Hoboken, NJ).

Giesecke, K., Portfolio credit risk: Top-down vs. bottom-up approaches. In *Frontiers in Quantitative Finance: Credit Risk and Volatility Modeling*, edited by R. Cont, 2008 (Wiley: Hoboken, NJ).

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

Giesecke, K., Goldberg, L. and Ding, X., A top down approach to multi-name credit. *Oper. Res.*, 2011, **59**(2), 283–300.

Gu, J., Ching, W., Siu, T. and Zheng, H., On pricing basket credit default swaps. *Quant. Finance*, 2013, **13**(12), 1845–1854.

Gu J., Ching W. and Zheng H., A hidden Markov reduced-form risk model, 2014 IEEE Conference Computational Intelligence for Financial Engineering & Economics, London, pp. 190–196, 2014.

Jarrow, R. and Turnbull, S., Pricing derivatives on financial securities subject to credit risk. *J. Finance*, 1995, **50**, 53–86.

Jarrow, R. and Yu, F., Counterparty risk and the pricing of defaultable securities. *J. Finance*, 2001, **56**(5), 555–576.

Lando, D., On Cox processes and credit risky securities. *Rev. Deriv. Res.*, 1998, **2**, 99–120.

Longstaff, F. and Rajan, A., An empirical analysis of collateralized debt obligations. *J. Finance*, 2008, **63**(2), 529–563.

Madan, D. and Unal, H., Pricing the risks of default. *Rev. Deriv. Res.*, 1998, **2**(2–3), 121–160.

Merton, R., On the pricing of corporate debt: The risk structure of interest rates. *J. Finance*, 1974, **29**(2), 449–470.

Norros, I., A compensator representation of multivariate life length distributions, with applications. *Scand. J. Stat.*, 1986, **13**, 99–112.

Schönbucher, P. and Schubert, D., Copula-dependent default risk in intensity models. Working Paper, Universität Bonn, 2001.

Shaked, M. and Shanthikumar, G., The multivariate hazard construction. *Stoch. Proc. Appl.*, 1987, **24**, 241–258.

Yu, F., Correlated defaults in intensity-based models. *Math. Finance*, 2007, **17**(2), 155–173.

Zheng, H. and Jiang, L., Basket CDS pricing with interacting intensities. *Finance Stoch.*, 2009, **13**, 445–469.