Law of Large Numbers for Self-Exciting Correlated Defaults

Jakša Cvitanić [*], Jin Ma [†], Jianfeng Zhang [‡]

Abstract. We consider a model of correlated defaults in which the default times of multiple entities depend not only on a common and specific factors, but also on the extent of past defaults in the market, via the average loss process, including the average number of defaults as a special case. The paper characterizes the average loss process when the number of entities becomes large, showing that under some monotonicity conditions the limiting average loss process can be determined by a fixed point problem. We also show that the Law of Large Numbers holds under certain compatibility conditions.

Keywords: Self-exciting, correlated defaults, credit risk, law of large numbers.
2010 Mathematics Subject Classification: 91G40, 60H10, 91G80

[*]Corresponding author: Caltech, 1200 E. California Blvd., M/C 228-77, Pasadena, CA 91125. Ph: (626) 395-1784. E-mail: cvitanic@hss.caltech.edu. Research supported in part by NSF grant DMS 06-31298 and 10-08873

[†]USC Department of Mathematics, 3620 S Vermont Ave, KAP 108, Los Angeles, CA 90089-1113. Ph: (213) 740-. E-mail: jinma@usc.edu. Research supported in part by NSF grant DMS #1816127.

[‡]USC Department of Mathematics, 3620 S Vermont Ave, KAP 108, Los Angeles, CA 90089-1113. Ph: (213) 740-9805. E-mail: jianfenz@usc.edu. Research supported in part by NSF grant DMS 10-08873.

1 Introduction

Modeling of correlation between default probabilities of multiple "names" (individuals, firms, countries, etc.) has been one of the central issues in the theory and applications of managing and pricing credit risk in the last several years. There have been dozens of models in the literature. While each of these models has its own advantages and disadvantages, lax use of such models in practice could in part affect the understanding of the risk of the credit default and consequently contribute to the extent of a potential crisis in the market.

In this paper we propose a "bottom-up" model for correlated defaults within the standard "reduced form" framework. In particular, we assume that in a large collection of defaultable entities, the intensity of each individual default depends on factors specific to the individual entity, and on a common factor. The main novelty of our model is that we further allow a part of the common factor to have a *self-exciting* nature, reflecting the general "health" of the market. More precisely, we assume that the self-exciting factor takes the form of an "average loss process", including the average number of defaults to-date as a special case. The self-exciting feature allows us, in the limiting case, to analyze the impact of such a "general health" index on the individual entities. However, it also generates a circular feedback phenomenon that is technically non-trivial.

The self-exciting structure of our model can be thought of as an example of the so-called "contagion" feature, which has been investigated by many authors using various approaches. These include Jarrow and Yu (2001), Davis and Lo (2001), Collin-Dufresne, Goldstein and Helwege (2003), Collin-Dufresne, Goldstein and Hugonnier (2004), Dembo, Deuschel and Duffie (2004), Giesecke and Goldberg (2004), Giesecke and Weber (2005) and (2006), Frey and Backhaus (2006) and (2007), Horst (2007), Yu (2007), and Dai Pra, Runggaldier, Sartori and Tolotti (2009). None of these models contains the circular nature presented in our model. In a recent work, Giesecke, Spiliopoulos and Sowers (2010) consider a model similar to ours. However, they impose a more special structure, which enables them to obtain large deviation type results, in addition to the Law of Large Numbers type results that we focus on. Self-exciting feature is also present in Filipovic, Overback and Schmidt (2011), in a "top-down" model. For an overview of standard default risk models, one can consult, among many others, the texts Duffie and Singleton (2003), Lando (2004), and Frey, Mc Neil and Embrecths (2005), and the references cited therein.

Assuming that all the factors are diffusion processes, we first show that the proposed self-exciting model is well-posed. Our next main objective is to identify conditions under which the average number of defaults (or more generally the average default loss), has a limit, in the sense of the Law of Large Numbers, as the number of names tends to infinity. Under such conditions, we show that for the average number of defaults the limiting process solves an ordinary differential equation, while for the average loss the limiting process solves a more general and complex equation. It is worth remarking that these results, being of asymptotic nature, are not directly applicable to individual credit risk derivatives, because they require a large number of names to be involved in the limiting process. However, our results should be useful for the risk management at a level of an institution, or a country, with large portfolio of defaultable claims, when the aim is to analyze potential total losses. For example, it has been stated that the next crisis might come from potentially numerous defaults of credit card holders. This paper provides a theoretical model which may prove useful for addressing such issues.

The rest of the paper is organized as follows. In Section 2 we formulate the problem and the model. In Section 3 we show that the self-exciting model that we are proposing is well-posed. In Section 4 we study the fixed point problem that determines the limiting process. In Section 5 we present some potential applications where the fixed point problem could be solved. Finally, Sections 6 and 7 are devoted to the main theorem involving the Law of Large Numbers and its proof.

2 Problem Formulation

2.1 Average loss in correlated default models

We consider n "names", which could be individual investors, financial firms, loans, etc. We denote their default times by $\tau_1, \cdots, \tau_n$. Let us associate to each default time τ_i a "loss process" L^i_t, $t \geq 0$, so that the loss due to default at any time t is given by $L^i_{\tau^i}\mathbf{1}_{\{\tau^i \leq t\}}$. We define the "average loss" of all defaults at time t by

$$\bar{L}_t \stackrel{\triangle}{=} \bar{L}^n_t \stackrel{\triangle}{=} \frac{1}{n}\sum_{i=1}^{n} L^i_{\tau_i} 1_{\{\tau_i \leq t\}}. \tag{2.1}$$

Clearly, one can have various interpretations for $\bar{L}$ by imposing various choices for L^i. For example, if we set $L^i \equiv 1$, then $\bar{L}$ is the average number of defaults (for example, the average number of foreclosures in a given region).

Our main purpose is to investigate the limiting behavior of $\bar{L}^n$ as $n \to \infty$, namely,

$$\bar{L}^*_t \stackrel{\triangle}{=} \lim_{n\to\infty} \bar{L}^n_t, \tag{2.2}$$

whenever the limit exists, and to characterize the limit $\bar{L}^*$. It is to be expected that $\bar{L}^*$ will depend substantially on the correlation of the default times and the loss processes. The following two examples are the extreme cases, whose limits are quite different in nature:

1° Assume that the sequence $\{(\tau_n, L^1_{\tau_n})\}_{n\geq 1}$ is i.i.d. Then, the Law of Large Numbers (LLN) implies that $\bar{L}^n_t \to \bar{L}^*_t = \mathbb{E}\{L^1_{\tau_1}\mathbf{1}_{\{\tau_1\leq t\}}\}$, $\mathbb{P}$-a.s.

2° Assume that the times and the losses are fully correlated, that is, $\tau_1 = \cdots = \tau_n = \tau$, $L^1_{\tau_1} = \cdots = L^n_{\tau_n} = L_\tau$. Then, $\bar{L}^n_t = \bar{L}^*_t = L_\tau \mathbf{1}_{\{\tau\leq t\}}$.

In this paper, we will provide quite a general model such that the default times $\tau_1, \cdots, \tau_n$ are correlated and the limit $\bar{L}^*$ exists. The main,"self-exciting" feature of the model is that the correlation of $\tau_1, \cdots, \tau_n$ is built via the average loss $\bar{L}^n$.

2.2 The model

Throughout this paper we fix an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$, endowed with a filtration $\mathbb{F} \stackrel{\triangle}{=} \{\mathcal{F}_t\}_{t\geq 0}$. We assume that the probability space is rich enough to support a sequence of independent standard Brownian motions $(B^0, B^1, \cdots, B^n, \cdots)$ and a sequence of exponential random variables $(E^1, \cdots, E^n, \cdots)$, all with rate 1 and are independent of the Brownian motions. We define the following sub-filtrations of $\mathbb{F}$:

$$\mathbb{F}^0 \stackrel{\triangle}{=} \mathbb{F}^{B^0}, \quad \mathbb{F}^i \stackrel{\triangle}{=} \mathbb{F}^{B^0, B^i}, \quad i = 1, 2, \cdots, \tag{2.3}$$

the filtrations generated by the Brownian motions B^0 and (B^0, B^i), respectively, and augmented by the $\mathbb{P}$-null sets. For simplicity, let us assume that $\mathbb{F} = \bigvee_{i=1}^\infty \Big(\mathbb{F}^i \vee \sigma(E^i)\Big)$.

We now fix n and the loss processes L^i, $i = 1, \cdots, n$. As in reduced form models, see, e.g., Bielecki and Rutkowski (2002), Duffie and Singleton (2003), Jeanblanc, Yor and Chesney (2009), we define

$$\tau_i \stackrel{\triangle}{=} \inf\Big\{t \geq 0 : Y^i_t \geq E^i\Big\}, \tag{2.4}$$

where, for process $\bar{L}$ defined by (2.1), process Y^i denotes the "hazard process"

$$Y^i_t \stackrel{\triangle}{=} \int_0^t \lambda_i(s, B^0_{\cdot\wedge s}, B^i_{\cdot\wedge s}, X^0_s, X^i_s, \bar{L}_s)ds, \tag{2.5}$$

and X^0, X^i, $i = 1, 2, \cdots$ are factor processes defined by

$$\begin{aligned} X^0_t &= x_0 + \int_0^t b_0(s, B^0_{\cdot\wedge s}, X^0_s, \bar{L}_s)ds + \int_0^t \sigma_0(s, B^0_{\cdot\wedge s}, X^0_s, \bar{L}_s)dB^0_s, \\ X^i_t &= x_i + \int_0^t b_i(s, B^0_{\cdot\wedge s}, B^i_{\cdot\wedge s}, X^0_s, X^i_s, \bar{L}_s)ds + \int_0^t \sigma_i(s, B^0_{\cdot\wedge s}, B^i_{\cdot\wedge s}, X^0_s, X^i_s, \bar{L}_s)dB^i_s. \end{aligned} \tag{2.6}$$

Throughout the paper, we assume the following *Standing Assumptions*:

Assumption 2.1 *For each i, the process L^i is $\mathbb{F}^i$-adapted; the coefficients $b_0, \sigma_0 : \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}) \times \mathbb{R} \times \mathbb{R}_+ \mapsto \mathbb{R}$ and $b_i, \sigma_i, \lambda_i : \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R})^2 \times \mathbb{R}^2 \times \mathbb{R}_+ \mapsto \mathbb{R}$ are Lebesgue measurable functions; and $\lambda_i \geq 0$.*

We note that here X^0 denotes the common factor in the market, that is observable by everyone; X^i is the firm i's specific factor, observable only by firm i. It is possible that each individual firm has risk factors that are observable by others in the market, and we include such factors into the common factor X^0. It is clear that each τ_i is an $\mathbb{F}$-stopping time, but not necessarily an $\mathbb{F}^i$-stopping time. As pointed above, the main feature of our model is that the correlation among the defaults depends on, in addition to the common exogenous factor X^0, the past defaults through the process $\bar{L}$, so that it has a self-exciting nature. Moreover, since we model each τ_i rather than $\bar{L}$ directly, our model is "bottom-up".

When there is no confusion, for $\psi = b, \sigma, \lambda$ and $i = 1, 2, \cdots$, with a slight abuse of notation we denote

$$\begin{aligned} &\psi_0(t, x_0, \alpha) := \psi_0(t, \omega, x_0, \alpha) := \psi_0(t, B^0_{\cdot\wedge t}(\omega), x_0, \alpha), \\ &\psi_i(t, x_0, x_i, \alpha) := \psi_i(t, \omega, x_0, x_i, \alpha) := \psi_i(t, B^0_{\cdot\wedge t}(\omega), B^i_{\cdot\wedge t}(\omega), x_0, x_i, \alpha). \end{aligned} \tag{2.7}$$

Then clearly $\psi_0(\cdot, x_0, \alpha)$ is $\mathbb{F}^0$-adapted and $\psi_i(\cdot, x_0, x_i, \alpha)$ is $\mathbb{F}^i$-adapted.

Remark 2.2 (i) If $b_0, \sigma_0, b_i, \sigma_i, \lambda_i$ do not depend on $\bar{L}$, then our model becomes a standard reduced form model where the defaults are conditionally independent, conditional on the common factor X^0, and it is straightforward to check that in this case λ_i is the $\mathbb{F}^i$-intensity of τ_i, in the sense that $\mathbb{P}\{\tau_i > t | \mathcal{F}^i_t\} = \exp\{-\int_0^t \lambda_i(s, X^0_s, X^i_s) ds\}$, $t \geq 0$; see, e.g., Bielecki and Rutkowski (2002), Duffie and Singleton (2003).

(ii) In the general case when λ_i depends on $\bar{L}$, λ_i is obviously no longer an $\mathbb{F}^i$-adapted process (hence cannot be an "$\mathbb{F}^i$-intensity" of τ_i in the aforementioned sense). Due to the self-exciting nature of our model, λ^i can be interpreted as the conditional intensity of τ^i, conditional on all the past defaults. See Proposition 3.3 for a more precise statement; see also Jeanblanc and Song (2011a,b) for more on construction of default times with given intensities. ■

2.3 The main results

Notice that the system (2.1), (2.4)–(2.6) is "circular", and thus its well-posedness is by no means obvious. Our first result, Theorem 3.2 below, is that this system is indeed well-posed.

We next characterize the limit process $\bar{L}^*$ via a fixed point problem. We first conjecture that, if exists, $\bar{L}^*$ should be $\mathbb{F}^0$-adapted. Now, for an $\mathbb{F}^0$-adapted process α, by replacing $\bar{L}$

with α in the system (2.1), (2.4)–(2.6) we define

$$
\begin{aligned}
X_t^{0,\alpha} &= x_0 + \int_0^t b_0(s, X_s^{0,\alpha}, \alpha_s)ds + \int_0^t \sigma_0(s, X_s^{0,\alpha}, \alpha_s)dB_s^0; \\
X_t^{i,\alpha} &= x_i + \int_0^t b_i(s, X_s^{0,\alpha}, X_s^{i,\alpha}, \alpha_s)ds + \int_0^t \sigma_i(s, X_s^{0,\alpha}, X_s^{i,\alpha}, \alpha_s)dB_s^i, \\
Y_t^{i,\alpha} &\triangleq \int_0^t \lambda_i(s, X_s^{0,\alpha}, X_s^{i,\alpha}, \alpha_s)ds, \\
\tau_i^\alpha &\triangleq \inf\Big\{t \geq 0 : Y_t^{i,\alpha} \geq E_i\Big\}, \qquad i = 1, \cdots, n; \\
\bar{L}_t^\alpha &\triangleq \bar{L}^{n,\alpha} \triangleq \frac{1}{n}\sum_{i=1}^n L_{\tau_i^\alpha}^i 1_{\{\tau_i^\alpha \leq t\}}.
\end{aligned}
\tag{2.8}
$$

Clearly, given the information $\mathbb{F}^0$, processes $(X^{i,\alpha}, Y^{i,\alpha}, \tau_i^\alpha), i = 1, \cdots, n$, are conditionally independent; see Remark 2.2. Thus, under conditional probability $\mathbb{P}\{\ \cdot\ |\mathbb{F}^0\}$, the standard Law of Large Numbers should imply, modulo some technical conditions, that

$$\bar{L}_t^{n,\alpha} - \mathbb{E}\{\bar{L}_t^{n,\alpha}|\mathcal{F}_t^0\} \to 0, \quad t \geq 0. \tag{2.9}$$

Now if $\bar{L}^* = \alpha$, that is $\bar{L}^n \to \alpha$, one expects that the system (2.1), (2.4)–(2.6) converges to the system (2.8) in certain sense. In particular, $\bar{L}^{n,\alpha}$ and $\bar{L}^n$ should have the same limit, that is, we should expect that the process α would have the following "fixed point" property:

$$\alpha_t = \lim_{n\to\infty} \mathbb{E}\{\bar{L}_t^{n,\alpha}|\mathcal{F}_t^0\}, \quad t \geq 0, \tag{2.10}$$

provided that the limit and the fixed point α both exist.

In Theorem 4.9 below, we will provide some sufficient conditions so that the fixed point problem (2.10) has a solution. Our main result of the paper, Theorem 2.11 below, proves the Law of Large Numbers in our model. That is, it shows that if α solves the fixed point problem (2.10), then under certain technical conditions, we have

$$\lim_{n\to\infty} \mathbb{E}\{|\bar{L}_t^n - \alpha_t|\} = 0, \quad \forall t. \tag{2.11}$$

We finish this section by presenting a simple example in which $\bar{L}$ is the average number of defaults.

Example 2.3 Assume $L^i \equiv 1$, $\lambda_i = \lambda$, $\forall i$, and λ is independent of X^i (i.e., a "zero-factor" scenario). Then, conditioning on the values of X^0, all τ_i^α's have the same (exponential) distribution and the right-hand side in (2.10) is equal to

$$\mathbb{P}\Big\{\tau_1^\alpha \leq t\Big|\mathcal{F}_t^0\Big\} = 1 - e^{-Y_t^\alpha} = 1 - e^{-\int_0^t \lambda(s, X_s^{0,\alpha}, \alpha_s)ds},$$

and the equation (2.10) for α becomes

$$\alpha_t = 1 - e^{-\int_0^t \lambda(s, X_s^{0,\alpha}, \alpha_s)ds}.$$

A simple calculation implies that α should satisfy the following ODE:

$$\alpha_t' = (1 - \alpha_t)\lambda(t, X_t^0, \alpha_t), \quad \alpha_0 = 0. \tag{2.12}$$

■

3 Well-posedness

In this section we verify that the system (2.1), (2.4)–(2.6) is indeed well-defined. In other words, we show that, for each $n \in \mathbb{N}$, there exists a unique solution $(X^0, \{X^i, Y^i\}_{i=1}^n)$ that satisfies (2.1), (2.4)–(2.6). For this purpose we impose the following technical conditions.

Assumption 3.1 *(i) The mappings $x_0 \mapsto b_0(t, \omega, x_0, \alpha), \sigma_0(t, \omega, x_0, \alpha)$ are uniformly Lipschitz, uniformly in (t, ω, α); and the mappings $x_i \mapsto b_i(t, \omega, x_0, x_i, \alpha)$ and $\sigma(t, \omega, x_0, x_i, \alpha)$ are uniformly Lipschitz, uniformly in (t, ω, x_0, α).*

(ii) Let $D_0 \subset \mathbb{R}$ denote domain of L^i, that is, L^i takes values in D_0. There exists a constant $K > 0$ such that, for any $\alpha \in D_0$, any $i = 1, \cdots, n$, and any (t, ω, x_0, x_i),

$$\begin{aligned}
&|b_i(t, \omega, x_0, 0, \alpha) - b_i(t, \omega, 0, 0, \alpha)| + |\sigma_i(t, \omega, x_0, 0, \alpha) - \sigma_i(t, \omega, 0, 0, \alpha)| \leq K(1 + |x_0|), \\
&|\lambda_i(t, \omega, x_0, x_i, \alpha) - \lambda_i(t, \omega, 0, 0, \alpha)| \leq K(1 + |x_0| + |x_i|); \\
&\mathbb{E}\Big\{\int_0^T \sup_{\alpha \in D_0} \Big[|b_0|^2 + |\sigma_0|^2 + |b_i|^2 + |\sigma_i|^2 + |\lambda_i|\Big](t, 0, 0, \alpha)dt\Big\} < \infty.
\end{aligned}$$

We then have the following theorem.

Theorem 3.2 *Assume Assumptions 2.1 and 3.1 hold. Then for each $n \in \mathbb{N}$, the system (2.1), (2.4)–(2.6) admits a unique $\mathbb{F}$-adapted solution $(X^0, \{X^i, Y^i\}_{i=1}^n)$.*

Proof. In this proof and in the sequel we denote by $\tau_1^* \leq \cdots \leq \tau_n^*$ the order statistics of stopping times $\tau_1, \cdots, \tau_n$. We construct a solution to the system in the following. It can be seen from the construction that the solution is unique.

Notice that, if there is a solution, one must have $\bar{L}_t = 0$ for $t < \tau_1^*$. We thus first consider the following system:

$$\begin{aligned}
X_t^{0,1} &= x_0 + \int_0^t b_0(s, X_s^{0,1}, 0)ds + \int_0^t \sigma_0(s, X_s^{0,1}, 0)dB_s^0; \\
X_t^{i,1} &= x_i + \int_0^t b_i(s, X_s^{0,1}, X_s^{i,1}, 0)ds + \int_0^t \sigma_i(s, X_s^{0,1}, X_s^{i,1}, 0)dB_s^i, \quad i = 1, \cdots, n.
\end{aligned}$$

This SDE obviously has a unique solution $(X^{0,1}, \{X^{i,1}\}_{i=1}^n)$ under Assumptions 2.1 and 3.1. We can then define

$$
\begin{aligned}
&Y_t^{i,1} \triangleq \int_0^t \lambda_i(s, X_s^{0,1}, X_s^{i,1}, 0)ds, \quad \tau_i^1 \triangleq \inf\Big\{t \geq 0 : Y_t^{i,1} \geq E^i\Big\}, \quad i = 1, \cdots, n; \\
&\bar{L}_t^{n,1} \triangleq \frac{1}{n}\sum_{i=1}^n L^i_{\tau_i^1} \mathbf{1}_{\{\tau_i^1 \leq t\}}.
\end{aligned}
$$

Suppose that we have defined processes $(X^{0,k}, X^{i,k}, Y^{i,k}, \bar{L}^{n,k})$ and stopping times τ_i^k for $i = 1, \cdots, n$. Now for $k+1$, recalling that $\tau_k^{k,*}$ is the k-th order statistic of $\tau_1^k, \cdots, \tau_n^k$, we define for $i = 1, \cdots, n$

$$
(X_t^{0,k+1}, X_t^{i,k+1}, Y_t^{i,k+1}, \bar{L}_t^{n,k+1}) \triangleq (X_t^{0,k}, X_t^{i,k}, Y_t^{i,k}, \bar{L}_t^{n,k}), \quad 0 \leq t \leq \tau_k^{k,*}, \tag{3.1}
$$

and for $t \geq \tau_k^{k,*}$ and $i = 1, \cdots, n$,

$$
\begin{aligned}
X_t^{0,k+1} &= X^{0,k}_{\tau_k^{k,*}} + \int_{\tau_k^{k,*}}^t b_0(s, X_s^{0,k+1}, \bar{L}^{n,k}_{\tau_k^{k,*}})ds + \int_{\tau_k^{k,*}}^t \sigma_0(s, X_s^{0,k+1}, \bar{L}^{n,k}_{\tau_k^{k,*}})dB_s^0; \\
X_t^{i,k+1} &= X^{i,k}_{\tau_k^{k,*}} + \int_{\tau_k^{k,*}}^t b_i(s, X_s^{0,k+1}, X_s^{i,k+1}, \bar{L}^{n,k}_{\tau_k^{k,*}})ds + \int_{\tau_k^{k,*}}^t \sigma_i(s, X_s^{0,k+1}, X_s^{i,k+1}, \bar{L}^{n,k}_{\tau_k^{k,*}})dB_s^i; \\
Y_t^{i,k+1} &\triangleq Y^{i,k}_{\tau_k^{k,*}} + \int_{\tau_k^{k,*}}^t \lambda_i(s, X_s^{0,k+1}, X_s^{i,k+1}, \bar{L}^{n,k}_{\tau_k^{k,*}})ds; \\
\tau_i^{k+1} &\triangleq \inf\Big\{t \geq 0 : Y_t^{i,k+1} \geq E^i\Big\}, \quad \bar{L}_t^{n,k+1} \triangleq \frac{1}{n}\sum_{i=1}^n L^i_{\tau_i^{k+1}} \mathbf{1}_{\{\tau_i^{k+1} \leq t\}}.
\end{aligned}
$$

This defines τ_i^{k+1}, $i = 1, , \cdots, n$. By (3.1), it is clear that

$$
\tau_j^{k+1,*} = \tau_j^{k,*}, \quad j = 1, \cdots, k. \tag{3.2}
$$

Repeating the same procedure, we may define $(X^{0,n}, X^{i,n}, Y^{i,n}, \bar{L}^{n,n})$ and τ_i^n for $i = 1, \cdots, n$. Finally, we define

$$
(X_t^0, X_t^i, Y_t^i, \bar{L}_t) \triangleq (X_t^{0,n}, X_t^{i,n}, Y_t^{i,n}, \bar{L}_t^{n,n}), \quad 0 \leq t \leq \tau_n^{n,*}, \tag{3.3}
$$

and for $t > \tau_n^{n,*}$,

$$
\begin{aligned}
X_t^0 &= X^0_{\tau_n^{n,*}} + \int_{\tau_n^{n,*}}^t b_0(s, X_s^0, \bar{L}_{\tau_n^{n,*}})ds + \int_{\tau_n^{n,*}}^t \sigma_0(s, X_s^0, \bar{L}_{\tau_n^{n,*}})dB_s^0; \\
X_t^i &= X^i_{\tau_n^{n,*}} + \int_{\tau_n^{n,*}}^t b_i(s, X_s^0, X_s^i, \bar{L}_{\tau_n^{n,*}})ds + \int_{\tau_n^{n,*}}^t \sigma_i(s, X_s^0, X_s^i, \bar{L}_{\tau_n^{n,*}})dB_s^i; \\
Y_t^i &\triangleq Y^i_{\tau_n^{n,*}} + \int_{\tau_n^{n,*}}^t \lambda_i(s, X_s^0, X_s^i, \bar{L}_{\tau_n^{n,*}})ds; \\
\bar{L}_t &\triangleq \bar{L}_{\tau_n^{n,*}}.
\end{aligned}
$$

This defines $(X^0_t, X^i_t, Y^i_t, \bar L_t)$ for $t \ge 0$. Moreover, define τ_i by (2.4), , $i = 1, \cdots, n$. One can check straightforwardly that $(X^0_t, X^i_t, Y^i_t, \bar L_t, \tau_i)$ satisfies the system (2.1), (2.4)–(2.6), and

$$\tau^*_j = \tau^{n,*}_j = \tau^{k,*}_j, \quad 1 \le j \le k \le n. \tag{3.4}$$

■

The next proposition gives the conditional distribution of stopping times τ^{k+1}_i, when the previous defaults are known. We say that random variables ξ_i are *conditionally independent on D* if $\xi_i \mathbf{1}_D$ are conditionally independent.

Proposition 3.3 *Assume Assumptions 2.1 and 3.1 hold, ans let $i_1, \cdots, i_k$ be given. In the framework of Theorem 3.2, and recalling (3.4), denote*

$$D_k \stackrel{\triangle}{=} \{\tau^*_1 = \tau^k_{i_1}, \cdots, \tau^*_k = \tau^k_{i_k}\}, \quad \mathcal{G}^k_t \stackrel{\triangle}{=} \Big(\bigvee_{l=1}^k \mathcal{F}^{i_l}_{\tau^*_k + t}\Big) \bigvee \Big(\bigvee_{j \ne i_1, \cdots, i_k} \mathcal{F}^j_{\tau^*_k}\Big). \tag{3.5}$$

Then, for $j \ne i_1, \cdots, i_k$ and $t \ge 0$,

$$\mathbb{P}\Big\{\tau^{k+1}_j > \tau^*_k + t \;\Big|\mathcal{G}^k_t, D_k\Big\} \;=\; \mathbb{E}\Big\{\exp(Y^{j,k+1}_{\tau^*_k} - Y^{j,k+1}_{\tau^*_k + t})\Big|\mathcal{G}^k_t, D_k\Big\} \quad \text{on} \quad D_k. \tag{3.6}$$

Moreover, conditional on $\mathcal{G}^k_t \vee \sigma(D_k)$, the random vectors $(X^{j,k+1}_{\tau^*_k + t}, Y^{j,k+1}_{\tau^*_k + t}, \mathbf{1}_{\{\tau^{k+1}_j > \tau^*_k + t\}})$, $j \ne i_1, \cdots, i_k$, are conditionally independent on D_k, and consequently,

$$\mathbb{P}\Big\{\tau^*_{k+1} > \tau^*_k + t\Big|\mathcal{G}^k_t, D_k\Big\} = \mathbb{E}\Big\{\exp\Big(\sum_{j \ne i_1, \cdots, i_k} (Y^{j,k+1}_{\tau^*_k} - Y^{j,k+1}_{\tau^*_k + t})\Big)\Big|\mathcal{G}^k_t, D_k\Big\} \quad \text{on} \quad D_k. \tag{3.7}$$

Proof. (i) We first prove (3.6). For arbitrarily given $t_1 < \cdots < t_k$, denote

$$\tilde D_k \stackrel{\triangle}{=} D_k \cap \Big\{\tau^*_1 = t_1, \cdots, \tau^*_k = t_k\Big\}, \tag{3.8}$$

and define

$$(\tilde X^{0,1}, \tilde X^{i,1}, \tilde Y^{i,1}) \stackrel{\triangle}{=} (X^{0,1}, X^{i,1}, Y^{i,1}), \quad \text{and} \quad \tilde L^{n,1}_{t_1} \stackrel{\triangle}{=} \frac{1}{n} L^{i_1}_{t_1}.$$

For $j = 1, \cdots, k$, define $(\tilde X^{0,j+1}_t, \tilde X^{i,j+1}_t, \tilde Y^{i,j+1}_t) \stackrel{\triangle}{=} (\tilde X^{0,j}_t, \tilde X^{i,j}_t, \tilde Y^{i,j}_t)$ for $t \le t_j$, and for $t \ge t_j$,

$$\begin{aligned}
\tilde X^{0,j+1}_t &= \tilde X^{0,j}_{t_j} + \int_{t_j}^t b_0(s, \tilde X^{0,j+1}_s, \tilde L^{n,j}_{t_j}) ds + \int_{t_j}^t \sigma_0(s, \tilde X^{0,j+1}_s, \tilde L^{n,j}_{t_j}) dB^0_s; \\
\tilde X^{i,j+1}_t &= \tilde X^{i,j}_{t_j} + \int_{t_j}^t b_i(s, \tilde X^{0,j+1}_s, \tilde X^{i,j+1}_s, \tilde L^{n,j}_{t_j}) ds + \int_{t_j}^t \sigma_i(s, \tilde X^{0,j+1}_s, \tilde X^{i,j+1}_s, \tilde L^{n,j}_{t_j}) dB^i_s; \\
\tilde Y^{i,j+1}_t &\stackrel{\triangle}{=} \tilde Y^{i,j}_{t_j} + \int_{t_j}^t \lambda_i(s, \tilde X^{0,j+1}_s, \tilde X^{i,j+1}_s, \tilde L^{n,j}_{t_j}) ds;
\end{aligned}$$

where, for $j > 1$,

$$\tilde{L}^{n,j}_{t_j} \triangleq \tilde{L}^{n,j-1}_{t_{j-1}} + \frac{1}{n} L^{i_j}_{t_j}.$$

Then, it is clear that

$$(X^{0,k+1}, X^{i,k+1}, Y^{i,k+1}) = (\tilde{X}^{0,k+1}, \tilde{X}^{i,k+1}, \tilde{Y}^{i,k+1}) \ \text{ on } \ \tilde{D}_k. \tag{3.9}$$

Note that, for any i and t,

$$\begin{aligned}
\{\tau_i^{k+1} > t\} &= \{E_i > Y_t^{i,k+1}\}, \\
\{\tau_i^{k+1} = t\} &= \Big\{Y_t^{i,k+1} = E_i \ \text{ and } \ Y_s^{i,k+1} < E_i \text{ for all } s < t\Big\}.
\end{aligned}$$

Then

$$\begin{aligned}
\tilde{D}_k &= \Big\{\tau_{i_1}^{k+1} = t_1, \cdots, \tau_{i_k}^{k+1} = t_k, Y_{t_k}^{i,k+1} < E_i, i \neq i_1, \cdots, i_k\Big\} \\
&= \Big\{\tau_{i_1}^{k+1} = t_1, \cdots, \tau_{i_k}^{k+1} = t_k, E_j > Y_{t_k}^{j,k+1}, E_i > Y_{t_k}^{i,k+1}, i \neq i_1, \cdots, i_k, j\Big\}.
\end{aligned}$$

and, for each j,

$$\begin{aligned}
&\mathcal{G}_t^k \vee \sigma\Big\{\tau_{i_1}^{k+1} = t_1, \cdots, \tau_{i_k}^{k+1} = t_k, E_i > Y_{t_k}^{i,k+1}, i \neq i_1, \cdots, i_k, j\Big\} \\
\subseteq \ & \tilde{\mathcal{G}}_t^{k,j} \triangleq \Big(\bigvee_{i=1}^n \mathcal{F}^i_{t_k+t}\Big) \bigvee \Big(\bigvee_{i\neq j} \sigma(E_i)\Big).
\end{aligned}$$

Then, by (3.9), on $\tilde{D}_k$ we have

$$\begin{aligned}
\mathbb{P}\Big\{\tau_j^{k+1} > \tau_k^* + t \Big| \mathcal{G}_t^k, \tilde{D}_k\Big\} &= \mathbb{E}\Big\{\mathbb{P}\Big\{E_j > Y_{t_k+t}^{j,k+1} \Big| \tilde{\mathcal{G}}_t^{k,j}, E_j > Y_{t_k}^{j,k+1}\Big\} \Big| \mathcal{G}_t^k, \tilde{D}_k\Big\} \\
&= \mathbb{E}\Big\{\mathbb{P}\Big\{E_j > \tilde{Y}_{t_k+t}^{j,k+1} \Big| \tilde{\mathcal{G}}_t^{k,j}, E_j > \tilde{Y}_{t_k}^{j,k+1}\Big\} \Big| \mathcal{G}_t^k, \tilde{D}_k\Big\} \qquad (3.10)
\end{aligned}$$

Given $\tilde{\mathcal{G}}_t^{k,j}$ and $E_j > \tilde{Y}_{t_k}^{j,k+1}$, one can evaluate the conditional probability of the set $E_j > \tilde{Y}_{t_k+t}^{j,k+1}$ in (3.10) as

$$\mathbb{P}\Big\{E_j > \tilde{Y}_{t_k+t}^{j,k+1} \Big| \tilde{\mathcal{G}}_t^{k,j}, E_j > \tilde{Y}_{t_k}^{j,k+1}\Big\} = \mathbb{E}\Big\{\exp(\tilde{Y}_{t_k}^{j,k+1} - \tilde{Y}_{t_k+t}^{j,k+1}) \Big| \tilde{\mathcal{G}}_t^{k,j}, E_j > \tilde{Y}_{t_k}^{j,k+1}\Big\}.$$

Thus, by (3.9) again, we can continue from (3.10) to get

$$\begin{aligned}
\mathbb{P}\Big\{\tau_j^{k+1} > \tau_k^* + t \Big| \mathcal{G}_t^k, \tilde{D}_k\Big\} &= \mathbb{E}\Big\{\mathbb{E}\Big\{\exp(\tilde{Y}_{t_k}^{j,k+1} - \tilde{Y}_{t_k+t}^{j,k+1}) \Big| \tilde{\mathcal{G}}_t^{k,j}, E_j > \tilde{Y}_{t_k}^{j,k+1}\Big\} \Big| \mathcal{G}_t^k, \tilde{D}_k\Big\} \\
&= \mathbb{E}\Big\{\mathbb{E}\Big\{\exp(Y_{t_k}^{j,k+1} - Y_{t_k+t}^{j,k+1}) \Big| \tilde{\mathcal{G}}_t^{k,j}, E_j > Y_{t_k}^{j,k+1}\Big\} \Big| \mathcal{G}_t^k, \tilde{D}_k\Big\} \\
&= \mathbb{E}\Big\{\exp(Y_{t_k}^{j,k+1} - Y_{t_k+t}^{j,k+1}) \Big| \mathcal{G}_t^k, \tilde{D}_k\Big\}. \qquad (3.11)
\end{aligned}$$

Since $t_1, \cdots, t_k$ are arbitrary, (3.6) follows.

(ii) By the arguments in (i), clearly $\bar{L}_{t_k}\mathbf{1}_{\tilde{D}_k}$, $X^{j,k+1}_{t_k}$, $Y^{j,k+1}_{t_k}$ are all $\mathcal{G}^k_0\vee\sigma(\tilde{D}_k)$-measurable, $j\neq i_1,\cdots,i_k$. Then conditional on the filtration $\{\mathcal{G}^k_t\vee\sigma(\tilde{D}_k),t\geq 0\}$, the processes $\{X^{j,k+1}_{t_k+\cdot},j\neq i_1,\cdots,i_k\}$ are conditionally independent on $\tilde{D}_k$. Thus so are $\{Y^{j,k+1}_{t_k+\cdot},j\neq i_1,\cdots,i_k\}$ and therefore all τ^{k+1}_j's are conditionally independent on $\tilde{D}_k$. Since $t_1,\cdots,t_k$ are arbitrary, we see that $(X^{j,k+1}_{\tau^*_k+\cdot},Y^{j,k+1}_{\tau^*_k+\cdot},\tau^{k+1}_j)$, $j\neq i_1,\cdots,i_k$, are conditionally independent on D_k, conditional on the filtration $\{\mathcal{G}^k_t\vee\sigma(D_k),t\geq 0\}$. Since $\tau^*_{k+1}=\tau^{k+1,*}_{k+1}=\min\{\tau^{k+1}_j:j\neq i_1,\cdots,i_k\}$ on the set D_k, (3.7) follows from (3.6) immediately. ■

We conclude this section by some monotonicity properties of the system (2.8).

Assumption 3.4 *b_0 is decreasing in α; for all i, b_i is increasing in x_0 and decreasing in α; λ_i is decreasing in x_0,x_i and increasing in α; $L^i\geq 0$ and is decreasing in t.*

Lemma 3.5 *Assume that Assumptions 2.1, 3.1 and 3.4 hold. Then for any $\mathbb{F}^0$-adapted process α taking values in D_0, the system (2.8) is well-posed. Moreover, τ^α_i is decreasing in α, $i=1,\cdots,n$, and $\bar{L}^\alpha$ is increasing in t and α.*

Proof. Under Assumptions 2.1 and 3.1, it is clear that the system (2.8) is well-posed. Since $L^i\geq 0$, we see immediately that $\bar{L}^\alpha$ is increasing in t.

We now assume $\alpha^1\leq\alpha^2$. By the standard comparison theorem of SDEs one can easily show that

$$X^{0,\alpha_1}\geq X^{0,\alpha^2},\quad X^{i,\alpha_1}\geq X^{i,\alpha_2},\quad Y^{i,\alpha_1}\leq Y^{i,\alpha_2}.$$

It follows immediately that $\tau^{\alpha_1}_i\geq\tau^{\alpha_2}_i$. Since L^i is decreasing in t, we see that $\bar{L}^{\alpha_1}\leq\bar{L}^{\alpha_2}$. ■

Remark 3.6 (i) If we interpret X^i as the performance of the i-th firm, then the monotonicity assumptions in Assumption 3.4 imply that the n firms are "partners" and are positively correlated to the common factor X^0, and thus they are all negatively correlated to the average past loss $\bar{L}$.

(ii) Assumption 3.4 can be replaced by

b_0 is increasing in α; and for all i, b_i is increasing in x^0 and α;

λ_i is decreasing in x^0,x^i and α; $L^i\geq 0$ and is decreasing in t.

In this case the firms are "competitors", and all the results in this paper will still hold true, after some obvious modifications. ■

4 The Fixed Point Theorem

Recall that the fixed point problem (2.10) provides the candidate for the limit process $\bar{L}^*$. We first have the following obvious result:

Proposition 4.1 *In the setting of Example 2.3, if λ is bounded and uniformly Lipschitz continuous in α, then ODE (2.12) has a unique solution α taking values in $[0,1]$, and thus (2.10) has a unique fixed point.*

In the rest of this section we consider a more general and non-trivial case, in which the fixed point argument works. First, recall the coefficients in (2.5) and (2.6). For simplicity, we assume in this section that

$$x_i = x, \quad b_i = b, \quad \sigma_i = \sigma, \quad \lambda_i = \lambda, \quad i \geq 1. \tag{4.1}$$

We next introduce assumptions on the loss processes L^i. Since L^i is $\mathbb{F}^i$-adapted, we can write

$$L^i_t = \varphi_i(t, B^0_{\cdot\wedge t}, B^i_{\cdot\wedge t}), \quad t \geq 0, \quad i = 1, 2, \cdots \tag{4.2}$$

where each $\varphi_i : \mathbb{R}_+ \times C([0,\infty);\mathbb{R})^2 \to \mathbb{R}$ is a measurable function. The simplest case is the one in which all φ_i's are identical. However, we may consider a more general case in which there is a classification over the possible level of losses. The basic idea is that there are different loss types, known to the public, and each firm's loss at default falls into a particular type with a certain "frequency." The following definition, albeit technical, reflects the essence of this idea in a general form.

Definition 4.2 *Let $\varphi \stackrel{\triangle}{=} \{\varphi(\theta)\}_{\theta\in[0,1]}$ be a family of measurable mappings $\varphi(\theta) : \mathbb{R}_+ \times C(\mathbb{R}_+)^2 \to \mathbb{R}$ and μ a probability measure on $[0,1]$. We say the sequence $\{\varphi_i, i \geq 1\}$ has distribution (φ, μ) if, for any $\varepsilon > 0$ and $T > 0$, there exist $k = k(\varepsilon, T)$, disjoint subsets $\Theta_1, \cdots, \Theta_k \subset [0,1]$, and disjoint subsets $D_1, \cdots, D_k \subset \mathbb{N}$ such that*

$$\begin{aligned}
&\mu\Big([0,1]\Big\backslash(\Theta_1 \cup \cdots \cup \Theta_k)\Big) < \varepsilon;\\
&\sup_{i\in D_j} \|\varphi_i - \frac{1}{\mu(\Theta_j)}\int_{\Theta_j} \varphi(\theta)d\mu(\theta)\|_{T,\infty} < \varepsilon, \quad j = 1, \cdots, k;\\
&\lim_{n\to\infty} \frac{\Big|D_j \cap \{1, \cdots, n\}\Big|}{n} = \mu(\Theta_j), \quad j = 1, \cdots, k.
\end{aligned} \tag{4.3}$$

Here $\|\varphi\|_{T,\infty} \stackrel{\triangle}{=} \sup\Big\{|\varphi(t, \mathbf{x}^0_{\cdot\wedge t}, \mathbf{x}_{\cdot\wedge t}) : 0 \leq t \leq T, \mathbf{x}^0, \mathbf{x} \in C(\mathbb{R}^+)\Big\}$.

To illustrate the idea behind Definition 4.2, we provide several examples.

Example 4.3 (Singleton case) *Let $\theta_0 \in [0,1]$ and $\mu(\{\theta_0\}) = 1$. Then $\{\varphi_i, i \geq 1\}$ has distribution (φ, μ) if and only if there exists a set $D \subset \mathbb{N}$ such that*

$$\lim_{n\to\infty} \frac{\Big|D \cap \{1, \cdots, n\}\Big|}{n} = 1 \quad \textit{and} \quad \lim_{i\in D, i\to\infty} \|\varphi_i - \varphi(\theta_0)\|_{T,\infty} = 0 \;\; \textit{for any} \;\; T > 0.$$

The simplest case for which $\{\varphi_i, i \geq 1\}$ has distribution (φ, μ) in this case is of course when $\varphi_i = \varphi(\theta_0)$ for all $i \geq 1$. That is, there is only one type of loss.

Example 4.4 (Discrete case) *Let $\{\theta_k, k \geq 1\} \subset [0,1]$ and $\mu(\{\theta_k, k \geq 1\}) = 1$. Then $\{\varphi_i, i \geq 1\}$ has distribution (φ, μ) if and only if there exist disjoint subsets $D_k \subset \mathbb{N}$, $k \geq 1$, such that*

$$\lim_{n\to\infty} \frac{\Big|D_k \cap \{1, \cdots, n\}\Big|}{n} = \mu(\theta_k) \quad \textit{and} \quad \lim_{i\in D_k, i\to\infty} \|\varphi_i - \varphi(\theta_k)\|_{T,\infty} = 0, \quad T > 0, k \geq 1.$$

In particular, if $k = 2$, $\mu(\theta_1) = \mu(\theta_2) = \frac{1}{2}$, then we could set $\varphi_i = \varphi(\theta_1)$ when i is odd and $\varphi_i = \varphi(\theta_2)$ when i is even, so that $\{\varphi_i, i \geq 1\}$ has distribution (φ, μ).

Example 4.5 (Continuous case) *Let μ be the Lebesgue measure on $[0,1]$ and $\varphi(\theta) = \theta\varphi_0$, where φ_0 is a given mapping: $\mathbb{R}_+ \times C(\mathbb{R}_+)^2 \to \mathbb{R}$. For each n and $2^{n-1} \leq i < 2^n$, assume $\varphi_i = (i2^{1-n} - 1)\varphi_0$. Then one can easily check that $\{\varphi_i, i \geq 1\}$ has distribution (φ, μ).*

We will need the following assumptions on the coefficients:

Assumption 4.6 *(i) (4.1) holds and $\sigma_0(s, x_0, \alpha) = \sigma_0(s, x_0)$, $\sigma(s, x_0, x_i, \alpha) = \sigma(s, x_i)$;*
(ii) (4.2) holds and $\{\varphi_i, i \geq 1\}$ has distribution (φ, μ), in the sense of Definition 4.2;
(iii) there exists a constant $K > 0$ such that $|\varphi_i| \leq K$ and $|\lambda| \leq K$.

We note that under Assumption 4.6 (i), the system (2.8) now becomes:

$$\begin{aligned}
X_t^{0,\alpha} &= x_0 + \int_0^t b_0(s, B^0_{\cdot\wedge s}, X_s^{0,\alpha}, \alpha_s)ds + \int_0^t \sigma_0(s, B^0_{\cdot\wedge s}, X_s^{0,\alpha})dB_s^0;\\
X_t^{i,\alpha} &= x + \int_0^t b(s, B^0_{\cdot\wedge s}, B^i_{\cdot\wedge s}, X_s^{0,\alpha}, X_s^{i,\alpha}, \alpha_s)ds + \int_0^t \sigma(s, B^0_{\cdot\wedge s}, B^i_{\cdot\wedge s}, X_s^{i,\alpha})dB_s^i;\\
Y_t^{i,\alpha} &\triangleq \int_0^t \lambda(s, B^0_{\cdot\wedge s}, B^i_{\cdot\wedge s}, X_s^{0,\alpha}, X_s^{i,\alpha}, \alpha_s)ds; \quad \tau_i^\alpha \triangleq \inf\Big\{t \geq 0 : Y_t^{i,\alpha} \geq E^i\Big\}.
\end{aligned} \tag{4.4}$$

The following lemma is useful.

Lemma 4.7 *Assume Assumptions 2.1, 3.1 and 4.6 hold, and let α be an $\mathbb{F}^0$-adapted process taking values in $D_0 \stackrel{\triangle}{=} [-K,K]$. Denote*

$$
\begin{aligned}
\bar{\varphi} &\stackrel{\triangle}{=} \int_0^1 \varphi(\theta) d\mu(\theta); &(4.5)\\
\Gamma_t(\alpha) &\stackrel{\triangle}{=} \mathbb{E}\Big\{\int_0^t \bar{\varphi}(s, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s})\lambda(s, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s}, X^{0,\alpha}_s, X^{1,\alpha}_s, \alpha_s) e^{-Y^{1,\alpha}_s} ds \Big| \mathcal{F}^0_t\Big\}. &(4.6)
\end{aligned}
$$

Then

(i) τ_i^α are conditionally i.i.d., conditional on $\mathbb{F}^0$, and

$$\lim_{n\to\infty} \mathbb{E}\{|\bar{L}^{n,\alpha}_t - \Gamma_t(\alpha)|\} = 0, \tag{4.7}$$

(ii) Moreover, if Assumption 3.4 also holds, then $\Gamma(\alpha)$ is continuous and increasing in t, increasing in α, and satisfies $0 \le \Gamma_t(\alpha) \le K$, a.s.

(iii) The process $\Gamma(\alpha)$ can be written as

$$\Gamma_t(\alpha) = \int_0^t \mathbb{E}\Big\{\bar{\varphi}(s, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s})\lambda(s, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s}, X^{0,\alpha}_s, X^{1,\alpha}_s, \alpha_s) e^{-Y^{1,\alpha}_s} \Big| \mathcal{F}^0_s\Big\} ds. \tag{4.8}$$

Proof. (i) By our assumptions, it is readily seen that $\{(B^i, X^{i,\alpha}, Y^{i,\alpha}, \tau_i^\alpha)\}_{i=1}^n$ are conditionally i.i.d., conditional on $\mathcal{F}^0_t$. So it suffices to prove (4.7).

For any $t > 0$ and $\varepsilon > 0$, let k, Θ_j, D_j, $j = 1, \cdots k$, be as in Definition 4.2. Denote

$$\Theta_{k+1} \stackrel{\triangle}{=} [0,1]\backslash(\Theta_1 \cup \cdots \cup \Theta_k), \quad D_{k+1} \stackrel{\triangle}{=} \mathbb{N}\backslash(D_1 \cup \cdots \cup D_k), \quad D^n_j \stackrel{\triangle}{=} D_j \cap \{1, \cdots, n\},$$

and

$$\bar{\varphi}_j \stackrel{\triangle}{=} \frac{1}{\mu(\Theta_j)} \int_{\Theta_j} \varphi(\theta) d\mu(\theta).$$

Note that, by denoting $\varphi_i(s) \stackrel{\triangle}{=} \varphi_i(s, B^0_{\cdot\wedge s}, B^i_{\cdot\wedge s})$,

$$
\begin{aligned}
&\bar{L}^{n,\alpha}_t = \frac{1}{n}\sum_{i=1}^n L^i_{\tau_i^\alpha} \mathbf{1}_{\{\tau_i^\alpha \le t\}} = \frac{1}{n}\sum_{i=1}^n \varphi_i(\tau_i^\alpha)\mathbf{1}_{\{\tau_i^\alpha \le t\}} = \frac{1}{n}\sum_{j=1}^{k+1}\sum_{i\in D^n_j} \varphi_i(\tau_i^\alpha)\mathbf{1}_{\{\tau_i^\alpha \le t\}}\\
= \;& \frac{1}{n}\Big[\sum_{j=1}^k \sum_{i\in D^n_j} \bar{\varphi}_j(\tau_i^\alpha)\mathbf{1}_{\{\tau_i^\alpha \le t\}} + \sum_{j=1}^k \sum_{i\in D^n_j} [\varphi_i(\tau_i^\alpha) - \bar{\varphi}_j(\tau_i^\alpha)]\mathbf{1}_{\{\tau_i^\alpha \le t\}} + \sum_{i\in D^n_{k+1}} \varphi_i(\tau_i^\alpha)\mathbf{1}_{\{\tau_i^\alpha \le t\}}\Big].
\end{aligned}
$$

and that

$$\begin{aligned}
\Gamma_t(\alpha) &= \mathbb{E}\Big\{\bar\varphi(\tau_1^\alpha)\mathbf{1}_{\{\tau_1^\alpha\le t\}}\Big|\mathcal{F}_t^1\Big|\mathcal{F}_t^0\Big\} = \mathbb{E}\Big\{\bar\varphi(\tau_1^\alpha)\mathbf{1}_{\{\tau_1^\alpha\le t\}}\Big|\mathcal{F}_t^0\Big\}\\
&= \mathbb{E}\Big\{\Big[\sum_{j=1}^{k+1}\bar\varphi_j(\tau_1^\alpha)\mu(\Theta_j)\Big]\mathbf{1}_{\{\tau_1^\alpha\le t\}}\Big|\mathcal{F}_t^0\Big\}\\
&= \sum_{j=1}^{k}\mu(\Theta_j)\mathbb{E}\Big\{\bar\varphi_j(\tau_j^\alpha)\mathbf{1}_{\{\tau_j^\alpha\le t\}}\Big|\mathcal{F}_t^0\Big\} + \mu(\Theta_{k+1})\mathbb{E}\Big\{\bar\varphi_{k+1}(\tau_1^\alpha)\mathbf{1}_{\{\tau_1^\alpha\le t\}}\Big|\mathcal{F}_t^0\Big\}.
\end{aligned}$$

Since $|\varphi_i|\le K$, it is obvious that $|\bar\varphi_i|\le K$. Then by (4.3) we have

$$\begin{aligned}
&\frac{1}{n}\sum_{j=1}^{k}\sum_{i\in D_j^n}|\varphi_i(\tau_i^\alpha)-\bar\varphi_j(\tau_i^\alpha)|\mathbf{1}_{\{\tau_i^\alpha\le t\}}\le\varepsilon;\\
&\frac{1}{n}\sum_{i\in D_{k+1}^n}|\varphi_i(\tau_i^\alpha)|\le\frac{K|D_{k+1}^n|}{n}\to K\mu(\Theta_{k+1})\le K\varepsilon;\\
&\mu(\Theta_{k+1})|\bar\varphi_{k+1}(\tau_1^\alpha)|\le K\mu(\Theta_{k+1})\le K\varepsilon.
\end{aligned}$$

Moreover, for each $j=1,\cdots,k$, by the standard Law of Large Numbers we have

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}\sum_{i\in D_j^n}\bar\varphi_j(\tau_i^\alpha)\mathbf{1}_{\{\tau_i^\alpha\le t\}} &= \lim_{n\to\infty}\frac{|D_j^n|}{n}\frac{1}{|D_j^n|}\sum_{i\in D_j^n}\bar\varphi_j(\tau_i^\alpha)\mathbf{1}_{\{\tau_i^\alpha\le t\}}\\
&= \mu(\Theta_j)\mathbb{E}\Big\{\bar\varphi_j(\tau_1^\alpha)\mathbf{1}_{\{\tau_1^\alpha\le t\}}\Big|\mathcal{F}_t^0\Big\}.
\end{aligned}$$

Thus

$$\overline{\lim_{n\to\infty}}\Big|\bar L_t^{n,\alpha}-\Gamma_t(\alpha)\Big|\le(2K+1)\varepsilon.$$

Since ε is arbitrary, we prove (4.7).

(ii) It follows directly from Lemma 3.5 and (4.7) that $\Gamma(\alpha)$ is increasing in t and α, and $0\le\Gamma_t(\alpha)\le K$. Moreover, denote

$$\gamma_t(\alpha)\stackrel{\triangle}{=}\bar\varphi(t,B^0_{\cdot\wedge t},B^1_{\cdot\wedge t})\lambda(t,B^0_{\cdot\wedge t},B^1_{\cdot\wedge t},X_t^{0,\alpha},X_t^{1,\alpha},\alpha_t)e^{-Y_t^{1,\alpha}}.$$

For any t and $\varepsilon>0$,

$$\begin{aligned}
&\Big|\Gamma_{t+\varepsilon}(\alpha)-\Gamma_t(\alpha)\Big|\\
\le\ &\Big|\mathbb{E}\Big\{\int_t^{t+\varepsilon}\gamma_s(\alpha)ds|\mathcal{F}_{t+\varepsilon}^0\Big\}\Big|+\Big|\mathbb{E}\Big\{\int_0^t\gamma_s(\alpha)ds|\mathcal{F}_{t+\varepsilon}^0\Big\}-\mathbb{E}\Big\{\int_0^t\gamma_s(\alpha)ds|\mathcal{F}_t^0\Big\}\Big|\\
\le\ &K^2\varepsilon+\Big|\mathbb{E}\Big\{\int_0^t\gamma_s(\alpha)ds|\mathcal{F}_{t+\varepsilon}^0\Big\}-\mathbb{E}\Big\{\int_0^t\gamma_s(\alpha)ds|\mathcal{F}_t^0\Big\}\Big|.
\end{aligned}$$

Since the filtration $\mathbb{F}^0$ is continuous, sending $\varepsilon \to 0$ we obtain immediately that $\lim_{\varepsilon\to 0}\Gamma_{t+\varepsilon}(\alpha) = \Gamma_t(\alpha)$. Similarly, one can show that $\lim_{\varepsilon\to 0}\Gamma_{t-\varepsilon}(\alpha) = \Gamma_t(\alpha)$. Therefore, $\Gamma(\alpha)$ is continuous in t.

(iii) First, by the Fubini theorem we can write (4.6) as

$$\Gamma(\alpha)_t = \mathbb{E}\Big\{\int_0^t \gamma_s(\alpha)ds\Big|\mathcal{F}_t^0\Big\} = \int_0^t \mathbb{E}\Big\{\gamma_s(\alpha)\Big|\mathcal{F}_t^0\Big\}ds.$$

Since for each $s \in [0,t]$, $\gamma_s(\alpha)$ is $\mathcal{F}_s$-measurable, and $\mathcal{F}_t^0 = \mathcal{F}_s^0 \vee \mathcal{F}_{s,t}^0$, where $\mathcal{F}_{s,t}^0 \stackrel{\triangle}{=} \bigvee_{s\le u\le t}\mathcal{F}_u^0$ is independent of $\mathcal{F}_s$, it can be fairly easily checked that

$$\mathbb{E}\{\gamma_s(\alpha)|\mathcal{F}_t^0\} = \mathbb{E}\{\gamma_s(\alpha)|\mathcal{F}_s^0 \vee \mathcal{F}_{s,t}^0\} = \mathbb{E}\{\gamma_s(\alpha)|\mathcal{F}_s^0\},$$

and (4.8) follows. ■

Remark 4.8 The condition (4.1) is to ensure that τ_i^α are conditionally i.i.d. and thus one may apply the standard Law of Large Numbers. It can be weakened slightly if one applies generalized Law of Large Numbers by using the Linderberg condition. ■

We conclude this section with the following important result.

Theorem 4.9 *Assume Assumptions 2.1, 3.1, 3.4, and 4.6 hold. Then there exists $\mathbb{F}^0$-adapted process such that $\alpha = \Gamma(\alpha)$.*

Proof. We will apply Zorn's lemma to prove the theorem. First, denote

$$\mathscr{L} \stackrel{\triangle}{=} \Big\{\alpha : \mathbb{F}^0\text{-adapted, increasing, càdlàg , and } \ 0 \le \alpha \le K\Big\}.$$

By Lemma 4.7, we see that $\Gamma(\alpha) \in \mathscr{L}$ for any $\alpha \in \mathscr{L}$. We introduce a partial order "$\preceq$" in $\mathscr{L}$, by $\alpha^1 \preceq \alpha^2$ if and only if $\alpha_t^1 \le \alpha_t^2$, $t \ge 0$, $\mathbb{P}$-a.s. Now consider the set

$$\mathscr{L}_0 \stackrel{\triangle}{=} \{\alpha \in \mathscr{L} : \alpha \le \Gamma(\alpha)\}.$$

Obviously $0 \in \mathscr{L}_0$, so $\mathscr{L}_0$ is not empty.

Assume that $\{\alpha^\theta\}_{\theta\in\Theta}$ is a totally ordered subset of $\mathscr{L}_0$. Define $\hat\alpha_r \stackrel{\triangle}{=} \mathrm{esssup}_{\theta\in\Theta}\,\alpha_r^\theta$ for all $r \in \mathbb{Q}_+$. Then clearly $\hat\alpha_r$ is increasing in r, a.s. Define

$$\hat\alpha_t \stackrel{\triangle}{=} \overline{\lim_{r\in\mathbb{Q}_+\cap(t,\infty), r\downarrow t}}\, \hat\alpha_r, \quad t \ge 0.$$

Then it is easy to check that $\hat\alpha \in \mathcal{L}$. Since α^θ is càdlàg , we have $\hat\alpha_t \ge \alpha_t^\theta$, $t \ge 0$, a.s. for all $\theta \in \Theta$. Furthermore, since Γ is increasing in α, $\Gamma(\hat\alpha) \ge \Gamma(\alpha^\theta) \ge \alpha^\theta$ for all θ. Then

$\Gamma_r(\hat{\alpha}) \geq \hat{\alpha}_r$, $r \in \mathbb{Q}_+$, a.s. Since $\Gamma(\hat{\alpha})$ is continuous, we have $\Gamma_t(\hat{\alpha}) \geq \hat{\alpha}_t$ for all $t \geq 0$, a.s. Thus $\hat{\alpha} \in \mathscr{L}_0$, and therefore, $\hat{\alpha}$ is an upper bound of $\{\alpha^\theta\}_{\theta\in\Theta}$ in $\mathscr{L}_0$.

Now applying Zorn's lemma we conclude that $\mathscr{L}_0$ has a maximum point α^* in $\mathscr{L}_0$. We claim that $\alpha^* = \Gamma(\alpha^*)$. Indeed, suppose that the equality fails. Then there exists $\varepsilon > 0$ such that $\mathbb{P}(\tau_1 < \infty) > 0$, where $\tau_1 \stackrel{\triangle}{=} \inf\Big\{t \geq 0 : \Gamma_t(\alpha^*) \geq \alpha^*_t + \varepsilon\Big\}$ is an $\mathbb{F}^0$-stopping time. Let $\tau_2 \stackrel{\triangle}{=} \inf\{t \geq \tau_1 : \alpha^*_t \geq \alpha^*_{\tau_1} + \varepsilon\Big\}$ be another $\mathbb{F}^0$-stopping time taking values in $[0,\infty]$, and define

$$\tilde{\alpha}^*_t \stackrel{\triangle}{=} \begin{cases} \alpha^*_t, & t < \tau_1 \ \text{ or } \ t \geq \tau_2; \\ \alpha^*_{\tau_1} + \varepsilon, & \tau_1 \leq t < \tau_2. \end{cases}$$

Since α^* is càdlàg , we see that $\tau_2 > \tau_1$ on $\{\tau_1 < \infty\}$, thus

$$\alpha^* \preceq \tilde{\alpha}^* \quad \text{and} \quad \alpha^* \neq \tilde{\alpha}^*. \tag{4.9}$$

On the other hand, by the definition of τ_2 we see that $\tilde{\alpha}^*$ is still increasing, then it is clear that $\tilde{\alpha}^* \in \mathscr{L}$. Moreover, since Γ is increasing in both α and t, then for $t < \tau_1$ or $t \geq \tau_2$, we have $\Gamma_t(\tilde{\alpha}^*) \geq \Gamma_t(\alpha^*) \geq \alpha^*_t$, and for $t \in [\tau_1, \tau_2)$, $\Gamma_t(\tilde{\alpha}^*) \geq \Gamma_t(\alpha^*) \geq \Gamma_{\tau_1}(\alpha^*) \geq \alpha^*_{\tau_1} + \varepsilon = \tilde{\alpha}^*_t$. This implies that $\tilde{\alpha}^* \in \mathscr{L}_0$, in contradiction with (4.9) and the assumption that α^* is a maximum point of $\mathscr{L}_0$. ■

5 Potential Applications

In this section we present some potentially useful applications under the "i.i.d." framework. To the best of our knowledge, these cases have not been fully analyzed in the literature.

5.1 Pricing a single name credit derivative

Suppose we are interested in pricing a credit derivative written on one firm, but the default intensity of the firm, λ, depends on the average number of defaults of many firms, as in our model. If our assumptions hold and that number is approximated by the process α_t, then we can find the price by using $\lambda(t, X^0_t, \alpha_t)$.

Specifically, consider the setting of Example 2.3. Recall that in this case the fixed point can be determined by a randomized ODE (2.12):

$$\alpha_t = \int_0^t (1 - \alpha_s)\lambda(s, X^0_s, \alpha_s)ds. \tag{5.1}$$

Let us assume further that λ is linear in α, that is,

$$\lambda(t, X^0_t, \alpha_t) = A(t, X^0_t) + B(t, X^0_t)\alpha_t,$$

where A and B are continuous functions, and are uniformly Lipschitz in x. Then the ODE (5.1) becomes (path-by-path) a Riccati equation:

$$\alpha_t' = (1-\alpha_t)\lambda(t, X_t^0, \alpha_t) = P(t, X_t^0) + Q(t, X_t^0)\alpha_t + R(t, X_t^0)\alpha_t^2,$$

where $P = A$, $Q = B - A$, and $R = B$. Since the equation clearly has a particular solution $\alpha_t \equiv 1$, the general solution can be written as

$$\alpha_t = 1 + 1/v_t$$

where v_t solves the linear equation

$$v_t' = [A(t, X_t^0) + B(t, X_t^0)]v_t + B(t, X_t^0).$$

Since $\alpha_0 = 0$, we have $v_0 = -1$. Solving this ODE we obtain

$$v_t = -e^{\int_0^t p_s ds} + \int_0^t e^{\int_s^t p_r dr} B(s, X_s^0) ds, \quad t \geq 0,$$

where $p \stackrel{\triangle}{=} A + B$. The process α is thus explicitly found, as a functional of X^0, and we then face a standard problem in credit derivatives pricing, in which the (limiting) intensity only depends on the factor X^0.

If we further assume that A and B are constant, it then follows that

$$\alpha_t = 1 - \frac{A+B}{Ae^{(A+B)t} + B}$$

Thus, the default intensity can be approximated by

$$\hat{\lambda}_t = A + B\alpha_t = (A+B)\left[1 - \frac{B}{Ae^{(A+B)t} + B}\right]$$

We have then shown the following: If the intensity is of the form $\lambda_t = A + BN_t$ where N_t is the average number of defaults of many firms, then we can price derivatives which depend on λ by replacing it by simple deterministic process $\hat{\lambda}$.

5.2 Finding expected loss

We now consider a problem of computing the expected loss of a portfolio of a large number of defaultable loans, for example credit card customers. We assume that the loss of entity i is given by (4.2). According to (4.6) and (4.8), we expect to have

$$\begin{aligned}\alpha_t &= \mathbb{E}\Big\{ \int_0^t \bar{\varphi}(s, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s})\lambda(s, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s}, X_s^{0,\alpha}, X_s^{1,\alpha}, \alpha_s)e^{-Y_s^{1,\alpha}} ds \Big| \mathcal{F}_t^0 \Big\} \\ &= \int_0^t \mathbb{E}\{\bar{\varphi}(s, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s})\lambda(s, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s}, X_s^{0,\alpha}, X_s^{1,\alpha}, \alpha_s)e^{-Y_s^{1,\alpha}} | \mathcal{F}_s^0\} ds. \qquad (5.2)\end{aligned}$$

Let us assume further that

$$\lambda(\cdots) = \lambda_0(t, B^0_{\cdot\wedge t}, \alpha_t) + \lambda_1(t, B^0_{\cdot\wedge t}, B^i_{\cdot\wedge t}).$$

Then, we can write (5.2) as

$$\alpha_t = \int_0^t \Big[F_s\lambda_0(s, B^0_{\cdot\wedge s}, \alpha_s) + G_s\Big] e^{-\int_0^s \lambda_0(u, B^0_{\cdot\wedge s}, \alpha_u)du} ds,$$

where

$$\begin{aligned} F_s &\stackrel{\triangle}{=} \mathbb{E}\Big\{\bar\varphi(s, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s}) e^{-\int_0^s \lambda_1(u, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s})du}\Big|\mathcal{F}^0_s\Big\}, \\ G_s &\stackrel{\triangle}{=} \mathbb{E}\Big\{\bar\varphi(s, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s})\lambda_1(s, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s}) e^{-\int_0^s \lambda_1(u, B^0_{\cdot\wedge s}, B^1_{\cdot\wedge s})du}\Big|\mathcal{F}^0_s\Big\}. \end{aligned}$$

Or equivalently, denoting $\beta^0_t(\alpha) \stackrel{\triangle}{=} e^{-\int_0^t \lambda_0(u, B^0_{\cdot\wedge s}, \alpha_u)du}$,

$$\alpha'_t = [F_t\lambda_0(t, B^0_{\cdot\wedge s}, \alpha_t) + G_t]\beta^0_t(\alpha). \tag{5.3}$$

If we assume, in addition, that

$$\lambda_0(t, B^0_{\cdot\wedge t}, \alpha_t) = \bar\lambda_0(t, B^0_{\cdot\wedge t}) + R(t, B^0_{\cdot\wedge t})\alpha_t = I_t + R_t\alpha_t,$$

then (5.3) becomes

$$\alpha'_t = [(F_tI_t + G_t) + F_tR_t\alpha_t]H_t\beta_t(\alpha) = [\tilde F_t + \tilde R_t\alpha_t]\beta_t(\alpha), \tag{5.4}$$

where $\beta_t(\alpha) = e^{-\int_0^t R_s\alpha_s ds}$, and

$$H_t = e^{-\int_0^t I_s ds},\ \ \tilde F_t = (F_tI_t + G_t)H_t,\ \ \tilde R_t = F_tR_tH_t.$$

Differentiating on both sides of (5.4) and using (5.3) we obtain the ODE for α:

$$\begin{aligned} \alpha''_t &= \{[\tilde F'_t + \tilde R'_t\alpha_t + \tilde R_t\alpha'_t] - [\tilde F_t + \tilde R_t\alpha_t]R_t\alpha_t\}\beta_t(\alpha) \\ &= \frac{[\tilde R_t\alpha'_t + (\tilde R'_t + \tilde F_tR_t)\alpha_t + \tilde R_tR_t\alpha_t^2 + \tilde F'_t]\alpha'_t}{\tilde F_t + \tilde R_t\alpha_t}. \end{aligned} \tag{5.5}$$

Moreover, by (5.2) and (5.4) we have

$$\alpha_0 = 0, \quad \alpha'_0 = \tilde F_0 = F_0I_0 + G_0 = \bar\varphi(0,0,0)I_0 + G_0. \tag{5.6}$$

The equation (5.5) with initial conditions (5.6) is a non-linear second order ODE, which in general can only be solved numerically.

To recap, we have shown that if we impose technical conditions to guarantee that the limiting average loss is indeed equal to α_t, then we should be able to compute this limiting loss, for all times t, in this fairly complex model for individual losses.

6 The Law of Large Numbers

In this section we present our main result. The aim is to show that in our strongly correlated self-exciting model, the Law of Large Numbers still holds, and the limit will be a fixed point discussed in the previous sections. Since the proof is quite lengthy, we defer a part of the proof to the next section.

To begin with we strengthen the technical conditions:

Assumption 6.1 *(i)* $\sigma_0(t, x_0, \alpha) = \sigma_0(t), \sigma_i(t, x_0, x_i, \alpha) = \sigma_i(t)$*;*

(ii) b_0, b_i, λ_i *are Lipschitz continuous in* x_0, x_i*, uniformly in* (t, ω, α)*, and* L^i *is Lipschitz continuous in* t*, with a common Lipschitz constant* K*;*

(iii) b_0, b_i, λ_i *are Lipschitz continuous in* α*, uniformly in* (t, ω, x_0, x_i)*, with a common Lipschitz constant* Λ_0*;*

(iv) $0 < \Lambda_1 \le \lambda_i \le \Lambda_2$*;* $0 \le L^i \le \Lambda_3$*;*

(v) $\Lambda_0 \le \frac{\Lambda_1^2}{3\Lambda_2\Lambda_3}$*.*

Remark 6.2 The condition (v) above implies that the system is "weakly" correlated to the average loss $\bar{L}$. ■

In this and next section, we denote by C a generic constant which depends only on the constants K, Λ_i, $i = 0, 1, 2, 3$ in Assumption 6.1, and it may vary from line to line. We emphasize in particular that C is independent of n. Moreover, we denote by C_ε (resp. $C_{\varepsilon,T}$) if the constant depends additionally on ε (resp. ε, T).

The main result of this paper is the following.

Theorem 6.3 *Assume Assumptions 2.1, 3.1, 3.4, and 6.1 hold. If the fixed point problem (2.10) has an* $\mathbb{F}^0$*-adapted solution* α *satisfying*

$$\lim_{n\to\infty} \mathbb{E}\Big\{|\bar{L}_t^{n,\alpha} - \alpha_t|\Big\} = 0. \tag{6.1}$$

Then the Law of Large Numbers (2.11) holds.

As a direct consequence of Theorems 4.9 and 6.3, and (4.7), we have

Corollary 6.4 *Assume Assumptions 2.1, 3.1, 3.4, 4.6, and 6.1 hold. Let* α *be the solution to the fixed point problem:* $\alpha = \Gamma(\alpha)$*. Then the Law of Large Numbers (2.11) holds.*

Before we prove Theorem 6.3, let us make a quick analysis. We fix some $T > 0$ and consider $t \le T$. First recall $(X^{0,\alpha}, X^{i,\alpha}, Y^{i,\alpha}, , \tau_i^\alpha, \bar{L}^\alpha)$ in (2.8). Since

$$|\bar{L}_t - \alpha_t| \le |\bar{L}_t - \bar{L}_t^\alpha| + |\bar{L}_t^\alpha - \alpha_t|, \tag{6.2}$$

by (6.1) it suffices to analyze the convergence of $\mathbb{E}\Big\{|\bar L_t-\bar L^\alpha_t|\Big\}$. Notice that

$$\mathbb{E}\Big\{|\bar L_t-\bar L^\alpha_t|\Big\}\le\frac{1}{n}\sum_{i=1}^n I_i\quad\text{where}\quad I_i\stackrel{\triangle}{=}\mathbb{E}\{|L_{\tau_i}\mathbf{1}_{\{\tau_i\le t\}}-L_{\tau_i^\alpha}\mathbf{1}_{\{\tau_i^\alpha\le t\}}|\Big\}.\tag{6.3}$$

Without loss of generality we only estimate I_n. Note that

$$I_n\quad\le\quad C\mathbb{E}\Big\{|\tau_n-\tau_n^\alpha|\mathbf{1}_{\{\tau_n\le t,\tau_n^\alpha\le t\}}+\mathbf{1}_{\{\tau_n<t<\tau_n^\alpha\}}+\mathbf{1}_{\{\tau_n^\alpha<t<\tau_n\}}\Big\}.\tag{6.4}$$

Therefore a crucial step is then to estimate

$$\begin{aligned}\mathbb{E}\{\mathbf{1}_{\{\tau_n<t<\tau_n^\alpha\}}\}&=\mathbb{P}\{\tau_n<t<\tau_n^\alpha\}=\mathbb{P}\{Y_t^n>E_n>Y_t^{\alpha,n}\},\\ \mathbb{E}\{\mathbf{1}_{\{\tau_n^\alpha<t<\tau_n\}}\}&=\mathbb{P}\{\tau_n^\alpha<t<\tau_n\}=\mathbb{P}\{Y_t^{\alpha,n}>E_n>Y_t^n\}.\end{aligned}\tag{6.5}$$

The main difficulty here is that $Y^{\alpha,n}$, Y^n, and E_n are <u>*not*</u> independent in general. But without knowing their joint distribution it is difficult to estimate these probabilities. We therefore introduce an approximating system, in which adding a new (n-th) "name" each time n increases, we use the bounds on the underlying processes, so that the probabilities in (6.5) can be estimated. To be more precise, let us consider the following approximating losses. For $i=1,\cdots,n$,

$$\begin{aligned}&X_t^{0,1}=x_0+\int_0^t b_0(s,X_s^{0,1},\hat L_s^1)ds+\int_0^t\sigma_0(s)dB_s^0;\\ &X_t^{i,1}=x_i+\int_0^t b_i(s,X_s^{0,1},X_s^{i,1},\hat L_s^1)ds+\int_0^t\sigma_i(s)dB_s^i;\\ &Y_t^{i,1}\stackrel{\triangle}{=}\int_0^t\lambda_i(s,X_s^{0,1},X_s^{i,1},\hat L_s^1)ds;\\ &\tau_i^1\stackrel{\triangle}{=}\inf\{t:Y_t^{i,1}\ge E_i\},\quad \hat L_t^1\stackrel{\triangle}{=}\frac{1}{n}\sum_{i=1}^{n-1}L^i_{\tau_i^1}1_{\{\tau_i^1\le t\}};\end{aligned}$$

and

$$\begin{aligned}&X_t^{0,2}=x_0+\int_0^t b_0(s,X_s^{0,2},\hat L_s^2)ds+\int_0^t\sigma_0(s)dB_s^0;\\ &X_t^{i,2}=x_i+\int_0^t b_i(s,X_s^{0,2},X_s^{i,2},\hat L_s^2)ds+\int_0^t\sigma_i(s)dB_s^i;\\ &Y_t^{i,2}\stackrel{\triangle}{=}\int_0^t\lambda_i(s,X_s^{0,2},X_s^{i,2},\hat L_s^2)ds;\\ &\tau_i^2\stackrel{\triangle}{=}\inf\{t:Y_t^{i,2}\ge E_i\},\quad \hat L_t^2\stackrel{\triangle}{=}\frac{\Lambda_3}{n}+\frac{1}{n}\sum_{i=1}^{n-1}L^i_{\tau_i^2}1_{\{\tau_i^2\le t\}}.\end{aligned}$$

We emphasize that $\hat L^1$ and $\hat L^2$ do not involve τ_n^1,τ_n^2. Consequently, except for τ_n^1,τ_n^2, the above systems are now independent of E_n. The following theorem is essential for our analysis. We defer its proof to the next section.

Theorem 6.5 *Assume Assumptions 2.1, 3.1, 3.4, and 6.1 hold. Then, for $i = 1, \cdots, n$, it holds that,*

$$\hat{L}_t^1 \le \bar{L}_t \le \hat{L}_t^2,\ X_t^{0,1} \ge X_t^0 \ge X_t^{0,2},\ X_t^{i,1} \ge X_t^i \ge X_t^{i,2},\ Y_t^{i,1} \le Y_t^i \le Y_t^{i,2},\ \tau_i^1 \ge \tau_i \ge \tau_i^2. \tag{6.6}$$

Moreover, for any $T > 0$, there exist constants $\varepsilon = \varepsilon_T \in [0, \frac{1}{2})$ and $C_T > 0$ such that

$$\mathbb{E}\Big\{\Delta X_t^i + \Delta Y_t^i\Big\} \le \frac{C_T}{n^\varepsilon \ln n}, \quad \forall t \in [0, T]. \tag{6.7}$$

where

$$\Delta X_t^i \stackrel{\triangle}{=} X_t^{i,1} - X_t^{i,2}, \quad \Delta Y_t^i \stackrel{\triangle}{=} Y_t^{i,2} - Y_t^{i,1}.$$

Proof. See Section 7. ■

We are now ready to prove Theorem 6.3.

[*Proof of Theorem 6.3.*]

In light of the previous argument and (6.1)-(6.4), we need only to obtain uniform estimates for each term on the right hand side of (6.4) as n goes to ∞.

To this end we first note that, with a simple application of the Gronwall inequality and the uniform Lipschitz conditions on the coefficients, it is readily seen that

$$|X_t^0 - X_t^{0,\alpha}| + |X_t^i - X_t^{i,\alpha}| + |Y_t^i - Y_t^{i,\alpha}| \le C \int_0^t |\bar{L}_s - \alpha_s| ds. \tag{6.8}$$

Now, for $\tau_n < \tau_n^\alpha \le t$, one has

$$Y_{\tau_n}^n = L_n = Y_{\tau_n^\alpha}^{\alpha,n} = Y_{\tau_n}^{\alpha,n} + \int_{\tau_n}^{\tau_n^\alpha} \lambda_n(s, X_s^{\alpha,0}, X_s^{\alpha,n}, \alpha_s) ds \ge Y_{\tau_n}^{\alpha,n} + \Lambda_1[\tau_n^\alpha - \tau_n].$$

Thus

$$\tau_n^\alpha - \tau_n \le \frac{1}{\Lambda_1} |Y_{\tau_n}^n - Y_{\tau_n}^{\alpha,n}| \le C \int_0^t |\bar{L}_s - \alpha_s| ds.$$

With a similar argument for the case $\tau_n \le \tau_n^\alpha$ we then obtain

$$|\tau_n^\alpha - \tau_n| \le \frac{1}{\Lambda_1} |Y_{\tau_n}^n - Y_{\tau_n}^{\alpha,n}| \le C \int_0^t |\bar{L}_s - \alpha_s| ds. \tag{6.9}$$

Next, recall (6.5). By Theorem 6.5 (i) one has

$$\mathbb{P}\{\tau_n < t < \tau_n^\alpha\} = \mathbb{P}\{Y_t^n > E_n > Y_t^{\alpha,n}\} \le \mathbb{P}\{Y_t^{n,2} > E_n > Y_t^{\alpha,n}\}.$$

However, since E_n is now *independent* of $Y_t^{n,2}, Y_t^{\alpha,n}$, we can use the fact that $E_n \sim \exp(1)$ to get

$$\begin{aligned} \mathbb{P}\{\tau_n < t < \tau_n^\alpha\} &\le \mathbb{E}\{|e^{-Y_t^{\alpha,n}} - e^{-Y_t^{n,2}}|\} \le \mathbb{E}\{|Y_t^{\alpha,n} - Y_t^{n,2}|\} \\ &\le \mathbb{E}\Big\{|Y_t^{\alpha,n} - Y_t^n| + |Y_t^n - Y_t^{n,2}|\Big\} \\ &\le C\mathbb{E}\Big\{\int_0^t |\bar{L}_s - \alpha_s| ds + \Delta Y_t^n\Big\}, \end{aligned} \tag{6.10}$$

thanks to (6.8). Similarly we can also derive that

$$\mathbb{P}\{\tau_n^\alpha < t < \tau_n\} \le C\mathbb{E}\Big\{\int_0^t |\bar{L}_s - \alpha_s| ds + \Delta Y_t^n\Big\}. \tag{6.11}$$

This, together with (6.8)-(6.10) and (6.4), as well as (6.8), leads to that

$$I_n \le C\int_0^t \mathbb{E}\{|\bar{L}_s - \alpha_s|\}ds + C\mathbb{E}\Big\{|\bar{L}_t^\alpha - \alpha_t| + \Delta Y_t^n\Big\}.$$

Next, fix $T > 0$. For all $0 \le t \le T$, by Theorem 6.5 we have

$$I_n \le C\int_0^t \mathbb{E}\{|\bar{L}_s - \alpha_s|\}ds + C\mathbb{E}\Big\{|\bar{L}_t^\alpha - \alpha_t|\Big\} + \frac{C_T}{\ln n}.$$

Similarly, for $i = 1, \cdots, n$, we have

$$I_i \le C\int_0^t \mathbb{E}\{|\bar{L}_s - \alpha_s|\}ds + C\mathbb{E}\Big\{|\bar{L}_t^\alpha - \alpha_t|\Big\} + \frac{C_T}{\ln n}.$$

Then (6.2) and (6.3) lead to

$$\mathbb{E}\Big\{|\bar{L}_t^n - \alpha_t|\Big\} \le C\int_0^t \mathbb{E}\{|\bar{L}_s^n - \alpha_s|\}ds + C\mathbb{E}\Big\{|\bar{L}_t^{n,\alpha} - \alpha_t|\Big\} + \frac{C_T}{\ln n}.$$

Applying Gronwall's inequality we obtain

$$\mathbb{E}\{|\bar{L}_t^n - \alpha_t|\} \le C_T\mathbb{E}\Big\{|\bar{L}_t^{n,\alpha} - \alpha_t|\Big\} + \frac{C_T}{\ln n}, \quad 0 \le t \le T.$$

The theorem then follows immediately from (6.1). ■

7 Proof of Theorem 6.5

In this section we prove Theorem 6.5. We begin with two technical lemmas. Both of them are fairly easy to prove.

Lemma 7.1 *Let $\{a_k^n\}_{k,n}^{n,\infty}$ be a two indices sequence of nonnegative numbers. Assume that the following recursive relation holds for some constant C:*

$$a_0^n = 0 \quad \text{and} \quad a_{k+1}^n = [1 + \frac{\ln k}{n} + \frac{C}{n}]a_k^n + \frac{C}{n(n-k)}, \quad k = 0, 1, \cdots, n-1.$$

Then, a_k^n is increasing in k and there exists $\tilde{C} \ge C$, such that for any $\varepsilon > 0$,

$$a_{[(1-\varepsilon)n]}^n \le \frac{\tilde{C}}{\varepsilon n^\varepsilon \ln n} \to 0, \quad \text{as} \quad n \to \infty,$$

where $[x] \le x$ is the largest integer smaller than x.

Proof. That a_k^n is increasing in k is obvious. Now for any $0<\varepsilon<1$, and $k\le[(1-\varepsilon)n]$, we have

$$a_{k+1}^n \le [1+\frac{\ln n}{n}+\frac{C}{n}]a_k^n+\frac{C}{\varepsilon n^2}.$$

This implies that

$$\frac{a_{k+1}^n}{[1+\frac{\ln n}{n}+\frac{C}{n}]^{k+1}} \le \frac{a_k^n}{[1+\frac{\ln n}{n}+\frac{C}{n}]^{k}}+\frac{C}{\varepsilon n^2[1+\frac{\ln n}{n}+\frac{C}{n}]^{k+1}}.$$

Then, for $k\le[(1-\varepsilon)n]$,

$$\frac{a_{k+1}^n}{[1+\frac{\ln n}{n}+\frac{C}{n}]^{k+1}} \le \sum_{j=0}^{k}\frac{C}{\varepsilon n^2[1+\frac{\ln n}{n}+\frac{C}{n}]^{j+1}},$$

and thus

$$a_{k+1}^n \le \frac{C}{\varepsilon n^2}\sum_{j=0}^{k}[1+\frac{\ln n}{n}+\frac{C}{n}]^j \le \frac{C}{\varepsilon n^2}\frac{[1+\frac{\ln n}{n}+\frac{C}{n}]^{k+1}}{\frac{\ln n}{n}+\frac{C}{n}}.$$

Therefore, for n large enough and for some $\tilde{C}\ge C$ which may vary from line to line,

$$\begin{aligned} a_{[(1-\varepsilon)n]}^n &\le \frac{C}{\varepsilon n^2}\frac{[1+\frac{\ln n}{n}+\frac{C}{n}]^{(1-\varepsilon)n}}{\frac{\ln n}{n}+\frac{C}{n}}=\frac{C}{\varepsilon n(C+\ln n)}e^{(1-\varepsilon)n\ln(1+\frac{\ln n}{n}+\frac{C}{n})}\\ &\le \frac{\tilde{C}}{\varepsilon n\ln n}e^{(1-\varepsilon)n(\frac{\ln n}{n}+\frac{\tilde{C}}{n})}=\frac{\tilde{C}}{\varepsilon n\ln n}e^{(1-\varepsilon)(\ln n+\tilde{C})}\le\frac{\tilde{C}}{\varepsilon n^{\varepsilon}\ln n}\to 0,\quad \text{as } n\to\infty. \end{aligned}$$

The proof is complete. ■

Lemma 7.2 *Let ξ and η be two random variables and ψ an increasing (resp. decreasing) function with $\mathbb{E}|\psi(\xi)|<\infty$ and $\mathbb{E}|\psi(\eta)|<\infty$. Assume $\mathbb{P}\{\xi>x\}\le\mathbb{P}\{\eta>x\}$ for any $x\in\mathbb{R}$. Then $\mathbb{E}\{\psi(\xi)\}\le$ (resp. $\ge$) $\mathbb{E}\{\psi(\eta)\}$.*

Proof. We prove only the case in which ψ is increasing. Denote $G_\xi(x)\triangleq\mathbb{P}(\xi>x)$ and $G_\eta(x)\triangleq\mathbb{P}(\eta>x)$, $x\in\mathbb{R}$. Since ψ is increasing, we have

$$\psi(x)G_\xi(x)=E\Big\{\psi(x)\mathbf{1}_{\{\xi>x\}}\Big\}\le E\Big\{\psi(\xi)\mathbf{1}_{\{\xi>x\}}\Big\}\to 0 \quad \text{as} \quad x\to\infty.$$

Similarly,

$$\lim_{x\to\infty}\psi(x)G_\eta(x)=0,\quad \lim_{x\to-\infty}\psi(x)[1-G_\xi(x)]=0,\quad \lim_{x\to-\infty}\psi(x)[1-G_\eta(x)]=0.$$

Then

$$\begin{aligned} \lim_{x\to\infty}\psi(x)[G_\xi(x)-G_\eta(x)] &= 0,\\ \lim_{x\to-\infty}\psi(x)[G_\xi(x)-G_\eta(x)] &= \lim_{x\to-\infty}\psi(x)\Big[[1-G_\eta(x)]-[1-G_\xi(x)]\Big]=0. \end{aligned}$$

Integrating by parts, we get

$$\mathbb{E}\{\psi(\xi)-\psi(\eta)\} = -\int_{-\infty}^{\infty}\psi(t)d[G_\xi(t)-G_\eta(t)] = \int_{-\infty}^{\infty}[G_\xi(t)-G_\eta(t)]d\psi(t).$$

The result follows immediately. ■

[*Proof of Theorem 6.5.*] (i) First, (6.6) follows immediately from the monotonicity assumptions and the construction of the solutions.

(ii) In this step we establish some important estimates. Recall that $\{\tau_k^{1,*}\}, \{\tau_k^{2,*}\}$ are the order statistics of $\{\tau_k^1\}, \{\tau_k^2\}$, respectively; and denote

$$\Delta\tau_k^* \stackrel{\triangle}{=} \tau_k^{1,*} - \tau_k^{2,*}, \quad \tau_0^{1,*} \stackrel{\triangle}{=} \tau_0^{2,*} \stackrel{\triangle}{=} 0.$$

Then clearly, $\Delta X_0^0 = \Delta X_0^i = \Delta Y_0^i = \Delta\tau_0 = 0$.

Fix $k \ge 0$, and assume $t \in [\tau_k^{2,*}, \tau_{k+1}^{2,*}]$. Note that in this interval one has

$$\begin{aligned}
X_t^{0,2} &= X_{\tau_k^{2,*}}^{0,2} + \int_{\tau_k^{2,*}}^t b_0(s, X_s^{0,2}, \hat L^2_{\tau_k^{2,*}})ds + \int_{\tau_k^{2,*}}^t \sigma_0(s)dB_s^0,\\
X_t^{i,2} &= X_{\tau_k^{2,*}}^{i,2} + \int_{\tau_k^{2,*}}^t b_i(s, X_s^{0,2}, X_s^{i,2}, \hat L^2_{\tau_k^{2,*}})ds + \int_{\tau_k^{2,*}}^t \sigma_i(s)dB_s^i,\\
Y_t^{i,2} &= Y_{\tau_k^{2,*}}^{i,2} + \int_{\tau_k^{2,*}}^t \lambda_i(s, X_s^{0,2}, X_s^{i,2}, \hat L^2_{\tau_k^{2,*}})ds.
\end{aligned}$$

Now, for $l = 0, \cdots, k+1$, we denote $\tilde\tau_l \stackrel{\triangle}{=} (\tau_l^{1,*} \vee \tau_k^{2,*}) \wedge \tau_{k+1}^{2,*}$. By (i) we deduce that $\tau_k^{1,*} \ge \tau_k^{2,*}$, for all k, thus we must have

$$\tau_k^{2,*} = \tilde\tau_0 \le \tilde\tau_1 \le \cdots \le \tilde\tau_{k+1} = \tau_{k+1}^{2,*}.$$

Let us now consider the sub-interval $[\tilde\tau_l, \tilde\tau_{l+1}]$, on which

$$\begin{aligned}
X_t^{0,1} &= X_{\tilde\tau_l}^{0,1} + \int_{\tilde\tau_l}^t b_0(s, X_s^{0,1}, \hat L^1_{\tilde\tau_l})ds + \int_{\tilde\tau_l}^t \sigma_0(s)dB_s^0,\\
X_t^{i,1} &= X_{\tilde\tau_l}^{i,1} + \int_{\tilde\tau_i}^t b_i(s, X_s^{0,1}, X_s^{i,1}, \hat L^1_{\tilde\tau_l})ds + \int_{\tilde\tau_l}^t \sigma_i(s)dB_s^i,\\
Y_t^{i,1} &= Y_{\tilde\tau_l}^{i,1} + \int_{\tilde\tau_l}^t \lambda_i(s, X_s^{0,1}, X_s^{i,1}, \hat L^1_{\tilde\tau_l})ds.
\end{aligned}$$

Note that on the set $\{\tilde\tau_l < \tilde\tau_{l+1}\}$, we must have $\tau_l^{1,*} < \tau_{k+1}^{2,*}$ and $\tau_{l+1}^{1,*} > \tau_k^{2,*}$. Assume that for each $j = 1, \cdots, l$, the ordered statistics is attained at $\tau_j^{1,*} = \tau_{i_j}^1$. Then, in light of (i) we have $\tau_{i_j}^2 \le \tau_{i_j}^1 < \tau_{k+1}^{2,*}$, and thus

$$\tau_{i_j}^2 \le \tau_k^{2,*}. \tag{7.1}$$

Then, we have

$$
\begin{aligned}
0 \le \hat L^2_{\tau^{2,*}_k} - \hat L^1_{\tilde\tau_l} &= \frac{1}{n}\sum_{j=1}^{l}[L^{i_j}_{\tau^2_{i_j}} - L^{i_j}_{\tau^1_{i_j}}] + \frac{1}{n}\sum_{1\le i\le k, \tau^{2,*}_i \neq \tau^2_{i_j}} L^i_{\tau^{2,*}_i} \\
&\le \frac{K}{n}\sum_{j=1}^{l}\Delta\tau_{i_j} + \frac{k-l+1}{n}\Lambda_3. \qquad (7.2)
\end{aligned}
$$

By the Lipschitz continuity, we have

$$
\begin{aligned}
d\Delta X^0_t &\le \Big[K\Delta X^0_t + \Lambda_0(\hat L^2_{\tau^{2,*}_k} - \hat L^1_{\tilde\tau_l})\Big]dt; \\
d\Delta X^i_t &\le \Big[K\Delta X^0_t + K\Delta X^i_t + \Lambda_0(\hat L^2_{\tau^{2,*}_k} - \hat L^1_{\tilde\tau_l})\Big]dt; \\
d\Delta Y^i_t &\le \Big[K\Delta X^0_t + K\Delta X^i_t + \Lambda_0(\hat L^2_{\tau^{2,*}_k} - \hat L^1_{\tilde\tau_l})\Big]dt.
\end{aligned}
$$

Then,

$$
d(\Delta X^0_t + \Delta X^i_t + \Delta Y^i_t) \le 3\Big[K(\Delta X^0_t + \Delta X^i_t + \Delta Y^i_t) + \Lambda_0(\hat L^2_{\tau^{2,*}_k} - \hat L^1_{\tilde\tau_l})\Big]dt,
$$

and thus

$$
\begin{aligned}
&e^{-3Kt}[\Delta X^0_t + \Delta X^i_t + \Delta Y^i_t] \\
&\le e^{-3K\tilde\tau_l}[\Delta X^0_{\tilde\tau_l} + \Delta X^i_{\tilde\tau_l} + \Delta Y^i_{\tilde\tau_l}] + \frac{\Lambda_0}{K}[e^{-3K\tilde\tau_l} - e^{-3Kt}][\hat L^2_{\tau^{2,*}_k} - \hat L^1_{\tilde\tau_l}].
\end{aligned}
$$

Let us define

$$
A_0 \triangleq 0, \quad A_t \triangleq A_{\tilde\tau_l} + \frac{\Lambda_0}{K}[e^{-3K\tilde\tau_l} - e^{-3Kt}][\hat L^2_{\tau^{2,*}_k} - \hat L^1_{\tilde\tau_l}], \quad t \in [\tilde\tau_l, \tilde\tau_{l+1}]. \qquad (7.3)
$$

Then, A is increasing, and by induction one can easily see that

$$
e^{-3Kt}[\Delta X^0_t + \Delta X^i_t + \Delta Y^i_t] \le A_t, \quad t \ge 0. \qquad (7.4)
$$

We now estimate A. First note that for any i,

$$
Y^{i,2}_{\tau^2_i} = E_i = Y^{i,1}_{\tau^1_i} = Y^{i,1}_{\tau^2_i} + \int_{\tau^2_i}^{\tau^1_i} \lambda_i(s, X^{0,1}_s, X^{i,1}_s, \hat L^1_s)ds \ge Y^{i,1}_{\tau^2_i} + \Lambda_1\Delta\tau_i.
$$

This, together with the monotonicity properties in (i) (for Y^i), shows that

$$
\Delta\tau_i \le \frac{1}{\Lambda_1}\Delta Y^i_{\tau^2_i} \le \frac{1}{\Lambda_1}e^{3K\tau^2_i}A_{\tau^2_i}. \qquad (7.5)
$$

Assume that the order statistics $\tau^{2,*}$'s are attained at $\tau^{2,*}_1 = \tau^2_{\tilde i_1}, \cdots, \tau^{2,*}_k = \tau^2_{\tilde i_k}$. Then for $j = 1, \cdots, k$, one has

$$
\tau^1_{\tilde i_j} = \tau^2_{\tilde i_j} + \Delta\tau_{\tilde i_j} \le \tau^2_{\tilde i_j} + \frac{1}{\Lambda_1}e^{3K\tau^2_{\tilde i_j}}A_{\tau^2_{\tilde i_j}} \le \tau^{2,*}_k + \frac{1}{\Lambda_1}e^{3K\tau^{2,*}_k}A_{\tau^{2,*}_k}.
$$

Since $\tau_k^{1,*} \le \max\limits_{1\le j\le k} \tau^1_{\tilde{i}_j}$, we obtain

$$\Delta\tau_k^* \le \frac{1}{\Lambda_1} e^{3K\tau_k^{2,*}} A_{\tau_k^{2,*}}. \tag{7.6}$$

Plugging (7.2) and (7.5) into (7.3) and recalling (7.1), we see that

$$\begin{aligned} A_{\tilde{\tau}_{l+1}} &\le A_{\tilde{\tau}_l} + \frac{\Lambda_0}{K}[e^{-3K\tilde{\tau}_l} - e^{-3K\tilde{\tau}_{l+1}}][\frac{K}{n\Lambda_1}\sum_{j=1}^{l} e^{3K\tau_i^2} A_{\tau^2_{i_j}} + \frac{k-l+1}{n}\Lambda_3] \\ &\le A_{\tilde{\tau}_l} + 3\Lambda_0[\tilde{\tau}_{l+1} - \tilde{\tau}_l]\Big[\frac{Kl}{n\Lambda_1} A_{\tau_k^{2,*}} + \frac{k-l+1}{n}\Lambda_3 e^{-3K\tau_k^{2,*}}\Big]. \end{aligned}$$

Summing over $l = 0, \cdots, k$, we obtain

$$\begin{aligned} & A_{\tau^{2,*}_{k+1}} - A_{\tau_k^{2,*}} \\ \le\ & \frac{C}{n} A_{\tau_k^{2,*}} \sum_{l=0}^{k} l[\tilde{\tau}_{l+1} - \tilde{\tau}_l] + \frac{3\Lambda_0\Lambda_3}{n} e^{-3K\tau_k^{2,*}} \sum_{l=0}^{k} (k-l+1)[\tilde{\tau}_{l+1} - \tilde{\tau}_l] \\ =\ & \frac{C}{n} A_{\tau_k^{2,*}} \sum_{l=1}^{k} [\tau^{2,*}_{k+1} - \tilde{\tau}_l] + \frac{3\Lambda_0\Lambda_3}{n} e^{-3K\tau_k^{2,*}} \sum_{l=1}^{k} [\tilde{\tau}_l - \tau_k^{2,*}] \\ \le\ & \frac{Ck}{n} A_{\tau_k^{2,*}}[\tau^{2,*}_{k+1} - \tau_k^{2,*}] + \frac{3\Lambda_0\Lambda_3}{n} e^{-3K\tau_k^{2,*}} \sum_{l=1}^{k} [(\tau_l^{1,*} - \tau_k^{2,*})^+ \wedge (\tau^{2,*}_{k+1} - \tau_k^{2,*})] \\ =\ & \frac{Ck}{n} A_{\tau_k^{2,*}}[\tau^{2,*}_{k+1} - \tau_k^{2,*}] + \frac{3\Lambda_0\Lambda_3}{n} e^{-3K\tau_k^{2,*}} \sum_{l=1}^{k+1} [(\Delta\tau_l^* + \tau_l^{2,*} - \tau_k^{2,*})^+ \wedge (\tau^{2,*}_{k+1} - \tau_k^{2,*})]. \end{aligned}$$

Note that, for any $x, \alpha, \beta > 0$, $(x-\alpha)^+ \wedge \beta \le \frac{\beta}{\alpha+\beta} x$. Then, by (7.6), we deduce from the above

$$\begin{aligned} & A_{\tau^{2,*}_{k+1}} - A_{\tau_k^{2,*}}\Big[1 + \frac{Ck}{n}[\tau^{2,*}_{k+1} - \tau_k^{2,*}]\Big] \\ \le\ & \frac{3\Lambda_0\Lambda_3}{n} e^{-3K\tau_k^{2,*}} \Big[\sum_{l=1}^{k} \Delta\tau_l^* \frac{\tau^{2,*}_{k+1} - \tau_k^{2,*}}{\tau^{2,*}_{k+1} - \tau_k^{2,*} + \tau_k^{2,*} - \tau_l^{2,*}} + \tau^{2,*}_{k+1} - \tau_k^{2,*}\Big] \\ \le\ & \frac{3\Lambda_0\Lambda_3}{n\Lambda_1} \sum_{l=1}^{k} A_{\tau_l^{2,*}} \frac{\tau^{2,*}_{k+1} - \tau_k^{2,*}}{\tau^{2,*}_{k+1} - \tau_k^{2,*} + \tau_k^{2,*} - \tau_l^{2,*}} + \frac{C}{n}[\tau^{2,*}_{k+1} - \tau_k^{2,*}]. \end{aligned} \tag{7.7}$$

For any $t_1 < \cdots < t_k$ and $i_1, \cdots, i_k$, recall (3.5) and (3.8). By Assumption 6.1 (iv) we derive from (3.7) that

$$e^{-(n-k)\Lambda_2 t} \le \mathbb{P}\Big\{\tau^{2,*}_{k+1} > t_k + t \Big| \mathcal{G}_t^k, \tau_l^{2,*} = \tau^2_{i_l} = t_l, l = 1, \cdots, k\Big\} \le e^{-(n-k)\Lambda_1 t}. \tag{7.8}$$

By (7.3), one can easily check that

$$A_{\tau_j^{2,*}} \mathbf{1}_{\{\tau_l^{2,*} = \tau^2_{i_l} = t_l, l=1,\cdots,k\}} \text{ is } \mathcal{G}_t^k \bigvee \Big(\bigvee_{l=1}^{k} \sigma(\tau_l^{2,*} = \tau^2_{i_l} = t_l)\Big)\text{-measurable}, \quad j = 1, \cdots, k.$$

Then (7.8) implies that

$$e^{-(n-k)\Lambda_2 t} \leq \mathbb{P}\Big\{\tau_{k+1}^{2,*} - \tau_k^{2,*} > t \Big| \sigma\Big(\tau_l^{2,*}, A_{\tau_l^{2,*}}, l = 1, \cdots, k\Big)\Big\} \leq e^{-(n-k)\Lambda_1 t}. \tag{7.9}$$

Next, using the second inequality in (7.9) and applying Lemma 7.2 (by setting $\xi = \tau_{k+1}^{2,*} - \tau_k^{2,*}$, $\eta \sim \exp\{(n-k)\Lambda_1\}$, and $\psi(x) = x$) we obtain

$$\mathbb{E}\Big\{\tau_{k+1}^{2,*} - \tau_k^{2,*} \Big| \sigma\Big(\tau_l^{2,*}, A_{\tau_l^{2,*}}, l = 1, \cdots, k\Big)\Big\} \leq \frac{1}{(n-k)\Lambda_1}.$$

Since $\frac{x}{a+x}$ is concave in x, applying Jensen's inequality we get

$$\begin{aligned} & \mathbb{E}\Big\{\frac{\tau_{k+1}^{2,*} - \tau_k^{2,*}}{\tau_{k+1}^{2,*} - \tau_k^{2,*} + \tau_k^{2,*} - \tau_l^{2,*}} \Big| \sigma\Big(\tau_l^{2,*}, A_{\tau_l^{2,*}}, l = 1, \cdots, k\Big)\Big\} \\ \leq \quad & \frac{\frac{1}{(n-k)\Lambda_1}}{\frac{1}{(n-k)\Lambda_1} + \tau_k^{2,*} - \tau_l^{2,*}} = \frac{1}{1 + (n-k)\Lambda_1(\tau_k^{2,*} - \tau_l^{2,*})} \end{aligned}$$

Let $\tilde{E}_1, \cdots, \tilde{E}_k$ be i.i.d. exponential random variables with rate 1 and independent of $\mathbb{F}$. Since $\frac{1}{1+ax}$ is decreasing in x for $a > 0$,we can apply Lemma 7.2 repeatedly by using the first inequality in (7.9) and setting $\xi \sim \exp\{(n-k)\Lambda_2\}$, $\eta = \tau_{k+1}^{2,*} - \tau_k^{2,*}$, and $\psi(x) = \frac{1}{1+ax}$ to get

$$\begin{aligned} & \mathbb{E}\Big\{\frac{1}{1 + (n-k)\Lambda_1(\tau_k^{2,*} - \tau_l^{2,*})} \Big| \sigma\Big(\tau_j^{2,*}, A_{\tau_j^{2,*}}, j = 1, \cdots, l\Big)\Big\} \\ \leq \quad & \mathbb{E}\Big\{\frac{1}{1 + (n-k)\Lambda_1 \sum_{j=l}^{k-1} \frac{\tilde{E}_j}{(n-j)\Lambda_2}}\Big\} \leq \mathbb{E}\Big\{\frac{1}{1 + \frac{(n-k)\Lambda_1}{(n-l)\Lambda_2} \sum_{j=l}^{k-1} \tilde{E}_j}\Big\}. \end{aligned}$$

For $k - l \geq 1$, noticing that $\sum_{j=l}^{k-1} \tilde{E}_j$ has exponential distribution with rate $k - l$, we have

$$\mathbb{E}\Big\{\frac{1}{1 + \frac{(n-k)\Lambda_1}{(n-l)\Lambda_2} \sum_{j=l}^{k-1} \tilde{E}_j}\Big\} \leq \frac{(n-l)\Lambda_2}{(n-k)\Lambda_1} \mathbb{E}\Big\{\frac{1}{\sum_{j=l}^{k-1} \tilde{E}_j}\Big\} = \frac{(n-l)\Lambda_2}{(n-k)\Lambda_1} \frac{1}{k-l}.$$

Plug all these into (7.7) and denote

$$a_k \overset{\triangle}{=} \mathbb{E}\{A_{\tau_k^{2,*}}\}, \quad a_k^* \overset{\triangle}{=} \max_{0 \leq i \leq k} a_i.$$

Then, we get

$$\begin{aligned} a_{k+1} \quad \leq \quad & \Big[1 + \frac{Ck}{n(n-k)}\Big] a_k + \frac{3\Lambda_0\Lambda_3}{n\Lambda_1}\Big[\sum_{l=1}^{k-1} a_l \frac{(n-l)\Lambda_2}{(n-k)\Lambda_1} \frac{1}{k-l} + a_k\Big] + \frac{C}{n(n-k)} \\ \leq \quad & a_k^*\Big[1 + \frac{Ck}{n(n-k)} + \frac{C}{n} + \frac{3\Lambda_0\Lambda_2\Lambda_3}{n\Lambda_1^2} \sum_{l=1}^{k-1} [\frac{1}{k-l} + \frac{1}{n-k}]\Big] + \frac{Ck}{n(n-k)} \\ \leq \quad & a_k^*\Big[1 + \frac{Ck}{n(n-k)} + \frac{C}{n} + \frac{3\Lambda_0\Lambda_2\Lambda_3}{\Lambda_1^2} \frac{\ln k}{n}\Big] + \frac{Ck}{n(n-k)}. \end{aligned}$$

For $k \le (1-\varepsilon)n$, thanks to Assumption 6.1 (v), we have

$$a_{k+1} \le \Big[1 + \frac{\ln k}{n} + \frac{C_\varepsilon}{n}\Big] a_k^* + \frac{C_\varepsilon}{n^2}.$$

This implies that

$$a_{k+1}^* \le \Big[1 + \frac{\ln k}{n} + \frac{C_\varepsilon}{n}\Big] a_k^* + \frac{C_\varepsilon}{n^2}, \quad k \le (1-\varepsilon)n.$$

Since $a_0 = 0$, applying Lemma 7.1 we obtain

$$a_{(1-\varepsilon)n}^* \le \frac{C_\varepsilon}{n^\varepsilon \ln n} \to 0, \qquad \text{as} \ \ n \to \infty,$$

for any $\varepsilon > 0$, proving (ii).

(iii) We now prove (6.7). Recall that A is increasing. For any $\varepsilon > 0$, by (7.4) we have

$$\begin{aligned}
\mathbb{E}\{\Delta X_t\} &= \mathbb{E}\Big\{\Delta X_t[\mathbf{1}_{\{\tau^{2,*}_{(1-\varepsilon)n} \ge t\}} + \mathbf{1}_{\{\tau^{2,*}_{(1-\varepsilon)n} < t\}}]\Big\} \\
&\le e^{3Kt}\mathbb{E}\Big\{A_t \mathbf{1}_{\{\tau^{2,*}_{(1-\varepsilon)n} \ge t\}}\Big\} + \mathbb{E}\Big\{\Delta X_t \mathbf{1}_{\{\tau^{2,*}_{(1-\varepsilon)n} < t\}}\Big\} \\
&\le C_T \mathbb{E}\{A_{\tau^{2,*}_{(1-\varepsilon)n}}\} + \mathbb{E}\{\Delta X_t \mathbf{1}_{\{\tau^{2,*}_{(1-\varepsilon)n} < t\}}\} \\
&\le \frac{C_{\varepsilon,T}}{n^\varepsilon \ln n} + \mathbb{E}^{\frac{1}{2}}\{|\Delta X_t|^2\}\mathbb{P}^{\frac{1}{2}}\{\tau^{2,*}_{(1-\varepsilon)n} < t\} \\
&\le \frac{C_{\varepsilon,T}}{n^\varepsilon \ln n} + C_T \mathbb{P}^{\frac{1}{2}}\{\tau^{2,*}_{(1-\varepsilon)n} < t\}.
\end{aligned}$$

However, from (7.9) and applying Lemma 7.2 we see that

$$\mathbb{P}\{\tau^{2,*}_{(1-\varepsilon)n} < t\} \le \mathbb{P}\Big\{\hat\tau^{2,*}_{(1-\varepsilon)n} < t\Big\}, \ \ \text{where} \ \ \hat\tau^{2,*}_{(1-\varepsilon)n} \stackrel{\triangle}{=} \sum_{i=1}^{(1-\varepsilon)n} \frac{\tilde E_i}{(n-i)\Lambda_2}.$$

Observe that

$$\begin{aligned}
\mathbb{E}\Big\{\hat\tau^{2,*}_{(1-\varepsilon)n}\Big\} &= \sum_{i=1}^{(1-\varepsilon)n} \frac{1}{(n-i)\Lambda_2} \ge \frac{1}{2\Lambda_2}\ln\frac{1}{\varepsilon}; \\
Var\Big\{\hat\tau^{2,*}_{(1-\varepsilon)n}\Big\} &= \sum_{i=1}^{(1-\varepsilon)n} \frac{1}{(n-i)^2\Lambda_2^2} \le \frac{2(1-\varepsilon)}{\varepsilon\Lambda_2^2 n}.
\end{aligned}$$

Choosing $\varepsilon \stackrel{\triangle}{=} \varepsilon_T > 0$ so that $\ln\frac{1}{\varepsilon} = 2\Lambda_2(T+1)$, we then have

$$\begin{aligned}
&\mathbb{P}\{\tau^{2,*}_{(1-\varepsilon)n} < t\} \le \mathbb{P}\Big\{\hat\tau^{2,*}_{(1-\varepsilon)n} < t\Big\} \le \mathbb{P}\Big\{\hat\tau^{2,*}_{(1-\varepsilon)n} - \mathbb{E}\{\hat\tau^{2,*}_{(1-\varepsilon)n}\} < t - \frac{1}{2\Lambda_2}\ln\frac{1}{\varepsilon}\Big\} \\
\le\ & \mathbb{P}\Big\{\hat\tau^{2,*}_{(1-\varepsilon)n} - \mathbb{E}\{\hat\tau^{2,*}_{(1-\varepsilon)n}\} < -1\Big\} \le Var(\hat\tau^{2,*}_{(1-\varepsilon)n}) \le \frac{2(1-\varepsilon_T)}{\varepsilon_T\Lambda_2^2 n} = \frac{C_T}{n}.
\end{aligned}$$

Thus,

$$\mathbb{E}\{\Delta X_t\} \le \frac{C_{\varepsilon,T}}{n^\varepsilon \ln n} + \frac{C_T}{\sqrt{n}}.$$

This proves (6.7) immediately, hence the theorem. ■

References

[1] Bielecki, T. R.; Rutkowski, M. (2002) *Credit risk: modelling, valuation and hedging.* Springer Finance, Springer-Verlag, Berlin.

[2] Collin-Dufresne, P., Goldstein, R. and Helwege, J. (2003) Is credit event risk priced? Modeling contagion via updating of beliefs. Working paper, Univ. California Berkeley.

[3] Collin-Dufresne, P., Goldstein, R. and Hugonnier, J. (2004) A general formula for valuing defaultable securities. *Econometrica* 72, 1377–1407.

[4] Dai Pra, P., Runggaldier, W. J., Sartori, E., and Tolotti, M. (2009) Large portfolio losses: A dynamic contagion model. *Ann. Appl. Probab.* 19, 347–394.

[5] Davis, M. and Lo, V. (2001) Infectious default. *Quantitative Finance* 1, 382–387.

[6] Dembo, A., Deuschel, J.D. and Duffie, D. (2004) Large portfolio losses. *Finance & Stochastics* 8, 3–16.

[7] Duffie, D. and Singleton, K. (2003) *Credit Risk: Pricing, Measurement, and Management.* Princeton University Press.

[8] Filipovic, D., Overbeck, L., and Schmidt, T., (2011), Dynamic CDO Term Structure Modelling, to appear in Mathematical Finance.

[9] Frey, R. and Backhaus, J. (2006), Credit derivatives in models with interacting default intensities: A Markovian approach. Preprint, Dept. of Mathematics, Universitat Leipzig.

[10] Frey, R. and Backhaus, J. (2007), Dynamic hedging of syntentic CDO tranches with spread risk and default contagion. Preprint, Dept. of Mathematics, Universitat Leipzig.

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

[12] Giesecke, K., Spiliopoulos, K., and Sowers, R. (2010), Default Clustering in Large Portfolios: Typical and Atypicla Events, Preprint.

[13] Giesecke, K. and Weber, S. (2005), Cyclical correlations, credit contagion and portfolio losses. *J. of Banking and Finance* 28, 3009–3036.

[14] Giesecke, K. and Weber, S. (2006), Credit contagion and aggregate losses. *J. Econom. Dynam. Control* 30, 741–767.

[15] Horst, U. (2007), Stochastic cascades, contagion and large portfolio losses. *Journal of Economic Behaviour and Organization* 63 25–54.

[16] Jarrow, R. A. and Yu, F. (2001) Counterparty risk and the pricing of defaultable securities. *Journal of Finance* 53, 2225–2243.

[17] Jeanblanc, M., Yor, M., and Chesney, M., (2009) *Mathematical methods for financial markets.* Springer Finance, Springer-Verlag London, Ltd., London.

[18] Jeanblanc, M. and Song. S., (2011a), An explicit model of default time with given survival probability. Preprint.

[19] Jeanblanc, M. and Song. S., (2011b), Random times with given survival probability and their $\mathbb{F}$-martingale decomposition formula. Preprint.

[20] Lando, D. (2004) *Credit Risk Modeling: Theory and Applications.* Princeton University Press.

[21] Mc Neil, A., Frey, R. and Embrechts, P. (2005) *Quantitative Risk Management: Concepts, Techniques, and Tools.* Princeton University Press.

[22] Yu, F. (2007) Correlated Defaults in Intensity-Based Models, *Mathematical Finance* 17, 155–173.