Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 119 (2009) 2913–2944
stochastic processes and their applications
www.elsevier.com/locate/spa

Heterogeneous credit portfolios and the dynamics of the aggregate losses

Paolo Dai Pra[a], Marco Tolotti[b,*]

[a] *Dipartimento di Matematica Pura ed Applicata, Università di Padova 63, Via Trieste, I - 35121 - Padova, Italy*
[b] *Dipartimento di Matematica Applicata, Università Ca' Foscari di Venezia, Dorsoduro 3825/e, I - 30123 - Venice, Italy*

Received 24 June 2008; received in revised form 4 March 2009; accepted 12 March 2009
Available online 27 March 2009

Abstract

We study the impact of contagion in a network of firms facing credit risk. We describe an intensity based model where the homogeneity assumption is broken by introducing a *random environment* that makes it possible to take into account the idiosyncratic characteristics of the firms. We shall see that our model goes behind the identification of groups of firms that can be considered basically exchangeable. Despite this heterogeneity assumption our model has the advantage of being totally tractable. The aim is to quantify the losses that a bank may suffer in a large credit portfolio. Relying on a large deviation principle on the trajectory space of the process, we state a suitable law of large numbers and a central limit theorem useful for studying large portfolio losses. Simulation results are provided as well as applications to portfolio loss distribution analysis.

MSC: 60K35; 91B70

Keywords: Central limit theorems in Banach spaces; Credit contagion; Intensity based models; Large deviations; Large portfolio losses; Random environment

0. Introduction

During the last few years the challenging issue of describing the *dynamics* of the loss process connected with portfolios of many obligors has received more and more attention. Applications can be found both for management purposes (see [1]) and in the literature dealing

* Corresponding author. Tel.: +39 0412346928.
E-mail addresses: daipra@math.unipd.it (P. Dai Pra), tolotti@unive.it (M. Tolotti).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2009.03.006

with pricing and hedging of risky derivatives such as CDOs (Collateralized Debt Obligations). For a discussion of this framework see [2,3].

When dealing with *portfolio* losses, the modeling of the dependence structure among the obligors becomes crucial. One standard procedure is to directly specify the *intensity of default* of the single obligors belonging to the portfolio in order to infer the dynamics of the global system and thus the distribution of the aggregate losses. In the context of reduced form models a rather general framework is the conditionally Markov modeling approach of Frey and Backhaus (see [4,5]). One drawback of the intensity based models is the difficulty in managing large heterogeneous portfolios because of the presence of many obligors with different specifications. In this case it is common practice to make homogeneity assumptions in order to reduce the complexity of the problem. A typical approach is to divide the portfolio into groups where the obligors may be considered exchangeable.

In this paper we describe an intensity based model where the homogeneity assumption is broken by introducing a *random environment* that makes it possible to take into account the idiosyncratic characteristics of the firms. We shall see that our model goes behind the identification of groups of firms that can be considered basically exchangeable. Despite this heterogeneity assumption our model has the advantage of being totally tractable.

The goal is to describe the evolution of the losses for a large portfolio where *heterogeneity* and *direct contagion* among the firms are taken into account. We denote by $L^N(t)$ the random variable describing the losses at time $t \in [0, T]$ for a portfolio of size N. Our approach works as follows. First we study the $N \to \infty$ limiting distributions on the path space of some aggregate variables useful for characterizing the evolution of $L^N(t)$ for $t \in [0, T]$. To this effect we shall derive an appropriate *law of large numbers* based on a large deviation principle in order to describe a limiting behavior that can be considered as a asymptotic regime with infinitely many firms. Finally, we study the *finite volume approximations* (for finite but large N) of the limiting distribution via a suitable version of the *central limit theorem* that describes the fluctuations around this limit. In most cases, these dynamical fluctuation theorems are proved by the method of weak convergence of processes; this approach has been widely applied to models close in spirit to this work. We quote [4,6] for applications to finance. The effectiveness of those methods for heterogeneous models is, however, unclear. We follow here a different approach, which allows us to prove a central limit theorem directly in the underlying trajectory space. This approach is based on a general central limit theorem in [7]. Although various applications of this theorem to fluctuations of Markov processes can be found in the literature (see e.g. [8,9]), to our knowledge this is the first application to a non-reversible Markov process.

In the risk management context, our model may be useful for the management of large portfolios, in the spirit of other models proposed in [10] or in [6]. It has been remarked that in many real world applications defaults are rather rare events, with the result that, for instance, the fraction of defaulted firms is close to zero and a normal approximation is not meaningful. Our models and results are only concerned with time scales for which a proper fraction of the portfolio is likely to be affected by the defaults.

We believe that our paper may be considered as an original contribution in the modeling of portfolio loss dynamics that accounts for both heterogeneity and contagion. On the other hand, to our knowledge, this is the first attempt to apply large deviations and normal fluctuation theory *on path spaces* (that is, in a dynamic fashion) for finance or credit management purposes, except for what is contained in [6]. Moreover, the model studied in [6] did not require the development of large deviations in Banach spaces as we are forced to use here. For a survey on more standard *non-dynamic* large deviation methods applied to finance and credit risk see [11].

The models that we propose in this paper are the simplest heterogeneous models describing systems comprised by many defaultable components, whose defaults are positively correlated via an interaction of mean-field type, i.e. with no geometric structure. Although we have been inspired by financial applications, we believe that the basic principles should apply to other contexts.

The outline of this paper is as follows. In Section 1 we illustrate the model and the main theorems. In Section 2 we apply these results to the large portfolio losses analysis. Some examples with explicit computations and simulations are also provided. In Section 3 we draw some conclusions on the proposed methods. Appendix is devoted to the proofs of the three main theorems stated in Section 1.

1. Model and main results

Consider a network of N defaultable firms, whose states are denoted by $y_1, y_2, \ldots, y_N$, $y_i \in \{0, 1\}$. The event $\{y_i = 1\}$ means that the i-th firm has defaulted. The values $y_1, y_2, \ldots, y_N$ give rise to an *aggregate variable* m_N which indicates the global state of the network:

$$m_N := \frac{1}{N} \sum_{i=1}^{N} \alpha_i y_i,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_N$ are given nonnegative numbers. α_i can be interpreted as the impact that the default of the i-th firm has on the aggregate variable m_N. In order to model contagion, we assume that the instantaneous rate of default of the i-th firm is an increasing function of m_N. More specifically, we assume that the rate of default of the i-th firm is given by

$$I_{\{y_i=0\}} \mathrm{e}^{\beta_i m_N - \gamma_i},$$

where I_A is the indicator function of the set A, and $\beta_i \geq 0$, $\gamma_i \in \mathbb{R}$ are given constants. β_i represents the sensitivity of the i-th firm to variations of the aggregate variable m_N, while γ_i can be interpreted as the "robustness" of the i-th firm: a large value of γ_i means that the i-th firm is very unlikely to default within a given time.

Thus, for any *fixed* values of $\underline{\alpha} := (\alpha_1, \alpha_2, \ldots, \alpha_N)$, $\underline{\beta} := (\beta_1, \beta_2, \ldots, \beta_N)$ and $\underline{\gamma} := (\gamma_1, \gamma_2, \ldots, \gamma_N)$, the variable $\underline{y} := (y_1, y_2, \ldots, y_N)$ evolves as a Markov chain in continuous time, with infinitesimal generator given by

$$\mathcal{L}_{\underline{\alpha},\underline{\beta},\underline{\gamma}} f(\underline{y}) := \sum_{i=1}^{N} I_{\{y_i=0\}} \mathrm{e}^{\beta_i m_N - \gamma_i} [f(\underline{y}^i) - f(\underline{y})], \tag{1.1}$$

where $\underline{y}^i$ denotes the *configuration* obtained from $\underline{y}$ by changing y_i from 0 to 1. We assume the system to start at time $t = 0$ from the configuration $\underline{y}(0) = (0, 0, \ldots, 0)$. The evolution randomly drives the network towards the trap state $(1, 1 \ldots, 1)$, which is reached in finite time.

From now on we denote by $\lambda_i := (\alpha_i, \beta_i, \gamma_i)$ the triple of the parameters corresponding to the i-th firm. The λ_is model the heterogeneity of the system. We consider here the point of view of *disordered models*, i.e. we assume $\lambda_1, \lambda_2, \ldots, \lambda_N$ to be i.i.d. random variables, with a given law μ. In order to avoid inessential difficulties, the law μ is assumed to have compact support in $\mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}$. Note that, for a given i, the random variables $\alpha_i, \beta_i, \gamma_i$ are *not* assumed to be independent. Sometimes, the vector $\underline{\lambda} := (\lambda_1, \lambda_2, \ldots, \lambda_N)$ will be referred to as *random environment*.

Consider a time $T > 0$, and denote by $\underline{y}[0,T] = (\underline{y}(t))_{t\in[0,T]}$ the trajectory described by the configuration under the stochastic evolution (1.1). Each component $y_i[0,T]$ is either identically 0 or it flips from 0 to 1 at the *default time*

$$\tau_i := \inf\{t > 0 : y_i(t) = 1\}. \tag{1.2}$$

By convention, we set $y_i(\tau_i) = 1$. This set of $\{0,1\}$-valued trajectories is denoted by $\mathcal{D}[0,T]$. Each trajectory in $\mathcal{D}[0,T]$ can be identified with its default time (which is set to be equal to T if there is no default); thus $\mathcal{D}[0,T]$ inherits the topology induced by the usual topology on $\mathbb{R}$ for the default time. Equivalently, the topology on $\mathcal{D}[0,T]$ is the one induced by the Skorohod topology on the set of $\mathbb{R}$-valued functions which are right-continuous and admit a limit from the left at any point of $[0,T]$ (see e.g. [12]).

In this paper we are interested in the asymptotic behavior, as $N \to +\infty$, of the empirical averages of the form

$$\frac{1}{N}\sum_{i=1}^{N} f(y_i[0,T]) =: \int f\,\mathrm{d}\rho_N(\underline{y}[0,T]),$$

where $f : \mathcal{D}[0,T] \to \mathbb{R}$ is a Borel measurable function, and

$$\rho_N(\underline{y}[0,T]) := \frac{1}{N}\sum_{i=1}^{N} \delta_{y_i[0,T]}$$

is called the *empirical measure*. More generally, we shall consider the empirical measure

$$\rho_N(\underline{y}[0,T],\underline{\lambda}) := \frac{1}{N}\sum_{i=1}^{N} \delta_{y_i[0,T],\lambda_i},$$

which is a random measure on $\mathcal{D}[0,T] \times \mathbb{R}^3$. Note that $\rho_N(\underline{y}[0,T])$ is the marginal of $\rho_N(\underline{y}[0,T],\underline{\lambda})$ on $\mathcal{D}[0,T]$.

In what follows, we denote by $\mathcal{M}_1$ the set of probability measures on $\mathcal{D}[0,T] \times \mathrm{Supp}(\mu)$, while $\mathcal{M}$ will denote the set of signed measures on $\mathcal{D}[0,T] \times \mathbb{R}^3$. Both sets are provided with the weak topology.

For $y[0,T] \in \mathcal{D}[0,T]$ with $y(T) = 1$, we set

$$\tau(y[0,T]) := \inf\{t > 0 : y(t) = 1\}.$$

For $Q \in \mathcal{M}$ we define

$$\begin{aligned} F(Q) := \int Q(\mathrm{d}y[0.T],\mathrm{d}\lambda)\Bigg\{ &\int_0^T (1-y(t))\left(1-\mathrm{e}^{-\gamma}\mathrm{e}^{\beta\int Q(\mathrm{d}\eta[0,T],\mathrm{d}\lambda')\alpha'\eta(t)}\right)\mathrm{d}t \\ &+ y(T)\left[-\gamma + \left(\beta\int Q(\mathrm{d}\eta[0,T],\mathrm{d}\lambda')\alpha'\eta(t^-)\right)\Bigg|_{t=\tau(y[0,T])}\right]\Bigg\}, \end{aligned} \tag{1.3}$$

where $\lambda = (\alpha,\beta,\gamma)$ and $\lambda' = (\alpha',\beta',\gamma')$. For a fixed $\underline{\lambda}$, the infinitesimal generator (1.1), together with the initial condition $\underline{y}(0) = (0,0,\ldots,0)$, induces a probability $P_N^{\underline{\lambda}}$ on $\mathcal{D}^N[0,T]$. We think of $P_N^{\underline{\lambda}}$ as the *conditional law* of the process given the random environment. We denote by

$$P_N(\mathrm{d}\underline{y}[0,T],\mathrm{d}\underline{\lambda}) := P_N^{\underline{\lambda}}(\mathrm{d}\underline{y}[0,T]) \otimes \mu^{\otimes N}(\mathrm{d}\underline{\lambda})$$

the joint law of the process and the environment. The distribution of $\rho_N(\underline{y}[0,T],\underline{\lambda})$ under P_N will be denoted by $P_N \circ \rho_N^{-1}$.

A special case is when all components of $\underline{\lambda}$ are zero. In this case each firm defaults with rate 1, independently of the others. We denote by W the law on $\mathcal{D}[0,T]$ of this process.

In what follows, for $Q_1, Q_2 \in \mathcal{M}_1$, we denote by

$$H(Q_1|Q_2) := \begin{cases} \int \mathrm{d}Q_2 \left(\frac{\mathrm{d}Q_1}{\mathrm{d}Q_2}\log\frac{\mathrm{d}Q_1}{\mathrm{d}Q_2}\right) & \text{if } Q_1 \ll Q_2 \text{ and } \frac{\mathrm{d}Q_1}{\mathrm{d}Q_2}\log\frac{\mathrm{d}Q_1}{\mathrm{d}Q_2} \in L^1(Q_2) \\ +\infty & \text{otherwise} \end{cases}$$

the relative entropy of Q_1 with respect to Q_2.

Theorem 1. *The sequence $P_N \circ \rho_N^{-1}$ of elements of $\mathcal{M}_1$ satisfies a Large Deviation Principle (LDP) with good rate function*

$$I(Q) := H(Q|W \otimes \mu) - F(Q).$$

The proof of Theorem 1, as well as of the other results stated in this section, is postponed to the Appendix.

We recall that the above statement means that, for each Borel subset A of $\mathcal{M}_1$,

$$\begin{aligned} -\inf_{Q\in \mathring{A}} I(Q) &\le \liminf_{N\to+\infty}\frac{1}{N}\log P_N\circ\rho_N^{-1}(A) \\ &\le \limsup_{N\to+\infty}\frac{1}{N}\log P_N\circ\rho_N^{-1}(A) \le -\inf_{Q\in\overline{A}} I(Q), \end{aligned}$$

where $\mathring{A}$ and $\overline{A}$ denote the interior and the closure of A respectively; moreover the function $I(\cdot)$ is nonnegative, lower semicontinuous, and the level sets $\{Q : I(Q) \le l\}$ are compact, for each $l > 0$.

Theorem 2. *The equation $I(Q) = 0$ has a unique solution Q_*, that can be identified as follows. Consider the nonlinear integro-differential equation*

$$\begin{cases} \frac{\partial}{\partial t} q_t(\lambda) = \mathrm{e}^{-\gamma}\exp\left[\beta\int\mu(\mathrm{d}\lambda')\alpha' q_t(\lambda')\right](1-q_t(\lambda)) \\ q_0(\lambda) \equiv 0 \end{cases} \tag{1.4}$$

for a real-valued $q_t(\lambda)$, $t \ge 0$, $\lambda = (\alpha,\beta,\gamma) \in \mathbb{R}^3$. This equation has a unique solution $0 \le q_t(\lambda) \le 1$. For every λ fixed, consider the Markov chain on $\{0,1\}$ with time-dependent infinitesimal generator

$$L_t^\lambda f(s) := c_t^\lambda(s)[f(1-s) - f(s)], \tag{1.5}$$

where

$$c_t^\lambda(s) := (1-s)\mathrm{e}^{-\gamma}\exp\left[\beta\int\mu(\mathrm{d}\lambda')\alpha' q_t(\lambda')\right]$$

and starting from $s = 0$ *(note that this process jumps only once, from* 0 *to*1*, and it is then trapped in*1*). Let* Q_*^λ *be the law of this process on* $\mathcal{D}[0, T]$*. Then*

$$Q_* = Q_*^\lambda \otimes \mu.$$

Moreover

$$q_t(\lambda) = Q_*^\lambda(y(t) = 1).$$

Theorems 1 and 2 have a simple consequence. Let U be an open neighborhood of Q^* in $\mathcal{M}_1$. By Theorem 2, lower semicontinuity of $I(\cdot)$ and compactness of its level sets, a standard argument shows that $k(U) := \inf_{Q \notin \overline{U}} I(Q) > 0$. By the upper bound in Theorem 1 there exists $C > 0$ such that

$$P_N(\rho_N \notin U) \le C\mathrm{e}^{-Nk(U)},$$

thus giving convergence to zero with exponential rate. We summarize this fact in the following *law of large numbers*.

Corollary 1. *Let* $d(\cdot, \cdot)$ *be any metric that induces the weak topology on* $\mathcal{M}_1$*. Then for every* $\epsilon > 0$*, the probability*

$$P_N(d(\rho_N, Q_*) \ge \epsilon)$$

converges to zero with exponential rate in N*.*

The next result concerns the fluctuations of ρ_N about Q_*, which has the form of a central limit theorem. In most cases, these dynamical fluctuation theorems are proved by the method of weak convergence of processes. A typical tool in this context is Theorem 1.6.1 in [12]; it has been widely applied to models close in spirit to this work; see for instance [13,6]. The effectiveness of those methods for heterogeneous models is, however, unclear. The main point is that, via Theorem 1.6.1 in [12], one obtains the dynamics of the fluctuation process, which is infinite dimensional; to get a *computable* expression for the asymptotic variance of a given function of the trajectory may be not feasible.

We follow here a different approach, which allows us to prove a central limit theorem directly in the space $\mathcal{M}$. This is inspired by the seminal work of Bolthausen [7], and it has been carried out in a context similar to ours in [8,9]. The main difference here is that the underlying stochastic dynamics P_N^λ are not reversible; the related difficulties have forced us to introduce a further assumption, that we call the *reciprocity condition*:

(R) There exists a deterministic $b > 0$ such that for all $i \ge 1$ the identity $\beta_i = b\alpha_i$ holds almost surely.

In economical terms, this means that the sensitivity of a firm to variation of the aggregate variable m_N is proportional to the impact that the default of that firm has in the network. In other words, the interaction between firms is *symmetric*: if the i-th firm strongly interacts with the network (i.e. α_i is large) then it has a large influence on the network but, symmetrically, it is also strongly influenced by the state of the network. This assumption may be reasonable in many situations.

From now on, whenever condition (R) is assumed, (λ_i) will denote the pair (α_i, γ_i); accordingly, $\mathcal{M}_1$ and $\mathcal{M}$ will be spaces of measures on $\mathcal{D}[0, T] \times \mathrm{Supp}(\mu)$, where $\mathrm{Supp}(\mu) \subset \mathbb{R}^+ \times \mathbb{R}$.

In order to state our central limit theorem, we need to introduce some notation. Let $\mathcal{M}_0$ be the subset of $\mathcal{M}$ comprised by the signed measures with zero total mass. Let ν_* be the law, induced by Q_*, of the $\mathcal{M}_0$-valued random variable $\delta_{(y[0,T],\lambda)} - Q_*$. We then denote by $\mathcal{C}_b$ the space of bounded, continuous, real-valued functions on $\mathcal{D}[0,T] \times \mathbb{R}^2$. For $\phi \in \mathcal{C}_b$, define $\hat{\phi} \in \mathcal{M}_0$ by

$$\hat{\phi}(A) := \int \nu_*(\mathrm{d}R) \left(R(A) \int \phi \mathrm{d}R \right), \tag{1.6}$$

for $A \subseteq \mathcal{D}[0,T] \times \mathbb{R}^2$ measurable.

Theorem 3. *Assume condition (R). Let $\phi_1, \phi_2, \ldots, \phi_n \in \mathcal{C}_b$. Then the P_N-law of the random vector*

$$\sqrt{N} \left(\int \phi_i \mathrm{d}\rho_N - \int \phi_i \mathrm{d}Q_* \right)_{i=1}^n$$

converges weakly as $N \to +\infty$ to an n-dimensional Gaussian probability measure with zero mean and covariance matrix $C = (C_{i,j})_{i,j=1}^n$ given by

$$C_{i,j} := \int (\phi_i - \phi_i^*)(\phi_j - \phi_j^*) \mathrm{d}Q_* - \mathrm{D}^2 F(Q_*)[\hat{\phi}_i, \hat{\phi}_j],$$

where $\phi_i^* := \int \phi_i \mathrm{d}Q_*$ and $\mathrm{D}^2 F(Q_*)[\hat{\phi}_i, \hat{\phi}_j]$ is the second Fréchet directional derivative of F at Q_* in the directions $\hat{\phi}_i, \hat{\phi}_j$:

$$\mathrm{D}F(Q_*)[\hat{\phi}_i] = \lim_{h \to 0} \frac{F(Q_* + h\hat{\phi}_i) - F(Q_*)}{h}$$

$$\mathrm{D}^2 F(Q_*)[\hat{\phi}_i, \hat{\phi}_j] = \lim_{h \to 0} \frac{\mathrm{D}F(Q_* + h\hat{\phi}_j)[\hat{\phi}_i] - \mathrm{D}F(Q_*)[\hat{\phi}_i]}{h}$$

(all these limits will be shown to exist).

Moreover the diagonal terms of the covariance $C_{i,j}$ can be written as

$$C_{i,i} = E^{Q_*} \left[\left((\phi_i - \phi_i^*) - \beta \int_0^T (1 - y(s)) \mathrm{Cov}_{Q_*}(\alpha y(s), \phi_i) \mathrm{d}M(s) \right)^2 \right], \tag{1.7}$$

where

$$M(t) := \mathbf{1}_{\{\tau \le t\}} - \int_0^t (1 - y(s)) \mathrm{e}^{-\gamma + \beta \int \alpha' \eta(s) Q_*(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')} \mathrm{d}s \tag{1.8}$$

is the compensated $(Q_*, \mathcal{F})$-martingale associated with the jump process of $y[0,T]$.

2. Applications to the portfolio analysis

2.1. Computation of large portfolio losses

We are now going to state a definition of *portfolio losses*. When speaking of portfolio losses, we mean the losses that a financial institution may suffer in a credit portfolio due to the default events. Many specifications may be chosen to this end. Some general rules are now stated. A rather general modeling framework is to consider the total loss that a bank may suffer due to a risky portfolio at time t as a random variable defined by $L^N(t) = \sum_i L_i(t)$, where $L_i(t)$ is the loss, called *marginal loss*, due to the obligor i. Different specifications for the marginal losses $L_i(t)$ can be chosen accounting for heterogeneity, time dependence, interaction, macroeconomic factors and so on. A detailed treatment of this general modeling framework can be found in the book by Embrechts, Frey and McNeil [14]. For a comparison with the most widely used industry examples of credit risk models see [15] or [16]. The same modeling insights are also developed in the most recent literature on risk management and large portfolio losses analysis; see [10,5] and [1] for different specifications.

Here we assume that

$$L_i(t) := \varphi(y_i[0,t], \lambda_i, t), \tag{2.1}$$

where $\varphi(\cdot, \cdot, t)$ is bounded and continuous in $\mathcal{D}[0,t] \times \mathrm{supp}(\mu)$. In other words the marginal loss depends explicitly on the realization of λ_i and on the *history* of y_i.

As a particular case of our general framework we obtain the most standard set-up commonly used in the literature of credit risk: consider $\varphi(y_i[0,t], \lambda_i, t) := e(\lambda_i, t) y_i(t)$ where $e : \mathbb{R}^3 \to \mathbb{R}^+$ is a continuous function of λ_i, and measures the *exposure* in case of default. Thus

$$L^{(N)}(t) = \sum_{i=1}^{N} e(\lambda_i, t) y_i(t).$$

We shall often speak of *asymptotic loss* or *asymptotic portfolio*. In this case we are referring to the $N \to \infty$ case of infinitely many obligors. The large portfolio is intended to be a *large* but finite approximation of this asymptotic regime.

As a consequence of the central limit theorem for the empirical measure, we obtain the following description for $L^N(t)$, the aggregate losses computed at time $t \in [0,T]$:

Corollary 2. *Assume that the reciprocity condition (R) is satisfied, and consider $L_i(t)$ as given in (2.1). In what follows, for $y[0,t] \in \mathcal{D}[0,t]$ and $\lambda \in \mathbb{R}^+ \times \mathbb{R}$, we write $L(t)$ for $\varphi(y[0,t], \lambda, t)$, and $l(t) = E^{Q_*}[L(t)]$. As $N \to \infty$ the sequence*

$$\sqrt{N} \left[\frac{\sum_{i=1}^{N} L_i(t)}{N} - l(t) \right]$$

converges weakly to a centered Gaussian random variable with variance $V(t)$, where

$$V(t) = E^{Q_*}\left[\left(L(t) - l(t) - \beta \int_0^t (1 - y(s)) \mathrm{Cov}_{Q_*}(\alpha y(s), L(t)) \, \mathrm{d}M(s) \right)^2 \right] \tag{2.2}$$

and where $M(t)$ has been defined in (1.8).

Proof. We apply Theorem 3 with $n = 1$ and $\phi_1 = \phi = \varphi(y[0,t], \lambda, t) = L(t)$. In this case $\int \phi \mathrm{d}\rho_N = \frac{L^N(t)}{N}$ and $\int \phi \mathrm{d}Q_* = E^{Q_*}[L(t)] = l(t)$. Notice that we can consider, without loss of generality, t as the final horizon of the time period, i.e. $t = T$. ■

Remark 1. By applying Theorem 3 with $(\phi_j)_{j=1}^n$, where $\phi_j := \varphi(y[0, t_j], \lambda, t_j)$, one could show that the *finite dimensional distributions* of the process

$$\left(\sqrt{N}\left[\frac{1}{N}\sum_{i=1}^{N} L_i(t) - l(t)\right]\right)_{t\geq 0}$$

converge to those of a Gaussian process, whose covariance can in principle be computed.

The asymptotic expected value $l(t)$ corresponds to the fraction of loss at time t in a benchmark portfolio of infinitely many firms. As regards Eq. (2.2), notice that in the case of no interaction (i.e., $\beta = 0$) we have

$$V(t) = E^{Q_*}\left[(L(t) - l(t))^2\right] = \mathrm{Var}_{Q_*}(L(t)).$$

In the case of $\beta > 0$ there is a supplementary noise given by the interaction. It depends on the past history of the process, and hence it produces a sort of "*memory*" of the variance $V(t)$.

The variance given in (2.2) involves the integral with respect to a martingale. A simpler form for the variance, more suitable for numerical computations, can be found in the following special but significant case.

Proposition 1. *Suppose that $L_i(T) = e(\lambda_i, T)y_i(T)$. Then as $N \to \infty$ we have that*

$$\sqrt{N}\left[\frac{L^N(T)}{N} - l(T)\right],$$

where $l(T) = E^{Q_*}[e(\lambda, T)y(T)]$, converges to a centered Gaussian random variable with variance

$$V(T) = \mathrm{Var}_{Q_*}(L(T)) + \int_0^T \left(E^{Q_*}[\alpha y(s)(e(\lambda, T) - l(T))]\right)^2 \times E^{Q_*}[\beta^2(1 - y(s))\mathrm{e}^{-\gamma+\beta m_{Q_*}(s)}]\,\mathrm{d}s, \tag{2.3}$$

where $m_{Q_*}(t) := \int \alpha' \eta(t) Q_*(\mathrm{d}\eta[0, T], \mathrm{d}\lambda')$.

Proof. We only need to show that the variance can be written in the form given in (2.3). From (2.2) it is easy to see that $V(T)$ can be written as

$$V(T) = E^{Q_*}\left[\left(L(T) - l(T) - \int_0^T \beta(1 - y(s)) \times E^{Q_*}[\alpha y(s)(e(\lambda, T)y(T) - l(T))]\,\mathrm{d}M(s)\right)^2\right]. \tag{2.4}$$

We now look at the expectation in the integral

$$E^{Q_*}[\alpha y(s)(e(\lambda, T)y(T) - l(T))] = E^{Q_*}[\alpha y(s)(e(\lambda, T) - l(T))]$$

where we have used the fact that $E^P[y(s)y(T)] = E^P[y(s)]$ for $s \leq T$ and for all $P \in \mathcal{M}_1$. Substituting in (2.4) we have

$$\begin{aligned}V(T) &= E^{Q_*}\left[\left(L(T)-l(T)-\int_0^T\beta(1-y(s))\right.\right.\\&\qquad\left.\left.\times\ E^{Q_*}[\alpha y(s)(e(\lambda,T)-l(T))]\,\mathrm{d}M(s)\right)^2\right]\\&= E^{Q_*}\left[(L(T)-l(T))^2\right]\\&\quad - E^{Q_*}\left[2(L(T)-l(T))\int_0^T\beta(1-y(s))E^{Q_*}[\alpha y(s)(e(\lambda,T)-l(T))]\,\mathrm{d}M(s)\right]\\&\quad + E^{Q_*}\left[\left(\int_0^T\beta(1-y(s))E^{Q_*}[\alpha y(s)(e(\lambda,T)-l(T))]\,\mathrm{d}M(s)\right)^2\right].\end{aligned}\tag{2.5}$$

The first expectation is the variance of $L(T)$ computed under Q_*. We show now that the second expectation is null. Indeed it is equal to

$$\begin{aligned}&2E^{Q_*}\left[L(T)\int_0^T\beta(1-y(s))E^{Q_*}[\alpha y(s)(e(\lambda,T)-l(T))]\,\mathrm{d}M(s)\right]\\&\quad-2l(T)E^{Q_*}\left[\int_0^T\beta(1-y(s))E^{Q_*}[\alpha y(s)(e(\lambda,T)-l(T))]\,\mathrm{d}M(s)\right].\end{aligned}$$

The second term is zero since M is a $(Q_*;\mathcal{F})$-martingale and the argument of the integral is $\mathcal{F}_s$ measurable. As regards the first one, we see that

$$\begin{aligned}E^{Q_*}\left[L(T)\int_0^T(1-y(s))(\cdot)\,\mathrm{d}M(s)\right] &= E^{Q_*}\left[e(\lambda,T)\mathbf{1}_{\{\tau\le T\}}\int_0^{T\wedge\tau}(\cdot)\,\mathrm{d}M(s)\right]\\&= E^{Q_*}\left[e(\lambda,T)\int_0^{\tau}(\cdot)\,\mathrm{d}M(s)\right]=0,\end{aligned}$$

where the last equality is due to the fact that $e(\lambda,T)$ is $\mathcal{F}_0$ measurable. As regards the last term in (2.5) we now show that

$$\begin{aligned}&E^{Q_*}\left[\left(\int_0^T\beta(1-y(s))E^{Q_*}[\alpha y(s)(e(\lambda,T)-l(T))]\,\mathrm{d}M(s)\right)^2\right]\\&\quad= E^{Q_*}\left[\int_0^T\left[\beta(1-y(s))E^{Q_*}[\alpha y(s)(e(\lambda,T)-l(T))]\right]^2\right.\\&\qquad\left.\times\ (1-y(s))\mathrm{e}^{-\gamma+\beta m_{Q_*}(s)}\mathrm{d}s\right].\end{aligned}\tag{2.6}$$

Indeed, as $M(t)=\mathbf{1}_{\{\tau\le t\}}-\int_0^t(1-y(s))\mathrm{e}^{-\gamma+\beta m_{Q_*}(s)}\mathrm{d}s$, it can be shown that the quadratic variation $\langle M\rangle$ of M is $\langle M\rangle_t=\int_0^t(1-y(s))\mathrm{e}^{-\gamma+\beta m_{Q_*}(s)}\mathrm{d}s$, for all $t\in[0,T]$. Thus for each progressively measurable process $X=(X(t))_{t\in[0,T]}\in L^2_M([0,T])$, where $L^2_M([0,T]):=\left\{X : E^{Q_*}\int_0^T|X(s)|^2\mathrm{d}\langle M\rangle_s<\infty\right\}$,

$$\|X\|_{L^2_M([0,T])}=E^{Q_*}\int_0^T|X(s)|^2\mathrm{d}\langle M\rangle_s=E^{Q_*}\int_0^T|X(s)|^2(1-y(s))\mathrm{e}^{-\gamma+\beta m_{Q_*}(s)}\mathrm{d}s.$$

As a consequence, applying this last result with $X = (X(t))_{t\in[0,T]}$ defined as

$$X(t) = \beta(1 - y(t))E^{Q_*}[\alpha y(t)(e(\lambda, T) - l(T))]\mathrm{d}t,$$

by the isometry between $L^2(\Omega, \mathcal{F}_T, Q_*)$ and $L^2_M([0, T])$, we have

$$\left\| \int_0^T X(s)\mathrm{d}M(s) \right\|_{L^2(\Omega,\mathcal{F}_T,Q_*)} = \|X\|_{L^2_M([0,T])}$$

which is exactly (2.6). Finally

$$\begin{aligned} E^{Q_*} &\int_0^T \left[\beta(1 - y(s))E^{Q_*}[\alpha y(s)(e(\lambda, T) - l(T))]\right]^2 (1 - y(s))\mathrm{e}^{-\gamma+\beta m_{Q_*}(s)}\mathrm{d}s \\ &= \int_0^T \left(E^{Q_*}[\alpha y(s)(e(\lambda, T) - l(T))]\right)^2 \cdot E^{Q_*}[\beta^2(1 - y(s))\mathrm{e}^{-\gamma+\beta m_{Q_*}(s)}]\,\mathrm{d}s, \end{aligned}$$

and this proves (2.3). ■

In the next section we show an application of this law of large numbers (Theorem 2) and central limit theorem (Proposition 1). Indeed, we show the evolution of the default probability of certain groups of obligors and infer the corresponding probability of suffering large losses in a credit portfolio.

2.2. *An example with simulation results*

Consider the simplified case in which $L_i(T) = y_i(T)$: the exposure at default is equal to 1 for each obligor. Moreover assume that $\mu = p_1\delta_{\lambda_1} + (1 - p_1)\delta_{\lambda_2}$. This means that the law of the random environment $\lambda = (\alpha, \beta, \gamma)$ puts mass on two possible outcomes, that is $\lambda \in \{\lambda_1; \lambda_2\}$. In this case the index $i = 1, 2$ identifies a group of obligors with the same marginal characteristics. In other words, we split the portfolio into two kinds of obligors with different specifications. More complex specifications could also be chosen. For the sake of simplicity we give this illustrative example, where interesting features of the dynamics of the state variables can be captured.

Recall that α_i specifies the relative weight of firm i in building the aggregate variable $m_N = \frac{1}{N}\sum_i \alpha_i y_i$ (see Eq. (1.1)). β_i is the parameter that measures how sensitive obligor i is with respect to the aggregate variable m_N; put differently, it is a measure of the contagion effect. γ_i is the idiosyncratic term in the marginal default probability.

As an example, we take a portfolio of $N = 125$ obligors. This is a typical size for CDO portfolios. We suppose that the portfolio consists of obligors of two types, in particular $\lambda_1 = (4, 4, 3)$, $\lambda_2 = (0.1, 0.1, 3)$. Notice that $\alpha_i = \beta_i$; hence the reciprocity condition (R) applies and we are allowed to place reliance on Theorem 3 and Corollary 2.

Obligors of type 1 are more sensitive to the aggregate variable, that is, their marginal default probability depends strongly on the default indicator of the other firms. Obligors of type 2 are less influenced by the aggregate variable m_N. The idiosyncratic term γ is the same for each obligor.

With this choice of the parameters, we want to stress the fact that even though the marginal default probabilities of the two types are rather similar for the very short horizon (where the impact of γ is higher), the contagion effect becomes preeminent as time increases, at least under certain specifications in the construction of the portfolio.

To illustrate this situation, in Fig. 1 (on the left), we show the dynamics of the marginal default probability of the two groups in two different scenarios. Scenario A mimics a portfolio

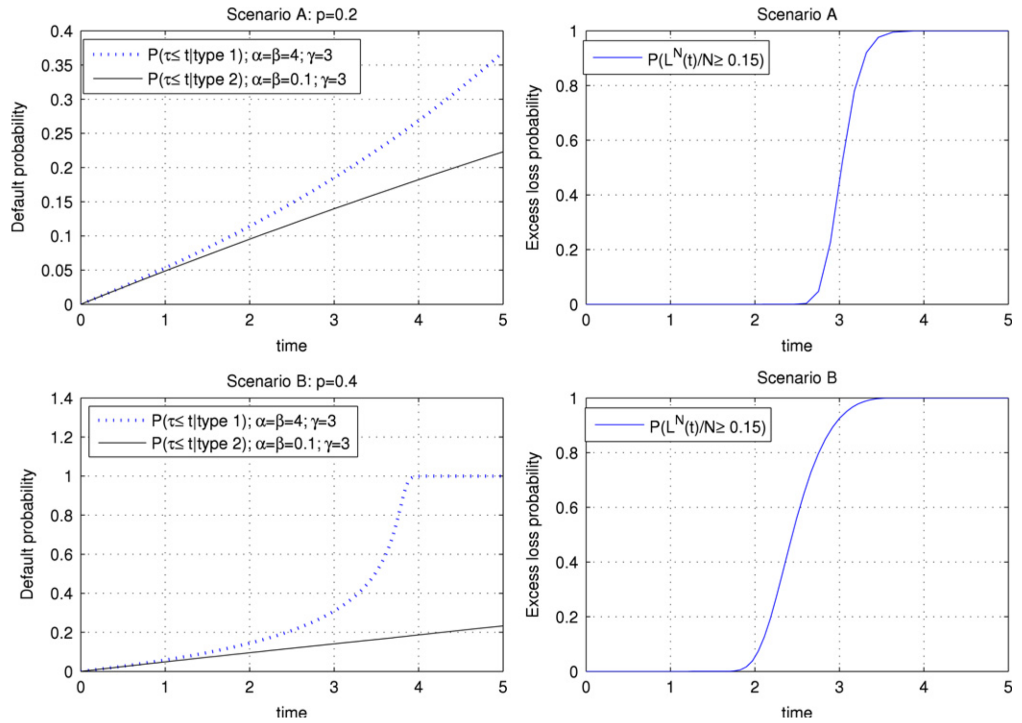

Fig. 1. Conditional probabilities of default and excess loss probabilities in a portfolio of $N = 125$ obligors where $L^N(t)$ is as defined in Section 2.2. Scenario A represents a portfolio where only the 20% of obligors are of type 1 whereas in Scenario B 40% of the obligors are of type 1. In both Scenario A and Scenario B we have $\lambda_1 = [4, 4, 3]$, $\lambda_2 = [0.1, 0.1, 3]$, where λ_i describes the idiosyncratic characteristics of obligors of type $i = 1, 2$.

where only 20% of the obligors are in group 1; this means that the proportion of firms exposed to *contagion risk* is lower. In scenario B the proportion is increased to 40%. Notice that in the second scenario the probability of default of the firms in the first group increases dramatically after the second year. Firms of type 2 are less sensitive.

Relying on Proposition 1, we compute the corresponding excess probabilities, that is, the probability of suffering a loss bigger than x as a function of time in the two scenarios. In Fig. 1 (on the right) we see these probabilities for $x = 0.15$. Notice that in this simple example, $L^N(t)$ counts the number of defaults up to time t, so $P(L^N(T)/N \geq 0.15)$ represents the probability of having at least 15% of defaults in the whole portfolio. Looking at the graphs, it is easy to see that in the first scenario the probability of having such a loss is smaller. At time $t = 2.5$, $P(\frac{L^N(2.5)}{N} \geq x) \sim 0$, whereas in the second scenario, $P(\frac{L^N(2.5)}{N} \geq x) \sim 0.55$.

This simple example suggests that the contagion effect can be very significant when looking at the probability of suffering a certain loss in a large portfolio. This is crucial for risk measurement purposes and for pricing tranches of CDOs.

It may be argued that an approximation via an infinite portfolio may not reproduce the real situation. In Fig. 2 we show a comparison between the evolution in time of the default probabilities in the two groups of obligors computed under Q_* (in the upper part) and simulated via the *real* Markov process with $N = 125$ (in the lower part). Here the parameters are

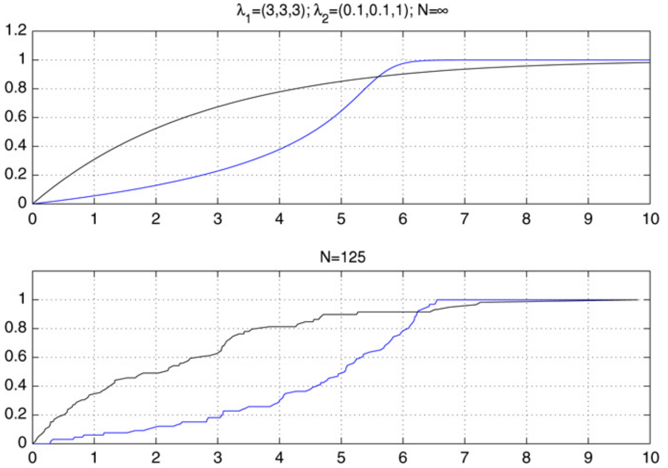

Fig. 2. Comparison between the dynamics of the default probabilities of two groups of firms computed under two different models. In the upper part we see the plot under the asymptotic model with infinite firms (under Q_*). In the lowest one we have implemented a simulation of the Markov process directly (here $N = 125$). The parameters are $\lambda_1 = [3, 3, 3]$, $\lambda_2 = [0.1, 0.1, 1]$, where λ_i describes the idiosyncratic characteristics of obligors of type $i = 1, 2$. The distribution of λ is in this case $\mu = \frac{1}{2}\delta_{\lambda_1} + \frac{1}{2}\delta_{\lambda_2}$.

$\lambda_1 = (3, 3, 3)$, $\lambda_2 = (0.1, 0.1, 1)$. These graphs show that the asymptotic equation is a good approximation for the Markov process even with $N = 125$.

Finally we show in Fig. 3 the comparison between the evolution of the aggregate loss (in black) evaluated as before in two different ways. In the upper part we see the aggregate loss computed relying on Proposition 1. Below we see the plot of a trajectory of the Markov process (with $N = 125$). In the same graphs we have also plotted the dynamics of the probability of suffering a loss over certain thresholds c_k. In other words we plot $t \mapsto P(L^N(t) \geq c_k)$ for $k = 1, 2, 3$ (in this case $c_1 = 5\%$, $c_2 = 15\%$, $c_3 = 25\%$). Those probabilities are the building blocks for computing the price of tranches of CDO contracts. For a description of these credit derivatives see for instance [14].

3. Conclusions

We have proposed a model for *credit contagion* where *heterogeneity* and *direct contagion* among the firms are taken into account. We then quantified the impact of contagion on the losses suffered by a financial institution holding a large portfolio with positions issued by the firms.

Unlike the existing literature on credit contagion, our work has proposed a *dynamic* model where it is possible to describe the evolution of the indicators of financial distress. In this way we are able to compute the distribution of the losses in a large portfolio for any time horizon T, via a suitable version of the central limit theorem.

The peculiarity of our model is the fact that the homogeneity assumption is broken by introducing a *random environment* that makes it possible to take into account the idiosyncratic characteristics of the firms. One drawback of the intensity based models commonly proposed in the literature is the difficulty in managing large heterogeneous portfolios because of the presence of many obligors with different specifications. In this case it is common practice to make homogeneity assumptions in order to reduce the complexity of the problem. A typical approach is to divide the portfolio into groups where the obligors may be considered exchangeable. We have

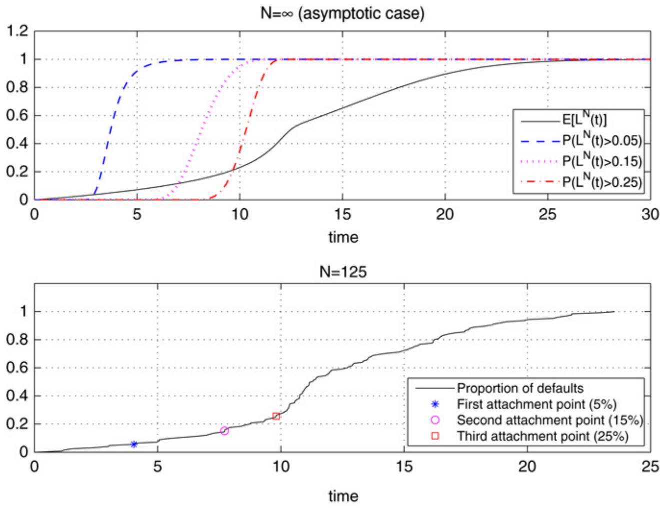

Fig. 3. Evolution in time of the aggregate losses (in black) computed under two different models: under the asymptotic model (upper) and the Markov process (lower). The parameters are $\lambda_1 = [2, 6, 4]$, $\lambda_2 = [1, 3, 5]$, where λ_i describes the idiosyncratic characteristics of obligors of type $i = 1, 2$. The distribution of λ is in this case $\mu = 0.4\delta_{\lambda_1} + 0.6\delta_{\lambda_2}$.

shown that our model goes behind the identification of groups of firms that can be considered basically exchangeable. Despite this heterogeneity assumption our model has the advantage of being totally tractable: it is possible to compute in closed form the mean and the variance of a central limit type approximation for the losses due to large portfolios in a dynamic fashion.

As an example of using the general theory, we have computed the default probabilities and different risk measures in a simple situation with only two groups of obligors. Moreover we have compared the numerical results obtained relying on the asymptotic model and on the central limit theorem (Corollary 2) with the results obtained in a simulation of the underlying Markov process with finite $N = 125$. These results show the validity of the approximation and are encouraging as regards a more involved analysis. This issue is left to future research: it is in fact beyond the scope of this work to pursue a detailed calibration of the model for real data.

Acknowledgments

The second author thanks the Department of Finance at Bocconi University (Milano) for financial support. This work was supported in part by the University of Padova under grant CPDA082105/08.

Appendix. Proofs of the main results

A.1. Proof of Theorem 1

We need to prove some technical lemmas.

Lemma 1. *Let $\mathcal{X}$ be a Polish space. Let $(P_N)_N$ satisfy the LDP with rate N and good rate function H. Let $F : \mathcal{X} \to \mathbb{R}$ be measurable, bounded from above and continuous on the set $\mathcal{X}_H := \{x : H(x) < \infty\}$. Then the sequence of probability measures $(P_N^F)_N$ defined by*

$$\frac{\mathrm{d}P_N^F}{\mathrm{d}P_N}(\cdot) = \frac{\exp(NF(\cdot))}{\int_{\mathcal{X}} \exp(NF(y))P_N(\mathrm{d}y)} \tag{A.1}$$

satisfies the LDP with the good rate function

$$I(x) = H(x) - F(x) - \inf_{y\in\mathcal{X}}[H(y) - F(y)]. \tag{A.2}$$

In particular

$$\lim_{N\to+\infty} \frac{1}{N} \log\left[\int_{\mathcal{X}} \exp\left(NF(y)\right) P_N(\mathrm{d}y)\right] = -\inf_{y\in\mathcal{X}} \left[H(y) - F(y)\right]. \tag{A.3}$$

For a proof see [17]. This is a relaxed version of the usual Varadhan lemma for tilted large deviation principles (see [18]). The statement is relaxed in the sense that we assume that a suitable function $F : \mathcal{X} \to \mathbb{R}$, instead of being continuous on all of its domain, is continuous only on a subset $\mathcal{X}_H \subsetneq \mathcal{X}$ in the following sense:

For any sequence $(x_n)_n \in \mathcal{X}$ such that $x_n \to x$, where $x \in \mathcal{X}_H$, we have $F(x_n) \to F(x)$.

We point out that this is a stronger assumption than assuming continuity of the restriction of F on the subset $\mathcal{X}_H$.

Lemma 2. *For given $\underline{\lambda} \in (\mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R})^N$,*

$$\frac{\mathrm{d}P_N^{\lambda}}{\mathrm{d}W^{\otimes N}}(\underline{y}[0,T]) = \exp\{NF(\rho_N(\underline{y}[0,T],\underline{\lambda}))\} \tag{A.4}$$

where $F(Q)$ has been defined in (1.3).

Proof. It basically follows from the Girsanov formula for point processes.

$$\begin{aligned}\frac{\mathrm{d}P_N^{\lambda}}{\mathrm{d}W^{\otimes N}} = \exp\Bigg\{&\sum_{i=1}^N \int_0^T \left[(1-y_i(t)) - (1-y_i(t))\mathrm{e}^{-\gamma_i+\beta_i\int \alpha'\eta(t)\rho_N(\mathrm{d}\eta[0,T],\mathrm{d}\lambda')}\right]\mathrm{d}t \\ &+ \sum_{i=1}^N y_i(T)\left[-\gamma_i + \left(\beta_i \int \alpha'\eta(t^-)\rho_N(\mathrm{d}\eta[0,T],\mathrm{d}\lambda')\right)\Big|_{t=\tau_i}\right]\Bigg\};\end{aligned}$$

where τ_i has been defined in (1.2).

The term in the { } brackets is exactly $F(Q)_{|Q=\rho_N}$ multiplied by N. ■

We now define for each $Q \in \mathcal{M}$ and $t \in [0,T]$

$$m_Q(t) := \int \alpha\eta(t)Q(\mathrm{d}\eta[0,T],\mathrm{d}\lambda). \tag{A.5}$$

We shall often use this notation in the rest of this Appendix.

Lemma 3. *$F(Q)$ is bounded on $\mathcal{M}_1$ and continuous on the subset $\mathcal{M}_A := \{Q \in \mathcal{M}_1 : Q \ll W \otimes \mu\}$.*

Proof. We rewrite $F(Q)$ as given in (1.3) using the notation introduced in (A.5):

$$\begin{aligned}F(Q) = \int Q(\mathrm{d}y[0.T],\mathrm{d}\lambda)\Bigg\{&\int_0^T (1-y(t))\left(1-\mathrm{e}^{-\gamma}\mathrm{e}^{\beta m_Q(t)}\right)\mathrm{d}t \\ &+ y(T)\left[-\gamma + \beta m_Q(t^-)\big|_{t=\tau(y[0,T])}\right]\Bigg\},\end{aligned} \tag{A.6}$$

The argument in the { } brackets is bounded; thus we are allowed to interchange the expectation with respect to Q and the time integral:

$$F(Q) = \int_0^T E^Q[1 - y(t)]\,\mathrm{d}t - \int_0^T E^Q\left[\mathrm{e}^{-\gamma}\mathrm{e}^{\beta m_Q(t)}(1 - y(t))\right]\mathrm{d}t \\ - E^Q[\gamma y(T)] + E^Q\left[\beta y(T) m_Q(t^-)\big|_{t=\tau(y[0,T])}\right].$$

We show now the boundedness and the continuity of F. The boundedness is easily proved since $y \in \{0; 1\}$, and the distribution of λ under $Q \in \mathcal{M}_1^b$ has bounded support.

In order to prove the continuity on $\mathcal{M}_A$, we consider a sequence of probabilities $(Q_n)_{n\geq 0} \in \mathcal{M}_1^b$ converging weakly to $Q \in \mathcal{M}_A$. We split the proof into different steps.

We show first that

$$\lim_n E^{Q_n}[f(\lambda)y(t)] = E^Q[f(\lambda)y(t)] \tag{A.7}$$

for all $t \in [0, T]$ and for any continuous function $f : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ bounded on the support of μ. This statement is not trivial, since the projection $y[0, T] \to y(t)$ is not continuous in $\mathcal{D}[0, T]$. However, we define for any $\epsilon > 0$ the functions

$$g_t^-(\epsilon; f) := \frac{1}{\epsilon}\int_{t-\epsilon}^t f(\lambda)y(s)\mathrm{d}s, \qquad g_t^+(\epsilon; f) := \frac{1}{\epsilon}\int_t^{t+\epsilon} f(\lambda)y(s)\mathrm{d}s;$$

where we suppose that the trajectory $y[0, T]$ can be extended to the larger interval $[0-\epsilon,\ T+\epsilon]$ by continuity.

These functions are continuous in $\mathcal{D}[0, T]$, bounded by $\|f\|_\infty$ for any ϵ and such that $g_t^-(\epsilon; f) \leq f(\lambda)y(t) \leq g_t^+(\epsilon; f)$ a.s. for any t. Thus, by the Lebesgue convergence theorem,

$$\limsup_n E^{Q_n}[f(\lambda)y(t)] \leq \lim_n E^{Q_n}[g^+(\epsilon; f)] = E^Q[g^+(\epsilon; f)], \quad \forall \epsilon > 0.$$

Letting $\epsilon \to 0$ and noticing that $\lim_{\epsilon\to 0} g_t^+(\epsilon; f) = f(\lambda)y(t)$ we get

$$\limsup_n E^{Q_n}[f(\lambda)y(t)] \leq E^Q[f(\lambda)y(t)].$$

The same argument holds for $g_t^-(\epsilon; f)$; here $\lim_{\epsilon\to 0} g_t^-(\epsilon; f) = f(\lambda)y(t^-)$. Thus

$$E^Q[f(\lambda)y(t^-)] \leq \liminf_n E^{Q_n}[f(\lambda)y(t)] \leq \limsup_n E^{Q_n}[f(\lambda)y(t)] \leq E^Q[f(\lambda)y(t)].$$

Notice that $f(\lambda)y(t)$ and $f(\lambda)y(t^-)$ may differ only in the event that $\{y(t^-) \neq y(t)\}$. But this event has measure zero for any $Q \in \mathcal{M}_A$, since $(W \otimes \eta)(\{y(t^-) \neq y(t)\}) = 0$. This implies that the corresponding expected values must coincide; as a consequence, $E^Q[f(\lambda)y(t)] - E^Q[f(\lambda)y(t^-)] = 0$. We have thus proved that

$$\lim_n E^{Q_n}[f(\lambda)y(t)] = E^Q[f(\lambda)y(t)] \quad \text{for all } t. \tag{A.8}$$

Notice that in saying that $(W\otimes\eta)(\{y(t^-) \neq y(t)\}) = 0$ we have used the fact that the distribution function of τ under $W \otimes \eta$ is exponential. In particular, it is absolutely continuous. A similar argument shows that if $Q \notin \mathcal{M}_A$ then $E^{Q_n}[y(t)]$ converges pointwise in t to $E^Q[y(t)]$, for all those t such that $Q(\tau = t) = 0$.

Taking $f(\lambda) \equiv 1$, we have that, for all t, $E^Q[y(t)]$ is a continuous mapping in $Q \in \mathcal{M}_A$. Choosing instead $f(\lambda) = \mathrm{e}^{-\gamma}$, $f(\lambda) = -\gamma$ and $f(\lambda) = \alpha$, we prove continuity for $E^Q\left[y(t)\mathrm{e}^{-\gamma}\right]$, $E^Q[-\gamma y(T)]$ and $m_Q(t)$ respectively.

The next step is to show that $Q_n \to Q$ implies that

$$\left| E^Q\left[\int_0^T (1-y(t))\,\mathrm{e}^{-\gamma}\mathrm{e}^{\beta m_Q(t)}\mathrm{d}t\right] - E^{Q_n}\left[\int_0^T (1-y(t))\,\mathrm{e}^{-\gamma}\mathrm{e}^{\beta m_{Q_n}(t)}\mathrm{d}t\right]\right| \tag{A.9}$$

converges to zero.

We add and subtract $E^{Q_n}\left[\int_0^T (1-y(t))\,\mathrm{e}^{-\gamma}\mathrm{e}^{\beta m_Q(t)}\mathrm{d}t\right]$:

$$\left| E^Q\left[\int_0^T (1-y(t))\,\mathrm{e}^{-\gamma}\mathrm{e}^{\beta m_Q(t)}\mathrm{d}t\right] - E^{Q_n}\left[\int_0^T (1-y(t))\,\mathrm{e}^{-\gamma}\mathrm{e}^{\beta m_Q(t)}\mathrm{d}t\right]\right.$$
$$\left. +\; E^{Q_n}\left[\int_0^T \left((1-y(t))\mathrm{e}^{-\gamma}\right)\left(\mathrm{e}^{\beta m_Q(t)} - \mathrm{e}^{\beta m_{Q_n}(t)}\right)\mathrm{d}t\right]\right| \le |a_n| + |b_n|$$

where

$$a_n = E^Q\left[\int_0^T (1-y(t))\,\mathrm{e}^{-\gamma}\mathrm{e}^{\beta m_Q(t)}\mathrm{d}t\right] - E^{Q_n}\left[\int_0^T (1-y(t))\,\mathrm{e}^{-\gamma}\mathrm{e}^{\beta m_Q(t)}\mathrm{d}t\right];$$
$$b_n = E^{Q_n}\left[\int_0^T \left((1-y(t))\,\mathrm{e}^{-\gamma}\right)\left(\mathrm{e}^{\beta m_Q(t)} - \mathrm{e}^{\beta m_{Q_n}(t)}\right)\mathrm{d}t\right].$$

$|a_n|$ goes to zero by weak convergence.

As regards b_n we see that

$$|b_n| \le \int_0^T E^{Q_n}\left[(1-y(t))\,\mathrm{e}^{-\gamma}\cdot\left|\mathrm{e}^{\beta m_Q(t)} - \mathrm{e}^{\beta m_{Q_n}(t)}\right|\right]\mathrm{d}t$$
$$\le \int_0^T E^{Q_n}\left[(1-y(t))\,\mathrm{e}^{-\gamma}\cdot\left|\mathrm{e}^{\beta m_Q(t)} - \mathrm{e}^{\beta m_{Q_n}(t)}\right|\right]\mathrm{d}t.$$

We now show that

$$E^{Q_n}\left[(1-y(t))\,\mathrm{e}^{-\gamma}\cdot\left|\mathrm{e}^{\beta m_Q(t)} - \mathrm{e}^{\beta m_{Q_n}(t)}\right|\right] \to 0.$$

We can rewrite it as

$$E^{Q_n}\left[(1-y(t))\,\mathrm{e}^{-\gamma}\mathrm{e}^{\beta m_Q(t)}\cdot\left|\mathrm{e}^{\beta[m_Q(t)-m_{Q_n}(t)]} - 1\right|\right]$$
$$\le K\,E^{Q_n}\left[\left|\mathrm{e}^{\beta[m_Q(t)-m_{Q_n}(t)]} - 1\right|\right] = (*)$$

for a suitable $K \in \mathbb{R}^+$, where we have used the fact that $(1-y(t))\,\mathrm{e}^{-\gamma}\mathrm{e}^{\beta m_Q(t)}$ is uniformly bounded. We now look at the term in the expectation

$$\left|\mathrm{e}^{\beta[m_Q(t)-m_{Q_n}(t)]} - 1\right| \le K_2|m_Q(t) - m_{Q_n}(t)|$$

again by uniformly boundedness, for a suitable K_2. Thus

$$(*) \le K_3|m_Q(t) - m_{Q_n}(t)|,$$

and this converges to zero thanks to what we have shown in (A.7).

As a consequence, $|b_n|$ goes to zero as well, since we are allowed to interchange the limit and the time integral, by dominated convergence.

It remains to show the continuity of the term

$$E^Q[\beta y(T) m_q(\tau^-)]. \tag{A.10}$$

Indeed, take a sequence $Q_n \to Q$, $Q \in \mathcal{M}_A$; then

$$\begin{aligned}&|E^{Q_n}[\beta y(T) m_{Q_n}(\tau^-)] - E^Q[\beta y(T) m_Q(\tau^-)]| \\ &\quad \le |E^{Q_n}[\beta y(T)\left\{m_{Q_n}(\tau^-) - m_Q(\tau^-)\right\}]| \\ &\qquad + |E^{Q_n}[\beta y(T) m_Q(\tau^-)] - E^Q[\beta y(T) m_Q(\tau^-)]|.\end{aligned}$$

The second term goes to zero by weak convergence, since the function m_Q is continuous. As regards the first term, it is enough to show that $\left\{m_{Q_n}(t) - m_Q(t)\right\}$ converges to zero uniformly on $[0, T]$. To show this, we fix $t \in [0, T]$; then the following facts hold true:

(a) For $\delta_1 > 0$ there exists $\epsilon > 0$ such that $|s - t| \le \epsilon \Rightarrow |m_Q(s) - m_Q(t)| \le \delta_1$.
(b) There exists $\bar{n}$ such that $\forall n \ge \bar{n}$ $|m_{Q_n}(t+\epsilon) - m_Q(t+\epsilon)| \le \delta_2$; $|m_{Q_n}(t-\epsilon) - m_Q(t-\epsilon)| \le \delta_3$.

Point (a) is due to the continuity of $m_Q(t)$ for $Q \in \mathcal{M}_A$ whereas (b) follows by the fact that $m_{Q_n}(t) \to m_Q(t)$ pointwise in t as shown in (A.7). Notice that when $t + \epsilon > T$ or $t - \epsilon < 0$ the inequalities in (b) are modified appropriately without loss of generality.

We now claim that fixing $s \in O_t := [t - \epsilon,\ t + \epsilon]$ we have

$$|m_{Q_n}(s) - m_Q(s)| \le \delta_1 + \delta_2 + \delta_3 = \bar{\delta}. \tag{A.11}$$

Fix s and n and suppose $m_{Q_n}(s) \le m_Q(s)$ (the other case is treated in the same way):

$$m_Q(s) - m_{Q_n}(s) \le m_Q(t+\epsilon) - m_Q(t-\epsilon) + m_Q(t-\epsilon) - m_{Q_n}(t-\epsilon) \le \delta_1 + \delta_3;$$

where we have used the fact that $m_Q(t)$ is increasing in t for all $Q \in \mathcal{M}_1$. Thus we can extract a finite covering $\{O_{t_k}\}$ of $[0, T]$ where (A.11) holds true; hence uniform convergence is proved. ∎

Proof of Theorem 1. We denote by $\mathcal{P}_N$ the distribution of ρ_N under P_N, i.e. $\mathcal{P}_N := P^N \circ \rho_N^{-1}$. We now state a LDP for the sequence $\{\mathcal{P}_N\}_N$. Thanks to Lemma 2, we have identified the Radon–Nikodym derivative that relates $W^{\otimes N}$ and $P_N^{\underline{\lambda}}$ (where $W^{\otimes N}$ plays the role of the reference measure). A natural way to develop a large deviation principle is now to rely on Lemma 1.

Since $(y_i[0, T]; \lambda_i)$ are i.i.d. random variables under $(W \otimes \mu)^{\otimes N}$, we can apply Sanov's theorem (see [18]) to the sequence of measures $(\mathcal{W}_N)_N$, where $\mathcal{W}_N$ represents the law of the empirical measure in the case of independence (i.e. under $(W \otimes \mu)^{\otimes N}$). Hence $(\mathcal{W}_N)_N$ obeys a large deviation principle with rate function $H(Q|W \otimes \mu)$.

As $F(Q)$ is bounded in the weak topology and continuous on $\mathcal{M}_A \supset \mathcal{M}_H$ where $\mathcal{M}_A = \{Q \in \mathcal{M}_1 | Q \ll W \otimes \mu\}$ and $\mathcal{M}_H = \{Q \in \mathcal{M}_1 : H(Q|W \otimes \mu) < \infty\}$, we can rely on Lemma 1 to conclude that the sequence $(\mathcal{P}_N)_N$ obeys a large deviation principle with good rate function

$$I(Q) = H(Q|W \otimes \mu) - F(Q) - \inf_{R \in \mathcal{M}_1} [H(R|W \otimes \mu) - F(R)].$$

We finish the proof by showing that

$$\inf_{R \in \mathcal{M}_1} [H(R|W \otimes \mu) - F(R)] = \lim_{N \to \infty} \frac{1}{N} \log \left[\int_{\mathcal{M}_1} \mathrm{e}^{NF(Q)} \mathcal{W}_N(\mathrm{d}Q)\right] = 0. \tag{A.12}$$

The first equality is simply a consequence of Eq. (A.3).

We are thus left to prove that $\int_{\mathcal{M}_1} \mathrm{e}^{NF(Q)} \mathcal{W}_N(\mathrm{d}Q) = 1$. Indeed,

$$\mathcal{P}_N(\,\cdot\,) = \int \mu^{\otimes N}(\mathrm{d}\underline{\lambda}) P_N^{\underline{\lambda}}(\rho_N \in \cdot\,)$$

$$= \int \mu^{\otimes N}(\mathrm{d}\underline{\lambda}) \int \mathbb{I}_{\{\rho_N \in \cdot\,\}} \frac{\mathrm{d}P_N^{\underline{\lambda}}}{\mathrm{d}W^{\otimes N}} \mathrm{d}W^{\otimes N} = \int \mathbb{I}_{\{\rho_N \in \cdot\,\}} \mathrm{e}^{NF(\rho_N)}\, \mathrm{d}(W^{\otimes N} \otimes \mu^{\otimes N})$$
$$= \int \mathbb{I}_{\{Q \in \cdot\,\}} \mathrm{e}^{NF(Q)} \mathcal{W}_N(\mathrm{d}Q).$$

As $\mathcal{P}_N(\mathcal{M}_1) = 1$, the theorem follows. ∎

A.2. *Proof of* Theorem 2

We need to define a new process and a technical lemma related to it.

We associate with any $Q \in \mathcal{M}_1$ the law of a time inhomogeneous Markov process on {0; 1} which evolves according to the following rules:

$$\begin{array}{lll} y = 0 \to y = 1 & \text{with intensity} & \mathrm{e}^{-\gamma} \mathrm{e}^{\beta \int \alpha' \eta(t) Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda)} \\ y = 1 \to y = 0 & \text{with intensity} & 0 \end{array}$$

and with $y_i(0) = 1$ for all $i = 1, \ldots, N$.

We denote by $P^{\lambda,Q}$ the law of this process and by $P^Q = P^{\lambda,Q} \otimes \eta$. In other words, $P^{\lambda,Q}$ is the law of the Markov process on {0; 1} with initial distribution δ_0 and time-dependent generator $L_t^{\lambda,Q}$ defined by

$$L_t^{\lambda,Q} f(s) = (1-s)\mathrm{e}^{-\gamma} \mathrm{e}^{\beta \int \alpha' \eta(t) Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda)} (f(1-s) - f(s)). \tag{A.13}$$

We show now an important property of P^Q.

Lemma 4. *For every $Q \in \mathcal{M}_1$, we have*

$$I(Q) = H(Q|P^Q).$$

Proof. We distinguish two cases:
Case 1. $Q : H(Q|W \otimes \eta) < \infty$. We have

$$I(Q) = \int \log \frac{\mathrm{d}Q}{\mathrm{d}(W \otimes \eta)} \mathrm{d}Q - F(Q).$$

By Girsanov's formula for continuous time Markov chains, we obtain

$$\log \frac{\mathrm{d}P^{\lambda,Q}}{\mathrm{d}W} = \int_0^T (1 - y(t)) \left(1 - \mathrm{e}^{-\gamma} \mathrm{e}^{\beta \int Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda') \alpha' \eta(t)}\right) \mathrm{d}t$$
$$+ y(T) \left[-\gamma + \left(\beta \int \alpha' \eta(t^-) Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda') \right) \Big|_{t = \tau(y[0,T])} \right];$$

hence, by the definition of F given in (1.3) we have

$$F(Q) = \int \log \frac{\mathrm{d}P^{\lambda,Q}}{\mathrm{d}W}\, \mathrm{d}Q$$

and so

$$I(Q) = \int \log \frac{\mathrm{d}Q}{\mathrm{d}(W \otimes \eta)} \mathrm{d}Q - \int \log \frac{\mathrm{d}P^{\lambda,Q}}{\mathrm{d}W} \mathrm{d}Q = \int \log \frac{\mathrm{d}Q}{\mathrm{d}P^Q} \mathrm{d}Q \tag{A.14}$$

where the last equality follows from

$$\frac{\mathrm{d}Q}{\mathrm{d}(W\otimes\eta)}\frac{\mathrm{d}W}{\mathrm{d}P^{\lambda,Q}}=\frac{\mathrm{d}Q}{\mathrm{d}(W\otimes\eta)}\frac{d(W\otimes\eta)}{\mathrm{d}P^{Q}}=\frac{\mathrm{d}Q}{\mathrm{d}P^{Q}}.$$

As $\int\log\frac{\mathrm{d}Q}{\mathrm{d}P^Q}\mathrm{d}Q=H(Q|P^Q)$, the theorem follows.

Case 2. $Q: H(Q|W\otimes\eta)=+\infty$. In this case $I(Q)=+\infty$.

Thus we have to check that $H(Q|P^Q)=+\infty$ as well. As

$$H(Q|P^Q)=\int\log\frac{\mathrm{d}Q}{\mathrm{d}(W\otimes\eta)}\mathrm{d}Q+\int\log\frac{\mathrm{d}W}{\mathrm{d}P^{\lambda,Q}}\mathrm{d}Q,$$

the theorem follows since $\int\log\frac{\mathrm{d}Q}{\mathrm{d}(W\otimes\eta)}\mathrm{d}Q=+\infty$, as $H(Q|W\otimes\eta)=+\infty$ and since $\int\log\frac{\mathrm{d}W}{\mathrm{d}P^{\lambda,Q}}\mathrm{d}Q=-F(Q)$ which is bounded. ■

Proof of Theorem 2. By the properness of the relative entropy ($H(\mu|\nu)=0\Rightarrow\mu=\nu$), from lemma (A.13) we have that the equation $I(Q)=0$ is equivalent to $Q=P^Q$. Suppose $\bar{Q}$ is a solution of this last equation. In this case $\bar{Q}=P^{\bar{Q}}=P^{\lambda,\bar{Q}}\otimes\mu$, where $P^{\lambda,\bar{Q}}$ is the law of the Markov process on {0; 1} with initial distribution δ_0 and time-dependent generator $L_t^{\lambda,\bar{Q}}$ as defined in (A.13). The marginals of a Markov process are solutions of the corresponding *forward equation*. This leads to the fact that $\bar{q}_t:=\Pi_t P^{\lambda,\bar{Q}}$, the t-projection of $P^{\lambda,\bar{Q}}$, is a solution of $\dot{q}_t=\mathcal{L}_t^{\lambda,\bar{Q}}$ where $\mathcal{L}_t^{\lambda,\bar{Q}}$ is the adjoint of $L_t^{\lambda,Q}$:

$$(\mathcal{L}_t^{\lambda,\bar{Q}}q)(x)=x\mathrm{e}^{-\gamma}\mathrm{e}^{\beta\int\alpha'\eta(t)\bar{Q}(\mathrm{d}\eta[0,T],\mathrm{d}\lambda')}q(-x)-(1-x)\mathrm{e}^{-\gamma}\mathrm{e}^{\beta\int\alpha'\eta(t)\bar{Q}(\mathrm{d}\eta[0,T],\mathrm{d}\lambda')}q(x).$$

More specifically, when $x=0$ we have

$$(\mathcal{L}_t^{\lambda,\bar{Q}}q)(0)=-\mathrm{e}^{-\gamma}\mathrm{e}^{\beta\int\alpha'\eta(t)\bar{Q}(\mathrm{d}\eta[0,T],\mathrm{d}\lambda')}q(0)$$

and when $x=1$

$$(\mathcal{L}_t^{\lambda,\bar{Q}}q)(1)=\mathrm{e}^{-\gamma}\mathrm{e}^{\beta\int\alpha'\eta(t)\bar{Q}(\mathrm{d}\eta[0,T],\mathrm{d}\lambda')}q(0).\tag{A.15}$$

We now prove that $\dot{q}_t=\mathcal{L}_t^{\lambda,\bar{Q}}q$ admits at most one solution for each initial condition. To see this, define $\bar{q}_t(\lambda):=P^{\lambda,\bar{Q}}(y(t)=1)$. Then $\frac{\mathrm{d}\bar{q}_t(\lambda)}{\mathrm{d}t}=G(\bar{q}_t(\lambda))$, where $G(q)=\mathrm{e}^{-\gamma}\mathrm{e}^{\beta\int\alpha'q(\lambda')\mu(\mathrm{d}\lambda')}(1-q(\lambda))$. Notice that $\bar{q}_t(\lambda)\in L^1(\mu)$ and $G(\cdot)$ is a locally Lipschitz operator on a Banach space. Thus $\frac{\mathrm{d}\bar{q}_t(\lambda)}{\mathrm{d}t}=G(\bar{q}_t(\lambda))$ has at most one solution in $[0,T]$, for a given initial condition (see [19], Theorem VII.3).

Since $P^{\bar{Q}}$ is totally determined by the flow $\bar{q}_t$, it follows that equation $Q=P^Q$ has at most one solution. The existence of a solution follows from the fact that $I(Q)$ is the rate function of a LDP, and therefore *must* have at least one zero. By what is shown in (A.15), $\bar{q}_t(\lambda)$ solves (1.4). Hence Q_* turns out to be the unique solution of the fixed point argument $Q=P^Q$. Moreover, it satisfies all the conditions of Theorem 2. ■

A.3. Proof of Theorem 3

The key technical tool for the proof of Theorem 3 is the following result due to Bolthausen [7].

Theorem 4. *Let $(B,\|\cdot\|)$ be a real separable Banach space. Let $(Z_k)_{k\geq1}$ be a sequence of B-valued, i.i.d. random variables, defined on the probability space $(\Omega,\mathcal{A},\mathbb{P})$, and denote by w*

their common law. Define $X_N := \frac{1}{N}\sum_{k=1}^{N} Z_k$ *and consider a continuous map* $\Psi : B \to \mathbb{R}$. *Suppose that the following conditions are satisfied:*

(B.1) $\int \exp(r|x|)w(\mathrm{d}x) < \infty$ *for all* $r \in \mathbb{R}$.

(B.2) *For any* $x \in B$, $\Psi(x) \le C_1 + C_2\|x\|$ *for some* $C_1, C_2 > 0$. *Moreover,* Ψ *is three times continuously Fréchet differentiable.*

(B.3) *Define, for* $h \in B'$ *(the topological dual of* B*),* $\Lambda(h) := \int \mathrm{e}^{h(z)} w(\mathrm{d}z)$, *and for* $x \in B$, $\Lambda^*(x) := \sup_{h\in B'}[h(x) - \Lambda(h)]$. *Assume that there exists a unique* $z^* \in B$ *such that* $\Lambda^*(z^*) - \Psi(z^*) = \inf_{z\in B}[\Lambda^*(z) - \Psi(z)]$.

(B.4) *Define the probability* p *on* B *by* $\frac{\mathrm{d}p}{\mathrm{d}w} = \frac{\mathrm{e}^{D\Psi(z_*)}}{c}$ *for a suitable normalizing factor* c. *This probability is well defined and* $\int zp(\mathrm{d}z) = z^*$. *Let* p_* *denote the centered version of* p, *i.e.,* $p_* = p \circ \theta_{x^*}^{-1}$, *where* $\theta_a : B \to B$ *is defined by* $\theta_a(x) = x - a$. *For* $h \in B'$ *define* $\tilde{h} \in B$ *by* $\tilde{h} = \int zh(z)p_*(\mathrm{d}z)$. *Then we assume that for every* $h \in B'$ *such that* $\tilde{h} \neq 0$

$$\int h^2(z)p_*(\mathrm{d}z) - \mathrm{D}^2\Psi(z^*)[\tilde{h}, \tilde{h}] > 0.$$

(B.5) *B is a Banach space of type 2.*[1]

Now, let π_N *be the probability on* $(\Omega, \mathcal{A})$ *given by*

$$\frac{\mathrm{d}\pi_N}{\mathrm{d}\mathbb{P}} = \frac{\mathrm{e}^{N\left(\Psi(X_N)+\frac{\Sigma(X_N)}{N}\right)}}{E^{\mathbb{P}}\left[\mathrm{e}^{N\left(\Psi(X_N)+\frac{\Sigma(X_N)}{N}\right)}\right]}, \tag{A.16}$$

where Σ *is linear, continuous and bounded on the support of the law of* X_N, *uniformly in* N. *Then, for every* $h_1, \dots, h_n \in B'$, *the* π_N*-law of the* n*-dimensional vector*

$$\sqrt{N}\,(h_i(X_N) - h_i(z_*))_{i=1}^n$$

converges weakly, as $N \to \infty$, *to the law of a centered Gaussian vector with covariance matrix* $\mathcal{C} \in \mathbb{R}^{n\times n}$, *such that for* $i, j = 1, \dots, n$

$$(\mathcal{C})_{i,j} = \int h_i(z)h_j(z)p_*(\mathrm{d}z) - \mathrm{D}^2\Psi(z_*)[\tilde{h}_i, \tilde{h}_j]. \tag{A.17}$$

Remark 2. (i) The theorem in [7] is stated for $\Sigma = 0$. The same proof applies with our assumptions on Σ without changes. It is likely that these assumptions can be weakened considerably.

(ii) The theorem in [7] contains a statement stronger than the one given here. Indeed, it is shown that the field $\sqrt{N}\,(h(X_N) - h(z_*))_{h\in B'}$ converges weakly to a Gaussian field, while we only stated convergence for finite dimensional distributions. The reason that we do not prove full convergence depends on the fact that our mapping to the Banach space depends of the choice of the observables. It is conceivable that one could get it by using a suitably chosen countable "convergence determining"class, but we have not succeeded in obtaining this result. The convergence that we get is all we need to prove Theorem 3.

[1] A Banach space B is said to be of type 2 if $\ell^2(B) \subseteq C(B)$. Here $\ell^2(B) = \{(x_n) \in B^\infty : \sum_i \|x_i\|^2 < \infty\}$ and $C(B) = \{(x_n) \in B^\infty : \sum_j \epsilon_j x_j$ *converges in probability*$\}$ where (ϵ_n) is a Bernoulli sequence, i.e., a sequence of independent random variables such that $P(\epsilon_n = \pm 1) = \frac{1}{2}$. For more details see [7].

The "natural" space for the central limit theorem in Theorem 3 is the set $\mathcal{M}$ of signed measures, which is not a Banach space. To apply Theorem 4, we need to map $\mathcal{M}$ to a Banach space of type 2.

Lemma 5. *The following properties hold true under the reciprocity condition (R):*

(i) *There exists a Banach space of type 2 $(B, \|\cdot\|)$, and a linear map $T : \mathcal{M} \to B$, continuous on the set $\{Q : Q(\tau = T) = 0\}$. Moreover there exist two continuous maps $\Psi, \Sigma : B \to \mathbb{R}$, where Ψ is bounded and three times Fréchet differentiable and Σ is linear, such that*

$$\frac{\mathrm{d}P_N^\lambda}{\mathrm{d}W^{\otimes N}} = \exp\left\{N\left[\Psi(T(\rho_N)) + \frac{\Sigma(T(\rho_N))}{N}\right]\right\}, \quad \text{a.s.} \tag{A.18}$$

(ii) *For any vector $\Phi = (\Phi_1, \ldots, \Phi_n) \in \mathcal{C}_b{}^n$ there exist $h = (h_1, \ldots, h_n) \subset B'$ such that $(h_i \circ T)(Q) = \int \Phi_i \mathrm{d}Q$, where B' stands for the topological dual of B.*

Proof. The first step consists in giving an alternative expression for $F(Q)$ given in (1.3). Look at the term

$$\int Q(\mathrm{d}y[0,T], \mathrm{d}\lambda)\left\{ y(T)\left(\beta \int Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')\alpha'\eta(t^-)\right)\Big|_{t=\tau(y[0,T])}\right\}.$$

Using the reciprocity condition (R), we obtain

$$\begin{aligned} &b\int Q(\mathrm{d}y[0,T], \mathrm{d}\lambda)\left\{ y(T)\alpha \int Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')\alpha'\eta(t^-)\Big|_{t=\tau(y[0,T])}\right\} \\ &\quad = b\int Q(\mathrm{d}y[0,T], \mathrm{d}\lambda)Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')\alpha\alpha'\mathbf{1}_{\{\tau(y[0,T])\le T\}}\mathbf{1}_{\{\tau(\eta[0,T])<\tau(y[0,T])\}} = (*). \end{aligned}$$

Note that $\mathbf{1}_{\{\tau(\eta[0,T])<\tau(y[0,T])\}} = \mathbf{1}_{\{\tau(\eta[0,T])\le T\}}$ unless $\tau(y[0,T]) \le \tau(\eta[0,T]) \le T$. Thus

$$\begin{aligned} (*) &= b\int Q(\mathrm{d}y[0,T], \mathrm{d}\lambda)Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')\alpha\alpha'\mathbf{1}_{\{\tau(y[0,T])\le T\}}\mathbf{1}_{\{\tau(\eta[0,T])\le T\}} \\ &\quad - b\int Q(\mathrm{d}y[0,T], \mathrm{d}\lambda)Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')\alpha\alpha'\mathbf{1}_{\{\tau(\eta[0,T])\le T\}}\mathbf{1}_{\{\tau(y[0,T])\le\tau(\eta[0,T])\}} \\ &= b\int Q(\mathrm{d}y[0,T], \mathrm{d}\lambda)Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')\alpha\alpha'\mathbf{1}_{\{\tau(y[0,T])\le T\}}\mathbf{1}_{\{\tau(\eta[0,T])\le T\}} \\ &\quad - b\int Q(\mathrm{d}y[0,T], \mathrm{d}\lambda)Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')\alpha\alpha'\mathbf{1}_{\{\tau(\eta[0,T])\le T\}}\mathbf{1}_{\{\tau(y[0,T])<\tau(\eta[0,T])\}} \\ &\quad - b\int Q(\mathrm{d}y[0,T], \mathrm{d}\lambda)Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')\alpha\alpha'\mathbf{1}_{\{\tau(\eta[0,T])\le T\}}\mathbf{1}_{\{\tau(y[0,T])=\tau(\eta[0,T])\}}. \end{aligned}$$

Therefore

$$\begin{aligned} &b\int Q(\mathrm{d}y[0,T], \mathrm{d}\lambda)\left\{ y(T)\alpha \int Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')\alpha'\eta(t^-)\Big|_{t=\tau(y[0,T])}\right\} \\ &\quad = \frac{b}{2}\int Q(\mathrm{d}y[0,T], \mathrm{d}\lambda)Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')\alpha\alpha'\mathbf{1}_{\{\tau(y[0,T])\le T\}}\mathbf{1}_{\{\tau(\eta[0,T])\le T\}} \\ &\qquad - \frac{b}{2}\int Q(\mathrm{d}y[0,T], \mathrm{d}\lambda)Q(\mathrm{d}\eta[0,T], \mathrm{d}\lambda')\alpha\alpha'\mathbf{1}_{\{\tau(\eta[0,T])\le T\}}\mathbf{1}_{\{\tau(y[0,T])=\tau(\eta[0,T])\}} \end{aligned}$$

$$
\begin{aligned}
&= \frac{b}{2}\left[\int Q(\mathrm{d}y[0,T],\mathrm{d}\lambda)\alpha \mathbf{1}_{\{\tau(y[0,T])\le T\}}\right]^2 \\
&\quad - \frac{b}{2}\sum_{t\in[0,T]}\left[\int Q(\mathrm{d}y[0,T],\mathrm{d}\lambda)\alpha \mathbf{1}_{\{\tau(y[0,T])=t\}}\right]^2 .
\end{aligned}
$$

Thus, defining

$$
\begin{aligned}
F_1(Q) := &\int Q(\mathrm{d}y[0,T],\mathrm{d}\lambda) \\
&\times \left\{\int_0^T (1-y(t))\left(1-\mathrm{e}^{-\gamma}\mathrm{e}^{\beta\int Q(\mathrm{d}\eta[0,T],\mathrm{d}\lambda')\alpha'\eta(t)}\right)\mathrm{d}t - \gamma y(T)\right\} \\
&+ \frac{b}{2}\left[\int Q(\mathrm{d}y[0,T],\mathrm{d}\lambda)\alpha y(T)\right]^2
\end{aligned}
\tag{A.19}
$$

and

$$
F_2(Q) := -\frac{b}{2}\sum_{t\in[0,T]}\left[\int Q(\mathrm{d}y[0,T],\mathrm{d}\lambda)\alpha \mathbf{1}_{\{\tau(y[0,T])=t\}}\right]^2, \tag{A.20}
$$

we have that $F(Q) = F_1(Q) + F_2(Q)$. Lemma 2 thus holds also after replacing F by $F_1 + F_2$. Let $M > 0$ be a constant such that, under η, the random parameters α and γ have absolute value less that $M/2$. Now we define the following maps:

$$
\begin{aligned}
T_1 :\ & \begin{array}{ccl} \mathcal{M} & \to & L^2[0,T] \\ Q & \mapsto & \dfrac{1}{2}\displaystyle\int Q(\mathrm{d}y[0,T],\mathrm{d}\lambda)[1-y(t)] \end{array} \\
T_2 :\ & \begin{array}{ccl} \mathcal{M} & \to & L^2([0,T]\times\mathbb{N}) \\ Q & \mapsto & C(M)\displaystyle\int Q(\mathrm{d}y[0,T],\mathrm{d}\lambda)\left[\mathrm{e}^{-\gamma}\frac{(M\beta)^n}{n!}(1-y(t))\right] \end{array} \\
T_3 :\ & \begin{array}{ccl} \mathcal{M} & \to & L^2[0,T] \\ Q & \mapsto & \displaystyle\int Q(\mathrm{d}y[0,T],\mathrm{d}\lambda)\frac{\alpha}{M}y(t) \end{array} \\
T_4 :\ & \begin{array}{ccl} \mathcal{M} & \to & \mathbb{R} \\ Q & \mapsto & \displaystyle\int Q(\mathrm{d}y[0,T],\mathrm{d}\lambda)\frac{\gamma}{M}y(T) \end{array} \\
T_5 :\ & \begin{array}{ccl} \mathcal{M} & \to & \mathbb{R} \\ Q & \mapsto & \displaystyle\int Q(\mathrm{d}y[0,T],\mathrm{d}\lambda)\frac{\alpha}{M}y(T) \end{array} \\
T_6 :\ & \begin{array}{ccl} \mathcal{M} & \to & \mathbb{R} \\ Q & \mapsto & \displaystyle\int Q(\mathrm{d}y[0,T],\mathrm{d}\lambda)\alpha^2 y(T), \end{array}
\end{aligned}
$$

where $C(M)$ is some positive constant such that $C(M)\mathrm{e}^{-\gamma}\mathrm{e}^{M\beta} \le \frac{1}{2}\eta$ almost surely. Note that, for $Q \in \mathcal{M}_1$ and $i = 1, 2, \ldots, 5$, we have that $|T_i(Q)| \le \frac{1}{2}$. Now, let $g : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^\infty$ function such that $g(x) = x$ for $|x| \le 1/2$, $g(x) = 0$ for $|x| > 3/4$ and $\|g\|_\infty \le 3/4$. For

$$
z = (z_1, z_2, z_3, z_4, z_5, z_6) \in L^2[0,T] \times L^2([0,T]\times\mathbb{N}) \times L^2[0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}
$$

we set

$$\Psi(z) := 2\int_0^T g(z_1(t))\mathrm{d}t - \frac{1}{C(M)}\sum_n \int_0^T g(z_2(t,n))g(z_3(t))^n \mathrm{d}t$$
$$- Mg(z_4) + \frac{bM^2}{2}g^2(z_5) \tag{A.21}$$

and

$$\Sigma(z) := -\frac{b}{2}z_6. \tag{A.22}$$

We now claim that, for $Q \in \mathcal{M}_1$ and setting $T = (T_1, T_2, T_3, T_4, T_5, T_6)$,

$$F_1(Q) = \Psi(T(Q)). \tag{A.23}$$

Moreover

$$F_2(\rho_N) = \frac{\Sigma(T(\rho_N))}{N}, \qquad W \otimes \mu\text{-a.s.} \tag{A.24}$$

so that (A.18) holds. Eq. (A.23) is straightforward; (A.24) follows since

$$F_2(\rho_N) = -\frac{b}{2}\sum_t \left(\frac{1}{N}\sum_i \alpha_i \Delta y_i(t)\right)^2 = -\frac{b}{2}\frac{1}{N^2}\sum_i \alpha_i^2 \sum_t (\Delta y_i(t))^2$$
$$= -\frac{b}{2}\frac{1}{N^2}\sum_i \alpha_i^2 y_i(T),$$

where the penultimate equality follows since simultaneous jumps may happen only with zero $(W \otimes \mu)^{\otimes N}$-probability. We thus have that $F_2(\rho_N) = -\frac{b}{2}\frac{1}{N}\int \alpha^2 y(T)\mathrm{d}\rho_N$ and the claim follows by the definition of Σ.

Now set

$$B := L^2[0,T] \times L^2([0,T]\times\mathbb{N}) \times L^2[0,T] \times \mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}^n.$$

Clearly B is a Hilbert space (and hence a Banach space of type 2), and the maps Ψ, Σ are trivially extended to B. Moreover, the map T can be completed to a B-valued map by letting, for $i = 1, 2, \ldots, n$,

$$T_{6+i}(Q) := \int \Phi_i \mathrm{d}Q,$$

where $\Phi = (\Phi_1, \ldots, \Phi_n) \in \mathcal{C}_b{}^n$ is given.

To complete the proof of part (i) of Lemma 5 one has to show the desired regularity of Ψ and Σ. The only nontrivial part is showing regularity of the term

$$\sum_n \int_0^T g(z_2(t,n))g(z_3(t))^n \mathrm{d}t.$$

However, the fact that $\|g\|_\infty \le 3/4$ allows us to control the tails of the sum above; continuity and Fréchet differentiability of any order are obtained by standard estimates, The details are omitted.

Finally, to prove part (ii), for $B \ni (z_1, \ldots, z_6, \ldots, z_{6+n})$, it is enough to define for $i = 1, 2, \ldots, n$

$$h_i(z) = z_{6+i}. \quad \blacksquare$$

Proof of Theorem 3. Now that we have identified a suitable Banach space, Theorem 3 immediately follows from Theorem 4 applied to the sequence $Z_i = T(\delta_{\{y_i[0,T],\lambda\}})$ taking values on $(B, \|\cdot\|)$. Notice that in our setting, $\Omega = (\mathcal{D}[0,T] \times \mathbb{R}^2)^N$ and $\mathbb{P} = (W \otimes \mu)^N$. Theorem 3 is guaranteed by the following three facts:

1. $P_N \equiv \pi_N$, where π_N is the probability appearing in Theorem 4.
2. $\left(\int \Phi_i \mathrm{d}\rho_N - \int \Phi_i \mathrm{d}Q_*\right)_{i=1}^n = (h_i(X_N) - h_i(z_*))_{i=1}^n$.
3. $\int(\phi_i - \phi_i^*)(\phi_j - \phi_j^*)\mathrm{d}Q_* - \mathrm{D}^2 F(Q_*)[\hat{\Phi}_i, \hat{\Phi}_j] = \int h_i(z)h_j(z)p_*(\mathrm{d}z) - \mathrm{D}^2\Psi(z_*)[\tilde{h}_i, \tilde{h}_j]$.

Point 1. follows from the definition of π_N and from Eqs. (A.23) and (A.24). Point 2. is a consequence of the fact that $z_* = T(Q_*)$ and $h_i \circ T(Q_*) = \int \Phi_i \mathrm{d}Q_*$. Point 3. will be proved in detail in Lemma 8 (see in particular Eq. (A.38)). An immediate application of Eqs. (A.38) and (A.33) finally guarantees the validity of (1.7).

Assuming point 3., it remains to show the validity of the central limit theorem in B. In other words we need to check the five assumptions of Theorem 4. (B.1), (B.2) and (B.5) are easy to see. (B.3) and (B.4) are not straightforward. The rest of this section is devoted to the proof that these two assumptions are satisfied.

We begin to prove (B.3). We define two sequences of measures on B as follows:

$$p_N(\cdot) = \mathcal{P}_N \circ T^{-1}(\cdot) \qquad w_N(\cdot) = \mathcal{W}_N \circ T^{-1}(\cdot).$$

From (A.18) it can be shown that

$$\frac{\mathrm{d}p_N}{\mathrm{d}w_N} = \mathrm{e}^{N\left(\Psi + \frac{\Sigma}{N}\right)} \tag{A.25}$$

for Ψ and Σ as defined in (A.21) and (A.22).

By the contraction principle (see Theorem 4.2.1 in [20]), the sequence $(p_N)_N$ satisfies a LDP with the good rate function $J(z) = \inf_{Q\in T^{-1}(z)} I(Q)$. Q_* being the unique zero for I, J has a unique zero $z_* = T(Q_*)$.

A LDP for the sequence $(p_N)_N$ can be obtained in an alternative way. Indeed, we notice that w_N is the law of the random variables

$$X_N = \frac{1}{N}\sum_{i=1}^N Z_i \ \in B,$$

where Z_i are i.i.d. B-valued random variables with law w. Thus we have that $(w_N)_N$ satisfies a (weak) LDP with rate function Λ^*, with $\Lambda^*(z) := \sup_{\varphi\in B'}\{\varphi(z) - \Lambda(\varphi)\}$ and $\Lambda(\varphi) := \ln\int \mathrm{e}^{\varphi(z)}\, w(\mathrm{d}z)$. Thus, applying Varadhan's lemma, $(p_N)_N$ satisfies a (weak) LDP with rate function $\Lambda^*(z) - \Psi(z)$. Since the rate function is unique, it follows that

$$J(z) = \Lambda^*(z) - \Psi(z).$$

Having proved already that $J(z)$ has a unique zero, the proof of (B.3) is completed.

We are thus left to show (B.4): for each $\lambda \in B'$ such that $\tilde{h} = \int zh(z)p_*(\mathrm{d}z) \neq 0$ we have

$$\int h^2(z)p_*(\mathrm{d}z) - \mathrm{D}^2\Psi(z^*)[\tilde{h}, \tilde{h}] > 0; \tag{A.26}$$

where p and p_* are defined in Theorem 4.

This proof is rather technical and long. We divide it into three steps. We first show that the measure p such that $\frac{dp}{dw} = e^{D\Psi(z_*)}$ is exactly the law of the random variable $T(\delta_{\{y[0,T],\lambda\}})$ induced by Q_*. This argument is then used in the second step to ensure the positivity of a suitable functional $\mathcal{H} : \mathcal{C}_b \times \mathcal{C}_b \to \mathbb{R}$. In the last part we see how to relate $\mathcal{H}$ to assumption (B.4).

Step 1: The key result of this first step is given in Lemma 6 below. We look at the measure p on B, defined by

$$\frac{dp}{dw}(z) = e^{D\Psi(z_*)[z]}, \quad \text{with } z_* = T(Q_*)$$

where, as already seen, w represents the law of $T(\delta_{\{y[0,T],\lambda\}})$ induced by $W \otimes \mu$.

We shall prove in Lemma 6 that p is the law of $T(\delta_{\{y[0,T],\lambda\}})$ induced by Q_*.

Lemma 6. *The measure p is the law of $T(\delta_{\{y[0,T],\lambda\}})$ induced by Q_*.*

Proof. We first prove the following two claims:

(i)

$$DF(Q_*)[\delta_{\{y[0,T],\lambda\}}] = \log \frac{dQ_*}{d(W \otimes \mu)}(y[0,T],\lambda), \tag{A.27}$$

for $W \otimes \mu$-almost all $(y[0,T],\lambda)$.

(ii)

$$DF_2(Q)[r] = 0 \tag{A.28}$$

for all $Q \in \mathcal{M}_1$ such that $\int \alpha' 1_{\{\tau(y[0,T])=t\}} dQ = 0$, where F_2 is defined in (A.20).

To prove the claim we need to compute $DF(Q_*)[\delta_{\{y[0,T],\lambda\}}]$, i.e. the Fréchet derivative of the function F at Q_* in the direction $\delta_{\{y[0,T],\lambda\}}$. An explicit computation reveals that for $Q \in \mathcal{M}_1$ and $r \in \mathcal{M}$, $DF(Q)[r]$ is well defined and in particular

$$\begin{aligned} DF(Q)[r] &= \lim_{h\to 0} \frac{F(Q+hr) - F(Q)}{h} \\ &= \int dQ \int_0^T (1-y(t)) m_r(t)\, e^{-\gamma+\beta m_Q(t)} dt + \int dr \int_0^T (1-y(t))(1-e^{-\gamma+\beta m_Q(t)}) dt \\ &\quad + \int dr\, \left[-y(T)(\gamma - \beta m_Q(\tau^-))\right] + \int dQ\, \left[y(T) m_r(\tau^-)\right], \end{aligned}$$

where we have put as usual $m_p(t) = \int \alpha' \eta(t) p(d\eta[0,T], d\alpha')$ for $p \in \mathcal{M}$, $t \in [0,T]$.

We now compute $DF(Q_*)[\delta_{\{y[0,T],\lambda\}}]$. Notice that $m_{\delta_{\{y[0,T],\lambda\}}}(t) = \alpha y(t) = 0$ for all $t < \tau(y[0,T])$ and so

$$\begin{aligned} DF(Q_*)[\delta_{\{y[0,T],\lambda\}}] = &\int_0^T (1-y(t))(1-e^{-\gamma+\beta m_{Q_*}(t)}) dt \\ &- \left[y(T)(\gamma - \beta m_{Q_*}(\tau^-))\right]. \end{aligned} \tag{A.29}$$

By virtue of Girsanov's formula for Markov chains it can be seen that

$$\int_0^T (1-y(t))(1-e^{-\gamma+\beta m_{Q_*}(t)}) dt - \left[y(T)(\gamma - \beta m_{Q_*}(\tau^-))\right] = \log \frac{dP^{Q_*}}{d(W \otimes \mu)}$$

where P^Q is the law of the Markov process with generator given in (A.13). (A.27) thus follows since $P^{Q_*} = Q_*$ as shown in the proof of Theorem 2. As regards (ii), notice that

$$\begin{aligned} F_2(Q+hr) &= -\frac{b}{2}\sum_t \left(\int \alpha' \mathbf{1}_{\{\tau(y[0,T])=t\}} \mathrm{d}\{Q+hr\}\right)^2 \\ &= -\frac{b}{2}\sum_t \left(\int \alpha' \mathbf{1}_{\{\tau(y[0,T])=t\}} \mathrm{d}Q\right)^2 \\ &\quad - bh \sum_t \left(\int \alpha' \mathbf{1}_{\{\tau(y[0,T])=t\}} \mathrm{d}Q \int \alpha' \mathbf{1}_{\{\tau(y[0,T])=t\}} \mathrm{d}r\right) \\ &\quad - \frac{b}{2} h^2 \sum_t \left(\int \alpha' \mathbf{1}_{\{\tau(y[0,T])=t\}} \mathrm{d}r\right)^2 . \end{aligned}$$

When $\int \alpha' \mathbf{1}_{\{\tau(y[0,T])=t\}} \mathrm{d}Q = 0$, the first two terms of the last expression vanish. Moreover under the same hypothesis $F_2(Q) = 0$. Hence

$$\begin{aligned} \mathrm{D}F_2(Q)[r] &= \lim_{h\to 0} \frac{1}{h} [F_2(Q+hr) - F_2(Q)] \\ &= \lim_{h\to 0} -\frac{b}{2}\frac{h^2}{h} \sum_t \left(\int \alpha' \mathbf{1}_{\{\tau(y[0,T])=t\}} \mathrm{d}r\right)^2 = 0. \end{aligned} \tag{A.30}$$

Notice that in writing the latter equality we have implicitly used the fact that $\sum_t \left(\int \alpha' \mathbf{1}_{\{\tau(y[0,T])=t\}} \mathrm{d}r\right)^2 < \infty$. This is true since for any $r \in \mathcal{M}$

$$0 \le \sum_t \left(\int \alpha' \mathbf{1}_{\{\tau(y[0,T])=t\}} \mathrm{d}r\right)^2 \le \|\alpha\|^2 \cdot |r|^2_{TV} < \infty$$

where $\|\alpha\|$ stands for the supremum of α in the support of μ and $|r|_{TV}$ denotes the *total variation* of r.

As a corollary of claim (ii) above we see that

$$\begin{aligned} D\Psi(T(Q_*))[T(r)] &= \lim_{h\to 0} \frac{\Psi(T(Q_* + hr)) - \Psi(T(Q_*))}{h} \\ &= \lim_{h\to 0} \frac{F_1(Q_* + hr) - F_1(Q_*)}{h} = \mathrm{D}F_1(Q_*)[r] = \mathrm{D}F(Q_*)[r]. \end{aligned}$$

Here we have used (A.30), the fact that $\int \alpha' \mathbf{1}_{\{\tau(y[0,T])=t\}} \mathrm{d}Q_* = 0$ since $Q_* \ll W \otimes \mu$ and Eq. (A.23).

Going back to the statement of the lemma, we see that for h measurable and bounded

$$\begin{aligned} \int h(z) p(\mathrm{d}z) &= \int h(z) \mathrm{e}^{D\Psi(z_*)[z]} w(\mathrm{d}z) \\ &= \int h(T(\delta_{\{y[0,T],\lambda\}})) \mathrm{e}^{\mathrm{D}F(Q_*)[\delta_{\{y[0,T],\lambda\}}]} (W \otimes \mu)(\mathrm{d}y[0,T], \mathrm{d}\lambda) \\ &= \int h(T(\delta_{\{y[0,T],\lambda\}})) \mathrm{d}Q_* \end{aligned}$$

where in the last equality we have used (A.27). ■

Step 2: The key result of the second step is given in Lemma 7 below. It involves the measures $\hat{\Phi}$ and $\hat{\Phi}^*$ defined in (1.6). First of all, it is not difficult to show that $\hat{\Phi}$ is absolutely continuous w.r.t. Q_* and in particular

$$\frac{\mathrm{d}\hat{\Phi}}{\mathrm{d}Q_*} = \Phi - \Phi^*. \tag{A.31}$$

Indeed, observe that, given $\hat{\Phi}$ as in (1.6), we have

$$\begin{aligned}\hat{\Phi}(S) &= \int_{\mathcal{M}_0} \left[R(S) \int \Phi \mathrm{d}R \right] \nu_*(\mathrm{d}R) \\ &= \int_{\mathcal{D}[0,T]\times \mathrm{Supp}(\mu)} (\mathbf{1}_{\{(y[0,T],\lambda)\in S\}} - Q_*(S)) \cdot \left(\int \Phi \, \mathrm{d}\delta_{\{y[0,T],\lambda\}} - \int \Phi \mathrm{d}Q_* \right) \mathrm{d}Q_* \\ &= \int_{\mathcal{D}[0,T]\times \mathrm{Supp}(\mu)} \left[(\mathbf{1}_{\{(y[0,T],\lambda)\in S\}} - Q_*(S)) \cdot (\Phi(y[0,T],\lambda) - \Phi^*) \right] \mathrm{d}Q_*\end{aligned}$$

for any $S \subset \mathcal{D}[0,T] \times \mathrm{Supp}(\mu)$. The second equality follows since ν_* is the law of the random variable $\delta_{\{y[0,T],\lambda\}} - Q_*$ induced by Q_*.

Notice that $Q_*(S) \int_{\mathcal{D}[0,T]\times \mathrm{Supp}(\mu)} (\Phi - \Phi^*) \mathrm{d}Q_* = 0$, Φ^* being the expectation under Q_* of $\Phi(\cdot)$. Hence

$$\hat{\Phi}(S) = \int_S (\Phi - \Phi^*) \mathrm{d}Q_*$$

and (A.31) follows.

Lemma 7. *Given Φ_1 and Φ_2 in $\mathcal{C}_b$, let*

$$\mathcal{H}(\Phi_1, \Phi_2) := \mathrm{Cov}_{Q_*}(\Phi_1, \Phi_2) - \mathrm{D}^2 F(Q_*)[\hat{\Phi}_1, \hat{\Phi}_2]; \tag{A.32}$$

where $\mathrm{Cov}_{Q_*}(\Phi_1, \Phi_2) := \int (\Phi_1 - \Phi_1^*)(\Phi_2 - \Phi_2^*) \mathrm{d}Q_*$. *Then*

$$\mathcal{H}(\Phi, \Phi) > 0, \quad \textit{for all } \Phi \textit{ such that } \hat{\Phi} \neq 0.$$

Proof. A tedious but straightforward computation provides the second-order derivative of F:

$$\begin{aligned}\mathrm{D}^2 F(Q_*)[\hat{\Phi}_i, \hat{\Phi}_j] = E^{Q_*} \Bigg[&\int_0^T -(1-y(t))\beta^2 m_{\hat{\Phi}_i}(t) m_{\hat{\Phi}_j}(t) \mathrm{e}^{-\gamma + \beta m_{Q_*}(t)} \mathrm{d}t \\ &- (\Phi_j - \Phi_j^*) \int_0^T (1-y(t))\beta m_{\hat{\Phi}_i}(t) \mathrm{e}^{-\gamma + \beta m_{Q_*}(t)} \mathrm{d}t \\ &- (\Phi_i - \Phi_i^*) \int_0^T (1-y(t))\beta m_{\hat{\Phi}_j}(t) \mathrm{e}^{-\gamma + \beta m_{Q_*}(t)} \mathrm{d}t \Bigg] \\ &+ E^{\hat{\Phi}_i} \left[y(T)\beta m_{\hat{\Phi}_j}(\tau^-) \right] + E^{\hat{\Phi}_j} \left[y(T)\beta m_{\hat{\Phi}_i}(\tau^-) \right].\end{aligned}$$

Notice that we have written β instead of $b\alpha$: the reciprocity condition is not necessary in this calculation. We now show that $\mathcal{H}(\Phi, \Phi)$ is the expected value of a square. Indeed

$$\begin{aligned}\mathcal{H}(\Phi, \Phi) &= \mathrm{Cov}_{Q_*}(\Phi, \Phi) - \mathrm{D}^2 F(Q_*)[\hat{\Phi}, \hat{\Phi}] \\ &= E^{Q_*}[(\Phi - \Phi^*)^2] + E^{Q_*} \left[\int_0^T (1-y(t))\beta^2 [m_{\hat{\Phi}}(t)]^2 \mathrm{e}^{-\gamma + \beta m_{Q_*}(t)} \mathrm{d}t \right.\end{aligned}$$

$$+ \; 2(\Phi - \Phi^*) \int_0^T (1 - y(t))\beta m_{\hat{\Phi}}(t) \mathrm{e}^{-\gamma + \beta m_{Q_*}(t)} \mathrm{d}t \Bigg] - 2E^{\hat{\Phi}} \left[y(T)\beta m_{\hat{\Phi}}(\tau^-) \right].$$

The latter expectation can be rewritten as

$$E^{\hat{\Phi}} \left[y(T)\beta m_{\hat{\Phi}}(\tau^-) \right] = E^{Q_*} \left[(\Phi - \Phi^*) \int_0^T (1 - y(t))\beta m_{\hat{\Phi}}(t^-) \mathrm{d}N(t) \right]$$

where we have used (A.31) and where $(N(t))_{t \in [0,T]}$ defined by $N(t) := \mathbf{1}_{\{\tau \geq t\}}$ is the Poisson process with intensity $\int_0^t (1 - y(s))\mathrm{e}^{-\gamma + \beta m_{Q_*}(s)} \mathrm{d}s$.

Recall that $M(t) = N(t) - \int_0^t (1 - y(s))\mathrm{e}^{-\gamma + \beta m_{Q_*}(s)} \mathrm{d}s$ defined in (1.8) is nothing but its *compensated* Q_*-martingale. Hence

$$\mathcal{H}(\Phi, \Phi) = E^{Q_*}[(\Phi - \Phi^*)^2] + E^{Q_*} \left[\int_0^T (1 - y(t))\beta^2 [m_{\hat{\Phi}}(t)]^2 \mathrm{e}^{-\gamma + \beta m_{Q_*}(t)} \mathrm{d}t \right]$$
$$- E^{Q_*} \left[2(\Phi - \Phi^*) \int_0^T (1 - y(t))\beta m_{\hat{\Phi}}(t) \mathrm{d}M(t) \right].$$

By the isometry property of square integrable martingales (and relying on the same argument as was used to prove (2.6)), we have

$$E^{Q_*} \left[\int_0^T (1 - y(t))\beta^2 m_{\hat{\Phi}}(t)^2 \mathrm{e}^{-\gamma + \beta m_{Q_*}(t)} \mathrm{d}t \right]$$
$$= E^{Q_*} \left[\left(\int_0^T (1 - y(t))\beta m_{\hat{\Phi}}(t) \mathrm{d}M(t) \right)^2 \right].$$

Hence

$$\mathcal{H}(\Phi, \Phi) = E^{Q_*} \left[\left((\Phi - \Phi^*) - \int_0^T (1 - y(t))\beta m_{\hat{\Phi}}(t) \mathrm{d}M(t) \right)^2 \right]. \tag{A.33}$$

$\mathcal{H}(\Phi, \Phi)$ is thus the expected value of a square; hence it cannot be negative. For this reason, we simply need to prove that it is non-zero. Without loss of generality we take $\Phi^* = 0$. Suppose by way of contradiction that $\mathcal{H}(\Phi, \Phi) = 0$. Then necessarily

$$\left(\Phi(y[0, T], \lambda) - \int_0^T (1 - y(t))\beta m_{\hat{\Phi}}(s) \mathrm{d}M(s) \right) = 0, \quad Q_* \text{ a.s.}$$

Using the fact that

$$m_{\hat{\Phi}}(s) = \int \alpha y(s) \; \hat{\Phi}(\mathrm{d}y[0, T], \mathrm{d}\lambda) = \int \alpha y(s) \; \Phi(y[0, T], \lambda) \; Q_*(\mathrm{d}y[0, T], \mathrm{d}\lambda),$$

where the last equality follows since $\frac{\mathrm{d}\hat{\Phi}}{\mathrm{d}Q_*} = \Phi$. We rewrite the expression above as

$$\Phi(y[0, T], \lambda) = \int_0^T (1 - y(t))\beta \left[\int \alpha y(s) \Phi(y[0, T], \lambda) \mathrm{d}Q_* \right] \mathrm{d}M(s) \,, \quad Q_*\text{- a.s.} \tag{A.34}$$

On the other hand, define $\Phi_t = E^{Q_*}[\Phi | \mathcal{F}_t]$, where

$$\mathcal{F}_t = \sigma\{y_s : 0 \leq s \leq t; \; \lambda\}.$$

Notice that

$$\int y(t)\Phi(\cdot)\mathrm{d}Q_* = E^{Q_*}[\alpha y(t)\Phi(\cdot)] = E^{Q_*}[\alpha y(t)E^{Q_*}[\Phi(\cdot)|\mathcal{F}_t]] = \int \alpha y(t)\Phi_t \mathrm{d}Q_*.$$

Taking the conditional expectation in (A.34), we obtain

$$\begin{aligned}\Phi_t &= E^{Q_*}\left[\int_0^T (1-y(t))\beta\left(\int \alpha y(s)\Phi_s \mathrm{d}Q_*\right)\mathrm{d}M(s)\middle|\mathcal{F}_t\right], \quad Q_* \text{ a.s.}\\ &= \int_0^t (1-y(s))\beta\left(\int \alpha y(s)\Phi_s \mathrm{d}Q_*\right)\mathrm{d}M(s), \quad Q_* \text{ a.s.}\end{aligned}$$

We now take the L^2-norm on both sides. For all $t \in [0,T]$ we have

$$\begin{aligned}\|\Phi_t\|^2_{L^2(Q_*)} &= \left\|\int_0^t (1-y(s))\beta\left(\int \alpha y(s)\Phi_s \mathrm{d}Q_*\right)\mathrm{d}M(s)\right\|^2_{L^2(Q_*)}\\ &= E^{Q_*}\left[\int_0^t (1-y(t))\beta^2\left(\int \alpha y(s)\Phi_s \mathrm{d}Q_*\right)^2 \mathrm{e}^{-\gamma+\beta m_{Q_*}(s)}\mathrm{d}s\right].\end{aligned}$$

Notice that $\left(\int \alpha y(s)\Phi_s \mathrm{d}Q_*\right)^2 \leq \|\alpha\|^2\left(\int \Phi_s \mathrm{d}Q_*\right)^2 \leq \|\alpha\|^2 \int \Phi_s^2 \mathrm{d}Q_* \leq \|\alpha\|^2 \int \Phi_t^2 \mathrm{d}Q_* = \|\alpha\|^2\|\Phi_t\|^2_{L^2(Q_*)}$, where $t \geq s$. The first inequality follows since $y \in \{0;1\}$; the second one is trivial and the latter one is due to the fact that $(\Phi_s^2)_s$ is a submartingale and thus its expected value is an increasing function of time. Then

$$\begin{aligned}\|\Phi_t\|^2_{L^2(Q_*)} &\leq \|\alpha\|^2 E^{Q_*}\left[\int_0^t (1-y(t))\beta^2\|\Phi_t\|^2_{L^2(Q_*)}\mathrm{e}^{-\gamma+\beta m_{Q_*}(s)}\mathrm{d}s\right]\\ &\leq t\epsilon^{-1}\ \|\Phi_t\|^2_{L^2(Q_*)}\end{aligned}$$

where $0 < \epsilon < \infty$ is a constant such that $\|\alpha\|^2\|\beta\|^2 E^{Q_*}[\mathrm{e}^{-\gamma+\beta m_{Q_*}(s)}] \leq \epsilon^{-1}$. As a consequence, $\Phi_s = 0$, Q_* a.s. for $s \in [0,\epsilon)$.

This argument can be iterated, defining $\Phi_t^{(2)} := \Phi_{t+\epsilon}$. The same argument shows that $\Phi_s^{(2)} = 0$, Q_* a.s. for $s \in [0,\epsilon)$; hence $\Phi_t = 0$, Q_* a.s. for $s \in [0,2\epsilon)$. Eventually we extend the statement to $s \in [0,T]$. As $\Phi_T = \Phi$, we would have $\hat{\Phi} = 0$ and this gives a contradiction. Hence the theorem follows. ■

Step 3: Consider $\lambda_1, \lambda_2 \in B'$. Since $\lambda_i \circ T$, for $i = 1,2$, are in the topological dual of $\mathcal{M}$, there exist $\Phi_1, \Phi_2 \in \mathcal{C}_b$ such that $\lambda_i \circ T(Q) = \int \Phi_i \mathrm{d}Q$. We define

$$\mathrm{Cov}_{p_*}(\lambda_1,\lambda_2) = \int \lambda_1(z)\lambda_2(z)p_*(\mathrm{d}z) \quad \text{and} \quad \tilde{\lambda}_i = \int z\lambda_i(z)p_*(\mathrm{d}z); \quad i = 1,2$$

where we recall that p_*, defined in (B.4) of Theorem 4, is the centered version of the law of $T(\delta_{\{y[0,T],\lambda\}})$ induced by Q_*. Then the following result holds true:

Lemma 8. (i)

$$\begin{aligned}&\mathrm{Cov}_{p_*}(\lambda_1,\lambda_2) = \mathrm{Cov}_{Q_*}(\Phi_1,\Phi_2);\\ &\mathrm{D}^2\Psi(z_*)[\tilde{\lambda}_1,\tilde{\lambda}_2] = \mathrm{D}^2 F(Q_*)[\hat{\Phi}_1,\hat{\Phi}_2].\end{aligned}$$

(ii) *For* λ_i, $i = 1,2$, *we have*

$$\mathrm{Cov}_{p_*}(\lambda_i,\lambda_i) - \mathrm{D}^2\Psi(z_*)[\tilde{\lambda}_i,\tilde{\lambda}_i] > 0.$$

Proof. Point (i). By the definition of p_* and $\lambda_i,\ i = 1, 2$, we see that

$$\begin{aligned}\mathrm{Cov}_{p_*}(\lambda_1, \lambda_2) &= \int [T_{6+1}(\delta_{\{y[0,T],\lambda\}}) - T_{6+1}(Q_*)][T_{6+2}(\delta_{\{y[0,T],\lambda\}}) - T_{6+2}(Q_*)]\mathrm{d}Q_* \\ &= \int [\Phi_1 - \Phi_1^*][\Phi_2 - \Phi_2^*]\mathrm{d}Q_*,\end{aligned}$$

where we have used the fact that $\lambda_i \circ T(Q) = T_{6+i}(Q) = \int \Phi_i \mathrm{d}Q$.

As regards the second statement, we first prove the following claim:

$$\tilde{\lambda}_i = T(\hat{\Phi}_i); \quad i = 1, 2. \tag{A.35}$$

To show the validity of (A.35), we use the following two facts:

$$\tilde{\lambda}_i = E^{Q_*}\left\{[T(\delta_{\{y[0,T],\lambda\}}) - T(Q_*)][\Phi_i(y[0,T],\lambda) - \Phi_i^*]\right\};$$
$$\hat{\Phi}_i = E^{Q_*}\left\{[\delta_{\{y[0,T],\lambda\}} - Q_*][\Phi_i(y[0,T],\lambda) - \Phi_i^*]\right\}.$$

The former follows by definition of p_*, λ and $T_{6+i}(Q)$, whereas the latter is a consequence of the definition of $\hat{\Phi}$ given in (1.6).

(A.35) is a consequence of the fact that T is both linear and continuous; hence we are allowed to interchange the operator T with the expectation.

Having proved (A.35), we compute the second-order Fréchet derivatives of the function Ψ as follows:

$$\mathrm{D}^2\Psi(z_*)[\tilde{\lambda}_1, \tilde{\lambda}_2] = \lim_{k\to 0} \frac{\mathrm{D}\Psi(z_* + k\tilde{\lambda}_2)[\tilde{\lambda}_1] - \mathrm{D}\Psi(z_*)[\tilde{\lambda}_1]}{k}. \tag{A.36}$$

Notice that, by the linearity of T and by (A.35), we have that

$$z_* + k\tilde{\lambda}_2 = T(Q_* + k\hat{\Phi}_2), \qquad z_* = T(Q_*).$$

Thus

$$\lim_{k\to 0} \frac{\mathrm{D}\Psi(z_* + k\tilde{\lambda}_2)[\tilde{\lambda}_1] - \mathrm{D}\Psi(z_*)[\tilde{\lambda}_1]}{k} = \lim_{k\to 0} \frac{\mathrm{D}\Psi(z_* + k\tilde{\lambda}_2)[\lambda_1] - \mathrm{D}\Psi(z_*)[\tilde{\lambda}_1]}{k}.$$

We now claim that

$$\begin{aligned}&\lim_{k\to 0} \frac{\mathrm{D}\Psi(z_* + k\tilde{\lambda}_2)[\lambda_1] - \mathrm{D}\Psi(z_*)[\tilde{\lambda}_1]}{k} \\ &\quad = \lim_{k\to 0} \frac{\mathrm{D}F(Q_* + k\hat{\Phi}_2)[\hat{\Phi}_1] - \mathrm{D}F(Q_*)[\hat{\Phi}_1]}{k}.\end{aligned} \tag{A.37}$$

By (A.28) we see that $\mathrm{D}F_2(Q_*)[\cdot] = 0$ since $Q_* \ll (W \otimes \eta)$. Moreover we have both $\mathrm{D}F_2(\hat{\Phi}_i)[\cdot] = 0$ and $\mathrm{D}F_2(Q_* + k\hat{\Phi}_i)[\cdot] = 0$ since $\hat{\Phi}_i$ is absolutely continuous with respect to Q_*. This proves (A.37). Finally we use the fact that F is Fréchet differentiable:

$$\lim_{k\to 0} \frac{\mathrm{D}F(Q_* + k\hat{\Phi}_2)[\hat{\Phi}_1] - \mathrm{D}F(Q_*)[\hat{\Phi}_1]}{k} = \mathrm{D}^2F(Q_*)[\hat{\Phi}_1, \hat{\Phi}_2].$$

We have thus proved that $\mathrm{D}^2\Psi(z_*)[\tilde{\lambda}_1, \tilde{\lambda}_2] = \mathrm{D}^2F(Q_*)[\hat{\Phi}_1, \hat{\Phi}_2]$.

As regards point (ii), we notice that

$$\begin{aligned}\mathrm{Cov}_{p_*}(\lambda_i, \lambda_i) - \mathrm{D}^2\Psi(z_*)[\tilde{\lambda}_i, \tilde{\lambda}_i] &= \mathrm{Cov}_{Q_*}(\Phi_i, \Phi_i) - \mathrm{D}^2F(Q_*)[\hat{\Phi}_i, \hat{\Phi}_i] \\ &= \mathcal{H}(\Phi_i, \Phi_i),\end{aligned} \tag{A.38}$$

where $\mathcal{H}$ has been defined in Eq. (A.32). Hence by Lemma 7 the positivity condition is ensured and the theorem follows. ■

By virtue of Lemma 8, for any $\lambda \in B'$ such that $\tilde{\lambda} \neq 0$, (A.26) holds true. As a consequence, assumption (B.4) is ensured and thus Theorem 3 is proved. ■

References

[1] D. Dembo, J.D. Deuschel, D. Duffie, Large portfolio losses, Finance and Stochastics 8 (2004) 3–16.
[2] K. Giesecke, L. Goldberg, A top-down approach to multi-name credit, Working paper, 2007.
[3] P. Schönbucher, Portfolio losses and the term structure of loss transition rates: A new methodology for the pricing of portfolio credit derivatives, Working paper, 2006.
[4] R. Frey, J. Backhaus, Credit derivatives in models with interacting default intensities: A Markovian approach, 2006. Preprint.
[5] R. Frey, J. Backhaus, Pricing and hedging of portfolio credit derivatives with interacting default intensities, International Journal of Theoretical and Applied Finance 11 (6) (2008) 611–634.
[6] P. Dai Pra, W.J. Runggaldier, E. Sartori, M. Tolotti, Large portfolio losses; A dynamic contagion model, Annals Applied Probability, 2008 (in press). arXiv:0704.1348v2.
[7] E. Bolthausen, Laplace approximations for sums of independent random vectors, Probability Theory and Related Fields 72 (1986) 305–318.
[8] G. Ben Arous, M. Brunaud, Méthode de Laplace: étude variationnelle des fluctuations de diffusions de type "champ moyens", Stochastics and Stochastic Reports 31 (1990) 79–144.
[9] P. Dai Pra, F. Den Hollander, McKean–Vlasov limit for interacting random processes in random media, Journal of Statistical Physics 84 (3–4) (1996) 735–772.
[10] K. Giesecke, S. Weber, Cyclical correlations, credit contagion and portfolio losses, Journal of Banking and Finance 28 (12) (2005) 3009–3036.
[11] H. Pham, Some applications and methods of large deviations in finance and insurance, in: Paris–Princeton Lectures on Mathematical Finance, in: Lectures notes in Mathematics, Springer, 2007.
[12] S.N. Ethier, T.G. Kurtz, Markov Processes, Characterization and Convergence, John Wiley and Sons, 1986.
[13] F. Comets, Th. Eisele, Asymptotic dynamics, non-critical and critical fluctuations for a geometric long-range interacting model, Communications in Mathematical Physics 118 (1988) 531–567.
[14] P. Embrechts, R. Frey, A. McNeil, Quantitative risk management: concepts, techniques and tools, Princeton Series in Finance (2005).
[15] R. Frey, A. McNeil, VaR and expected shortfall in portfolios of dependent credit risks: Conceptual and practical insights, Journal of Banking and Finance 26 (2002) 1317–1334.
[16] M. Crouhy, D. Galai, R. Mark, A comparative analysis of current credit risk models, Journal of Banking and Finance 24 (2000) 59–117.
[17] M. Tolotti, The impact of contagion on large portfolios. Modeling aspects, Ph.D. Thesis, Scuola Normale Superiore, 2008.
[18] S.R.S. Varadhan, Large Deviations and Applications, Society for Industrial and Applied Mathematics, Philadelphia, 1984.
[19] H. Brezis, Analyse fonctionnelle — Théorie et applications, Masson Editeur, Paris, 1983.
[20] A. Dembo, O. Zeitouni, Large Deviations Techniques, Jones and Bartlett Publishers, Boston, London, 1993.