

DEFAULT CLUSTERING IN LARGE POOLS: LARGE DEVIATIONS

KONSTANTINOS SPILIOPOULOS

*Department of Mathematics & Statistics
Boston University
Boston, MA 02215*

RICHARD B. SOWERS

*Department of Mathematics
University of Illinois at Urbana–Champaign
Urbana, IL 61801*

ABSTRACT. We study large deviations and rare default clustering events in a dynamic large heterogeneous portfolio of interconnected components. Defaults come as Poisson events and the default intensities of the different components in the system interact through the empirical default rate and via systematic effects that are common to all components. We establish the large deviations principle for the empirical default rate for such an interacting particle system. The rate function is derived in an explicit form that is amenable to numerical computations and derivation of the most likely path to failure for the system itself. Numerical studies illustrate the theoretical findings. An understanding of the role of the preferred paths to large default rates and the most likely ways in which contagion and systematic risk combine to lead to large default rates would give useful insights into how to optimally safeguard against such events.

1. INTRODUCTION

The financial crisis of 2007-2008 challenged the mathematical finance community to understand *connect- edness* in financial systems. Appropriate models need to be developed to understand how risk can propagate between financial objects which might have heretofore been modeled as closed systems.

It is possible that initial shocks, such as changes in interest rate values, changes of commodities prices or reduction in global economic growth, could trigger contagion effects, e.g., [19]. It is likely then that some transmission mechanism, such as financial linkages or simply investor irrationality, could cause components of the system to be affected by the initial shock. Reduce-form point process models of correlated default, that are usually based on counting processes, are often used to access portfolio credit risk in portfolios of defaultable assets such as loans and corporate defaults. In these models defaults arrive at intensities that are governed by a given system of stochastic differential equations. Due to the size of the portfolios computing the distribution of the loss from default in these models is often challenging. Main US banks for example may easily have 20,000 wholesale loans and 50,000 – 100,000 mid-market and commercial loans. Mortgage pools with size of 10,000 are often common. Simulation and analysis of such pools is non-trivial and often quite burdensome.

In this work we focus on using dynamic portfolio credit risk models to study rare events in large portfolios and default clustering. We statistically model failure via the classical framework of point processes. However, we include several distinct and meaningful sources of randomness that can lead to failure of the system itself.

E-mail addresses: kspiliop@math.bu.edu, r-sowers@illinois.edu.

Date: November 4, 2018.

1991 Mathematics Subject Classification. 60F10-60G55-91G40-91G80.

Key words and phrases. Large Deviations, Heterogeneous Networks, Large Pools, Rare Events, Default Clustering.

The authors would like to thank Kay Giesecke for helpful comments.

The intensities form an interacting particle system, where the interaction happens through feedback terms and exposure to common systematic risks. The foundation of the empirically motivated model that we study lies in [17]. In this work various properties of such models were investigated, finding that contagion and systematic risk give meaningful insights into clustering of defaults.

Typical (law of large numbers) and central limit type behavior for such models as the number of components in the system grows have been studied in [17, 18, 20]. Our interest here is tail behavior and rare events. One of the principle theoretical attractions of structured finance was that, supposedly, large pools of assets were regular enough that one could accurately price tail events. In hindsight, however, interaction in large pools proved an Achilles' heel. Interconnections often make a system robust, but they can also act as conduits for failure. In this paper, we seek to understand the mathematics of rare systemic collapse in large interacting systems.

We consider a large system with interacting components that is influenced by an exogenous source of randomness. There is a central source of interconnection (a central “bus”). Failure of any component stresses the central bus, which in turn can cause other components to fail (a feedback effect). In particular, we want to understand how the system can catastrophically fail, and how the feedback mechanism and the exogenous factor interact to produce large failure clusters. Along with an understanding of the likelihood of such a failure, we want to understand the structure of pathways to failure. Given a statistical description of the interacting system, an understanding of the “most likely” pathways to failure would naturally lead to an understanding of how to control the system to minimize failure and how to sense the onset of failure.

Of particular interest is *dynamic* interactions. The extra dimension of time allows one to compare different possible ways in which a system can end up at a given point. This may help in early-stage identification of certain phenomena. Our goal here is to develop some of these insights in the model of [17]. In particular, we would like to understand *pathways to systemic collapse*. In the presence of contagion, what are the most likely ways that a pool can suffer large losses? As in other engineered systems, characterization of most likely paths to failure can better allow effective intervention.

The heart of our analysis is the theory of *large deviations*, e.g., see the classical manuscripts [6, 11]. The theory of large deviations gives a rich framework in which to identify and then prove rates of decay of exponential tails. Two simple examples are well-known; Sanov's theorem, which in its simplest captures the tail behavior of a large collection of i.i.d coin flips, and the Freidlin-Wentzell theorem, which identifies the most likely way that diffusive perturbations can drive a stable differential equation out of equilibrium. Both of these core examples play a role in our analysis, even though the situation is more complex here because the coin flips are neither identically distributed, nor independent. Defaults can be modeled as coin flips, and the effects of contagion can be thought of as random perturbations of a dynamical system. Our main theorems, Theorems 3.1 and 3.10, give the large deviations principle. In the full case, where both contagion and systematic effects are present, the rate function is a combination of a relative entropy (a “nonlinear” Sanov's type theorem) and the Freidlin-Wentzell action functional.

Tail behavior in static pools is investigated in [5, 14] and in [21] the authors study the effect of stochastic recovery on the tail of the loss distribution when the recovery rate depends on the default rate. In [15, 16] rare event asymptotics for the loss distribution in the Gaussian copula model of portfolio credit risk and related importance sampling questions are studied. In a large deviations analysis of a mean field model in [3] the authors take the default intensity of a component in the pool to be a deterministic function of the percentage pool loss due to defaults, see also [4]. In [23], the authors establish a large time large deviations principle for an interacting system of affine point processes. In [12], the authors study systematic risk [from endogenous aspect] via a mean field model of interacting agents. Using a model of a two-well potential, agents can move freely from a healthy state to a failed state. The authors study probabilities of transition from the healthy to the failed state using large deviations ideas.

In this paper, we consider tail behavior for dynamic heterogeneous pools with stochastic intensity and we are interested in the behavior for large pools of interconnected components. Our contribution is two-fold. Firstly, we develop a large deviations principle for a dynamic point process model of correlated default timing that takes into account both effects of contagion and of systematic, exogenous, risks. We study the case that the number of constituent firms or components in the network grows. Secondly, we numerically explore the large deviations results which can help understand how default clusters occur in such systems.

In the theoretical side, we see that the rate function governing the tail events is given explicitly as an additive functional of a Freidlin-Wentzell action functional and a nonlinear relative entropy, where the nonlinearity originates from the fact that defaults are not independent. The dependence structure and the heterogeneity of the environment, i.e., the fact that defaults are not identically distributed, complicate the mathematical analysis. Nevertheless, we rigorously obtain a fairly explicit form of the large deviations principle which is amenable to numerical investigations. In particular, we derive a representation for the rate function which is suitable for numerically computing both the rate function and the most likely path to failure, i.e., the extremals, in both homogeneous and heterogeneous environments.

In the numerical side, the numerical experiments illustrate the effect of systematic (exogenous) risks and contagion on tail events of such systems. For instance, as we shall see in the numerical investigations of Section 4, if a large default cluster occurs, the systematic risk is most likely to play a large role in the initial phase, but then its importance decreases (and thus the contagion effects become more important). Moreover, the analysis of the extremals related to the large deviations rate function (i.e., the most likely paths to failure) in heterogeneous pools with two or more types can help understand which types in the pool are more vulnerable to default due to the impact of contagion, and as a consequence the pathway to creation of default clusters.

We end this introduction by mentioning that, even though the interacting system that we study is primarily motivated by issues of systemic risk in large financial networks, its formulation is sufficiently generic to make the analysis and results of broader interest. The physical phenomenon of the failure of an individual component in a network, which could have been caused by an external force, increasing the stress to other components of the network, making the system more likely to fail, is of broader interest and applicability.

The paper is organized as follows. In Section 2, we describe the model, the assumption and recall the law of large numbers result proven in [17]. In Section 3 we describe our main results, the large deviations principle for the empirical measure and empirical default rate. Then, in Section 4 we perform numerical experiments supporting the theoretical findings. Section 5 contains the proofs of the large deviations results of Section 3. We conclude with our conclusions in Section 6.

2. MODEL AND LAW OF LARGE NUMBERS

We assume that our overall system contains N components or subsystems (where N is large). We assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is an underlying probability triple on which all random variables are defined.

To start, let τ_n^N be the stopping time at which the n -th component (or particle) in our system fails. A failure time τ_n^N has intensity process $\lambda^{N,n}$, which satisfies

$$\mathbb{P}\{\tau_n^N \in (t, t + \delta] | \mathcal{F}_t, \tau_n^N > t\} \approx \lambda_t^{N,n} \delta$$

as $\delta \searrow 0$, where \mathcal{F}_t is the sigma-algebra generated by the entire system up to time t . That is, the process defined by $\chi_{\{\tau_n^N \leq t\}} - \int_0^t \lambda_s^{N,n} \mathbf{1}_{\{\tau_n^N > s\}} ds$ is a martingale with respect to \mathcal{F}_t .

We want our model to capture several important phenomena. Each component will be affected by three sources of randomness, one of which is unique to the component itself, one that is responsible for contagious effects, and one of which reflects the external environment. We assume that each component has been engineered to be stable. We also assume, however, the system is subject to cascading (or ‘‘contagious’’) behavior; failure of any component is likely to lead to failure of more components. In particular, fixing $N \in \mathbb{N}$ and $n \in \{1, \dots, N\}$ we consider the model

$$\begin{aligned} (1) \quad & d\lambda_t^{N,n} = -\alpha_{N,n}(\lambda_t^{N,n} - \bar{\lambda}_{N,n})dt + \sigma_{N,n}\sqrt{\lambda_t^{N,n}}dW_t^n + \beta_{N,n}^C dL_t^N + \varepsilon_N \beta_{N,n}^S \lambda_t^{N,n} dX_t \quad t > 0 \\ & \lambda_0^{N,n} = \lambda_{\circ, N, n} \\ & dX_t = b(X_t)dt + \kappa(X_t)dV_t \quad t > 0 \\ & X_0 = x_{\circ} \\ & L_t^N = \frac{1}{N} \sum_{n=1}^N \chi_{\{\tau_n^N \leq t\}}, \end{aligned}$$

where

$$(2) \quad \tau_n^N \stackrel{\text{def}}{=} \inf \left\{ t \geq 0 : \int_0^t \lambda_s^{N,n} ds \geq \epsilon_n \right\}.$$

Here the W^n 's are an independent collection of standard Brownian motions, V is also a standard Brownian motion, and the ϵ_n 's are standard exponential random variables; the W^n 's, V and the ϵ_n 's are all independent of each other.

In the case $\beta_{N,n}^C = \beta_{N,n}^S = 0$ for all $n \in \{1, \dots, N\}$, one recovers the classical CIR model in credit risk, e.g., [7]. The term $\beta_{N,n}^C dL_t^N$ reflects the contagious effects and the X_t term is the external source of randomness (systematic risk). Notice that movements in X cause correlated changes in each intensity $\lambda^{N,n}$, which then provides a channel for default clustering. Each system's component sensitivity to X is measured by the parameter $\beta_{N,n}^S \in \mathbb{R}$. Then, a default causes a jump of size $\beta_{N,n}^C/N$ in the intensity $\lambda^{N,n}$, where $\beta_{N,n}^C \in \mathbb{R}_+ = [0, \infty)$. The mean-reversion of $\lambda^{N,n}$ implies that the impact of a default fades away with time, exponentially with rate $\alpha_{N,n} \in \mathbb{R}_+$. The linear dependence of λ in the dX component guarantees that $\lambda_t^{N,n} \geq 0$ for every $N \in \mathbb{N}$, $n \in \{1, \dots, N\}$ and $t \geq 0$ (Proposition 3.3 in [17]), and makes the analytic computations easier. As it is demonstrated in [1], an important channel for clustering of defaults are self-exciting effects of such types.

Proposition 3.3 in [17] guarantees that under the assumption of an existence of a unique strong solution for the SDE for X process, the system (1) has a unique strong solution such that $\lambda_t^{N,n} \geq 0$ for every $N \in \mathbb{N}$, $n \in \{1, \dots, N\}$ and $t \geq 0$. The structure of the feedback term, i.e., the empirical average L_t^N , is of mean field type, which brings the system (1) roughly within the class of McKean-Vlasov models, e.g., [13]. However, as it is also demonstrated in [17, 18], the structure of (1) presents several difficulties that bring the analysis of such systems outside the scope of the standard setup. In particular, a key structural difference is that the empirical measure is not on the level of the intensities $\lambda_t^{N,n}$, but on the level of the default times τ_n^N , i.e. there is an extra level of randomness. Moreover, the interacting particle system is heterogeneous and not homogeneous, there is the presence of the additional term X , the coefficients of the intensity dynamics are not bounded and there is a square root degeneracy.

Let's define a 'type' space which describes the $\lambda^{N,n}$'s. Define $\mathcal{P} \stackrel{\text{def}}{=} \mathbb{R}_+^4 \times \mathbb{R}^2$ and we denote by $\mathbf{p}^{N,n} = (\alpha_{N,n}, \bar{\lambda}_{N,n}, \sigma_{N,n}, \beta_{N,n}^C, \beta_{N,n}^S, \lambda_{\circ, N,n}) \in \mathcal{P}$. Then $\mathbf{p}^{N,n}$ gives the dynamics of $\lambda^{N,n}$, the response to the loss process L^N , and the initial condition. Let's also define $\mathcal{P}_t^{N,n} \stackrel{\text{def}}{=} (\alpha_{N,n}, \bar{\lambda}_{N,n}, \sigma_{N,n}, \beta_{N,n}^C, \beta_{N,n}^S, \lambda_t^{N,n})$; this is a \mathcal{P} -valued process which keeps track of the dynamics of $\lambda^{N,n}$ and its current value. Of course, we have that $\mathcal{P}_0^{N,n} = \mathbf{p}_0^{N,n}$.

For any Polish space S , let $\mathcal{M}_1(S)$ be the collection of subprobability Borel measures on S ; i.e., S consists of Borel measures ν on S such that $\nu(S) \leq 1$. Then $\mathcal{M}_1(S)$ is itself a Polish space; [10]. Define here $E \stackrel{\text{def}}{=} \mathcal{M}_1(\mathcal{P})$, and let

$$\boldsymbol{\mu}_t^N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \delta_{\mathcal{P}_t^{N,n}} \chi_{\{\tau_n^N > t\}};$$

this is the empirical distribution of $\mathcal{P}_t^{N,n}$ for those components which are still "alive". We note that $\boldsymbol{\mu}_t^N$ is a random trajectory in E . Also,

$$L_t^N = 1 - \boldsymbol{\mu}_t^N(\mathcal{P}),$$

We shall denote by $\mathcal{P}(\mathcal{P})$ the set of probability measures on \mathcal{P} . In [17, 18, 20], the authors have developed law of large numbers and central limit theorem approximations to $\boldsymbol{\mu}_t^N$ and as a consequence to L_t^N as well. The goal of this paper is to establish a large deviations principle for the empirical loss L_t^N , namely to study its tail behavior. In this paper we mainly concentrate on the case $\lim_{N \rightarrow \infty} \epsilon_N = 0$. To focus our investigation, let's first recall the law of large numbers results obtained in [17]; we can only understand 'rare' events relative to the 'typical' event.

Assumption 2.1. *We assume that there is a $K_{2,1} > 0$ such that the $\alpha_{N,n}$'s, $\lambda_{N,n}$'s, $\sigma_{N,n}$'s, $|\beta_{N,n}^C|$'s, $|\beta_{N,n}^S|$'s, and $\lambda_{\circ, N,n}$'s are all bounded by $K_{2,1}$ for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$.*

Thus the types are bounded. In fact we want them to have a macroscopic distribution.

Assumption 2.2. For $N \in \mathbb{N}$, define

$$U_N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \delta_{\mathbf{p}^{N,n}}.$$

We assume that $U \stackrel{\text{def}}{=} \lim_{N \rightarrow \infty} U_N$ exists in $\mathcal{P}(\mathcal{P})$, in the sense of weak convergence of probability measures on \mathcal{P} .

Under these assumptions, $\mu \stackrel{\text{def}}{=} \lim_{N \rightarrow \infty} \mu^N$ is a well-defined measure-valued process. We can identify μ via the martingale problem. For $\mathbf{p} = (\hat{\mathbf{p}}, \lambda)$ where $\hat{\mathbf{p}} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S) \in \mathcal{P}$ and $f \in C^\infty(\mathcal{P})$, define the operators

$$\begin{aligned} (\mathcal{L}_1 f)(\mathbf{p}) &= \frac{1}{2} \sigma^2 \lambda \frac{\partial^2 f}{\partial \lambda^2}(\mathbf{p}) - \alpha (\lambda - \bar{\lambda}) \frac{\partial f}{\partial \lambda}(\mathbf{p}) - \lambda f(\mathbf{p}) \\ (\mathcal{L}_2 f)(\mathbf{p}) &= \beta^C \frac{\partial f}{\partial \lambda}(\mathbf{p}). \end{aligned}$$

Define also

$$\mathcal{Q}(\mathbf{p}) \stackrel{\text{def}}{=} \lambda$$

for $\mathbf{p} = (\hat{\mathbf{p}}, \lambda)$ where $\hat{\mathbf{p}} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S) \in \mathcal{P}$. The generator \mathcal{L}_1 corresponds to the diffusive part of the intensity with killing rate λ , and \mathcal{L}_2 is the macroscopic effect of contagion on the surviving intensities at any given time. For every $f \in C^\infty(\mathcal{P})$ and $\mu \in E$, define

$$\langle f, \mu \rangle_E \stackrel{\text{def}}{=} \int_{\mathbf{p} \in \mathcal{P}} f(\mathbf{p}) \mu(d\mathbf{p}).$$

Let \mathcal{S} be the collection of elements Φ in $B(\mathcal{P}(\mathcal{P}))$ of the form

$$\Phi(\mu) = \varphi(\langle f_1, \mu \rangle_E, \langle f_2, \mu \rangle_E \dots \langle f_M, \mu \rangle_E)$$

for some $M \in \mathbb{N}$, some $\varphi \in C^\infty(\mathbb{R}^M)$ and some $\{f_m\}_{m=1}^M$. For $\Phi \in \mathcal{S}$ of the form (2), define

$$(\mathcal{A}\Phi)(\mu) \stackrel{\text{def}}{=} \sum_{m=1}^M \frac{\partial \varphi}{\partial x_m}(\langle f_1, \mu \rangle_E, \langle f_2, \mu \rangle_E \dots \langle f_M, \mu \rangle_E) \{ \langle \mathcal{L}_1 f_m, \mu \rangle_E + \langle \mathcal{Q}, \mu \rangle_E \langle \mathcal{L}_2 f_m, \mu \rangle_E \}.$$

We claim that \mathcal{A} will be the generator of the limiting martingale problem (see [17]).

Lemma 2.3. ¹[Weak Convergence] Let $\lim_{N \rightarrow \infty} \varepsilon_N = 0$. The sequence $\{\mu^N\}_N$ is tight in $D_E([0, T])$. Moreover, for any $\Phi \in \mathcal{S}$ and $0 \leq r_1 \leq r_2 \dots r_J = s < t < T$ and $\{\psi_j\}_{j=1}^J \subset B(E)$, we have that

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E} \left[\left\{ \Phi(\mu_t^N) - \Phi(\mu_s^N) - \int_{r=s}^t (\mathcal{A}\Phi)(\mu_r^N) dr \right\} \prod_{j=1}^J \psi_j(\mu_{r_j}^N) \right] &= 0 \\ \lim_{N \rightarrow \infty} \mathbb{E} [\Phi(\mu_0^N)] &= \Phi(U). \end{aligned}$$

The limit $\mu \stackrel{\text{def}}{=} \lim_N \mu^N$ uniquely exists and is deterministic, this limit being in probability in $C([0, T]; E)$. Moreover, letting $L_t = 1 - \mu_t(\mathcal{P})$, we get that for every $\delta > 0$

$$\lim_{N \rightarrow \infty} \mathbb{P} \left\{ \sup_{0 \leq t \leq T} |L_t^N - L_t| \geq \delta \right\} = 0.$$

Corollary 3.3 shows that L_t can be expressed as the unique solution to a fixed point equation. As in the law of large numbers, this gives us a reference trajectory for defining rare events. For any closed subset F of $C([0, T]; [0, 1])$ which does not contain L , we have that

$$(3) \quad \lim_{N \rightarrow \infty} \mathbb{P}\{L^N \in F\} = 0.$$

¹At this point we would like to remark that there is a typo in the formulation of the particular result in [17]. In particular, it is mentioned there that $(\mathcal{L}_2 f)(\mathbf{p}) = \frac{\partial f}{\partial \lambda}(\mathbf{p})$ and $\mathcal{Q}(\mathbf{p}) = \beta^C \lambda$, where it should have been $(\mathcal{L}_2 f)(\mathbf{p}) = \beta^C \frac{\partial f}{\partial \lambda}(\mathbf{p})$ and $\mathcal{Q}(\mathbf{p}) = \lambda$.

The theory of large deviations gives us the *rate* of decay of probabilities like in (3), namely the tail of the distribution of large portfolio losses. It identifies a function $I : C([0, T]; [0, 1]) \rightarrow [0, \infty]$ such that, informally,

$$\mathbb{P}\{L^N \approx L_{\text{ref}}\} \stackrel{N \rightarrow \infty}{\asymp} \exp[-NI(L_{\text{ref}})]$$

While rare events are (by definition) unlikely to happen, large deviations theory gives a rigorous framework to compare the rarity of different rare events. It thus helps us understand the ‘most likely’ rare event in a given set of rare events. As the numerical experiments in Section 4 indicate, larger sensitivity to contagion and systematic risk leads to fatter tails, which leads to larger likelihood of large losses in the system. Insights of this type can help understand the role of contagion and systematic risk and how they interact to produce atypically large failure rates. This can then provide guidance in guarding against dangerous rare system behavior.

3. PROBLEM FORMULATION AND MAIN RESULTS

We here present our main results on large deviations for the empirical loss L_t^N . We start with some notation and preliminary computations in Subsection 3.1. The first result is Theorem 3.1 in Subsection 3.2, where we derive the large deviations principle in the case $\varepsilon_N = 0$. Lemma 2.3 with $\varepsilon_N = 0$ follows as a result; see Corollary 3.3. Corollary 3.2 gives an alternate formulation of the rate function of Theorem 3.1; this representation is useful for numerical studies. Then in Subsection 3.3, we present, in Theorem 3.10, the second main result of this paper, which is the large deviations principle when $\varepsilon_N \neq 0$, i.e., when both contagion and systematic effects are present.

The large deviations principle of Theorem 3.1 is obtained by first identifying a large deviations principle for the empirical measure of defaults in the heterogeneous pool, ν^N , (properly defined in (8)), and then using the contraction principle. The large deviations principle for ν^N is proved in two steps. First, we derive the large deviations principle in the independent (i.e., when all $\beta_{N,n}^C = 0$), but heterogeneous case. Varadhan’s transfer lemma (recalled as Theorem 5.7) then implies the LDP for the general case. The large deviations principle of Theorem 3.10 is obtained via a conditioning argument from the large deviations principle of Theorem 3.1 and that of small noise diffusion processes, [11]. In this section we present statement of theorems and the corresponding proofs are in Section 5.

Let’s recall the concept of large deviations and the associated rate function.

Definition 3.1. *If S is a Polish space and \mathbb{P} is a probability measure on $(S, \mathcal{B}(S))$, we say that a collection $(\xi_n)_{n \in \mathbb{N}}$ of S -valued random variables has a large deviations principle with rate function $I : S \rightarrow [0, \infty]$ and speed n if*

- (i) *For each $s \geq 0$, the set $\Phi(s) = \{s \in S : I(s) \leq s\}$ is a compact subset of S .*
- (ii) *For every open $G \subset S$,*

$$\underline{\lim}_{n \nearrow \infty} \frac{1}{n} \ln \mathbb{P}\{\xi_n \in G\} \geq - \inf_{s \in G} I(s)$$

- (iii) *For every closed $F \subset S$,*

$$\overline{\lim}_{n \nearrow \infty} \frac{1}{n} \ln \mathbb{P}\{\xi_n \in F\} \leq - \inf_{s \in F} I(s).$$

3.1. Preliminary computations. Let’s set up some notation. Fix $\mathbf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S, \lambda_o) \in \mathcal{P}$. Fix also a time horizon $T > 0$. Let a Polish space S and denote by $C(S; \mathbb{R})$ and $AC(S; \mathbb{R})$ to be the collection of continuous and respectively absolutely continuous paths from S to \mathbb{R} . For notational convenience we will sometimes write $C(S; \mathbb{R})$ and $AC(S; \mathbb{R})$ if no confusion arises. For φ and ψ in $AC([0, T]; \mathbb{R})$ and W^* a reference Brownian motion, let $\lambda^{\varphi, \psi}$ be the solution of the SDE

$$(4) \quad \begin{aligned} \lambda_t^{\varphi, \psi}(\mathbf{p}) = & \lambda_o - \alpha \int_0^t (\lambda_s^{\varphi, \psi}(\mathbf{p}) - \bar{\lambda}) ds + \sigma \int_0^t \sqrt{\lambda_s^{\varphi, \psi}(\mathbf{p})} dW_s^* \\ & + \beta^C \varphi(t) \chi_{[0, T]}(t) + \beta^S \int_0^t \lambda_s^{\varphi, \psi}(\mathbf{p}) \chi_{[0, T]}(s) d\psi(s) \quad t > 0 \end{aligned}$$

This will represent the conditional intensity of a “randomly-selected” name of type \mathbf{p} in our pool. Let the random variable $\tau^{\varphi, \psi}$ be such that

$$(5) \quad \mathbb{P} \{ \tau^{\varphi, \psi} \leq t \} = \mathbb{P} \left\{ \int_0^t \lambda_s^{\varphi, \psi}(\mathbf{p}) ds \geq \epsilon \right\} = 1 - \mathbb{E} \left[\exp \left[- \int_0^t \lambda_s^{\varphi, \psi}(\mathbf{p}) ds \right] \right]$$

for all $t > 0$, where ϵ is an exponential (1) random variable which is independent of W^* . Define

$$f_{\varphi, \psi}^{\mathbf{p}}(t) \stackrel{\text{def}}{=} \mathbb{E} \left[\lambda_t^{\varphi, \psi}(\mathbf{p}) \exp \left[- \int_0^t \lambda_s^{\varphi, \psi}(\mathbf{p}) ds \right] \right].$$

Then for $t \in [0, T]$

$$(6) \quad \mathbb{P} \{ \tau^{\varphi, \psi} \leq t \} = \int_0^t f_{\varphi, \psi}^{\mathbf{p}}(s) ds = 1 - \mathbb{E} \left[\exp \left[- \int_0^t \lambda_s^{\varphi, \psi}(\mathbf{p}) ds \right] \right].$$

Let’s now collect together all scenarios where $\tau^{\varphi, \psi} > T$. Fix an abstract point \star not in $[0, T]$ and define

$$\mathcal{T} \stackrel{\text{def}}{=} [0, T] \cup \{\star\}$$

(which is a Polish space). For given trajectories φ and ψ in $AC([0, T]; \mathbb{R})$, define $\mu_{\varphi, \psi}^{\mathbf{p}} \in \mathcal{P}(\mathcal{T})$ as

$$(7) \quad \mu_{\varphi, \psi}^{\mathbf{p}}(A) \stackrel{\text{def}}{=} \int_{t \in A \cap [0, T]} f_{\varphi, \psi}^{\mathbf{p}}(t) dt + \delta_{\star}(A) \left\{ 1 - \int_0^T f_{\varphi, \psi}^{\mathbf{p}}(t) dt \right\}$$

for all $A \in \mathcal{B}(\mathcal{T})$. In other words, $\mu_{\varphi, \psi}^{\mathbf{p}}$ is the corresponding probability measure on \mathcal{T} .

The large deviations principles appearing in Theorems 3.1 and 3.10 are in terms $\mu_{\varphi, \psi}^{\mathbf{p}}$. Moreover, taking advantage of the special structure of the problem, Lemma 4.1 shows that $\mu_{\varphi, \psi}^{\mathbf{p}}$ can be computed in closed form.

3.2. Large deviations principle for the case $\varepsilon_N = 0$. In this case the exogenous source of randomness, X plays no role in the computations. Define the probability measures $\nu^N \in \mathcal{P}(\mathcal{P} \times \mathcal{T})$

$$(8) \quad \nu^N = \frac{1}{N} \sum_{n=1}^N \delta_{\mathbf{p}^{N,n}, \tau_n} \chi_{\{\tau_n \leq T\}} + \frac{1}{N} \sum_{n=1}^N \delta_{\mathbf{p}^{N,n}, \star} \chi_{\{\tau_n > T\}}$$

This captures the distribution of $\mathbf{p}^{N,n}$ (the ‘type’ of the asset) and $\tau^{N,n}$ (the default time).

For notational convenience set

$$\mathsf{X} \stackrel{\text{def}}{=} \mathcal{P} \times \mathcal{T}$$

then X is Polish. We can of course recover the empirical loss from ν^N ; $L_t^N = \nu^N(\mathcal{P} \times [0, t])$.

Let’s make the following definition.

Definition 3.2. Fix $\omega \in \mathcal{P}(\mathsf{X})$ and $\nu \in \mathcal{P}(\mathcal{P})$; we say that $\omega = \xi \otimes \nu$, where ξ is a measurable map from \mathcal{P} to $\mathcal{P}(\mathcal{S})$ (i.e., ξ is a stochastic kernel) if

$$\omega(A \times B) = \int_{\mathbf{p} \in A} \xi(\mathbf{p})(B) \nu(d\mathbf{p}).$$

for all $A \in \mathcal{B}(\mathcal{P})$ and $B \in \mathcal{B}(\mathcal{T})$. In this case, we write

$$\xi = \frac{\partial \omega}{\partial \nu}.$$

Let us define

$$(9) \quad H(\nu, \mu) = \begin{cases} \int_{t \in \mathcal{T}} \ln \frac{d\nu}{d\mu}(t) \nu(dt), & \text{if } \nu \ll \mu \\ \infty, & \text{otherwise} \end{cases}$$

and then

$$(10) \quad \bar{H}(\nu) = \begin{cases} \int_{\mathbf{p} \in \mathcal{P}} H\left(\frac{\partial \nu}{\partial U}(\mathbf{p}), \mu_{\bar{\nu}, 0}^{\mathbf{p}}\right) U(d\mathbf{p}) & \text{if } \frac{\partial \nu}{\partial U}(\mathbf{p}) \text{ exists} \\ \infty & \text{otherwise} \end{cases}$$

and where $\mu_{\bar{\nu},0}^{\mathbf{p}}$ is defined in (7), U is given by Assumption 2.2, the superscript \mathbf{p} denotes the dependence on the particular element $\mathbf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S, \lambda) \in \mathcal{P}$, and for $B \in \mathcal{B}(\mathcal{T})$, we set

$$\bar{\nu}(B) = \nu(\mathcal{P} \times B).$$

The form of the action functional is fairly easy to understand, at least in the homogeneous case. Suppose that there is a fixed $\mathbf{p}^* \in \mathcal{P}$ and $\mathbf{p}^{N,n} = \mathbf{p}^*$ for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$ (and thus $U = \delta_{\mathbf{p}}$). If $\nu^* = \delta_{\mathbf{p}^*} \times \omega$ for some $\omega \in \mathcal{P}(\mathcal{T})$, then $\frac{\partial \nu^*}{\partial U}(\mathbf{p}^*) = \omega$ and one expects that

$$\mathbb{P}_N(d\nu^N \in d\nu^*) \stackrel{N \rightarrow \infty}{\underset{\sim}{\asymp}} e^{-NH(\omega, \mu_{\bar{\nu},0}^{\mathbf{p}^*})} = e^{-N\bar{H}(\nu^*)}.$$

Consider next a heterogeneous pool with two fixed types \mathbf{p}_A^* and \mathbf{p}_B^* . Assume that, in a pool of size N , every third name is of type \mathbf{p}_A^* the remaining names are of type \mathbf{p}_B^* . If

$$\nu^* = \frac{1}{3}\delta_{\mathbf{p}_A^*} \times \omega_A + \frac{2}{3}\delta_{\mathbf{p}_B^*} \times \omega_B,$$

then

$$\frac{\partial \nu^*}{\partial U}(\mathbf{p}) = \begin{cases} \omega_A & \text{if } \mathbf{p} = \mathbf{p}_A^* \\ \omega_B & \text{if } \mathbf{p} = \mathbf{p}_B^* \end{cases}$$

Then

$$\begin{aligned} \frac{\partial \nu^N}{\partial U_N}(\mathbf{p}_A^*) &= \frac{1}{N/3} \sum_{\substack{1 \leq n \leq N \\ n \in 3\mathbb{N}}} \delta_{\tau_n} \chi_{\{\tau_n \leq T\}} + \frac{1}{N/3} \sum_{\substack{1 \leq n \leq N \\ n \in 3\mathbb{N}}} \delta_* \chi_{\{\tau_n > T\}} \\ \frac{\partial \nu^N}{\partial U_N}(\mathbf{p}_B^*) &= \frac{1}{2N/3} \sum_{\substack{1 \leq n \leq N \\ n \notin 3\mathbb{N}}} \delta_{\tau_n} \chi_{\{\tau_n \leq T\}} + \frac{1}{2N/3} \sum_{\substack{1 \leq n \leq N \\ n \notin 3\mathbb{N}}} \delta_* \chi_{\{\tau_n > T\}} \end{aligned}$$

At a heuristic level, one expects that

$$\begin{aligned} \mathbb{P}_N(d\nu^N \in d\nu^*) &= \mathbb{P}_N \left\{ \frac{\partial \nu^N}{\partial U_N}(\mathbf{p}_A^*) \approx d\omega_A, \frac{\partial \nu^N}{\partial U_N}(\mathbf{p}_B^*) \approx d\omega_B \right\} \\ &\stackrel{N \rightarrow \infty}{\underset{\sim}{\asymp}} \exp \left[-\frac{N}{3} H(\omega_A, \mu_{\bar{\nu},0}^{\mathbf{p}_A^*}) - \frac{2N}{3} H(\omega_B, \mu_{\bar{\nu},0}^{\mathbf{p}_B^*}) \right] = e^{-N\bar{H}(\nu^*)}. \end{aligned}$$

Theorem 5.8 gives the rigorous proof that, in general, $\bar{H}(\nu)$ is the rate function for $\{\nu^N, N < \infty\}$ in the heterogeneous case. An immediate consequence of Theorem 5.8 and contraction principle is the large deviations principle for $L_t^N = \nu^N(\mathcal{P} \times [0, t])$.

Theorem 3.1. *Consider the system defined in (1) with $\varepsilon_N = 0$ and let $T < \infty$. Under Assumptions 2.1 and 2.2, the family $\{L_T^N, N \in \mathbb{N}\}$ satisfies the large deviation bounds of Definition 3.1, with rate function*

$$I(\ell) = \inf \{ \bar{H}(\nu) : \nu \in \mathcal{P}(\mathcal{X}) \text{ and } \nu(\mathcal{P} \times [0, T]) = \ell \}.$$

and speed N . The rate function I is lower semicontinuous and has compact level sets.

As expected the rate function is essentially defined via an entropy functional, which in the present setting takes the form $H\left(\frac{\partial \nu}{\partial U}(\mathbf{p}), \mu_{\bar{\nu},0}^{\mathbf{p}}\right)$. However, due to the heterogeneity of the environment and due to the feedback term dL_t^N in the system, new phenomena appear. The effect of heterogeneity is to essentially integrate over all the different types in the \mathcal{P} space, whereas the feedback term is responsible for the subscript $\bar{\nu}$ in the $\mu_{\bar{\nu},0}^{\mathbf{p}}$, which is a non-linear effect in the entropy.

For $\xi, \varphi \in AC([0, T]; \mathbb{R})$ let us define the functional

$$g(\xi, f_{\varphi,0}^{\mathbf{p}}) = \int_0^T \ln \left(\frac{\dot{\xi}(t)}{f_{\varphi,0}^{\mathbf{p}}(t)} \right) \dot{\xi}(t) dt + \ln \left(\frac{1 - \xi(T)}{1 - \int_0^T f_{\varphi,0}^{\mathbf{p}}(t) dt} \right) (1 - \xi(T))$$

Next, in Corollary 3.2, we note that the rate function I has a straightforward alternate representation which is a bit more suited to numerical investigations.

Corollary 3.2. Consider the system defined in (1) with $\varepsilon_N = 0$. Define

$$I'(\ell) = \inf_{\varphi \in AC(\mathcal{P} \times [0, T]; \mathbb{R}), \varphi \geq 0} \left\{ \int_{\mathcal{P}} g(\varphi(\mathbf{p}), f_{\bar{\varphi}, 0}^{\mathbf{p}}(t)) U(d\mathbf{p}) : \varphi(\mathbf{p}, 0) = 0, \bar{\varphi}(s) = \int_{\mathcal{P}} \varphi(\mathbf{p}, s) U(d\mathbf{p}), \text{ and } \bar{\varphi}(T) = \ell \right\}$$

Under the assumptions and notation of Theorem 3.1, $I = I'$.

An immediate consequence of Theorem 3.1 is the law of large numbers result for L_t^N obtained in [17].

Corollary 3.3. Assume that $\varepsilon_N = 0$. There is a unique measure ν^* such that for every $t \in [0, T]$

$$\bar{\nu}^*([0, t]) = 1 - \int_{\mathcal{P}} e^{-(b^{\mathbf{p}}(t)\lambda_0 + \int_0^t \alpha \bar{\lambda} b^{\mathbf{p}}(t-u) du + \beta^C \int_0^t b^{\mathbf{p}}(t-u) \bar{\nu}^*(du))} U(d\mathbf{p});$$

this is the unique solution of $\bar{H}(\nu) = 0$. Finally, $\bar{\nu}^*([0, t]) = L_t$ for all $t \in [0, T]$, where L was given in Lemma 2.3.

The proof of Corollary 3.3 is in Section 5.2.

As an example of possible use of Theorem 3.1 and Corollary 3.2, let's consider homogenous pools and heterogeneous pools composed of K different bins that are homogeneous within each bin.

Example 3.4 (Homogeneous). Fix $\mathbf{p}^* \in \mathcal{P}$ and assume that $\mathbf{p}^{N,n} = \mathbf{p}^*$ for all N and n . Then

$$I(\ell) = \inf \{ g(\varphi, f_{\varphi, 0}^{\mathbf{p}}) : \varphi(0) = 0, \varphi(T) = \ell, \varphi \geq 0, \varphi \in AC([0, T]; \mathbb{R}) \}.$$

Example 3.5 (Heterogeneous). Let us assume that the pool is composed of K different bins. Assume that $\kappa_i\%$ of the names are of type A_i with $i = 1, \dots, K$ and $\sum_{i=1}^K \kappa_i = 100$. Setting $\varphi(\mathbf{p}, s) = \sum_{i=1}^K \frac{\kappa_i}{100} \varphi_{A_i}(s) \chi_{\{\mathbf{p}_{A_i}\}}$, we get that

$$I(\ell) = \inf \left\{ \sum_{i=1}^K \frac{\kappa_i}{100} g(\varphi_{A_i}, f_{\varphi, 0}^{\mathbf{p}_{A_i}}) : \varphi(t) = \sum_{i=1}^K \frac{\kappa_i}{100} \varphi_{A_i}(t) \text{ for every } t \in [0, T] \right. \\ \left. \varphi(T) = \ell, \varphi_{A_i}(0) = 0, \varphi_{A_i} \geq 0, \varphi_{A_i} \in AC([0, T]; \mathbb{R}) \text{ for every } i = 1, \dots, K \right\}.$$

3.3. Main result: Large deviations principle for the case $\lim_{N \rightarrow \infty} \varepsilon_N = 0$. In this subsection, we study the case where systematic effects are present. When $\varepsilon_N \neq 0$, the large deviations of process $\{X_t^N = \varepsilon_N X_t, N \in \mathbb{N}\}$ affect the large deviations of the empirical default rate process $\{L_t^N, N \in \mathbb{N}\}$. In order to properly formulate our result we need to make some assumptions on the scaling properties of the coefficients of the X process. These are minimal assumptions that guarantee the existence of a large deviations principle for the family $\{X_t^N = \varepsilon_N X_t, N \in \mathbb{N}\}$. In particular:

Assumption 3.6. We assume that the following limits exist uniformly on bounded subsets of \mathbb{R}

- $\bar{b}(x) = \lim_{\varepsilon \downarrow 0} b_\varepsilon(x) = \lim_{\varepsilon \downarrow 0} \varepsilon b(x/\varepsilon)$.
- For some $\zeta \in (0, 1]$, $\bar{\kappa}(x) = \lim_{\varepsilon \downarrow 0} \kappa_\varepsilon(x) = \lim_{\varepsilon \downarrow 0} \varepsilon^{1-\zeta} \kappa(x/\varepsilon)$.

Moreover, the coefficients $\bar{b}(x), \bar{\kappa}(x)$ are uniformly continuous on compact subsets of \mathbb{R} and we assume that the SDE with drift coefficient $\bar{b}(x)$ and diffusion coefficient $\bar{\kappa}(x)$ has a unique strong solution.

For any $u \in L^2([0, T]; \mathbb{R})$ define the map $\Gamma : L^2([0, T]; \mathbb{R}) \mapsto C([0, T]; \mathbb{R})$ by the equation

$$(11) \quad \psi(t) = \int_0^t \bar{b}(\psi(s)) ds + \int_0^t \bar{\kappa}(\psi(s)) u(s) ds$$

Assumption 3.7. We assume that for any $u \in L^2([0, T]; \mathbb{R})$ the map $\Gamma : L^2([0, T]; \mathbb{R}) \mapsto C([0, T]; \mathbb{R})$ defined by (11) is well defined and (11) has a unique solution. Moreover, we assume that for every $N \in \mathbb{N}$, the map Γ is continuous when it is restricted to the set $\{u \in L^2([0, T]; \mathbb{R}) : \int_0^T |u(s)|^2 ds \leq N\}$ endowed with the weak topology of $L^2[0, T]$.

We need one more assumption.

Assumption 3.8. Let $u \in \mathcal{A}$ the set of square integrable on $[0, T]$, \mathbb{R} -valued and \mathfrak{F}_t predictable processes and consider the controlled sde

$$dX_t^{\varepsilon, u} = [b_\varepsilon(X_t^{\varepsilon, u}) + \kappa_\varepsilon(X_t^{\varepsilon, u})u(t)] dt + \varepsilon^\zeta \kappa_\varepsilon(X_t^{\varepsilon, u}) dV_t, \quad X_0^{\varepsilon, u} = 0.$$

If $\lim_{n \rightarrow \infty} \varepsilon_n = 0$ and $\{u_n\}_{n \in \mathbb{N}} \subset \mathcal{A}$ such that $\sup_{n \in \mathbb{N}} \int_0^T |u_n(s)|^2 ds \leq N$ almost surely, then X^{ε_n, u_n} is tight in $C([0, T]; \mathbb{R})$ and

$$\sup_{n \in \mathbb{N}} \mathbb{E} \int_0^T |\bar{\kappa}(X_s^{\varepsilon_n, u_n})|^2 ds < \infty.$$

As we shall see in Lemma 5.10 of Subsection 5.3, under Assumptions 3.6, 3.7 and 3.8, the large deviations principle for the family $\{X_t^N = \varepsilon_N X_t, N \in \mathbb{N}\}$ on $C([0, T]; \mathbb{R})$ with speed $1/\varepsilon_N^{2\zeta}$ is the same as the large deviations principle and with the same speed for the family $\{\bar{X}_t^N, N \in \mathbb{N}\}$, where

$$(12) \quad d\bar{X}_t^N = \bar{b}(\bar{X}_t^N) dt + \varepsilon_N^\zeta \bar{\kappa}(\bar{X}_t^N) dV_t, \quad \bar{X}_0^N = 0.$$

In this case, the large deviations action functional for $\{\bar{X}_t^N, N \in \mathbb{N}\}$ in $C([0, T]; \mathbb{R})$ is

$$(13) \quad J_X(\psi) = \inf \left\{ \frac{1}{2} \int_0^T |u(s)|^2 ds : u \in L^2([0, T]; \mathbb{R}), \Gamma(u) = \psi \right\}$$

whenever $\{u \in L^2([0, T]; \mathbb{R}), \Gamma(u) = \psi\} \neq \emptyset$ and $J_X(\psi) = \infty$ otherwise.

We remark here that if $\bar{\kappa}(x) \neq 0$ for all $x \in \mathbb{R}$, then for $\psi \in AC([0, T]; \mathbb{R})$ with $\psi(0) = 0$, we have the simplified well known form, see Section 5.3 of [11],

$$J_X(\psi) = \frac{1}{2} \int_0^T \left| \frac{\dot{\psi}(t) - \bar{b}(\psi(s))}{\bar{\kappa}(\psi(s))} \right|^2 ds$$

and $J_X(\psi) = \infty$ otherwise. However, in general, the form of the rate function is given by (13).

Example 3.9. Two classical examples, where Assumptions 3.6, 3.7 and 3.8 hold and thus the LDP for the process $\{X^N, N \in \mathbb{N}\}$ holds, are (a): the Ornstein-Uhlenbeck process with $b(x) = -\gamma x$ and $\kappa(x) = 1$, where then $\bar{b}(x) = -\gamma x$, $\bar{\kappa}(x) = 1$ and $\zeta = 1$, and (b): the CIR (or square-root) process with $b(x) = -\gamma(x - \bar{x}/\varepsilon)$ and $\kappa(x) = \sqrt{x}$, where then $\bar{b}(x) = -\gamma(x - \bar{x})$, $\bar{\kappa}(x) = \sqrt{x}$ and $\zeta = 1/2$.

Then, we have the following theorem.

Theorem 3.10. Consider the system defined in (1) with $\lim_{N \rightarrow \infty} \varepsilon_N = 0$ such that $\lim_{N \rightarrow \infty} N\varepsilon_N^{2\zeta} = c \in (0, \infty)$ and let $T < \infty$. Under Assumptions 2.1, 2.2, 3.6, 3.7 and 3.8 the family $\{L_T^N, N \in \mathbb{N}\}$ satisfies the large deviation bounds of Definition 3.1, with speed N and with rate function

$$I(\ell) = \inf \{S(\varphi, \psi) : \varphi \in C(\mathcal{P} \times [0, T]; \mathbb{R}), \psi \in C([0, T]; \mathbb{R}), \bar{\varphi}(T) = \ell\}$$

where

$$S(\varphi, \psi) = \begin{cases} \int_{\mathcal{P}} g(\varphi(\mathbf{p}), f_{\bar{\varphi}, \psi}^{\mathbf{p}}(t)) U(d\mathbf{p}) + \frac{1}{c} J_X(\psi), & \text{if } \varphi \in AC(\mathcal{P} \times [0, T]; \mathbb{R}), \psi \in AC([0, T]; \mathbb{R}), \psi(0) = 0, \\ & \varphi(\mathbf{p}, 0) = 0, \varphi \geq 0, \bar{\varphi}(s) = \int_{\mathcal{P}} \varphi(\mathbf{p}, s) U(d\mathbf{p}) \\ \infty, & \text{otherwise} \end{cases}$$

Here, $J_X(\psi)$ is the rate function for the process $\{X^N, N < \infty\}$, as defined by (13). $I(\ell)$ has compact level sets.

The definition of $S(\varphi, \psi)$ suggests that if $c = \lim_{N \rightarrow \infty} N\varepsilon_N^{2\zeta} = \infty$ then the $\int_{\mathcal{P}} g(\varphi(\mathbf{p}), f_{\bar{\varphi}, \psi}^{\mathbf{p}}(t)) U(d\mathbf{p})$ integral in $S(\varphi, \psi)$ will be the dominant factor, whereas if $c = \lim_{N \rightarrow \infty} N\varepsilon_N^{2\zeta} = 0$, then the $J_X(\psi)$ entropy term will be the dominant factor. This will become clearer in Section 5.3. Hence both effects are present if $c = \lim_{N \rightarrow \infty} N\varepsilon_N^{2\zeta} \in (0, \infty)$ and this is the case that we focus on in this paper. However, in the course of the proof of Theorem 3.10 we prove Lemma 5.9, which is the large deviations principle for $\{L_t^N, N \in \mathbb{N}, t \in [0, T]\}$ conditional a given path of the process $t \mapsto X_t$ in $C_c([0, T]; \mathbb{R})$, when $\varepsilon_N = 1$.

4. NUMERICAL EXPLORATION OF THE LARGE DEVIATIONS PRINCIPLE

In this section we illustrate our theoretical results by some numerical computations. In particular, in Subsection 4.1 we investigate how the rate function, the tails of the probability loss distribution and the extremals behave in specific situations for a homogeneous pool. We compare two different portfolios and qualitatively compare the two cases. Then, in Subsection 4.2 we perform some numerical experiments for a heterogeneous pool composed of two types, where the difference in the two types is in the level of influence of contagion and systematic effects.

Understanding the most likely ways in which contagion and systematic risk combine to lead to large default rates, gives useful insights into how to optimally hedge against such events. In particular, the numerical experiments show that if a large cluster were to occur, the effect of the systematic factor would be more significant in the initial phase, but then its importance decreases and contagion effects become more important. Moreover, in the case of heterogeneous pools, the large deviations analysis allows us to make statements about which types of names in the pool are likely to be affected more and in which order.

In the numerical results that follow we have computed numerically the rate function $I(\ell)$ and the corresponding extremals φ and ψ solving the variational problems of Theorem 3.10 for the homogeneous and heterogeneous cases of interest. In the numerical computation of the rate function, one needs to compute $f_{\varphi,\psi}^{\text{P}}(t)$ and this is done with the help of Lemma 4.1.

4.1. Numerics for homogeneous pool. Let us now understand how our calculations look like in some specific cases in the case of homogeneous portfolios. We consider the systematic risk X_t to be of Ornstein-Uhlenbeck type and in particular

$$(14) \quad \begin{aligned} dX_t &= -\gamma X_t dt + dV_t \\ X_0 &= 0 \end{aligned}$$

For the numerical experiments below we have taken $\varepsilon_N = \frac{1}{\sqrt{N}}$. The LDP is given by Theorem 3.10 with $c = 1$. The main difficulty in evaluating $I(\ell)$ and the extremals is the computation of $\mu_{\varphi,\psi}^{\text{P}}$ of (7). The virtue of the CIR-based evolution of $\lambda^{N,n}$ of (1) is that fairly explicit formulae are available. For $t \in [0, T]$, let θ_t^{P} solve

$$(15) \quad \theta_t^{\text{P}}(s) = \int_0^s \left(1 - \frac{1}{2} \sigma^2 (\theta_t^{\text{P}}(r))^2 - \alpha \theta_t^{\text{P}}(r) \right) dr + \beta^S \int_0^s \theta_t^{\text{P}}(t-r) d\psi(r). \quad s \in [0, t]$$

Then define

$$\Gamma_{\varphi,\psi}^{\text{P}}(t) = \theta_t^{\text{P}}(t) \lambda_0 + \alpha \bar{\lambda} \int_0^t \theta_t^{\text{P}}(r) dr + \beta^C \int_0^t \theta_t^{\text{P}}(t-r) d\varphi(r)$$

Lemma 4.1. *We have that*

$$f_{\varphi,\psi}^{\text{P}}(t) = \dot{\Gamma}_{\varphi,\psi}^{\text{P}}(t) \exp \left[-\Gamma_{\varphi,\psi}^{\text{P}}(t) \right].$$

Proof. Define

$$M_s \stackrel{\text{def}}{=} \exp \left[-\theta_t^{\text{P}}(t-s) \lambda_s^{\varphi,\psi} - \alpha \bar{\lambda} \int_s^t \theta_t^{\text{P}}(t-r) dr - \beta^C \int_s^t \theta_t^{\text{P}}(t-r) d\varphi(r) - \int_0^s \lambda_r^{\varphi,\psi} dr \right]$$

for $s \in [0, t]$. Note that $\theta_t^{\text{P}}(0) = 0$. In differential form,

$$\begin{aligned} dM_s &= \left\{ \dot{\theta}_t^{\text{P}}(t-s) \lambda_s^{\varphi,\psi} - \theta_t^{\text{P}}(t-s) \left\{ -\alpha (\lambda_s^{\varphi,\psi} - \bar{\lambda}) + \beta^C \dot{\varphi}(s) + \beta^S \dot{\psi}(s) \lambda_s^{\varphi,\psi} \right\} - \lambda_s^{\varphi,\psi} \right. \\ &\quad \left. + \frac{1}{2} (\theta_t^{\text{P}}(t-s))^2 \sigma^2 \lambda_s^{\varphi,\psi} + \theta_t^{\text{P}}(t-s) \left\{ \alpha \bar{\lambda} + \beta^C \dot{\varphi}(s) \right\} \right\} M_s ds + d\mathcal{M}_s \end{aligned}$$

where \mathcal{M} is a martingale. The ODE (15) implies that the ds term is identically zero, so M is a martingale. Noting that

$$M_0 = \exp \left[-\Gamma_{\varphi,\psi}^{\text{P}}(t) \right] \quad \text{and} \quad M_t = \exp \left[-\int_0^t \lambda_r^{\varphi,\psi} dr \right],$$

we get that

$$e^{-\Gamma_{\varphi,\psi}^{\text{P}}(t)} = M_0 = \mathbb{E}[M_t] = \mathbb{E} \left[\exp \left[-\int_0^t \lambda_r^{\varphi,\psi} dr \right] \right].$$

Comparing this with (5) and differentiating, the claim follows. \square

We consider two test portfolios; see Table 1. For each test portfolio, we compare four different cases, (a) Independence: $\beta^S = \beta^C = 0$, (b) Contagion only: $\beta^S = 0, \beta^C \neq 0$, (c) Systematic risk only: $\beta^S \neq 0, \beta^C = 0$, and (d) Systematic risk and contagion: $\beta^S \neq 0, \beta^C \neq 0$. In each case, the time horizon is $T = 1$.

	N	α	λ	σ	λ_0	γ	β^S	β^C
Portfolio I	200	1	1	0.9	0.5	1	10	3
Portfolio II	200	5	1	1	0.5	0.1	28	1

TABLE 1. Model parameter values for two test portfolios.

In Table 2 we report the losses in the pool at time $T = 1$ for the two portfolios, both in the case that there is a contagion effect (i.e. $\beta^C \neq 0$ and taking the values of Table 1) and in the case that there is no contagion effect (i.e., setting $\beta^C = 0$).

	$L_c(1)$	$L_{nc}(1)$
Portfolio I	0.804	0.470
Portfolio II	0.650	0.589

TABLE 2. Typical default rate $\bar{\ell} = L_c(T)$ at $T = 1$ when contagion is present, i.e., $\beta^C \neq 0$ and $\bar{\ell} = L_{nc}(T)$ at $T = 1$ when contagion is not present, i.e., $\beta^C = 0$.

In Figure 1 we see a comparison of all rate functions I . Note that in both cases $\bar{\ell}$ satisfies $I(\bar{\ell}) = 0$, where the law of large numbers is numerically given by Table 2. Of course, this is expected and is in accordance with large deviations theory.

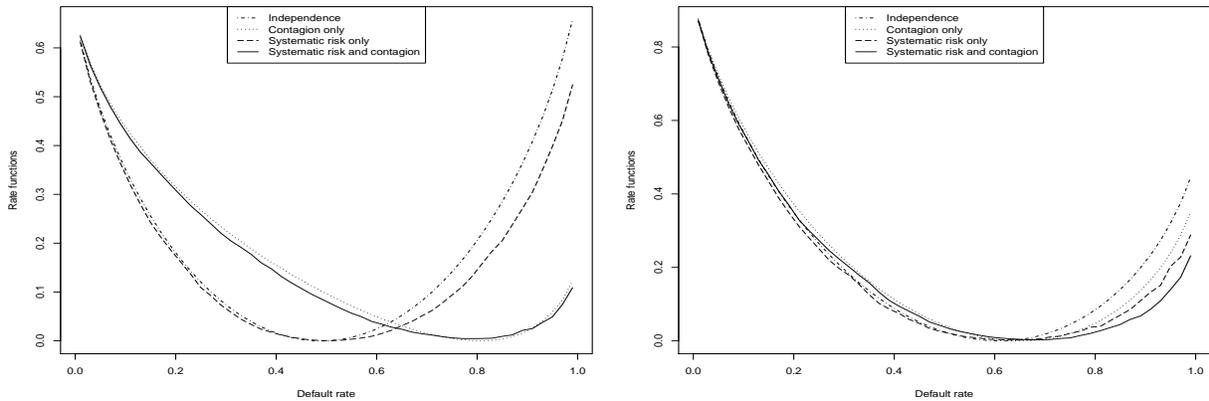


FIGURE 1. The rate function $I'(\ell)$. *Left panel*: Portfolio I. *Right panel*: Portfolio II.

In Figure 2 we see a large deviations approximation to the tail of L_T^N . That is, we plot $\mathbb{P}\{L_T^N \approx \ell\} \asymp \exp[-NI(\ell)]$ where $\ell > \bar{\ell} = L_T$. The exposure to contagion and the systematic risk has significant implications for the tail of L_T^N . Moreover, we see here that, in terms of the tail of the distribution, in Portfolio I, the effect of contagion dominates the effect of systematic risk, whereas, in Portfolio II, the effect of systematic risk dominates the effect of contagion. This is mainly due to the differences in the values of β^S and β^C in the two portfolios, Table 1.

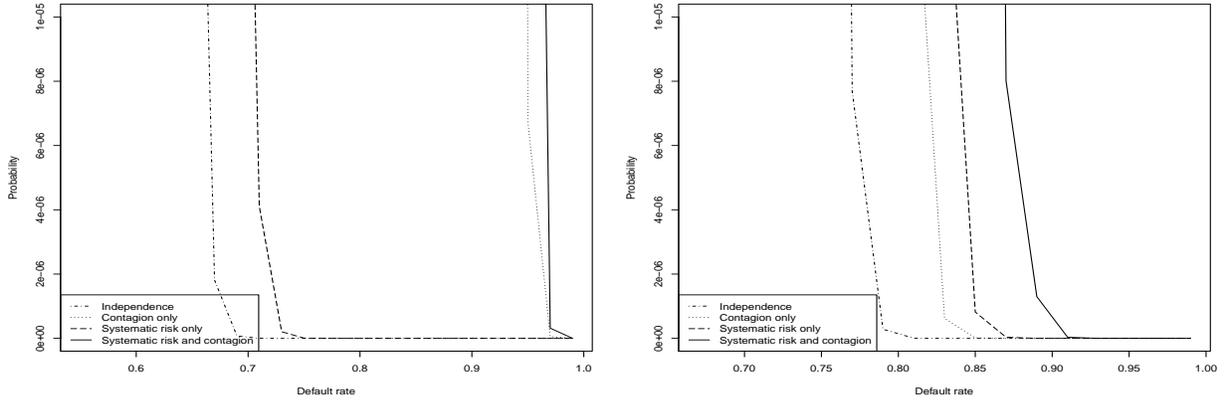


FIGURE 2. Large deviations approximation to the tail of the default rate L_T^N . *Left panel:* Portfolio I. *Right panel:* Portfolio II.

In Figures 3 and 4 we see a comparison of the optimal φ and ψ for a level $\ell = 0.85$. The effect of contagion and of the systematic risk alter the behavior of the extremals (the most likely path to failure). Comparing the tails of the loss distributions of portfolios I and II in Figure 2, we conclude that in Portfolio I, the effect of contagion is in general more profound than the effect of systematic risk. On the other hand in the case of Portfolio II, the effect of systematic risk is more profound than the effect of contagion. Moreover, in the left panel of Figure 4, we see that if a large default cluster occurs, the systematic risk is most likely to play a large role in the initial phase, but then its importance decreases (and thus the contagion effect becomes more important).

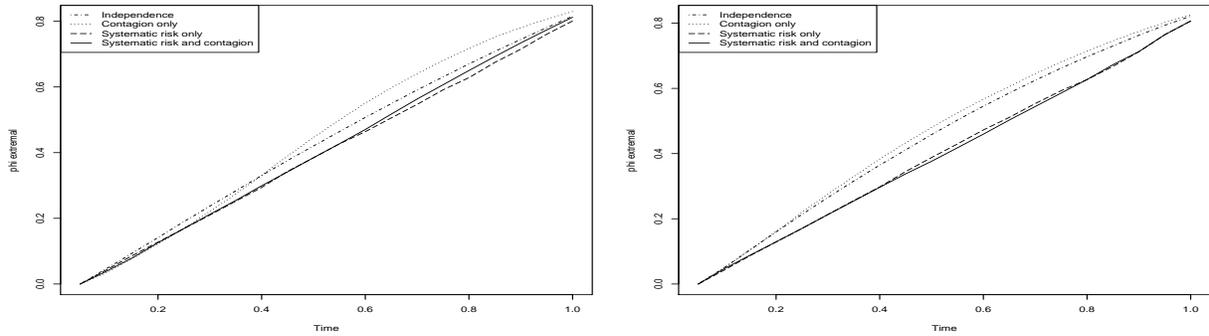


FIGURE 3. Optimal $\phi(t)$ for $t \in [0, 1]$ and $\ell = 0.85$. *Left panel:* Portfolio I. *Right panel:* Portfolio II.

4.2. Numerics for heterogeneous pool. In this subsection, we illustrate numerically the case of a heterogeneous portfolio, which is composed of two types. Type A is 1/3 of the names and type B is 2/3 of the names. In order to illustrate the effect of contagion and systematic risk in such a heterogeneous portfolio, we keep all parameters the same, except for the parameters (β_A^C, β_A^S) and (β_B^C, β_B^S) for types A and B respectively. The systematic risk is assumed to be given by (14).

In particular, we assume that initially we have $N = 200$ names in the pool and we consider a portfolio composed of two types with the following parameters. In such a portfolio we compute that the typical loss at time $T = 1$ is 0.62 if there is no contagion and 0.81 if there is contagion. Below, we see plots comparing the rate functions, tails of the distribution and extremals in all cases, depending on whether both contagion and systematic risk effects are present or not.

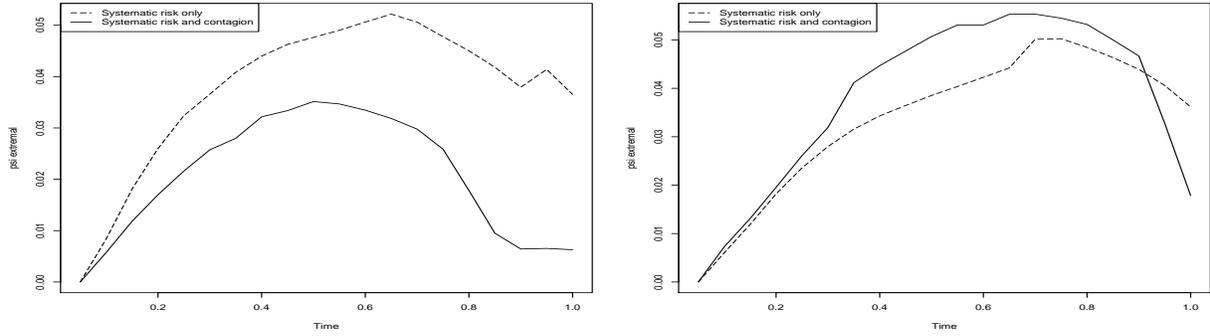


FIGURE 4. Optimal $\psi(t)$ for $t \in [0, 1]$ and $\ell = 0.85$. *Left panel:* Portfolio I. *Right panel:* Portfolio II.

	α	λ	σ	λ_0	γ	β^S	β^C
Type A	1	2	1	0.5	1	5	10
Type B	1	2	1	0.5	1	1	2

TABLE 3. Model parameter values for two types in the portfolios.

In Figure 5 we see a comparison of all rate functions I and of the tails of the loss distribution.

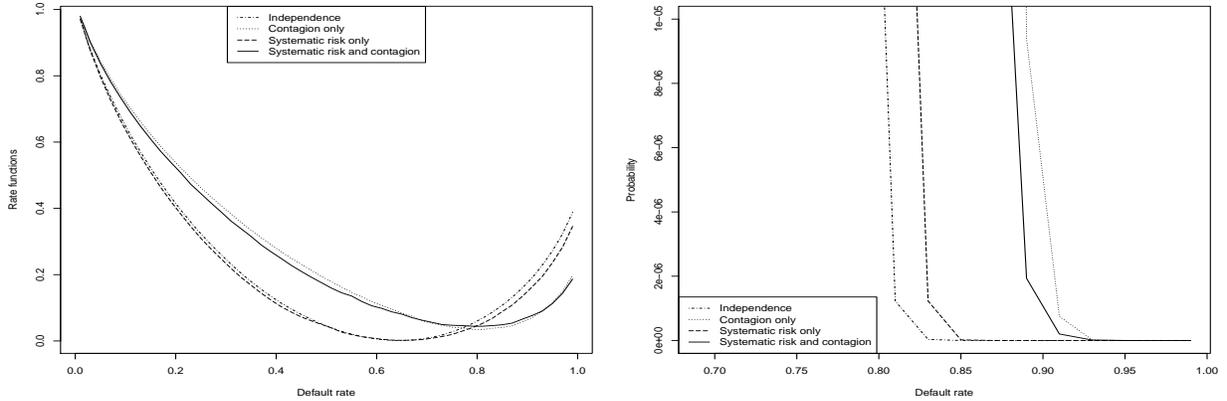


FIGURE 5. The rate function $I(\ell)$ and the tails of the loss distribution

The effect of contagion and of the systematic risk alter the behavior of the extremals (the most likely path to failure). In Figure 6 we see a comparison of the optimal φ for a level of loss $\ell = 0.85$ for types A and B under the different scenarios of contagion effects or not and systematic risk effects or not. In order to illustrate the difference between the behavior of the most likely path to failure between types A and B, we compare the corresponding φ extremals in the left panel of Figure 7 when both effects of contagion and systematic risk are included in the model. We notice that at any given time t , the extremal for type A is bigger than the extremal for type B. This implies that unlikely large losses for components of type A are more likely than unlikely large losses for components of type B. Thus, components of type A affect the pool more than components of type B even though type A composes 1/3 of the pool, whereas type B, composes 2/3 of the pool.

Then, in the right of Figure 7 we see a comparison of the optimal ψ extremals for a level $\ell = 0.85$ for the cases of systematic risk effects are present or not. We see that in the presence of contagion, if a large default cluster occurs, the systematic risk is most likely to play a large role in the initial phase, but then its importance decreases, which means that contagion effect becomes more important. So, in conclusion the analysis of the extremals shows that in such a pool if a large cluster were to occur, it would most likely be due to effects of the systematic risk factor which then affects more names of type A and less names of type B.

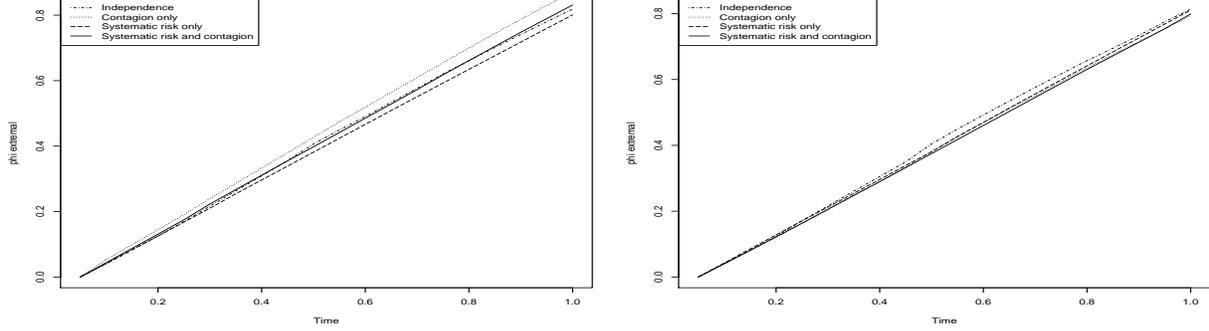


FIGURE 6. Optimal $\phi(t)$ for $t \in [0, 1]$ and $\ell = 0.85$. *Left panel:* ϕ extremal for type A. *Right panel:* ϕ extremal for type B.

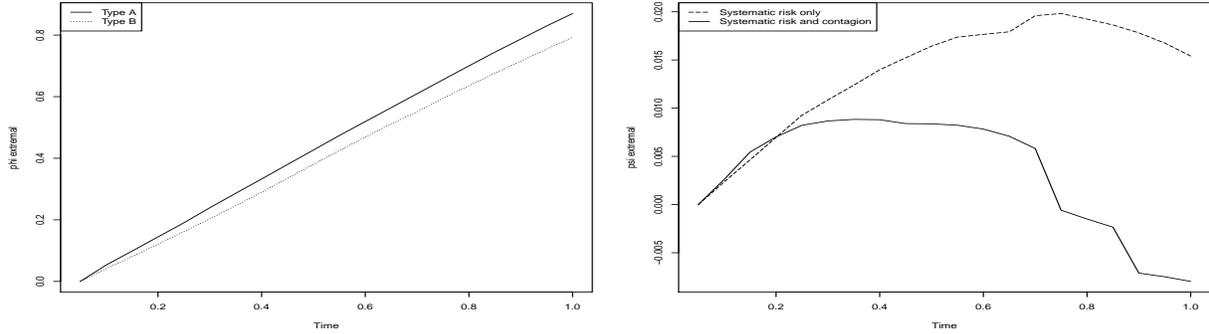


FIGURE 7. *Left panel:* Comparison of ϕ extremals for types A and B when systematic risk effects are present. *Right panel:* Optimal $\psi(t)$ for $t \in [0, 1]$ and $\ell = 0.85$.

5. PROOF OF THE LARGE DEVIATIONS PRINCIPLE

The proof of Theorem 3.1 goes in two steps. First we prove the corresponding result when $\beta^C = 0$, namely we prove the LDP in the heterogeneous case when defaults are independent of each other (Subsection 5.1). Then, based on this result combined with Theorem 5.7, we shall obtain the desired LDP of of Theorem 3.1 in Subsection 5.2. In Subsection 5.3, we prove Theorem 3.10.

5.1. Large deviations under independence and heterogeneity. For each $N \in \mathbb{N}$, let $\{\tau_n^{\text{ind},N}\}_{n=1}^N$ be the independent collection of default times with intensities

$$(16) \quad \begin{aligned} d\lambda_t^{\text{ind},N,n} &= -\alpha_{N,n}(\lambda_t^{\text{ind},N,n} - \bar{\lambda}_{N,n})dt + \sigma_{N,n}\sqrt{\lambda_t^{\text{ind},N,n}}dW_t^n \quad t > 0 \\ \lambda_0^{\text{ind},N,n} &= \lambda_{o,N,n}. \end{aligned}$$

In other words,

$$\tau_n^{\text{ind},N} \stackrel{\text{def}}{=} \inf \left\{ t \geq 0 : \int_0^t \lambda_s^{\text{ind},N,n} \geq \epsilon_n \right\}$$

where the ϵ_n 's are as in (2). Then $\{\tau_n^{\text{ind},N}\}_{n=1}^N$ is a collection of independent random variables (although they are not identically distributed). Define the empirical measure of the types and default times:

$$\nu^{\text{ind},N} = \frac{1}{N} \sum_{n=1}^N \delta_{\mathbf{p}^{N,n}, \tau_n^{\text{ind},N}}.$$

Moreover, let

$$(17) \quad \bar{H}^{\text{ind}}(\nu) = \begin{cases} \int_{\mathbf{p} \in \mathcal{P}} H\left(\frac{d\nu}{dU}(\mathbf{p}), \mu_{0,0}^{\mathbf{p}}\right) U(d\mathbf{p}) & \text{if } \frac{\partial \nu}{\partial U} \text{ exists} \\ \infty & \text{otherwise} \end{cases}$$

where H is as in (9). The definition of \bar{H}^{ind} differs from the definition (10) of \bar{H} in that we replace $\mu_{\nu,0}^{\mathbf{p}}$ in (10) by $\mu_{0,0}^{\mathbf{p}}$. This corresponds to no contagion or systematic effects, which is the same as setting $\beta^C = \beta^S = 0$ leading to independent default rates.

Then, the corresponding large deviations principle reads as follows.

Theorem 5.1. *The family $\{\nu^{\text{ind},N}, N \in \mathbb{N}\}$ satisfies the large deviations principle with rate function $\bar{H}^{\text{ind}}(\nu) : \mathcal{P}(\mathcal{P} \times \mathcal{T}) \mapsto [0, \infty]$.*

Proof. The upper and lower bound follow by Lemmas 5.4 and 5.6 respectively. Let us now prove that this is a lower-semicontinuous functional with compact level sets. Lower-semicontinuity of $H\left(\frac{d\nu}{dU}(\mathbf{p}), \mu^{\mathbf{p}}\right)$ follows immediately, since $H(\nu, \mu)$ is a convex, lower-semicontinuous function of each variable ν or μ separately (e.g., Lemma 1.4.3 in [8]). Compactness of level sets follows again as a consequence of the corresponding property of the relative entropy $H(\nu, \mu)$ (e.g., Lemma 1.4.3 in [8]). These properties are being inherited to $\bar{H}^{\text{ind}}(\nu)$, by Lemma 6.2.16 of [6]. This completes the proof of the Theorem. \square

Remark 5.2. *Under homogeneity, namely if $\mathbf{p}_0^{N,n} = \mathbf{p}_0$ for all $N \in \mathbb{N}$ and $n \in \{1, \dots, N\}$ then the LDP is a direct consequence of Sanov's theorem and is given by (9). The fact that the defaults are not identically distributed poses some additional difficulties in the proof, as it is seen below.*

5.1.1. *Compact Support.* It is often useful (in passing from a large deviations upper bound for compact sets to an upper bound for closed sets) to show sufficient tightness. In our case, Assumption 2.1 actually gives us compact support. Assumption 2.1 implies that there is a bounded subset K of \mathcal{P} such that $\mathbf{p}_0^{N,n} \in K$ for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$. Since \mathcal{T} is itself compact, $[-K_{2,1}, K_{2,1}]^6 \times \mathcal{T}$ is compact, and fact

$$\nu^{\text{ind},N}([-K_{2,1}, K_{2,1}]^6 \times \mathcal{T}) = 1.$$

Moving this to measure space, let's define

$$(18) \quad \mathcal{K} \stackrel{\text{def}}{=} \left\{ \omega \in \mathcal{P}(\mathbf{X}) : \omega([-K_{2,1}, K_{2,1}]^6 \times \mathcal{T}) = 1 \right\}.$$

Then $\mathcal{K} \subset \subset \mathcal{P}(\mathbf{X})$, and \mathbb{P} -a.s. $\nu^{\text{ind},N} \in \mathcal{K}$ for all $N \in \mathbb{N}$.

5.1.2. *Large deviations upper bound.* In this section, we prove the large deviations upper bound.

Let's start with an equivalent characterization of $\bar{H}(\nu)$.

Lemma 5.3. *For every $\nu \in \mathcal{P}(\mathbf{X})$ we have that*

$$\bar{H}(\nu) = \sup_{\phi \in C(\mathcal{P} \times \mathcal{T})} \left\{ \int_{(\mathbf{p}, t) \in \mathbf{X}} \phi(\mathbf{p}, t) \nu(d\mathbf{p}, dt) - \int_{\mathbf{p} \in \mathcal{P}} \log \int_{t \in \mathcal{T}} e^{\phi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) U(d\mathbf{p}) \right\}.$$

Proof. First we assume that the stochastic kernel $\xi = \frac{\partial \nu}{\partial U}$ exists. For $\phi \in C(\mathcal{T})$ and $\mathbf{p} \in \mathcal{P}$, let's define

$$G(\phi, \mathbf{p}) \stackrel{\text{def}}{=} \int_{t \in \mathcal{T}} \phi(t) \xi(\mathbf{p})(dt) - \log \int_{t \in \mathcal{T}} e^{\phi(t)} \mu_{0,0}^{\mathbf{p}}(dt).$$

The Donsker-Varadhan variational representation of the relative entropy [8, Theorem 1.4.3] is that

$$H(\xi(\mathbf{p}), \mu_{0,0}^{\mathbf{p}}) = \sup_{\phi \in C(\mathcal{T})} G(\phi, \mathbf{p}).$$

For any $\phi \in C(\mathcal{X}; \mathbb{R})$, we have that

$$\begin{aligned} & \int_{(\mathbf{p}, t) \in \mathcal{P} \times \mathcal{T}} \phi(\mathbf{p}, t) \nu(d\mathbf{p}, dt) - \int_{\mathbf{p} \in \mathcal{P}} \log \int_{t \in \mathcal{T}} e^{\phi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) U(d\mathbf{p}) \\ &= \int_{\mathbf{p} \in \mathcal{P}} \left\{ \int_{t \in \mathcal{T}} \phi(\mathbf{p}, t) \xi(\mathbf{p})(dt) - \log \int_{t \in \mathcal{T}} e^{\phi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) \\ &= \int_{\mathbf{p} \in \mathcal{P}} G(\phi(\mathbf{p}, \cdot), \mathbf{p}) U(d\mathbf{p}) \leq \int_{\mathbf{p} \in \mathcal{P}} H(\xi(\mathbf{p}), \mu_{0,0}^{\mathbf{p}}) U(d\mathbf{p}) = \bar{H}^{\text{ind}}(\nu). \end{aligned}$$

To prove the opposite inequality, define

$$F_N(\mathbf{p}) = \sup_{\phi \in C(\mathcal{X}), \|\phi\| \leq N} G(\phi(\mathbf{p}, \cdot), \mathbf{p})$$

for all positive integers N . Fix now $\eta > 0$. For each N , let $\phi_N^* \in C(\mathcal{X}; \mathbb{R})$ be such that $\|\phi_N^*\| \leq N$ and $G(\phi_N^*(\mathbf{p}, \cdot), \mathbf{p}) \geq F_N(\mathbf{p}) - \eta$. Since $N \mapsto F_N(\mathbf{p})$ is non-decreasing for each $\mathbf{p} \in \mathcal{P}$, monotone convergence implies that

$$\begin{aligned} \bar{H}(\nu) &= \int_{\mathbf{p} \in \mathcal{P}} \lim_{N \rightarrow \infty} F_N(\mathbf{p}) U(d\mathbf{p}) \leq \lim_{N \rightarrow \infty} \int_{\mathbf{p} \in \mathcal{P}} F_N(\mathbf{p}) U(d\mathbf{p}) \\ &\leq \lim_{N \rightarrow \infty} \int_{\mathbf{p} \in \mathcal{P}} G(\phi_N^*(\mathbf{p}), \mathbf{p}) U(d\mathbf{p}) + \eta \leq \sup_{\phi \in C(\mathcal{X})} \int_{\mathcal{P}} G(\phi, \mathbf{p}) U(d\mathbf{p}) + \eta \end{aligned}$$

Let $\eta \searrow 0$; combining things together, we have the claim when $\frac{\partial \nu}{\partial U}$ exists.

Let's next consider the case where $\frac{\partial \nu}{\partial U}$ is not well defined; thus $\bar{H}^{\text{ind}}(\nu) = \infty$. Then there is an $A \in \mathcal{B}(\mathcal{P})$ and $B \in \mathcal{B}(\mathcal{T})$ such that $\nu(A \times B) > 0$ and $U(B) = 0$. Since \mathcal{P} is Polish, there is a closed subset F of A such that $\nu(F \times B) > 0$. For $c > 0$ and $N \in \mathbb{N}$, define

$$\phi_N^c(\mathbf{p}, t) = c \exp[-N \text{dist}_{\mathcal{X}}((\mathbf{p}, t), F \times B)]$$

where $\text{dist}_{\mathcal{X}}(\cdot, F \times B)$ is the distance (in \mathcal{X}) to $F \times B$. Note that

$$\lim_{N \rightarrow \infty} \phi_N^c = c \chi_{F \times B}$$

pointwise. By dominated convergence, we then have that

$$\begin{aligned} & \sup_{\phi \in C(\mathcal{X})} \left\{ \int_{(\mathbf{p}, t) \in \mathcal{X}} \phi(\mathbf{p}, t) \nu(d\mathbf{p}, dt) - \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\phi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) \right\} \\ & \geq \overline{\lim}_{N \rightarrow \infty} \left\{ \int_{(\mathbf{p}, t) \in \mathcal{X}} \phi_N^c(\mathbf{p}, t) \nu(d\mathbf{p}, dt) - \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\phi_N^c(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) \right\} \\ & = \int_{(\mathbf{p}, t) \in \mathcal{X}} c \chi_{F \times B}(\mathbf{p}, t) \nu(d\mathbf{p}, dt) - \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{c \chi_{F \times B}(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) \end{aligned}$$

If $\mathbf{p} \in F^c$, then

$$\int_{t \in \mathcal{T}} e^{c \chi_{F \times B}(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) = \int_{t \in \mathcal{T}} 1 \mu_{0,0}^{\mathbf{p}}(dt) = \mu_{0,0}^{\mathbf{p}}(\mathcal{T}) = 1.$$

Thus

$$\sup_{\phi \in C(\mathcal{X})} \left\{ \int_{(\mathbf{p}, t) \in \mathcal{X}} \phi(\mathbf{p}, t) \nu(d\mathbf{p}, dt) - \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\phi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) \right\} \geq c \nu(F \times B).$$

Letting $c \nearrow \infty$, we get the claim when $\frac{\partial \nu}{\partial U}$ is not well defined, finishing the proof. \square

We now can prove the upper bound.

Lemma 5.4. For any closed set $F \subset \mathcal{P}(\mathsf{X})$, we have

$$\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P} \{ \nu^{\text{ind}, N} \in F \} \leq - \inf_{\nu \in F} H^{\text{ind}}(\nu)$$

Proof. Using (18), we start with the fact that

$$(19) \quad \mathbb{P} \{ \nu^{\text{ind}, N} \in F \} = \mathbb{P} \{ \nu^{\text{ind}, N} \in F \cap \mathcal{K} \}.$$

Fix $\ell < \inf_{\nu \in F \cap \mathcal{K}} \bar{H}^{\text{ind}}(\nu)$. For every $\phi \in C(\mathsf{X}; \mathbb{R})$, define

$$A_\phi = \left\{ \nu \in \mathcal{P}(\mathsf{X}) : \int_{(\mathbf{p}, t) \in \mathsf{X}} \phi(\mathbf{p}, t) \nu(d\mathbf{p}, dt) - \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\phi(\mathbf{p}, t)} \mu^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) > \ell \right\}$$

By Lemma 5.3, we have that

$$F \cap \mathcal{K} \subset \cup_{\phi \in C(\mathsf{X})} A_\phi.$$

Since $F \cap \mathcal{K}$ is compact, we in fact have that

$$F \cap \mathcal{K} \subset \cup_{\phi \in \Phi} A_\phi$$

for some finite subset Φ of $C(\mathsf{X}; \mathbb{R})$. Consequently

$$(20) \quad \mathbb{P} \{ \nu^{\text{ind}, N} \in F \cap \mathcal{K} \} \leq \sum_{\phi \in \Phi} \mathbb{P} \{ \nu^{\text{ind}, N} \in A_\phi \}.$$

Using the exponential Chebychev inequality, we calculate that

$$\begin{aligned} \mathbb{P} \{ \nu^{\text{ind}, N} \in A_\phi \} &\leq \mathbb{P} \left\{ \int_{(\mathbf{p}, t) \in \mathsf{X}} \phi(\mathbf{p}, t) \nu^{\text{ind}, N}(d\mathbf{p}, dt) \right. \\ &\quad \left. > \ell + \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\phi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) \right\} \\ &= \mathbb{P} \left\{ N \int_{(\mathbf{p}, t) \in \mathsf{X}} \phi(\mathbf{p}, t) \nu^{\text{ind}, N}(d\mathbf{p}, dt) \right. \\ &\quad \left. > N\ell + N \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\phi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) \right\} \\ &\leq \exp[-N\ell] \mathbb{E} \exp \left[N \int_{(\mathbf{p}, t) \in \mathsf{X}} \phi(\mathbf{p}, t) \nu^{\text{ind}, N}(d\mathbf{p}, dt) \right] \\ &\quad \times \exp \left[-N \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\phi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) \right] \\ &\leq \exp[-N\ell] \exp \left[N \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\phi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U_N(d\mathbf{p}) \right] \\ &\quad \times \exp \left[-N \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\phi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) \right] \end{aligned}$$

Thus

$$\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P} \{ \nu^{\text{ind}, N} \in A_\phi \} \leq -\ell.$$

Using this a finite number of times in (20), we get that

$$\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P} \{ \nu^{\text{ind}, N} \in F \cap \mathcal{K} \} \leq -\ell.$$

Letting $\ell \nearrow \inf_{\nu \in F \cap \mathcal{K}} \bar{H}^{\text{ind}}(\nu)$, we have that

$$\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P} \{ \nu^{\text{ind}, N} \in F \cap \mathcal{K} \} \leq - \inf_{\nu \in F \cap \mathcal{K}} \bar{H}^{\text{ind}}(\nu).$$

Finally returning to (19), we calculate that

$$\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P} \{ \nu^{\text{ind}, N} \in F \} \leq - \inf_{\nu \in F \cap \mathcal{K}} \bar{H}^{\text{ind}}(\nu) \leq - \inf_{\nu \in F} \bar{H}^{\text{ind}}(\nu)$$

giving us the claim. \square

5.1.3. *Large deviations lower bound.* In this section, we prove the large deviations lower bound.

We need the following continuity result.

Lemma 5.5. *The map $\mathbf{p} \mapsto \mu_{0,0}^{\mathbf{p}}$ is continuous as a map from \mathcal{P} to $\mathcal{P}(\mathcal{T})$.*

Proof. This follows fairly easily from the explicit formula (7). \square

We can now prove the lower bound.

Lemma 5.6. *Let G be an open subset of $\mathcal{P}(\mathcal{X})$. Then*

$$\underline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P} \{ \nu^{\text{ind}, N} \in G \} \geq - \inf_{\nu \in G} \bar{H}^{\text{ind}}(\nu)$$

Proof. We proceed in a standard way. It suffices to fix a $\nu^* \in G$ such that $\bar{H}^{\text{ind}}(\nu^*) < \infty$ and a $\eta > 0$ and show that

$$(21) \quad \underline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{P} \{ \nu^{\text{ind}, N} \in G \} \geq -\bar{H}^{\text{ind}}(\nu^*) - \eta.$$

Let's understand the implications of $\bar{H}^{\text{ind}}(\nu^*) < \infty$. Then $\xi \stackrel{\text{def}}{=} \frac{\partial \nu^*}{\partial U}$ exists and

$$\int_{\mathbf{p} \in \mathcal{P}} H(\xi(\mathbf{p}), \mu_{0,0}^{\mathbf{p}}) U(d\mathbf{p}) < \infty$$

so $H(\xi(\mathbf{p}), \mu_{0,0}^{\mathbf{p}}) < \infty$ for U -a.e. $\mathbf{p} \in \mathcal{P}$. Thus $\xi(\mathbf{p}) \ll \mu_{0,0}^{\mathbf{p}}$ for U -a.e. $\mathbf{p} \in \mathcal{P}$. Finally, this and Tonelli's theorem ensures that

$$\phi(\mathbf{p}, t) \stackrel{\text{def}}{=} \frac{d\xi(\mathbf{p})}{d\mu_{0,0}^{\mathbf{p}}}(t)$$

is well-defined. In fact, defining the stochastic kernel $\mu_{0,0}^{\bullet}(p) \stackrel{\text{def}}{=} \mu_{0,0}^{\mathbf{p}}$, we have that

$$(22) \quad \phi = \frac{d\nu^*}{d(\mu_{0,0}^{\bullet} \otimes U)}$$

so ϕ is measurable. Using (22), we furthermore note that

$$\nu^* \{ (\mathbf{p}, t) \in \mathcal{X} : \phi(\mathbf{p}, t) = 0 \} = \int_{(\mathbf{p}, t) \in \mathcal{X}} \chi_{\{0\}}(\phi(\mathbf{p}, t)) \phi(\mathbf{p}, t) (\mu_{0,0}^{\bullet} \otimes U)(d\mathbf{p}, dt) = 0.$$

Let's thus define

$$\psi(\mathbf{p}, t) \stackrel{\text{def}}{=} \begin{cases} \ln \phi(\mathbf{p}, t) & \text{if } \phi(\mathbf{p}, t) > 0 \\ -\infty & \text{if } \phi(\mathbf{p}, t) = 0 \end{cases}$$

Defining, for convenience $e^{-\infty} \stackrel{\text{def}}{=} 0$; we know that

$$(23) \quad \int_{t \in \mathcal{T}} e^{\psi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) = \xi(\mathbf{p})(B) = 1.$$

for U -a.e. $\mathbf{p} \in \mathcal{P}$. We also note that

$$\bar{H}^{\text{ind}}(\nu^*) = \int_{(\mathbf{p}, t) \in \mathcal{X}} \psi(\mathbf{p}, t) \nu^*(d\mathbf{p}, dt) - \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\psi(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}).$$

Fix some $\mu^\circ \in \mathcal{P}(\mathcal{T})$. For $\Psi \in C(\mathcal{X}; \mathbb{R})$, define now the stochastic kernel

$$\Xi_\Psi(\mathbf{p})(B) \stackrel{\text{def}}{=} \begin{cases} \frac{\int_{t \in B} \exp[\Psi(\mathbf{p}, t)] \mu_{0,0}^{\mathbf{p}}(dt)}{\int_{t \in \mathcal{T}} \exp[\Psi(\mathbf{p}, t)] \mu_{0,0}^{\mathbf{p}}(dt)} & \text{if } \int_{t \in \mathcal{T}} \exp[\Psi(\mathbf{p}, t)] \mu_{0,0}^{\mathbf{p}}(dt) > 0 \\ \mu^\circ(B) & \text{if } \int_{t \in \mathcal{T}} \exp[\Psi(\mathbf{p}, t)] \mu_{0,0}^{\mathbf{p}}(dt) = 0 \end{cases} \quad B \in \mathcal{B}(\mathcal{T})$$

Let's now approximate, keeping in mind Lemma 5.5, (22), and the fact that $x \mapsto e^{-x}$ is continuous on $[-\infty, \infty)$. We then have that there is a $\hat{\psi} \in C(\mathbf{X}; \mathbb{R})$ such that

$$\left| \bar{H}^{\text{ind}}(\nu^*) - \int_{(\mathbf{p}, t) \in \mathbf{X}} \hat{\psi}(\mathbf{p}, t) \nu^*(d\mathbf{p}, dt) - \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\hat{\psi}(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) \right| < \eta/3$$

and such that the measure $\hat{\nu} \stackrel{\text{def}}{=} \Xi_{\hat{\psi}} \otimes U$ is in G .

Let's now define

$$(24) \quad S_N \stackrel{\text{def}}{=} \int_{(\mathbf{p}, t) \in \mathbf{X}} \hat{\psi}(\mathbf{p}, t) \nu^{\text{ind}, N}(d\mathbf{p}, dt) - \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\hat{\psi}(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U_N(d\mathbf{p}).$$

Then $\mathbb{E}[\exp[NS_N]] = 1$, and we thus define

$$\hat{\mathbb{P}}_N(A) \stackrel{\text{def}}{=} \mathbb{E}[\chi_A \exp[NS_N]] \quad A \in \mathcal{F}$$

The next step is to reduce G . Let $B \subset \mathbf{X}$ be an open subset of G which contains $\hat{\nu}$ and such that

$$\left| \int_{(\mathbf{p}, t) \in \mathbf{X}} \hat{\psi}(\mathbf{p}, t) \nu(d\mathbf{p}, dt) - \int_{(\mathbf{p}, t) \in \mathbf{X}} \hat{\psi}(\mathbf{p}, t) \nu^*(d\mathbf{p}, dt) \right| < \eta/2$$

for all $\nu \in B$. Thus

$$\mathbb{P}\{\nu^{\text{ind}, N} \in G\} \geq \mathbb{P}\{\nu^{\text{ind}, N} \in B\} = \hat{\mathbb{E}}_N[\chi_{\{\nu^{\text{ind}, N} \in B\}} \exp[-NS_N]].$$

Thanks to Lemma 5.5, the map

$$\mathbf{p} \mapsto \int_{t \in \mathcal{T}} e^{\hat{\psi}(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt)$$

is in $C(\mathcal{P}; \mathbb{R})$. Thus we can find an $N^* \in \mathbb{N}$ such that

$$\left| \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\hat{\psi}(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U(d\mathbf{p}) - \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\hat{\psi}(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U_N(d\mathbf{p}) \right| \leq \eta/2$$

if $N \geq N^*$. Combining things together, we have that

$$\left| \int_{(\mathbf{p}, t) \in \mathbf{X}} \hat{\psi}(\mathbf{p}, t) \nu^{\text{ind}, N}(d\mathbf{p}, dt) - \int_{\mathbf{p} \in \mathcal{P}} \left\{ \log \int_{t \in \mathcal{T}} e^{\hat{\psi}(\mathbf{p}, t)} \mu_{0,0}^{\mathbf{p}}(dt) \right\} U_N(d\mathbf{p}) - \bar{H}^{\text{ind}}(\nu^*) \right| < \eta$$

if $\nu^{\text{ind}, N} \in B$ and $N \geq N^*$.

Assume now that $N \geq N^*$. Then

$$\begin{aligned} \mathbb{P}\{\nu^{\text{ind}, N} \in G\} &\geq \mathbb{P}\{\nu^{\text{ind}, N} \in B\} = \hat{\mathbb{E}}_N[\chi_{\{\nu^{\text{ind}, N} \in B\}} \exp[-NS_N]] \\ &\geq \hat{\mathbb{P}}_N\{\nu^{\text{ind}, N} \in B\} \exp[-N\{\bar{H}^{\text{ind}}(\nu^*) + \eta\}]. \end{aligned}$$

To finish the proof, let's show that

$$(25) \quad \varliminf_{N \rightarrow \infty} \hat{\mathbb{P}}_N\{\nu^{\text{ind}, N} \in B\} > 0.$$

To do so, let's construct a metric on $\mathcal{S}(\mathbf{X})$. Define $d(x) \stackrel{\text{def}}{=} |x|/(1+|x|)$; then the map $(x, y) \mapsto d(x-y)$ is a metric on \mathbb{R} . Let $\{f_n\}_{n \in \mathbb{N}}$ be a countable and dense subset of $C(\mathbf{X}; \mathbb{R})$, and define

$$(26) \quad \rho(\nu_1, \nu_2) \stackrel{\text{def}}{=} \sum_{n \in \mathbb{N}} 2^{-n} d \left(\int_{(\mathbf{p}, t) \in \mathbf{X}} f_n(\mathbf{p}, t) \nu_1(d\mathbf{p}, dt) - \int_{(\mathbf{p}, t) \in \mathbf{X}} f_n(\mathbf{p}, t) \nu_2(d\mathbf{p}, dt) \right)$$

for all ν_1 and ν_2 in $\mathcal{S}(\mathbf{X})$. Then ρ is a metric on $\mathcal{S}(\mathbf{X})$. Since $\hat{\nu} \in B$, there is a $\delta_{(27)} > 0$ small enough such that

$$(27) \quad \{\nu' \in \mathcal{S}(\mathbf{X}) : \rho(\nu', \hat{\nu}) < \delta_{(27)}\} \subset B.$$

Let's now understand the statistics of $\nu^{\text{ind}, N}$ under $\hat{\mathbb{P}}_N$. From the structure (24) of S_N , we see that, under $\hat{\mathbb{P}}_N$, $\{\tau_n^{\text{ind}, N}\}_{n=1}^N$ are independent and τ_n^N has law $\Xi_{\hat{\psi}}(\mathbf{p}_n^N)$.

Fix $f \in C(X; \mathbb{R})$. Let's write

$$\begin{aligned} & \left| \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) \nu^{\text{ind}, N}(d\mathbf{p}, dt) - \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) \hat{\nu}(d\mathbf{p}, dt) \right| \\ & \leq \left| \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) \nu^{\text{ind}, N}(d\mathbf{p}, dt) - \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) (\Xi_{\hat{\psi}} \otimes U_N)(d\mathbf{p}, dt) \right| \\ & \quad + \left| \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) (\Xi_{\hat{\psi}} \otimes U_N)(d\mathbf{p}, dt) - \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) (\Xi_{\hat{\psi}} \otimes U)(d\mathbf{p}, dt) \right|. \end{aligned}$$

Keeping Lemma 5.5 in mind, we have that

$$\lim_{N \rightarrow \infty} \left| \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) (\Xi_{\hat{\psi}} \otimes U_N)(d\mathbf{p}, dt) - \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) (\Xi_{\hat{\psi}} \otimes U)(d\mathbf{p}, dt) \right| = 0.$$

Let's next use Chebychev's inequality; we have arranged things to take advantage of the fact that

$$\hat{\mathbb{E}}_N [f(\mathbf{p}_n^N, \sigma_n^N)] = \int_{t \in \mathcal{T}} f(\mathbf{p}_n^N, t) \Xi_{\hat{\psi}}^{\mathbf{p}_n^N}(dt)$$

for all N and n . We then calculate that

$$\begin{aligned} & \hat{\mathbb{E}}_N \left[\left| \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) \nu^{\text{ind}, N}(d\mathbf{p}, dt) - \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) (\Xi_{\hat{\psi}} \otimes U_N)(d\mathbf{p}, dt) \right|^2 \right] \\ & = \hat{\mathbb{E}}_N \left[\left| \frac{1}{N} \sum_{n=1}^N \left\{ f(\mathbf{p}_n^N, \sigma_n^N) - \int_{t \in \mathcal{T}} f(\mathbf{p}_n^N, t) \Xi_{\hat{\psi}}^{\mathbf{p}_n^N}(dt) \right\} \right|^2 \right] \\ & = \frac{1}{N^2} \sum_{n=1}^N \hat{\mathbb{E}}_N \left[\left\{ f(\mathbf{p}_n^N, \sigma_n^N) - \int_{t \in \mathcal{T}} f(\mathbf{p}_n^N, t) \Xi_{\hat{\psi}}^{\mathbf{p}_n^N}(dt) \right\}^2 \right] \\ & \leq \frac{4}{N} \sup_{(\mathbf{p}, t) \in X} |f(\mathbf{p}, t)|^2. \end{aligned}$$

Let's collect things together. For each $f \in C(X; \mathbb{R})$, we have that

$$\lim_{N \rightarrow \infty} \hat{\mathbb{E}}_N \left[\left| \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) \nu^{\text{ind}, N}(d\mathbf{p}, dt) - \int_{(\mathbf{p}, t) \in X} f(\mathbf{p}, t) \hat{\nu}(d\mathbf{p}, dt) \right|^2 \right] = 0.$$

The structure of (26) then implies that

$$\lim_{N \rightarrow \infty} \hat{\mathbb{P}}_N \{ \rho(\nu^{\text{ind}, N}, \hat{\nu}) > \delta_{(27)} \} = 0.$$

This then implies (25), finishing the proof. \square

5.2. Proof of Theorem 3.1. In this section we prove Theorem 3.1, using Varadhan's integral lemma [22] as it appears in Theorem II.7.2 of [9].

Theorem 5.7 (Varadhan's integral lemma). *Let S be a Polish space. Suppose that $\{\xi^N\}_{N \in \mathbb{N}}$ is a sequence of S -valued random variables with large deviations principle with rate function I . Let $\{\tilde{\xi}^N\}_{N \in \mathbb{N}}$ be another sequence of S -valued random variables and assume that there is a continuous function $G : S \rightarrow \mathbb{R}$ such that*

$$\mathbb{P}\{\tilde{\xi}^N \in A\} = \mathbb{E}[\chi_A(\xi^N) \exp[NG(\xi^N)]]$$

for all $A \in \mathcal{B}(S)$. If

$$\lim_{\alpha \rightarrow \infty} \overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \ln \mathbb{E} \left[\chi_{[\alpha, \infty)}(G(\xi^N)) e^{NG(\xi^N)} \right] = -\infty$$

then $\{\tilde{\xi}^N\}_{N \in \mathbb{N}}$ has a large deviation principle with good rate function $I - G$.

In our case, we take $S \stackrel{\text{def}}{=} \mathcal{P}(X)$ and $\xi_N \stackrel{\text{def}}{=} \nu^{\text{ind},N}$. Lemma 4.1 allows us to define the Radon-Nikodym derivative e^{NG} .

For $\mathbf{p} \in \mathcal{P}$, define

$$\begin{aligned} \dot{b}^{\mathbf{p}}(t) &= 1 - \frac{1}{2}\sigma^2 (b^{\mathbf{p}}(t))^2 - \alpha b^{\mathbf{p}}(t) \quad t > 0 \\ b^{\mathbf{p}}(0) &= 0. \end{aligned}$$

Note that since $\beta^S = 0$ then $\theta_t^{\mathbf{p}}(s) = b^{\mathbf{p}}(s)$ for all $0 \leq s \leq t \leq T$, where $\theta_t^{\mathbf{p}}(s)$ is defined in (15). It is easy to see that $b^{\mathbf{p}}$ and $\dot{b}^{\mathbf{p}}$ are uniformly bounded in $\mathcal{P} \times [0, T]$. Using (6) and Lemma 4.1 we can fairly easily compute the ratio of densities of τ_n^N and $\tau_n^{\text{ind},N}$ and consequently of ν^N and $\nu^{\text{ind},N}$. For $\mathbf{p} \in \mathcal{P}$, $r \in [0, T]$, and $\nu \in \mathcal{P}(\mathcal{P} \times [0, T])$, define

$$\begin{aligned} g_\nu(\mathbf{p}, r) &\stackrel{\text{def}}{=} \log \frac{d\mu_{\nu,0}^{\mathbf{p}}([0, r])}{d\mu_{0,0}^{\mathbf{p}}([0, r])} = \log \left\{ \dot{b}^{\mathbf{p}}(r)\lambda_0 + \alpha \bar{\lambda} b^{\mathbf{p}}(r) + \beta^C \int_0^r \dot{b}^{\mathbf{p}}(r-u)\nu(\mathcal{P}, du) \right\} \\ &\quad - \log \left(\dot{b}^{\mathbf{p}}(r)\lambda_0 + \alpha \bar{\lambda} b^{\mathbf{p}}(r) \right) - \beta^C \int_0^r b^{\mathbf{p}}(r-u)\nu(\mathcal{P}, du) \end{aligned}$$

For $\mathbf{p} \in \mathcal{P}$, let's also define

$$g_\nu(\mathbf{p}, \star) = \beta^C \int_0^T b^{\mathbf{p}}(T-u)\nu(\mathcal{P}, du)$$

For $\nu \in \mathcal{P}(X)$, let's finally define

$$G(\nu) \stackrel{\text{def}}{=} \int_{(\mathbf{p}, t) \in X} g_\nu(\mathbf{p}, t)\nu(d\mathbf{p}, dt).$$

Then

$$\mathbb{P} \{ \nu^N \in A \} = \mathbb{E} \left[\chi_A(\nu^{\text{ind},N}) \exp [NG(\nu^{\text{ind},N})] \right].$$

Theorem 3.1 follows by Theorem 5.8 and contraction principle.

Theorem 5.8. *The family $(\nu^N)_{N \in \mathbb{N}}$ of (8) satisfies the large deviations principle with rate function $\bar{H}(\nu) : \mathcal{P}(\mathcal{P} \times \mathcal{T}) \mapsto [0, \infty]$.*

Proof. Note that $G : \mathcal{P}(X) \rightarrow \mathbb{R}$ is continuous in the weak topology and bounded. Thus

$$\lim_{\alpha \rightarrow \infty} \overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E} \left[\chi_{[\alpha, \infty)}(G(\nu^{\text{ind},N})) \exp [NG(\nu^{\text{ind},N})] \right] = -\infty$$

holds. Thus $\{\nu^N, N \in \mathbb{N}\}$ has a large deviations principle with rate function $\bar{H} \stackrel{\text{def}}{=} \bar{H}^{\text{ind}} - G$. Fix $\nu \in \mathcal{P}(X)$. If $\frac{\partial \nu}{\partial U}$ does not exist, then $\bar{H}^{\text{ind}}(\nu) = \infty$ so $\bar{H}(\nu) = \infty$. If $\frac{\partial \nu}{\partial U}$ exists, then

$$\begin{aligned} \bar{H}(\nu) &= \int_{\mathbf{p} \in \mathcal{P}} \left\{ \int_{t \in \mathcal{T}} \ln \frac{\frac{\partial \nu}{\partial U}(\mathbf{p})}{d\mu_{0,0}^{\mathbf{p}}(t)}(t) \frac{d\nu}{dU}(\mathbf{p}, dt) \right\} U(d\mathbf{p}) - \int_{(\mathbf{p}, t) \in X} \ln \frac{d\mu_{\nu,0}^{\mathbf{p}}(s)}{d\mu_{0,0}^{\mathbf{p}}(s)} \nu(d\mathbf{p}, ds) \\ &= \int_{\mathbf{p} \in \mathcal{P}} \left\{ \int_{t \in \mathcal{T}} \left\{ \ln \frac{\frac{\partial \nu}{\partial U}(\mathbf{p})}{d\mu_{0,0}^{\mathbf{p}}(t)}(t) - \ln \frac{d\mu_{\nu,0}^{\mathbf{p}}(s)}{d\mu_{0,0}^{\mathbf{p}}(s)} \right\} \frac{\partial \nu}{\partial U}(\mathbf{p}, ds) \right\} U(d\mathbf{p}) \\ &= \int_{\mathbf{p} \in \mathcal{P}} \left\{ \int_{t \in \mathcal{T}} \ln \frac{\frac{\partial \nu}{\partial U}(\mathbf{p})}{d\mu_{\nu,0}^{\mathbf{p}}(t)}(t) \frac{d\nu}{dU}(\mathbf{p}, dt) \right\} U(d\mathbf{p}) \\ &= \int_{\mathbf{p} \in \mathcal{P}} H \left(\frac{d\nu}{dU}(\mathbf{p}), \mu_{\nu,0}^{\mathbf{p}} \right) U(d\mathbf{p}) \end{aligned}$$

□

We conclude with the proof of Corollary 3.3.

Proof of Corollary 3.3. Clearly, it is enough to consider the case $\bar{H}(\nu) < \infty$. $H(\cdot, \cdot)$ is relative entropy, which implies that it is a convex and lower semicontinuous function with respect to both arguments and strictly convex as a function of the first argument for each fixed second argument. Thus $\bar{H}(\nu) = 0$ implies that $H\left(\frac{d\nu}{dU}(\mathbf{p}), \mu_{\bar{\nu}, 0}^{\mathbf{p}}\right) = 0$, which due to the relative entropy nature of H , it is true if $\frac{\partial\nu}{\partial U}(\mathbf{p}) = \mu_{\bar{\nu}, 0}^{\mathbf{p}}$ U -a.e. This implies the relation

$$\nu(\mathcal{P}, B) = \int_{\mathcal{P} \times B} \mu_{\bar{\nu}, 0}^{\mathbf{p}}(ds) U(d\mathbf{p}), \text{ for all } B \in \mathcal{B}(\mathcal{T})$$

which by (6) and Lemma 4.1, translates to the equation

$$\begin{aligned} \bar{\nu}([0, t]) &= \int_{\mathcal{P}} \left[1 - \mathbb{E} \left[e^{-\int_0^t \lambda_s^{\bar{\nu}, 0}(\mathbf{p}) ds} \right] \right] U(d\mathbf{p}) \\ (28) \quad &= 1 - \int_{\mathcal{P}} e^{-(b^{\mathbf{p}}(t)\lambda_0 + \int_0^t \alpha \bar{\lambda} b^{\mathbf{p}}(t-u) du + \beta^C \int_0^t b^{\mathbf{p}}(t-u) \bar{\nu}(du))} U(d\mathbf{p}) \end{aligned}$$

where $\lambda^{\nu, 0}(\mathbf{p})$ satisfies (4) with $\varphi(t) = \bar{\nu}([0, t])$ and $\psi(t) = 0$. Let us next argue that this equation has a unique solution which coincides with L_t , the limit in probability of the empirical loss L_t^N as $N \rightarrow \infty$.

By Lemma 4.1 of [17] we get that

$$(29) \quad L_t = 1 - \mu_t(\mathcal{P}) = 1 - \int_{\mathcal{P}} \mathbb{E} \left[e^{-\int_0^t \lambda_s^*(\mathbf{p}) ds} \right] U(d\mathbf{p})$$

where for each $\mathbf{p} \in \mathcal{P}$, there is a unique pair $\{(Q(t), \lambda_t^*(\mathbf{p})) : t \in [0, T]\}$ taking values in $\mathbb{R}_+ \times \mathbb{R}_+$ such that

$$Q(t) = \int_{\mathcal{P}} \mathbb{E} \left\{ \lambda_t^*(\mathbf{p}) \exp \left[-\int_0^t \lambda_s^*(\mathbf{p}) ds \right] \right\} U(d\mathbf{p}).$$

and

$$\lambda_t^*(\mathbf{p}) = \lambda_0 - \alpha \int_0^t (\lambda_s^*(\mathbf{p}) - \bar{\lambda}) ds + \sigma \int_0^t \sqrt{\lambda_s^*(\mathbf{p})} dW_s^* + \beta^C \int_0^t Q(s) ds.$$

Then, using Lemma 4.1, one easily computes from (29) that $L_t = \bar{\nu}([0, t])$, the unique solution of (28). \square

5.3. Proof of Theorem 3.10. In this Subsection, we consider Theorem 3.10. The proof is a direct consequence of (a): Theorem 5.8 and (b): the LDP for the $X_t^N = \varepsilon_N X_t$ process, Lemma 5.10. Thus, we only sketch the main arguments.

Firstly, we notice that an immediate consequence of Theorem 5.8, is the following conditional LDP.

Lemma 5.9. *Consider the system defined in (1) with $\varepsilon_N = 1$. Under Assumptions 2.1 and 2.2, and given a path $t \mapsto X_t$ in $C_c([0, T]; \mathbb{R})$, the family $\{\nu^N, N \in \mathbb{N}\}$ satisfies the conditional large deviation principle with rate function $\bar{H}(\nu; \mathbf{X}) : \mathcal{P}(\mathcal{P} \times \mathcal{T}) \mapsto [0, \infty]$ and speed N , where*

$$\bar{H}(\nu; \mathbf{X}) = \begin{cases} \int_{\mathcal{P}} H\left(\frac{\partial\nu}{\partial U}(\mathbf{p}), \mu_{\bar{\nu}, \mathbf{X}}^{\mathbf{p}}\right) U(d\mathbf{p}), & \text{if } \frac{\partial\nu}{\partial U}(\mathbf{p}) \text{ exists} \\ +\infty, & \text{otherwise.} \end{cases}$$

Secondly, we notice that under Assumptions 3.6, 3.7 and 3.8 the family $\{X^N = \varepsilon_N X, N \in \mathbb{N}\}$ satisfies a large deviations principle as given by Lemma 5.10.

Lemma 5.10. *Let $X^N = \varepsilon_N X$ and assume that Assumptions 3.6, 3.7 and 3.8 hold. Then, the process $\{X^N, N \in \mathbb{N}\}$ satisfies the large deviations principle with rate function $J_X(\cdot)$ as given by (13) and speed $1/\varepsilon_N^{2\zeta}$.*

Proof. Notice that $X^N = \varepsilon_N X$ is the unique strong solution of the SDE

$$dX_t^N = \varepsilon_N b\left(\frac{X_t^N}{\varepsilon_N}\right) dt + \varepsilon_N^\zeta \left[\varepsilon_N^{1-\zeta} \kappa\left(\frac{X_t^N}{\varepsilon_N}\right) \right] dV_t, \quad X_0^N = \varepsilon_N x_0$$

It follows then by Theorem 2.1 in [2], or in the case of non-degenerate bounded diffusion from the classical results in Section 5.3. of [11], that under Assumptions 3.6, 3.7 and 3.8 the processes $\{\varepsilon_N X, N < \infty\}$ and $\{\bar{X}^N, N < \infty\}$ defined in (12) have the same large deviations principle with rate function $J_X(\cdot)$ and where speed of the large deviations asymptotic is $1/\varepsilon_N^{2\zeta}$. \square

By Lemma 5.9 and contraction principle we get that, given a path $t \mapsto X_t$, the large deviations rate function for the family $\{L_t^N = \nu^N(\mathcal{P} \times [0, t]), N \in \mathbb{N}, t \in [0, T]\}$ is given by

$$(30) \quad J(\varphi; \mathbf{X}) = \begin{cases} \int_{\mathcal{P}} g\left(\varphi(\mathbf{p}), f_{\bar{\varphi}, \mathbf{X}}^{\mathbf{p}}(t)\right) U(d\mathbf{p}), & \varphi \in AC(\mathcal{P} \times [0, T]; \mathbb{R}), \varphi(\mathbf{p}, 0) = 0, \varphi \geq 0, \\ & \text{and } \bar{\varphi}(s) = \int_{\mathcal{P}} \varphi(\mathbf{p}, s) U(d\mathbf{p}) \\ +\infty, & \text{otherwise} \end{cases}$$

The latter statement and Lemma 5.10 imply that the rate function for the pair $\{(L_t^N, X_t^N), N \in \mathbb{N}, t \in [0, T]\}$ is given by $S(\varphi, \psi)$, as defined in Theorem 3.10. Indeed, by Lemma 1.2.18 and Theorem 4.1.11 of [6], it is enough to establish a local large deviations principle and exponential tightness. Exponential tightness is a direct consequence of the compact support computations of Section 5.1.1 and the exponential tightness of the small noise diffusion process X^N . Denoting by $B(\cdot, \delta)$ the ball in the space of continuous functions in $[0, T]$ with uniform norm, proving a local large deviations principle amounts to proving

$$\begin{aligned} \lim_{\delta \downarrow 0} \limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P} \{(L^N, X^N) \in B((\varphi, \psi), \delta)\} &\leq -S(\varphi, \psi), \quad \text{and} \\ \lim_{\delta \downarrow 0} \liminf_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P} \{(L^N, X^N) \in B((\varphi, \psi), \delta)\} &\geq -S(\varphi, \psi) \end{aligned}$$

In the case $\lim_{N \rightarrow \infty} N\varepsilon_N^{2\zeta} = c \in (0, \infty)$, both statements are seen to be true due to the conditional probability relation $\mathbb{P}(A \cap \Gamma) = \mathbb{P}(A|\Gamma)\mathbb{P}(\Gamma)$ and using the conditional large deviations principle discussed at Lemma 5.9 and the large deviations principle of Lemma 5.10. Basically, to exponential order, we have as $N \rightarrow \infty$

$$(31) \quad \mathbb{P}\{L^N \approx \varphi, X^N \approx \psi\} \approx \exp \left[-N J(\varphi; \psi) - \frac{1}{\varepsilon_N^{2\zeta}} J_X(\psi) \right].$$

which means that if $\lim_{N \rightarrow \infty} N\varepsilon_N^{2\zeta} = c \in (0, \infty)$, the rate function for the pair $\{(L_t^N, X_t^N), N \in \mathbb{N}, t \in [0, T]\}$ is given by $S(\varphi, \psi)$, as defined in Theorem 3.10. Then, by varying over $\psi \in C([0, T]; \mathbb{R})$ and $\varphi \in C(\mathcal{P} \times [0, T]; \mathbb{R})$ such that $\bar{\varphi}(T) = \ell$, we get that the rate function for the loss $\{L_T^N, N < \infty\}$ at time T is $I(\cdot)$, which is the statement of Theorem 3.10, concluding its proof.

Display (31) also shows that if $\lim_{N \rightarrow \infty} N\varepsilon_N^{2\zeta} = \infty$, then the term $J(\varphi; \psi)$ is the dominant factor, while if $\lim_{N \rightarrow \infty} N\varepsilon_N^{2\zeta} = 0$, then the $J_X(\psi)$ entropy term is the dominant factor.

6. CONCLUSION

In this paper we studied large deviations for an empirically motivated interacting particle system of default clustering. The components in the system interact through the empirical default rate in the pool and through a systematic risk which is common to all of the them. An explicit large deviations principle is derived and probabilities of tail events are then studied. One can compute the extremals of the rate function, which characterize the most likely path to failure of the system. The numerical experiments reveal the role that contagion and systematic risk play in the failure of the system. In the numerical examples that we performed, we saw that if a large default cluster occurs, the systematic risk is most likely to play a large role in the initial phase, but then its importance decreases, and then contagion effect become more important. Then, the analysis of the extremals in heterogeneous pools can help understanding which components in the pool are more vulnerable to the effect of contagion.

7. ACKNOWLEDGEMENTS

During preparation of this work, K.S. was partially supported by the National Science Foundation (DMS 1312124) and R.S. was partially supported by the National Science Foundation (DMS 1312071).

REFERENCES

- [1] S. Azizpour, K. Giesecke and G. Schwenkler, Exploring the Sources of Default Clustering, working paper, Stanford University, (2014).
- [2] A. Chiarini and M. Fischer, On large deviations for small noise Itô processes, *Advances in Applied Probability*, (2014), to appear.

- [3] P. Dai Pra, W. Runggaldier, E. Sartori, and M. Tolotti, Large portfolio losses: A dynamic contagion model, *The Annals of Applied Probability*, Vol. 19, (2009), pp. 347–394.
- [4] P. Dai Pra and M. Tolotti, Heterogeneous credit portfolios and the dynamics of the aggregate losses, *Stochastic Processes and their Applications*, Vol. 119, (2009), pp. 2913–2944.
- [5] A. Dembo, J.-D. Deuschel, and D. Duffie, Large portfolio losses, *Finance Stoch.*, Vol. , No. 1, (2004), pp. 3–16.
- [6] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics*, Springer-Verlag, New York, 2nd ed., (1998).
- [7] D. Duffie, J. Pan, and K. Singleton, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica*, 68, (2000), pp. 1343–1376.
- [8] P. Dupuis and R.S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations* Wiley series in probability and statistics, (1997).
- [9] R.S. Ellis, *Entropy, Large Deviations and Statistical Mechanics*, Volume 271 of Grundlehren der Mathematischen Wissenschaften, Springer, New York, (1985).
- [10] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, John Wiley & Sons Inc., New York, (1986).
- [11] M.I. Freidlin and A.D. Wentzell, *Random Perturbations of Dynamical Systems*, 3rd Edition, Springer, (2012).
- [12] J. Garnier, G. Papanicolaou and T.-W. Yang, Large deviations for a mean field model of systemic risk. *SIAM Journal on Financial Mathematics*, Vol. 4, (2012), pp. 151–184.
- [13] J. Gärtner, On the McKean-Vlasov Limit for interacting diffusions, *Mathematische Nachrichten*, Volume 137, Issue 1, (1988), pp. 197–248
- [14] P. Glasserman, Tail approximations for portfolio credit risk, *Journal of Derivatives*, Vol. 1, No. 2, (2004), pp. 24–42.
- [15] P. Glasserman and J. Li, Importance sampling for portfolio credit risk, *Management Science*, Vol. 51, (2005), pp. 1643–1656.
- [16] P. Glasserman, W. Kang and P. Shahabuddin, Large deviations in multifactor portfolio credit risk, *Mathematical Finance*, Vol. 17, No. 3, (2007), pp. 345–379
- [17] K. Giesecke, K. Spiliopoulos and R. Sowers, Default clustering in large portfolios: typical events, *Annals of Applied Probability*, Vol. 23, No. 1, (2013), pp. 348–385.
- [18] K. Giesecke, K. Spiliopoulos, R. Sowers and J. A. Sirignano, Large portfolio asymptotics for losses from default, *Mathematical Finance*, (2013), to appear.
- [19] C. Meinerding, Asset allocation and asset pricing in the face of systemic risk: a literature overview and assessment, *International Journal of Theoretical and Applied Finance (IJTAF)*, Vol. 15, No. 03, (2012), pp. 1250023-1–1250023-27.
- [20] K. Spiliopoulos, J. A. Sirignano and K. Giesecke, Fluctuation analysis for the loss from default, *Stochastic Processes and their Applications*, Vol. 124, No. 7, (2014), pp. 2322–2362.
- [21] K. Spiliopoulos and R. Sowers, Recovery rates in investment-grade pools of credit assets: A large deviations analysis, *Stochastic Processes and their Applications*, Vol. 121, No. 12, (2011), pp. 2861–2898.
- [22] S.R.S. Varadhan, Asymptotic probabilities and differential equations. *Communications in Pure and Applied Mathematics*, Vol. 19, (1966), pp. 261–286.
- [23] X. Zhang, J. Blanchet, K. Giesecke and P. W. Glynn, Affine point processes: approximation and efficient simulation, *Mathematics of Operations Research*, (2014), to appear.