The Annals of Applied Probability
2013, Vol. 23, No. 1, 348–385
DOI: 10.1214/12-AAP845

DEFAULT CLUSTERING IN LARGE PORTFOLIOS: TYPICAL EVENTS

BY KAY GIESECKE, KONSTANTINOS SPILIOPOULOS AND RICHARD B. SOWERS

Stanford University, Brown University and University of Illinois at Urbana, Champaign

We develop a dynamic point process model of correlated default timing in a portfolio of firms, and analyze typical default profiles in the limit as the size of the pool grows. In our model, a firm defaults at a stochastic intensity that is influenced by an idiosyncratic risk process, a systematic risk process common to all firms, and past defaults. We prove a law of large numbers for the default rate in the pool, which describes the "typical" behavior of defaults.

1. Introduction. The financial crisis of 2007–09 has made clear the need to better understand the diversification of risk in financial systems with interacting entities. Prior to the crisis, the common belief was that risk had been diversified away by using the tools of structured finance. As it turned out, the correlation between assets was much larger than supposed. The collapse fed on itself and created a spiral.

We study the behavior of defaults in a large portfolio of interacting firms. We develop a dynamic point process model of correlated default timing, and then analyze typical default profiles in the limit as the number of constituent firms grows. Our empirically motivated model incorporates two distinct sources of default clustering. First, the firms are exposed to a risk factor process that is common to all entities in the pool. Variations in this systematic risk factor generate correlated movements in firms' conditional default probabilities. Das, Duffie, Kapadia and Saita [5] show that this mechanism is responsible for a large amount of corporate default clustering in the U.S. Second, a default has a contagious impact on the health of other firms. This impact fades away with time. Azizpour, Giesecke and Schwenkler [1] provide statistical evidence for the presence of such self-exciting effects in U.S. corporate defaults, after controlling for the exposure of firms to systematic risk factors.

More precisely, we assume that a firm defaults at an intensity, or conditional arrival rate, which follows a mean-reverting jump-diffusion process that is driven by several terms. The first term, a square root diffusion, represents an independent,

Received January 2011; revised January 2012.
MSC2010 subject classifications. 91G40, 60H10, 60G55, 60G57, 91680.
Key words and phrases. Interacting point processes, law of large numbers, portfolio credit risk, contagion.

firm-specific source of risk. The second term is a systematic risk factor that influences all firms, and that generates diffusive correlation between the intensities. For simplicity, we take this systematic risk factor to be an Ornstein–Uhlenbeck process. The third term affecting the intensity is the default rate in the pool. Defaults cause jumps in the intensity; they are common to all surviving firms. We thus have two sources of correlation between the firms: the dependence on the systematic risk factor and the influence of past defaults. While this formulation parsimoniously captures several of the sources of default correlation identified in empirical research, the intricate event dependence structure presents a challenge for the mathematical analysis of the system.

Our goal is to understand the behavior of the default rate in the portfolio in the limit as the number of firms in the pool grows. Large stochastic systems often tend to have macroscopic organization due to limit theorems such as the law of large numbers. This allows us to identify typical behavior. Our main result is a law of large numbers for the default rate in the pool; this describes the macroscopically typical profile. The limiting default rate satisfies an integral equation that makes explicit the role of the contagion exposure for the behavior of default clustering in the pool. The result depends heavily on the analysis of Markov processes via the martingale problem; see Ethier and Kurtz [11]. We will have more to say on the mathematical aspects of this in a moment. Once the typical behavior has been identified, one can then study Gaussian fluctuations and the structure of atypically large default clusters in the portfolio. We plan to pursue these directions in a future work.

Previous studies have analyzed the behavior of defaults in large portfolios. Dembo, Deuschel and Duffie [9] examine a doubly-stochastic model of default timing. In their model, default correlation is due to the exposure of firms to a common systematic risk factor which is represented by a random variable. Conditional on this variable, defaults are independent. A large deviation argument leads to an approximation of the tail of the conditional portfolio loss distribution. Glasserman, Kang and Shahabuddin [14] study a copula model of default timing using large deviation techniques. In that formulation, default events are conditionally independent given a set of common risk factors. Bush, Hambly, Haworth, Jin and Reisinger [2] prove a law of large numbers for a related dynamic model. Davis and Rodriguez [6] develop a law of large numbers and a central limit theorem for the default rate in a stochastic network setting, in which firms default independently of one another conditional on the realization of a systematic factor governed by a finite state Markov chain. Sircar and Zariphopoulou [20] examine large portfolio asymptotics for utility indifference valuation of securities exposed to the losses in the pool. As with these papers, our model includes exposure to a common systematic risk factor. In contrast, however, our model captures the self-exciting nature of defaults. Therefore, the firms in the pool are correlated even after conditioning on the path of the systematic factor process.

The use of interacting particle systems to study the behavior of default clustering in large portfolios is a growing area. In a mean-field model, Dai Pra, Runggaldier, Sartori and Tolotti [3] and Dai Pra and Tolotti [4] take the intensity of a constituent firm as a deterministic function of the percentage portfolio loss due to defaults. In a model with local interaction, Giesecke and Weber [13] take the intensity of a constituent firm as a deterministic function of the state of the firms in a specified neighborhood of that firm. The interacting particle perspective leads to the study of the convergence of interacting Markov processes, laws of large numbers for the percentage portfolio loss, and Gaussian approximations to the portfolio loss distribution based on central limit theorems. The interacting particle system which we propose and study incorporates an additional source of clustering, namely, the exposure of a firm to a systematic risk factor process. Moreover, firm-specific sources of default risk are present in our system. Also, the nature of mean-field interaction in our system is different. In [3] and [4], a constituent intensity is a function of the current default rate in the pool. In that formulation, the impact of a default on the dynamics of the surviving firms is permanent. In our work, a constituent intensity depends on the path of the default rate. The impact of a default on the surviving firms is transient, and fades away exponentially with time. There is a recovery effect.

As we were finishing this work, we learned of a related law of large numbers type result by Cvitanić, Ma and Zhang [16]. They take the intensity of a constituent firm as a function of a firm-specific risk factor, a systematic risk factor and the percentage portfolio loss due to defaults. The risk factors follow diffusion processes whose coefficients may depend on the portfolio loss. Our model of the risk factors is more specific than theirs, and thus we are able to arrive at slightly more explicit results. Moreover, the effect of defaults in [16] is permanent, as in [3] and [4].

There are several mathematical contributions in our efforts. Our analysis of typical events (a weak convergence result) is somewhat similar to that of certain genetic models (most notably the Fleming–Viot process; see Chapter 10 of [11], Fleming and Viot [12] and Dawson and Hochberg [7]), but the specific form of our intensity processes imply both complications and simplifications. Our work is centered on a jump-diffusion intensity process which is driven by Ornstein–Uhlenbeck and square root diffusion terms. This formulation allows some explicit simplifying calculations which are not available in a more abstract framework. On the other hand, due to the square root singularity, certain technical estimates need to be developed from scratch (see Section 10). A final point of interest is heterogeneity. Interacting particle systems are often assumed to have homogeneous dynamics, where various parameters are the same for each particle. This allows the main mathematical arguments to take their simplest form. Practitioners in credit risk, however, in reality face an extra problem in data aggregation, where each firm in a portfolio has its own statistical parameters. We have framed our weak convergence result to allow for a distribution of "types," that is, a frequency count of the different model parameters. This leads us to the correct effective dynamics of the

portfolio and, in particular, to a precise formulation of the effects of self-excitation (see Remark 5.2).

The rest of this paper is organized as follows. Section 2 formulates our model of default timing. We establish that our model is well-posed via the results of Section 3. In Section 4 we identify the limit as the number of firms in the portfolio goes to infinity—a law of large numbers result. The proof of this result is in Section 8, but depends upon the technical calculations of Sections 5, 6 and 7. Section 9 concludes and discusses extensions. Section 10 contains a number of technical results on square-root-like processes which are used in our calculations.

2. Model, assumptions and notation. We construct a point process model of correlated default timing in a portfolio of firms. We assume that $(\Omega, \mathscr{F}, \mathbb{P})$ is an underlying probability triple on which all random variables are defined. Let $\{W^n\}_{n\in\mathbb{N}}$ be a countable collection of standard Brownian motions. Let $\{\mathfrak{e}_n\}_{n\in\mathbb{N}}$ be an i.i.d. collection of standard exponential random variables. Finally, let V be a standard Brownian motion which is independent of the W^n's and $\mathfrak{e}_n$'s. Each W^n will represent a source of risk which is idiosyncratic to a specific firm. Each $\mathfrak{e}_n$ will represent a normalized default time for a specific firm. The process V will drive a systematic risk factor process to which all firms are exposed.

Fix an $N \in \mathbb{N}$, $n \in \{1, 2, \dots, N\}$ and consider the following system:

$$
\begin{aligned}
d\lambda_t^{N,n} &= -\alpha_{N,n}(\lambda_t^{N,n} - \bar{\lambda}_{N,n})\,dt + \sigma_{N,n}\sqrt{\lambda_t^{N,n}}\,dW_t^n \\
&\qquad + \beta_{N,n}^C\,dL_t^N + \varepsilon_N \beta_{N,n}^S \lambda_t^{N,n}\,dX_t, \qquad t > 0, \\
\lambda_0^{N,n} &= \lambda_{\circ,N,n}, \\
dX_t &= -\gamma X_t\,dt + dV_t, \qquad t > 0, \\
X_0 &= x_\circ, \\
L_t^N &= \frac{1}{N}\sum_{n=1}^N \chi_{[\mathfrak{e}_n,\infty)}\left(\int_{s=0}^t \lambda_s^{N,n}\,ds\right).
\end{aligned}
\tag{2.1}
$$

Here, $\beta_{N,n}^C \in \mathbb{R}_+ \stackrel{\text{def}}{=} [0,\infty)$ and $\beta_{N,n}^S \in \mathbb{R}$ are constants which represent the exposure of the nth firm in the pool to L^N and X, respectively. The $\alpha_{N,n}$'s, $\bar{\lambda}_{N,n}$'s and $\sigma_{N,n}$'s are in $\mathbb{R}_+$ and characterize the dynamics of the firms. We will address the role of ε_N in a moment. The initial condition $x_\circ$ of X is fixed and $\gamma > 0$. We use χ to represent the indicator function here and throughout the paper. The description of L^N is equivalent to a more standard construction. In particular, define

$$
\tau^{N,n} \stackrel{\text{def}}{=} \inf\left\{t \ge 0 : \int_{s=0}^t \lambda_s^{N,n}\,ds \ge \mathfrak{e}_n\right\}. \tag{2.2}
$$

Then

$$
\chi_{[\mathfrak{e}_n,\infty)}\left(\int_{s=0}^t \lambda_s^{N,n}\,ds\right) = \chi_{\{\tau^{N,n}\le t\}} \tag{2.3}
$$

and, consequently,

$$L_t^N = \frac{1}{N} \sum_{n=1}^{N} \chi_{\{\tau^{N,n} \le t\}}.$$

The process $\lambda^{N,n}$ represents the intensity, or conditional event rate, of the nth firm in a portfolio of N firms. More precisely, $\lambda^{N,n}$ is the instantaneous Doob–Meyer compensator to the default indicator process (2.3); see (4.1). We will see in Proposition 3.3 in Section 3 that the $\lambda^{N,n}$'s are indeed nonnegative. The process X represents a source of systematic risk; in our model this is a stable Ornstein–Uhlenbeck process. The process L^N is the default rate in the pool. The jump-diffusion model for $\lambda^{N,n}$ captures several sources of default clustering. A firm's intensity is driven by an idiosyncratic source of risk represented by a Brownian motion W^n, and a source of systematic risk common to all firms—the process X. Movements in X cause correlated changes in firms' intensities and thus provide a source of default clustering emphasized by [5] for corporate defaults in the U.S. The sensitivity of $\lambda^{N,n}$ to changes in X is measured by the parameter $\beta^S_{N,n}$. The second source of default clustering is through the feedback ("contagion") term $\beta^C_{N,n}\, dL_t^N$. A default causes an upward jump of size $\frac{1}{N}\beta^C_{N,n}$ in the intensity $\lambda^{N,n}$. Due to the mean-reversion of $\lambda^{N,n}$, the impact of a default fades away with time, exponentially with rate $\alpha_{N,n}$. Self-exciting effects of this type have been found to be an important source of the clustering of defaults in the U.S., over and above any clustering caused by the exposure of firms to systematic risk factors [1].

In the special case that $\beta^C_{N,n} = \beta^S_{N,n} = 0$ for all $n \in \{1, 2, \dots, N\}$, the intensities $\lambda^{N,n}$ follow independent square root processes so firms default independently of one another. The formulation (2.1) is a natural generalization of the widely used square root model to address the clustering between defaults.

The interest in large pools of assets is that they provide *diversification*; they allow one to construct portfolios which have small variance. The dynamics of X imply that X is stochastically of order 1, that is, it is stable.[1] Thus, the only way for the pool to have small variance in our model is for each of the constituent firms to have small exposure to X. We thus assume that

$$\lim_{N\to\infty} \varepsilon_N = 0.$$

If ε_N is not small, the influence of the systematic risk factor X will be of order 1, and the "typical" behavior of the pool will strongly depend on X (and the tail behavior of the whole system will be strongly determined by the tail of X).

[1]Regulatory agencies, for example, are charged with preventing systematic factors from spiraling out of control.

REMARK 2.1. Given the simple structure of X, our model is equivalent, if $x_\circ = 0$, to a model where each intensity has exposure of order 1 to a small systematic risk. Namely, if $x_\circ = 0$, then $\varepsilon_N X = \tilde{X}^N$ where

$$d\tilde{X}_t^N = -\gamma \tilde{X}_t^N \, dt + \varepsilon_N \, dV_t.$$

Our model allows for a significant amount of bottom-up heterogeneity; the intensity dynamics of each firm can be different. We capture these different dynamics by defining the "types"

$$\mathsf{p}^{N,n} \stackrel{\text{def}}{=} (\alpha_{N,n}, \bar{\lambda}_{N,n}, \sigma_{N,n}, \beta_{N,n}^C, \beta_{N,n}^S); \tag{2.4}$$

the $\mathsf{p}^{N,n}$'s take values in parameter space $\mathcal{P} \stackrel{\text{def}}{=} \mathbb{R}_+^4 \times \mathbb{R}$. In order to expect regular macroscopic behavior of L^N as $N \to \infty$, the $\mathsf{p}^{N,n}$'s and the $\lambda_{\circ,N,n}$'s should have enough regularity as $N \to \infty$. For each $N \in \mathbb{N}$, define

$$\pi^N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \delta_{\mathsf{p}^{N,n}} \quad \text{and} \quad \Lambda_\circ^N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \delta_{\lambda_{\circ,N,n}};$$

these are elements of $\mathscr{P}(\mathcal{P})$ and $\mathscr{P}(\mathbb{R}_+)$, respectively.[2]

We need two main assumptions. First, we assume that the types of (2.4) and the initial distributions (the $\lambda_{\circ,N,n}$'s) are sufficiently regular.

ASSUMPTION 2.2. We assume that

$$\pi \stackrel{\text{def}}{=} \lim_{N\to\infty} \pi^N \quad \text{and} \quad \Lambda_\circ \stackrel{\text{def}}{=} \lim_{N\to\infty} \Lambda_\circ^N$$

exist [in $\mathscr{P}(\mathcal{P})$ and $\mathscr{P}(\mathbb{R}_+)$, resp.].

Note that this is what happens in practice; one constructs a frequency count of the parameters of the different assets in a large pool and uses this to seek aggregate dynamics for the pool itself. For a large pool, one hopes that this frequency count will have some simpler macroscopic description. Second, we assume that the types are bounded.

ASSUMPTION 2.3. We assume that there is a $K_{2.3} > 0$ such that the $\alpha_{N,n}$'s, $\bar{\lambda}_{N,n}$'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 $K_{2.3}$ for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$.

Equivalently, we require that the π_N's and $\Lambda_\circ^N$'s all (uniformly in N) have compact support. We could relax this requirement, at the cost of a much more careful error analysis.

We are interested in the *typical behavior* of $\{L^N\}$. In Section 3 we consider the well-posedness of the model (2.1), while in Section 4 we state the law of large numbers result, Theorem 4.2.

[2]As usual, if E is a topological space, $\mathscr{P}(E)$ is the collection of Borel probability measures on E.

3. Well-posedness of the model. We here state several technical results concerning the intensities which are a central part of our model. We want to understand the structure of the $\lambda^{N,n}$'s a bit more. The complications which require our attention are the square root singularity, and the fact that the $\lambda_t\,dX_t$ term contains the term $\lambda_t X_t\,dt$, implying that the dynamics of the $\mathbb{R}^2$-valued process (λ, X) contain a superlinear drift. The proofs of the results here will be given in Section 10.

Let W^* be a reference Brownian motion with respect to a filtration $\{\mathscr{G}_t\}_{t\geq 0}$. Assume also that V is adapted to $\{\mathscr{G}_t\}_{t\geq 0}$. Let ξ be a $\{\mathscr{G}_t\}_{t\geq 0}$-adapted, point process which takes values in $[0, 1]$ and such that $\xi_0 = 0$. Fix $\mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S) \in \mathcal{P}$ and $\lambda_\circ$ in $\mathbb{R}_+$. Consider the SDE

$$\begin{aligned} d\lambda_t &= -\alpha(\lambda_t - \bar{\lambda})\,dt + \sigma\sqrt{\lambda_t \vee 0}\,dW_t^* + \beta^C\,d\xi_t + \beta^S\lambda_t\,dX_t, \qquad t > 0, \\ \lambda_0 &= \lambda_\circ. \end{aligned} \tag{3.1}$$

Note that by expanding the dynamics of dX and rearranging a bit, we get that

$$d\lambda_t = -\{\alpha + \beta^S\gamma X_t\}\lambda_t\,dt + \alpha\bar{\lambda}\,dt + \beta^C\,d\xi_t + \sigma\sqrt{\lambda_t \vee 0}\,dW_t^* + \beta^S\lambda_t\,dV_t.$$

Also, we have for the moment subsumed the small parameter ε_N into the β^S term; see the proof of Proposition 3.3.

We will use a number of ideas from [15] (see also [8]).

LEMMA 3.1. *There is a nonnegative solution λ of the $\mathbb{R}$-valued SDE* (3.1). *Furthermore*, $\sup_{t\in[0,T]}\mathbb{E}[|\lambda_t|^p] < \infty$ *for all $T > 0$ and $p \geq 1$.*

We also have uniqueness.

LEMMA 3.2. *The solution of* (3.1) *is unique.*

The model (2.1) is thus well posed.

PROPOSITION 3.3. *The system* (2.1) *has a unique solution such that $\lambda_t^{N,n} \geq 0$ for every $N \in \mathbb{N}$, $n \in \{1, 2, \dots, N\}$ and $t \geq 0$.*

PROOF. Using Lemmas 3.1 and 3.2, solve (2.1) between the default times. Replace β^S by $\varepsilon_N\beta^S$ in applying Lemma 3.1. $\square$

We shall also need a macroscopic bound on the intensities.

LEMMA 3.4. *For each $p \geq 1$ and $T \geq 0$,*

$$K_{p,T,3.4} \stackrel{def}{=} \sup_{\substack{0\leq t\leq T \\ N\in\mathbb{N}}} \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}[|\lambda_t^{N,n}|^p]$$

is finite.

4. Typical events: A law of large numbers. Our first task is to understand the "typical" behavior of our system. To do so, we need to understand a system which contains a bit more information than the default rate L^N. For each $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$, define

$$\mathrm{M}_t^{N,n} \stackrel{\text{def}}{=} \chi_{[0,\mathfrak{e}_n)}\left(\int_{s=0}^t \lambda_s^{N,n}\, ds\right) = \chi_{\{\tau^{N,n} > t\}} \tag{4.1}$$

[where $\tau^{N,n}$ is as in (2.2)]. In other words, $\mathrm{M}_t^{N,n} = 1$ if and only if the nth firm is still alive at time t; otherwise $\mathrm{M}_t^{N,n} = 0$. Thus, $\mathrm{M}^{N,n}$ is nonincreasing and right-continuous. It is easy to see that

$$\mathrm{M}_t^{N,n} + \int_{s=0}^t \lambda_s^{N,n} \mathrm{M}_s^{N,n}\, ds$$

is a martingale. Define $\hat{\mathcal{P}} \stackrel{\text{def}}{=} \mathcal{P} \times \mathbb{R}_+$. For each $N \in \mathbb{N}$, define $\hat{\mathsf{p}}_t^{N,n} \stackrel{\text{def}}{=} (\mathsf{p}^{N,n}, \lambda_t^{N,n})$ for all $n \in \{1, 2, \dots, N\}$ and $t \geq 1$. For each $t \geq 0$, define

$$\mu_t^N \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^N \delta_{\hat{\mathsf{p}}_t^{N,n}} \mathrm{M}_t^{N,n};$$

in other words, we keep track of the empirical distribution of the type and credit spread for those assets which are still "alive." We note that

$$L_t^N = 1 - \mu_t^N(\hat{\mathcal{P}})$$

for all $t \geq 0$.

We want to understand the dynamics of μ_t^N for large N (this will then imply the "typical" behavior for L_t^N). To understand what our main result is, let's first set up a topological framework to understand convergence of μ^N. Let E be the collection of sub-probability measures (i.e., defective probability measures) on $\hat{\mathcal{P}}$, that is, E consists of those Borel measures ν on $\hat{\mathcal{P}}$ such that $\nu(\hat{\mathcal{P}}) \leq 1$. We can topologize E in the usual way (by projecting onto the one-point compactification of $\hat{\mathcal{P}}$; see [19], Chapter 9.5). In particular, fix a point $\star$ that is not in $\hat{\mathcal{P}}$ and define $\hat{\mathcal{P}}^+ \stackrel{\text{def}}{=} \hat{\mathcal{P}} \cup \{\star\}$. Give $\hat{\mathcal{P}}^+$ the standard topology; open sets are those which are open subsets of $\hat{\mathcal{P}}$ (with its original topology) or complements in $\hat{\mathcal{P}}^+$ of closed subsets of $\hat{\mathcal{P}}$ (again, in the original topology of $\hat{\mathcal{P}}$). Define a bijection ι from E to $\mathscr{P}(\hat{\mathcal{P}}^+)$ (the collection of Borel probability measures on $\hat{\mathcal{P}}^+$) by setting

$$(\iota\nu)(A) \stackrel{\text{def}}{=} \nu(A \cap \hat{\mathcal{P}}) + \big(1 - \nu(\hat{\mathcal{P}})\big)\delta_\star(A)$$

for all $A \in \mathscr{B}(\hat{\mathcal{P}}^+)$. We can define the Skorohod topology on $\mathscr{P}(\hat{\mathcal{P}}^+)$, and define a corresponding metric on E by requiring ι to be an isometry. This makes E a Polish space. We thus have that μ^N is an element[3] of $D_E[0, \infty)$.

[3]If S is a Polish space, then $D_S[0, \infty)$ is the collection of maps from $[0, \infty)$ into S which are right-continuous and which have left-hand limits. The space $D_S[0, \infty)$ can be topologized by the Skorohod metric, which we will denote by d_S; see Chapter 3.5 of [11].

The main theorem of this section is Theorem 4.2, essentially a law of large numbers. The construction of the limiting process will take several steps. First, for each $\mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S) \in \mathcal{P}$, let b^{p} satisfy

$$\begin{aligned} &\dot{b}^{\mathsf{p}}(t) = 1 - \tfrac{1}{2}\sigma^2 (b^{\mathsf{p}}(t))^2 - \alpha b^{\mathsf{p}}(t), \qquad t > 0, \\ &b^{\mathsf{p}}(0) = 0. \end{aligned} \tag{4.2}$$

Note that if $b^{\mathsf{p}}(t) = 0$, then $\dot{b}^{\mathsf{p}}(t) = 1 > 0$. Thus, $b^{\mathsf{p}}(t) > 0$ for all $t > 0$.

The next lemma is essential for the characterization of the limit. Its proof is deferred to Section 10.

LEMMA 4.1. *There is a unique $\mathbb{R}_+$-valued trajectory $\{Q(t); t \geq 0\}$ which satisfies the equation*

$$\begin{aligned} Q(t) = \int_{\substack{\hat{\mathsf{p}} = (\mathsf{p}, \lambda) \in \hat{\mathcal{P}} \\ \mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S)}} & \beta^C \Big[\dot{b}^{\mathsf{p}}(t)\lambda + \int_{r=0}^{t} \dot{b}^{\mathsf{p}}(t-r)\{Q(r) + \alpha\bar{\lambda}\}\, dr \Big] \\ & \times \exp\Big[-b^{\mathsf{p}}(t)\lambda - \int_{r=0}^{t} b^{\mathsf{p}}(t-r)\{Q(r) + \alpha\bar{\lambda}\}\, dr \Big] \\ & \times \pi(d\mathsf{p})\Lambda_\circ(d\lambda). \end{aligned} \tag{4.3}$$

Here, π and $\Lambda_\circ$ are as in Assumption 2.2.

Now let W^* be a reference Brownian motion. For each $\hat{\mathsf{p}} = (\mathsf{p}, \lambda_\circ) \in \hat{\mathcal{P}}$ where $\mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S)$, let $\lambda_t^*(\mathsf{p})$ be the unique solution to

$$\begin{aligned} \lambda_t^*(\hat{\mathsf{p}}) = \lambda_\circ - \alpha \int_{s=0}^{t} (\lambda_s^*(\hat{\mathsf{p}}) - \bar{\lambda})\, ds + \sigma \int_{s=0}^{t} \sqrt{\lambda_s^*(\hat{\mathsf{p}})}\, dW_s^* \\ + \int_{s=0}^{t} Q(s)\, ds. \end{aligned} \tag{4.4}$$

We now have our main result.

THEOREM 4.2. *For all $A \in \mathscr{B}(\mathcal{P})$ and $B \in \mathscr{B}(\mathbb{R}_+)$, define*

$$\begin{aligned} \bar{\mu}_t(A \times B) \stackrel{def}{=} \int_{\hat{\mathsf{p}} = (\mathsf{p}, \lambda) \in \hat{\mathcal{P}}} & \chi_A(\mathsf{p}) \mathbb{E}\Big[\chi_B(\lambda_t^*(\hat{\mathsf{p}})) \exp\Big[-\int_{s=0}^{t} \lambda_s^*(\hat{\mathsf{p}})\, ds \Big] \Big] \\ & \times \pi(d\mathsf{p})\Lambda_\circ(d\lambda). \end{aligned} \tag{4.5}$$

Then

$$\lim_{N \to \infty} \mathbb{P}\{ d_{\mathscr{P}(\hat{\mathcal{P}})}(\mu^N, \bar{\mu}) \geq \delta \} = 0 \tag{4.6}$$

for every $\delta > 0$. *Define*

$$
\begin{aligned}
F(t) &\stackrel{def}{=} 1 - \bar{\mu}_t(\hat{\mathcal{P}}) \\
&= 1 - \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}} \mathbb{E}\Big[\exp\Big[-\int_{s=0}^{t} \lambda_s^*(\hat{\mathsf{p}})\,ds\Big]\Big]\pi(d\mathsf{p})\Lambda_\circ(d\lambda).
\end{aligned} \tag{4.7}
$$

Then, for all $\delta > 0$ *and* $T > 0$,

$$
\lim_{N\to\infty} \mathbb{P}\Big\{\sup_{0\le t\le T} |L_t^N - F(t)| \ge \delta\Big\} = 0.
$$

The parts of the proof of this result will be given in Sections 5, 6 and 7. In particular, in Section 5 we identify a candidate limit for $\{\mu^N\}$ using the martingale problem formulation. Then in Section 6 we prove that $\{\mu^N\}$ is tight, which ensures that the laws of $\{\mu^N\}$'s have at least one limit point. In Section 7 we prove that the limit is necessarily unique. Then, in Section 8 we collect things together to prove Theorem 4.2.

With this result in hand, we can rewrite (4.4) to see the exposure of a typical firm to the contagion factor.

REMARK 4.3. We have that

$$
\begin{aligned}
\dot{F}(t) &= \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}} \mathbb{E}\Big[\lambda_t^*(\hat{\mathsf{p}})\exp\Big[-\int_{s=0}^{t} \lambda_s^*(\hat{\mathsf{p}})\,ds\Big]\Big]\pi(d\mathsf{p})\Lambda_\circ(d\lambda) \\
&= \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}} \lambda\bar{\mu}_t(d\hat{\mathsf{p}}).
\end{aligned} \tag{4.8}
$$

Thus,

$$
\begin{aligned}
\lambda_t^*(\hat{\mathsf{p}}) &= \lambda_\circ - \alpha\int_{s=0}^{t} \big(\lambda_s^*(\hat{\mathsf{p}}) - \bar{\lambda}\big)\,ds + \sigma\int_0^t \sqrt{\lambda_s^*(\hat{\mathsf{p}})}\,dW_s^* \\
&\quad + \int_0^t B(\bar{\mu}_s)\dot{F}(s)\,ds,
\end{aligned} \tag{4.9}
$$

where

$$
B(\mu) \stackrel{\text{def}}{=} \int_{\substack{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}} \\ \mathsf{p}=(\alpha,\bar{\lambda},\sigma,\beta^C,\beta^S)}} \beta^C\lambda\mu(d\hat{\mathsf{p}}) \Big/ \int_{\substack{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}} \\ \mathsf{p}=(\alpha,\bar{\lambda},\sigma,\beta^C,\beta^S)}} \lambda\mu(d\hat{\mathsf{p}})
$$

for all $\mu \in E$. In other words, the effective sensitivity of a typical intensity to the contagion is given by an average weighted by the instantaneous intensities. Note that $0 \le B(\mu) \le K_{2.3}$.

The homogeneous case provides more explicit insights into the role of the contagion exposure for the behavior of default clustering in the pool.

REMARK 4.4. Fix $\hat{\mathsf{p}} = (\mathsf{p}, \lambda_\circ) \in \hat{\mathcal{P}}$ where $\mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S)$. Assume that the pool is homogeneous, that is, $\hat{\mathsf{p}}^{N,n} = \hat{\mathsf{p}}$ for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$. By

the relation (4.9), we then have that $Q(t) = \beta^C \dot{F}(t)$. In this case, F is given by the unique solution to the integral equation

$$F(t) = 1 - \exp\left[-\alpha\bar{\lambda}\int_{r=0}^{t} b^{\mathsf{p}}(t-r)\,dr - \beta^C \int_{r=0}^{t} F(r)\dot{b}^{\mathsf{p}}(t-r)\,dr - b^{\mathsf{p}}(t)\lambda_\circ\right].$$

Furthermore, if there is no exposure to contagion, that is, $\beta^C = 0$, then this integral equation reduces to the well-known explicit formula

$$F(t) = 1 - \exp\left[-\alpha\bar{\lambda}\int_{r=0}^{t} b^{\mathsf{p}}(t-r)\,dr - b^{\mathsf{p}}(t)\lambda_\circ\right].$$

Figure 1 shows the limiting default rate $F(t)$ for different values of the contagion sensitivity β^C. The default rate increases with β^C. Figure 2 shows the limiting default rate $F(t)$ for different values of the parameter α, which specifies the reversion speed of the intensity. The default rate is relatively insensitive to changes in α for shorter horizons; for longer horizons it decreases with α. The limiting default rate is more sensitive to variation in the reversion level $\bar{\lambda}$, as indicated in Figure 3. Variations in the diffusive volatility σ of the intensity have little effect on $F(t)$.

We finally note that the structure of the unperturbed (i.e., $\beta^C = \beta^S = 0$) dynamics of the intensity (2.1) was crucial in singling out the equation (4.3) as the

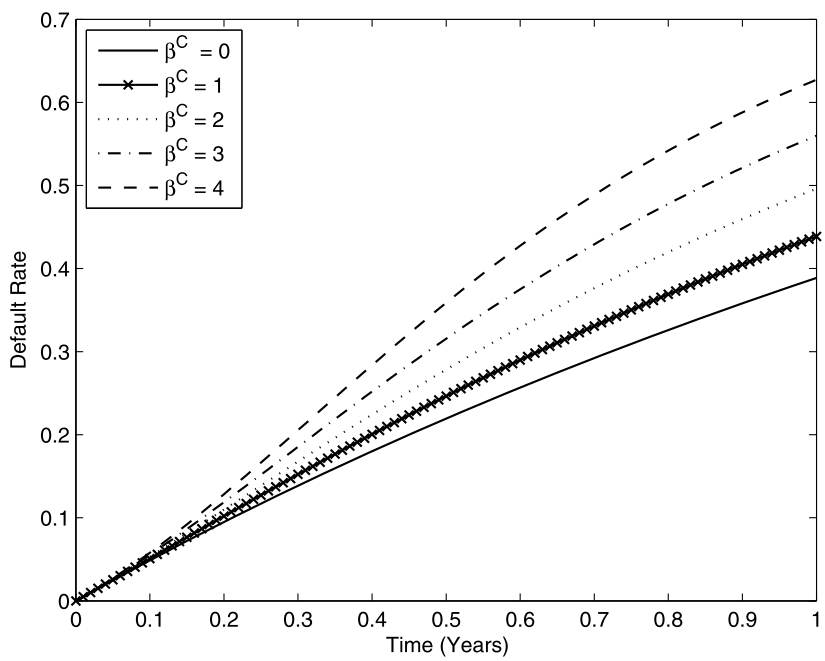

FIG. 1. *Comparison of limiting default rate $F(t)$ for different values of the contagion sensitivity β^C. The parameter case is $\sigma = 0.9$, $\alpha = 4$, $\bar{\lambda} = 0.5$ and $\lambda_0 = 0.5$.*

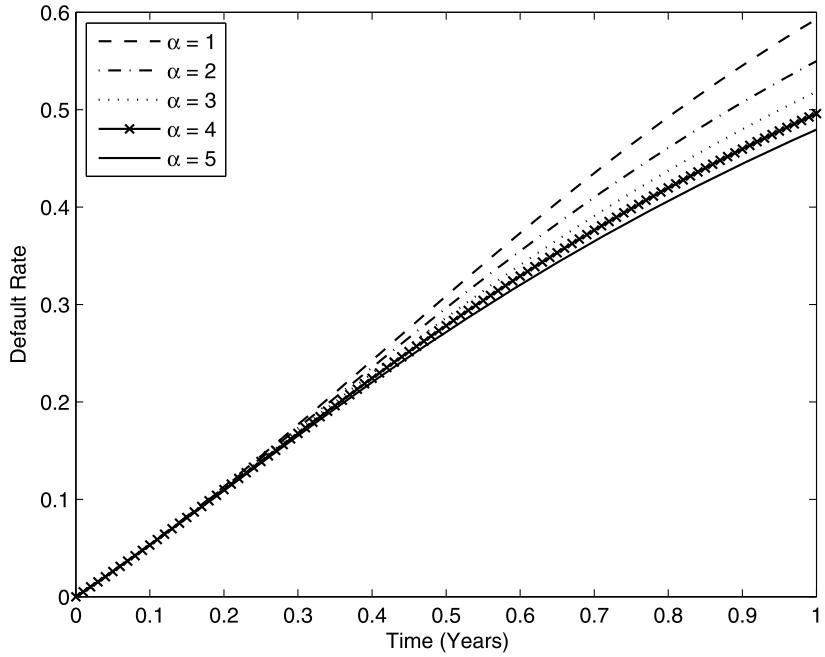

FIG. 2. *Comparison of limiting default rate $F(t)$ for different values of the reversion speed α. The parameter case is $\sigma = 0.9$, $\beta^C = 2$, $\bar{\lambda} = 0.5$ and $\lambda_0 = 0.5$.*

proper macroscopic effect of the contagion (see the proof of Lemma 8.2). The calculations in fact hinge upon the explicit formulae for affine jump diffusions developed in [10]. In a more general setting we would need a more abstract framework (see [16]).

5. Identification of the limit. We want to use the martingale problem (see Chapter 4 of [11]) to show that μ^N's converge to a limiting process. For every $f \in C^\infty(\hat{\mathcal{P}})$ and $\mu \in E$, define

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

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

$$\Phi(\mu) = \varphi(\langle f_1, \mu \rangle_E, \langle f_2, \mu \rangle_E, \dots, \langle f_M, \mu \rangle_E) \tag{5.1}$$

for some $M \in \mathbb{N}$, some $\varphi \in C^\infty(\mathbb{R}^M)$ and some $\{f_m\}_{m=1}^M$. Then $\mathcal{S}$ separates $\mathscr{P}(\hat{\mathcal{P}})$ (see Chapter 3.4 of [11]). It thus suffices to show convergence of the martingale problem for functions of the form (5.1).

Let's fix $f \in C^\infty(\hat{\mathcal{P}})$ and understand exactly what happens to $\langle f, \mu^N \rangle_E$ when one of the firms defaults. Suppose that the nth firm defaults at time t and that none

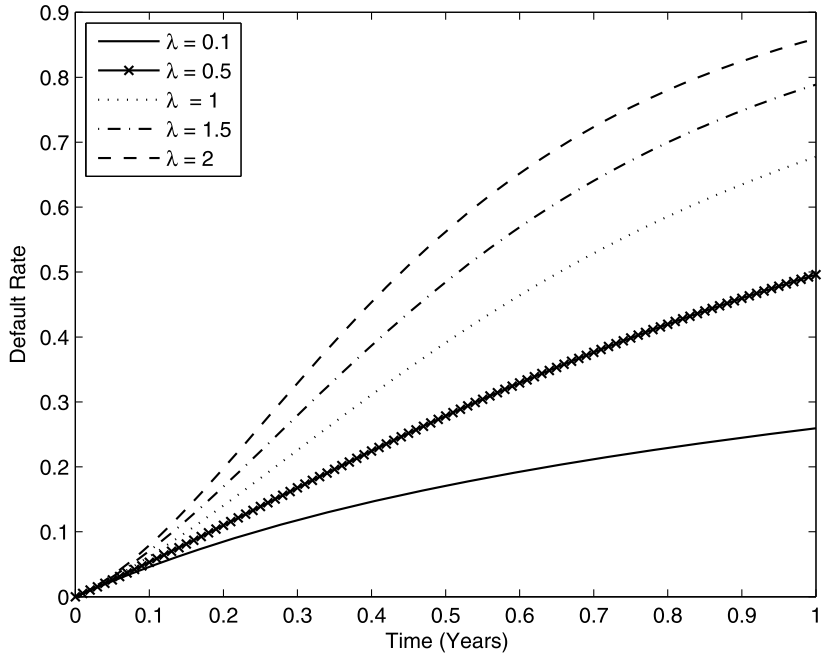

FIG. 3. *Comparison of limiting default rate $F(t)$ for different values of the reversion level $\bar{\lambda}$. The parameter case is $\sigma = 0.9$, $\beta^C = 2$, $\alpha = 4$ and $\lambda_0 = 0.5$.*

of the other firms default at time t (defaults occur simultaneously with probability zero). Then

$$\langle f, \mu_t^N \rangle_E = \frac{1}{N} \sum_{\substack{1 \le n' \le N \\ n' \ne n}} f\left(\mathsf{p}^{N,n'}, \lambda_t^{N,n'} + \frac{\beta_{N,n}^C}{N}\right) \mathrm{M}_t^{N,n'},$$

$$\langle f, \mu_{t-}^N \rangle_E = \frac{1}{N} f(\mathsf{p}^{N,n'}, \lambda_t^{N,n'}) \mathrm{M}_t^{N,n'} + \frac{1}{N} f(\mathsf{p}^{N,n}, \lambda_t^{N,n}).$$

Note, furthermore, that the default at time t means that $\int_{s=0}^t \lambda_s^{N,n}\, ds = \mathfrak{e}_n$, so $\mathrm{M}_t^{N,n} = 0$. Hence,

$$\langle f, \mu_t^N \rangle_E - \langle f, \mu_{t-}^N \rangle_E = \mathcal{J}_{N,n}^f(t), \tag{5.2}$$

where

$$\mathcal{J}_{N,n}^f(t) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n'=1}^{N} \left\{ f\left(\mathsf{p}^{N,n'}, \lambda_t^{N,n'} + \frac{\beta_{N,n}^C}{N}\right) - f(\mathsf{p}^{N,n'}, \lambda_t^{N,n'}) \right\} \mathrm{M}_t^{N,n'}$$
$$- \frac{1}{N} f(\mathsf{p}^{N,n}, \lambda_t^{N,n})$$

for all $t \geq 0$, $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$.

We now identify the limiting martingale problem for μ^N. For $\hat{\mathsf{p}} = (\mathsf{p}, \lambda)$ where $\mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S) \in \mathcal{P}$ and $f \in C^\infty(\hat{\mathcal{P}})$, define the operators

$$
\begin{aligned}
(\mathcal{L}_1 f)(\hat{\mathsf{p}}) &= \frac{1}{2}\sigma^2 \lambda \frac{\partial^2 f}{\partial \lambda^2}(\hat{\mathsf{p}}) - \alpha(\lambda - \bar{\lambda}) \frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}) - \lambda f(\hat{\mathsf{p}}), \\
(\mathcal{L}_2 f)(\hat{\mathsf{p}}) &= \frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}).
\end{aligned} \tag{5.3}
$$

Define also

$$
\mathcal{Q}(\hat{\mathsf{p}}) \stackrel{\text{def}}{=} \lambda \beta^C
$$

for $\hat{\mathsf{p}} = (\mathsf{p}, \lambda)$ where $\mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S) \in \mathcal{P}$. 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 $\Phi \in \mathcal{S}$ of the form (5.1), define

$$
\begin{aligned}
(\mathcal{A}\Phi)(\mu) \stackrel{\text{def}}{=} \sum_{m=1}^{M} \frac{\partial \varphi}{\partial x_m}(\langle f_1, \mu\rangle_E, \langle f_2, \mu\rangle_E, \dots, \langle f_M, \mu\rangle_E) \\
\times \{\langle \mathcal{L}_1 f_m, \mu\rangle_E + \langle \mathcal{Q}, \mu\rangle_E \langle \mathcal{L}_2 f_m, \mu\rangle_E\}.
\end{aligned} \tag{5.4}
$$

We claim that $\mathcal{A}$ will be the generator of the limiting martingale problem.

LEMMA 5.1 (Weak convergence). *For any $\Phi \in \mathcal{S}$ and $0 \leq r_1 \leq r_2 \cdots r_J = s < t < T$ and $\{\psi_j\}_{j=1}^J \subset B(E)$, we have that*

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

PROOF. For $\hat{\mathsf{p}} = (\mathsf{p}, \lambda)$ where $\mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S) \in \mathcal{P}$, define

$$
\begin{aligned}
(\mathcal{L}^a f)(\hat{\mathsf{p}}) &= \frac{1}{2}\sigma^2 \lambda \frac{\partial^2 f}{\partial \lambda^2}(\hat{\mathsf{p}}) - \alpha(\lambda - \bar{\lambda}) \frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}), \\
(\mathcal{L}_x^b f)(\hat{\mathsf{p}}) &= \lambda\left\{\frac{1}{2} \frac{\partial^2 f}{\partial \lambda^2}(\hat{\mathsf{p}}) - \gamma x \frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}})\right\}, \qquad x \in \mathbb{R}.
\end{aligned}
$$

Then $\mathcal{L}^a$ is the generator of the idiosyncratic part of the intensity and $\mathcal{L}_x^b$ is the generator of the systematic risk.

We start by writing that

$$\Phi(\mu_t^N) = \Phi(\mu_0^N) + \int_{r=0}^t \{A_r^{N,1} + A_r^{N,2}\}\, dr + \mathcal{M}_t,$$

where $\mathcal{M}$ is a martingale and

$$\begin{aligned}
A_t^{N,1} &= \sum_{m=1}^M \frac{\partial \varphi}{\partial x_m}(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E, \dots, \langle f_M, \mu_t^N\rangle_E) \\
&\qquad \times \frac{1}{N}\sum_{n=1}^N \{(\mathcal{L}^a f_m)(\hat{\mathsf{p}}_t^{N,n}) + \varepsilon_N(\mathcal{L}_{X_t}^b f_m)(\hat{\mathsf{p}}_t^{N,n})\}\mathrm{M}_t^{N,n} \\
&= \sum_{m=1}^M \frac{\partial \varphi}{\partial x_m}(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E, \dots, \langle f_M, \mu_t^N\rangle_E) \\
&\qquad \times \{\langle \mathcal{L}^a f_m, \mu_t^N\rangle_E + \varepsilon_N \langle \mathcal{L}_{X_t}^b f_m, \mu_t^N\rangle_E\}, \\
A_t^{N,2} &= \sum_{n=1}^N \lambda_t^{N,n}\big\{\varphi\big(\langle f_1, \mu_t^N\rangle_E + \mathcal{J}_{N,n}^{f_1}(t), \\
&\qquad \langle f_2, \mu_t^N\rangle_E + \mathcal{J}_{N,n}^{f_2}(t), \dots, \langle f_M, \mu_t^N\rangle_E + \mathcal{J}_{N,n}^{f_M}(t)\big) \\
&\qquad - \varphi(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E, \dots, \langle f_M, \mu_t^N\rangle_E)\big\}\mathrm{M}_t^{N,n}.
\end{aligned}$$

Using Lemma 3.4, it is fairly easy to see that for all $f \in C^\infty(\hat{\mathcal{P}})$,

$$\lim_{N\to\infty} \mathbb{E}\Big[\varepsilon_N \int_{r=0}^t |\langle \mathcal{L}_{X_r}^b f, \mu_r^N\rangle_E|\, dr\Big] = 0.$$

To proceed, let's simplify $\mathcal{J}_{N,n}^f$. For each $f \in C^\infty(\hat{\mathcal{P}})$, $t \geq 0$, $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$, define

$$\begin{aligned}
\tilde{\mathcal{J}}_{N,n}^f(t) &\stackrel{\text{def}}{=} \frac{\beta_{N,n}^C}{N}\sum_{m=1}^N \frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}_t^{N,n})\mathrm{M}_t^{N,m} - f(\mathsf{p}^{N,n}, \lambda_t^{N,n}) \\
&= \beta_{N,n}^C \langle \mathcal{L}_2 f, \mu_t^N\rangle_E - f(\hat{\mathsf{p}}_t^N).
\end{aligned}$$

Then

$$\left|\mathcal{J}_{N,n}^f(t) - \frac{1}{N}\tilde{\mathcal{J}}_{N,n}^f(t)\right| \leq \frac{K_{2.3}^2}{N^2}\left\|\frac{\partial^2 f}{\partial \lambda^2}\right\|_C,$$

where $K_{2.3}$ is the constant from Assumption 2.3.

Define $\iota(\hat{\mathsf{p}}) \stackrel{\text{def}}{=} \lambda$ for $\hat{\mathsf{p}} = (\mathsf{p}, \lambda) \in \hat{\mathcal{P}}$. Setting

$$\begin{aligned}\tilde{A}_t^{N,2} &\stackrel{\text{def}}{=} \sum_{m=1}^{M} \frac{\partial \varphi}{\partial x_m}(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E, \dots, \langle f_M, \mu_t^N\rangle_E) \\ &\qquad \times \frac{1}{N}\sum_{n=1}^{N} \lambda_t^{N,n} \tilde{\mathcal{J}}_{N,n}^{f_m}(t) \mathrm{M}_t^{N,n} \\ &= \sum_{m=1}^{M} \frac{\partial \varphi}{\partial x_m}(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E, \dots, \langle f_M, \mu_t^N\rangle_E) \\ &\qquad \times \{\langle \mathcal{Q}, \mu_t^N\rangle_E \langle \mathcal{L}_2 f_m, \mu_t^N\rangle_E - \langle \iota f, \mu_t^N\rangle_E\},\end{aligned} \tag{5.5}$$

we have that

$$\lim_{N\to\infty} \mathbb{E}\left[\int_{r=0}^{t} |A_r^{N,2} - \tilde{A}_r^{N,2}|\, dr\right] = 0.$$

Collecting things together, we have that

$$\lim_{N\to\infty} \mathbb{E}\left[\left\{\int_{r=s}^{t} A_r^{N,1}\, dr + \int_{r=s}^{t} A_r^{N,2}\, dr - \int_{r=s}^{t} (\mathcal{A}\Phi)(\mu_r^N)\, dr\right\} \prod_{j=1}^{J} \psi_j(\mu_{r_j}^N)\right] = 0,$$

which implies the claim. □

We, in particular, note the macroscopic effect of the contagion.

REMARK 5.2. The key step in quantifying the coarse-grained effect of contagion was (5.5). Namely, we average the combination of the jump rate and the exposure to contagion across the pool.

6. Tightness. In this subsection we verify that the family $\{\mu^N\}_{N\in\mathbb{N}}$ is relatively compact (as a $D_E[0,\infty)$-valued random variable); this of course is necessary to ensure that the laws of the μ^N's have at least one limit point. The complication of course is the feedback through contagion. We need to show that the system is unlikely to "explode" via feedback. Our calculations are framed by Theorem 8.6 of Chapter 3 of [11]; we need to show compact containment and regularity of the μ^N's.

In particular, compact containment ensures that there is a compact set $\mathcal{K}$ such that μ_t^N will belong to $\mathcal{K}$ for all $N \in \mathbb{N}$ and $t \in [0, T]$ with high probability; see Lemma 6.1. Regularity shows, roughly speaking, that $\mu_t^N - \mu_s^N$ is bounded in a certain sense by a function of the time interval $t - s$, that goes to zero as the length of the time interval goes to zero; see Lemma 6.3. By Theorem 8.6 of Chapter 3 of [11], these two statements imply relative compactness of the family $\{\mu^N\}_{N\in\mathbb{N}}$ in $D_E[0,\infty)$; see Lemma 6.4.

Let's first address compact containment.

LEMMA 6.1. *For each $\eta > 0$ and $t \geq 0$, there is a compact subset $\mathcal{K}$ of E such that*

$$\sup_{\substack{N\in\mathbb{N}\\ 0\leq t<T}} \mathbb{P}\{\mu_t^N \notin \mathcal{K}\} < \eta.$$

PROOF. For each $L > 0$, define $K_L \stackrel{\text{def}}{=} [-K_{2.3}, K_{2.3}]^3 \times [0, K_{2.3}]^2 \times [0, L]$. Then $K_L \subset\subset \hat{\mathcal{P}}$, and for each $t \geq 0$ and $N \in \mathbb{N}$,

$$\mathbb{E}[\mu_t^N(\hat{\mathcal{P}} \setminus K_L)] = \frac{1}{N}\sum_{n=1}^{N} \mathbb{P}\{\lambda_t^{N,n} \geq L\} \leq \frac{K_{1,T,3.4}}{L}.$$

Here $K_{2.3}$ and $K_{1,T,3.4}$ are the constants from Assumption 2.3 and Lemma 3.4. Let's next define

$$\mathcal{K}_L^* \stackrel{\text{def}}{=} \overline{\left\{\nu \in E : \nu\big(\hat{\mathcal{P}} \setminus K_{(L+j)^2}\big) < \frac{1}{\sqrt{L+j}} \text{ for all } j \in \mathbb{N}\right\}};$$

these are compact subsets of E. We have that

$$\begin{aligned}
\mathbb{P}\{\mu_t^N \notin \mathcal{K}_L^*\} &\leq \sum_{j=1}^{\infty} \mathbb{P}\left\{\mu_t^N\big(\hat{\mathcal{P}} \setminus K_{(L+j)^2}\big) > \frac{1}{\sqrt{L+j}}\right\} \\
&\leq \sum_{j=1}^{\infty} \frac{\mathbb{E}[\mu_t^N(\hat{\mathcal{P}} \setminus K_{(L+j)^2})]}{1/\sqrt{L+j}} \\
&\leq \sum_{j=1}^{\infty} \frac{K_{1,T,3.4}}{(L+j)^2/\sqrt{L+j}} \leq \sum_{j=1}^{\infty} \frac{K_{1,T,3.4}}{(L+j)^{3/2}}.
\end{aligned}$$

Since

$$\lim_{L\to\infty} \sum_{j=1}^{\infty} \frac{K_{1,T,3.4}}{(L+j)^{3/2}} = 0,$$

the result follows. □

We next need to understand the regularity of the μ^N's. For each $t \geq 0$ and $N \in \mathbb{N}$, we define

$$\mathscr{F}_t^N \stackrel{\text{def}}{=} \sigma\{\lambda_s^{N,n}; 0 \leq s \leq t, n \in \{1, 2, \dots, N\}\}.$$

Let's also define $q(x, y) \stackrel{\text{def}}{=} \min\{|x - y|, 1\}$ for all x and y in $\mathbb{R}$.

To proceed, let's first consider the L^N's. A useful tool will be the following integral bound. Fix $T > 0$ and suppose that f is a square-integrable function on

$[0, T]$. Then for any $0 \le s \le t \le T$,

$$\begin{aligned}\int_{r=s}^{t} f(r)\,dr &\le \sqrt{t-s}\sqrt{\int_{r=0}^{T} f^2(r)\,dr} \\ &\le \frac{1}{2}\left\{\frac{\sqrt{t-s}}{(t-s)^{1/4}} + (t-s)^{1/4}\int_{r=0}^{T} f^2(r)\,dr\right\} \\ &= \frac{1}{2}(t-s)^{1/4}\left\{1 + \int_{r=0}^{T} f^2(r)\,dr\right\}.\end{aligned} \tag{6.1}$$

LEMMA 6.2. *Define*

$$\Xi_N \stackrel{def}{=} \frac{1}{2N}\sum_{n=1}^{N}\left\{1 + \int_{r=0}^{t}(\lambda_r^{N,n})^2\,dr\right\} = \frac{1}{2}\left\{1 + \frac{1}{N}\sum_{n=1}^{N}\int_{r=0}^{T}(\lambda_r^{N,n})^2\,dr\right\}.$$

Then $\mathbb{E}[\Xi_N] \le \frac{1}{2}\{1 + K_{2,T,3.4}\}$ (*where* $K_{2,T,3.4}$ *is the constant from Lemma* 3.4) *and*

$$\mathbb{E}[|L_t^N - L_s^N||\mathscr{F}_s^N] \le (t-s)^{1/4}\mathbb{E}[\Xi_N|\mathscr{F}_s^N]$$

for all $0 \le s \le t \le T$.

PROOF. The bound $\mathbb{E}[\Xi_N] \le \frac{1}{2}\{1 + K_{2,T,3.4}\}$ is clear from Lemma 3.4. To proceed, let's write

$$L_t^N = 1 - \frac{1}{N}\sum_{n=1}^{N}\mathrm{M}_t^{N,n}.$$

By the martingale problem for L^N, we have that $L_t^N = A_t^N + \mathcal{M}_t$ where $\mathcal{M}$ is a martingale and where

$$A_t^N = \frac{1}{N}\sum_{n=1}^{N}\int_{r=0}^{t}\lambda_r^{N,n}\mathrm{M}_r^{N,n}\,dr.$$

Thus, for $0 \le s \le t$, we have (keeping in mind that L^N is nondecreasing)

$$|L_t^N - L_s^N| = L_t^N - L_s^N = A_t^N - A_s^N + \mathcal{M}_t - \mathcal{M}_s.$$

We then can use (6.1) to see that

$$A_t^N - A_s^N \le \frac{1}{N}\sum_{n=1}^{N}\int_{r=0}^{t}\lambda_r^{N,n}dr \le (t-s)^{1/4}\Xi_N.$$

The claimed bound follows. □

Of course, $\mathbb{P}\{L_t^N \in [0,1]\} = 1$ for all $t \ge 0$ and $N \in \mathbb{N}$, so compact containment (i.e., condition (a) of Theorem 7.2 of Chapter 2 of [11]) definitely holds.

Moreover, by Lemma 6.2 we have that for any $0 \le t \le T$, $0 \le u \le \delta$, and $0 \le v \le \delta \wedge t$,

$$\mathbb{E}[q(L^N_{t+u}, L^N_t)|\mathscr{F}^N_t]q(L^N_t, L^N_{t-v}) \le \mathbb{E}[L^N_{t+u} - L^N_t|\mathscr{F}^N_t] \le \delta^{1/4}\mathbb{E}[\Xi_N|\mathscr{F}^N_t].$$

Theorem 8.6 of Chapter 3 of [11] thus implies that $\{L^N\}_{N\in\mathbb{N}}$ is relatively compact.

LEMMA 6.3. *There is a random variable H_N with $\sup_{N\in\mathbb{N}}\mathbb{E}[H_N] < \infty$, such that for any $0 \le t \le T$, $0 \le u \le \delta$, and $0 \le v \le \delta \wedge t$,*

$$\mathbb{E}[q^2(\langle f, \mu^N_{t+u}\rangle_E, \langle f, \mu^N_t\rangle_E)q^2(\langle f, \mu^N_t\rangle_E, \langle f, \mu^N_{t-v}\rangle_E)|\mathscr{F}^N_t] \le \delta^{1/4}\mathbb{E}[H_N|\mathscr{F}^N_t].$$

PROOF. We start by using (5.2) to see that

$$\langle f, \mu^N_t\rangle_E = \langle f, \mu^N_0\rangle_E + A^{1,N}_t + A^{2,N}_t + B^{1,N}_t + B^{2,N}_t,$$

where

$$A^{1,N}_t = \frac{1}{N}\sum_{n=1}^N \int_{s=0}^t a^{1,N,n}_s\,ds,$$

$$A^{2,N}_t = \sum_{n=1}^N \int_{s=0}^t \mathcal{J}^f_{N,n}(s)\,d(1-\mathrm{M}^{N,n}_s),$$

$$B^{1,N}_t = \frac{1}{N}\sum_{n=1}^N \int_{s=0}^t \sigma_{N,n}\frac{\partial f}{\partial\lambda}(\hat{\rho}^{N,n}_s)\sqrt{\lambda^{N,n}_s}\,\mathrm{M}^{N,n}_s\,dW^n_s,$$

$$B^{2,N}_t = \varepsilon_N\frac{1}{N}\sum_{n=1}^N \int_{s=0}^t \beta^S_{N,n}\lambda^{N,n}_s\frac{\partial f}{\partial\lambda}(\hat{\rho}^{N,n}_s)\mathrm{M}^{N,n}_s\,dV_s,$$

where, for simplicity, we have defined

$$a^{1,N,n}_s \stackrel{\text{def}}{=} \frac{1}{2}\Big\{\big(\sigma^2_{N,n}\lambda^{N,n}_s + \varepsilon^2_N(\beta^S_{N,n})^2(\lambda^{N,n}_s)^2\big)\frac{\partial^2 f}{\partial\lambda^2}(\hat{\rho}^{N,n}_s) + \big(-\alpha_{N,n}(\lambda^{N,n}_s - \bar{\lambda}_{N,n}) - \varepsilon_N\beta^S_{N,n}\lambda^{N,n}_s X_s\big)\frac{\partial f}{\partial\lambda}(\hat{\rho}^{N,n}_s)\Big\}\mathrm{M}^{N,n}_s.$$

Thus, for any $0 \le s \le t \le T$,

$$\begin{aligned}
&\mathbb{E}[q^2(\langle f, \mu^N_t\rangle_E, \langle f, \mu^N_s\rangle_E)|\mathscr{F}^N_s]\\
&\quad\le 4\{\mathbb{E}[q^2(A^{1,N}_t, A^{1,N}_s)|\mathscr{F}^N_s] + \mathbb{E}[q^2(A^{2,N}_t, A^{2,N}_s)|\mathscr{F}^N_s]\\
&\qquad + \mathbb{E}[q^2(B^{1,N}_t, B^{1,N}_s)|\mathscr{F}^N_s] + \mathbb{E}[q^2(B^{2,N}_t, B^{2,N}_s)|\mathscr{F}^N_s]\}\\
&\quad\le 4\{\mathbb{E}[|A^{1,N}_t - A^{1,N}_s||\mathscr{F}^N_s] + \mathbb{E}[|A^{2,N}_t - A^{2,N}_s||\mathscr{F}^N_s]\\
&\qquad + \mathbb{E}[|B^{1,N}_t - B^{1,N}_s|^2|\mathscr{F}^N_s] + \mathbb{E}[|B^{2,N}_t - B^{2,N}_s|^2|\mathscr{F}^N_s]\}.
\end{aligned}$$

We now need to get some bounds. Due to Assumption 2.3, for any $0 \le s \le t < T$, we have that

$$|\mathcal{J}^f_{N,n}(t)| \le \frac{1}{N}\left\{K_{2.3}\left\|\frac{\partial f}{\partial \lambda}\right\|_C + \|f\|\right\}.$$

This implies

$$|A^{2,N}_t - A^{2,N}_s| \le \left\{K_{2.3}\left\|\frac{\partial f}{\partial \lambda}\right\|_C + \|f\|\right\}|L^N_t - L^N_s|;$$

thus, by Lemma 6.2 we have that

$$\mathbb{E}[|A^{2,N}_t - A^{2,N}_s||\mathscr{F}^N_s] \le (t-s)^{1/4}\left\{K_{2.3}\left\|\frac{\partial f}{\partial \lambda}\right\|_C + \|f\|\right\}\mathbb{E}[\Xi_N|\mathscr{F}^N_s]$$

for all $0 \le s \le t \le T$. To bound the increments of $A^{1,N}$, define

$$\Xi^{(1)}_N \stackrel{\text{def}}{=} \frac{1}{2}\left\{1 + \frac{1}{N}\sum_{n=1}^N \int_{r=0}^t (a^{1,N,n}_r)^2\,dr\right\}.$$

By Lemmata 3.4 and 10.1 we have that $\sup_{N\in\mathbb{N}} \mathbb{E}[\Xi^{(1)}_N] < \infty$. By (6.1),

$$|A^{1,N}_t - A^{1,N}_s| \le (t-s)^{1/4}\mathbb{E}\big[\Xi^{(1)}_N|\mathscr{F}^N_s\big].$$

We next turn to the martingale terms. We have that

$$\begin{aligned}
&\mathbb{E}[|B^{1,N}_t - B^{1,N}_s|^2|\mathscr{F}^N_s] \\
&\quad = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^N \int_{r=s}^t \left(\sigma_{N,n}\frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}^{N,n}_r)\sqrt{\lambda^{N,n}_r}\,\mathrm{M}^{N,n}_r\right)^2 dr\,\Big|\mathscr{F}^N_s\right] \\
&\quad \le \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^N \int_{r=s}^t \left(\sigma_{N,n}\frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}^{N,n}_r)\right)^2 \lambda^{N,n}_r\,dr\,\Big|\mathscr{F}^N_s\right] \\
&\quad \le (t-s)^{1/4}\mathbb{E}\big[\Xi^{(2)}_N|\mathscr{F}^N_s\big], \\
&\mathbb{E}[|B^{2,N}_t - B^{2,N}_s|^2|\mathscr{F}^N_s] \\
&\quad = \varepsilon^2_N\mathbb{E}\left[\int_{r=s}^t \left(\frac{1}{N}\sum_{n=1}^N \beta^S_{N,n}\lambda^{N,n}_r\frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}^{N,n}_s)\mathrm{M}^{N,n}_r\right)^2 dr\,\Big|\mathscr{F}^N_s\right] \\
&\quad \le \varepsilon^2_N\mathbb{E}\left[\frac{1}{N}\sum_{n=1}^N \int_{r=s}^t \left(\beta^S_{N,n}\frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}^{N,n}_r)\right)^2 (\lambda^{N,n}_r)^2\,dr\,\Big|\mathscr{F}^N_s\right] \\
&\quad \le \varepsilon^2_N(t-s)^{1/4}\mathbb{E}\big[\Xi^{(2)}_N|\mathscr{F}^N_s\big],
\end{aligned}$$

where

$$\Xi_N^{(2)} \stackrel{\text{def}}{=} \frac{1}{2}\left\{1 + \frac{1}{N}\sum_{n=1}^{N}\int_{r=0}^{T}\left(\sigma_{N,n}\frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}_r^{N,n})\right)^4 (\lambda_r^{N,n})^2\,dr\right\},$$

$$\Xi_N^{(3)} \stackrel{\text{def}}{=} \frac{1}{2}\left\{1 + \frac{1}{N}\sum_{n=1}^{N}\int_{r=0}^{T}\left(\beta_{N,n}^{S}\frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}_r^{N,n})\right)^4 (\lambda_r^{N,n})^4\,dr\right\}.$$

We have that

$$\sup_{N\in\mathbb{N}} \mathbb{E}[\Xi_N^{(2)}] < \infty \quad \text{and} \quad \sup_{N\in\mathbb{N}} \mathbb{E}[\Xi_N^{(3)}] < \infty.$$

Collecting things together, we get that for any $0 \le t \le T$, $0 \le u \le \delta$, and $0 \le v \le \delta \wedge t$,

$$\begin{aligned}
&\mathbb{E}[q^2(\langle f, \mu_{t+u}^N\rangle_E, \langle f, \mu_t^N\rangle_E)|\mathscr{F}_t^N]q^2(\langle f, \mu_t^N\rangle_E, \langle f, \mu_{t-v}^N\rangle_E)\\
&\quad \le \mathbb{E}[q^2(\langle f, \mu_{t+u}^N\rangle_E, \langle f, \mu_t^N\rangle_E)|\mathscr{F}_t^N]\\
&\quad \le 4\delta^{1/4}\mathbb{E}\left[\left\{\Xi_N^{(1)} + \left\{K_{2.3}\left\|\frac{\partial f}{\partial \lambda}\right\|_C + \|f\|\right\}\Xi_N + \Xi_N^{(2)} + \varepsilon_N^2\,\Xi_N^{(3)}\right\}\Big|\mathscr{F}_t^N\right]. \qquad \square
\end{aligned}$$

We can now prove the desired relative compactness.

LEMMA 6.4. *The sequence $\{\mu^N\}_{N\in\mathbb{N}}$ is relatively compact in $D_E[0,\infty)$.*

PROOF. Given Lemmas 6.1 and 6.3, the statement follows by Theorem 8.6 of Chapter 3 of [11]. $\square$

7. Uniqueness. We next verify that the solution of the resulting martingale problem is *unique*. We will use a duality argument (cf. Chapter 4.4 of [11]). In particular, here duality means that existence of a solution to a dual problem ensures uniqueness to the original problem.

LEMMA 7.1 (Uniqueness). *There is at most one solution of the martingale problem for $\mathcal{A}$ of (5.4) with initial condition $\pi \times \Lambda_\circ$.*

PROOF. We will use the duality arguments of Chapter 4.4 of [11]. Define $E^* \stackrel{\text{def}}{=} \bigcup_{M=1}^{\infty} C^\infty(\hat{P}^M)$. Let's begin by defining a flow on E^* as follows. Fix $f \in E^*$. Then $f \in C^\infty(\hat{P}^M)$ for some $M \in \mathbb{N}$. Fix next $(\hat{\mathsf{p}}_1, \hat{\mathsf{p}}_2, \dots, \hat{\mathsf{p}}_M) \in \hat{\mathcal{P}}^M$ where $\hat{\mathsf{p}}_m = (\mathsf{p}_m, \lambda_m)$ and $\mathsf{p}_m = (\alpha_m, \bar{\lambda}_m, \sigma_m, \beta_m^C, \beta_m^S)$ for $m \in \{1, 2, \dots, M\}$. Define

$$(T_t f)(\hat{\mathsf{p}}_1, \hat{\mathsf{p}}_2, \dots, \hat{\mathsf{p}}_M) \stackrel{\text{def}}{=} \mathbb{E}\left[f(\hat{\mathsf{p}}_t^{*,1}, \hat{\mathsf{p}}_t^{*,2}, \dots, \hat{\mathsf{p}}_t^{*,M})\exp\left[-\sum_{m=1}^{M}\int_{s=0}^{t}\lambda_s^{*,m}\,ds\right]\right],$$

where $\hat{\mathsf{p}}_t^{*,m} = (\mathsf{p}, \lambda_t^{*,m})$ and

$$\lambda_t^{*,m} = \lambda_m - \alpha_m \int_{s=0}^{t} (\lambda_s^{*,m} - \bar{\lambda}_m)\, ds + \sigma_m \int_{s=0}^{t} \sqrt{\lambda_s^{*,m}}\, dW_s^m$$

for all $m \in \{1, 2, \dots, M\}$. We also define

$$(H_m f)(\hat{\mathsf{p}}_1, \hat{\mathsf{p}}_2, \dots, \hat{\mathsf{p}}_M, \hat{\mathsf{p}}_{M+1}) = M\beta_{M+1}^C \lambda_{M+1} \frac{\partial f}{\partial \lambda_m}(\hat{\mathsf{p}}_1, \hat{\mathsf{p}}_2, \dots, \hat{\mathsf{p}}_M)$$

for $m \in \{1, 2, \dots, M\}$ and $\hat{\mathsf{p}}_{M+1} = (\mathsf{p}_{M+1}, \lambda_{M+1}) \in \hat{\mathcal{P}}$ where $\mathsf{p}_{M+1} = (\alpha_{M+1}, \bar{\lambda}_{M+1}, \sigma_{M+1}, \beta_{M+1}^C, \beta_{M+1}^S)$. Suppose that $f \in E^*$ and that in fact $f \in C^\infty(\hat{P}^M)$ for some $M \in \mathbb{N}$. Let $\mathfrak{e}$ be an exponential(1) random variable. Set $F_t \stackrel{\text{def}}{=} T_t f$ for $t < \mathfrak{e}$. Select $m \in \{1, 2, \dots, M\}$ according to a uniform distribution on $\{1, 2, \dots, M\}$ and set $F_{\mathfrak{e}} \stackrel{\text{def}}{=} H_m(T_{\mathfrak{e}} f)$. Restart the system.

Let's now connect F to μ. Fix $f \in E^*$ and $\mu \in E$. Then $f \in C^\infty(\hat{\mathcal{P}}^M)$ for some $M \in \mathbb{N}$, and we define

$$\phi(\mu, f) \stackrel{\text{def}}{=} \int_{(\hat{\mathsf{p}}_1, \hat{\mathsf{p}}_2, \dots, \hat{\mathsf{p}}_M) \in \hat{\mathcal{P}}^M} f(\hat{\mathsf{p}}_1, \hat{\mathsf{p}}_2, \dots, \hat{\mathsf{p}}_M) \mu(d\hat{\mathsf{p}}_1)\mu(d\hat{\mathsf{p}}_2) \cdots \mu(d\hat{\mathsf{p}}_M). \tag{7.1}$$

If we fix $1 = m_1 < m_2 < m_3 < \cdots < m_{L+1} = M + 1$ and $\{\tilde{f}_l\}_{l=1}^L \subset C^\infty(\hat{\mathcal{P}})$ and assume that

$$f(\hat{\mathsf{p}}_1, \hat{\mathsf{p}}_2, \dots, \hat{\mathsf{p}}_M) = \prod_{1 \le l \le L} \left\{ \prod_{m_l \le m < m_{l+1} - 1} \tilde{f}_l(\hat{\mathsf{p}}_m) \right\}$$

for all $(\hat{\mathsf{p}}_1, \hat{\mathsf{p}}_2, \dots, \hat{\mathsf{p}}_M) \in \hat{\mathcal{P}}^M$, then

$$\phi(\mu, f) = \prod_{l=1}^{L} \langle \tilde{f}_l, \mu \rangle_E^{m_{l+1} - m_l - 1}.$$

By Stone–Weierstrass, we can thus approximate Φ in $\mathcal{S}$ by linear combinations of functions of the form $\phi(\cdot, f)$ of (7.1) for some f's in E.

To proceed, let's fix $f \in E$ and apply $\mathcal{A}$ to the function $\mu \mapsto \phi(\mu, f)$ given by (7.1). It is fairly easy to see that if $\{\bar{\mu}_t^*\}_{t \ge 0}$ satisfies the martingale problem for $\mathcal{A}$, then for each $f \in E$,

$$\varphi(\bar{\mu}_t^*, f) = \int_{s=0}^{t} h_1(\bar{\mu}_s^*, f)\, ds + \mathcal{M}_t^{(1)},$$

where $\mathcal{M}^{(1)}$ is a martingale and where, if $f \in C^\infty(\hat{\mathcal{P}}^M)$,

$$h_1(\mu, f) = \sum_{m=1}^{M} \int_{\hat{\mathsf{p}} = (\hat{\mathsf{p}}_1, \hat{\mathsf{p}}_2, \dots, \hat{\mathsf{p}}_M) \in \mathcal{P}^M} \{(\mathcal{L}_{1,m} f)(\hat{\mathsf{p}}) + \langle \mathcal{Q}, \mu \rangle_E (\mathcal{L}_{2,m} f)(\hat{\mathsf{p}})\}$$
$$\times \mu(d\hat{\mathsf{p}}_1)\mu(d\hat{\mathsf{p}}_2) \cdots \mu(d\hat{\mathsf{p}}_M),$$

where $\mathcal{L}_{1,m}$ and $\mathcal{L}_{2,m}$ denote, respectively, the actions of $\mathcal{L}_1$ and $\mathcal{L}_2$ defined by (5.3) on the mth coordinate of f. On the other hand, we also have that for $\mu \in E$,

$$\varphi(\mu, F_t) = \int_{s=0}^{t} h_2(\mu, F_s)\, ds + \mathcal{M}_t^{(2)},$$

where $\mathcal{M}^{(2)}$ is a martingale and

$$h_2(\mu, f) = \sum_{m=1}^{M} \int_{\hat{\mathfrak{p}}=(\hat{\mathfrak{p}}_1,\hat{\mathfrak{p}}_2,\dots,\hat{\mathfrak{p}}_M)\in\mathcal{P}^M} (\mathcal{L}_{1,m} f)(\hat{\mathfrak{p}})\mu(d\hat{\mathfrak{p}}_2)\cdots\mu(d\hat{\mathfrak{p}}_M)$$
$$+ \frac{1}{M}\sum_{m=1}^{M}\{\varphi(\mu, H_m f) - \varphi(\mu, f)\}.$$

Note that

$$\begin{aligned}
&\frac{1}{M}\sum_{m=1}^{M}\varphi(\mu, H_m f)\\
&\quad= \sum_{m=1}^{M}\int_{\hat{\mathfrak{p}}=(\hat{\mathfrak{p}}_1,\hat{\mathfrak{p}}_2,\dots,\hat{\mathfrak{p}}_M,\hat{\mathfrak{p}}_{M+1})\in\mathcal{P}^{M+1}} \beta_{M+1}^{C}\lambda_{M+1}\frac{\partial f}{\partial\lambda_m}(\hat{\mathfrak{p}})\mu(d\hat{\mathfrak{p}}_2)\cdots\mu(d\hat{\mathfrak{p}}_{M+1})\\
&\quad= \sum_{m=1}^{M}\int_{\hat{\mathfrak{p}}=(\hat{\mathfrak{p}}_1,\hat{\mathfrak{p}}_2,\dots,\hat{\mathfrak{p}}_M)\in\mathcal{P}^{M}} \langle\mathcal{Q},\mu\rangle_E(\mathcal{L}_{2,m} f)(\hat{\mathfrak{p}})\mu(d\hat{\mathfrak{p}}_2)\cdots\mu(d\hat{\mathfrak{p}}_M).
\end{aligned}$$

Collecting things together, we have that

$$h_1(\mu, f) = h_2(\mu, f) + \varphi(\mu, f)$$

and this implies uniqueness. □

8. Proof of main theorem. We now have our first convergence result. Let $\mathbb{Q}_N$ be the $\mathbb{P}$-law of μ^N, that is,

$$\mathbb{Q}_N(A) \stackrel{\text{def}}{=} \mathbb{P}\{\mu^N \in A\}$$

for all $A \in \mathscr{B}(D_E[0,\infty))$. Thus, $\mathbb{Q}_N \in \mathscr{P}(D_E[0,\infty))$ for all $N \in \mathbb{N}$. For $\omega \in D_E[0,\infty)$, define $X_t(\omega) \stackrel{\text{def}}{=} \omega(t)$ for all $t \geq 0$.

PROPOSITION 8.1. *We have that $\mathbb{Q}_N$ converges [in the topology of $\mathscr{P}(D_E[0,\infty))$] to the solution $\mathbb{Q}$ of the martingale problem generated by $\mathcal{A}$ of (5.4) and such that $\mathbb{Q}X_0^{-1} = \delta_{\pi\times\Lambda_\circ}$. In other words, $\mathbb{Q}\{X_0 = \pi \times \Lambda_\circ\} = 1$ and for all $\Phi \in \mathcal{S}$ and $0 \leq r_1 \leq r_2 \leq \cdots \leq r_J = s < t < T$ and $\{\psi_j\}_{j=1}^{J} \subset B(E)$, we have that*

$$\lim_{N\to\infty}\mathbb{E}^{\mathbb{Q}}\left[\left\{\Phi(X_t) - \Phi(X_s) - \int_{r=s}^{t}(\mathcal{A}\Phi)(X_r)\, dr\right\}\prod_{j=1}^{J}\psi_j(X_{r_j})\right] = 0,$$

where $\mathbb{E}^{\mathbb{Q}}$ is the expectation operator defined by $\mathbb{Q}$.

PROOF. The result follows from Lemmata 5.1, 6.4 and 7.1. Of course, we also have that for any $\Phi \in \mathcal{S}$,

$$\mathbb{E}^{\mathbb{Q}}[\Phi(X_0)] = \lim_{N\to\infty} \mathbb{E}^{\mathbb{Q}}[\Phi(\mu_0^N)] = \Phi(\pi \times \Lambda_\circ),$$

which implies the claimed initial condition. □

We next want to identify $\mathbb{Q}$.

LEMMA 8.2. *We have that* $\mathbb{Q} = \delta_{\bar{\mu}}$, *where* $\bar{\mu}$ *is given by* (4.5).

PROOF. Recall (4.4) and the operators $\mathcal{L}_1, \mathcal{L}_2$ from (5.3) and the definition of Q in (4.3). For any $f \in C^\infty(\hat{\mathcal{P}})$,

$$\langle f, \bar{\mu}_t\rangle_E = \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}} \mathbb{E}\Big[f(\mathsf{p}, \lambda_t^*(\hat{\mathsf{p}}))\exp\Big[-\int_{s=0}^t \lambda_s^*(\hat{\mathsf{p}})\,ds\Big]\Big]\pi(d\mathsf{p})\Lambda_\circ(d\lambda).$$

Thus,

$$\begin{aligned}\frac{d}{dt}\langle f, \bar{\mu}_t\rangle_E &= \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}} \mathbb{E}\Big[(\mathcal{L}_1 f)(\mathsf{p}, \lambda_t^*(\hat{\mathsf{p}}))\exp\Big[-\int_{s=0}^t \lambda_s^*(\hat{\mathsf{p}})\,ds\Big]\Big]\pi(d\mathsf{p})\Lambda_\circ(d\lambda)\\ &\quad + \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}} \mathbb{E}\Big[(\mathcal{L}_2 f)(\mathsf{p}, \lambda_t^*(\hat{\mathsf{p}}))Q(t)\exp\Big[-\int_{s=0}^t \lambda_s^*(\hat{\mathsf{p}})\,ds\Big]\Big]\\ &\qquad\qquad \times \pi(d\mathsf{p})\Lambda_\circ(d\lambda)\\ &= \langle \mathcal{L}_1 f, \bar{\mu}_t\rangle_E + Q(t)\langle \mathcal{L}_2 f, \bar{\mu}_t\rangle_E.\end{aligned}$$

To proceed, define

$$G(t) \stackrel{\text{def}}{=} \int_{\substack{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}\\ \mathsf{p}=(\alpha,\bar{\lambda},\sigma,\beta^C,\beta^S)}} \beta^C \mathbb{E}\Big[\exp\Big[-\int_{s=0}^t \lambda_s^*(\hat{\mathsf{p}})\,ds\Big]\Big]\pi(d\mathsf{p})\Lambda_\circ(d\lambda).$$

On the one hand, we have that

$$\begin{aligned}\dot{G}(t) &= -\int_{\substack{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}\\ \mathsf{p}=(\alpha,\bar{\lambda},\sigma,\beta^C,\beta^S)}} \beta^C \mathbb{E}\Big[\lambda_t^*(\hat{\mathsf{p}})\exp\Big[-\int_{s=0}^t \lambda_s^*(\hat{\mathsf{p}})\,ds\Big]\Big]\pi(d\mathsf{p})\Lambda_\circ(d\lambda)\\ &= -\int_{\substack{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}\\ \mathsf{p}=(\alpha,\bar{\lambda},\sigma,\beta^C,\beta^S)}} \beta^C \lambda \bar{\mu}_t(d\hat{\mathsf{p}}) = -\langle \mathcal{Q}, \bar{\mu}_t\rangle_E.\end{aligned}$$

We want to show that

$$\dot{G}(t) = -Q(t). \tag{8.1}$$

Indeed, fix $\hat{\mathsf{p}} = (\mathsf{p}, \lambda) \in \hat{\mathcal{P}}$ where $\mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S)$. Define

$$M_s \stackrel{\text{def}}{=} \exp\Big[-b^{\mathsf{p}}(t-s)\lambda_s^*(\hat{\mathsf{p}}) - \int_{r=s}^t b^{\mathsf{p}}(t-r)\{Q(r) + \alpha\bar{\lambda}\}\,dr - \int_{r=0}^s \lambda_r^*(\hat{\mathsf{p}})\,dr\Big]$$

for $0 \le s \le t$. Using the calculations of [10],

$$\begin{aligned} dM_s &= d\mathcal{M}_s + \big\{\dot{b}^{\mathsf{p}}(t-s)\lambda_s^*(\hat{\mathsf{p}}) - b^{\mathsf{p}}(t-s)\big\{-\alpha\big(\lambda_s^*(\hat{\mathsf{p}}) - \bar{\lambda}\big) + Q(s)\big\} \\ &\qquad + \tfrac{1}{2}\sigma^2\big(b^{\mathsf{p}}(t-s)\big)^2\lambda_s^*(\hat{\mathsf{p}}) + b^{\mathsf{p}}(t-s)\big(Q(s) + \alpha\bar{\lambda}\big) - \lambda_s^*(\hat{\mathsf{p}})\big\}M_s\,ds \\ &= d\mathcal{M}_s, \end{aligned}$$

where $\mathcal{M}$ is a martingale [we use here the ODE (4.2)]. Noting that

$$\begin{aligned} M_0 &= \exp\Big[-b^{\mathsf{p}}(t)\lambda - \int_{r=0}^t b^{\mathsf{p}}(t-r)\{Q(r) + \alpha\bar{\lambda}\}\,dr\Big], \\ M_t &= \exp\Big[-\int_{r=0}^t \lambda_r^*(\hat{\mathsf{p}})\,dr\Big], \end{aligned}$$

we have that

$$\begin{aligned} G(t) = \int_{\substack{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}} \\ \mathsf{p}=(\alpha,\bar{\lambda},\sigma,\beta^C,\beta^S)}} \beta^C \exp\Big[-b^{\mathsf{p}}(t)\lambda \\ - \int_{r=0}^t b^{\mathsf{p}}(t-r)\{Q(r) + \alpha\bar{\lambda}\}\,dr\Big] \\ \times \pi(d\mathsf{p})\Lambda_\circ(d\lambda). \end{aligned}$$

Differentiating this, we get that

$$\begin{aligned} \dot{G}(t) = -\int_{\substack{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}} \\ \mathsf{p}=(\alpha,\bar{\lambda},\sigma,\beta^C,\beta^S)}} \beta^C\Big[\dot{b}^{\mathsf{p}}(t)\lambda + \int_{r=0}^t \dot{b}^{\mathsf{p}}(t-r)\{Q(r) + \alpha\bar{\lambda}\}\,dr\Big] \\ \times \exp\Big[-b^{\mathsf{p}}(t)\lambda - \int_{r=0}^t b^{\mathsf{p}}(t-r)\{Q(r) + \alpha\bar{\lambda}\}\,dr\Big] \\ \times \pi(d\mathsf{p})\Lambda_\circ(d\lambda) \\ = -Q(t), \end{aligned}$$

where we have used the defining equation (4.3) for Q. Thus, (8.1) holds, so we have that

$$\frac{d}{dt}\langle f, \bar{\mu}_t\rangle_E = \langle \mathcal{L}_1 f, \bar{\mu}_t\rangle_E + \langle \mathcal{Q}, \bar{\mu}_t\rangle_E \langle \mathcal{L}_2 f, \bar{\mu}_t\rangle_E.$$

Thus,

$$\Phi(\bar{\mu}_t) = \Phi(\bar{\mu}_0) + \int_{s=0}^t (\mathcal{A}\Phi)(\bar{\mu}_s)\,ds,$$

and, hence, $\delta_{\bar{\mu}}$ satisfies the martingale problem generated by $\mathcal{A}$. Of course, we also have that $\bar{\mu}_0 = \pi \times \Lambda_\circ$. By uniqueness, the claim follows. $\square$

We now can finish the proof of our main result.

PROOF OF THEOREM 4.2. Since weak convergence to a constant implies convergence in probability, we have (4.6). Using the fact that the map $\varphi : \hat{\mathcal{P}} \mapsto 1$ is in $C(\hat{\mathcal{P}})$, L^N is a continuous transformation of μ^N into $D_{\mathbb{R}}[0, \infty)$. From (4.7) we have that

$$\lim_{N\to\infty} \mathbb{P}\{d_{\mathbb{R}}(L^N, F) \geq \delta\} = 0$$

for each $\delta > 0$. To finish the proof, we need to replace the Skorohod norm $d_{\mathbb{R}}$ by the supremum norm. From (4.8) we have that $K_T \stackrel{\text{def}}{=} \sup_{0\leq t\leq T} \dot{F}(t)$ is finite for each $T > 0$. To get the claimed convergence, we adopt the notation of Chapter 3.5 of [11]. For any nondecreasing and differentiable map g of $[0, T]$ into itself and any $t \in [0, T]$, we have that

$$\begin{aligned}
|L_t^N - F(t)| &\leq |L_t^N - F(g(t))| + |F(g(t)) - F(t)| \\
&\leq \sup_{0\leq t\leq T} |L_t^N - F(g(t))| + K_T |g(t) - t| \\
&\leq \sup_{0\leq t\leq T} |L_t^N - F(g(t))| + K_T T \sup_{0\leq t\leq T} |\dot{g}(t) - 1| \\
&\leq \sup_{0\leq t\leq T} |L_t^N - F(g(t))| \\
&\quad + K_T T \max\Big\{\Big|\exp\Big[\sup_{0\leq t\leq T} |\log \dot{g}(t)|\Big] - 1\Big|, \\
&\qquad\qquad \Big|\exp\Big[-\sup_{0\leq t\leq T} |\log \dot{g}(t)|\Big] - 1\Big|\Big\}.
\end{aligned}$$

Varying g, we get that

$$\begin{aligned}
&\sup_{0\leq t\leq T} |L_t^N - F(t)| \\
&\quad \leq d_{\mathbb{R}}(L^N, F) + K_T T \max\{|\exp[d_{\mathbb{R}}(L^N, F)] - 1|, \\
&\qquad\qquad |\exp[-d_{\mathbb{R}}(L^N, F)] - 1|\}.
\end{aligned}$$

The claim now follows; note that F and L^N both take values in $[0, 1]$. $\square$

9. Conclusion and extensions. We have developed a point process model of correlated default timing in a portfolio of firms, and have analyzed typical default profiles in the limit as the size of the pool grows. Our empirically motivated model

captures two important sources of default clustering, namely, the exposure of firms to a systematic risk process, and contagion. We have proved a law of large numbers for the default rate in the pool.

There are several potential extensions of our work. For example, the default intensity dynamics (2.1) can be generalized to include a dependence on the systematic risk process of the magnitude of the jump at a default. Then, the impact of a default on the surviving firms depends on the state of the systematic risk: intuitively, if the economy is weak, firms are fragile and more susceptible to contagion. This generalization of the intensity dynamics is empirically plausible, and can be treated with arguments similar to the ones we currently use.

10. Proofs of Lemmas 3.1, 3.2, 3.4 and 4.1. In this section we prove Lemmas 3.1, 3.2, 3.4 and 4.1. For presentation purposes, we first collect in Lemma 10.1 some a-priori bounds that will be useful in the proof of these lemmas. Then, in Section 10.2 we proceed with the proof of Lemmas 3.1, 3.2 and 3.4. We mention here that the square-root singularity unavoidably complicates the analysis. The theory behind CIR-like processes is a bit delicate due to the square root singularity in the diffusion, so we need to develop some new modifications to existing results (cf. [15, 17, 18]). Last, in Section 10.3 we prove Lemma 4.1.

10.1. *Effect of systematic risk.* Our first step is to get some usable bounds on the systematic risk X. We need these bounds since, as we mentioned in Section 3, the $\lambda_t\,dX_t$ term contains the term $\lambda_t X_t\,dt$, implying that the dynamics of the $\mathbb{R}^2$-valued process (λ, X) contain a superlinear drift. Note that the systematic risk process X of course has an explicit form:

$$X_t = e^{-\gamma t}x_\circ + \int_{s=0}^{t} e^{-\gamma(t-s)}\,dV_s, \qquad t > 0.$$

Fix $\mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S) \in \mathcal{P}$, $\lambda_\circ$ in $\mathbb{R}_+$, and ξ as required in the beginning of Section 3. Define

$$\begin{aligned}
\Gamma_t &\stackrel{\text{def}}{=} \alpha t + \beta^S\gamma \int_{s=0}^{t} X_s\,ds,\\
Z_t &\stackrel{\text{def}}{=} \lambda_\circ + \alpha\bar{\lambda}\int_{s=0}^{t} e^{\Gamma_s}\,ds + \beta^C\int_{s=0}^{t} e^{\Gamma_s}\,d\xi_s\\
&= \lambda_\circ + \alpha\bar{\lambda}\int_{s=0}^{t} e^{\Gamma_s}\,ds + \beta^C\left\{e^{\Gamma_t}\xi_t - \int_{s=0}^{t} e^{\Gamma_s}\xi_s(\alpha+\beta^S\gamma X_s)\,ds\right\}\\
&= \lambda_\circ + \int_{s=0}^{t} e^{\Gamma_s}\{\alpha\bar{\lambda} - \beta^C\xi_s(\alpha+\beta^S\gamma X_s)\}\,ds + \beta^C\xi_t e^{\Gamma_t}
\end{aligned}$$

for all $t \geq 0$. The alternate representations of Z will allow us bounds which are independent of ξ.

Our first result is a bounds on X, Γ and Z which explicitly depend on various coefficients. The importance of the bound on the moments of Z_t is that they do not depend on ξ.

LEMMA 10.1. *For each $p \geq 1$ and $t \geq 0$,*

$$\mathbb{E}[X_t^{2p}]^{1/(2p)} \leq |x_\circ| + \frac{1}{2\sqrt{\gamma}}\left(\frac{(2p)!}{p!}\right)^{1/(2p)},$$

$$\mathbb{E}[\exp[p\Gamma_t]] \leq \exp\big[|p|\{\alpha t + |\beta^C x_\circ|\} + \tfrac{1}{2}(p\beta^C)^2 t\big],$$

$$\begin{aligned}\mathbb{E}[Z_t^{2p}]^{1/(2p)} &\leq \lambda_\circ + |\beta^C|\mathbb{E}[e^{-2p\Gamma_t}]^{1/(2p)} \\ &\quad + t^{1-1/(2p)}\left(\int_{s=0}^t \mathbb{E}[e^{-4p\Gamma_s}]\,ds\right)^{1/(4p)} \\ &\quad \times \Bigg\{\alpha\bar{\lambda}t^{1/(4p)} + \alpha|\beta^C|t^{1/(4p)} \\ &\quad + |\beta^C\beta^S\gamma|\left(\int_{s=0}^t \mathbb{E}[X_s^{4p}]\,ds\right)^{1/(4p)}\Bigg\}.\end{aligned}$$

PROOF. We first bound X. For every $p \geq 1$ and $t \geq 0$

$$\begin{aligned}\mathbb{E}[X_t^{2p}]^{1/(2p)} &\leq |x_\circ e^{-\gamma t}| + \left\{\mathbb{E}\left[\left|\int_{s=0}^t e^{-\gamma(t-s)}\,dV_s\right|^{2p}\right]\right\}^{1/(2p)} \\ &= |x_\circ e^{-\gamma t}| + \sqrt{\int_{s=0}^t e^{-2\gamma(t-s)}\,ds}\left(\frac{(2p)!}{2^p p!}\right)^{1/(2p)} \\ &\leq |x_\circ| + \frac{1}{\sqrt{2\gamma}}\left(\frac{(2p)!}{2^p p!}\right)^{1/(2p)} \\ &= |x_\circ| + \frac{1}{2\sqrt{\gamma}}\left(\frac{(2p)!}{p!}\right)^{1/(2p)}.\end{aligned}$$

Next note that

$$\begin{aligned}\Gamma_t &= \alpha t - \beta^S\gamma\int_{s=0}^t x_\circ e^{-\gamma s}\,ds \\ &\quad - \beta^S\gamma\int_{s=0}^t\left\{\int_{r=0}^s e^{-\gamma(s-r)}\,dV_r\right\}ds \\ &= \alpha t - \beta^S x_\circ\{1-e^{-\gamma t}\} - \beta^S\gamma\int_{r=0}^t\left\{\int_{s=r}^t e^{-\gamma(s-r)}\,ds\right\}dV_r \\ &= \alpha t - \beta^S x_\circ\{1-e^{-\gamma t}\} - \beta^S\int_{r=0}^t\{1-e^{-\gamma(t-r)}\}\,dV_r.\end{aligned}$$

Thus, for any $p \in \mathbb{R}$

$$\begin{aligned}\mathbb{E}[\exp[p\Gamma_t]] &= \exp\Big[p\{\alpha t + \beta^C x_\circ(1-e^{-\gamma t})\} \\ &\qquad + \frac{(p\beta^C)^2}{2}\int_{r=0}^t \{1-e^{-\gamma(t-r)}\}^2\,dr\Big] \\ &\le \exp\Big[|p|\{\alpha t + |\beta^C x_\circ|\} + \frac{1}{2}(p\beta^C)^2 t\Big].\end{aligned}$$

We can finally bound Z. We have that

$$\begin{aligned}\mathbb{E}[Z_t^{2p}]^{1/(2p)} &\le \lambda_\circ + \mathbb{E}\Big[\Big(\int_{s=0}^t e^{-\Gamma_s}\{\alpha\bar{\lambda} - \beta^C\xi_s(\alpha+\beta^S\gamma X_s)\}\,ds\Big)^{2p}\Big]^{1/(2p)} \\ &\qquad + |\beta^C|\mathbb{E}[e^{-2p\Gamma_t}]^{1/(2p)}.\end{aligned}$$

We also have that

$$\begin{aligned}&\mathbb{E}\Big[\Big(\int_{s=0}^t e^{\Gamma_s}\{\alpha\bar{\lambda} - \beta^C\xi_s(\alpha+\beta^S\gamma X_s)\}\,ds\Big)^{2p}\Big]^{1/(2p)} \\ &\quad\le \mathbb{E}\Big[\Big(\int_{s=0}^t e^{2\Gamma_s}\,ds\Big)^p\Big(\int_{s=0}^t \{\alpha\bar{\lambda} - \beta^C\xi_s(\alpha+\beta^S\gamma X_s)\}^2\,ds\Big)^p\Big]^{1/(2p)} \\ &\quad\le \mathbb{E}\Big[\Big(\int_{s=0}^t e^{2\Gamma_s}\,ds\Big)^{2p}\Big]^{1/(4p)} \\ &\qquad\times \mathbb{E}\Big[\Big(\int_{s=0}^t \{\alpha\bar{\lambda} - \beta^C\xi_s(\alpha+\beta^S\gamma X_s)\}\,ds\Big)^{2p}\Big]^{1/(4p)} \\ &\quad\le t^{1-1/(2p)}\mathbb{E}\Big[\int_{s=0}^t e^{-4p\Gamma_s}\,ds\Big]^{1/(4p)} \\ &\qquad\times \mathbb{E}\Big[\int_{s=0}^t \{\alpha\bar{\lambda} - \beta^C\xi_s(\alpha+\beta^S\gamma X_s)\}^{2p}\,ds\Big]^{1/(4p)} \\ &\quad\le t^{1-1/(2p)}\Big(\int_{s=0}^t \mathbb{E}[e^{-4p\Gamma_s}]\,ds\Big)^{1/(4p)} \\ &\qquad\times \Big\{\alpha\bar{\lambda}t^{1/(4p)} + \alpha|\beta^C|t^{1/(4p)} + |\beta^C\beta^S\gamma|\mathbb{E}\Big[\int_{s=0}^t X_s^{4p}\,ds\Big]^{1/(4p)}\Big\}.\end{aligned}$$

Combine things together to get the bound on Z. $\square$

10.2. *Proofs of Lemmas* 3.1, 3.2 *and* 3.4. Let's next understand the regularity of various CIR-like processes which we use. Before proceeding with the proofs, we define a function $\psi_\eta(x)$ that will be essential for the proofs. It is introduced in

order to deal with the square-root singularity. In particular, let

$$\psi_\eta(x) \stackrel{\text{def}}{=} \frac{2}{\ln \eta^{-1}} \int_{y=0}^{|x|} \left\{ \int_{z=0}^{y} \frac{1}{z} \chi_{[\eta,\eta^{1/2}]}(z)\, dz \right\} dy \quad \text{and} \quad g_\eta(x) \stackrel{\text{def}}{=} |x| - \psi_\eta(x)$$

for all $x \in \mathbb{R}$.

Let us then study some important properties of $\psi_\eta(x)$ that will be repeatedly used in the proofs. First, we note that ψ_η is even, so g_η is also even. Taking derivatives, we have that

$$\dot{\psi}_\eta(x) = \frac{2}{\ln \eta^{-1}} \int_{z=0}^{x} \frac{1}{z} \chi_{[\eta,\eta^{1/2}]}(z)\, dz \quad \text{and} \quad \ddot{\psi}_\eta(x) = \frac{2}{\ln \eta^{-1}} \frac{1}{x} \chi_{[\eta,\eta^{1/2}]}(x)$$

for all $x > 0$. Since $\ddot{g}_\eta = -\ddot{\psi}_\eta \le 0$, $\dot{g}_\eta$ is nonincreasing. For $x > \sqrt{\eta}$,

$$\dot{g}_\eta(x) = 1 - 2\frac{\ln \eta^{1/2} - \ln \eta}{\ln(1/\eta)} = 0,$$

so in fact $\dot{g}_\eta$ is nonnegative on $(0, \infty)$ and it vanishes on $[\sqrt{\eta}, \infty)$. Thus, g_η is nondecreasing and reaches its maximum at $\sqrt{\eta}$. Since $g_\eta(0) = 0$, we in fact have that

$$0 \le g_\eta(x) \le g_\eta(\sqrt{\eta})$$

for all $x \ge 0$. Since $\dot{g}_\eta$ is nonincreasing on $(0, \infty)$ and $\dot{g}_\eta(x) = 1$ for $x \in (0, \eta)$, we have that $\dot{g}_\eta(x) \le 1$ for all $x \in (0, \sqrt{\eta})$, so $g_\eta(\sqrt{\eta}) \le \sqrt{\eta}$. Since g_η is even, we in fact must have that $|g_\eta(x)| \le \sqrt{\eta}$ for all $x \in \mathbb{R}$. Hence,

$$|x| \le \psi_\eta(x) + \sqrt{\eta}$$

for all $x \in \mathbb{R}$. We finally note that

$$|\ddot{\psi}_\eta(x)| \le \frac{2}{\ln \eta^{-1}} \frac{1}{|x|} \chi_{[\eta,\infty)}(|x|) \le \frac{2}{\ln \eta^{-1}} \min\left\{ \frac{1}{|x|}, \frac{1}{\eta} \right\}$$

for all $x \in \mathbb{R}$.

Now we have all the necessary tools to proceed with the proof of the lemmas.

PROOF OF LEMMA 3.1. For each $N \in \mathbb{N}$, define

$$\varrho_N(t) \stackrel{\text{def}}{=} \frac{\lfloor tN \rfloor}{N}$$

for all $t \in [0, T]$. For each $N \in \mathbb{N}$, define

$$Y_t^N = \sigma \int_{s=0}^{t} e^{\Gamma_s/2} \sqrt{(Y_{\varrho_N(s)}^N + Z_s) \vee 0}\, dW_s^* + \beta \int_{s=0}^{t} \big((Y_{\varrho^N(s)}^N + Z_s) \vee 0\big)\, dV_s.$$

We will show that $(Z_t + Y_t^N)e^{\Gamma_t}$ converges to a solution of (3.1) (as $N \nearrow \infty$).

As a first step, let's bound some moments. Fix $p > 1$. For $0 \le s \le t \le T$, [17], Exercise 3.25, gives us that

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\int_{r=s}^{t} \big((Y^N_{\varrho^N(r)} + Z_r) \vee 0\big)\, dV_r\Big|^{2p}\Big] \\
&\quad \le \big(p(2p-1)\big)^p (t-s)^{p-1} \int_{r=s}^{t} \mathbb{E}\big[\big((Y^N_{\varrho^N(r)} + Z_r) \vee 0\big)^{2p}\, dr\big] \\
&\quad \le \big(p(2p-1)\big)^p (t-s)^{p-1} \int_{r=s}^{t} \mathbb{E}\big[|Y^N_{\varrho^N(r)} + Z_r|^{2p}\big]\, dr \\
&\quad \le 2^{2p-1}\big(p(2p-1)\big)^p (t-s)^{p-1} \\
&\qquad \times \Big\{\int_{r=s}^{t} \mathbb{E}\big[|Y^N_{\varrho^N(r)}|^{2p}\big]\, dr + \int_{r=s}^{t} \mathbb{E}\big[|Z_r|^{2p}\big]\, dr\Big\}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\int_{r=s}^{t} e^{\Gamma_r/2}\sqrt{(Y^N_{\varrho_N(r)} + Z_r) \vee 0}\, dW^*_r\Big|^{2p}\Big] \\
&\quad \le \big(p(2p-1)\big)^p (t-s)^{p-1} \int_{r=s}^{t} \mathbb{E}\big[e^{p\Gamma_r}\big|(Y^N_{\varrho_N(r)} + Z_r) \vee 0\big|^p\big]\, dr \\
&\quad \le \frac{1}{2}\big(p(2p-1)\big)^p (t-s)^{p-1}\Big\{\int_{r=s}^{t} \mathbb{E}[e^{2p\Gamma_r}] + \mathbb{E}\big[|Y^N_{\varrho_N(r)} + Z_r|^{2p}\big]\, dr\Big\} \\
&\quad \le \frac{1}{2}\big(p(2p-1)\big)^p (t-s)^{p-1} \\
&\qquad \times \Big\{\int_{r=s}^{t} \mathbb{E}[e^{2p\Gamma_r}]\, dr + 2^{2p-1}\int_{r=s}^{t} \mathbb{E}\big[|Y^N_{\varrho_N(r)}|^{2p}\big]\, dr \\
&\qquad\qquad + 2^{2p-1}\int_{r=s}^{t} \mathbb{E}\big[|Z_r|^{2p}\big]\, dr\Big\}.
\end{aligned}$$

We can bound the effect of Z by Lemma 10.1. Collecting things together, and using the fact that $\varrho_N(t) \le t$, we have that there is a $K_A > 0$ such that

$$\begin{aligned}
\mathbb{E}\big[|Y^N_{\varrho_N(t)}|^{2p}\big] &\le K_A + K_A \int_{s=0}^{\varrho_N(t)} \mathbb{E}\big[|Y^N_{\varrho_N(s)}|^{2p}\big]\, dr \\
&\le K_A + K_A \int_{s=0}^{t} \mathbb{E}\big[|Y^N_{\varrho_N(s)}|^{2p}\big]\, dr
\end{aligned}$$

for all $N \in \mathbb{N}$ and $t \in [0, T]$, which in turn implies that

$$\sup_{0 \le t \le T} \mathbb{E}\big[|Y^N_{\varrho_N(t)}|^{2p}\big] \le K_A e^{K_A T} \tag{10.1}$$

for $0 \le t \le T$. This in turn implies that there is a $K_B > 0$ such that

$$\mathbb{E}\big[|Y^N_t - Y^N_{\varrho^N(t)}|^{2p}\big] \le K_B |t - \varrho_N(t)|^p \le K_B \frac{1}{N^p} \tag{10.2}$$

for all $0 \le t \le T$.

We next want to show that Y^N converges in L^1. Fix N and N' in $\mathbb{N}$ and define

$$\nu_t^{N,N'} \stackrel{\text{def}}{=} Y_t^N - Y_t^{N'}.$$

Fix also $\eta > 0$. We have that

$$|\nu_t^{N,N'}| \le \psi_\eta(\nu_t^{N,N'}) + \sqrt{\eta} = \sigma^2 A_t^{1,N,N'} + \beta^2 A_t^{2,N,N'} + \mathcal{M}_t + \sqrt{\eta},$$

where $\mathcal{M}$ is a martingale and

$$\begin{aligned}
A_t^{1,N,N'} &= \frac{1}{2}\int_{s=0}^t \ddot{\psi}_\eta(\nu_s^{N,N'}) e^{\Gamma_s}\{\sqrt{(Y_{\varrho_N(s)}^{N'} + Z_s)\vee 0} - \sqrt{(Y_{\varrho_{N'}(s)}^{N} + Z_s)\vee 0}\}^2\,ds \\
&\le \frac{1}{2}\int_{s=0}^t \ddot{\psi}_\eta(\nu_s^{N,N'}) e^{\Gamma_s}|Y_{\varrho_N(s)}^N - Y_{\varrho_{N'}(s)}^{N'}|\,ds \\
&\le \frac{1}{2}\int_{s=0}^t e^{\Gamma_s}\ddot{\psi}_\eta(\nu_s^{N,N'})\{|\nu_s^{N,N'}| + |Y_s^N - Y_{\varrho_N(s)}^N| + |Y_s^{N'} - Y_{\varrho_{N'}(s)}^{N'}|\}\,ds \\
&\le \frac{1}{2\ln\eta^{-1}}\int_{s=0}^t e^{\Gamma_s}\left\{1 + \frac{1}{\eta}|Y_s^N - Y_{\varrho_N(s)}^N| + \frac{1}{\eta}|Y_s^{N'} - Y_{\varrho_{N'}(s)}^{N'}|\right\}ds \\
&\le \frac{1}{4\ln\eta^{-1}}\int_{s=0}^t \left\{e^{2\Gamma_s} + \left\{1 + \frac{1}{\eta}|Y_s^N - Y_{\varrho_N(s)}^N| + \frac{1}{\eta}|Y_s^{N'} - Y_{\varrho_{N'}(s)}^{N'}|\right\}^2\right\}ds \\
&\le \frac{1}{4\ln\eta^{-1}}\int_{s=0}^t \left\{e^{2\Gamma_s} + 3 + \frac{3}{\eta^2}|Y_s^N - Y_{\varrho_N(s)}^N|^2 + \frac{3}{\eta^2}|Y_s^{N'} - Y_{\varrho_{N'}(s)}^{N'}|^2\right\}ds, \\
A_t^{2,N,N'} &= \frac{1}{2}\int_{s=0}^t \ddot{\psi}_\eta(\nu_s^{N,N'})\{((Y_{\varrho^N(s)}^N + Z_s)\vee 0) - ((Y_{\varrho^{N'}(s)}^{N'} + Z_s)\vee 0)\}^2\,ds \\
&\le \frac{1}{2}\int_{s=0}^t \ddot{\psi}_\eta(\nu_s^{N,N'})|Y_{\varrho_N(s)}^N - Y_{\varrho_{N'}(s)}^{N'}|^2\,ds \\
&\le \frac{3}{2}\int_{s=0}^t \ddot{\psi}_\eta(\nu_s^{N,N'})\{|\nu_s^{N,N'}|^2 + |Y_s^N - Y_{\varrho_N(s)}^N|^2 + |Y_s^{N'} - Y_{\varrho_{N'}(s)}^{N'}|^2\}\,ds \\
&\le \frac{3}{2\ln\eta^{-1}}\int_{s=0}^t \left\{\eta^{1/2} + \frac{1}{\eta}|Y_s^N - Y_{\varrho_N(s)}^N|^2 + \frac{1}{\eta}|Y_s^{N'} - Y_{\varrho_{N'}(s)}^{N'}|^2\right\}ds.
\end{aligned}$$

In the bound on $A^{1,N,N'}$, we have used Young's inequality, and in the bound on $A^{2,N,N'}$ we have used the fact that the support of $\ddot{\psi}_\eta$ is contained in $[0, \sqrt{\eta}]$. Collecting things together, we have that there is a $K > 0$ such that

$$\begin{aligned}
\mathbb{E}[A_t^{1,N,N'}] &\le \frac{K}{\ln\eta^{-1}}\left\{1 + \frac{1}{N\eta^2} + \frac{1}{N'\eta^2}\right\}, \\
\mathbb{E}[A_t^{2,N,N'}] &\le \frac{K}{\ln\eta^{-1}}\left\{\eta^{1/2} + \frac{1}{N\eta} + \frac{1}{N'\eta}\right\}
\end{aligned}$$

for all $t \in [0, T]$. Thus,

$$\lim_{N,N'\to\infty} \mathbb{E}[|\nu_t^{N,N'}|] \le \sqrt{\eta} + \frac{K\sigma^2}{\ln \eta^{-1}} + \frac{K\beta^2\eta^{1/2}}{\ln \eta^{-1}}$$

for all $t \in [0, T]$. Letting $\eta \searrow 0$, we indeed get that $\lim_{N,N'\to\infty} \mathbb{E}[|\nu_t^{N,N'}|] = 0$.

We thus have that

$$\overline{\lim_{N,N'\to\infty}} \mathbb{E}[|Y_t^N - Y_t^{N'}|] = 0.$$

For any $p > 1$, we also have by interpolation and (10.1) and (10.2) that

$$\begin{aligned}\overline{\lim_{N,N'\to\infty}} \mathbb{E}[|Y_t^N - Y_t^{N'}|^p] &\le \overline{\lim_{N,N'\to\infty}} \sqrt{\mathbb{E}[|Y_t^N - Y_t^{N'}|]\mathbb{E}[|Y_t^N - Y_t^{N'}|^{2p-1}]} \\ &= 0.\end{aligned}$$

Thus, there is a solution Y of the integral equation

$$Y_t = \sigma \int_{s=0}^t e^{\Gamma_s/2}\sqrt{(Y_s + Z_s) \vee 0}\, dW_s^* + \beta \int_{s=0}^t \big((Y_s + Z_s) \vee 0\big)\, dV_s$$

such that $\sup_{t\in[0,T]} \mathbb{E}[|Y_t|^p] < \infty$ for all $T > 0$ and $p \ge 1$. Setting $\bar{Y}_t \stackrel{\text{def}}{=} Z_t + Y_t$, we have that $\bar{Y}_t \in \bigcap_{p\ge1} L^p$ and that

$$\bar{Y}_t = Z_t + \sigma \int_{s=0}^t e^{\Gamma_s/2}\sqrt{\bar{Y}_s \vee 0}\, dW_s^* + \beta \int_{s=0}^t (\bar{Y}_s \vee 0)\, dV_s.$$

We claim that $\bar{Y}$ is nonnegative. For each $\eta > 0$ we have that

$$\begin{aligned}\psi_\eta(\bar{Y}_t)\chi_{\mathbb{R}-}(\bar{Y}_t) = \psi_\eta(\lambda_\circ)\chi_{\mathbb{R}-}(\lambda_\circ) &+ \frac{\sigma^2}{2}\int_{s=0}^t \ddot{\psi}_\eta(\bar{Y}_s)\chi_{\mathbb{R}-}(\bar{Y}_s)e^{\Gamma_s/2}(\bar{Y}_s \vee 0)\, ds \\ &+ \frac{\beta^2}{2}\int_{s=0}^t \ddot{\psi}_\eta(\bar{Y}_s)\chi_{\mathbb{R}-}(\bar{Y}_s)(\bar{Y}_s \vee 0)^2\, ds + \mathcal{M}_t,\end{aligned}$$

where $\mathcal{M}$ is a martingale. Taking expectations and then letting $\eta \searrow 0$, we have that $\mathbb{E}[\bar{Y}_t^-] = 0$. We finally set $\lambda_t \stackrel{\text{def}}{=} e^{-\Gamma_t}\bar{Y}_t$. The claim follows. $\square$

PROOF OF LEMMA 3.2. Let λ and λ' be two solutions of (3.1). Define $Y_t \stackrel{\text{def}}{=} \lambda_t e^{\Gamma_t} - Z_t$ and $Y_t' \stackrel{\text{def}}{=} \lambda_t' e^{\Gamma_t} - Z_t$. Since λ and λ' are assumed to be nonnegative, Y and Y' satisfy

$$\begin{aligned}Y_t &= \sigma \int_{s=0}^t e^{\Gamma_s/2}\sqrt{Y_s + Z_s}\, dW_s^* + \beta^S \int_{s=0}^t (Y_s + Z_s)\, dV_s, \\ Y_t' &= \sigma \int_{s=0}^t e^{\Gamma_s/2}\sqrt{Y_s' + Z_s}\, dW_s^* + \beta^S \int_{s=0}^t (Y_s' + Z_s)\, dV_s.\end{aligned}$$

Set $\nu_t \stackrel{\text{def}}{=} Y_t - Y'_t$. For each $\eta > 0$,

$$|\nu_t| \le \psi_\eta(\nu_t) + \sqrt{\eta} = \sigma^2 A_t^1 + (\beta^S)^2 A_t^2 + \mathcal{M}_t + \sqrt{\eta},$$

where $\mathcal{M}$ is a martingale and where

$$A_t^1 = \frac{1}{2}\int_{s=0}^t \ddot{\psi}_\eta(\nu_s) e^{\Gamma_s}\{\sqrt{Y_s + Z_s} - \sqrt{Y'_s + Z_s}\}^2\, ds \le \frac{1}{\ln \eta^{-1}}\int_{s=0}^t e^{\Gamma_s}\, ds,$$

$$A_t^2 = \frac{1}{2}\int_{s=0}^t \ddot{\psi}_\eta(\nu_s)\nu_s^2\, ds \le \frac{\eta^{1/2}}{\ln \eta^{-1}} t.$$

Collecting things together, we have that

$$\mathbb{E}[|\nu_t|] \le \sqrt{\eta} + \frac{1}{\ln \eta^{-1}}\left\{\sqrt{\eta} t + \int_{s=0}^t \mathbb{E}[e^{\Gamma_s}]\, ds\right\}.$$

Let $\eta \searrow 0$ to get that $Y = Y'$. The claim follows. $\square$

Let's next prove the needed macroscopic bound on the $\lambda^{N,n}$'s.

PROOF OF LEMMA 3.4. For each $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$, define

$$\Gamma_t^{N,n} \stackrel{\text{def}}{=} \alpha_{N,n} t + \beta_{N,n}^S \gamma \int_{s=0}^t X_s\, ds,$$

$$Z_t^{N,n} \stackrel{\text{def}}{=} \lambda_{N,n,\circ} + \alpha_{N,n}\bar{\lambda}_{N,n}\int_{s=0}^t e^{\Gamma_s^{N,n}}\, ds + \beta_{N,n}^C \int_{s=0}^t e^{\Gamma_s}\, dL_s^N$$

and let $Y^{N,n}$ satisfy the equation

$$Y_t^{N,n} = \sigma_{N,n}\int_{s=0}^t e^{\Gamma_s^{N,n}/2}\sqrt{Y_s^{N,n} + Z_s^{N,n}}\, dW_s^n + \varepsilon_N \beta_{N,n}^S \int_{s=0}^t (Y_s^{N,n} + Z_s^{N,n})\, dV_s;$$

then $\lambda_t^{N,n} = e^{\Gamma_t^{N,n}}(Y_t^{N,n} + Z_t^{N,n})$. We calculate that

$$\begin{aligned}|\lambda_t^{N,n}|^p &\le \tfrac{1}{2}\{e^{-2p\Gamma_t^{N,n}} + |Y_t^{N,n} + Z_t^{N,n}|^{2p}\}\\ &\le \tfrac{1}{2}\{e^{-2p\Gamma_t^{N,n}} + 2^{2p-1}(|Y_t^{N,n}|^{2p} + |Z_t^{N,n}|^{2p})\}.\end{aligned}$$

From Lemma 10.1, we have that

$$\sup_{\substack{0\le t\le T\\ N\in\mathbb{N}}} \frac{1}{N}\sum_{n=1}^N \mathbb{E}[e^{2p\Gamma_t^{N,n}}] \quad \text{and} \quad \sup_{\substack{0\le t\le T\\ N\in\mathbb{N}}} \frac{1}{N}\sum_{n=1}^N \mathbb{E}[|Z_t^{N,n}|^{2p}]$$

are both finite.

For each $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$, we compute that

$$\mathbb{E}[|Y_t^{N,n}|^{2p}] = p(2p-1)\Big\{\sigma_{N,n}^2 \int_{s=0}^t \mathbb{E}[|Y_s^{N,n}|^{2p-2} e^{\Gamma_s^{N,n}} |Y_s^{N,n} + Z_s^{N,n}|]\,ds + \varepsilon_N^2 (\beta_{N,n}^S)^2 \int_{s=0}^t \mathbb{E}[|Y_s^{N,n}|^{2p-2} |Y_s^{N,n} + Z_s^{N,n}|^2]\,ds\Big\}.$$

To bound the integrals, we have that

$$\begin{aligned}
&|Y_s^{N,n}|^{2p-2} e^{\Gamma_s^{N,n}} |Y_s^{N,n} + Z_s^{N,n}| \\
&\quad \le \frac{1}{2p} e^{2p\Gamma_t^{N,n}} + \frac{p-1}{p} |Y_t^{N,n}|^{2p} + \frac{1}{2p} |Y_s^{N,n} + Z_s^{N,n}|^{2p} \\
&\quad \le \frac{1}{2p} e^{2p\Gamma_t^{N,n}} + \frac{p-1}{2p} |Y_t^{N,n}|^{2p} + \frac{2^{2p-1}}{2p} \{|Y_s^{N,n}|^{2p} + |Z_s^{N,n}|^{2p}\}, \\
&|Y_s^{N,n}|^{2p-2} |Y_s^{N,n} + Z_s^{N,n}|^2 \\
&\quad \le \frac{p-1}{p} |Y_s^{N,n}|^{2p} + \frac{1}{p} |Y_s^{N,n} + Z_s^{N,n}|^{2p} \\
&\quad \le \frac{p-1}{p} |Y_s^{N,n}|^{2p} + \frac{2^{2p-1}}{p} \{|Y_s^{N,n}|^{2p} + |Z_s^{N,n}|^{2p}\}.
\end{aligned}$$

Combining things together, we have that there is a $K > 0$ such that

$$\begin{aligned}
\mathbb{E}[|Y_t^{N,n}|^{2p}] \le K\{\sigma_{N,n}^2 + \varepsilon_N^2 (\beta_{N,n}^S)^2\} &\times \Big\{\int_{s=0}^t \mathbb{E}[|Y_s^{N,n}|^{2p}]\,ds + \int_{s=0}^t \mathbb{E}[e^{2p\Gamma_s^{N,n}}]\,ds \\
&\quad + \int_{s=0}^t \mathbb{E}[|Z_s^{N,n}|^{2p}]\,ds\Big\}
\end{aligned}$$

for all $N \in \mathbb{N}$ and $n \in \{1, 2, \dots, N\}$. Using Assumption 2.3 and averaging over n, we get the claimed result. □

10.3. *Proof of Lemma* 4.1. Define a homeomorphism Φ of $C[0, \infty)$ as

$$\begin{aligned}
\Phi(q)(t) \stackrel{\text{def}}{=} \int_{\substack{\hat{\mathsf{p}} = (\mathsf{p}, \lambda) \in \hat{\mathcal{P}} \\ \mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S)}} &\beta^C \Big[\dot{b}^{\mathsf{p}}(t)\lambda + \int_{r=0}^t \dot{b}^{\mathsf{p}}(t-r)\{q(r) + \alpha\bar{\lambda}\}\,dr\Big] \\
&\times \exp\Big[-b^{\mathsf{p}}(t)\lambda - \int_{r=0}^t b^{\mathsf{p}}(t-r)\{q(r) + \alpha\bar{\lambda}\}\,dr\Big] \\
&\times \pi(d\mathsf{p})\Lambda_\circ(d\lambda)
\end{aligned}$$

for all $q \in C[0, \infty)$ and $t \ge 0$. Note that since b, q and λ are all nonnegative,

$$0 \le \exp\Big[-b^{\mathsf{p}}(t)\lambda - \int_{r=0}^t b^{\mathsf{p}}(t-r)\{q(r) + \alpha\bar{\lambda}\}\,dr\Big] \le 1.$$

We can then set up a recursion; we want to solve $Q = \Phi(Q)$. Note that there is a $K > 0$ such that

$$|\Phi(q)(t)| \le K \int_{s=0}^{t} q(r)\,dr$$

for all nonnegative $q \in C[0,\infty)$.

For any q_1 and q_2 in $C[0,\infty)$, we have that

$$\Phi(q_1)(t) - \Phi(q_2)(t) = \Gamma_t^a(q_1,q_2) + \Gamma_t^b(q_1,q_2),$$

where

$$\begin{aligned}
&\Gamma_t^a(q_1,q_2)\\
&\quad \stackrel{\text{def}}{=} \int_{s=0}^{t}\Bigg\{\int_{\theta=0}^{1}\int_{\substack{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}\\ \mathsf{p}=(\alpha,\bar{\lambda},\sigma,\beta^C,\beta^S)}} \beta^C \dot{b}^{\mathsf{p}}(t-s)\{q_1(s)-q_2(s)\}\\
&\qquad\times \exp\Bigg[-b^{\mathsf{p}}(t)\lambda\\
&\qquad\qquad - \int_{r=0}^{t} b^{\mathsf{p}}(t-r)\big[\{q_2(r)+\theta\big(q_1(r)-q_2(r)\big)\}+\alpha\bar{\lambda}\big]\,dr\Bigg]\\
&\qquad\qquad\qquad\times \pi(d\mathsf{p})\Lambda_\circ(d\lambda)\,d\theta\Bigg\}\,ds,
\end{aligned}$$

$$\begin{aligned}
&\Gamma_t^b(q_1,q_2)\\
&\quad \stackrel{\text{def}}{=} -\int_{s=0}^{t}\Bigg\{\int_{\theta=0}^{1}\int_{\substack{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}\\ \mathsf{p}=(\alpha,\bar{\lambda},\sigma,\beta^C,\beta^S)}} \beta^C\Bigg\{\dot{b}^{\mathsf{p}}(t)\lambda + \int_{r=0}^{t}\dot{b}^{\mathsf{p}}(t-r)\\
&\qquad\qquad\qquad\times\big[\{q_2(r)+\theta\big(q_1(r)-q_2(r)\big)\}+\alpha\bar{\lambda}\big]\,dr\Bigg\}\\
&\qquad\times\big\{b^{\mathsf{p}}(t-s)\big(q_1(s)-q_2(s)\big)\big\}\\
&\qquad\times\exp\Bigg[-b^{\mathsf{p}}(t)\lambda\\
&\qquad\qquad - \int_{r=0}^{t} b^{\mathsf{p}}(t-r)\big[\{q_2(r)+\theta\big(q_1(r)-q_2(r)\big)\}+\alpha\bar{\lambda}\big]\,dr\Bigg]\\
&\qquad\qquad\qquad\times \pi(d\mathsf{p})\Lambda_\circ(d\lambda)\,d\theta\Bigg\}\,ds.
\end{aligned}$$

Standard techniques from Picard iterations give us the result.

Acknowledgments. R. B. Sowers would like to thank the Departments of Mathematics and Statistics of Stanford University for their hospitality in the Spring

of 2010 during a sabbatical stay. The authors are grateful to Thomas Kurtz for some insight into the literature on mean field models, and to Michael Gordy and participants of the 3rd SIAM Conference on Financial Mathematics and Engineering in San Francisco and the 2010 Annual INFORMS Meeting in Austin for comments.

REFERENCES

[1] AZIZPOUR, S., GIESECKE, K. and SCHWENKLER, G. (2010). Exploring the sources of default clustering. Working paper, Stanford Univ.

[2] BUSH, N., HAMBLY, B. M., HAWORTH, H., JIN, L. and REISINGER, C. (2011). Stochastic evolution equations in portfolio credit modelling. *SIAM J. Financial Math.* **2** 627–664. MR2836495

[3] DAI PRA, P., RUNGGALDIER, W. J., SARTORI, E. and TOLOTTI, M. (2009). Large portfolio losses: A dynamic contagion model. *Ann. Appl. Probab.* **19** 347–394. MR2498681

[4] DAI PRA, P. and TOLOTTI, M. (2009). Heterogeneous credit portfolios and the dynamics of the aggregate losses. *Stochastic Process. Appl.* **119** 2913–2944. MR2554033

[5] DAS, S., DUFFIE, D., KAPADIA, N. and SAITA, L. (2007). Common failings: How corporate defaults are correlated. *J. Finance* **62** 93–117.

[6] DAVIS, M. H. A. and ESPARRAGOZA-RODRIGUEZ, J. C. (2007). Large portfolio credit risk modeling. *Int. J. Theor. Appl. Finance* **10** 653–678. MR2341487

[7] DAWSON, D. A. and HOCHBERG, K. J. (1982). Wandering random measures in the Fleming–Viot model. *Ann. Probab.* **10** 554–580. MR0659528

[8] DEELSTRA, G. and DELBAEN, F. (1994). Existence of solutions of stochastic differential equations related to the Bessel process. Working paper. Dept. Mathematics, ETH, Zürich, Switzerland.

[9] DEMBO, A., DEUSCHEL, J.-D. and DUFFIE, D. (2004). Large portfolio losses. *Finance Stoch.* **8** 3–16. MR2022976

[10] DUFFIE, D., PAN, J. and SINGLETON, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* **68** 1343–1376. MR1793362

[11] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York. MR0838085

[12] FLEMING, W. H. and VIOT, M. (1979). Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.* **28** 817–843. MR0542340

[13] GIESECKE, K. and WEBER, S. (2006). Credit contagion and aggregate losses. *J. Econom. Dynam. Control* **30** 741–767. MR2224986

[14] GLASSERMAN, P., KANG, W. and SHAHABUDDIN, P. (2007). Large deviations in multifactor portfolio credit risk. *Math. Finance* **17** 345–379. MR2332261

[15] IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. North-Holland, Amsterdam. MR1011252

[16] JAKŠA CVITANIĆ, J. M. and ZHANG, J. (2012). Law of large numbers for self-exciting correlated defaults. *Stochastic Process. Appl.* To appear.

[17] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. MR1121940

[18] REVUZ, D. and YOR, M. (1991). *Continuous Martingales and Brownian Motion*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Springer, Berlin. MR1083357

[19] ROYDEN, H. L. (1988). *Real Analysis*, 3rd ed. Macmillan Publishing Company, New York. MR1013117

[20] SIRCAR, R. and ZARIPHOPOULOU, T. (2010). Utility valuation of multi-name credit derivatives and application to CDOs. *Quant. Finance* **10** 195–208. MR2642963

K. GIESECKE
DEPARTMENT OF MANAGEMENT SCIENCE
AND ENGINEERING
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4026
USA
E-MAIL: giesecke@stanford.edu

K. SPILIOPOULOS
DIVISION OF APPLIED MATHEMATICS
BROWN UNIVERSITY
PROVIDENCE, RHODE ISLAND 029126
USA
E-MAIL: kspiliop@dam.brown.edu

R. B. SOWERS
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS
AT URBANA, CHAMPAIGN
URBANA, ILLINOIS 61801
USA
E-MAIL: r-sowers@illinois.edu