MATHEMATICAL FINANCE
An International Journal of Mathematics, Statistics and Financial Economics

Mathematical Finance, Vol. 25, No. 1 (January 2015), 77–114

LARGE PORTFOLIO ASYMPTOTICS FOR LOSS FROM DEFAULT

KAY GIESECKE
Stanford University

KONSTANTINOS SPILIOPOULOS
Boston University

RICHARD B. SOWERS
University of Illinois at Urbana–Champaign

JUSTIN A. SIRIGNANO
Stanford University

We prove a law of large numbers for the loss from default and use it for approximating the distribution of the loss from default in large, potentially heterogeneous portfolios. The density of the limiting measure is shown to solve a nonlinear stochastic partial differential equation, and certain moments of the limiting measure are shown to satisfy an infinite system of stochastic differential equations. The solution to this system leads to the distribution of the limiting portfolio loss, which we propose as an approximation to the loss distribution for a large portfolio. Numerical tests illustrate the accuracy of the approximation, and highlight its computational advantages over a direct Monte Carlo simulation of the original stochastic system.

KEY WORDS: law of large numbers, interacting point process, credit risk.

1. INTRODUCTION

Reduced-form point process models of correlated default timing are widely used to measure portfolio credit risk and to value securities exposed to correlated default risk. Computing the distribution of the loss from default in these models tends to be difficult, however, especially in bottom-up formulations with many names. Semi-analytical transform methods have limited scope. Monte Carlo simulation is much more broadly applicable but can be slow for the large portfolios and longer time horizons common in practice.

We are grateful to two anonymous referees and to Michael Gordy, Monique Jeanblanc, and participants of the 5th Financial Risks International Forum in Paris, the 2011 Annual INFORMS Meeting in Charlotte, the 2012 Workshop on Mathematical Finance in Mumbai, the 2011 International Congress on Industrial and Applied Mathematics in Vancouver, the Mathematics Seminars at the University of Southern California and Santa Clara University, and the SEEM Seminar at the Chinese University of Hong Kong for comments.

Manuscript received August 2011; final revision received May 2012.

Address correspondence to Konstantinos Spiliopoulos, Department of Mathematics & Statistics, Boston University, 111 Cummington Mall, Boston, MA 02215, USA; e-mail: kspiliop@math.bu.edu.

[Correction added on 24 October 2013 after first online publication on 31 October 2012. Equations affecting the key theorem have been corrected throughout the article.]

DOI: 10.1111/mafi.12011
© 2012 Wiley Periodicals, Inc.

This paper develops an approximation to the distribution of the loss from default in large portfolios that may have a heterogeneous structure. The approximation is valid for a class of reduced-form models in which a name defaults at a stochastic intensity that is influenced by an idiosyncratic risk factor process, a systematic risk factor process X common to all names in the pool, and the portfolio loss rate. It is based on a law of large numbers for the portfolio loss rate. The limiting portfolio loss is not deterministic but follows a stochastic process driven by X, indicating that the exposure to the systematic risk cannot be diversified. We show that the density of the limiting measure, if it exists, satisfies a nonlinear stochastic partial differential equation (SPDE) driven by X. We develop a numerical method for solving this equation. The method is based on the observation that certain moments of the limiting measure satisfy an infinite system of SDEs. These SDEs are driven by the systematic risk factor X; a truncated system can be solved using a discretization scheme, for example. The solution to the SDE system leads to the solution to the SPDE through an inverse moment problem. It also leads to the distribution of the limiting portfolio loss, which we propose as an approximation to the distribution of the loss from default for a large portfolio. Estimators of portfolio value at risk and other risk measures are immediate from the limiting loss distribution.

Numerical tests illustrate the accuracy and computational efficiency of the approximation for large but finite portfolios. We find a substantial reduction in computational effort over the alternative of direct Monte Carlo simulation of the high-dimensional original stochastic system. The accuracy of the approximation mainly depends on the portfolio size N and the sensitivity to the systematic risk factor. For a given sensitivity, the accuracy increases with N, as expected. The higher the sensitivity to the systematic risk, the higher the variance of the loss distribution and the more accurate is the approximation for fixed N. The approximation is remarkably accurate in the tail of the loss distribution. This renders it particularly suitable for the estimation of risk measures for the large pools of loans commonly held by banks.

Large portfolio approximations were first studied by Vasicek (1991). In Vasicek's static model of a homogeneous pool, firms default independently of one another conditional on a normally distributed random variable representing a systematic risk factor. Because the losses from defaults are conditionally i.i.d., the classical law of large numbers ensures the convergence of the portfolio loss rate to its conditional mean, from which the limiting loss distribution is immediate. Schloegl and O'Kane (2005) examine alternative distributions of the systematic factor, and Lucas et al. (2001) and Gordy (2003) study the limiting loss in a heterogeneous portfolio. Bush et al. (2011) analyze a dynamic extension of Vasicek's homogeneous pool model in which the systematic risk factor follows a Brownian motion. They obtain an SPDE driven by that Brownian motion for the density of the limiting measure, which they solve using a finite element method. The conditional independence of defaults can also be exploited to analyze the tail behavior of the losses in large, not necessarily homogeneous, portfolios using large deviations arguments; see Dembo, Deuschel, and Duffie (2004) and Glasserman, Kang, and Shahabuddin (2007).

The analysis in this paper differs from that in the aforementioned articles in several important respects. We study a class of dynamic point process models of correlated default timing in which a firm defaults at a stochastic intensity process. The intensity is influenced by an idiosyncratic risk factor process following a square-root diffusion, a systematic risk factor process following a diffusion with arbitrary coefficient functions, and the portfolio loss rate. To address the heterogeneity of a portfolio, the intensity parameters of each name are allowed to be different. The choice for dependence of the intensity on idiosyncratic and systematic risk factor processes is motivated by the

empirical findings of Duffie, Saita, and Wang (2006). The choice for dependence of the intensity on the portfolio loss is motivated by the empirical results of Azizpour, Giesecke, and Schwenkler (2010), who find that defaults have a statistically significant feedback effect on the surviving firms. The self-exciting behavior of defaults violates the conditional independence property that is exploited in the aforementioned articles. It complicates the asymptotic analysis and induces an integral term in the drift of the SPDE governing the density of the limiting measure. The exposure to the systematic risk leads to the noise term in the SPDE, which is given by an Itô integral against the Brownian motion driving the systematic risk diffusion. The solution to the SPDE governs the distribution of the limiting loss at all future horizons, facilitating the computation of the "loss surface." This dynamic perspective is absent in the static formulations considered in most of the aforementioned articles.

The law of large numbers (LLN) proved in this paper significantly extends an earlier result in Giesecke, Spiliopoulos and Sowers (2011), which assumes Ornstein-Uhlenbeck dynamics for the systematic risk factor and the sensitivity of the intensity to the systematic risk to vanish in the large-portfolio limit. The LLN developed here allows for general diffusion dynamics for the systematic risk factor. Moreover, the exposure of an intensity to this diffusion is not required to vanish, generating a much richer, nondeterministic limiting behavior governed by an SPDE rather than a PDE. The treatment of these features requires additional arguments. Our main result (Theorem 3.1) develops the stochastic evolution equation that the limiting empirical measure satisfies. If the limiting empirical measure admits a density, then an integration by parts argument shows that the density satisfies an SPDE. The filtered martingale problem is used to identify the limit and prove the LLN. A major difficulty in the identification of the limit is the solution of a coupled system of SDEs, which we address using fixed-point arguments. In contrast to Giesecke et al. (2011), this system does not easily decouple in the more general setting considered here. The analysis of the fixed-point arguments is complicated due to the square-root singularity.

Cvitanić, Ma, and Zhang (2011) prove a LLN for a related system, taking an intensity as a function of an idiosyncratic risk factor, a systematic risk factor, and the portfolio loss rate. The risk factors follow diffusion processes whose coefficients may depend on the portfolio loss. In that formulation, the impact of a default on the dynamics of the surviving firms is permanent. In our work, an intensity depends on the path of the portfolio loss. Therefore, the impact of a default on the surviving firms may be transient and fade away with time. There is a recovery effect. Other interacting particle systems with permanent default impact are analyzed by Dai Pra et al. (2009) and Dai Pra and Tolotti (2009), who take an intensity as a function of the portfolio loss rate. In a model with local interaction, Giesecke and Weber (2006) take the intensity of a name as a function of the state of the names in a specified neighborhood of that name. These papers prove LLNs for the portfolio loss and develop Gaussian approximations to the portfolio loss distribution based on central limit theorems. The interacting particle system which we propose and study includes firm-specific sources of default risk and addresses an additional source of default clustering, namely the exposure of a firm to a systematic risk factor process. This exposure generates a random limiting behavior.

There are several other related articles. Davis and Rodriguez (2007) develop large portfolio approximations based on a law of large numbers and a central limit theorem 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 (2010) examine large portfolio asymptotics for utility

indifference valuation of securities exposed to the losses in the pool. Our formulation addresses the dependence of the intensity on a systematic diffusion factor and the portfolio loss. It allows for self-exciting effects that violate the conditional independence assumption. Gaussian and large deviation approximations to the distribution of portfolio losses for an affine point process system with features similar to those of our, not necessarily affine, system are provided by Zhang et al. (2011). Their asymptotic analysis is based on a "large horizon" rather than a "large portfolio" regime considered here and in the aforementioned articles. The scope of their approximations differs from that of ours.

The rest of the paper is organized as follows. Section 2 describes a class of point process models of correlated default timing in a pool of names. Section 3 states our main result, Theorem 3.1, a law of large numbers for the loss rate in the pool. Section 4 provides further insights into the limiting behavior of the loss for the special case of a homogeneous pool. Section 5 develops, implements, and tests a moment method for the numerical solution of the SPDE. Numerical results illustrate the method and demonstrate the accuracy of the approximation of the portfolio loss by the limiting loss. Section 6 discusses the extension of our results to more general intensity dynamics. Sections 7 and 8 are devoted to the proof of Theorem 3.1. Appendices provide auxiliary results.

2. MODEL AND ASSUMPTIONS

We provide a dynamic point process model of correlated default timing in a portfolio of names. We assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is an underlying probability space on which all random variables are defined. Let $\{W^n\}_{n\in\mathbb{N}}$ be a countable collection of independent standard Brownian motions. Let $\{\mathfrak{e}_n\}_{n\in\mathbb{N}}$ be an i.i.d. collection of standard exponential random variables which are independent of the W^n's. 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 name. Each $\mathfrak{e}_n$ will represent a normalized default time for a specific name. The process V will drive a systematic risk factor process to which all names are exposed. Define $\mathcal{V}_t = \sigma(V_s, 0 \le s \le t) \vee \mathcal{N}$ and $\mathcal{F}_t = \sigma((V_s, W_s^n), 0 \le s \le t, n \in \mathbb{N}) \vee \mathcal{N}$, where $\mathcal{N}$ contains the $\mathbb{P}$-null sets.

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

(2.1)

$$
\begin{aligned}
d\lambda_t^{N,n} &= -\alpha_{N,n}\big(\lambda_t^{N,n} - \bar{\lambda}_{N,n}\big)\,dt + \sigma_{N,n}\sqrt{\lambda_t^{N,n}}\,dW_t^n + \beta_{N,n}^C\,dL_t^N + \beta_{N,n}^S\lambda_t^{N,n}\,dX_t \quad t > 0\\
\lambda_0^{N,n} &= \lambda_{\circ,N,n}\\
dX_t &= b_0(X_t)\,dt + \sigma_0(X_t)\,dV_t \quad t > 0\\
X_0 &= x_\circ\\
L_t^N &= \frac{1}{N}\sum_{n=1}^{N} \chi_{[\mathfrak{e}_n,\infty)}\left(\int_{s=0}^{t} \lambda_s^{N,n}\,ds\right).
\end{aligned}
$$

Here, χ is the indicator function. The initial condition $x_\circ$ of X is fixed. The $\alpha_{N,n}, \bar{\lambda}_{N,n}, \sigma_{N,n}, \beta_{N,n}^C, \beta_{N,n}^S$ are constant parameters, for each N and n. We discuss their meaning below. 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 $\chi_{[\epsilon_n,\infty)}(\int_{s=0}^{t} \lambda_s^{N,n}\,ds) = \chi_{\{\tau^{N,n}\le t\}}$ and consequently

$$L_t^N = \frac{1}{N}\sum_{n=1}^{N} \chi_{\{\tau^{N,n}\le t\}}. \tag{2.3}$$

The process L^N represents the loss rate in a portfolio of N names, assuming a loss given default of one unit. The process $\lambda^{N,n}$ represents the intensity, or conditional event rate, of the nth name in the pool. More precisely, $\lambda^{N,n}$ is the density of the Doob–Meyer compensator to the default indicator $\chi_{\{\tau^{N,n}\le t\}}$; see (3.1). The results in Section 3 of Giesecke et al. (2011) imply that the system (2.1) has a unique solution such that $\lambda_t^{N,n} \ge 0$ for every $N \in \mathbb{N}$, $n \in \{1, 2, \ldots, N\}$ and $t \ge 0$. Thus, the model is well posed.

The jump-diffusion intensity model (2.1) is empirically motivated. It addresses several channels of default clustering. An 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 diffusion process X. Movements in X cause correlated changes in firms' intensities and thus provide a channel for default clustering emphasized by Das et al. (2007) for corporate defaults in the United States. The sensitivity of $\lambda^{N,n}$ to changes in X is measured by the parameter $\beta_{N,n}^S \in \mathbb{R}$. The second channel for default clustering is modeled through the feedback ("contagion") term $\beta_{N,n}^C dL_t^N$. A default causes a jump of size $\frac{1}{N}\beta_{N,n}^C$ in the intensity $\lambda^{N,n}$, where $\beta_{N,n}^C \in \mathbb{R}_+ = [0,\infty)$. Due to the mean-reversion of $\lambda^{N,n}$, the impact of a default fades away with time, exponentially with rate $\alpha_{N,n} \in \mathbb{R}_+$. Azizpour et al. (2010) have found self-exciting effects of this type to be an important channel for the clustering of defaults in the United States, over and above any clustering caused by the exposure of firms to systematic risk factors. Giesecke and Schwenkler (2011) develop and analyze likelihood estimators of the parameters of point process models such as (2.1).

We allow for a heterogeneous pool; the intensity dynamics of each name can be different. We capture these different dynamics by defining the "types"

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

the $\mathsf{p}^{N,n}$'s take values in parameter space $\mathcal{P} \stackrel{\text{def}}{=} \mathbb{R}_+^4 \times \mathbb{R}$. 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.[1]

We require three main conditions. These conditions are in force throughout the paper, even though this may not always be stated explicitly. First, we assume that the types of (2.4) and the initial distributions (the $\lambda_{\circ,N,n}$'s) are sufficiently regular.

CONDITION 2.1. $\pi \stackrel{\text{def}}{=} \lim_{N\to\infty} \pi^N$ and $\Lambda_\circ \stackrel{\text{def}}{=} \lim_{N\to\infty} \Lambda_\circ^N$ *exist (in* $\mathscr{P}(\mathcal{P})$ *and* $\mathscr{P}(\mathbb{R}_+)$*, respectively).*

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

We also 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.

CONDITION 2.2. *There is a $K > 0$ such that the $\alpha_{N,n}$'s, $\bar{\lambda}_{N,n}$'s, $\sigma_{N,n}$'s, $\beta^C_{N,n}$'s, $|\beta^S_{N,n}|$'s, and $\lambda_{\circ,N,n}$'s are all bounded by K for all $N \in \mathbb{N}$ and $n \in \{1, 2, \ldots, N\}$.*

Regarding the systematic risk process X, we assume

CONDITION 2.3. *The functions b_0 and σ_0 that govern the systematic risk diffusion X are such that the corresponding SDE has a unique strong solution. Moreover, there is a function $u(x)$ such that $\sigma_0(x)u(x) = -b_0(x)$ and for every $T > 0$ we have*

$$\mathbb{E}\left[e^{\frac{1}{2}\int_0^T |u(X_s)|^2\, ds}\right] < \infty. \tag{2.5}$$

The Novikov condition (2.5) may not be necessary. Lemma 8.2 is the key step for the proof of a law of large numbers for the loss rate L^N in the system (2.1), which is stated as Theorem 3.1 below. Its proof is based on a fixed point argument and uses Girsanov's Theorem; this is where (2.5) is required.[2]

Our basic formulation significantly extends that of Giesecke et al. (2011). First, we allow the systematic risk X to follow a general diffusion process with coefficients satisfying Condition 2.3 rather than a simple Ornstein–Uhlenbeck process. Second, we no longer require the exposure to the systematic risk, $\beta^S_{N,n}$, to vanish in the limit as $N \to \infty$. This implies a richer, nondeterministic limiting behavior. The analysis of this behavior is more challenging and requires new arguments.

Section 6 discusses further extensions of our basic formulation, including stochastic position losses and more general intensity dynamics.

3. LAW OF LARGE NUMBERS

We develop a law of large numbers for the portfolio loss rate L^N in the system (2.1). To this end, we need to understand a system which contains a bit more information than the loss rate L^N. For each $N \in \mathbb{N}$ and $n \in \{1, 2, \ldots, N\}$, define

$$\mathrm{M}_t^{N,n} \stackrel{\text{def}}{=} \chi_{[0,\mathrm{e}_n)}\left(\int_{s=0}^t \lambda_s^{N,n}\, ds\right) = \chi_{\{\tau^{N,n} > t\}} \tag{3.1}$$

(where $\tau^{N,n}$ is as in (2.2)). In other words, $\mathrm{M}_t^{N,n} = 1$ if and only if the nth name 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

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

[2] If condition (2.5) is required to hold only for some $T > 0$, then the statement of Theorem 3.1 below will hold for $t \in [0, T]$ instead of $t > 0$.

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 $n \in \{1, 2, \ldots, N\}$ and $t \geq 0$. 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 intensity for those assets which are still "alive." We note that

$$L_t^N = 1 - \mu_t^N(\hat{\mathcal{P}}), \quad t \geq 0.$$

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 $\hat{\mathcal{P}}$; i.e., 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 Royden 1988, ch. 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

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

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. Thus, μ^N is an element[3] of $D_E[0, \infty)$.

The main result of this paper is Theorem 3.1, essentially a law of large numbers for μ_t^N as $N \uparrow \infty$. 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

$$\begin{aligned}
(\mathcal{L}_1 f)(\hat{\mathsf{p}}) &= \frac{1}{2}\sigma^2\lambda \frac{\partial^2 f}{\partial \lambda^2}(\hat{\mathsf{p}}) - \alpha(\lambda - \bar{\lambda})\frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}) - \lambda f(\hat{\mathsf{p}}) \\
(\mathcal{L}_2 f)(\hat{\mathsf{p}}) &= \frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}).\beta^c \\
(\mathcal{L}_3^x f)(\hat{\mathsf{p}}) &= \beta^S \lambda b_0(x) \frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}) + \frac{1}{2}(\beta^S)^2 \lambda^2 \sigma_0^2(x) \frac{\partial^2 f}{\partial \lambda^2}(\hat{\mathsf{p}}) \\
(\mathcal{L}_4^x f)(\hat{\mathsf{p}}) &= \beta^S \lambda \sigma_0(x) \frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}).
\end{aligned} \tag{3.2}$$

Also define

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

[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 Ethier and Kurtz (1986).

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. Operators $\mathcal{L}_3^x$ and $\mathcal{L}_4^x$ are related to the exogenous systematic risk X.

For every $f \in C^\infty(\hat{\mathcal{P}})$ and $\mu \in E$, define

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

THEOREM 3.1. *We have that $\mu_\cdot^N$ converges in distribution to $\bar{\mu}_\cdot$ in $D_E[0, T]$. The evolution of $\bar{\mu}_\cdot$ is given by the measure evolution equation*

$$d\langle f, \bar{\mu}_t \rangle_E = \left\{ \langle \mathcal{L}_1 f, \bar{\mu}_t \rangle_E + \langle \mathcal{Q}, \bar{\mu}_t \rangle_E \langle \mathcal{L}_2 f, \bar{\mu}_t \rangle_E + \left\langle \mathcal{L}_3^{X_t} f, \bar{\mu}_t \right\rangle_E \right\} dt + \left\langle \mathcal{L}_4^{X_t} f, \bar{\mu}_t \right\rangle_E dV_t, \quad \forall f \in C^\infty(\hat{\mathcal{P}}) \text{ a.s.}$$

Suppose there is a solution of the nonlinear SPDE

$$d\upsilon(t, \hat{\mathsf{p}}) = \left\{ \mathcal{L}_1^* \upsilon(t, \hat{\mathsf{p}}) + \mathcal{L}_3^{*, X_t} \upsilon(t, \hat{\mathsf{p}}) + \left(\int_{\hat{\mathsf{p}}' \in \hat{\mathcal{P}}} \mathcal{Q}(\hat{\mathsf{p}}') \upsilon(t, \hat{\mathsf{p}}') \, d\hat{\mathsf{p}}' \right) \mathcal{L}_2^* \upsilon(t, \hat{\mathsf{p}}) \right\} dt + \mathcal{L}_4^{*, X_t} \upsilon(t, \hat{\mathsf{p}}) dV_t, \quad t > 0, \quad \hat{\mathsf{p}} \in \hat{\mathcal{P}}, \tag{3.3}$$

where $\mathcal{L}_i^*$ denote adjoint operators, with initial condition

$$\lim_{t \searrow 0} \upsilon(t, \hat{\mathsf{p}}) \, d\hat{\mathsf{p}} = \pi \times \Lambda_\circ.$$

Then

$$\bar{\mu}_t = \upsilon(t, \hat{\mathsf{p}}) \, d\hat{\mathsf{p}}.$$

REMARK 3.2. The SPDE (3.3) should be supplied with appropriate boundary conditions. In Section 4.2, we will justify the conditions

$$\upsilon(t, \lambda = 0, \mathsf{p}) = \upsilon(t, \lambda = \infty, \mathsf{p}) = 0.$$

The proof of Theorem 3.1 is given in Sections 7 and 8. Lemma 8.4 provides an alternative characterization of the limit $\bar{\mu}$. Auxiliary results are given in the appendices.

REMARK 3.3. Equation (3.3) is a stochastic partial integral differential equation (SPIDE) in the half line that degenerates at the boundary $\lambda = 0$. Due to Lemma 8.2, (3.3) can be viewed as a linear SPDE in the half line that degenerates at the boundary. Indeed, by Lemma 8.2, there is a unique pair $\{(Q(t), \lambda_t(\hat{\mathsf{p}})) : t \geq 0\}$ taking values in $\mathbb{R}_+ \times \mathbb{R}_+$ satisfying the coupled system (8.3)–(8.4). Noting that $Q(t) = \langle \mathcal{Q}, \bar{\mu}_t \rangle_E$, we see that the SPDE (3.3) can be written as

$$d\upsilon(t, \hat{\mathsf{p}}) = \left\{ \mathcal{L}_1^* \upsilon(t, \hat{\mathsf{p}}) + \mathcal{L}_3^{*, X_t} \upsilon(t, \hat{\mathsf{p}}) + Q(t) \mathcal{L}_2^* \upsilon(t, \hat{\mathsf{p}}) \right\} dt + \mathcal{L}_4^{*, X_t} \upsilon(t, \hat{\mathsf{p}}) \, dV_t, \quad t > 0, \quad \hat{\mathsf{p}} \in \hat{\mathcal{P}}. \tag{3.4}$$

Notice also that by Remark 4.1, $Q(t)$ is bounded for all $t \in \mathbb{R}_+$. Linear SPDEs in the half line that degenerate at the boundary are treated, under alternative assumptions, by Krylov and Lototsky (1998), Kim (2009, 2008), and Lototsky (2001).

4. HOMOGENEOUS POOL

We develop further insights into the SPDE governing the limit density (if it exists) in the case that the portfolio is homogeneous. Let $\hat{\mathsf{p}} = (\mathsf{p}, \lambda)$ where $\mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S) \in \mathcal{P}$. For a homogeneous pool, $\hat{\mathsf{p}}^{N,n} = \hat{\mathsf{p}}$ for all $N \in \mathbb{N}$ and $n \in \{1, 2, \ldots, N\}$. In this case, we write $\upsilon(t, \lambda)$ for the solution of the SPDE (3.3), suppressing the dependence on the fixed $\hat{\mathsf{p}}$.

4.1. Limiting Portfolio Loss

The SPDE takes the form

$$
\begin{aligned}
d\upsilon(t,\lambda) &= \left\{ \mathcal{L}_1^* \upsilon(t,\lambda) + \mathcal{L}_3^{*,X_t} \upsilon(t,\lambda) + \left(\int_0^\infty \lambda \upsilon(t,\lambda)\, d\lambda \right) \mathcal{L}_2^* \upsilon(t,\lambda) \right\} dt \\
&\quad + \mathcal{L}_4^{*,X_t} \upsilon(t,\lambda) dV_t, \; t, \lambda > 0 \\
\upsilon(0,\lambda) &= \Lambda_\circ(\lambda), \\
\upsilon(t,0) &= \lim_{\lambda \nearrow \infty} \upsilon(t,\lambda) = 0,
\end{aligned} \tag{4.1}
$$

where the adjoint operators are given by

$$
\begin{aligned}
\mathcal{L}_1^* \upsilon(t,\lambda) &= \frac{\partial^2}{\partial \lambda^2} \left(\frac{1}{2} \sigma^2 \lambda \upsilon(t,\lambda) \right) + \frac{\partial}{\partial \lambda} (\alpha(\lambda - \bar{\lambda}) \upsilon(t,\lambda)) - \lambda \upsilon(t,\lambda) \\
\mathcal{L}_2^* \upsilon(t,\lambda) &= -\frac{\partial \upsilon(t,\lambda)}{\partial \lambda} . \beta^c \\
\mathcal{L}_3^{*,x} \upsilon(t,\lambda) &= \frac{\partial^2}{\partial \lambda^2} \left(\frac{1}{2} (\beta^S)^2 \lambda^2 \sigma_0^2(x) \upsilon(t,\lambda) \right) - \frac{\partial}{\partial \lambda} (\beta^S \lambda b_0(x) \upsilon(t,\lambda)) \\
\mathcal{L}_4^{*,x} \upsilon(t,\lambda) &= -\beta^S \sigma_0(x) \frac{\partial}{\partial \lambda} (\lambda \upsilon(t,\lambda)) .
\end{aligned} \tag{4.2}
$$

Define the *limiting portfolio loss* L by

$$
L_t \stackrel{\text{def}}{=} 1 - \int_0^\infty \upsilon(t,\lambda)\, d\lambda, \quad t \geq 0; \tag{4.3}
$$

this is a random quantity since $\upsilon(t, \lambda)$ depends on the systematic risk X_t. For large N, Theorem 3.1 suggests the "large-portfolio approximation"

$$
L_t^N \approx L_t, \quad t \geq 0. \tag{4.4}
$$

4.2. Justification of Boundary Conditions

Note that $\int_0^\infty \upsilon(t,\lambda)\, d\lambda \leq 1$ and $\upsilon(t, \lambda) \geq 0$ for all $t \geq 0$ and $\lambda \in \mathbb{R}^+$. Assuming that $\upsilon(t, \lambda)$ is continuous in λ, this gives $\lim_{\lambda \to \infty} \upsilon(t, \lambda) = 0$, for otherwise the integral $\int_0^\infty \upsilon(t,\lambda)\, d\lambda$ diverges.

The boundary condition of $\upsilon(t, \lambda = 0)$ is implied by the intensities $\lambda_t^{N,n}$ being positive almost surely. For the deterministic case of $\beta^S = 0$, it is sufficient to stipulate Feller's condition of $\alpha\bar{\lambda} > \frac{1}{2}\sigma^2$, see Feller (1951). For the case of $\beta^S > 0$, let us assume that

$\upsilon(t, \lambda = 0) < \infty$. Then, Feller's condition is again sufficient to imply $\upsilon(t, \lambda = 0) = 0$. Let the flux be $f(t)$ where $f(t) \stackrel{\text{def}}{=} \frac{\partial}{\partial t} \int_0^\infty \upsilon(t, \lambda)\, d\lambda = -\int_0^\infty \lambda \upsilon(t, \lambda)\, d\lambda$. This follows from the fact that the $\lambda_t^{N,n}$ stay nonnegative almost surely by Lemma 3.1 in Giesecke et al. (2011) and therefore, according to the empirical measure μ_t^N, only leave $[0, \infty)$ by defaulting. In the asymptotic case described by the SPDE, defaults occur via the sink term $-\lambda\upsilon$. Then, probability mass only leaves $[0, \infty)$ via the sink term and does not flow across the boundaries at $\lambda = 0$ and $\lambda = \infty$. Integrating (4.1) over $\mathbb{R}^+$ and using the aforementioned flux condition along with the boundary condition at $\lambda = \infty$, we have that

$$0 = \alpha\bar{\lambda}\upsilon(t, \lambda = 0) - \frac{1}{2}\sigma^2\upsilon(t, \lambda = 0) + \beta^C \left(\int_0^\infty \lambda \upsilon(t, \lambda)\, d\lambda \right) \upsilon(t, \lambda = 0),$$

which is only satisfied by $\upsilon(t, \lambda = 0) = 0$.

This discussion provides a justification for the choice of the boundary conditions for a homogeneous pool. The treatment of a heterogeneous pool is analogous.

4.3. Alternative Representation of Limiting Loss

As we shall see below, if the density $\upsilon(t, \lambda)$ has sufficiently fast decay at $\lambda = \infty$ and at $\lambda = 0$, then one can justify an alternative representation of the limiting loss (4.3). By integration by parts, we have

(4.5)

$$\begin{aligned}
L_t &= 1 - \int_0^\infty \upsilon(t, \lambda)\, d\lambda \\
&= 1 - \Bigg[\int_0^\infty \upsilon(0, \lambda)\, d\lambda + \int_0^t \left[\alpha(\lambda - \bar{\lambda})\upsilon(s, \lambda) - \beta^S \lambda b_0(X_s)\upsilon(s, \lambda) \right]_{\lambda=0}^{\lambda=\infty} ds \\
&\quad + \int_0^t \left[\frac{1}{2}\sigma^2 \lambda \upsilon_\lambda(s, \lambda) + \frac{1}{2}\sigma^2 \upsilon(s, \lambda) + \frac{1}{2}(\beta^S)^2 \lambda^2 \sigma_0^2 \upsilon_\lambda(s, \lambda) + (\beta^S)^2 \lambda \sigma_0^2 \upsilon(s, \lambda) \right]_{\lambda=0}^{\lambda=\infty} ds \\
&\quad - \int_0^t \int_0^\infty \lambda \upsilon(s, \lambda)\, d\lambda\, ds - \beta^C \int_0^t \langle \iota, \upsilon(s, \cdot) \rangle \left[\upsilon(s, \lambda) \right]_{\lambda=0}^{\lambda=\infty} ds \\
&\quad - \int_0^t \left[\beta^S \lambda \sigma_0(X_s) \upsilon(s, \lambda) \right]_{\lambda=0}^{\lambda=\infty} dV_s \Bigg].
\end{aligned}$$

Now, observe that

$$\mu_0^N(\hat{\mathcal{P}}) = \frac{1}{N} \sum_{n=1}^N \delta_{\hat{\mathsf{p}}_0^{N,n}}(\hat{\mathcal{P}}) \mathrm{M}_0^{N,n} = \frac{1}{N} \sum_{n=1}^N \delta_{\hat{\mathsf{p}}_0^{N,n}}(\hat{\mathcal{P}}) \chi_{\{\tau^{N,n} > 0\}} = 1.$$

This implies that

$$\int_0^\infty \upsilon(0, \lambda)\, d\lambda = 1. \tag{4.6}$$

Recall the boundary conditions $\upsilon(t, 0) = \lim_{\lambda \nearrow \infty} \upsilon(t, \lambda) = 0$. Then, if we assume that for any $t \in \mathbb{R}_+$, $\upsilon(t, \cdot)$ and $\upsilon_\lambda(t, \cdot)$ decay fast at infinity, in the sense that $\lim_{\lambda \uparrow \infty} \lambda^2 \upsilon_\lambda(t, \lambda) =$

$\lim_{\lambda\uparrow\infty}\lambda\upsilon(t,\lambda)=0$, the integration by parts formula (4.5) and equation (4.6) imply

$$L_t = 1 - \int_0^\infty \upsilon(t,\lambda)\,d\lambda = \int_0^t \int_0^\infty \lambda\upsilon(s,\lambda)\,d\lambda\,ds. \tag{4.7}$$

A particularly interesting consequence of (4.7) is summarized in the following remark.

REMARK 4.1. Relation (4.7) implies that the integral term in (4.1) is bounded. Indeed, the term is $Q(t)=\int_0^\infty \lambda\upsilon(t,\lambda)\,d\lambda$. Then, (4.7) can be rewritten as

$$\int_0^\infty \upsilon(t,\lambda)\,d\lambda + \int_0^t Q(s)\,ds = 1. \tag{4.8}$$

Since $\int_0^\infty \upsilon(t,\lambda)\,d\lambda \le 1$, relation (4.8) implies that for any $t\in\mathbb{R}_+$ we have $Q(t)<\infty$. It is easy to see that the corresponding conclusion also holds for the heterogeneous pool due to Condition 2.2.

Moreover, (4.7) implies for the rate of change

$$\dot{L}_t = -\int_0^\infty \upsilon_t(t,\lambda)\,d\lambda = \int_0^\infty \lambda\upsilon(t,\lambda)\,d\lambda = \int_0^\infty \lambda\bar{\mu}_t(d\lambda).$$

Finally, we mention that one can view $(L_t,\upsilon(t,\lambda))$ as a pair satisfying

$$\begin{aligned} d\upsilon(t,\lambda) &= \left\{\mathcal{L}_1^*\upsilon(t,\lambda) + \mathcal{L}_3^{*,X_t}\upsilon(t,\lambda)\right\}dt + \mathcal{L}_2^*\upsilon(t,\lambda)\,dL_t + \mathcal{L}_4^{*,X_t}\upsilon(t,\lambda)\,dV_t \quad t,\lambda>0\\ \upsilon(0,\lambda) &= \Lambda_\circ(\lambda)\\ L_t &= 1-\int_0^\infty \upsilon(t,\lambda)\,d\lambda = \int_0^t\int_0^\infty \lambda\upsilon(s,\lambda)\,d\lambda\,ds. \end{aligned}$$

5. NUMERICAL METHOD AND RESULTS

We develop and implement a method for the numerical solution of the SPDE (4.1) governing the limiting portfolio loss (4.3) in a homogeneous pool. We obtain the distribution of the limiting loss L_t, which we propose as an approximation to the loss L_t^N when N is large. Numerical results illustrate the method, as well as the accuracy and computational efficiency of the approximation.

5.1. Numerical Approaches in the Deterministic Case

In the case of $\beta^S=0$, (4.1) becomes a deterministic, quasi-linear PDE. For completeness, we outline some numerical methods for this case.

5.1.1. *No Feedback*. If in addition $\beta^C=0$, the default times $\tau^{N,n}$ are independent and we have the analytic solution $L_t=1-\int_0^\infty \exp(A(t)+B(t)\lambda)\Lambda_\circ(\lambda)\,d\lambda$ where $A(t)$

and $B(t)$ satisfy

$$B(t) = \frac{1}{\sigma^2}\left(\alpha + \gamma \tanh\left(-\frac{1}{2}\gamma t + c_1\right)\right)$$

$$A(t) = -c_2 d_2 t + \frac{2d_1 d_2}{\gamma} \log\left(\frac{\cosh\left(-\frac{1}{2}\gamma t + c_1\right)}{\cosh(c_1)}\right),$$

where $\gamma = \sqrt{\kappa^2 + 2\sigma^2}$, $c_1 = \tanh^{-1}(\frac{-\alpha}{\gamma})$, $c_2 = \frac{\alpha}{\sigma^2}$, $d_1 = \frac{\gamma}{\sigma^2}$, and $d_2 = -\alpha\bar{\lambda}$.

5.1.2. *Finite Difference*. For the deterministic case where $\beta^C > 0$, a finite difference scheme can efficiently solve the PDE. We devise a scheme which is implicit in the differential operators and explicit in the integral operator. A predictor–corrector iteration is employed to increase accuracy for the integral term. This finite difference scheme is second-order accurate. Let Δ be the time-step for the scheme. Also, denote $\upsilon_j = \upsilon(j\Delta, \lambda)$ and let $\upsilon_{j+\frac{m}{k}}$ for $m = 1, 2, \ldots, k-1$ be predictor steps. Formally,

$$\frac{\upsilon_{j+\frac{1}{k}} - \upsilon_j}{\Delta} = \mathcal{L}\left[\frac{1}{2}\left(\upsilon_{j+\frac{1}{k}} + \upsilon_j\right)\right] + \mathcal{I}[\upsilon_j]\mathcal{L}_2^*\left[\frac{1}{2}\left(\upsilon_{j+\frac{1}{k}} + \upsilon_j\right)\right],$$

$$\frac{\upsilon_{j+\frac{2}{k}} - \upsilon_j}{\Delta} = \mathcal{L}\left[\frac{1}{2}\left(\upsilon_{j+\frac{2}{k}} + \upsilon_j\right)\right] + \mathcal{I}\left[\frac{1}{2}\left(\upsilon_{j+\frac{1}{k}} + v_j\right)\right]\mathcal{L}_2^*\left[\frac{1}{2}\left(\upsilon_{j+\frac{2}{k}} + \upsilon_j\right)\right],$$

$$\vdots$$

$$\frac{\upsilon_{j+1} - \upsilon_j}{\Delta} = \mathcal{L}\left[\frac{1}{2}(\upsilon_{j+1} + \upsilon_j)\right] + \mathcal{I}\left[\frac{1}{2}\left(\upsilon_{j+\frac{k-1}{k}} + \upsilon_j\right)\right]\mathcal{L}_2^*\left[\frac{1}{2}(\upsilon_{j+1} + \upsilon_j)\right],$$

where $\mathcal{L} = \mathcal{L}_1^* + \mathcal{L}_3^*$ and $\mathcal{I}[\upsilon(t, \lambda)] = \beta^C \int_0^\infty \lambda \upsilon(t, \lambda)\, d\lambda$.

5.2. Method of Moments

We provide a method for the numerical solution of the SPDE (4.1) that applies in the case that $\beta^S \geq 0$. Suppose that the boundary conditions for the SPDE are $\upsilon(t, \lambda = 0) = 0$ and $\lim_{\lambda\to\infty}\upsilon(t, \lambda) = 0$, as justified in Section 4.2 above. (Note that the latter boundary condition also implies $\lim_{\lambda\to\infty}\upsilon_\lambda(t, \lambda) = 0$.) Furthermore, suppose that for each $k \in \mathbb{N}$, $\lambda^k\upsilon(t, \lambda)$ is integrable on $\mathbb{R}^+$, almost surely. A sufficient condition is that the solution $\upsilon(t, \lambda)$ decays exponentially in λ; that is, there exist constants C_1, $C_2 > 0$ such that $\upsilon(t, \lambda) < C_1 e^{-C_2\lambda}$ almost surely for $t \geq 0$, $\lambda \in \mathbb{R}^+$. (We note that it was shown in Remark 4.1 that u_0 and u_1 exist.) Then, the moments $u_k(t) = \int_0^\infty \lambda^k \upsilon(t, \lambda)\, d\lambda$ exist almost surely. They follow the SDE system

(5.1)

$$du_k(t) = \big\{u_k(t)\big(-\alpha k + \beta^S b_0(X_t)k + 0.5(\beta^S)^2\sigma_0^2(X_t)k(k-1)\big)$$
$$+ u_{k-1}(t)\big(0.5\sigma^2 k(k-1) + \alpha\bar{\lambda}k + \beta^C k u_1(t)\big) - u_{k+1}(t)\big\}dt + \beta^S\sigma_0(X_t)k u_k(t)\, dV_t,$$

$$u_k(0) = \int_0^\infty \lambda^k \Lambda_\circ(\lambda)\, d\lambda.$$

To find $u_k(t)$, multiply (4.1) by λ^k and integrate by parts over $[0, \infty)$. Also, use the boundary conditions at $\lambda = 0$ and $\lambda = \infty$. Note that the limiting loss $L_t = 1 - u_0(t)$. A standard discretization method can be used to numerically solve the system (5.1).[4]

Moment methods have previously been applied to deterministic PDEs such as the Boltzmann equation, see Vincenti and Kruger (1965), for example. The moment SDE system in our case is not closed since the kth equation introduces the $(k+1)$-th moment. So, in practice one must perform a truncation at some level $k = K$ where we let $u_{K+1} = u_K$ (that is, we use the first $K+1$ moments). In the asymptotic time limit of the case where $\beta^S = \beta^C = 0$, the sensitivity of u_0 to u_{K+1} is of the order $1/K!$. The sensitivity in the more general case $\beta^S > 0$, $\beta^C > 0$ is more difficult to analyze. Numerical results, reported below, indicate that the convergence in terms of the number of moments is very rapid.

We remark that β^C plays a pivotal role in the truncated version of system (5.1). If $\beta^C = 0$ and appropriate choices are made for the coefficient functions $b_0(\cdot)$ and $\sigma_0(\cdot)$ of the systematic risk, the truncated system satisfies the global Lipschitz condition and we have the standard existence and uniqueness results. If $\beta^C > 0$, the truncated system is only locally Lipschitz and therefore there exists a unique solution up to a stopping time $\zeta(u_0(0), \ldots, u_K(0), \omega) : \mathbb{R}^{K+1} \times \Omega \to [0, \infty)$.

In addition, the moments can be inverted to yield $\upsilon(t, \lambda)$. However, although a distribution uniquely determines its moments, the converse may not be true. See Ang et al. (2002) for some numerical methods for moment inversion.

There is an alternative approach to viewing the moment system. The system (5.1) is driven by a single diffusion, which suggests there should be a canonical form where only one SDE has a diffusion term and the other SDEs only have drift terms. Define $\eta_k(t) = X_t - \frac{1}{k\beta^S}\log(u_k(t))$ for $k \geq 1$ and $\beta^S > 0$. Then

$$
\begin{aligned}
du_0(t) &= -e^{\beta^S(X_t - \eta_1(t))}\,dt,\\
d\eta_k(t) &= \Big\{ b_0(X_t) + 0.5k\beta^S\sigma_0^2(X_t) - \frac{1}{\beta^S}\big[-\alpha + \beta^S b_0(X_t) + 0.5(\beta^S)^2\sigma_0^2(X_t)(k-1)\\
&\quad + e^{(k-1)\beta^S(X_t - \eta_{k-1}(t)) - k\beta^S(X_t - \eta_k(t))}\big(0.5\sigma^2(k-1) + \alpha\bar{\lambda} + \beta^C e^{\beta^S(X_t - \eta_1(t))}\big)\\
&\quad - k^{-1}e^{(k+1)\beta^S(X_t - \eta_{k+1}(t)) - k\beta^S(X_t - \eta_k(t))}\big]\Big\}\,dt.
\end{aligned}
$$

The moment u_0 and the "canonical" moments η_k solve a system of random ODEs. They depend on the path of the systematic risk X. A skeleton of X can be generated exactly (without discretization bias) using the methods of Beskos and Roberts (2005), Chen (2013), or Giesecke and Smelov (2013).

[4] Since the solution υ of the SPDE is nonnegative, the moments should also be nonnegative. However, due to the time discretization, the moments may become negative when simulated. This can cause instability in the numerical scheme. To avoid these problems, one could immediately set a moment to zero if it ever goes negative. In particular, instability in the higher moments may occur for large β^S due to the exponential growth term $\frac{1}{2}(\beta^S)^2\sigma_0^2(X_t)k(k-1)u_k(t)dt$. To reduce this instability, one could solve the transformed moments $w_k(t) = \exp(-\frac{1}{2}(\beta_S)^2k(k-1)\int_0^t \sigma_0^2(X_s)ds)$. The SDEs for $\{w_k\}_{k=0}^K$ will not have the exponential growth term anymore. Note also that $w_0(t) = u_0(t)$.

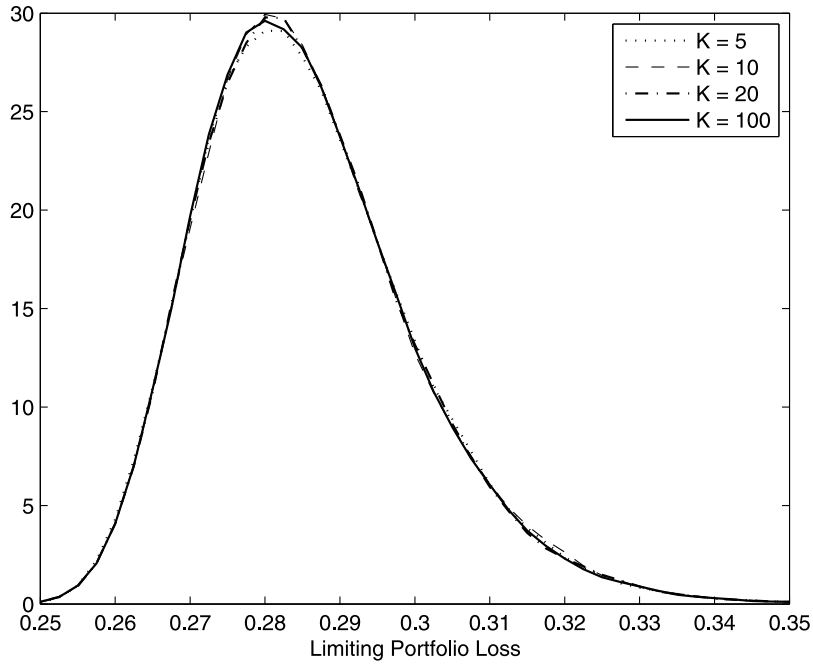


FIGURE 5.1. Comparison of distribution of limiting portfolio loss L_t for different truncation levels K at $t = 1$. The parameter case is $\sigma = .9$, $\alpha = 4$, $\bar{\lambda} = .2$, $\lambda_0 = .2$, $\beta^C = 2$, and $\beta^S = 2$. A total of 50,000 Monte Carlo trials were used.

5.3. Behavior of Limiting Loss Distribution

We provide some numerical results. Here and below, we choose the systematic risk process to be the CIR process $dX_t = \kappa(\theta - X_t)\,dt + \epsilon\sqrt{X_t}dV_t$ where $\kappa = 4$, $\theta = .5$, $\epsilon = .5$, and $X_0 = .5$. We choose the initial condition $\Lambda_\circ(\lambda) = \delta(\lambda - \lambda_0)$. The moment system (5.1) is solved using an Euler scheme with time-step of .01. Figure 5.1 shows the rapid convergence of the moment system solution. Even using as few as six moments ($K = 5$), one can achieve a very accurate distribution for the limiting loss L_t. Also, it is noteworthy that in the deterministic case of $\beta^S = 0$, the moment method is faster for the same accuracy than the finite difference approach outlined in Section 5.1.2.

We report some salient features of the limiting loss distribution. Figure 5.2 shows the effect of the feedback sensitivity parameter β^C on the distribution of L_t. As β^C increases, the mean of the losses increases and a heavy tail develops on the right (indicating a greater probability of extreme losses). Larger β^C also causes a wider or more spread-out distribution, indicating a higher variance. An important ramification is that greater connectivity between firms, modeled here through a nonlinear term, can increase the volatility of their ensemble behavior. Similarly, increasing the systematic risk sensitivity parameter β^S causes heavy tails on the right, see Figure 5.3. A hypothesis is that the initial losses can be sparked by the systematic risk factor (whose influence is determined by the parameter β^S) and then are later magnified by the contagion risk factor (determined by the parameter β^C). This has significant economic implications for the spread of risk through macro credit markets. The joint presence and interaction of systematic and contagion risk greatly magnifies the likelihood of extreme default events.

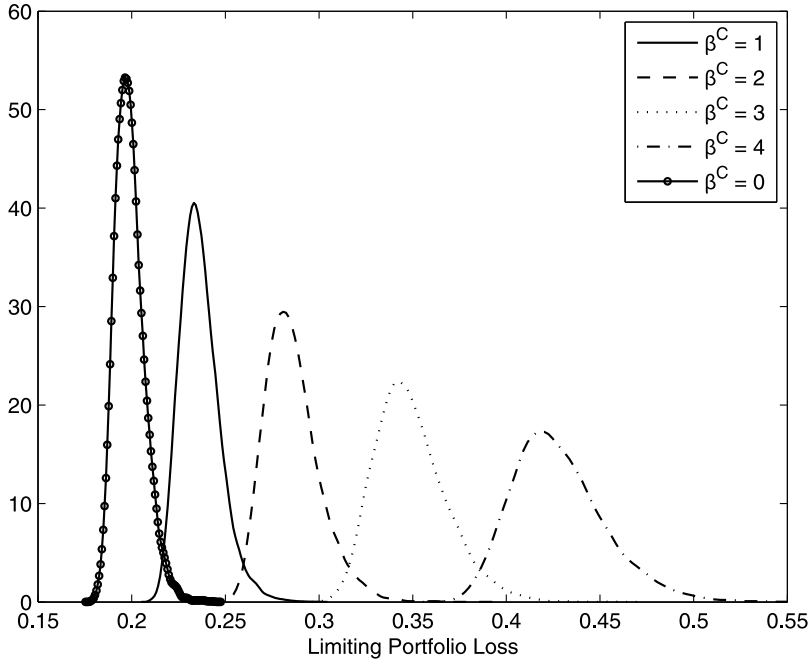


FIGURE 5.2. Comparison of distribution of limiting portfolio loss L_t for different values of the contagion sensitivity β^C at $t = 1$. The parameter case is $\sigma = .9$, $\alpha = 4$, $\bar{\lambda} = .2$, $\lambda_0 = .2$, and $\beta^S = 2$. A total of 15,000 Monte Carlo trials and 16 moments ($K = 15$) were used.

To shed more light on this issue, we calculate the Spearman correlation between X_t and L_t. Figure 5.4 shows that the correlation over a short time increases with the parameter β^C while holding β^S fixed. This indicates that the larger the exposure of firms in the pool to contagion effects, the more susceptible the system is to shocks from the systematic risk. This relationship is demonstrated in Figure 5.2 by the loss distribution's heavy right tails for large β^C. Our finding quantifies the central feature of the model: the complex interaction between systematic risk and contagion. The system becomes increasingly vulnerable to stresses from the systematic risk as the contagion channel in the system becomes stronger.

Another interesting observation is that the loss distribution can demonstrate nonmonotonic behavior in time; for example, the distribution of losses may widen and then later tighten. Figure 5.5 shows the evolution of the distribution of L_t over time t, demonstrating this nonmonotonicity. It is noteworthy that our numerical method yields the limiting loss distribution for all horizons simultaneously; this is useful for some applications, including the analysis of portfolio risk measures such as value at risk.

5.4. Accuracy of Large-Portfolio Approximation

We analyze the accuracy of the approximation (4.4). To this end, we estimate the distribution of the loss L_t^N in a pool of N names by Monte Carlo simulation of the default times $\tau^{N,n}$. The simulation uses a time-scaling method, which is based on a discretization of the intensities $\lambda^{N,n}$ on a time grid $\{t_j\}_{j=1}^J$ where $t_j = j\Delta$ for some $\Delta > 0$.

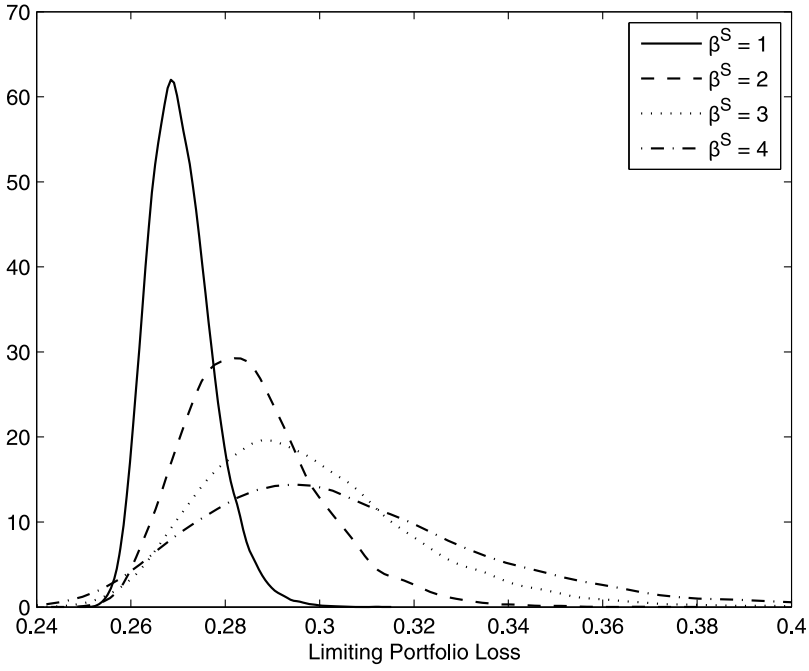


FIGURE 5.3. Comparison of distribution of limiting portfolio loss L_t for different values of the systematic risk sensitivity β^S at $t = 1$. The parameter case is $\sigma = .9$, $\alpha = 4$, $\bar{\lambda} = .2$, $\lambda_0 = .2$, and $\beta^C = 2$. A total of 15,000 Monte Carlo trials and 16 moments ($K = 15$) were used.

Intensities are simulated on $[t_j, t_{j+1})$ using a truncated Euler scheme so that they remain nonnegative. Firm defaults (as well as jumps in intensity due to defaults) occur at the grid points according to a discretized version of (2.2):

- Generate independent $\mathfrak{e}_n \sim \text{Exp}(1)$ for each firm $n = 1, 2, \ldots, N$,
- For $j = 0, 1, 2, \ldots, J-1$:
  (1) $\tilde{\lambda}^{N,n}_{t_{j+1}} = \max\left[0, \alpha(\bar{\lambda} - \lambda^{N,n}_{t_j})\Delta + \sigma(\lambda^{N,n}_{t_j})^{1/2}\sqrt{\Delta}\mathcal{N}(0,1) + \beta^S\lambda^{N,n}_{t_j}(X_{t_{j+1}} - X_{t_j})\right]$,
  (2) If $\tau^{N,n} > t_j$ and $\Delta(\sum_{i=1}^{j}\lambda^{N,n}_{t_i} + \tilde{\lambda}^{N,n}_{t_{j+1}}) \geq \mathfrak{e}_n$, then $\tau^{N,n} = t_{j+1}$,
  (3) $\lambda^{N,n}_{t_{j+1}} = \tilde{\lambda}^{N,n}_{t_{j+1}} + \beta^C\frac{1}{N}\sum_{n=1}^{N}\chi_{\{\tau^{N,n}=t_{j+1}\}}$,
- end.

The increments $X_{t_{j+1}} - X_{t_j}$ can be simulated using an Euler scheme similar to the one described above or by an exact scheme. For the results presented here, we again choose X to be a CIR process with the same parameters as stated earlier. The CIR process X is simulated using an Euler scheme (truncated at zero as shown above). We choose a time-step of $\Delta = .01$.

Convergence of the distribution of L^N_t to that of the limiting loss L_t tends be more rapid when β^S is larger. When the variance of the losses is very small (i.e., the limiting losses are close to a deterministic solution), the convergence rate is slower. Convergence is most rapid in the tails of the loss distribution. Figures 5.6 and 5.7 show the convergence of the distribution of L^N_t to that of L_t for two different parameter cases. Convergence is relatively slow, but does indicate that the asymptotic solution is applicable for a portfolio

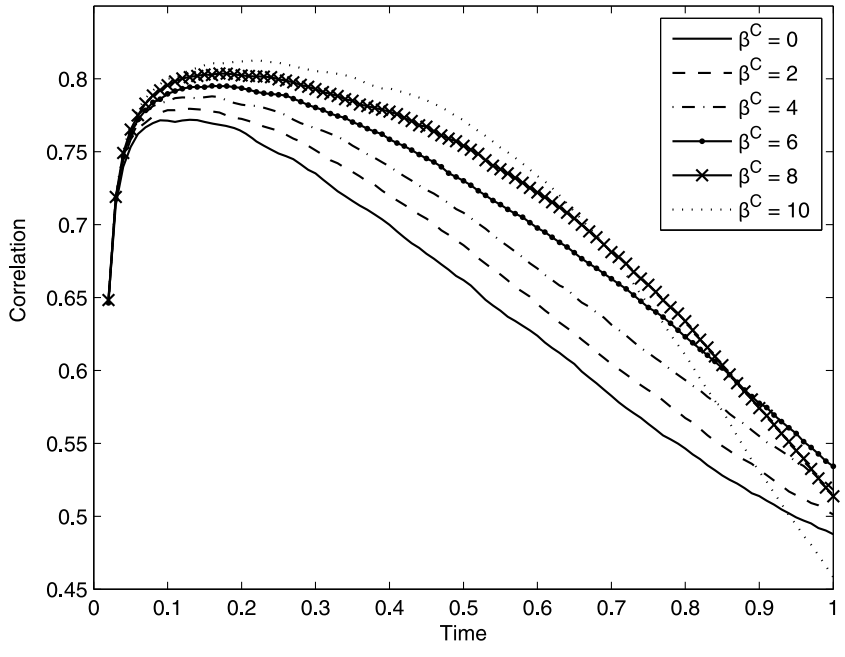


FIGURE 5.4. Evolution of the Spearman correlation between X_t and L_t over time t. The parameters are $\sigma = .9$, $\alpha = 4$, $\bar{\lambda} = .1$, $\lambda_0 = .1$, and $\beta^S = 4$. A total of 200,000 Monte Carlo trials and 16 moments ($K = 15$) were used.

consisting of several thousand names, a portfolio size not unusual in practice. Figure 5.8 compares the value at risk (VaR) of L_t^N and L_t at the 95% and 99% levels. The limiting VaR is surprisingly accurate even for moderately sized portfolios or several thousand firms. If N is relatively small, then the VaR of L_t tends to understate the VaR of L_t^N. This is because fluctuations of L^N due to the idiosyncratic noise terms in the intensity processes have not completely averaged out for small N.

We evaluate the computational efficiency of the approximation. The parameter case is $\sigma = 0.9, \alpha = 4, \bar{\lambda} = 0.2, \beta^C = 2, \beta^S = 3, \lambda_0 = 0.2$, and $t = 1$. A discrete time-step of 0.01 is used. One thousand Monte Carlo samples are produced. The table below indicates computation times for L_t^N, for each of several N. The computation time for L_t using $K = 200$ is 6.25 seconds. (Here and below, the computation times are based on a Matlab implementation on a computer running Mac OS X with a 2.7 GHz dual-core processor.) In general, the computational effort will be of an order $\frac{N}{K}$ greater for the generation of L_t^N, where N is the number of firms in the pool and K is the truncation level of the moment method.

Portfolio size, N	Computation time (seconds)
500	13.58
1,000	22.12
5,000	236.24
10,000	441.80
25,000	1048.80

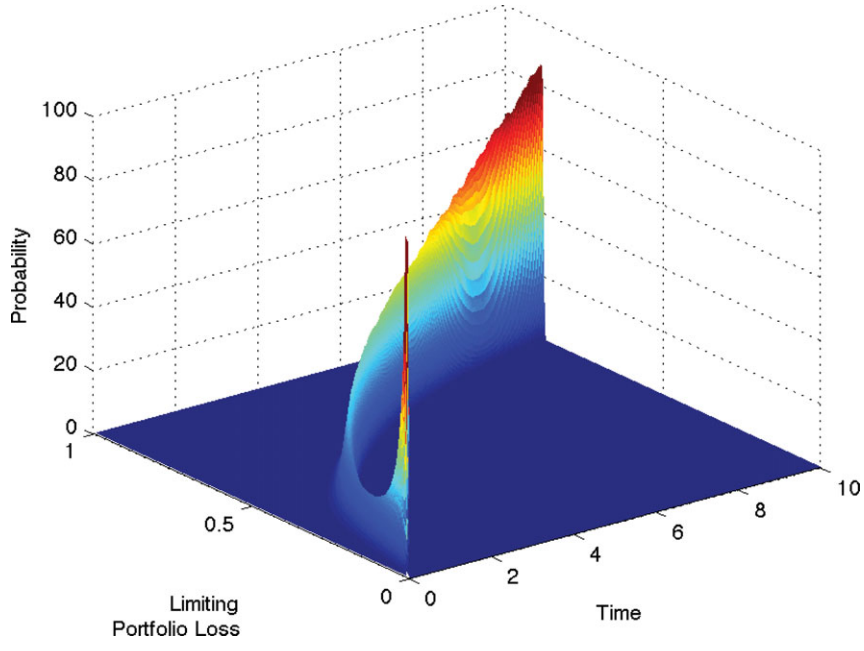


FIGURE 5.5. Evolution of distribution of limiting portfolio loss L_t over time t. The parameter case is $\sigma = .9$, $\alpha = 4$, $\bar{\lambda} = .2$, $\lambda_0 = .2$, $\beta^S = 4$, and $\beta^C = 2$. A total of 15,000 Monte Carlo trials and 16 moments ($K = 15$) were used.

5.5. Comparison of Method of Moments and Explicit Finite Difference

An alternate approach to the method of moments is a direct finite difference of the SPDE. An implicit method such as Crank–Nicholson cannot be used since the term $\beta^S \sigma_0(X_t)(\lambda v)_\lambda dV_t$ must be nonanticipating. We therefore use explicit finite difference, which has the disadvantages of only first-order accuracy in time and conditional stability.

We provide the explicit finite difference scheme in the case of the general diffusion $dX_t = b_0(X_t)\,dt + \sigma_0(X_t)\,dV_t$. The time-step is denoted Δ and the mesh-size by δ. Let $v_{i,j} = v(i\Delta, j\delta)$, $\lambda_j = j\delta$, $X_i = X_{i\Delta}$, and $\Delta V_i = V_{i\Delta} - V_{(i-1)\Delta}$, for $i = 0, \ldots, N$ and $j = 0, \ldots, J$. Then, the explicit finite difference scheme is

$$
\begin{aligned}
v_{i,j} = \Delta\Bigg[& \frac{\mathcal{I}_{i-1}}{2\delta} - \frac{\sigma^2}{2\delta} + \frac{\sigma^2\lambda_j}{2\delta^2} + (\beta^S\sigma_0(X_{i-1}))^2\frac{\lambda_j^2}{2\delta^2} + \frac{\alpha(\bar{\lambda} - \lambda_j)}{2\delta} \\
& + \beta^S\frac{\lambda_j}{2\delta}\left(b_0(X_{i-1}) + \sigma_0(X_{i-1})\frac{\Delta V_i}{\Delta}\right) - (\beta^S\sigma_0(X_{i-1}))^2\frac{\lambda_j}{\delta}\Bigg]v_{i-1,j-1} \\
& + \Delta\Bigg[\frac{1}{\Delta} + \alpha - \frac{\sigma^2\lambda_j}{\delta^2} - (\beta^S\sigma_0(X_{i-1}))^2\frac{\lambda_j^2}{\delta^2} - \lambda_j \\
& - \beta^S\left(b_0(X_{i-1}) + \sigma_0(X_{i-1})\frac{\Delta V_i}{\Delta}\right) + (\beta^S\sigma_0(X_{i-1}))^2\Bigg]v_{i-1,j} \\
& + \Delta\Bigg[-\frac{\mathcal{I}_{i-1}}{2\delta} + \frac{\sigma^2}{2\delta} + \sigma^2\frac{\lambda_j}{2\delta^2} + (\beta^S\sigma_0(X_{i-1}))^2\frac{\lambda_j^2}{2\delta^2} - \frac{\alpha(\bar{\lambda} - \lambda_j)}{2\delta} \\
& - \beta^S\frac{\lambda_j}{2\delta}\left(b_0(X_{i-1}) + \sigma_0(X_{i-1})\frac{\Delta V_i}{\Delta}\right) + (\beta^S\sigma_0(X_{i-1}))^2\frac{\lambda_j}{\delta}\Bigg]v_{i-1,j+1},
\end{aligned}
$$

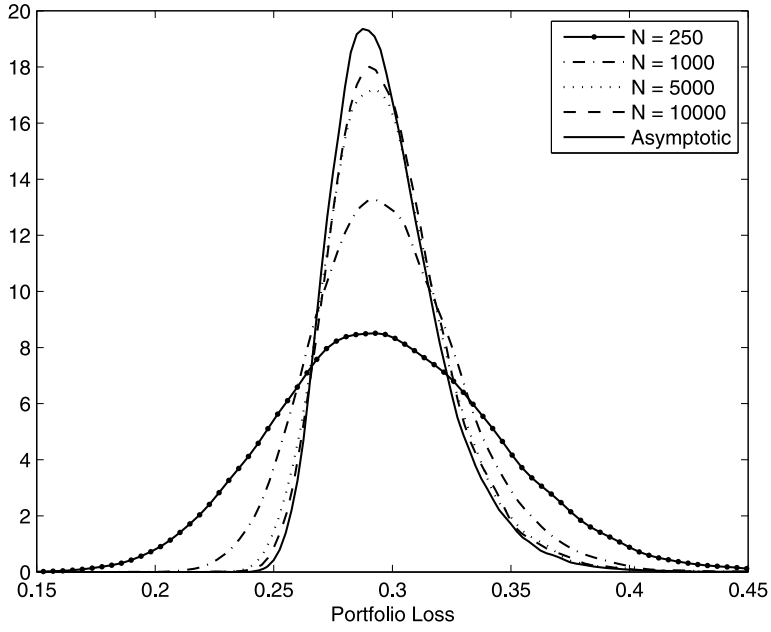


FIGURE 5.6. Comparison of distributions of limiting portfolio loss L_t (using 16 moments) and portfolio loss of finite system L_t^N for different N at $t = 1$. A total of 25,000 Monte Carlo trials were used for the finite system and 100,000 for the asymptotic solution. The parameter case is $\sigma = .9$, $\alpha = 4$, $\bar{\lambda} = .2$, $\lambda_0 = .2$, $\beta^C = 2$, and $\beta^S = 3$.

with boundary condition $\upsilon_{i,0} = \upsilon_{i,J} = 0$ and $\mathcal{I}_{i-1} = \sum_{j=1}^{J} \delta \frac{\upsilon(i-1,j)+\upsilon(i-1,j-1)}{2}$. For $\sigma_0 = 1$, the criterion for conditional stability for a deterministic diffusion PDE with constant coefficients leads us to propose the approximate criterion $\Delta \lesssim \frac{\delta^2}{(\beta^S \lambda_{\max})^2}$ for stability of the above numerical scheme, where $\lambda_{\max} = J\delta$. Numerical studies confirm that this condition for stability is a good approximation. Note that the time-step must become very small as the effect of systematic risk (i.e., the stochastic terms in the SPDE proportional to the parameter β^C) increases. For a general diffusion coefficient $\sigma_0(x)$, we expect instability to generally increase if $\sigma_0(X_t) > 1$ with high probability (and to decrease if $\sigma_0(X_t) < 1$ with high probability).

Accuracy of explicit finite difference and the method of moments is comparable since we also use a first-order accurate scheme, the Euler method, for the SDE moment system. However, the great advantage of the SDE moment system is its extremely low computational cost in comparison with explicit finite difference of the SPDE. If one chooses a mesh with J points for a finite difference scheme, then the finite difference scheme has an order of complexity of at least J coupled SDEs whereas the method of moments can achieve highly accurate results with as little as half a dozen SDEs. As an example comparison, if the finite difference scheme has a mesh ranging from 0 to 10 with a mesh-size of $\delta = 0.1$, the finite difference scheme has at least the order of complexity of 100 coupled SDEs. Furthermore, the explicit finite difference scheme is only conditionally stable. This means that even if one is satisfied with a time error of $\mathcal{O}(\bar{\Delta})$, one may have to choose a much smaller time-step Δ to avoid instability. For instance, if $\delta = 0.1$, $\beta^S = 5$, and $\lambda_{\max} = 10$, our approximate criterion indicates Δ must be less than 4×10^{-6}. Given a desired accuracy in time of $\mathcal{O}(\bar{\Delta})$, we estimate the ratio of the computational cost of

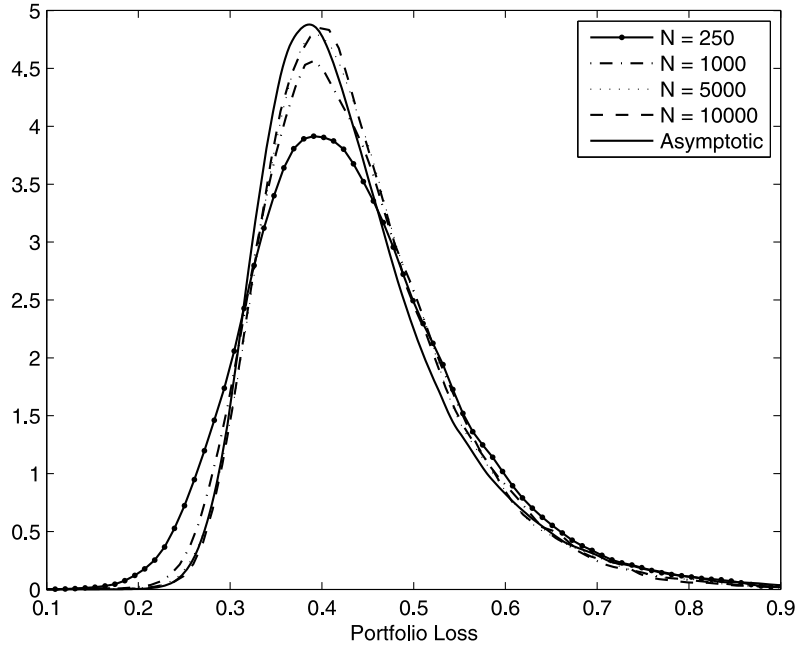


FIGURE 5.7. Comparison of distributions of limiting portfolio loss L_t (using 16 moments) and portfolio loss of finite system L_t^N for different N at $t = 1$. A total of 25,000 Monte Carlo trials were used for the finite system and 100,000 for the asymptotic solution. The parameter case is $\sigma = .9$, $\alpha = 4$, $\bar{\lambda} = .2$, $\lambda_0 = .2$, $\beta^C = 4$, and $\beta^S = 8$.

the explicit finite difference to that of the method of moments for the case of $\sigma_0 = 1$ to be

$$\frac{\text{Cost of Explicit FD}}{\text{Cost of Method of Moments}} \approx \frac{J}{K} \min\left(\bar{\Delta}\left(\beta^S \lambda_{\max}\right)^2 / \delta^2, 1\right). \tag{5.2}$$

As a concrete example, we report computational times for computing the loss at $t = 0.5$ for the case of $\sigma = 1, \alpha = 4, \bar{\lambda} = 1, \beta^C = 1.5, \beta^S = 2$, and $\lambda_0 = 2$. The systematic risk X_t is just a Brownian motion (i.e., $b_0 = 0$ and $\sigma_0 = 1$); we can then compare the observed numerical instability with our approximate criterion for stability. The finite difference method uses a mesh-size of $\delta = 0.1$ and $\lambda_{\max} = 10$ while the method of moments is truncated at level $K = 100$. This parameter case was chosen to demonstrate the instability of the explicit finite difference scheme when β^S becomes reasonably large and the mesh includes large values. Computational times are reported for 1,000 Monte Carlo trials.

Time step	Method of moments (seconds)	Explicit finite difference
10^{-2}	2.4890	Unstable
10^{-3}	25.7241	Unstable
10^{-4}	254.6494	Unstable
10^{-5}	2561.9614	8512.3172 seconds

The approximate criterion for stability gives $\Delta \leq 1.1111 \times 10^{-5}$, which matches well with the numerical results. In this example, we used a rather large number of moments

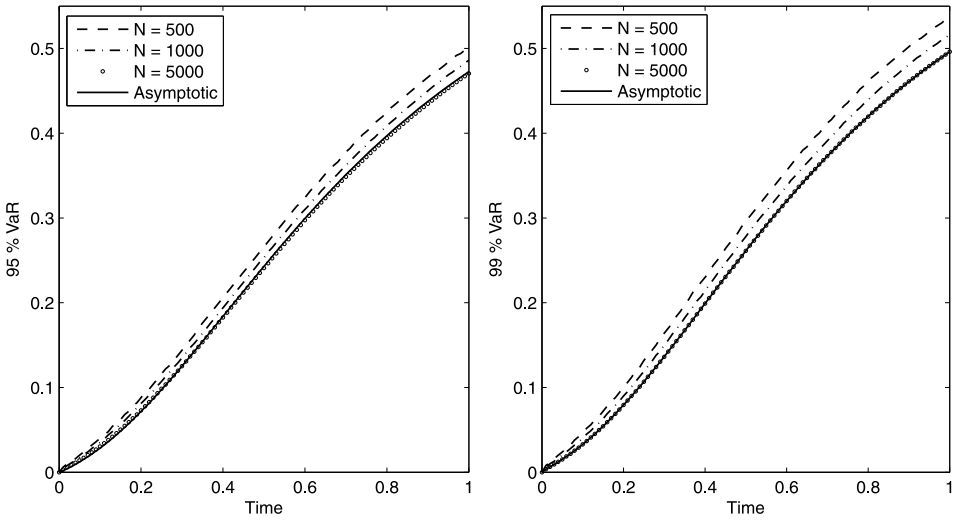


FIGURE 5.8. Comparison of 95% and 99% VaR of distributions of limiting portfolio loss L_t (using 16 moments) and portfolio loss of finite system L_t^N for different N and several horizons t. The parameter case is $\sigma = 0.9$, $\alpha = 4$, $\bar{\lambda} = 0.2$, $\lambda_0 = 0.2$, $\beta^S = 2$, and $\beta^C = 4$. A total of 50,000 Monte Carlo samples are used for the finite system and 250,000 for the asymptotic solution.

(100, to match the number of mesh points in the finite difference scheme). As we remarked earlier, the computational advantage of the method of moments over finite difference increases substantially if we take a small number of moments since highly accurate results are still achievable even using only a few moments.

6. EXTENSIONS

6.1. Extending the Model

We can extend the system (2.1) to the case of more general coefficient functions for the intensity as well as stochastic position losses. Fix $N \in \mathbb{N}$ and $n \in \{1, 2, \ldots, N\}$. Let $\{\ell^{N,n}\}_{n=1}^N$ be a family of i.i.d. random variables with support (0, 1). The variable $\ell^{N,n}$ represents the loss rate at default of the nth name in a pool of size N. Consider the following system:

(6.1)
$$\begin{aligned}
d\lambda_t^{N,n} &= b\big(\lambda_t^{N,n}\big)\,dt + \sigma\big(\lambda_t^{N,n}\big)\,dW_t^n + \beta_{N,n}^C \theta(X_t)\,dL_t^N + \beta_{N,n}^S \gamma\big(\lambda_t^{N,n}\big) dX_t \quad t > 0\\
\lambda_0^{N,n} &= \lambda_{\circ,N,n}\\
dX_t &= b_0(X_t)\,dt + \sigma_0(X_t) dV_t \quad t > 0\\
X_0 &= x_\circ\\
L_t^N &= \frac{1}{N}\sum_{n=1}^N \ell^{N,n} \chi_{\{\tau^{N,n} \le t\}},
\end{aligned}$$

where the coefficient functions $b(\cdot)$, $\sigma(\cdot)$, $\theta(\cdot)$, and $\gamma(\cdot)$ satisfy suitable regularity conditions guaranteeing the existence of a unique nonnegative solution $\lambda_t^{N,n}$.

Define the operators

$$
\begin{aligned}
(\mathcal{L}_1 f)(\hat{\mathsf{p}}) &= \frac{1}{2}\sigma^2(\lambda)\frac{\partial^2 f}{\partial \lambda^2}(\hat{\mathsf{p}}) + b(\lambda)\frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}) - \lambda f(\hat{\mathsf{p}}) \\
(\mathcal{L}_2 f)(\hat{\mathsf{p}}) &= \frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}).\beta^C \\
(\mathcal{L}_3^x f)(\hat{\mathsf{p}}) &= \beta^S \gamma(\lambda) b_0(x)\frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}) + \frac{1}{2}(\beta^S)^2\gamma^2(\lambda)\sigma_0^2(x)\frac{\partial^2 f}{\partial \lambda^2}(\hat{\mathsf{p}}) \\
(\mathcal{L}_4^x f)(\hat{\mathsf{p}}) &= \beta^S \gamma(\lambda)\sigma_0(x)\frac{\partial f}{\partial \lambda}(\hat{\mathsf{p}}) \\
\mathcal{Q}(\hat{\mathsf{p}}) &= \lambda \ell.
\end{aligned}
$$

Then, following the arguments applied to the system (2.1), we can show that the SPDE governing the limiting density takes the following form:

$$
\begin{aligned}
d\upsilon(t, \hat{\mathsf{p}}) = &\left\{ \mathcal{L}_1^* \upsilon(t, \hat{\mathsf{p}}) + \mathcal{L}_3^{*, X_t}\upsilon(t, \hat{\mathsf{p}}) + \theta(X_t)\left(\int_{\hat{\mathsf{p}}' \in \hat{\mathcal{P}}} \mathcal{Q}(\hat{\mathsf{p}}')\upsilon(t, \hat{\mathsf{p}}')\, d\hat{\mathsf{p}}'\right)\mathcal{L}_2^*\upsilon(t, \hat{\mathsf{p}})\right\} dt \\
&+ \mathcal{L}_4^{*, X_t}\upsilon(t, \hat{\mathsf{p}}) dV_t,
\end{aligned}
$$

for $t > 0$ and $\hat{\mathsf{p}} \in \hat{\mathcal{P}}$, where $\mathsf{p} = (\beta^C, \beta^S, \ell) \in \mathcal{P} \stackrel{\text{def}}{=} \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}$ and $\hat{\mathcal{P}} \stackrel{\text{def}}{=} \mathcal{P} \times \mathbb{R}_+$.

Naturally, one expects that, for large N and for every $t \geq 0$, $L_t^N \approx \ell(1 - \int_0^\infty \upsilon(t, \mathsf{p}, \lambda)\, d\lambda)$.

6.2. Extending the Moment Method

The moment SDE system can be extended to the case of general coefficient functions as well as the nonhomogeneous parameter case. By Theorem 3.1, in the nonhomogeneous case the SPDE takes the form

$$
\begin{aligned}
d\upsilon(t, \hat{\mathsf{p}}) = &\left\{ \frac{1}{2}\sigma^2(\lambda\upsilon(t, \hat{\mathsf{p}}))_{\lambda\lambda} + \alpha((\lambda - \bar{\lambda})\upsilon(t, \hat{\mathsf{p}}))_\lambda - \lambda\upsilon(t, \hat{\mathsf{p}}) - \beta^S b_0(X_t)(\lambda\upsilon(t, \hat{\mathsf{p}}))_\lambda \right. \\
&+ \frac{1}{2}(\beta^S)^2\sigma_0^2(X_t)(\lambda^2\upsilon(t, \hat{\mathsf{p}}))_{\lambda\lambda} - \beta^C \int_{\lambda \in \mathbb{R}^+, \mathsf{p} \in \mathcal{P}} \lambda\upsilon(t, \hat{\mathsf{p}})\, d\lambda d\mathsf{p}(\upsilon(t, \hat{\mathsf{p}}))_\lambda \Big\} dt \\
&- \beta^S\sigma_0(X_t)(\lambda\upsilon(t, \hat{\mathsf{p}}))_\lambda dV_t, \\
\upsilon(t, \lambda = 0, \mathsf{p}) &= \upsilon(t, \lambda = \infty, \mathsf{p}) = 0, \\
\upsilon(t = 0, \hat{\mathsf{p}}) &= h(\hat{\mathsf{p}}).
\end{aligned}
$$

The SDE moment system for the nonhomogeneous case follows:

$$
\begin{aligned}
du_k(t, \mathsf{p}) = &u_k(t, \mathsf{p})[-\alpha k + \beta^S b_0(X_t)k + \frac{1}{2}(\beta^S)^2\sigma_0^2(X_t)k(k-1)]\, dt - u_{k+1}(t, \mathsf{p})\, dt \\
&+ u_{k-1}(t, \mathsf{p})[0.5\sigma^2 k(k-1) + \alpha\bar{\lambda}k + \beta^C k \int_{\mathsf{p} \in \mathcal{P}} u_1(t, \mathsf{p})\, d\mathsf{p}]\, dt \\
&+ \beta^S\sigma_0(X_t)k u_k(t, \mathsf{p})\, dV_t, \\
u_k(t = 0, \mathsf{p}) = &\int_0^\infty \lambda^k h(\hat{\mathsf{p}})\, d\lambda,
\end{aligned}
$$

where $u_k(t, \mathsf{p}) = \int_0^\infty \lambda^k u_k(t, \hat{\mathsf{p}})\, d\lambda$. The SDE moment system is coupled across the parameter space $\mathcal{P}$. For numerical implementation, the parameter space $\mathcal{P}$ must be discretized.

For the case of intensity processes with general coefficients $b(\cdot)$, $\sigma(\cdot)$, and $\gamma(\cdot)$, we have an SPDE of the form stated in Section 6.1. A general moment method can be applied to this class of SPDEs, provided that we prescribe $b(\cdot)$, $\sigma(\cdot)$, and $\gamma(\cdot)$ such that the processes $\lambda^{N,n}$ stay positive almost surely. Also, assume $b(\cdot)$, $\sigma(\cdot)$, and $\gamma(\cdot)$ are analytic on $\mathbb{R}^+$. Then, the generalized moment

$$u_{k,k_1,k_2,k_3,k_4,k_5,k_6}(t) = \int_0^\infty \lambda^k b(\lambda)^{k_1} \frac{\partial^{k_2} b}{\partial \lambda^{k_2}} \big(\sigma(\lambda)^2\big)_3^k \frac{\partial^{k_4} \sigma}{\partial \lambda^{k_4}} \gamma(\lambda)^{k_5} \frac{\partial^{k_6} b}{\partial \lambda^{k_6}} \upsilon(t, \lambda)\, d\lambda$$

solves a moment system similar (albeit more complicated) to (5.1).

7. TIGHTNESS AND IDENTIFICATION OF THE LIMIT

We start by discussing relative compactness of the sequence $\{\mu^N\}_{N\in\mathbb{N}}$.

LEMMA 7.1. *The sequence $\{\mu^N\}_{N\in\mathbb{N}}$ is relatively compact in $D_E[0, \infty)$.*

Proof. The proof follows exactly as in section 6 of Giesecke et al. (2011). We omit the details. □

Next, we want to use the martingale problem (see Ethier and Kurtz 1986) to identify the limit of μ^N's.

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

$$\Phi(x, \mu) = \varphi_1(x)\varphi_2\left(\langle f_1, \mu\rangle_E, \langle f_2, \mu\rangle_E \ldots \langle f_M, \mu\rangle_E\right) \tag{7.1}$$

for some $M \in \mathbb{N}$, some $\varphi_1 \in C^\infty(\mathbb{R})$, some $\varphi_2 \in C^\infty(\mathbb{R}^M)$, and some $\{f_m\}_{m=1}^M$. Then $\mathcal{S}$ separates $\mathscr{P}(\mathbb{R} \times \hat{\mathcal{P}})$ (Ethier and Kurtz 1986). It thus suffices to show convergence of the martingale problem for functions of the form (7.1).

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

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

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

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

$$\left\langle f, \mu_t^N\right\rangle_E - \left\langle f, \mu_{t-}^N\right\rangle_E = \mathcal{J}_{N,n}^f(t), \tag{7.2}$$

where

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

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

For convenience, let's define

$$\mathcal{G}(f)(x) \stackrel{\text{def}}{=} b_0(x)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma_0^2(x)\frac{\partial^2 f}{\partial x^2}$$

for all $f \in C^2(\mathbb{R})$. This is the generator of the systematic risk.

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 $f \in C^\infty(\hat{\mathcal{P}})$, recall the definitions of the operators in (3.2).

For $\Phi \in \mathcal{S}$ of the form (7.1) we define the following operators.

$$\begin{aligned}(\mathcal{A}\Phi)(x,\mu) \stackrel{\text{def}}{=}\ & (\mathcal{G}(\varphi_1))(x)\varphi_2\left(\langle f_1,\mu\rangle_E, \langle f_2,\mu\rangle_E \ldots \langle f_M,\mu\rangle_E\right)\\ &+ \varphi_1(x)\sum_{m=1}^{M}\frac{\partial\varphi_2}{\partial x_m}\left(\langle f_1,\mu\rangle_E, \langle f_2,\mu\rangle_E \ldots \langle f_M,\mu\rangle_E\right)\big\{\langle\mathcal{L}_1 f_m,\mu\rangle_E\\ &+ \big\langle\mathcal{L}_3^x f_m,\mu\big\rangle_E + \langle\mathcal{Q},\mu\rangle_E\langle\mathcal{L}_2 f_m,\mu\rangle_E\big\}\\ &+ \frac{\partial\varphi_1}{\partial x}(x)\sum_{m=1}^{M}\frac{\partial\varphi_2}{\partial x_m}\left(\langle f_1,\mu\rangle_E, \langle f_2,\mu\rangle_E \ldots \langle f_M,\mu\rangle_E\right)\big\{\langle\sigma_0(x)\mathcal{L}_4^x f_m,\mu\rangle_E\big\}\end{aligned}$$

and

(7.3)

$$\begin{aligned}(\mathcal{B}\Phi)(x,\mu) \stackrel{\text{def}}{=}\ & \sigma_0(x)\frac{\partial\varphi_1}{\partial x}(x)\varphi_2\left(\langle f_1,\mu\rangle_E, \langle f_2,\mu\rangle_E \ldots \langle f_M,\mu\rangle_E\right)\\ &+ \varphi_1(x)\sum_{m=1}^{M}\frac{\partial\varphi_2}{\partial x_m}\left(\langle f_1,\mu\rangle_E, \langle f_2,\mu\rangle_E \ldots \langle f_M,\mu\rangle_E\right)\big\{\big\langle\mathcal{L}_4^x f_m,\mu\big\rangle_E\big\}.\end{aligned}$$

Moreover, we define the following processes

$$\begin{aligned}(\hat{\mathcal{A}}^{N,n}\Phi)(X_t,\mu_t^N) \stackrel{\text{def}}{=}\ & \varphi_1(X_t)\sum_{n=1}^{N}\lambda_t^{N,n}\big\{\varphi_2\big(\big\langle f_1,\mu_t^N\big\rangle_E + \mathcal{J}^{f_1}_{N,n}(t), \big\langle f_2,\mu_t^N\big\rangle_E\\ &+ \mathcal{J}^{f_2}_{N,n}(t)\ldots\big\langle f_M,\mu_t^N\big\rangle_E + \mathcal{J}^{f_M}_{N,n}(t)\big)\\ &- \varphi_2\big(\big\langle f_1,\mu_t^N\big\rangle_E, \big\langle f_2,\mu_t^N\big\rangle_E \ldots \big\langle f_M,\mu_t^N\big\rangle_E\big)\big\}\mathrm{M}_t^{N,n}\\ &- \varphi_1(X_t)\sum_{m=1}^{M}\frac{\partial\varphi_2}{\partial x_m}\big(\big\langle f_1,\mu_t^N\big\rangle_E, \big\langle f_2,\mu_t^N\big\rangle_E \ldots \big\langle f_M,\mu_t^N\big\rangle_E\big)\\ &\times\big\{\big\langle\mathcal{Q},\mu_t^N\big\rangle_E\big\langle\mathcal{L}_2 f_m,\mu_t^N\big\rangle_E - \big\langle\iota f,\mu_t^N\big\rangle_E\big\}\end{aligned} \tag{7.4}$$

and

$$\mathcal{M}_t^N \stackrel{\text{def}}{=} \mathcal{M}_t^{N,W} + \mathcal{M}_t^{N,J}, \tag{7.5}$$

where

$$\mathcal{M}_t^{N,W} = \frac{1}{N}\sum_{m=1}^{M}\sum_{n=1}^{N}\int_0^t \left[\varphi_1(X_r)\frac{\partial\varphi_2}{\partial x_m}\Big(\langle f_1,\mu_r^N\rangle_E,\langle f_2,\mu_r^N\rangle_E\ldots\langle f_M,\mu_r^N\rangle_E\Big)\right.$$
$$\left.\times\left\{\sigma_{N,n}\sqrt{\lambda_r^{N,n}}\frac{\partial f_m}{\partial\lambda}(\hat{\mathsf{p}}_r^{N,n})\mathrm{M}_r^{N,n}\right\}\right]dW_r^n$$

is the Brownian martingale and $\mathcal{M}^{N,J}$ is the martingale

$$\mathcal{M}_t^{N,J}$$
$$= \sum_{n=1}^{N}\int_0^t \varphi_1(X_r)\big\{\varphi_2\big(\langle f_1,\mu_r^N\rangle_E + \mathcal{J}_{N,n}^{f_1}(r),\langle f_2,\mu_r^N\rangle_E + \mathcal{J}_{N,n}^{f_2}(r)\ldots\langle f_M,\mu_r^N\rangle_E + \mathcal{J}_{N,n}^{f_M}(r)\big)$$
$$-\varphi_2\big(\langle f_1,\mu_r^N\rangle_E,\langle f_2,\mu_r^N\rangle_E\ldots\langle f_M,\mu_r^N\rangle_E\big)\big\}\big(-d\mathrm{M}_r^{N,n}-\lambda_r^{N,n}\mathrm{M}_r^{N,n}\big).$$

With these definitions we have the following lemma.

LEMMA 7.2. *For any $\Phi\in\mathcal{S}$ and any $t>0$ we have that*

$$\Phi\big(X_t,\mu_t^N\big) = \Phi\big(X_0,\mu_0^N\big) + \int_0^t(\mathcal{A}\Phi)\big(X_r,\mu_r^N\big)\,dr + \int_0^t(\mathcal{B}\Phi)\big(X_r,\mu_r^N\big)dV_r$$
$$+\int_0^t(\hat{\mathcal{A}}^{N,n}\Phi)\big(X_r,\mu_r^N\big)\,dr + \mathcal{M}_t^N.$$

Moreover, for any $T>0$, the following limits hold

$$\lim_{N\to\infty}\mathbb{E}\left[\int_0^t\big|(\hat{\mathcal{A}}^{N,n}\Phi)\big(X_r,\mu_r^N\big)\big|dr\right]=0 \text{ and } \lim_{N\to\infty}\sup_{0\le t\le T}\mathbb{E}\big[\mathcal{M}_t^N\big]^2=0.$$

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

$$(\mathcal{L}^a f)(\hat{\mathsf{p}}) = \frac{1}{2}\sigma^2\lambda\frac{\partial^2 f}{\partial\lambda^2}(\hat{\mathsf{p}}) - \alpha(\lambda-\bar{\lambda})\frac{\partial f}{\partial\lambda}(\hat{\mathsf{p}}).$$

Then $\mathcal{L}^a$ is the generator of the idiosyncratic part of the intensity.

We start by writing that

$$\Phi\big(X_t,\mu_t^N\big) = \Phi\big(X_0,\mu_0^N\big) + \sum_{k=1}^{4}\int_{r=0}^t A_r^{N,k}dr + \sum_{k=1}^{2}\int_{r=0}^t B_r^{N,k}dV_r + \mathcal{M}_t,$$

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

$$
\begin{aligned}
A_t^{N,1} &= \mathcal{G}(\varphi_1)(X_t)\varphi_2\big(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E \ldots \langle f_M, \mu_t^N\rangle_E\big)\\
A_t^{N,2} &= \varphi_1(X_t)\sum_{m=1}^{M}\frac{\partial\varphi_2}{\partial x_m}\big(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E \ldots \langle f_M, \mu_t^N\rangle_E\big)\\
&\quad\times\frac{1}{N}\sum_{n=1}^{N}\big\{(\mathcal{L}^a f_m)\big(\hat{\mathsf{p}}_t^{N,n}\big) + \big(\mathcal{L}_3^{X_t} f_m\big)\big(\hat{\mathsf{p}}_t^{N,n}\big)\big\}\mathrm{M}_t^{N,n}\\
&= \varphi_1(X_t)\sum_{m=1}^{M}\frac{\partial\varphi_2}{\partial x_m}\big(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E \ldots \langle f_M, \mu_t^N\rangle_E\big)\\
&\quad\times\big\{\langle\mathcal{L}^a f_m, \mu_t^N\rangle_E + \langle\mathcal{L}_3^{X_t} f_m, \mu_t^N\rangle_E\big\}\\
A_t^{N,3} &= \varphi_1(X_t)\sum_{n=1}^{N}\lambda_t^{N,n}\big\{\varphi_2\big(\langle f_1, \mu_t^N\rangle_E + \mathcal{J}_{N,n}^{f_1}(t), \langle f_2, \mu_t^N\rangle_E\\
&\quad+\mathcal{J}_{N,n}^{f_2}(t)\ldots\langle f_M, \mu_t^N\rangle_E + \mathcal{J}_{N,n}^{f_M}(t)\big)\\
&\quad-\varphi_2\big(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E \ldots \langle f_M, \mu_t^N\rangle_E\big)\big\}\mathrm{M}_t^{N,n}.\\
A_t^{N,4} &= \frac{\partial\varphi_1}{\partial x}(X_t)\sum_{m=1}^{M}\frac{\partial\varphi_2}{\partial x_m}\big(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E \ldots \langle f_M, \mu_t^N\rangle_E\big)\big\{\langle\sigma_0(X_t)\mathcal{L}_4^{X_t} f_m, \mu_t^N\rangle_E\big\}.\\
B_t^{N,1} &= \sigma_0(X_t)\frac{\partial\varphi_1}{\partial x}(X_t)\varphi_2\big(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E \ldots \langle f_M, \mu_t^N\rangle_E\big)\\
B_t^{N,2} &= \varphi_1(X_t)\sum_{m=1}^{M}\frac{\partial\varphi_2}{\partial x_m}\big(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E \ldots \langle f_M, \mu_t^N\rangle_E\big)\big\{\langle\mathcal{L}_4^{X_t} f_m, \mu_t^N\rangle_E\big\}.
\end{aligned}
$$

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 $n \in \{1, 2, \ldots, N\}$, define

$$
\tilde{\mathcal{J}}_{N,n}^f(t) \stackrel{\text{def}}{=} N\sum_{m=1}^{N}\frac{\partial f}{\partial\lambda}\big(\hat{\mathsf{p}}_t^{N,m}\big)\beta_{N,m}^C\mathrm{M}_t^{N,m} - f\big(\mathsf{p}^{N,n}, \lambda_t^{N,n}\big) = \big\langle\mathcal{L}_2 f, \mu_t^N\big\rangle_E - f\big(\hat{\mathsf{p}}_t^N\big). \tag{7.6}
$$

Then

$$
\left|\mathcal{J}_{N,n}^f(t) - \frac{1}{N}\tilde{\mathcal{J}}_{N,n}^f(t)\right| \leq \frac{K_{2.2}^2}{N^2}\left\|\frac{\partial^2 f}{\partial\lambda^2}\right\|_C,
$$

where $K_{2.2}$ is the constant from Condition 2.2.

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

$$
\begin{aligned}
\tilde{A}_t^{N,3} &\stackrel{\text{def}}{=} \varphi_1(X_t)\sum_{m=1}^{M}\frac{\partial\varphi_2}{\partial x_m}\big(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E \ldots \langle f_M, \mu_t^N\rangle_E\big)\frac{1}{N}\sum_{n=1}^{N}\lambda_t^{N,n}\tilde{\mathcal{J}}_{N,n}^{f_m}(t)\mathrm{M}_t^{N,n}\\
&= \varphi_1(X_t)\sum_{m=1}^{M}\frac{\partial\varphi_2}{\partial x_m}\big(\langle f_1, \mu_t^N\rangle_E, \langle f_2, \mu_t^N\rangle_E \ldots \langle f_M, \mu_t^N\rangle_E\big)\\
&\quad\times\big\{\langle\mathcal{Q}, \mu_t^N\rangle_E\langle\mathcal{L}_2 f_m, \mu_t^N\rangle_E - \langle\iota f, \mu_t^N\rangle_E\big\},
\end{aligned}
$$

we have that

$$\lim_{N\to\infty} \mathbb{E}\left[\int_{r=0}^{t} \left|(\hat{\mathcal{A}}^{N,n}\Phi)\big(X_r, \mu_r^N\big)\right| dr\right] = \lim_{N\to\infty} \mathbb{E}\left[\int_{r=0}^{t} \left| A_r^{N,3} - \tilde{A}_r^{N,3}\right| dr\right] = 0.$$

Moreover, by lemma 3.4 in Giesecke et al. (2011) we have that for any $T > 0$ and any $p \geq 1$, there is a constant C_0 such that $\sup_{\substack{0 \leq t \leq T \\ N \in \mathbb{N}}} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[|\lambda_t^{N,n}|^p] < C_0$. This and Condition 2.2 imply that

$$\lim_{N\to\infty} \sup_{0\leq t\leq T} \mathbb{E}\left[\mathcal{M}_t^N\right]^2 = 0.$$

Collecting things together, we get the statements of the lemma. □

We in particular note the macroscopic effect of the contagion.

REMARK 7.3. The key step in quantifying the coarse-grained effect of contagion was (7.6). Namely, we average the combination of the jump rate and the exposure to contagion across the pool.

8. PROOF OF THEOREM 3.1

Let $\mathbb{Q}_N$ be the $\mathbb{P}$-law of (X, μ^N); i.e.,

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

for all $A \in \mathcal{B}(D_{\mathbb{R}\times E}[0,\infty))$. Thus $\mathbb{Q}_N \in \mathcal{P}(D_{\mathbb{R}\times E}[0,\infty))$ for all $N \in \mathbb{N}$. For $\omega \in D_{\mathbb{R}\times E}[0,\infty)$, define $Y_t(\omega) \stackrel{\text{def}}{=} \omega(t)$ for all $t \geq 0$.

Also for $\Phi \in \mathcal{S}$, define the quantity

$$\Lambda_t^{\Phi}(Y) \stackrel{\text{def}}{=} \Phi(Y_t) - \Phi(Y_0) - \int_{r=0}^{t} (\mathcal{A}\Phi)(Y_r)\, dr - \int_{r=0}^{t} (\mathcal{B}\Phi)(Y_r)\, dV_r. \tag{8.1}$$

Let $\mathcal{V} = \bigcup_{t\in\mathbb{R}_+} \mathcal{V}_t$.

PROPOSITION 8.1. *We have that $\mathbb{Q}_N$ converges (in the topology of $\mathcal{P}(D_{\mathbb{R}\times E}[0,\infty))$) to the solution $\mathbb{Q}$ of the (filtered) martingale problem for $\Lambda_t^{\Phi}(Y)$ by (8.1) and such that $\mathbb{Q}\, Y_0^{-1} = \delta_{x_\circ\times\pi\times\Lambda_\circ}$. In particular, $\mathbb{Q}\{Y_0 = x_\circ \times \pi \times \Lambda_\circ\} = 1$ and for all $\Phi \in \mathcal{S}$ and $0 \leq r_1 \leq r_2 \ldots r_J = s < t < T$ and $\{\psi_j\}_{j=1}^{J} \subset B(\mathbb{R}\times E)$, we have that Λ_t^{Φ} is a square integrable martingale with respect to both $\mathcal{F}_t \bigvee \mathcal{V}$ and $\mathcal{F}_t$. Namely,*

$$\mathbb{E}\left[\big(\Lambda_t^{\Phi}(Y) - \Lambda_s^{\Phi}(Y)\big) \prod_{j=1}^{J} \psi_j\big(X_{r_j}, \mu_{r_j}^N\big)\right] = 0 \quad \text{and} \quad \sup_{0\leq t\leq T} \mathbb{E}\left[\Lambda_t^{\Phi}(Y)\right]^2 < \infty. \tag{8.2}$$

Lastly, $\mathbb{E}[\Lambda_t^{\Phi}(Y) \mid \mathcal{V}] = 0$.

Proof. The family $\{\mu^N\}_{N\in\mathbb{N}}$ is relatively compact (as a $D_E[0,\infty)$-valued random variable) by Lemma 7.1. Hence $\{X, \mu^N\}_{N\in\mathbb{N}}$ is also relatively compact. Let $(X, \bar{\mu})$ be an accumulation point of one of its convergent subsequences. Then, $(X, \mu^N, \int_0^{\cdot}(\mathcal{B}\Phi)(X_r, \mu_r^N)dV_r)$ will converge in distribution to $(X, \bar{\mu}, \int_0^{\cdot}(\mathcal{B}\Phi)(X_r, \bar{\mu}_r)\, dV_r)$. This and Lemma 7.2 imply

that the process $\Lambda_t^{\Phi}(X, \bar{\mu})$ will satisfy (8.2). Uniqueness of this martingale problem can also be shown as in Giesecke et al. (2011).

Of course, we also have that for any $\Phi \in \mathcal{S}$,

$$\mathbb{E}^{\mathbb{Q}}[\Phi(Y_0)] = \lim_{N\to\infty} \mathbb{E}^{\mathbb{Q}}\big[\Phi\big(X_0, \mu_0^N\big)\big] = \Phi(x_\circ \times \pi \times \Lambda_\circ)$$

which implies the claimed initial condition. The rest of the statements are easily seen to be true. □

We next want to identify $\mathbb{Q}$. This will take a couple of steps.

For notational convenience we shall write

$$\mathbb{E}_{\mathcal{V}}[\cdot] \stackrel{\text{def}}{=} \mathbb{E}[\cdot \mid \mathcal{V}].$$

The next lemma is essential for the characterization of the limit. Its proof is deferred to Appendix A.

LEMMA 8.2. *Let W^* be a reference Brownian motion and assume that Condition 2.3 is satisfied. For each $\hat{\mathsf{p}} = (\mathsf{p}, \lambda_\circ) \in \hat{\mathcal{P}}$ where $\mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S)$, there is a unique pair $\{(Q(t), \lambda_t(\hat{\mathsf{p}})) : t \in [0, T]\}$ taking values in $\mathbb{R}_+ \times \mathbb{R}_+$ such that*

$$Q(t) = \int_{\substack{\hat{\mathsf{p}} = (\mathsf{p}, \lambda) \in \hat{\mathcal{P}} \\ \mathsf{p} = (\alpha, \bar{\lambda}, \sigma, \beta^C, \beta^S)}} \mathbb{E}_{\mathcal{V}}\left\{\lambda_t^*(\hat{\mathsf{p}}) \exp\left[-\int_{s=0}^t \lambda_s^*(\hat{\mathsf{p}})\, ds\right]\right\} \pi(d\mathsf{p})\Lambda_\circ(d\lambda) \tag{8.3}$$

and

$$\begin{aligned}\lambda_t^*(\hat{\mathsf{p}}) = \lambda_\circ &- \alpha \int_{s=0}^t (\lambda_s^*(\hat{\mathsf{p}}) - \bar{\lambda})\, ds + \sigma \int_{s=0}^t \sqrt{\lambda_s^*(\hat{\mathsf{p}})}\, d W_s^* + \beta^C \int_{s=0}^t Q(s)\, ds \\ &+ \beta^S \int_{s=0}^t \lambda_s^*(\hat{\mathsf{p}})\, dX_s. \quad t \geq 0\end{aligned} \tag{8.4}$$

(where π and $\Lambda_\circ$ are as in Condition 2.1).

REMARK 8.3. For notational convenience we do not write the dependence of Q and λ^* on X but this is always assumed.

LEMMA 8.4. *We have that $\mathbb{Q} = \delta_{(X,\bar{\mu})}$, where for all $A \in \mathcal{B}(\mathcal{P})$ and $B \in \mathcal{B}(\mathbb{R}_+)$, $\bar{\mu}$ is given by*

$$\bar{\mu}_t(A \times B) \stackrel{\text{def}}{=} \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}} \chi_A(\mathsf{p}) \mathbb{E}_{\mathcal{V}_t}\left[\chi_B(\lambda_t^*(\hat{\mathsf{p}})) \exp\left[-\int_{s=0}^t \lambda_s^*(\hat{\mathsf{p}})\, ds\right]\right] \pi(d\mathsf{p})\Lambda_\circ(d\lambda).$$

Proof. Recall (8.4) and the operators $\mathcal{L}_1$, $\mathcal{L}_2$, $\mathcal{L}_3^x$, $\mathcal{L}_4^x$ from (3.2) and the definition of Q in (8.3).

For any $f \in C^\infty(\hat{\mathcal{P}})$,

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

Using Lemmas B.1 and B.2 we obtain

$$
\begin{aligned}
&d\langle f, \bar{\mu}_t\rangle_E \\
&= \left\{ \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}} \mathbb{E}_{\mathcal{V}_t}\left[\left[(\mathcal{L}_1 f)(\mathsf{p},\lambda_t^*(\hat{\mathsf{p}})) + \left(\mathcal{L}_3^{X_t} f\right)(\mathsf{p},\lambda_t^*(\hat{\mathsf{p}}))\right]\right.\right. \\
&\quad \left.\left. \times \exp\left[-\int_{s=0}^t \lambda_s^*(\hat{\mathsf{p}})\,ds\right]\right]\pi(d\mathsf{p})\Lambda_\circ(d\lambda)\right\} dt \\
&\quad + \left\{\int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}} \mathbb{E}_{\mathcal{V}_t}\left[(\mathcal{L}_2 f)(\mathsf{p},\lambda_t^*(\hat{\mathsf{p}}))\,Q(t)\exp\left[-\int_{s=0}^t \lambda_s^*(\hat{\mathsf{p}})\,ds\right]\right]\pi(d\mathsf{p})\Lambda_\circ(d\lambda)\right\} dt \\
&\quad + \left\{\int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)\in\hat{\mathcal{P}}} \mathbb{E}_{\mathcal{V}_t}\left[\left(\mathcal{L}_4^{X_t} f\right)(\mathsf{p},\lambda_t^*(\hat{\mathsf{p}}))\exp\left[-\int_{s=0}^t \lambda_s^*(\hat{\mathsf{p}})\,ds\right]\right]\pi(d\mathsf{p})\Lambda_\circ(d\lambda)\right\} dV_t \\
&= \left\{\langle \mathcal{L}_1 f, \bar{\mu}_t\rangle_E + Q(t)\langle \mathcal{L}_2 f, \bar{\mu}_t\rangle_E + \left\langle \mathcal{L}_3^{X_t} f, \bar{\mu}_t\right\rangle_E\right\} dt + \left\langle \mathcal{L}_4^{X_t} f, \bar{\mu}_t\right\rangle_E dV_t.
\end{aligned}
$$

To proceed, define

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

On the one hand, we have that

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

On the other hand, by Lemma 8.2 we have

$$
\dot{G}(t) = -Q(t). \tag{8.5}
$$

Thus, we have that

$$
d\langle f, \bar{\mu}_t\rangle_E = \left\{\langle \mathcal{L}_1 f, \bar{\mu}_t\rangle_E + Q(t)\langle \mathcal{L}_2 f, \bar{\mu}_t\rangle_E + \left\langle \mathcal{L}_3^{X_t} f, \bar{\mu}_t\right\rangle_E\right\} dt + \left\langle \mathcal{L}_4^{X_t} f, \bar{\mu}_t\right\rangle_E dV_t.
$$

Thus

$$
\Phi(X_t, \bar{\mu}_t) = \Phi(X_0, \bar{\mu}_0) + \int_{s=0}^t (\mathcal{A}\Phi)(X_s, \bar{\mu}_s)\,ds + \int_{s=0}^t (\mathcal{B}\Phi)(X_s, \bar{\mu}_s)\,dV_s,
$$

and hence $\delta_{(X,\bar{\mu})}$ 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. □

Now we collect our results to prove the law of large numbers given in Theorem 3.1.

Proof. [Proof of Theorem 3.1] In Lemma 8.4, we proved that, for any $f \in C^\infty(\hat{\mathcal{P}})$, the limiting measure $\bar{\mu}$ satisfies the measure evolution equation

$$
d\langle f, \bar{\mu}_t\rangle_E = \left\{\langle \mathcal{L}_1 f, \bar{\mu}_t\rangle_E + \langle \mathcal{Q}, \bar{\mu}_t\rangle_E\langle \mathcal{L}_2 f, \bar{\mu}_t\rangle_E + \left\langle \mathcal{L}_3^{X_t} f, \bar{\mu}_t\right\rangle_E\right\} dt + \left\langle \mathcal{L}_4^{X_t} f, \bar{\mu}_t\right\rangle_E dV_t.
$$

From this expression it is immediately derived by integration by parts that, if there exists a solution to the nonlinear SPDE (3.3), then the density of $\bar{\mu}$ should satisfy (3.3). This concludes the proof of the theorem. □

APPENDIX A: PROOF OF LEMMA 8.2

In this section, we prove Lemma 8.2. The proof uses a fixed point theorem argument. For notational convenience we sometimes drop the superscript $*$ and $\hat{\rho}$ from the notation of $\lambda_t^*(\hat{\rho})$ and simply write λ_t.

The square root singularity imposes some technical difficulties in the proof. For this purpose we introduce in Section A.1 an auxiliary function $\psi_\eta(x)$ that will be used later on and study its properties. In Section A.2, we study existence and uniqueness and properties of λ satisfying (8.4) with a given $Q(t)$.

In Section A.3, we prove the lemma under the additional condition that b_0 is bounded. Then, in Section A.4 we prove the lemma using Girsanov's Theorem for the X process.

A.1. An Auxiliary Function

Let $0 < \eta \ll 1$ and define

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

for all $x \in \mathbb{R}$. 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 $x > \sqrt{\eta}$,

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

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

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

and that $x\dot{\psi}_\eta(x) \ge 0$ for all $x \in \mathbb{R}$.

A.2. The Uncoupled Linear Case

Let ξ be a $\{\mathcal{G}_t\}_{t\geq 0}$-predictable, nondecreasing, bounded, and right-continuous process such that $\xi_0 = 0$. Consider the SDE

$$\text{(A.2)} \qquad \begin{aligned} d\lambda_t &= -\alpha(\lambda_t - \bar{\lambda})\,dt + \sigma\sqrt{\lambda_t \vee 0}\,dW_t + \beta^C d\xi_t + \beta^S \lambda_t d X_t \quad t > 0 \\ \lambda_0 &= \lambda_\circ. \end{aligned}$$

Then, as in lemmas 3.1 and 3.2 of Giesecke et al. (2011), we get that (A.2) has a unique nonnegative solution such that $\sup_{t\in[0,T]} \mathbb{E}[|\lambda_t|^q] < \infty$ for all $T > 0$ and $q \geq 1$.

A.3. Proof Assuming that $b_0(x)$ Is Bounded

In this subsection, we prove Lemma 8.2 assuming that $b_0(x)$ is bounded. The proof uses a fixed point theorem argument.

The main condition of this subsection is that there exists a $M < \infty$ such that

$$\text{(A.3)} \qquad \sup_{x\in\mathbb{R}} |b_0(x)| \leq M.$$

Let $q \geq 1$ and $S^q(\mathbb{R}_+)$ be the set of $\mathbb{R}_+$ valued, adapted, continuous processes $\{\lambda_t\}_{t\in[0,T]}$ such that

$$\|\lambda\|_{T,q} = \Big(\sup_{0\leq t\leq T} \mathbb{E}|\lambda_t|^q \Big)^{1/q} < \infty$$

The space $S^q(\mathbb{R}_+)$ endowed with the norm $\|\cdot\|_{T,q}$ is a Banach space.

Let us consider now a nonnegative process $U_t(\hat{\mathsf{p}}) \in S^q(\mathbb{R}_+)$. Set

$$\text{(A.4)} \qquad \xi(U)_t = \int_{\substack{\hat{\mathsf{p}} = (\mathsf{p},\,\lambda)\,\in\,\hat{\mathcal{P}} \\ \mathsf{p} = (\alpha,\,\bar{\lambda},\,\sigma,\,\beta^C,\,\beta^S)}} \left(1 - \mathbb{E}_{\mathcal{V}}\left\{\exp\left[-\int_{s=0}^{t} U_s(\hat{\mathsf{p}})\,ds\right]\right\}\right) \pi(d\mathsf{p})\Lambda_\circ(d\lambda)$$

and consider (A.2) with ξ in place of $\xi(U)$. We are going to prove that the map Φ defined by $\lambda = \Phi(U)$ through (A.2)–(A.4) with $U \in S^1(\mathbb{R}_+)$ is a contraction on $S^1(\mathbb{R}_+)$ equipped with the norm

$$\|\lambda\|_{t,1} < \infty$$

locally in t which can also be extended to any arbitrary T.

We collect in the following lemma some important properties of ξ as defined through (A.4).

LEMMA A.1. *Let a nonnegative process $U_t(\hat{\mathsf{p}}) \in S^q(\mathbb{R}_+)$ be given and let K be the constant from Condition 2.2. The map $\xi(U)_t$, as a function of t, is continuous, nondecreasing, positive, bounded uniformly in $t \in \mathbb{R}_+$ by K, and satisfies $\xi(U)_0 = 0$.*

Proof. All the statements are obvious. □

Fix $U_t(\hat{\mathsf{p}}), U'_t(\hat{\mathsf{p}}) \in S^1(\mathbb{R}_+)$ and write for notational convenience $\xi_t = \xi(U)_t$ and $\xi'_t = \xi(U')_t$. The process $Z_t = \lambda_t - \lambda'_t$ satisfies

$$Z_t = -\alpha \int_0^t Z_s\, ds + \sigma \int_0^t (\sqrt{\lambda_s} - \sqrt{\lambda'_s})\, dW_s + \beta^C \int_0^t (d\xi_s - d\xi'_s) + \beta^S \int_0^t Z_s\, dX_s.$$

Next, we apply Itô's formula to $\psi_\eta(x)$ defined in Section A.1 with $x = Z_t$, getting

(A.5)

$$\begin{aligned}
&\psi_\eta(Z_t) \\
&= -\alpha \int_0^t Z_s \psi'_\eta(Z_s)\, ds + \int_0^t \psi'_\eta(Z_s)(d\xi_s - d\xi'_s) \\
&\quad + \frac{\sigma^2}{2} \int_0^t (\sqrt{\lambda_s} - \sqrt{\lambda'_s})^2 \psi''_\eta(Z_s)\, ds + \sigma \int_0^t (\sqrt{\lambda_s} - \sqrt{\lambda'_s}) \psi'_\eta(Z_s)\, dW_s \\
&\quad + \beta^S \int_0^t \beta_o(X_s) Z_s \psi'_\eta(Z_s)\, ds + \beta^S \int_0^t \sigma_o(X_s) Z_s \psi'_\eta(Z_s) dV_s + \int_0^t (\beta^S \sigma_o(X_s) Z_s)^2 \psi''_\eta(Z_s)\, ds.
\end{aligned}$$

For notational convenience we define the quantity $\Delta_t Y = Y_t - Y'_t$ for any given couple of stochastic processes Y, Y'.

We have the following three technical Lemmas A.2–A.4.

LEMMA A.2. *For any $q \geq 1$ and $t > 0$ we have*

$$\begin{aligned}
\mathbb{E}|\Delta_t \xi|^q &\leq K^q \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \mathbb{E} \int_0^t |\Delta_s U(\hat{\mathsf{p}})|^q\, ds \pi(d\mathsf{p}) \Lambda_\circ(d\lambda) \\
&\leq t K^q \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \|\Delta_\cdot U(\hat{\mathsf{p}})\|^q_{t,q} \pi(d\mathsf{p}) \Lambda_\circ(d\lambda),
\end{aligned}$$

where K is the constant from Condition 2.2.

Proof. By the definition of $\xi(U)$ from (A.4) we get

$$\Delta_t \xi = \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \left(\mathbb{E}_V \left\{ \exp\left[-\int_{s=0}^t U'_s(\hat{\mathsf{p}})\, ds \right] - \exp\left[-\int_{s=0}^t U_s(\hat{\mathsf{p}})\, ds \right] \right\} \right) \pi(d\mathsf{p}) \Lambda_\circ(d\lambda).$$

Recall the trivial inequality $|e^a - e^b| \leq \max\{e^a, e^b\}|a - b|$. Since $U_s(\hat{\mathsf{p}}), U'_s(\hat{\mathsf{p}}) \geq 0$, we get that

$$\left| \exp\left[-\int_{s=0}^t U'_s(\hat{\mathsf{p}})\, ds \right] - \exp\left[-\int_{s=0}^t U_s(\hat{\mathsf{p}})\, ds \right] \right| \leq \int_0^t \left| U_s(\hat{\mathsf{p}}) - U'_s(\hat{\mathsf{p}}) \right| ds.$$

The latter, Hölder inequality and Fubini's Theorem imply

$$\begin{aligned}
\mathbb{E}\,|\Delta_t \xi|^q &\leq K^q \int_0^t \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \mathbb{E}|\Delta_s U(\hat{\mathsf{p}})|^q \pi(d\mathsf{p}) \Lambda_\circ(d\lambda)\, ds \\
&\leq t K^q \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \sup_{0 \leq s \leq t} \mathbb{E}|\Delta_s U(\hat{\mathsf{p}})|^q \pi(d\mathsf{p}) \Lambda_\circ(d\lambda).
\end{aligned}$$

This concludes the proof of the lemma. □

LEMMA A.3. *The following bound holds*

$$\mathbb{E}\psi_\eta(Z_t) \le C_1 \int_0^t \mathbb{E}|Z_s|\,ds + C_2 \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \|\Delta_\cdot U(\hat{\mathsf{p}})\|_{t,1}\pi(d\mathsf{p})\Lambda_\circ(d\lambda) + C_3 \tag{A.6}$$

where $C_1 = C_1(K, M)$, $C_2 = C_2(K, t)$ and $C_3(K, t, \eta)\downarrow 0$ as $\eta\downarrow 0$. Here K is the constant from Condition 2.2 and M as in (A.3).

Proof. The proof of this lemma follows by bounding each term on the right-hand side of (A.5) separately using the bound from Lemma A.2 and the bounds for the first- and second-order derivatives of ψ_η, i.e. $|\psi'_\eta(x)| \le 1$ and (A.1), respectively. In particular we have the following.

Taking expected value in (A.5) we obtain

$$\begin{aligned}\mathbb{E}\psi_\eta(Z_t) = {} & -\alpha\mathbb{E}\int_0^t Z_s\psi'_\eta(Z_s)\,ds + \mathbb{E}\int_0^t \beta^S b_o(X_s)Z_s\psi'_\eta(Z_s)\,ds + \mathbb{E}\int_0^t \beta^C\psi'_\eta(Z_s)(d\xi_s - d\xi'_s)\\ & + \frac{\sigma^2}{2}\mathbb{E}\int_0^t (\sqrt{\lambda_s} - \sqrt{\lambda'_s})^2\psi''_\eta(Z_s)\,ds + \mathbb{E}\int_0^t (\beta^S\sigma_o(X_s)Z_s)^2\psi''_\eta(Z_s)\,ds.\end{aligned} \tag{A.7}$$

Let us now bound each term on the right-hand side of (A.7).

Due to the boundedness condition on b_0 we have for the first and second term

$$-\alpha\mathbb{E}\int_0^t Z_s\psi'_\eta(Z_s)\,ds + \mathbb{E}\int_0^t \beta^S b_o(X_s)Z_s\psi'_\eta(Z_s)\,ds \le \left(\alpha + |\beta^S|M\right)\int_0^t \mathbb{E}|Z_s|\,ds. \tag{A.8}$$

Regarding the third term we recall the properties of ξ outlined in Lemma A.1 and that $|\psi'_\eta(x)| \le 1$ for all $\eta \ge 0$ and $x \in \mathbb{R}$. Then, approximating by simple processes, we get that there exists a constant C_0 such that

$$\left|\mathbb{E}\int_0^t \psi'_\eta(Z_s)\left(d\xi_s - d\xi'_s\right)\right| \le C_0 \sup_{0\le s\le t}\mathbb{E}\left|\xi_s - \xi'_s\right|.$$

The latter display and Lemma A.2 imply that

$$\left|\mathbb{E}\int_0^t \psi'_\eta(Z_s)(d\xi_s - d\xi'_s)\right| \le C_0 t K \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \|\Delta_\cdot U(\hat{\mathsf{p}})\|_{t,1}\pi(d\mathsf{p})\Lambda_\circ(d\lambda). \tag{A.9}$$

Relation (A.1) gives for the fourth term

$$\frac{\sigma^2}{2}\mathbb{E}\left|\int_0^t (\sqrt{\lambda_s} - \sqrt{\lambda'_s})^2\,\psi''_\eta(Z_s)\,ds\right| \le \frac{\sigma^2}{2}\int_0^t (\mathbb{E}|Z_s|\psi''_\eta(Z_s))\,ds \le \frac{\sigma^2}{\ln\eta^{-1}}. \tag{A.10}$$

Relation (A.1) and the condition $\mathbb{E}\int_0^t (\sigma_o(X_s))^2\,ds < \infty$ give for the fifth term

$$\mathbb{E}\left|\int_0^t (\beta^S\sigma_o(X_s)Z_s)^2\psi''_\eta(Z_s)\,ds\right| \le (\beta^S)^2\left(\frac{\sqrt{\eta}}{\ln\eta^{-1}}t\right)\mathbb{E}\int_0^t (\sigma_o(X_s))^2\,ds. \tag{A.11}$$

Collecting now terms (A.8)–(A.11) we obtain the statement of the lemma with

$$C_1 = \alpha + |\beta^S| M$$
$$C_2 = C_0 t K$$
$$C_3 = \left(\frac{\sigma^2}{\ln \eta^{-1}}\right) + \left(\beta^S\right)^2 \left(\frac{\sqrt{\eta}}{\ln \eta^{-1}} t\right) \mathbb{E} \int_0^t (\sigma_o(X_s))^2 \, ds. \qquad \square$$

LEMMA A.4. *For any $t > 0$, there exists a positive $C(t)$ such that*

$$\|\Delta_\cdot \lambda\|_{t,1} \leq C(t) \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \|\Delta_\cdot U(\hat{\mathsf{p}})\|_{t,1} \pi(d\mathsf{p}) \Lambda_\circ(d\lambda).$$

Moreover, $C(t)$ is continuous, increasing in t, and satisfies $\lim_{t \downarrow 0} C(t) = 0$.

Proof. Applying $|x| \leq \psi_\eta(x) + \sqrt{\eta}$ to $x = Z_t$ and using Gronwall's Lemma we obtain

$$\begin{aligned}
\mathbb{E}|Z_t| \leq \Bigg[& C_2 \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \|\Delta_\cdot U(\hat{\mathsf{p}})\|_{t,1} \pi(d\mathsf{p}) \Lambda_\circ(d\lambda) + C_3(t,\eta) + \eta^{1/2} \\
& + C_1 \int_0^t e^{C_1(t-s)} (C_3(s,\eta) + \eta^{1/2}) \, ds \\
& + C_1 \int_0^t e^{C_1(t-s)} \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} C_2 \|\Delta_\cdot U(\hat{\mathsf{p}})\|_{s,1} \pi(d\mathsf{p}) \Lambda_\circ(d\lambda) \, ds \Bigg]
\end{aligned}$$

where the constants C_1, C_2, C_3 are as in Lemma A.3. Taking $\eta \downarrow 0$ we obtain

$$\mathbb{E}|Z_t| \leq C(t) \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \|\Delta_\cdot U(\hat{\mathsf{p}})\|_{t,1} \pi(d\mathsf{p}) \Lambda_\circ(d\lambda)$$

where $C(t) = C_0[1 + e^{C_1 t} - e^{C_1 0}]t = C_0 e^{C_1 t} t$ for some constant $C_0 > 0$. This concludes the proof of the lemma. $\square$

We collect the previous results for the proof of Lemma 8.2 in the case of a bounded drift b_0.

LEMMA A.5. *If $b_0(x)$ satisfies (A.3), then the statement of Lemma 8.2 is true.*

Proof. Using Lemma A.4, a standard Picard iteration procedure shows that there exists a fixed point λ^* of Φ, i.e. $\lambda^*_t = \Phi_t(\lambda^*)$, for $t \in [0, t_1]$ such that $C(t_1) < 1$.

Let us show that this fixed point is necessarily unique. Indeed, suppose that λ^*, λ' are both fixed points. Then, Lemma A.4 implies that

$$\|\lambda^*(\hat{\mathsf{p}}) - \lambda'(\hat{\mathsf{p}})\|_{t,1} \leq C(t) \int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \|\lambda^*(\hat{\mathsf{p}}) - \lambda'(\hat{\mathsf{p}})\|_{t,1} \pi(d\mathsf{p}) \Lambda_\circ(d\lambda).$$

By integrating, and using the condition $C(t) < 1$ for $t \in [0, t_1]$, we immediately obtain that

$$\int_{\hat{\mathsf{p}}=(\mathsf{p},\lambda)} \|\lambda^*(\hat{\mathsf{p}}) - \lambda'(\hat{\mathsf{p}})\|_{t,1} \, \pi(d\mathsf{p}) \Lambda_\circ(d\lambda) = 0.$$

Therefore, for every $t \in [0, t_1]$ and $\hat{\mathfrak{p}} \in \hat{\mathcal{P}}$ we should have $\lambda_t^*(\hat{\mathfrak{p}}) = \lambda_t'(\hat{\mathfrak{p}})$ almost surely. This gives uniqueness.

For the general case, we only have to subdivide the interval $[0, T]$ into a finite number of small intervals. Thus, there is a unique fixed point to the equation $\lambda_t^* = \Phi_t(\lambda^*)$ for all $t \in [0, T]$. Notice that the equation $\lambda_t^* = \Phi_t(\lambda^*)$ can be trivially written as the pair of the coupled equations (8.3)–(8.4). Clearly $Q(t)$ is nonnegative since λ^* is nonnegative. This concludes the proof of the lemma. □

A.4. Proof of Lemma 8.2

In this section we prove Lemma 8.2. By Lemma A.5 we know that the statement of Lemma 8.2 is true if $b_0(x)$ is bounded. Our strategy is to apply this result for the case $b_0(x) = 0$ and then to generalize using Girsanov's Theorem.

Recall the map Φ defined by $\lambda_t = \Phi_t(\lambda)$ through (A.2)–(A.4). For notational convenience we shall write λ° for $\lambda_t^\circ = \Phi_t(\lambda^\circ)$ to emphasize the fact that we are considering (A.2)–(A.4) with $b_0(x) = 0$. We keep the notation λ for the case of a general $b_0(x)$.

Let $u(x)$ be such that $\sigma_0(x)u(x) = -b_0(x)$ and define the quantity

$$M_T = e^{-\int_0^T u(X_s)dV_s - \frac{1}{2}\int_0^T |u(X_s)|^2 ds}.$$

For given $\lambda_0 \in S^1(\mathbb{R}_+)$ we define iteratively the sequence

$$\lambda_{n+1,t} = \Phi_t(\lambda_n).$$

We shall prove that the sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in probability uniformly in $t \in [0, T]$. In particular we have the following lemma.

LEMMA A.6. *Assume that there is a $p > 1$ such that $\mathbb{E}\, M_T^p < \infty$ and let $q > 1$ be such that $1/p + 1/q = 1$. For every $\epsilon > 0$, there exists $N < \infty$ such that for all $N < n < m$*

$$\sup_{0\le t\le T} \mathbb{P}\left[\left|\lambda_t^m - \lambda_t^n\right| > \delta\right] \le C_0\,(\epsilon/\delta)^{1/q}$$

for every $\delta > 0$ and for some constant $C_0 > 0$ independent of n, m.

Proof. As we saw in the proof of Lemma A.5, it is enough to consider the equation $\lambda_t = \Phi_t(\lambda)$. Let λ^*, λ' be two solutions of the equation $\lambda_t = \Phi_t(\lambda)$.

By Girsanov's Theorem on the absolute continuous change of measure in the space of trajectories, Hölder and Chebychev inequality, we have the following

$$\begin{aligned}
\sup_{0\le t\le T} \mathbb{P}\left[\left|\lambda_t^m - \lambda_t^n\right| > \delta\right] &= \sup_{0\le t\le T} \mathbb{E}[\chi_{\{|\lambda_t^m - \lambda_t^n|>\delta\}}] \\
&= \sup_{0\le t\le T} \mathbb{E}[\chi_{\{|\lambda_t^{\circ,m} - \lambda_t^{\circ,n}|>\delta\}} M_T] \\
&\le \left(\sup_{0\le t\le T} \mathbb{P}\left[\left|\lambda_t^{\circ,m} - \lambda_t^{\circ,n}\right| > \delta\right]\right)^{1/q} \left(\mathbb{E}\, M_T^p\right)^{1/p} \\
&\le \left(\frac{\sup_{0\le t\le T} \mathbb{E}\left[\left|\lambda_t^{\circ,m} - \lambda_t^{\circ,n}\right|\right]}{\delta}\right)^{1/q} \left(\mathbb{E}\, M_T^p\right)^{1/p}
\end{aligned}$$

where $\lambda^{\circ,m}$, $\lambda^{\circ,n}$ satisfy $\lambda_t^{\circ,m} = \Phi_t(\lambda^{\circ,m})$ and $\lambda_t^{\circ,n} = \Phi_t(\lambda^{\circ,n})$.

Then Lemma A.4 implies that $\sup_{0\le t\le T}\mathbb{E}[|\lambda_t^{\circ,m}-\lambda_t^{\circ,n}|]$ can be made arbitrarily small by choosing n, m large enough. This together with the boundedness of $\mathbb{E}\,M_T^p$ conclude the proof of the lemma. □

Now we are in position to prove Lemma 8.2.

Proof. [Proof of Lemma 8.2] By Lemma A.6 we know that there is a nonnegative solution to the equation $\lambda_t = \Phi_t(\lambda)$. Uniqueness is shown in a similar fashion. This concludes the proof of the lemma. □

APPENDIX B: SOME AUXILIARY RESULTS

In this section, we prove some auxiliary lemmata that were needed in various places in the paper. These results are of independent interest.

LEMMA B.1. *If Y is an integrable $\mathcal{F}_t$-measurable random variable, then* $\mathbb{E}[Y|\mathcal{V}_t] = \mathbb{E}[Y|\mathcal{V}]$.

Proof. Notice that $\mathcal{V} = \mathcal{V}_t \bigvee \mathcal{V}_t^+$ where

$$\mathcal{V}_t^+ = \sigma\left(V_{t+s} - V_t;\ s \ge 0\right).$$

Since V is an $(\mathcal{F}_t)$-adapted Brownian motion, we get that $\mathcal{V}_t^+ \subset \mathcal{V}$ is independent of $\mathcal{F}_t$. Therefore, we have

$$\mathbb{E}\left[Y|\mathcal{V}_t\right] = \mathbb{E}\left[Y|\mathcal{V}_t \bigvee \mathcal{V}_t^+\right] = \mathbb{E}\left[Y|\mathcal{V}\right]. \qquad \square$$

LEMMA B.2. *Consider a càdlàg and $(\mathcal{F}_t)$-adapted process $\{Y_t;\ t \ge 0\}$ such that $\mathbb{E}[\int_0^T Y_t^2\,dt] < \infty$ for every $T \ge 0$. Then we have*

$$\mathbb{E}\left[\int_0^t Y_s dV_s \middle| \mathcal{V}_t\right] = \int_0^t \mathbb{E}[Y_s \mid \mathcal{V}_s]\, dV_s.$$

Proof. Given an arbitrary $u \in L^\infty([0,t];\mathbb{R})$, i.e. $\sup_{0\le s\le t}|u_s| < \infty$, consider the quantity

$$\mathcal{Z}_t = \exp\left(\frac{1}{2}\int_0^t |u_s|^2\,ds + i\int_0^t u_s\,dV_s\right).$$

It is easy to see that $\mathcal{Z}_t$ satisfies the SDE

$$\mathcal{Z}_t = 1 + i\int_0^t \mathcal{Z}_s u_s\,dV_s.$$

Therefore, on the one hand we have

$$\mathbb{E}\left[\mathcal{Z}_t\mathbb{E}\left[\int_0^t Y_s dV_s \,\middle|\, \mathcal{V}_t\right]\right] = \mathbb{E}\left[\mathcal{Z}_t\int_0^t Y_s dV_s\right] = \mathbb{E}\left[i\int_0^t \mathcal{Z}_s u_s Y_s\, ds\right]$$
$$= \mathbb{E}\left[i\int_0^t \mathcal{Z}_s u_s \mathbb{E}[Y_s \mid \mathcal{V}_s]\, ds\right].$$

On the other hand,

$$\mathbb{E}\left[\mathcal{Z}_t\int_0^t \mathbb{E}[Y_s|\mathcal{V}_s]\, dV_s\right] = \mathbb{E}\left[i\int_0^t \mathcal{Z}_s u_s \mathbb{E}[Y_s|\mathcal{V}_s]\, ds\right].$$

Thus we arrive at the equality

$$\mathbb{E}\left[\mathcal{Z}_t\mathbb{E}\left[\int_0^t Y_s dV_s \,\middle|\, \mathcal{V}_t\right]\right] = \mathbb{E}\left[\mathcal{Z}_t\int_0^t \mathbb{E}[Y_s|\mathcal{V}_s]\, dV_s\right].$$

Since this statement is true for every $u \in L^\infty([0, t]; \mathbb{R})$, we are done. □

REFERENCES

Ang, D. D., R. Gorenflo, V. K. Le, and D. D. Trong (2002): *Moment Theory and Some Inverse Problems in Potential Theory and Heat Conduction*, Lecture Notes in Mathematics, Berlin: Springer-Verlag.

Azizpour, S., K. Giesecke, and G. Schwenkler (2010): Exploring the Sources of Default Clustering, Working Paper, Stanford University.

Beskos, A., and G. Roberts (2005): Exact Simulation of Diffusions, *Ann. Appl. Probab.* 15(4), 2422–2444.

Bush, N., B. Hambly, H. Haworth, L. Jin, and C. Reisinger (2011): Stochastic Evolution Equations in Portfolio Credit Modelling, *SIAM J. Finan. Math.* 2, 627–664.

Chen, N. (2013): Localization and Exact Simulation of Brownian Motion Driven Stochastic Differential Equations, *Math. Operat. Res.* 38(3), 591–616.

Cvitanić, J., J. Ma, and J. Zhang (2011): Law of Large Numbers for Self-Exciting Correlated Defaults, *Stoch. Process. Appl.* 122, 2781–2810.

Dai Pra, P., and M. Tolotti (2009): Heterogeneous Credit Portfolios and the Dynamics of the Aggregate Losses, *Stoch. Process. Appl.* 119, 2913–2944.

Dai Pra, P., W. Runggaldier, E. Sartori, and M. Tolotti (2009): Large Portfolio Losses: A Dynamic Contagion Model, *Ann. Appl. Probab.* 19, 347–394.

Das, S., D. Duffie, N. Kapadia, and L. Saita (2007): Common Failings: How Corporate Defaults Are Correlated, *J. Finance* 62, 93–117.

Davis, M., and J. C. E. Rodriguez (2007): Large Portfolio Credit Risk Modelling, *Int. J. Theor. Appl. Finance* 10, 653–678.

Dembo, A., J.-D. Deuschel, and D. Duffie (2004): Large Portfolio Losses, *Finance Stoch.* 8, 3–16.

Duffie, D., L. Saita, and K. Wang (2006): Multi-Period Corporate Default Prediction with Stochastic Covariates, *J. Finan. Econ.* 83(3), 635–665.

Ethier, S. N., and T. G. Kurtz (1986): *Markov Processes: Characterization and Convergence*, New York: John Wiley & Sons Inc.

Feller, W. (1951): Two Singular Diffusion Problems, *Ann. Math.* 54(1), 173–182.

GIESECKE, K., and D. SMELOV (2013): Exact Sampling of Jump-Diffusions, *Operat. Res.* 61(4), 894–907.

GIESECKE, K., and G. SCHWENKLER (2011): Filtered Likelihood for Point Processes. Working Paper, Stanford University.

GIESECKE, K., K. SPILIOPOULOS, and R. SOWERS (2011): Default Clustering in Large Portfolios: Typical Events, *Ann. Appl. Probab.* Forthcoming.

GIESECKE, K., and S. WEBER (2006): Credit Contagion and Aggregate Losses, *J. Econ. Dyn. Contl.* 30, 741–767.

GLASSERMAN, P., W. KANG, and P. SHAHABUDDIN (2007): Large Deviations in Multifactor Portfolio Credit Risk, *Math. Finance* 17(3), 345–379.

GORDY, M. B. (2003): A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules, *J. Finan. Inter.* 12, 199–232.

KIM, K.-H. (2008): l^p-theory of Parabolic SPDEs Degenerating on the Boundary of c^1, *J. Theor. Probab.* 21(1), 169–192.

KIM, K.-H. (2009): Sobolev Space Theory of SPDEs with Continuous or Measurable Leading Coefficients, *Stoch. Process. Appl.* 119(1), 16–44.

KRYLOV, N., and S. LOTOTSKY (1998): A Sobolev Space Theory of SPDEs with Constant Coefficients on a Half Line, *SIAM J. Math. Anal.* 30(2), 298–325.

LOTOTSKY, S. (2001): Linear Stochastic Parabolic Equations Degenerating on the Boundary of a Domain, *Electr. J. Probab.* 6(24), 1–14.

LUCAS, A., P. KLAASSEN, P. SPREIJ, and S. STRAETMANS (2001): An Analytic Approach to Credit Risk of Large Corporate Bond and Loan Portfolios, *J. Bank. Finance* 25, 1635–1664.

ROYDEN, H. (1988): *Real Analysis*, 3rd ed., New York: Macmillan Publishing Company.

SCHLOEGL, L., and D. O'KANE (2005): A Note on the Large Homogeneous Portfolio Approximation with the Student-t Copula, *Finance Stoch.* 9(4), 577–584.

SIRCAR, R., and T. ZARIPHOPOULOU (2010): Utility Valuation of Credit Derivatives and Application to CDOs, *Quantit. Finance* 10(2), 195–208.

VASICEK, O. (1991): Limiting Loan Loss Probability Distribution, Technical Report, KMV Corporation.

VINCENTI, W. G., and C. H. KRUGER (1965): *Introduction to Physical Gas Dynamics*, New York: John Wiley & Sons.

ZHANG, X., J. BLANCHET, K. GIESECKE, and P. GLYNN (2011): Affine Point Processes: Asymptotic Analysis and Efficient Rare-Event Simulation, Working Paper, Stanford University.