Numerical Valuation of Basket Credit Derivatives in Structural Jump-Diffusion Models

Karolina Bujok[*][†] and Christoph Reisinger[‡]

February 8, 2012

Abstract

We consider a model where each company's asset value follows a jump-diffusion process, and is connected with other companies via global factors. Motivated by ideas in Bush et al. (2011), where the joint density of asset values is evolved in a large basket approximation, we develop an algorithm for the efficient estimation of CDO index and tranche spreads consistent with underlying CDSs, through a finite difference simulation of the resulting SPDE. We verify the validity of this approximation numerically by comparison to results obtained by direct Monte Carlo simulation of the basket constituents. A calibration exercise assesses the flexibility of the model and its extensions to match CDO spreads from pre-crisis and crisis periods.

Key words: basket credit derivatives, structural models, jump-diffusion, SPDE, Monte Carlo methods

1 Introduction

Basket credit derivatives are financial instruments which help manage the credit risk of portfolios of defaultable assets. The credit crunch of 2008 highlighted the need for a more robust and sophisticated methodology to price and hedge complex credit instruments, such as collateralized debt obligations (CDOs).

Most of the currently used models fall into one of two categories, reduced-form and structural. In the first approach, a random time of default is modelled directly, typically as the jump of a Poisson process. Due to simplicity and ease of implementation, reduced-form models are commonly used in practice, even though they do not have direct economic grounding.

An alternative are structural models, which are based on the natural observation that default depends on the value of a company's assets versus its debt. In these models, default happens when the asset value falls below a given barrier. Merton (1974) assumes that a firm's asset value follows geometric Brownian motion, and that the company can

default only at the maturity of the debt. Using first passage time theory, Black & Cox (1976) extend the model by allowing default at any time.

A survey of the literature on reduced-form and structural models can be found, eg, in Schönbucher (2003) and Bielecki & Rutkowski (2002).

Bush et al. (2011) propose a new approach to the pricing of large portfolios of credit risky assets. The authors consider a structural model where firm values are driven by exchangeable diffusion processes where the dependence between firms is due to exposure to a common market factor. The innovation is that the joint density of the portfolio, which determines portfolio losses, is described by a stochastic partial differential equation (SPDE), the solution of which is shown to be the limit of the empirical measure of firm values for large basket size. This allows a macroscopic look at the basket, in the spirit of conditionally independent factor models, and has increasing computational advantages for large baskets. The model is by definition dynamic, and can handle both standard products, like CDOs, and more exotic ones, like forward starting CDOs.

It is found, however, that the fit of the basic diffusion model to two data sets, pre-crisis and during the crisis, is not accurate enough for practical use. There are well documented improvements of structural credit models, for instance by considering: unobservable default barriers (Duffie & Lando (2001)); asset value processes modelled by jump-diffusions (Zhou (1997); Hilberink & Rogers (2002); Hu & Ye (2007); Kiesel & Scherer (2007); and Lipton & Sepp (2009)) or more general Levy processes (Fang et al. (2010)); stochastic volatility (Fouque et al. (2008)); contagion effects (Haworth & Reisinger (2007)); stochastic recovery rate (Finger et al. (2002)).

In this work, we focus on the numerical implementation and calibration of large basket models as outlined above, which allows us to address questions of practical feasibility, market fit, and to assess the justification of the large basket approximation. We focus here on jump-diffusion models, in which a default can happen both expectedly, due to the diffusion part, and unexpectedly, due to the jump part, by a sudden fall in a company's value. We will discuss at the end of this article how the other extensions listed above might be incorporated in this framework.

The article is organised as follows. In Section 2, we briefly explain the credit instruments that we analyse later. In Sections 3 and 4, we present our structural model for single- and multi-name credit derivatives and derive an SPDE model for the joint density of a large credit portfolio. Additionally, in Subsection 4.3, we check the validity of the large basket approximation for a large but finite number of companies. Section 5 gives details of the numerical implementation of the model. Finally, in Section 6, we outline how to calibrate the model efficiently and analyse if the model is flexible enough to match market spreads both in pre-crisis and in crisis periods, while Section 7 concludes with remarks on further improvements of the model.

2   Basket Credit Derivatives

A collateralized debt obligation (CDO) is a set of fixed income securities whose payments depend on credit events in a pool of defaultable assets, like loans, bonds or credit default swaps (CDSs). In this article, we analyse a *synthetic* CDO, whose portfolio consists of CDSs.

A CDO offers products with different risk profiles, based on so-called *tranches*. The

risk connected to tranches varies from equity tranches, which are typically unrated or speculative-grade, investment-grade mezzanine tranches, to senior tranches, which can even have AAA rating. By dividing the pool of defaultable assets into tranches, for instance investors who would not be allowed to invest directly in the underlying assets, because of their too low ratings, can invest in senior tranches.

Each tranche of a CDO is defined by an attachment point, a, and a detachment point, $d > a$, which are typically given as a percentage of a portfolio notional, eg, for the equity tranche typically $a = 0\%$ and $d = 3\%$. The tranche notional is given as the difference between a and d.

Let the reference portfolio for a synthetic CDO consist of N CDSs, each with notional $N_0 = 1/N$. The total loss of the portfolio is then given by

$$L_t = N_0(1 - R) \sum_{i=1}^{N} 1_{\{\tau_i \leq t\}}, \tag{2.1}$$

where R is a recovery rate, τ_i is the default time of the i-th entity. The outstanding notional, Z_t, of a single tranche is given by

$$Z_t = [d - L_t]^+ - [a - L_t]^+.$$

The STCDO (Single Tranche CDO) protection buyer pays a running spread c^{CDO} on the outstanding tranche notional only, at a set of payment dates T_j, where $1 \leq j \leq n$. Denote the payment intervals by $\varrho_j = T_j - T_{j-1}$, and the value of a bank account at time t by b_t.

We now derive a formula for c^{CDO}, following Bush et al. (2011). It is assumed that default risk is hedgeable, eg, with defaultable bonds, which translates into the existence of a pricing measure $\mathbb{Q}$ under which the discounted values of traded credit risky assets are martingales. For an overview of pricing approaches see, eg, Schönbucher (2003), Brigo & Mercurio (2007).

The value of the fee leg of the contract, Bc^{CDO}, is then given by

$$Bc^{CDO} = c^{CDO} \sum_{j=1}^{n} \frac{\varrho_j}{b_{T_j}} \mathbb{E}^{\mathbb{Q}}[Z_{T_j}].$$

The value of the protection leg, D, is a function of tranche losses incurred in the intervals $(T_{j-1}, T_j]$, and can be written as

$$D = \sum_{j=1}^{n} \frac{1}{b_{T_j}} \mathbb{E}^{\mathbb{Q}}[Z_{T_{j-1}} - Z_{T_j}].$$

Since, at time $t = 0$, the fair values of fee and protection legs have to be equal, $Bc^{CDO} = D$, the STCDO spread is given by

$$c^{CDO} = \frac{\sum_{j=1}^{n} \frac{1}{b_{T_j}} \mathbb{E}^{\mathbb{Q}}[Z_{T_{j-1}} - Z_{T_j}]}{\sum_{j=1}^{n} \frac{\varrho_j}{b_{T_j}} \mathbb{E}^{\mathbb{Q}}[Z_{T_j}]}.$$

Similarly, for a synthetic CDO index with outstanding index notional

$$Z_t^I = N_0 \sum_{i=1}^{N} 1_{\{\tau_i > t\}},$$

the pricing formula is

$$c^I = \frac{(1-R)\sum_{j=1}^n \frac{1}{b_{T_j}}\mathbb{E}^{\mathbb{Q}}[Z_{T_{j-1}}^I - Z_{T_j}^I]}{\sum_{j=1}^n \frac{\varrho_j}{b_{T_j}}\mathbb{E}^{\mathbb{Q}}[Z_{T_j}^I]},$$

For a single-name CDS, we have

$$c^{CDS} = \frac{(1-R)\sum_{j=1}^n \frac{1}{b_{T_j}}\mathbb{E}^{\mathbb{Q}}[1_{\{T_{j-1}>\tau\}} - 1_{\{T_j>\tau\}}]}{\sum_{j=1}^n \frac{\varrho_j}{b_{T_j}}\mathbb{E}^{\mathbb{Q}}[1_{\{\tau>t\}}]}, \qquad (2.2)$$

where $\mathbb{E}^{\mathbb{Q}}[1_{\{\tau>t\}}] = P(\tau > t)$ is the survival probability of the reference entity with default time τ.

Tranche spreads depend crucially on the joint default probability of CDSs and hence on the dependence structure of the underlying firms. In contrast, index spreads only depend on expected losses in the entire portfolio.

3 A Single-Name Structural Model

We begin by discussing a model and pricing formulae for the individual CDSs in this section, and present the full multi-name model in Section 4. The results here will also be used for the calibration of multi-name models.

3.1 The Model Setup

We model a company's asset value, A_t, by a jump-diffusion process

$$\frac{dA_t}{A_t} = \alpha\, dt + \sigma\, dW_t + (Y-1)\, dN_t, \qquad (3.1)$$

where σ is the asset volatility, W is a standard Brownian motion, N a Poisson process with intensity λ, $(Y_k - 1)$ is the size of the kth jump of the compound Poisson process $\sum_{k=1}^{N_t}(Y_k - 1)$. We assume that W, N, Y are mutually independent, and that $\{Y_k\}$ are independent and identically distributed (i.i.d.). Under the risk-neutral pricing measure $\mathbb{Q}$, $\mathbb{E}^{\mathbb{Q}}[dA_t/A_t] = r\, dt$. Since $\mathbb{E}^{\mathbb{Q}}[(Y-1)\, dN_t] = \lambda\nu\, dt$ with $\nu = \mathbb{E}^{\mathbb{Q}}[Y_k - 1]$, $\alpha = r - \lambda\nu$.

A solution to (3.1) is given by

$$A_t = A_0 \exp\left\{\left(\alpha - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}\prod_{k=1}^{N_t} Y_k, \qquad A_0 > 0. \qquad (3.2)$$

If $\lambda = 0$, then $N_t = 0$, and we obtain geometric Brownian motion. As Merton (1976) notes, $Y - 1$ is the relative change of A_t through jumps. What is more, in the special case when $\{Y_k\}$ are log-normally distributed, conditional on N_t, A_t is also log-normally distributed.

Following Black & Cox (1976), one can define the default time τ as the first passage time of the company's asset value of a constant default barrier B,

$$\tau = \inf\{t > 0 : A_t \le B\}.$$

We assume that $A_0 > B$ and $P(Y_k > 0) = 1$, such that the asset process remains positive. In order to eliminate explicit dependence on B, we introduce the *distance-to-default*

$$X_t = \frac{1}{\sigma}(\ln(A_t) - \ln(B)). \tag{3.3}$$

Since, $A_0 > B$, $X_0 > 0$. By applying Itô's lemma to (3.3) and using (3.1), we obtain

$$dX_t = \beta \, dt + dW_t + \Pi \, dN_t, \tag{3.4}$$

where $\beta = \frac{1}{\sigma}(r - \lambda\nu - \frac{1}{2}\sigma^2)$, $\Pi = \ln Y / \sigma$. The first passage time τ of X_t is now

$$\tau = \inf\{t > 0 : X_t \le 0\}.$$

Implicit here is the assumption that default can be monitored continuously, while we will later work in a framework where default can only be detected at a discrete set of times $t_1, \ldots, t_n$, and then set

$$\tau = \min\{t \in \{t_1, \ldots, t_n\} : X_{t-} \le 0\}.$$

We account for default in X by setting

$$X_t = 0 \quad \forall \, t \ge \tau. \tag{3.5}$$

3.2 Distribution of the Jump Amplitude

We are aiming towards a computationally tractable, yet economically convincing structural default model. To meet these conditions, we need a distribution of jump amplitudes that enables sufficiently fast calculation of the survival probability of individual firms, yet gives realistic market dynamics.

Zhou (1997) suggests a jump-diffusion model with log-normal jumps, as first introduced in the general context of financial modelling by Merton (1976) as an extension to the Black-Scholes approach. In contrast, Kou (2002), Hilberink & Rogers (2002), Hu & Ye (2007) and Lipton & Sepp (2009) propose a double exponential jump size distribution. Kou (2002) specifically argues that although both types of jumps can lead to the leptokurtic feature of equity returns observed in the market, in continuous time double-exponential jump-diffusion models have better analytical tractability. As Kou & Wang (2003) show, due to the lack of memory associated with the exponential distribution, there exists an analytical solution, via Laplace transforms, to the distribution of first passage times for the double exponential case. Such a solution does not exists for the log-normal model. However, as Ramezani & Zeng (2007) note, the transition density of log-normal jump-diffusions has a more convenient form than of double-exponential ones.

What is more, Ramezani & Zeng (2007) assess empirically the performance of double exponential jump-diffusion compared to log-normal jump-diffusion and geometric Brownian motion in matching stock prices. Because of the inherent link between equity and credit in structural models (Merton (1976)), empirical findings in equity markets are relevant for the present setting. They find that both double-exponential and log-normal jump-diffusions give a better fit to market data than geometric Brownian motion. In their study, double-exponential jump-diffusions outperform log-normal ones when stock indices are concerned, however, for individual stocks the results are inconclusive. It should be

noted that a slightly better performance of the double-exponential jump-diffusion model can be attributed to the presence of an extra parameter.

We focus in the following on the log-normal case, for reasons of convenience, which will become clear in Section 3.3, however, the overall framework is independent of the jump size distribution.

3.3 Computing Survival Probabilities for Log-Normal Jumps

The initial distance-to-default X_0, together with the jumps, determines default probabilities at the very short end, while diffusion plays a role over the medium time range and the drift only for long time horizons. One could therefore use CDS spreads with different maturities to back out the different parameters. We will follow a different tack later on and use information from tranche spreads, and use CDS spreads only to infer X_0.

A key ingredient in the pricing of CDSs, which are used for the calibration of CDO models in our framework, is the survival probability $P(\tau > t)$. For a log-normal jump-diffusion model, an analytical formula for the survival probability does not exist, hence Zhou Zhou (1997) applies a Monte Carlo algorithm. Since this approach is computationally demanding when a portfolio of CDSs is considered, in order to obtain an analytical approximation to survival probabilities, Willemann (2007) assumes that a firm has survived to a given monitoring time out of $t_1, \ldots, t_n$, if at this particular time the asset value is above the barrier. As the author notes, this is a crude approximation, and the survival probability in the model is overestimated, especially when t is high,

$$P(\tau > t_n) = P(X_{t_n} > 0, X_{t_{n-1}} > 0, \ldots, X_{t_1} > 0) < P(X_{t_n} > 0).$$

We now derive a practically tractable algorithm to calculate the survival probability for log-normally distributed jumps. The proofs of the following Propositions are given in Appendix 3.3.

Proposition 3.1. *Let $0 < t_n \leq T$ be a monitoring date, X a jump-diffusion process driven by (3.4) with $\Pi \sim N(\mu_\Pi, \sigma_\Pi^2)$. The survival probability at time t_n, for $n \geq 2$ is given recursively by*

$$
\begin{aligned}
p(X_{t_n} = x_{t_n}, \tau > t_{n-1}) &= \int_0^\infty p(X_{t_n} = x_{t_n} \mid X_{t_{n-1}} = x_{t_{n-1}})\, p(X_{t_{n-1}} = x_{t_{n-1}}, \tau > t_{n-2})\, dx_{t_{n-1}}, \\
P(\tau > t_n) &= \int_0^\infty p(X_{t_n} = x_{t_n}, \tau > t_{n-1})\, dx_{t_n}.
\end{aligned}
\tag{3.6}
$$

Proposition 3.2. *The probability of no default at t_1, and the density of X_{t_1}, are given*

6

by

$$P(\tau > t_1) = \sum_{c=0}^{\infty} P(X_{t_1} > 0 \mid N_{t_1} = c) \cdot p(N_{t_1} = c)$$

$$= e^{-\lambda t_1} \sum_{c=0}^{\infty} \Phi\left(\frac{\mu_{X_{t_1}}}{\sigma_{X_{t_1}}}\right) \frac{(\lambda t_1)^c}{c!}, \qquad (3.7)$$

$$p(X_{t_1} = x_{t_1}) = \sum_{c=0}^{\infty} p(X_{t_1} > 0 \mid N_{t_1} = c) \cdot p(N_{t_1} = c)$$

$$= e^{-\lambda t_1} \sum_{c=0}^{\infty} \frac{1}{\sigma_{X_{t_1}}} \phi\left(\frac{x_{t_1} - \mu_{X_{t_1}}}{\sigma_{X_{t_1}}}\right) \frac{(\lambda t_1)^c}{c!}, \qquad (3.8)$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution, $\phi(\cdot)$ is the standard normal density, $\mu_{X_{t_1}} = x_0 + \beta t_1 + c\mu_\Pi$, $\sigma^2_{X_{t_1}} = t_1 + c\sigma^2_\Pi$.

Proposition 3.3. *The density of X_{t_n}, conditional on $X_{t_{n-1}}$, for $n \geq 2$, is given by*

$$p(X_{t_n} = x_{t_n} \mid X_{t_{n-1}} = x_{t_{n-1}}) = \sum_{c=0}^{\infty} p(X_{t_n} = x_{t_n} \mid X_{t_{n-1}} = x_{t_{n-1}}, \Delta N_{t_n} = c) \cdot p(\Delta N_{t_n} = c)$$

$$= e^{-\lambda \Delta t} \sum_{c=0}^{\infty} \frac{1}{\sigma_{X_{t_n} \mid X_{t_{n-1}}}} \phi\left(\frac{x_{t_n} - \mu_{X_{t_n} \mid X_{t_{n-1}}}}{\sigma_{X_{t_n} \mid X_{t_{n-1}}}}\right) \frac{(\lambda \Delta t)^c}{c!}, \ (3.9)$$

where $\Delta t = t_n - t_{n-1}$, $\Delta N_{t_n} = N_{t_n} - N_{t_{n-1}}$, $\phi(\cdot)$ is the standard normal density, $\mu_{X_{t_n} \mid X_{t_{n-1}}} = x_{t_n} + \beta \Delta t + c\mu_\Pi$, $\sigma^2_{X_{t_n} \mid X_{t_{n-1}}} = \Delta t + c\sigma^2_\Pi$.

Remark 3.1. *The recursion given in above Propositions, can be computed using straightforward numerical integration, as the integrands are sums of standard normal densities. What is more, the sums converge quickly, since $\frac{(\lambda \Delta t)^c}{c!} \to 0$ rapidly as $c \to \infty$, especially for $\lambda \ll 1$, as is the case for the calibration results given later in the paper, and for $\Delta t = 0.25$, ie, for quarterly monitored defaults.*

Remark 3.2. *A similar recursion is suggested in Fang et al. (2010). The difference is that the recursion given by Fang et al. is backwards, ie, first $p(X_{t_{n-1}} = x_{t_{n-1}}, X_{t_n} > 0)$ is obtained, then $p(X_{t_{n-2}} = x_{t_{n-2}}, X_{t_{n-1}} > 0, X_{t_n} > 0)$ and finally $p(X_{t_1} = x_{t_1}, X_{t_2} > 0, \ldots, X_{t_{n-1}} > 0, X_{t_n} > 0)$. In our case, we first calculate $p(X_{t_1} = x_{t_1})$, then $p(X_{t_2} = x_{t_2}, X_{t_1} > 0)$ and, in the end, $p(X_{t_n} = x_{t_n}, X_{t_{n-1}} > 0, \ldots, X_{t_1} > 0)$, which gives immediately $P(\tau > t_n) = P(X_{t_n} > 0, \ldots, X_{t_1} > 0)$ for each t_n, $0 \leq t_n < T$. The latter approach is more efficient for pricing CDSs, since for all monitoring times, t_n, the survival probability is needed.*

3.4 Continuous vs Discrete Default Monitoring

Both continuous monitoring of default (eg, Black & Cox (1976) and Hu & Ye (2007)) and discrete monitoring (eg, Hull et al. (2010) and Fang et al. (2010)) is assumed in applications of structural default models. In practice, a default is announced on a daily basis, as Fang et al. (2010) argue.
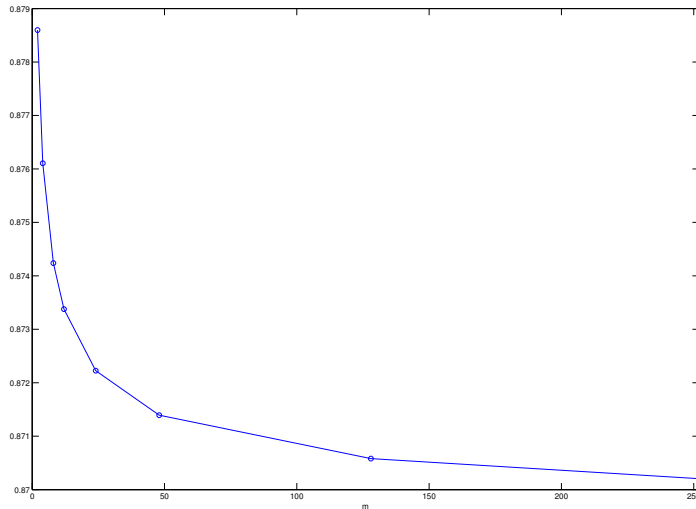
Figure 1: Survival probability at $t = 5$ in the log-normal jump-diffusion model with m default monitoring dates per year, where $m = 2, \ldots, 252$. Parameters are taken from a calibration to market data from 5 December 2008 (see Section 6 for details).

In order to make models computationally more tractable, some studies, eg, Hull et al. (2010) and Bush et al. (2011), assume that default is detected only on spread payment dates. We follow this line and thus assume that defaults are monitored quarterly.

In order to check the impact of this approximation, we calculate survival probabilities for $m = 2, \ldots, 252$ monitoring dates per year, ie, from half-yearly to daily monitoring, at time $t = 5$, ie, at the maturity of CDSs that we will use later to calibrate CDO pricing models. The results are shown in Fig. 1. As expected, the survival probabilities converge and are a decreasing function of m. The difference between the survival probability calculated for $m = 4$, ie, for the number of monitoring times we will assume in the computations, compared to $m = 252$, ie, the maximum realistic monitoring times per year, is 0.59 percentage points.

Remark 3.3. *The question of continuous versus discrete monitoring of defaults is mathematically equivalent to that of continuous versus discrete barrier monitoring when pricing barrier options.Broadie et al. (1997), in a diffusion model, approximate prices of discrete barrier options by formulae for continuously observed but shifted barrier, precisely (Theorem 1.1 in Broadie et al. (1997))*

$$V_n(H) = V\left(H\, e^{-\beta\sigma\sqrt{\Delta T}}\right) + o\left(\frac{1}{\sqrt{n}}\right), \tag{3.10}$$

where $V_n(H)$ is the price of a discretely monitored knock-out call option, $V(H)$ is the price of a corresponding continuously monitored barrier option, β is a generic constant derived by the authors, n is the number of monitoring times at intervals ΔT. This correction accounts for the probability of undetected barrier crossings between monitoring dates.

In a finite activity jump model, as the probability of a jump in an interval of length ΔT

is $O(\Delta T)$, and that of an undetected down-up combination even $O(\Delta T^2)$, we conjecture similar asymptotic behaviour.

As we calibrate the initial distance-to-default, X_0, to CDS spreads, which are functions of the survival probabilities, the barrier correction (3.10) is implicitly invoked. This is seen from (3.3), which shows that a shift of the initial position is equivalent to an opposite shift of the default barrier. When we use the calibrated model later to price basket derivatives based on the same monitoring dates, the difference between monitoring frequencies should be small.

4   A Multi-Name Structural Model

In this section, we extend the single-name jump-diffusion to a multi-name model, with the view of pricing basket credit derivatives.

The works closest to our basic setup are Willemann (2007) and Kiesel & Scherer (2007), where Willemann (2007) uses a log-normal jump-diffusion multi-name model to price CDOs, while Kiesel & Scherer (2007) use a double exponential jump distribution.

4.1   The Model Setup

We extend the single-name model to the setting of a portfolio of CDSs on N different companies, which is consistent with (3.1) for the individual names, but makes a special assumption on the dependence structure, namely,

$$\frac{dA_t^i}{A_t^i} = (r - \lambda\nu)\,dt + \sqrt{1-\rho}\,\sigma dW_t^i + \sqrt{\rho}\,\sigma dM_t + (Y-1)\,dN_t, \qquad (4.1)$$

where M is a standard Brownian motion, $\rho \in [0,1)$, $\ln Y \sim N(\mu_Y, \sigma_Y^2)$, (W^i) is a standard Brownian motion, W^i, M, N, Y are mutually independent.

W^i is an idiosyncratic factor, specific to each company, which affects the value of each company's assets alone, such as the management of a company. M, N and Y represent global factors that affect the default environment of all companies. As Hull et al. (2010) note, such a global factor could be for instance the S&P 500 index. Other candidates could be Gross Domestic Product, or more specifically investment spending, central bank interest rates or the unemployment rate, since they indicate the phase of the business cycle of the economy.

The jumps N together with Y, model sudden effects by global factors on the situation of companies, while M models more gradual influences. We aim at taking into account predominantly negative jumps and consider only $\mu_Y < 0$. The intensity of N, λ, then measures the frequency of economic shocks. If the absolute value of the jump mean is high, then λ can be interpreted as the frequency of economic crises. A similar interpretation can be found in Willemann (2007).

Like in the univariate case, we consider the *distance-to-default* $X_t^i = \frac{1}{\sigma}(\ln(A_t^i) - \ln(B_i))$, where B_i is a constant default barrier. We assume $X_0^i > 0$ and the dynamics of X_t^i are then given by

$$dX_t^i = \beta\,dt + \sqrt{1-\rho}\,dW_t^i + \sqrt{\rho}\,dM_t + \Pi\,dN_t, \qquad (4.2)$$

where $\beta = (r - \lambda\nu - \frac{1}{2}\sigma^2)/\sigma$, $\Pi \sim N(\mu_\Pi, \sigma_\Pi^2)$, $\mu_\Pi = \mu_Y/\sigma$, $\sigma_\Pi^2 = \sigma_Y^2/\sigma^2$.
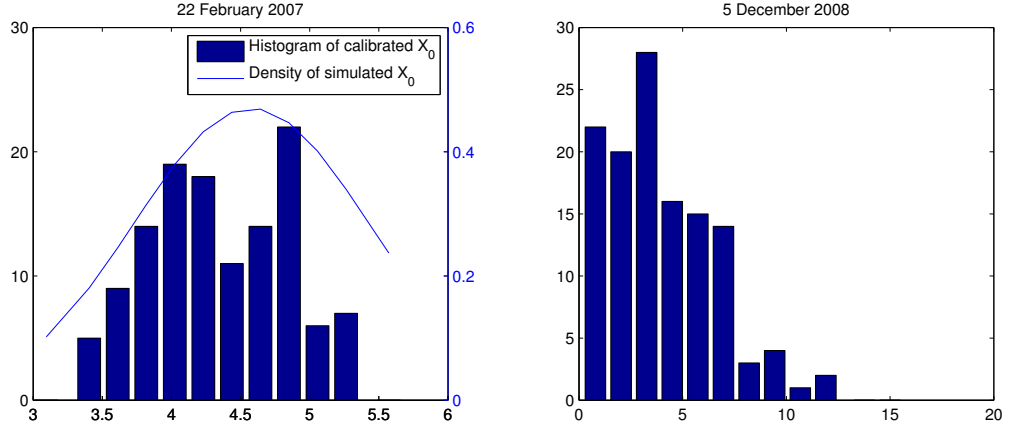
Figure 2: Histogram of X_0^i obtained by calibration to 5 year CDS spreads from the iTraxx Main Series, on two different dates pre- and during the crisis. On the left also shown is the density of X_0^i used in Subsection 4.3 for the simulation of synthetic credit baskets.

The distribution of *distances-to-default* X_0^i indicates the current state of the market. If the mass of the distribution is close to zero, as shown in Fig. 2, right, for the iTraxx constituents in December 2008, then the financial situation of companies is relatively bad; if the mass is concentrated far away from zero, as shown in Fig. 2, left, again for the iTraxx but in February 2007, the economic climate is good.

It is worth discussing in more detail a main premise of the model, namely that the volatility parameter σ is identical for all firms. Credit indices usually consist of companies from different industry sectors and different regions, and clearly there will be differences between the variances of asset values. If we allowed individual σ_i for each company in the portfolio, the dynamics would be, instead of (4.1),

$$\frac{dA_t^i}{A_t^i} = (r - \lambda \nu)\, dt + \sqrt{1-\rho}\, \sigma_i dW_t^i + \sqrt{\rho}\, \sigma_i dM_t + (e^{\sigma_i \Pi} - 1)\, dN_t, \qquad (4.3)$$

where σ_Π and μ_Π still have the same values for all companies. In terms of X_t^i,

$$dX_t^i = \beta_i\, dt + \sqrt{1-\rho}\, dW_t^i + \sqrt{\rho}\, dM_t + \Pi\, dN_t, \qquad (4.4)$$

where $\beta_i = (r - \lambda \nu - \frac{1}{2}\sigma_i^2)/\sigma_i$, and the initial *distance-to-default* is $X_0^i = (\ln(A_0^i) - \ln(B_i))/\sigma_i$. The volatility dependence of the default probabilities, is therefore to a large extent captured by the initial credit quality through X_0^i, and, especially for not too long maturities, the effect of the drift will be negligible compared to the diffusive and jump components.

The consequence of identical driving processes for all firms is that, as long as X_0^i are *exchangeable* random variables, ie, their joint law is invariant under permutation, the firm values X_t^i are also exchangeable. This allows us to study a large basket approximation in the following Section 4.2. Furthermore, conditional on X_0^i, M and J, X_t^i are independent.

For practical applications, the initial X_0^i will be calibrated to individual CDS spreads allowing through this heterogeneity in the portfolio.

10

The dependence between companies in the model is determined by *diffusion correlation*, represented by a single parameter ρ, together with perfect *jump correlation* for all companies, governed by common jump times of N_t and the same jump sizes, Π. A simple calculation shows that the overall correlation between X_t^i and X_t^j, $i \neq j$, driven by (4.2), is

$$\rho_{X_t^i X_t^j} = \frac{\rho + \zeta}{1 + \zeta}, \tag{4.5}$$

where $\zeta = \lambda(\sigma_\Pi^2 + \mu_\Pi^2)$.

In the present setting, the dependence between companies can vary from slight, if the values of ρ, λ, $|\mu_\Pi|$, σ_Π are small, to extremely strong, if the values are high. A few years ago, taking into account such a strong dependence might have been regarded as unrealistic, but the recent credit crunch, which affected not only financial markets but also the global real economy, showed that the dependence between companies can indeed be very strong. One might consider to allow different exposure to market factors by individual companies, which we do not do here for ease of calibration and computation, and since we are more concerned with the macroscopic behaviour of the basket.

4.2 The Large Basket Limit

Willemann (2007) derives analytical approximations to the survival probability for each company in the basket, and semi-analytical approximations to the portfolio loss distribution. In contrast, Kiesel & Scherer (2007) simulate the asset value process for each company to estimate joint default probabilities, which is not subject to model simplifications, but computationally expensive, since the basket size is typically large, eg, 125 for the iTraxx index. We make a virtue of this by using a large basket approximation in the spirit of Bush et al. (2011), extending the model there to jump-diffusions. Using this approach, the computational effort is essentially independent of the number of companies. We will test the validity of the large basket approximation in Section 4.3, and its numerical solution in Section 5.

Following Bush et al. (2011), we study the evolution of the empirical measure of $\{X_t^i, i = 1, \ldots, N\}$,

$$\nu_{N,t} = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^i}, \tag{4.6}$$

where δ_x is a Dirac measure centred at x. We are particularly interested in the limiting behaviour of $\nu_{N,t}$ for large N,

$$\nu_t = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^i}. \tag{4.7}$$

In the diffusion case, $\lambda = 0$ in (4.2), Kurtz & Xiong (1999) show that for exchangeable X_t^i with

$$E|X_0^1|^2 < \infty \tag{4.8}$$

the limit ν_t exists and its density is the unique solution of the stochastic partial differential equation

$$dv(t,x) = -\beta v_x(t,x)\, dt + \frac{1}{2} v_{xx}(t,x)\, dt - \sqrt{\rho} v_x(t,x)\, dM_t. \tag{4.9}$$

11

It is a straightforward application of Itô's lemma (Bush et al. (2011)) to see that a solution is given by
$$v(t, x) = u(t, x - \sqrt{\rho}M_t),$$
if $u(x, t)$ satisfies the deterministic PDE
$$u_t = \frac{1}{2}(1 - \rho)u_{xx} - \beta u_x \tag{4.10}$$
with initial condition $u(0, x) = v_0(x)$, the density of X_0^i.

The intuitive explanation is that in the large basket limit, the independent drivers have averaged to a deterministic diffusion, while the market factor M drives the whole basket.

Adding finite activity jumps, as per (4.2), we get the following result.

Proposition 4.1. *For X_t^i as per (4.2) with X_0^i satisfying (4.8), the limit empirical measure ν_t defined in (4.7) exists and has a density v of the form*
$$v(t, x) = u(t, x - \sqrt{\rho}M_t - J_t), \tag{4.11}$$
$$J_t = \sum_{k=1}^{N_t} \Pi_k, \tag{4.12}$$

and u satisfies (4.10). Moreover, v satisfies the SPDE
$$\begin{aligned} dv(t, x) = & -\beta v_x(t, x)\, dt + \frac{1}{2}v_{xx}(t, x)\, dt - \sqrt{\rho}v_x(t, x)\, dM_t \\ & + v(t_-, x - \Pi)\, dN_t - v(t_-, x)\, dN_t. \end{aligned} \tag{4.13}$$

Proof. See Appendix B. $\qquad\square$

In the present setting, we are especially interested in losses, ie, barrier crossings of X_t^i. Bush et al. (2011) derive an initial boundary value problem for the purely diffusive SPDE for the case of a continuously active default barrier.

For computational convenience, we make the assumption that defaults are only observed at a discrete set of times, which we take quarterly to coincide with the payment dates (see discussion in Section 3.4). If a firm's value is below the default barrier on one of the observation dates T_i, it is considered defaulted and removed from the basket, so we use (4.2) in $[0, \tau^i)$, and set
$$X_t^i = 0 \quad t \geq \tau^i, \tag{4.14}$$
$$\tau^i = \min\{t \in \{T_1, \ldots, T_n\} : X_{t-}^i \leq 0\}. \tag{4.15}$$

Then, the empirical measure can be written in the form
$$\bar{\nu}_{N,t} = \nu_{N,t} + \delta_0 L_t^N/(1 - R), \tag{4.16}$$
$$L_t^N = \frac{1}{N}(1 - R)\sum_{i=1}^{N} 1_{\{\tau_i \leq t\}}, \tag{4.17}$$
$$\nu_{N,t} = \frac{1}{N}\sum_{\substack{1 \leq i \leq N \\ \tau^i > t}} \delta_{X_t^i}. \tag{4.18}$$

Proposition 4.2. *For X_t^i as per (4.2) in $(0, \tau_i)$, together with (4.14), where X_0^i satisfy (4.8), the limit empirical measure*

$$\bar{\nu}_t = \lim_{N \to \infty} \bar{\nu}_{N,t} \tag{4.19}$$

exists and has the form $\bar{\nu}_t = \nu_t + \delta_0 L_t / (1 - R)$, where ν_t has a density v satisfying (4.13) for $t \in (T_k, T_{k+1}), 0 \le k < n$, and

$$L_t = (1 - R) \left(1 - \int_{-\infty}^{\infty} v(t, x) \, dx \right). \tag{4.20}$$

Moreover, with probability one,

$$\lim_{t \downarrow T_k} v(t, x) = \begin{cases} \lim_{t \uparrow T_k} v(t, x) & x > 0, \\ 0 & x < 0. \end{cases} \tag{4.21}$$

Proof. The existence of the limit (4.19) and of a density which solves the SPDE in (T_k, T_{k+1}) is a repeated application of Proposition 4.1. The form of the loss variable (4.20) is a direct consequence of the definitions and of the convergence of the measures. Together with (4.14), this implies (4.21). Details will be given in Bujok (2012). □

We discuss the numerical computation of (4.20) via (4.13) in Section 5.

4.3 Validity of the Large Basket Approximation

In this section we analyse if the size of common CDO baskets, typically $N = 125$, is large enough to use the (numerical) solution to the limiting SPDE (4.13) obtained for $N \to \infty$, as approximation to the evolution of firm values. We construct "nested" baskets of size $N_k = 5^k$, $k = 1, \ldots, 9$, ie, the N_k-basket contains all firms of the N_{k-1}-basket etc. For each basket, we calculate expected tranche losses, $\mathbb{E}[Y_t]$, for $t = 5$ years, where

$$Y_t = [L_t - a]^+ - [L_t - d]^+,$$

and the losses are either calculated from (2.1) by "direct" Monte Carlo simulation of (4.15), (4.2), or by (4.20) where v is the solution to the SPDE.

To create hypothetical baskets of different sizes, we simulate initial *distances-to-default*, X_0^i. For direct Monte Carlo simulation, the samples serve directly as starting values for (4.2), whereas for the SPDE, we construct the empirical measure (4.16) at $t = 0$. For simplicity, we draw X_0^i from a normal distribution, where the mean $\mu_{X_0} = 4.6$ and standard deviation $\sigma_{X_0} = 0.8$ are roughly fit to X_0^i obtained in a calibration exercise from data of 22 February 2007, see Fig. 2, left.

We calculate expected tranche losses conditional on these initial positions, ie, we average tranche losses over a large number of sample paths, but do *not* resample the initial positions. This is theoretically slightly less elegant but closer to the practical situation where initial values are backed out from implied default probabilities.

The differences between the direct Monte Carlo and SPDE results are due to: error of the finite difference discretisation of the SPDE; simulation error, both for the SPDE and direct Monte Carlo; large basket approximation of the dynamics. We want to focus on the latter effect, therefore we choose a high number of grid points and time steps to

reduce the discretisation bias, and a high number of simulations to reduce the Monte Carlo error to a negligible size.

As seen in Fig. 3, the difference between the SPDE results and those from direct Monte Carlo simulation of the basket vanishes (within the sampling error of X_0^i, which we will investigate in the next subsection) for large N.

The SPDE results depend on N only via the initial sample X_0^i, which is chosen to be the same here as for the direct Monte Carlo simulation. Therefore, the results in Fig. 3 indicate that the approximation error of tranche spreads due to approximating the evolution of the firm value distribution by the SPDE is noticeably smaller than the effect of the finite sample approximating the continuous density at $t = 0$.

Closer inspection reveals that this is due to the fact that the tails of the initial distribution of X_0^i are not well resolved for moderate N, which affects junior tranches, which measure losses in the left tail, more than senior tranches. To verify this numerically, we repeated the experiment with all mass centred in a single point $X_0^i = z$, where z is a constant, for all firms in the portfolio. Thus the initial condition becomes independent of the basket size. This gave a significantly smaller difference, especially for junior tranches.

For $N = 125$, ie, $k = 3$, the results of both methods are close, hence we argue that this number of companies is sufficiently large to use the large basket approximation.

4.4 Computational Issues of Direct Simulation

The cost of simulating the basket directly increases linearly with the basket size, while the difference of expected tranche losses to the ones obtained in the large basket limit becomes smaller. We carried out a preliminary numerical experiment with 10^6 paths, but the obtained confidence intervals for Monte Carlo estimators were too wide to draw reliable conclusions about the difference between direct Monte Carlo and SPDE methods for large baskets. This can be observed in Fig. 4.A, where we present results for the equity tranche. In this preliminary experiment, we calculated expected losses for $k \leq 7$, which was the largest of the baskets that was computable with the above number of simulations within an hour on a standard desktop computer.

To calculate expected tranche losses with confidence intervals as shown in Fig.4.B for the final experiment (standard deviation 4×10^{-6}), requires between 10^7 to 10^8 paths, which multiplied by a basket size from $N_k = 5$ to $N_k = 5^9$, and $n = 20$ monitoring times, gives a computational complexity from 10^{10} to 10^{14}, measured by the number of random variables to be drawn. This can be seen in Fig. 4.C and 4.D. It turned out in practice that a computational complexity of order 10^{10} was feasible on a standard computer within an hour, but 10^{14} was not.

To this end, we developed a nested variance reduction method, drawing on ideas from the *Multi-level* Monte Carlo method, proposed in Giles (2008). The main idea is that the expected tranche loss for a basket of size N_{k-1}, serves as control variate for the basket containing N_k companies, and so on. Writing $Y_t^{(k)}$ for the tranche loss of an N_k-basket,

$$E[Y_t^{(k)}] = E[Y_t^{(1)}] + \sum_{j=2}^{k} E[Y_t^{(j)} - Y_t^{(j-1)}], \qquad (4.22)$$

and we define independent estimators $\widehat{W}^{(j)}$ for the expectations in the sum, using $N_{sims}^{(j)}$
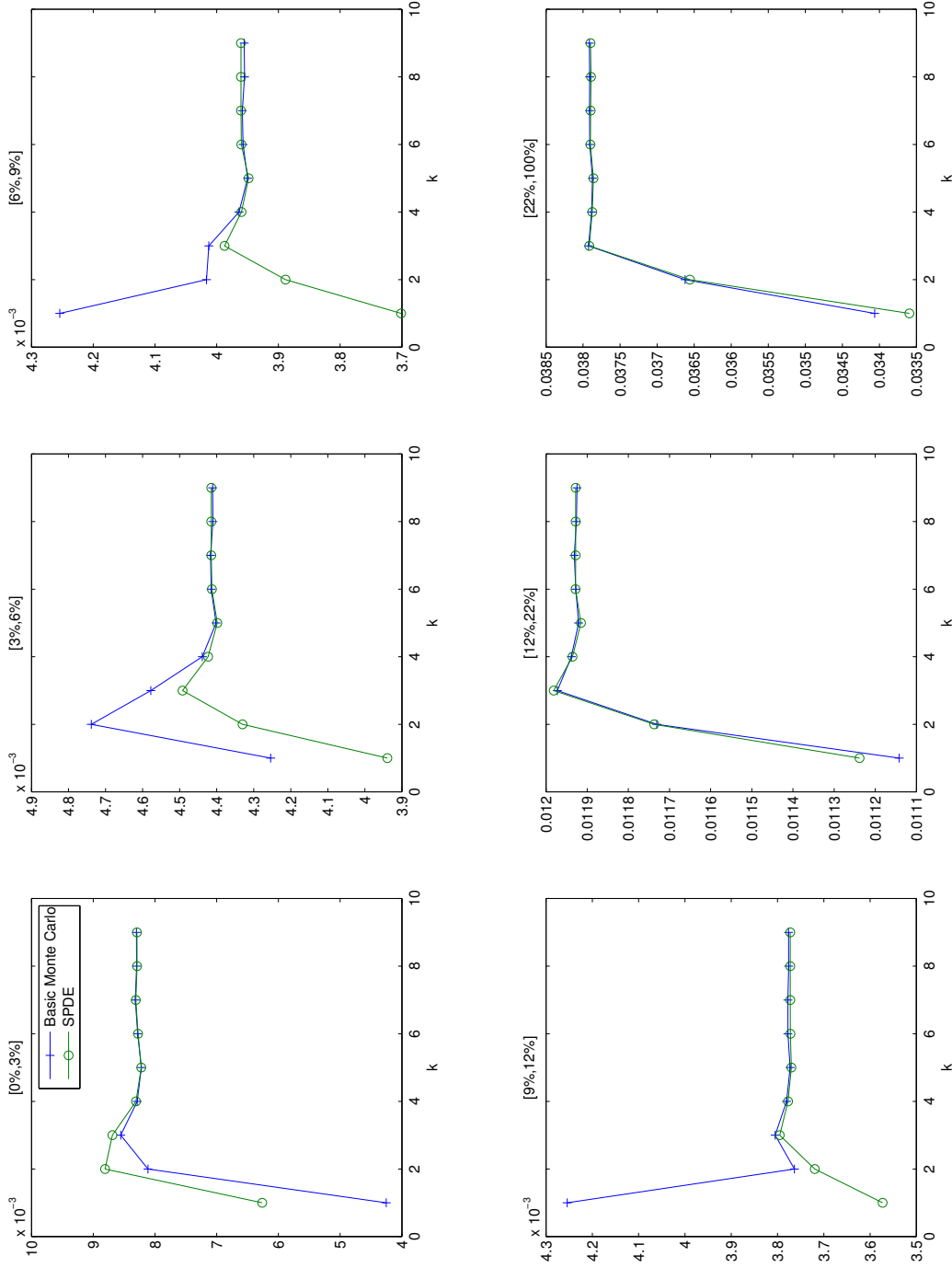
Figure 3: Expected losses in tranches [0%-3%], [3%-6%], [6%-9%], [9%-12%], [12%-22%], [22%-100%], calculated by direct Monte Carlo simulation of $N_k = 5^k$ asset value processes, $k = 1, \ldots, 9$, and the SPDE approximation, t=5 years, $X_0 \sim N(\mu_{X_0}, \sigma_{X_0}^2)$, where $\mu_{X_0} = 4.6$, $\sigma_{X_0} = 0.8$, and, for $t > 0$, X_t^i follows a jump diffusion process (4.2) with parameters obtained in the calibration exercise for 5 December 2008 from Section 6.

paths. We find empirically that

$$Var[\widehat{W}^{(j)}] \sim N_j^{-1}(N_{sims}^{(j)})^{-1}, \tag{4.23}$$

which is what one would expect: the $N_{sims}^{(j)}$ simulations are independent, and for each simulation, *conditional* on the common factors, the losses are a sum of independent random variables. Expressions (4.22) and (4.23) motivate to keep the product of the number of companies and number of samples asymptotically constant across all levels in order to obtain optimal overall complexity for fixed a variance ϵ^2. This means most samples are used on the coarsest levels (ie, small N_j) and the complexity spent on each level is approximately the same. With this choice, as the number k of levels and therefore estimators increases, and the required variance needed on each level is ϵ^2/k for overall variance ϵ^2, the overall complexity only grows proportional to $k^2 \sim (\log N_k)^2$ rather than proportional to N_k.

In fact, the optimal number of simulations can be obtained a posteriori for each level similar to the formulae in Giles (2008). The theoretical underpinning for this method and complexity bounds are the subject of current research and will be given in Bujok (2012).

5 SPDE Simulation of the Large Basket Model

The key ingredient in the calculation of index and tranche spreads are the basket losses. We ascertained in Section 4.3 that the large basket approximation from Section 4.2 is sufficiently accurate for the considered basket size. Therefore, the losses may be obtained from the large basket density v as per (4.20), where v has to be approximated by a numerical solution to the SPDE. In this section, we outline a numerical method based on Monte Carlo simulation on top of a PDE solver. In the presence of non-smooth (initial) data, and with senior tranche prices depending on very low probability events, a carefully adapted numerical scheme is necessary for good accuracy.

We discuss the calibration of model parameters, including initial conditions, in Section 6, and assume these as given here.

5.1 Setup and Finite Difference Discretisation

We can take advantage of the representation (4.11) of the solution to the SPDE problem (4.13) piecewise in time, as the boundary condition is not active in intervals (T_k, T_{k+1}), and therefore the Brownian driver and the jump part only introduce a shift to the solution, accumulated over the entire interval. Conditional on the number of jumps that occur in the interval, the shift is normally distributed. Using Proposition 4.2, a different density with identical losses can therefore defined by

$$v(t,x) = \begin{cases} 0 & x \leq 0, t = T_{k+1}, \\ v^{(k)}(t - T_k, x - \sqrt{\rho}(M_t - M_{T_k}) - (J_t - J_{T_k})), & \text{else if } t \in (T_k, T_{k+1}], \end{cases} \tag{5.1}$$

for $0 \leq k < n$, where $J_t = \sum_{i=1}^{N_t} \Pi_i$, and $v^{(k)}$ is the solution to the (deterministic) problem

$$v_t^{(k)} = \frac{1}{2}(1-\rho)v_{xx}^{(k)} - \beta v_x^{(k)}, \quad t \in (0, \Delta T) = (0, T_{k+1} - T_k) \tag{5.2}$$
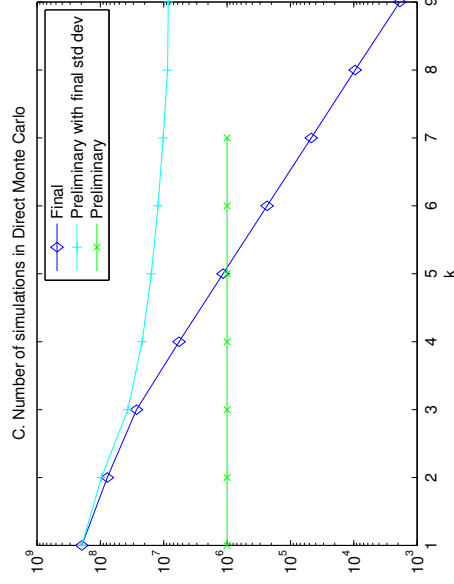
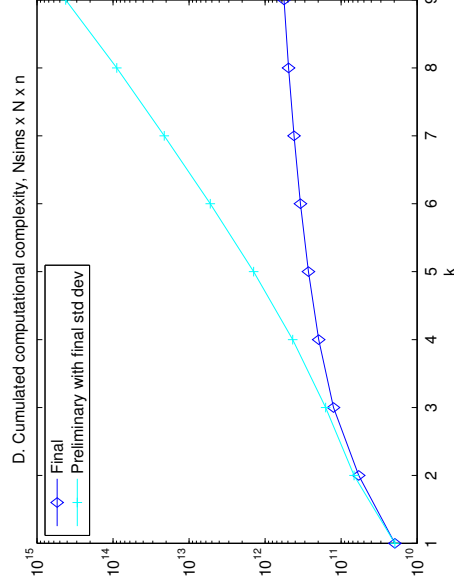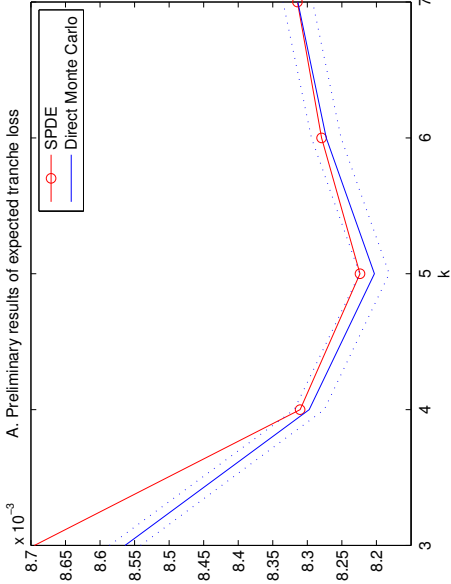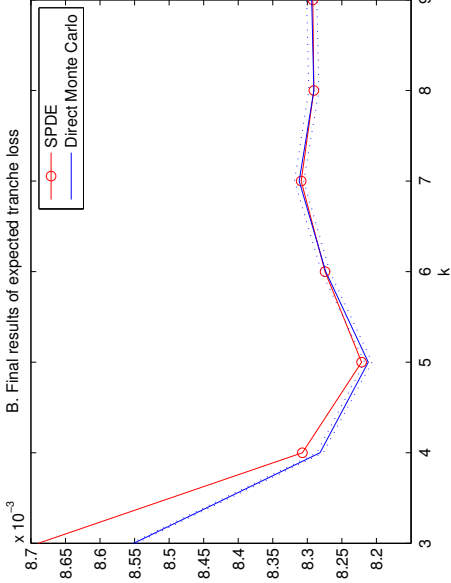$$v^{(k)}(0,x) = v(T_k, x) \tag{5.3}$$

16

Figure 4: "Preliminary" (ie, fixed number of simulations) and "final" (ie, adaptive number of simulations) results for the expected loss in the tranche [0%–3%] for baskets of size $N_k = 5^k$. A. For $N_{sims} = 10^6$, estimators together with 95% confidence intervals for the direct Monte Carlo method, compared to the SPDE results with similar confidence intervals. B. Results of the method based on direct *Multi-level* Monte Carlo, again with 95% confidence intervals, for a standard deviation of 4×10^{-6}. C. "Preliminary with final std dev" denotes the number of simulations needed for direct Monte Carlo (without *Multi-level* sampling) to achieve the standard deviation from above; "Final" denotes the number of simulations for each level of the multi-level approach. D. Number of simulations from C. times N_k times n.

assuming monitoring dates are equally spaced with intervals $\Delta T = T_{k+1} - T_k$.

This suggests the following inductive strategy for $k = 0, \ldots, n - 1$:

1. Start with $v^{(0)}(0, x) = v_0(x)$, the initial empirical measure.

2. Solve the PDE (5.2) numerically in the interval $(0, T_1)$, to obtain $v^{(0)}(T_1, x)$.

3. Simulate M_{T_1}, J_{T_1} evaluate $v(T_1, x)$ according to (5.1).

4. For $k > 0$, having computed $v(T_k, x)$ in the previous step, use this as initial condition for $v^{(k)}$, and repeat until $k = n$.

To solve the PDEs (5.2) by a finite difference method, we approximate the measure by one with support $[x_{min}, x_{max}]$, and set zero boundary conditions at x_{min} and x_{max}.

Proposition 5.1. *If v^b is the solution to (5.1)-(5.3), where (5.2) holds only on $(-b, b)$, with $v^{(k)}(t, -b) = v^{(k)}(t, -b) = 0$, then*

$$E[Z_t^b] \to E[Z_t] \quad \text{for} \quad b \to \infty,$$

where Z_t^b is the outstanding tranche notional derived from losses L_t^b of v^b.

Proof. See Appendix B. □

In practice, we pick x_{min} and x_{max} experimentally large enough such that the effect of the truncation is negligible.

Then, introduce a grid $x_0 = x_{min}, x_1 = x_{min} + \Delta x, \ldots, x_j = x_{min} + j\Delta x, \ldots, x_J = x_{min} + J\Delta x$, where $\Delta x = (x_{max} - x_{min})/J$, timesteps $t_0 = 0, t_1 = \Delta t, \ldots, t_I = I\Delta t$, where $\Delta t = \Delta T / I$, and define an approximation v_j^i to $v(t_i, x_j)$ by the difference scheme

$$\frac{v_j^i - v_j^{i-1}}{\Delta t} = \theta \left\{ \frac{1}{2}(1 - \rho) \frac{v_{j+1}^i - 2v_j^i + v_{j-1}^i}{\Delta x^2} - \beta \frac{v_{j+1}^i - v_{j-1}^i}{2\Delta x} \right\} \tag{5.4}$$

$$+ (1 - \theta) \left\{ \frac{1}{2}(1 - \rho) \frac{v_{j+1}^{i-1} - 2v_j^{i-1} + v_{j-1}^{i-1}}{\Delta x^2} - \beta \frac{v_{j+1}^{i-1} - v_{j-1}^{i-1}}{2\Delta x} \right\}. \tag{5.5}$$

For an introduction to finite differences in financial instrument pricing see, eg, Tavella & Randall (2000). The scheme is of second order accurate in Δx for smooth solutions. The Crank-Nicolson scheme $\theta = \frac{1}{2}$ is of second order accurate in Δt, and is unconditionally stable for sufficiently smooth solutions, but does not converge for initial conditions comprising of δ-distributions as in the present setting. We address this, together with aspects of approximating non-smooth initial and interface conditions (4.21) accurately on the grid, in the following section.

5.2   Initial and Interface Conditions

The initial condition has the form

$$v(0, x) = \frac{1}{N} \sum_{i=1}^{N} \delta(x - x^i),$$

where x^i is the observed initial distance-to-default of firm i.

To achieve second order accuracy in Δx, the δ's have to be "split" between adjacent grid points. The correct weighting can be written as integral of linear splines

$$\Phi_k(x) = \frac{1}{\Delta x} \min\left(\max(x - x_k + \Delta x, 0), \max(x_k + \Delta x - x, 0)\right)$$

over the initial condition, ie,

$$v_k^0 = \frac{1}{\Delta x} \int_{x_{min}}^{x_{max}} \Phi_k(x) v(0, x) \, dx.$$

Note that the initial condition is a probability measure and in particular the above definition is well defined for Dirac measures. By this construction, $\Delta x \sum_{k=0}^{J} v_k^0 = 1$. See also Pooley et al. (2003) for applications of this idea to option pricing and estimation of sensitivities.

To incorporate the interface condition (5.1) at $t = T_k$, one has to evaluate the grid function at shifted arguments $x_j - \sqrt{\rho}(M_t - M_{T_k}) - (J_t - J_{T_k})$, which do not normally coincide with grid points. To deal with this, we first define a piecewise linear interpolant $\widehat{v}$, obtained in the last step over the previous interval $[T_{k-1}, T_k]$, $v_j^{(k),I}$, I the number of timesteps, by

$$\widehat{v}(T_k, x) = \sum_{j=0}^{J} \Phi_j(x - \sqrt{\rho}\Delta M - \Delta J) v_j^{(k),I}, \tag{5.6}$$

where $\Delta M = M_{T_k} - M_{T_{k-1}}$, $\Delta J = J_{T_k} - J_{T_{k-1}}$. Then, approximate (5.1) by setting

$$v_j^{(k+1),0} = \int_{\max(x_j - \Delta x/2, 0)}^{\max(x_j + \Delta x/2, 0)} \widehat{v}(T_k, x) \, dx, \tag{5.7}$$

and use this as initial condition for the next interval. This ensures that

$$\Delta x \sum_{j=0}^{J} v_j^{(k+1),0} = \int_{0}^{x_{max}} \widehat{v}(T_k, x) \, dx,$$

so the cumulative density of firms with firm values greater than 0 is preserved. It also has the effect that the solution is smoothed at $x = 0$. In contrast to simpler, eg, pointwise application of the interface condition, this procedure guarantees second order convergence in Δx. See Pooley et al. (2003) for averaging procedures to restore higher order convergence for non-smooth payoffs in option pricing.

Finally, it is well-known that Crank-Nicolson timestepping has reduced convergence order for discontinuous initial conditions, and does not converge at all for Dirac initial conditions, unless the timesteps are chosen very small. This severely slows down the performance. A simple and well-established remedy is to replace the first Crank-Nicolson steps with backward Euler steps, a practice now known as "Rannacher start-up" Rannacher (1984). The analysis in Giles & Carter (2006) shows that the optimal balance between accuracy and stability is achieved by replacing the first two Crank-Nicolson steps by four backward Euler steps of half the stepsize. We do this at $t = 0$, and also at $t = T_k$, where the interface conditions introduce discontinuities at $x = 0$.

5.3 Loss Simulation

For a given realisation of the market factors, we can approximate the loss functional (4.20) at time T_k, using (5.6) and (5.7), by

$$\widehat{L}_{T_k} = (1 - R)\left(1 - \int_0^{x_{max}} \widehat{v}(T_k, x)\,dx\right) = (1 - R)\left(1 - \Delta x \sum_{j=1}^{J-1} v_j^{(k+1),0}\right). \qquad (5.8)$$

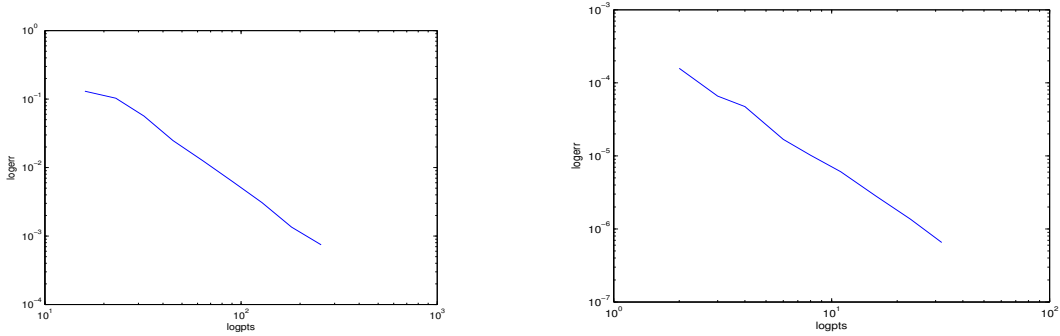We first study the discretisation error of (5.8) in Δt and Δx numerically, first for a single realisation of the path of the market factors. Fig. 5 shows, for a typical set of

Figure 5: Estimated discretisation error of $\widehat{L}_{T_n}$ for increasing J (left) and I (right) for a single realisation of the path of the market factors. The error estimator is based on Richardson extrapolation of the numerical solutions on subsequent refinement levels.

model parameters, how many grid points (I) and how many timesteps between payment dates (J) are necessary for a desired accuracy ϵ. We clearly see second order convergence in Δt and Δx. Note that the time smoothing scheme, weighted approximation of initial positions and adapted averaging, as per Section 5.2, are essential to achieve this.

Next, we want to compute expected losses and outstanding tranche notionals. If we explicitly include the dependence on the Monte Carlo samples ϕ_i of $\sqrt{\rho}(M_{T_i} - M_{T_{i-1}}) + J_{T_i} - J_{T_{i-1}}$ in (5.8) by writing $\widehat{L}_{T_k}(\phi)$, $\phi = (\phi_i)_{1 \le i \le n}$, where ϕ_i are independent, then for N_{sims} simulations with samples $\phi^l = (\phi_i^l)_{1 \le i \le n}$, $1 \le l \le N_{sims}$,

$$\mathbb{E}^{\mathbb{Q}}[Z_{T_k}] \approx \mathbb{E}^{\mathbb{Q}}[\max(d - \widehat{L}_{T_k}, 0) - \max(a - \widehat{L}_{T_k}, 0)]$$

$$\approx \frac{1}{N_{sims}} \sum_{l=1}^{N_{sims}} \left(\max(d - \widehat{L}_{T_k}(\phi^l), 0) - \max(a - \widehat{L}_{T_k}(\phi^l), 0)\right).$$

We expect (weak and strong) convergence for this estimator of order $\Delta x^2 + \Delta t^2$, which is confirmed by numerical tests. In the following, we fix Δx and Δt at values which have proven empirically to give negligible discretisation error.

For the following simulations, we have used two data sets: from 22 February 2007 and 5 December 2008. Initial positions for individual firms were calibrated to their individual CDS spreads, and were well within the range $[x_{min}, x_{max}] = [-10, 20]$.

We now analyse the convergence of the obtained Monte Carlo estimator. The estimated expected tranche losses with confidence intervals are given in Figure 6. For 22
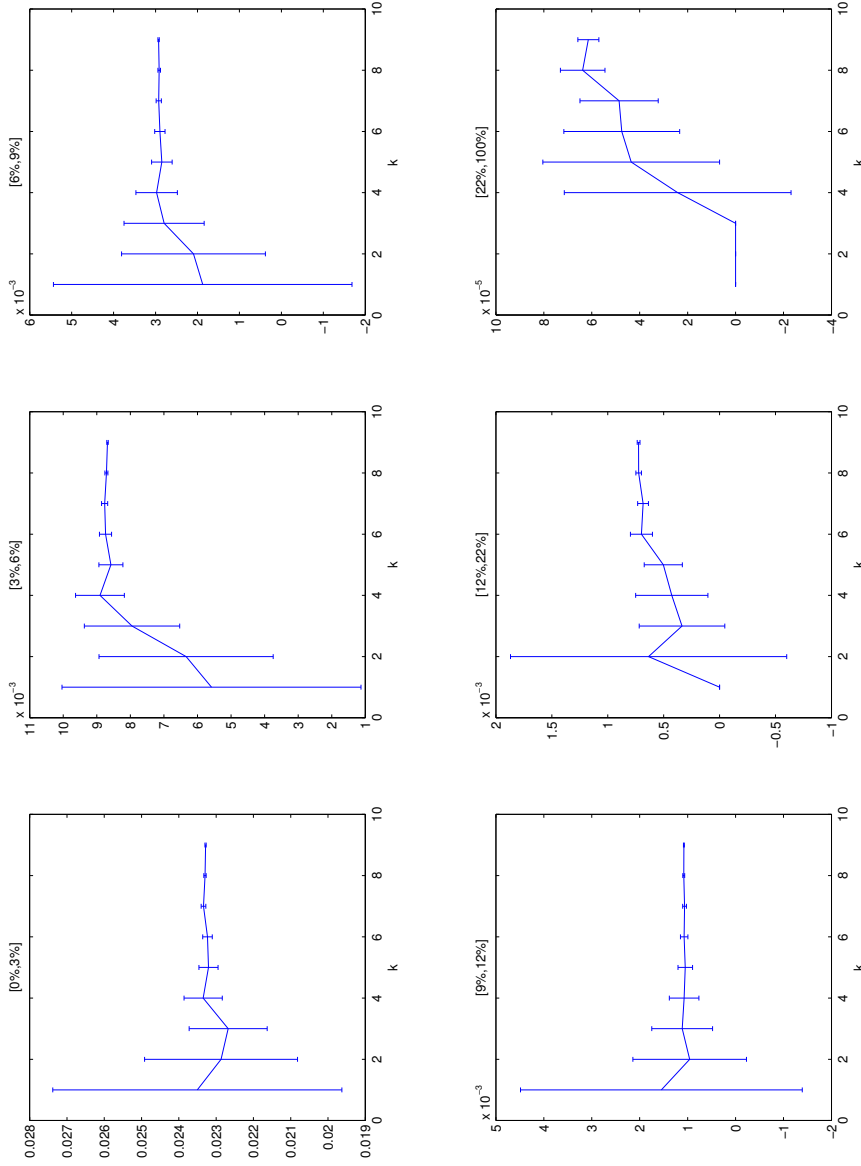
Figure 6: Monte Carlo estimators with 95%-confidence intervals for expected losses in tranches [0%-3%], [3%-6%], [6%-9%], [9%-12%], [12%-22%], [22%-100%], for $N = 4^{k+1}$ simulations, $k = 1, \ldots, 9$. The results are for parameters from a calibration of the model to data from 22 February 2007 (see Section 6).

February 2007, for super senior tranches especially, the standard error is higher relative to the values than for equity and mezzanine ones, especially. The reason behind this is that a large number of companies have to default in order to affect senior tranches and such an event is rare. For this date, the estimated parameters imply normal market conditions, where multiple defaults are highly unlikely (see Section 6 for a discussion of calibration results). In order to price senior tranches more accurately, variance reduction techniques such as importance sampling should be applied, see, eg, Carmona et al. (2009), however the accuracy of the results for the basic scheme was found sufficient for the purposes of this study.

6   Calibration and Results

The first step in using a model is to estimate its parameters by calibrating the model to market data. We fit the jump-diffusion model and, for comparison, the pure diffusion model, to CDO index and tranche spreads.

6.1   Stating the Calibration Problem

The initial distances-to-default of the individual names are thereby calibrated to their short-term CDS spreads, to ensure consistency of the multi-name CDO model with single-name CDS models. We fit the remaining parameters defined in Section 4.1 to index spreads for M maturities (usually 5-, 7- and 10-year, ie, $M = 3$) and G tranche spreads for each maturity (usually $G = 6$). The number of calibration prices, $M \cdot (G+1)$, is then typically much larger than the number of parameters, here 5. We therefore cannot expect the model to fit all prices exactly, and formulate the calibration exercise as a weighted least-squares problem.

Problem 1. *Given market spreads at time $t = 0$, of CDO tranches, $C_0^j(T_i)$, and the CDO index, $CI_0(T_i)$, for maturities T_i, $i = 1, \ldots, M$, tranches $j = 1, \ldots, G$, and given spreads $c_0 = \left(c_0^1, \ldots, c_0^N\right)$ for CDSs written on N underlying companies, solve the minimisation problem*

$$\sum_{i=1}^{M} \sum_{j=1}^{G} \alpha_i^j \left(C_0^j(T_i) - C_0^{j,\theta,x_0}(T_i) \right)^2 + \sum_{i=1}^{M} \alpha_i \left(CI_0(T_i) - CI_0^{\theta,x_0}(T_i) \right)^2 \quad \rightarrow \quad \min_\theta, \quad (6.1)$$

where $\theta = (\rho, \sigma, \lambda, \sigma_\Pi, \mu_\Pi)$, subject to

(i) $x_0 = \left(x_0^1, \ldots, x_0^N\right)$ is a solution to

$$c_0 = c_0^\theta(x_0), \quad (6.2)$$

(ii) $\rho \in [0, 1)$, $\sigma > 0$, $\lambda > 0$, $\sigma_\Pi > 0$, $\mu_\Pi < 0$,

where c^θ, C_0^{j,θ,x_0}, CI_0^{θ,x_0} denote CDS, CDO tranche and index spreads calculated in the model with parameter vector θ, x_0 is a vector of initial distances-to-default, $\alpha = (\alpha_i^j, \alpha_i)$ is a scaling vector.

22

The minimisation problem has a scaled least-squares objective function, linear inequality constraints (ii) and non-linear equality constraints (i).

Since CDO tranche and index spreads have different orders of magnitudes, we scale the data by α to make each observation roughly equally important, precisely, we choose the weights α as powers of ten such that the scaled market prices lie between 0.1 and 1. Alternative choices are discussed in Section 6.4.

The non-linear constraints are incorporated directly into (6.1) by numerical inversion of (6.2) giving $x_0 = x_0(\sigma, \lambda, \sigma_\Pi, \mu_\Pi) = x_0(\theta)$. We then solve the minimisation problem (6.1) with bounds on the parameters. A robust algorithm developed specially for this kind of minimisation problems is the interior-reflective Newton method in Coleman & Li (1994) and Coleman & Li (1996). We use this algorithm as implemented in the Matlab Optimisation Toolbox.

6.2 Computational Issues

For each calibration, a number γ of Newton-type iterations is needed to solve (6.1) to a specified accuracy, and in each iteration, an approximation to the objective function and its derivatives is required. We address here ways of improving the computational efficiency of the most costly components.

Search for Initial *Distances-To-Default*

Since we use finite differences to approximate the derivatives, x_0^n is searched $K = (1 + 2b) \times N \times \gamma$ times, where $n = 1, \ldots, N = 125$, the number of CDSs in the portfolio, $b = 4$ is the number of parameters (excluding ρ, which does not affect x_0), and γ is in the range of $10 - 20$. Since $K \approx 1.5 \times 10^4$, an efficient way of finding x_0 from (6.2) is crucial.

Given CDS data, $c_0 = (c_0^i)_{1 \leq i \leq N}$, we search for initial *distances-to-default*, $x_0 = (x_0^i)_{1 \leq i \leq N}$, taking advantage of properties of c_0. CDS spreads decrease with the initial *distance-to-default*, since higher x_0 leads to higher survival probability, which entails lower spreads, ie, if $c_0^i > c_0^j$ then $x_0^i < x_0^j$. Hence, for a CDS spread, c_0^s, where $c_0^{\min} \leq c_0^{s-1} < c_0^s < c_0^{s+1} \leq c_0^{\max}$, the initial *distance-to-default*, x_0^s, is in (x_0^{s+1}, x_0^{s-1}). Hence, a good starting point for x_0^s is either x_0^{s+1} or x_0^{s-1}. Sorting spreads, then choosing starting points as stated above, greatly decreases the computation time of x_0.

Monte Carlo "Inside" the Objective Function

What is more, the objective function (6.1) is estimated by a Monte Carlo method, and in order to obtain accurate results, a high number of simulations is needed, as seen from the accuracy of expected tranche losses in Fig. 6. Given the numerical evaluation of the objective function and its derivatives is part of an iterative solution method for the optimisation problem, where the initial parameter iterates will be inaccurate, it is unnecessary to evaluate the spreads there with high accuracy.

Let $N_i, i = 1, \ldots, Z$, be an increasing number of paths, and let $\theta_{N_i}^i$ be a sequence of parameters obtained using N_i simulations, with an iterative scheme with starting value $\theta_{N_{i-1}}^{i-1}$ and γ_i iterations. Since the Monte Carlo estimator and Newton's method both converge, it is hoped that the $\theta_{N_i}^i$ converge to the optimiser, while only few iterations are needed for large N_i. This idea is a simple version of the *Multi-layer method* in Kaebe et al. (2009).

In the calibration exercises, we heuristically picked $N_1 = 6 \times 10^3$, since for this number of simulations the estimator "starts converging", as can be observed in Fig. 6, and the computation time per iteration is very low. For N_1, γ_1 is about 15, whereas for $N_3 = N_{Z-1} = 10^5$, it is about 2-4, and, finally, for $N_4 = 10^6$, it is only 1. The overall computation time of the calibration algorithm is significantly reduced compared to calibration with uniformly N_Z simulations, here by a factor of roughly 15.

It would be possible to automatise this procedure and find an optimal sequence N_i to minimize the overall computational time, based on the variance of the Monte Carlo estimator and the convergence speed of the optimisation algorithm.

6.3 Calibration Results

In order to check if the model is flexible enough to match market spreads both in quiet and extreme market conditions, we calibrate it to pre-crisis data, from 22 February 2007, and during the crisis, 5 December 2008. We do not use current data since the market is not sufficiently liquid.

For the first data set (see Table 1), spread curves both for index and tranches are increasing with time, super senior tranches (22%-100%) are close to zero, indicating that the market does not expect any shocks in the near future. For the second period (see Table 4), all quotes are much higher (for example prices of super senior tranches are higher than junior mezzanine (3%-6%) with maturity 5 years for 22 February 2007), the index spread curve is inverted, curves for tranches are almost flat, hence the market anticipates that the situation will get worse in the near future. We anticipate that parameters indicating market turbulences will be considerably higher for the second data set.

We compare the calibration results by the ARPE (Average Relative Percentage Error),

$$ARPE = \frac{1}{s}\sum_{i=1}^{s}\frac{|y_i - y_i^\theta|}{y_i}, \tag{6.3}$$

and the RMSE (Root Mean Square Error),

$$RMSE = \sqrt{\frac{1}{s}\sum_{i=1}^{s}(y_i - y_i^\theta)^2}, \tag{6.4}$$

where y is a vector of observed prices, y^θ is a vector of prices obtained from the model, s the total number of calibration prices. It is important to note that the ARPE can be used to compare calibration results between data sets with different orders of magnitudes, while the RMSE can be used for comparing calibration results within the same or similar data sets and will attach disproportionate weight to the equity tranche spreads. Also recall that the calibration optimises (6.1).

Diffusion Model

Calibration results for the diffusion model are given in Tables 1 to 6. Root mean square errors (RMSE) are very high, especially for 5 December 2008, but as was noted before, this is partially a result of higher spreads for 5 Dec 2008 compared to 22 Feb 2007.

For 5 Dec 2008, the estimated correlation coefficient, ρ, is much higher than the one obtained for 22 Feb, 2007. High correlation not only increases the probability of

T=5 Years			
	Market	Jump-Diffusion	Diffusion
Index	21	24	17
Tranche 0%-3%	7.19%	11.07%	1.22%
Tranche 3%-6%	41	61	48
Tranche 6%-9%	10.8	15.2	10
Tranche 9%-12%	5	5.9	3
Tranche 12%-22%	1.8	2.5	0
Tranche 22%-100%	0.9	1.8	0
T=7 Years			
	Market	Jump-Diffusion	Diffusion
Index	30	32	25
Tranche 0%-3%	22.1%	24.7%	12.1%
Tranche 3%-6%	110	155	137
Tranche 6%-9%	32.5	39.5	41
Tranche 9%-12%	15	14	14
Tranche 12%-22%	4.9	3.9	2.0
Tranche 22%-100%	2	1.8	0
T=10 Years			
	Market	Jump-Diffusion	Diffusion
Index	41	37	32
Tranche 0%-3%	38%	35.4%	21.9%
Tranche 3%-6%	302.5	281.5	247
Tranche 6%-9%	83	87	97
Tranche 9%-12%	37	32	41
Tranche 12%-22%	12.5	7.3	9
Tranche 22%-100%	3.6	1.8	0

Table 1: Calibration results for 22 February 2007, iTraxx Main Series 6 index. Units are basis points (bps) if not stated otherwise. Estimated parameters are given in Table 2, measures of fit in Table 3. We assume that $r = 0.042$, $R = 0.4$.

Parameters	Jump-Diffusion	Diffusion
σ	0.16	0.18
ρ	0.11	0.22
λ	0.04	-
E[Y-1]	-0.07	-
Var[Y-1]	0.01	-
$\rho_{X_t^i X_t^j}$	0.13	-

Table 2: Parameters estimated for 22 February 2007, overall correlation $\rho_{X_t^i X_t^j}$ as in (4.5).

Measure of Fit	Jump-Diffusion	Diffusion
RMSE	1.17	4.34
ARPE	0.27	0.42

Table 3: Measures of fit for 22 February 2007. Root Mean Square Error (RMSE) is defined by (6.4), Average Relative Percentage Error (ARPE) is given by (6.3).

companies to default together, but also to survive together. Hence, in order to fit equity tranches, a low correlation is needed, whereas high correlation for the senior tranches. This can be observed in Figure 3 in Bush et al. (2011), showing the implied correlation for each tranche. Hence, in a diffusion model, there is usually a trade-off between matching equity and senior tranches. The low correlation coefficient for 22 Feb 2007 results in zero spreads for super senior tranches ($22\% - 100\%$), while equity tranches are underpriced. For 5 Dec 2008, the model produces non-zero spreads for super senior tranches, owing to a higher correlation coefficient, but still both super senior and equity tranches are underpriced.

Jump-Diffusion Model

Tables 1 to 6 show calibration results for the jump-diffusion model. Compared to the diffusion model, error measures for both data sets are much lower. What is more, the jump-diffusion model gives spreads of roughly the correct magnitude for all tranches. However, still some tranches are slightly underpriced, whereas other tranches are somewhat overpriced, albeit to a lesser extent than in the diffusion case.

In the jump-diffusion model, apart from ρ and σ, also λ, μ_Y and σ_Y affect multiple defaults. Most importantly, unlike ρ, higher absolute values of jump parameters, with $\mu_X < 0$, always lead to more expected defaults. Therefore, the trade-off between matching particular tranches is smaller and the overall fit is better in the jump-diffusion model.

Let us now analyse the situation on the market in 22 Feb 2007 and 5 Dec 2008, implied by the model parameters. As mentioned in 4.1, the initial *distances-to-default*, obtained from CDS spreads, indicate the current state of the market. As expected, on average, x_0 is lower for December 2008, around 4.0, whereas for February 2007, it is 4.3. It may seem a bit surprising that the difference is so small, though a better picture is given by the tails of the distribution of initial *distances-to-default*, as seen back in Fig. 2. For the former data set, for 6% of the companies x_0 is in (0,1], ie, 6% of the companies in the portfolio are already in a bad financial situation, and for 26% x_0 is in (0,2], whereas for none of the firms in the latter data set x_0 is in (0,2]. This well reflects the situation in the market during a crisis and in normal conditions.

The jump intensity, λ, is 0.042 for 22 Feb 2007 and 0.035 for 5 Dec 2008 respectively. This implies that under the risk neutral measure a market crash is expected to happen every 25 years for 22 Feb 2007, and every 34 years for 5 Dec 2008. For the former data set, market turbulences will be benign: if a jump occurs, on average a company's assets value will fall by 7%. For the latter, market conditions will worsen much more severely: if a jump occurs, a company's asset value will drop on average by 50%. Both results agree with the adaptive expectation hypothesis, according to which in a quiet economic conditions market participants anticipate that the good economic climate will be continued, while after a serious market crash, another serious crash is expected.

T=5 Years			
	Market	Jump-Diffusion	Diffusion
Index	215	245	200
Tranche 0%-3%	71.5%	76.5%	47%
Tranche 3%-6%	1576.3	1652	1158
Tranche 6%-9%	811.5	822	825
Tranche 9%-12%	506.1	504	636
Tranche 12%-22%	180.3	305	415
Tranche 22%-100%	77.9	77	56
T=7 Years			
	Market	Jump-Diffusion	Diffusion
Index	195	224	198
Tranche 0%-3%	72.9%	78.8%	50%
Tranche 3%-6%	1473.2	1557	1101
Tranche 6%-9%	804.2	805	805
Tranche 9%-12%	512.4	511	637
Tranche 12%-22%	182.6	311	434
Tranche 22%-100%	75.8	74	66
T=10 Years			
	Market	Jump-Diffusion	Diffusion
Index	175	204	188
Tranche 0%-3%	73.8%	81.2%	51.7%
Tranche 3%-6%	1385.5	1462	1016
Tranche 6%-9%	824.7	768	755
Tranche 9%-12%	526.1	500	607
Tranche 12%-22%	174.1	310	427
Tranche 22%-100%	76.3	69	70

Table 4: Calibration results for 5 December 2008, iTraxx Main Series 10 index. Units are basis points (bps) if not stated otherwise. Estimated parameters are given in Table 5, measures of fit in Table 6. We assume that $r = 0.033$, $R = 0.4$.

Parameters	Jump-Diffusion	Diffusion
σ	0.13	0.18
ρ	0.35	0.8
λ	0.04	-
E[Y-1]	-0.5	-
Var[Y-1]	0.17	-
$\rho_{X_t^i X_t^j}$	0.83	-

Table 5: Parameters estimated for 5 Dec, 2008, overall correlation $\rho_{X_t^i X_t^j}$ as in (4.5).

Measure of Fit	Jump-Diffusion	Diffusion
RMSE	2.42	8.95
ARPE	0.16	0.35

Table 6: Measures of fit for 5 December 2008. Root Mean Square Error (RMSE) is defined by (6.4), Average Relative Percentage Error (ARPE) is given by (6.3).

To our knowledge, a similar calibration exercise is found only in Willemann (2007) and Kiesel & Scherer (2007), however the latter study does not discuss the parameters obtained. In Willemann (2007), a jump-diffusion model with log-normally distributed jumps was calibrated to market data from 8 August 2005. The obtained jump intensity is very small, $\lambda = 0.0012$, while the mean jump size is -96%. Under these settings, on average there is a complete market wipe-out every 870 years. Such a market situation would appear hard to reconcile with the timeframes market participants operate in, or even the age of credit markets and the economy as we know it.

6.4 Further Results and Extensions

Here, we comment on some aspects of the practical use of the model.

First, we remark that in the course of the calibration we calculate sensitivities of the prices with respect to the input parameters, which can be used for hedging with other products exposed to the same variables.

In practice, a common hedge is the index hedge, and in that case it is important to be able to calibrate the model accurately to index spreads. The results of Tables 1 and 6 are unlikely to be good enough for this, in contrast to those in Bush et al. (2011), where the volatility is first calibrated to the index spreads, and *then* the correlation (as single remaining parameter) is fit to tranche spreads, rather than performing an overall fit. A better calibration to specific important data, say the index spreads, can be enforced by increasing those weights in the least-squares functional. An interpretation of this is as a penalty method for solving a least-squares optimisation problem for the remaining data, where the penalty terms approximate equality constraints for the important data.

Table 7 shows that by increasing the weights of the index in the objective function (Cases 1 and 2), the market observed spreads can be fit accurately, especially when decreasing the weight of the super-senior tranche (Case 1). This is of course at the expense of worse fit of tranche spreads (especially the last tranche in Case 1). Calibrated parameter values and measures of fit are given in Tables 8 and 9, respectively.

7 Conclusions

In this article, we present an extension of the method proposed in Bush et al. (2011) and discuss in detail its numerical implementation.

We hereby consider a multi-name jump-diffusion, show how the model can be calibrated to individual CDSs by means of semi-analytical formulae for survival probabilities, derive a numerical method for basket credit derivatives based on a large basket approximation and a Monte Carlo finite difference solution of the resulting SPDE, and outline an algorithm for calibration to CDO index and tranche spreads. We also show that the

T=5 Years			
	Market	Case 1	Case 2
Index	215	214.6	222.0
Tranche 0%-3%	71.5%	81.6%	80.2%
Tranche 3%-6%	1576.3	1775.1	1675.4 9
Tranche 6%-9%	811.5	763.7	726.4
Tranche 9%-12%	506.1	419.9	417.2
Tranche 12%-22%	180.3	231.9	252.4
Tranche 22%-100%	77.9	12.7	36.2
T=7 Years			
	Market	Case 1	Case 2
Index	195	193.8	198.1
Tranche 0%-3%	72.9%	83.7%	82.0%
Tranche 3%-6%	1473.2	1736.4	1575.1
Tranche 6%-9%	804.2	781.6	710.7
Tranche 9%-12%	512.4	448.2	422.5
Tranche 12%-22%	182.6	241.6	255.1
Tranche 22%-100%	75.8	7.6	30.2
T=10 Years			
	Market	Case 1	Case 2
Index	175	174.6	175.3
Tranche 0%-3%	73.8%	85.4%	84.2%
Tranche 3%-6%	1385.5	1691.9	1488.8
Tranche 6%-9%	824.7	781.3	679.7
Tranche 9%-12%	526.1	452.6	409.3
Tranche 12%-22%	174.1	246.3	247.9
Tranche 22%-100%	76.3	3.4	23.7

Table 7: Calibration results for different weight vectors for 5 December 2008, iTraxx Main Series 10 index. Case 1: $\alpha_i = 10\,(CI_0(T_i))^{-2}$, $\alpha_i^j = (C_0^j(T_i))^{-2}$, $j = 1,\ldots,5$, $\alpha_i^6 = 0.01\,(C_0^6(T_i))^{-2}$. Case 2: $\alpha_i = 10\,(CI_0(T_i))^{-2}$, $\alpha_i^j = (C_0^j(T_i))^{-2}$, $j = 1,\ldots,6$. Estimated parameters are given in Table 8, measures of fit in Table 9. We assume that $r = 0.033$, $R = 0.4$.

Parameters	Case 1	Case 2
σ	0.11	0.10
ρ	0.25	0.27
λ	0.042	0.0397
E[Y-1]	-0.16	-0.34 6
Var[Y-1]	0.82	0.24
$\rho_{X_t^i X_t^j}$	0.85	0.82

Table 8: Parameters estimated for 5 December 2008, overall correlation $\rho_{X_t^i X_t^j}$ as in (4.5).

Measure of Fit	Case 1	Case 2
RMSE	4.24	3.64
ARPE	0.25	0.22

Table 9: Measures of fit for 5 December 2008. Root Mean Square Error (RMSE) is defined by (6.4), Average Relative Percentage Error (ARPE) is given by (6.3).

large basket approximation is valid for a typical CDO portfolio. We provide an economic interpretation of the model and its pricing performance.

The jump-diffusion model proves more flexible than the pure diffusion model to fit market data in vastly different scenarios, with only a small number of parameters. We conclude that the method proposed in Bush et al. (2011) has potential for practical applications. However, to employ the model for pricing and risk management in practice, it is necessary to further extend the model to allow a richer dependence and term structure. This can be achieved in a number of ways.

The extension to different finite activity jump processes, or even infinite activity, assuming discrete default monitoring, is straightforward. The latter could potentially be achieved by a time-change of the SPDE, were applicable. A stochastic volatility common to all assets, correlated, eg, to the driving Brownian market factor, can be included straightforwardly and would only require the simulation of an SDE for the variance process and the solution of a PDE of the type (4.10) with this simulated variance as diffusion coefficient. The inclusion of stochastic interest rates would follow similar lines.

Contagion effects can be introduced by making the correlation coefficient loss dependent. Similar to local volatility models, such a "local correlation model" could be specified to provide a perfect fit to tranche spreads, while the correlation does not impact index and CDS spreads, which aids calibration. Contagion could also manifest in loss dependent (and hence stochastic) recovery rates. The jump component will give a more realistic term-structure for short maturities. The model could be further extended by a random default barrier, which has been shown to give more realistic short-term credit spreads and would be an obvious and simple extension to the model.

With such improved fit to standard tranches, the calibrated model would be suitable for interpolating bespoke tranches on the same underlying pool. Applications to over-the-counter basket products consisting of illiquid names would be conceivable if information on the individual default probabilities could be obtained from, eg, balance sheets; if the company is listed, parameters may be obtained from equity time series; finally, parameters for global factors may be obtained from liquid products of names operating in similar sectors.

References

Bielecki, T. & Rutkowski, M. (2002), *Credit Risk: Modeling, Valuation and Hedging*, Springer Verlag, Berlin Heidelberg New York.

Black, F. & Cox, J. (1976), 'Valuing corporate securities: Some effects of bond indenture provisions', *Journal of Finance* **31**(2).

Brigo, D. & Mercurio, F. (2007), *Interest Rate Models – Theory and Practice*, Springer.

Broadie, M., Glasserman, P. & Kou, S. (1997), 'A continuity correction for discrete barrier options', *Mathematical Finance* **7**(4).

Bujok, K. (2012), Numerical Solution of a Class of SPDEs in Finance, PhD thesis, University of Oxford. In preparation.

Bush, N., Hambly, B., Haworth, H., Jin, L. & Reisinger, C. (2011), 'Stochastic evolution equations in portfolio credit modelling', *SIAM Journal on Financial Mathematics* **2**(1), 627–664.

Carmona, R., Fouque, J.-P. & Vestal, D. (2009), 'Interacting particle systems for the computation of rare credit portfolio losses', *Finance and Stochastics* **13**(4), 613–633.

Coleman, T. & Li, Y. (1994), 'On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds', *Mathematical Programming* **67**(2).

Coleman, T. & Li, Y. (1996), 'An interior trust region approach for nonlinear minimization subject to bounds', *SIAM Journal on Optimization* **6**(2).

Duffie, D. & Lando, D. (2001), 'Term structures of credit spreads with incomplete accounting information', *Econometrica* **69**(3), 633–664.

Fang, F., Jönsson, H., Oosterlee, C. & Schoutens, W. (2010), 'Fast valuation and calibration of credit default swaps under Lévy dynamics', *Journal of Computational Finance* **14**(2).

Finger, C., Finkelstein, V., Pan, G., Lardy, J.-P., Ta, T. & Tierney, J. (2002), 'Creditgrades. Technical document', *Risk Metrics Group, Inc.* .

Fouque, J., Wignall, B. & Zhou, X. (2008), 'Modeling correlated defaults: first passage model under stochastic volatility', *Journal of Computational Finance* **11**(3), 43–78.

Giles, M. (2008), 'Multi-level Monte Carlo path simulation', *Operations Research* **56**(3).

Giles, M. & Carter, R. (2006), 'Convergence analysis of Crank-Nicolson and Rannacher time-marching', *Journal of Computational Finance* **9**(4).

Haworth, H. & Reisinger, C. (2007), 'Modelling basket credit default swaps with default contagion', *Journal of Credit Risk* **3**(4), 31–67.

Hilberink, B. & Rogers, L. (2002), 'Optimal capital structure and endogenous default', *Finance and Stochastics* **6**(2).

Hu, X. & Ye, Z. (2007), 'Valuing credit derivatives in a jump-diffusion model', *Applied Mathematics and Computation* **190**(1).

Hull, J., Predescu, M. & White, A. (2010), 'The valuation of correlation-dependent credit derivatives using a structural model', *Journal of Credit Risk* **6**(3).

Kaebe, C., Maruhn, J. H. & Sachs, E. W. (2009), 'Adjoint based Monte Carlo calibration of financial market models', *Journal of Finance and Stochastics* **13**, 351–379.

Kiesel, R. & Scherer, M. (2007), Dynamic credit portfolio modelling in structural models with jumps, Preprint, Universität Ulm.

Kou, S. (2002), 'A jump-diffusion model for option pricing', *Management Science* **48**(8).

Kou, S. & Wang, H. (2003), 'First passage times of a jump diffusion process', *Advances in Applied Probability* **35**(2).

Kurtz, T. & Xiong, J. (1999), 'Particle representations for a class of nonlinear SPDEs', *Stochastic Processes and Applications* **83**, 103–126.

Lipton, A. & Sepp, A. (2009), 'Credit value adjustment for credit default swaps via the structural default model', *The Jounal of Credit Risk* **5**(2).

Merton, R. (1974), 'On the pricing of corporate debt: the risk structure of interest rates', *Journal of Finance* **29**(2).

Merton, R. (1976), 'Option pricing when underlying stock returns are discontinuous', *Journal of Financial Economics* **3**(1-2).

Pooley, D. M., Vetzal, K. R. & Forsyth, P. A. (2003), 'Remedies for non-smooth payoffs in option pricing', *Journal of Computational Finance* **6**(4).

Ramezani, C. & Zeng, Y. (2007), 'Maximum likelihood estimation of the double exponential jump-diffusion process', *Annals of Finance* **3**(4).

Rannacher, R. (1984), 'Finite element solution of diffusion problems with irregular data', *Numerische Mathematik* **43**(2).

Schönbucher, P. (2003), *Credit Derivatives Pricing Models. Models, Pricing and Implementation*, John Wiley & Sons Ltd.

Tavella, D. & Randall, C. (2000), *Pricing Financial Instruments – The Finite Difference Method*, Wiley.

Willemann, S. (2007), 'Fitting the CDO correlation skew: A tractable structural jump-diffusion model', *The Journal of Credit Risk* **3**(1).

Zhou, C. (1997), 'A jump-diffusion approach to modelling credit risk and valuing defaultable securities', *Finance and Economics Discussion Series, Board of Governors of Federal Reserve System* .

A   Proofs of Results in Section 3.3

Proof of Proposition 3.1.

$$
\begin{aligned}
p(X_{t_n} = x_{t_n}, \tau > t_{n-1}) &= p(X_{t_n} = x_{t_n}, X_{t_{n-1}} > 0, \ldots, X_{t_1} > 0) \\
&= p(X_{t_n} = x_{t_n}, X_{t_{n-1}} > 0, \tau > t_{n-2}) \\
&= \int_0^\infty p(X_{t_n} = x_{t_n}, X_{t_{n-1}} = x_{t_{n-1}}, \tau > t_{n-2}) \, dx_{t_{n-1}}.
\end{aligned}
$$

Since

$$p(X_{t_n} = x_{t_n} \mid X_{t_{n-1}} = x_{t_{n-1}}, \tau > t_{n-2}) = \frac{p(X_{t_n} = x_{t_n}, X_{t_{n-1}} = x_{t_{n-1}}, \tau > t_{n-2})}{p(X_{t_{n-1}} = x_{t_{n-1}}, \tau > t_{n-2})},$$

we obtain

$$p(X_{t_n} = x_{t_n}, \tau > t_{n-1}) = \int_0^\infty p(X_{t_n} = x_{t_n} \mid X_{t_{n-1}} = x_{t_{n-1}}, \tau > t_{n-2}) p(X_{t_{n-1}} = x_{t_{n-1}}, \tau > t_{n-2}) dx_{t_{n-1}}.$$

Observe now that for two subsequent monitoring times t_{n-1} and t_n, such that $0 \leq t_n \leq T$, and $x_0 > 0$, (3.4) can be written as

$$\begin{aligned} X_{t_n} &= x_0 + \beta t_n + W_{t_n} + \sum_{k=1}^{N_{t_n}} \Pi_k \\ &= X_{t_{n-1}} + \beta \Delta t_n + \Delta W_{t_n} + \sum_{k=1}^{\Delta N_{t_n}} \Pi_k, \end{aligned} \qquad (A.1)$$

where $\Delta t_n = t_n - t_{n-1}$, $\Delta W_{t_n} = W_{t_n} - W_{t_{n-1}}$, $\Delta N_{t_n} = N_{t_n} - N_{t_{n-1}}$. By independence of increments of Brownian motion and Poisson process respectively, ΔW_{t_n} is independent of $W_{t_{n-1}}$, ΔN_{t_n} is independent of $N_{t_{n-1}}$. What is more, $\Delta W_{t_n} \sim N(0, \Delta t_n)$ and ΔN_{t_n} is a Poisson process with intensity $\lambda \Delta t_n$. Hence, knowing that $X_{t_{n-1}} = x_{t_{n-1}}$, we have

$$X_{t_n} = x_{t_{n-1}} + \beta \Delta t_n + \Delta W_{t_n} + \sum_{k=1}^{\Delta N_{t_n}} \Pi_k. \qquad (A.2)$$

Then, by (3.5) and (A.2), we obtain

$$X_{t_n} = \begin{cases} x_{t_{n-1}} + \beta \Delta t_n + \Delta W_{t_n} + \sum_{k=1}^{\Delta N_{t_n}} \Pi_k, & x_{t_{n-1}} > 0 \\ 0, & x_{t_{n-1}} = 0. \end{cases} \qquad (A.3)$$

Hence,

$$p(X_{t_n} = x_{t_n} \mid X_{t_{n-1}} = x_{t_{n-1}}, \tau > t_{n-2}) = p(X_{t_n} = x_{t_n} \mid X_{t_{n-1}} = x_{t_{n-1}}),$$

and we get

$$p(X_{t_n} = x_{t_n}, \tau > t_{n-1}) = \int_0^\infty p(X_{t_n} = x_{t_n} \mid X_{t_{n-1}} = x_{t_{n-1}}) p(X_{t_{n-1}} = x_{t_{n-1}}, \tau > t_{n-2}) dx_{t_{n-1}}.$$

Similarly,

$$P(\tau > t_n) = P(X_{t_n} > 0, \tau > t_{n-1}) = \int_0^\infty p(X_{t_n} = x_{t_n}, \tau > t_{n-1}) \, dx_{t_n}.$$

Proof of Proposition 3.2. *Since, t_1 is the first monitoring time, we have*

$$P(\tau > t_1) = P(X_{t_1} > 0).$$

33

By total probability

$$P(X_{t_1} > 0) = \sum_{c=0}^{\infty} P(X_{t_1} > 0 \mid N_{t_1} = c) \cdot p(N_{t_1} = c).$$

We observe that $\sum_{k=1}^{c} \Pi_k$, where $\{\Pi_k\}$ are i.i.d. and $\Pi \sim N(\mu_\Pi, \sigma_\Pi^2)$, is normally distributed with mean $c\mu_\Pi$ and variance $c\sigma_\Pi^2$. By (A.1), condtional on $N_{t_1} = c$, $X_{t_1} \sim N(\mu_{X_{t_1}}, \sigma_{X_{t_1}}^2)$, where $\mu_{X_{t_1}} = x_0 + \beta t_1 + c\mu_\Pi$, $\sigma_{X_{t_1}}^2 = t_1 + c\sigma_\Pi^2$. Then,

$$P(X_{t_1} > 0 \mid N_{t_1} = c) = \Phi\left(\frac{\mu_{X_{t_1}}}{\sigma_{X_{t_1}}}\right).$$

What is more, N_{t_1} is a Poisson process with intensity λt_1, hence

$$p(N_{t_1} = c) = e^{-\lambda t_1} \frac{(\lambda t_1)^c}{c!},$$

and we obtain (3.7). Observe now that,

$$P(X_{t_1} \le x_{t_1} \mid N_{t_1} = c) = P\left(Z \le \frac{x_{t_1} - \mu_{X_{t_1}}}{\sigma_{X_{t_1}}}\right) = \Phi\left(\frac{x_{t_1} - \mu_{X_{t_1}}}{\sigma_{X_{t_1}}}\right),$$

where $Z \sim N(0,1)$, $\Phi(\cdot)$ is the cumulative standard normal distribution, and

$$p(X_{t_1} = x_{t_1} \mid N_{t_1} = c) = \frac{1}{\sigma_{X_{t_1}}} \phi\left(\frac{x_{t_1} - \mu_{X_{t_1}}}{\sigma_{X_{t_1}}}\right),$$

where $\phi(\cdot)$ is standard normal density.

Proof of Proposition 3.3. *By (A.1), conditional on $\Delta N_{t_n} = c$ and on $X_{t_{n-1}} = x_{t_{n-1}}$, X_{t_n} is normaly distributed with mean $\mu_{X_{t_n}|X_{t_{n-1}}} = x_{t_n} + \beta\Delta t + c\mu_\Pi$ and variance $\sigma_{X_{t_n}|X_{t_{n-1}}}^2 = \Delta t + c\sigma_\Pi^2$, where $\Delta t = t_n - t_{n-1}$, $\Delta N_{t_n} = N_{t_n} - N_{t_{n-1}}$. Similarly to the case when $t = t_1$,*

$$p(X_{t_n} = x_{t_n} \mid X_{t_{n-1}} = x_{t_{n-1}}, \Delta N_{t_n} = c) = \frac{1}{\sigma_{X_{t_n}|X_{t_{n-1}}}} \phi\left(\frac{x_{t_n} - \mu_{X_{t_n}|X_{t_{n-1}}}}{\sigma_{X_{t_n}|X_{t_{n-1}}}}\right),$$

ΔN_{t_n} is a Poisson process with intensity $\lambda\Delta t$, therefore

$$p(\Delta N_{t_n} = c) = e^{-\lambda\Delta t} \frac{(\lambda\Delta t)^c}{c!}.$$

B Proofs of Results in Sections 4.2 and 5.1

Outline of Proof of Proposition 4.1. *The proof follows the same lines as in Bush et al. (2011), where an SPDE for the diffusion case with absorbing boundary is derived, and also builds on the results in Kurtz & Xiong (1999) for diffusion on the infinite axis.*

For a test function ϕ and empirical measure $\nu_{N,t}$, we define

$$F_t^{N,\phi} = \langle \phi, \nu_{N,t} \rangle = \frac{1}{N} \sum_{i=1}^{N} \phi(X_t^i). \tag{B.1}$$

By applying Itô's lemma to (B.1) we obtain

$$F_t^{N,\phi} = F_0^{N,\phi} + \frac{1}{N} \sum_{i=1}^{N} \int_0^t \phi'(X_s^i) \left[\beta \, ds + \sqrt{1-\rho} \, dW_s^i + \sqrt{\rho} \, dM_s \right]$$
$$+ \frac{1}{N} \sum_{i=1}^{N} \int_0^t \frac{1}{2} \phi''(X_s^i) \, ds + \frac{1}{N} \sum_{i=1}^{N} \int_0^t \left[\phi(X_{s-}^i + \Pi) - \phi(X_{s-}^i) \right] dN_s. \tag{B.2}$$

In terms of $\nu_{N,t}$, (B.2) can be written as

$$F_t^{N,\phi} = F_0^{N,\phi} + \int_0^t \langle \beta \phi'(x) + \frac{1}{2} \phi''(x), \nu_{N,s} \rangle \, ds + \int_0^t \frac{1}{N} \sum_{i=1}^{N} \phi'(X_s^i) \sqrt{1-\rho} \, dW_s^i$$
$$+ \int_0^t \langle \sqrt{\rho} \phi'(x), \nu_{N,s} \rangle \, dM_s + \int_0^t \langle \phi(X_{s-}^i + \Pi) - \phi(X_{s-}^i), \nu_{N,s-} \rangle \, dN_s. \tag{B.3}$$

We are interested in the behaviour of $F_t^{N,\phi}$ as $N \to \infty$. Similar to the diffusion case, the idiosyncratic term given by

$$I_t^{N,\phi} = \frac{1}{N} \sum_{i=1}^{N} \int_0^t \phi'(X_s^i) \sqrt{1-\rho} \, dW_s^i \tag{B.4}$$

is a martingale with zero mean and quadratic variation

$$[I^{N,\phi}]_t = \int_0^t \frac{1}{N^2} \sum_{i=1}^{N} |\phi'(X_s^i)|^2 (1-\rho) \, ds \leq t \frac{1}{N^2} N \max_x |\phi'(x)|^2 \sqrt{1-\rho} \to 0 \quad \text{as} \quad N \to \infty. \tag{B.5}$$

Therefore, in the limit, the idiosyncratic term vanishes.

Let $\nu_{N,t}^0$ the empirical measure in the diffusion case (jump intensity $\lambda = 0$), and $F_t^{N,\phi,0}$ defined as in (B.1), with $\nu_{N,t}$ replaced by $\nu_{N,t}^0$. Then by inspection of (4.6) with X_t^i as per (4.2), the two measures in the diffusion and jump-diffusion case are explicitly related by

$$F_t^{N,\phi,0} = F_t^{N,\phi(\cdot - J_t)},$$

where J_t given in (4.12) is a process which is piecewise constant in t.

As $N \to \infty$, it is known that the empirical measure $\nu_{N,t}^0$ tends to some ν_t^0 with density v^0, hence, including jumps, $\nu_{N,t} \to \nu_t$ with density v, which is related to v^0 by $v(t,x) = v^0(t, x - J_t)$. Therefore, (4.11) follows. Details will be given in Bujok (2012).

Then, $F_t^{N,\phi} \to F_t^{\phi}$, where

$$F_t^{\phi} = \langle \phi, \nu_t \rangle = \int \phi(x) \nu_t(dx) = \int \phi(x) v(t,x) dx,$$

and we can write

$$\langle \phi, \nu_t \rangle = \langle \phi, \nu_0 \rangle + \int_0^t \langle \beta \phi'(x) + \frac{1}{2} \phi''(x), \nu_s \rangle ds$$
$$+ \int_0^t \langle \sqrt{\rho} \phi'(x), \nu_s \rangle dM_s + \int_0^t \langle \phi(x + \Pi), \nu_{s_-} \rangle dN_s \qquad \text{(B.6)}$$
$$+ \int_0^t \langle \phi(x), \nu_{s_-} \rangle dN_s,$$

or, in terms of v,

$$\int \phi(x) v(t,x)\, dx = \int \phi(x) v(0,x)\, dx + \int_0^t \int \left(\beta \phi'(x) + \frac{1}{2} \phi''(x) \right) v(s,x)\, dx\, ds$$
$$+ \int_0^t \int \sqrt{\rho}\, \phi'(x) v(s,x)\, dx\, dM_s + \int_0^t \int \phi(x) v(s_-,x)\, dx\, dN_s \quad \text{(B.7)}$$
$$+ \int_0^t \int \phi(x + \Pi) v(s_-,x)\, dx\, dN_s.$$

Integrating by parts and using that ϕ has compact support, we obtain

$$\int \phi(x) v(t,x)\, dx = \int \phi(x) \Big(v(0,x) - \int_0^t \beta \frac{\partial v}{\partial x}(s,x)\, ds$$
$$+ \int_0^t \frac{1}{2} \frac{\partial^2 v}{\partial x^2}(s,x)\, ds - \int_0^t \sqrt{\rho}\, \frac{\partial v}{\partial x}(s,x)\, dM_s \qquad \text{(B.8)}$$
$$+ \int_0^t v(s_-, x - \Pi)\, dN_s - \int_0^t v(s_-, x)\, dN_s \Big)\, dx.$$

Since ϕ is arbitrary, we have

$$v(t,x) = v(0,x) - \int_0^t \left(\beta \frac{\partial v}{\partial x}(s,x) + \frac{1}{2} \frac{\partial^2 v}{\partial x^2}(s,x) \right) ds - \int_0^t \sqrt{\rho} \frac{\partial v}{\partial x}(s,x)\, dM_s$$
$$+ \int_0^t v(s_-, x - \Pi)\, dN_s - \int_0^t -v(s_-, x)\, dN_s, \qquad \text{(B.9)}$$

which can be written in differential form as (4.13).

Proof of Proposition 5.1. *We want to show*

$$E[Z_t - Z_t^b] \quad \to \quad 0 \qquad \text{as} \qquad b \to \infty.$$

As $0 \le Z_t - Z_t^b \le L_t^b - L_t$ it suffices to show $E[L_t^b - L_t] \to 0$.

Define μ_t and μ_t^b by

$$0 \le \mu_t^b = \int_{-b}^b v^b(t,x)\, dx \le \int_{-\infty}^\infty v(t,x)\, dx = \mu_t \le 1.$$

If we can show that for any given $\epsilon > 0$, for sufficiently large b,

$$P[\mu_t - \mu_t^b \ge \epsilon/2] \le \epsilon/2, \qquad \text{(B.10)}$$

it follows

$$\begin{aligned}
E[L_t^b - L_t] &= E[\mu_t - \mu_t^b] \\
&= \underbrace{E[\mu_t - \mu_t^b | \mu_t - \mu_t^b \geq \epsilon/2]}_{\leq 1} \cdot \underbrace{P[\mu_t - \mu_t^b \geq \epsilon/2]}_{\leq \epsilon/2} + \underbrace{E[\mu_t - \mu_t^b | \mu_t - \mu_t^b \leq \epsilon/2]}_{\leq \epsilon/2} \cdot \underbrace{P[\mu_t - \mu_t^b \leq \epsilon/2]}_{\leq 1} \\
&\leq \epsilon.
\end{aligned}$$

We show by induction in k for $t = T_k$ that (B.10) can be satisfied with $\epsilon/2$ replaced by $\epsilon_k = 3^{-(n-k)} \epsilon/2$. Then, for $k = n$, we recover (B.10).

This is certainly true for $k = t = 0$, where $u = u_0$ deterministic (given by the observable X_0^i) and integrable on $\mathbb{R}$.

Assume now this is true for $t = T_k$ for all $b \geq b_k$ for some b_k. Write the event $B_k = \{\mu_{T_k} - \mu_{T_k}^{b_k} \geq \epsilon_k\}$, CB_k its complement, then

$$\begin{aligned}
P[B_{k+1}] &= P(B_{k+1}|B_k)P(B_k) + P(B_{k+1}|CB_k)P(CB_k) \\
&\leq \epsilon_k + P(B_{k+1}|CB_k).
\end{aligned}$$

First, pick a_k large enough such that

$$\mu_{T_k} - \mu_{T_{k+1}-}^{a_k} \leq 2\epsilon_k,$$

if CB_k. This is a standard result for the PDE (5.2). Then, pick b_{k+1} such that

$$P(A_k) \leq \epsilon_k,$$

for

$$A_k = \{a_k + \Delta M_k \geq b_{k+1}\} \cup \{-a_k + \Delta M_k \leq -b_{k+1}\},$$

such that

$$\begin{aligned}
P(B_{k+1}|CB_k) &= P(B_{k+1}|CB_k, A_k)P(A_k) + P(B_{k+1}|CB_k, CA_k)P(CA_k) \\
&\leq \epsilon_k + \epsilon_k.
\end{aligned}$$

This gives $P[B_{k+1}] \leq 3\epsilon_k = \epsilon_{k+1}$.