# Stochastic Finite Differences and Multilevel Monte Carlo for a Class of SPDEs in Finance[*]

Michael B. Giles[†] and Christoph Reisinger[†]

**Abstract.** In this article, we propose a Milstein finite difference scheme for a stochastic partial differential equation (SPDE) describing a large particle system. We show, by means of Fourier analysis, that the discretization on an unbounded domain is convergent of first order in the timestep and second order in the spatial grid size, and that the discretization is stable with respect to boundary data. Numerical experiments clearly indicate that the same convergence order also holds for boundary value problems. Multilevel path simulation, previously used for SDEs, is shown to give substantial complexity gains compared to a standard discretization of the SPDE or direct simulation of the particle system. We derive complexity bounds and illustrate the results by an application to basket credit derivatives.

**Key words.** parabolic stochastic partial differential equations, multilevel Monte Carlo simulation, error and complexity analysis, credit portfolio models

**AMS subject classifications.** 35R60, 60H15, 65C05, 65N06, 65N12, 91G40, 91G60

**DOI.** 10.1137/110841916

## 1. Introduction.

Various stochastic partial differential equations (SPDEs) have emerged over the last two decades in different areas of mathematical finance. A classical example is the Heath–Jarrow–Morton interest rate model in Heath, Jarrow, and Morton (1992) of the form

$$(1.1) \qquad \mathrm{d}r(x,t) = \frac{\partial}{\partial x}\left( r(t,x) + \frac{1}{2}\left| \int_t^{t+x} \sigma(t,u)\,\mathrm{d}u \right|^2 \right)\mathrm{d}t + \sigma(t,t+x)\,\mathrm{d}W_t,$$

where $r(x,t)$ is the forward rate of tenor $x$ at time $t$ and $\sigma(t,t+x)$ its instantaneous volatility. In Benth and Koekebakker (2008), a similar equation was proposed, more recently to model electricity forwards. Most of the SPDEs studied share with (1.1) the property that the derivatives of the solution appear only in the drift term; in the case of (1.1) the volatility of the Brownian driver does not depend on the solution $r$ at all. Numerical methods for hyperbolic SPDEs of the type (1.1) have been studied, for example, in Roth (2002) and Roth (2006).

This article, in contrast, considers the parabolic SPDE

$$(1.2) \qquad \mathrm{d}v = -\mu\,\frac{\partial v}{\partial x}\,\mathrm{d}t + \frac{1}{2}\,\frac{\partial^2 v}{\partial x^2}\,\mathrm{d}t - \sqrt{\rho}\,\frac{\partial v}{\partial x}\,\mathrm{d}M_t,$$

where $M$ is a standard Brownian motion, and $\mu$ and $0 \leq \rho < 1$ are real-valued parameters. It is clear that the behavior of this equation is fundamentally different from those with additive or multiplicative noise.

[†]Oxford-Man Institute of Quantitative Finance, and Mathematical Institute, University of Oxford, OX1 3LB Oxford, UK (mike.giles@maths.ox.ac.uk, christoph.reisinger@maths.ox.ac.uk).

The significance of (1.2) for the following applications is that it describes the limiting density of a large system of exchangeable particles. Specifically, if we consider the system of SDEs

$$\text{(1.3)} \qquad \mathrm{d}X_t^i = \mu\, \mathrm{d}t + \sqrt{1-\rho}\, \mathrm{d}W_t^i + \sqrt{\rho}\, \mathrm{d}M_t$$

for $1 \le i \le N$, with $\langle \mathrm{d}W_t^i, \mathrm{d}W_t^j \rangle = \delta_{ij}$ and $\langle \mathrm{d}W_t^i, \mathrm{d}M_t \rangle = 0$, where $X_0^i$ are assumed independent and identically distributed with finite second moment, the empirical measure

$$\nu_t^N = \frac{1}{N} \sum_{i=1}^{N} \delta(\cdot - X_t^i)$$

has a limit $\nu_t$ for $N \to \infty$, whose density $v$ satisfies (1.2) in a weak sense. For a derivation of this result in the more general context of quasi-linear PDEs see Kurtz and Xiong (1999). While the motivation in Kurtz and Xiong (1999) is to use a large particle system (1.3) to approximate the solution to the SPDE (1.2), our viewpoint is to use (1.2) as an approximate model for a large particle system, and we will argue later the (computational) advantages of this approach in situations when the number of particles is large.

As a first possible application, one may consider $X^i$ as the log price processes of a basket of equities, which have idiosyncratic components $W^i$ and share a common driver $M$ (the "market"). If the size of the basket is large enough, the solution to the SPDE can be used to find the values of basket derivatives. In this paper, we study an application of a similar model to basket credit derivatives.

We mention in passing that equations of the form (1.2) arise also in stochastic filtering. To be precise, (1.2) is the Zakai equation for the distribution of a signal $X$ given observation of $M$; see, e.g., Bain and Crisan (2009).

It is interesting to note that the solution to the SPDE (1.2) without boundary conditions can be written as the solution of the PDE

$$\text{(1.4)} \qquad \frac{\partial u}{\partial t} = \frac{1}{2}(1-\rho)\frac{\partial^2 u}{\partial x^2} - \mu\frac{\partial u}{\partial x},$$

shifted by the current value of the Brownian driver,

$$\text{(1.5)} \qquad v(t,x) = u(t, x - \sqrt{\rho}\, M_t).$$

In particular, if $v(0,x) = \delta(x_0 - x)$, then

$$\text{(1.6)} \qquad v(T,x) = \frac{1}{\sqrt{2\pi\,(1-\rho)\,T}}\, \exp\left(-\frac{(x - x_0 - \mu\,T - \sqrt{\rho}\, M_T)^2}{2\,(1-\rho)\,T}\right).$$

The intuitive interpretation of this result is that the independent Brownian motions $W^i$ have averaged into a deterministic diffusion in the infinite particle limit, whereas the common factor $M$, which moves all processes in parallel, shifts the whole profile (and also adds to the diffusion via the Itô term).

In Bush et al. (2011), the analysis of the large particle system is extended to cases with absorption at the boundary ($x = 0$):

(1.7)
$$X_t^i = 0, \quad t \geq T_0^i,$$
$$T_0^i = \inf\{t : X_t^i = 0\}.$$

It is shown that there is still a limit measure $\nu_t$, which may now be decomposed as

$$\nu_t = \nu_t^+ + L_t \delta_0,$$

where

(1.8)
$$L_t = \lim_{N \to \infty} L_t^N = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} 1_{T_0^i \leq t}$$

is the proportion of absorbed particles (the "loss function"), and the density $v$ of $\nu_t^+$ satisfies (1.2) in $(0, \infty)$ with absorbing boundary condition

(1.9)
$$v(t, 0) = 0.$$

Bush et al. (2011) consider applications to basket credit derivatives. For the market pricing examples, they consider a simplified model, where defaults are monitored only at a discrete set of dates. Between these times, the default barrier is inactive and (1.2) is solved on the real line by using (1.4) and (1.5).

For the initial-boundary value problem (1.2), (1.9), however, such a semianalytic solution strategy is no longer possible and an efficient numerical method is needed. Moreover, there is a loss of regularity at the boundary in this case, such that $xu_{xx} \in L_2$ but not $u_{xx}$, as is documented in Krylov (1994).

Recent papers on the numerical solution of SPDEs deal with cases relevant to ours, yet they are structurally crucially different. A comprehensive analysis of finite difference and finite element discretizations of the stochastic heat equation with multiplicative white noise and nonlinear driving term is given in Gyöngy and Nualart (1997), Gyöngy (1999), and Walsh (2005), respectively. Lang (2010) shows a Lax equivalence theorem for the SDE

$$\mathrm{d}X_t = AX_t \, \mathrm{d}t + G(X_t) \, \mathrm{d}M_t$$

in a Hilbert space, driven by a process $M$ from a class including Brownian motion, where $A$ is a suitable (e.g., elliptic differential) operator and $G$ a Lipschitz function.

In this paper, we propose a Milstein finite difference discretization for (1.2) and analyze its stability and convergence in the mean-square sense by Fourier analysis. A main consideration of this paper is the computational complexity of the proposed methods, and we will demonstrate that a multilevel approach achieves a cost for the SPDE simulation no larger than that of direct Monte Carlo sampling from a known univariate distribution, $O(\varepsilon^{-2})$ for root mean-square (r.m.s.) accuracy $\epsilon$, and is in that sense optimal.

Multilevel Monte Carlo path simulation, first introduced in Giles (2008a), is an efficient technique for computing expected values of path-dependent payoffs arising from the solution

of SDEs. It is based on a multilevel decomposition of Brownian paths, similar to a Brownian bridge construction. The complexity gain can be explained by the observation that the variance of high-level corrections—involving a large number of timesteps—is typically small, and consequently only a relatively small number of Monte Carlo samples is required to estimate these contributions to an acceptable accuracy. Overall, for SDEs, if an r.m.s. accuracy of $\varepsilon$ is required, the standard Monte Carlo method requires $O(\varepsilon^{-3})$ operations, whereas the multilevel method based on the Milstein discretization in Giles (2008b) requires $O(\varepsilon^{-2})$ operations.

The first extension of the multilevel approach to SPDEs was for parabolic PDEs with a multiplicative noise term in Graubner (2008). There have also been recent extensions to elliptic PDEs with random coefficients in Barth, Schwab, and Zollinger (2010) and Cliffe et al. (2011).

Our approach, for a rather different parabolic SPDE, is similar to the previous work on SDEs and SPDEs in that the solution is decomposed into a hierarchy with increasing resolution in both time and space. Provided the variance of the multilevel corrections decreases at a sufficiently high rate as one moves to higher levels of refinement, the number of fine grid Monte Carlo simulations which is required is greatly reduced. Indeed, the total cost is only $O(\varepsilon^{-2})$ to achieve an r.m.s. accuracy of $\varepsilon$ compared to an $O(\varepsilon^{-7/2})$ cost for the standard approach which combines a finite difference discretization of the spatial derivative terms and a Milstein discretization of the stochastic integrals.

The rest of the paper is structured as follows. Section 2 outlines the finite difference scheme used and analyzes its accuracy and stability in the standard Monte Carlo approach. Section 3 presents the modification to multilevel path simulation of functionals of the solution. Numerical experiments for a CDO tranche pricing application are given in section 4, providing empirical support for the postulated properties of the scheme and demonstrating the computational gains achieved. Section 5 discusses the benefits over standard Monte Carlo simulation of particle systems and outlines extensions.

## 2. Discretization and convergence analysis.

### 2.1. Milstein finite differences. Integrating (1.2) over the time interval $[t, t+k]$ gives

$$v(t+k, x) = v(t, x) + \int_t^{t+k} \left( -\mu \frac{\partial v}{\partial x} + \frac{1}{2} \frac{\partial^2 v}{\partial x^2} \right) \mathrm{d}s - \int_t^{t+k} \sqrt{\rho} \, \frac{\partial v}{\partial x} \, \mathrm{d}M_s.$$

Making the approximation $v(s, x) \approx v(t, x)$ for $t < s < t+k$ in the first integral and

$$v(s, x) \approx v(t, x) - \sqrt{\rho} \, \frac{\partial v}{\partial x}\bigg|_{(t,x)} (M_s - M_t)$$

in the second, and noting the standard Itô calculus result that

$$\int_t^{t+k} (M_s - M_t) \, \mathrm{d}M_s = \frac{1}{2} \left( (\Delta M^n)^2 - k \right),$$

where $\Delta M^n \equiv M_{t+k} - M_t = \sqrt{k} \, Z_n$ with $Z_n \sim N(0, 1)$ (see Glasserman (2004) or Kloeden and Platen (1992)), leads to the Milstein semidiscrete approximation

$$v^{n+1}(x) = v^n(x) - (\mu k + \sqrt{\rho k} \, Z_n) \frac{\partial v^n}{\partial x} + \frac{1}{2} \left( (1-\rho) k + \rho k Z_n^2 \right) \frac{\partial^2 v^n}{\partial x^2}.$$

Using a spatial grid with uniform spacing $h$, standard central difference approximations to the spatial derivatives as in Richtmyer and Morton (1967) then give the finite difference equation

$$
(2.1) \quad
\begin{aligned}
v_j^{n+1} = v_j^n \; &- \; \frac{\mu\,k + \sqrt{\rho\,k}\,Z_n}{2h}\left(v_{j+1}^n - v_{j-1}^n\right) \\
&+ \; \frac{(1-\rho)\,k + \rho\,k\,Z_n^2}{2h^2}\left(v_{j+1}^n - 2v_j^n + v_{j-1}^n\right),
\end{aligned}
$$

in which $v_j^n$ is an approximation to $v(nk, jh)$.

The spatial domain is truncated by introducing an upper boundary at $x_{max} = J\,h$ and using the boundary condition $v_J^n = 0$. Since the initial distribution will be assumed localized, both the localization error for a given path $M$ and the expected error of functionals of the solution can be made as small as needed by a large enough choice $x_{max} > 0$.

The system of SDEs can then be written in matrix-vector form as

$$
(2.2) \quad V_{n+1} = V_n \; - \; \frac{\mu\,k + \sqrt{\rho\,k}\,Z_n}{2h}D_1 V_n + \; \frac{(1-\rho)\,k + \rho\,k\,Z_n^2}{2h^2}D_2 V_n,
$$

where $V_n$ is the vector with elements $v_j^n, j = 1, \ldots, J-1$, and $D_1$ and $D_2$ are the matrices corresponding to first and second central differences as explicitly given in Appendix A.

*Remark* 2.1. An alternative discretization arises if the spatial derivatives are discretized first and the Milstein scheme is applied to the resulting system of SDEs. The practical implication is that the Itô term then contains a second finite difference with twice the step size, $D_1^2$ instead of $D_2$, resulting in pentadiagonal discretization matrices instead of tridiagonal ones, specifically

$$
(2.3) \quad V_{n+1} = V_n - \frac{\mu\,k + \sqrt{\rho\,k}\,Z_n}{2h}D_1 V_n + \frac{k}{2h^2}D_2 V_n + \frac{(\rho\,k\,(Z_n^2 - 1)}{2h^2}D_1^2 V_n.
$$

The accuracy and stability of the two schemes are similar, and hence we will not go into details.

To approximate the initial condition on the grid, we apply the initial measure $\nu_0^N$ to a basis of "hat functions" $\langle \Psi_j \rangle_{0 < j < J}$, where

$$
\Psi_j(x) = \frac{1}{h}\max\left(h - |x - x_j|, 0\right),
$$

giving

$$
v_j^0 = \langle \Psi_j, v_0 \rangle = \int_{-\infty}^{\infty} \Psi_j(x)\,v_0(x)\,\mathrm{d}x.
$$

For the corresponding PDE ($\rho = 0$), this is known to result in $O(h^2)$ convergence provided that $k$ satisfies a certain stability limit, even when the initial data are not smooth; see Pooley, Vetzal, and Forsyth (2003) and Carter and Giles (2007). We will see that an extension of this analysis holds for SPDEs.

**2.2. Fourier stability analysis.** Finite difference Fourier stability analysis ignores the boundary conditions and considers Fourier modes of the form

(2.4) $$v_j^n = g_n \exp(ij\theta), \quad |\theta| \le \pi,$$

which satisfy (2.1) provided that

$$g_{n+1} = \left(a(\theta) + b(\theta)\, Z_n + c(\theta)\, Z_n^2\right) g_n,$$

where

$$a(\theta) = 1 - \frac{i\,\mu\,k}{h}\,\sin\theta - \frac{2\,(1-\rho)\,k}{h^2}\,\sin^2\tfrac{\theta}{2},$$

$$b(\theta) = -\frac{i\sqrt{\rho\,k}}{h}\,\sin\theta,$$

$$c(\theta) = -\frac{2\,\rho\,k}{h^2}\,\sin^2\tfrac{\theta}{2}.$$

Following the mean-square stability analysis approach in Higham (2000), Saito and Mitsui (1996), and Buckwar and Sickenberger (2011), we obtain

$$\mathbb{E}[\,|g_{n+1}|^2] = \mathbb{E}\left[(a + b\,Z_n + c\,Z_n^2)(a^* + b^*Z_n + c^*Z_n^2)\,|g_n|^2\right]$$

$$= \left(\,|a+c|^2 + |b|^2 + 2|c|^2\,\right)\,\mathbb{E}\left[\,|g_n|^2\right],$$

where $a^*$ denotes the complex conjugate of $a$. Mean-square stability therefore requires that

$$|a|^2 + |b|^2 + 3|c|^2 + a\,c^* + a^*c$$

$$= 1 - 4\sin^2\tfrac{\theta}{2}\left\{\frac{k}{h^2} - (1+2\rho^2)\left(\frac{k}{h^2}\right)^2\sin^2\tfrac{\theta}{2} - \left(\left(\frac{\mu\,k}{h}\right)^2 + \frac{\rho\,k}{h^2}\right)\cos^2\tfrac{\theta}{2}\right\}$$

$$\le\ 1,$$

and enforcing this for all $\theta$ leads to the two conditions summarized in the following theorem.

**Theorem 2.1.** *The Milstein finite difference scheme* (2.2) *is stable in the mean-square sense provided that*

(2.5) $$\mu^2 k \le 1 - \rho,$$

(2.6) $$\frac{k}{h^2} \le (1 + 2\rho^2)^{-1}.$$

The analysis in Appendix A combines mean-square and matrix stability analysis to prove that in the limit $k, h \to 0$ the condition $k/h^2 \le (1 + 2\rho^2)^{-1}$ is also a sufficient condition for mean-square stability of the initial-boundary value problem with the boundary conditions at $j=0$ and $j=J$.

**2.3. Fourier analysis of accuracy.** Fourier analysis can also be used to examine the accuracy of the finite difference approximation in the absence of boundary conditions. Considering a Fourier mode of the form $g(t) \exp(i\kappa x)$, the PDE (1.4) reveals that

$$g(t) = g(0) \, \exp\left( -\tfrac{1}{2}(1-\rho)\,\kappa^2\,t - i\,\kappa\,(\mu\,t + \sqrt{\rho}\,M_t) \right),$$

and therefore

$$g(t_{n+1}) = g(t_n) \, \exp\left( -\tfrac{1}{2}(1-\rho)\,\kappa^2\,k - i\,\kappa\,(\mu\,k + \sqrt{\rho\,k}\,Z_n) \right),$$

where $M_{t_{n+1}} - M_{t_n} = \sqrt{k}\,Z_n$. Fourier analysis of the discretization gives

$$g_{n+1} = \left( a(\kappa h) + b(\kappa h)\,Z_n + c(\kappa h)\,Z_n^2 \right) g_n,$$

where $a(\theta)$, $b(\theta)$, $c(\theta)$ are as defined before. Writing

$$a(\kappa h) + b(\kappa h)\,Z_n + c(\kappa h)\,Z_n^2 = \exp\left( -\tfrac{1}{2}(1-\rho)\,\kappa^2\,t - i\,\kappa\,(\mu\,k + \sqrt{\rho\,k}\,Z_n) + e_n \right)$$

we obtain, after performing lengthy expansions using the Symbolic Toolbox of MATLAB,

$$e_n = \sqrt{\rho\,k}\,\kappa^2\mu\,k\,Z_n - \tfrac{1}{6}i\,\sqrt{\rho\,k}\,\kappa^3 Z_n \left( 3\,(1-\rho)\,k + \rho\,k\,Z_n^2 - h^2 \right) + r_n,$$

where, for all $p \geq 1$,

$$\mathbb{E}[|r_n|^p] \leq c(\kappa, p)\left( k^2 + kh^2 \right)^p.$$

When summing over $T/k$ timesteps, $\sum Z_n$ and $\sum Z_n^3$ are both $O(k^{-1/2})$ since they have zero expectation and $O(k^{-1})$ variance. Hence, it follows that

$$\left( \mathbb{E}\left[ \left( \sum_n e_n \right)^2 \right] \right)^{1/2} = O(k, h^2),$$

so the r.m.s. error in the Fourier mode over the full simulation interval is $O(k, h^2)$.

Following the method of analysis in Carter and Giles (2007), it can be deduced from this that the r.m.s. $L_2$ and $L_\infty$ errors for $V_n$ are both $O(k, h^2)$. This is consistent with the usual $O(k, h^2)$ accuracy of the forward-time central space discretization of a parabolic PDE (see, e.g., Richtmyer and Morton (1967)) and the $O(k)$ strong accuracy of the Milstein discretization of an SDE (see, e.g., Kloeden and Platen (1992)).

**2.4. Convergence tests.** We now test the accuracy and stability of the scheme numerically.

The chosen parameters for (1.2) are $\sigma = 0.22$, $\rho = 0.2$, $r = 0.042$, $\mu = (r - \sigma^2/2)/\sigma$, $v(0, x) = \delta(x - x_0)$ with $x_0 = 5$. These values are typical for the applications later on.

The upper boundary for the computation is $x_{max} = 16$, and it is chosen to ensure that the use of zero Dirichlet data has negligible influence on the solution.
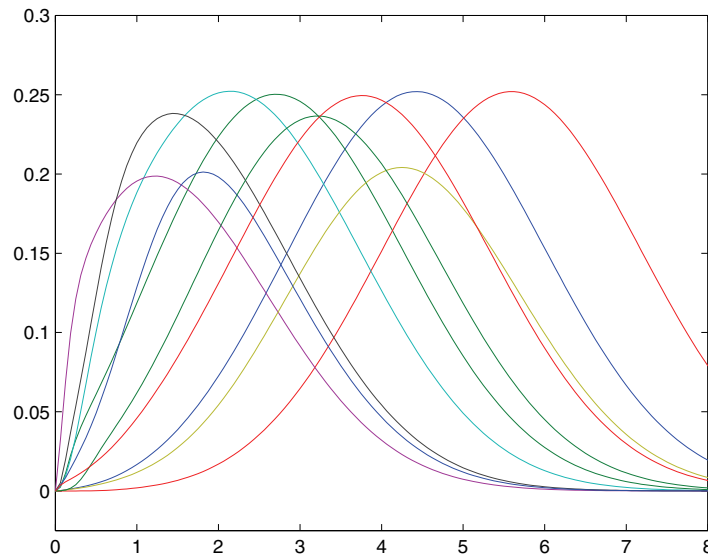
**Figure 1.** *Some solutions $v(T, x)$ for different driving Brownian motions $W_t$.*

We approximate the initial value problem without boundary conditions on $[-16/3, 16]$ (note that $x_0 = 5$ is roughly in the center) and also the initial-boundary value problem with zero Dirichlet conditions on $[0, 16]$.

Figure 1 plots several solutions to the latter problem at $T = 5$, each corresponding to a different Brownian path $M_t$. It can be seen that for those realizations for which $M_t$ has been largely positive, the boundary at $x = 0$ has had negligible influence, and so the solution is approximately equal to the displaced normal distribution given by (1.6).

For the unbounded case, we approximate the mean-square $L_2$ error by

$$E(h, k)^2 = \mathbb{E}\left[\sum_{j=0}^{J}(v_j^N(\omega) - v(Nk, jh; \omega))^2\, h\right]$$

(2.7)
$$\approx \frac{1}{M}\sum_{m=1}^{M}\sum_{j=0}^{J}(v_j^N(\omega_m) - v(Nk, jh; \omega_m))^2\, h,$$

where $J$ is the number of grid intervals, $N$ is the number of timesteps, and the expectation is taken over Brownian paths $\omega$, of which $\omega_m$ are $M$ samples.

To study the convergence, we introduce decreasing grid sizes $h_l = h_0\, 2^{-l}$ and timesteps $k_l = k_0\, 4^{-l}$, motivated by the second order convergence in $h$ and first order convergence in $k$ as well as the stability limit for the explicit scheme, and denote $E_l = E(h_l, k_l)$ the mean-square $L_2$ error at level $l$.

For the initial-boundary value problem, no analytical solution is known, but we can compute error indicators via the difference between a fine grid solution $f$ with mesh parameters
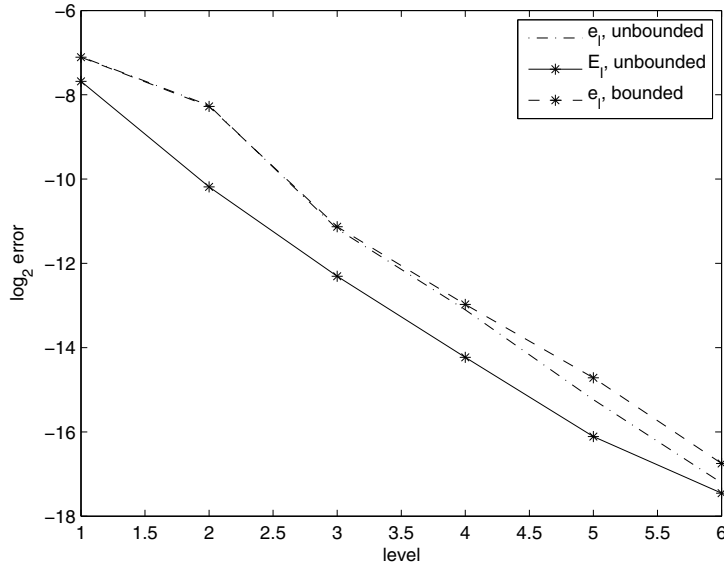
**Figure 2.** *Mean-square error estimators for the Milstein scheme. For the unbounded case, $L_2$ difference between true and numerical solution, $E_L$ as in (2.7), and coarse and fine grid solutions, $e_l$ as in (2.8); for the bounded case, difference between coarse and fine grid solutions $e_l$.*

$k$ and $h$, and a coarse solution $c$ with mesh parameters $4k$ and $2h$,

$$e(h,k)^2 = \mathbb{E}\left[\sum_{j=0}^{J/2}(f_{2j}^N(\omega) - c_j^{N/4}(\omega))^2\, h\right]$$

(2.8)
$$\approx \frac{1}{M}\sum_{m=1}^{M}\sum_{j=0}^{J/2}(f_{2j}^N(\omega_m) - c_j^{N/4}(\omega_m))^2\, h,$$

and define $e_l = e(h_l, k_l)$.

On the coarsest level, $h_0 = 4/3$, $k_0 = 1/4$ such that $x_0$ does not coincide with a grid point.

Figure 2 shows both the computed values of $E_l^2$ and $e_l^2$ for the unbounded case and $e_l^2$ for the bounded case.

The results confirm the theoretical $O(k, h^2)$ convergence in the unbounded case and support the conjecture that the convergence order is unchanged in the bounded case.

We now formalize this conjecture about the error due to the SPDE discretization. Denote $U_T = u(T, \cdot)$ the solution to the initial-boundary value problem at time $T$ for a given Brownian path and $\widehat{U}_T$ its numerical approximation with grid size $h$ and timestep $k \propto h^2$.

Conjecture 2.1. *The error in the solution at time $T$ satisfies the strong error estimate in the $L_2$ norm $\|\cdot\|$:*

$$\sqrt{\mathbb{E}[\|\widehat{U}_T - U_T\|^2]} = O(h^2).$$

**Corollary 2.2.** *A Lipschitz payoff function $P(U_T)$ has a similar strong error bound:*

$$\sqrt{\mathbb{E}[(P(\widehat{U}_T) - P(U_T))^2]} = O(h^2).$$

We know Conjecture 2.1 is true for the initial value problem on $\mathbb{R}$, and it conjectures that the introduction of the boundary condition at $x = 0$ does not affect the (weak and) strong error, which is supported by the numerical results.

The conjecture assumes $x_{max} = \infty$; in practice, a finite value for $x_{max}$ will introduce an additional truncation error which will decay exponentially in $x_{max}$.

Numerical tests with different parameters, not reproduced here, indicate that the error in the bounded case increases for values of $x_0$ close to 0, in which case more paths contribute solutions with large higher derivatives close to the zero boundary, even if the asymptotic convergence order is still $O(k, h^2)$. We analyze this loss of regularity further in Appendix B.

**3. Multilevel Monte Carlo simulation.** We now consider estimating the expectation of a scalar functional of the SPDE solution, an expected "payoff" $P$ in a computational finance context.

Defining $\widehat{P}$ to be the payoff arising from a single SPDE approximation, the standard Monte Carlo approach is to average the payoff from $N$ independent SPDE simulations, each one using an independent vector of $N(0,1)$ normal variables $Z = (Z^0, Z^1, \ldots, Z^{n-1})$. Thus the estimator for the expected value is

$$\widehat{Y} = \frac{1}{N} \sum_{i=1}^{N} \widehat{P}^{(i)}.$$

The mean-square error for this estimator can be expressed as the sum of two contributions, one due to the variance of the estimator and the other due to the error in its expectation,

$$\mathbb{E}\left[\left(\widehat{Y} - \mathbb{E}[P]\right)^2\right] = N^{-1}\mathbb{V}[\widehat{P}] + \left(\mathbb{E}[\widehat{P}] - \mathbb{E}[P]\right)^2.$$

To achieve an r.m.s. error of $\varepsilon$ requires that both of these terms be $O(\varepsilon^2)$. This in turn requires that $N = O(\varepsilon^{-2})$, $k = O(\varepsilon)$, and $h = O(\varepsilon^{1/2})$, based on the conjecture that the weak error $\mathbb{E}[\widehat{P}] - \mathbb{E}[P]$ is $O(k, h^2)$. Since the computational cost is proportional to $Nk^{-1}h^{-1}$, this implies an overall cost which is $O(\varepsilon^{-7/2})$. The aim of the multilevel Monte Carlo simulation is to reduce this complexity to $O(\varepsilon^{-2})$.

Consider Monte Carlo simulations with different levels of refinement, $l = 0, 1, \ldots, L$, with $l = 0$ being the coarsest level (i.e., the largest values for $k$ and $h$) and level $L$ being the finest level corresponding to that used by the standard Monte Carlo method.

Let $\widehat{P}_l$ denote an approximation to payoff $P$ using a numerical discretization with parameters $k_l$ and $h_l$. Because of the linearity of the expectation operator, it is clearly true that

$$(3.1) \qquad\qquad \mathbb{E}[\widehat{P}_L] = \mathbb{E}[\widehat{P}_0] + \sum_{l=1}^{L} \mathbb{E}[\widehat{P}_l - \widehat{P}_{l-1}].$$

This expresses the expectation on the finest level as being equal to the expectation on the coarsest level plus a sum of corrections which give the difference in expectation between simulations using different numbers of timesteps. The multilevel idea is to independently estimate each of the expectations on the right-hand side in a way which minimizes the overall variance for a given computational cost.

Let $\widehat{Y}_0$ be an estimator for $\mathbb{E}[\widehat{P}_0]$ using $N_0$ samples, and let $\widehat{Y}_l$ for $l>0$ be an estimator for $\mathbb{E}[\widehat{P}_l-\widehat{P}_{l-1}]$ using $N_l$ samples. Each estimator is an average of $N_l$ independent samples, which for $l>0$ is

$$(3.2) \qquad \widehat{Y}_l = N_l^{-1} \sum_{i=1}^{N_l} \left( \widehat{P}_l^{(i)} - \widehat{P}_{l-1}^{(i)} \right).$$

The key point here is that the quantity $\widehat{P}_l^{(i)} - \widehat{P}_{l-1}^{(i)}$ comes from two discrete approximations using the same Brownian path. The variance of this simple estimator is $\mathbb{V}[\widehat{Y}_l] = N_l^{-1}V_l$, where $V_l$ is the variance of a single sample. Combining this with independent estimators for each of the other levels, the variance of the combined estimator $\sum_{l=0}^{L} \widehat{Y}_l$ is $\sum_{l=0}^{L} N_l^{-1}V_l$. The corresponding computational cost is $\sum_{l=0}^{L} N_l\, C_l$, where $C_l$ represents the cost of a single sample on level $l$. Treating the $N_l$ as continuous variables, the variance is minimized for a fixed computational cost by choosing $N_l$ to be proportional to $\sqrt{V_l/C_l}$, with the constant of proportionality chosen so that the overall variance is $O(\varepsilon^{-2})$.

The total cost on level $l$ is proportional to $\sqrt{V_l\, C_l}$. If the variance $V_l$ decays more rapidly as a function of the level $l$ than the cost $C_l$ increases, the dominant cost is on level 0. The number of samples on that level will be $O(\varepsilon^{-2})$, and the cost savings compared to standard Monte Carlo will be approximately $C_0/C_L$, reflecting the different costs of samples on level 0 compared to level $L$. On the other hand, if the variance $V_l$ decays more slowly than the cost $C_l$ increases, the dominant cost will be on the finest level $L$, and the cost savings compared to standard Monte Carlo will be approximately $V_L/V_0$, reflecting the difference between the variance of the finest grid correction compared to the variance of the standard Monte Carlo estimator, which is similar to $V_0$.

This outline analysis is made more precise in the following theorem.

**Theorem 3.1.** *Let $P$ denote a functional of the solution of an SPDE for a given Brownian path $M_t$, and let $\widehat{P}_l$ denote the corresponding level $l$ numerical approximation.*

*If there exist independent estimators $\widehat{Y}_l$ based on $N_l$ Monte Carlo samples, and positive constants $\alpha, \beta, \gamma, c_1, c_2, c_3$ such that $\alpha \geq \frac{1}{2}\gamma$ and*

(i) $|\mathbb{E}[\widehat{P}_l - P]| \leq c_1\, 2^{-\alpha l}$,

(ii) $\mathbb{E}[\widehat{Y}_l] = \begin{cases} \mathbb{E}[\widehat{P}_0], & l = 0, \\ \mathbb{E}[\widehat{P}_l - \widehat{P}_{l-1}], & l > 0, \end{cases}$

(iii) $\mathbb{V}[\widehat{Y}_l] \leq c_2\, N_l^{-1}2^{-\beta l}$,

(iv) $C_l \leq c_3\, N_l\, 2^{\gamma l}$, *where $C_l$ is the computational complexity of $\widehat{Y}_l$,*

*then there exists a positive constant $c_4$ such that for any $\varepsilon < e^{-1}$ there are values $L$ and $N_l$ for*

*which the multilevel estimator*

$$\widehat{Y} = \sum_{l=0}^{L} \widehat{Y}_l$$

*has a mean-square error with bound*

$$MSE \equiv \mathbb{E}\left[\left(\widehat{Y} - E[P]\right)^2\right] < \varepsilon^2$$

*with a computational complexity $C$ with bound*

$$C \le \begin{cases} c_4\,\varepsilon^{-2}, & \beta > \gamma, \\ c_4\,\varepsilon^{-2}(\log\varepsilon)^2, & \beta = \gamma, \\ c_4\,\varepsilon^{-2-(\gamma-\beta)/\alpha}, & 0 < \beta < \gamma. \end{cases}$$

*Proof.* The proof is a slight generalization of the proof in Giles (2008a). ∎

In our application, we choose $h_l \propto 2^{-l}$ and $k_l \propto 4^{-l}$ so that the ratio $k_l/h_l^2$ is held fixed to satisfy the finite difference stability condition. The computational cost increases by factor 8 in moving from level $l$ to $l+1$, so $\gamma=3$.

Given that the payoff is a Lipschitz function of the loss approximations at various dates, Conjecture 2.1 implies that the weak error is also $O(h^2)$, and so $\alpha=2 > \frac{1}{2}\gamma$. Also, due to the triangle inequality

$$\sqrt{\mathbb{V}[\widehat{P}_l - \widehat{P}_{l-1}]} \le \sqrt{\mathbb{V}[\widehat{P}_l - P]} + \sqrt{\mathbb{V}[\widehat{P}_{l-1} - P]}$$
$$\le \sqrt{\mathbb{E}[(\widehat{P}_l - P)^2]} + \sqrt{\mathbb{E}[(\widehat{P}_{l-1} - P)^2]},$$

Conjecture 2.1 and its corollary give $\beta=4$. Consequently, the computational cost to achieve an r.m.s. error of $\varepsilon$ is $O(\varepsilon^{-2})$.

**4. Numerical experiments.** In this section we study the numerical performance of the algorithms presented earlier on the example of CDO tranche pricing in a large basket limit.

**4.1. CDO pricing in the structural credit model.** Basket credit derivatives provide protection against the default of a certain segment ("tranche") of a basket of firms. A typical example is that of a *collateralized debt obligation* where the protection buyer receives a notional amount, minus some recovery proportion $0 \le R \le 1$, if firms in a specified tranche of the basket default, and in return pays a regular spread until the default event occurs.

The arbitrage-free spread depends crucially on the (risk-neutral, when hedged with defaultable bonds) probability of joint defaults. For a tranche with attachment point $0 \le a < 1$ and detachment point $1 \ge d > a$, the outstanding tranche notional

(4.1) $$P(L_t) = \max(d - L_t, 0) - \max(a - L_t, 0),$$

where the loss variable $L_t$ is the proportion of losses in the basket up to time $t$, determines the spread and default payments related to that tranche. The risk-neutral value of the tranche

spread can be derived as

$$(4.2) \qquad s = \frac{\sum_{i=1}^{n} e^{-rT_i} \mathbb{E}^{\mathbb{Q}}[P(L_{T_{i-1}}) - P(L_{T_i})]}{\delta \sum_{i=1}^{n} e^{-rT_i} \mathbb{E}^{\mathbb{Q}}[P(L_{T_i})]};$$

see, e.g., Bush et al. (2011). Here, $n$ is the maximum number of spread payments, $T$ the expiry, $T_i$ the payment dates for $1 \leq i \leq n$, and $\delta = 0.25$ the interval between payments. Spreads are quoted as an annual payment, as a ratio of the notional, but assumed to be paid quarterly.[1]

For an extensive survey of credit derivatives and pricing models we refer the reader to Schönbucher (2003) and note only that they typically fall into one of two classes: so-called reduced-form models, which model default times of firms directly as (dependent) random variables; and structural models, which capture the evolution of the firms' values and model default events as the first passage of a lower default barrier. We will consider the latter here.

In a structural model in the spirit of the classical works of Merton (1974) and Black and Cox (1976), as extended to multiple firms, e.g., in the work of Hull, Predescu, and White (2005), a company $i$'s firm value, $1 \leq i \leq N$, is assumed to follow a model of the type

$$\frac{\mathrm{d}A_t^i}{A_t^i} = r \, \mathrm{d}t + \sqrt{1-\rho_i} \, \sigma_i \, \mathrm{d}W_t^i + \sqrt{\rho_i} \, \sigma_i \, \mathrm{d}M_t, \qquad A_0^i = a^i,$$

where $0 \leq \rho_i \leq 1$ is a correlation parameter, $r$ is the risk-free interest rate, $\sigma_i$ is the volatility, and $M$, $W^i$ are standard Brownian motions.

Here, the individual firms are correlated through a common "market" factor $M$ but independent conditionally on $M$. We make in the following the assumption that the firms are *exchangeable* in the sense that their dynamics is governed by the same set of parameters, specifically $\rho = \rho_i$, $\sigma = \sigma_i$. Their initial values $a_i$ are not necessarily identical, which allows for different default probabilities for individual firms, consistent with their CDS spreads.

In this framework, the default time $T_0^i$ of the $i$th firm is modeled as the first hitting time of a default barrier $B^i$, for simplicity constant, and the *distance-to-default*

$$(4.3) \qquad X_t^i = \frac{1}{\sigma} \left( \log A_t^i - \log B^i \right)$$

evolves according to (1.3), where $\mu = \left( r - \frac{1}{2}\sigma^2 \right)/\sigma$, $x^i = \left( \log a^i - \log B^i \right)/\sigma$. $T_0^i$ as in (1.7) is precisely the default time.

In the majority of applications, Monte Carlo simulation of the firm value processes is used for the estimation of tranche spreads; see, e.g., Hull, Predescu, and White (2005) or Fouque, Wignall, and Zhou (2008) and Carmona, Fouque, and Vestal (2009). This is largely enforced by the size of $N$, for instance in the case of index tranches $N = 125$.

This, however, puts the model precisely in the realm of the large basket approximation (1.2). See Bujok and Reisinger (2012) for a numerical study of this large basket approxima-

---

[1]There is a variation for the equity tranche, and recently sometimes the mezzanine tranche, but we do not go into details here.

tion. The loss functional is thereby approximated by

$$\widehat{L}_{T_i} = (1 - R) \left( 1 - h \sum_{j=1}^{J-1} v_j^{T_i/k} \right)$$

in terms of the numerical SPDE solution $v_j^{T_i/k}$ at time $T_i$, which feeds into the estimator for the outstanding tranche notional (4.1) and subsequently the tranche spreads (4.2).

**4.2. Pricing results.** All the following results are for a representative set of parameter values, taken from a calibration performed in Bush et al. (2011) to 2007 market data, $\sigma = 0.22$, $\rho = 0.2$, $r = 0.042$, $\mu = (r - \sigma^2/2)/\sigma$. While the initial distribution used in Bush et al. (2011) is somewhat spread out to match individual CDS spreads of obligors, most of the mass is around $x_0 = 5$, and for simplicity we center all firms there for the numerical tests, $v(0, x) = \delta(x - x_0)$. The upper boundary is chosen as $x_{max} = 16$.

We consider a maturity $T = 5$ and tranches $[a, d] = [0, 0.03]$, $[0.03, 0.06]$, $[0.06, 0.09]$, $[0.09, 0.12]$, $[0.12, 0.22]$, $[0.22, 1]$.

Figure 3 shows the multilevel results for the expected protection payment from the first tranche,

$$\sum_{i=1}^{n} e^{-rT_i} \mathbb{E}^{\mathbb{Q}}[P(\widehat{L}_{T_{i-1}}) - P(\widehat{L}_{T_i})],$$

expressed as a fraction of the initial tranche notional.

We pick mesh sizes $h_l = h_0 \, 2^{-l}$ and $k_l = k_0 \, 4^{-l}$ for $l > 0$ and $h_0 = 8/5$, $k_0 = 1/4$. This is motivated by the second and first order consistency in $h$ and $k$, respectively, and within the stability limit (2.6).

The top-left plot shows the convergence of the variance $\mathbb{V}[\widehat{P}_l - \widehat{P}_{l-1}]$ as well as the variance of the standard single level estimate, while the top-right plot shows the convergence of the expectation $\mathbb{E}[\widehat{P}_l - \widehat{P}_{l-1}]$.

The bottom-left plot shows results for different multilevel calculations, each with a different user-specified accuracy requirement. For each value of the accuracy $\varepsilon$, the multilevel algorithm determines (by formula (12) from Giles (2008a)) the number of levels of refinement which are needed to ensure that the contribution to the mean-square error (MSE) due to the weak error on the finest grid is less than $\frac{1}{2}\varepsilon^2$, and it determines (by formula (10) from Giles (2008a)) the optimal number of samples on each level so that the combined variance of the multilevel estimate is also less than $\frac{1}{2}\varepsilon^2$, and hence the MSE is less than $\varepsilon^2$.

Treating a single finite difference calculation on the coarsest level as having a unit cost, the bottom-right plot compares the total cost of the standard and multilevel algorithms. Since the objective is to achieve a computational cost which is approximately proportional to $\varepsilon^{-2}$, it is the cost $C$ multiplied by $\varepsilon^2$ which is plotted versus $\varepsilon$. The results confirm that $\varepsilon^2 C$ does not vary much as $\varepsilon \to 0$ for the multilevel method, in line with the prediction that $\varepsilon^2 C \propto 1$, whereas it grows significantly for the standard method since $\varepsilon^2 C \propto 8^L$, where $L$ is the index of the finest level. Hence, there is a big jump in the cost of the standard method each time it is necessary to switch to a finer level to ensure the weak error is less than $\varepsilon/\sqrt{2}$, whereas the jump is minimal for the multilevel method.
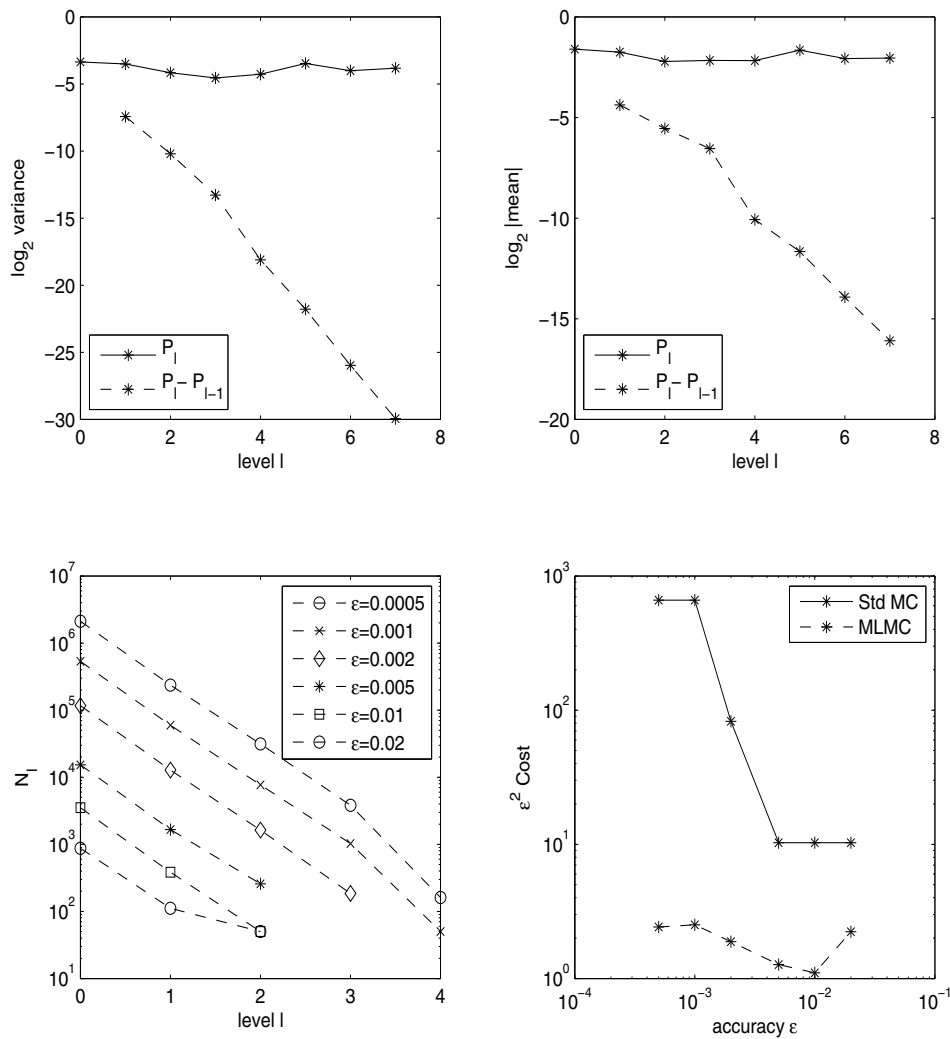
**Figure 3.** *Multilevel results for the expected loss from the first tranche.*

In practice, default is not monitored continuously but only at a discrete set of times, and for pricing often the simplifying assumption is made that default is determined only at the spread payment dates $T_l$.

This can be incorporated as follows. In the time intervals $(T_{l-1}, T_l)$, we solve (1.2) on a sufficiently large domain (e.g., $[-4, 16]$) and apply the following interface conditions at $T_l$:

$$(4.4) \qquad \lim_{t \downarrow T_l} v(t, x) = \begin{cases} 0, & x \leq 0, \\ \lim_{t \uparrow T_l} v(t, x), & x > 0. \end{cases}$$
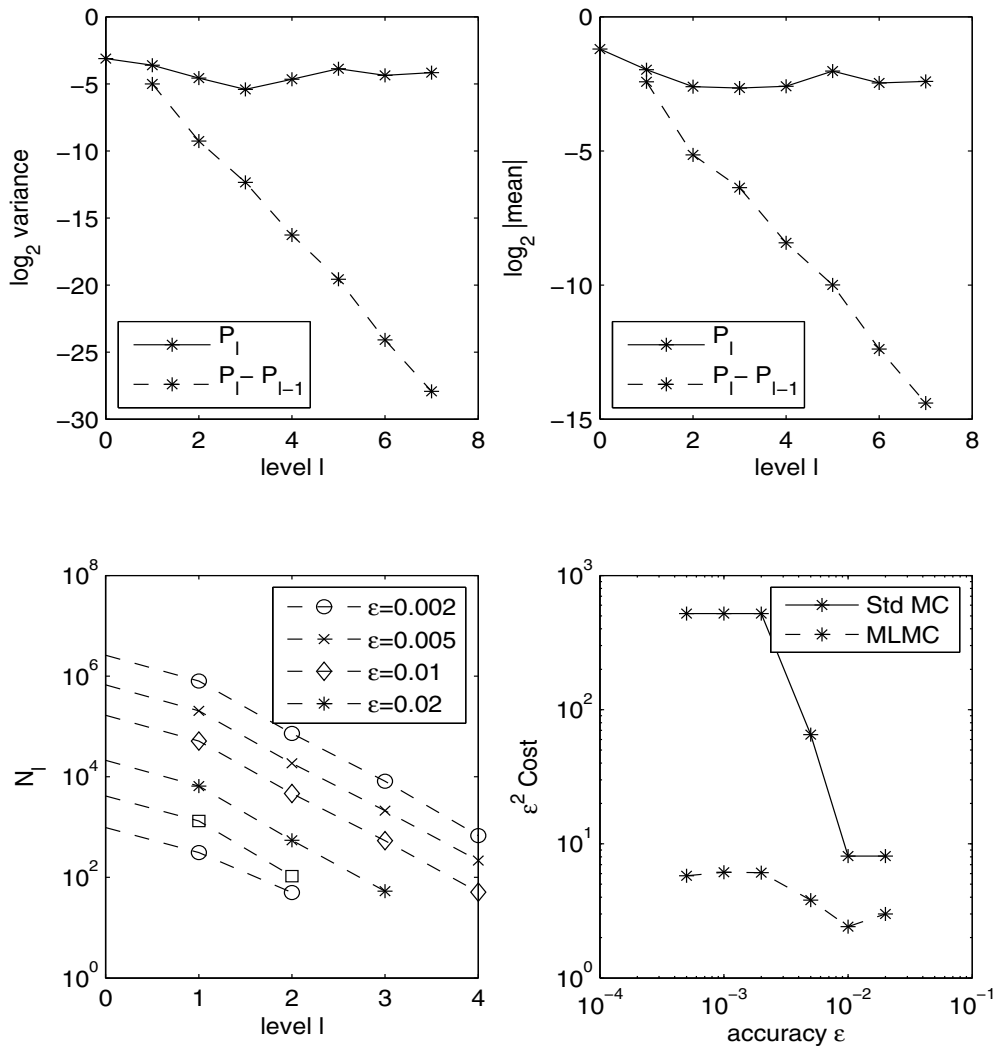
**Figure 4.** *Multilevel results for the expected loss from the first tranche, with discrete monitoring.*

To maintain quadratic grid convergence in spite of the discontinuity introduced by (4.4), we choose the mesh such that a grid point coincides with the 0 boundary (e.g., by setting $h_0 = 2$ in the above example), and set the numerical solution after default monitoring to 0 for grid coordinates below 0 and to $1/2$ its previous value at 0; see, e.g., Pooley, Vetzal, and Forsyth (2003).

It is seen in Figure 4 that convergence is very similar to the previous case.

| Method/model | SDE | SPDE |
|---|---|---|
| Standard MC | $O(\varepsilon^{-4})$ | $O(\varepsilon^{-7/2})$ |
| Multilevel MC | $O(\varepsilon^{-3})$ | $O(\varepsilon^{-2})$ |

## 5. Conclusions.

**5.1. Complexity and cost.** Here we discuss the computational complexity of the multilevel solution of the SPDE compared with the alternative use of the multilevel method for solving the SDEs which arise from directly simulating a large number of SDEs.

We have already explained that to achieve an r.m.s. accuracy of $\varepsilon$ requires $O(\varepsilon^{-7/2})$ work when solving the SPDE by the standard Monte Carlo method, but the cost is $O(\varepsilon^{-2})$ using the multilevel method, provided Conjecture 2.1 is correct.

Consider now the alternative of using a finite number of particles (firms), $M$, to estimate the tranche loss in the limit of an infinite number of particles (firms). In this case, empirical results suggest that there is an additional $O(M^{-1})$ error (see also Bujok and Reisinger (2012)), and the proof of this convergence order is the subject of current research. Taking this to be the case, the optimal choice of $M$ to minimize the computational complexity to achieve an r.m.s. error of $\varepsilon$ is $O(\varepsilon^{-1})$. Using the standard Monte Carlo method, the optimal timestep is $O(\varepsilon)$, and the optimal number of paths is $O(\varepsilon^{-2})$, so the overall cost is $O(\varepsilon^{-4})$. Using the multilevel method for the SDEs reduces the cost per company to $O(\varepsilon^{-2})$, so the total cost is $O(\varepsilon^{-3})$.

This complexity information is summarized in Table 1. There is also a practical implementation aspect to note. The computational cost per grid point in the finite difference approximation of the SPDE is minimal, requiring just three floating point multiply-add operations if (2.1) is recast as

$$v_j^{n+1} = a v_{j-1}^n + b v_j^n + c v_{j+1}^n$$

with the coefficients $a, b, c$ computed once per timestep for all $j$.

If we let $C$ be the cost of generating all of the Gaussian random numbers $Z_n$ for a single SPDE simulation, then the cost of the rest of the finite difference calculation with 20 points in $x$ (as used on the coarsest level of our multilevel calculations) is probably similar, giving a total cost of $2C$ for each SPDE. On the other hand, each SDE needs its own Gaussian random numbers for the idiosyncratic risk, and so the cost of simulating *each* SDE is approximately $C$, roughly half of the cost of the SPDEs on the coarsest level of approximation, giving a total cost of $MC$.

**5.2. Further work.** We have shown that stochastic finite differences combined with a multilevel simulation approach achieve optimal complexity for the computation of expected payoffs of an SPDE model. In the case of an absorbing boundary, the complexity estimate is a conjecture in so far as it relies on the convergence order of the finite difference scheme, which does not follow from the Fourier analysis of the unbounded case. The matrix stability analysis in Appendix A could form part of a rigorous analysis if a Lax equivalence theorem

could be proved. In the case of multiplicative white noise, this is shown in Lang (2010). One difficulty in the present case of an SPDE with stochastic drift is the loss of regularity towards the boundary, which may be accounted for by weighted Sobolev norms of the solution, but even then it is not clear that convergence of the functionals of interest follows.

Several possible extensions of the present basic model exist as shown in Bujok and Reisinger (2012), ranging from stochastic volatility and jump-diffusion to contagion models. The methods developed in this paper should be of use there also, building, for example, on multilevel versions of jump-adapted discretizations for jump-diffusion SDEs in Bruti-Liberati and Platen (2010) and Xia and Giles (2012).

**Appendix A. Mean-square matrix stability analysis.** If $V_n$ is the vector with elements $v_j^n, j = 1, \dots, J-1$, then the finite difference equation can be expressed as

$$V_{n+1} = (A + B\,Z_n + C\,Z_n^2)\,V_n,$$

$$A = I - \frac{\mu\,k}{2h}\,D_1 + \frac{(1-\rho)\,k}{2h^2}\,D_2,$$

$$B = -\frac{\sqrt{\rho\,k}}{2h}\,D_1,$$

$$C = \frac{\rho\,k}{2h^2}\,D_2,$$

where $I$ is the identity matrix and $D_1$ and $D_2$ are the matrices corresponding to central first and second differences, which for $J = 6$ are

$$D_1 = \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & -1 & 0 & 1 & \\ & & -1 & 0 & 1 \\ & & & -1 & 0 \end{pmatrix}, \quad D_2 = \begin{pmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{pmatrix}.$$

From the recurrence relation we get

$$\mathbb{E}[\,V_{n+1}^T V_{n+1}] = \mathbb{E}\left[V_n^T(A^T + B^T Z_n + C^T Z_n^2)(A + B\,Z_n + C\,Z_n^2)\,V_n\right]$$

$$= \mathbb{E}\left[V_n^T\left((A+C)^T(A+C) + B^T B + 2\,C^T C\right)V_n\right].$$

Noting that $D_1$ is antisymmetric and $D_2$ is symmetric, and that

$$D_1 D_2 - D_2 D_1 = E_1 - E_2, \quad D_1^2 = D_3 + E_1 + E_2,$$

where $D_3$ corresponds to a central second difference with twice the usual span,

$$D_3 = \begin{pmatrix} -3 & 0 & 1 & & \\ 0 & -2 & 0 & 1 & \\ 1 & 0 & -2 & 0 & 1 \\ & 1 & 0 & -2 & 0 \\ & & 1 & 0 & -3 \end{pmatrix}$$

(with the end values of $-3$ being chosen to correspond to $V_{-1} \equiv -V_1$ and $V_{J+1} \equiv -V_{J-1}$), and $E_1$ and $E_2$ are each entirely zero apart from one corner element,

$$
E_1 = \begin{pmatrix} 2 & & \\ & & \\ & & \\ & & \end{pmatrix}, \qquad E_2 = \begin{pmatrix} & & \\ & & \\ & & \\ & & 2 \end{pmatrix},
$$

then after some lengthy algebra we get

$$
\mathbb{E}\left[ V_n^T \left( (A+C)^T(A+C) + B^T B + 2\,C^T C \right) V_n \right]
$$
$$
= \mathbb{E}\left[ V_n^T M V_n \right] - (e_1 + e_2)\,\mathbb{E}[(v_1^n)^2] - (e_1 - e_2)\,\mathbb{E}[(v_{J-1}^n)^2],
$$

where

$$
M = I - \frac{k}{h^2}\,D_2 + \frac{k^2}{4\,h^4}\,D_2^2 - \left( \frac{\rho k}{4\,h^2} + \frac{\mu^2 k^2}{4\,h^2} \right) D_3,
$$

and

$$
e_1 = \frac{\rho k}{2\,h^2} + \frac{\mu^2 k^2}{2\,h^2}, \qquad e_2 = \frac{\mu k^2}{2\,h^3}.
$$

It can be verified that the $m$th eigenvector of $M$ has elements $\sin j\theta_m$ for $\theta_m = m\,\pi/J$, and the associated eigenvalue is

$$
|a(\theta_m) + c(\theta_m)|^2 + |b(\theta_m)|^2 + 2|c(\theta_m)|^2,
$$

where $a(\theta), b(\theta), c(\theta)$ are the same functions as defined in the mean-square Fourier analysis. In addition, in the limit $h, k/h \to 0$, $e_1 \pm e_2 > 0$, and therefore in this limit the Fourier stability condition

$$
\sup_{\theta} \left\{ |a(\theta) + c(\theta)|^2 + |b(\theta)|^2 + 2|c(\theta)|^2 \right\} \leq 1
$$

is also a sufficient condition for mean-square matrix stability.

**Appendix B. Regularity considerations.** Figure 5 shows the convergence behavior as the computational grid is refined. Level 0 has $h = 1/4$, $k = T/4$; $h$ is reduced by factor 2 and $k$ by factor 4 in moving to finer levels.

The top-left plot shows the convergence of $\mathbb{E}[\widehat{P}_l - \widehat{P}_{l-1}]$, with notation as in sections 3 and 4.2; the top-right plot shows the convergence of $\mathbb{V}[\widehat{P}_l - \widehat{P}_{l-1}]$.

The bottom two plots show the behavior of $\partial^2 v/\partial x^2(T, 0)$, which is estimated on each grid level using the standard second difference

$$
\frac{\partial^2 v}{\partial x^2}(0) \approx \frac{v_2 - 2v_1 + v_0}{h^2}.
$$

The left plot indicates that the mean of this quantity is well behaved, but the right plot indicates a singular behavior of its variance, with the value increasing rapidly with increased grid resolution.

This is in accordance with the result shown in Krylov (1994), that for the unique solution $v$ to the SPDE, $x\,v_{xx}$ is square integrable, but $v_{xx}$ has a singularity at 0.
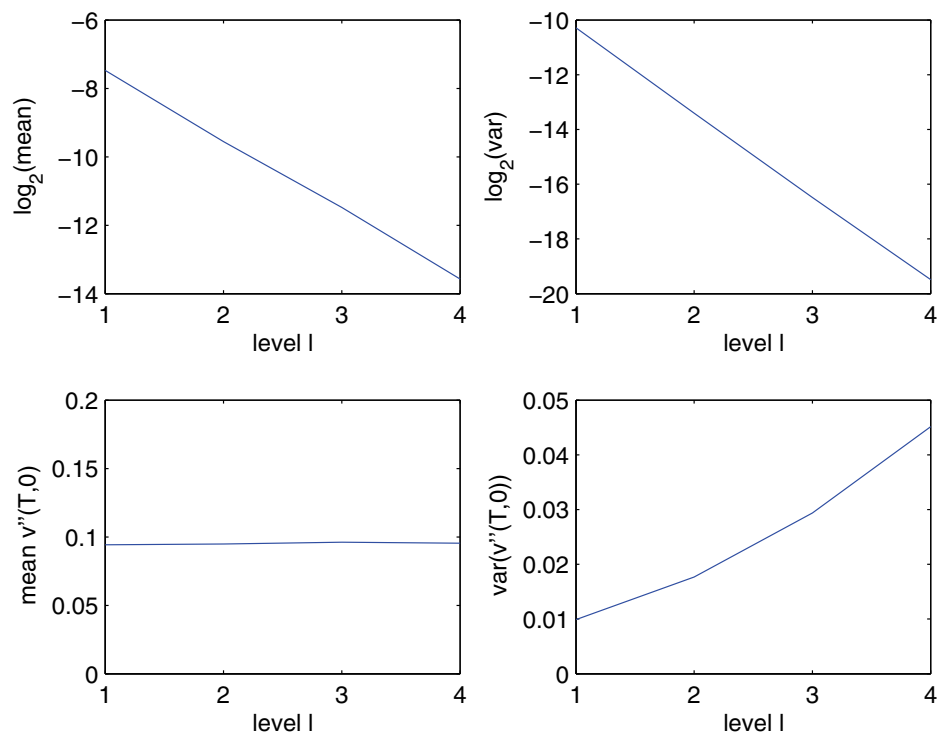
**Figure 5.** *Convergence plots as a function of grid level.*

## REFERENCES

A. BAIN AND D. CRISAN (2009), *Fundamentals of Stochastic Filtering*, Springer, New York.

A. BARTH, C. SCHWAB, AND N. ZOLLINGER (2010), *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*, Numer. Math., 119, pp. 123–161.

F. E. BENTH AND S. KOEKEBAKKER (2008), *Stochastic modeling of financial electricity contracts*, Energ. Econ., 30, pp. 1116–1157.

F. BLACK AND J. COX (1976), *Valuing corporate securities: Some effects of bond indenture provision*, J. Finance, 31, pp. 351–367.

N. BRUTI-LIBERATI AND E. PLATEN (2010), *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*, Springer, Berlin.

E. BUCKWAR AND T. SICKENBERGER (2011), *A comparative linear mean-square stability analysis of Maruyama- and Milstein-type methods*, Math. Comput. Simulation, 81, pp. 1110–1127.

K. BUJOK AND C. REISINGER (2012), *Numerical valuation of basket credit derivatives in structural jump-diffusion models*, J. Comput. Finance, 15, pp. 115–158.

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

R. CARMONA, J.-P. FOUQUE, AND D. VESTAL (2009), *Interacting particle systems for the computation of rare credit portfolio losses*, Finance Stoch., 13, pp. 613–633.

R. CARTER AND M. B. GILES (2007), *Sharp error estimates for discretisations of the 1D convection/diffusion equation with Dirac initial data*, IMA J. Numer. Anal., 27, pp. 406–425.

K. A. CLIFFE, M. B. GILES, R. SCHEICHL, AND A. L. TECKENTRUP (2011), *Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients*, Comput. Vis. Sci., 14, pp. 3–15.

J. FOUQUE, B. WIGNALL, AND X. ZHOU (2008), *Modeling correlated defaults: First passage model under*

*stochastic volatility*, J. Comput. Finance, 11, pp. 43–78.

M. B. GILES (2008a), *Multi-level Monte Carlo path simulation*, Oper. Res., 56, pp. 981–986.

M. B. GILES (2008b), *Improved multilevel Monte Carlo convergence using the Milstein scheme*, in Monte Carlo and Quasi-Monte Carlo Methods 2006, A. Keller, S. Heinrich, and H. Niederreiter, eds., Springer, Berlin, pp. 343–358.

P. GLASSERMAN (2004), *Monte Carlo Methods in Financial Engineering*, Springer, New York.

S. GRAUBNER (2008), *Multi-level Monte Carlo Methoden für stochastische partielle Differentialgleichungen*, Diplomarbeit, TU Darmstadt, Darmstadt, Germany.

I. GYÖNGY (1999), *Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise* II, Potential Anal., 11, pp. 1–37.

I. GYÖNGY AND D. NUALART (1997), *Implicit schemes for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise*, Potential Anal., 7, pp. 725–757.

D. HEATH, R. JARROW, AND A. MORTON (1992), *Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation*, Econometrica, 60, pp. 77–105.

D. J. HIGHAM (2000), *Mean-square and asymptotic stability of the stochastic theta method*, SIAM J. Numer. Anal., 38, pp. 753–769.

J. HULL, M. PREDESCU, AND A. WHITE (2005), *The valuation of correlation-dependent credit derivatives using a structural model*, J. Credit Risk, 6, pp. 99–132.

P. E. KLOEDEN AND E. PLATEN (1992), *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin.

N. V. KRYLOV (1994), *A $W_2^n$-theory of the Dirichlet problem for SPDEs in general smooth domains*, Probab. Theory Related Fields, 98, pp. 389–421.

T. G. KURTZ AND J. XIONG (1999), *Particle representations for a class of nonlinear SPDEs*, Stochastic Process. Appl., 83, pp. 103–126.

A. LANG (2010), *A Lax equivalence theorem for stochastic differential equations*, J. Comput. Appl. Math., 234, pp. 3387–3396.

R. MERTON (1974), *On the pricing of corporate debt: The risk structure of interest rates*, J. Finance, 29, pp. 449–470.

D. M. POOLEY, K. R. VETZAL, AND P. A. FORSYTH (2003), *Remedies for non-smooth payoffs in option pricing*, J. Comput. Finance, 6, pp. 25–40.

R. D. RICHTMYER AND K. W. MORTON (1967), *Difference Methods for Initial-Value Problems*, Wiley-Interscience, New York.

C. ROTH (2002), *Difference methods for stochastic partial differential equations*, ZAMM Z. Angew. Math. Mech., 82, pp. 821–830.

C. ROTH (2006), *A combination of finite difference and Wong-Zakai methods for hyperbolic stochastic partial differential equations*, Stoch. Anal. Appl., 24, pp. 221–240.

Y. SAITO AND T. MITSUI (1996), *Stability analysis of numerical schemes for stochastic differential equations*, SIAM J. Numer. Anal., 33, pp. 2254–2267.

P. J. SCHÖNBUCHER (2003), *Credit Derivatives Pricing Models*, Wiley, New York.

J. B. WALSH (2005), *Finite element methods for parabolic stochastic PDEs*, Potential Anal., 23, pp. 1–43.

Y. XIA AND M. B. GILES (2012), *Multilevel path simulation for jump-diffusion SDEs*, in Monte Carlo and Quasi-Monte Carlo Methods 2010, H. Wozniakowski and L. Plaskota, eds., Springer, Berlin, pp. 657–670.

C. ZHOU (2001), *An analysis of default correlations and multiple defaults*, Rev. Financ. Stud., 14, pp. 555–576.