Two SPDE applications using multilevel Monte Carlo

Mike Giles

mike.giles@maths.ox.ac.uk

Oxford-Man Institute of Quantitative Finance

Mathematical Institute, University of Oxford

with Rob Scheichl, Aretha Teckentrup (Bath),

Andrew Cliffe (Nottingham) and Christoph Reisinger (Oxford)

NIM11 Workshop on Rough Paths and Numerical Integration Methods

Philipps University, Marburg, September 21-23, 2011

Outline

- standard Monte Carlo simulation
- multilevel Monte Carlo simulation
- elliptic SPDE application
- parabolic SPDE application
- conclusions

Monte Carlo simulation

In many stochastic applications want to estimate $\mathbb{E}[P(\omega)]$ where $\omega \in \Omega$ is an infinite-dimensional random variable.

In standard MC simulation we estimate it using

$$\widehat{Y} \;=\; N^{-1} \sum_{n=1}^{N} \widehat{P}(\omega^{(n)})$$

where $\omega^{(n)}$ are N independent samples

Monte Carlo simulation

The mean square error is

$$\begin{aligned}
\mathbb{E}\left[\left(\widehat{Y}-\mathbb{E}[P]\right)^2\right] &= \mathbb{E}\left[\left(\widehat{Y}-\mathbb{E}[\widehat{Y}]\;+\;\mathbb{E}[\widehat{Y}]-\mathbb{E}[P]\right)^2\right] \\
&= \mathbb{E}\left[(\widehat{Y}-\mathbb{E}[\widehat{Y}])^2\right]+\left(\mathbb{E}[\widehat{Y}]-\mathbb{E}[P]\right)^2 \\
&= \mathbb{V}[\widehat{Y}]+\left(\mathbb{E}[\widehat{Y}]-\mathbb{E}[P]\right)^2 \\
&= N^{-1}\mathbb{V}[\widehat{P}]+\left(\mathbb{E}[\widehat{P}]-\mathbb{E}[P]\right)^2
\end{aligned}$$

- first term is Monte Carlo sampling error
- second term is due to discretisation bias

Monte Carlo simulation

To achieve RMS accuracy of ε requires:

- $N = O(\varepsilon^{-2})$
- bias $= O(\varepsilon)$

In an elliptic SPDE application with grid spacing h, if the bias is $O(h^{\alpha})$ then need $h = O(\varepsilon^{1/\alpha})$, and total cost is $O(\varepsilon^{-(2+d/\alpha)})$ in d dimensions

(This assumes efficient multigrid solution – very challenging because of very rough coefficients (Graham & Scheichl))

To get acceptable accuracy in 3D applications may need 10,000 simulations on a 128^3 grid $\Longrightarrow$ very expensive

Multilevel Monte Carlo

The multilevel objective is to greatly reduce this cost:

	$\alpha = 1$		$\alpha = 2$	
d	MC	MLMC	MC	MLMC
1	ε^{-3}	ε^{-2}	$\varepsilon^{-2.5}$	ε^{-2}
2	ε^{-4}	$\varepsilon^{-2}(\log \varepsilon)^2$	ε^{-3}	$\varepsilon^{-2}(\log \varepsilon)^2$
3	ε^{-5}	ε^{-3}	$\varepsilon^{-3.5}$	$\varepsilon^{-2.5}$

How does this compare to theoretical lower bound?

- ε^{-2} calculations on coarsest grid costs $O(\varepsilon^{-2})$
- 1 calculation on finest grid costs $O(\varepsilon^{-d/\alpha})$

so minimum cost is $O(\varepsilon^{-\max(2,d/\alpha)})$

Multilevel Monte Carlo

Consider Monte Carlo simulations with different levels of refinement, $\ell = 0, 1, \ldots, L$, with level L being the finest.

If $\widehat{P}_\ell$ is the approximation of P on level ℓ, then

$$\mathbb{E}[\widehat{P}_\ell] = \mathbb{E}[\widehat{P}_0] + \sum_{l=1}^{L} \mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}].$$

Idea is to independently estimate each of the terms on the r.h.s., in a way which minimises the overall variance for a fixed computational cost.

Finest level is still the same, but will use very few samples at that level.

Multilevel Monte Carlo

Simplest estimator for $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$ for $\ell > 0$ is

$$\widehat{Y}_\ell = N_\ell^{-1} \sum_{n=1}^{N_\ell} \left(\widehat{P}_\ell^{(n)} - \widehat{P}_{\ell-1}^{(n)} \right)$$

using same stochastic sample $\omega^{(n)}$ for both levels

Variance is $N_\ell^{-1} V_\ell$ where $V_\ell = \mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$

If the cost of one sample on level ℓ is C_ℓ then optimally use $N_\ell \propto \sqrt{V_\ell / C_\ell}$, and cost on level ℓ is proportional to $\sqrt{V_\ell \, C_\ell}$.

MLMC Theorem

If there exist independent estimators $\widehat{Y}_\ell$ based on N_ℓ Monte Carlo samples, each costing C_ℓ, and positive constants $\alpha, \beta, \gamma, c_1, c_2, c_3$ such that $\alpha \geq \frac{1}{2}\min(\beta, \gamma)$ and

i) $\left|\mathbb{E}[\widehat{P}_\ell - P]\right| \leq c_1\, 2^{-\alpha\,\ell}$

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

iii) $\mathbb{V}[\widehat{Y}_\ell] \leq c_2\, N_\ell^{-1} 2^{-\beta\,\ell}$

iv) $C_\ell \leq c_3\, 2^{\gamma\,\ell}$

MLMC Theorem

then there exists a positive constant c_4 such that for any $\varepsilon < 1$ there exist L and N_ℓ for which the multilevel estimator

$$\widehat{Y} = \sum_{l=0}^{L} \widehat{Y}_\ell,$$

has a mean-square-error with bound $\mathbb{E}\left[\left(\widehat{Y} - \mathbb{E}[P]\right)^2\right] < \varepsilon^2$

with a computational cost C with bound

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

SPDEs

- (Achi Brandt arguably did first multilevel calculations in statistical physics in 1994 – multi-dimensional but not an SPDE)
- Tom Hou post-doc – first elliptic SPDE work in 2006/7
- Klaus Ritter & Simone Graubner – first parabolic SPDE calculations in 2007
- G and Reisinger – “unusual” parabolic SPDE, 2008-11
- Cliffe, G, Scheichl, Teckentrup – elliptic SPDE, 2009-11
- Barth, Lang, Schwab, Zollinger – elliptic and parabolic SPDEs, 2009-11

SPDEs

As far as I know, all methods use the simple multilevel estimator – unlike SDEs, none use advanced techniques to improve the multilevel correction variance

Hard part with SPDEs is the numerical analysis – recent work on elliptic SPDEs by Charrier, Scheichl & Teckentrup is particularly noteworthy

Elliptic SPDE

We consider the elliptic PDE

$$-\nabla.\left(k(\mathbf{x},\omega)\nabla p(\mathbf{x},\omega)\right)=0, \qquad \mathbf{x}\in D,$$

with **random coefficient** $k(\mathbf{x},\omega)$**.**

We model k as a **lognormal random field**, i.e. $\log k$ is a Gaussian field with mean 0 and covariance function

$$R(\mathbf{x},\mathbf{y})=\sigma^2\exp\left(-\|\mathbf{x}-\mathbf{y}\|_1/\lambda\right)$$

2D numerical experiments use $\sigma=1$ and $\lambda=0.2$ on unit square $[0,1]^2$

Elliptic SPDE

Samples of $\log k$ are provided by a Karhunen-Loève expansion:

$$\log k(\mathbf{x}, \omega) = \sum_{n=0}^{\infty} \sqrt{\theta_n}\ \xi_n(\omega)\ f_n(\mathbf{x}),$$

where θ_n, f_n are eigenvalues / eigenfunctions of the correlation function:

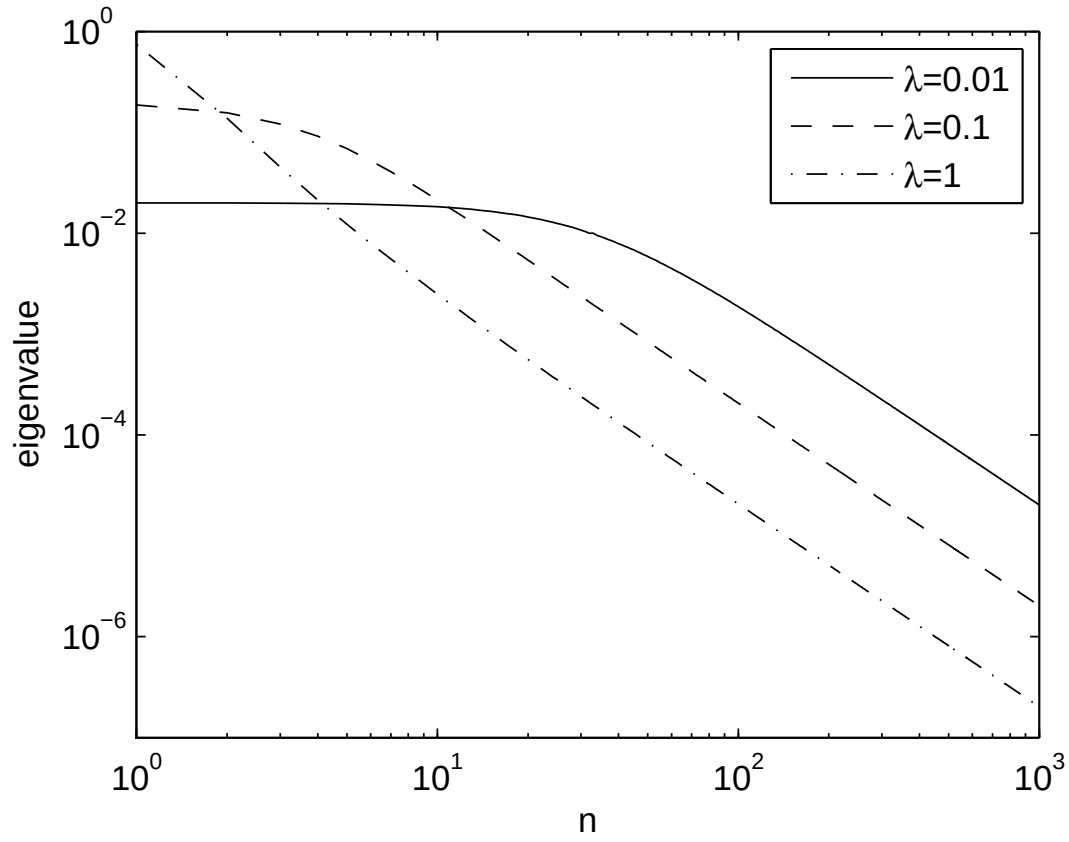
$$\int R(\mathbf{x}, \mathbf{y})\ f_n(\mathbf{y})\ \mathrm{d}\mathbf{y} = \theta_n\, f_n(\mathbf{x})$$

and $\xi_n(\omega)$ are standard Normal random variables.

Numerical experiments truncate the expansion.

Elliptic SPDE

Decay of 1D eigenvalues



When $\lambda = 1$, can use a low-dimensional polynomial chaos approach, but it's impractical for smaller λ.

Elliptic SPDE

Discretisation:

- cell-centred finite volume discretisation on a uniform grid – for rough coefficients we need to make grid spacing very small on finest grid
- each level of refinement has twice as many grid points in each direction
- current numerical experiments use a direct solver for simplicity, but in the future will use an efficient multigrid solver and so “computational cost” is defined to be proportional to the total number of grid points

Elliptic SPDE

Boundary conditions for unit square $[0,1]^2$:
– fixed pressure: $p(0,x_2)=1,\ p(1,x_2)=0$
– Neumann b.c.: $\partial p/\partial x_2(x_1,0)=\partial p/\partial x_2(x_1,1)=0$

Output quantity – mass flux: $-\int k\,\dfrac{\partial p}{\partial x_1}\,\mathrm{d}x_2$

Correlation length: $\lambda = 0.2$

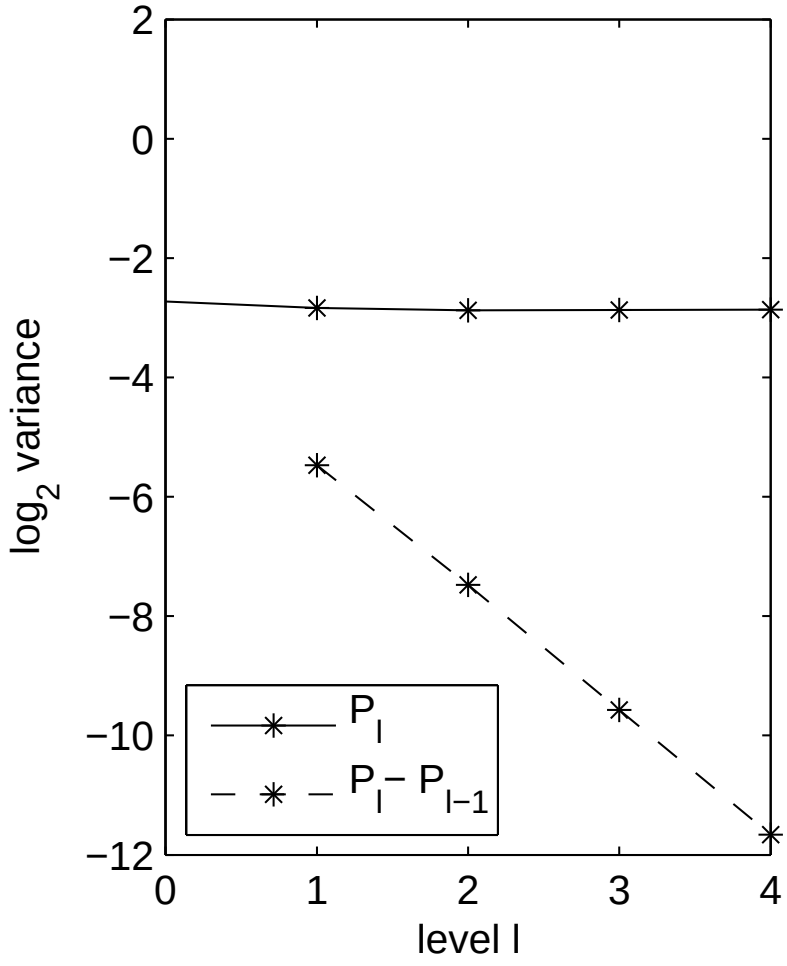
Coarsest grid: $h = 1/8$ (comparable to λ)

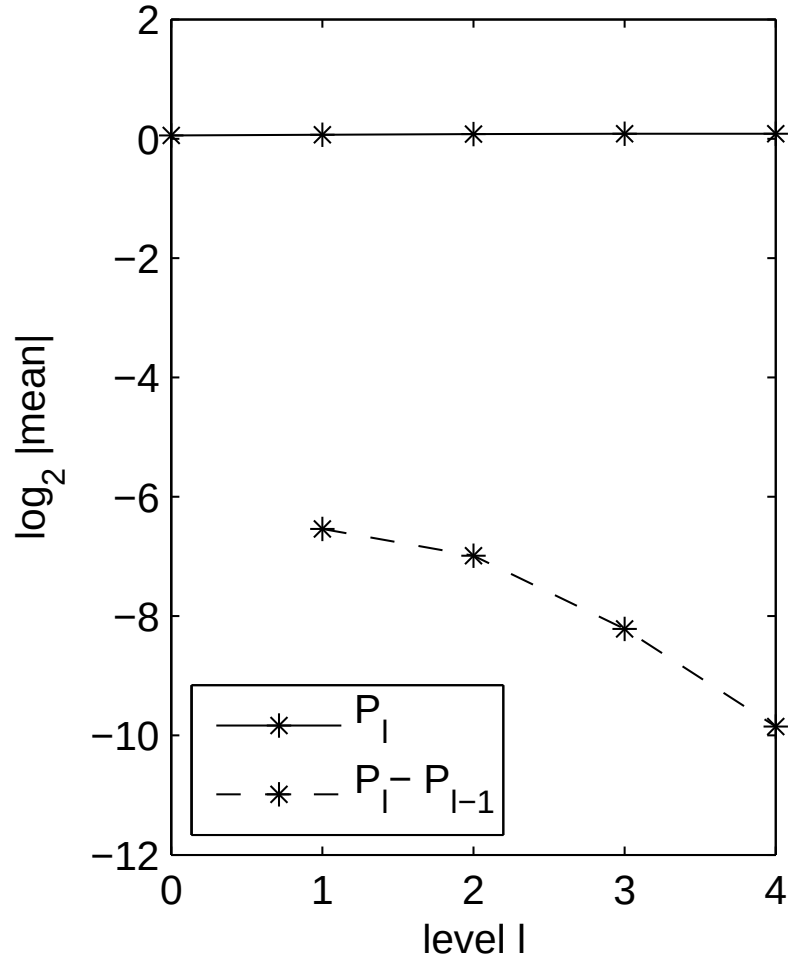
Finest grid: $h = 1/128$

Karhunen-Loève truncation: $m_{KL} = 4000$

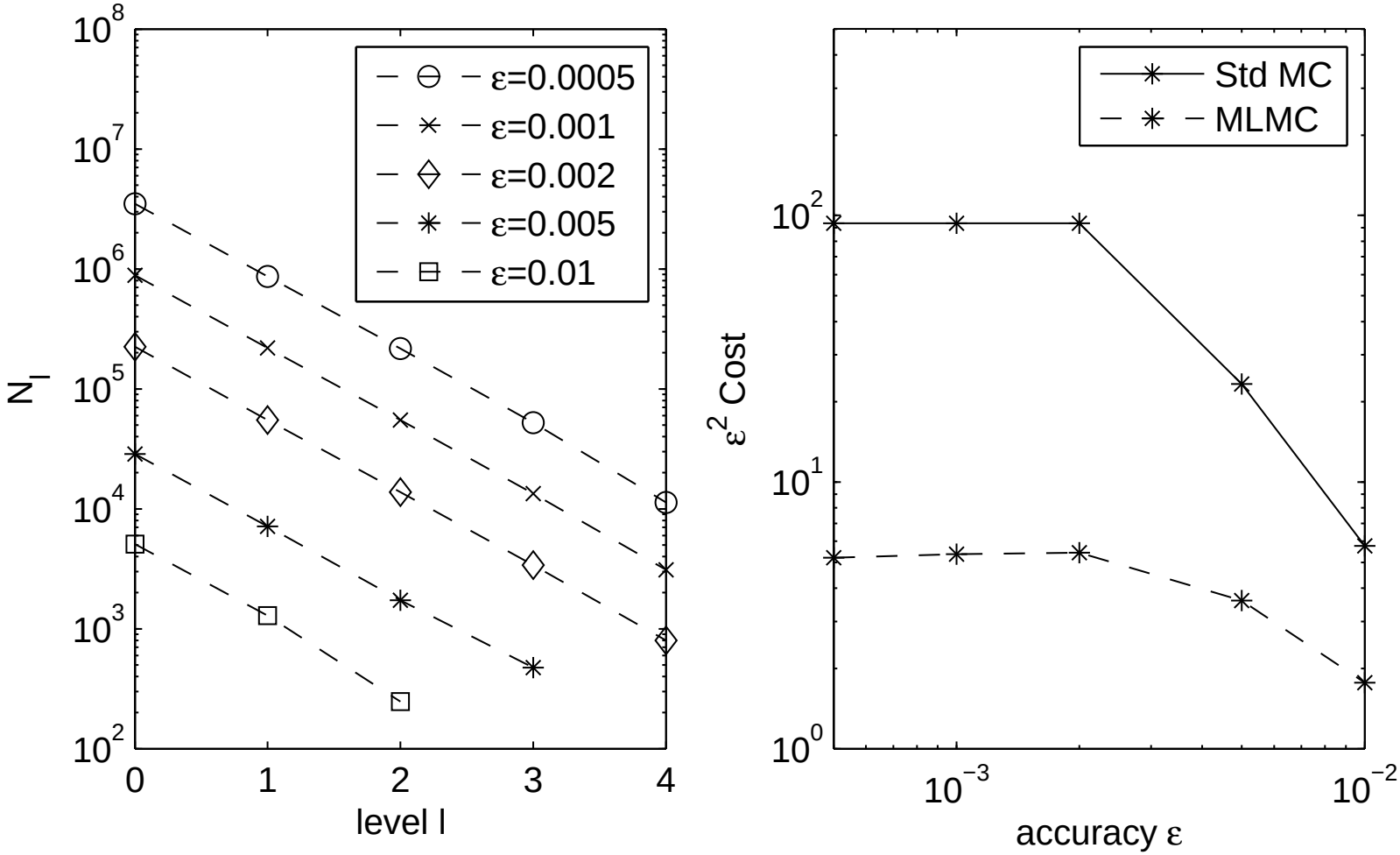
Cost taken to be proportional to number of nodes

Elliptic SPDE





Elliptic SPDE



Greater savings because of greater cost on finer grids

Parabolic SPDE

Unusual parabolic SPDE arises in CDO modelling (Bush, Hambly, Haworth & Reisinger)

$$\mathrm{d}p = -\mu \frac{\partial p}{\partial x}\,\mathrm{d}t + \frac{1}{2}\frac{\partial^2 p}{\partial x^2}\,\mathrm{d}t + \sqrt{\rho}\frac{\partial p}{\partial x}\,\mathrm{d}W$$

with absorbing boundary $p(0,t) = 0$

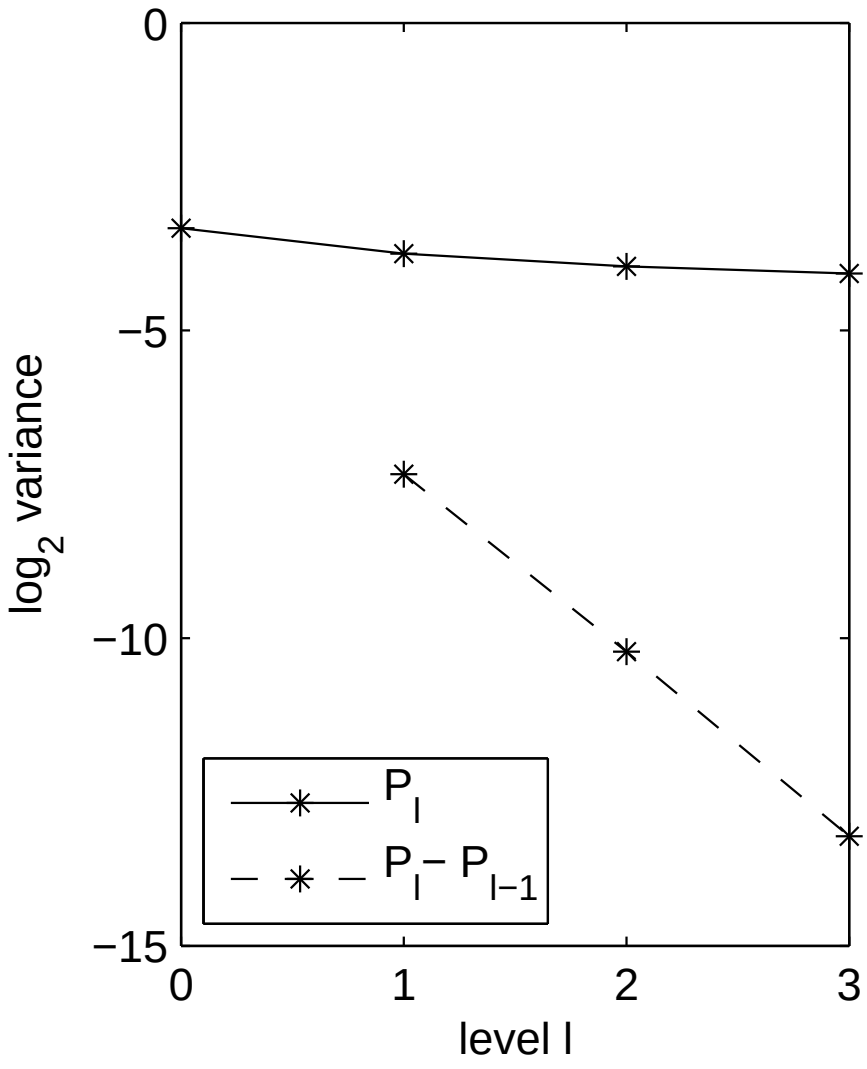
- derived in limit as number of firms $\longrightarrow \infty$
- x is distance to default
- $p(x,t)$ is probability density function
- $\mathrm{d}W$ term corresponds to systemic risk
- $\partial^2 p/\partial x^2$ comes from idiosyncratic risk

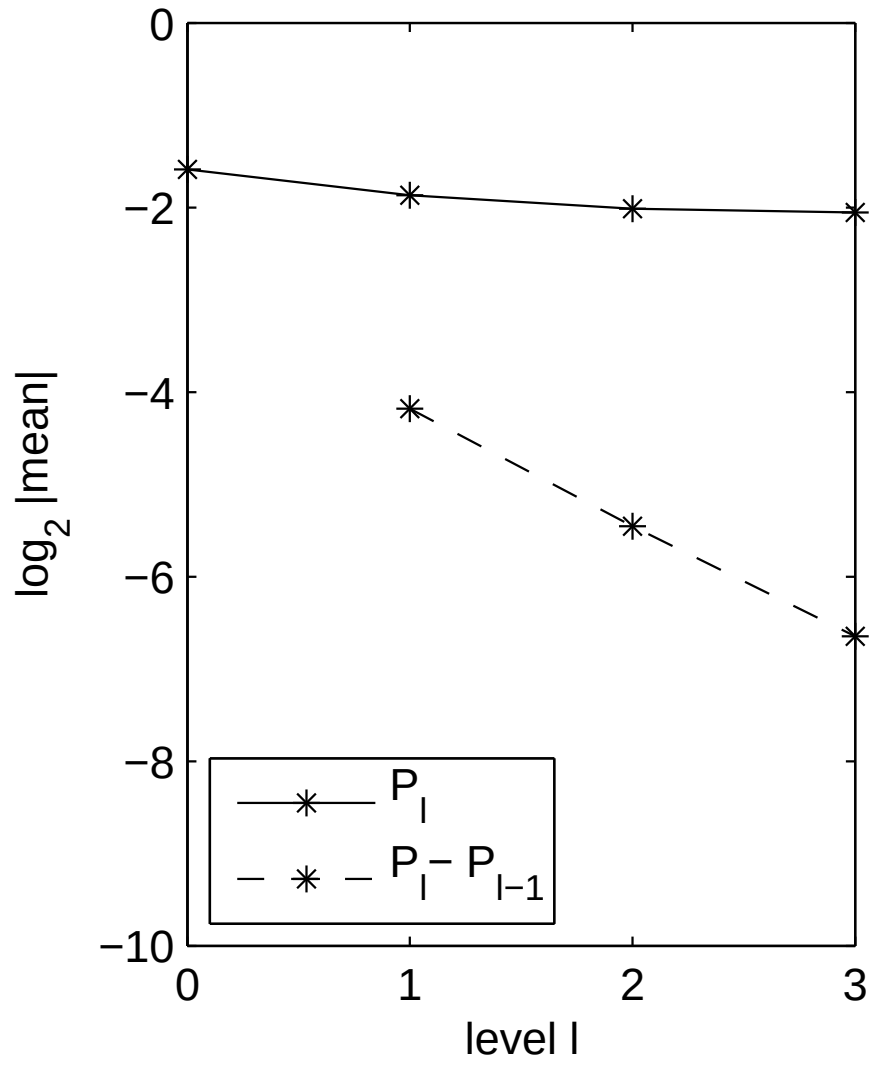
Parabolic SPDE

- numerical discretisation combines Milstein time-marching with central difference approximations
- coarsest level of approximation uses 1 timestep per quarter, and 10 spatial points
- each finer level uses four times as many timesteps, and twice as many spatial points – ratio is due to numerical stability constraints
- mean-square stability theory, with and without absorbing boundary
- computational cost $C_\ell \propto 8^\ell$
- numerical results suggest variance $V_\ell \propto 8^{-\ell}$
- can prove $V_\ell \propto 16^{-\ell}$ when no absorbing boundary

Parabolic SPDE

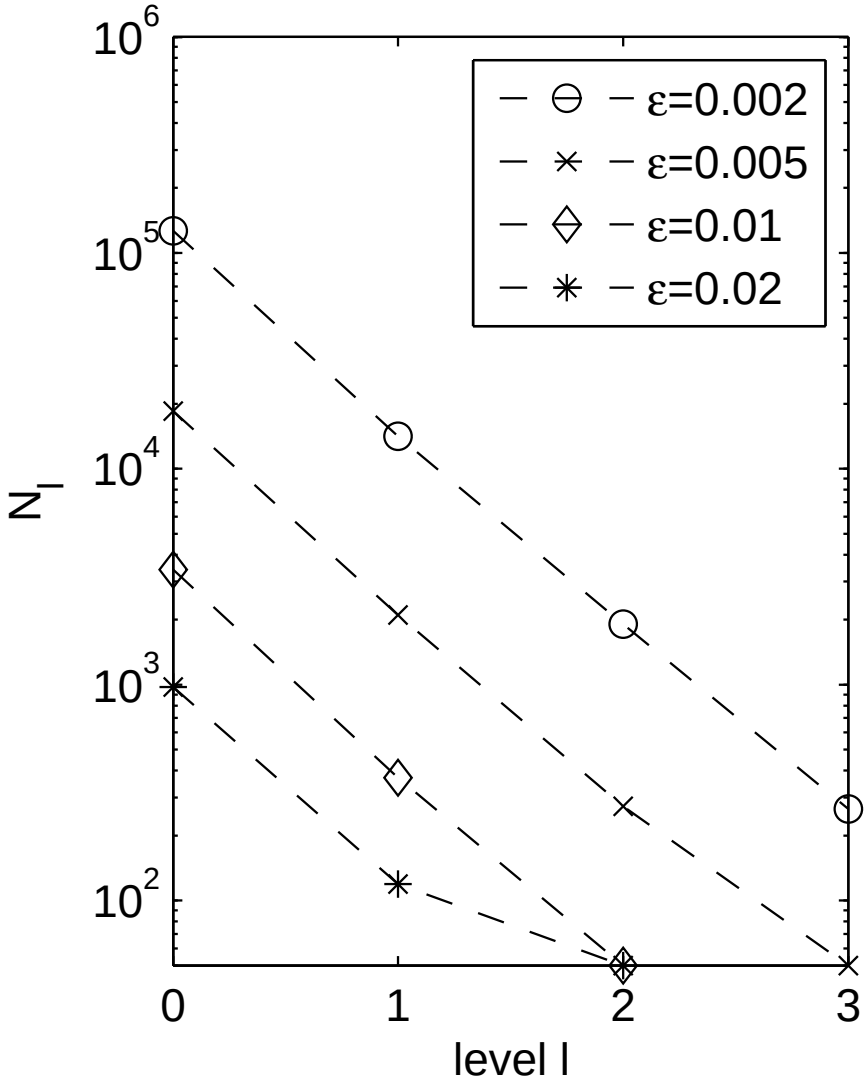
Fractional loss on equity tranche of a 5-year CDO:

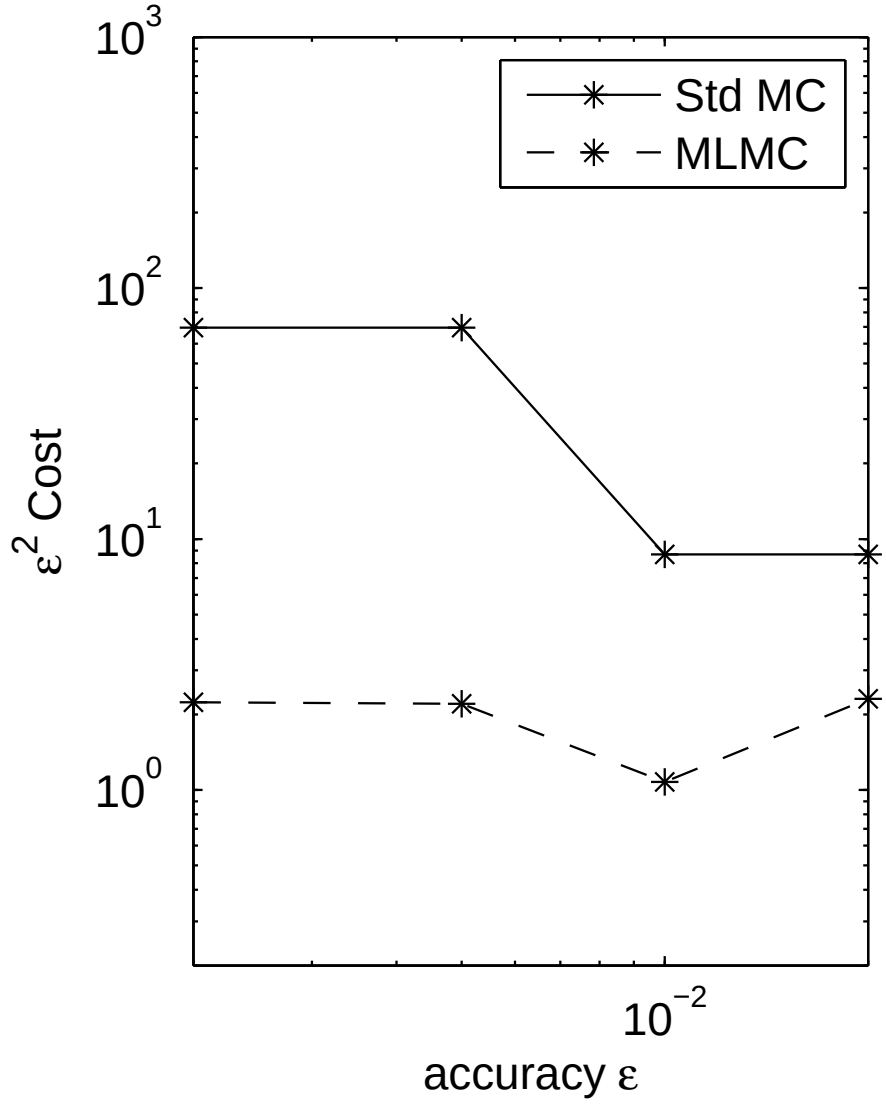




Parabolic SPDE

Fractional loss on equity tranche of a 5-year CDO:





Parabolic SPDE

Milstein and central difference discretisation leads to

$$v_j^{n+1} = v_j^n - \frac{\mu\, k + \sqrt{\rho\, k}\, Z_n}{2h}\left(v_{j+1}^n - v_{j-1}^n\right) + \frac{(1-\rho)\, k + \rho\, k\, Z_n^2}{2h^2}\left(v_{j+1}^n - 2v_j^n + v_{j-1}^n\right)$$

where $Z_n \sim N(0,1)$.

Considering a Fourier mode

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

leads to ...

Parabolic SPDE

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

where

$$\begin{aligned} a(\theta) &= 1 - \frac{i\,\mu\,k}{h}\,\sin\theta - \frac{2\,(1-\rho)\,k}{h^2}\,\sin^2\tfrac{\theta}{2}, \\ b(\theta) &= -\frac{i\sqrt{\rho\,k}}{h}\,\sin\theta, \\ c(\theta) &= -\frac{2\,\rho\,k}{h^2}\,\sin^2\tfrac{\theta}{2}. \end{aligned}$$

Parabolic SPDE

Following the approach of mean-square stability analysis (e.g. see Higham)

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

so stability requires $|a+c|^2+|b|^2+2|c|^2 \;\leq\; 1$ for all θ, which leads to a timestep stability limit:

$$\begin{aligned}\mu^2 k &\leq 1-\rho, \\ \frac{k}{h^2} &\leq (1+2\rho^2)^{-1}.\end{aligned}$$

Parabolic SPDE

This can be extended to finite domains using matrix stability analysis, writing the discrete equations as

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

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

and D_1 and D_2 look like

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

Parabolic SPDE

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

D_1 is anti-symmetric and D_2 is symmetric, and

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

where D_3 looks like

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

Parabolic SPDE

and E_1 and E_2 are zero apart from one corner element,

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

This leads to

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

where e_1 and e_2 are scalars and

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

Parabolic SPDE

It can be verified that the m^{th} eigenvector of M is a Fourier mode 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 before.

In the limit $h, k/h \to 0$, $e_1 \pm e_2 > 0$, and therefore 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.

Conclusions

- multilevel Monte Carlo greatly reduces the cost of SPDE simulations, making it feasible for the first time for engineering applications
- future work on elliptic SPDEs will move to 3D (which requires efficient solvers for each PDE solution), and less idealistic modelling (which requires work on creating stochastic instances of the permeability field)
- numerical analysis is tough, but Charrier, Scheichl & Teckentrup have made excellent progress on finite element analysis of the variance of the multilevel correction
- most interesting aspect of unusual parabolic SPDE application is the numerical stability analysis

Papers

- K.A. Cliffe, M.B. Giles, R. Scheichl and A.L. Teckentrup, Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients, Computing and Visualisation in Science, 2011.
- J. Charrier, R. Scheichl and A.L. Teckentrup, Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods, submitted, 2011.
- M.B. Giles and C. Reisinger, Stochastic finite differences and multilevel Monte Carlo for a class of SPDEs in finance, submitted, 2011.

`people.maths.ox.ac.uk/gilesm/mlmc.html`

MLMC Community Webpage

`people.maths.ox.ac.uk/gilesm/mlmc_community.html`

I will try to keep this updated with papers and presentations by key researchers in MLMC from around the world.

Germany is well represented, partly due to DFG SPP 1324 which has sponsored this workshop.