Multilevel Monte Carlo method for SPDEs and PDEs with uncertain data

Mike Giles

Mathematical Institute, University of Oxford

International Conference on Partial Differential Equations

Academy of Mathematics and System Sciences,
Chinese Academy of Sciences

August 15, 2015

PDEs with Uncertainty

Looking at the history of numerical methods for PDEs, the first steps were about improving the modelling:

- 1D $\rightarrow$ 2D $\rightarrow$ 3D
- steady $\rightarrow$ unsteady
- laminar flow $\rightarrow$ turbulence modelling $\rightarrow$ large eddy simulation $\rightarrow$ direct Navier-Stokes
- simple geometries (e.g. a wing) $\rightarrow$ complex geometries (e.g. an aircraft in landing configuration)
- adding new features such as combustion, coupling to structural / thermal analyses, etc.

. . . and then engineering switched from analysis to design.

PDEs with Uncertainty

There is now a big move towards handling uncertainty:

- uncertainty in modelling parameters
- uncertainty in geometry
- uncertainty in initial conditions
- uncertainty in spatially-varying material properties
- inclusion of stochastic source terms

Engineering wants to move to "robust design" taking into account the effects of uncertainty.

Other areas want to move into Bayesian inference, starting with an *a priori* distribution for the uncertainty, and then using data to derive an improved *a posteriori* distribution.

PDEs with Uncertainty

Examples:

- Long-term climate modelling:

  Lots of sources of uncertainty including the effects of aerosols, clouds, carbon cycle, ocean circulation (`http://climate.nasa.gov/uncertainties`)

- Short-range weather prediction

  Considerable uncertainty in the initial data due to limited measurements

- Engineering analysis

  Perhaps the biggest uncertainty is geometric due to manufacturing tolerances

- Nuclear waste repository and oil reservoir modelling

  Considerable uncertainty about porosity of rock

PDEs with Uncertainty

In the past, Monte Carlo simulation has been viewed as impractical due to its expense, and so people have used other methods:

- stochastic collocation
- polynomial chaos

Because of Multilevel Monte Carlo, this is changing and there are now several research groups using MLMC for PDE applications

Monte Carlo method

Given a function f of a random input ω, to estimate the value of $\mathbb{E}[f]$ we can use the Monte Carlo estimate

$$N^{-1} \sum_{n=1}^{N} f(\omega^{(n)}).$$

based on N independent samples $\omega^{(n)}$.

By the Central Limit Theorem, as $N \to \infty$, the error in this estimate becomes Normally distributed, with variance $N^{-1}\mathbb{V}[f]$.

The error lies within 3 s.d. with probability 99.7%, giving us a confidence interval.

Control variate

Classic approach to variance reduction: approximate $\mathbb{E}[f]$ using

$$N^{-1}\sum_{n=1}^{N}\left\{f(\omega^{(n)})-\lambda\left(g(\omega^{(n)})-\mathbb{E}[g]\right)\right\}$$

where

- control variate g has known expectation $\mathbb{E}[g]$
- g is well correlated with f, and optimal value for λ can be estimated by a few samples

For the optimal value of λ, the variance is reduced by factor $(1-\rho^2)$, where ρ is the correlation between f and g.

Two-level Monte Carlo

If we want to estimate $\mathbb{E}[f_1]$ but it is much cheaper to simulate $f_0 \approx f_1$, then since

$$\mathbb{E}[f_1] = \mathbb{E}[f_0] + \mathbb{E}[f_1 - f_0]$$

we can use the estimator

$$N_0^{-1} \sum_{n=1}^{N_0} f_0^{(0,n)} \;+\; N_1^{-1} \sum_{n=1}^{N_1} \left(f_1^{(1,n)} - f_0^{(1,n)} \right)$$

Two differences from standard control variate method:

- $\mathbb{E}[f_0]$ is not known, so has to be estimated
- $\lambda = 1$

Benefit: if $f_1 - f_0$ is small, won't need many samples to accurately estimate $\mathbb{E}[f_1 - f_0]$, so cost will be reduced greatly.

Multilevel Monte Carlo

Natural generalisation: given a sequence $f_0, f_1, \ldots, f_L$

$$\mathbb{E}[f_L] = \mathbb{E}[f_0] + \sum_{\ell=1}^{L} \mathbb{E}[f_\ell - f_{\ell-1}]$$

we can use the estimator

$$N_0^{-1} \sum_{n=1}^{N_0} f_0^{(0,n)} \; + \; \sum_{\ell=1}^{L} \left\{ N_\ell^{-1} \sum_{n=1}^{N_\ell} \left(f_\ell^{(\ell,n)} - f_{\ell-1}^{(\ell,n)} \right) \right\}$$

with independent estimation for each level

Multilevel Monte Carlo

If we define

- C_0, V_0 to be cost and variance of f_0
- C_ℓ, V_ℓ to be cost and variance of $f_\ell - f_{\ell-1}$

then the total cost is $\sum_{\ell=0}^{L} N_\ell \, C_\ell$ and the variance is $\sum_{\ell=0}^{L} N_\ell^{-1} V_\ell$.

Using a Lagrange multiplier μ^2 to minimise the cost for a fixed variance

$$\frac{\partial}{\partial N_\ell} \sum_{k=0}^{L} \left(N_k \, C_k + \mu^2 N_k^{-1} V_k \right) = 0$$

gives

$$N_\ell = \mu \sqrt{V_\ell / C_\ell} \quad \Longrightarrow \quad N_\ell \, C_\ell = \mu \sqrt{V_\ell \, C_\ell}$$

Multilevel Monte Carlo

Setting the total variance equal to ε^2 gives

$$\mu = \varepsilon^{-2} \left(\sum_{\ell=0}^{L} \sqrt{V_\ell \, C_\ell} \right)$$

and hence, the total cost is

$$\sum_{\ell=0}^{L} N_\ell \; C_\ell \;=\; \varepsilon^{-2} \left(\sum_{\ell=0}^{L} \sqrt{V_\ell C_\ell} \right)^2$$

in contrast to the standard cost which is approximately $\varepsilon^{-2} \, V_0 \, C_L$.

The MLMC cost savings are therefore:

- V_L/V_0, if $\sqrt{V_\ell C_\ell}$ increases with level
- C_0/C_L, if $\sqrt{V_\ell C_\ell}$ decreases with level

General 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) $\mathbb{E}[C_\ell] \leq c_3\, 2^{\gamma\,\ell}$

General 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

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

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

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

Proof: M.B. Giles, *Operations Research* (2008)

Discussion

If $\beta > \gamma$, the cost is $O(\varepsilon^{-2})$:

- Monte Carlo requires $O(\varepsilon^{-2})$ samples to get a RMS error of ε
- average cost per sample is $O(1)$
- can't do any better without going to Quasi-Monte Carlo

When $\beta < \gamma$, the cost is $O(\varepsilon^{-2-(\gamma-\beta)/\alpha})$:

- cost of one calculation on the finest level is $O(\varepsilon^{-\gamma/\alpha})$
- MLMC is equivalent to $O(\varepsilon^{-2+\beta/\alpha})$ calcs on finest level
- if $\beta = 2\,\alpha$, that's $O(1)$ calcs on finest level — can't do any better

Engineering Uncertainty Quantification

- consider a 3D elliptic PDE, with uncertain boundary data
- use grid spacing proportional to $2^{-\ell}$ on level ℓ
- cost is $O(2^{+3\ell})$ if using an efficient multigrid solver
- assuming 2nd order accuracy for output P means that

$$\widehat{P}_\ell(\omega) - P(\omega) \approx c(\omega)\, 2^{-2\ell}$$

$$\Longrightarrow \quad \widehat{P}_{\ell-1}(\omega) - \widehat{P}_\ell(\omega) \approx 3\, c(\omega)\, 2^{-2\ell}$$

- hence, $\alpha=2,\ \beta=4,\ \gamma=3$
- cost is $O(\varepsilon^{-2})$ to obtain ε RMS accuracy
- cost is $O(\varepsilon^{-1.5})$ for a single deterministic PDE simulation
- cost would be $O(\varepsilon^{-3.5})$ for ordinary Monte Carlo

Parabolic SPDE

Unusual parabolic SPDE arises in a financial setting
(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
- stochastic $\mathrm{d}W$ term corresponds to systemic risk
  (a market crash or recession hits all firms)
- $\partial^2 p/\partial x^2$ comes from idiosyncratic risk
  (each firm is affected by its own unique circumstances)

Parabolic SPDE

Numerical discretisation combines Milstein time-marching with central difference spatial approximation:

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

where k is the timestep, h is the grid spacing, and $\Delta W_n \sim N(0,k)$.

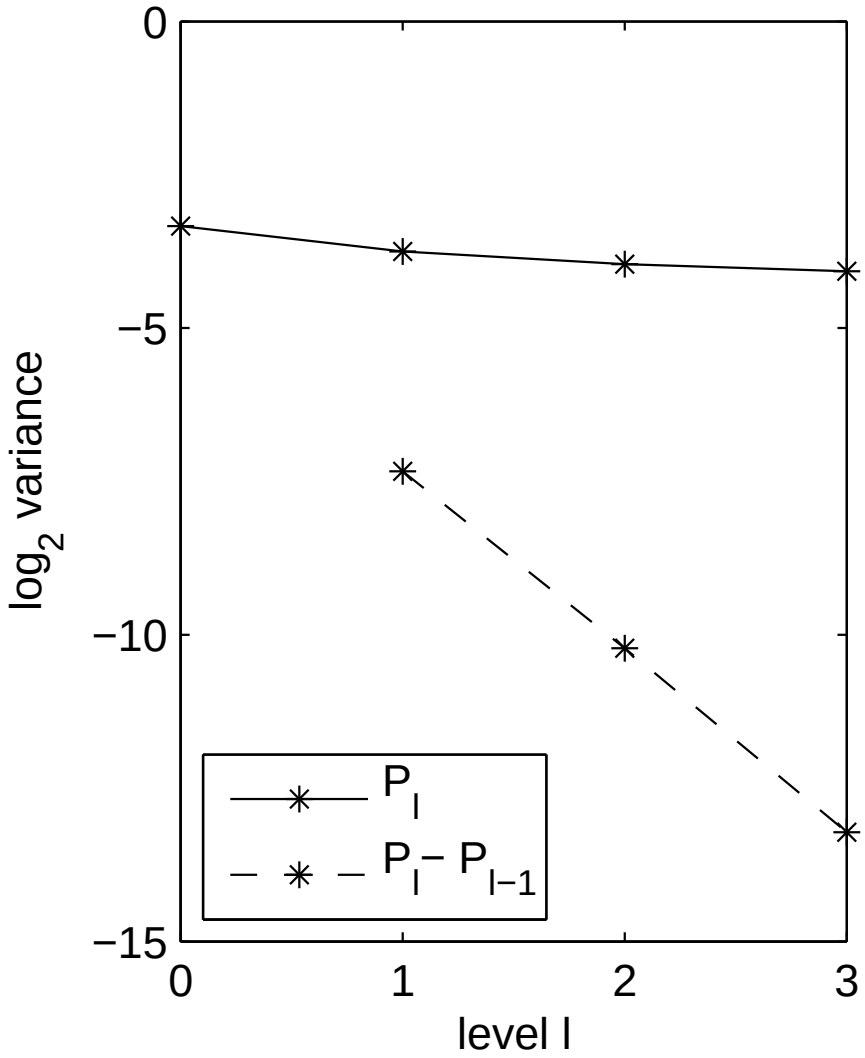
Each finer level uses four times as many timesteps, and twice as many spatial points, due to numerical stability constraints.

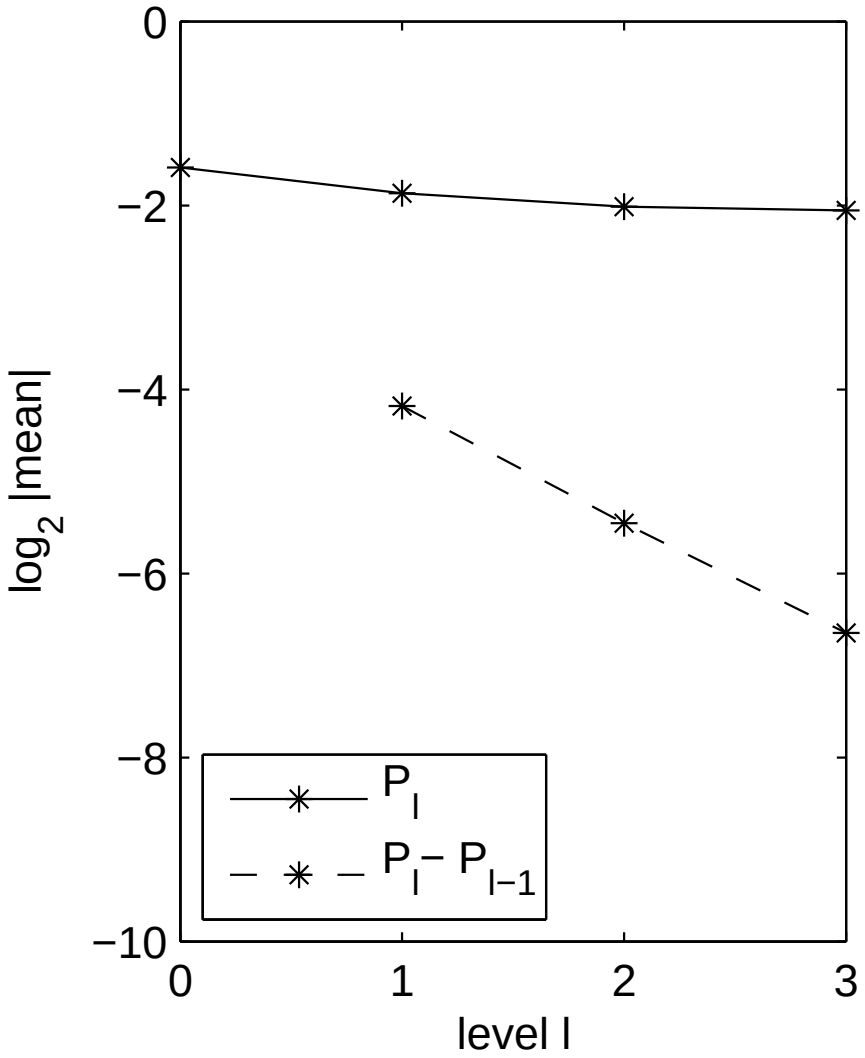
Parabolic SPDE

- coarsest level of approximation uses 1 timestep per quarter, and 10 spatial points
- implementation was really very easy – most interesting part of research was 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

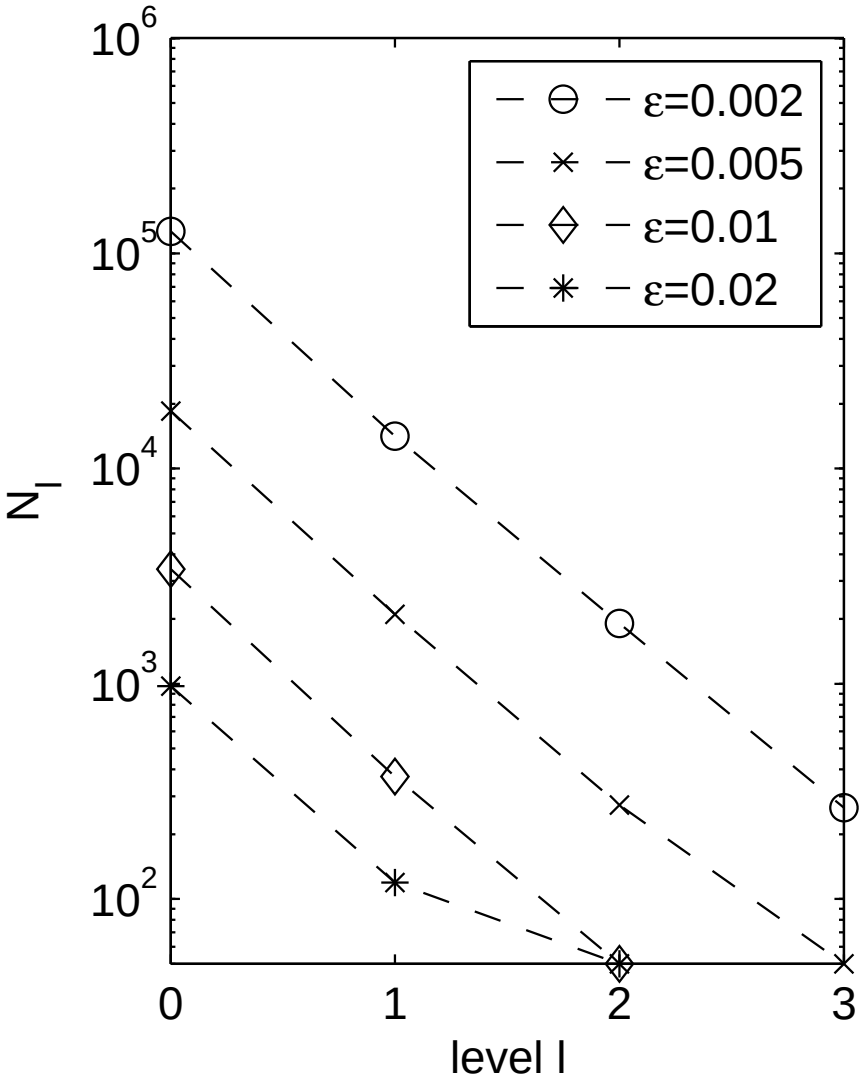
Fractional loss on equity tranche of a 5-year CDO:

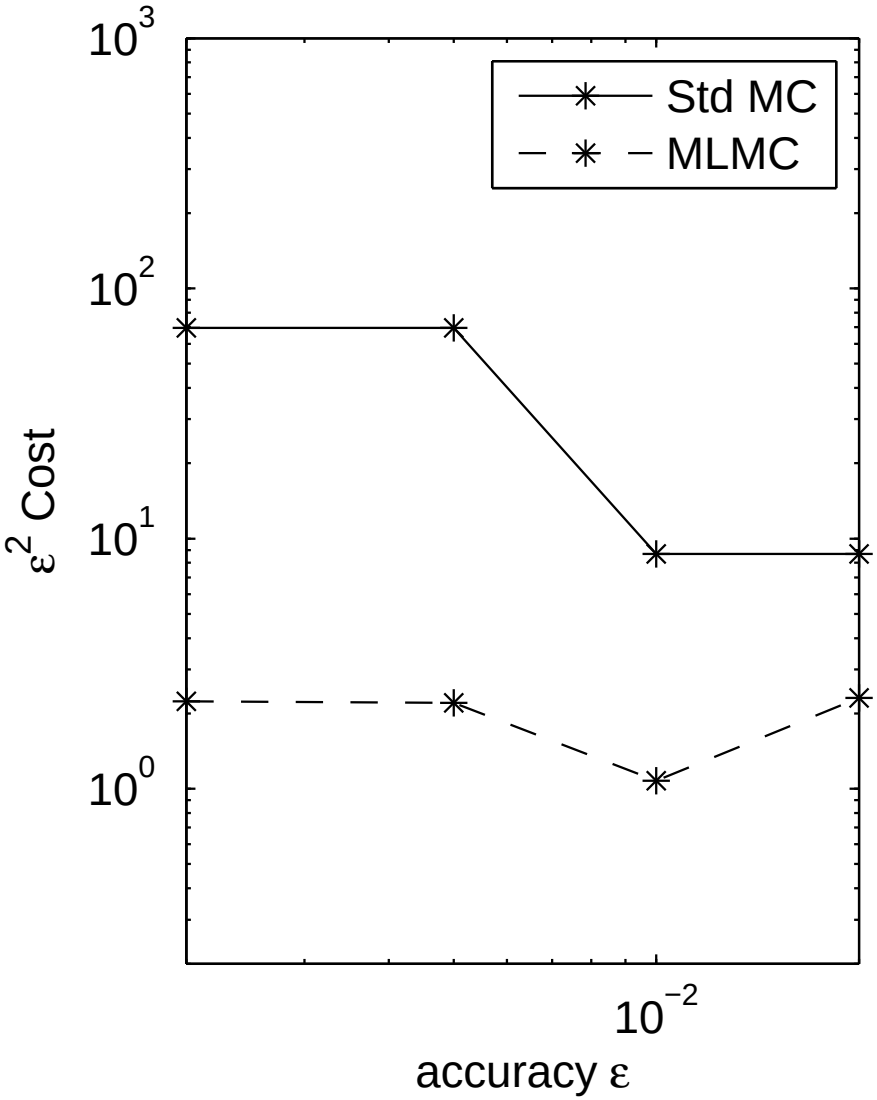




Parabolic SPDE

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





Elliptic SPDEs

I worked with Rob Scheichl (Bath) and Andrew Cliffe (Nottingham) on multilevel Monte Carlo for the modelling of oil reservoirs and groundwater contamination in nuclear waste repositories.

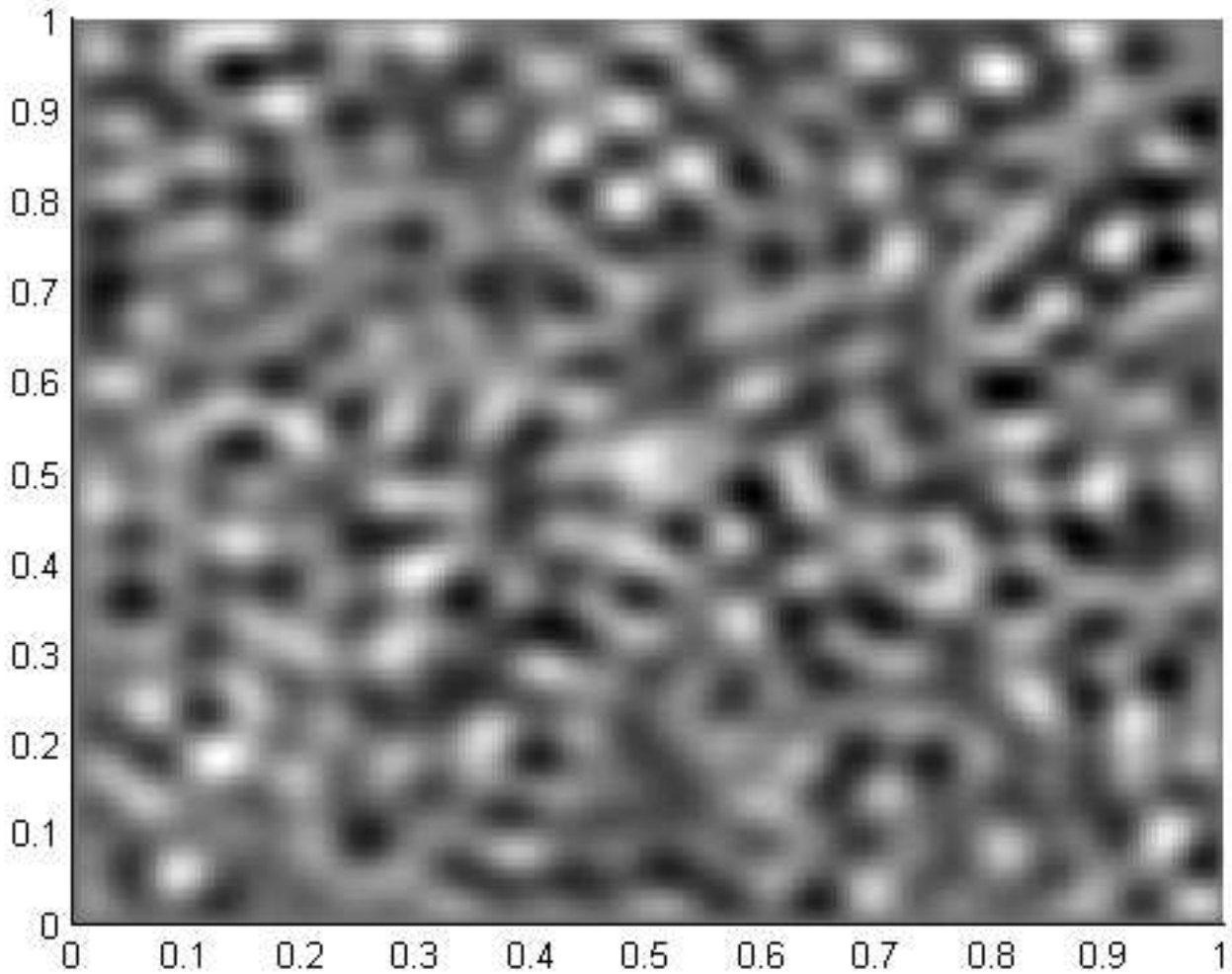
Here we have an elliptic SPDE coming from Darcy's law:

$$\nabla \cdot \Big(\kappa(x) \nabla p \Big) = 0$$

where the permeability $\kappa(x)$ is uncertain, and $\log \kappa(x)$ is often modelled as being Normally distributed with a spatial covariance such as

$$\mathrm{cov}(\log \kappa(x_1), \log \kappa(x_2)) = \sigma^2 \exp(-\|x_1 - x_2\|/\lambda)$$

Elliptic SPDE



A typical realisation of κ for $\lambda = 0.001,\ \sigma = 1$.

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:

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

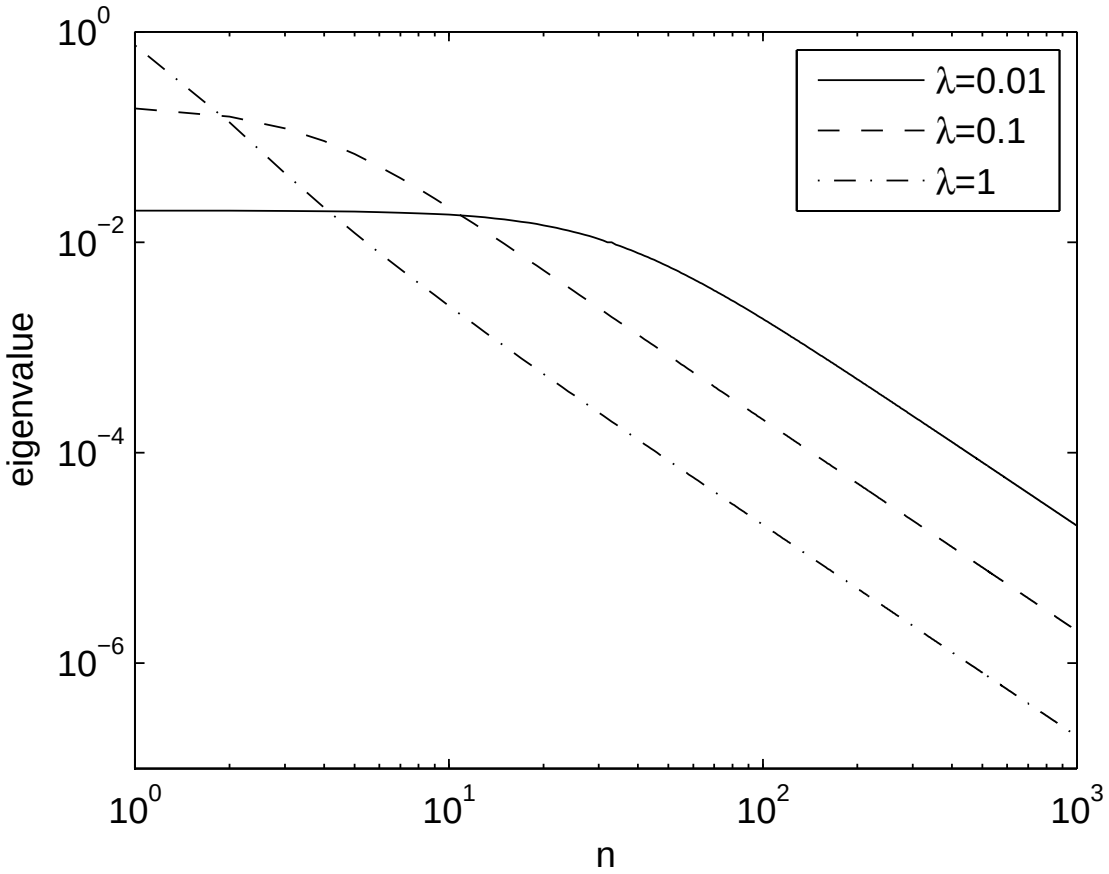
and $\xi_n(\omega)$ are standard Normal random variables.

Numerical experiments truncate the expansion.

(Latest 2D/3D work uses an efficient FFT construction based on a circulant embedding.)

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
- old numerical experiments used a direct solver for simplicity, but later 3D work used an efficient AMG multigrid solver with a cost roughly proportional to the total number of grid points

2D Results

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: $\lambda = 0.2$

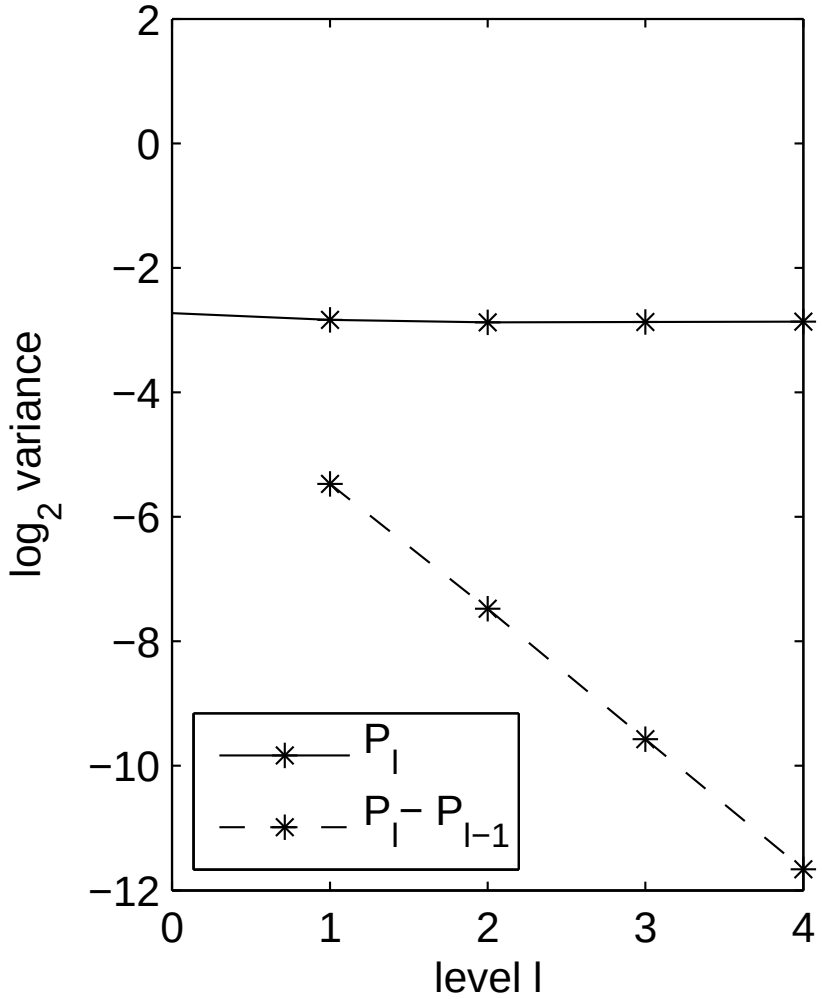
Coarsest grid: $h = 1/8$ (comparable to λ)

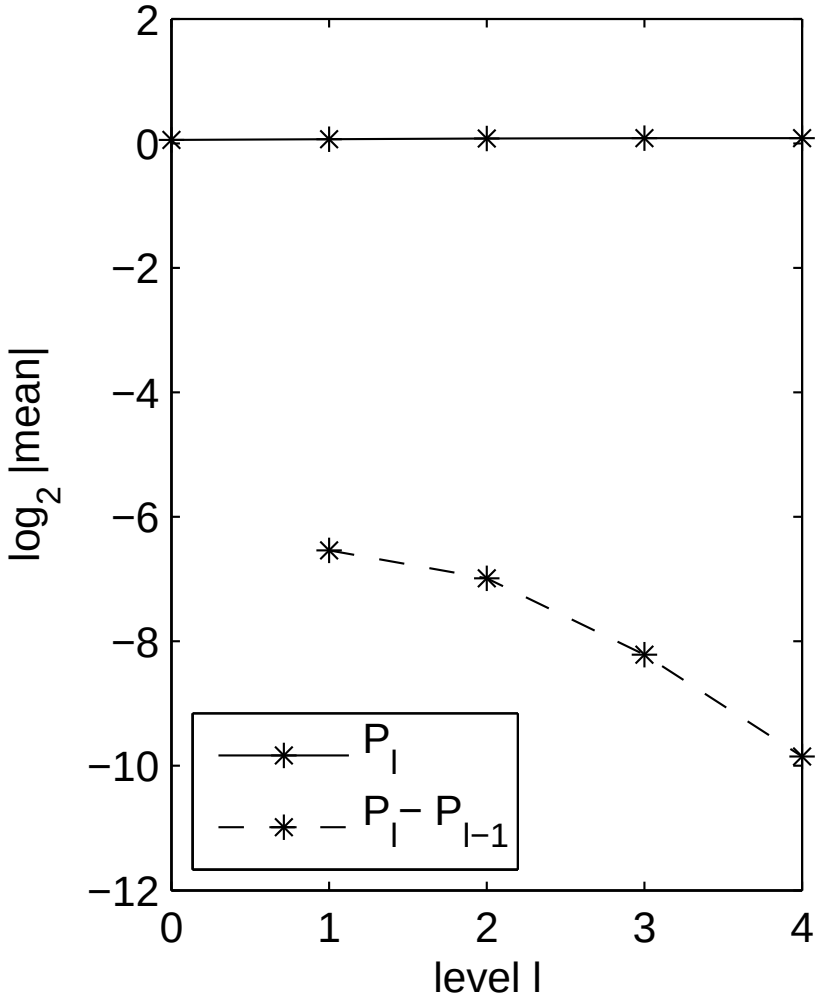
Finest grid: $h = 1/128$

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

Cost taken to be proportional to number of nodes

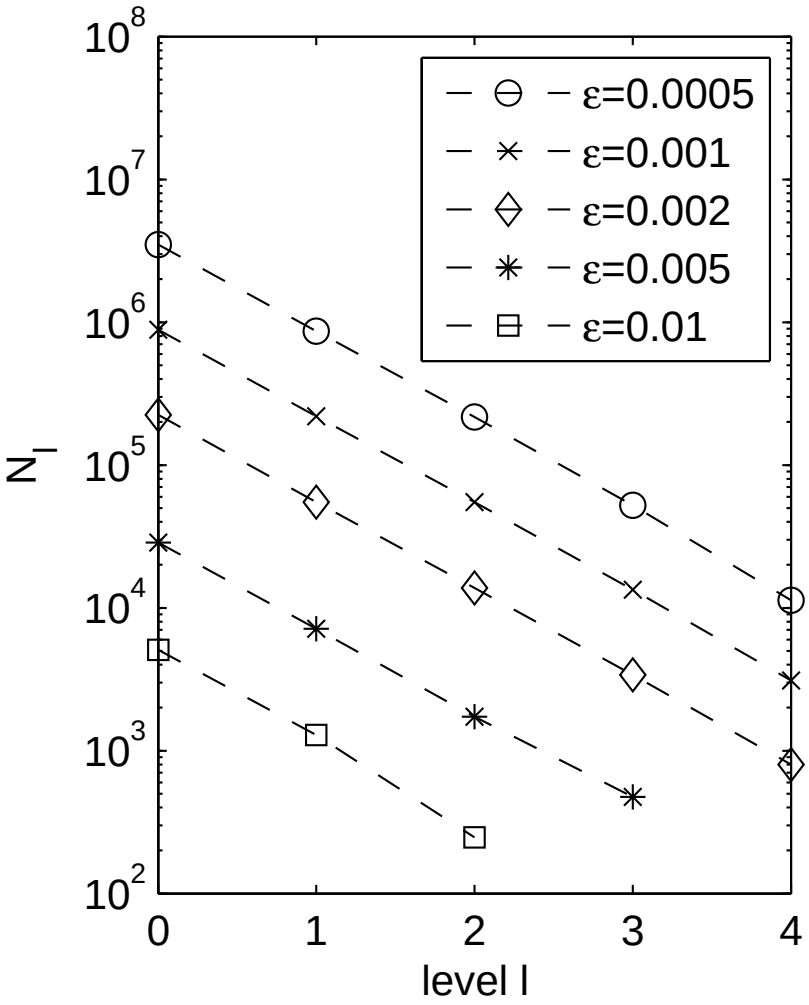
2D Results

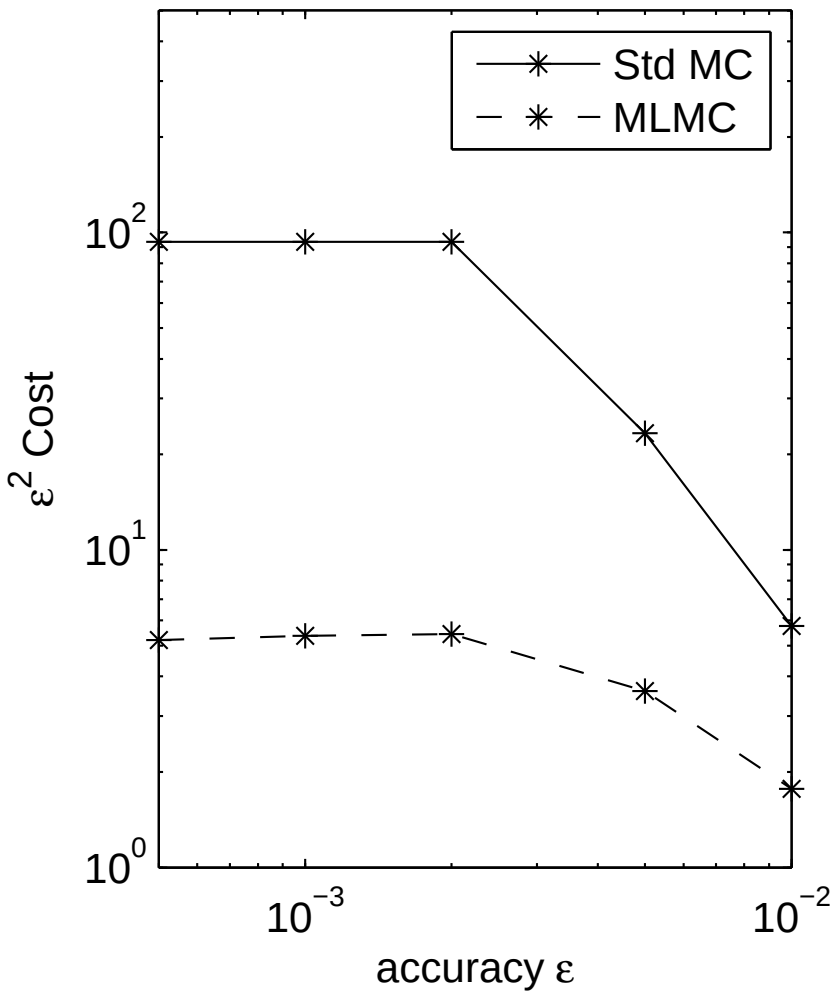




$$\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell^2 \qquad \mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell^2$$

2D Results


ε=0.0005
ε=0.001
ε=0.002
ε=0.005
ε=0.01
N_l
level l


Std MC
MLMC
ε^2 Cost
accuracy ε

Complexity analysis

Relating things back to the MLMC theorem:

$$\begin{aligned} \mathbb{E}[\widehat{P}_\ell - P] &\sim 2^{-2\ell} \implies \alpha = 2 \\ V_\ell &\sim 2^{-2\ell} \implies \beta = 2 \\ C_\ell &\sim 2^{d\ell} \implies \gamma = d \quad \text{(dimension of PDE)} \end{aligned}$$

To achieve r.m.s. accuracy ε requires finest level grid spacing $h \sim \varepsilon^{1/2}$ and hence we get the following complexity:

dim	MC	MLMC
1	$\varepsilon^{-2.5}$	ε^{-2}
2	ε^{-3}	$\varepsilon^{-2}(\log \varepsilon)^2$
3	$\varepsilon^{-3.5}$	$\varepsilon^{-2.5}$

Non-geometric multilevel

Almost all applications of multilevel in the literature so far use a geometric sequence of levels, refining the timestep (or the spatial discretisation for PDEs) by a constant factor when going from level ℓ to level $\ell + 1$.

Coming from a multigrid background, this is very natural, but it is **NOT** a requirement of the multilevel Monte Carlo approach.

All MLMC needs is a sequence of levels with

- increasing accuracy
- increasing cost
- increasingly small difference between outputs on successive levels

Reduced Basis PDE approximation

Vidal-Codina, Nguyen, G, Peraire (2014) take a fine FE discretisation:

$$A(\omega)\, u = f(\omega)$$

and use a reduced basis approximation

$$u \approx \sum_{k=1}^{K} v_k\, u_k$$

to obtain a low-dimensional reduced system

$$A_r(\omega)\, v = f_r(\omega)$$

- larger $K \Longrightarrow$ greater accuracy at greater cost
- in multilevel treatment, K_ℓ varies with level
- brute force optimisation determines the optimal number of levels, and reduced basis size on each level

Other SPDE applications

- Schwab (ETH Zürich) – elliptic, parabolic, hyperbolic
- Jenny (ETH Zürich) – fluids, two-phase flow
- Efendiev (Texas A&M) & Iliev (ITWM)– two-phase flow, numerical homogenization
- Hou (Caltech) – numerical homogenization
- Harbrecht (Basel) – elliptic
- Tempone (KAUST) & Nobile (EPFL) – elliptic, Multi-Index Monte Carlo (important generalisation of MLMC)

For other papers on multilevel, see my MLMC community homepage:

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

Final words

- multilevel Monte Carlo has made the Monte Carlo approach viable for engineering applications which require the solution of SPDEs or PDEs with random data.
- the implementation is often very straightforward, using the same stochastic inputs for simulations on two levels of refinement
- it is particularly effective for applications with strong nonlinearity and high stochastic dimensionality, where the alternative methods are not appropriate

For my multilevel papers see:
`people.maths.ox.ac.uk/gilesm/mlmc.html`
`people.maths.ox.ac.uk/gilesm/acta/`

References

M.B. Giles, "Multilevel Monte Carlo path simulation", *Operations Research*, 56(3):607-617, 2008.

K.A. Cliffe, M.B. Giles, R. Scheichl, A.L. Teckentrup, "Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients", *Computing and Visualization in Science*, 14(1):3-15, 2011

M.B. Giles, C. Reisinger. "Stochastic finite differences and multilevel Monte Carlo for a class of SPDEs in finance", *SIAM Journal of Financial Mathematics*, 3(1):572-592, 2012.

F. Vidal-Codina, N.-C. Nguyen, M.B. Giles, J. Peraire. "A model and variance reduction method for computing statistical outputs of stochastic elliptic partial differential equations". *Journal of Computational Physics*, 2015.

M.B. Giles. 'Multilevel Monte Carlo methods', *Acta Numerica 2015*.