Multilevel Monte Carlo for PDEs with random coefficients

Aretha Teckentrup

University of Bath

Joint work with Robert Scheichl (University of Bath),
Andrew Cliffe (Nottingham) and Mike Giles (Oxford)

May 25, 2010

Model problem

We consider the elliptic partial differential equation (PDE)

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

with **random coefficient** $k(x,\omega)$ and random data $f(x,\omega)$.

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

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

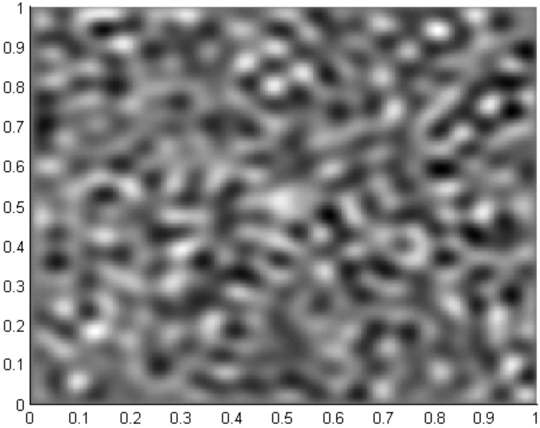


Figure: A typical realisation for $D = [0, 1]^2, \lambda = 0.001, \sigma^2 = 1$.

Discretisation

We employ a **Monte Carlo type method**, accelerated by a novel variance reduction technique.

- **Discretisation in space:** Cell-centred Finite Volume discretisation on a uniform grid $\mathcal{T}_h$. For rough coefficients we need to make h very small.
- **Sampling from the random coefficient:** Currently based on truncated Karhunen-Lòeve expansion, evaluated at cell centres

  But the method of sampling is not essential to the algorithm.

Standard Monte Carlo

Suppose we are interested in the expected value of an output functional $Q = \mathcal{G}(p)$. The standard **Monte Carlo estimator** for this is

$$\mathbb{E}\left[Q\right] \;\approx\; \hat{Q}_h^{\mathrm{MC}} := \frac{1}{N}\sum_{i=1}^{N} Q_h^{(i)},$$

where $Q_h^{(i)}$ is the ith sample of Q approximated on grid $\mathcal{T}_h$.

The **mean square error** can be shown to equal

$$\mathbb{E}\Big[\big(\hat{Q}_h^{\mathrm{MC}} - \mathbb{E}[Q]\big)^2\Big] = \underbrace{\mathbb{V}[Q_h^{\hat{\mathrm{MC}}}]}_{\text{Variance of MC estimator}} + \underbrace{\big(\mathbb{E}[\hat{Q}_h^{\mathrm{MC}}] - \mathbb{E}[Q]\big)^2}_{\text{(spatial) discretisation error}}$$

$$= \quad N^{-1}\mathbb{V}[Q_h] \quad + \quad \big(\mathbb{E}[\hat{Q}_h^{\mathrm{MC}}] - \mathbb{E}[Q]\big)^2$$

.

Multilevel Monte Carlo

The key ingredient in the multilevel method is **linearity of expectation**, which gives us

$$\mathbb{E}\left[Q_h\right] = \mathbb{E}\left[Q_h - Q_{2h}\right] + \mathbb{E}\left[Q_{2h}\right].$$

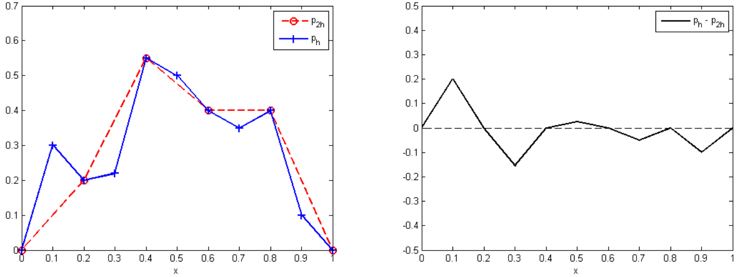


Figure: Example solutions on $\mathcal{T}_h$ and $\mathcal{T}_{2h}$ (left), and the difference between the two (right).

We can use this idea to define a **2 level MC estimator**

$$\hat{Q}_h^{\mathrm{TL}} := \hat{Q}_{2h}^{\mathrm{MC}} \; + \; \left(\widehat{Q_h - Q_{2h}}\right)^{MC}$$

Consider approximations of the PDE on a **sequence of levels**, s.t. $h_\ell = 2^{-\ell} h_0$, $\ell = 0, 1, \ldots, L$, and set $Q_\ell = Q_{h_\ell}$

We can define a **multilevel MC estimator**

$$\hat{Q}_L^{\mathrm{ML}} := \hat{Q}_0^{\mathrm{MC}} \; + \; \sum_{\ell=1}^{L} \left(\widehat{Q_\ell - Q_{\ell-1}}\right)^{\mathrm{MC}}$$

The **mean square error** of the multilevel estimator is

$$\mathbb{E}\Big[\big(\hat{Q}_L^{\mathrm{ML}} - \mathbb{E}[Q]\big)^2\Big] = \underbrace{\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}]}_{\text{Variance of ML estimator}} + \underbrace{\big(\mathbb{E}[\hat{Q}_L^{\mathrm{ML}}] - \mathbb{E}[Q]\big)^2}_{\text{(spatial) discretisation error}}$$

$$= \mathbb{V}[Q_0]N_0^{-1} + \sum_{\ell=1}^{L} \mathbb{V}[Q_\ell - Q_{\ell-1}]N_\ell^{-1} + \big(\mathbb{E}[\hat{Q}_L^{\mathrm{ML}}] - \mathbb{E}[Q]\big)^2$$

- N_0 still needs to be large, **but** samples are much cheaper to obtain on coarser grid
- N_ℓ $(\ell > 0)$ much smaller, **since** $\mathbb{V}[Q_\ell - Q_{\ell-1}] \to 0$ as $h_\ell \to 0$

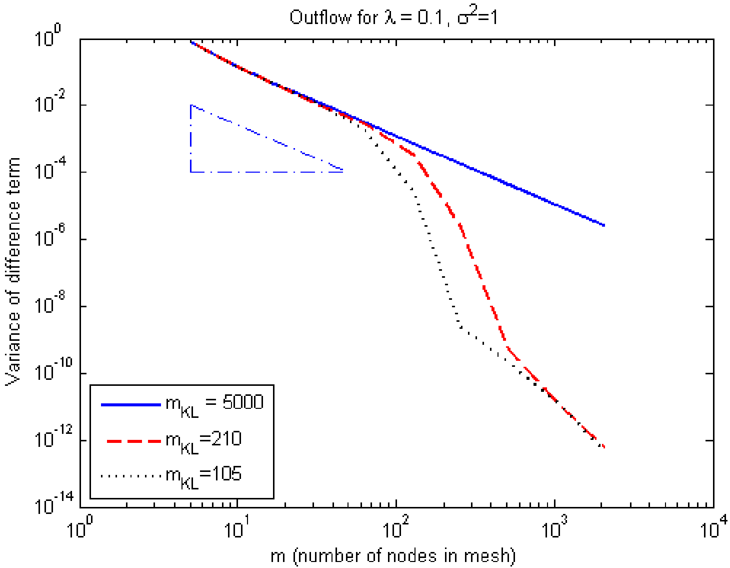


Figure: The variance of the difference terms as a function of h.

Theorem (Multilevel Monte Carlo)

If there exist $\alpha, \beta, \gamma > 0$ such that $\alpha \geq \frac{1}{2}\min(\beta,\gamma)$ and

(A1) $\left|\mathbb{E}[\hat{Q}_\ell^{\mathrm{MC}} - Q]\right| = \mathcal{O}(2^{-\alpha\ell})$

(A2) $\mathbb{V}\left[\left(Q_\ell \widehat{- Q_{\ell-1}}\right)^{\mathrm{MC}}\right] = \mathcal{O}(N_\ell^{-1} 2^{-\beta\ell})$

(A3) $C_\ell = \mathcal{O}(2^{\gamma\ell})$ (cost of one sample on level ℓ)

then for any $\varepsilon < 1$ there exist L and $\{N_\ell\}$ such that

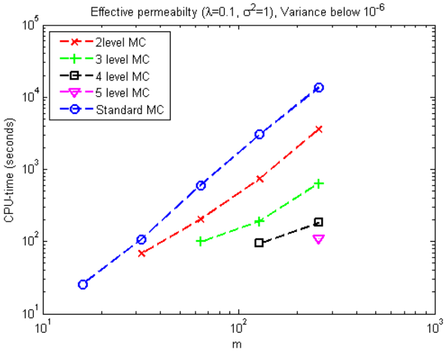
$$\mathbb{E}\left[\left(\hat{Q}_L^{\mathrm{ML}} - \mathbb{E}[Q]\right)^2\right] = \mathcal{O}(\varepsilon^2)$$

and the **total computational cost** is

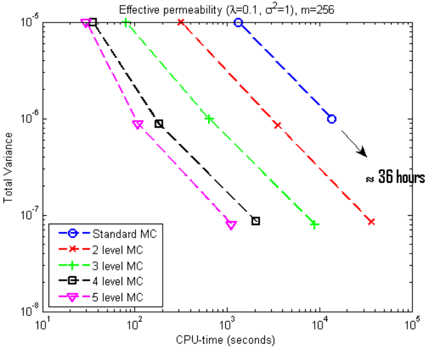
$$C^{\mathrm{ML}} = \begin{cases} \mathcal{O}(\varepsilon^{-2}), & \text{if } \beta > \gamma, \\ \mathcal{O}(\varepsilon^{-2}(\log\varepsilon)^2), & \text{if } \beta = \gamma, \\ \mathcal{O}(\varepsilon^{-2-(\gamma-\beta)/\alpha}), & \text{if } \beta < \gamma. \end{cases}$$

Performance of the Multilevel Method

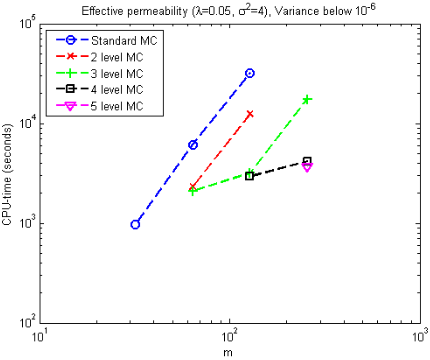
for $Q := k_{\mathrm{eff},1}$ (effective permeability), $\lambda = 0.1$, $\sigma^2 = 1$ and 500 KL–modes



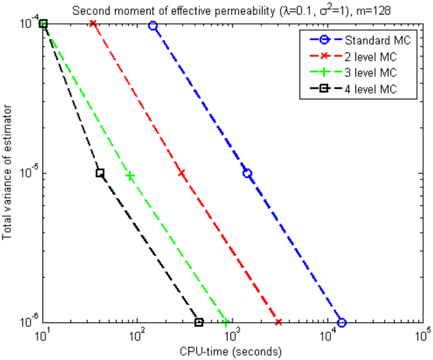
CPU-time versus $m := h_L^{-1}$ to attain $\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}] < 10^{-6}$.



$\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}]$ versus CPU-time for fixed $m = 256$.



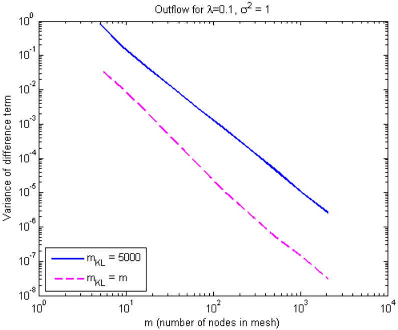
CPU-time versus $m := h_L^{-1}$
to attain $\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}] < 10^{-6}$
for a **rougher problem**.



$\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}]$ versus CPU-time
for the **second moment** of $k_{\mathrm{eff},1}$
for fixed $m = 128$.

Further work.

- Confirm predicted rate for computational cost w.r.t. ε.
- Theoretically bound $\mathbb{E}[\hat{Q}_\ell^{\mathrm{MC}} - Q]$ and $\mathbb{V}\Big[\Big(\widehat{Q_\ell - Q_{\ell-1}}\Big)^{\mathrm{MC}}\Big]$.
- Make further approximations of k on coarser levels, e.g. drop KL–modes:



- Combine with Quasi–MC sampling (complementary gains!)