MULTILEVEL MONTE CARLO FOR PDES WITH RANDOM COEFFICIENTS

Aretha Teckentrup and Robert Scheichl
University of Bath
(joint work with Andrew Cliffe (Nottingham) and Mike Giles (Oxford))

Motivation: Uncertainty in Groundwater Flow

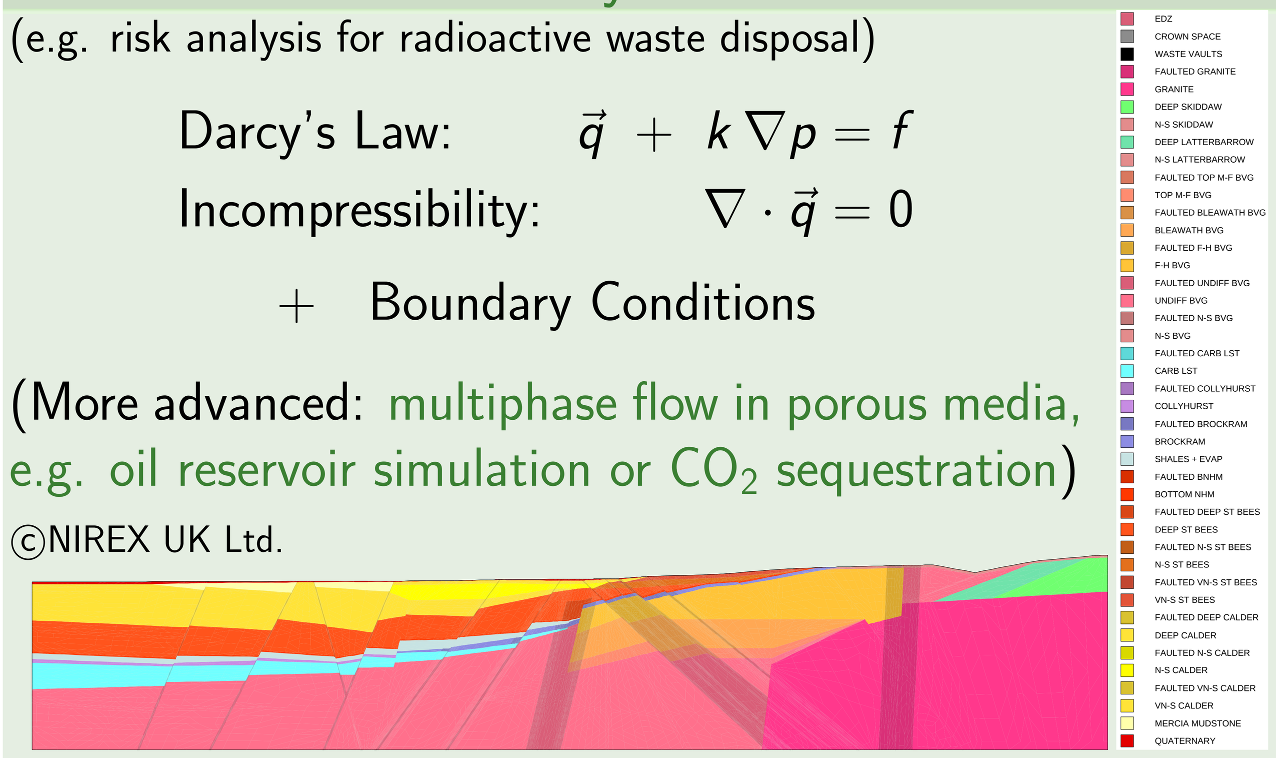
(e.g. risk analysis for radioactive waste disposal)

Darcy's Law: $\vec{q} \; + \; k\nabla p = f$

Incompressibility: $\nabla \cdot \vec{q} = 0$

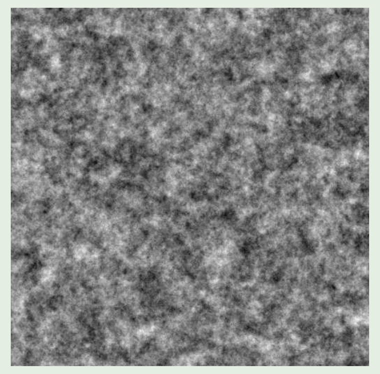
$+$ Boundary Conditions

(More advanced: multiphase flow in porous media, e.g. oil reservoir simulation or CO_2 sequestration)



Stochastic Modelling of Uncertainty (lognormal)

Typical Realisation (here with $n = 512^2$, $\lambda = 1/64$, $\sigma^2 = 8$)



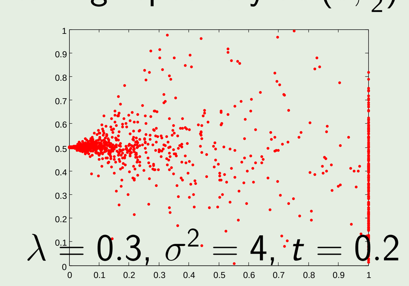
$\log k(x,\omega) \;=\;$ Gaussian with mean 0 & covariance function

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

(Data suggests this is reasonable representation of reality [Gelhar, 1975], [Hoeksema et al, 1985])

Quantities of Interest

Leaking repository at $(0,\frac{1}{2})$.



- flow field $\vec{q}$
- particle position at time t
- breakthrough time (to right)
- effective permeability: $k_{\mathrm{eff},1} := \int_D q_1\, d\vec{x} \big/ \int_D p_{x_1}\, d\vec{x}$

Key Computational Challenges

- Discretisation w.r.t. ω – "Curse of Dimensionality":
  - stochastic Galerkin (+ sparse versions)
  - stochastic collocation (+ anisotropic versions)
  - Monte Carlo (+ variants), etc... ← used here!
- Sample from random field k efficiently (& accurately):
  - truncated Karhunen-Lòeve expansion ← used here!
  - Matrix factorisation, e.g. circulant embedding, etc...
- Solve huge number of deterministic PDEs (with rough coefficients and small mesh size) Here: FVM + Sparse LU

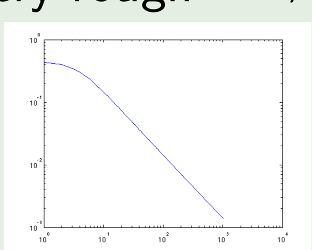
Model problem

Elliptic PDE in bounded domain $D \subset \mathbb{R}^d$, $d = 1,2,3$

$$-\nabla.\left(k(\mathbf{x},\omega)\nabla p(\mathbf{x},\omega)\right) = f(\mathbf{x},\omega), \qquad \mathbf{x} \in D, \qquad (1)$$

where the random coefficient $k(x,\omega)$ is very rough $\Longrightarrow$

- large number of KL–modes $\gg 100$!!
- small (spatial) mesh width $h \ll 1$!!



Standard Monte Carlo

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

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

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

The **mean square error** of this estimator is given by

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

It is well known that $\mathbb{V}[\hat{Q}_h^{\mathrm{MC}}] = \mathbb{V}[Q_h]/N$ (independent of dimension, i.e. independent of the number of KL modes).

Multilevel Monte Carlo (MLMC)

Consider approximations of (1) on a **sequence of levels**, s.t. $h_\ell = 2^{-\ell}h_0$, $\ell = 0,1,\ldots,L$, and set $Q_\ell = Q_{h_\ell}$. The key idea is now to use the following **telescoping sum**

$$\mathbb{E}[Q_L] = \mathbb{E}[Q_0] \; + \; \sum_{\ell=1}^{L}\mathbb{E}[Q_\ell - Q_{\ell-1}].$$

Let $Y_\ell := Q_\ell - Q_{\ell-1}$, $\ell > 0$, and $Y_0 := Q_0$. Then we can define the following **multilevel MC** estimator for $\mathbb{E}[Q]$:

$$\hat{Q}_L^{\mathrm{ML}} := \hat{Q}_0^{\mathrm{MC}} \; + \; \sum_{\ell=1}^{L}\hat{Y}_\ell^{\mathrm{MC}}$$

using the same sample $k(x,\omega^{(i)})$ for $Q_\ell^{(i)}$ **and** $Q_{\ell-1}^{(i)}$ in $Y_\ell^{(i)}$ and N_ℓ samples on level ℓ. For the **variance** we have

$$\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}] := \mathbb{V}[Q_0]N_0^{-1} \; + \; \sum_{\ell=1}^{L}\mathbb{V}[Y_\ell]N_\ell^{-1}$$

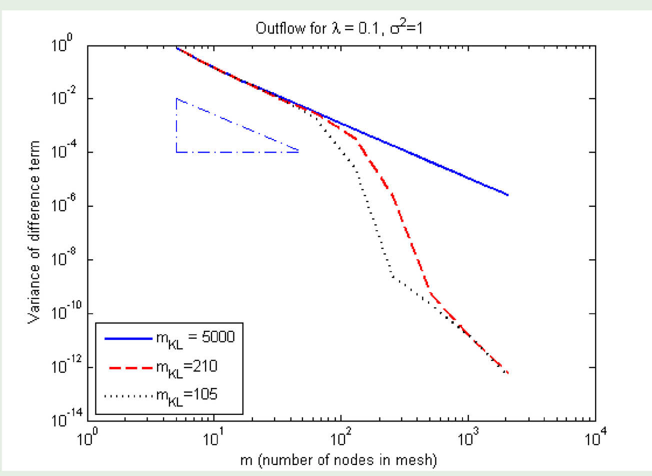
Minimising this over all choices of $\{N_\ell\}$ we get

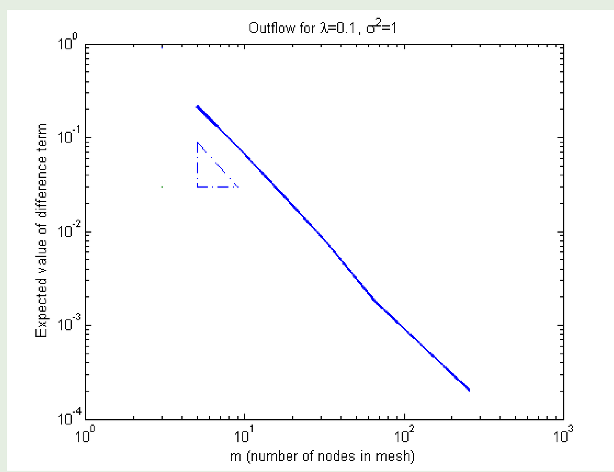
$$N_\ell \;\sim\; \sqrt{\mathbb{V}[Y_\ell]/C_\ell}$$

where C_ℓ is the computational cost to get one sample $Y_\ell^{(i)}$.

Decay of $\mathbb{V}[Y_\ell]$ and $\mathbb{E}[Y_\ell]$ (1D Numerics)

The **key observation** is now that $\mathbb{V}[Y_\ell]$ decays with $\ell \to \infty$ (or equivalently with $h_\ell \to 0$):





$\mathbb{V}[Y_\ell]$ (left, with 105, 210, 5000 KL–modes) and $\mathbb{E}[Y_\ell]$ (right, 5000 KL–modes) for the functional $Q = k_{\mathrm{eff},1}$ in 1D with $f = 0$, $\lambda = 0.1$, $\sigma^2 = 1$, $p(0) = 1$, $p(1) = 0$ ($m = h_\ell^{-1}$).

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) $\mathbb{V}[\hat{Y}_\ell^{\mathrm{MC}}] = \mathcal{O}(N_\ell^{-1}2^{-\beta\ell})$

(A3) $C_\ell = \mathcal{O}(2^{\gamma\ell})$

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 C^{ML} satisfies

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

Note for comparison that the cost of standard Monte Carlo to achieve the same mean square error of $\mathcal{O}(\varepsilon^2)$ is $C^{\mathrm{MC}} = \mathcal{O}(\varepsilon^{-2-\gamma/\alpha})$.

Application of MLMC Theorem

From **plots on left** we see that **in 1D** for $Q = k_{\mathrm{eff},1}$ we observe (numerically) $\alpha \approx 1.5$ and $\beta \approx 2$. In 1D $\gamma = 1$.

The behaviour of Y_ℓ in **2D is similar** (in our experiments) and with optimal linear solver (e.g. AMG) $\gamma \approx 2$.

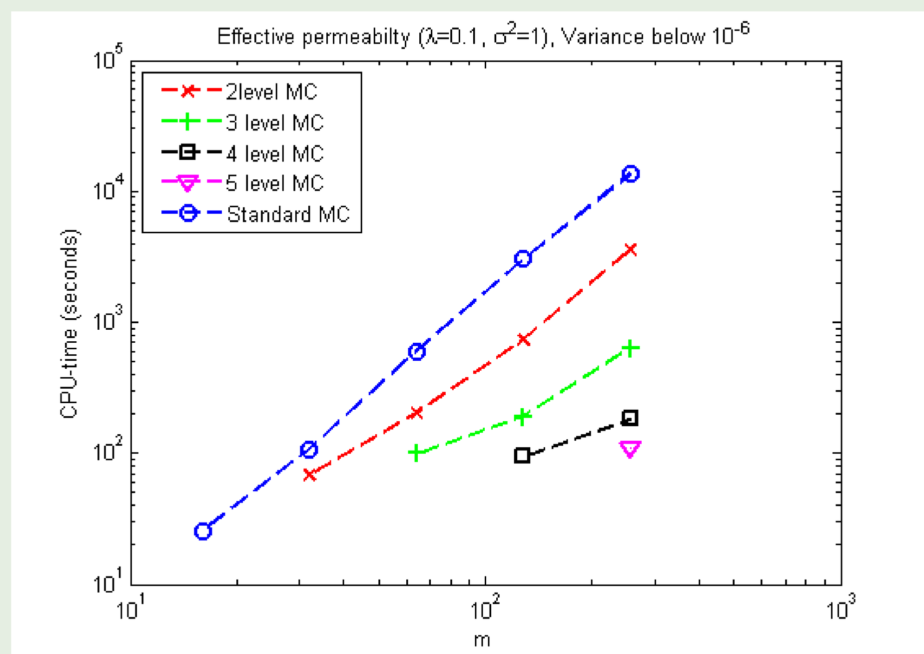
Hence we expect the following **relative costs** to achieve a root mean square error (RMSE) of ε ("extrapolating" to 3D):

dim	C^{MC}	C^{ML}
1	$\varepsilon^{-8/3}$	ε^{-2}
2	$\varepsilon^{-10/3}$	$\varepsilon^{-2}\log(\varepsilon)^2$
3	ε^{-4}	$\varepsilon^{-8/3}$

Improvement even bigger for quantities where discretisation error is bigger: in 2D if $\alpha = 1$ and $\beta = \gamma = 2$ (as before) then $C^{\mathrm{MC}} = \mathcal{O}(\varepsilon^{-4})$ while $C^{\mathrm{ML}} = \mathcal{O}(\varepsilon^{-2}\log(\varepsilon)^2)$.

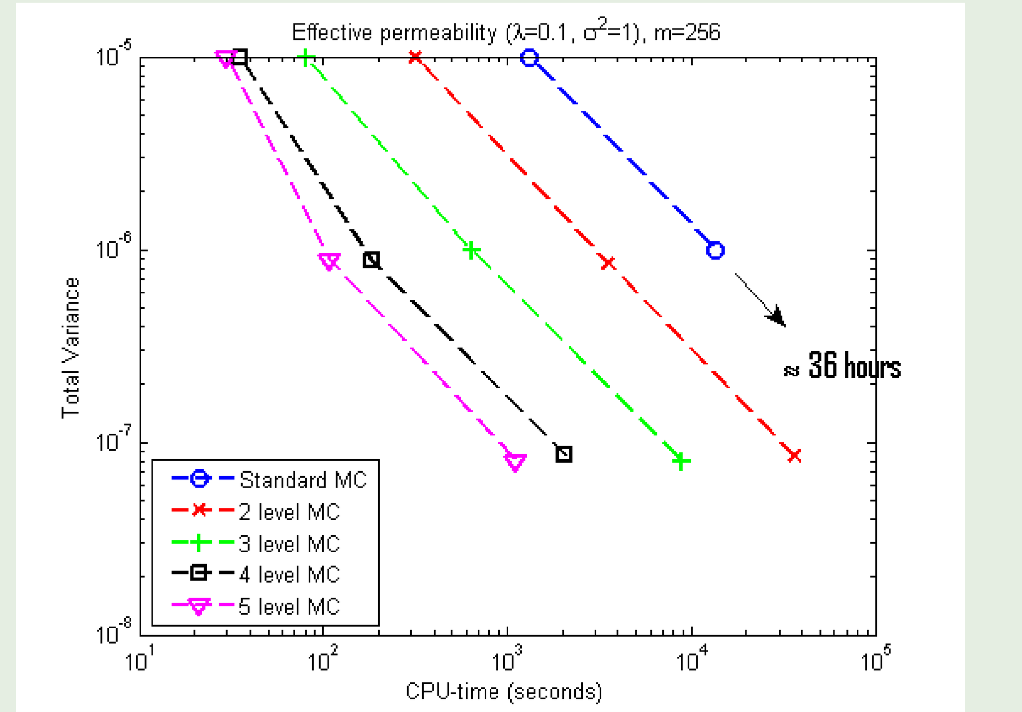
(If $Q_\ell \to Q$ linearly with h_ℓ (i.e. $\alpha = 1$) then solving for 1 sample $Q_\ell^{(i)}$ costs $\mathcal{O}(\varepsilon^{-2})$.)

Preliminary Numerical Results in 2D



CPU-time to achieve $\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}] < 10^{-6}$ versus $m = h_L^{-1}$ for $Q = k_{\mathrm{eff},1}$, $\lambda = 0.1$, $\sigma^2 = 1$, and 500 KL-modes.

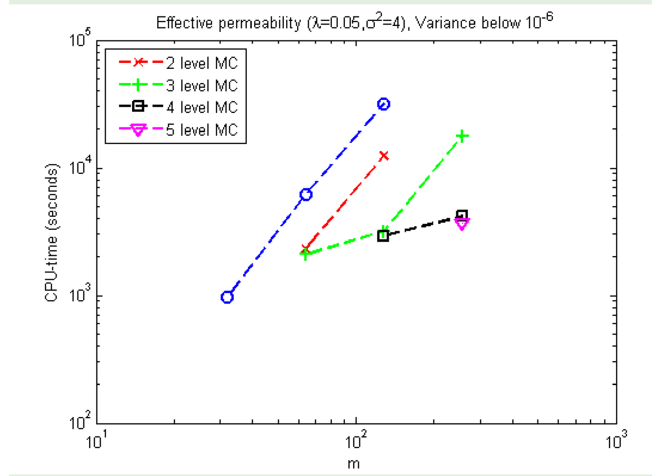
Preliminary Numerical Results in 2D



$\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}]$ versus CPU–time for fixed mesh size $m = h_L^{-1} = 256$, for $Q = k_{\mathrm{eff},1}$, $\lambda = 0.1$, $\sigma^2 = 1$ and 500 KL-modes.

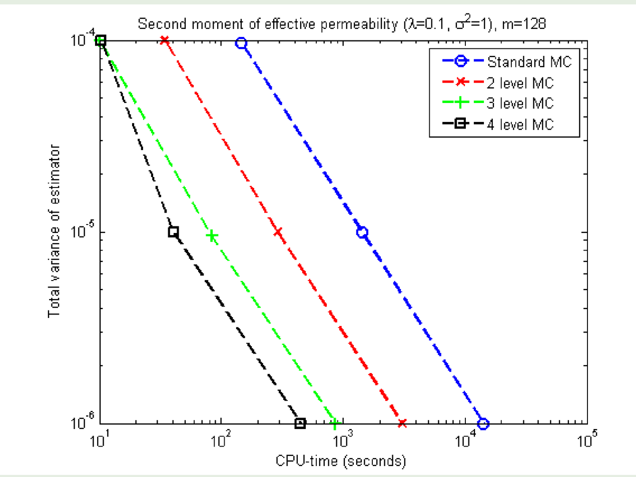
Further 2D Results

$\mathbb{E}[k_{\mathrm{eff},1}]$



CPU-time to get $\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}] < 10^{-6}$ vs. m for $\lambda = 0.05$, $\sigma^2 = 4$, and 500 KL-modes

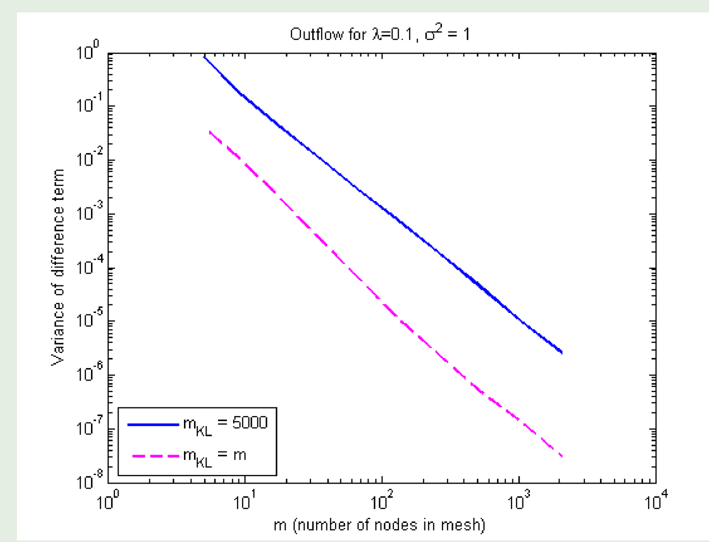
$\mathbb{E}[(k_{\mathrm{eff},1})^2]$



$\mathbb{V}[\hat{Q}_L^{\mathrm{ML}}]$ versus CPU–time for $m = 256$, $\lambda = 0.1$, $\sigma^2 = 4$ and 500 KL-modes.

Further work

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



$\mathbb{V}[Y_\ell]$ for $Q = k_{\mathrm{eff},1}$ in 1D with $f = 0$, $\lambda = 0.1$, $\sigma^2 = 1$.
Dashed line: 5000 KL-modes; solid line: m KL-modes.

- Circulant embedding instead of truncated KL-series.
- Combine with Quasi–MC sampling (gains complement!)
- Combine with other variance reduction techniques, such as antithetic sampling.
- Change measure/importance sampling for "rare events".