Stochastic Simulation: Lecture 12

Prof. Mike Giles

Oxford University Mathematical Institute

PDEs with Uncertainty

Motivated by the modelling of oil reservoirs and groundwater contamination, there has been a lot of MLMC/MLQMC research on an elliptic SPDE coming from Darcy's law:

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

where the permeability $\kappa(x)$ is uncertain due to lack of knowledge.

A log-Normal model uses

$$\log \kappa(x) = u_0(x) + u(x)$$

where $u_0(x)$ is the mean, and $u(x)$ is Normally distributed pointwise, with zero mean and spatial covariance

$$\mathbb{E}[u(x_1) \, u(x_2)] = K(x_1, x_2).$$

Karhunen-Loève expansion

The linear operator

$$u(x) \longrightarrow \int_D K(x,y)\, u(y)\, \mathrm{d}y$$

has a complete set of orthonormal eigenfunctions $e_j(x)$, and associated eigenvalues $\lambda_j > 0$ such that

$$\int_D K(x,y)\, e_j(y)\, \mathrm{d}y = \lambda_j\, e_j(x)$$

Hence, $u(x)$ can be expressed as

$$u(x) = \sum_j Z_j\, e_j(x)$$

where

$$Z_j = \int_D e_j(x)\, u(x)\, \mathrm{d}x$$

Karhunen-Loève expansion

It follows that Z_j is Normally distributed with zero mean, and

$$\begin{aligned}
\mathbb{E}[Z_j\, Z_k] &= \int_D \int_D e_j(x)\; \mathbb{E}[u(x)\, u(y)]\, e_k(y)\; \mathrm{d}x\, \mathrm{d}y \\
&= \int_D \int_D e_j(x)\; K(x,y)\, e_k(y)\; \mathrm{d}x\, \mathrm{d}y \\
&= \lambda_k \int_D e_j(x)\, e_k(x) \mathrm{d}x \\
&= \lambda_k \delta_{j,k}
\end{aligned}$$

Thus the Z_j are independent, and have variance λ_j.

Re-scaling, we can make the Z_j have unit variance and define $u(x)$ as

$$u(x) = \sum_j \sqrt{\lambda_j}\, Z_j\, e_j(x)$$

which is the Karhunen-Loève expansion.

Karhunen-Loève expansion

Note that this is simply the continuous generalisation of the PCA factorisation in which we ended up with

$$u = U \Lambda^{1/2} Z \equiv \sum_j \sqrt{\lambda_j}\, Z_j\, U_j$$

which U_j, λ_j being the eigenvectors and eigenvalues of the covariance matrix Σ.

They both have a very similar computational drawback: if there are M grid points in the domain D, then PCA has M eigenmodes so the cost per sample is $O(M^2)$, in addition to an $O(M^3)$ setup cost and an $O(M^2)$ memory requirement.

The K-L expansion is even worse as there are an infinite number of eigenmodes, but in both cases the cost can be reduced by truncating the expansion – keeping only the leading terms. Even so, it can still be the dominant cost in a calculation, much more than the solution of the elliptic PDE!

Elliptic SPDE

Often the covariance function is assumed to be stationary – i.e. a function of the separation $x_1 - x_2$.

"Exponential" covariance:

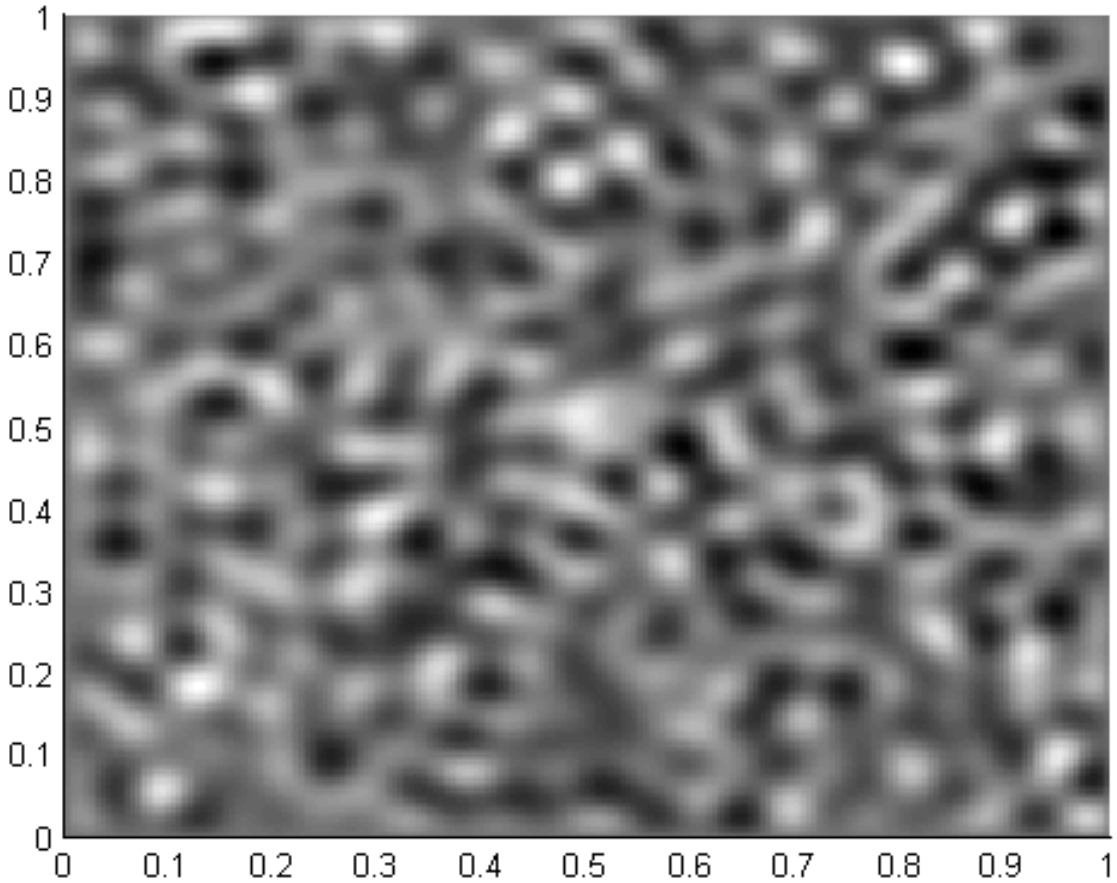
$$\mathrm{cov}(\log \kappa(x_1), \log \kappa(x_2)) = \sigma^2 \exp(-\|x_1 - x_2\|/\lambda)$$

"Gaussian" covariance:

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

Typically in real applications σ is large, and λ is small – both of these mean that PDE methods such as stochastic collocation are very expensive

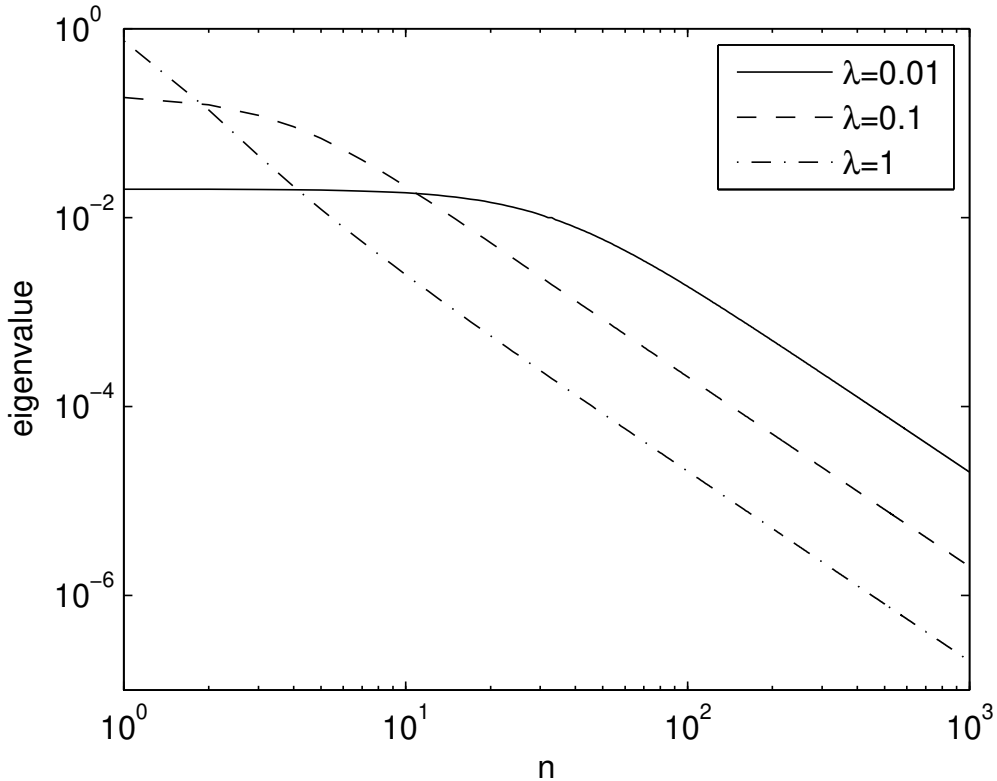
Elliptic SPDE



A typical realisation of κ for exponential covariance with $\lambda = 0.01,\ \sigma = 1$.

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

Some very old 2D experiments performed by Rob Scheichl and Aretha Teckentrup at the University of Bath:

- 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
- these experiments used a direct solver for simplicity, but later 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: $\lambda = 0.2$

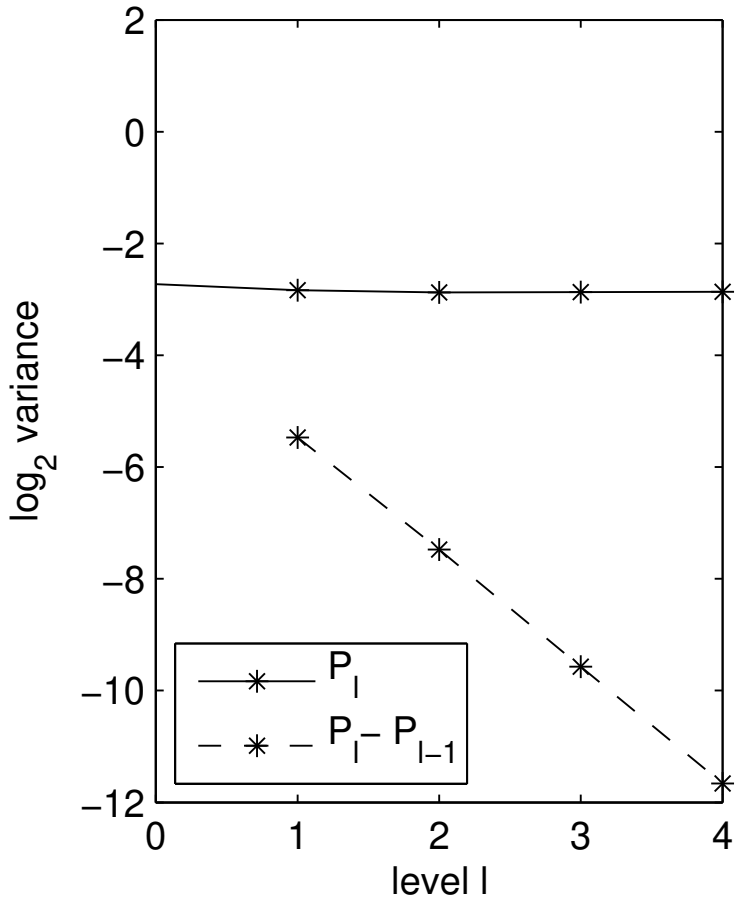
Coarsest grid: $h = 1/8$ (comparable to λ)

Finest grid: $h = 1/128$

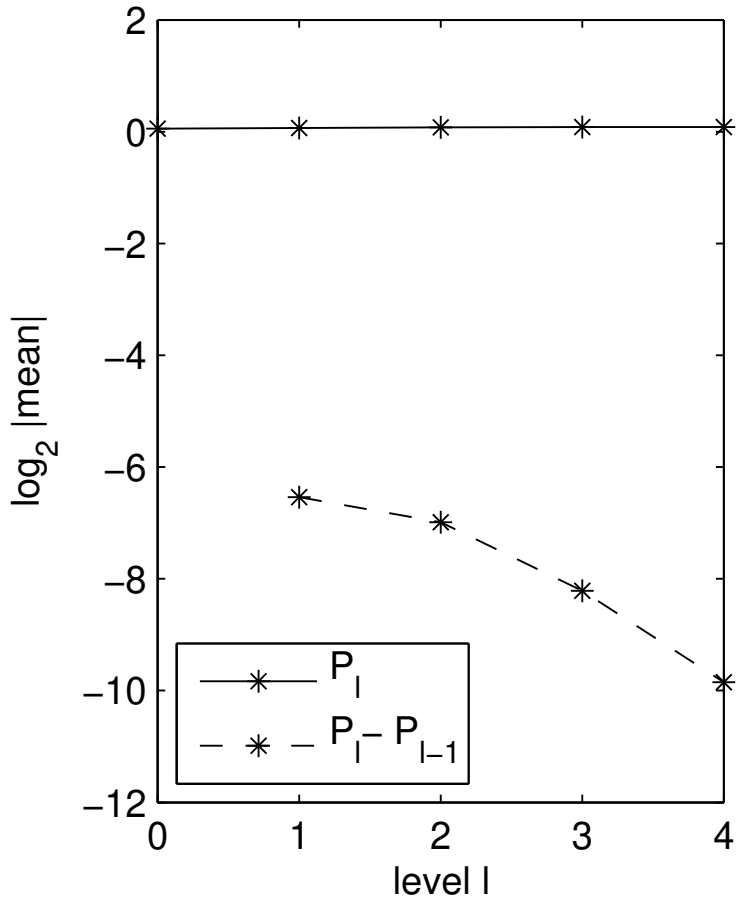
Karhunen-Loève truncation: used the leading 4000 modes

Cost taken to be proportional to number of nodes

2D Results

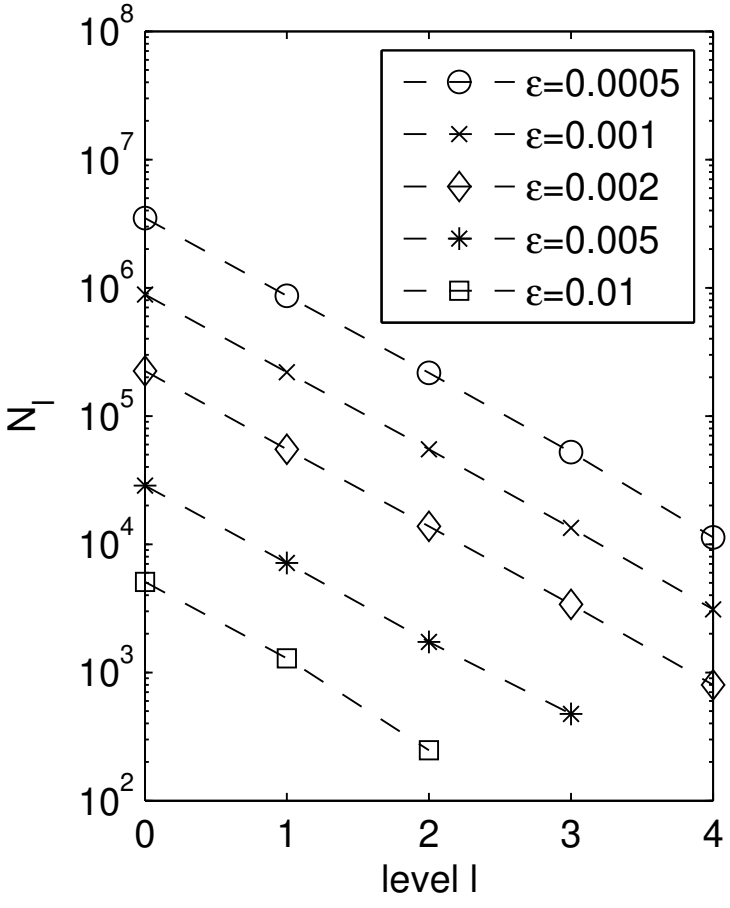


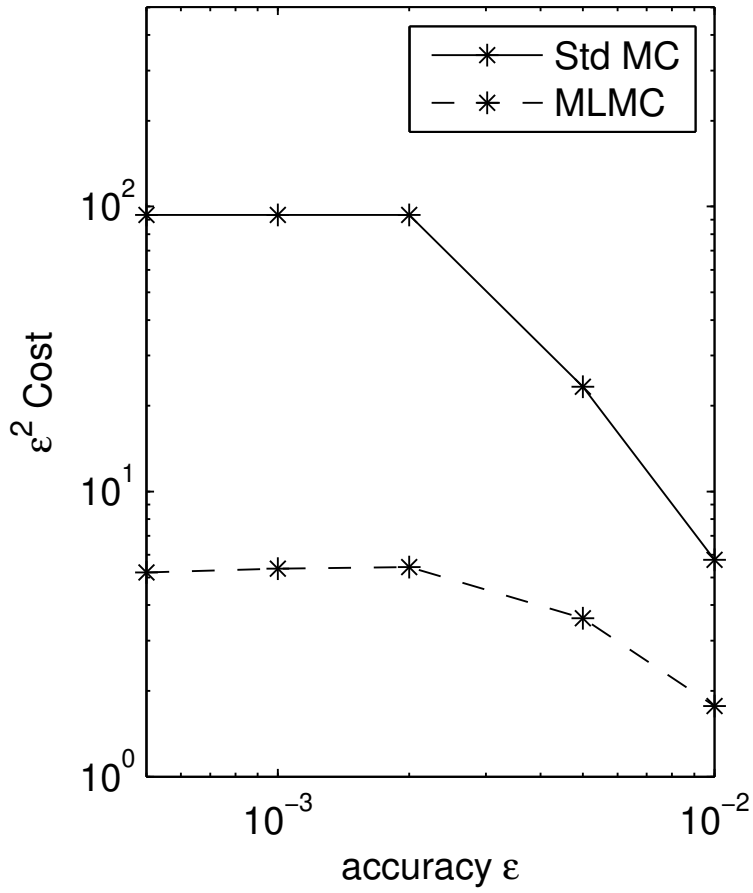
$$\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] \sim h_\ell^2$$



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

2D Results


10^8
10^7
10^6
10^5
10^4
10^3
10^2
N_l
0
1
2
3
4
level l
ε=0.0005
ε=0.001
ε=0.002
ε=0.005
ε=0.01


10^2
10^1
10^0
ε^2 Cost
10^{-3}
10^{-2}
accuracy ε
Std MC
MLMC

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	$\varepsilon^{-3.5}$	$\varepsilon^{-2.5}$

Alternative ways to generate Gaussian fields

a) Cholesky factorisation of covariance matrix

Same bad points as PCA factorisation: $O(M^3)$ setup cost, $O(M^2)$ memory requirement, $O(M^2)$ cost per sample

b) use of H-matrices (H stands for Hierarchical)

Promising new research direction:

`https://en.wikipedia.org/wiki/Hierarchical_matrix`

J. Dölz, H. Harbrecht, and C. Schwab, Covariance regularity and H-matrix approximation for rough random fields, Numerische Mathematik, 135:1045-1071, 2017.

Alternative ways to generate Gaussian fields

c) circulant embedding

In 1D, for a uniform grid and a stationary covariance function, the covariance matrix looks like:

$$\begin{pmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 \\ a_1 & a_0 & a_1 & a_2 & a_3 & a_4 \\ a_2 & a_1 & a_0 & a_1 & a_2 & a_3 \\ a_3 & a_2 & a_1 & a_0 & a_1 & a_2 \\ a_4 & a_3 & a_2 & a_1 & a_0 & a_1 \\ a_5 & a_4 & a_3 & a_2 & a_1 & a_0 \end{pmatrix}$$

which can be inserted into a larger circulant matrix . . .

Alternative ways to generate Gaussian fields

$$A = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & a_4 & a_3 & a_2 & a_1 \\ a_1 & a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & a_4 & a_3 & a_2 \\ a_2 & a_1 & a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & a_4 & a_3 \\ a_3 & a_2 & a_1 & a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & a_4 \\ a_4 & a_3 & a_2 & a_1 & a_0 & a_1 & a_2 & a_3 & a_4 & a_5 \\ a_5 & a_4 & a_3 & a_2 & a_1 & a_0 & a_1 & a_2 & a_3 & a_4 \\ a_4 & a_5 & a_4 & a_3 & a_2 & a_1 & a_0 & a_1 & a_2 & a_3 \\ a_3 & a_4 & a_5 & a_4 & a_3 & a_2 & a_1 & a_0 & a_1 & a_2 \\ a_2 & a_3 & a_4 & a_5 & a_4 & a_3 & a_2 & a_1 & a_0 & a_1 \\ a_1 & a_2 & a_3 & a_4 & a_5 & a_4 & a_3 & a_2 & a_1 & a_0 \end{pmatrix}$$

in which each row (and column) is the same as the previous, but rotated by 1 position

Hence the eigenvectors are Fourier modes, and the corresponding eigenvalues come from a Fourier transform of the first row.

Alternative ways to generate Gaussian fields

This give

$$A = F \Lambda F^T = L L^T$$

where

$$L = F \Lambda^{1/2}$$

and therefore

$$L Z = \sum_j \sqrt{\lambda_j} Z_j F_j$$

which can be computed in $O(M \log M)$ cost.

There are some minor technical difficulties (making sure the embedded matrix has positive eigenvalues), but it extends to 2D and 3D and works well when doing MLMC on regular grids.

Alternative ways to generate Gaussian fields

d) spatial white noise

If $\dot{W}$ is spatial white noise. then the solution of

$$(I - \kappa^{-2}\nabla^2)^k u = \dot{W}$$

is a Matérn field with covariance of the form

$$K(x, y) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\kappa r)^\nu \mathcal{K}_\nu(\kappa r), \quad r = \|x - y\|_2,$$

where $\nu = 2k - d/2$, and $\mathcal{K}_\nu$ is a Bessel function of the second kind – the Matérn class includes the exponential and Gaussian cases mentioned previously.

Alternative ways to generate Gaussian fields

What is white noise $\dot{W}$?

It is a generalised stochastic field defined by its effect on L_2-integrable test functions ϕ_j so that

$$\langle \dot{W}, \phi_j \rangle \sim N\left(0, \int_D \phi_j^2 \, \mathrm{d}x\right)$$

and

$$\mathbb{E}\left[\langle \dot{W}, \phi_j \rangle \, \langle \dot{W}, \phi_k \rangle\right] = \int_D \phi_j \, \phi_k \, \mathrm{d}x$$

If a domain D is split up into a number of disjoint pieces D_1, D_2, D_3, ..., then $\dot{W}$ can be decomposed into the sum of its restrictions onto each of those pieces, and the effects of each are independent.

Hence, can independently simulate the effect of each, and then sum them up.

Alternative ways to generate Gaussian fields

In a MLMC setting, working with coarse and fine grids composed of triangles (2D) or tetrahedra (3D), can create a finer supermesh of triangles/tetrahedra such that each new cell has a non-zero intersection with one and only one coarse and fine cell.

If $\dot{W}\Big|_{\Delta}$ is the restriction of $\dot{W}$ to this cell, then we can create a small covariance matrix for the test functions which are non-zero on this cell:

$$A_{\Delta} \equiv \mathbb{E}\left[\langle \dot{W}\Big|_{\Delta} \phi_j \rangle \, \langle \dot{W}\Big|_{\Delta} \phi_k \rangle\right] = \int_{\Delta} \phi_j \, \phi_k \, \mathrm{d}x$$

A_{Δ} is small, so can use Cholesky factorisation to generate required samples of $\langle \dot{W}\Big|_{\Delta} \phi_j \rangle$ on both grids.

Alternative ways to generate Gaussian fields

Additional complication: for some values of the Matérn parameter ν, need to solve

$$(I - \kappa^{-2}\nabla^2)^k u = \dot{W}$$

for non-integer values of k.

How do we do that? What does it even mean?

We rely on a complex contour integral representation of generalised matrix functions:

$$f(A) = \frac{1}{2\pi \mathrm{i}} \int_\Gamma f(z)\,(zI - A)^{-1}\,\mathrm{d}z$$

and approximate the integral.

Final comments

- PDEs with random inputs / boundary data have been well explored
- PDEs with random coefficients have also been well explored – growing body of literature on numerical analysis too
- Stochastic PDEs with white noise or Brownian noise inputs have received much less attention, so still more to be done?

  I'm not even sure how much is proven concerning wellposedness and numerical analysis

References

K.A. Cliffe, MBG, 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.

A. Barth, Ch. Schwab, N. Zollinger. "Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients". Numerische Mathematik, 119(1):123-161, 2011.

M. Croci, MBG, M.E. Rognes, P.E. Farrell. "Efficient white noise sampling and coupling for multilevel Monte Carlo with non-nested meshes". SIAM/ASA Journal on Uncertainty Quantification, 6(4):1454-1474, 2018.

N. Hale, N.J. Higham, L.N. Trefethen. "Computing $A^\alpha, \log(A)$, and related matrix functions by contour integrals". SIAM Journal on Numerical Analysis, 46(5):2505-2523, 2008.