University of Oxford

Mathematical Institute

InFoMM
Industrially Focused
Mathematical Modelling

Efficient white noise sampling and coupling for multilevel Monte Carlo

M. Croci (Oxford), M. B. Giles (Oxford), P. E. Farrell (Oxford), M. E. Rognes (Simula)

MCQMC2018 - July 4, 2018

EPSRC
Engineering and Physical Sciences
Research Council

simula

Overview

Overview

Motivation

The motivation of our research is the sampling of lognormal Gaussian fields. A Matérn Gaussian field (approximately) satisfies a linear elliptic SPDE of the form

$$Lu = \dot{W}, \quad x \in D, \quad \omega \in \Omega \quad + \ \text{BCs},$$

where $u = u(x, \omega)$ and $\dot{W}$ is **spatial white noise**. Other approaches can be used (with pros and cons), but we will not discuss them here.

Motivation

The motivation of our research is the sampling of lognormal Gaussian fields. A Matérn Gaussian field (approximately) satisfies a linear elliptic SPDE of the form

$$Lu = \dot{W}, \quad x \in D, \quad \omega \in \Omega \quad + \text{ BCs},$$

where $u = u(x, \omega)$ and $\dot{W}$ is **spatial white noise**. Other approaches can be used (with pros and cons), but we will not discuss them here.

The same techniques can be used to solve a more general class of SPDEs, e.g.

$$N(u) + Lu = \dot{W}, \quad x \in D, \quad \omega \in \Omega \quad + \text{ BCs}.$$

In this case solving means to compute $\mathbb{E}[P(u)]$ for some functional P of the solution.

Motivation

The motivation of our research is the sampling of lognormal Gaussian fields. A Matérn Gaussian field (approximately) satisfies a linear elliptic SPDE of the form

$$Lu = \dot{W}, \quad x \in D, \quad \omega \in \Omega \quad + \text{ BCs},$$

where $u = u(x, \omega)$ and $\dot{W}$ is **spatial white noise**. Other approaches can be used (with pros and cons), but we will not discuss them here.

The same techniques can be used to solve a more general class of SPDEs, e.g.

$$N(u) + Lu = \dot{W}, \quad x \in D, \quad \omega \in \Omega \quad + \text{ BCs}.$$

In this case solving means to compute $\mathbb{E}[P(u)]$ for some functional P of the solution.

Common applications: finance, geology, meteorology, **biology**...

Motivation

The motivation of our research is the sampling of lognormal Gaussian fields. A Matérn Gaussian field (approximately) satisfies a linear elliptic SPDE of the form

$$Lu = \dot{W}, \quad x \in D, \quad \omega \in \Omega \quad + \text{ BCs},$$

where $u = u(x, \omega)$ and $\dot{W}$ is **spatial white noise**. Other approaches can be used (with pros and cons), but we will not discuss them here.

The same techniques can be used to solve a more general class of SPDEs, e.g.

$$N(u) + Lu = \dot{W}, \quad x \in D, \quad \omega \in \Omega \quad + \text{ BCs}.$$

In this case solving means to compute $\mathbb{E}[P(u)]$ for some functional P of the solution.

Common applications: finance, geology, meteorology, **biology**...

Main issue: sampling $\dot{W}$ is hard!

Motivation

The motivation of our research is the sampling of lognormal Gaussian fields. A Matérn Gaussian field (approximately) satisfies a linear elliptic SPDE of the form

$$Lu = \dot{W}, \quad x \in D, \quad \omega \in \Omega \quad + \text{ BCs},$$

where $u = u(x, \omega)$ and $\dot{W}$ is **spatial white noise**. Other approaches can be used (with pros and cons), but we will not discuss them here.

The same techniques can be used to solve a more general class of SPDEs, e.g.

$$N(u) + Lu = \dot{W}, \quad x \in D, \quad \omega \in \Omega \quad + \text{ BCs}.$$

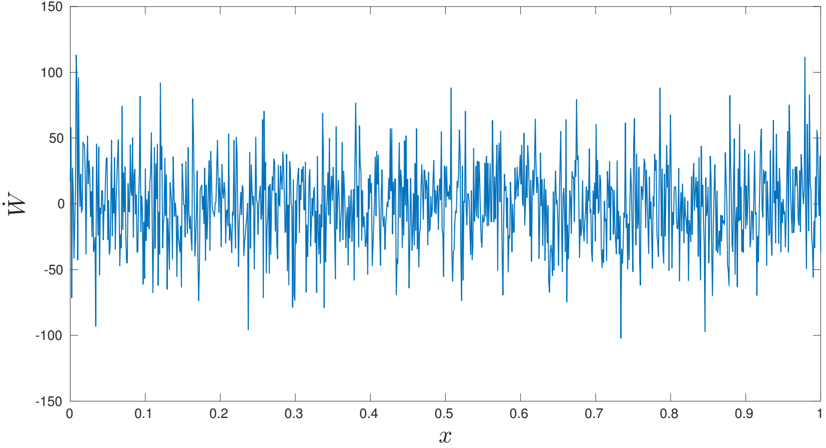
In this case solving means to compute $\mathbb{E}[P(u)]$ for some functional P of the solution.

Common applications: finance, geology, meteorology, **biology**...

Main issue: sampling $\dot{W}$ is hard!

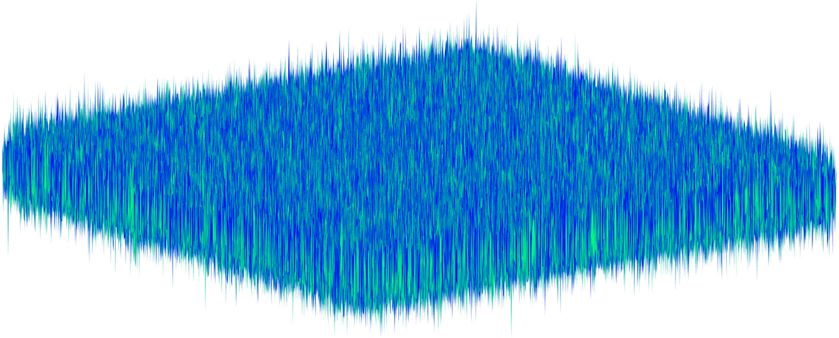
The efficient sampling of $\dot{W}$ is the focus of this talk.

White noise (1D)



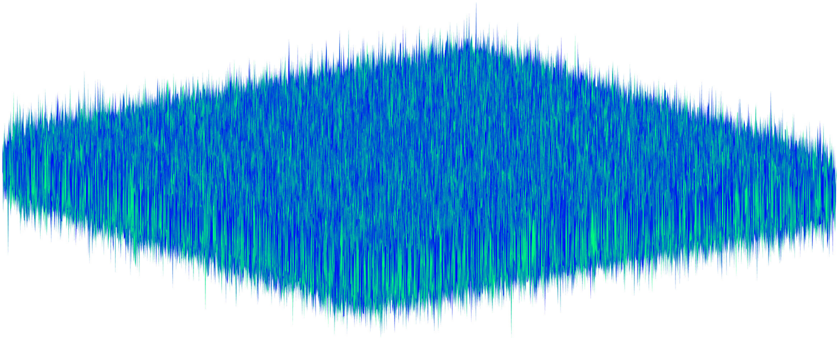
WARNING! Point evaluation not defined!

White noise (2D)



WARNING! Point evaluation not defined!

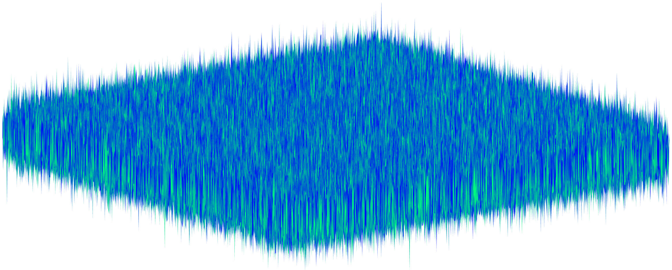
White noise (2D)



WARNING! Point evaluation not defined!

IDEA! Avoid point evaluation by integrating $\dot{W}$.

White Noise (practical definition)

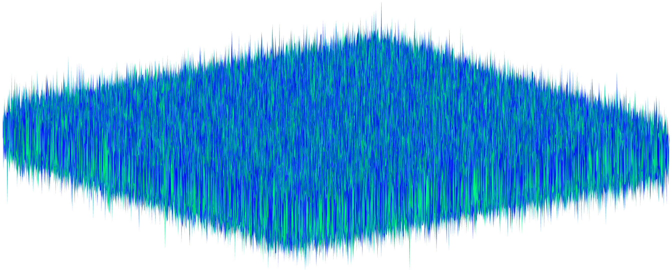


Definition (Spatial White Noise $\dot{W}$)

For any $\phi \in L^2(D)$, define $\langle \dot{W}, \phi \rangle := \int_D \dot{W}\phi \,\mathrm{d}x$. For any $\phi_i,\ \phi_j \in L^2(D)$, $b_i = \langle \dot{W}, \phi_i \rangle$, $b_j = \langle \dot{W}, \phi_j \rangle$ are zero-mean Gaussian random variables, with,

$$\mathbb{E}[b_i b_j] = \int_D \phi_i \phi_j \,\mathrm{d}x =: M_{ij}, \quad \mathbf{b} \sim \mathcal{N}(0, M). \tag{1}$$

White Noise (practical definition)



IMPORTANT NOTE: Generalised random fields of type I and of type II.

Finite element (FEM) framework

When solving SPDEs (see 1^{st} slide) with FEM, we get (for **linear problems**)

Discrete weak form: find $u_h \in V_h$ s.t. for all $v_h \in V_h$,

$$a(u_h, v_h) = \langle \dot{W}, v_h \rangle, \tag{2}$$

Where $V_h = \text{span}(\{\phi_i\}_{i=0}^{n})$, (e.g. with Lagrange elements).

Finite element (FEM) framework

When solving SPDEs (see 1^{st} slide) with FEM, we get (for **linear problems**)

Discrete weak form: find $u_h \in V_h$ s.t. for all $v_h \in V_h$,

$$a(u_h, v_h) = \langle \dot{W}, v_h \rangle, \tag{2}$$

Where $V_h = \text{span}(\{\phi_i\}_{i=0}^n)$, (e.g. with Lagrange elements).

FEM linear system: $u_h = \sum_i u_i \phi_i$, $\mathbf{u} = [u_0, \dots, u_n]^T$,

$$A\mathbf{u} = \mathbf{b}(\omega), \tag{3}$$

where the entries of **b** are given by,

$$\langle \dot{W}, \phi_i \rangle(\omega) = b_i(\omega), \tag{4}$$

with $\mathbf{b} \sim \mathcal{N}(0, M)$ as before. M is the **mass matrix** of V_h.

Multilevel Monte Carlo [Giles 2008] + FEM framework

For MLMC, we have two approximation levels ℓ and $\ell - 1$. For any particular $\omega \in \Omega$, we need to solve: find $u_h^\ell \in V_h^\ell$, $u_h^{\ell-1} \in V_h^{\ell-1}$ s.t. for all $v_h^\ell \in V_h^\ell$, $v_h^{\ell-1} \in V_h^{\ell-1}$,

$$a(u_h^\ell, v_h^\ell) = \langle \dot{W}, v_h^\ell \rangle(\omega), \tag{5}$$

$$a(u_h^{\ell-1}, v_h^{\ell-1}) = \langle \dot{W}, v_h^{\ell-1} \rangle(\omega). \tag{6}$$

Multilevel Monte Carlo [Giles 2008] + FEM framework

For MLMC, we have two approximation levels ℓ and $\ell-1$. For any particular $\omega \in \Omega$, we need to solve: find $u_h^\ell \in V_h^\ell$, $u_h^{\ell-1} \in V_h^{\ell-1}$ s.t. for all $v_h^\ell \in V_h^\ell$, $v_h^{\ell-1} \in V_h^{\ell-1}$,

$$a(u_h^\ell, v_h^\ell) = \langle \dot{W}, v_h^\ell \rangle(\omega), \tag{5}$$

$$a(u_h^{\ell-1}, v_h^{\ell-1}) = \langle \dot{W}, v_h^{\ell-1} \rangle(\omega). \tag{6}$$

This yields the linear system

$$\left[\begin{array}{c|c} A^\ell & 0 \\ \hline 0 & A^{\ell-1} \end{array}\right] \left[\begin{array}{c} \mathbf{u}^\ell \\ \hline \mathbf{u}^{\ell-1} \end{array}\right] = \left[\begin{array}{c} \mathbf{b}^\ell \\ \hline \mathbf{b}^{\ell-1} \end{array}\right] = \mathbf{b},$$

Multilevel Monte Carlo [Giles 2008] + FEM framework

For MLMC, we have two approximation levels ℓ and $\ell-1$. For any particular $\omega \in \Omega$, we need to solve: find $u_h^\ell \in V_h^\ell$, $u_h^{\ell-1} \in V_h^{\ell-1}$ s.t. for all $v_h^\ell \in V_h^\ell$, $v_h^{\ell-1} \in V_h^{\ell-1}$,

$$a(u_h^\ell, v_h^\ell) = \langle \dot{W}, v_h^\ell \rangle(\omega), \tag{5}$$

$$a(u_h^{\ell-1}, v_h^{\ell-1}) = \langle \dot{W}, v_h^{\ell-1} \rangle(\omega). \tag{6}$$

This yields the linear system

$$\left[\begin{array}{c|c} A^\ell & 0 \\ \hline 0 & A^{\ell-1} \end{array}\right] \left[\begin{array}{c} \mathbf{u}^\ell \\ \hline \mathbf{u}^{\ell-1} \end{array}\right] = \left[\begin{array}{c} \mathbf{b}^\ell \\ \hline \mathbf{b}^{\ell-1} \end{array}\right] = \mathbf{b},$$

where $\mathbf{b} \sim \mathcal{N}(0, M)$. Let $V_h^\ell = \mathrm{span}(\{\phi_i^\ell\}_{i=0}^{n_\ell})$ and $V_h^{\ell-1} = \mathrm{span}(\{\phi_i^{\ell-1}\}_{i=0}^{n_{\ell-1}})$, then

$$M = \left[\begin{array}{c|c} M^\ell & M^{\ell,\ell-1} \\ \hline (M^{\ell,\ell-1})^T & M^{\ell-1} \end{array}\right], \quad M_{ij}^{\ell,\ell-1} = \int \phi_i^\ell \phi_j^{\ell-1}\, \mathrm{d}x.$$

Multilevel Monte Carlo [Giles 2008] + FEM framework

For MLMC, we have two approximation levels ℓ and $\ell-1$. For any particular $\omega \in \Omega$, we need to solve: find $u_h^\ell \in V_h^\ell$, $u_h^{\ell-1} \in V_h^{\ell-1}$ s.t. for all $v_h^\ell \in V_h^\ell$, $v_h^{\ell-1} \in V_h^{\ell-1}$,

$$a(u_h^\ell, v_h^\ell) = \langle \dot{W}, v_h^\ell \rangle(\omega), \tag{5}$$
$$a(u_h^{\ell-1}, v_h^{\ell-1}) = \langle \dot{W}, v_h^{\ell-1} \rangle(\omega). \tag{6}$$

This yields the linear system

$$\left[\begin{array}{c|c} A^\ell & 0 \\ \hline 0 & A^{\ell-1} \end{array}\right] \left[\begin{array}{c} \mathbf{u}^\ell \\ \hline \mathbf{u}^{\ell-1} \end{array}\right] = \left[\begin{array}{c} \mathbf{b}^\ell \\ \hline \mathbf{b}^{\ell-1} \end{array}\right] = \mathbf{b},$$

where $\mathbf{b} \sim \mathcal{N}(0, M)$. Let $V_h^\ell = \mathrm{span}(\{\phi_i^\ell\}_{i=0}^{n_\ell})$ and $V_h^{\ell-1} = \mathrm{span}(\{\phi_i^{\ell-1}\}_{i=0}^{n_{\ell-1}})$, then

$$M = \left[\begin{array}{c|c} M^\ell & M^{\ell,\ell-1} \\ \hline (M^{\ell,\ell-1})^T & M^{\ell-1} \end{array}\right], \quad M_{ij}^{\ell,\ell-1} = \int \phi_i^\ell \phi_j^{\ell-1} \,\mathrm{dx}.$$

NOTE: we do not require the FEM approximation subspaces to be nested!

Overview

The challenge

SAMPLING PROBLEM 1: single level realisations:
sample $b \sim \mathcal{N}(0, M)$, where M is the mass matrix of V_h.

SAMPLING PROBLEM 2: coupled realisations:
sample $b \sim \mathcal{N}(0, M)$, where M is the block mass matrix given by V_h^ℓ and $V_h^{\ell-1}$.

How to sample b?

Sampling b is hard!

How to sample b?

Sampling b is hard!

Naïve approach

- Factorise $M = HH^T$ (**cubic complexity!**) and set $\mathbf{b} = H\mathbf{z}$, with $\mathbf{z} \sim \mathcal{N}(0, I)$.

$$\Rightarrow \quad \mathbb{E}[\mathbf{b}\mathbf{b}^T] = \mathbb{E}[H\mathbf{z}(H\mathbf{z})^T] = H\mathbb{E}[\mathbf{z}\mathbf{z}^T]H^T = HIH^T = M.$$

How to sample b?

Sampling b is hard!

Naïve approach

- Factorise $M = HH^T$ (**cubic complexity!**) and set $\mathbf{b} = H\mathbf{z}$, with $\mathbf{z} \sim \mathcal{N}(0, I)$.

$$\Rightarrow \quad \mathbb{E}[\mathbf{b}\mathbf{b}^T] = \mathbb{E}[H\mathbf{z}(H\mathbf{z})^T] = H\mathbb{E}[\mathbf{z}\mathbf{z}^T]H^T = HIH^T = M.$$

- Works well if M **diagonal**: previous work used either mass-lumping [Lindgren, Rue and Lindström 2009], piecewise constant elements [Osborn, Vassilevski and Villa 2017] or a piecewise constant approximation of white noise [Drzisga, et al. 2017, Du and Zhang 2002].
- We do not require M to be diagonal (and we do not approximate white noise).
- We can sample $\mathbf{b}$ with **linear complexity**.

How to sample b?

InFoMM
Industrially Focused Mathematical Modelling

Sampling b is hard!

Naïve approach

- Factorise $M = HH^T$ (**cubic complexity!**) and set $\mathbf{b} = H\mathbf{z}$, with $\mathbf{z} \sim \mathcal{N}(0, I)$.

$$\Rightarrow \quad \mathbb{E}[\mathbf{b}\mathbf{b}^T] = \mathbb{E}[H\mathbf{z}(H\mathbf{z})^T] = H\mathbb{E}[\mathbf{z}\mathbf{z}^T]H^T = HIH^T = M.$$

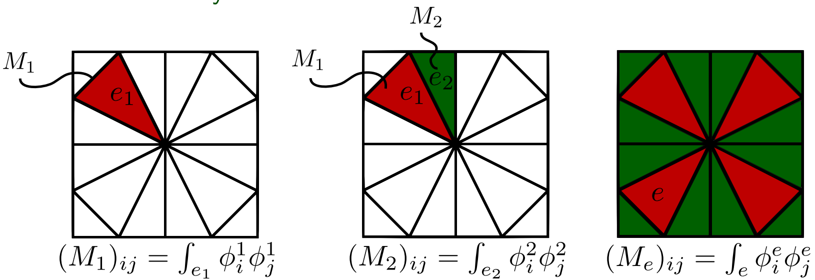
- Works well if M **diagonal**: previous work used either mass-lumping [Lindgren, Rue and Lindström 2009], piecewise constant elements [Osborn, Vassilevski and Villa 2017] or a piecewise constant approximation of white noise [Drzisga, et al. 2017, Du and Zhang 2002].
- We do not require M to be diagonal (and we do not approximate white noise).
- We can sample $\mathbf{b}$ with **linear complexity**.

IDEA! H does not need to be square, maybe we can find a more efficient factorisation!

White noise sampling: single level realisations

SAMPLING PROBLEM 1: need to sample $\mathbf{b} \sim \mathcal{N}(0, M)$.

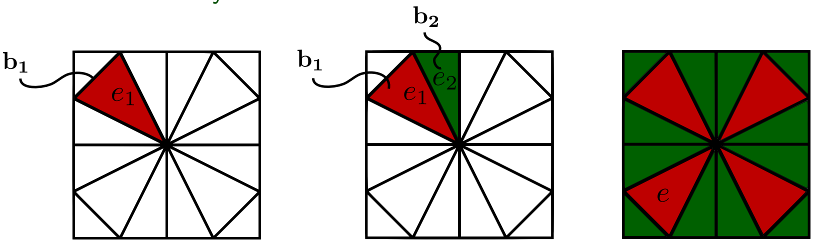
Exploit the FEM assembly



$$M = L^T \begin{bmatrix} M_1 & 0 & \cdots \\ 0 & M_2 & \ddots \\ \vdots & \ddots & \ddots \end{bmatrix} L = L^T \mathrm{diag}_e(M_e) L. \tag{7}$$

SAMPLING PROBLEM 1: need to sample $\mathbf{b} \sim \mathcal{N}(0, M)$.

Exploit the FEM assembly



$\mathbf{b_1} \sim \mathcal{N}(0, M_1)$ $\mathbf{b_2} \sim \mathcal{N}(0, M_2)$ $\mathbf{b_e} \sim \mathcal{N}(0, M_e)$

$$\mathcal{N}(0, M) \sim \mathbf{b} = L^T \begin{bmatrix} \mathbf{b_1} \\ \mathbf{b_2} \\ \vdots \end{bmatrix} = L^T \mathrm{vstack}_e(\mathbf{b}_e) \tag{8}$$

White noise sampling: single level realisations

SAMPLING PROBLEM 1: need to sample $\mathbf{b} \sim \mathcal{N}(0, M)$.

Exploit the FEM assembly

- Each $\mathbf{b}_e$ can be sampled as $\mathbf{b}_e = H_e \mathbf{z}_e$ with $\mathbf{z}_e \sim \mathcal{N}(0, I)$ and $H_e H_e^T = M_e$.
- $\mathbf{b} = L^T \text{vstack}_e(\mathbf{b}_e)$ is $\mathcal{N}(0, M)$ since

$$\begin{aligned}\mathbb{E}[\mathbf{b}\mathbf{b}^T] &= L^T \mathbb{E}[\text{vstack}_e(\mathbf{b}_e)\text{vstack}_e(\mathbf{b}_e)^T] L \\ &= L^T \text{diag}_e(H_e)\text{diag}_e(H_e^T) L = L^T \text{diag}_e(M_e) L = M.\end{aligned}$$

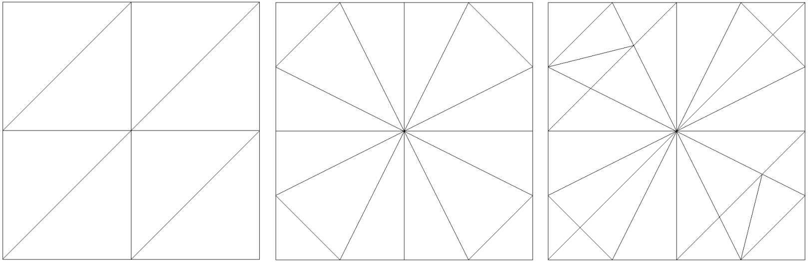
- If the mapping to the FEM reference element is **affine** (e.g. Lagrange elements on simplices) we have that $M_e/|e| =$ const on each element and **only one local factorisation is needed**.

This approach is trivially parallelisable!

SAMPLING PROBLEM 2: need to sample $\mathbf{b} \sim \mathcal{N}(0, M)$, where M is now the block mixed mass matrix.

Definition (Supermesh, [Farrell 2009])

Let A and B be two (possibly non-nested) meshes. Their supermesh S is one of their common refinements. A and B are both nested within S.



A $\qquad$ B $\qquad$ S

White noise sampling: coupled realisations

SAMPLING PROBLEM 2: need to sample $\mathbf{b} \sim \mathcal{N}(0, M)$, where M is now the block mixed mass matrix.

- Factorise locally, this time on each **supermesh element**.
- Sample $\mathbf{b}$ on S, then interpolate/project the result onto A and B (this step can be performed locally).
- Since A and B are nested within S, this operation is exact. Note that A and B **need not** be nested.

Previous work on white noise coupling for MLMC used either a nested hierarchy [Drzisga et al. 2017, Osborn et al. 2017] or an algebraically constructed hierarchy of agglomerated meshes [Osborn, Vassilevski and Villa 2017].

Complexity overview

	offline cost	online cost (per sample)	memory storage
single level	0 (or $O(m^3N)$)	$O(m^3N)$ (or $O(m^2N)$)	$O(m^2)$ (or $O(m^2N)$)
single l. (affine)	$O(m^3)$	$O(m^2N)$	$O(m^2)$
coupled	0 (or $O(m^3N_S)$)	$O(m^3N_S)$ (or $O(m^2N_S)$)	$O(m^2)$ (or $O(m^2N_S)$)
coupled (affine)	$O(m^3)$	$O(m^2N_S)$	$O(m^2)$

Table: *Memory and cost complexity of our white noise sampling strategy. In the non-affine case the cost per sample can be lowered by precomputing and storing the local factorisations (see entries in blue). N_S is the number of supermesh elements. In our experience with MLMC, $N_S \leq c_d N_\ell$ and $c_d = 2$ (1D), $c_d = 2.5$ (2D), $c_d = 45$ (3D).*

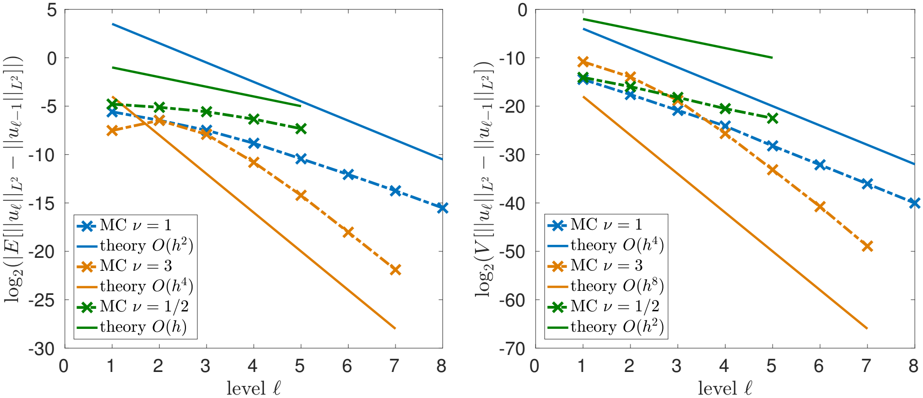
Overview

Numerical results: convergence of $P(u) = \|u\|^2_{L^2(D)}$

Consider the linear elliptic SPDE [Lindgren, Rue and Lindström 2009], [Bolin, Kirchner and Kovács 2017],

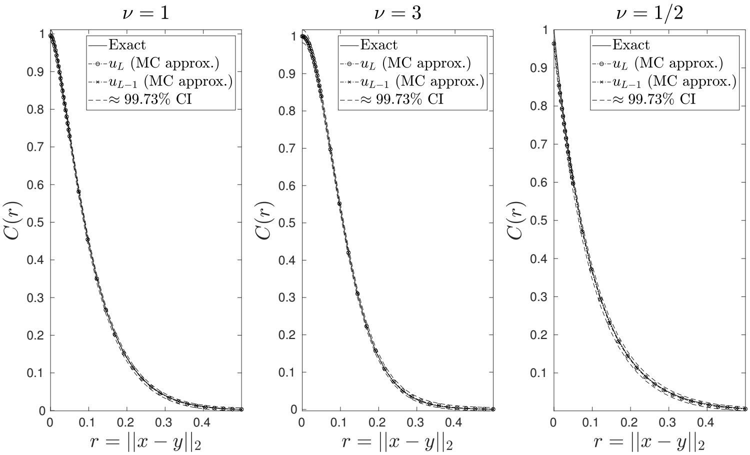
$$\left(\mathcal{I} - \kappa^{-2}\Delta\right)^k u(x,\omega) = \eta \dot{W}, \quad x \in D \subseteq \mathbb{R}^d, \quad \omega \in \Omega, \quad \nu = 2k - d/2 > 0.$$

We compute FEM solutions $\{u_h^\ell\}_{\ell=1}^{\ell=8}$ with a non-nested hierarchy of subspaces $\{V_h^\ell\}_{\ell=1}^{\ell=8}$.



Numerical results: covariance convergence

$$\mathcal{C}(r) = \mathbb{E}[u(x)u(y)] = \frac{1}{2^{\nu-1}\Gamma(\nu)}(\kappa r)^{\nu}\mathcal{K}_{\nu}(\kappa r), \quad r = \|x-y\|_2, \quad \kappa = \frac{\sqrt{8\nu}}{\lambda}, \quad x, y \in D,$$



Overview

Outlook

- White noise is an extremely non-smooth object and is defined through its integral.
- We can sample single level white noise realisations efficiently.
- We can couple white noise between different FEM approximation subspaces. A supermesh construction is not needed in the nested case.
- The overall order of **complexity is linear** in the number of elements of the supermesh and it can be **trivially parallelised**. Standard techniques usually have cubic complexity.

Further work: extensions to QMC and MLQMC.

Paper: `https://arxiv.org/abs/1803.04857`

References - Thank you for listening!

[1] M. Croci, M. B. Giles, M. E. Rognes, and P. E. Farrell. Efficient white noise sampling and coupling for multilevel Monte Carlo with non-nested meshes. *Preprint*, 2018. URL `https://arxiv.org/abs/1803.04857`.

[2] D. Bolin, K. Kirchner, and M. Kovács. Numerical solutions of fractional elliptic stochastic PDEs with spatial white noise. *Preprint*, 2017.

[3] D. Drzisga, B. Gmeiner, U. Ude, R. Scheichl, and B. Wohlmuth. Scheduling massively parallel multigrid for multilevel Monte Carlo methods. *SIAM Journal of Scientific Computing*, 39(5): 873–897, 2017.

[4] Q. Du and T. Zhang. Numerical approximation of some linear stochastic partial differential equations driven by special additive noises. *SIAM Journal on Numerical Analysis*, 40(4):1421–1445, 2002.

[5] M. B. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.

[6] F. Lindgren, H. Rue, and J. Lindström. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2009.

[7] S. Osborn, P. S. Vassilevski, and U. Villa. A multilevel, hierarchical sampling technique for spatially correlated random fields. *Preprint*, 2017.

[8] S. Osborn, P. Zulian, T. Benson, U. Villa, R. Krause, and P. S. Vassilevski. Scalable hierarchical PDE sampler for generating spatially correlated random fields using non-matching meshes. *Preprint*, 2017.

[9] A. J. Wathen. Realistic eigenvalue bounds for the Galerkin mass matrix. *IMA Journal of Numerical Analysis*, 7:449–457, 1987.