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
Overview

Introduction

White noise sampling

Numerical results

Conclusions and further work
Overview

Introduction

White noise sampling

Numerical results

Conclusions and further work
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

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

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

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

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.

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

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

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

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.

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

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

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

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

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.

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

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

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

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.

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

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.
WARNING! Point evaluation not defined!
WARNING! Point evaluation not defined!
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 $\langle \dot{W}, \phi \rangle := \int_D \dot{W}\phi \, dx$. 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,

$$\mathbb{E}[b_ib_j] = \int_D \phi_i\phi_j \, dx =: M_{ij}, \quad b \sim \mathcal{N}(0, M).$$

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

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

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

(2)

Where $V_h = \text{span}(\{\phi_i\}_{i=0}^n)$, (e.g. with Lagrange elements).
When solving SPDEs (see 1st slide) with FEM, we get (for linear problems)

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

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

(2)

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

**FEM linear system:** $u_h = \sum_i u_i \phi_i$, $u = [u_0, \ldots, u_n]^T$,

$$Au = b(\omega),$$

(3)

where the entries of $b$ are given by,

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

(4)

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


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

$$a(u_\ell^h, v_\ell^h) = \langle \dot{W}, v_\ell^h \rangle (\omega),$$

$$a(u_{\ell - 1}^h, v_{\ell - 1}^h) = \langle \dot{W}, v_{\ell - 1}^h \rangle (\omega).$$
For MLMC, we have two approximation levels $\ell$ 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 $v_h^\ell \in V_h^\ell$, $v_h^{\ell-1} \in V_h^{\ell-1}$,

\[ a(u_h^\ell, v_h^\ell) = \langle \hat{W}, v_h^\ell \rangle(\omega), \]

\[ a(u_h^{\ell-1}, v_h^{\ell-1}) = \langle \hat{W}, v_h^{\ell-1} \rangle(\omega). \]

This yields the linear system

\[
\begin{bmatrix}
A^\ell & 0 \\
0 & A^{\ell-1}
\end{bmatrix}
\begin{bmatrix}
u^\ell \\
u^{\ell-1}
\end{bmatrix}
= \begin{bmatrix}
b^\ell \\
b^{\ell-1}
\end{bmatrix} = b,
\]
For MLMC, we have two approximation levels $\ell$ and $\ell - 1$. For any particular $\omega \in \Omega$, we need to solve: find $u_\ell^h \in V_\ell^h$, $u_{\ell - 1}^h \in V_{\ell - 1}^h$ s.t. for all $v_\ell^h \in V_\ell^h$, $v_{\ell - 1}^h \in V_{\ell - 1}^h$,

$$a(u_\ell^h, v_\ell^h) = \langle \dot{W}, v_\ell^h \rangle(\omega),$$  
(5)

$$a(u_{\ell - 1}^h, v_{\ell - 1}^h) = \langle \dot{W}, v_{\ell - 1}^h \rangle(\omega).$$  
(6)

This yields the linear system

$$
\begin{bmatrix}
A_\ell & 0 \\
0 & A_{\ell - 1}
\end{bmatrix}
\begin{bmatrix}
u_\ell \\
u_{\ell - 1}
\end{bmatrix}
= 
\begin{bmatrix}
b_\ell \\
b_{\ell - 1}
\end{bmatrix}
= b,
$$

where $b \sim \mathcal{N}(0, M)$. Let $V_\ell^h = \text{span}(\{\phi_\ell^i\}_{i=0}^{n_\ell})$ and $V_{\ell - 1}^h = \text{span}(\{\phi_{\ell - 1}^i\}_{i=0}^{n_{\ell - 1}})$, then

$$M = 
\begin{bmatrix}
M_\ell & M_{\ell, \ell - 1} \\
(M_{\ell, \ell - 1})^T & M_{\ell - 1}
\end{bmatrix},
M_{ij}^{\ell, \ell - 1} = \int \phi_i^\ell \phi_j^{\ell - 1} \, dx.$$
For MLMC, we have two approximation levels $\ell$ 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 $v_h^\ell \in V_h^\ell$, $v_h^{\ell - 1} \in V_h^{\ell - 1}$,

$$a(u_h^\ell, v_h^\ell) = \langle \dot{W}, v_h^\ell \rangle(\omega),$$  \hspace{1cm} (5)

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

This yields the linear system

$$\begin{bmatrix} A^\ell & 0 \\ 0 & A^{\ell - 1} \end{bmatrix} \begin{bmatrix} u^\ell \\ u^{\ell - 1} \end{bmatrix} = \begin{bmatrix} b^\ell \\ b^{\ell - 1} \end{bmatrix} = b,$$

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

$$M = \begin{bmatrix} M^\ell & M_{\ell,\ell - 1} \\ (M_{\ell,\ell - 1})^T & M^{\ell - 1} \end{bmatrix}, \quad M_{ij}^{\ell,\ell - 1} = \int \phi_i^\ell \phi_j^{\ell - 1} \, dx.$$

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

Introduction

White noise sampling

Numerical results

Conclusions and further work
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 $V_h^\ell$ and $V_h^{\ell-1}$. 
Sampling $b$ is hard!

**Na"ıve approach**

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

$$E[bb^T] = E[Hz(Hz)^T] = HE[zz^T]H^T = HIH^T = M.$$ 


- We do not require $M$ to be diagonal (and we do not approximate white noise).

- We can sample $b$ with linear complexity.

**IDEA!**

$H$ does not need to be square, maybe we can find a more efficient factorisation!
Sampling $b$ is hard!

Naïve approach

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

$$\Rightarrow \quad \mathbb{E}[b b^T] = \mathbb{E}[H z (H z)^T] = H \mathbb{E}[z z^T] H^T = H I H^T = M.$$
How to sample $b$?

Sampling $b$ is hard!

Naïve approach

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

\[ \mathbb{E}[bb^T] = \mathbb{E}[Hz(Hz)^T] = HE[zz^T]H^T = HIH^T = M. \]


- We do not require $M$ to be diagonal (and we do not approximate white noise).

- We can sample $b$ with linear complexity.
How to sample $b$?

Sampling $b$ is hard!

Naïve approach

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

\[
\Rightarrow \quad \mathbb{E}[bb^T] = \mathbb{E}[Hz(Hz)^T] = H\mathbb{E}[zz^T]H^T = HH^T = M.
\]


- We do not require $M$ to be diagonal (and we do not approximate white noise).

- We can sample $b$ with linear complexity.

IDEA! $H$ does not need to be square, maybe we can find a more efficient factorisation!
SAMPLING PROBLEM 1: need to sample $\mathbf{b} \sim \mathcal{N}(0, M)$.

Exploit the FEM assembly

\[
(M_1)_{ij} = \int_{e_1} \phi_i^1 \phi_j^1, \quad (M_2)_{ij} = \int_{e_2} \phi_i^2 \phi_j^2, \quad (M_e)_{ij} = \int_e \phi_i^e \phi_j^e
\]

\[
M = L^T \begin{bmatrix}
M_1 & 0 & \cdots \\
0 & M_2 & \ddots \\
\vdots & \ddots & \ddots 
\end{bmatrix} L = L^T \text{diag}_e(M_e) L.
\]
White noise sampling: single level realisations

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

Exploit the FEM assembly

\begin{align*}
\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)
\end{align*}

\[ \mathcal{N}(0, M) \sim \mathbf{b} = L^T \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \end{bmatrix} = L^T \text{vstack}_e(\mathbf{b}_e) \quad (8) \]
SAMPLING PROBLEM 1: need to sample $b \sim \mathcal{N}(0, M)$.

Exploit the FEM assembly

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

\[
\mathbb{E}[bb^T] = L^T \mathbb{E}[\text{vstack}_e(b_e)\text{vstack}_e(b_e)^T]L \\
= L^T \text{diag}_e(H_e)\text{diag}_e(H_e^T)L = L^T \text{diag}_e(M_e)L = M.
\]

- 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 $b \sim \mathcal{N}(0, M)$, 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 $S$. 

![Diagram showing meshes A, B, and S]
SAMPLING PROBLEM 2: need to sample $b \sim \mathcal{N}(0, M)$, where $M$ is now the block mixed mass matrix.

- Factorise locally, this time on each supermesh element.
- Sample $b$ 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

<table>
<thead>
<tr>
<th></th>
<th>offline cost</th>
<th>online cost (per sample)</th>
<th>memory storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>single level</td>
<td>0 (or $O(m^3N)$)</td>
<td>$O(m^3N)$ (or $O(m^2N)$)</td>
<td>$O(m^2)$ (or $O(m^2N)$)</td>
</tr>
<tr>
<td>single l. (affine)</td>
<td>$O(m^3)$</td>
<td>$O(m^2N)$</td>
<td>$O(m^2)$</td>
</tr>
<tr>
<td>coupled</td>
<td>0 (or $O(m^3N_S)$)</td>
<td>$O(m^3N_S)$ (or $O(m^2N_S)$)</td>
<td>$O(m^2N_S)$</td>
</tr>
<tr>
<td>coupled (affine)</td>
<td>$O(m^3)$</td>
<td>$O(m^2N_S)$</td>
<td>$O(m^2)$</td>
</tr>
</tbody>
</table>

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

Introduction

White noise sampling

Numerical results

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

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

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

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

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

Introduction

White noise sampling

Numerical results

Conclusions and further work
Conclusions and further work

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.

References - Thank you for listening!


