MeQUR: Uncertainty in THM Coupled Analyses of Barrier Integrity

Numerical Approximation of Random Fields

Oliver Ernst and Chao Zhang
Department of Mathematics, TU Chemnitz

URS Retreat 17.-19.04.2023
Annweiler am Trifels

TECHNISCHE UNIVERSITÄT
IN DER KULTURHAUPTSTADT EUROPAS
CHEMNITZ

BUNDESGESELLSCHAFT FÜR ENDLAGERUNG

Mathematik!

Random Fields in UQ for THM Simulations

Context

Many **uncertain input parameters** vary in space:

- intrinsic permeability (in mass and energy balance, T/H)
- thermal conductivity (in energy balance, T)
- Young's modulus (in momentum balance, M)

Model these as **random variables**

- homogeneous: each uncertain quantity Z is a scalar random variable $Z = Z(\omega)$, each realization constant throughout computational domain D;
- heterogeneous: each uncertain quantity Z is **random field** (a.k.a. random function, stochastic process) $Z = Z(\boldsymbol{x}, \omega)$, $\boldsymbol{x}$ spatial coordinate.

Some uncertain/random quantities not scalar but tensors (symmetric positive definite matrices). In these cases the random fields are tensor-valued.

Random Fields in UQ for THM Simulations

Context

Many **uncertain input parameters** vary in space:

- intrinsic permeability (in mass and energy balance, T/H)
- thermal conductivity (in energy balance, T)
- Young's modulus (in momentum balance, M)

Model these as **random variables**

- homogeneous: each uncertain quantity Z is a scalar random variable $Z = Z(\omega)$, each realization constant throughout computational domain D;
- heterogeneous: each uncertain quantity Z is **random field** (a.k.a. random function, stochastic process) $Z = Z(\boldsymbol{x}, \omega)$, $\boldsymbol{x}$ spatial coordinate.

Some uncertain/random quantities not scalar but tensors (symmetric positive definite matrices). In these cases the random fields are tensor-valued.

Random Fields in UQ for THM Simulations
Context

Many **uncertain input parameters** vary in space:

- intrinsic permeability (in mass and energy balance, T/H)
- thermal conductivity (in energy balance, T)
- Young's modulus (in momentum balance, M)

Model these as **random variables**

- homogeneous: each uncertain quantity Z is a scalar random variable $Z = Z(\omega)$, each realization constant throughout computational domain D;
- heterogeneous: each uncertain quantity Z is **random field** (a.k.a. random function, stochastic process) $Z = Z(\boldsymbol{x}, \omega)$, $\boldsymbol{x}$ spatial coordinate.

Some uncertain/random quantities not scalar but tensors (symmetric positive definite matrices). In these cases the random fields are tensor-valued.

Common assumptions/notation:

- finite mean

$$f_0(\boldsymbol{x}) = \mathbf{E}\left[Z(\boldsymbol{x}, \cdot)\right], \qquad \boldsymbol{x} \in D$$

- finite covariances

$$c(\boldsymbol{x}, \boldsymbol{y}) := \mathbf{Cov}(Z(\boldsymbol{x}, \cdot), Z(\boldsymbol{y}, \cdot)) = \mathbf{E}\left[(Z(\boldsymbol{x}, \cdot) - f_0(\boldsymbol{x}))\,(Z(\boldsymbol{y}, \cdot) - f_0(\boldsymbol{y}))\right]$$

- (Wide-sense) stationarity, i.e., $c(\boldsymbol{x}, \boldsymbol{y}) = c(\boldsymbol{x} - \boldsymbol{y})$
- Isotropy

$$c(\boldsymbol{x}, \boldsymbol{y}) = c(|\boldsymbol{x} - \boldsymbol{y}|) \qquad \text{(Euclidean distance)}$$

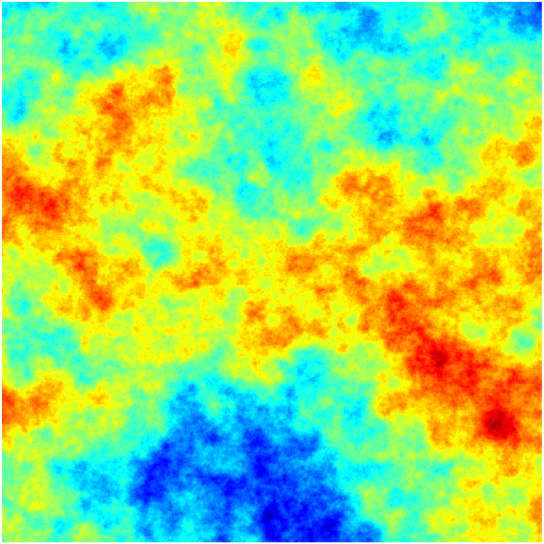
- (Statistical) anisotropy

$$c(\boldsymbol{x}, \boldsymbol{y}) = c((\boldsymbol{x} - \boldsymbol{y})^\top \boldsymbol{M}(\boldsymbol{x} - \boldsymbol{y})) \qquad \boldsymbol{M} \text{ symmetric positive definite.}$$

- Gaussian RF: $[Z(\boldsymbol{x}_1), \dots, Z(\boldsymbol{x}_n)]$ multivariate Gaussian for any n, $\{\boldsymbol{x}_j\}_{j=1}^n$.

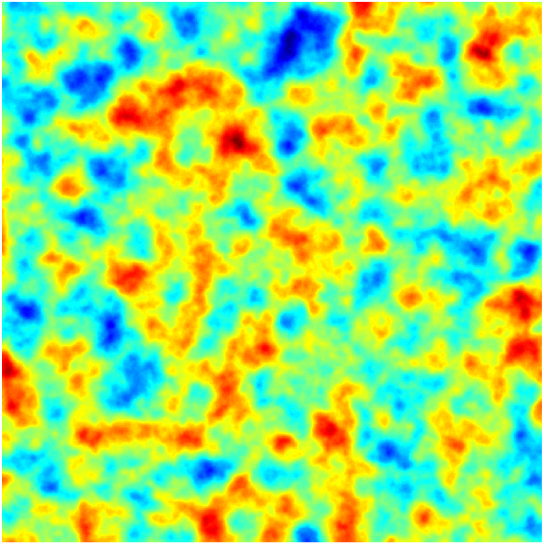
Random Fields in UQ for THM Simulations

Gauss-Matérn random fields



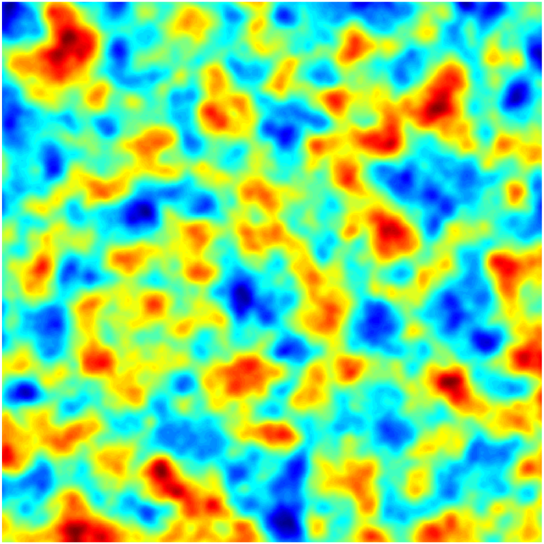
Random Fields in UQ for THM Simulations

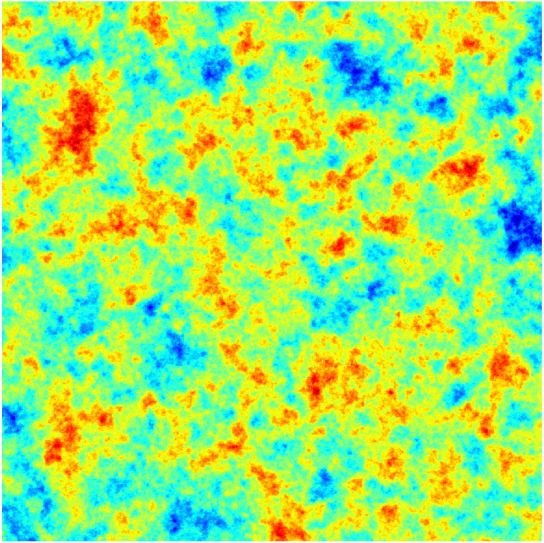
Gauss-Matérn random fields



Random Fields in UQ for THM Simulations

Gauss-Matérn random fields





Any second-order RF possesses expansion (convergent in $L^2(D)$ and $L^2(\Omega; L^2(D))$)

$$Z(\boldsymbol{x},\omega) = f_0(\boldsymbol{x}) + \sum_{m=1}^{\infty} \sqrt{\lambda_m}\,\xi_m(\omega)\, f_m(\boldsymbol{x}), \qquad \boldsymbol{x} \in D,$$

[Karhunen(1947)], [Loève(1978)]

- $f_0(\boldsymbol{x}) = \mathbf{E}[Z(\boldsymbol{x},\cdot)]$ mean of Z at $\boldsymbol{x} \in D$;
- $\{\xi_m\}_{m\in\mathbb{N}}$ pairwise uncorrelated random variables, $\mathbf{Var}\,\xi_m = 1$;
- $\{(\lambda_m, f_m)\}_{m\in\mathbb{N}}$ eigenvalues/eigenfunctions of covariance operator

$$C : L^2(D) \to L^2(D), \qquad u \mapsto Cu, \quad (Cu)(x) = \int_D u(\boldsymbol{y}) c(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}.$$

Truncated KL expansion

$$Z(\boldsymbol{x},\omega) \approx Z_M(\boldsymbol{x},\omega) := f_0(\boldsymbol{x}) + \sum_{m=1}^{M} \sqrt{\lambda_m}\,\xi_m(\omega)\, f_m(\boldsymbol{x}), \qquad M \in \mathbb{N}_0,$$

is optimal in $L^2(D)$.

Any second-order RF possesses expansion (convergent in $L^2(D)$ and $L^2(\Omega; L^2(D))$)

$$Z(\boldsymbol{x},\omega) = f_0(\boldsymbol{x}) + \sum_{m=1}^{\infty} \sqrt{\lambda_m}\,\xi_m(\omega)\, f_m(\boldsymbol{x}), \qquad \boldsymbol{x} \in D,$$

[Karhunen(1947)], [Loève(1978)]

- $f_0(\boldsymbol{x}) = \mathbf{E}\left[Z(\boldsymbol{x},\cdot)\right]$ mean of Z at $\boldsymbol{x} \in D$;
- $\{\xi_m\}_{m\in\mathbb{N}}$ pairwise uncorrelated random variables, $\mathbf{Var}\,\xi_m = 1$;
- $\{(\lambda_m, f_m)\}_{m\in\mathbb{N}}$ eigenvalues/eigenfunctions of covariance operator

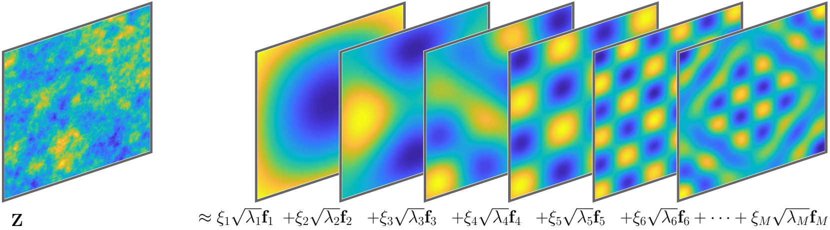
$$C : L^2(D) \to L^2(D), \qquad u \mapsto Cu, \quad (Cu)(x) = \int_D u(\boldsymbol{y})c(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}.$$

Truncated KL expansion

$$Z(\boldsymbol{x},\omega) \approx Z_M(\boldsymbol{x},\omega) := f_0(\boldsymbol{x}) + \sum_{m=1}^{M} \sqrt{\lambda_m}\,\xi_m(\omega)\, f_m(\boldsymbol{x}), \qquad M \in \mathbb{N}_0,$$

is optimal in $L^2(D)$.

- For Gaussian RF we have $\xi_m \sim \mathrm{N}(0,1)$ for all m.
- Truncated Gaussian RF: randomness completely contained in i.i.d. Gaussian random vector $\boldsymbol{\xi} = (\xi_1, \dots, \xi_M)$.
- Spatial discretization determined by that of covariance operator.



Random Fields in UQ for THM Simulations

Numerical solution of covariance eigenproblem

- Galerkin discretization $C_h \approx C$ based on piecewise constant approximation based on triangular mesh of domain D, mesh width h.
- Results in generalized eigenvalue problem $\boldsymbol{C}\boldsymbol{f} = \lambda \boldsymbol{M}\boldsymbol{f}$

$$[\boldsymbol{C}]_{i,j} = \int_D \phi_j(\boldsymbol{x}) \int_D c(\boldsymbol{x}, \boldsymbol{y}) \phi_i(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \, \mathrm{d}\boldsymbol{x}$$

$$[\boldsymbol{M}]_{i,j} = \int_D \phi_j(\boldsymbol{x}) \phi_i(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$$

Difficulties

1. $[\mathbf{C}]_{i,j}$ involves integration of **nonsmooth** kernels $c(\boldsymbol{x}, \boldsymbol{y})$
2. **C** **dense** of size $N \times N$ for N DoFs
   - storage expensive
   - matrix-vector product expensive
3. **eigen** solver

Solutions

1. **Schauter-Schwab** quadrature from BEM community to alleviate singularity
2. **hierarchical matrices**
3. **thick-restart Lanczos process**

- Galerkin discretization $C_h \approx C$ based on piecewise constant approximation based on triangular mesh of domain D, mesh width h.
- Results in generalized eigenvalue problem $\boldsymbol{C}\boldsymbol{f} = \lambda \boldsymbol{M}\boldsymbol{f}$

$$[\boldsymbol{C}]_{i,j} = \int_D \phi_j(\boldsymbol{x}) \int_D c(\boldsymbol{x}, \boldsymbol{y}) \phi_i(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \, \mathrm{d}\boldsymbol{x}$$

$$[\boldsymbol{M}]_{i,j} = \int_D \phi_j(\boldsymbol{x}) \phi_i(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$$

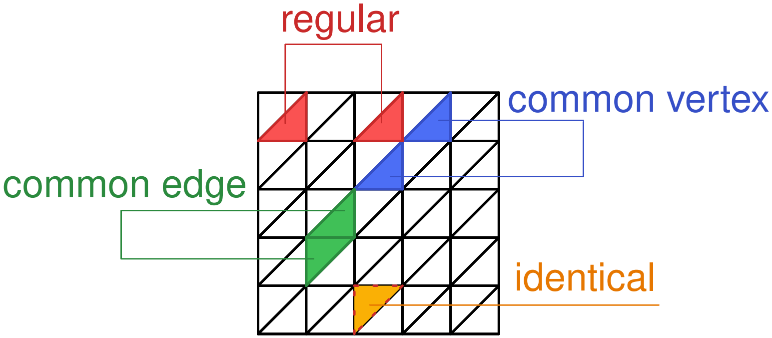
Difficulties

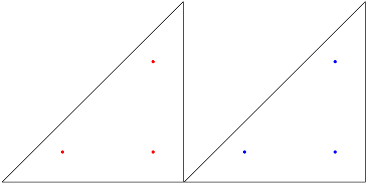
1. $[\mathbf{C}]_{i,j}$ involves integration of **nons-mooth** kernels $c(\boldsymbol{x}, \boldsymbol{y})$
2. **C** **dense** of size $N \times N$ for N DoFs
   - storage expensive
   - matrix-vector product expensive
3. **eigen** solver

Solutions

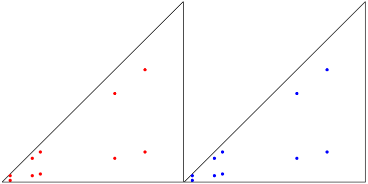
1. **Schauter-Schwab** quadrature from BEM community to allevia-te singularity
2. **hierarchical matrices**
3. **thick-restart Lanczos process**

$$[\mathbf{C}]_{i,j} = \int_D \phi_j(\boldsymbol{x}) \int_D c(\boldsymbol{x}, \boldsymbol{y}) \phi_i(\boldsymbol{y}) \, d\boldsymbol{y} \, d\boldsymbol{x}$$

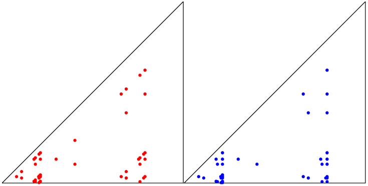




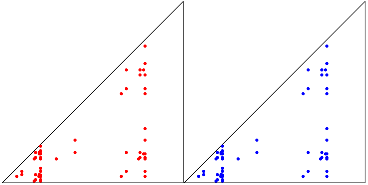
regular (Gauss rule for triangles)



common vertex



common edge

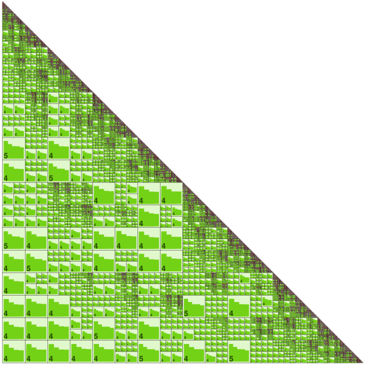


identical

Quadrature nodes on reference element pairs [Sauter & Schwab(2011)].

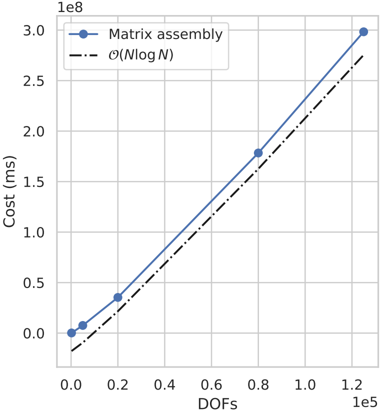
Hierarchical matrix approximation of $\boldsymbol{C}$

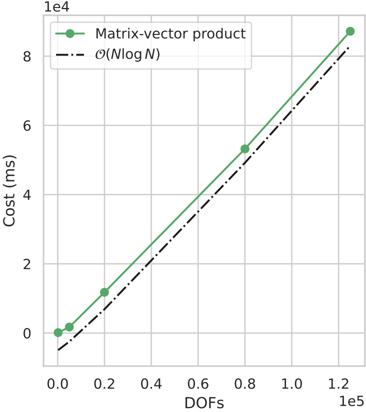
- Arrange DoF in clusters.
- Low-rank approximation of far-field cluster blocks.
- For N DoF results on $O(N \log N)$ complexity for assembly, storage and matrix-vector product. [Hackbusch(2015)]



Random Fields in UQ for THM Simulations

Hierarchical matrix approximation of $\boldsymbol{C}$





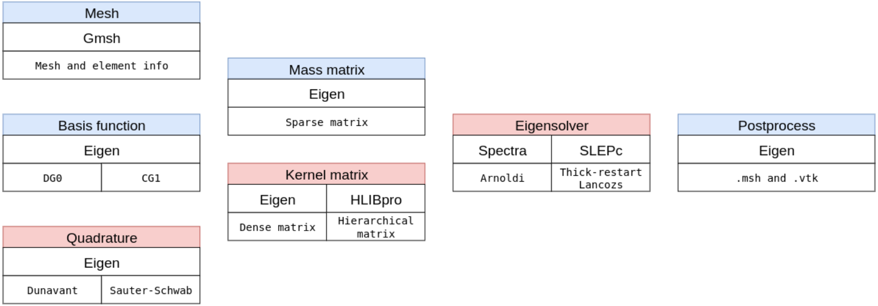
Complexity of matrix assembly and matrix vector multiplication reduced from $O(N^2)$ to $O(N \log N)$.

- For KL expansion: need only M dominant eigenpairs of $\boldsymbol{C}$.
- Hierarchical matrix approximation makes assembly and matrix-vector products inexpensive.
- Ideal setting for Krylov subspace-based eigensolvers.
- Covariance operators symmetric positive definite: can use Lanczos process based on short recurrences.
- Thick-restart variant [Wu & Simon(2000)] allows iterative refinement of target eigenspaces.
- Developed extensions for generalized eigenvalue problem, block methods for multiple eigenvalues.

Random Fields in UQ for THM Simulations

Software tool `kleme`

- `kleme` (KL expansion made easy) C++ software library for efficient solution of covariance eigenvalue problems for use with THM simulator OGS.
- Based on highly optimized packages `Eigen`, `HLIBpro`, `SpectrA`, `SlepC`.
- Available on TUC Gitlab server, includes DOXYGEN online documentation.



`kleme` components.

```cpp
#include <kleme.h>
int main(int argc,char **argv)
{
  // parse mesh
  kleme::Mesh mesh(argv[1]);

  // define dofhandler to handle mesh and basis function
  kleme::DofHandler dofhandler(mesh, 0);

  // create and assemble mass matrix
  kleme::Mass m_matrix(&dofhandler);
  m_matrix.assemble_matrix();

  // prepare quadrature rule and kernel for later use
  // in assembling stiffness matrix
  kleme::Quadrature quad(mesh.get_dim(), 2);
  kleme::ExponentialKernel kernel(1, 0.1, 45, 0.5);

  // create and assemble stiffness matrix
  kleme::StiffnessHmatrix k_hmatrix(&dofhandler, &kernel, &quad);
  k_hmatrix.assemble_matrix();

  // create solver
  kleme::SLEPc_Solver solver_hmatrix(&k_hmatrix, &m_matrix);
  int no_of_eigens = 100;
  solver_hmatrix.solve(no_of_eigens);

  // postprocess
  kleme::Postprocess postprocess(&dofhandler);
  postprocess.write_vtk(slepc_solver.eigen_vectors, "slepc_eigens.vtk");

  return 0;
}
```

Summary and Outlook

Summary

- Extend to multiple RF in THM simulation, 3D
- KL expansion versatile approach for generation of RF samples of non-homogeneous spatial inputs.
- Scalable computational approach requires adapted quadrature, hierarchical matrices and Krylov-based eigensolver.
- Software tool `kleme` for incorporation in THM simulator OGS.

Current Work

- RF sampling by solving linear fractional SPDEs [Lindgren & al.(2011)]:

$$(-\Delta + \kappa^2)^{\alpha/2} Z(x) = G(x), \quad G \text{ Gaussian white noise}$$

- "Data-free inference"[Berry & al.(2012)] for accounting for correlations in input parameters in the absence of distribution information.

Summary and Outlook

Summary

- Extend to multiple RF in THM simulation, 3D
- KL expansion versatile approach for generation of RF samples of non-homogeneous spatial inputs.
- Scalable computational approach requires adapted quadrature, hierarchical matrices and Krylov-based eigensolver.
- Software tool `kleme` for incorporation in THM simulator OGS.

Current Work

- RF sampling by solving linear fractional SPDEs [Lindgren & al.(2011)]:

$$(-\Delta + \kappa^2)^{\alpha/2} Z(\boldsymbol{x}) = G(\boldsymbol{x}), \quad G \text{ Gaussian white noise}$$

- "Data-free inference"[Berry & al.(2012)] for accounting for correlations in input parameters in the absence of distribution information.

References

Kari Karhunen.
Über lineare Methoden in der Wahrscheinlichkeitsrechnung.
Ann. Acad. Sci. Fenn., Ser. A.I. 37:3–79, 1947.

Michel Loève.
Probability Theory II, volume 46 of *Graduate Texts in Mathematics*.
Springer-Verlag, New York, 4th edition, 1978.

Stefan A. Sauter and Christoph Schwab.
Boundary Element Methods, volume 39 of *Springer Series in Computational Mathematics*.
Springer, 2011.

W Hackbusch.
Hierarchical Matrices: Algorithms and Analysis.
Springer, Berlin Heidelberg, 2015.

Finn Lindgren, Håvard Rue, and Johan 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, 73(Part 4):423–498, 2011.

Robert D. Berry, Habib N. Najm, Bert J. Debusschere, Youssef M. Marzouk, and Helgi Adalsteinsson.
Data-free inference of the joint distribution of uncertain model parameters.
Journal of Computational Physics, 231:2180–2198, 2012.

Oliver G. Ernst, Björn Sprungk, and Chao Zhang.
Uncertainty modeling and propagation for groundwater flow: A comparative study of surrogates.
(submitted), 2023.