Formulation and Stochastic Galerkin Methods for Stochastic Partial Differential Equations I

Hermann G. Matthies
with Andreas Keese

Institute of Scientific Computing
Technische Universität Braunschweig, Germany

wire@tu-bs.de
http://www.wire.tu-bs.de

Overview I

1. Background: Motivation and Model Problem

2. Formulating a Well-Posed Problem

3. Discretisation of Random Fields

4. General and Computational Approaches

5. Stochastic processes and random fileds

6. Karhunen-Loève Expansion

Why Probabilistic or Stochastic Models?

Many descriptions (especially of future events) contain elements, which are uncertain and not precisely known.

- For example future rainfall, or discharge from a river.
- More generally, action from surrounding environment.
- The system itself may contain only incompletely known parameters, processes or fields (not possible or too costly to measure)
- There may be small, unresolved scales in the model, they act as a kind of background noise.

All these introduce some uncertainty in the model.

Ontology of Uncertainty

A bit of ontology:

- Uncertainty may be aleatoric, which means random and not reducible, or
- epistemic, which means due to incomplete knowledge.

Stochastic models can give quantitative information about uncertainty, they are used for both types of uncertainty.

Possible areas of use: Reliability, heterogeneous materials, upscaling, incomplete knowledge of details, uncertain [inter-]action with environment, random loading, etc.

Quantification of Uncertainty

Uncertainty may be modelled in different ways:

Intervals / convex sets do not give a degree of uncertainty, quantification only through size of sets.

Fuzzy and possibilistic approaches model quantitative possibility with certain rules. Generalisation of set membership.

Evidence theory models basic probability, but also (as a generalisation) plausability (a kind of lower bound) and belief (a kind of upper bound) in a quantitative way. Mathematically no measures.

Stochastic / probabilistic methods model probability quantitatively, have most developed theory.

Probability

What is probability? We may understand probability as

- A mathematical concept — theory of a finite measure.
- Applies to aleatoric phenomena, i.e. frequencies of occurance — Bernoulli's weak law of large numbers.
- Applies also to epistemic concepts — extension of Aristotelian propositional logic to uncertain propositions — Cox's theorem. Realm of Bayesian and maximum entropy methods.

First view is today often labeled classical, historically Bernoulli and Laplace had the latter view.

Physical Models

Models for a system $\mathcal{S}$ may be stationary with state u, exterior action f and random model description (realisation) $\omega \in \Omega$, with probability measure $\mathbb{P}$: $\mathcal{S}(u, \omega) = f(\omega)$.

Evolution in time may be discrete (e.g. Markov chain), may be driven by discrete random process $u_{n+1} = \mathcal{F}(u_n, \omega)$,

or continuous, (e.g. Markov process, stochastic differential equation), may be driven by random processes

$$du = (\mathcal{S}(u, \omega) - f(\omega, t))dt + \mathcal{B}(u, \omega)dW(\omega, t) + \mathcal{P}(u, \omega)dQ(\omega, t)$$

In this Itô evolution equation, $W(\omega, t)$ is a Wiener process, and $Q(\omega, t)$ is a (compensated) Poisson process.

References (Incomplete)

Formulation of PDEs with random coefficients,
i.e. Stochastic Partial Differntial Equations (SPDEs):

Babuška, Tempone; Glimm; Holden, Øksendal; Xiu, Karniadakis; M., Keese; Schwab, Tudor; P.-L. Lions

Spatial/temporal expansion of stochastic processes/ random fields:

Adler; Fourier; Karhunen, Loève; Krée, Soize; Wiener

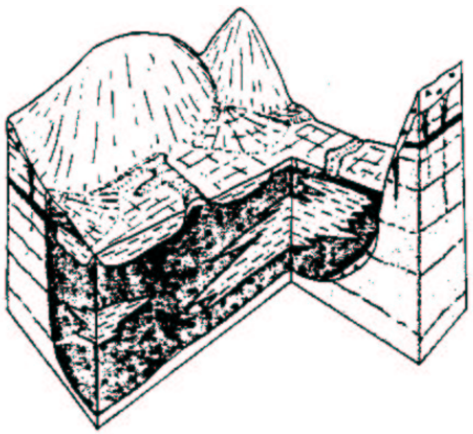
White noise analysis/ polynomial chaos (PCE)/ multiple Itô integrals:

Wiener; Cameron, Martin; Hida, Potthoff; Holden, Øksendal; Itô; Kondratiev; Malliavin

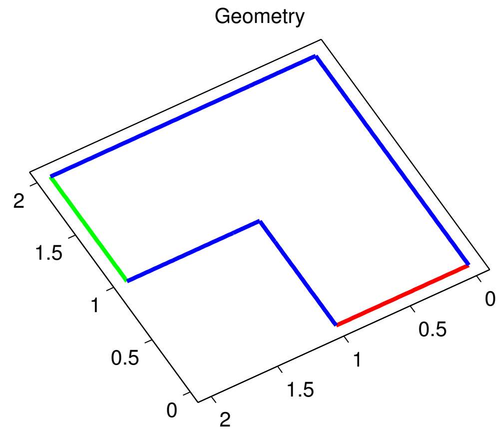
Galerkin methods for SPDEs:

Babuška, Tempone; Benth, Gjerde; Cao; Ghanem, Spanos; Xiu, Karniadakis, Lucor,; M., Keese; Schwab, Tudor

Model Problem



Aquifier



2D Model

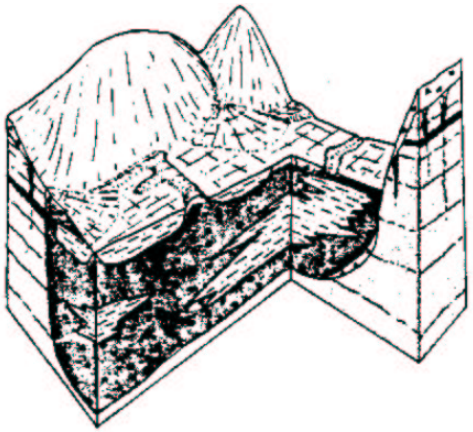
simple stationary model of groundwater flow

$$-\nabla \cdot (\kappa(x)\nabla u(x)) = f(x) \quad \text{\& b.c.,} \qquad x \in \mathcal{G} \subset \mathbb{R}^d$$

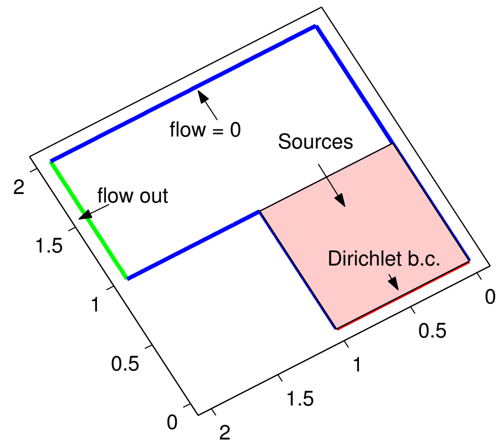
$$(\kappa(x)\nabla u(x)) \cdot \boldsymbol{n} = g(x), \qquad x \in \Gamma \subset \partial\mathcal{G},$$

u hydraulic head, κ conductivity, f and g sinks and sources.

Model Stochastic Problem



Aquifer



2D Model

simple stationary model of groundwater flow with stochastic data

$$-\nabla \cdot (\kappa(x,\omega)\nabla u(x,\omega)) = f(x,\omega) \quad \text{\& b.c.,} \qquad x \in \mathcal{G} \subset \mathbb{R}^d$$

$$(\kappa(x)\nabla u(x,\omega)) \cdot \boldsymbol{n} = g(x,\omega), \qquad x \in \Gamma \subset \partial\mathcal{G}, \quad \omega \in \Omega$$

κ stochastic conductivity, f and g stochastic sinks and sources.

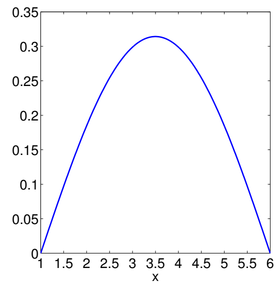
Stochastic Model

- Uncertainty of system parameters—e.g. $\kappa = \kappa(x, \omega)$ stochastic field $\omega \in \Omega$ = probability space with measure $\mathbb{P}$.

- Assumption: $0 < \kappa_0 \leq \kappa(x, \omega) < \kappa_1$. (Rather κ random tensor field).
Better $\|\kappa\|_{L_\infty(\mathcal{G}\times\Omega)} < \kappa_1 \ \wedge \ \|\kappa^{-1}\|_{L_\infty(\mathcal{G}\times\Omega)} < \kappa_0^{-1}$
Possibilities: Transformation

$$\kappa(x, \omega) = \phi(x, \gamma(x, \omega)) := F^{-1}_{\kappa(x)} \circ \Phi(\gamma(x, \omega))$$

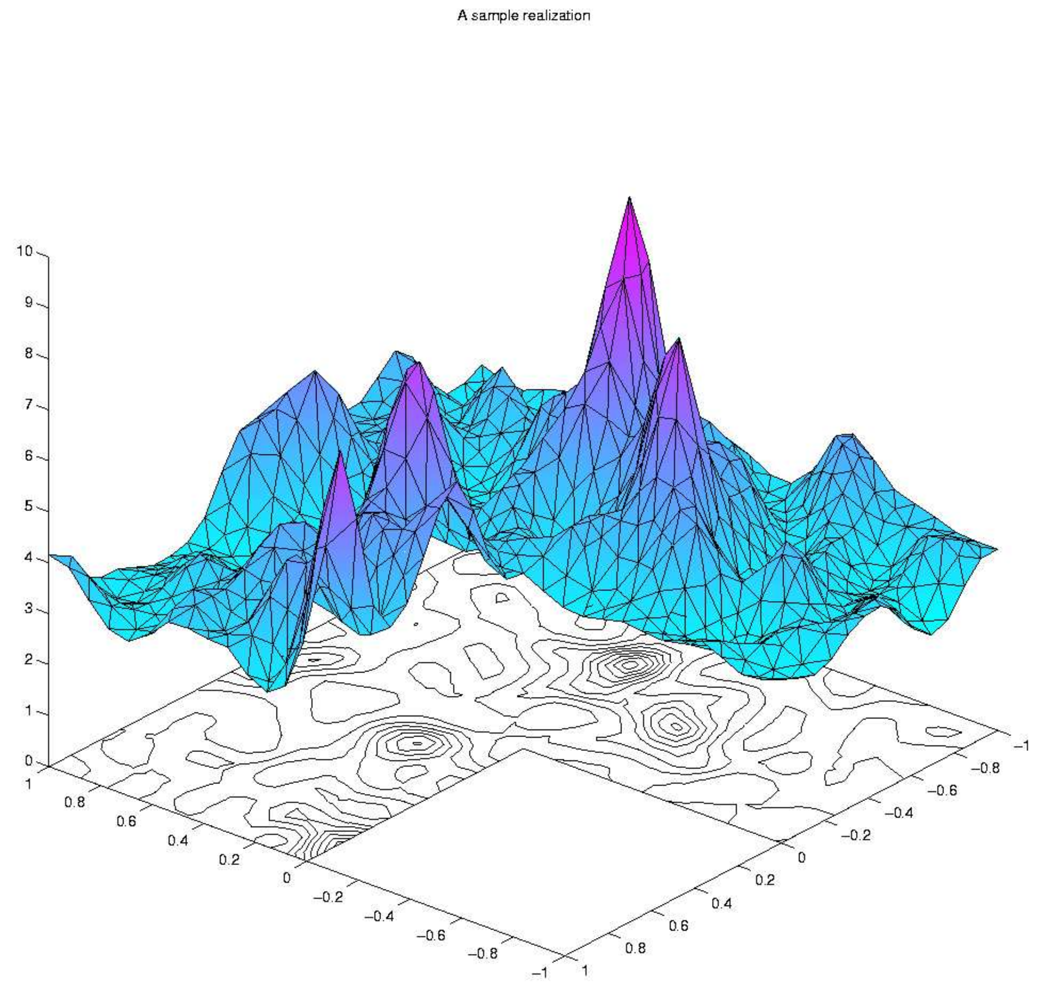
of Gaussian field γ with given 2nd order statistic.
e.g. $\kappa(x, \omega)$ has marginal
$\beta(1/2, 1/2)$-distribution



or log-normal distribution

$$\kappa(x, \omega) = a(x) + \exp(\gamma(x, \omega))$$

Realisation of $\kappa(x, \omega)$



Stochastic PDE and Variational Form

Solution $u(x,\omega)$ is a stochastic field—in a tensor product space $\mathscr{W}$ is a Sobolev space of spatial functions, $\mathscr{S}$ a space of random variables (e.g. $\mathscr{W} = H^1_{eb}(\mathcal{G}), \mathscr{S} = L_2(\Omega)$):

$$\mathscr{W} \otimes \mathscr{S} \ni u(x,\omega) = \sum_\mu v_\mu(x) u^{(\mu)}(\omega)$$

Variational formulation: Find $u \in \mathscr{W} \otimes \mathscr{S}$, such that $\forall v \in \mathscr{W} \otimes \mathscr{S}$:

$$\mathsf{a}(v,u) := \int_\Omega \int_{\mathcal{G}} \nabla v(x,\omega) \cdot (\kappa(x,\omega) \nabla u(x,\omega))\; dx \,\mathbb{P}(d\omega) =$$

$$\int_\Omega \left[\int_{\mathcal{G}} v(x,\omega) f(x,\omega)\; dx + \int_{\partial\mathcal{G}} v(x,\omega) g(x,\omega)\; dS(x) \right] \mathbb{P}(d\omega) =: \langle\!\langle f, v \rangle\!\rangle.$$

Mathematical Results

To find a solution $u \in \mathscr{W} \otimes \mathscr{S}$ such that for $\forall v:$ $\quad \mathsf{a}(v,u) = \langle\!\langle f, v \rangle\!\rangle$ under certain conditions

- is guaranteed by Lax-Milgram lemma, problem is well-posed in the sense of Hadamard (existence, uniqueness, continuous dependence on data f, g in L_2- and on κ in L_∞-norm).
- may be achieved by Galerkin methods, convergence established with Céa's lemma
- Galerkin methods are stable, if no variational crimes are committed

Good approximating subspaces of $\mathscr{W} \otimes \mathscr{S}$ have to be found, as well as efficient numerical procedures worked out.

Functionals of Interest

Desirable: Uncertainty Quantification or Optimisation under uncertainty:

The goal is to compute functionals of the solution:

$$\boldsymbol{\Psi}_u = \langle \boldsymbol{\Psi}(u) \rangle := \mathbb{E}\left(\boldsymbol{\Psi}(u)\right) := \int_{\Omega} \int_{\mathcal{G}} \Psi(u(x,\omega), x, \omega) \, dx \, \mathbb{P}(d\omega)$$

e.g.: $\bar{u} = \mathbb{E}\left(u\right)$, or $\mathrm{var}_u = \mathbb{E}\left((\tilde{u})^2\right)$, where $\tilde{u} = u - \bar{u}$,
or $\mathbb{P}\{u \leq u_0\} = \mathbb{P}(\{\omega \in \Omega |\, u(\omega) \leq u_0\}) = \mathbb{E}\left(\chi_{\{u \leq u_0\}}\right)$

All desirables are usually expected values of some functional, to be computed via (high dimensional) integration over Ω.

General Approaches

Alternative Formulations / Approaches

- Moments: Derive equations for the moments of the quantities of interest.
- Probablity distributions / densities: Derive equations for the probability densities, e.g. Master-Equation, Fokker-Planck.
- Direct Integration: Compute desired statistics via direct integration over Ω—high dimensional (e.g. Monte Carlo, Quasi Monte Carlo, Smolyak ($=$ sparse grids)).
- Direct Approximation: Compute an approximation to $u(x, \omega)$, use this to compute everything else (traditional response surface methods, stochastic Galerkin, stochastic collocation)

General Computational Approach

Principal Approach:

1. Discretise / approximate physical model (e.g. via finite elements, finite differences), and approximate stochastic model (processes, fields) in finitely many independent random variables (RVs), $\Rightarrow$ stochastic discretisation.

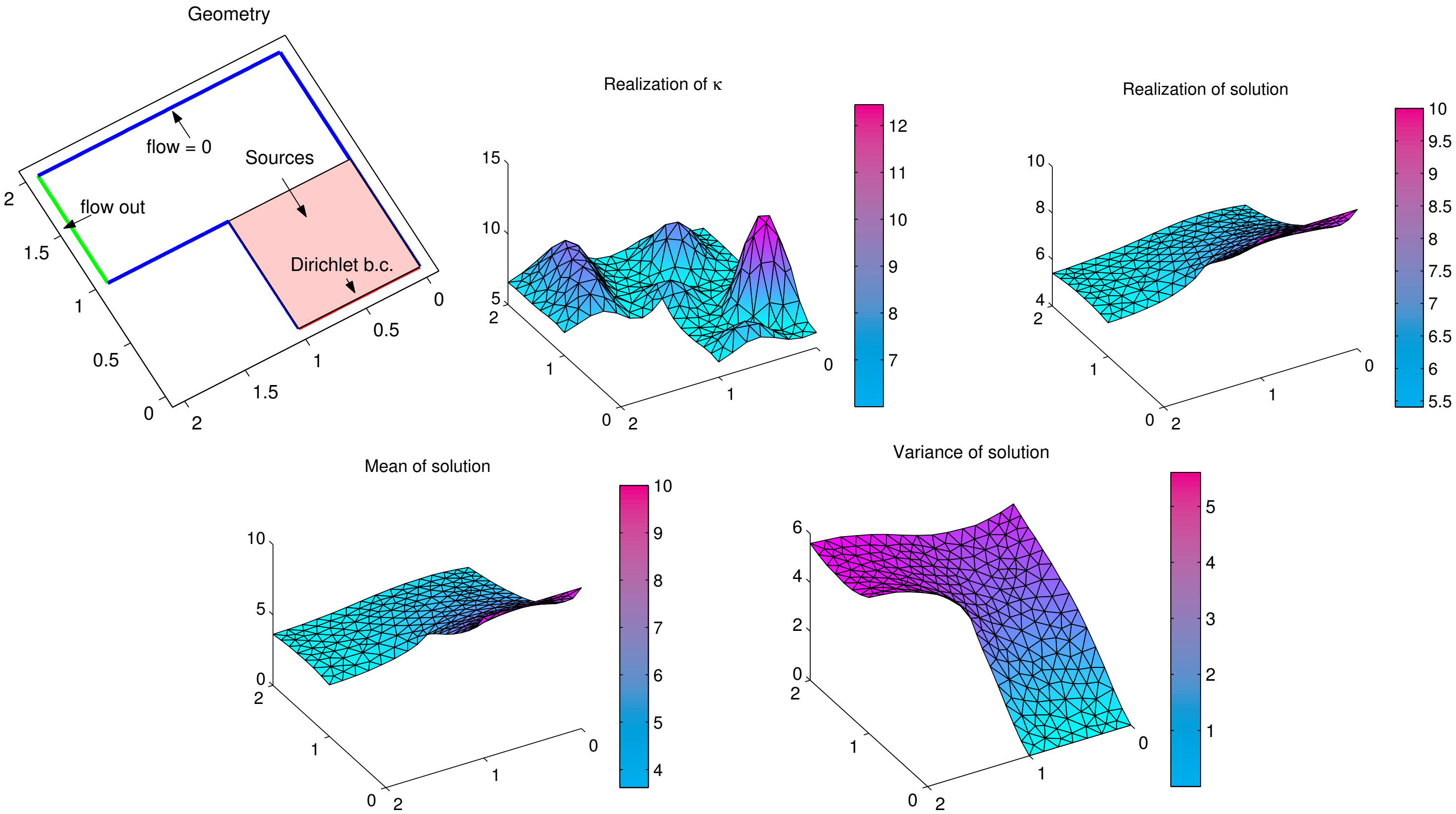
2. Compute statistics via integration over Ω—high dimensional (e.g. Monte Carlo, Quasi Monte Carlo, Smolyak (= sparse grids)):
   - Via direct integration. Each integration point $\omega_z \in \Omega$ requires one expensive PDE solution (with rough data).
   - Or approximate solution with some response-surface, then integration by sampling a cheap expression at each integration point.

Computational Requirements

- How to represent a stochastic process for computation, both simulation or otherwise?
- Best would be as some combination of countably many independent random variables (RVs).
- How to compute the required integrals or expectations numerically?
- Best would be to have probability measure as a product measure $\mathbb{P} = \mathbb{P}_1 \otimes \ldots \otimes \mathbb{P}_\ell$, then integrals can be computed as iterated one-dimensional integrals via Fubini's theorem,

$$\int_\Omega \Psi \, \mathbb{P}(d\omega) = \int_{\Omega_1} \ldots \int_{\Omega_\ell} \Psi \, \mathbb{P}_1(d\omega_1) \ldots \mathbb{P}_\ell(d\omega_\ell)$$

Example Solution



Tools of the Trade

A $\mathcal{V}$-valued random variable (RV) r is a map $\Omega \mapsto \mathcal{V}$ (mostly $\mathcal{V} = \mathbb{R}$) completely specified by its distribution function

$$\forall r \in \mathbb{R}: \; F_{\mathsf{r}}(r) := \Pr\{\mathsf{r}(\omega) \le r\} := \int_{\{\mathsf{r}(\omega)\le r\}} \mathbb{P}(d\omega) = \mathbb{E}\left(\chi_{\{\mathsf{r}(\omega)\le r\}}\right).$$

Mean $\bar{\mathsf{r}} = \mathbb{E}\left(\mathsf{r}(\cdot)\right)$, [auto-]covariance $C_{\mathsf{r}} := \mathbb{E}\left(\tilde{\mathsf{r}} \otimes \tilde{\mathsf{r}}\right)$, and fluctuating part $\tilde{\mathsf{r}}(\omega) = \mathsf{r}(\omega) - \bar{\mathsf{r}}$, with $\mathbb{E}\left(\tilde{\mathsf{r}}\right) = 0$.

Two RVs r_1 and r_2 are

uncorrelated If the [cross-]covariance $C_{\mathsf{r}1,2} := \mathbb{E}\left(\tilde{\mathsf{r}}_1 \otimes \tilde{\mathsf{r}}_2\right) = 0$, or if in case $\mathcal{V} = \mathbb{R}$: $\langle \tilde{\mathsf{r}}_1, \tilde{\mathsf{r}}_2 \rangle := \mathbb{E}\left(\tilde{\mathsf{r}}_1 \tilde{\mathsf{r}}_2\right) = 0$ (orthogonal).

independent if for all functions ϕ_1 and ϕ_2 it holds that $\mathbb{E}\left(\phi_1(\mathsf{r}_1)\phi_2(\mathsf{r}_2)\right) \equiv \mathbb{E}\left(\phi_1(\mathsf{r}_1)\right)\mathbb{E}\left(\phi_2(\mathsf{r}_2)\right)$.

Stochastic Processes I

Consider interval $\mathscr{T} = [0, T]$, stochastic process is $\forall t \in \mathscr{T}$ a RV $\mathsf{s}_t(\omega)$

alternatively $\forall \omega \in \Omega$ random function—a realisation— $\mathsf{s}_\omega(t)$ on $\mathscr{T}$

Often only second order information—mean and covariance—known.

Mean $\bar{\mathsf{s}}(t) = \mathbb{E}\left(\mathsf{s}_\omega(t)\right)$—now a function of t—and fluctuating part $\tilde{\mathsf{s}}(t, \omega)$.

Covariance may be considered at different times

$$\mathsf{C}_\mathsf{s}(t_1, t_2) := \mathbb{E}\left(\tilde{\mathsf{s}}(t_1, \cdot) \otimes \tilde{\mathsf{s}}(t_2, \cdot)\right)$$

Stochastic Processes II

If $\bar{\mathsf{s}}(t) \equiv \bar{\mathsf{s}}$, and $\mathsf{C_s}(t_1, t_2) = c_\mathsf{s}(t_1 - t_2)$,
process is (weakly) stationary, with spectrum

$$S_\mathsf{s}(\nu_k) = \int_0^T c_\mathsf{s}(t) \exp(-i2\pi\nu_k t)\, dt, \quad \nu_k = \frac{k}{T};\ k \in \mathbb{Z}.$$

Process s may be realised (Fourier synthesised) by

$$\mathsf{s}(t, \omega) = \bar{\mathsf{s}} + \sum_{k=-\infty}^{\infty} \varsigma_k(\omega) \sqrt{S_\mathsf{s}(\nu_k)} \exp(i2\pi\nu_k t)$$

$\varsigma_k(\omega)$ are zero mean, unit variance uncorrelated RVs
$(\mathbb{E}\left(\varsigma_k \varsigma_\ell\right) = \langle \varsigma_k, \varsigma_\ell \rangle = \delta_{k\ell})$.

Random Fields

Mean $\bar{\kappa}(x) = \mathbb{E}\left(\kappa_\omega(x)\right)$ and fluctuating part $\tilde{\kappa}(x,\omega)$.

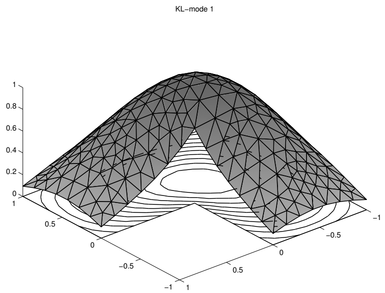
Covariance may be considered at different positions
$\mathsf{C}_\kappa(x_1,x_2) := \mathbb{E}\left(\tilde{\kappa}(x_1,\cdot)\otimes\tilde{\kappa}(x_2,\cdot)\right)$

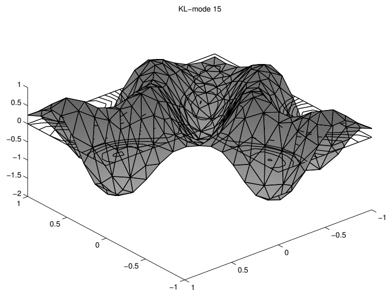
If $\bar{\kappa}(x) \equiv \bar{\kappa}$, and $\mathsf{C}_\kappa(x_1,x_2) = c_\kappa(x_1-x_2)$, process is homogeneous.

Here representation through spectrum as a Fourier sum is well known.

- Need to discretise spatial aspect (generalise Fourier representation). One possibility is the Karhunen-Loève expansion (KLE).
- Need to discretise each of the random variables in Fourier synthesis. One possibility is Wiener's polynomial chaos expansion (PCE).

Karhunen-Loève Expansion I

mode 1: 

mode 15: 

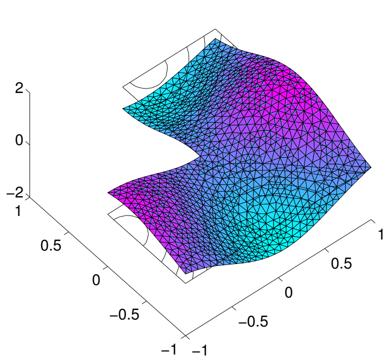
KLE: Other names: Proper Orthogonal Decomposition (POD), Singular Value Decomposition (SVD), Principal Component Analalysis (PCA): spectrum of $\{\varkappa_\jmath^2\} \subset \mathbb{R}_+$ and orthogonal KLE eigenfunctions $g_\jmath(x)$:

$$\int_{\mathcal{G}} \mathsf{C}_\kappa(x,y) g_\jmath(y)\, dy = \varkappa_\jmath^2\, g_\jmath(x) \quad \text{with} \quad \int_{\mathcal{G}} g_\jmath(x) g_k(x)\, dx = \delta_{\jmath k}.$$

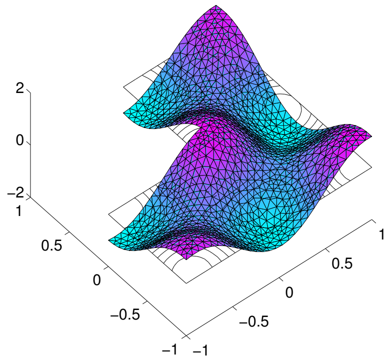
$\Rightarrow$ Mercer's representation of C_κ:

$$\mathsf{C}_\kappa(x,y) = \sum_{\jmath=1}^{\infty} \varkappa_\jmath^2\, g_\jmath(x) g_\jmath(y)$$

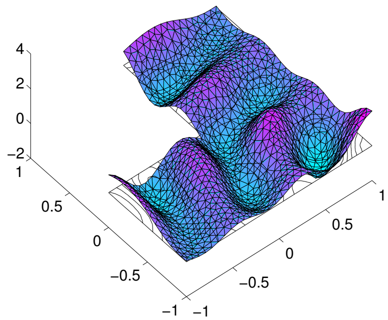
Karhunen-Loève Expansion II



mode 5



mode 10



mode 25

Representation of κ:

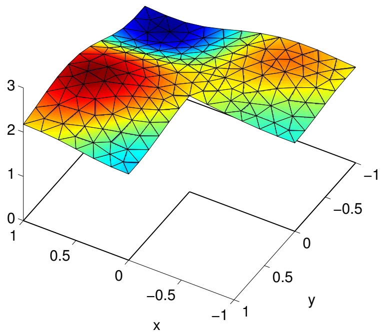
$$\kappa(x,\omega) = \bar{\kappa}(x) + \sum_{\jmath=1}^{\infty} \varkappa_\jmath \, g_\jmath(x)\xi_\jmath(\omega) =: \sum_{\jmath=0}^{\infty} \varkappa_\jmath \, g_\jmath(x)\xi_\jmath(\omega)$$

with centred, normalised, uncorrelated random variables $\xi_\jmath(\omega)$:

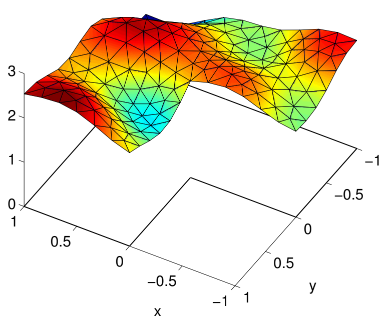
$$\mathbb{E}\left(\xi_\jmath\right) = 0, \quad \mathbb{E}\left(\xi_\jmath \xi_k\right) =: \langle \xi_\jmath, \xi_k \rangle_{L_2(\Omega)} = \delta_{\jmath k}.$$

Karhunen-Loève Expansion III

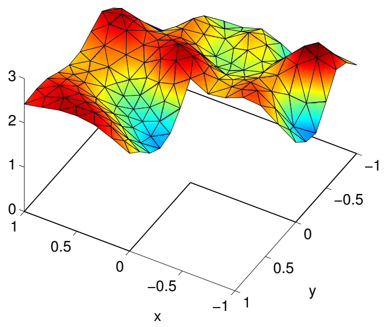
Realisation with:



6 modes



15 modes



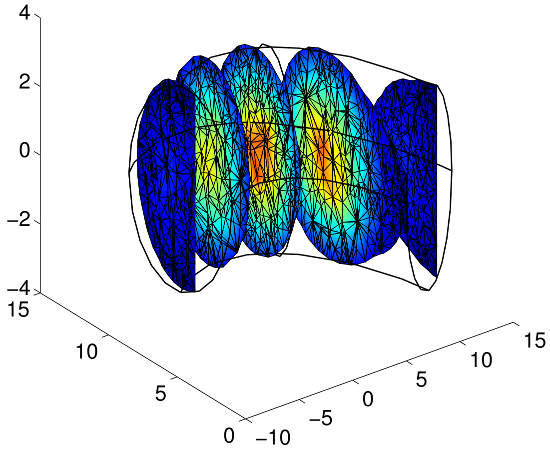
40 modes

Truncate after m largest eigenvalues

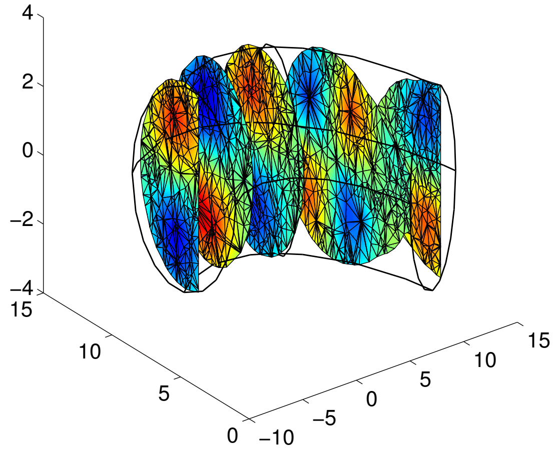
$\Rightarrow$ optimal—in variance—expansion in m RVs.

Karhunen-Loève Expansion IV

Modes for a 3-D domain



mode 1



mode 19

Karhunen-Loève Expansion V

Reminder: SVD of a matrix $\boldsymbol{W} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \sum_\jmath \sigma_\jmath \boldsymbol{u}_\jmath \boldsymbol{v}_\jmath^T$ with $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}$, $\boldsymbol{V}^T\boldsymbol{V} = \boldsymbol{I}$, and $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_\jmath)$. The $\sigma_\jmath$ are singular values of $\boldsymbol{W}$ and $\sigma_\jmath^2$ are eigenvalues of $\boldsymbol{W}^T\boldsymbol{W}$ or $\boldsymbol{W}\boldsymbol{W}^T$.

To every random field $w(x, \omega) \in L_2(\mathcal{G}) \otimes L_2(\Omega)$ associate a linear map $W : L_2(\mathcal{G}) \to L_2(\Omega)$

$$W : L_2(\mathcal{G}) \ni v \mapsto W(v)(\omega) = \langle v(\cdot), w(\cdot, \omega)\rangle_{L_2(\mathcal{G})} = \int_{\mathcal{G}} v(x) w(x, \omega)\, dx \in L_2(\Omega).$$

KLE is SVD of the map W, the covariance operator is $C_w := W^* W$,

Karhunen-Loève Expansion VI

$$\langle u, C_w v\rangle_{L_2(\mathcal{G})} = \langle u, W^* W v\rangle_{L_2(\mathcal{G})} = \langle W(u), W(v)\rangle_{L_2(\Omega)} =$$

$$\mathbb{E}\left(W(u)W(v)\right) = \mathbb{E}\left(\int_{\mathcal{G}} u(x)w(x,\omega)\,dx \int_{\mathcal{G}} v(y)w(y,\omega)\,dy\right) =$$

$$\int_{\mathcal{G}}\int_{\mathcal{G}} u(x)\mathbb{E}\left(w(x,\omega)w(y,\omega)\right)v(y)\,dy\,dx =$$

$$\int_{\mathcal{G}} u(x) \int_{\mathcal{G}} \mathsf{C}_w(x,y)v(y)\,dy\,dx$$

Covariance operator C_w is represented by covariance kernel $\mathsf{C}_w(x,y)$.

Truncating the KLE is therefore the same as what is done when truncating a SVD, finding a sparse representation (model reduction).

First Summary

- Motivation, Probabiliy, aleatoric and epistemic Uncertainty
- Formulation as a well-posed problem
- RVs, Stochastic Processes and Random Fields
- Spectral Expansion, Karhunen-Loève Expansion
- Still open:
  - How to discretise RVs ?
  - How to actually compute $u(\omega)$?
  - How to perform integration ?