Methods for Uncertainty Propagation and Quantification in Numerical Models

Olivier Le Maître

1LIMSI, CNRS
UPR-3251, Orsay, France
http://www.limsi.fr/Individu/olm
Methods for Uncertainty Propagation and Quantification in Numerical Models

Olivier Le Maître

1LIMSI, CNRS
UPR-3251, Orsay, France
http://www.limsi.fr/Individu/olm

Simurex 2012
Materials of this lecture are taken from the book: "Spectral Methods for Uncertainty Quantification with applications in computational fluid dynamics" with Omar Knio, Springer (2010).
Simulation framework.

**Basic ingredients**

- Selection of a **mathematical model**:
  retain essential physical processes.

- Selection of a **numerical method**:
  to solve the model equations.

- Define all **input-data** needed:
  select a specific system in the class spanned by the model.
Simulation and errors

**Simulation framework.**

### Basic ingredients

- **Selection of a mathematical model:** retain essential physical processes.
- **Selection of a numerical method:** to solve the model equations.
- **Define all input-data needed:** select a specific system in the class spanned by the model.

### Simulation errors

- **Model errors:** physical approximations and simplifications.
- **Numerical errors:** discretization, approximate solvers, finite arithmetics, . . .
- **Input-data error:** boundary/initial conditions, model constants and parameters, external forcings, . . .
Sources of data uncertainty

- Inherent **variability** (e.g. industrial processes).
- **Epistemic** uncertainty (e.g. model constants).
- May not be fully reducible, even theoretically.
Sources of data uncertainty

- Inherent variability (e.g. industrial processes).
- Epistemic uncertainty (e.g. model constants).
- May not be fully reducible, even theoretically.

Probabilistic framework

- Define an abstract probability space $(\Omega, \mathcal{A}, d\mu)$.
- Consider input-data $D$ as random quantity: $D(\omega)$, $\omega \in \Omega$.
- Simulation output $S$ is random and on $(\Omega, \mathcal{A}, d\mu)$. 
Input-data uncertainty

### Sources of data uncertainty

- **Inherent variability** (e.g. industrial processes).
- **Epistemic uncertainty** (e.g. model constants).
- **May not be fully reducible, even theoretically.**

### Probabilistic framework

- Define an abstract probability space \((\Omega, \mathcal{A}, d\mu)\).
- Consider input-data \(D\) as random quantity: \(D(\omega), \omega \in \Omega\).
- Simulation output \(S\) is random and on \((\Omega, \mathcal{A}, d\mu)\).
- Data \(D\) and simulation output \(S\) are **dependent** random quantities (through the mathematical model \(M\)):

\[
M(S(\omega), D(\omega)) = 0, \quad \forall \omega \in \Omega.
\]
Introduction

Spectral UQ

Solution methods

Application to conduction

Application to natural convection

Conclusion

Input-data uncertainty

**Propagation of data uncertainty**

Data density

Solution density

\[ \mathcal{M}(S, D) = 0 \]
Introduction

Spectral UQ

Solution methods

Application to conduction

Application to natural convection

Conclusion

Input-data uncertainty

Propagation of data uncertainty

Data density

\[ M(S, D) = 0 \]

Solution density

- Variability in model output: numerical error bars.
- Assessment of predictability.
- Support decision making process.
- What type of information (abstract quantities, confidence intervals, density estimations, structure of dependencies, ...) one needs?
Alternative UQ methods

**Deterministic methods**

- **Sensitivity analysis** (adjoint based, AD, …) : local.
- **Perturbation techniques** : limited to low order and simple data uncertainty.
- **Neumann expansions** : limited to low expansion order.
- **Moments** method : closure problem (non-Gaussian / non-linear problems).

**Simulation techniques**

- Monte-Carlo

**Spectral Methods**
Deterministic methods

Simulation techniques

Monte-Carlo

- Generate a sample set of data realizations and compute the corresponding sample set of model output.

- Use sample set based random estimates of abstract characterizations (moments, correlations, ...).

- Plus: Very robust and re-use deterministic codes: (parallelization, complex data uncertainty).

- Minus: slow convergence of the random estimates with the sample set dimension.

Spectral Methods
Deterministic methods

Simulation techniques

Spectral Methods

- Parameterization of the data with random variables (RVs).
- Projection of solution on the ($L^2$) space spanned by the RVs.
- Plus: arbitrary level of uncertainty, deterministic approach, convergence rate, information contained.
- Minus: parameterizations (limited # of RVs), adaptation of simulation tools (legacy codes), robustness (non-linear problems, non-smooth output, ...).
- Not suited for model uncertainty
1 Introduction
   - Simulation and errors
   - Input-data uncertainty
   - Alternative UQ methods

2 Spectral UQ
   - Polynomial Chaos expansions
   - Application to spectral UQ

3 Solution methods
   - Non-Intrusive methods
   - Galerkin projection

4 Application to conduction
   - Stochastic conduction problem
   - Heat exchanger

5 Application to natural convection
   - Boussinesq equations
   - Results

6 Conclusion
Polynomial Chaos expansions

Any well behaved RV $U(\omega)$ (e.g. 2nd-order one) defined on $(\Omega, \mathcal{A}, d\mu)$ has a convergent expansion of the form:

$$U(\omega) = u_0 \Gamma_0 + \sum_{i_1=1}^{\infty} u_{i_1} \Gamma_1(\xi_{i_1}(\omega)) + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} u_{i_1,i_2} \Gamma_2(\xi_{i_1}(\omega), \xi_{i_2}(\omega)) + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} u_{i_1,i_2,i_3} \Gamma_3(\xi_{i_1}(\omega), \xi_{i_2}(\omega), \xi_{i_3}(\omega)) + \ldots$$

- $\{\xi_1, \xi_2, \ldots\}$: independent normalized Gaussian RVs.
- $\Gamma_p$ polynomials with degree $p$, orthogonal to $\Gamma_q, \forall q < p$.
- Convergence in the mean square sense [Cameron & Martin, 1947].

[Wiener, 1938]
Polynomial Chaos expansions

Polynomial Chaos expansions

Truncated PC expansion at order \( N_0 \) and using \( N \) RVs:

\[
U(\omega) \approx \sum_{k=0}^{P} u_k \psi_k(\xi(\omega)), \quad \xi = \{\xi_1, \ldots, \xi_N\}, \quad P = \frac{(N + N_0)!}{N!N_0!}.
\]

- \( \{u_k\}_{k=0,...,P} \): deterministic expansion coefficients,
- \( \{\psi_k\}_{k=0,...,P} \): \( \perp \) random polynomials wrt the inner product involving the density of \( \xi \):

\[
\mathbb{E} \{\psi_k \psi_l\} = \langle \psi_k, \psi_l \rangle = \int_{\Omega} \psi_k(\xi(\omega))\psi_l(\xi(\omega))d\mu(\omega)
= \int \psi_k(\xi)\psi_l(\xi)\rho(\xi)d\xi = \delta_{kl} \langle \psi_k, \psi_k \rangle.
\]
Polynomial Chaos expansions

Truncated PC expansion at order $N_0$ and using $N$ RVs:

$$U(\omega) \approx \sum_{k=0}^{P} u_k \Psi_k(\xi(\omega)), \quad \xi = \{\xi_1, \ldots, \xi_N\}, \quad P = \frac{(N + N_0)!}{N!N_0!}.$$

- $\{u_k\}_{k=0,...,P}$: deterministic expansion coefficients,
- $\{\Psi_k\}_{k=0,...,P}$: $\perp$ random polynomials wrt the inner product involving the density of $\xi$:

$$\mathbb{E} \{\Psi_k \Psi_l\} = \langle \Psi_k, \Psi_l \rangle \equiv \int_{\Omega} \Psi_k(\xi(\omega))\Psi_l(\xi(\omega))d\mu(\omega)$$

$$= \int_{\Omega} \Psi_k(\xi)\Psi_l(\xi)\rho(\xi)d\xi = \delta_{kl} \langle \Psi_k, \Psi_k \rangle.$$

- Gaussian RVs: $\rho(\xi) = \prod_{i=1}^{N} \frac{\exp(-\xi_i^2/2)}{\sqrt{2\pi}} \quad \rightarrow \text{Hermite polynomials (Wiener-Hermite expansions)}$
- $\{\Psi_0, \ldots, \Psi_P\}$ is an orthogonal basis of $S^P \subset L^2(\mathbb{R}^N, \rho(\xi))$. 

[Wiener, 1938]
Polynomial Chaos expansions

Truncated PC expansion:

\[ U(\omega) \approx \sum_{k=0}^{P} u_k \Psi_k(\xi(\omega)). \]

- Convention \( \Psi_0 \equiv 1 \): mean mode.

- Expectation of \( U \):

\[ E\{U\} \equiv \int_{\Omega} U(\omega) d\mu(\omega) \approx \sum_{k=0}^{P} u_k \int_{\Xi} \Psi_k(\xi)p(\xi) d\xi = u_0. \]

- Variance of \( U \):

\[ V[U] = E\{U^2\} - E\{U\}^2 \approx \sum_{k=1}^{P} u_k^2 \langle \Psi_k, \Psi_k \rangle. \]

- Extension to random vectors & stochastic processes:

\[
\begin{pmatrix}
U_1 \\
\vdots \\
U_m
\end{pmatrix}
(\omega, \mathbf{x}, t) \approx \sum_{k=0}^{P} \begin{pmatrix}
\vdots \\
u_1 \\
\vdots \\
u_m
\end{pmatrix}
(\mathbf{x}, t) \Psi_k(\xi(\omega)).
\]
Generalized PC expansion

[Xiu & Karniadakis, 2002]

Askey scheme

| Distribution of $\xi_i$ | Polynomial family |
|------------------------|-------------------|
| Gaussian               | Hermite           |
| Uniform                | Legendre          |
| Exponential            | Laguerre          |
| $\beta$-distribution   | Jacobi            |

Also: discrete RVs (Poisson process).

\[ U(\omega) \approx \sum_{k=0}^{P} u_k \psi_k(\xi(\omega)) \]

where $\psi_k$: classical (or product of) polynomials: spectral expansions
Instead of a spectral expansion over $\Xi$ one can use **Piecewise polynomial expansion** on a mesh $\Sigma$ of $\Xi$

- $\Xi = \bigcup_{SE \in \Sigma} \Xi_{SE}$, $\Xi_{SE} \cap \Xi_{SE'} = \emptyset$ for $SE \neq SE'$
- $S = \left\{ U \in L^2(\Xi, p_\xi), U(\xi \in \Xi_{SE}) \in P^N_{No}(\Xi_{SE}) \right\}$
- $U(\omega) \approx \sum_{k=0}^{P} u_k \psi_k(\xi(\omega))$

- $\psi_k$ are orthogonal with:
  1. **Support of $\psi_k$ limited to an element**: **Stochastic multi-element method** [Deb et al., 2001], [Wang and Karniadakis, 2005]
    - Fully decouple the approximation problem for $\neq$ elements
  2. **Hierarchical orthogonal $\psi_k$**: **Stochastic Multiwavelet method** [olm et al., 2004, 2006, 2009]
    - Couple the approximation problem, well suited for adaptive strategy (MRA)
Application to spectral UQ

Input-data parametrization

Parametrization of $D$ using $N < \infty$ independent RVs with prescribed distribution $p(\xi) :$

$$D(\omega) \approx D(\xi(\omega)), \quad \xi = (\xi_1, \ldots, \xi_N) \in \Xi.$$  

- Iso-probabilistic transformations of RVs,
- Karhunen-Loève expansion: $D(x, \omega)$ stochastic field/process,
- Identification (e.g. Bayesian).

Model

Solution expansion
Input-data parametrization

Model

We assume that for a.e. $\xi \in \Xi$, the problem $M(S, D(\xi)) = 0$

1. is well-posed,
2. has a unique solution

and that

the random solution $S(\xi) \in L^2(\Xi, p_\xi)$:

$$
\mathbb{E} \{ S^2 \} = \int_\Omega S^2(\xi(\omega)) d\mu(\omega) = \int_{\Xi} S^2(\xi)p(\xi) d\xi < +\infty.
$$

Solution expansion
Input-data parametrization

Model

Solution expansion

Let \( \{\psi_0, \psi_1, \ldots\} \) be a basis of \( L^2(\Xi, p_\xi) \) then

\[
S(\xi) = \sum_k s_k \psi_k(\xi).
\]

Knowledge of the spectral coefficients \( s_k \) fully determine the random solution.

Makes explicit the dependence between \( D(\xi) \) and \( S(\xi) \).
Input-data parametrization

Model

Solution expansion

Let \( \{\psi_0, \psi_1, \ldots\} \) be a basis of \( L^2(\Xi, p_\xi) \) then

\[
S(\xi) = \sum_k s_k \psi_k(\xi).
\]

- Knowledge of the spectral coefficients \( s_k \) fully determine the random solution.
- Makes explicit the dependence between \( D(\xi) \) and \( S(\xi) \).
- Need efficient procedure(s) to compute the \( s_k \).

[Ghanem & Spanos, 1991]
Non-Intrusive methods

- Compute/estimate spectral coefficients via a set of deterministic model solutions
- Requires a deterministic solver only
- Overcome issues related to non-linearities.
- Suffers from the curse of dimensionality
Non-Intrusive methods

Use code as a black-box

- Compute/estimate spectral coefficients via a set of deterministic model solutions
- Requires a deterministic solver only

1. $S_{\xi} \equiv \{\xi^{(1)}, \ldots, \xi^{(m)}\}$ sample set of $\xi$
2. Let $s^{(i)}$ be the solution of the deterministic problem
   \[ M \left( s^{(i)}, D(\xi^{(i)}) \right) = 0 \]
3. $S_S \equiv \{s^{(1)}, \ldots, s^{(m)}\}$ sample set of model solutions
4. Estimate expansion coefficients $s_k$ from this sample set.

- Complex models, reuse of determinisitic codes, planification, . . .
- Error control and computational complexity (curse of dimensionality), . . .
Least square fit

- Best approximation is defined by minimizing a (weighted) sum of squares of residuals:

\[ R^2(s_0, \ldots, s_P) \equiv \sum_{i=1}^{m} w_i \left( s^{(i)} - \sum_{k=0}^{P} s_k \psi_k (\xi^{(i)}) \right)^2. \]

Advantages/issues

- Convergence with number of regression points \( m \)
- Selection of the regression points and “regressors” \( \psi_k \)
- Error estimate
Non-intrusive spectral projection: NISP

Exploit the orthogonality of the basis:

\[
\mathbb{E} \{ \psi_k^2 \} s_k = \langle S, \psi_k \rangle = \int_{\Xi} S(\xi) \psi_k(\xi) \rho(\xi) d\xi.
\]

Computation of \((P+1)\) N-dimensional integrals

\[
\langle S, \psi_k \rangle \approx \sum_{i=1}^{N_Q} w^{(i)} S(\xi^{(i)}) \psi_k(\xi^{(i)}).
\]
Non intrusive projection

Approximate integrals from a (pseudo) random sample set $S_S$:

$$\langle S, \psi_k \rangle \approx \frac{1}{m} \sum_{i=1}^{m} w^{(i)} s^{(i)} \psi_k \left( \xi^{(i)} \right).$$

| MC | LHS | QMC |
|----|-----|-----|
| ![MC samples](image1.png) | ![LHS samples](image2.png) | ![QMC samples](image3.png) |

- Convergence rate
- Error estimate
- Optimal sampling strategy
Non intrusive projection

Approximate integrals by $N$-dimensional quadratures:

$$\langle S, \psi_k \rangle \approx \sum_{i=1}^{N_Q} w^{(i)} s^{(i)} \psi_k \left( \xi^{(i)} \right).$$

Quadrature points $\xi^{(i)}$ and weights $w^{(i)}$ obtained by

- full tensorization of $n$ points 1-D quadrature (i.e. Gauss):
  $$N_Q = n^N$$

- partial tensorization of nested 1-D quadrature formula (Féjer, Clenshaw-Curtis):
  $$N_Q << n^N$$
Non-Intrusive methods

### Non intrusive projection

| $l = 4$ | $l = 5$ | $l = 6$ |
|---------|---------|---------|
| ![Diagram](image1.png) | ![Diagram](image2.png) | ![Diagram](image3.png) |

### Deterministic Quadratures

| $l = 4$ | $l = 5$ | $l = 6$ |
|---------|---------|---------|
| ![Diagram](image4.png) | ![Diagram](image5.png) | ![Diagram](image6.png) |

See other lecture(s)

- Important development of sparse-grid methods
- Anisotropic and adaptivity
- Extension to collocation approach (N-dimensional interpolation)
**Galerkin projection**

- Weak solution of the stochastic problem $\mathcal{M}(S, D) = 0$
- Needs adaptation of deterministic codes
- Potentially more efficient than NI techniques.
- Better suited to improvement (error estimate, optimal and basis reduction, . . .), thanks to functional analysis.
Galerkin projection

Method of weighted residual

1. Introduce truncated expansions in model equations
2. Require residual to be $\perp$ to the stochastic subspace $S^p$

$$\left\langle \mathcal{M} \left( \sum_{k=0}^{P} s_k \psi_k(\xi), D(\xi) \right), \psi_m(\xi) \right\rangle = 0 \quad \text{for } m = 0, \ldots, P.$$ 

Set of $P + 1$ coupled problems.

**Plus**
- Implicitly account for modes’ coupling
- Often inherit properties of the deterministic model

**Minus**
- Requires adaptation of deterministic solvers
- Treatment of non-linearities.
1 Introduction
   - Simulation and errors
   - Input-data uncertainty
   - Alternative UQ methods

2 Spectral UQ
   - Polynomial Chaos expansions
   - Application to spectral UQ

3 Solution methods
   - Non-Intrusive methods
   - Galerkin projection

4 Application to conduction
   - Stochastic conduction problem
   - Heat exchanger

5 Application to natural convection
   - Boussinesq equations
   - Results

6 Conclusion
Consider the linear steady heat equation in an isotropic two-dimensional domain $\Omega$, with boundary $\partial \Omega$.

$x \in \Omega \mapsto u(x) \in \mathbb{R}$ is the temperature field satisfying:

$$\nabla \cdot (\nu(x) \nabla u(x)) = -f(x) + BC$$

where $\nu > 0$ is the thermal conductivity and $f \in L_2(\Omega)$ is a source term.

We consider homogeneous Dirichlet and Neumann conditions over the respective portions $\Gamma_d$ and $\Gamma_n$ of the domain boundary $\partial \Omega = \Gamma_d \cup \Gamma_n$, i.e.

$$u(x) = 0, \quad x \in \Gamma_d \quad \frac{\partial u}{\partial n} = 0, \quad x \in \Gamma_n.$$
Let $\mathcal{V}$ be the functionals space on $\Omega$ such that:

$$\mathcal{V} = \{ u \in H^1(\Omega) : u = 0 \text{ on } \Gamma_d \},$$

where $H^1(\Omega)$ is the Sobolev space of square integrable functionals whose first order derivatives are also square integrable.

The variational problem is:

Find $u \in \mathcal{V}$ such that

$$a(u, v) = b(v) \quad \forall v \in \mathcal{V},$$

where $a(u, v)$ and $b(v)$ are bilinear and linear forms respectively defined as:

$$a(u, v) \equiv \int_{\Omega} \nu(x) \nabla u(x) \cdot \nabla v(x) dx, \quad b(v) \equiv \int_{\Omega} f(x) v(x) dx.$$
Stochastic conduction problem

Let \( T = \{ \Sigma_1, \ldots, \Sigma_{ne} \} \) be a triangulation of \( \Omega \) with \( ne \) non-overlapping triangular elements \( \Sigma_i \). The \( P - 1 \) finite element space \( \mathcal{V}^h \) consists in linear functions in each \( \Sigma_i \), that are continuous across inter-element boundaries. A function \( \nu \in \mathcal{V}^h \) is completely defined by its values at the mesh nodes, and \( \nu \) can be expressed as

\[
\nu^h(x) = \sum_{i \in \mathcal{N}} v_i^h \Phi_i(x),
\]

where \( \mathcal{N} \) is the set of nodes which are not lying on \( \Gamma_d \) and \( \Phi_i(x) \) are the shape functions associated to these nodes.

\[
\mathcal{V}^h = \text{span} \{ \Phi_i \}_{i \in \mathcal{N}}.
\]
The Galerkin formulation in $\mathcal{V}^h$ is:
Find $u_i, i \in \mathcal{N}$ such that
\[
\sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} a_{i,j} u_i v_j = \sum_{j \in \mathcal{N}} b_j v_j,
\]
where
\[
a_{i,j} = \int_\Omega \nu(x) \nabla \Phi_i(x) \cdot \nabla \Phi_j(x) \, dx, \quad b_i = \int_\Omega f(x) \Phi_i(x) \, dx.
\]
The problem can be recast as a system of linear equations
\[
\begin{pmatrix}
  a_{1,1} & \ldots & a_{1,n} \\
  \vdots & \ddots & \vdots \\
  a_{n,1} & \ldots & a_{n,n}
\end{pmatrix}
\begin{pmatrix}
  u_1 \\
  \vdots \\
  u_n
\end{pmatrix}
= \begin{pmatrix}
  b_1 \\
  \vdots \\
  b_n
\end{pmatrix},
\]
where $n = \text{Card}(\mathcal{N})$. \([a]\) is a (sparse) symmetric positive definite matrix.
Stochastic conduction problem

Left: sketch of the domain $\Omega$ and decomposition of the boundary $\partial\Omega$ into Dirichlet $\Gamma_d$ and Neumann $\Gamma_n$ regions. Right: example of a finite-element mesh with 508 elements and 284 nodes.
Consider the case of random conductivity and source term, defined on an abstract probability space \((\Theta, \Sigma, P)\):

\[
\nu \rightarrow \nu(x, \theta), \quad f \rightarrow F(x, \theta).
\]

Then, \(u \rightarrow U(x, \theta)\) satisfies almost surely the stochastic problem

\[
\begin{aligned}
\nabla \cdot (\nu(x, \theta) \nabla U(x, \theta)) &= -F(x, \theta) & \quad x &\in \Omega \\
U(x, \theta) &= 0 & \quad x &\in \Gamma_d, \\
\frac{\partial U(x, \theta)}{\partial n} &= 0 & \quad x &\in \Gamma_n.
\end{aligned}
\]
The functional space for $U(x, \theta)$ will be $\mathcal{V} \otimes L_2(\Theta, P)$. In other words,

$$U(\cdot, \theta) \in \mathcal{V}, \quad U(x, \cdot) \in L_2(\Theta, P),$$

The variational form of the stochastic problem is:
Find $U \in \mathcal{V} \otimes L_2(\Theta, P)$ such that

$$A(U, V) = B(V) \quad \forall V \in \mathcal{V} \otimes L_2(\Theta, P),$$

where

$$A(U, V) \equiv \mathbb{E}\{a(U, V)\} = \int_{\Theta} \left[ \int_{\Omega} \nu(x, \theta) \nabla U(x, \theta) \cdot \nabla V(x, \theta) \, dx \right] \, dP(\theta),$$

and

$$B(V) \equiv \mathbb{E}\{b(V)\} = \int_{\Theta} \left[ \int_{\Omega} F(x, \theta) V(x, \theta) \, dx \right] \, dP(\theta).$$
Stochastic conduction problem

introducing the deterministic discretization in $\mathcal{V}^h$ it comes

$$U^h(x, \theta) = \sum_{i \in \mathcal{N}} U_i(\theta) \Phi_i(x) \in (\mathcal{V}^h \otimes L_2(\Theta, P)) .$$

It shows that the semi-discrete solution consists in $n = \text{Card}(\mathcal{N})$ random variables $U_i(\theta)$. They satisfy

$$\sum \sum \mathbb{E} \{A_{i,j}(\theta) U_i(\theta) V_j(\theta)\} = \sum \mathbb{E} \{B_i(\theta) V_i(\theta)\}, \ \forall V_i(\theta) \in L_2(\Theta, P), i \in \mathcal{N},$$

where

$$A_{i,j}(\theta) = \int_{\Omega} \nu(x, \theta) \nabla \Phi_i(x) \cdot \nabla \Phi_j(x) dx,$$

and

$$B_i(\theta) = \int_{\Omega} f(x, \theta) \Phi_i(x) dx.$$
We assume $\nu$ and $F$ parameterized with $N$ independent r.v. $\xi = \{\xi_1 \cdots \xi_N\}$ defined on $(\Theta, \Sigma, P)$:

$$
\nu(x, \theta) = \nu(x, \xi(\theta)), \quad F(x, \theta) = F(x, \xi(\theta)).
$$

Examples of parameterization will be shown later. The space of second-order random functionals in $\xi$ is spanned by the Polynomial Chaos basis:

$$
S = \text{span}\{\psi_k(\xi)\}_{k=0}^{k=\infty} = L_2(\mathbb{R}^2, \rho_\xi),
$$

where the $\psi_i$'s form a set of orthogonal multidimensional polynomials in $\xi$:

$$
\langle \psi_i, \psi_j \rangle = \int_{\Xi} \psi_i(\eta)\psi_j(\eta)\rho_\xi(\eta)d\eta = \delta_{ij} \langle \psi_i^2 \rangle.
$$

Provided that $\nu$ and $F$ are second-order quantities, they have orthogonal representations:

$$
\nu(x, \xi) = \sum_{k=0}^{\infty} \nu_k(x)\psi_k(\xi), \quad F(x, \xi) = \sum_{k=0}^{\infty} f_k(x)\psi_k(\xi).
$$
Similarly, the expansion of the discrete solution $U^h$ is

$$U^h(x, \xi) = \sum_{i \in \mathcal{N}} \left( \sum_{k=0}^{\infty} u_{i,k} \psi_k(\xi) \right) \Phi_i(x).$$

The stochastic expansions are truncated to a finite polynomial order $N_0$. Different orders of truncation may be considered for the conductivity, source and solution. For simplicity, we use the same truncation order $N_0$. It corresponds to a stochastic approximation space

$$S^p \equiv \text{span}\{\psi_0, \ldots, \psi_p\} \subset S, \quad P + 1 = \frac{(N_0 + N)!}{N_0!N!}.$$
The Galerkin problem is obtained by inserting the expansions of $\nu$, $F$, $U^h$ and test functions $V \in \mathcal{V}^h \otimes S^p$ into the variational form of the semi discrete stochastic problem. This results in:

Find $u_{i,k}$, $i \in \mathcal{N}$ and $k = 0, \ldots, P$, such that

$$\sum_{i,j \in \mathcal{N}} \sum_{k,l,m=0}^{P} \langle \Psi_k \Psi_l \Psi_m \rangle A_{i,j}^{k} u_{i,l} v_{j,m} = \sum_{i \in \mathcal{N}} \sum_{k=0}^{P} b_{i}^{k} v_{i,k},$$

$\forall v_{i,k}$, $i \in \mathcal{N}$, $k = 0, \ldots, P$

where

$$A_{i,j}^{k} \equiv \int_{\Omega} \nu_{k}(x) \nabla \Phi_{i}(x) \cdot \nabla \Phi_{j}(x) dx, \quad b_{i}^{k} \equiv \langle \Psi_{k}^{2} \rangle \int_{\Omega} f_{k}(x) \Phi_{i}(x) dx.$$

It involves $n \times (P + 1)$ deterministic quantities.
Denote \( \mathbf{u}_k := (u_{1,k} \ldots u_{n,k})^t \in \mathbb{R}^n \) the vector of nodal values of the \( k \)-th stochastic mode of the solution.

With this notation, the Galerkin problem becomes:

Find \( \mathbf{u}_0, \ldots, \mathbf{u}_P \) such that for all \( k = 0, \ldots, P \)

\[
\sum_{l=0}^{P} \sum_{m=0}^{P} \langle \psi_k \psi_l \psi_m \rangle \left[ A^l \right] \mathbf{u}_m = \mathbf{b}_k,
\]

where the matrix \( \left[ A^l \right] \) has for coefficients \( A_{i,j}^l \) and the vector \( \mathbf{b}_k = (b_1^k \ldots b_n^k)^t \).
Denote \( u_k := (u_{1,k} \ldots u_{n,k})^t \in \mathbb{R}^n \) the vector of nodal values of the \( k \)-th stochastic mode of the solution.

This set of systems can be formally expressed as a single system \([A]u = B\) where the global system matrix \([A]\) has the block structure, corresponding to:

\[
\begin{pmatrix}
A_{0,0} & \cdots & A_{0,P} \\
\vdots & \ddots & \vdots \\
A_{P,0} & \cdots & A_{P,P}
\end{pmatrix}
\begin{pmatrix}
u_0 \\
\vdots \\
u_P
\end{pmatrix} =
\begin{pmatrix}
b_0 \\
\vdots \\
b_P
\end{pmatrix}.
\]

The matrix blocks are given by:

\[
A_{i,j} = \sum_{m=0}^{P} [A^m] \langle \psi_i, \psi_j, \psi_m \rangle \quad 0 \leq i, j \leq P.
\]

The system \([A]u = B\) is called the spectral or Galerkin problem.
Solution method:

- The matrix $[A]$ of the Galerkin problem has a block symmetric structure, $A_{i,j} = A_{j,i}$, since $\langle \psi_i \psi_j \psi_m \rangle = \langle \psi_j \psi_i \psi_m \rangle$.

- The blocks are in fact symmetric because $A_{i,j}^k = A_{j,i}^k$, so the matrix $[A]$ is symmetric.

- Standard solution techniques for (large) symmetric linear systems can be reused.

- Due to the size of the system, sparse storage is mandatory, even-though many blocks are zero.
Existence and uniqueness of the solution:
Properties of the Galerkin system have been the focus of many works. e.g. [Babuška, 2002], [Babuška, 2005], [Frauenfelder, 2005], [Matthies, 2005]

- For Dirichlet boundary conditions, the Galerkin system for stochastic elliptic problems has a unique solution provided that the random conductivity field satisfies some probabilistic (sufficient) conditions.

- For the deterministic discretization with $P - 1$ finite-elements, these probabilistic conditions reduce to

$$
\frac{1}{\nu(x, \xi)} \in L_2(\Xi, P_\Xi), \forall x \in \Omega
$$

- For Neumann boundary conditions only, $U(x, \xi)$ is defined up to an arbitrary random variable and an integral constraint on the source term is necessary for homogeneous conditions,

$$
\int_\Omega F(x, \xi)dx = 0 \quad a.s.
$$
We consider $\Omega = [0, 1]^2$, with Dirichlet boundary conditions over 3 edges and a Neumann condition over the left edge $x = 1$.

Left : computational domain $\Omega$ and decomposition of the boundary $\partial\Omega$ into Dirichlet $\Gamma_d$ and Neumann $\Gamma_n$ parts.

Right : typical finite-element triangulation of $\Omega$ using 512 elements and 289 nodes.
Consider first the case of a uniform deterministic source term and constant random conductivity

\[ F(x, \theta) = f(x) = 1, \quad \nu(x, \theta) = \beta(\theta). \]

- The random conductivity \( \beta \) is assumed to be log-normal, with unit median value \( \bar{\beta} = 1 \) and coefficient of variation \( C \geq 1 \).
- \( \beta \) is parametrized with a unique normalized Gaussian variable \( \xi_1(\theta) \) so \( N = 1 \), and the PC basis is made of the one-dimensional Hermite polynomials.

\[ \beta(\xi_1) = \exp(\mu_\beta + \sigma_\beta \xi_1), \quad \mu_\beta = \log(\bar{\beta}) \text{ and } \sigma_\beta = \frac{\log C}{2.85}. \]

- The PC coefficients \( \beta_k \) have closed form expressions [Ghanem, 1999]:

\[ \beta(\xi_1) = \sum_{k=0}^{\infty} \beta_k \psi_k(\xi_1), \quad \beta_k = \exp(\mu_\beta + \sigma_\beta^2/2) \frac{\sigma_\beta^k}{\langle \psi_k^2 \rangle}. \]
### Stochastic modes of the solution for $\text{No} = 4$

| $U_h^0$ | $U_h^1$ | $U_h^2$ |
|---------|---------|---------|
| $u_h^k$ | $u_h^k$ | $u_h^k$ |

- $u_h^k$ values for $U_h^0$: 
  - $0.12$, $0.08$, $0.04$, $0.00$, $0.00$...
  - $-0.05$, $-0.04$, $-0.03$, $-0.02$, $-0.01$...

- $u_h^k$ values for $U_h^1$: 
  - $0.12$, $0.08$, $0.04$, $0.00$, $0.00$...
  - $-0.05$, $-0.04$, $-0.03$, $-0.02$, $-0.01$...

- $u_h^k$ values for $U_h^2$: 
  - $0.12$, $0.08$, $0.04$, $0.00$, $0.00$...
  - $-0.05$, $-0.04$, $-0.03$, $-0.02$, $-0.01$...

| $U_h^3$ | $U_h^4$ | Standard deviation |
|---------|---------|-------------------|
| $u_h^k$ | $u_h^k$ | STD of $u_h^k$ |

- $u_h^k$ values for $U_h^3$: 
  - $-4.0e-4$, $-8.0e-4$, $1.2e-3$, $0.00$, $0.00$...
  - $-2.0e-05$, $-4.0e-05$, $6.0e-05$, $8.0e-05$...

- $u_h^k$ values for $U_h^4$: 
  - $-4.0e-4$, $-8.0e-4$, $1.2e-3$, $0.00$, $0.00$...
  - $-2.0e-05$, $-4.0e-05$, $6.0e-05$, $8.0e-05$...

- STD of $u_h$: 
  - $0.06$, $0.04$, $0.02$, $0.00$, $0.00$...
  - $0.06$, $0.04$, $0.02$, $0.00$, $0.00$...
  - $0.06$, $0.04$, $0.02$, $0.00$, $0.00$...
Convergence with the expansion order

$\beta$ being log-normal, so is its inverse, and the expansion of $1/\beta$ is consequently given by:

\[
\left(\frac{1}{\beta}\right)_k = \exp\left(-\mu_\beta + \frac{\sigma_\beta^2}{2}\right) \frac{(-\sigma_\beta)^k}{\langle \psi_k^2 \rangle}.
\]

The spectrum of the numerical solution should decay as $|\sigma_\beta|^k/k!$.

Normalized spectra of the random solution $u_h^k$ at node $x = (1, 0.5)$ as computed using different expansion orders.
Computed probability density functions of $U^h$ at $x = (1, 0.5)$ for different expansion orders $N_o$ as indicated. Top plot: $N_o = 1, \ldots, 6$. Bottom plot: same pdfs in log scale for $N_o = 2, \ldots, 6$ together with the theoretical pdf.
Consider the random conductivity field defined as:

\[ \nu(x, \theta) = \begin{cases} 
\nu^1(\theta), & x \leq 0.5 \\
\nu^2(\theta), & x > 0.5 
\end{cases} \]

- \( \nu^1 \) and \( \nu^2 \) are two independent log-normal random variables with respective medians \( \bar{\nu}^1 \) and \( \bar{\nu}^2 \), and coefficients of variation \( C^1 \) and \( C^2 \).

- Two normalized Gaussian variables \( \xi_1 \) and \( \xi_2 \) are used to parametrize the conductivity field.

- The stochastic dimension is \( N = 2 \), and the stochastic basis consists of two-dimensional Hermite polynomials.
The expansion on $S^P$ of the random conductivity field,

$$\nu(x, \xi) = \sum_{k=0}^{P} \nu_k(x) \psi_k(\xi),$$

has many zero modes $\nu_k(x)$ (due to the independence over distinct sub-domain). Consequently, some elementary matrices $[A']$ are zero, resulting in a sparse block structure for the Galerkin system.

The sparsity of the full Galerkin matrix system $[A]$ for $No = 4, \ldots, 6$ (dim $S = P + 1 = 15, 21$ and 28).
Stochastic conduction problem

Expectations (top) and standard deviations (bottom) of $U^h$ for $\text{No} = 5$. Left: two random conductivities ($N = 2$, $P = 20$). Right: single random conductivity ($N = 1$, $P = 5$).
Uncertainty in $U^h$ along the line $y = 0.5$. Left: two random conductivities ($N = 2, P = 20$). Right: single random conductivity ($N = 1, P = 5$).
Stochastic modes

\[ \Psi_0 = \psi_0(\xi_1)\psi_0(\xi_2) \]

\[ \Psi_1 = \psi_1(\xi_1)\psi_0(\xi_2) \]

\[ \Psi_2 = \psi_0(\xi_1)\psi_1(\xi_2) \]
Stochastic conduction problem

\[ \psi_3 = \psi_2(\xi_1)\psi_0(\xi_2) \quad \psi_4 = \psi_1(\xi_1)\psi_1(\xi_2) \quad \psi_5 = \psi_0(\xi_1)\psi_2(\xi_2) \]

\[ \psi_6 = \psi_3(\xi_1)\psi_0(\xi_2) \quad \psi_7 = \psi_2(\xi_1)\psi_1(\xi_2) \quad \psi_8 = \psi_1(\xi_1)\psi_2(\xi_2) \quad \psi_9 = \psi_0(\xi_1)\psi_3(\xi_2) \]

Modes \( u_k(x) \) of the stochastic solution for the nonuniform conductivity problem.
Heat exchanger

Table 5.1 Distributions of the random quantities for the problem of Sect. 5.1.5, and associated random variables

| Log normal | Median value | COV | Gaussian variable |
|------------|--------------|-----|-------------------|
| $v^1$      | 1            | $\sqrt{5}$ | $\xi_1$           |
| $v^2$      | $10^{-2}$    | $\sqrt{5}$ | $\xi_2$           |
| $W_n$      | 125          | $\sqrt{2}$ | $\xi_3$           |

| Uniform    | Lower bound | Upper bound | Uniform variable |
|------------|-------------|-------------|------------------|
| $T_c$      | 5           | 10          | $\xi_4$         |
| $T_e$      | 10          | 30          | $\xi_5$         |
Heat exchanger : FE meshes and solution

Fig. 5.13 Coarse (left) and fine (right) finite-element meshes used for the resolution of the problem of Sect. 5.1.5. The coarse mesh has 4,326 elements and 2,365 nodes, while the fine mesh has 9,140 elements and 4,947 nodes.

Fig. 5.14 Mean temperature field for the problem of Sect. 5.1.5. The computations are performed using $N_o = 4$ (126 spectral modes) and the fine finite-element mesh.

Fig. 5.15 Standard deviation of the temperature field for the problem of Sect. 5.1.5. The computations are performed using $N_o = 4$ (126 spectral modes) and the fine finite-element mesh.
Heat exchanger: FE meshes and solution

Fig. 5.16 Probability density function of temperature at $x_{\text{max}}$: left: linear scale; right: logarithmic scale. Results are shown for the coarse and fine finite-element meshes. In both cases, an expansion with No = 4 expansion (126 spectral modes) is used.

Fig. 5.17 Probability density functions of $W^h$ across the external boundary, $\Gamma_x$, of the casing for the problem of Sect. 5.1.5: left: linear scale; right: logarithmic scale. Results are shown for the coarse and fine finite-element meshes. In both cases, an expansion with No = 4 expansion (126 spectral modes) is used.

Fig. 5.20 Probability density functions of $\eta^h_c$ and $\eta^h_e$ of problem in Sect. 5.1.5 for the coarse (left) and fine (right) finite-element meshes. Solutions with No = 4 (126 spectral modes) are used.

Pdfs of $T_{\text{max}}$, flux at "cold" boundary and of efficiencies.
Analysis of the variance (ANOVA)
Any functional $G(\xi_1, \ldots, \xi_N) \in L_2(\Xi, P_\Xi)$ has unique orthogonal hierarchical decomposition of the form (the so-called Sobol or Hoeffding’s decomposition)

$$G(\xi_1, \ldots, \xi_N) = \sum_{s \subseteq \{1, \ldots, N\}} G_s(\xi_s)$$

$$= G_{\emptyset} + \sum_{i=1}^{N} G_{\{i\}}(\xi_i) + \sum_{i=1}^{N} \sum_{j=i+1}^{N} G_{\{i,j\}}(\xi_i, \xi_j)$$

$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \sum_{k=j+1}^{N} G_{\{i,j,k\}}(\xi_i, \xi_j, \xi_k) + \cdots + G_{\{1,\ldots,N\}}(\xi).$$

Orthogonality:

$$G_{\emptyset} = \mathbb{E}\{G\}, \quad \mathbb{E}\{G_s G_{s'}\} = \delta_{ss'}.$$
Variance decomposition:
The Sobol decomposition leads to a natural decomposition of the variance:

$$\text{Var} \{ G \} = \sum_{s \in \{1, \ldots, N\} \setminus \emptyset} \text{Var} \{ G_s \},$$

from which one defines the 1-st order sensitivity indices

$$S_i = \frac{\text{Var} \{ G \{i\} \}}{\text{Var} \{ G \}} \quad i = 1, \ldots, N,$$

and the total sensitivity indices

$$T_i = \frac{1}{\text{Var} \{ G \}} \sum_{s \ni \{i\}} \text{Var} \{ G_s \} \quad i = 1, \ldots, N.$$

Sensitivity indices are easily computed from the PC expansion of $G(\xi)$ [Crestaux et al, 2009].
Heat exchanger

### Table 5.2 Total sensitivity indices of the random fluxes and efficiencies

| Quantity | $I_{v1}$ | $I_{v2}$ | $I_{W_n}$ | $I_{T_e}$ | $I_{T_c}$ | Total   |
|----------|----------|----------|-----------|-----------|-----------|---------|
| $W_n$    | 0.0000   | 0.0000   | 1.0000    | 0.0000    | 0.0000    | 1.0000  |
| $W_c$    | 0.0254   | 0.0481   | 0.9185    | 0.0006    | 0.0114    | 1.0041  |
| $W_e$    | 0.2729   | 0.5116   | 0.1298    | 0.0076    | 0.1216    | 1.0435  |
| $T_{max}$| 0.0132   | 0.8082   | 0.1910    | 0.0016    | 0.0008    | 1.0147  |
| $\eta_1$ | 0.2980   | 0.5747   | 0.0121    | 0.0080    | 0.1437    | 1.0364  |
| $\eta_2$ | 0.3137   | 0.6000   | 0.0128    | 0.0095    | 0.1510    | 1.0869  |
1 Introduction
   • Simulation and errors
   • Input-data uncertainty
   • Alternative UQ methods

2 Spectral UQ
   • Polynomial Chaos expansions
   • Application to spectral UQ

3 Solution methods
   • Non-Intrusive methods
   • Galerkin projection

4 Application to conduction
   • Stochastic conduction problem
   • Heat exchanger

5 Application to natural convection
   • Boussinesq equations
   • Results

6 Conclusion
**Natural convection**

**Boussinesq approximation**

**Governing equations**

- **Momentum**:\[
\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \frac{\text{Pr}}{\sqrt{\text{Ra}}} \nabla^2 \mathbf{u} + \text{Pr} \theta \mathbf{y} \]

- **Mass**:\[
\nabla \cdot \mathbf{u} = 0
\]

- **Energy**:\[
\frac{\partial \theta}{\partial t} + \mathbf{u} \cdot \nabla \theta = \frac{1}{\sqrt{\text{Ra}}} \nabla^2 \theta
\]

**Uncertain boundary conditions**
Boussinesq equations

Natural convection

Boussinesq approximation

Governing equations

Uncertain boundary conditions
Boussinesq equations

Natural convection

Boussinesq approximation

Governing equations

Uncertain boundary conditions

- $u = 0$ on $\Gamma$.
- $\partial \theta(x, y = 0, 1)/\partial y = 0.$
- $\theta(x = 0, y) = 1/2.$
Natural convection

Boussinesq approximation

Governing equations

Uncertain boundary conditions

\[ u = 0 \text{ on } \Gamma. \]

\[ \partial \theta(x, y = 0, 1)/\partial y = 0. \]

\[ \theta(x = 0, y) = 1/2. \]

\[ \theta(x = 1, y, \omega) = -1/2 + \theta'(y, \omega). \]
Natural convection

Boussinesq approximation

Governing equations

Uncertain boundary conditions

\[ u = 0 \text{ on } \Gamma. \]
\[ \partial \theta(x, y = 0, 1) / \partial y = 0. \]
\[ \theta(x = 0, y) = 1/2. \]
\[ \theta(x = 1, y, \omega) = -1/2 + \theta'(y, \omega). \]

\[ \mathbb{E} \{ \theta'(y)\theta'(y') \} = \sigma_{\theta}^2 \exp[-|y - y'|/L], \quad \theta' \sim \mathcal{N}(0, \sigma_{\theta}^2). \]
Boussinesq equations

\[
\theta'(y, \xi) = \sum_{i=1}^{N} \sqrt{\lambda_i} \tilde{\theta}_i(y) \xi_i = \sum_{k=0}^{P} \theta_k(y) \psi_k(\xi),
\]

and the solution expresses as

\[
(u, p, \theta)(\xi) = \sum_{k=0}^{P} (u, p, \theta)_k \psi_k(\xi).
\]

- \( \xi_i \sim N(0, 1) \) \( \rightarrow \) Wiener-Hermite expansion.
- Stochastic dimension is \( N \).
- Expansion order \( \text{No} \) \( \rightarrow \) \( P + 1 = (N + \text{No})!/(N!\text{No}!) \).
Boussinesq equations

**BC and solution representations**

\[
\frac{\partial u_k}{\partial t} + \sum_{i,j=0}^{P} u_i \cdot \nabla u_j \frac{\langle \psi_i \psi_j, \psi_k \rangle}{\langle \psi_k, \psi_k \rangle} = -\nabla p_i + \frac{\text{Pr}}{\sqrt{\text{Ra}}} \nabla^2 u_k + \text{Pr} \theta_k y
\]

\[
\frac{\partial \theta_k}{\partial t} + \sum_{i,j=0}^{P} u_i \cdot \nabla \theta_j \frac{\langle \psi_i \psi_j, \psi_k \rangle}{\langle \psi_k, \psi_k \rangle} = \frac{1}{\sqrt{\text{Ra}}} \nabla^2 \theta_k
\]

\[
\nabla \cdot u_k = 0
\]

- \(P + 1\) **coupled** momentum and energy equations.
- \(P + 1\) **uncoupled** divergence constraints and BCs.

**Implementation and solver**
**BC and solution representations**

**Galerkin projection**

**Implementation and solver**

**Discretization**
- Uniform grid, staggered arrangement and 2nd order F-Diff.
- Semi-explicit second order Adams-Bashford time-scheme

**Incompressibility Treatment**
- Prediction / Projection method [Chorin, 1971]
- FFT based solver for the elliptic pressure equations

**CPU**: essentially projection of **uncoupled** modes:

\[ \text{Stochastic CPU} \approx (P + 1) \times \text{deterministic CPU}. \]
Boussinesq equations

Convergence and performance (unsteady solver)

- $N = 4 \sim 6$ is enough for $L \geq 1/3$
- $No = 3 \rightarrow$ relative error on variance $< 10^{-4}$
- $\sim 1000$ times more efficient than MC (LHS)
- $\sim 10$ times more efficient than NISP + GH quadrature (sparse grid ?)
- Parallelization

[olm et al., 2001]
Parallelization

Structure of $\langle \psi_m, \psi_k \rangle$

- Distribution of modes resolution
- Not scalable with increasing $P$
- Assembly of rhs requires too many communications
- Load balancing
- Domain decomposition?
**Velocity modes**

Ra = $10^6$, $L = 1$, $\sigma_\theta = 0.25$.

**Uncertainty bars**

$\sigma_\theta = 0.125$

$\sigma_\theta = 0.25$

$\sigma_\theta = 0.5$

[olm et al., 2002]
Results

**Temperature modes**

\[
Ra = 10^6, L = 1 - \sigma_\theta = 0.25.
\]

**Heat-transfer density**

\[
L = 1 - \neq \sigma_\theta
\]

\[
\neq L - \sigma_\theta = 0.25
\]

[olm et al., 2002]
Functional approximation provides rich information

For complex & very high-dimensional problems Monte-Carlo and sampling approaches remain the only alternative

Non smooth dependences require other type of bases: piecewise polynomial approximations & multiwavelets

High-dimensional problems calls for adaptive methods

Reduced basis approaches for complexity reduction: Greedy, PGD, ...

Do not address model uncertainty
Collaborators:
O. Knio (Duke University), H. Najm and B. Debusschere (Sandia, Livermore), R. Ghanem (USC), A. Ern (CEMICS, ENPC), M. Ndjinga and J.-M. Martinez (CEA, Saclay), A. Nouy (GeM, Centrale Nantes), L. Mathelin (LIMSI, CNRS), . . .

Students and Post-Docs:
J. Tryoen, Th. Cresteau, L. Tamellini, M. Schick, P. Tassi, A. Alexandian, . . .

Current Fundings:
GNR MoMaS and Agence Nationale pour la Recherche (grants: TYCHE & ASRMEI)