An adaptive sparse grid algorithm for elliptic PDEs with lognormal diffusion coefficient

F. Nobile[†], L. Tamellini[†], R. Tempone[♭], F. Tesei[†]

[†] CSQI - MATHICSE, EPFL, Switzerland
[♭] SRI UQ Center, KAUST, Saudi Arabia

Swiss Numerical Analysis Day 2015
Genève, April 17, 2015

Outline

Outline

The uncertainty quantification problem

math. model
$\mathcal{P}(\text{state vars.}) = 0$ $\rightarrow$ quantity of interest u

The uncertainty quantification problem

parameters $\mathbf{y}$ $\rightarrow$ math. model $\mathcal{P}(\text{state vars.}, \mathbf{y}) = 0$ $\rightarrow$ quantity of interest $u(\mathbf{y})$

The uncertainty quantification problem

parameters $\mathbf{y}$	$\rightarrow$	math. model $\mathcal{P}(\text{state vars.}, \mathbf{y}) = 0$	$\rightarrow$	quantity of interest $u(\mathbf{y})$

- The parameters $\mathbf{y}$ of the model may be *affected by uncertainty* (experimental measures, limited knowledge on system properties).

The uncertainty quantification problem

parameters $\mathbf{y}$	$\rightarrow$	math. model $\mathcal{P}(\text{state vars.}, \mathbf{y}) = 0$	$\rightarrow$	quantity of interest $u(\mathbf{y})$

- The parameters $\mathbf{y}$ of the model may be *affected by uncertainty* (experimental measures, limited knowledge on system properties).
- $\mathbf{y}$ can be modeled as a **random vector** with N components, taking values in $\Gamma \subseteq \mathbb{R}^N$, with joint probability density function $\varrho(\mathbf{y})$.

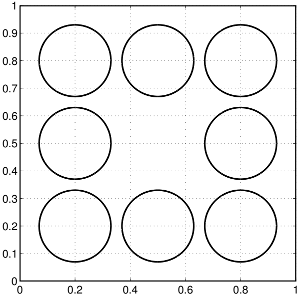
Therefore u is a **random function**, $u(\mathbf{y})$.

Goal: Compute statistical quantities for u, i.e. assess how the uncertainty on the parameters reflects on u: $\mathbb{E}[u]$, $\mathbb{V}\mathrm{ar}[u]$, $Pr(u > u_0)$.

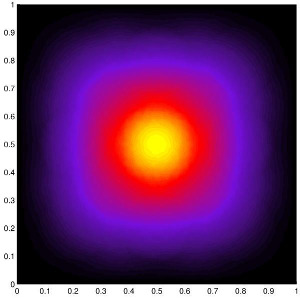
Method: Use sparse grids collocation to exploit regularity of $u(\mathbf{y})$.

Some examples on what can be done

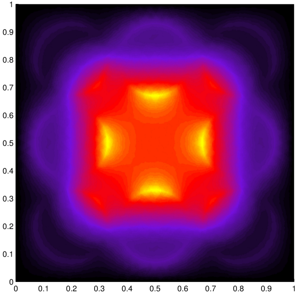
Diffusion problem with random inclusions ("the cookies problem")



geometry

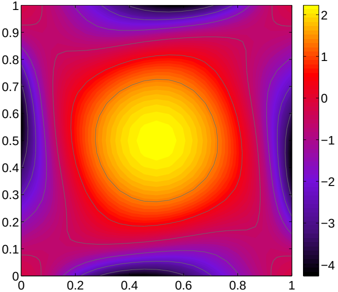


mean of u

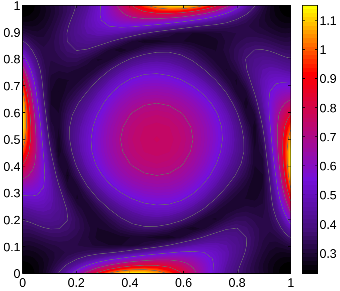


standard deviation of u

Steady Navier-Stokes with uncertain Reynolds and forcing term



mean vorticity field



st. dev. of the vorticity field

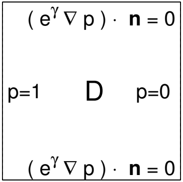
Outline

The Darcy problem

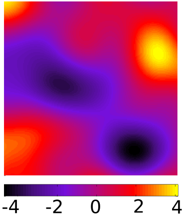
Find a pressure $p : \overline{D} \to \mathbb{R}$, such that

$$\begin{cases} -\nabla \cdot (e^{\gamma} \nabla p) = f & \text{in } D, \\ +B.C. \text{ (see plot on the right).} \end{cases}$$

u = outward flux from the right-hand boundary.



the log-permeability field γ is not constant in space (see right) but in practice we know its value only at log-points (drill locations). How to fill the gaps?



$\Downarrow$

model γ as a random field

Random fields

- instead of $\gamma = \gamma(\mathbf{x})$, let $\gamma = \gamma(\mathbf{x}, \mathbf{y})$ with y_n random variables, $\mathbf{y} \in \Gamma$:
  - for each realization $\gamma(\cdot, \mathbf{y})$ is a function in $L^\infty(D)$
  - for each physical point $\gamma(\mathbf{x}, \cdot)$ is a random variable

- a covariance function describes the interaction between any couple of points, e.g. $Cov\left[\mathbf{x}_0, \mathbf{x}_1\right] = \exp\left(-\frac{\|\mathbf{x}_0 - \mathbf{x}_1\|^2}{L_C^2}\right)$

- represented by a (truncated) Karhunen-Loève expansion

$$\gamma(\mathbf{x}, \mathbf{y}) = \sigma \sum_{k=1}^{N} y_k \gamma_k \phi_k(\mathbf{x}), \quad y_i \sim \mathcal{N}(0,1) \text{ i.i.d.} \Rightarrow \begin{cases} \Gamma = \mathbb{R}^N \\ \varrho(\mathbf{y}) = \frac{1}{\sqrt{(2\pi)^N}} e^{-\frac{\sum_{n=1}^N y_n^2}{2}} \end{cases}$$

Find a pressure $p(\mathbf{x}, \mathbf{y}) : \overline{D} \times \Gamma \to \mathbb{R}$, such that ϱ-a.e.:

$$\begin{cases} -\nabla \cdot \left(e^{\gamma(\mathbf{x}, \mathbf{y})} \nabla p(\mathbf{x}, \mathbf{y})\right) = f(\mathbf{x}) & \mathbf{x} \in D, \\ +B.C. \end{cases}$$

$u(\mathbf{y}) =$ outward flux is a random fun. $\to$ approximate e.g. $\mathbb{E}[u]$

What strategy?

Problem: we may need **tens-hundreds of random variables** to represent accurately the field! How can we tackle this?

Monte Carlo

- Generate a sample $\{\mathbf{y}_i\}_{i=1}^M$
- $\mathbb{E}[u] \approx u_{MC} = \frac{1}{M}\sum_i u(\mathbf{y}_i)$
- accuracy $\mathcal{O}(1/\sqrt{M})$. Slow but independent of N (no "*curse of dimensionality*")

Polynomial approximation

- The map $\mathbf{y} \to u(\mathbf{x}, \mathbf{y})$ is actually very smooth
- Approximate it by a polynomial surrogate model, $u(\mathbf{y}) \approx u_{pol} = \sum_i u_i \phi_i(\mathbf{y})$
- $\mathbb{E}[u_{pol}]$ by post-process
- convergence of $\|u - u_{pol}\|$ may deteriorate with N

In this talk: Polynomial approximation by adaptive sparse grids (reduced "curse of dimensionality") + Monte Carlo with control variate

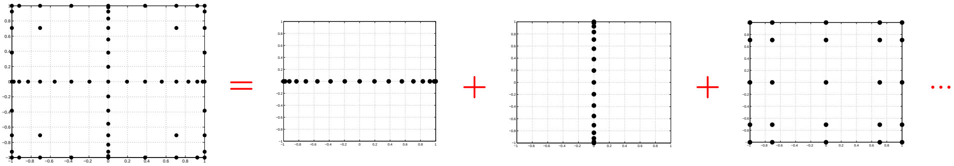
Outline

Sparse grid approximation of $u(\mathbf{y})$

Let $\mathbf{i} \in \mathbb{N}_+^N$ and denote by $\mathcal{T}^{m(\mathbf{i})}[u](\mathbf{y})$ the tensor Lagrangian interpolant of $u(\mathbf{y})$ over Γ with $m(i_1) \times m(i_1) \times \ldots m(i_N)$ points.

A sparse grid approximation is a linear combination of tensor interpolants:

$$\mathcal{S}_{\mathcal{I}}[u](\mathbf{y}) := \sum_{\mathbf{i} \in \mathcal{I}} c_{\mathbf{i}} \mathcal{T}^{m(\mathbf{i})}[u](\mathbf{y}), \qquad \mathbb{E}[u] \approx \mathcal{Q}_{\mathcal{I}}[u] := \sum_{\mathbf{y}_i} \omega_i u(\mathbf{y}_i)$$



$$\mathcal{S}_{\mathcal{I}}[u](\mathbf{y}) = \mathcal{T}^{[15\ 1]}[u](\mathbf{y}) + \mathcal{T}^{[1\ 15]}[u](\mathbf{y}) + \mathcal{T}^{[5\ 5]}[u](\mathbf{y}) \ldots$$

In practice, solve the Darcy problem for each $\mathbf{y}_i$ in the grid and combine them according to the formula above

Sparse grid approximation of $u(\mathbf{y})$

Let $\mathbf{i} \in \mathbb{N}_+^N$ and denote by $\mathcal{T}^{m(\mathbf{i})}[u](\mathbf{y})$ the tensor Lagrangian interpolant of $u(\mathbf{y})$ over Γ with $m(i_1) \times m(i_1) \times \ldots m(i_N)$ points.

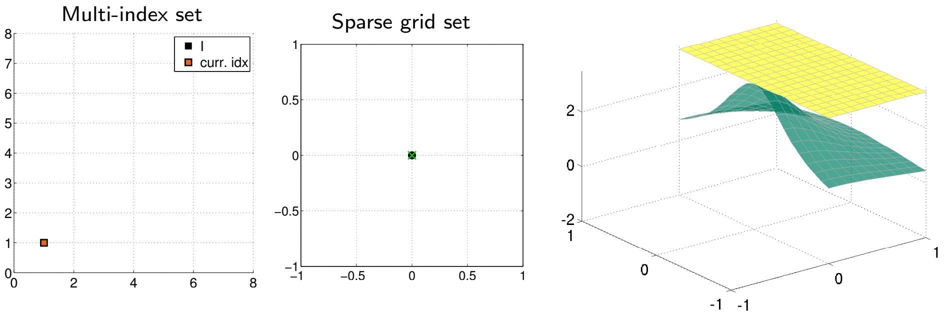
A sparse grid approximation is a linear combination of tensor interpolants:

$$\mathcal{S}_{\mathcal{I}}[u](\mathbf{y}) := \sum_{\mathbf{i} \in \mathcal{I}} c_{\mathbf{i}} \mathcal{T}^{m(\mathbf{i})}[u](\mathbf{y}), \qquad \mathbb{E}[u] \approx \mathcal{Q}_{\mathcal{I}}[u] := \sum_{\mathbf{y}_i} \omega_i u(\mathbf{y}_i)$$

- **Sparse grids idea**: cheaper than full tensor grids, but similar accuracy
- **Univariate points:** $y_i \sim \mathcal{N}(0,1) \rightarrow$ Gauss–Hermite, Genz–Keister, gen. Leja
- **The efficiency of the sparse grids depends on $\mathcal{I}$.**
- **Admissibility condition for $\mathcal{I}$:** $\quad \forall \mathbf{i} \in \mathcal{I}, \quad \mathbf{i} - \mathbf{e}_j \in \mathcal{I}$ if $i_j > 1$.
- **The coefficients** $c_{\mathbf{i}}$ are uniquely defined given $\mathcal{I}$

How to select $\mathcal{I}$: adaptive sparse grids

Gerstner & Griebel, 2003



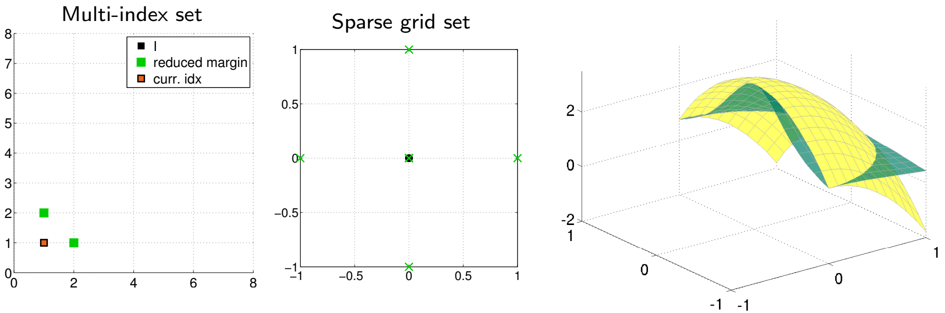
Given $\mathbf{i} = \mathbf{1}$, $\mathcal{I} = \{\mathbf{i}\}$ and $\mathcal{R} = \emptyset$ repeat:

1. Add to $\mathcal{R}$ the neighbors of $\mathbf{i}$ feasible wrt to $\mathcal{I}$
2. Compute $\mathcal{S}_{\mathcal{I} \cup \mathcal{B}}[u]$
3. find the index $\mathbf{j} \in \mathcal{R}$ that improved the most the approximation (e.g. check the difference in approximation of the mean or in L^∞-norm)
4. set $\mathbf{i} = \mathbf{j}$ and move it from $\mathcal{R}$ to $\mathcal{I}$

NB: omitting technicalities using non-nested points and unbounded Γ.

How to select $\mathcal{I}$: adaptive sparse grids

Gerstner & Griebel, 2003



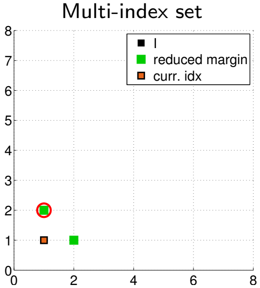
Given $\mathbf{i} = \mathbf{1}$, $\mathcal{I} = \{\mathbf{i}\}$ and $\mathcal{R} = \varnothing$ repeat:

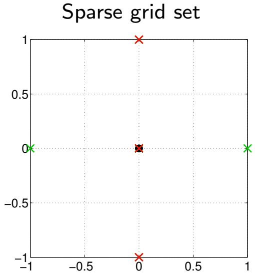
1. Add to $\mathcal{R}$ the neighbors of $\mathbf{i}$ feasible wrt to $\mathcal{I}$
2. Compute $\mathcal{S}_{\mathcal{I}\cup\mathcal{B}}[u]$
3. find the index $\mathbf{j} \in \mathcal{R}$ that improved the most the approximation (e.g. check the difference in approximation of the mean or in L^∞-norm)
4. set $\mathbf{i} = \mathbf{j}$ and move it from $\mathcal{R}$ to $\mathcal{I}$

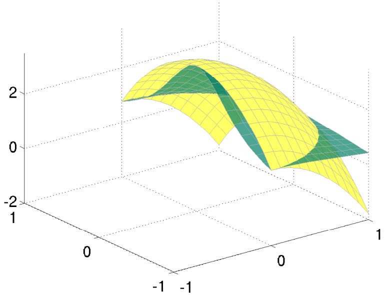
NB: omitting technicalities using non-nested points and unbounded Γ.

How to select $\mathcal{I}$: adaptive sparse grids

Gerstner & Griebel, 2003







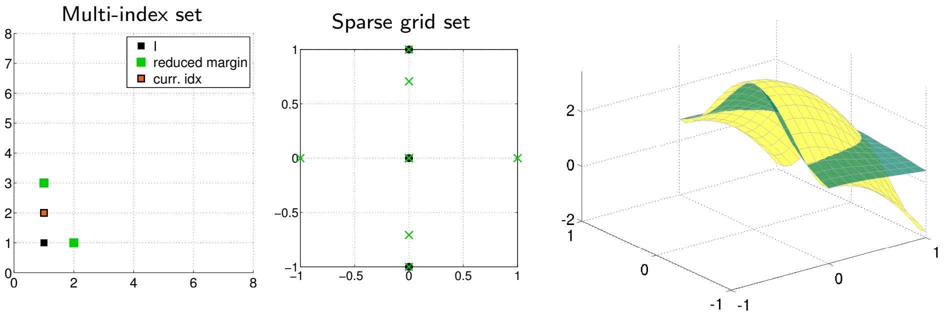
Given $\mathbf{i} = \mathbf{1}$, $\mathcal{I} = \{\mathbf{i}\}$ and $\mathcal{R} = \varnothing$ repeat:

1. Add to $\mathcal{R}$ the neighbors of $\mathbf{i}$ feasible wrt to $\mathcal{I}$
2. Compute $\mathcal{S}_{\mathcal{I}\cup\mathcal{B}}[u]$
3. find the index $\mathbf{j} \in \mathcal{R}$ that improved the most the approximation (e.g. check the difference in approximation of the mean or in L^∞-norm)
4. set $\mathbf{i} = \mathbf{j}$ and move it from $\mathcal{R}$ to $\mathcal{I}$

NB: omitting technicalities using non-nested points and unbounded Γ.

How to select $\mathcal{I}$: adaptive sparse grids

Gerstner & Griebel, 2003



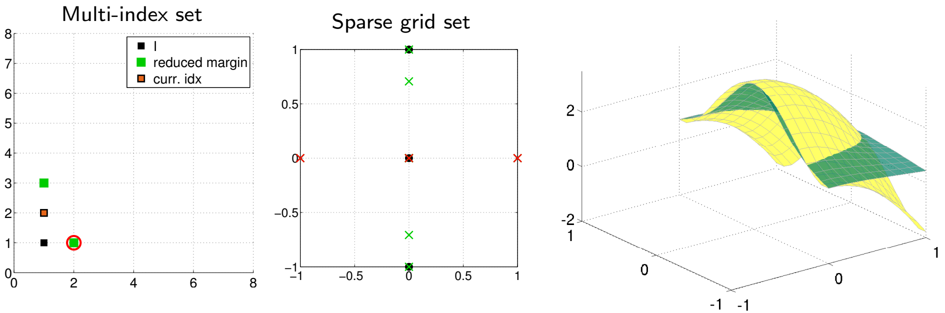
Given $\mathbf{i} = \mathbf{1}$, $\mathcal{I} = \{\mathbf{i}\}$ and $\mathcal{R} = \varnothing$ repeat:

1. Add to $\mathcal{R}$ the neighbors of $\mathbf{i}$ feasible wrt to $\mathcal{I}$
2. Compute $\mathcal{S}_{\mathcal{I} \cup \mathcal{B}}[u]$
3. find the index $\mathbf{j} \in \mathcal{R}$ that improved the most the approximation (e.g. check the difference in approximation of the mean or in L^∞-norm)
4. set $\mathbf{i} = \mathbf{j}$ and move it from $\mathcal{R}$ to $\mathcal{I}$

NB: omitting technicalities using non-nested points and unbounded Γ.

How to select $\mathcal{I}$: adaptive sparse grids

Gerstner & Griebel, 2003



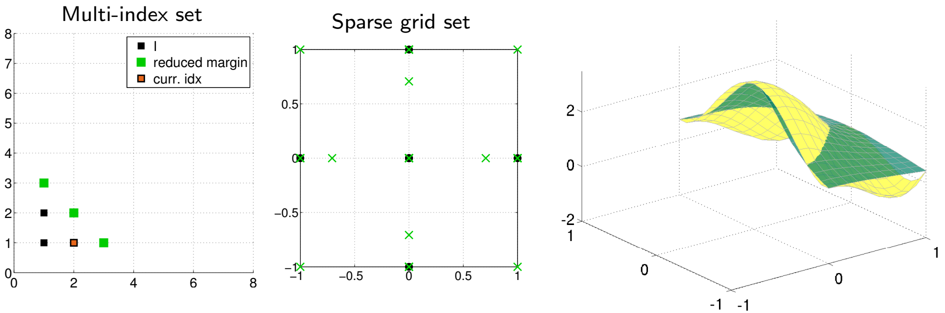
Given $\mathbf{i} = \mathbf{1}$, $\mathcal{I} = \{\mathbf{i}\}$ and $\mathcal{R} = \varnothing$ repeat:

1. Add to $\mathcal{R}$ the neighbors of $\mathbf{i}$ feasible wrt to $\mathcal{I}$
2. Compute $\mathcal{S}_{\mathcal{I}\cup\mathcal{B}}[u]$
3. find the index $\mathbf{j} \in \mathcal{R}$ that improved the most the approximation (e.g. check the difference in approximation of the mean or in L^∞-norm)
4. set $\mathbf{i} = \mathbf{j}$ and move it from $\mathcal{R}$ to $\mathcal{I}$

NB: omitting technicalities using non-nested points and unbounded Γ.

How to select $\mathcal{I}$: adaptive sparse grids

Gerstner & Griebel, 2003



Given $\mathbf{i} = \mathbf{1}$, $\mathcal{I} = \{\mathbf{i}\}$ and $\mathcal{R} = \emptyset$ repeat:

1. Add to $\mathcal{R}$ the neighbors of $\mathbf{i}$ feasible wrt to $\mathcal{I}$
2. Compute $\mathcal{S}_{\mathcal{I}\cup\mathcal{B}}[u]$
3. find the index $\mathbf{j} \in \mathcal{R}$ that improved the most the approximation (e.g. check the difference in approximation of the mean or in L^∞-norm)
4. set $\mathbf{i} = \mathbf{j}$ and move it from $\mathcal{R}$ to $\mathcal{I}$

NB: omitting technicalities using non-nested points and unbounded Γ.

To refine or to add rand. var.? Dimension-adaptivity

- **Problem:** generating $\mathcal{R}$ in high-dimensional spaces is too expensive.

To refine or to add rand. var.? Dimension-adaptivity

- **Problem:** generating $\mathcal{R}$ in high-dimensional spaces is too expensive.
- **Assume** that the Karhunen–Loève expansion $\gamma = \sigma \sum_{k=1}^{N} y_k \gamma_k \phi_k$ introduces a "weak ordering" of random variables, i.e. there exists $N_b \geq 1$ (buffer) s.t. y_{n+N_b} is guaranteed to be less important than y_n.

To refine or to add rand. var.? Dimension-adaptivity

- **Problem:** generating $\mathcal{R}$ in high-dimensional spaces is too expensive.
- **Assume** that the Karhunen–Loève expansion $\gamma = \sigma \sum_{k=1}^{N} y_k \gamma_k \phi_k$ introduces a "weak ordering" of random variables, i.e. there exists $N_b \geq 1$ (buffer) s.t. y_{n+N_b} is guaranteed to be less important than y_n.
- **Idea:** Add random variables **gradually** (balance refinement and addition of variables). Note that this means we don't need to truncate a-priori the Karhunen–Loève expansion of γ!

To refine or to add rand. var.? Dimension-adaptivity

- **Problem:** generating $\mathcal{R}$ in high-dimensional spaces is too expensive.
- **Assume** that the Karhunen–Loève expansion $\gamma = \sigma \sum_{k=1}^{N} y_k \gamma_k \phi_k$ introduces a "weak ordering" of random variables, i.e. there exists $N_b \geq 1$ (buffer) s.t. y_{n+N_b} is guaranteed to be less important than y_n.
- **Idea:** Add random variables **gradually** (balance refinement and addition of variables). Note that this means we don't need to truncate a-priori the Karhunen–Loève expansion of γ!
- **Define** a random variable y_n as **activated** if $\min_{\mathbf{i} \in \mathcal{I}} i_n > 1$.

To refine or to add rand. var.? Dimension-adaptivity

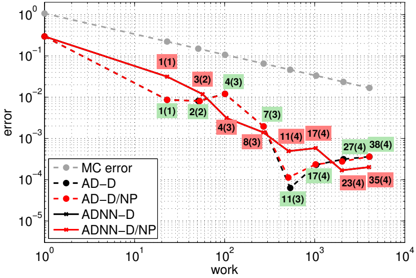
- **Problem:** generating $\mathcal{R}$ in high-dimensional spaces is too expensive.
- **Assume** that the Karhunen–Loève expansion $\gamma = \sigma \sum_{k=1}^{N} y_k \gamma_k \phi_k$ introduces a "weak ordering" of random variables, i.e. there exists $N_b \geq 1$ (buffer) s.t. y_{n+N_b} is guaranteed to be less important than y_n.
- **Idea:** Add random variables **gradually** (balance refinement and addition of variables). Note that this means we don't need to truncate a-priori the Karhunen–Loève expansion of γ!
- **Define** a random variable y_n as **activated** if $\min_{\mathbf{i} \in \mathcal{I}} i_n > 1$.

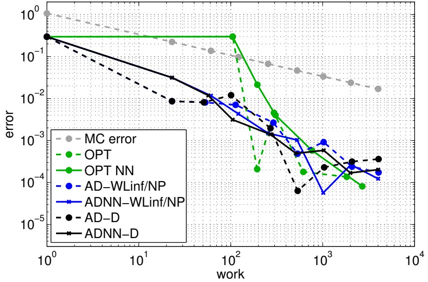
A simple dimension-adaptive algorithm

1. start the adaptive algorithm using N_b random variables
2. As soon as one of these "buffer variables" gets activated, add a new random variable to the approximaton.

The uncertain Darcy problem – results 1

Field data: $\sigma = 1$, corr. length $L_c = 0.5$, $\gamma(\mathbf{x}, \mathbf{y})$ smooth wrt $\mathbf{x}$





- moderate number of random variables needed
- convergence robust wrt. type of points and $\mathbb{E}[]/L^\infty$-driven adaptation

Smoothness Warning!

If $\gamma(\mathbf{x}, \mathbf{y})$ is not smooth wrt $\mathbf{x}$ (depends on the covariance function), a larger number of random variables is needed and even the adaptive sparse grids may not be effective!

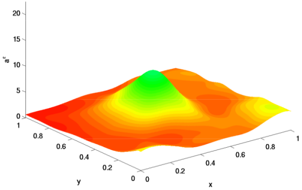
Outline

Rough random fields γ: Monte Carlo Control Variate

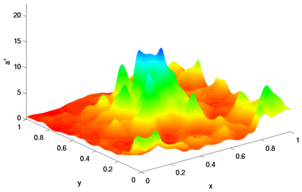
γ non-differentiable wrt $\mathbf{x} \Rightarrow$ sparse grids may be non-effective.

Remedy: use sparse grids as **control var.** (preconditioner) for MC

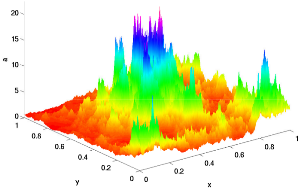
1. Consider a **smoothed field** γ^ϵ, such that $\mathcal{Q}_\mathcal{I}[u^\epsilon] \to \mathbb{E}[u^\epsilon]$ quickly.



smoothed field, $\epsilon = 1/2^4$



smoothed field $\epsilon = 1/2^6$



non-smoothed field, $\epsilon = 0$

Rough random fields γ: Monte Carlo Control Variate

γ non-differentiable wrt $\mathbf{x} \Rightarrow$ sparse grids may be non-effective.

Remedy: use sparse grids as **control var.** (preconditioner) for MC

1. Consider a **smoothed field** γ^ϵ, such that $\mathcal{Q}_\mathcal{I}[u^\epsilon] \to \mathbb{E}[u^\epsilon]$ quickly.
2. Define $u_{CV} = u - u^\epsilon + \mathcal{Q}_\mathcal{I}[u^\epsilon]$. There holds

$$\mathbb{E}[u_{CV}] = \mathbb{E}[u], \quad \mathbb{V}\mathrm{ar}(u_{CV}) = \mathbb{V}\mathrm{ar}(u) + \mathbb{V}\mathrm{ar}(u^\epsilon) - 2\mathrm{cov}(u, u^\epsilon)$$

Thus, the smaller ϵ, the smaller the MC error, but slower the convergence $\mathcal{Q}_\mathcal{I}[u^\epsilon] \to \mathbb{E}[u^\epsilon]$.

Rough random fields γ: Monte Carlo Control Variate

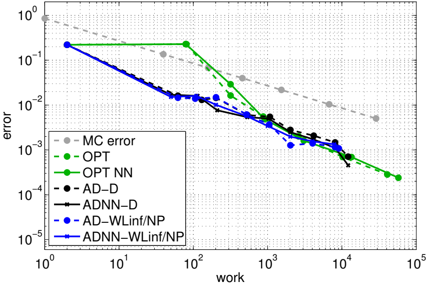
γ non-differentiable wrt $\mathbf{x} \Rightarrow$ sparse grids may be non-effective.

Remedy: use sparse grids as **control var.** (preconditioner) for MC

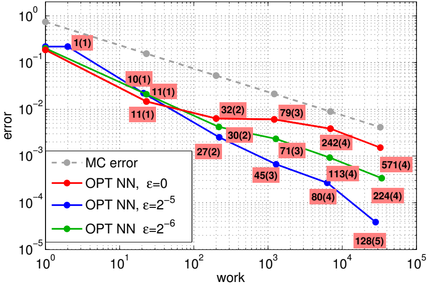
1. Consider a **smoothed field** γ^ϵ, such that $\mathcal{Q}_\mathcal{I}[u^\epsilon] \to \mathbb{E}[u^\epsilon]$ quickly.
2. Define $u_{CV} = u - u^\epsilon + \mathcal{Q}_\mathcal{I}[u^\epsilon]$. There holds
$$\mathbb{E}[u_{CV}] = \mathbb{E}[u], \quad \mathbb{V}\mathrm{ar}(u_{CV}) = \mathbb{V}\mathrm{ar}(u) + \mathbb{V}\mathrm{ar}(u^\epsilon) - 2\mathrm{cov}(u, u^\epsilon)$$
Thus, the smaller ϵ, the smaller the MC error, but slower the convergence $\mathcal{Q}_\mathcal{I}[u^\epsilon] \to \mathbb{E}[u^\epsilon]$.
3. Set $\mathbb{E}[u_{CV}] \approx \frac{1}{M}\sum_{i=1}^{M} u^{CV}(\omega_i) = \frac{1}{M}\sum_{i=1}^{M} (u(\omega_i) - u^\epsilon(\omega_i)) + \mathcal{Q}_\mathcal{I}^m[u^\epsilon]$.
Here we simply choose $M = card(pts(\mathcal{S}_\mathcal{I}[u]))$ (work balance). Other strategies are possible.

The uncertain Darcy problem – results 2

Field data: $\sigma = 1$, corr. length $L_c = 0.5$, rough field realizations (Hölder continuous only)



MCCV error. ∼ 30 r.v. activated.



The sparse grid error worsens as $\epsilon \to 0$.

Conclusions

1. Uncertainty Quantification is a fast-growing area at the interface between Scientific Computing and Statistics;
2. Whenever the quantity of interest is smooth wrt the random parameters, adaptive sparse grids schemes can be used as an effective alternative to the Monte Carlo strategy;
3. The dimension-adaptive implementation allows to work without a-priori truncation of the random field;
4. If the random field has rough realizations, using adaptive sparse grids in a Monte Carlo Control Variate framework can improve results.

Thank you for your attention!

Thank you for your attention!

Make sure to attend

SIAM-UQ 2016

April 5-8, 2016

SwissTech Convention Center

EPFL, Lausanne, Switzerland

Bibliography

- F. Nobile, L. Tamellini, F. Tesei and R. Tempone. An adaptive sparse grid algorithm for elliptic PDEs with lognormal diffusion coefficient. Mathicse report 4/2015, EPFL, 2015.

- F. Tesei and F. Nobile. A Multi Level Monte Carlo Method with Control Variate for elliptic PDEs with log-normal coefficients. Mathicse report 49/2014, EPFL, 2014.

- F. Nobile, L. Tamellini, and R. Tempone. Convergence of quasi-optimal sparse grids approximation of Hilbert-valued functions: application to random elliptic PDEs. Mathicse report 12/2014, EPFL, 2014.

- L. Tamellini, O. Le Maître, A. Nouy Model Reduction Based on Proper Generalized Decomposition for the Stochastic Steady Incompressible Navier–Stokes Equations. SIAM Journal on Scientific Computing 36(3), 2014

- J. Beck, F. Nobile, L. Tamellini, and R. Tempone. A Quasi-optimal Sparse Grids Procedure for Groundwater Flows. Selected papers from the ICOSAHOM '12 conference.

- J. Bäck, F. Nobile, L. Tamellini, and R. Tempone. Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: a numerical comparison. Selected papers from the ICOSAHOM '09 conference.