An introduction to optimal control problems under uncertainty

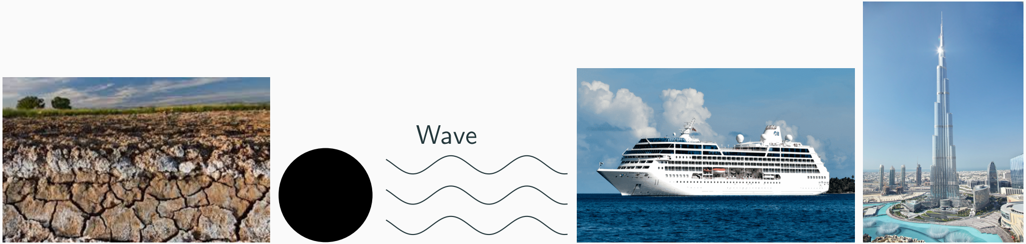
Tommaso Vanzan
CSQI Chair, EPFL

Research School - Domain Decomposition for Optimal Control Problems
CIRM (Marseille), 6th September 2022

Table of contents

Motivation

Uncertainty is ubiquitous



Role of Uncertainty Quantification (UQ)

- Forward UQ: study the propagation of the randomness from the mathematical model to the output quantities.
- Inverse UQ: incorporate uncertainty in engineering design processes or reconstruct the uncertain input from measurements.

Anatomy of a PDE-constrained Optimal Control Problem Under Uncertainty (OCPUU)

General formulation

$(\Omega, \mathcal{F}, \mathbb{P})$ complete probability space.

$$\min_{y \in \mathcal{Y}, u \in U_{ad} \subset U} \mathcal{R}\left[Q(y)\right] + \frac{\nu}{2}\|u\|_U^2,$$
$$\text{s.t. } \langle e(y, u, \omega), v \rangle = 0, \quad \forall v \in Y, \text{for a.e. } \omega \in \Omega.$$

General formulation

$(\Omega, \mathcal{F}, \mathbb{P})$ complete probability space.

$$\min_{y\in\mathcal{Y}, u\in U_{ad}\subset U} \mathcal{R}\left[Q(y)\right] + \frac{\nu}{2}\|u\|_U^2,$$
$$\text{s.t. } \langle e(y(\omega), u, \omega), v\rangle = 0, \quad \forall v \in Y, \text{for a.e. } \omega \in \Omega.$$

- A control/design variable unknown $u \in U_{ad} \subset U$, a state variable $y \in \mathcal{Y}$. U, $\mathcal{Y}$ and Y suitable Hilbert spaces. $\mathcal{Y}$ and Y are linked:

$$\mathcal{Y} = L^p(\Omega; Y) = \left\{ v : \Omega \to Y : \text{ v strongly measurable s.t } \int_\Omega \|v(\omega)\|_Y^p d\mathbb{P}(\omega) < \infty \right\}.$$

General formulation

$(\Omega, \mathcal{F}, \mathbb{P})$ complete probability space.

$$\min_{y \in \mathcal{Y}, u \in U_{ad} \subset U} \mathcal{R}\left[Q(y(\omega))\right] + \frac{\nu}{2}\|u\|_U^2,$$

$$\text{s.t. } \langle e(y, u, \omega), v\rangle = 0, \quad \forall v \in Y, \text{for a.e. } \omega \in \Omega.$$

- A control/design variable unknown $u \in U_{ad} \subset U$, a state variable $y \in \mathcal{Y}$. U, $\mathcal{Y}$ and Y(e.g. $H_0^1(D)$) suitable Hilbert spaces. $\mathcal{Y}$ and Y are linked:

$$\mathcal{Y} = L^p(\Omega; Y) = \left\{ v : \Omega \to Y : \text{ v strongly measurable s.t } \int_\Omega \|v(\omega)\|_Y^p d\mathbb{P}(\omega) < \infty \right\}.$$

- random partial differential equation expressed in weak form.

General formulation

$(\Omega, \mathcal{F}, \mathbb{P})$ complete probability space.

$$\min_{y \in \mathcal{Y}, u \in U_{ad} \subset U} \mathcal{R}\left[Q(y(\omega))\right] + \frac{\nu}{2}\|u\|_U^2,$$

$$\text{s.t. } \langle e(y, u, \omega), v \rangle = 0, \quad \forall v \in Y, \text{for a.e. } \omega \in \Omega.$$

- A control/design variable unknown $u \in U_{ad} \subset U$, a state variable $y \in \mathcal{Y}$. U, $\mathcal{Y}$ and Y(e.g. $H_0^1(D)$) suitable Hilbert spaces. $\mathcal{Y}$ and Y are linked:

$$\mathcal{Y} = L^p(\Omega; Y) = \left\{ v : \Omega \to Y : \text{ v strongly measurable s.t } \int_\Omega \|v(\omega)\|_Y^p d\mathbb{P}(\omega) < \infty \right\}.$$

- random partial differential equation expressed in weak form.
- Quantity of interest $Q(\cdot)$, risk measure $\mathcal{R}(\cdot)$.

General formulation

$(\Omega, \mathcal{F}, \mathbb{P})$ complete probability space.

$$\min_{y \in \mathcal{Y}, u \in U_{ad} \subset U} \mathcal{R}\left[Q(y(\omega))\right] + \frac{\nu}{2}\|u\|_U^2,$$
$$\text{s.t. } \langle e(y,u,\omega), v\rangle = 0, \quad \forall v \in Y, \text{ for a.e. } \omega \in \Omega.$$

- A control/design variable unknown $u \in U_{ad} \subset U$, a state variable $y \in \mathcal{Y}$. U, $\mathcal{Y}$ and Y(e.g. $H_0^1(D)$) suitable Hilbert spaces. $\mathcal{Y}$ and Y are linked:
$$\mathcal{Y} = L^p(\Omega; Y) = \left\{ v : \Omega \to Y : \text{ v strongly measurable s.t } \int_\Omega \|v(\omega)\|_Y^p d\mathbb{P}(\omega) < \infty \right\}.$$
- random partial differential equation expressed in weak form.
- Quantity of interest $Q(\cdot)$, risk measure $\mathcal{R}(\cdot)$.
- regularization term on u, $\nu \geq 0$.

The control saught is <u>deterministic</u>! (a stochastic control would be $u = u(\omega)$.)

Random PDE: a canonical problem

$$
\begin{aligned}
-\nabla \cdot (\kappa(x,\omega)\nabla y(x,\omega)) &= u(x), \quad x \in D, \\
y(x,\omega) &= g(x), \quad x \in \partial D.
\end{aligned}
$$

For each realization of ω, different $\kappa(x,\omega) \implies$ different solution $y(x,\omega)$.

$$\int_D \kappa(x,\omega)\nabla y(x,\omega) \cdot \nabla v(x) dx = \int_D u(x)v(x) dx \quad \forall v \in Y, \text{for a.e. } \omega \in \Omega..$$

$$\downarrow$$

$$\langle e(y, u, \omega), v\rangle = 0, \quad \forall v \in Y, \text{for a.e. } \omega \in \Omega.$$

Random PDE: a canonical problem

$$
\begin{aligned}
-\nabla \cdot (\kappa(x,\omega)\nabla y(x,\omega)) &= u(x), \quad x \in D, \\
y(x,\omega) &= g(x), \quad x \in \partial D.
\end{aligned}
$$

For each realization of ω, different $\kappa(x,\omega) \implies$ different solution $y(x,\omega)$.

$$
\int_D \kappa(x,\omega)\nabla y(x,\omega) \cdot \nabla v(x) dx = \int_D u(x)v(x) dx \quad \forall v \in Y, \text{for a.e. } \omega \in \Omega..
$$

$$
\downarrow
$$

$$
\langle e(y,u,\omega), v\rangle = 0, \quad \forall v \in Y, \text{for a.e. } \omega \in \Omega.
$$

Questions:

- What is a random field $\kappa(x,\omega)$?
- What does it mean to "solve" a random PDE?

Main reference: Lord, Powell, Shardlow, *An introduction to computational Stochastic PDEs*, Cambridge Press, 2014.

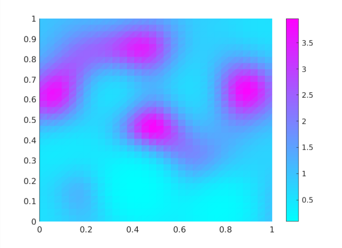
Random fields - I

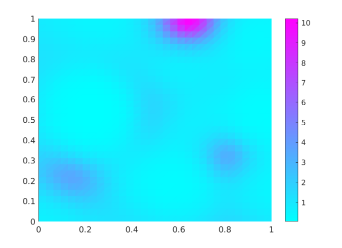
Definition (Random field)

Given a $D \subset \mathbb{R}^d$, a random field $\{\kappa(x) : \ x \in D\}$ is a collection of real-valued random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A random field is said to be of second order if $\kappa(x, \cdot) \in L^2(\Omega)$, $\forall x \in D$.

Definition (Realizations)

For a fixed $\omega \in \Omega$, a realization of $\{\kappa(x) : \ x \in D\}$ is a deterministic function $\kappa : D \to \mathbb{R}$ defined by $\kappa(x) = \kappa(x, \omega)$ for every $x \in D$.





How to generate realizations of a random field? Let's take one more abstract step..

Random fields - II

Definition (Covariance)

The covariance function of a second order random field $\{\kappa(x) : \; x \in D\}$ is $\mathrm{Cov}(x_1, x_2) : D \times D \to \mathbb{R}$:

$$\mathrm{Cov}(x_1, x_2) = \mathbb{E}\left[(\kappa(x_1, \cdot) - \mathbb{E}\left[\kappa(x_1, \cdot)\right])(\kappa(x_2, \cdot) - \mathbb{E}\left[\kappa(x_2, \cdot)\right])\right].$$

If $C(\cdot, \cdot)$ is continuous, let $T_C : L^2(D) \to L^2(D)$, with $T_C(f)(x) = \int_D Cov(x, y) f(y) dy$.

Theorem (Properties of Covariance function)

Let $D \subset \mathbb{R}^d$ be compact and $Cov : D \times D \to \mathbb{R}$ continuous. There exists a sequence of values $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq \cdots \geq 0$, with $\lim_{k \to 0} \lambda_k = 0$, and corresponding functions $b_j(x) : D \to \mathbb{R}$, $b_j \in L^2(D)$, such that

$$\int_D Cov(x, y) b_j(y) dy = \lambda_j b_j(x), \; \text{and} \; \int_D b_i(x) b_j(x) dx = \delta_{i,j}.$$

Proof: Show T_c is a self-adjoint compact operator (symmetry of $Cov(\cdot, \cdot)$ + Ascoli Arzelá).

Theorem (Karhunen-Loéve expansion)

A random field $\{\kappa(x) : \ x \in D\}$ with continuous covariance $Cov(x, y)$ can be represented as the infinite series

$$\kappa(x, \omega) = \mathbb{E}\left[\kappa(x, \omega)\right] + \sum_{j=1}^{\infty} \sqrt{\lambda_j} b_j(x) z_j(\omega),$$

where the random variables $z_j(\omega)$ are uncorrelated with zero mean, unit variance and defined by

$$y_j(\omega) = \frac{1}{\sqrt{\lambda_j}} \int_D (\kappa(x, \omega) - \mathbb{E}\left[\kappa(x, \omega)\right]) b_j(x) dx$$

Simple computational approach: compute covariance matrix over finite element mesh and compute its eigendecomposition to get $b_j(x)$ and λ_j.

Random fields - IV

The Karhunen-Loéve expansion "justifies"

$$\kappa(x,\omega) \approx \bar{\kappa}(x) + \sum_{j=1}^{N} \sqrt{\lambda_j} b_j(x) \widetilde{z}_j(\omega),$$

where $\widetilde{z}_j(\omega)$ are uniform,Gaussian,etc... random variables.
To generate κ, we only need to generate $\{\widetilde{z}_j\}$ according to their (simple) distribution.

A well-studied random field is the log-normal field

$$\kappa(x,\omega) = \exp(g(x,\omega)) \approx \exp\left(\mathbb{E}\left[g(x,\omega)\right] + \sum_{j=1}^{N} \sqrt{\lambda_j} b_j(x) N_j(\omega)\right).$$

Basic assumptions:

- $\kappa(\cdot,\omega), \kappa^{-1}(\cdot,\omega) \in L^\infty(D)$ for $\mathbb{P}$-a.e. ω.
- Further, $\kappa_{\min}(\omega) := \mathrm{essinf}_{x\in D}\kappa(x,\omega)$ and $\kappa_{\max}(\omega) := \mathrm{esssup}_{x\in D}\kappa(x,\omega)$ satisfy

$$\kappa_{\min}, \kappa_{\max} \in L^p(\Omega), \text{ for some } p \in [1,\infty].$$

Solve a Random PDE - - Monte Carlo

Aim to compute $\bar{y}(x) := \mathbb{E}[y(x,\omega)]$ and $\sigma_y(x) := \mathbb{E}\left[(y(x,\omega) - \bar{y}(x))^2\right]$.

Generate M samples of $\kappa(x,\omega)$: $\{\kappa_r(x,\omega_r)\}_{r=1}^M$. Solve

$$\int_D \kappa_r(x,\omega_r)\nabla y(x,\omega_r)\nabla v(x)dx = \int_D u(x)v(x), \forall v \in Y, \;\; r = 1,\dots,M.$$

Approximate:

$$\bar{y}(x) \approx \bar{y}^{MC}(x) := \frac{1}{M}\sum_{r=1}^{M} y(x,\omega_r), \quad \sigma_y(x) \approx \frac{1}{N-1}\sum_{r=1}^{M}(y(x,\omega_r) - \bar{y}^{MC}(x))^2.$$

Error estimate:

$$\|\bar{y} - \bar{y}^{MC}\|_{L^2(\Omega;Y)} = \sqrt{\mathbb{E}\left[\|\bar{y} - \bar{y}^{MC}\|_Y^2\right]} \le \frac{C}{\sqrt{M}}.$$

Proof uses standard Monte Carlo arguments

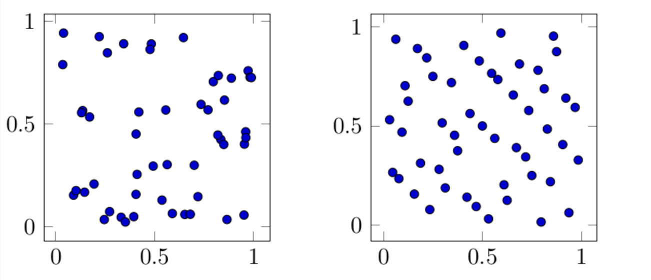
$$\mathbb{E}\left[\|\bar{y} - \bar{y}^{MC}\|_Y^2\right] = \mathbb{E}\left[\left\|\frac{1}{M}\sum_{r=1}^{M}\bar{y} - \frac{1}{M}\sum_{r=1}^{M} y(\omega_r)\right\|_Y^2\right] = \frac{1}{M^2}\mathbb{E}\left[\left\|\sum_{r=1}^{M}(\bar{y} - y(\omega_r))\right\|_Y^2\right]$$

$$\frac{1}{M^2}\sum_{r=1}^{M}\mathbb{E}\left[\|(\bar{y} - y(\omega_r))\|_Y^2\right] \le \frac{1}{M^2}\sum_{r=1}^{M}\mathbb{E}\left[\|y(\omega)\|_Y^2\right] \le \frac{C^2}{M}.$$

Solve a Random PDE - - Quasi Monte Carlo

Assumption: randomness is modelled through a random vector of $\zeta(\omega) = (\zeta_1(\omega), \dots, \zeta_N(\omega)) \in \Gamma := (\times_{j=1}^N \Gamma_j) \subset \mathbb{R}^N$ of indipendent random variables ζ_j.

Quasi-Monte Carlo uses better distributed points and (+ some regularity conditions on the map $\zeta \to y(\zeta)$) improves convergence rate up to $M^{-1+\delta}$, $\delta > 0$ (see Kuo, Schwab, Sloan, SINUM, 2012).



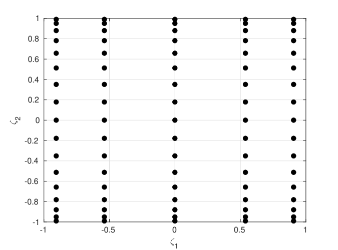
Monte Carlo (left) vs Quasi-Monte Carlo (right) nodes

Solve a Random PDE - - Stochastic Collocation

Idea: build interpolation of the map $\zeta \in \Gamma \subset \mathbb{R}^N \to y(\zeta) \in Y$, where $y(\zeta)$ solves the PDE.

$$y_{SC}(x, \zeta) = \sum_{k=1}^{Q} y(\zeta_k, x) l_k(\zeta),$$

where $\{l_k\}$ is the Lagrangian basis associated to $\{\zeta_k\}_{k=1}^{Q}$.



Under suitable regularity conditions, exponential convergence (but dimension dependent)

$$\|y - y_{SC}\|_{L^2(\Gamma;Y)} \leq C_{SC} \sum_{n=1}^{N} e^{-r_n \beta_n}$$

$\boldsymbol{\beta} = (\beta_1, \dots, \beta_N)$: β_n is # of quadrature points for the n-th random variable. $\{r_j\}_{j=1}^{N}$ parameters linked to the holomorphic extension of $\zeta \to y(\zeta)$. (See Babuska, Nobile, Tempone, SIREV, 2010).

Solve a Random PDE - - Stochastic Galerkin

Stochastic Galerkin methods are based on the global space-probability weak formulation

$$\int_\Gamma \int_D \kappa(x,\zeta)\nabla y(x,\zeta)\nabla v(x,\zeta)\rho(\zeta)dxd\zeta = \int_\Gamma \int_D u(x)v(x,\zeta)\rho(\zeta)dxd\zeta, \quad \forall v(x,\zeta) \in L^2(\Gamma; Y)$$

- Intrusive method: cannot recycle standard FEM codes. Need for tailor solvers.
- Need to define an appropriate Galerkin subspace $V_{h,k} \subset L^2(\Gamma; Y)$.
- Weak formulation not always well-defined: $\kappa_{\min}^{-1}(\omega)$ and $\kappa_{\max}(\omega)$ are not in $L^\infty(\Omega)$ (See Schwab, Gittelson, Acta, 2011).
- Equivalence between global weak formulation and pathwise formulation if $\kappa_{\min}^{-1}(\omega)$ and $\kappa_{\max}(\omega)$ are in $L^\infty(\Omega)$.
- Under suitable assumptions, convergence is similar to Stochastic Collocation (see Back, Nobile, Tamellini, Tempone, (2011) for a comparison).

Quantity of interest

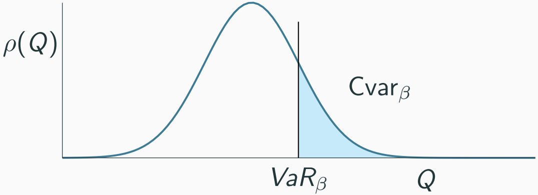
Let $y(x,\zeta) \in L^p(\Omega; Y)(i.e.\|y\|^p_{L^p(\Omega;Y)} = \int_\Gamma \|y(\cdot,\zeta)\|^p_Y d\rho(\zeta) < \infty)$ $(Y \subset L^2(D))$.
Quantity of interest: $Q : L^p(\Omega; Y) \to L^q(\Omega; \mathbb{R})$.
Examples:

- Tracking term: $Q(y(\zeta)) = \|y(\cdot,\zeta) - y_d\|^2_{L^2(D)}$ with $y_d \in L^2(D)$.
  Exercise: $Q : L^p(\Omega; L^2(D)) \to L^q(\Omega; \mathbb{R})$, with $p \geq 2$ and $q \leq p/2$.
- Elastic energy: $Q(y(\zeta)) = \int_D f y(x,\zeta) dx$.
  Exercise: $Q : L^p(\Omega; L^2(D)) \to L^q(\Omega; \mathbb{R})$, with $p \geq 1$ and $q \leq p$.

Risk measures

$\mathcal{R} : L^q(\Omega; \mathbb{R}) \to \mathbb{R}.$

- $\mathcal{R}(Q) = \mathbb{E}[Q]$ (mean),
- $\mathcal{R}(Q) = \mathbb{E}[Q] + \gamma \mathrm{Var}[Q]$ (mean plus variance),
- $\mathcal{R}(Q) = VaR_\beta[Q] = \inf\{x \in \mathbb{R} : P(Q < x) \geq \beta\}$ (Value at Risk),
- $\mathcal{R}(Q) = CVaR_\beta(Q) := \mathbb{E}[Q | Q \geq VaR_\beta(Q)]$ (Conditional Value at Risk),
- $\mathcal{R}(Q) = \sigma^{-1} \log(\mathbb{E}[exp(\sigma Q)])$ (Entropy risk measure).



Canonical problem

Set $Y = H^1(\Omega)$, $U = L^2(D)$, $\mathcal{Y} = L^2(\Omega; Y)$.

$$\min_{y \in \mathcal{Y}, u \in U} \frac{1}{2}\mathbb{E}\left[\|y - y_d\|^2_{L^2(D)}\right] + \frac{\nu}{2}\|u\|^2_U,$$
$$\text{s.t. } \langle e(y, u, \zeta), v\rangle = 0, \quad \forall v \in Y, \text{for a.e. } \zeta \in \Gamma.$$

Existence of a control u can be derived using calculus of variation, see:

- Frutos, Esparza, *Optimal control of PDEs under uncertainty*, Springer, 2018.
- Antil, Kouri, Lacasse, Ridzal, *Frontiers in PDE-Constrained optimization*, Springer, 2018.
- Kouri, Surowiec, *Existence and optimality conditions for risk-averse PDE-constrained optimization*, SIAM/ASA J. UQ., 2018.

Numerical approximation and solution strategies

Sample Average Approximation (SAA)

Replace $\mathbb{E}[X]$ with $\widehat{\mathbb{E}}[X] \approx \sum_{j=1}^{M} w_j X(\zeta_j)$:

- Monte Carlo: $\{\zeta_j\}_{j=1}^{M}$ drawn randomly, $w_j = \frac{1}{M}$.
- Quasi Monte Carlo: $\{\zeta_j\}_{j=1}^{M}$ drawn according to low discrepancy sets, $w_j = \frac{1}{M}$.
- Stochastic Collocation: $\{\zeta_j\}_{j=1}^{M}$ according to zeros of orthogonal polynomials. $w_j = \int_\Gamma l_j(\zeta) d\zeta$ (Gaussian weights).

Semi-discrete OCPUU:

$$\min_{y \in (\times_{j=1}^{M} Y), u \in U} \frac{1}{2} \widehat{\mathbb{E}} \left[\|y - y_d\|_{L^2(D)}^2 \right] + \frac{\nu}{2} \|u\|_U^2,$$

$$\text{s.t. } \langle e(y, u, \zeta_j), v \rangle = 0, \quad \forall v \in Y, \ j = 1, \ldots, M.$$

Variables are $(\boldsymbol{y}, u) = (y_1, y_2, \ldots, y_M, u) \in (\times_{j=1}^{M} Y) \times U$.

Optimality conditions: adjoint calculus

$\mathcal{L}(\boldsymbol{y}, u, \boldsymbol{p}) = \frac{1}{2}\widehat{\mathbb{E}}\left[\|y - y_d\|^2_{L^2(D)}\right] + \frac{\nu}{2}\|u\|^2_U + \sum_{j=1}^{M}\langle e(y, u, \zeta_j), p_j\rangle.$
Compute $\partial \boldsymbol{y}\mathcal{L}$, $\partial u\mathcal{L}$ and $\partial_{\boldsymbol{p}}\mathcal{L}$,

$$
\begin{aligned}
\langle e(y, u, \zeta_j), v\rangle &= 0, \quad \forall v \in Y, \qquad i = 1, \dots, M && \text{(state equations)}\\
(\nu u, v) - (\widehat{\mathbb{E}}\left[\boldsymbol{p}\right], v)_{L^2(D)} &= 0, \quad \forall v \in U && \text{(optimality condition)}\\
\langle e(p_j, u, \zeta_j), v\rangle &= (y_j - y_d, v)_{L^2(D)}, \quad \forall v \in Y, \qquad i = 1, \dots, M && \text{(adjoint equations)}
\end{aligned}
$$

NB: We used that PDE is linear and self-adjoint.

Linear optimality systems

- Full-space strategies involve a linear system of dimension $(2M+1)N_h$ (Nobile, V., arXiv, 2110.07362).

$$\begin{pmatrix} \mathcal{M} & 0 & \mathcal{E} \\ 0 & \nu M_s & -M_s I^\top \\ \mathcal{E} & -IM_s & 0 \end{pmatrix} \begin{pmatrix} \mathrm{y} \\ \mathrm{u} \\ \mathrm{p} \end{pmatrix} = \begin{pmatrix} \mathrm{f}_p \\ 0 \\ \mathrm{f}_y \end{pmatrix},$$

  $\mathcal{E} = \mathrm{diag}\left(w_1 E(\zeta_1), \dots, w_M E(\zeta_M)\right), \mathcal{M} = \mathrm{diag}\left(w_1 M_s, \dots, w_M M_s\right)$, and $I = \left(w_1 I_s, \dots, w_M I_s\right)^\top$.

- Iterative reduced strategies requires to invert $2M$ PDE at each iteration. (Kouri PhD thesis, 2014 - Kouri et al., SISC 2013).

$$\left(\nu M_s + \sum_{j=1}^{M} w_j M_s E(\zeta_j)^{-1} M_s E(\zeta_j)^{-1} M_s\right) \mathrm{u} = \mathrm{f}.$$

Preconditioners for the full-space linear system

$$\mathcal{S} = \begin{pmatrix} \mathcal{M} & 0 & \mathcal{E} \\ 0 & \nu M_s & -M_s I^\top \\ \mathcal{E} & -I M_s & 0 \end{pmatrix} = \begin{pmatrix} C & B^\top \\ B & 0 \end{pmatrix},$$

with $C = \operatorname{diag}(\mathcal{M}, \nu M_s)$, $\mathcal{B} = (\mathcal{E}, -I M_s)^\top$.

Optimal preconditioner[1]: $P = \begin{pmatrix} C & \\ & B C^{-1} B^\top \end{pmatrix} = \begin{pmatrix} \mathcal{M} & & \\ & \nu M_s & \\ & & \mathcal{E}\mathcal{M}^{-1}\mathcal{E} + \frac{1}{\nu} I M_s I^\top \end{pmatrix}$

Feasible preconditioner: $\widetilde{P} = \begin{pmatrix} \mathcal{M} & & \\ & \nu M_s & \\ & & \mathcal{E}\mathcal{M}^{-1}\mathcal{E} \end{pmatrix}$ For the random diffusion equation:

- Spectrum of $\widetilde{P}^{-1}\mathcal{S}$ depends on the moments $\|\kappa_{\min}^{-1}\|_{L^p(\Omega)}$, $\|\kappa_{\max}\|_{L^p(\Omega)}$.
- The preconditioner is robust w.r.t. finer discretization in probability, but not w.r.t to ν.
- Highly parallelizable since $\mathcal{M}$ and $\mathcal{E}$ are block-diagonal.

[1] Murphy, Golub, Wathen, SISC, 2000. MINRES converges in at most 4 iterations.

Approximation properties - I

Let $u_\star$ solution of the original OCP, and $\widehat{u_\star}$ be the solution of the semi-discrete *OCP*.
Let $J(u) = \frac{1}{2}\mathbb{E}\left[\|y(\omega,u)-y_d\|^2_{L^2(D)}\right] + \frac{\nu}{2}\|u\|^2_U$ and $\widehat{J}(u) = \frac{1}{2}\widehat{\mathbb{E}}\left[\|y(\omega,u)-y_d\|^2_{L^2(D)}\right] + \frac{\nu}{2}\|u\|^2_U$.
Optimal solutions satisfy

$$\langle \nabla J(u_\star), v\rangle = \langle \nu u_\star + \mathbb{E}\left[p(u_\star)\right], v\rangle = 0, \quad \langle \nabla \widehat{J}(\widehat{u}_\star), v\rangle = \langle \nu \widehat{u}_\star + \widehat{\mathbb{E}}\left[p(\widehat{u}_\star)\right], v\rangle = 0.$$

Hence choosing $v = \widehat{u}_\star - u_\star$,

$$\langle \nu(u_\star - \widehat{u}_\star) + \mathbb{E}\left[p(u_\star,\omega)\right] - \widehat{\mathbb{E}}\left[p(\widehat{u}_\star,\omega)\right], \widehat{u}_\star - u_\star\rangle = 0, \quad \text{leads to}$$

$$\nu\|u_\star - \widehat{u}_\star\|^2_{L^2(D)} = \langle \mathbb{E}\left[p(u_\star)\right] - \widehat{\mathbb{E}}\left[p(u_\star)\right] + \widehat{\mathbb{E}}\left[p(u_\star)\right] - \widehat{\mathbb{E}}\left[p(\widehat{u}_\star)\right], \widehat{u}_\star - u_\star\rangle$$

First term: $\langle \mathbb{E}\left[p(u_\star)\right] - \widehat{\mathbb{E}}\left[p(u_\star)\right], \widehat{u} - u_\star\rangle \leq \frac{1}{2\nu}\|\mathbb{E}\left[p(u_\star)\right] - \widehat{\mathbb{E}}\left[p(u_\star)\right]\|^2_{L^2(D)} + \frac{\nu}{2}\|u_\star - \widehat{u}_\star\|^2_{L^2(D)}$.
Second term: $\langle \widehat{\mathbb{E}}\left[p(u_\star)\right] - \widehat{\mathbb{E}}\left[p(\widehat{u}_\star)\right], \widehat{u} - u_\star\rangle = \cdots = -\widehat{\mathbb{E}}\left[\|y(u_\star) - y(\widehat{u}_\star)\|^2_{L^2(D)}\right] \leq 0$ (believe me![2]).
Thus,

$$\frac{\nu}{2}\|u_\star - \widehat{u}_\star\|^2_{L^2(D)} \leq \frac{1}{2\nu}\|\mathbb{E}\left[p(u_\star)\right] - \widehat{\mathbb{E}}\left[p(u_\star)\right]\|^2_{L^2(D)}.$$

[2] Use state and adjoint equations path-wise. It holds only if quadrature weights are positive

Approximation properties - II

From $\frac{\nu}{2}\|u_\star - \widehat{u}_\star\|^2_{L^2(D)} \leq \frac{1}{2\nu}\|\mathbb{E}\left[p(u_\star)\right] - \widehat{\mathbb{E}}\left[p(u_\star)\right]\|^2_{L^2(D)} \overset{\text{Jensen+properties Q.F.}}{\leq} \|p - \widehat{p}\|^2_{L^2(\Omega;L^2(D))}$, we conclude according to the parameter regularity of the adjoint equation

- Monte Carlo: $\mathbb{E}\left[\|u^\star - u^\star_{MC}\|^2_{L^2(D)}\right] \leq C(u^\star)\frac{1}{M} \rightarrow$ standard $\frac{1}{\sqrt{M}}$ in the root mean squared sense.
- Quasi Monte Carlo[3]: $\|u^\star - u^\star_{QMC}\|_{L^2(D)} \leq C(u^\star)M^{-1+\delta}$.
- Stochastic Collocation[4]: $\|u^\star - u^\star_{SC}\|_{L^2(D)} \leq C(u^\star)\sum_{j=1}^N e^{-r_j\beta_j}$.

[3] Guth, Kaarnioja, Kuo, Schillings, Sloan, SIAM/ASA J. UQ, 2021
[4] Martin, Krumscheid, Nobile, ESAIM, 2021

Stochastic Approximation[5]

Define the solution operator $S_\zeta : U \to Y$ s.t. $\langle e(S_\zeta u, u, \zeta, v\rangle = 0, \quad \forall v \in Y$, for a.e. $\zeta \in \Gamma$.

$$\min_{u\in U} \frac{1}{2}\sum_{j=1}^{M} \|S_{\zeta_j} u - y_d\|^2_{L^2(D)} + \frac{\nu}{2}\|u\|^2_U = \min_{u\in U} \frac{1}{2}\sum_{j=1}^{M} f_j(\zeta) + \frac{\nu}{2}\|u\|^2_U$$

Stochastic gradient algorithm:

1 Draw randomly a sample $\zeta_j \sim \rho(\zeta)d\zeta$.

2 Solve state and adjoint equation for sample ζ_j with control u^k.

3 $\nabla J_j = \nu u^k + p(\zeta_j, u^k)$.

4 Update: $u^{k+1} = u^k - \frac{\tau_0}{j}\nabla J_j$.

Properties

- Mini-batch variants.
- No approximation error $u^{k+1} \to u_\star$!
- $\mathbb{E}\left[\|u^{k+1} - u_\star\|^2_{L^2(D)}\right] \leq \frac{D}{k}$.

[5] Martin, Krumscheid, Nobile, ESAIM 2021- Geiersbach, Wollner, SISC 2020 -Geiersbach, Plug, SIAM J. Opt, 2019

Other topics

Approximations properties

- Quantify the truncation error: $\kappa \approx \kappa_N(x,\omega) \to \|y_N(x,\omega) - y(x,\omega)\|$.
- Break the curse of dimensionality: $\zeta = \{\zeta_j\}_{j=1}^{\infty}$.
- Techniques to reduce computational cost *sparse grids*, *Multi Level Monte Carlo*, *Combination technique*..

General theory of existence and uniqueness

- measurability issues, well-posedness, loss of compactness, etc..

Nonlinear OCPUU

- Semi-discretization in probability $\to$ use common strategies from nonlinear optimization (e.g. trust-region, SQP).
- Computational cost may be prohibitive $\to$ tailored solvers needed!
- Risk adverse measures can easily introduce nonsmoothness (e.g. Cvar).

References I

- Lord, Powell, Shardlow, *An introduction to computational Stochastic PDEs*, Springer, 2014.
- Kuo, Schwab, Sloan, *Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients*, SINUM, 2012.
- Babuska, Nobile, Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIREV, 2010.
- Schwab, Gittelson, *Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs*, Acta Num., 2011.
- Back, Nobile, Tamellini, Tempone, *Stochastic spectral Galerkin and Collocation methods for PDEs with random coefficients: a numerical comparison*, in Spectral and High Order Methods for Partial Differential Equations, 2011.
- Kouri, Surowiec, *Existence and optimality conditions for risk-averse PDE-constrained optimization*, SIAM/ASA J. UQ., 2018.
- Frutos, Esparza, *Optimal control of PDEs under uncertainty*, Springer, 2018.
- Antil, Kouri, Lacasse, Ridzal, *Frontiers in PDE-Constrained optimization*, Springer, 2018.

- Kouri, Surowiec, *Existence and optimality conditions for risk-averse PDE-constrained optimization*, SIAM/ASA J. UQ., 2018.

- Nobile, Vanzan, *Preconditioners for robust optimal control problems under uncertainty*, arXiv, 2110.07362, 2021.

References II

- Kouri, *An Approach for the Adaptive Solution of Optimization Problems Governed by Partial Differential Equations with Uncertain Coefficients*, PhD thesis, 2012.
- Kouri, Heinekenschloss, Ridzal, Van Bloemen Waanders, *A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty*, SISC, 2013.
- Murphy, Golub, Wathen, *A note on preconditioning for indefinite linear systems*, SISC, 2000.
- Martin, Krumscheid, Nobile, *A complexity analysis of stochastic gradient methods for PDE-constrained optimal control problems with uncertain parameters*, ESAIM, 2021.
- Guth, Kaarnioja, Kuo, Schillings, Sloan, *A Quasi-Monte Carlo method for optimal control under uncertainty*, SIAM/ASA J. UQ, 2021.
- Geiersbach, Wollner, *A stochastic gradient method with mesh refinement for PDE-constrained optimization under uncertainty*, SISC, 2020.

Thank you!

Questions?