Sparse polynomial chaos expansions

Benchmark of compressive sensing solvers and experimental design techniques

N. Lüthen, B. Sudret

ETH Zürich, Chair of Risk, Safety and Uncertainty Quantification, luethen@ibk.baug.ethz.ch

Surrogate modelling by Regression-based PCE

Our objectives

For complex and costly models from science and engineering, find a **surrogate model** that...

- **approximates the model well**
- requires only **few model evaluations**
- is **nonintrusive**

PCE

- **Input** random vector $\boldsymbol{X} = (X_1, \dots, X_d)$ with $X_k \sim f_{X_k}$ and range $\mathcal{D} \subset \mathbb{R}^d$
- **Model** $\mathcal{M} \in L^2_{f_{\boldsymbol{X}}}(\mathcal{D})$
- **Output** random variable $Y = \mathcal{M}(\boldsymbol{X})$
- **Polynomial orthonormal basis** for $L^2_{f_{\boldsymbol{X}}}(\mathcal{D})$: $\{\psi_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathbb{N}^d\}$
- **Truncation** of the **spectral expansion**: Use a subset of multi-indices $\mathcal{A} \subset \mathbb{N}^d$ to approximate

$$Y \approx Y_{\mathrm{PCE}} = \mathcal{M}^{\mathrm{PCE}}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} \boldsymbol{c}_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{X}).$$

Computing a PCE by regression

- **Experimental design (ED)** $\{\boldsymbol{x}^{(i)}\}_{i=1}^{N_{\mathrm{ED}}} \subset \mathcal{D}$
- Model evaluations $\{y^{(i)} = \mathcal{M}(\boldsymbol{x}^{(i)})\}_{i=1}^{N_{\mathrm{ED}}}$
- Multi-index set $\mathcal{A}$ defining the set of **candidate polynomials**
- Matrix $\boldsymbol{\Psi}$ containing evaluations of the polynomials

Solve $\boldsymbol{y} \approx \boldsymbol{\Psi}\boldsymbol{c}$.

> Many real-world models are compressible.

Find a sparse solution $\boldsymbol{c}$, e.g. by Basis Pursuit Denoising:

$$\min_{\boldsymbol{c}} \|\boldsymbol{c}\|_1 \text{ s.t. } \|\boldsymbol{\Psi}\boldsymbol{c} - \boldsymbol{y}\|_2^2 \leq \epsilon$$

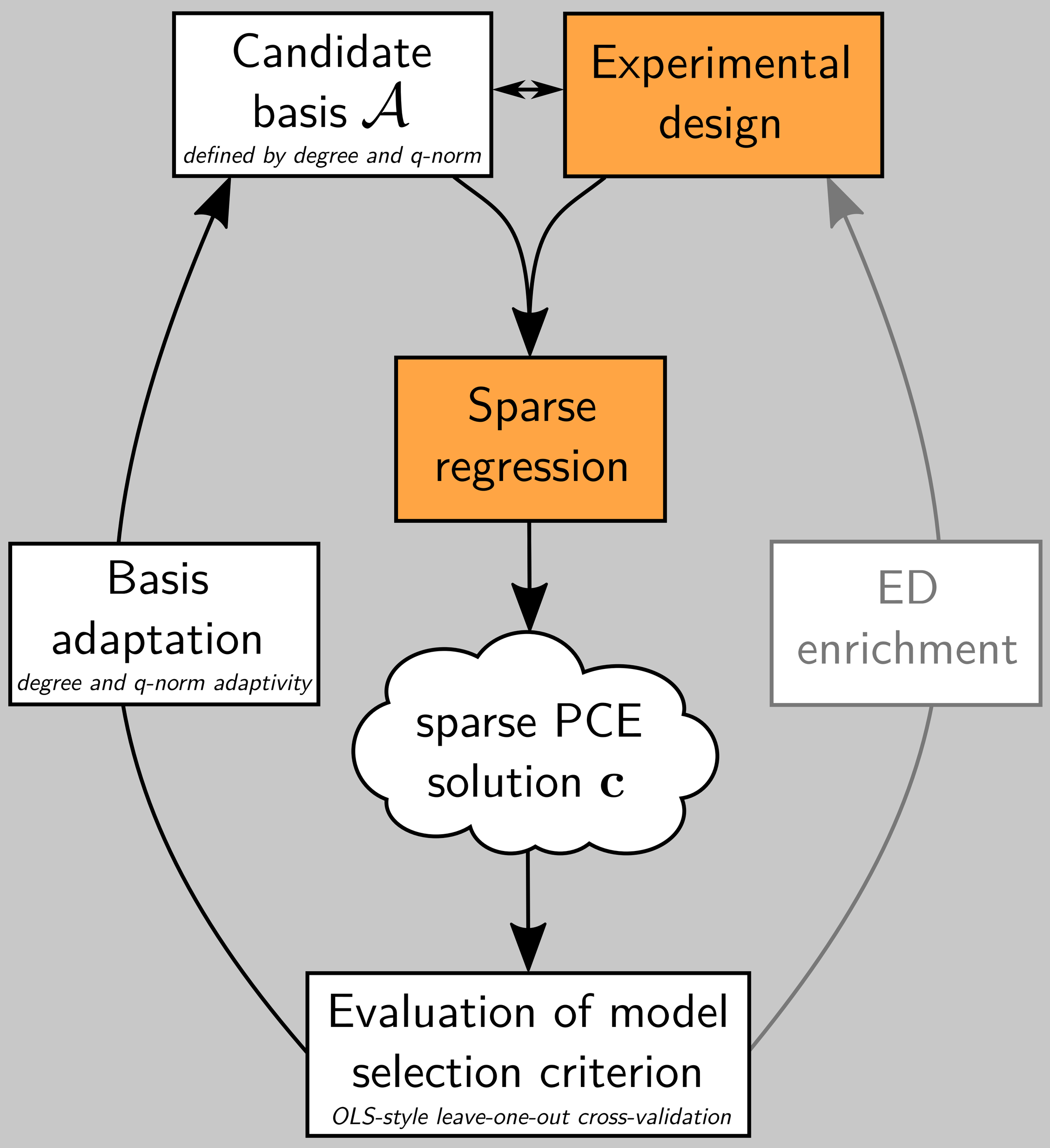
Sparse regression solvers

Sparse regression: Find $\boldsymbol{c}$ sparse with $\|\boldsymbol{\Psi}\boldsymbol{c} - \boldsymbol{y}\|_2$ small.

Solver options:

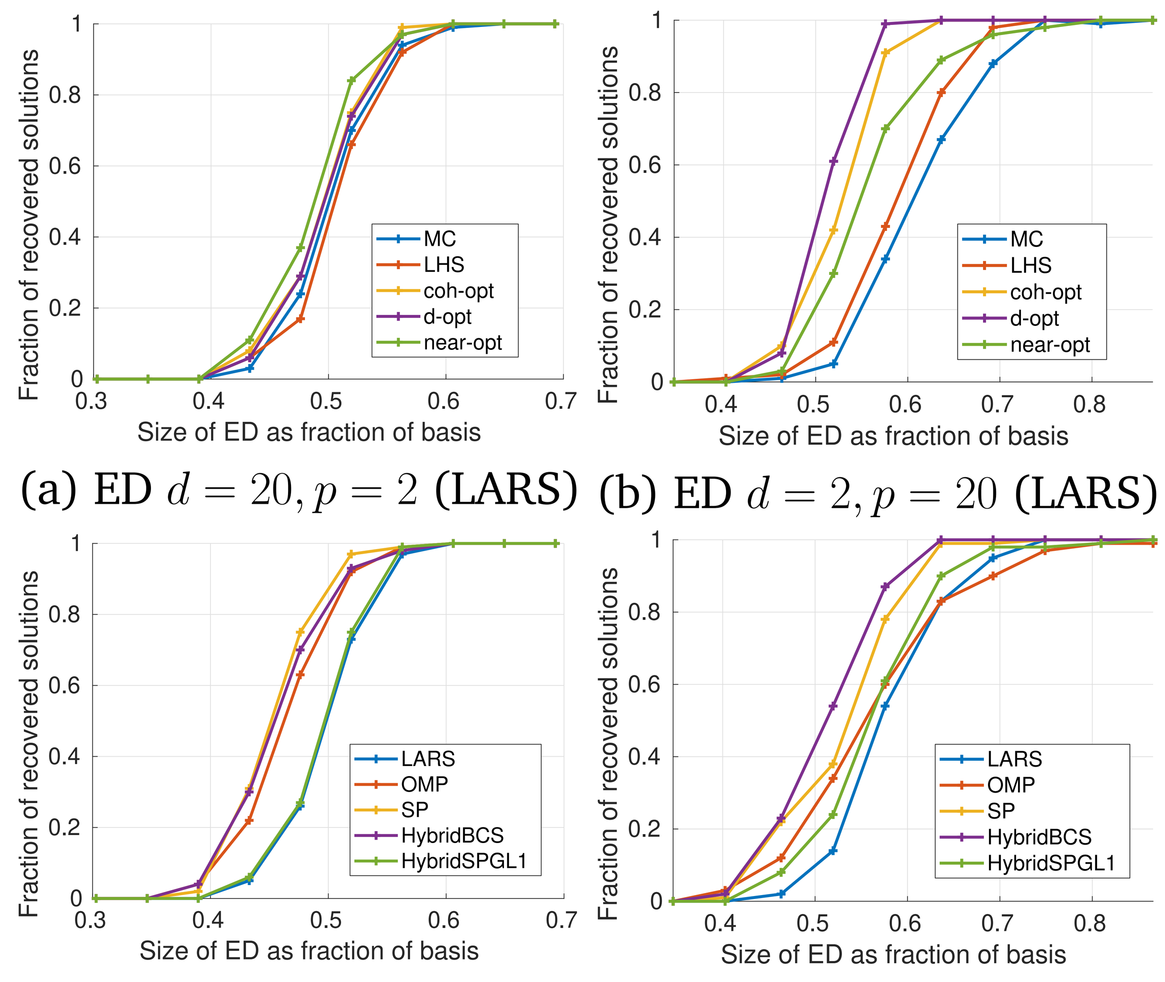
- **Least angle regression (LARS)** Efron, Hastie, Johnstone, Tibshirani 2003; Blatman, Sudret 2011
  - greedy method based on correlation between regressors and residual; using equiangularity
- **Orthogonal matching pursuit (OMP)** Pati, Rezaiifar, Krishnaprasad 1993; Tropp, Gilbert 2007; Berchier 2015
  - "stagewise regression": greedy method based on correlation between regressors and residual
- **Subspace pursuit** Dai, Milenkovic 2003; Diaz, Doostan, Hampton 2018
  - iterative updating of a size K regressor set based on correlation and coefficient magnitude
- **FastLaplace** (Bayesian compressive sensing) Tipping 2001; Babacan, Molina, Katsaggelos 2010
  - Hierarchical Bayesian setting: sparsity-inducing prior on coefficients
- **SPGL1** Van den Berg, Friedlander 2008
  - convex optimization: solving LASSO with spectral projected gradient descent

and more.



Benchmark results

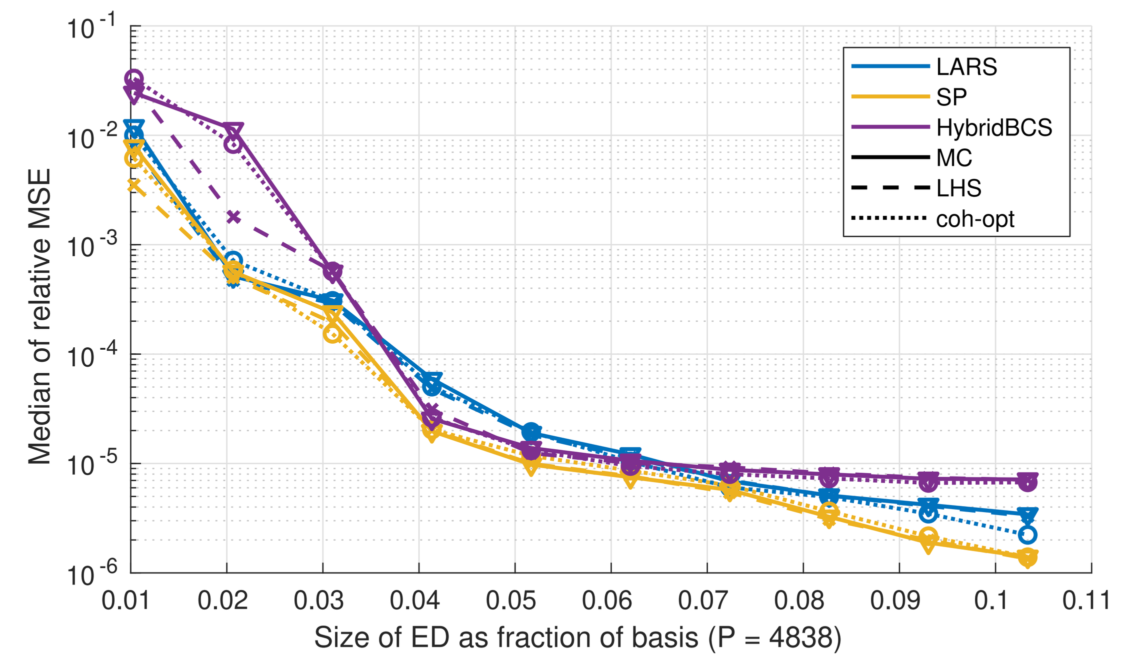
1) Fraction of correctly identified coefficient vector for a manufactured sparse PCE (uniform input, sparsity = 20%)



(a) ED $d = 20, p = 2$ (LARS) (b) ED $d = 2, p = 20$ (LARS)

(c) Solvers $d = 20, p = 2$ (LHS) (d) Solvers $d = 2, p = 20$ (LHS)

2) Validation error for the wingweight function (analytical model, uniform input, $d = 10, p \leq 10, q \leq 0.7$)



Conclusion

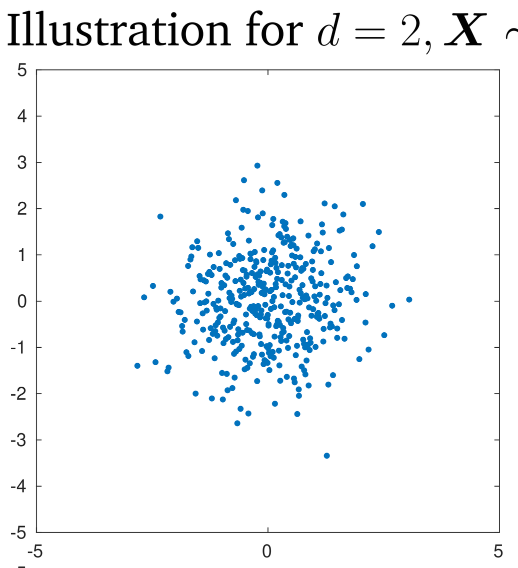
- Solver and sampling type make a difference in the quality of the solution
- Model characteristics influence the solver and sampling performance

Further work

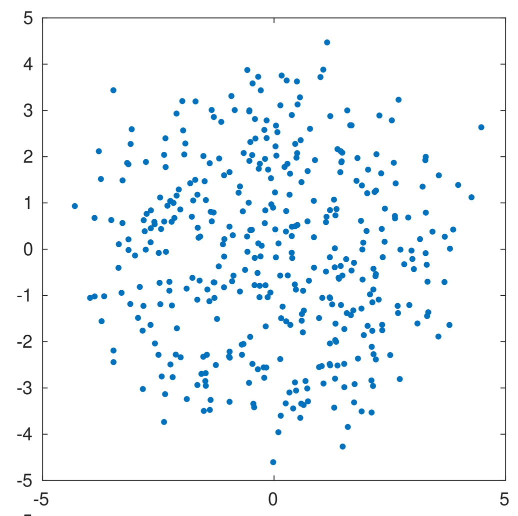
- Characterize the influence of input and model characteristics on the choice of best components for the sparse PCE procedure
- Investigate impact of sequential ED enrichment

Experimental design sampling

Illustration for $d = 2$, $\boldsymbol{X} \sim \mathcal{N}(0, \mathbb{1})$, $p = 20$



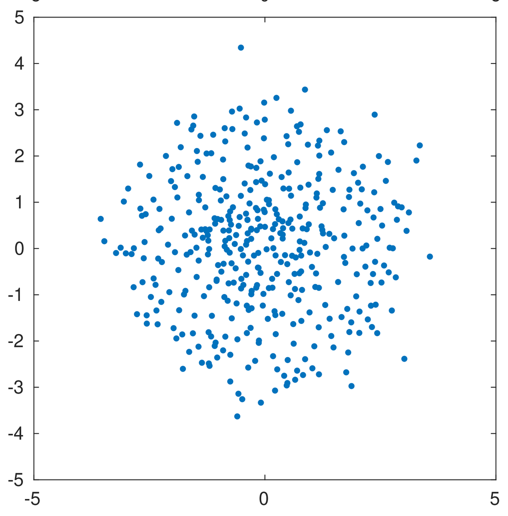
Latin Hypercube Sampling (LHS)



Coherence-optimal

$$\mu(\mathcal{A}) = \max_{\boldsymbol{\alpha} \in \mathcal{A}} \sup_{\boldsymbol{x} \in \mathcal{D}} |\psi_{\boldsymbol{\alpha}}(x)|$$

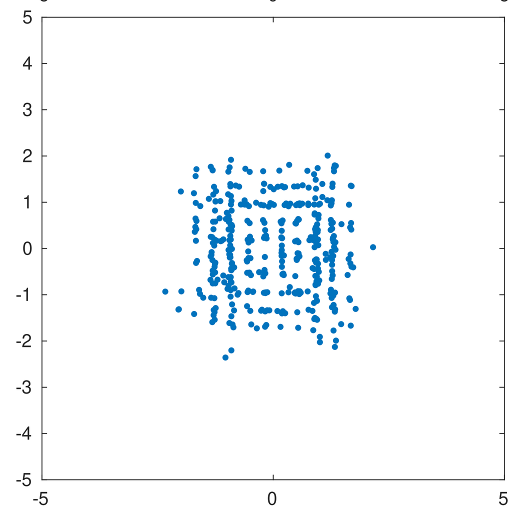
Hampton, Doostan 2015



D-optimal

$D(\boldsymbol{\Psi}) = \det(\boldsymbol{\Psi}^T \boldsymbol{\Psi})$

Diaz, Hampton, Doostan 2019



Near-optimal

$$\eta(\boldsymbol{\Psi}) = \max_{i \neq j} \frac{|\boldsymbol{\Psi}_i^{\mathrm{col}\,T} \boldsymbol{\Psi}_j^{\mathrm{col}}|}{\|\boldsymbol{\Psi}_i^{\mathrm{col}}\|_2 \|\boldsymbol{\Psi}_j^{\mathrm{col}}\|_2}$$

$$\gamma(\boldsymbol{\Psi}) = \frac{1}{P(P-1)} \sum_{i \neq j} \frac{|\boldsymbol{\Psi}_i^{\mathrm{col}\,T} \boldsymbol{\Psi}_j^{\mathrm{col}}|}{\|\boldsymbol{\Psi}_i^{\mathrm{col}}\|_2 \|\boldsymbol{\Psi}_j^{\mathrm{col}}\|_2}$$

Alemazkoor, Meidani 2018

Chair of Risk, Safety and Uncertainty Quantification
Stefano-Franscini-Platz 5
CH-8093 Zürich

Risk, Safety & Uncertainty Quantification