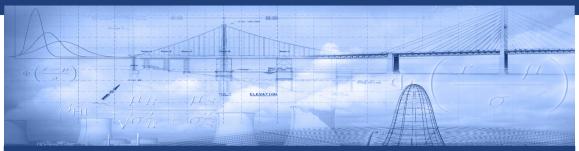
ETHzürich



Poincaré chaos expansions for derivative-enhanced surrogate modelling and sensitivity analysis

UQSay #41

N. Lüthen with O. Roustant, F. Gamboa, B. looss, S. Marelli, and B. Sudret

ETH zürich





Olivier Roustant Berti

Bertrand looss

Fabrice Gamboa

Institut de Mathématiques de Toulouse, France

- · Poincaré inequalities
- · Sensitivity analysis
- ...







Nora Lüthen

• ...

Bruno Sudret

Chair of Risk, Safety and Uncertainty Quantification, ETH Zürich, Switzerland

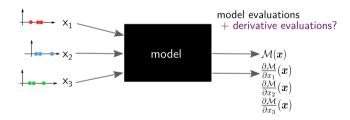
- UQ for engineering models
- Surrogate modelling in particular, sparse polynomial chaos expansions

Global sensitivity analysis using derivative-based sparse Poincaré chaos expansions, https://arxiv.org/abs/2107.00394



Poincaré chaos expansions: Topics

Black-box model Surrogate modelling Sensitivity analysis Using derivative information





Outline

Spectral expansions as surrogate models

Variance-based and derivative-based sensitivity analysis

Poincaré constants and the associated differential operator

Computing Poincaré chaos expansions

Numerical example

Conclusion & Outlook



Chaos expansions as surrogate models

Setting:

- Input random vector \boldsymbol{X} with d independent components and joint distribution $f_{\boldsymbol{X}}$
- Model $\mathcal{M} \in L^2_{f_{\mathcal{X}}}$ (square-integrable)
- Output random variable $Y = \mathcal{M}(\mathbf{X})$

We want to model random variable Y

Let $(\psi_k)_{k\in\mathbb{N}}$ be a basis of $L^2_{f_X}$. Then:

$$Y = \mathcal{M}(\mathbf{X}) = \sum_{k \in \mathbb{N}} c_k \psi_k(\mathbf{X})$$

surrogate model

For example:

Risk, Safety &

- (Fourier expansion)
- · Polynomial chaos expansion
- Poincaré chaos expansion

Approximation of Y by orthogonal polynomials in X (d = 1)

Theorem: Density of polynomials in $L^2_{f_X}(\mathcal{D})$

Assume that X possesses finite moments of all orders, and that F_X is continuous. If the distribution function is uniquely defined by the sequence of its moments, then the polynomials are dense in $L^2_{f_X}(\mathcal{D})$.

Hermite chaos

Wiener (1938); Ghanem, Spanos (1991)

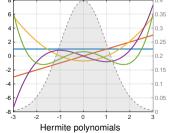
• X Gaussian \rightarrow Hermite polynomials:

$$\psi_0(x) = 1, \ \psi_1(x) = x, \ \psi_2(x) = \frac{x^2 - 1}{2}, \dots$$

Generalized chaos

Risk, Safety 6

- Xiu, Karniadakis (2002)
- X uniform \rightarrow Legendre polynomials
- * $X \operatorname{Beta} \to \operatorname{Jacobi polynomials}$
- X Gamma \rightarrow Laguerre polynomials



Approximation of Y by orthogonal polynomials in X (d = 1)

Ernst et al (2012)

Theorem: Density of polynomials in $L^2_{f_X}(\mathcal{D})$

Assume that X possesses finite moments of all orders, and that F_X is continuous. If the distribution function is uniquely defined by the sequence of its moments, then the polynomials are dense in $L^2_{f_X}(\mathcal{D})$.

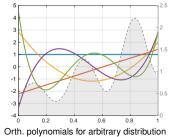
Notable exception: lognormal distribution!

Arbitrary chaos Wan and Karniadakis (2006); Oladyshkin and Nowak (2012)

 One can compute an orthogonal polynomial basis for any distribution that fulfills the assumptions (e.g., with compact support)

By the way: the term *polynomial chaos* goes back to Wiener (1938)

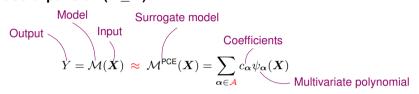
 \rightarrow Use of the word "chaos" older than Chaos theory in mathematics! (1938 vs 1977)





Risk, Safety &

Polynomial chaos expansion ($d \ge 1$)



with tensor product basis functions

 $\psi_{\alpha}(x) = \prod_{i=1}^{d} \psi_{\alpha_i}^{(i)}(x_i),$ where the multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$ defines the degree

and set of multi-indices \mathcal{A} , e.g., total-degree basis of degree p:

$$\mathcal{A} = \{ \boldsymbol{\alpha} \in \mathbb{N}^d : \sum_{i=1}^d \alpha_i \leq p \}$$

• If for each X_i the moment problem is uniquely solvable, then the multivariate polynomials are dense in $L^2_{f_X}(\mathcal{D})$ and this approximation converges in mean-square to Y Ernst et al (2012)

How to compute a PCE?

$$Y = \mathcal{M}(\boldsymbol{X}) \approx \mathcal{M}^{\mathsf{PCE}}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} c_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$$

Ingredients of a PCE:

- (Basis functions $\{\psi_{oldsymbol lpha}: oldsymbol lpha \in \mathbb{N}^d\}$ defined by input distribution.)
- Need to decide subset of multi-indices $\mathcal{A} \subset \mathbb{N}^d$
- Need to choose points $x \in \mathcal{X} \subset \mathcal{D}$ (experimental design) and collect the corresponding model evaluations $y = \mathcal{M}(x)$
- Need to compute the coefficients c
 - Projection:

$$c_{\alpha} = \langle \mathcal{M}, \psi_{\alpha} \rangle$$
 integration in *d* dimensions

- Regression:

Risk, Safety &

$$c = \min_{c'} \left\| oldsymbol{y} - oldsymbol{\Psi} c'
ight\|_2 (+ ext{ regularization}) ext{ properties of } oldsymbol{\Psi} ext{ are crucia}$$

Poincaré chaos expansions

Outline

Spectral expansions as surrogate models

Variance-based and derivative-based sensitivity analysis

Poincaré constants and the associated differential operator

Computing Poincaré chaos expansions

Numerical example

Conclusion & Outlook



Sobol'-Hoeffding / ANOVA decomposition

Any $\mathcal{M} \in L^2_{f_X}$ can be decomposed uniquely as a sum of terms of increasing complexity

$$\mathcal{M}(\mathbf{X}) = m_0 + \sum_{1 \le i \le d} m_i(X_i) + \sum_{1 \le i < j \le d} m_{i,j}(X_i, X_j) + \dots + m_{1,\dots,d}(X_1, \dots, X_d)$$

where the terms satisfy $\int m_I(\boldsymbol{X}_I) f_{X_k}(x_k) dx_k = 0$ for all $k \in I \subset \{1, \ldots, d\}$.

Variance decomposition

$$\operatorname{Var}\left[\mathcal{M}(\boldsymbol{X})\right] = \sum_{1 \le i \le d} \operatorname{Var}\left[m_i(X_i)\right] + \sum_{1 \le i < j \le d} \operatorname{Var}\left[m_{i,j}(X_i, X_j)\right] + \dots + \operatorname{Var}\left[m_{1,\dots,d}(X_1, \dots, X_d)\right]$$

 \rightarrow ANalysis Of VAriance decomposition



Sobol' indices

Variance decomposition

$$\underbrace{\operatorname{Var}\left[\mathcal{M}(\boldsymbol{X})\right]}_{i=D} = \sum_{1 \le i \le d} \operatorname{Var}\left[m_i(X_i)\right] + \sum_{1 \le i < j \le d} \operatorname{Var}\left[m_{i,j}(X_i, X_j)\right] + \dots + \operatorname{Var}\left[m_{1,\dots,d}(X_1, \dots, X_d)\right]$$

total variance

First-order Sobol' index:

$$S_i^1 = \frac{\operatorname{Var}\left[m_i(X_i)\right]}{D},$$

Total Sobol' index:

$$S_{i}^{\mathsf{tot}} = \frac{1}{D} \sum_{J: i \in J} \operatorname{Var}\left[m_{J}(X_{J})\right]$$



ETHzürich

Risk, Safety &

PCE \heartsuit Sobol' indices

Sudret (2008)

The ANOVA decomposition of a PCE $\mathcal{M}^{\mathsf{PCE}}(X) = \sum_{\pmb{lpha} \in \mathbb{N}^d} c_{\pmb{lpha}} \psi_{\pmb{lpha}}(X)$ is given by

$$m_{I}(\boldsymbol{X}) := \sum_{\substack{\boldsymbol{\alpha}: \alpha_{i} > 0, i \in I \\ \alpha_{j} = 0, j \notin I}} c_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$$

From orthonormality in $L_{f_{\mathcal{X}}}^2$ it follows that

$$\operatorname{Var}\left[m_{I}(\boldsymbol{X})\right] = \sum_{\substack{\boldsymbol{\alpha}:\alpha_{i} > 0, i \in I\\\alpha_{i} = 0, j \notin I}} c_{\boldsymbol{\alpha}}^{2}$$

and the total variance and the Sobol' indices are given by

$$D = \sum_{\alpha \neq 0} c_{\alpha}^{2}, \qquad \qquad S_{i}^{1} = \frac{1}{D} \sum_{\substack{\alpha:\alpha_{i} > 0, \\ \alpha_{j} = 0, j \neq i}} c_{\alpha}^{2}, \qquad \qquad S_{i}^{\text{tot}} = \frac{1}{D} \sum_{\alpha:\alpha_{i} > 0} c_{\alpha}^{2}.$$

Any tensor-product orthonormal basis, made from 1D bases that each contain the constant function, allows the same construction

Derivative-based global sensitivity measure (DGSM)

Another sensitivity measure: DGSM

$$\nu_i = \mathbb{E}\left[\left(\frac{\partial \mathcal{M}}{\partial x_i}(\boldsymbol{X})\right)^2\right] = \int_{\mathcal{D}} \left(\frac{\partial \mathcal{M}}{\partial x_i}(\boldsymbol{x})\right)^2 f_{\boldsymbol{X}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \left\|\frac{\partial \mathcal{M}}{\partial x_i}\right\|^2.$$

Relation to Sobol' indices:

Sobol and Kucherenko (2009); Lamboni et al. (2013)

$$S_i^{\text{tot}} D \leq C_P \nu_i$$

with Poincaré constant C_P of measure $f_{X_i} dx_i$ \rightarrow low-cost variable screening



Outline

Spectral expansions as surrogate models

Variance-based and derivative-based sensitivity analysis

Poincaré constants and the associated differential operator

Computing Poincaré chaos expansions

Numerical example

Conclusion & Outlook



The Poincaré constant in 1D

Definition: The Poincaré constant C_P associated to a measure μ is the best possible constant C with

$$\int g^2 d\mu \le C \int (g')^2 d\mu \tag{1}$$

for all $g \in H^1_\mu$ with $\int g d\mu = 0$.

"A function with a small (weak) derivative (in the sense of μ) is close to a constant function (in the sense of μ)."

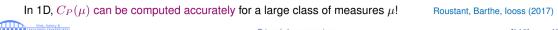
Useful for:

• . . .

- · Bounding total Sobol' indices
- Convergence rate of Markov chains Quantifying multimodality of μ

Lamboni et al (2013)

Pillaud-Vivien et al (2020)



Eigenproblem for Poincaré differential operator

Assumption

Assume that f_X is supported on a bounded interval (a, b) and that $f_X(x) = e^{-V(x)}$ with V continuous and piecewise C^1 on [a, b].

Theorem: 1D Poincaré basis

Under this assumption, for the solutions of the eigenproblem

$$egin{aligned} L\psi &:= \psi^{\prime\prime} - V^{\prime}\psi^{\prime} = -\lambda\psi, \ \psi^{\prime}(a) &= \psi^{\prime}(b) = 0 \end{aligned}$$

it holds that

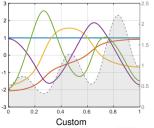
Risk, Safety &

- The eigenfunctions $(\psi_k)_{k\geq 0}$ form an orthonormal basis of $L^2_{f_X}$
- Eigenvalues: $0 = \lambda_0 < \lambda_1 < \ldots \rightarrow \infty$
- $\lambda_0 = 0$ and $\psi_0(x) = 1$
- $C_P(f_X) = \frac{1}{\lambda_1}$, and ψ_1 attains equality in Eq. (1)

Poincaré basis (1D)

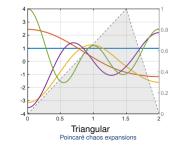
Roustant, Gamboa, looss (2020)





- · In general not polynomial
 - Exception: f_X Gaussian \rightarrow Hermite polynomials
 - f_X uniform leads to cosine basis functions (Fourier basis)
- · Behavior similar to polynomials:
 - ψ_k has k zeros, i.e., higher-order functions oscillate more
- If $f_X \in C^m$, $\psi_k \in C^{m+1}$ ($f_X \in C^0$ by assumption)





Poincaré basis (1D)

Special property

$$\left\langle g',\psi_{k}'
ight
angle _{f_{X}}=\lambda_{k}\left\langle g,\psi_{k}
ight
angle _{f_{X}}$$
 for all $g\in H_{f_{X}}^{1}$

• Consequence: the derivatives of the Poincaré basis form again an orthogonal basis of $L_{f_X}^2$, i.e., an orthogonal system that is dense in $L_{f_X}^2$:

$$\left\langle \psi_{j}^{\prime},\psi_{k}^{\prime}\right\rangle _{f_{X}}=\lambda_{k}\left\langle \psi_{j},\psi_{k}\right\rangle _{f_{X}}=\begin{cases}\lambda_{k},\text{ if }j=k\\0\text{ else}\end{cases}$$

- \rightarrow Well suited for dealing with derivatives.
- · The Poincaré basis is the only basis with this property!

Lüthen et al (2021, preprint)



Outline

Spectral expansions as surrogate models

Variance-based and derivative-based sensitivity analysis

Poincaré constants and the associated differential operator

Computing Poincaré chaos expansions

Numerical example

Conclusion & Outlook



Poincaré chaos expansion (\geq 1D)

Define the Poincaré chaos expansion (PoinCE) by

$$\mathcal{M}(\boldsymbol{X}) \; \approx \; \mathcal{M}^{\mathsf{PoinCE}}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} c_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\alpha}}(\boldsymbol{X})$$

with $(\psi_{\alpha})_{\alpha \in \mathcal{A}}$ tensorized Poincaré basis associated to $f_{\boldsymbol{X}} = \prod_{i=1}^{d} f_{X_i}$

- Basis for $L^2_{f_X}(\mathcal{D}) \to \text{chaos expansion just like PCE}$
- · In particular, coefficients of PoinCE yield moments and Sobol' sensitivity indices
- Partial derivatives of multivariate basis functions are orthogonal: for all $g \in H^1_{f_x}$,

$$\left\langle \frac{\partial}{\partial x_{i}}\psi_{\boldsymbol{\alpha}}, \frac{\partial}{\partial x_{i}}g \right\rangle_{f_{\boldsymbol{X}}} = \left. \lambda_{i,\alpha_{i}} \left\langle \psi_{\boldsymbol{\alpha}}, g \right\rangle_{f_{\boldsymbol{X}}} \right.$$



Computation of DGSM from Poincaré coefficients

Proposition: DGSM formula for Poincaré chaos

Let $\mathcal{M} = \sum_{\alpha} c_{\alpha} \psi_{\alpha}$ be the expansion of $\mathcal{M} \in H^1_{f_X}$ in the Poincaré basis. Then the DGSM index of \mathcal{M} with respect to X_i is

$$\nu_i = \sum_{\alpha:\alpha_i > 0} \lambda_{i,\alpha_i} (c_{\alpha})^2$$

This is an extension of a previous result for Hermite PCE.

Sudret and Mai (2015)

We obtain lower and upper bounds to total partial variances:

$$\sum_{\boldsymbol{\alpha} \in \boldsymbol{\mathcal{A}}: \alpha_i > 0} (c_{\boldsymbol{\alpha}})^2 \leq S_i^{\text{tot}} D \leq C_P(f_{X_i}) \nu_i = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^d: \alpha_i > 0} \frac{\lambda_{i,\alpha_i}}{\lambda_{i,1}} (c_{\boldsymbol{\alpha}})^2.$$



Poincaré derivative expansion

Roustant, Gamboa, looss (2020)

PoinCE

$$\mathcal{M}(oldsymbol{x})pprox \sum_{oldsymbol{lpha}\in\mathcal{A}}c_{oldsymbol{lpha}}\psi_{oldsymbol{lpha}}(oldsymbol{x})$$

An alternative way to compute the coefficients $(c_{\alpha})_{\alpha \in \mathcal{A}}$: Make use of model partial derivatives

PoinCE-der

Risk, Safety &

$$\frac{\partial}{\partial x_i}\mathcal{M}(\boldsymbol{x}) \approx \sum_{\boldsymbol{\alpha}\in\mathcal{A}, \alpha_i > 0} \tilde{c}_{\boldsymbol{\alpha}} \; \frac{\partial}{\partial x_i} \psi_{\boldsymbol{\alpha}}(\boldsymbol{x})$$

- Theoretically $c_{\alpha} = \tilde{c}_{\alpha}$ for $\alpha \in \mathcal{A}$ with $\alpha_i > 0$
- · In practice (when computed from data) they are not equal!
- Terms with $\alpha_i = 0$ vanish when differentiated w.r.t. x_i
 - + Fewer coefficients to estimate from same number of data points!
 - Some coefficients cannot be estimated from a single partial derivative expansion: unnormalized Sobol indices can be computed, but not e.g. variance
- · Aggregate the coefficients from the partial derivative expansions to get the full picture

Poincaré derivative expansion: Aggregation of coefficients

PoinCE-der-i

$$\frac{\partial}{\partial x_i}\mathcal{M}(\boldsymbol{x}) \approx \sum_{\boldsymbol{\alpha}\in\mathcal{A}, \alpha_i > 0} \tilde{c}_{\boldsymbol{\alpha}}^{\partial,i} \; \frac{\partial}{\partial x_i} \psi_{\boldsymbol{\alpha}}(\boldsymbol{x})$$

- PoinCE-der-*i* computes the coefficients corresponding to the multi-indices $\{\alpha \in A, \alpha_i > 0\}$
- For each multi-index $\alpha \neq 0$, average over all active variables:

$$\tilde{c}^{\partial, \text{avg}}_{\alpha} = \frac{1}{\#\{i: \underbrace{1 \le i \le d, \alpha_i > 0}_{\psi_{\alpha} \text{ is not constant in } x_i}} \sum_{\substack{i: 1 \le i \le d, \\ \alpha_i > 0}} \tilde{c}^{\partial, i}_{\alpha}$$

 \rightarrow The coefficients $(\tilde{c}^{\partial,\text{avg}}_{\alpha})_{\alpha\in\mathcal{A}\setminus\mathbf{0}}$ can be used for computing total variance and sensitivity indices

• For surrogate modelling: Compute the remaining coefficient $\tilde{c}_0^{\partial, avg}$ of the constant term by OLS on the residual

Outline

Spectral expansions as surrogate models

Variance-based and derivative-based sensitivity analysis

Poincaré constants and the associated differential operator

Computing Poincaré chaos expansions

Numerical example

Conclusion & Outlook



Application to dyke cost toy model

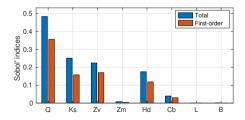
The model describes the cost in million euros given by

$$Y = \mathbb{1}_{S>0} + \left[0.2 + 0.8\left(1 - \exp^{-\frac{1000}{S^4}}\right)\right] \mathbb{1}_{S\le 0} + \frac{1}{20}\left(8\mathbb{1}_{H_d\le 8} + H_d\mathbb{1}_{H_d>8}\right)$$

where H_d is the dyke height and S is the maximal annual overflow given by

$$S = \left(\frac{Q}{BK_s\sqrt{\frac{Z_m - Z_v}{L}}}\right)^{0.6} + Z_v - H_d - C_b$$

Q	Maximal annual flowrate	truncated Gumbel
K_s	Strickler coefficient	truncated Gaussian
Z_v	River downstream level	Triangular
Z_m	River upstream level	Triangular
H_d	Dyke height	Uniform
C_b	Bank level	Triangular
L	Length of river stretch	Triangular
B	River width	Triangular





Methods for estimating sensitivity indices

Sobol' indices

- Sample-based estimation
- Through ANOVA decomposition
 - Poincaré basis, coefficients computed through MC-projection
 Roustant, Gamboa, looss (2020)
 - sparse regression
 Poincaré derivative basis, coefficients computed through MC-projection
 - PCE basis, coefficients computed through sparse regression

DGSM indices

- Sample-based estimation
- Analytical computation from Poincaré derivative basis, coefficients computed through sparse regression
 Lüthen et al (2021, preprint)



Janon et al (2014)

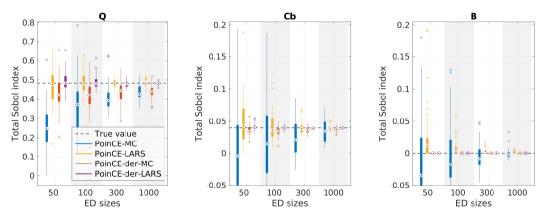
Lüthen et al (2021, preprint)

Kucherenko et al (2009)

ETHzürich

Risk, Safety 6

Results: Sobol' indices – MC-projection vs sparse regression

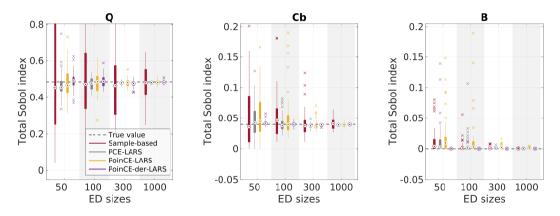


- · Projection underestimates the Sobol' index, sparse regression gives more accurate estimates
- · PoinCE-der estimates have a smaller variance than PoinCE estimates

ETH zürich

Risk, Safety &

Results: Sobol' indices - PoinCE vs PCE and sample-based

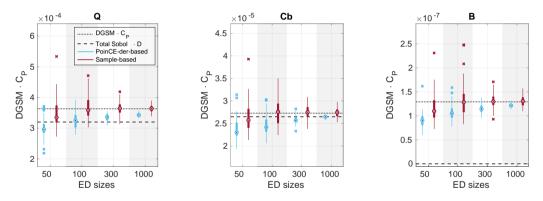


- PoinCE-der outperforms PCE especially for low-importance variables (→ screening)
- · Sample-based estimation shows large variability for important variables

Poincaré chaos expansions

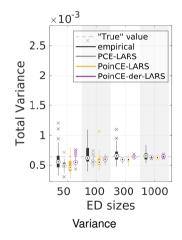
Risk, Safety 6

Results: DGSM indices



- Poincaré-based estimate for DGSM underestimates the true value $(C_P \nu_i = \sum_{\alpha;\alpha_i>0} \frac{\lambda_{i,\alpha_i}}{\lambda_{i,1}} (c_{\alpha})^2)$
- Sample-based DGSM more accurate

Results: Total variance

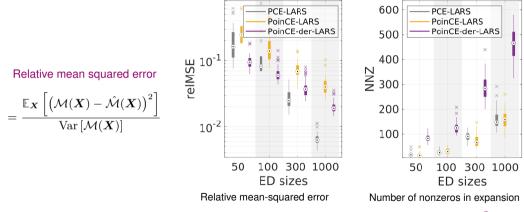


· PoinCE-der (aggregated coefficients) estimates the variance well



ETHzürich

Results: Performance as global surrogate model



• PoinCE-der estimates the variance well, but PCE gives a better approximation in terms of $L_{f_x}^2$ -error



Conclusion

- Poincaré chaos expansions (PoinCE) are like PCE, but with a different basis consisting of the eigenfunctions of the Poincaré differential operator
- · Sobol' indices and DGSM can be computed analytically from the PoinCE coefficients
- The Poincaré basis is the only orthogonal basis for $L^2_{f_X}(\mathcal{D})$ for which the partial derivatives form again an orthogonal basis for the same space
- · PoinCE is well suited to sensitivity analysis and to utilizing derivatives

Outlook:

- Work in progress: Use model evaluations and derivatives at once for computing the coefficients (in the spirit of gradient-enhanced PCE)
- · Non-polynomial basis with special derivative property usefulness for UQ in practice?

Thank you for your attention!



Literature

Polynomial chaos expansion

- Wiener (1938): The homogeneous chaos
- Ghanem, Spanos (1991): Stochastic finite element: A spectral approach
- Xiu, Karniadakis (2002): The Wiener-Askey polynomial chaos for stochastic differential equations
- Wan, Karniadakis (2006): Beyond Wiener-Askey expansions: Handling arbitrary PDFs
- Ernst, Mugler, Starkloff, Ullmann (2012): On the convergence of generalized polynomial chaos expansions
- Oladyshkin, Nowak (2012): Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion

Sensitivity analysis

- Sudret (2008): Global sensitivity analysis using polynomial chaos expansion
- Kucherenko, Rodriguez-Fernandez, Pantelides, Shah (2009): Monte Carlo evaluation of derivative-based global sensitivity measures



Literature

Sensitivity analysis (cont.)

- Lamboni, looss, Popelin, Gamboa (2013): Derivative-based global sensitivity measures: General links with Sobol' indices and numerical tests
- · Janon, Klein, Lagnoux, Nodet, Prieur (2014): Asymptotic normality and efficiency of two Sobol index estimators
- Sudret, Mai (2015): Computing derivative-based global sensitivity measures using polynomial chaos expansions

Poincaré constant and Poincaré chaos expansion

- Roustant, Barthe, looss (2017): Poincaré inequalities on intervals application to sensitivity analysis
- · Roustant, Gamboa, looss (2020): Parseval inequalities and lower bounds for variance-based sensitivity indices
- Pillaud-Vivien, Bach, Lelièvre, Rudi, Stoltz (2020): Statistical Estimation of the Poincaré constant and Application to Sampling Multimodal Distributions
- Lüthen, Roustant, Gamboa, Iooss, Marelli, Sudret (2021): *Global sensitivity analysis using derivative-based sparse Poincaré chaos expansions* (preprint)

