Joint PCE Surrogate Construction with Uncertainty Quantification for Parameterized Stochastic Processes

Joy N. Mueller, Khachik Sargsyan, and Habib N. Najm

14th International Conference on Application of Statistics and Probability in Civil Engineering

9-13 July 2023
Dublin, Ireland

Sandia National Laboratories is a multimission laboratory managed and opera
owned subsidiary of Honeywell International Inc., for the U.S. Department of E
DE-NA0003525.

Acknowledgements

U.S. DEPARTMENT OF ENERGY

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Exascale Catalytic Chemistry

This work was done within the Exascale Catalytic Chemistry (ECC) Project, which is supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Chemical Sciences, Geosciences and Biosciences Division, as part of the Computational Chemistry Sciences Program.

We would like to thank our colleagues ***Judit Zádor, Craig Daniels, Kyungjoo Kim****, and* ***Nathan Roberts*** *for their contributions.*

Sandia National Laboratories

Outline

1. Motivation and Background
2. Surrogate Construction
3. Application to Catalysis
4. Closure

Overview

- Stochastic models: viewed as functions

$$\mathcal{M}_s \; : \; \mathcal{D}_\Lambda \times \Omega \to \mathbb{R}^d$$
$$(\boldsymbol{\lambda}, \omega) \mapsto \mathcal{M}_s(\boldsymbol{\lambda}, \omega)$$

$\mathcal{D}_\Lambda$: domain of input parameters $\boldsymbol{\lambda}$
Ω: event space for intrinsic stochasticity ω

- Want to perform forward/inverse UQ studies of stochastic model $\mathcal{M}_s(\boldsymbol{\lambda}, \omega)$ when parameters are *uncertain*
  - Build approximation, e.g. $f(\boldsymbol{\lambda}, \omega)$, for model response (output QoI)
  - Polynomial chaos (PC) surrogate is the main workhorse
    - treat $\boldsymbol{\lambda}$, ω, $f(\boldsymbol{\lambda}, \omega)$ as random variables/vectors; represent as PCEs
    - global sensitivity analysis (GSA) is free bi-product

Challenges constructing PCEs & tools to tackle them

- Infinite dimensionality: temporal and/or spatial fields $f(\boldsymbol{\lambda}, \omega; \boldsymbol{t})$
  - ➤ Karhunen-Loève expansion
- High-dimensional input (parametric) space
  - ➤ sparse PC regression
- Multiple uncertainty sources in extracted QoIs
  - – assume input parameters $\boldsymbol{\lambda}$ are controllable but ω is not
  - – direct use of conventional PCEs not ideal:
    - ○ usually relies on *smoothness* of input/output map
    - ○ typically constructed for ensemble means/integrated QoIs (expensive, fails to represent intrinsic noise)
    - ○ ignores/fails to represent residual noise and influence of residual noise on model output

➤ represent random variables with stochastic/parametric PCEs[*,**]

[*] Lüthen *et. al.* A spectral surrogate model for stochastic simulators computed from trajectory samples. *Comput. Method. in Appl. Mech. Eng.*, 406:115875, 2023
[**] Zhu *et. al.* Stochastic polynomial chaos expansions to emulate stochastic simulators. *Int. J Uncert. Quantif.*, 2023

Challenges constructing PCEs & tools to tackle them

Idea of joint stochastic-parametric surrogate models is not new

- see Lüthen *et. al.**, Zhu *et. al.***
  - applied to scalar/vector random variables and random fields
  - give overview of general methodology and summary of GSA for stochastic surrogates
- Our focus
  - ensure surrogate remains in PCE framework
  - represent mapping from input to response PDF (generative)
  - compute Sobol indices for parametric/ noise contributions
  - apply methodology to catalysis

* Lüthen *et. al.* A spectral surrogate model for stochastic simulators computed from trajectory samples. *Comput. Method. in Appl. Mech. Eng.*, 406:115875, 2023

** Zhu *et. al.* Stochastic polynomial chaos expansions to emulate stochastic simulators. *Int. J Uncert. Quantif.*, 2023

From Parametric to Stochastic-Parametric PCEs

PCE surrogate *combines* representations of stochastic *and* parametric components to form *joint PCE* surrogate*

$$f(\boldsymbol{\lambda}, \omega) \approx f^{\text{PCE}}(\xi, \zeta) = \sum_{s=0}^{S-1} \underbrace{\left(\sum_{p=0}^{P-1} \boldsymbol{a}_{sp} \Psi_p(\xi) \right)}_{\boldsymbol{z}_s(\boldsymbol{\lambda})} \Psi_s(\zeta)$$

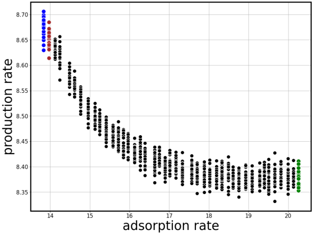
* **JNM** *et. al.* A joint PCE surrogate construction with uncertainty quantification for parameterized stochastic processes applied to heterogeneous catalysis. *In prep.*

From Parametric to Stochastic-Parametric PCEs

PCE surrogate *combines* representations of stochastic *and* parametric components to form *joint PCE* surrogate

$$f(\boldsymbol{\lambda}, \omega) \approx f^{\text{PCE}}(\xi, \zeta) = \sum_{s=0}^{S-1} \underbrace{\left(\sum_{p=0}^{P-1} \boldsymbol{a}_{sp} \Psi_p(\xi) \right)}_{\boldsymbol{z}_s(\boldsymbol{\lambda}^{(n)})} \Psi_s(\zeta)$$

1. Build PCE representations of training samples $f(\boldsymbol{\lambda}^{(n)}, \omega^{(n,m)})$ ($m = 1, ..., M$, n fixed) using *inverse Rosenblatt transformation** *:*



Sandia National Laboratories

From Parametric to Stochastic-Parametric PCEs

PCE surrogate *combines* representations of stochastic *and* parametric components to form *joint PCE* surrogate

$$f(\boldsymbol{\lambda},\omega) \approx f^{\text{PCE}}(\xi,\zeta) = \sum_{s=0}^{S-1} \underbrace{\left(\sum_{p=0}^{P-1} \boldsymbol{a}_{sp} \Psi_p(\xi) \right)}_{\boldsymbol{z}_s(\boldsymbol{\lambda}^{(n)})} \Psi_s(\zeta)$$

1. Build PCE representations of training samples $f(\boldsymbol{\lambda}^{(n)}, \omega^{(n,m)})$ ($m = 1, \ldots, M$, n fixed) using *inverse Rosenblatt transformation**:
   - ➤ inverse CDF functions $\mathcal{F}^{-1}(f)$ capture intrinsic noise ω influence (mapping through $\mathcal{F}$ defined using standard germ ζ)

$$\mathcal{R}^{-1}(\zeta) \; : \; f_i = \mathcal{F}^{-1}_{i|i=1,\ldots,1}(\zeta_i|\zeta_{i-1},\ldots,\zeta_1), \quad i = 1,\ldots,d$$

$$\Longrightarrow f(\boldsymbol{\lambda}^{(n)},\omega) \approx \mathcal{R}^{-1}(\zeta) = \sum_{s=0}^{S-1} \boldsymbol{z}_s(\boldsymbol{\lambda}^{(n)}) \Psi_s(\zeta)$$

* Rosenblatt. Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, 23(3):470 – 472, 1952

Sandia National Laboratories

From Parametric to Stochastic-Parametric PCEs

PCE surrogate *combines* representations of stochastic *and* parametric components to form *joint PCE* surrogate

$$f(\boldsymbol{\lambda}, \omega) \approx f^{\mathrm{PCE}}(\xi, \zeta) = \sum_{s=0}^{S-1} \underbrace{\left(\sum_{p=0}^{P-1} \boldsymbol{a}_{sp} \Psi_p(\xi) \right)}_{\boldsymbol{z}_s(\boldsymbol{\lambda})} \Psi_s(\zeta)$$

2. Define mapping between model inputs $\boldsymbol{\lambda}$ and output $\boldsymbol{z}_s(\boldsymbol{\lambda})$:

$$\boldsymbol{\lambda} \approx \sum_{p=0}^{P'-1} \boldsymbol{c}_p \Psi_p(\xi), \quad \boldsymbol{z}_s(\boldsymbol{\lambda}) \approx \sum_{p=0}^{P-1} \boldsymbol{a}_{sp} \Psi_p(\xi(\boldsymbol{\lambda}))$$

- cast $\boldsymbol{\lambda}$ in terms of germ ξ
- constructed associated PCE for $\boldsymbol{z}_s(\boldsymbol{\lambda})$
- select PCE orders using model evidence
- construct PCE via sparse regression*

* Babacan *et. al.* Bayesian compressive sensing using Laplace priors. *IEEE Trans. Img. Proc.*, 19(1):53–63, 2010

From Parametric to Stochastic-Parametric PCEs

PCE surrogate *combines* representations of stochastic *and* parametric components to form *joint PCE* surrogate*

$$f(\boldsymbol{\lambda}, \omega) \approx f^{\text{PCE}}(\xi, \zeta) = \sum_{s=0}^{S-1} \underbrace{\left(\sum_{p=0}^{P-1} \boldsymbol{a}_{sp} \Psi_p(\xi) \right)}_{\boldsymbol{z}_s(\boldsymbol{\lambda})} \Psi_s(\zeta)$$

- represents *random variables* $f(\boldsymbol{\lambda}, \omega)$ as polynomial functions of *two* sources of randomness ξ (maps to $\boldsymbol{\lambda}$) and ζ
- PCE surrogate can be *resampled*
- analytic expressions for moments and sensitivity indices, GSA for *both* parametric elements and intrinsic noise
- extends to random fields via Karhunen-Loève expansions (KLE)

* **JNM** *et. al.* A joint PCE surrogate construction with uncertainty quantification for parameterized stochastic processes applied to heterogeneous catalysis. *In prep.*

Sandia National Laboratories

Extending joint PCE to random field QoIs

- Find representation for fields, e.g. Karhunen-Loève expansions*

$$f(\boldsymbol{\lambda}, \omega; \boldsymbol{t}) = f_0(\boldsymbol{t}) + \sum_{\ell=1}^{L} \eta_\ell(\boldsymbol{\lambda}, \omega) \sqrt{\mu_\ell} \phi_\ell(\boldsymbol{t})$$

Represent with joint PCE $\eta_\ell(\boldsymbol{\lambda}, \omega) = \sum_{j=0}^{J-1} \tilde{\boldsymbol{a}}_j \Psi_j(\boldsymbol{\xi}, \zeta)$

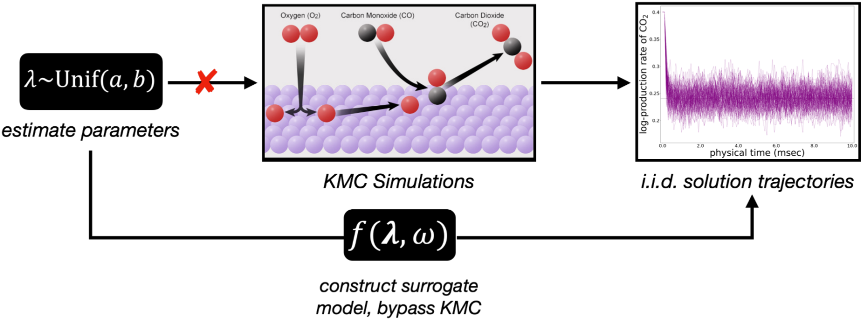
- Construct joint PCE for random coefficients $\eta_\ell(\boldsymbol{\lambda}, \omega)$, rewrite

$$f(\boldsymbol{\lambda}, \omega; \boldsymbol{t}) = \sum_{j=0}^{J-1} \underbrace{\left(f_0(\boldsymbol{t})\delta_{0j} + \sum_{\ell=1}^{L} \tilde{\boldsymbol{a}}_{\ell j} \sqrt{\mu_\ell} \phi_\ell(\boldsymbol{t}) \right)}_{\boldsymbol{c}_j(t)} \Psi_j(\boldsymbol{\xi}, \zeta)$$

* Sargsyan *et. al.* Spectral representation and reduced order modeling of the dynamics of stochastic reaction networks via adaptive data partitioning. *SIAM J Sci. Comput.*, 31(6):4395–4421, 2010

Sandia National Laboratories

Modeling oxidation of CO on RuO_2 with joint PCEs

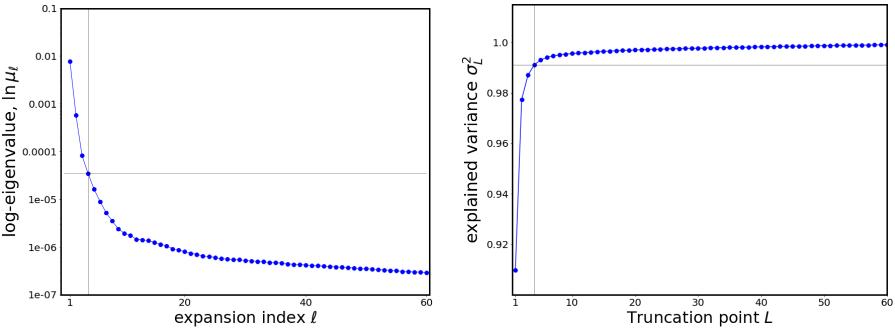


- Multiple molecular species CO, O_2, and CO_2
- *Stochastic* system due to randomness of chemical rections
  - uncertain forward rates collected in vector $\boldsymbol{\lambda} = [\ln k_1, \ldots, \ln k_{15}]$
  - intrinsic stochasticity ω due to randomness of chemical reactions
  - QoI: consumption and production rates over time

Image: https://derekcarrsavvy-chemist.blogspot.com/2017/05/transition-metals-heterogeneous.html

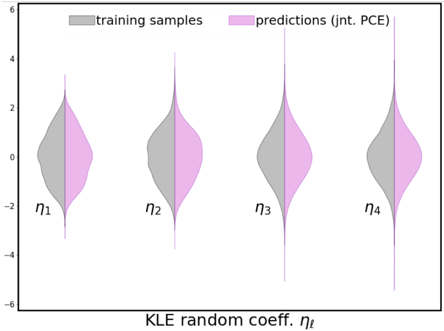
KLE eigenvalue decay indicates room for significant dimension reduction

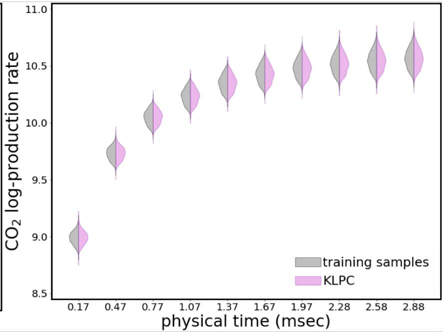
Explained variance in KLE computed from eigenvalues: $\frac{\sum_{\ell=1}^{L} \mu_\ell}{\sum_{\ell=1}^{\mathcal{L}} \mu_\ell}, \quad L \ll \mathcal{L}$



Joint PCE predictions agree in both spectral space and physical (QoI) space

PDF comparisons of combined stochastic-parametric samples

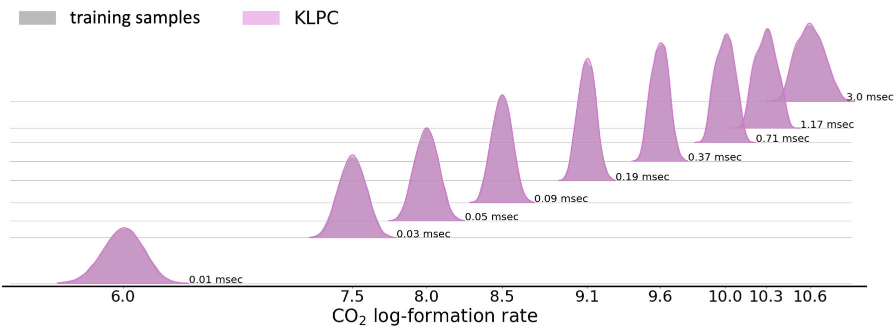




Spectral samples $\eta_\ell(\boldsymbol{\lambda}, \omega)$ *Physical samples* $f(\boldsymbol{\lambda}, \omega; \boldsymbol{t})$

Sandia National Laboratories

Predictions agree with training samples across time



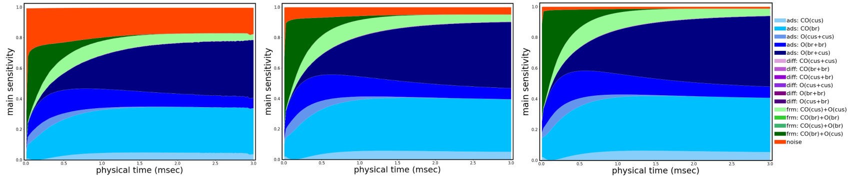
KLPC surrogate model

- provides *reliable* representations physical model
- is *generative*, i.e. can be resampled
- provides samples from correct distribution *at fraction of cost* (sample-intensive UQ tasks are feasible)

Sandia National Laboratories

Sobol indices for output QoI reveal reduction in noise as lattice size increases.

$$S_i = \frac{1}{\mathbb{V}[f(\boldsymbol{t})]} \sum_{j \in \mathcal{J}_{S_i}} c_j(t)^2 ||\Psi_j||^2 = \frac{1}{\mathbb{V}[f(\boldsymbol{t})]} \sum_{j \in \mathcal{J}_{S_i}} \left(\sum_{\ell \le L} \tilde{\boldsymbol{a}}_{\ell j} \sqrt{\mu_\ell}\, \phi_\ell(\boldsymbol{t}) \right)^2 ||\Psi_j||^2$$



25 × 25 lattice   50 × 50 lattice   100 × 100 lattice

Sobol indices over time for adsorption rates (blues), diffusion rates (purples), formation rates (greens), and noise (red).

Zhu *et. al.* Global sensitivity analysis for stochastic simulators based on generalized lambda surrogate models. *Reliab. Eng. Syst Safe.*, 214:107815, 2021

Conclusion

Summary:

- PC-based method of constructing surrogate for stochastic model with parametric uncertainty
- Useful for representing both vector and field QoIs
- GSA attributes output uncertainty to *both* parameters and noise
- Successfully applied to a catalysis problem

On-going Work:

- Improvements to method
  - adaptive basis selection for parametric PCE construction
- Extend to chemical systems with correlated parameters λ
- Utilize information ‘thrown away’ in KLE dimension reduction step

...Thank you!