Predictability and Reduced Order Modeling in Stochastic Reaction Networks

Khachik Sargsyan, Bert Debusschere, Habib Najm, Youssef Marzouk

Biological and Energy Sciences Center
Sandia National Laboratories
Livermore, CA, USA

Acknowledgement

DOE Office of Science, Advanced Scientific Computing Research, Applied Mathematics Program

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

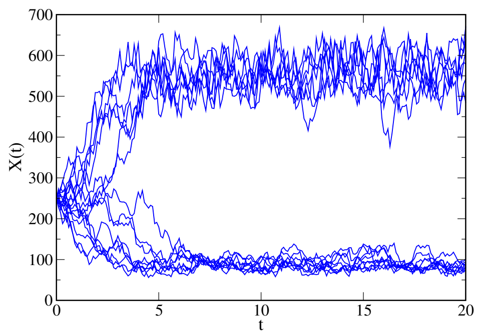
Sandia National Laboratories

Outline

- Motivation: Stochastic Reaction Networks
- Predictability: Parametric Uncertainty Propagation using Polynomial Chaos Expansion
  - Polynomial Chaos
  - Bayesian Inference
  - Markov Chain Monte Carlo
  - Adaptive Domain Decomposition
- Dynamical Analysis: Reduced Order Modeling via Karhunen-Loève Decomposition
  - Karhunen-Loève Decomposition
  - Rosenblatt Transformation
  - Quadrature Integration
  - Adaptive Data Clustering
- Application: Schlögl Model (a benchmark bistable process)

Motivation: Stochastic Reaction Networks (SRNs)

- Reaction networks involving small number of molecules necessitate the use of *stochastic* modeling instead of the *deterministic* one. E.g.
  - Immune system signaling reactions
  - Microbial reactions
  - Surface catalytic reactions



- SRNs are modeled as Jump Markov Processes
  - Governed by Chemical Master Equation $\dot{P}(X(t)=n)=\sum_m A_{nm}P(X(t)=n)$
  - Reduces to deterministic Rate Equations in the large volume limit
  - Numerically, Gillespie's Stochastic Simulation Algorithm (SSA, Gillespie, 1977)

Problem Definition and Methods

- Develop tools for *predictability*($\boldsymbol{\lambda}$) and *dynamical analysis*(t) of SRNs accounting for
  - Inherent stochasticity ($\boldsymbol{\theta}$)
  - Model/parameter variability ($\boldsymbol{\lambda}$)
  - Limited data

$$\mathcal{D} = \{X_i\}_{i=1}^{N}$$

- Techniques employed:
  - Polynomial chaos expansion; Bayesian inference; Domain decomposition
  - Karhunen-Loève decomposition; Rosenblatt transformation; Data clusttering

$$X(t, \boldsymbol{\theta}, \boldsymbol{\lambda})$$

Polynomial Chaos Expansion (PCE) - Intro

- A second order random variable $X(\boldsymbol{\theta})$ can be described by a PCE in terms of standard orthogonal polynomials Ψ_k, of associated standard random variables $\{\eta_i\}_{i=1}^{\infty}$, and spectral mode strengths c_k.

  (Wiener, 1938)(Cameron & Martin, 1947)(Ghanem & Spanos, 1991)

- Truncated PCE: finite dimension n and order p

$$X(\boldsymbol{\theta}) \simeq \sum_{k=0}^{P} c_k \Psi_k(\eta_1, \cdots, \eta_n)$$

  with the number of terms $P+1 = \frac{(n+p)!}{n!p!}$.

- Most common standard Polynomial-Variable pairs: (continuous) Gauss-Hermite, Legendre-Uniform, (discrete) Poisson-Charlier. (Askey Scheme: Xiu & Karniadakis, 2002)

Galerkin Projection is challenged by the intrinsic noise

$$X(\boldsymbol{\theta}) \simeq \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\eta}) = g_{\mathcal{D}}(\boldsymbol{\eta})$$

$$c_k = \frac{\langle X(\boldsymbol{\theta}) \Psi_k(\boldsymbol{\eta}) \rangle}{\langle \Psi_k^2(\boldsymbol{\eta}) \rangle}$$

- Intrusive Spectral Projection (ISP)
  - ⋆ Direct projection of governing equations
  - ⋆ Leads to deterministic equations for PC coefficients
  - ∗ No explicit governing equation for SRNs
- Non-intrusive Spectral Projection (NISP)
  - ⋆ Sampling based
  - ⋆ No explicit evolution equation for X needed
  - ∗ Galerkin projection not well-defined for SRNs

Bayesian inference handles the intrinsic stochasticity well

$$\overbrace{P(\boldsymbol{c}|\mathcal{D})}^{\text{Posterior}} \propto \overbrace{P(\mathcal{D}|\boldsymbol{c})}^{\text{Likelihood}}\ \overbrace{P(\boldsymbol{c})}^{\text{Prior}}$$

$$L(\boldsymbol{c}) = P(\mathcal{D}|\boldsymbol{c}) = \prod_{i=1}^{N} \mathsf{pdf}_g(X_i)$$

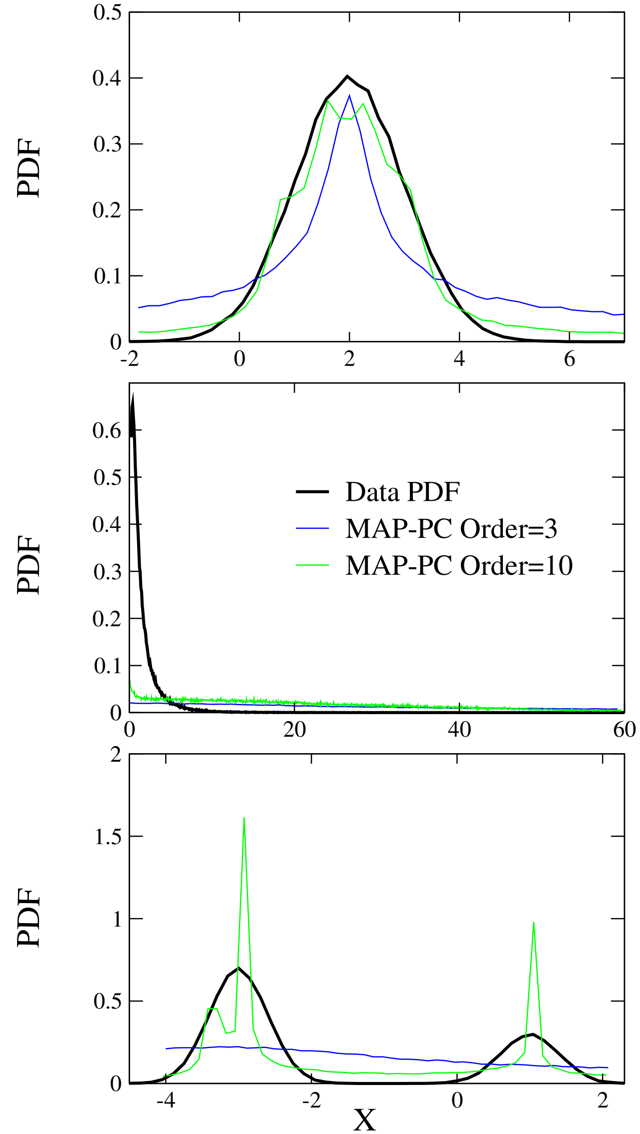
$$X \simeq \sum_{k=0}^{P} c_k \Psi_k(\eta) = g_{\mathcal{D}}(\eta)$$

- Noise model is inherent in SSA data $\mathcal{D} = \{X_i\}_{i=1}^{N}$
- Uniformly distributed priors
- Posterior exploration using Markov Chain Monte Carlo (MCMC)
- The whole posterior distribution is accessible
- Maximum a posteriori (MAP) estimate: $\boldsymbol{c}^{MAP} = \mathrm{argmax}_{\boldsymbol{c}} P(\boldsymbol{c}|\mathcal{D})$

Global PCE can fail for strongly non-linear or bimodal variables

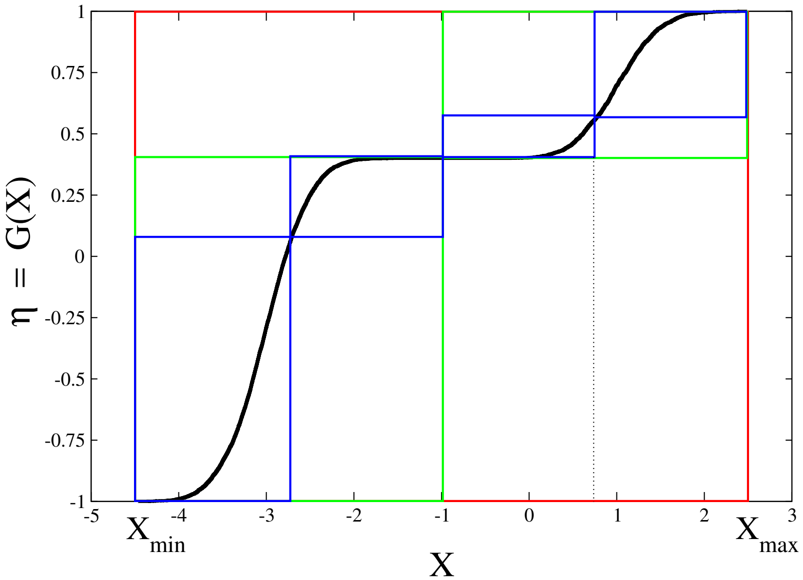
Legendre-Uniform PC

- Normal
- Lognormal
- Bimodal

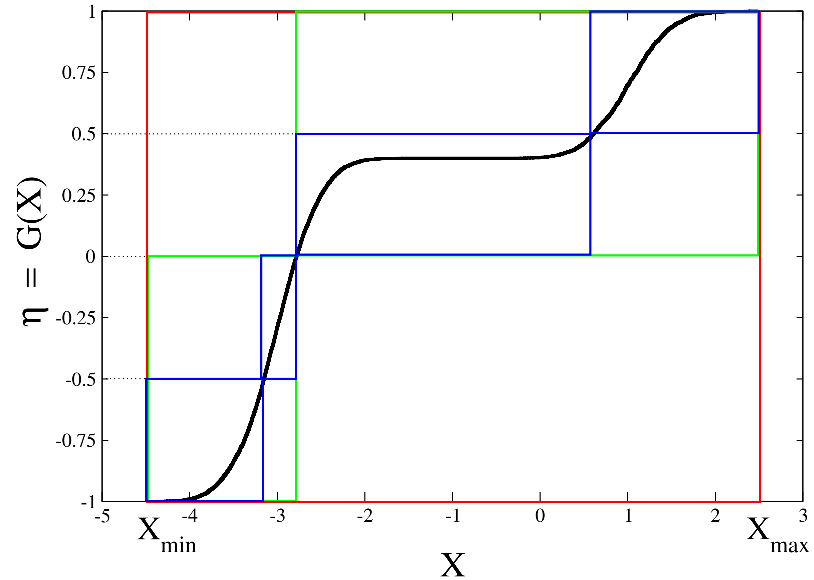


Two domain decomposition regimes available

Cumulative Distribution Function (CDF): $F(x) = P(X < x)$.
Rescaled CDF: $\eta = G(X) \equiv 2F(X) - 1$ is Uniform[-1,1].



X-domain decomposition



η-domain decomposition

Le Maître et al., 2004 - Adaptive multi-wavelet basis.

Wan & Karniadakis, 2005 - Adaptive domain decomposition.

Adaptive criterion is consistent with Kullback-Leibler 'distance'

Data: $\mathcal{D} = \{X_i\}_{i=1}^N$

Model: $X \simeq \sum_{k=0}^P c_k \Psi_k(\eta) = g_{\mathcal{D}}(\eta)$

MAP-PC samples: $\{Y_i\}_{i=1}^N$, where $Y_i = g_{\mathcal{D}}(\eta_i)$

- Log-likelihood:

$$\log L = \log P(\text{Data}|\text{Model}) = \sum_{i=1}^N \log P_Y(X_i)$$

- Target log-likelihood (the *perfect match* log-likelihood, i.e. for $\{Y_i\}_{i=1}^N = \mathcal{D}$):

$$\log L_T = \sum_{i=1}^N \log P_X(X_i)$$

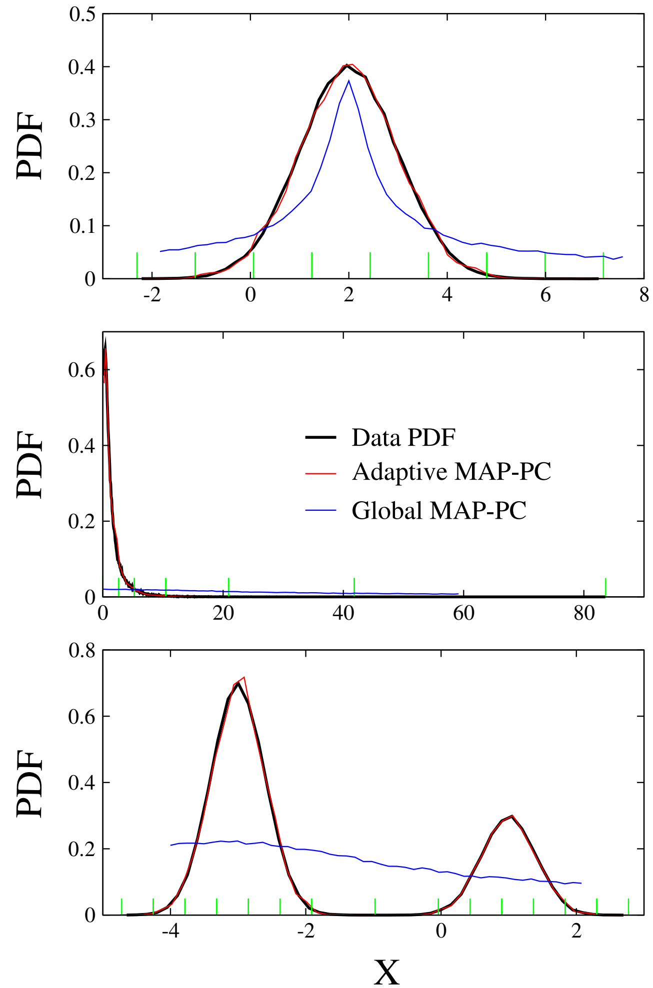
- Kullback-Leibler divergence:

$$\rho(P_X, P_Y) = \int P(z) \log \frac{P_X(z)}{P_Y(z)} dz \simeq \frac{1}{N} \sum_{i=1}^N \log \frac{P_X(X_i)}{P_Y(X_i)}$$

Multi-domain PCE captures non-linearities and bimodalities well

Legendre-Uniform multi-domain PC

- Normal
- Lognormal
- Bimodal

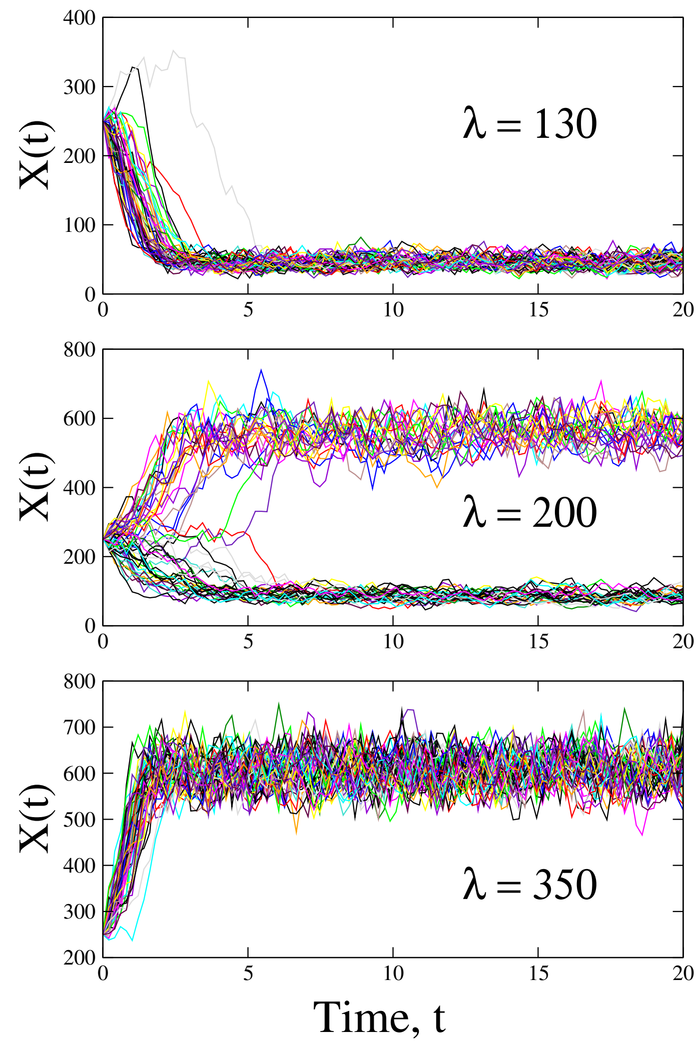


Schlögl model is a benchmark bistable process

- Reactions

$$A + 2X \underset{a_2}{\overset{a_1}{\rightleftarrows}} 3X$$

$$B \underset{a_4}{\overset{a_3}{\rightleftarrows}} X$$



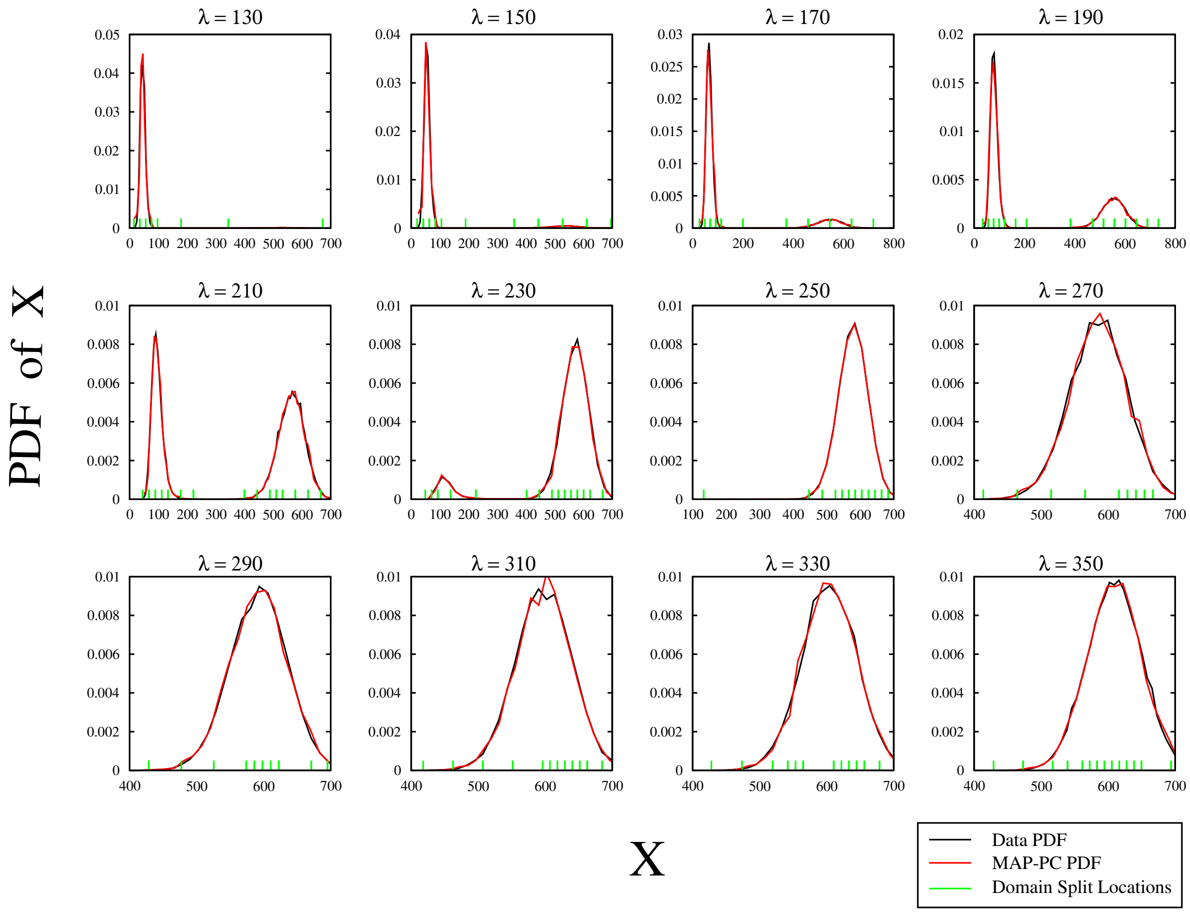
- Propensities

$$a_1 = k_1 A X(X-1)/2,$$
$$a_2 = k_2 X(X-1)(X-2)/6,$$
$$a_3 = k_3 B,$$
$$a_4 = k_4 X.$$

- Nominal parameters

$k_1 A$	0.03
k_2	0.0001
$k_3 B = \lambda$	200
k_4	3.5
A	10^5
B	$2 \cdot 10^5$
$X(0)$	250

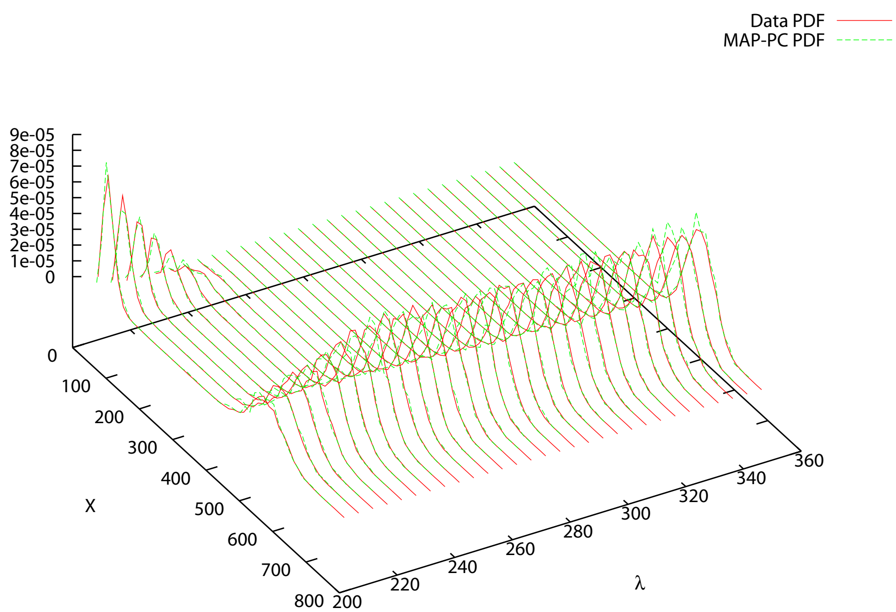
PC Inference for fixed parameter values

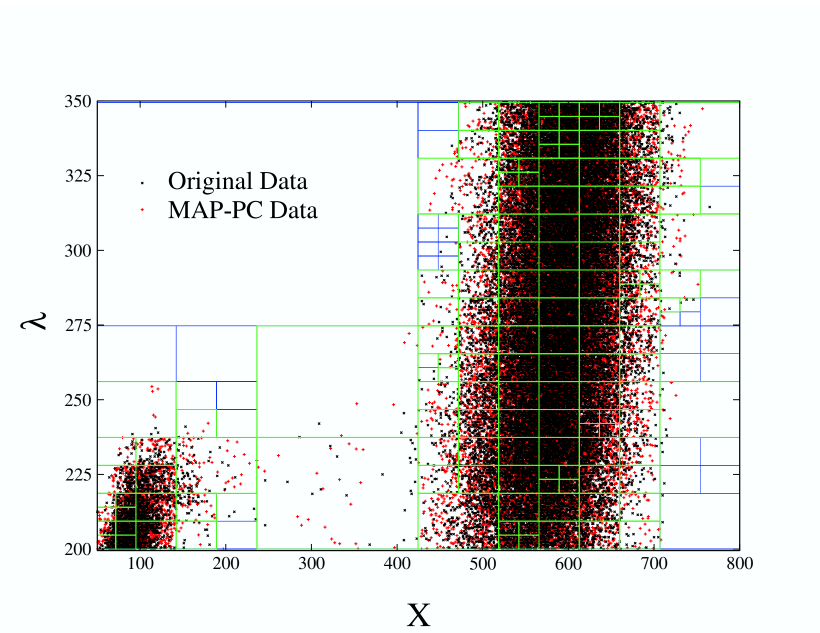


Qualitatively different behaviors across a range of λ values necessitates the parametric uncertainty introduction.

Predictability: Parametric uncertainty propagation through PCE

- Postulate parametric uncertainty $\qquad \lambda = \lambda_0 + \Delta\lambda\eta_1$
- Gather two-dimensional data $\mathcal{D} = \{(X_i, \lambda_i)\}_{i=1}^N$
- Infer the model parameters c_k's, where $X = \sum_{k=0}^P c_k \Psi_k(\eta_1, \eta_2)$
- If the representation is not satisfactory (see the criterion), split the data domain and proceed recursively





Dynamical Analysis: Big Picture

Fix the parameter $X(t,\boldsymbol{\theta},\boldsymbol{\Lambda}) \equiv X(t,\boldsymbol{\theta})$

SSA $\longleftrightarrow$ KL $\longleftrightarrow$ PCE

$$X(t,\boldsymbol{\theta}) \longleftrightarrow \xi_i(\boldsymbol{\theta})(i=\overline{1,L}) \longleftrightarrow c_{ik}(i=\overline{1,L}, k=\overline{0,P})$$

$$X(t,\boldsymbol{\theta})-\bar{X}(t) \simeq \sum_{i=1}^{L} \xi_i(\boldsymbol{\theta})\sqrt{\lambda_i} f_i(t) \simeq \sum_{i=1}^{L}\left(\sum_{k=0}^{P} c_{ik}\Psi_k(\boldsymbol{\eta})\right)\sqrt{\lambda_i} f_i(t)$$

SSA $\longleftrightarrow$ KL : Karhunen-Loève (KL) decomposition of the stochastic process

KL $\longleftrightarrow$ PCE: Polynomial Chaos expansion of each KL random variable

Karhunen-Loève Decomposition: Intro

- Separate the average:

$$X_0(t,\theta) = X(t,\theta) - \bar{X}(t)$$

- The covariance function is symmetric, bounded and positive definite. Hence, it can be expanded as a sum

$$C(t_1,t_2) = \langle X_0(t_1,\theta) X_0(t_2,\theta) \rangle = \sum_{n=1}^{\infty} \lambda_n f_n(t_1) f_n(t_2)$$

- Positive eigenvalues:

$$\int_0^T C(t_1,t_2) f_n(t_1) dt_1 = \lambda_n f_n(t_2).$$

- KL decomposition:

$$X(t,\theta) = \bar{X}(t) + \sum_{n=1}^{\infty} \xi_n(\theta) \sqrt{\lambda_n} f_n(t)$$

Karhunen-Loève decomposition leads to reduced order modeling

- KL decomposition:

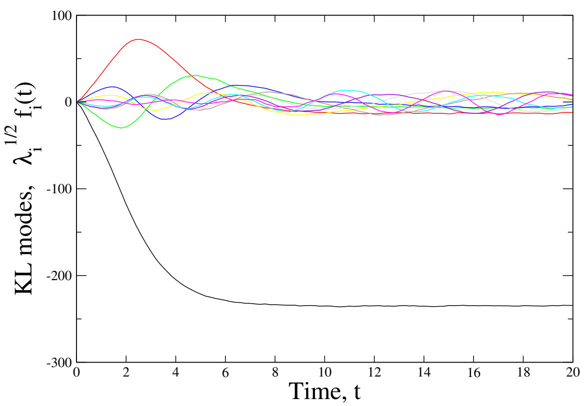
$$X(t,\theta) = \bar{X}(t) + \sum_{n=1}^{\infty} \xi_n(\theta)\sqrt{\lambda_n} f_n(t)$$

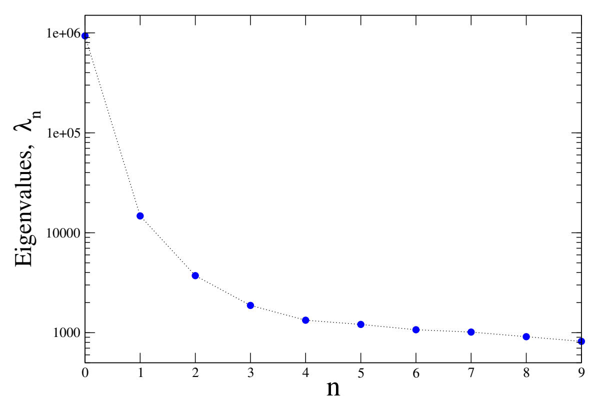
- Uncorrelated, zero-mean KL variables:

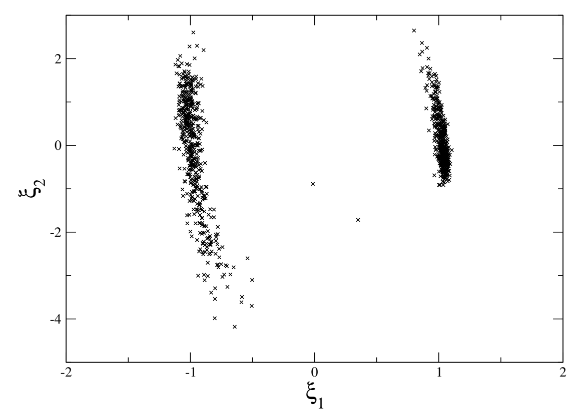
$$\langle \xi_n \rangle = 0, \qquad \langle \xi_n \xi_m \rangle = \delta_{nm}$$

- SSA(continuum) $\longleftrightarrow$ KL(discrete)

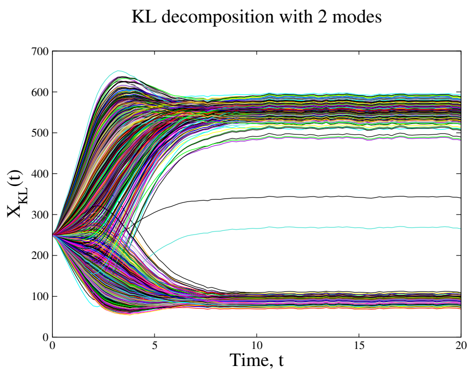
$$X(t) \longleftrightarrow \boldsymbol{\xi} = (\xi_1, \xi_2, \dots)$$

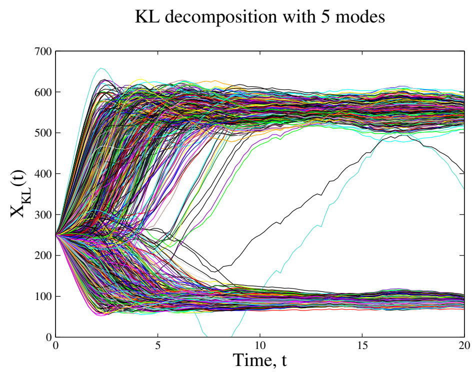


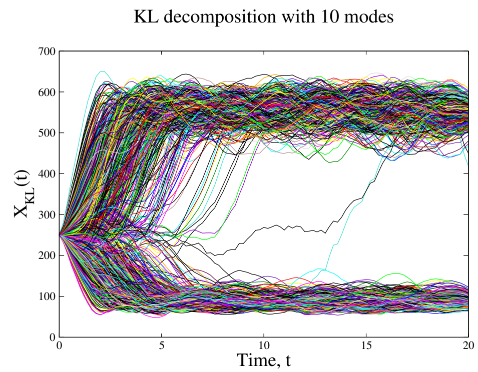


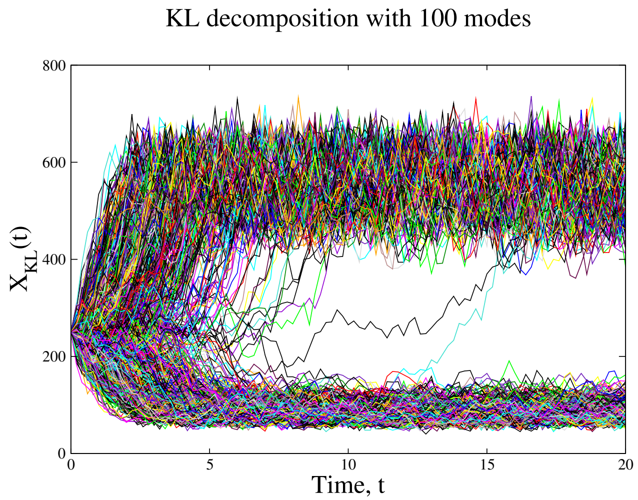


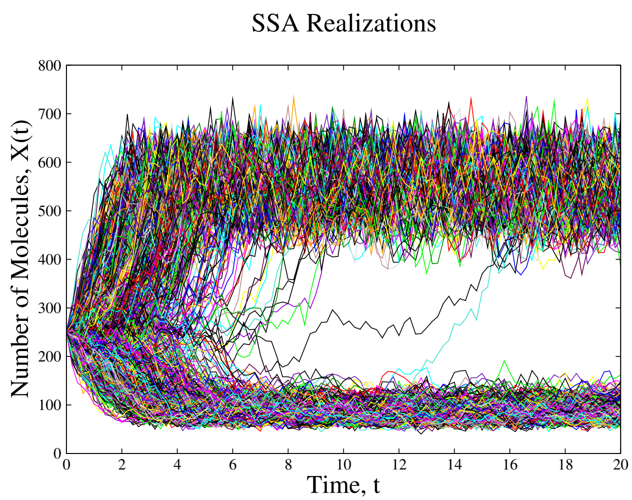
Karhunen-Loève decomposition captures each realization



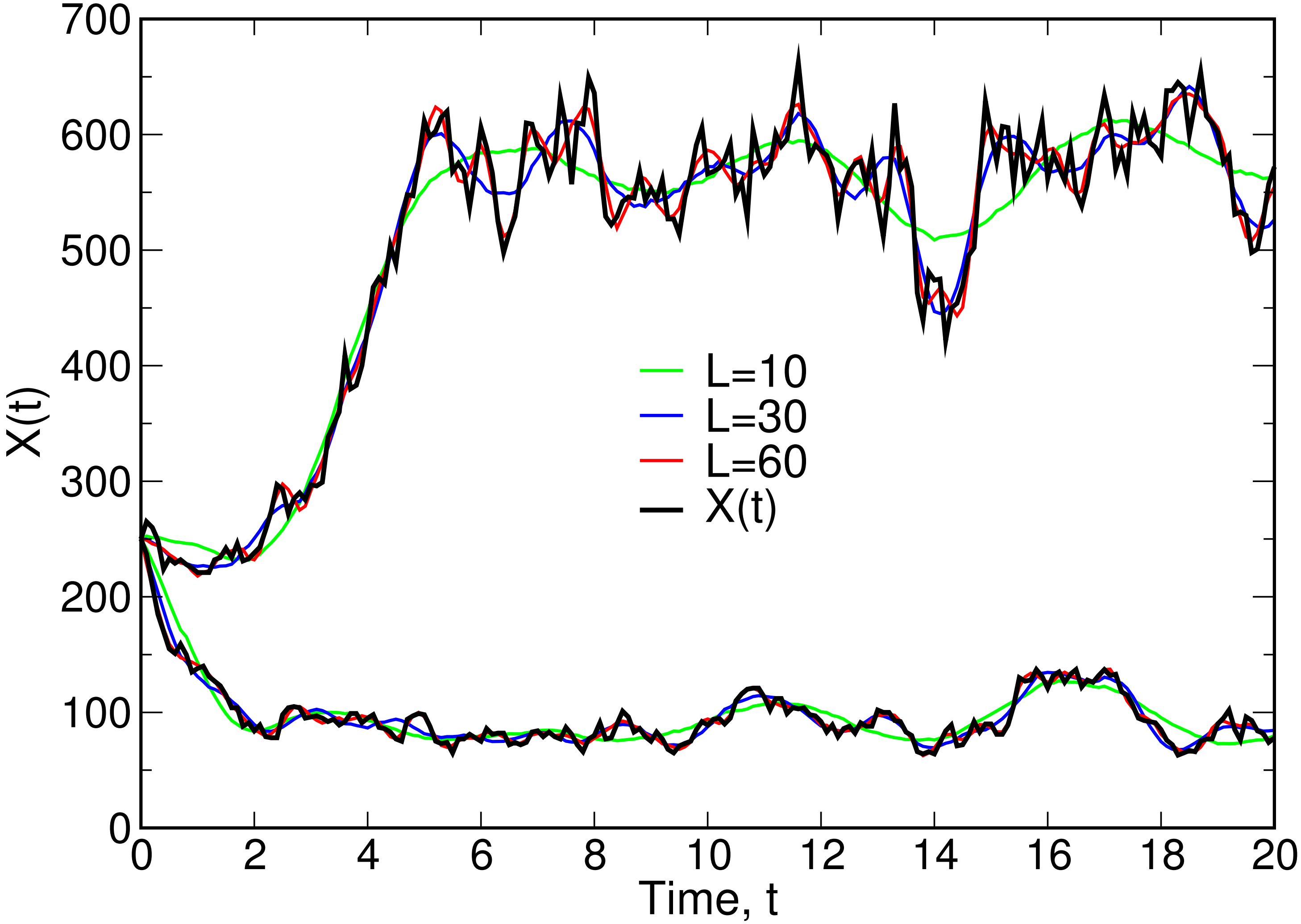








Karhunen-Loève decomposition captures each realization



Sandia National Laboratories

Challenges in PC expansions of KL random variables

Need to PC-expand each of the KL random variables

$$\xi_i = \sum_{k=0}^{P} c_{ik} \Psi_k(\boldsymbol{\eta}), \text{ for } i = 1, \ldots, L$$

- Quadrature-based non-intrusive spectral projection is not well-defined

$$c_{ik} = \frac{\langle \xi_i \Psi_k(\boldsymbol{\eta}) \rangle}{\langle \Psi_k^2(\boldsymbol{\eta}) \rangle}$$

  - Employ (inverse) Rosenblatt transformation
- Multimodal variables not captured well
  - Use data clustering

Rosenblatt Transformation

- Rosenblatt transformation maps any (not necessarily independent) set of random variables $(\xi_1, \ldots, \xi_n)$ to uniform i.i.d.'s $\{\eta_i\}_{i=1}^n$ (Rosenblatt, 1952).

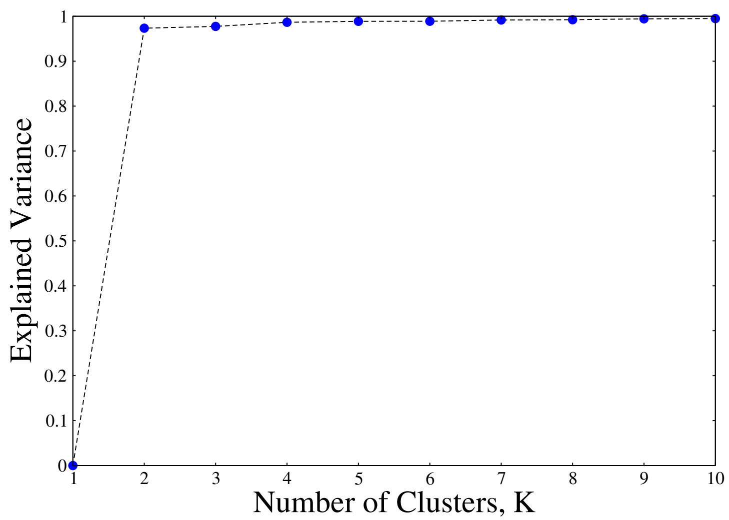
$$\begin{aligned} \eta_1 &= F_1(\xi_1) \\ \eta_2 &= F_{2|1}(\xi_2|\xi_1) \\ \eta_3 &= F_{3|2,1}(\xi_3|\xi_2, \xi_1) \\ &\vdots \\ \eta_n &= F_{n|n-1,\ldots,1}(\xi_n|\xi_{n-1}, \ldots, \xi_1) \end{aligned}$$

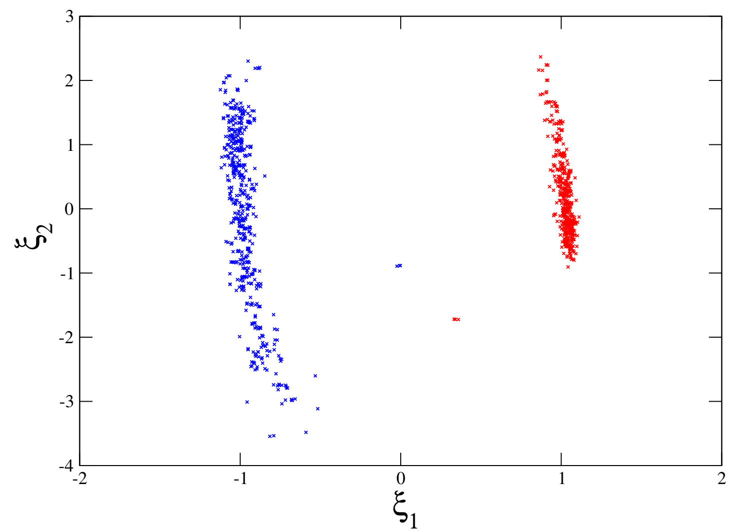
- Inverse Rosenblatt transformation $\boldsymbol{\xi} = R^{-1}(\boldsymbol{\eta})$ ensures a well-defined integration

$$\langle \xi_i \Psi_k(\boldsymbol{\eta}) \rangle = \int (R^{-1}(\boldsymbol{\eta}))_i \Psi_k(\boldsymbol{\eta}) d\boldsymbol{\eta}$$

Adaptive Data Clustering

- Finite number of KL variables: $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_L)$
- Multidimensional data: $\{\boldsymbol{\xi}^{(i)}\}_{i=1}^{N}$
- K-Center clustering
- 'Elbow' criterion to pick initial number of clusters
- Adaptive clustering: split clusters if output is not good enough (check the Kullback-Leibler distance)





PC Mixture Model

- Divide data into K clusters with fractions p_j:

$$p_1 + p_2 + \cdots + p_K = 1$$

- Find PC expansion for ξ in each cluster:

$$\xi_{PC}^{(j)} = \sum_{k=0}^{P} \xi_k^{(j)} \Psi_k(\eta^{(j)})$$

- Superpose the results to obtain PC mixture model:

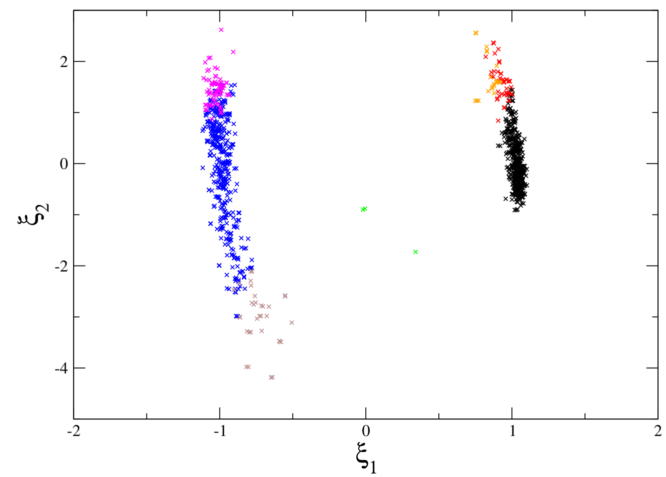
$$\xi = \xi_{PC}^{(j)} \text{ w. prob. } p_j$$

- Probability distribution function is a mixture of PC PDFs:

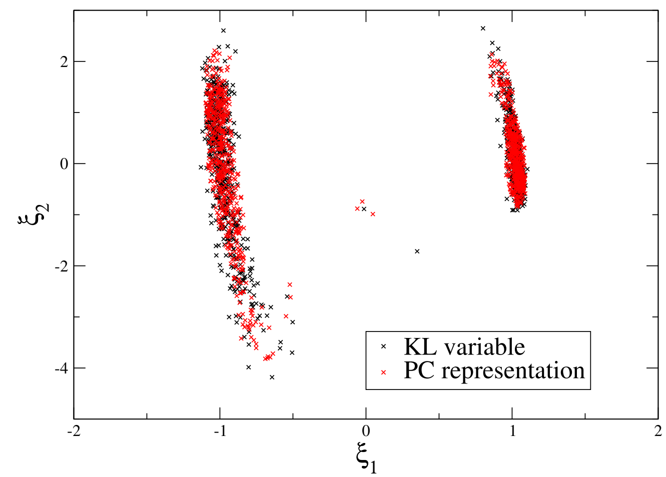
$$\mathsf{Pdf}_{\xi}(x) = p_1 \mathsf{Pdf}_{\xi_{PC}^{(1)}}(x) + \cdots + p_K \mathsf{Pdf}_{\xi_{PC}^{(K)}}(x)$$

Results for the Schlögl Model

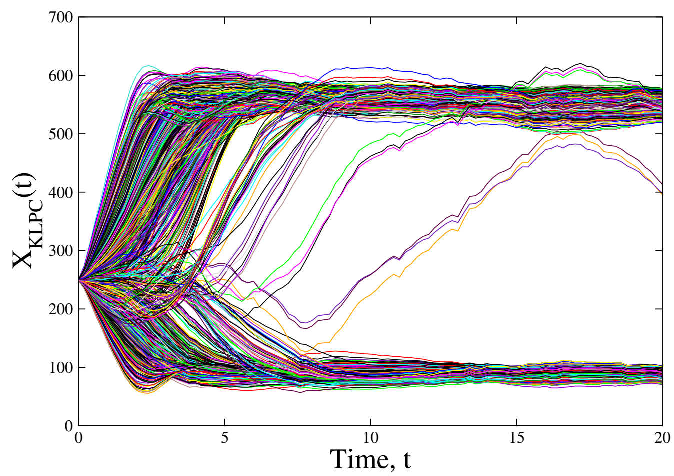
Adaptive Clustering



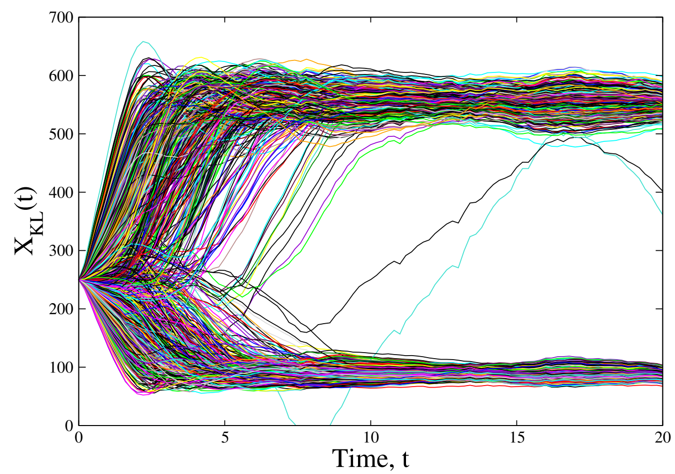
PC representation of first two KL variables



KL-PC representation, 5 KL modes, 3rd PC order



KL decomposition with 5 modes



Conclusions and Ongoing Work

- Adaptive domain decomposition with Bayesian inference allows representation of the state $X(T, \boldsymbol{\theta}, \boldsymbol{\lambda})$ across large range of parameter values
- PC mixture model with Karhunen-Loève decomposition represents the dynamics of the state $X(t, \boldsymbol{\theta}, \boldsymbol{\Lambda})$

- Dimensionality (complexity increase) studies
- Adaptive PC order
- Sparce quadrature integration or Latin Hypercube Sampling
- Combination of parameter uncertainties and time evolution

Literature

- N. Wiener : 1938 : *The homogeneous chaos* – Hermite polynomials of a Gaussian process
- R. Cameron & W. Martin : 1947 : *The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals* – L^2 Convergence for any L^2 stochastic process
- M. Rosenblatt : 1952 : *Remarks on a multivariate transformation* – Rosenblatt transformation
- D. Gillespie : 1977 : *Exact stochastic simulation of coupled chem. reactions* – SSA algorithm
- R. Ghanem & P. Spanos : 1991 : *Stochastic Finite Elements* – PCE and KL in the context of stochastic finite element method
- D. Xiu & G. Karniadakis : 2002 : *The Wiener-Askey polynomial chaos for stochastic differential equations* – The Askey scheme
- O. Le Maître, R. Ghanem, O. Knio & H. Najm : 2004 : *Uncertainty propagation using Wiener-Haar expansions* and *Multi-resolution analysis of Wiener-type uncertainty propagation schemes* – Adaptive multi-resolution analysis
- X. Wan & G. Karniadakis : 2005 : *An adaptive multi-element generalized polynomial chaos method for stochastic differential equations* – Adaptive domain decomposition
- Y. Marzouk, H. Najm & L. Rahn : 2007 : *Stochastic spectral methods for efficient Bayesian solution of inverse problems* – Inference+PCE
- D. Kim, B. Debusschere & H. Najm : 2007 : *Spectral methods for parametric sensitivity in stochastic dynamical systems* – Parametric sensitivity analysis
- K. Sargsyan, B. Debusschere, H. Najm & Y. Marzouk : 2008 : *Predictability assessment in stochastic reaction networks* – Submitted to JCTN UQ special issue

Multi-domain PC expansion

- Partition: $-1 = a_1 < b_1 = a_2 < b_2 = \cdots = a_n < b_n = 1$

$$\mathcal{P} = \{[a_1, b_1), [a_2, b_2), \ldots, [a_n, b_n]\}$$

- Linear map: $f^I : I \equiv [a, b] \mapsto [-1, 1]$ from an interval $[a, b]$ (subscripts dropped for simplicity) to $[-1, 1]$, by

$$f^I(\xi) = \tilde{\xi} = \frac{2}{b-a}\left(\xi - \frac{a+b}{2}\right)$$

- Multi-domain PC expansion

$$X \simeq g(\xi) = \sum_{I \in \mathcal{P}} \sum_{k=0}^{P} c_k^I \Psi_k^I(\xi),$$

where

$$\Psi_k^I(\xi) \equiv 0, \text{ if } \xi \notin I$$
$$\Psi_k^I(\xi) = \Psi_k(f^I(\xi)), \text{ if } \xi \in I$$