Reduced Order Modeling and Dynamical Analysis in Stochastic Reaction Networks

Khachik Sargsyan, Bert Debusschere, Habib Najm

Transportation Energy 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.

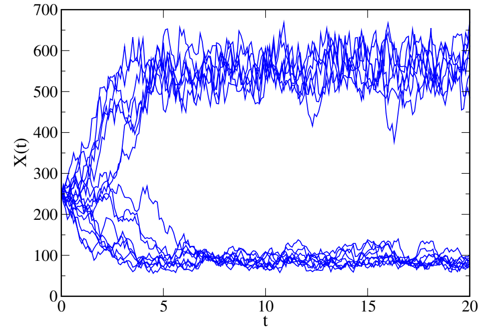
Sandia National Laboratories

Outline

- Motivation: stochastic reaction networks
- Application: Schlögl Model (a benchmark bistable process)
- Reduced order modeling (Karhunen-Loève decomposition)
- Spectral representation (Polynomial Chaos expansion)
  - Non-intrusive orthogonal projection
  - Rosenblatt transformation
- Adaptive data partitioning
  - K-center clustering
  - Data range bisection
- Mixture PC model

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

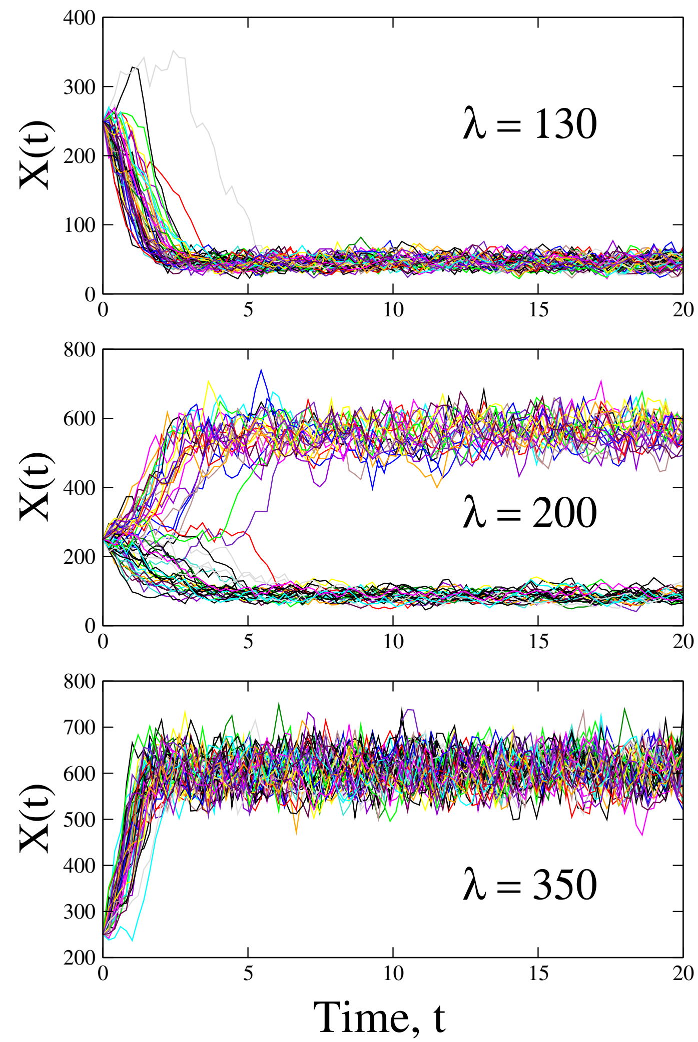
- SRNs are modeled as Jump Markov Processes
  - Governed by Chemical Master Equation $\dot{P}(X(t)=n)=\sum_m A_{nm}P(X(t)=n)$
  - Reduces to deterministic Rate Equations in the large volume limit
  - Trajectories simulated by Gillespie's Stochastic Simulation Algorithm (SSA, Gillespie, 1977)

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

Sandia National Laboratories

Problem Definition and Methods

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

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

- Techniques employed:
  - Karhunen-Loève decomposition
  - Polynomial chaos expansion
  - Rosenblatt transformation
  - Data partitioning/clustering

- Obtain a surrogate model for the dynamics of $X(t, \theta, \boldsymbol{\lambda} = \boldsymbol{\Lambda})$

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:

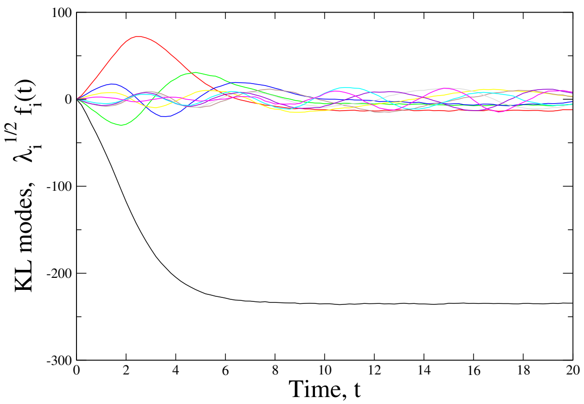
$$X(t,\theta) = \bar{X}(t) + \sum_{n=1}^{\infty} \xi_n(\theta)\sqrt{\lambda_n} f_n(t)$$

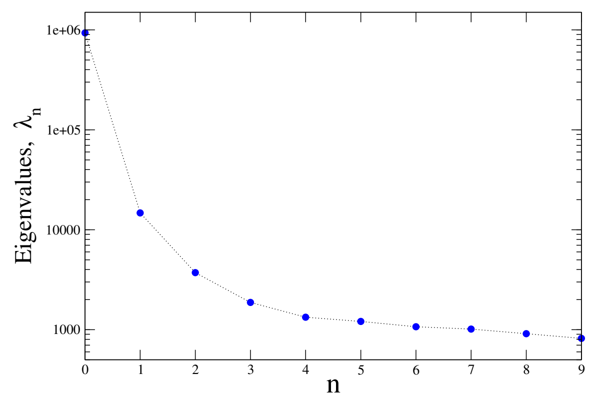
- Uncorrelated, zero-mean KL variables:

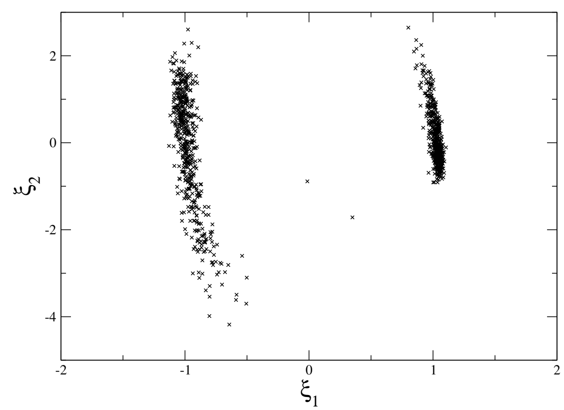
$$\langle \xi_n \rangle = 0, \qquad \langle \xi_n \xi_m \rangle = \delta_{nm}$$

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

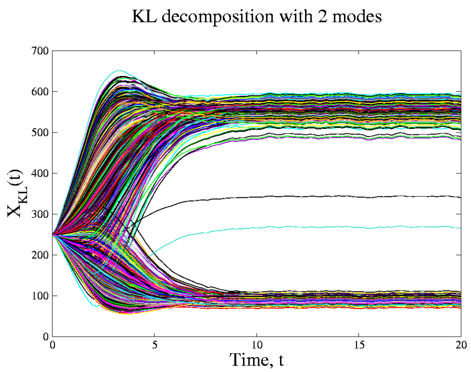
$$X(t) \longleftrightarrow \boldsymbol{\xi} = (\xi_1, \xi_2, \dots)$$

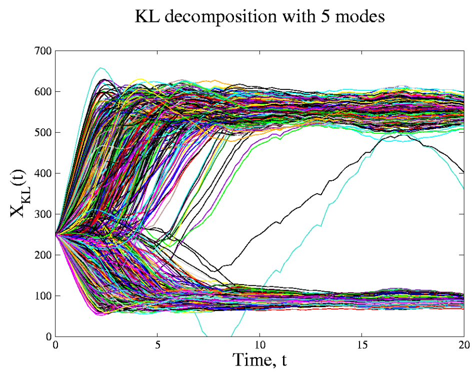


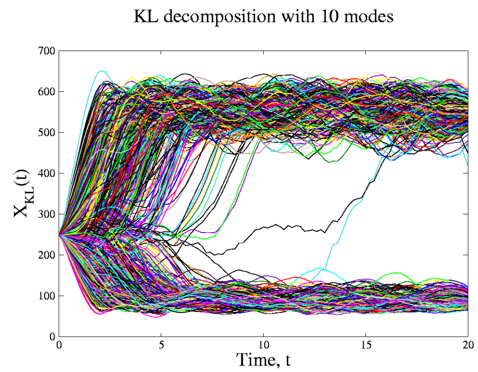


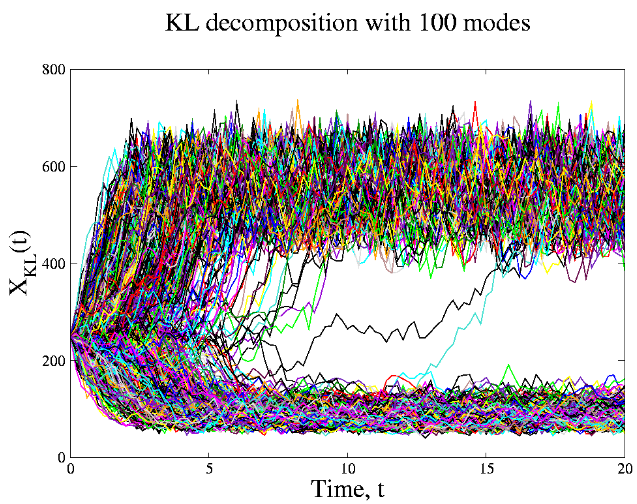


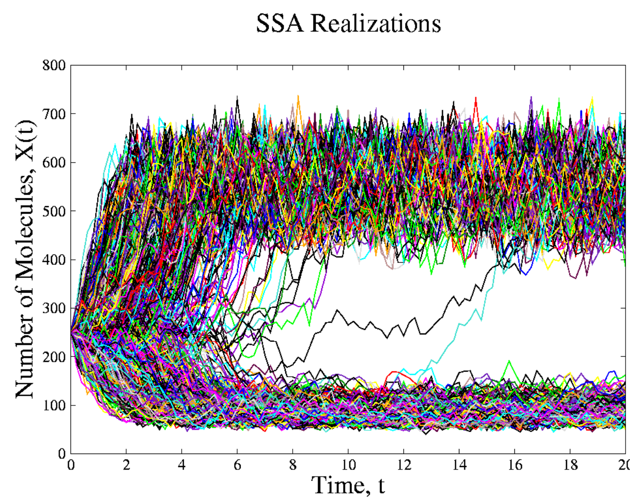
Karhunen-Loève decomposition captures each realization



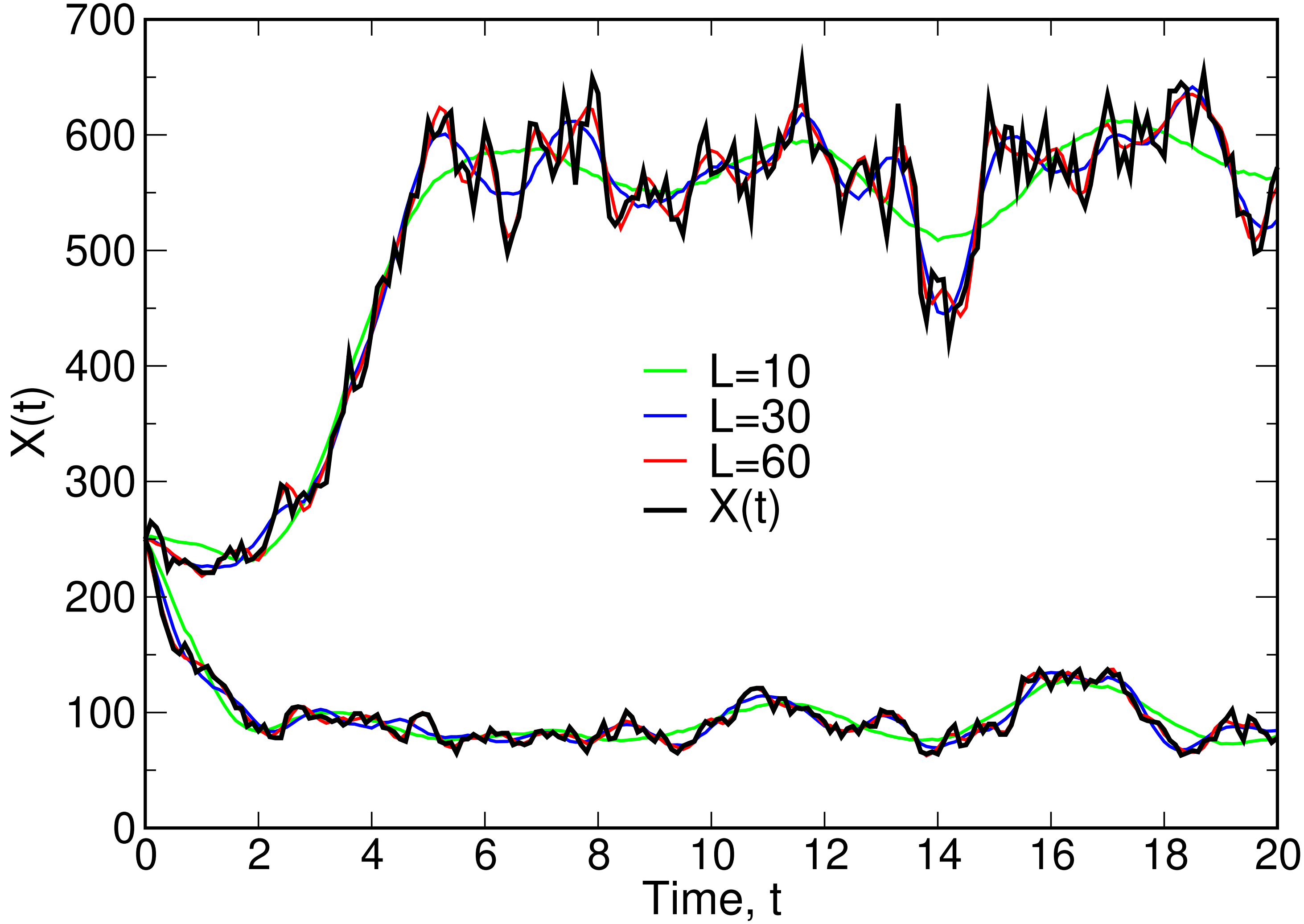








Karhunen-Loève decomposition captures each realization



Dynamical Analysis: Big Picture

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

SSA $\longrightarrow$ KL $\longrightarrow$ PCE

Random process $\longrightarrow$ L random variables $\longrightarrow$ L(P+1) deterministic variables

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

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

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

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

Polynomial Chaos Expansion (PCE) - Intro

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

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

- Truncated PCE: finite dimension n and order p

$$X(\theta) \simeq \sum_{k=0}^{P} c_k \Psi_k(\zeta_1, \cdots, \zeta_n)$$

  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)

Challenges in PC expansions of KL random variables

Need to PC-expand each of the KL random variables

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

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

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

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

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).

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

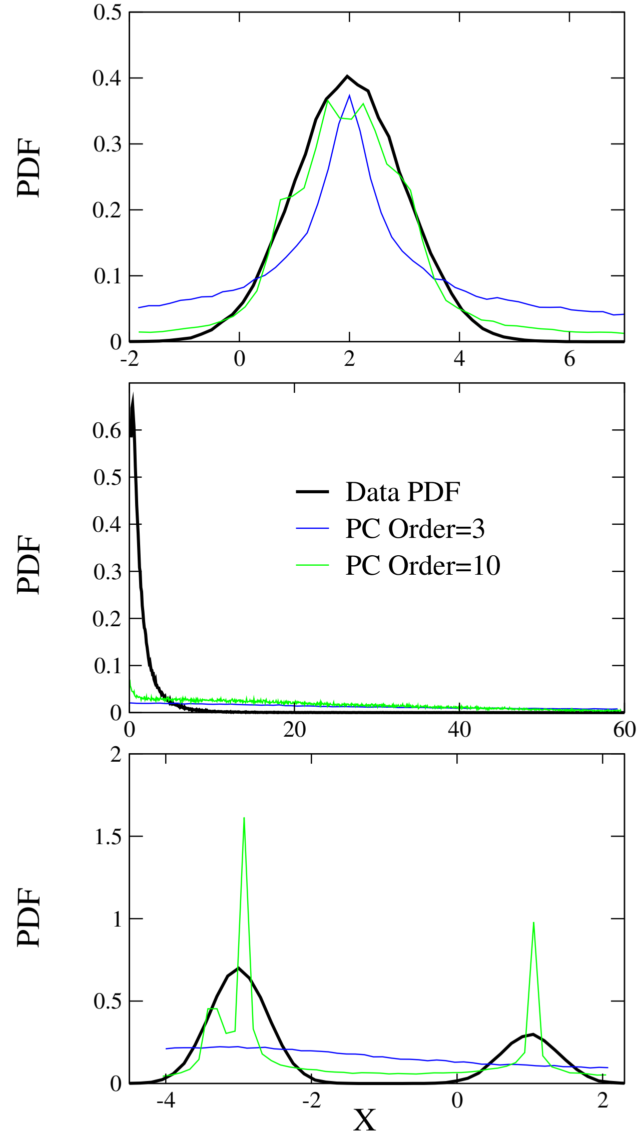
- Inverse Rosenblatt transformation $\boldsymbol{\xi} = R^{-1}(\boldsymbol{\eta})$ (with standard normal CDF $\Phi(\cdot)$) ensures a well-defined quadrature integration

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

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

Legendre-Uniform PC

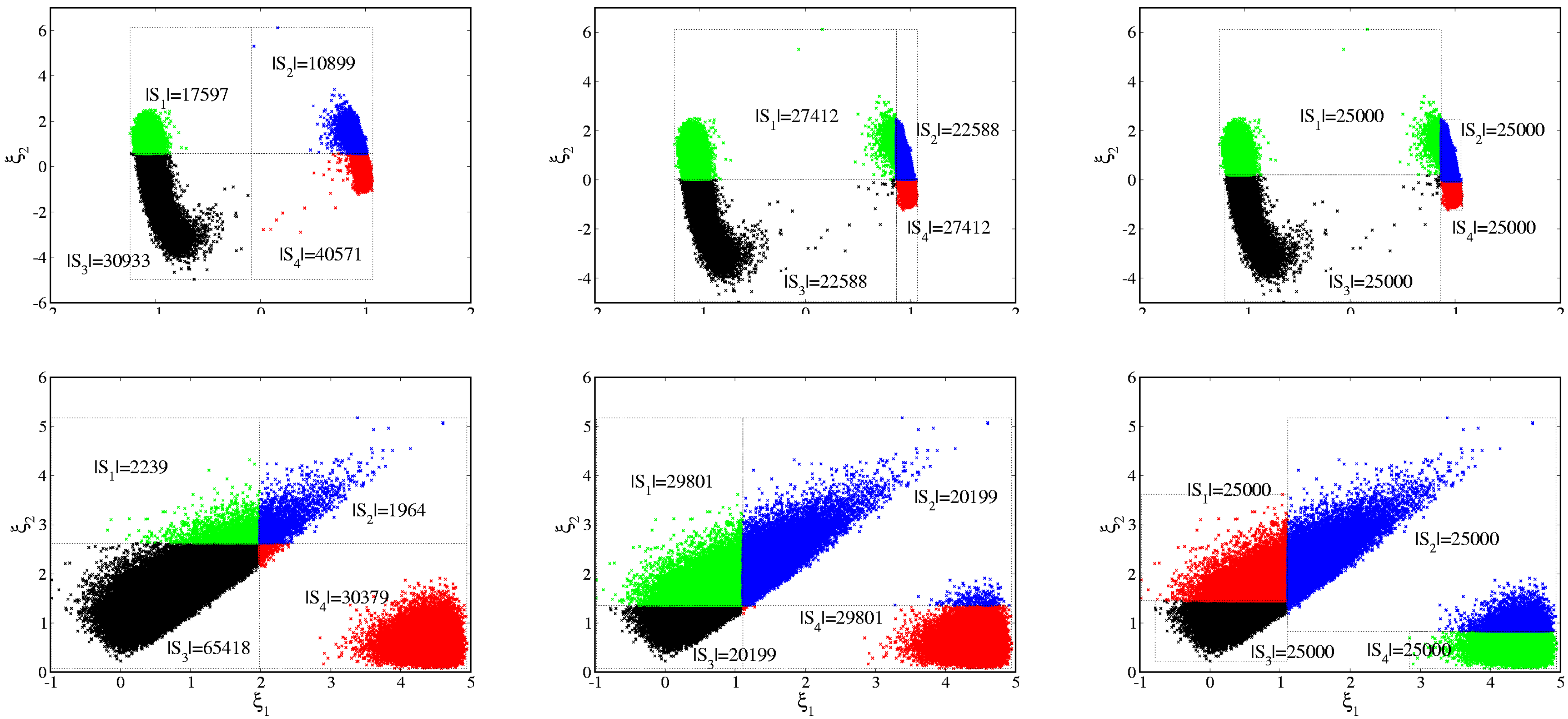
- Normal
- Lognormal
- Bimodal



Sandia National Laboratories

Domain-based data partitioning methods do not detect bimodalities

Data range bisection | Data median bisection | Data size bisection

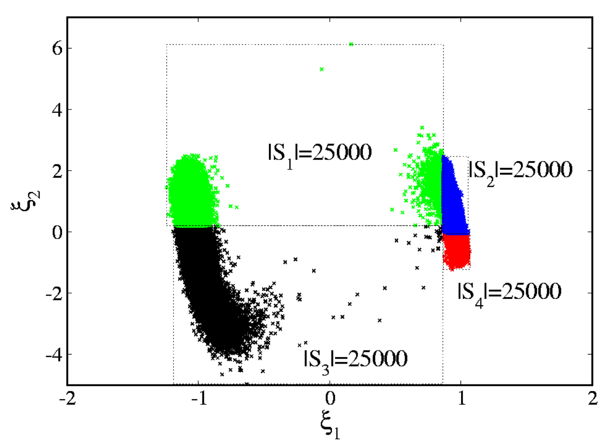


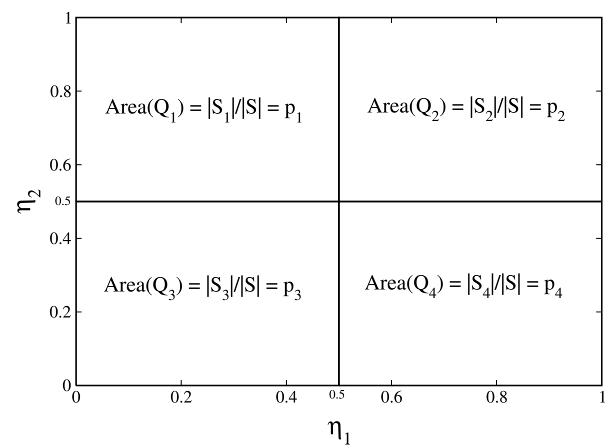
Relation to multidomain expansions

Given a stochastic space partition $\cup_{i=1}^{K_f} Q_i = [0,1]^L$

$$\xi_n = \sum_{i=1}^{K_f} \sum_{k=0}^{P} c_{nk}^{(i)} \tilde{\Psi}_k^{(i)}(\boldsymbol{\eta}), \text{ for } n = 1, \ldots, L,$$

with $\boldsymbol{\eta} \in [0,1]^L$, and where the basis functions $\tilde{\Psi}_k^{(i)}(\cdot)$ vanish outside their support Q_i.



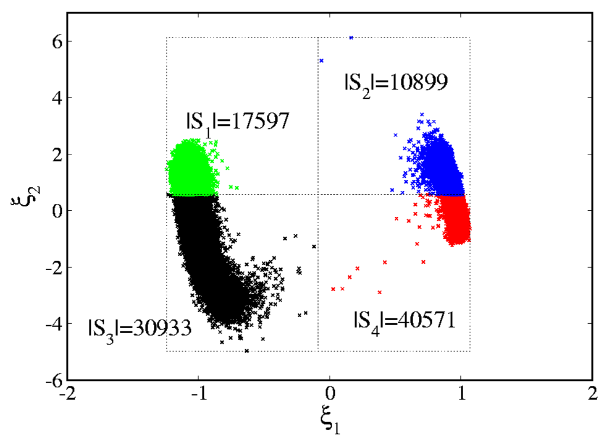


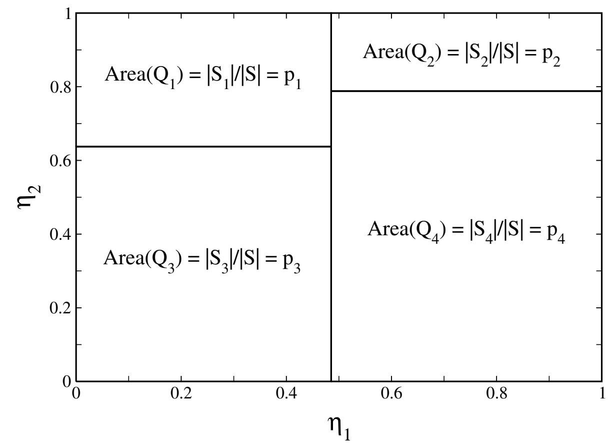
Relation to multidomain expansions

Given a stochastic space partition $\cup_{i=1}^{K_f} Q_i = [0,1]^L$

$$\xi_n = \sum_{i=1}^{K_f} \sum_{k=0}^{P} c_{nk}^{(i)} \tilde{\Psi}_k^{(i)}(\boldsymbol{\eta}), \text{ for } n = 1, \ldots, L,$$

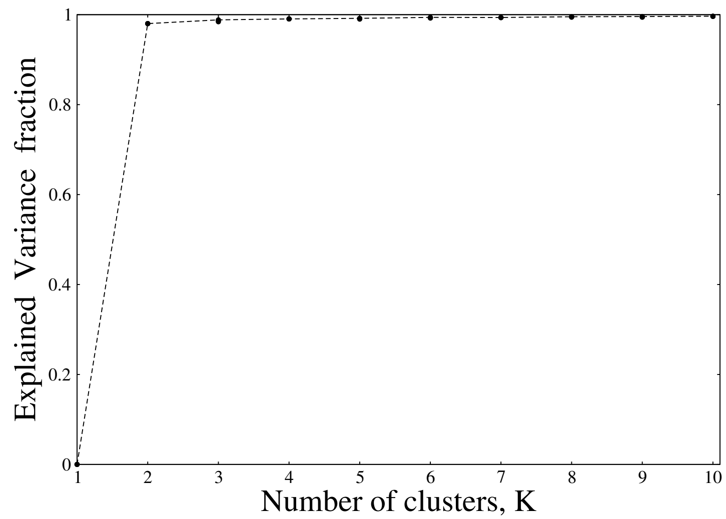
with $\boldsymbol{\eta} \in [0,1]^L$, and where the basis functions $\tilde{\Psi}_k^{(i)}(\cdot)$ vanish outside their support Q_i.

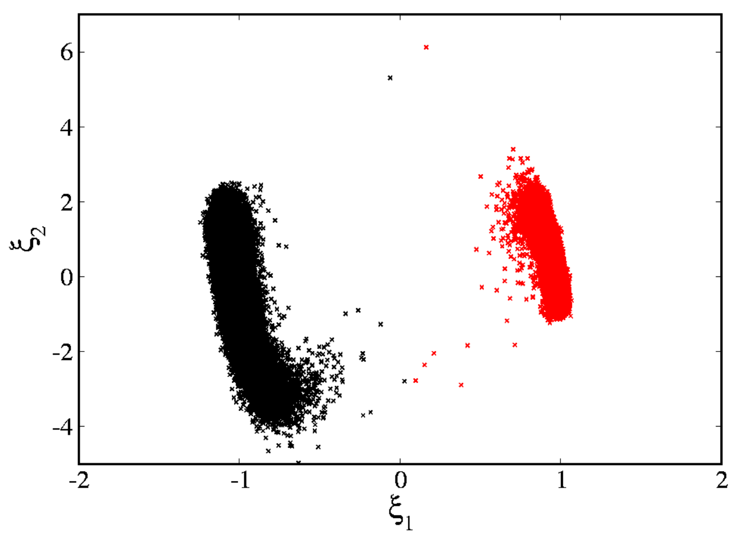




Data Clustering

- 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 (Gonzalez, 1985)
- Distance measure scaled with KL eigenvalues
- 'Elbow' criterion with Explained Variance to pick the optimal number of clusters
- E.V. = Variance of dataset with all points replaced with their corresponding cluster's center





- Let A be a set of n objects
- Partition A into K sets $C_1, C_2, \ldots, C_K$
- *Cluster size* of C_k: the least value D_k s.t. all points in C_k are
  a) within D_k of each other, or
  b) within $D_k/2$ of some point (called *cluster center*)
- Goal:

$$\min_{S} \max_{k=1,\ldots,K} D_k$$

Which distance?

- L_∞: $d(\boldsymbol{x}, \boldsymbol{y}) = \max_i |x_i - y_i|$
- L_2: $d(\boldsymbol{x}, \boldsymbol{y})^2 = \sum_i (x_i - y_i)^2$
- Rescaled L_2: $d(\boldsymbol{x}, \boldsymbol{y})^2 = \sum_i \lambda_i (x_i - y_i)^2$

Domain-based bisection VS clustering

Approximate k-center clustering	Data range bisection
Polythetic (more effective use of data structure)	Monothetic
Detects multimodalities and outliers	Blind to multimodalities and outliers
No curse of dimensionality	Number of new partitions scales exponentially with dimensions
Dimension-specific weight measure	No weight measure
Non-unique (randomized) partitioning	Unique partitioning
Not effective for unimodal data	Performs well for unimodal data
New subset sizes are of a similar order of magnitude	New subset sizes are extremely reduced

We propose an adaptive, hybrid approach:

- Start with clustering to detect the modalities
- Continue with data range bisection

- Refinement criterion: *Kullback-Leibler divergence* between data PDF and model PDF

$$\rho(P_{\mathcal{D}}, P_M) = \int P_{\mathcal{D}}(\boldsymbol{\xi}) \log \frac{P_{\mathcal{D}}(\boldsymbol{\xi})}{P_M(\boldsymbol{\xi})} d\boldsymbol{\xi} \simeq \frac{1}{N} \sum_{i=1}^{N} \log \frac{P_{\mathcal{D}}(\boldsymbol{\xi}_i)}{P_M(\boldsymbol{\xi}_i)}$$

Adaptive Hybrid Algorithm

- (**Simulation**) Obtain N SSA realizations $X(t)$.
- (**Reduced order model - KL**) Perform KL decomposition up to the eigenmode L: $X_{KL}(t) = \bar{X}(t) + \sum_{n=1}^{L} \xi_n \sqrt{\lambda_n} f_n(t)$.
  - As a result, obtain a set of N data samples of the random vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_L)$ and call it the current data set S.
  - *Only if* the explained variance criterion detects modalities, cluster the data into the optimal number of clusters and proceed considering each cluster as a new data set.
- (**Spectral expansion - PC**) Find a finite order PC representation for the current data samples: $\boldsymbol{\xi} = \sum_{k=0}^{P} \boldsymbol{c}_k \Psi_k(\boldsymbol{\zeta})$.
- (**Adaptive refinement**) *If* the number of samples $|S| > N_{\text{thr}}$ *and* the Kullback-Leibler divergence $\rho_{KL} > d_{\text{thr}}$, partition the current data set according to data range bisection.
  Else keep the current PC representation.
  - Move to the next untreated data set.

Final representation is a mixture model of PC PDFs

- Divide data into K partitions with fractions p_j:

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

- Find PC expansion for ξ in each partition:

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

- Superpose the results to obtain PC mixture model (assuming data points are of equal importance/weight):

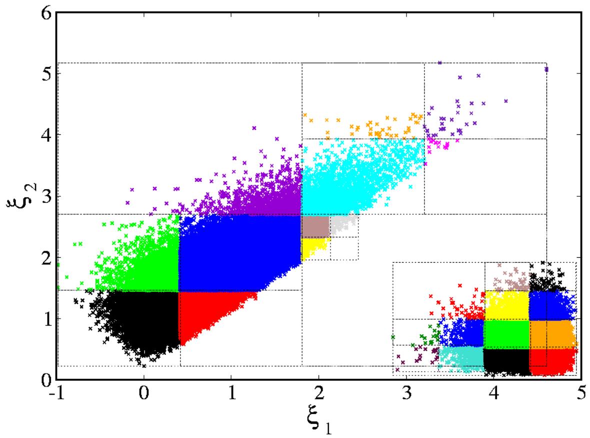
$$\xi = \xi_{PC}^{(j)} \text{ w. prob. } p_j$$

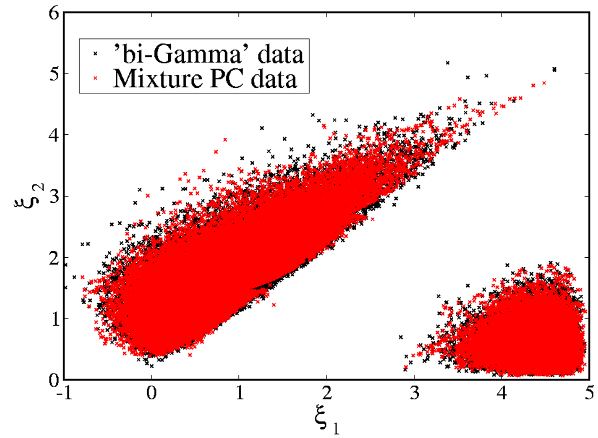
- Probability distribution function is a mixture of PC PDFs:

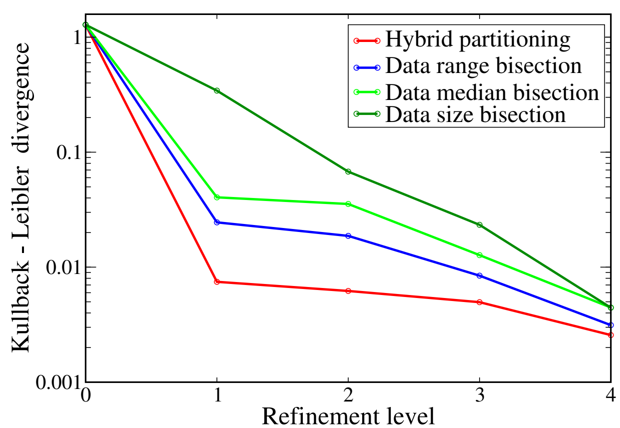
$$\mathsf{Pdf}_{\xi}(x) = p_1 \mathsf{Pdf}_{\xi_{PC}^{(1)}}(x) + \cdots + p_K \mathsf{Pdf}_{\xi_{PC}^{(K)}}(x)$$

Convergence results confirm the hybrid approach superiority

Bi-Gamma test distribution

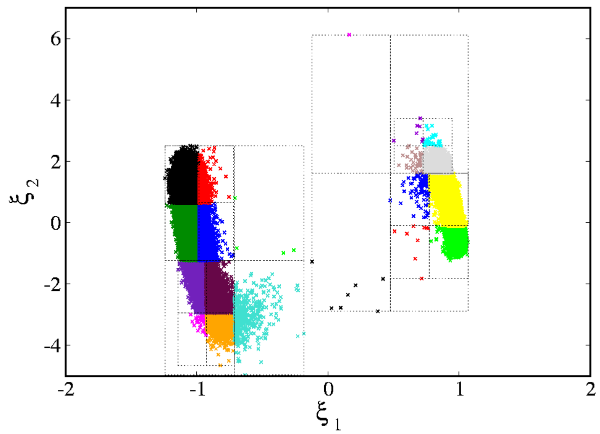


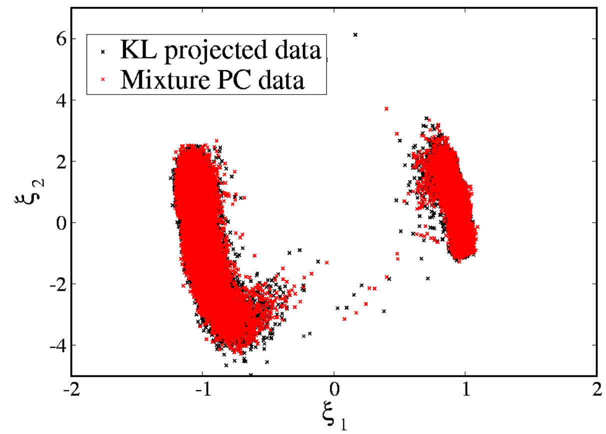


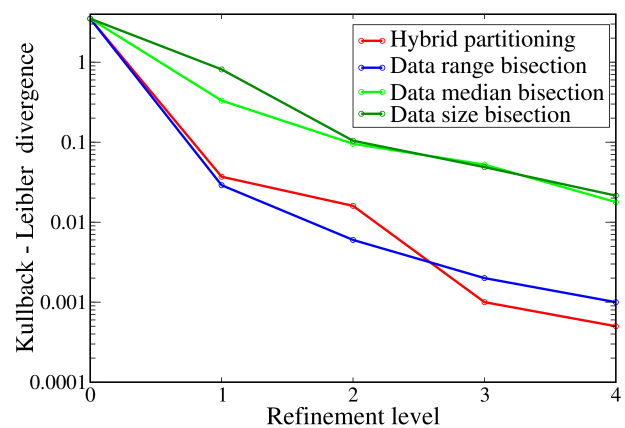


Convergence results confirm the hybrid approach superiority

Karhunen-Loève projection of the Schlogl model







Dynamical Analysis: Big Picture

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

$$\text{SSA} \longrightarrow \text{KL} \longrightarrow \text{PCE}$$

Random process $\longrightarrow$ L random variables $\longrightarrow$ L(P+1) deterministic variables

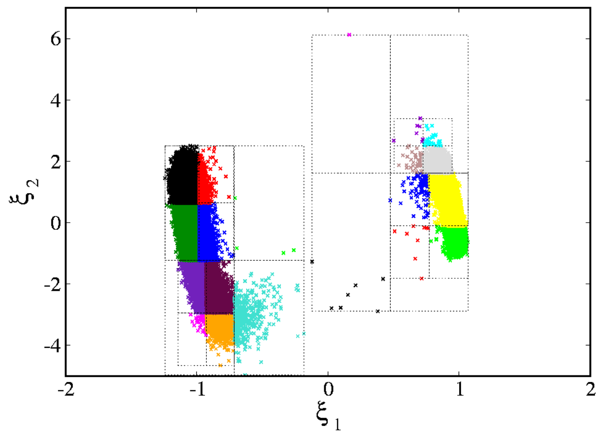
$$X(t,\theta) \longrightarrow \xi_i(\theta)(i=\overline{1,L}) \longrightarrow c_{ik}(i=\overline{1,L}, k=\overline{0,P})$$

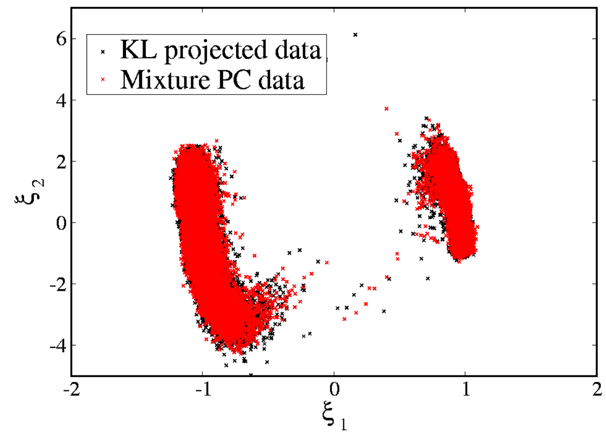
$$X(t,\theta)-\bar{X}(t) \simeq \sum_{i=1}^{L} \xi_i(\theta)\sqrt{\lambda_i} f_i(t) \simeq \sum_{i=1}^{L}\left(\sum_{k=0}^{P} c_{ik}\Psi_k(\boldsymbol{\eta})\right)\sqrt{\lambda_i} f_i(t)$$

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

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

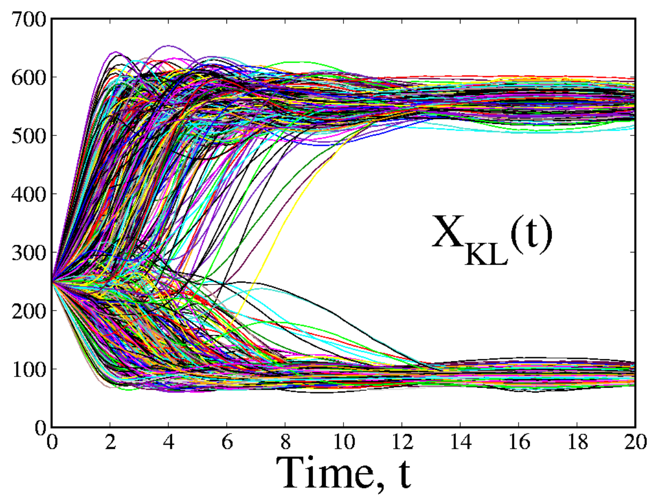
Results for the Schlögl model with Karhunen-Loève expansion

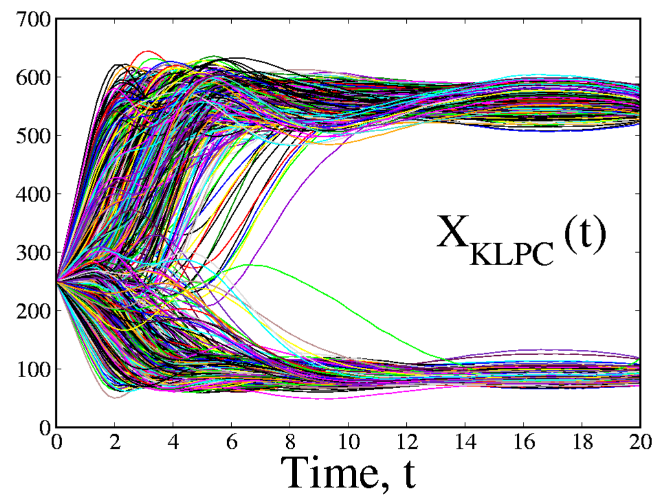




5 Mode KL Representation

3-rd Order PC Expansion





Conclusions and ongoing work

- Mixture model of PC PDFs with Karhunen-Loève decomposition represents the dynamics of the state $X(t, \boldsymbol{\theta}, \boldsymbol{\Lambda})$
- Karhunen-Loève expansion removes small-scale fluctuations
- Rosenblatt transformation maps to standard random variables
- Hybrid adaptive data partitioning for multimodal distributions

- Dimensionality (complexity increase) studies
- Sparse quadrature integration or Latin Hypercube Sampling
- Adaptive PC order
- 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
- K. Sargsyan, B. Debusschere, H. Najm & Y. Marzouk : 2009 : *Bayesian inference of spectral expansions for predictability assessment in stochastic reaction networks* – Predictability analysis in SRNs
- 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 & O. Le Maître : 2009 : *Spectral representation and reduced order modeling of the dynamics of stochastic reaction networks via adaptive data partitioning* – Submitted to SISC

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 $[-1, 1]$:

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

- Multi-domain PC expansion

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

where

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