Surrogate Model Development using Arbitrary Polynomial Chaos Expansion : Independent Variables

HyeongUk Lim and Lance Manuel

Outline

Use of Traditional PCE

Estimations of structural vibration responses are studied using traditional PCE:

- Estimation of VIV-induced fatigue damage
  Inputs: ΔA_{max} and $\Delta\omega$—Shifted Generalized Log-normal Distribution (SGLD)
- Long-term extreme responses of a moored floating structure
  Inputs: H_s and T_p—Lognormal and Weibull
- Wood floor vibrations
  Inputs: seven ρ_i—Gaussians

Hermite polynomial family has been applied
with a mapping from $\boldsymbol{X}$ (physical) to $\boldsymbol{Q}$ (standard normal);
also with Nataf and Rosenblatt transformation if needed.

Outside Askey Scheme

- For certain distributions, there exist a best representation which guarantees fast convergence rate: Askey scheme - Xiu, 2003.
- But, for inputs outside the Askey scheme, the Askey polynomial families converge, but not exponentially - Witteveen, 2006.

  Also, the other reasons are:
  - correlations between inputs - Eldred, 2009.
  - combinations of standard inputs - Oladyshkin, 2012.
- These problems cannot be handled efficiently by using the Askey polynomial families.

Transformation of Input Domain

Flowchart of Askey PCE

$\boldsymbol{X}$: an input vector

↓

$\boldsymbol{Q} = \xi(\boldsymbol{X})$

ξ: a mapping for the Askey scheme

↓

$$\widehat{Y}(\boldsymbol{Q}) = \sum_{i=1}^{P} c_i \Psi_i(\boldsymbol{Q})$$

$$\underbrace{Y(\boldsymbol{X})}_{\text{truth model in } \boldsymbol{X} \text{ domain}} = \underbrace{Y(\boldsymbol{Q})}_{\text{truth model in } \boldsymbol{Q} \text{ domain}} \equiv \underbrace{\widehat{Y}(\boldsymbol{Q})}_{\text{surrogate in } \boldsymbol{Q} \text{ domain}}$$

- ξ can be a non-linear
- Y and $\boldsymbol{X}$ vs. Y and $\boldsymbol{Q}$
- $\widehat{Y}(\boldsymbol{Q})$ aims to fit $Y(\boldsymbol{Q})$, not $Y(\boldsymbol{X})$

Ex. 1: Legendre PCE for a Gaussian input

A stupid example:

Truth model: $Y_X(x) = x^2, \quad X \sim N(0,1), \quad Y \sim \chi^2_{\text{dof}=1}$

What if we try Legendre PCE (instead of Hermite) to the truth model?

Q: Unif[-1,1] $\quad X$: N(0,1)

$$\boxed{\begin{gathered}\boldsymbol{Q} \;=\; \xi(\boldsymbol{X}): \\ \text{Askey scheme}\end{gathered}}$$

$$\begin{aligned} q &= F_Q^{-1}(F_X(x)) \\ &= F_Q^{-1}\left(\frac{1}{2}(1 + \text{erf}(\frac{x}{\sqrt{2}}))\right) \\ &= \text{erf}\left(\frac{x}{\sqrt{2}}\right) \end{aligned}$$

$\boldsymbol{Q}$		Ψ
Normal	$\Rightarrow$	Hermite
Uniform	$\Rightarrow$	Legendre

$$Y_X(x) = x^2 = Y_Q(q) = 2\left(\text{erf}^{-1}(q)\right)^2$$

Ex. 1: Legendre PCE for a Gaussian input (cont'd)

$$Y_X(x) = x^2$$

- $X \sim N(0, 1)$
- $X : \infty < x < \infty$
- A quadratic polynomial

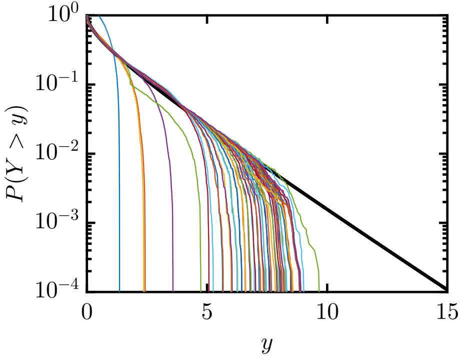
$$Y_Q(q) = 2\Big(\text{erf}^{-1}(q)\Big)^2$$

- $Q \sim \text{Unif}[-1, 1]$
- $Q : -1 < q < 1$
- Became complicated

$\widehat{Y}_Q(q) \approx \sum_{i=0}^{M} c_i \Psi_i(q)$: M should be larger than 2

Ex. 1: Legendre PCE for a Gaussian input (cont'd)

Comparison of PoEs by $1 - F_Y(y)$ and MCS of surrogate models



- Black solid line: $Y_X(x)$
- Others: $\widehat{Y}_Q(q)$ of orders from 1 to 60 with $3 \times \#$ LR samples
- $\widehat{Y}_Q(q)$ show slow convergence to estimate low PoEs (quiet inefficient)

But Hermite PCE with an order 2 shows the convergence to $Y_X(x) = x^2$

Ex. 1: Comparison of Coefficients and Basis Functions

Hermite PCE ($M = 2$), $X \sim N(0,1)$, $c_i = \frac{\mathbb{E}[Y_X(x)H_i(x)]}{\mathbb{E}[H_i^2(x)]}$

$$\widehat{Y}_X(x) = \sum_{i=0}^{M} c_i H_i(x) = \underset{c_0}{1} \cdot \underset{H_0}{1} + \underset{c_1}{0} \cdot \underset{H_1}{x} + \underset{c_2}{1} \cdot \underset{H_2}{(x^2 - 1)}$$

Legendre PCE ($M = 2$), $Q \sim \text{Unif}[-1,1]$, $c_i = \frac{\mathbb{E}[Y_Q(q)L_i(q)]}{\mathbb{E}[L_i^2(q)]}$

$$\widehat{Y}_Q(q) = \sum_{i=0}^{M} c_i L_i(q) = \underset{c_0}{1} \cdot \underset{L_0}{1} + \underset{c_1}{\left[3\int_{-1}^{1} \{\text{erf}^{-1}(q)\}^2 q \, dq\right]} \cdot \underset{L_1}{q}$$

$$+ \underset{c_2}{\left[\frac{5}{2}\int_{-1}^{1} \{\text{erf}^{-1}(q)\}^2 (3q^2 - 1) \, dq\right]} \cdot \underset{L_2}{\left\{\frac{1}{2}(3q^2 - 1)\right\}}$$

Ex. 2: Rosenbrock Function with Lognormal Variables

Truth model: $Y_{\boldsymbol{X}}(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \quad X_{1,2} \sim LN(1, 0.5)$

No Askey scheme for lognormal variables → try Hermite

$\boldsymbol{Q} = \xi(\boldsymbol{X})$:
Askey scheme

$$Q_1\colon\ N(0,1) \qquad X_1\colon\ LN(1, 0.5)$$

$$q_1 = F_Q^{-1}(F_X(x)) = \frac{\ln(x_1) - 1}{0.5}$$

$$Y_{\boldsymbol{Q}}(q_1, q_2) = 100(e^{0.5q_2+1} - e^{q_1+2})^2 + (1 - e^{0.5q_1+1})^2$$

Ex. 2: Comparison of Coefficients and Basis Functions

Truth model: $Y_{\boldsymbol{X}}(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \quad X_{1,2} \sim LN(1, 0.5)$

multi-index, α_i

	X_1	X_2	P_i
$i=0$	0	0	$P_{(0,0)}$
1	1	0	$P_{(1,0)}$
2	0	1	$P_{(0,1)}$
3	2	0	$P_{(2,0)}$
4	1	1	$P_{(1,1)}$
5	0	2	$P_{(0,2)}$

An illustrative case: $p = 2$

$$\text{PCE}: \quad \widehat{Y}_{\boldsymbol{Q}}(q_1, q_2) = \sum_{i=0}^{5} \overbrace{c_i}^{\text{LR}} \underbrace{H_i(q_1, q_2)}_{\text{pre-defined}}$$

$$\text{aPCE}: \quad \widehat{Y}_{\boldsymbol{X}}(x_1, x_2) = \sum_{i=0}^{5} \overbrace{c_i}^{\text{LR}} \underbrace{\Phi_i(x_1, x_2)}_{\text{Gram-Schmidt}}$$

E.g.)

$$H_4(q_1, q_2) = \underbrace{\widehat{H}_1(q_1)}_{Q_1} \underbrace{\widehat{H}_1(q_2)}_{Q_2}$$

$$L^2 \equiv \int_{\boldsymbol{X}} \Phi_i \Phi_j \underbrace{f_{\boldsymbol{X}}}_{\text{LN jpdf}} d\boldsymbol{x} = 0, \quad \text{for } i \neq j$$

Ex. 2: Univariate Gram-Schmidt Polynomials: $\widehat{\Phi}$

make $\{1,\ x_1,\ x_1^2\}$ orthogonal each other $\rightarrow \{1,\ \widehat{\Phi}_1,\ \widehat{\Phi}_2\}$

$$\widehat{\Phi}_0(x_1) = 1$$

$$\widehat{\Phi}_1(x_1) = x_1 - \frac{\langle x_1, \widehat{\Phi}_0 \rangle}{\langle \widehat{\Phi}_0, \widehat{\Phi}_0 \rangle} \underbrace{\widehat{\Phi}_0(x_1)}_{=1}$$

$$\widehat{\Phi}_2(x_1) = x_1^2 - \frac{\langle x_1^2, \widehat{\Phi}_0 \rangle}{\langle \widehat{\Phi}_0, \widehat{\Phi}_0 \rangle} \underbrace{\widehat{\Phi}_0(x_1)}_{=1} - \frac{\langle x_1^2, \widehat{\Phi}_1 \rangle}{\langle \widehat{\Phi}_1, \widehat{\Phi}_1 \rangle} \widehat{\Phi}_1(x_1)$$

$$\langle \widehat{\Phi}_i, \widehat{\Phi}_i \rangle \equiv \int_{\boldsymbol{X}} (\widehat{\Phi}_i)^2 f_{\boldsymbol{X}} \, d\boldsymbol{x}$$

$$m_p = \int x_1^p f_{X_1} \, dx_1$$

$$\widehat{\Phi}_1(x_1) = x_1 - \frac{\langle x_1, 1 \rangle}{\langle 1, 1 \rangle} \cdot 1 = x_1 - \int_{X_1} x_1 f_{X_1} \, dx_1 = x_1 - m_1$$

$$\widehat{\Phi}_2(x_1) = x_1^2 - \frac{\langle x_1, 1 \rangle}{\langle 1, 1 \rangle} \cdot 1 - \frac{\langle x_1^2, x_1 - m_1 \rangle}{\langle x_1 - m_1, x_1 - m_1 \rangle} (x_1 - m_1)$$

$$= \underbrace{1}_{a_2^{(2)}} \cdot x_1^2 + \underbrace{\left(\frac{m_1 m_2 - m_3}{m_2 - m_1^2} \right)}_{a_1^{(2)}} x_1 + \underbrace{\left(\frac{m_1 m_3 - m_2^2}{m_2 - m_1^2} \right)}_{a_0^{(2)}}$$

$$\widehat{\Phi}_i(x_1) = \sum_{k=0}^{i} a_k^{(i)} x_1^k$$

for $a_k^{(i)}$
a matrix form
is compact

Ex. 2: Comparison of Multi-indices

$$
\begin{aligned}
Y_{\boldsymbol{X}}(x_1, x_2) &= 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \\
&= \underbrace{1}_{(0,0)} - \underbrace{2x_1}_{(1,0)} + \underbrace{x_1^2}_{(2,0)} + \underbrace{100x_2^2}_{(0,2)} - \underbrace{200x_1^2 x_2}_{(2,1)} + \underbrace{100x_1^4}_{(4,0)}
\end{aligned}
$$

$$
\begin{aligned}
Y_{\boldsymbol{Q}}(q_1, q_2) &= 100(e^{0.5q_2+1} - e^{q_1+2})^2 + (1 - e^{0.5q_1+1})^2 \\
&= 1 - 2e^{0.5q_1+2} + e^{q_1+2} + 100e^{q_1+2} - 200e^{q_1+0.5q_2+3} + 100e^{2q_1+4}
\end{aligned}
$$

multi-index inconveniently applicable

$e^{q_1} = 1 + q_1 + \frac{q_1^2}{2!} + \frac{q_1^3}{3!} + \cdots$

Ex. 2: Parameter Estimation

Same 45 random samples are used as a part of LR; $3 \times \binom{4+2}{2} = 45$.

$$Y_{\boldsymbol{X}}(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \approx \widehat{Y}_{\boldsymbol{X}}(x_1, x_2) = \sum_{i=0}^{14} c_i \Phi_i(x_1, x_2)$$

$$Y_{\boldsymbol{Q}}(q_1, q_2) = 100(e^{0.5q_2+1} - e^{q_1+2})^2 + (1 - e^{0.5q_1+1})^2 \approx \widehat{Y}_{\boldsymbol{Q}}(q_1, q_2) = \sum_{i=0}^{14} c_i H_i(q_1, q_2)$$

c_i	H_i (pre-defined)	c_i	Φ_i (Gram-Schmidt)
24323	1	34063	1
41547	q_1	73678	$x_1 - 3.08$
-1985	q_2	-1533	$x_2 - 3.08$
36366	$q_1^2 - 1$	33880	$x_1^2 - 9.03x_1 + 15.64$
-2488	$q_1 q_2$	-1807	$(x_1 - 3.08)(x_2 - 3.08)$
-2170	$q_2^2 - 1$	100	$x_2^2 - 9.03x_2 + 15.64$
13018	$q_1^3 - 3q_1$	3945	$x_1^3 - 19.97x_1^2 + 101.43x_1 - 130.97$
-29	$q_2(q_1^2 - 1)$	-200	$(x_1^2 - 9.03x_1 + 15.64)(x_2 - 3.08)$
-942	$q_1(q_2^2 - 1)$	0	$(x_1 - 3.08)(x_2^2 - 9.03x_2 + 15.64)$
-47	$q_2^3 - 3q_2$	0	$x_2^3 - 19.97x_2^2 + 101.43x_2 - 130.97$
4411	$q_1^4 - 6q_1^2 + 3$	100	$x_1^4 - 39.4x_1^3 + 442.9x_1^2 - 1677.4x_1 + 1808$
10	$q_2(q_1^3 - 3q_1)$	0	$(x_1^3 - 19.97x_1^2 + 101.43x_1 - 130.97)(x_2 - 3.08)$
-117	$(q_1^2 - 1)(q_2^2 - 1)$	0	$(x_1^2 - 9.03x_1 + 15.64)(x_2^2 - 9.03x_2 + 15.64)$
-298	$q_1(q_2^3 - 3q_2)$	0	$(x_1 - 3.08)(x_2^3 - 19.97x_2^2 + 101.43x_2 - 130.97)$
-560	$q_2^4 - 6q_2^2 + 3$	0	$x_2^4 - 39.4x_2^3 + 442.9x_2^2 - 1677.4x_2 + 1808$

Ex. 2: Coefficients by LR: OLS

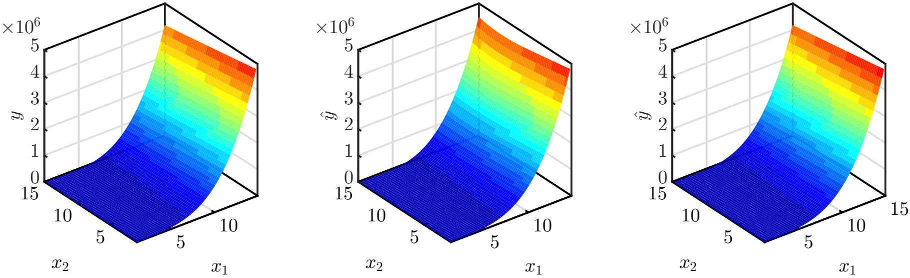
For c_i, the same samples are drawn for LR:

$$\underset{\substack{\text{normal}\\\text{space}}}{\overset{\text{PCE}}{\boldsymbol{Q}}} \Leftarrow \underset{\substack{\text{uniform}\\\text{random numbers}}}{\boldsymbol{U}} \Rightarrow \underset{\substack{\text{lognormal}\\\text{space}}}{\overset{\text{aPCE}}{\boldsymbol{X}}}$$

$$\begin{bmatrix} H_0(\boldsymbol{q}^{(1)}) & \cdots & H_{14}(\boldsymbol{q}^{(1)}) \\ \vdots & \ddots & \vdots \\ H_0(\boldsymbol{q}^{(45)}) & \cdots & H_{14}(\boldsymbol{q}^{(45)}) \end{bmatrix} \begin{Bmatrix} c_0 \\ \vdots \\ c_{14} \end{Bmatrix} = \overbrace{\begin{Bmatrix} Y_{\boldsymbol{X}}(\boldsymbol{x}^{(1)}) \\ \vdots \\ Y_{\boldsymbol{X}}(\boldsymbol{x}^{(45)}) \end{Bmatrix}}^{\text{4th order polynomial}} = \overbrace{\begin{Bmatrix} Y_{\boldsymbol{Q}}(\boldsymbol{q}^{(1)}) \\ \vdots \\ Y_{\boldsymbol{Q}}(\boldsymbol{q}^{(45)}) \end{Bmatrix}}^{\substack{\text{complicated than}\\ Y_{\boldsymbol{X}}}}$$

$$\begin{bmatrix} \Phi_0(\boldsymbol{x}^{(1)}) & \cdots & \Phi_{14}(\boldsymbol{x}^{(1)}) \\ \vdots & \ddots & \vdots \\ \Phi_0(\boldsymbol{x}^{(45)}) & \cdots & \Phi_{14}(\boldsymbol{x}^{(45)}) \end{bmatrix} \begin{Bmatrix} c_0 \\ \vdots \\ c_{14} \end{Bmatrix} = \begin{Bmatrix} Y_{\boldsymbol{X}}(\boldsymbol{x}^{(1)}) \\ \vdots \\ Y_{\boldsymbol{X}}(\boldsymbol{x}^{(45)}) \end{Bmatrix}$$

Ex. 2: Response Surfaces



Truth

PCE
Order= 13

aPCE
Order= 4

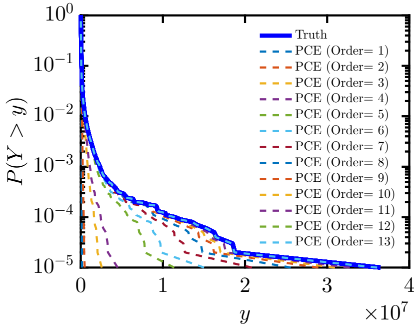
- An order-4 aPCE shows convergence to the truth.
- An order-13 PCE still has inaccuracy at some regions.

Ex. 2: Exceedance Probabilities

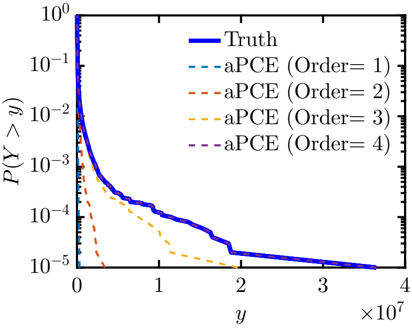
Truth: $Y_{\boldsymbol{X}}(x_1, x_2)$ PCE: $\widehat{Y}_{\boldsymbol{Q}}(q_1, q_2)$ aPCE: $\widehat{Y}_{\boldsymbol{X}}(x_1, x_2)$

$P(Y > y)$ obtained empirically by using MCS of truth and surrogates

Identical random seeds are used for the comparison

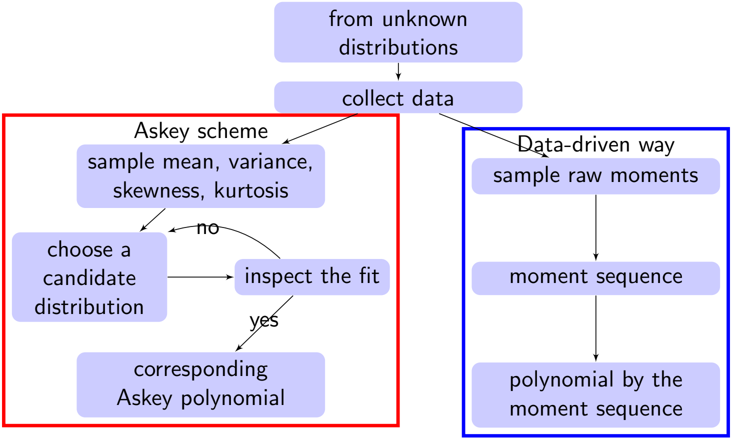


PCE



aPCE

Framework: APCE



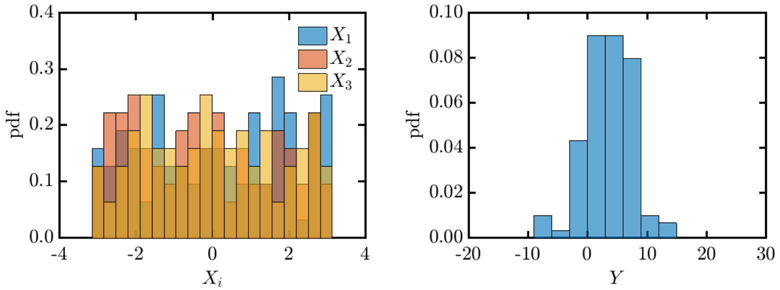
Ishigami function

The Ishigami function:

$$Y_{\boldsymbol{X}}(\boldsymbol{x}) = \sin(x_1) + a\big(\sin(x_2)\big)^2 + bx_3^4\sin(x_1).$$

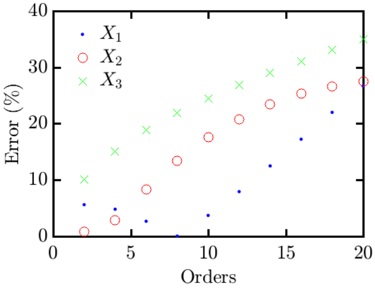
X_1, X_2, and X_3 follow uniform distributions over $[-\pi, \pi]^3$.
a and b are 7 and 0.1, respectively.

Data is given: $\mathcal{D} = \{\boldsymbol{x}_i, y_i\}_{i=1}^{N}$, $N = 100$



Sample Raw Moments

$$\text{Error (\%)} = \frac{|\text{Exact} - \text{SRM}|}{\text{Exact}} \times 100$$

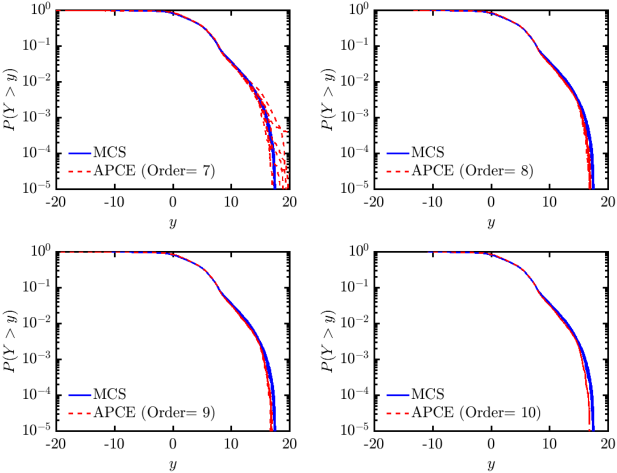


APCE

$$\widehat{Y}_{\boldsymbol{X}}(\boldsymbol{x}) = \sum_{i=0}^{P} c_i \Psi_i(\boldsymbol{x})$$

Ψ is constructed based on $\mathcal{D}$;
c_i is also constructed based on $\mathcal{D}$ with resampling (if needed);
Uncertainty propagation is also done by resampling in $\mathcal{D}$.

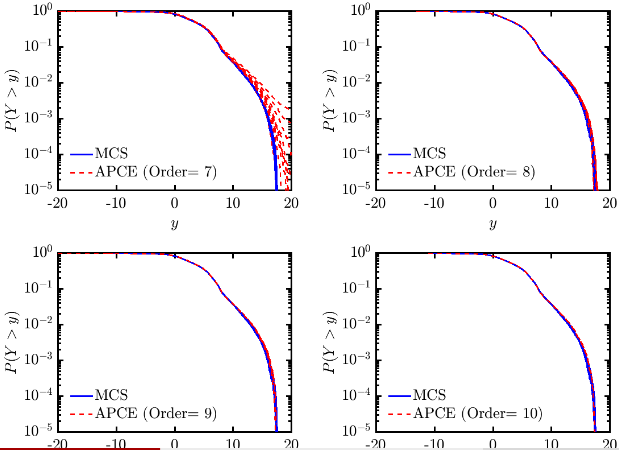
Exceedance Plots



An order-8 APCE models look convergent to the truth.
The estimations at low $P(Y > y)$ levels exhibit inaccuracy.

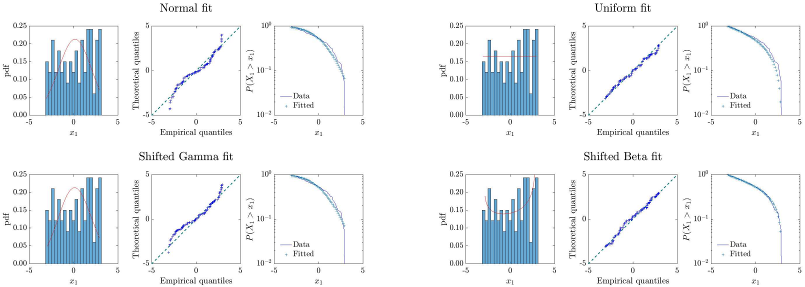
Reason for Inaccuracy

Inaccurate estimation of high-order raw moments ⇒ inaccurate estimation of high-order basis functions
Accuracy is improved when $N = 1000$.
High variability of order-7 APCE models might be due to high sampling variability when $N = 1000$.



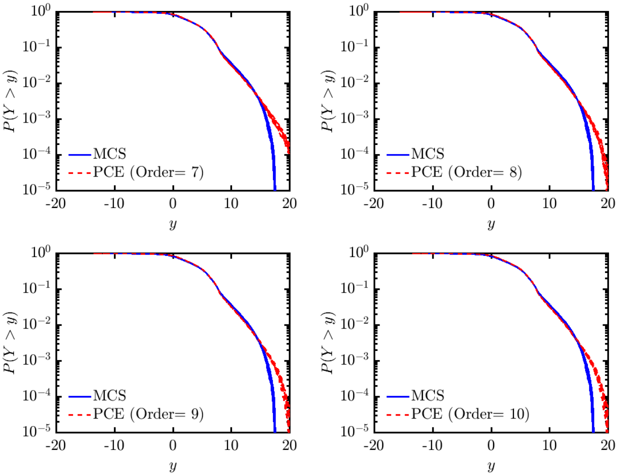
Classic PCE: Askey Scheme

Candidate distributions are investigated for the Askey scheme PCE.
Normal (Hermite), Uniform (Legendre), Gamma (Laguerre), Beta (Jacobi)



A Shifted Beta fit is selected; Jacobi polynomials are used.

Exceedance Plots



Pros and Cons: APCE

Pros:

- Can make a surrogate without fitting input variables to certain distribution types.
- Fast convergence to the truth model (regarding response surface).

Cons:

- Surrogate model accuracy depends on the estimation of raw moments.
- High-order sample raw moments can be inaccurate with a small number of samples.
- A high-order basis function can be inaccurate, too.

Future work:

- To improve accuracy of high-order raw moment approximation.

Thank you very much for your attention!