A Spline Dimensional Decomposition for High-Dimensional Uncertainty Quantification

Sharif Rahman and Ramin Jahanbin
The University of Iowa, Iowa City, IA 52242

HDA 2019: 8th Workshop on HDA
Zurich, Switzerland
September 9-13, 2019

Acknowledgment: NSF (CMMI-1607398)

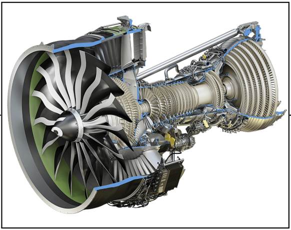
Outline

1. INTRODUCTION
2. SDD
3. EXAMPLES
4. CLOSURE

Uncertainty Quantification

Complex System
(jet engine)

Input $\mathbf{X} = (X_1, \ldots, X_N)$
$\mathbf{X} : (\Omega, \mathcal{F}) \to (\mathbb{A}^N, \mathcal{B}^N) \to$
$\mathbb{A}^N \subseteq \mathbb{R}^N, \; N \in \mathbb{N}$



Output $Y = y(\mathbf{X})$
$\to \; Y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$
$y \in L^2(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}} d\mathbf{x})$

- **Goals & Objectives**
  - Moments: $\mathbb{E}\left[Y^l\right] := \int_\Omega Y^l d\mathbb{P} = \int_{\mathbb{A}^N} y^l(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$
  - Probability distribution: $\mathbb{P}\left[Y \leq y_0\right] := \int_{\{\mathbf{x}: y(\mathbf{x}) \leq y_0\}} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$
  - Stochastic design optimization (RDO/RBDO)

UQ Challenges & Methods

- **Challenges (Works at Iowa)**
  - High-dimensional random input ($N > 10$)
  - Locally prominent (nonsmoothness, discontinuity) responses
  - Statistical dependence among random input
  - Data-driven problems
- **Polynomial Expansion Methods (PCE & PDD)**

$$y_{\mathbf{p}}(\mathbf{X}) = \sum_{\mathbf{0} \le \mathbf{i} \le \mathbf{p}} C_{\mathbf{i}} \Psi_{\mathbf{i}}(\mathbf{X}) \quad \text{(PCE)}$$

$$y_{S,\mathbf{p}}(\mathbf{X}) = y_{\emptyset} + \sum_{\substack{\emptyset \ne u \subseteq \{1,\ldots,N\} \\ 1 \le |u| \le S}} \sum_{\mathbf{0} \le \mathbf{i}_u \le \mathbf{p}_u} C^u_{\mathbf{i}_u} \Psi^u_{\mathbf{i}_u}(\mathbf{X}_u) \quad \text{(PDD)}$$

Explore spline basis equipped with local support

Assumptions

The random vector $\mathbf{X} := (X_1, \ldots, X_N)^\intercal : (\Omega, \mathcal{F}) \to (\mathbb{A}^N, \mathcal{B}^N)$ satisfies the following conditions:

1. All component random variables X_k, $k = 1, \ldots, N$, are statistically independent, but not necessarily identical.
2. Each input random variable X_k has absolute continuous marginal CDF and continuous marginal PDF.
3. Each input random variable X_k is defined on a closed bounded interval $[a_k, b_k] \subset \mathbb{R}$, $b_k > a_k$, so that all moments exist, *i.e.*, for $l \in \mathbb{N}_0$,

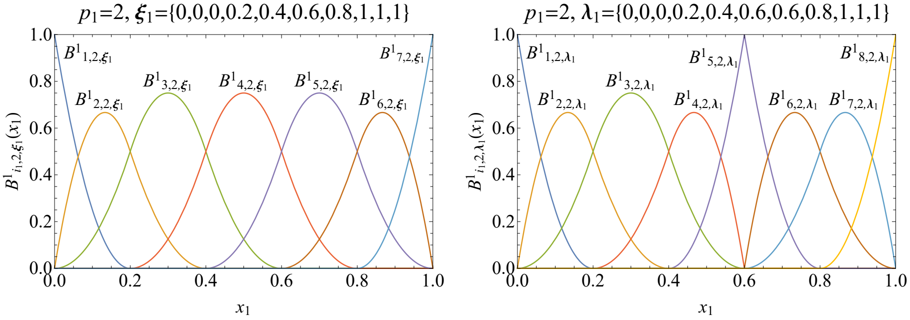
$$\mathbb{E}\left[X_k^l\right] := \int_\Omega X_k^l(\omega)\, d\mathbb{P}(\omega) = \int_{a_k}^{b_k} x_k^l f_{X_k}(x_k)\, dx_k < \infty.$$

Univariate B-Splines (Cox & de Boor, 1972)

For a knot sequence $\boldsymbol{\xi}_k = \{a_k = \xi_{k,1}, \ldots, \xi_{k,n_k+p_k+1} = b_k\}$, where $\xi_{k,1} \leq \cdots \leq \xi_{k,n_k+p_k+1}$, $n_k > p_k \geq 0$, the B-splines are

$$B^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k) := \frac{(x_k - \xi_{k,i_k})B^k_{i_k,p_k-1,\boldsymbol{\xi}_k}(x_k)}{\xi_{k,i_k+p_k} - \xi_{k,i_k}} + \frac{(\xi_{k,i_k+p_k+1} - x_k)B^k_{i_k+1,p_k-1,\boldsymbol{\xi}_k}(x_k)}{\xi_{k,i_k+p_k+1} - \xi_{k,i_k+1}},$$

$$1 \leq k \leq N, 1 \leq i_k \leq n_k, 1 \leq p_k < \infty.$$



Univariate ON B-Splines

- **Auxiliary B-Spline Vector**

$$\mathbf{P}_k(x_k) := \left(1, B^k_{2,p_k,\boldsymbol{\xi}_k}(x_k), \ldots, B^k_{n_k,p_k,\boldsymbol{\xi}_k}(x_k)\right)^\intercal$$

- **Spline Moment Matrix**

$$\mathbf{G}_k := \mathbb{E}[\mathbf{P}_k(X_k)\mathbf{P}_k^\intercal(X_k)] \in \mathbb{R}^{n_k \times n_k}$$

$$\mathbf{G}_k \to \text{symmetric, positive}-\text{definite}$$

- **Whitening Transformation**

$$\boldsymbol{\psi}_k(x_k) = \mathbf{Q}_k^{-1}\mathbf{P}_k(x_k), \text{ where } \mathbf{G}_k = \mathbf{Q}_k\mathbf{Q}_k^\intercal$$

For $k = 1, \ldots, N$, let $\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k}$ be a space real-valued splines in x_k of degree p_k and knot sequence $\boldsymbol{\xi}_k$. Then

$$\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k} = \text{span}\left\{\psi^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k)\right\}_{i_k=1,\ldots,n_k}.$$

Multivariate ON B-Splines

Given $N \in \mathbb{N}$, let $\emptyset \neq u \subseteq \{1, \ldots, N\}$. For $\mathbf{i}_u := (i_{k_1}, \ldots, i_{k_{|u|}})$, $\mathbf{p}_u := (p_{k_1}, \ldots, p_{k_{|u|}})$, $\boldsymbol{\Xi}_u := (\boldsymbol{\xi}_{k_1}, \ldots, \boldsymbol{\xi}_{k_{|u|}})$, the tensor-product ON B-splines in $\mathbf{x}_u = (x_{k_1}, \ldots, x_{k_{|u|}})$ are

$$\boxed{\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{x}_u) = \prod_{k \in u} \psi^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k), \;\; \mathbf{i}_u \in \bar{\mathcal{I}}_{u,\mathbf{n}_u}.}$$

$$\bar{\mathcal{I}}_{u,\mathbf{n}_u} := \left\{ \mathbf{i}_u = (i_{k_1}, \ldots, i_{k_{|u|}}) : 2 \leq i_{k_l} \leq n_{k_l}, l = 1, \ldots, |u| \right\}$$

The second-moment properties are

$$\mathbb{E}\left[\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{X}_u)\right] = 0,$$

$$\mathbb{E}\left[\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{X}_u)\Psi^v_{\mathbf{j}_v,\mathbf{p}_v,\boldsymbol{\Xi}_v}(\mathbf{X}_v)\right] = \begin{cases} 1, & u = v \text{ and } \mathbf{i}_u = \mathbf{j}_v, \\ 0, & \text{otherwise.} \end{cases}$$

Dimensionwise Spline Space Splitting

For $\mathbf{p} = (p_1, \ldots, p_N) \in \mathbb{N}_0^N$ & $\boldsymbol{\Xi} = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N\}$, let $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$ be the space of all real-valued splines of degree $\mathbf{p}$ in $\mathbf{x} = (x_1, \ldots, x_N)$. Then

$$
\begin{aligned}
\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}} &= \bigotimes_{k=1}^{N} \left(\mathbf{1} \oplus \bar{\mathcal{S}}_{k,p_k,\boldsymbol{\xi}_k}\right) \\
&= \mathbf{1} \oplus \bigoplus_{\emptyset \neq u \subseteq \{1,\ldots,N\}} \bar{\mathcal{S}}^u_{\mathbf{p}_u,\boldsymbol{\Xi}_u} \\
&= \mathbf{1} \oplus \bigoplus_{\emptyset \neq u \subseteq \{1,\ldots,N\}} \operatorname{span}\left\{\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{x}_u)\right\}_{\mathbf{i}_u \in \bar{\mathcal{I}}_{u,\mathbf{n}_u}} .
\end{aligned}
$$

$$
\bar{\mathcal{S}}^u_{\mathbf{p}_u,\boldsymbol{\Xi}_u} = \bigotimes_{k \in u} \bar{\mathcal{S}}_{k,p_k,\boldsymbol{\xi}_k} = \operatorname{span}\left\{\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{x}_u)\right\}_{\mathbf{i}_u \in \bar{\mathcal{I}}_{u,\mathbf{n}_u}} \quad \text{(zero mean)}
$$

$$
\bar{\mathcal{S}}_{k,p_k,\boldsymbol{\xi}_k} = \operatorname{span}\left\{\psi^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k)\right\}_{i_k = 2,\ldots,n_k} \quad \text{(zero mean)}
$$

Spline Dimensional Decomposition

Theorem

Under Assumptions 1-3, a random variable $y(\mathbf{X}) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ *admits a hierarchical orthogonal expansion in multivariate ON spline basis* $\{\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\mathbf{\Xi}_u}(\mathbf{X}_u)\}$, *referred to as the SDD of*

$$y_{\mathbf{p},\mathbf{\Xi}}(\mathbf{X}) := y_\emptyset + \sum_{\emptyset \neq u \subseteq \{1,\ldots,N\}} \sum_{\mathbf{i}_u \in \bar{\mathcal{I}}_{u,\mathbf{n}_u}} C^u_{\mathbf{i}_u,\mathbf{p}_u,\mathbf{\Xi}_u} \Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\mathbf{\Xi}_u}(\mathbf{X}_u),$$

$$\text{where } y_\emptyset := \int_{\mathbb{A}^N} y(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x},$$

$$C^u_{\mathbf{i}_u,\mathbf{p}_u,\mathbf{\Xi}_u} := \int_{\mathbb{A}^N} y(\mathbf{x}) \Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\mathbf{\Xi}_u}(\mathbf{x}_u) f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}.$$

Moreover, the SDD of $y(\mathbf{X})$ *is the best approximation, i.e.,*

$$\mathbb{E}\left[y(\mathbf{X}) - y_{\mathbf{p},\mathbf{\Xi}}(\mathbf{X})\right]^2 = \inf_{g \in \mathcal{S}_{\mathbf{p},\mathbf{\Xi}}} \mathbb{E}\left[y(\mathbf{X}) - g(\mathbf{X})\right]^2.$$

Error Bound & Convergence

- **Modulus of smoothness** $(\alpha_k \geq 1)$

$$\omega_{\alpha_k}(y;h_k)_{L^2[a_k,b_k]} := \sup_{0\leq u_k\leq h_k} \left\| \Delta_{u_k}^{\alpha_k} y(x_k) \right\|_{L^2[a_k,b_k-\alpha_k u_k]}, \quad h_k \geq 0,$$

$$\omega_{\boldsymbol{\alpha}}(y;\mathbf{h})_{L^2[\mathbb{A}^N]} := \sup_{\mathbf{0}\leq\mathbf{u}\leq\mathbf{h}} \left\| \Delta_{\mathbf{u}}^{\boldsymbol{\alpha}} y(\mathbf{x}) \right\|_{L^2[\mathbb{A}^N_{\boldsymbol{\alpha},\mathbf{u}}]}, \quad \mathbf{h} \geq \mathbf{0}$$

- L^2**-error**

$$\mathbb{E}\left[\left|y(\mathbf{X}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right|^2\right] \leq C\omega_{\mathbf{p+1}}(y;\mathbf{h})_{L^2(\mathbb{A}^N)}$$

$$\lim_{\mathbf{h}\to\mathbf{0}} \mathbb{E}\left[\left|y(\mathbf{X}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right|^2\right] = 0$$

SDD converges in m.s., in probability and in distribution.

Truncation

- S**-variate, SDD Approximation** (Poly. Complexity)

$$y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X}) := y_\emptyset + \sum_{\substack{\emptyset\neq u\subseteq\{1,\ldots,N\}\\ 1\leq|u|\leq S}} \sum_{\mathbf{i}_u\in\bar{\mathcal{I}}_{u,\mathbf{n}_u}} C^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u} \Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{X}_u)$$

$$\text{No. of coeff., } L_{S,\mathbf{p},\boldsymbol{\Xi}} = 1 + \sum_{s=1}^{S} \binom{N}{s} \prod_{k=1}^{s} (n_k - 1) \leq \prod_{k=1}^{N} n_k$$

$$(N = 15,\ n_k = 5,\ S = 1 \text{ or } 2:\ L_{S,\mathbf{p},\boldsymbol{\Xi}} = 61 \text{ or } 1741 \ll 5^{15})$$

- **Second-Moment Statistics**

$$\mathbb{E}\left[y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] = y_\emptyset = \mathbb{E}\left[y(\mathbf{X})\right]$$

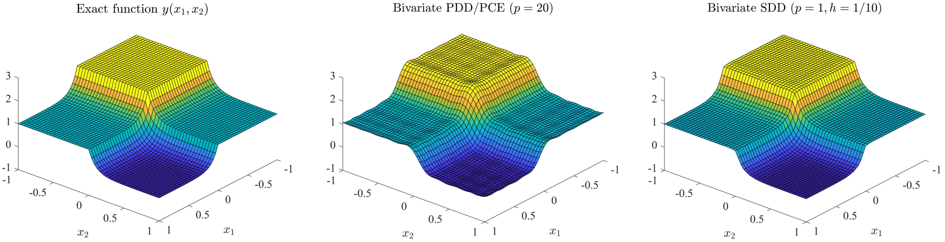
$$\operatorname{var}\left[y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] = \sum_{\substack{\emptyset\neq u\subseteq\{1,\ldots,N\}\\ 1\leq|u|\leq S}} \sum_{\mathbf{i}_u\in\bar{\mathcal{I}}_{u,\mathbf{n}_u}} {C^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}}^2 \leq \operatorname{var}\left[y(\mathbf{X})\right]$$

Example 1: A Nonsmooth Function ($N = 2$)

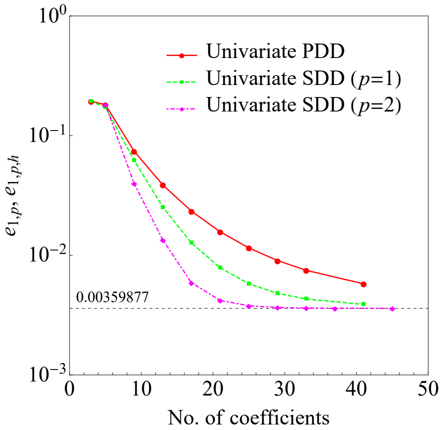
Defined on the square $\mathbb{A}^2 = [-1, 1]^2$, consider a nonsmooth function

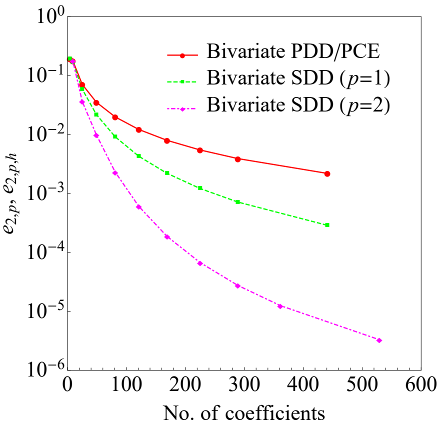
$$y(X_1, X_2) = g(X_1) + g(X_2) + \frac{1}{5} g(X_1) g(X_2), \quad X_1, X_2 \sim \text{i.i.d. } U[-1, 1],$$

$$g(x_i) = \begin{cases} 1, & -1 \le x_i \le 0, \\ \exp(-10 x_i), & 0 < x_i \le 1. \end{cases}$$



Example 1: Variance Errors





Example 2: A Linear Elasticity Problem ($N = 15$)

A twisting horseshoe

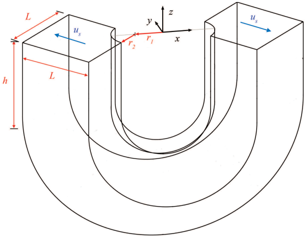
- Stochastic PDE (Elliptical)

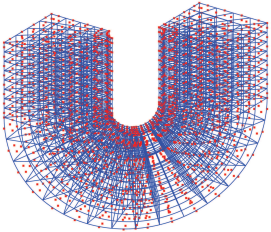
$$\nabla \cdot \boldsymbol{\sigma}(\mathbf{z};\mathbf{X}) = \mathbf{0} \text{ in } \mathcal{D} \subset \mathbb{R}^3,$$
$$\boldsymbol{\sigma}(\mathbf{z};\mathbf{X}) \cdot \mathbf{n}(\mathbf{z};\mathbf{X}) = \bar{\mathbf{t}}(\mathbf{z};\mathbf{X}) \text{ on } \partial\mathcal{D}_t,$$
$$\mathbf{u}(\mathbf{z};\mathbf{X}) = \bar{\mathbf{u}}(\mathbf{z};\mathbf{X}) \text{ on } \partial\mathcal{D}_u,$$
$$\partial\mathcal{D}_t \cup \partial\mathcal{D}_u = \partial\mathcal{D},\ \partial\mathcal{D}_t \cap \partial\mathcal{D}_u = \emptyset.$$

- Random Input ($N = 15$)

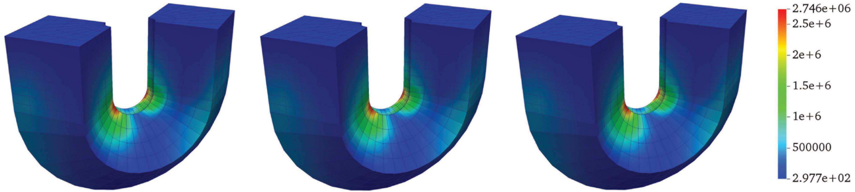
$$E(\mathbf{z};\cdot) = C_\alpha \exp[\alpha(\mathbf{z};\cdot)],$$
$$C_\alpha = \mu_E / \sqrt{1 + \nu_E^2},$$
$$\alpha(\mathbf{z};\cdot) \rightarrow \text{homogen. Gaussian RF},$$
$$\Gamma_\alpha(\mathbf{z},\mathbf{z}') = \sigma^2 \exp(-||\mathbf{z} - \mathbf{z}'||/bL),$$
$$\alpha(\mathbf{z};\cdot) = \sum_{i=1}^{15} \sqrt{\lambda_i}\phi_i(\mathbf{z}) X_i.$$

SDD coeffs. est. by dim.-red. integ.



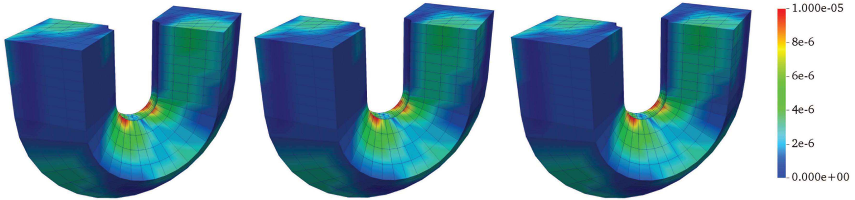


Example 2: St. Dev. of Displacement Field



Univariate SDD ($p = I = 2$) Bivariate SDD ($p = I = 2$) MCS (10,000 samples)

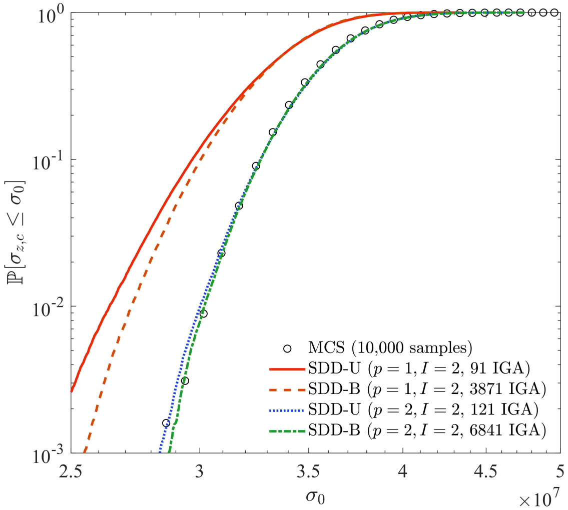
normal stress σ_z



Univariate SDD ($p = I = 2$) Bivariate SDD ($p = I = 2$) MCS (10,000 samples)

shear strain ϵ_{xz}

Example 2: Probability Distribution of a Critical Stress



Conclusion

- A new ON spline expansion (SDD) is introduced.
- Comp. effort scales polynomially, not exponentially.
- SDD converges in m.s. and others weaker modes.
- A low-order SDD is more accurate than high-order PDD/PCE for nonsmooth functions.

Future work

- Explore nonuniform knot sequences.
- Study unbounded domains without transformation.