
STOCHASTIC GALERKIN METHODS

STOCHASTIC GALERKIN METHODS

- Functions of the parameters have to be discretized in much the same way functions of the (finite number of) spatial variables have to be discretized
 - spatial discretization is effected via a standard finite element discretization in the usual manner by choosing a J -dimensional subspace $X_J \subset X$
 - let $\{\phi_j(\vec{y})\}_{j=1}^J$ denote a basis for X_J
- Stochastic Galerkin methods are methods for which discretization with respect to parameter space is also effected using a Galerkin approach, i.e.,
 - we choose a K -dimensional subspace $Z_K \subset Z$
 - let $\{\psi_k(\vec{y})\}_{k=1}^K$ denote a basis for the parameter approximating space Z_K

- Due to the product nature of the domain $\mathcal{D} \otimes \Gamma$ and of the space $X \otimes Z$, it is natural to seek approximations that use this structure, i.e.,
 - approximations are defined as a sum of products of the spatial and probabilistic basis functions
- Thus, we seek an approximate solution of the SPDE of the form[†]

$$u_{JK} = \sum_{j=1}^J \sum_{k=1}^K c_{jk} \phi_j(\mathbf{x}) \psi_k(\vec{y}) \in X_J \times Z_K$$

- The coefficients c_{jk} , and therefore u_{JK} , are determined by solving the problem

$$\int_{\mathcal{D}} \int_{\Gamma} \rho(\vec{y}) S(u_{JK}, \vec{y}) T(v) d\mathbf{x} d\vec{y} = \int_{\mathcal{D}} \int_{\Gamma} \rho(\vec{y}) v f(\vec{y}) d\mathbf{x} d\vec{y} \quad \forall v \in X_J \times Z_K$$

[†]Potentially, some economies can be effected if one also approximates the data functions (e.g., coefficients) appearing in the problem in the same way one approximates the solution, e.g., for a data function $a(\mathbf{x}; \vec{y})$, one determines $a_k(\mathbf{x})$, $k = 1, \dots, K$, such that

$$\sum_{k=1}^K a_k(\mathbf{x}) \psi_k(\vec{y}) \approx a(\mathbf{x}; y_1, \dots, y_N);$$

in actuality, these economies can be realized only in very limited settings; more on this later

- We then have that the discretized problem

$$\int_{\mathcal{D}} \int_{\Gamma} \rho(\vec{y}) S \left(\sum_{j=1}^J \sum_{k=1}^K c_{jk} \phi_j(\mathbf{x}) \psi_k(\vec{y}), \vec{y} \right) T \left(\phi_{j'}(\mathbf{x}) \right) \psi_{k'}(\vec{y}) d\mathbf{x} d\vec{y}$$

$$= \int_{\mathcal{D}} \int_{\Gamma} \rho(\vec{y}) \phi_{j'}(\mathbf{x}) \psi_{k'}(\vec{y}) f(\vec{y}) d\mathbf{x} d\vec{y}$$

for $j' \in \{1, \dots, J\}$ and $k' \in \{1, \dots, K\}$

- Of course, the solution

$$u_{JK}(\mathbf{x}; \vec{y}) = \sum_{j=1}^J \sum_{k=1}^K c_{jk} \phi_j(\mathbf{x}) \psi_k(\vec{y})$$

of this problem is independent of the basis set used

- although the coefficients c_{jk} do depend on the choice of basis

- In general, the integrals cannot be evaluated exactly
 - quadrature rules must be invoked to effect approximate evaluations
 - thus, the integrals with respect to the parameter domain[†] Γ are approximated by a quadrature rule to obtain

$$\begin{aligned}
 & \sum_{r=1}^R \hat{w}_r \rho(\hat{\mathbf{y}}_r) \psi_{k'}(\hat{\mathbf{y}}_r) \int_{\mathcal{D}} S \left(\sum_{j=1}^J \sum_{k=1}^K c_{jk} \phi_j(\mathbf{x}) \psi_k(\hat{\mathbf{y}}_r), \hat{\mathbf{y}}_r \right) T \left(\phi_{j'}(\mathbf{x}) \right) d\mathbf{x} \\
 &= \sum_{r=1}^R \hat{w}_r \rho(\hat{\mathbf{y}}_r) \psi_{k'}(\hat{\mathbf{y}}_r) \int_{\mathcal{D}} \phi_{j'}(\mathbf{x}) f(\hat{\mathbf{y}}_r) d\mathbf{x} \\
 & \quad \text{for } j' \in \{1, \dots, J\} \text{ and } k' \in \{1, \dots, K\}
 \end{aligned}$$

for some choice of **quadrature weights** $\{\hat{w}_r\}_{r=1}^R$ and **quadrature points** $\{\hat{\mathbf{y}}_r\}_{r=1}^R$ in Γ

[†]Integrals with respect to the spatial domain \mathcal{D} must also be approximated using quadrature rules; we do not need to consider this issue since we assume that all methods discussed treat all aspects of the spatial discretization in the same manner

- this quadrature rule need not be the same as the quadrature rule $\{w_q, \vec{y}_q\}_{r=1}^Q$ used to obtain the approximation of a quantity of interest
- In general, the discrete problem is a **fully coupled** (in physical and parameter spaces) **$JK \times JK$ system**
 - there are JK equations and JK degrees of freedom c_{jk}^\dagger
- On the other hand, **one can solve for the approximate dependence of the solution** $u_{JK}(\mathbf{x}, \vec{y})$ on both the spatial coordinates \mathbf{x} and the random parameters \vec{y} **by solving a single deterministic problem** of size JK
 - in particular
 - one does not have to explicitly sample the random parameters \vec{y}
 - one does not have to determine multiple solutions of the SPDE

[†]Economies are possible for **linear** SPDEs; more on this later

- Note that, once the c_{jk} 's are determined, one has obtained the **explicit formula**

$$u_{JK}(\mathbf{x}; \vec{y}) = \sum_{j=1}^J \sum_{k=1}^K c_{jk} \phi_j(\mathbf{x}) \psi_k(\vec{y})$$

for the approximate solution of the SPDE that can be evaluated at any point $\mathbf{x} \in \mathcal{D}$ in the spatial domain and for any value $\vec{y} \in \Gamma$ of the random parameters

- in particular, one can determine, by straightforward evaluation, $u_{JK}(\mathbf{x}, \vec{y}_q)$ at any quadrature point \vec{y}_q appearing in a quadrature rule approximation of a quantity of interest

- Thus, we obtain the stochastic Galerkin approximation to the quantity of interest

$$\begin{aligned} \int_{\Gamma} G(u(\mathbf{x}; \vec{y})) \rho(\vec{y}) d\vec{y} &\approx \sum_{q=1}^Q w_q \rho(\vec{y}_q) G(u(\mathbf{x}; \vec{y}_q)) \\ &\approx \sum_{q=1}^Q w_q \rho(\vec{y}_q) G(u_{JK}(\mathbf{x}; \vec{y}_q)) \\ &= \sum_{q=1}^Q w_q \rho(\vec{y}_q) G\left(\sum_{j=1}^J \sum_{k=1}^K c_{jk} \phi_j(\mathbf{x}) \psi_k(\vec{y}_q)\right) \end{aligned}$$

- To complete the description of the problem actually solved on a computer, one has to make specific choices[†]
 - for an approximating subspace $Z_K \subset Z$
 - for a basis $\{\psi_k(\vec{y})\}_{k=1}^K$ for Z_K
 - for a quadrature rule $\{\widehat{w}_r, \widehat{\vec{y}}_r\}_R$ used to approximate the parameter integrals in the discretized SPDE
 - for a quadrature rule $\{w_q, \vec{y}_q\}_{q=1}^Q$ used to approximate the parameter integrals in the discretized quantity of interest

- We arrange our discussion according to the first two choices
 - for each choice for the approximating space and the basis set, we will make choices for the two quadrature rules

[†]We assume that the approximating subspace $S_J \subset S$ and a basis $\{\phi_j(\mathbf{x})\}_{j=1}^J$ used for spatial discretization have been already chosen

- For parameter approximating spaces Z_K , one can use
 - locally-supported piecewise polynomial spaces
 - i.e., a finite element-type method
 - globally-supported polynomial spaces
 - i.e., a spectral-type method
- Following this plan will enable us to show that many (if not all) numerical methods for SPDEs can be derived from the stochastic Galerkin framework

**GLOBAL POLYNOMIAL APPROXIMATING SPACES –
POLYNOMIAL CHAOS AND
LAGRANGE INTERPOLATORY METHODS**

GLOBAL POLYNOMIAL APPROXIMATING SPACES FOR PARAMETER APPROXIMATION

- Let P_r denote the set of all polynomials of degree less than or equal to r
- Let $\{\Theta_i(y)\}_{i=0}^r$ denote a basis for P_r
 - of course, there are an infinite number of possible bases
 - the simplest is the monomial basis for which
$$\Theta_i(y) = y^i \text{ for } i = 0, 1, \dots, r$$
 - we will discuss several bases later
- Let $p = (p_1, p_2, \dots, p_N)$ be a multi-index, i.e.,
 - an N -vector whose components are non-negative integers

and let $|p| = \sum_{n=1}^N p_n$

- For each parameter y_n , we use polynomials of degree M and a basis $\{\Theta_{n,k}(y_n)\}_{k=1}^{K_n}$
 - for the sake of simplicity, we assume that $M_n = M$ for all n
 - there are good reasons for sometimes choosing different degree polynomials for each parameter
 - we will point out some instances for which this is the case

- For a given integer $M \geq 0$, let $\{\psi_k(\vec{y})\}_{k=1}^K$ denote the set of distinct multivariate polynomials such that

$$\left\{ \psi_k(\vec{y}) \right\}_{k=1}^K = \left\{ \prod_{n=1}^N \Theta_{n,i_n}(y_n) \right\}$$

where

$$\Theta_{n,i_n}(y_n) \in P_M \quad \text{and} \quad |p| \leq M$$

- the highest degree term in any of the multivariate polynomials is M

- thus, if $N = 2$ and $M = 2$, we have terms like

$$y_1^2 \text{ and } y_1 y_2 \text{ but not terms like } y_1^2 y_2$$

- the number of probabilistic degrees of freedom is given by

$$K = \frac{(N + M)!}{N! M!}$$

where N = number of random parameters

M = maximal degree of any of the

N -dimensional global polynomials used

– for example, if $N = 2$ and $M = 3$, we have

$$|p| = p_1 + p_2 \leq M = 3$$

and

$$K = \frac{(N + M)!}{N! M!} = \frac{(2 + 3)!}{2! 3!} = 10$$

and we have the set of 10 basis functions

$$\left\{ \psi_1(y_1, y_2), \dots, \psi_{10}(y_1, y_2) \right\} = \left\{ \begin{array}{l} \Theta_{1,0}(y_1) \Theta_{2,0}(y_2) \\ \Theta_{1,1}(y_1) \Theta_{2,0}(y_2) \\ \Theta_{1,0}(y_1) \Theta_{2,1}(y_2) \\ \Theta_{1,1}(y_1) \Theta_{2,1}(y_2) \\ \Theta_{1,2}(y_1) \Theta_{2,0}(y_2) \\ \Theta_{1,0}(y_1) \Theta_{2,2}(y_2) \\ \Theta_{1,2}(y_1) \Theta_{2,1}(y_2) \\ \Theta_{1,1}(y_1) \Theta_{2,2}(y_2) \\ \Theta_{1,3}(y_1) \Theta_{2,0}(y_2) \\ \Theta_{1,0}(y_1) \Theta_{2,3}(y_2) \end{array} \right\}$$

- Alternately, one could use the **tensor product basis**

$$\left\{ \psi_k(\vec{y}) \right\}_{k=1}^K = \left\{ \prod_{n=1}^N \Theta_{n,i_n}(y_n) \right\}$$

where

$$\Theta_{n,i_n}(y_n) \in P_M \quad \text{and} \quad p_n \leq M \quad \text{for all } n$$

- now the highest degree term in any of the polynomials is M in each y_n
 - thus, if $M = 2$, we have not only have terms like y_1^2 and $y_1 y_2$, but we also have terms like $y_1^2 y_2$ and $y_1^2 y_2^2$
- the number of probabilistic degrees of freedom is now given by

$$K = (M + 1)^N$$

where N = number of random parameters
 M = maximal degree in any variable y_n of any of the N -dimensional global polynomials used

– for example, if $N = 2$ and $M = 3$, we have

$$K = (M + 1)^N = (3 + 1)^2 = 16$$

$$\left\{ \psi_1(y_1, y_2), \dots, \psi_{16}(y_1, y_2) \right\} = \left\{ \begin{array}{l} \Theta_{1,0}(y_1) \Theta_{2,0}(y_2) \\ \Theta_{1,1}(y_1) \Theta_{2,0}(y_2) \\ \Theta_{1,0}(y_1) \Theta_{2,1}(y_2) \\ \Theta_{1,1}(y_1) \Theta_{2,1}(y_2) \\ \Theta_{1,2}(y_1) \Theta_{2,0}(y_2) \\ \Theta_{1,0}(y_1) \Theta_{2,2}(y_2) \\ \Theta_{1,2}(y_1) \Theta_{2,1}(y_2) \\ \Theta_{1,1}(y_1) \Theta_{2,2}(y_2) \\ \Theta_{1,3}(y_1) \Theta_{2,0}(y_2) \\ \Theta_{1,0}(y_1) \Theta_{2,3}(y_2) \\ \Theta_{1,1}(y_1) \Theta_{2,3}(y_2) \\ \Theta_{1,2}(y_1) \Theta_{2,3}(y_2) \\ \Theta_{1,3}(y_1) \Theta_{2,3}(y_2) \\ \Theta_{1,2}(y_1) \Theta_{2,2}(y_2) \\ \Theta_{1,3}(y_1) \Theta_{2,1}(y_2) \\ \Theta_{1,3}(y_1) \Theta_{2,2}(y_2) \end{array} \right\}$$

Global polynomial approximation in parameter space

$N =$ no. random parameters	$M =$ maximal degree of polynomials	$K =$ no. of probabilistic degrees of freedom	
		using complete polynomial basis	using tensor product basis
3	3	20	64
	5	56	216
5	3	56	1,024
	5	252	7,776
10	3	286	1,048,576
	5	3,003	60,046,176
20	3	1,771	$> 1 \times 10^{12}$
	5	53,130	$> 3 \times 10^{15}$
100	3	176,851	$> 1 \times 10^{60}$
	5	96,560,646	$> 6 \times 10^{77}$

- It seems that using tensor product bases is a **bad idea**

- Once a basis set $\{\psi_k(\vec{y})\}_{k=1}^K$ is chosen, we use the approximation

$$u_{J,K} = \sum_{j=1}^J \sum_{k=1}^K c_{j,k} \phi_j(\mathbf{x}) \psi_k(\vec{y})$$

- the probabilistic basis functions $\{\psi_k(\vec{y})\}_{k=1}^K$ are multivariate global polynomials
- The discrete system involves JK equations in JK unknowns, where
 - J = the number of finite element degrees of freedom used to discretize in physical space
 - K = the number of global polynomials used to discretize in parameter space

GLOBAL ORTHOGONAL POLYNOMIAL BASES

- For $n = 1, \dots, N$, let $\{H_{n,m_n}(y_n)\}_{m_n=0}^M$ denote the set of polynomials in \mathbb{R} of degree less than or equal to M that are **orthonormal** with respect to the function $\rho_n(y_n)$

– we have that

$$\int_{\mathcal{I}_n} H_{n,m_n}(y_n) H_{n,m'_n}(y_n) \rho_n(y_n) dy_n = \delta_{mm'} \quad \text{for } m_n, m'_n \in \{0, \dots, M\}$$

– note that the set $\{H_{n,m_n}(y_n)\}_{m_n=0}^M$ is hierarchical in the sense that

$$\text{degree}(H_{n,m_n}) = m_n$$

- Let

$$\Psi_k(\vec{y}) = \prod_{n=1}^N H_{n,m_n}(y_n) \quad \text{for all } m_n \in \{0, \dots, M\} \text{ such that } \sum_{n=1}^N m_n \leq M$$

- We then have that $k \in \left\{ 1, \dots, K_{PC} = \frac{(N + M)!}{N!M!} \right\}$

- For example, if $M = 1$ and $N = 3$ we have the $K_{PC} = 4$ basis functions[†]

$$\begin{aligned}
 & H_{1,0}(y_1)H_{2,0}(y_2)H_{3,0}(y_3) \\
 & H_{1,1}(y_1)H_{2,0}(y_2)H_{3,0}(y_3) \quad H_{1,0}(y_1)H_{2,1}(y_2)H_{3,0}(y_3) \quad H_{1,0}(y_1)H_{2,0}(y_2)H_{3,1}(y_3)
 \end{aligned}$$

while for if $M = 2$ and $N = 3$ we have the $K_{PC} = 10$ basis functions (suppressing noting the explicit dependences on the \vec{y}_n 's)

$$\begin{aligned}
 & H_{1,0}H_{2,0}H_{3,0} \\
 & H_{1,1}H_{2,0}H_{3,0} \quad H_{1,0}H_{2,1}H_{3,0} \quad H_{1,0}H_{2,0}H_{3,1} \\
 & H_{1,2}H_{2,0}H_{3,0} \quad H_{1,1}H_{2,1}H_{3,0} \quad H_{1,1}H_{2,0}H_{3,1} \quad H_{1,0}H_{2,2}H_{3,0} \quad H_{1,0}H_{2,1}H_{3,1} \quad H_{1,0}H_{2,0}H_{3,2}
 \end{aligned}$$

[†]It is convenient to write the N -dimensional polynomials so that each row contains the polynomials of the same total degree $\sum_{n=1}^N m_n$; thus the first row contains all possible products of the N one-dimensional polynomials of total degree 0, the second row has total degree 1, etc.

- We see that the functions $\Psi_k(\vec{y})$'s are **products of one-dimensional orthonormal polynomials** and have total degree less than or equal to M

– we then have that

$$\begin{aligned} \int_{\Gamma} \Psi_k(\vec{y}) \Psi_{k'}(\vec{y}) \rho(\vec{y}) d\vec{y} &= \int_{\Gamma} \Psi_k(\vec{y}) \Psi_{k'}(\vec{y}) \prod_{n=1}^N \rho_n(y_n) d\vec{y} \\ &= \prod_{n=1}^N \int_{\mathcal{I}_n} H_{n,m_n}(y_n) H_{n,m'_n}(y_n) \rho_n(y_n) dy_n = \delta_{kk'} \end{aligned}$$

– note that we need to write $\rho(\vec{y}) = \prod_{n=1}^N \rho_n(y_n)$, i.e., as a product as well, so that we know what $H_{n,m}(\cdot)$ is orthonormal with respect to

– thus, we are restricted to **independent random variables and to parameter domains Γ that are (possibly infinite) hypercubes**

- It is easy to see that the set $\{\Psi_k\}_{k=1}^{K_{PC}}$ of N -dimensional polynomials is a basis for the **complete polynomial space** of degree M , i.e.,

$$\text{span}\{\Psi_k\}_{k=1}^{K_{PC}} = \text{all polynomials of total degree } \leq M$$

- The **stochastic Galerkin-global orthogonal polynomial approximation** of the solution of the SPDE is then defined by setting

$$Z_{PC} = \text{span}\{\Psi_k\}_{k=1}^{K_{PC}}$$

so that

$$u_{PC}(\mathbf{x}, \vec{y}) = \sum_{j=1}^J \sum_{k=1}^{K_{PC}} c_{jk} \phi_j(\mathbf{x}) \Psi_k(\vec{y})$$

- This is better known under another name[†]

(stochastic Galerkin) polynomial chaos approximation (SG-PC)
= complete, global orthonormal polynomial approximation

[†]**Polynomial chaos approximations** usually refer to the case for which, for all n , $\rho_n(y_n)$ is a **Gaussian PDF** so that, for all n , $\{H_{n,m}(y_n)\}_{m=0}^M$ are sets of Hermite polynomials; for **other PDFs**, the SC-PC approximation is usually referred to as a **generalized polynomial chaos approximation**; here we do not differentiate between the two and refer to all cases as polynomial chaos approximations

- The implementation of the SG-PC method is simpler if one instead uses a tensor product polynomial space; however, as we have seen, such a choice leads to hugely more costly approximations[†]

[†]The tensor product basis is given by

$$\Psi_k(\vec{y}) = \prod_{n=1}^N H_{n,m_n}(y_n) \quad \text{for all } m_n \in \{0, \dots, M\} \text{ such that } m_n \leq M$$

in this case, $\text{span}\{\Psi_k\}_{k=1}^K$ is the **tensor product** space of polynomials such that the degree in any coordinate y_n is less than or equal to M ; if we do this, we end up with $K = (M + 1)^N$ basis functions; for example, if $M = 1$ and $N = 3$, we have the 8 polynomials (the 4 we had before plus 4 additional ones)

$$\begin{array}{ccc} & H_{1,0}H_{2,0}H_{3,0} & \\ H_{1,1}H_{2,0}H_{3,0} & H_{1,0}H_{2,1}H_{3,0} & H_{1,0}H_{2,0}H_{3,1} \\ H_{1,1}H_{2,1}H_{3,0} & H_{1,1}H_{2,0}H_{3,1} & H_{1,0}H_{2,1}H_{3,1} \\ & H_{1,1}H_{2,1}H_{3,1} & \end{array}$$

for $N > 1$ and $M > 0$ we have that $(M + 1)^N > \frac{(N+M)!}{N!M!}$; for a moderate number of parameters or for a moderately high degree polynomial, we in fact have that $(M + 1)^N \gg \frac{(N+M)!}{N!M!}$; for example,

$$\text{if } M = 6 \text{ and } N = 3 \implies (N + M)!/(N!M!) = 84 \text{ and } (M + 1)^N = 343$$

$$\text{if } M = 4 \text{ and } N = 5 \implies (N + M)!/(N!M!) = 126 \text{ and } (M + 1)^N = 3125$$

$$\text{if } M = 2 \text{ and } N = 7 \implies (N + M)!/(N!M!) = 36 \text{ and } (M + 1)^N = 2187$$

the disparity gets worse as, say, N increases; for example,

$$\text{if } M = 2 \text{ and } N = 10 \implies (N + M)!/(N!M!) = 66 \text{ and } (M + 1)^N = 59059$$

on the other hand, since the accuracy, i.e., the rate of convergence of global polynomial approximation, is determined by the degree of the largest complete polynomial space contained in the approximate space, for the same M , the accuracy obtained using a tensor product space is the same as that obtained using a complete polynomial space; as a result, by using the latter one can obtain the same accuracy with substantially fewer degrees of freedom

SG-PC approximations of quantities of interest

- The SG-PC approximation of a quantity of interest is then defined by

$$\int_{\Gamma} G(u(\mathbf{x}; \vec{y})) \rho(\vec{y}) d\vec{y} \approx \sum_{q=1}^Q w_q \rho(\vec{y}_q) G(u_{PC}(\mathbf{x}; \vec{y}_q))$$

where

- $u_{PC}(\mathbf{x}; \vec{y}_q)$, $q = 1, \dots, Q$, is obtained by evaluation of the SG-PC approximation of the stochastic SPDE at the quadrature points
 - i.e., we have that

$$u_{PC}(\mathbf{x}, \vec{y}_q) = \sum_{j=1}^J \sum_{k=1}^{K_{PC}} c_{jk} \phi_j(\mathbf{x}) \Psi_k(\vec{y}_q) \quad \text{for } q = 1, \dots, Q$$

- Thus, the SG-PC approximation of a quantity of interest can be determined by
 1. first solving a single $JK_{PC} \times JK_{PC}$ system of equations to determine the SG-PC approximation of the solution of the SPDE;
 2. then evaluating the SG-PC approximation at the Q quadrature points;
 3. substituting the results of Step 2 into the quadrature rule approximation of the quantity of interest
- The cost of obtaining an SG-PC approximation of a quantity of interest is dominated by the first step

GLOBAL LAGRANGE INTERPOLATORY BASES

- Instead of using global orthogonal polynomials to define a stochastic Galerkin method, one can use **interpolatory** polynomials
- Given a set of points $\{\tilde{\vec{y}}_k\}_{k=1}^{K_{LI}}$ in Γ
 - for $k \in \{1, \dots, K_{LI}\}$, let $L_k(\vec{y})$ denote the set of **Lagrange interpolating polynomials** for these points
 - we have that

$$L_k(\vec{y}_{k'}) = \delta_{kk'} \quad \text{for all } k, k' \in \{1, \dots, K_{LI}\}$$

- Set $\psi_k(\vec{y}) = L_k(\vec{y})$ for $k \in \{1, \dots, K_{LI}\}$ so that

$$Z_{K_{LI}} = \text{span}\{L_k\}_{k=1}^{K_{LI}}$$

- Then, the **stochastic Galerkin-Lagrange interpolant (SG-LI)** approximation of the solution of the SPDE takes the form

$$u_{LI}(\mathbf{x}, \vec{y}) = \sum_{j=1}^J \sum_{k=1}^{K_{LI}} c_{jk} \phi_j(\mathbf{x}) L_k(\vec{y})$$

- In general, the SG-LI approximation to the solution of an SPDE can be obtained by solving a **single** $JK_{LI} \times JK_{LI}$ system
 - this would also be the dominant cost encountered in obtaining an SG-LI approximation of a quantity of interest

- If we choose a point set $\{\tilde{\vec{y}}_k\}_{k=1}^{K_{LI}}$ that can be used to define a complete interpolating polynomial of degree less than or equal M , we have that

$$Z_{K_{LI}} = Z_{K_{PC}} \quad \text{and} \quad K_{LI} = K_{PC} = \frac{(N + M)!}{N!M!}$$

- In this case, it is clear that

**the polynomial chaos approximation $u_{PC}(\mathbf{x}; \vec{y})$
= global Lagrange interpolant approximation
 $u_{LI}(\mathbf{x}; \vec{y})$ based on a complete polynomial space**

- the only differences between the two approximations result from the choices of bases

- Unfortunately, even for a moderate number of parameters, it may not be easy to define a “good” set of interpolation points that can be used to determine a complete Lagrange interpolant
 - it is easy to define a set of interpolation points that can be used to define a **tensor product** Lagrange interpolant[†]
 - however, as we have seen, this leads to a very inefficient approximation compared to complete polynomial approximation
- There exists intermediate choices, e.g., **Smolyak point sets**, that can be systematically defined in any dimension
 - for the Smolyak point sets, $K_{LI} > \frac{(M + N)!}{N!M!}$ so that they require more points compared to complete polynomial interpolation
 - however, we have that $K_{LI} \ll (M + 1)^N$ so that it requires much fewer points compared to tensor product interpolation

[†]Unlike the case for orthogonal polynomials, for Lagrange polynomials it is not easy to define a complete polynomial basis from the tensor product basis; for the Lagrange case, the tensor product basis is not hierarchical since all Lagrange polynomials are of the same degree

- We therefore conclude that

in general, for the same accuracy, a stochastic Galerkin-Lagrange polynomial approximation is (a little) more costly to obtain than is a stochastic Galerkin-polynomial chaos approximation

- However, as we shall now see, a **judicious** choice for the interpolation points can lead to great efficiency improvements in stochastic Galerkin-Lagrange interpolation methods
 - we defer discussion of how one obtains the LI-approximation of a quantity of interest until after we consider this special case of the SG-LI method
 - we also defer further discussion of Smolyak point sets until later

STOCHASTIC COLLOCATION METHODS

- For the SG-LI method, the discretized SPDE looks like

$$\begin{aligned} & \sum_{r=1}^R \widehat{w}_r \rho(\widehat{\mathbf{y}}_r) \mathbf{L}_{k'}(\widehat{\mathbf{y}}_r) \int_{\mathcal{D}} S\left(\sum_{j=1}^J \sum_{k=1}^K c_{jk} \phi_j(\mathbf{x}) \mathbf{L}_k(\widehat{\mathbf{y}}_r), \widehat{\mathbf{y}}_r\right) T\left(\phi_{j'}(\mathbf{x})\right) d\mathbf{x} \\ &= \sum_{r=1}^R \widehat{w}_r \rho(\widehat{\mathbf{y}}_r) \mathbf{L}_{k'}(\widehat{\mathbf{y}}_r) \int_{\mathcal{D}} \phi_{j'}(\mathbf{x}) f(\widehat{\mathbf{y}}_r) d\mathbf{x} \\ & \quad \text{for } j' \in \{1, \dots, J\} \text{ and } k' \in \{1, \dots, K\} \end{aligned}$$

- Suppose we choose

the interpolating points $\{\widetilde{\mathbf{y}}_k\}_{k=1}^{K_{LI}}$ for the SG-LI method

to be the same as

the quadrature points $\{\widehat{\mathbf{y}}_r\}_{r=1}^R$ used in the discretized SPDE

- We then have that

$$\mathbf{L}_k(\widehat{\mathbf{y}}_r) = \delta_{kr} \quad \forall r, k \in \{1, \dots, R = K_{LI}\}$$

- As a result, the discretized SPDE reduces to

$$\int_{\mathcal{D}} S\left(\sum_{j=1}^J c_{jr} \phi_j(\mathbf{x}), \widehat{\mathbf{y}}_r\right) T\left(\phi_{j'}(\mathbf{x})\right) d\mathbf{x} = \int_{\mathcal{D}} \phi_{j'}(\mathbf{x}) f(\widehat{\mathbf{y}}_r) d\mathbf{x}$$

for $j' \in \{1, \dots, J\}$, $r \in \{1, \dots, R = K_{LI}\}$

- Thus, we have **total uncoupling in parameter space**

- for each $r \in \{1, \dots, R\}$, we can solve the separate standard, deterministic finite element problem for $\{c_{jr}\}_{j=1}^J$

for $r \in \{1, \dots, R\}$, determine $u_r(\mathbf{x}) = \sum_{j=1}^J c_{jr} \phi_j(\mathbf{x})$ satisfying

$$\int_{\mathcal{D}} S\left(u_r(\mathbf{x}), \widehat{\mathbf{y}}_r\right) T\left(\phi_{j'}(\mathbf{x})\right) d\mathbf{x} = \int_{\mathcal{D}} \phi_{j'}(\mathbf{x}) f(\widehat{\mathbf{y}}_r) d\mathbf{x}$$

for $j' \in \{1, \dots, J\}$

- Such a method is referred to as a **stochastic collocation (SC) method** so that

stochastic collocation methods are stochastic Galerkin-Lagrange interpolation methods for which the interpolation points are the same as the quadrature points of the quadrature rule used to discretize the SPDE

- It is important to note that **for stochastic collocation methods,**
the uncoupling of the spatial and probabilistic degrees of freedom
occurs for

general nonlinear PDEs

general joint probability distributions

and

general random field data

- If desired, the stochastic collocation approximation to the solution $u(\mathbf{x}, \vec{y})$ of the SPDE is then given by

$$u_{SC}(\mathbf{x}, \vec{y}) = \sum_{r=1}^R u_r(\mathbf{x}) L_r(\vec{y}) = \sum_{j=1}^J \sum_{r=1}^R c_{jr} \phi_j(\mathbf{x}) L_r(\vec{y})$$

- however, as we will now see, one does not need to form this expression to determine an approximation of a quantity of interest
- this is unlike the case for general stochastic Galerkin methods, including polynomial chaos methods, for which one must evaluate the approximation of the solution of the SPDE at the quadrature points of the approximation of a quantity of interest

SC-approximations of quantities of interest

- It is also convenient to use the same quadrature rule

- to approximate a quantity of interest

as was used to

- approximate the integrals in the discretized SPDE

and that was also used as

- the Lagrange interpolations points,

i.e., we choose

$$K_{LI} = R = Q$$

$$\{\tilde{y}_k\}_{k=1}^{K_{LI}} = \{\hat{y}_r\}_{r=1}^R = \{\vec{y}_q\}_{q=1}^Q \quad \text{and} \quad \{\hat{w}_r\}_{r=1}^R = \{w_q\}_{q=1}^Q$$

- We then have that

$$L_r(\vec{y}_q) = \delta_{rq} \quad \text{for all } r, q \in \{1, \dots, K_{LI} = R = Q\}$$

- Using this in the expression for the approximation of a quantity of interest results in

$$\begin{aligned} \int_{\Gamma} G(u(\mathbf{x}; \vec{y})) \rho(\vec{y}) d\vec{y} &\approx \sum_{q=1}^Q w_q \rho(\vec{y}_q) G(u_{SC}(\mathbf{x})) \\ &= \sum_{q=1}^Q w_q \rho(\vec{y}_q) G\left(\sum_{r=1}^R u_r(\mathbf{x}) L_r(\vec{y}_q)\right) = \sum_{q=1}^Q w_q \rho(\vec{y}_q) G(u_q(\mathbf{x})) \end{aligned}$$

i.e.,

$$\int_{\Gamma} G(u(\mathbf{x}; \vec{y})) \rho(\vec{y}) d\vec{y} \approx \sum_{q=1}^Q w_q \rho(\vec{y}_q) G(u_q(\mathbf{x}))$$

where, for $q \in \{1, \dots, Q = R = K_{LI}\}$, $u_q(\mathbf{x}) = \sum_{j=1}^J c_{jq} \phi_j(\mathbf{x})$

is determined from

$$\int_{\mathcal{D}} S(u_q(\mathbf{x}), \vec{y}_q) T(\phi_{j'}(\mathbf{x})) d\mathbf{x} = \int_{\mathcal{D}} \phi_{j'}(\mathbf{x}) f(\vec{y}_q) d\mathbf{x} \quad \text{for } j' \in \{1, \dots, J\}$$

- Note that
 - we do not have to explicitly determine the Lagrange interpolating polynomials $\{L_k(\vec{y})\}_{k=1}^{K_{LI}}$ to determine the approximation of a quantity of interest
 - nor do we have to form and evaluate, at quadrature points, the SC-approximation[†]
- Thus, we see that the SC-approximation of a quantity of interest can be determined by
 1. first solving $Q = K_{LI}$ systems of equations of size J to determine $u_q(\mathbf{x})$ for $q = 1, \dots, Q = K_{LI}$;
 2. then substituting the results of Step 1 into the approximation of the quantity of interest

[†]In contrast, for PC approximations of quantities of interest one must explicitly evaluate the PC approximation at quadrature points

- The cost of obtaining the SC-approximation of a quantity of interest is dominated by the first step which requires the solution of K_{LI} systems of size J
 - recall that the cost of obtaining the PC-approximation of a quantity of interest is dominated by the cost of solving a single deterministic system of size JK_{PC}
 - for general, nonlinear problems, the SC-approximation can be obtained at much less cost[†]

[†]In the best-case scenario for which the PC-system of size JK_{PC} and each of the $Q = R = K_{LI}$ SC-systems of size J can be solved in linear time, the solution cost associated with the PC-approximation of a quantity of interest will be of $O(JK_{PC})$ while the corresponding solution cost for the SC-approximation of a quantity of interest is of $O(JK_{LI})$; for the same accuracy, in practice $K_{LI} > K_{PC}$ so that in this best-case scenario, the SC-approximation of a quantity of interest is more costly to obtain than is the PC-approximation; for more general problems for which solution costs are not linear in the number of degrees of freedom, the PC-approximation is more costly to obtain than is the SC-approximation since, for some $\alpha > 1$, one must compare the cost of $O(JK_{PC})^\alpha$ for the PC case to the cost of $O(J^\alpha K_{LI})$ for the SC case, keeping in mind that although $K_{LI} > K_{PC}$, using Smolyak points as collocation points we have that $K_{LI} \approx K_{PC}$

NON-INTRUSIVE POLYNOMIAL CHAOS METHODS

- Can the uncoupling of parameter and spatial degrees of freedom be effected in a polynomial chaos setting?
- The PC approximation is given by

$$u_{PC}(\mathbf{x}, \vec{y}) = \sum_{j=1}^J \sum_{k=1}^{K_{PC}} c_{jk} \phi_j(\mathbf{x}) \Psi_k(\vec{y}) = \sum_{k=1}^{K_{PC}} \tilde{u}_k(\mathbf{x}) \Psi_k(\vec{y})$$

where for $k \in \{1, \dots, K_{PC}\}$,

$$\tilde{u}_k(\mathbf{x}) = \sum_{j=1}^J c_{jk} \phi_j(\mathbf{x})$$

and $\{\Psi_k(\vec{y})\}_{k=1}^{K_{PC}}$ is a set of orthonormal polynomials with respect to weight $\rho(\vec{y}) = \prod_{n=1}^N \rho_n(y_n)$

- As a result, we have that, for $k' \in \{1, \dots, K_{PC}\}$,

$$\int_{\Gamma} u_{PC}(\mathbf{x}, \vec{y}) \Psi_{k'}(\vec{y}) \rho(\vec{y}) d\vec{y} = \sum_{k=1}^{K_{PC}} u_k(\mathbf{x}) \int_{\Gamma} \Psi_k(\vec{y}) \Psi_{k'}(\vec{y}) \rho(\vec{y}) d\vec{y} = \tilde{u}_{k'}(\mathbf{x})$$

- We view this as a formula for $\tilde{u}_{k'}(\mathbf{x})$, i.e.,

$$\tilde{u}_{k'}(\mathbf{x}) = \sum_{j=1}^J c_{jk'} \phi_j(\mathbf{x}) = \int_{\Gamma} u_{PC}(\mathbf{x}, \vec{y}) \Psi_{k'}(\vec{y}) \rho(\vec{y}) d\vec{y}$$

- We use a quadrature rule[†] $\{\hat{w}_r, \hat{\vec{y}}_r\}_{r=1}^R$ to approximate the integral to obtain

$$\tilde{u}_{k'}(\mathbf{x}) \approx \sum_{r=1}^R \hat{w}_r u_{PC}(\mathbf{x}, \hat{\vec{y}}_r) \Psi_{k'}(\hat{\vec{y}}_r) \rho(\hat{\vec{y}}_r) \quad \text{for } k' \in \{1, \dots, K_{PC}\}$$

- For $r \in \{1, \dots, R\}$, we replace $u_{PC}(\mathbf{x}, \hat{\vec{y}}_r)$ by the solution $u_r(\mathbf{x})$ of[‡]

$$\int_{\mathcal{D}} S(u_r(\mathbf{x}), \hat{\vec{y}}_r) T(\phi_{j'}(\mathbf{x})) d\mathbf{x} = \int_{\mathcal{D}} \phi_{j'}(\mathbf{x}) f(\hat{\vec{y}}_r) d\mathbf{x} \quad \text{for } j' \in \{1, \dots, J\}$$

[†]This quadrature rule may be the same or may be different from the quadrature rule used to approximate a quantity of interest

[‡]Note that this is exactly the same set of R equations that is solved for in the stochastic collocation case

- We thus obtain

$$\tilde{u}_{k'}(\mathbf{x}) \approx \sum_{r=1}^R \hat{w}_r u_r(\mathbf{x}) \Psi_{k'}(\hat{\vec{y}}_r) \rho(\hat{\vec{y}}_r)$$

- We use this approximation to define the[†] **non-intrusive polynomial chaos (NIPC) approximation** to the solution $u(\mathbf{x}, \vec{y})$ of the SPDE:[‡]

$$\begin{aligned} u(\mathbf{x}, \vec{y}) &\approx u_{PC}(\mathbf{x}, \vec{y}) = \sum_{k=1}^{K_{PC}} \tilde{u}_k(\mathbf{x}) \Psi_k(\vec{y}) \\ &\approx u_{NIPC}(\mathbf{x}, \vec{y}) = \sum_{k=1}^{K_{PC}} \sum_{r=1}^R \hat{w}_r u_r(\mathbf{x}) \Psi_k(\hat{\vec{y}}_r) \rho(\hat{\vec{y}}_r) \Psi_k(\vec{y}) \end{aligned}$$

[†]Nowadays, the polynomial chaos method previously discussed is often referred as the **intrusive** polynomial chaos method to differentiate it from the non-intrusive polynomial chaos method defined here

[‡]In comparison, the stochastic collocation approximation takes the simpler form

$$u_{SC}(\mathbf{x}, \vec{y}) = \sum_{r=1}^R u_r(\mathbf{x}) L_r(\vec{y})$$

due to the fact that $L_k(\hat{\vec{y}}_r) = \delta_{kr}$ in the SC case while $\Psi_k(\hat{\vec{y}}_r) \neq 0$ for all k and r in the NIPC case

- Thus, the NIPC approximation can be obtained by solving

R deterministic problems of size J to obtain $u_r(\mathbf{x})$ for $r = 1, \dots, R$ instead of the

one deterministic problem of size JK_{PC}

that is solved in the intrusive polynomial chaos method

- All K_{PC} “coefficients” $\sum_{r=1}^R \hat{w}_r u_r(\mathbf{x}) \Psi_k(\hat{\vec{y}}_r) \rho(\hat{\vec{y}}_r)$, $k \in \{1, \dots, K_{PC}\}$, in the NIPC expansion

$$\begin{aligned} u_{NIPC}(\mathbf{x}, \vec{y}) &= \sum_{k=1}^{K_{PC}} \sum_{r=1}^R \hat{w}_r u_r(\mathbf{x}) \Psi_k(\hat{\vec{y}}_r) \rho(\hat{\vec{y}}_r) \Psi_k(\vec{y}) \\ &= \sum_{r=1}^R \hat{w}_r \rho(\hat{\vec{y}}_r) u_r(\mathbf{x}) \sum_{k=1}^{K_{PC}} \Psi_k(\hat{\vec{y}}_r) \Psi_k(\vec{y}) \end{aligned}$$

can be obtained from the same R solutions $u_r(\mathbf{x})$, $r \in \{1, \dots, R\}$, of the SPDE

- The cost of obtaining the NIPC-approximation is dominated by the need to solve[†] R systems of size J
- For non-intrusive-polynomial chaos approximations,
the uncoupling of the spatial and probabilistic degrees of freedom occurs for
general nonlinear PDEs
but only for
independent random variables[‡]
and
Gaussian random field data[‡]

[†]This is just the same as for the stochastic collocation approximation

[‡]This is unlike the case for stochastic collocation methods for which similar uncouplings are possible for general joint probability distributions and general random fields

- Thus, it is clear that

non-intrusive polynomial chaos approximations are stochastic Galerkin-global orthogonal polynomial approximations obtained by approximating the coefficients of the orthogonal polynomials via a quadrature rule

- It is also clear that, for the same accuracy

the costs of obtaining stochastic collocation and non-intrusive polynomial chaos approximations are comparable and, in general, both are much lower than the cost of obtaining the intrusive polynomial chaos approximation

NIPC-approximations of quantities of interest

- Unlike the stochastic collocation case, there is no great advantage to using the same quadrature rule for approximating a quantity of interest as is used to construct the non-intrusive polynomial chaos approximation
 - on the other hand, there is no reason not to do so
 - so, we choose

$$Q = R, \quad \{w_q\}_{q=1}^Q = \{\widehat{w}_r\}_{r=1}^R, \quad \text{and} \quad \{\vec{y}_q\}_{q=1}^Q = \{\widehat{\vec{y}}_r\}_{r=1}^R$$

- Then, the NIPC approximation of a quantity of interest has the form[†]

$$\begin{aligned} \int_{\Gamma} G(u(\mathbf{x}; \vec{y})) \rho(\vec{y}) d\vec{y} &\approx \sum_{q=1}^Q w_q \rho(\vec{y}_q) G(u_{NIPC}(\mathbf{x})) \\ &= \sum_{q=1}^Q w_q \rho(\vec{y}_q) G\left(\sum_{k=1}^{K_{PC}} \left(\sum_{q=1}^Q w_q u_q(\mathbf{x}) \Psi_k(\vec{y}_q) \rho(\vec{y}_q)\right) \Psi_k(\vec{y}_q)\right) \end{aligned}$$

where, for $q \in \{1, \dots, Q = R\}$, $u_q(\mathbf{x}) = \sum_{j=1}^J c_{jq} \phi_j(\mathbf{x})$ is determined from

$$\int_{\mathcal{D}} S(u_q(\mathbf{x}), \vec{y}_q) T(\phi_{j'}(\mathbf{x})) d\mathbf{x} = \int_{\mathcal{D}} \phi_{j'}(\mathbf{x}) f(\vec{y}_q) d\mathbf{x} \quad \text{for } j' \in \{1, \dots, J\}$$

[†]In comparison, the stochastic collocation approximation of the quantity of interest takes the simpler form

$$\int_{\Gamma} G(u(\mathbf{x}; \vec{y})) \rho(\vec{y}) d\vec{y} \approx \sum_{q=1}^Q w_q \rho(\vec{y}_q) G(u_q(\mathbf{x}))$$

again due to the fact that $L_k(\hat{\vec{y}}_r) = \delta_{kr}$ in the SC case while $\Psi_k(\hat{\vec{y}}_r) \neq 0$ for all k and r in the NIPC case

- Thus, we see that the NIPC approximation of a quantity of interest can be determined by
 1. first solving Q systems of equations of size J to determine $u_q(\mathbf{x})$ for $q = 1, \dots, Q$;
 2. then substituting the results of Step 1 into the NIPC-approximation of the quantity of interest
- Note that one is not restricted to use of any particular quadrature rule, either to determine the NIPC approximation of the solution of the SPDE or the NIPC approximation to a quantity of interest
 - in particular, one does not have to use interpolatory quadrature rules
 - one can use, e.g., any of the rules to be discussed in connection with stochastic sampling methods

- Note also that to obtain this approximation, one has to explicitly construct and evaluate, at the quadrature points \vec{y}_q , the non-intrusive polynomial chaos approximation
 - this includes having to explicitly evaluate the orthogonal polynomial basis functions $\Psi_k(\cdot)$ at the quadrature points
 - this should be contrasted with the SC approximation of a quantity of interest that does not need the explicit construction or evaluation of the SC approximation nor of the the Lagrange interpolatory polynomial basis functions $L_k(\cdot)$
 - again, these differences between the two methods are due to the facts that $L_k(\vec{y}_q) = \delta_{kq}$ while $\Psi_k(\vec{y}_q) \neq 0$ for all k and q