Error analysis of the Wiener-Askey polynomial chaos with hyperbolic cross approximation and its application to differential equations with random input

Xue Luo

School of Mathematics and Systems Science, Beihang University, Beijing, P. R. China, 100191 (`xluo@buaa.edu.cn`*)*

Abstract

s It is well-known that sparse grid algorithm has been widely accepted as an efficient tool to overcome the "curse of dimensionality" in some degree. In this note, we give the error estimate of hyperbolic cross (HC) approximations with all sorts of Askey polynomials. These polynomials are useful in generalized polynomial chaos (gPC) in the field of uncertainty quantification. The exponential convergences in both regular and optimized HC approximations have been shown under the condition that the random variable depends on the random inputs smoothly in some degree. Moreover, we apply gPC to numerically solve the ordinary differential equations with slightly higher dimensional random inputs. Both regular and optimized HC have been investigated with Laguerre-chaos, Charlier-chaos and Hermite-chaos in the numerical experiment. The discussion of the connection between the standard ANOVA approximation and Galerkin approximation is in the appendix.

Keywords: generalized polynomial chaos, hyperbolic cross approximation, differential equations with random inputs, spectral method
2000 MSC: 65M70

1. Introduction

Uncertainty is ubiquitous. It is usually related to the lack of knowledge about the processes involved. Although this kind of uncertainty can be reduced by obtaining more observations or by improving the accuracy of the measurements, it is quite impractical to measure at all the points, or even at a relatively large number of points. Mathematically, one usually models the uncertainty by random variables or processes, with a realistic probability distribution. The main goal in the field of uncertainty quantification is to predict the quantities of physical interest by mathematical and computational analysis. Usually, the quantity of physical interests are the real-valued functionals of the solution to certain partial/ordinary differential equations with random inputs. Generally speaking, the random inputs in the system can be expanded by an infinite combinations of random variables, say the Karhunen-Loeve expansion [10, 11] or generalized polynomial chaos (gPC). In particular, the gPC is one of the most popular approximation in the literature. The name *polynomial chaos* (PC) is coined by N. Wiener [21] in 1938, in which he studied the decomposition of Gaussian stochastic processes. The convergence of the Hermite-chaos expansion of arbitrary random processes with finite second-order moments has been shown rigorously by Cameron and Martin [5]. The study of the original PC was started by Ghanem and his coworkers. He represented the random processes by the Hermite polynomials and used this technique with finite element method to many different practical problems, see [7]. Although the Hermite-chaos is mathematically sound, the convergence rate of non-Gaussian problems are far from optimal. It is Xiu and Karniadakis [24] who for the first time generalized the Hermite polynomials to the Wiener-Askey polynomials, and numerical experiments verified the optimal convergence by choosing proper polynomial basis according to the distribution. This is so-called gPC in the literature. Later, the gPC has been further generalized to other set of complete basis, for instance the piecewise polynomial basis [2], the wavelet basis [14], and multi-element gPC [20].

After choosing an appropriate set of polynomials basis, the partial/ordinary differential equations with random inputs yields a set of coupled deterministic equations with the stochastic Galerkin (SG) procedure. Most of the early works are based on this method, which minimizes the error of Galerkin projection onto the linear subspace spanned by a finite-order gPC, see [7, 25, 2, 14] and references therein. Another alternative numerical approach is the stochastic collocation (SC) method, which originates from the idea of deterministic sampling. Usually the nodes of the quadrature rule are selected to be a set of realizations of the random variables. An ensemble of repetitive deterministic codes with the realizations has been executed, and a synchronization has been processed to get the desired quantity of interest from the deterministic solution ensemble.

However, when the dimension of random inputs is high, no matter the SG or the SC method will inevitably encounter the so-called "curse of dimensionality". As in the case of SG method, if the linear subspace is spanned by tensor product of the polynomials basis, and assume that the first N-order polynomials are used in each direction, then the total number of polynomial basis is $M = N^d$, where d is the dimension of the random inputs. Let X_N be the subspace spanned by the tensor product of polynomials basis. A typical error estimate is of the form

$$\inf_{u_N \in X_N} ||u - u_N||_{L^2} \lesssim N^{-r} ||u||_{H^r} \lesssim M^{-\frac{r}{d}} ||u||_{H^r},$$

where H^r is the Sobolev space, and the notation $\lesssim$ represents $\leq$ up to a positive generic constant independent of N. It is clear to see that the error of Galerkin projection deteriorates exponentially with respect to the dimension d. As in the case of SC method, the total number of the nodes grows exponentially with respect to the dimension d. Indeed, if N represents the number of the nodes in each direction, then the total number from the tensor product is $M = N^d$. It indicates that the deterministic simulations should be executed repetitively for M times. It is almost impractical for problems with 5 or even higher dimensional random inputs.

One alleviation of the "curse of dimensionality" is the so-called sparse grid, which can be dated back to Smolyak [18], and has been further investigated by many researchers, see [6, 3, 23], among which [23] proposed a high-order SC approach. Much work after that has been focused on further reduction of the nodes, see [13, 15] and references therein. Meanwhile, the sparse grid applied to the SG method is to reduce the total number of polynomial basis spanning the linear subspace. Approximations by hyperbolic cross (HC) have recently been received much attention, see [27, 4] and references therein which can further reduce the total number of polynomial basis. To the best of our knowledge, the error analysis of the HC approximations based on polynomials is first investigated in [17] to the Jacobi polynomials in spectral method. More recently, Yau and the author [12] showed the error analysis of HC approximations based on the generalized Hermite functions and studied its application of solving deterministic parabolic PDEs.

The main goal of our paper is to investigate the error analysis of the HC approximation with the orthogonal polynomials of Askey scheme. For any second order random variable $u(\theta)$, it can be approximated by

$$u(\theta) \approx \sum_{\boldsymbol{i} \in \boldsymbol{\Omega}_N} \hat{u}_{\boldsymbol{i}} \Phi_{\boldsymbol{i}}(\boldsymbol{\xi}(\theta)) =: u_N(\theta),$$

where $\boldsymbol{\Omega}_N$ is an index set with $\mathrm{card}(\boldsymbol{\Omega}_N) < \infty$, $\boldsymbol{\xi}(\theta) \in \mathbb{R}^d / \boldsymbol{N}_0^d / \cdots$ is a $d-$dimensional random variables, and $\Phi_{\boldsymbol{i}}$ are the orthogonal polynomials of Askey scheme. In this paper, we shall derive the typical error estimate of the form:

$$\inf_{u_N \in X_N} ||u - u_N||_{\mathcal{K}^l} = ||u - P_N u||_{\mathcal{K}^l} \lesssim N^{c(l,m)} |u|_{\mathcal{K}^m},$$

for $0 \leq l < m$, where $\mathcal{K}^l$ is the Koborov space, $X_N := \mathrm{span}\{\Phi_{\boldsymbol{i}} : \boldsymbol{i} \in \boldsymbol{\Omega}_N\}$, P_N is the projection operator onto the linear subspace X_N, and $c(l,m)$ is a negative constant.

The paper is organized as following. The notations and the orthogonal polynomials of Askey scheme have been introduced in section 2. For the readers' convenience, we include the frequently used orthogonal polynomials of Askey scheme and their properties in appendix B. Section 3 is devoted to the error analysis of the projection with HC approximations using Laguerre-chaos and Charlier-chaos, as the representatives. The results of the error estimates with other orthogonal polynomials of Askey scheme have been stated without proofs in section 4. The applications of the gPC to Galerkin method of ordinary differential equations with random inputs are numerically investigated in section 5. The conclusion is in section 6. Appendix A is devoted to discuss the connection between the standard ANOVA approximation and Galerkin approximation. It shows in theory that HC can be naturally combined with ANOVA approaches.

2. Preliminaries

2.1. Notations

Let us first clarify the notations to be used throughout the paper.

- ⋄ Let $\mathbb{R}$(resp., $\mathbb{N}$) denotes all the real numbers (resp., natural numbers), $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, and $\mathbb{N}_N = \{0, 1, \cdots, N\}$.
- ⋄ For any $d \in \mathbb{N}$, we use boldface lowercase letters to denote d-dimensional multi-indices and vectors. For example, $\boldsymbol{k} = (k_1, k_2, \ldots, k_d) \in \mathbb{N}_0^d$.
- ⋄ Denote $\mathbf{1} = (1, 1, \ldots, 1) \in \mathbb{N}^d$, and let $\boldsymbol{e}_i = (0, \ldots, 1, \ldots, 0)$ be the i^{th} unit vector in $\mathbb{R}^d$. For any scalar $s \in \mathbb{R}$, we define the component-wise operations:

$$\begin{aligned} \boldsymbol{\alpha} \pm \boldsymbol{k} =& (\alpha_1 \pm k_1, \ldots, \alpha_d \pm k_d), \quad \boldsymbol{\alpha} \pm s := \boldsymbol{\alpha} \pm s\mathbf{1} = (\alpha_1 \pm s, \ldots, \alpha_d \pm s), \\ \boldsymbol{\alpha}^s =& (\alpha_1^s, \ldots, \alpha_d^s), \quad \boldsymbol{\alpha}^{\boldsymbol{k}} = \alpha_1^{k_1} \cdots \alpha_d^{k_d}, \quad \boldsymbol{\alpha}! = \alpha_1! \cdots \alpha_d!, \end{aligned}$$

  and

$$\boldsymbol{\alpha} \geq \boldsymbol{k} \Leftrightarrow \alpha_j \geq k_j, \quad \forall\, 1 \leq j \leq d, \quad \boldsymbol{\alpha} \geq s \Leftrightarrow \alpha_j \geq s, \quad \forall\, 1 \leq j \leq d.$$

- ⋄ The frequently used norms are denoted as

$$\begin{aligned} &|\boldsymbol{k}|_0 = \sharp \text{ of nonzero elements in } \boldsymbol{k}, \quad |\boldsymbol{k}|_1 = \sum_{j=1}^{d} k_j, \quad |\boldsymbol{k}|_\infty = \max_{1 \leq j \leq d} k_j, \\ &|\boldsymbol{k}|_{\min} = \min\{k_j : 1 \leq j \leq d\}, \quad \text{and } |\boldsymbol{k}|_{\text{mix}} = \prod_{j=1}^{d} \bar{k}_j, \end{aligned} \tag{2.1}$$

  where $\bar{k}_j = \max\{1, k_j\}$.
- ⋄ Given a multivariate function $u(\boldsymbol{x})$, we denote the $\boldsymbol{k}^{\text{th}}$ mixed partial derivative by

$$\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}} u = \frac{\partial^{|\boldsymbol{k}|_1} u}{\partial x_1^{k_1} \cdots \partial x_d^{k_d}} = \partial_{x_1}^{k_1} \cdots \partial_{x_d}^{k_d} u.$$

  In particular, we denote $\boldsymbol{\partial}_{\boldsymbol{x}}^s u = \boldsymbol{\partial}_{\boldsymbol{x}}^{s\mathbf{1}} u = \boldsymbol{\partial}_{\boldsymbol{x}}^{(s,s,\ldots,s)} u$. Similarly, the $\boldsymbol{k}^{\text{th}}$ mixed forward difference is denoted by

$$\triangle_{\boldsymbol{x}}^{\boldsymbol{k}} u = \triangle_{x_1}^{k_1} \cdots \triangle_{x_d}^{k_d} u,$$

  where the forward differences is defined as

$$\triangle_x^k u(x) = \triangle_x \left(\triangle_x^{k-1} u(x) \right),$$

  for $k \geq 1$, where $\triangle_x u(x) = u(x+1) - u(x)$, with the convention that $\triangle_x^0 u(x) = u(x)$.
- ⋄ We follow the convention in the asymptotic analysis that $a \lesssim b$ means that there exists some constant $C > 0$ such that $a \leq Cb$, and $N \gg 1$ means that N is sufficiently large.
- ⋄ We denote C as some generic positive constant, which may vary from line to line.

2.2. The orthogonal polynomials of Askey scheme and polynomial chaos

Wiener-Askey polynomials are the orthogonal polynomials which can be expressed by using hypergeometric series. In general, a system of orthogonal polynomials $\{Q_n(x)\}_{n=0}^{\infty}$ holds the orthogonality relation with respect to a real positive measure ω, i.e.

$$\int_S Q_m(x)Q_n(x)d\omega(x) = \gamma_n \delta_{mn},$$

for $m, n = 0, 1, 2\cdots$, where $\delta_{mn} = 1$ if $m = n$, otherwise $\delta_{mn} = 0$, S is the support of the measure $\omega(x)$, and γ_n are normalization constants. Besides the orthogonality relation, all orthogonal polynomials on the real line satisfy a three-term recurrence relation:

$$-xQ_n(x) = b_n Q_{n+1}(x) + a_n Q_n(x) + c_n Q_{n-1}(x),$$

for $n \geq 1$, where $b_n, c_n \neq 0$ and $\frac{c_n}{b_{n-1}} > 0$, with $Q_{-1}(x) = 0$ and $Q_0(x) = 1$. Askey and Wilson [1] for the first time generalized the Jacobi polynomials to the Askey polynomials. The generalized hypergeometric series ${}_rF_s$ is defined by

$${}_rF_s(a_1, \cdots, a_r; b_1, \cdots, b_r; z) = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_r)_k}{(b_1)_k \cdots (b_s)_k} \frac{z^k}{k!},$$

where $b_i \neq 0$ for $i = 1, \cdots, s$, and $(a)_n$ is the Pochhammer symbol defined as

$$(a)_n = \begin{cases} 1, & n = 0 \\ a(a+1)\cdots(a+n-1), & n = 1, 2, \cdots \end{cases}. \tag{2.2}$$

For details about hypergeometric polynomials and the Askey scheme, we refer interested readers to [16]. The orthogonal polynomials of Askey scheme (namely the ones in [24]) and their properties will be used frequently in this paper, which can be found in the appendix of [22] and references therein. For the readers' convenience, we include them in appendix B. The continuous ones are Hermite, Laguerre and Jacobi polynomials, while Charlier, Meixner, Krawtchouk and Hahn polynomials are the discrete ones.

The gPC has been proposed for the first time in [24] to get the optimal convergence with the non-Gaussian random inputs. It is a generalization of the Wiener PC expansion. The expansion basis is a set of complete orthogonal polynomials of Askey scheme introduced before. For any second-order random variable $u(\theta)$, it can be expanded as

$$\begin{aligned} u(\theta) =& a_0 I_0 + \sum_{i_1=1}^{\infty} c_{i_1} I_1(\xi_{i_1}(\theta)) + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} c_{i_1 i_2} I_2(\xi_{i_1}(\theta), \xi_{i_2}(\theta)) \\ &+ \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} c_{i_1 i_2 i_3} I_3(\xi_{i_1}(\theta), \xi_{i_2}(\theta), \xi_{i_3}(\theta)) + \cdots, \end{aligned}$$

where $I_n(\xi_{i_1}, \cdots, \xi_{i_n})$ denotes the PC of order n in terms of random vector $\boldsymbol{\xi} = (\xi_{i_1}, \cdots, \xi_{i_n})$. It is clear to see that the more independent random variables ξ_{i_j}s used, the higher order of PC applied, the more terms appear in the expansion, and intuitively the closer the expansion to the random process $u(\theta)$ is. For the sake of conciseness, we rewrite the expansion as

$$u(\theta) = \sum_{|\boldsymbol{i}|=0}^{\infty} c_{\boldsymbol{i}} \boldsymbol{\Phi}_{\boldsymbol{i}}(\boldsymbol{\xi}(\theta)), \tag{2.3}$$

where $|\boldsymbol{i}|$ can be $|\boldsymbol{i}|_\infty$ or other norms, $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots)$, and $\boldsymbol{\Phi}_{\boldsymbol{i}} = \Phi_{i_1}(\xi_1)\Phi_{i_2}(\xi_2)\cdots$, where Φ_i are orthogonal polynomials of Askey scheme of degree i.

3. Multivariate orthogonal projection and approximations

As shown in (2.3), the gPC is an expansion with infinite many terms. It is also shown that in the standard ANOVA approximation with certain measure is exactly the expansion (2.3) with $|\boldsymbol{i}|_0 \leq \nu$, which also contains infinite many terms, see details in appendix A. It only becomes practical when certain truncation has been made. To be specific, suppose there are $d \geq 1$ independent random variables ξ_i, $i = 1, \cdots, d$, denoted briefly as $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_d)$, and we choose orthogonal polynomials of Askey scheme $\boldsymbol{\Phi}_{\boldsymbol{i}}(\boldsymbol{\xi})$ of certain order such that $\boldsymbol{i} \in \boldsymbol{\Omega}_N$, where $\boldsymbol{\Omega}_N$ is an index set parametrized by $N \in \mathbb{N}$ such that $\text{card}(\boldsymbol{\Omega}_N) < \infty$, then the gPC (2.3) becomes an approximation with finite terms:

$$u(\theta) \approx \sum_{\boldsymbol{i} \in \boldsymbol{\Omega}_N} c_{\boldsymbol{i}} \boldsymbol{\Phi}_{\boldsymbol{i}}(\boldsymbol{\xi}(\theta)).$$

Therefore, it is natural to ask how to choose such index set $\boldsymbol{\Omega}_N$ so that $\text{card}(\boldsymbol{\Omega}_N)$ can be as small as possible without sacrificing the accuracy too much, and how the error changes with respect to the parameter N and the dimension d.

In this paper, we shall focus on the error analysis of the projection with three type of index sets: tensor product, regular hyperbolic cross (RHC) and optimal hyperbolic cross (OHC) with parameter $\gamma \in [-\infty, 1)$, which are defined as

$$\begin{aligned} \boldsymbol{\Omega}_{N,\text{tensor}} &:= \left\{ \boldsymbol{i} \in \mathbb{N}_0^d : \, |\boldsymbol{i}|_\infty \leq N \right\}, \\ \boldsymbol{\Omega}_{N,RHC} &:= \left\{ \boldsymbol{i} \in \mathbb{N}_0^d : \, |\boldsymbol{i}|_{\text{mix}} \leq N \right\}, \\ \boldsymbol{\Omega}_{N,OHC,\gamma} &:= \left\{ \boldsymbol{i} \in \mathbb{N}_0^d : \, |\boldsymbol{i}|_{\text{mix}} |\boldsymbol{i}|_\infty^{-\gamma} \leq N^{1-\gamma} \right\}, \quad -\infty \leq \gamma < 1. \end{aligned} \tag{3.1}$$

The typical error estimate is of the form

$$\inf_{U_N \in X_N} ||u - U_N||_l = ||u - P_N u||_l \leq C N^{-c(l,r)} ||u||_r,$$

where C is a generic constant independent of N, but it may depend on d, $c(l, r)$ is some positive constant depending on l and r, $|| \cdot ||_l$ is the norm of some functional space, l indicates the regularity in some sense, X_N is a linear subspace spanned by the orthogonal polynomials of Askey scheme $\Phi_{\boldsymbol{i}}$, i.e.

$$X_N := \text{span} \left\{ \boldsymbol{\Phi}_{\boldsymbol{i}} : \, \boldsymbol{i} \in \boldsymbol{\Omega}_N \right\},$$

and P_N projects u onto the subspace X_N, i.e.

$$P_N u(\theta) = \sum_{\boldsymbol{i} \in \boldsymbol{\Omega}_N} \hat{u}_{\boldsymbol{i}} \boldsymbol{\Phi}_{\boldsymbol{i}}(\boldsymbol{\xi}(\theta)).$$

In this section, we only include the detailed proofs of the error analysis using Laguerre polynomials as a representative of the continuous ones and the Charlier polynomails as that of the discrete ones. All the results by using other orthogonal polynomials of Askey scheme will be stated without proofs in section 4.

3.1. Approximation by using Laguerre polynomials

In this subsection, we shall show the approximation by Laguerre polynomials in detail. The subspace X_N is defined as

$$X_N^{\boldsymbol{\alpha}} = \text{span} \left\{ \boldsymbol{L}_{\boldsymbol{n}}^{(\boldsymbol{\alpha})} : \, \boldsymbol{n} \in \boldsymbol{\Omega}_N \right\}, \tag{3.2}$$

for some $\boldsymbol{\alpha} > -1$, where $\boldsymbol{\Omega}_N \subset \mathbb{N}_0^d$ is one of the index sets $\boldsymbol{\Omega}_{N,\text{tensor}}$, $\boldsymbol{\Omega}_{N,RHC}$ and $\boldsymbol{\Omega}_{N,OHC,\gamma}$ in (3.1). Let us denote the orthogonal projection operator $P_N^{\boldsymbol{\alpha}} : L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}}\left(\mathbb{R}_+^d\right) \to X_N^{\boldsymbol{\alpha}}$, i.e., for any $u \in L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}}\left(\mathbb{R}_+^d\right)$,

$$\langle (u - P_N^{\boldsymbol{\alpha}} u), v \rangle_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}} = 0, \quad \forall \, v \in X_N^{\boldsymbol{\alpha}}, \tag{3.3}$$

where $L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}}\left(\mathbb{R}^d_+\right)$ is the weighted L^2 space, and $\langle u, v\rangle_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}} = \int_{\mathbb{R}^d_+} uv\boldsymbol{\omega}_{\boldsymbol{\alpha}}dx$ is the weighted inner product in $L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}}\left(\mathbb{R}^d_+\right)$. Or equivalently,

$$P_N^{\boldsymbol{\alpha}}u(\theta) = \sum_{\boldsymbol{n}\in\boldsymbol{\Omega}_N} \hat{u}_{\boldsymbol{n}}\boldsymbol{L}^{(\boldsymbol{\alpha})}_{\boldsymbol{n}}(\boldsymbol{\xi}(\theta)),$$

where $\hat{u}_n$ is the Fourier-Laguerre coefficient, which can be computed by

$$\hat{u}_{\boldsymbol{n}} = \frac{1}{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{\alpha}}}\langle u(\theta), L^{(\boldsymbol{\alpha})}_{\boldsymbol{n}}(\boldsymbol{\xi}(\theta))\rangle_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}},$$

with $\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{\alpha}}$ specified in (B.9) and (B.11).

We shall estimate how close the projection $P_N^{\boldsymbol{\alpha}}u$ is to u, with respect to various norms and index sets $\boldsymbol{\Omega}_N$.

3.1.1. Tensor product

The index set $\boldsymbol{\Omega}_N$ corresponding to the d-dimensional tensor product is

$$\boldsymbol{\Omega}_{N,\text{tensor}} = \left\{\boldsymbol{n} \in \mathbb{N}_0^d : \ |\boldsymbol{n}|_\infty \leq N\right\},$$

and $X_N^{\boldsymbol{\alpha}}$ is defined in (3.2) with $\boldsymbol{\Omega}_N = \boldsymbol{\Omega}_{N,\text{tensor}}$. Let us define the Sobolev-type space as

$$\mathcal{W}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right) = \left\{u : \ \boldsymbol{\partial}^{\boldsymbol{k}}_{\boldsymbol{x}}u \in L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k}}}\left(\mathbb{R}^d_+\right), 0 \leq |\boldsymbol{k}|_1 \leq m\right\}, \quad \forall\, m \in \mathbb{N}_0, \tag{3.4}$$

equipped with the norm and seminorm

$$||u||_{\mathcal{W}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)} = \left(\sum_{0\leq|\boldsymbol{k}|_1\leq m} \left|\left|\boldsymbol{\partial}^{\boldsymbol{k}}_{\boldsymbol{x}}u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k}},\mathbb{R}^d_+}\right)^{\frac{1}{2}},$$

$$|u|_{\mathcal{W}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)} = \left(\sum_{j=1}^d \left|\left|\partial^m_{x_j}u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+m\boldsymbol{e}_j},\mathbb{R}^d_+}\right)^{\frac{1}{2}}.$$

It is clear that $\mathcal{W}^0_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right) = L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}}\left(\mathbb{R}^d_+\right)$, and

$$|u|^2_{\mathcal{W}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)} = \sum_{j=1}^d \sum_{\boldsymbol{n}\in\mathbb{N}_0^d} \boldsymbol{\rho}_{\boldsymbol{n}-m\boldsymbol{e}_j,\boldsymbol{\alpha}+m\boldsymbol{e}_j}\left|\hat{u}_{\boldsymbol{n}}\right|^2, \tag{3.5}$$

where $\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{\alpha}} = \prod_{j=1}^d \rho_{n_j,\alpha_j}$, $\rho_{n,\alpha} = \frac{(\alpha+1)_n}{n!}$, with the Pochhammer symbol $(a)_n$ defined in (2.2). It is followed from the orthogonality of Laguerre polynomials with respect to the gamma distribution, i.e.

$$\int_{\mathbb{R}_+} L^{(\alpha)}_m(x)L^{(\alpha)}_n(x)\omega_\alpha(x)dx = \frac{(\alpha+1)_n}{n!}\delta_{mn} =: \rho_{n,\alpha}\delta_{mn}.$$

We also define the Koborov-type space as

$$\mathcal{K}^r_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right) = \left\{u : \ \boldsymbol{\partial}^{\boldsymbol{k}}_{\boldsymbol{x}}u \in L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k}}}\left(\mathbb{R}^d_+\right), \ 0 \leq |\boldsymbol{k}|_\infty \leq r\right\}, \quad \forall\, r \in \mathbb{N}_0^d, \tag{3.6}$$

equipped with the norm and seminorm

$$||u||_{\mathcal{K}^r_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)} = \left(\sum_{0\leq|\boldsymbol{k}|_\infty\leq r} \left|\left|\boldsymbol{\partial}^{\boldsymbol{k}}_{\boldsymbol{x}}u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k}},\mathbb{R}^d_+}\right)^{\frac{1}{2}},$$

$$|u|_{\mathcal{K}^r_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)} = \left(\sum_{|\boldsymbol{k}|_\infty=r} \left|\left|\boldsymbol{\partial}^{\boldsymbol{k}}_{\boldsymbol{x}}u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k}},\mathbb{R}^d_+}\right)^{\frac{1}{2}}.$$

It is easy to see from the definitions that $\mathcal{K}^0_{\boldsymbol{\alpha}}(\mathbb{R}^d_+) = L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}}\left(\mathbb{R}^d_+\right)$ and $\mathcal{W}^{dl}_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right) \subset \mathcal{K}^l_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right) \subset \mathcal{W}^l_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)$.

Theorem 3.1 (tensor product with Laguerre polynomials). *For any $0 \le l < m$, if $u \in \mathcal{W}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)$, we have*

$$
\begin{aligned}
&|P_N^{\boldsymbol{\alpha}} u - u|_{\mathcal{W}^l_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)} \qquad (3.7)\\
\le &\left[(|\boldsymbol{\alpha}|_\infty + l + 1)_{m-l} + d(|\boldsymbol{\alpha}|_\infty + 1)_m \max\left\{\frac{1}{(|\boldsymbol{\alpha}|_{\min} + 1)_l}, 1\right\}\right]^{\frac{1}{2}} (N - m + 1)^{\frac{l-m}{2}} |u|_{\mathcal{W}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)},
\end{aligned}
$$

for $N \gg 1$. Furthermore, if $u \in \mathcal{K}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)$, for $0 \le l < m$, we have

$$
|P_N^{\boldsymbol{\alpha}} u - u|_{\mathcal{K}^l_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)} \le d^{\frac{1}{2}} (l+1)^{\frac{d-1}{2}} \left[(|\boldsymbol{\alpha}|_\infty + l + 1)_m\right]^{\frac{d}{2}} (N - m + 1)^{\frac{l-m}{2}} |u|_{\mathcal{K}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)}. \qquad (3.8)
$$

Proof. To obtain the estimate (3.7) in Sobolev space, we proceed as that in [17]. Let $\boldsymbol{\Omega}^c_{N,\text{tensor}} = \left\{\boldsymbol{n} \in \mathbb{N}^d_0 :\ |\boldsymbol{n}|_\infty > N\right\}$. By (3.5), we have

$$
|P_N^{\boldsymbol{\alpha}} u - u|^2_{\mathcal{W}^l_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)} = \sum_{j=1}^d \sum_{\boldsymbol{n} \in \boldsymbol{\Omega}^c_N} \boldsymbol{\rho}_{\boldsymbol{n} - l\boldsymbol{e}_j, \boldsymbol{\alpha} + l\boldsymbol{e}_j} \left|\hat{u}_{\boldsymbol{n}}\right|^2. \qquad (3.9)
$$

For any $1 \le j \le d$,

$$
\sum_{\boldsymbol{n} \in \boldsymbol{\Omega}^c_N} \boldsymbol{\rho}_{\boldsymbol{n} - l\boldsymbol{e}_j, \boldsymbol{\alpha} + l\boldsymbol{e}_j} \left|\hat{u}_{\boldsymbol{n}}\right|^2 = \sum_{\boldsymbol{n} \in \boldsymbol{\Lambda}^{1,j}_N} \boldsymbol{\rho}_{\boldsymbol{n} - l\boldsymbol{e}_j, \boldsymbol{\alpha} + l\boldsymbol{e}_j} \left|\hat{u}_{\boldsymbol{n}}\right|^2 + \sum_{\boldsymbol{n} \in \boldsymbol{\Lambda}^{2,j}_N} \boldsymbol{\rho}_{\boldsymbol{n} - l\boldsymbol{e}_j, \boldsymbol{\alpha} + l\boldsymbol{e}_j} \left|\hat{u}_{\boldsymbol{n}}\right|^2 := I_1 + I_2, \qquad (3.10)
$$

where

$$
\boldsymbol{\Lambda}^{1,j}_N = \{\boldsymbol{n} \in \boldsymbol{\Omega}^c_N :\ n_j > N\}; \quad \boldsymbol{\Lambda}^{2,j}_N = \{\boldsymbol{n} \in \boldsymbol{\Omega}^c_N :\ n_j \le N\}. \qquad (3.11)
$$

For I_1:

$$
I_1 \le \max_{\boldsymbol{n} \in \boldsymbol{\Lambda}^{1,j}_N} \left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n} - l\boldsymbol{e}_j, \boldsymbol{\alpha} + l\boldsymbol{e}_j}}{\boldsymbol{\rho}_{\boldsymbol{n} - m\boldsymbol{e}_j, \boldsymbol{\alpha} + m\boldsymbol{e}_j}}\right\} \sum_{\boldsymbol{n} \in \boldsymbol{\Lambda}^{1,j}_N} \boldsymbol{\rho}_{\boldsymbol{n} - m\boldsymbol{e}_j, \boldsymbol{\alpha} + m\boldsymbol{e}_j} \left|\hat{u}_{\boldsymbol{n}}\right|^2, \qquad (3.12)
$$

where

$$
\begin{aligned}
\max_{\boldsymbol{n} \in \boldsymbol{\Lambda}^{1,j}_N} \left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n} - l\boldsymbol{e}_j, \boldsymbol{\alpha} + l\boldsymbol{e}_j}}{\boldsymbol{\rho}_{\boldsymbol{n} - m\boldsymbol{e}_j, \boldsymbol{\alpha} + m\boldsymbol{e}_j}}\right\} &= \max_{\boldsymbol{n} \in \boldsymbol{\Lambda}^{1,j}_N} \left\{\frac{\rho_{n_j - l, \alpha_j + l}}{\rho_{n_j - m, \alpha_j + m}}\right\} = \max_{\boldsymbol{n} \in \boldsymbol{\Lambda}^{1,j}_N} \left\{\frac{(\alpha_j + l + 1)_{m-l}}{(n_j - l)(n_j - l - 1) \cdots (n_j - m + 1)}\right\}\\
&\le (\alpha_j + l + 1)_{m-l} (N - m + 1)^{l-m} \le (|\boldsymbol{\alpha}|_\infty + l + 1)_{m-l} (N - m + 1)^{l-m}, \qquad (3.13)
\end{aligned}
$$

and

$$
\sum_{j=1}^d \sum_{\boldsymbol{n} \in \Lambda^{1,j}_N} \boldsymbol{\rho}_{\boldsymbol{n} - m\boldsymbol{e}_j, \boldsymbol{\alpha} + m\boldsymbol{e}_j} \left|\hat{u}_{\boldsymbol{n}}\right|^2 \overset{(3.5)}{\le} |u|^2_{\mathcal{W}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)}. \qquad (3.14)
$$

For I_2, if $\boldsymbol{n} \in \boldsymbol{\Lambda}^{2,j}_N$, then there exists some $k \ne j$, such that $n_k > N$.

$$
I_2 \le \max_{\boldsymbol{n} \in \boldsymbol{\Lambda}^{2,j}_N} \left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n} - l\boldsymbol{e}_j, \boldsymbol{\alpha} + l\boldsymbol{e}_j}}{\boldsymbol{\rho}_{\boldsymbol{n} - m\boldsymbol{e}_k, \boldsymbol{\alpha} + m\boldsymbol{e}_k}}\right\} \sum_{\boldsymbol{n} \in \boldsymbol{\Lambda}^{2,j}_N} \boldsymbol{\rho}_{\boldsymbol{n} - m\boldsymbol{e}_k, \boldsymbol{\alpha} + m\boldsymbol{e}_k} \left|\hat{u}_{\boldsymbol{n}}\right|^2, \qquad (3.15)
$$

since

$$
\begin{aligned}
\max_{\boldsymbol{n}\in\boldsymbol{\Lambda}_N^{2,j}}\left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n}-l\boldsymbol{e}_j,\boldsymbol{\alpha}+l\boldsymbol{e}_j}}{\boldsymbol{\rho}_{\boldsymbol{n}-m\boldsymbol{e}_k,\boldsymbol{\alpha}+m\boldsymbol{e}_k}}\right\} &= \max_{\boldsymbol{n}\in\boldsymbol{\Lambda}_N^{2,j}}\left\{\frac{\rho_{n_j-l,\alpha_j+l}\rho_{n_k,\alpha_k}}{\rho_{n_j,\alpha_j}\rho_{n_k-m,\alpha_k+m}}\right\}\\
&= \begin{cases} \max\limits_{\boldsymbol{n}\in\boldsymbol{\Lambda}_N^{2,j}}\left\{\dfrac{(n_j-l+1)\cdots n_j}{(n_k-m+1)\cdots n_k}\cdot\dfrac{(\alpha_k+1)\cdots(\alpha_k+m)}{(\alpha_j+1)\cdots(\alpha_j+l)}\right\}, & \text{if } l\geq 1\\ \max\limits_{\boldsymbol{n}\in\boldsymbol{\Lambda}_N^{2,j}}\left\{\dfrac{(\alpha_k+1)_m}{n_k\cdots(n_k-m+1)}\right\}, & \text{if } l=0\end{cases}\\
&\leq \begin{cases} \dfrac{(\alpha_k+1)_m}{(\alpha_j+1)_l}\cdot\dfrac{1}{(N-m+1)\cdots(N-l)}, & \text{if } l\geq 1\\ \dfrac{(\alpha_k+1)_m}{(N-m+1)\cdots N}, & \text{if } l=0\end{cases}\\
&\leq \begin{cases} \dfrac{(|\boldsymbol{\alpha}|_\infty+1)_m}{(|\boldsymbol{\alpha}|_{\min}+1)_l}(N-m+1)^{l-m}, & \text{if } l\geq 1\\ (|\boldsymbol{\alpha}|_\infty+1)_m(N-m+1)^{-m}, & \text{if } l=0\end{cases},
\end{aligned}
\tag{3.16}
$$

and

$$
\sum_{j=1}^{d}\sum_{\boldsymbol{n}\in\Lambda_N^{2,j}}\boldsymbol{\rho}_{\boldsymbol{n}-m\boldsymbol{e}_k,\boldsymbol{\alpha}+m\boldsymbol{e}_k}\left|\hat{u}_{\boldsymbol{n}}\right|^2\leq\sum_{j=1}^{d}\sum_{k=1}^{d}\sum_{\boldsymbol{n}\in\Lambda_N^{2,j}}\boldsymbol{\rho}_{\boldsymbol{n}-m\boldsymbol{e}_k,\boldsymbol{\alpha}+m\boldsymbol{e}_k}\left|\hat{u}_{\boldsymbol{n}}\right|^2\leq d|u|^2_{\mathcal{W}^m_{\boldsymbol{\alpha}}(\mathbb{R}^d_+)}. \tag{3.17}
$$

Combine (3.9)-(3.17), we obtain the result (3.7).

To obtain the estimate (3.8) in Koborov space, we need to estimate $\left|\left|\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{l}}\left(P_N^{\boldsymbol{\alpha}}u-u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{l}},\mathbb{R}^d_+}$, for $0\leq\boldsymbol{l}<m$. For given $\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,\text{tensor}}$, we split the index $1\leq j\leq d$ into two parts

$$
\mathcal{N}:=\{j:\ l_j\leq n_j<m,\,1\leq j\leq d\},\quad \mathcal{N}^c:=\{j:\ n_j\geq m,\,1\leq j\leq d\}. \tag{3.18}
$$

It is easy to see that $\mathcal{N}^c$ cannot be empty, due to the fact that $|\boldsymbol{n}|_\infty>N>m$. We denote

$$
\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{\alpha}}:=\left(\prod_{j\in\mathcal{N}}\rho_{n_j-l_j,\alpha_j+l_j}\right)\left(\prod_{i\in\mathcal{N}^c}\rho_{n_i-m,\alpha_i+m}\right). \tag{3.19}
$$

From the orthogonality and the property that

$$
\frac{d^k}{dx^k}L_n^{(\alpha)}(x)=(-1)^kL_{n-k}^{(\alpha+k)}(x),
$$

if $n\geq k\geq 0$, we have

$$
\begin{aligned}
\left|\left|\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{l}}(P_N^{\boldsymbol{\alpha}}u-u)\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{l}},\mathbb{R}^d_+} &= \sum_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,\text{tensor}}}\boldsymbol{\rho}_{\boldsymbol{n}-\boldsymbol{l},\boldsymbol{\alpha}+\boldsymbol{l}}\left|\hat{u}_{\boldsymbol{n}}\right|^2\\
&\leq \max_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,\text{tensor}}}\left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n}-\boldsymbol{l},\boldsymbol{\alpha}+\boldsymbol{l}}}{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{\alpha}}}\right\}\sum_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,\text{tensor}}}\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{\alpha}}\left|\hat{u}_{\boldsymbol{n}}\right|^2.
\end{aligned}
\tag{3.20}
$$

It remains to estimate the maximum in (3.20). Similarly as in (3.13), we get

$$
\frac{\boldsymbol{\rho}_{\boldsymbol{n}-\boldsymbol{l},\boldsymbol{\alpha}+\boldsymbol{l}}}{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{\alpha}}}=\prod_{j\in\mathcal{N}^c}\frac{\rho_{n_j-l_j,\alpha_j+l_j}}{\rho_{n_j-m,\alpha_j+m}}=\prod_{j\in\mathcal{N}^c}\frac{(\alpha_j+l_j+1)_{m-l_j}}{(n_j-l_j)\cdots(n_j-m+1)}.
$$

Notice that $\boldsymbol{n} \in \boldsymbol{\Omega}_{N,\text{tensor}}^{c}$, there exists at least one j_0 such that $n_{j_0} > N$. Therefore, we have

$$\begin{aligned}
&\max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^{c}} \left\{ \frac{\boldsymbol{\rho}_{\boldsymbol{n}-\boldsymbol{l},\boldsymbol{\alpha}+\boldsymbol{l}}}{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{\alpha}}} \right\} \qquad (3.21)\\
\leq & \max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^{c}} \left\{ \prod_{j\in\mathcal{N}^c} (\alpha_j + l_j + 1)_{m-l_j} \right\} \max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^{c}} \left\{ \prod_{j\in\mathcal{N}^c, j\neq j_0} \frac{1}{(n_j - l_j)\cdots(n_j - m + 1)} \right\}\\
&\cdot \max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^{c}} \left\{ \frac{1}{(n_{j_0} - l_{j_0})\cdots(n_{j_0} - m + 1)} \right\}\\
\leq & \left[(|\boldsymbol{\alpha}|_\infty + |\boldsymbol{l}|_\infty + 1)_{m-|\boldsymbol{l}|_{\min}} \right]^d (N - m + 1)^{|\boldsymbol{l}|_\infty - m},
\end{aligned}$$

since

$$\begin{aligned}
\max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^{c}} \left\{ \prod_{j\in\mathcal{N}^c} (\alpha_j + l_j + 1)_{m-l_j} \right\} &\leq \max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^{c}} \left\{ \prod_{j\in\mathcal{N}^c} (|\boldsymbol{\alpha}|_\infty + |\boldsymbol{l}|_\infty + 1)_{m-|\boldsymbol{l}|_{\min}} \right\}\\
&\leq \left[(|\boldsymbol{\alpha}|_\infty + |\boldsymbol{l}|_\infty + 1)_{m-|\boldsymbol{l}|_{\min}} \right]^d, \qquad (3.22)\\
\max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^{c}} \left\{ \frac{1}{(n_{j_0} - l_{j_0})\cdots(n_{j_0} - m + 1)} \right\} &\leq (N - m + 1)^{|\boldsymbol{l}|_\infty - m},
\end{aligned}$$

and the fact that the second maximum on the right-hand side of (3.21) is less than or equal to 1. Therefore, we have

$$\sum_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^{c}} \boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{\alpha}} |\hat{u}_{\boldsymbol{n}}|^2 \leq \left|\left|\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}} u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k}},\mathbb{R}_+^d} \leq |u|^2_{\mathcal{K}_{\boldsymbol{\alpha}}^m(\mathbb{R}_+^d)}, \qquad (3.23)$$

where $\boldsymbol{k}$ is a d-dimensional index consisting of l_j for $j \in \mathcal{N}$ and m for $j \in \mathcal{N}^c$, with $|\boldsymbol{k}|_\infty = m$. The result (3.8) follows immediately from (3.20)-(3.23) and the fact that

$$|P_N^{\boldsymbol{\alpha}} u - u|^2_{\mathcal{K}_{\boldsymbol{\alpha}}^l(\mathbb{R}_+^d)} = \sum_{|\boldsymbol{l}|_\infty = l} \left|\left|\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{l}} (P_N^{\boldsymbol{\alpha}} u - u)\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{l}},\mathbb{R}_+^d},$$

with $\text{card}\left(\{\boldsymbol{l} : |\boldsymbol{l}|_\infty = l\}\right) = d(l+1)^{d-1}$. □

It is clear that the convergence rate deteriorates rapidly with respect to the dimension d. That is,

$$||P_N^{\alpha} u - u||^2_{\mathcal{K}_{\boldsymbol{\alpha}}^l(\mathbb{R}_+^d)} = \sum_{r=0}^{l} |P_N^{\alpha} u - u|^2_{\mathcal{K}_{\boldsymbol{\alpha}}^r(\mathbb{R}_+^d)} \lesssim \text{card}(\boldsymbol{\Omega}_{N,\text{tensor}})^{\frac{l-m}{d}} |u|^2_{\mathcal{K}_{\boldsymbol{\alpha}}^m(\mathbb{R}_+^d)},$$

since $\text{card}(\boldsymbol{\Omega}_{N,\text{tensor}}) = (N+1)^d$.

3.1.2. RHC approximation

As we mentioned in the introduction, the HC approximation is an efficient tool to overcome the "curse of dimensionality" in some degree. The index set of RHC approximation is $\boldsymbol{\Omega}_{N,RHC} = \left\{\boldsymbol{n} \in \mathbb{N}_0^d : |\boldsymbol{n}|_{\text{mix}} \leq N\right\}$. It is known that the cardinality of $\boldsymbol{\Omega}_{N,RHC}$ is $\mathcal{O}\left(N(\ln N)^{d-1}\right)$ [8]. Correspondingly, the finite dimensional subspace $X_N^{\boldsymbol{\alpha}}$ is

$$X_N^{\boldsymbol{\alpha}} = \text{span}\left\{\boldsymbol{L}_{\boldsymbol{n}}^{\boldsymbol{\alpha}} : \boldsymbol{n} \in \boldsymbol{\Omega}_{N,RHC}\right\}. \qquad (3.24)$$

Let the orthogonal projection operator $P_N^{\boldsymbol{\alpha}} : L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}}\left(\mathbb{R}_+^d\right) \to X_N^{\boldsymbol{\alpha}}$ be defined in (3.3). The similar result as Theorem 3.2 and 3.3 below for Jacobi polynomials have been obtained in [17] for the first time with a gap. Yau and the author [12] made it rigorous for generalized Hermite functions. In this paper, the error analysis in [12] has been further simplified, see detailed discussion in Remark 3.1.

Theorem 3.2 (RHC with Laguerre polynomials). *Given $u \in \mathcal{K}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)$, for $0 \le \boldsymbol{l} < m$, we have*

$$\left|\left|\boldsymbol{\partial}^{\boldsymbol{l}}_{\boldsymbol{x}}\left(P^{\boldsymbol{\alpha}}_N u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{l}},\mathbb{R}^d_+} \le \left[(|\boldsymbol{\alpha}|_\infty + |\boldsymbol{l}|_\infty + 1)_{m-|\boldsymbol{l}|_{\min}}\right]^{\frac{d}{2}} m^{\frac{d(2m-|\boldsymbol{l}|_\infty)-|\boldsymbol{l}|_{\min}}{2}} N^{\frac{|\boldsymbol{l}|_\infty - m}{2}} |u|_{\mathcal{K}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)},$$

for $N \gg 1$.

Proof. We proceed as the second part in the proof of Theorem 3.1. For given $\boldsymbol{n} \in \boldsymbol{\Omega}^c_{N,RHC}$, we split the index $1 \le j \le d$ into $\mathcal{N}$ and $\mathcal{N}^c$ two parts as in (3.18). It is easy to see that $\mathcal{N}^c$ cannot be empty, otherwise, $|\boldsymbol{n}|_{\text{mix}} \le m^d < N$, for $N \gg 1$, which contradicts with $\boldsymbol{n} \in \boldsymbol{\Omega}^c_{N,RHC}$. As before, we need to estimate the maximum in (3.20) within $\boldsymbol{n} \in \boldsymbol{\Omega}^c_{N,RHC}$. Similarly as in (3.13), we get

$$\begin{aligned}\frac{\boldsymbol{\rho}_{\boldsymbol{n}-\boldsymbol{l},\boldsymbol{\alpha}+\boldsymbol{l}}}{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{\alpha}}} &= \prod_{j\in\mathcal{N}^c} \frac{\rho_{n_j-l_j,\alpha_j+l_j}}{\rho_{n_j-m,\alpha_j+m}} = \prod_{j\in\mathcal{N}^c} \frac{(\alpha_j+l_j+1)_{m-l_j}}{(n_j-l_j)\cdots(n_j-m+1)} \\ &= \prod_{j\in\mathcal{N}^c} n_j^{l_j-m} \prod_{j\in\mathcal{N}^c}\left(1-\frac{l_j}{n_j}\right)^{-1}\cdots\left(1-\frac{m-1}{n_j}\right)^{-1} \prod_{j\in\mathcal{N}^c}(\alpha_j+l_j+1)_{m-l_j}. \end{aligned} \tag{3.25}$$

Observe that $j \in \mathcal{N}^c$ implies $n_j \ge m > \boldsymbol{l} \ge 0$. That is, $n_j \ge 1$. Hence, $\bar{n}_j = n_j$, for all $j = 1, \cdots, d$. Given any $\boldsymbol{n} \in \boldsymbol{\Omega}^c_{N,RHC}$, we deduce that

$$\prod_{j\in\mathcal{N}^c} \bar{n}_j > \frac{N}{\prod_{j\in\mathcal{N}} \bar{n}_j} > \frac{N}{\prod_{j\in\mathcal{N}} m} \ge m^{-d}N.$$

Thus,

$$\begin{aligned}\max_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,RHC}}\left\{\prod_{j\in\mathcal{N}^c} n_j^{l_j-m}\right\} &\le \max_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,RHC}}\left\{\prod_{j\in\mathcal{N}^c} n_j^{|\boldsymbol{l}|_\infty-m}\right\} = \max_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,RHC}}\left(\prod_{j\in\mathcal{N}^c} n_j\right)^{|\boldsymbol{l}|_\infty-m} \\ &\le \left(\min_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,RHC}}\prod_{j\in\mathcal{N}^c} n_j\right)^{|\boldsymbol{l}|_\infty-m} \le \left(m^{-d}N\right)^{|\boldsymbol{l}|_\infty-m} = m^{d(m-|\boldsymbol{l}|_\infty)}N^{|\boldsymbol{l}|_\infty-m}.\end{aligned} \tag{3.26}$$

Furthermore, we have

$$\begin{aligned}\max_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,RHC}}\left\{\prod_{j\in\mathcal{N}^c}\left(1-\frac{l_j}{n_j}\right)^{-1}\cdots\left(1-\frac{m-1}{n_j}\right)^{-1}\right\} &\le \max_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,RHC}}\left\{\prod_{j\in\mathcal{N}^c}\left(1-\frac{m-1}{n_j}\right)^{l_j-m}\right\} \\ &\le \prod_{j\in\mathcal{N}^c} m^{m-l_j} \le m^{dm-|\boldsymbol{l}|_{\min}}.\end{aligned} \tag{3.27}$$

The desired result follows immediately from (3.20), (3.25)-(3.27), (3.22) and (3.23). □

It is clear to see that

$$||P^{\boldsymbol{\alpha}}_N u - u||_{\mathcal{K}^l_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)} \lesssim N^{\frac{l-m}{2}}|u|_{\mathcal{K}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)} \le \text{card}(\boldsymbol{\Omega}_{N,RHC})^{\frac{l-m}{2(1+\epsilon(d-1))}}|u|_{\mathcal{K}^m_{\boldsymbol{\alpha}}\left(\mathbb{R}^d_+\right)}, \quad \forall\, 0 \le l < m,$$

where $\text{card}(\boldsymbol{\Omega}_{N,RHC}) = \mathcal{O}\left(N(\ln N)^{d-1}\right) \le CN^{1+\epsilon(d-1)}$, for arbitrary small $\epsilon > 0$. Here, the convergence rate deteriorates slightly with increasing d.

Remark 3.1. *In [12], the authors estimate (3.20) by splitting the index set $\boldsymbol{\Omega}^c_N$*

$$\sum_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_N} = \sum_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,m}} + \sum_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,\boldsymbol{l}}\setminus\boldsymbol{\Omega}^c_{N,m}} =: II_1 + II_2,$$

where $\boldsymbol{\Omega}^c_{N,\boldsymbol{k}} := \left\{\boldsymbol{n} \in \mathbb{N}^d_0 : |\boldsymbol{n}|_{\text{mix}} > N \text{ and } \boldsymbol{n} \ge \boldsymbol{k}\right\}$, for some given $\boldsymbol{k} \in \mathbb{N}^d_0$. In this paper, one realizes that the method used to estimate II_2 in [12] is also applicable to II_1 with $\mathcal{N} = \emptyset$. Therefore, it is rebundant to analyze II_1 seperately. This is also true in the proof of Theorem 3.3 for OHC approximation.

3.1.3. OHC approximation

In order to alleviate the curse of dimensionality further, we consider the OHC index set introduced in [8]:

$$\boldsymbol{\Omega}_{N,OHC,\gamma} := \left\{\boldsymbol{n} \in \mathbb{N}_0^d : \ |\boldsymbol{n}|_{\text{mix}} |\boldsymbol{n}|_{\infty}^{-\gamma} \le N^{1-\gamma}\right\}, \quad -\infty \le \gamma < 1. \tag{3.28}$$

The cardinality of $\boldsymbol{\Omega}_{N,OHC,\gamma}$ is $\mathcal{O}(N)$, for $\gamma \in (0,1)$, where the dependence of dimension is in the big-O, see [8]. The family of spaces are defined as

$$X_{N,\gamma}^{\boldsymbol{\alpha}} := \text{span}\left\{\boldsymbol{L}_{\boldsymbol{n}}^{(\boldsymbol{\alpha})} : \ \boldsymbol{n} \in \boldsymbol{\Omega}_{N,OHC,\gamma}\right\}. \tag{3.29}$$

Remark 3.2. *Actually, OHC is a generalization of RHC and tensor product. In particular,* $X_{N,0}^{\boldsymbol{\alpha}} = X_N^{\boldsymbol{\alpha}}$ *in* (3.24) *corresponds to RHC approximation, while* $X_{N,-\infty}^{\boldsymbol{\alpha}} = \text{span}\left\{\boldsymbol{L}_{\boldsymbol{n}}^{(\boldsymbol{\alpha})} : |\boldsymbol{n}|_\infty \le N\right\}$ *describes the tensor product.*

We denote the projection operator as $P_{N,\gamma}^{\boldsymbol{\alpha}} : \ L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}}\left(\mathbb{R}_+^d\right) \to X_{N,\gamma}^{\boldsymbol{\alpha}}$.

Theorem 3.3 (OHC with Laguerre polynomials). *For any* $u \in \mathcal{K}_{\boldsymbol{\alpha}}^m\left(\mathbb{R}_+^d\right)$, $d \ge 2$, *and* $0 \le |\boldsymbol{l}|_1 < m$,

$$\begin{aligned}\left|\left|\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{l}}\left(P_{N,\gamma}^{\boldsymbol{\alpha}} u - u\right)\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{l}},\mathbb{R}_+^d} &\le \left[(|\boldsymbol{\alpha}|_\infty + |\boldsymbol{l}|_\infty + 1)_{m-|\boldsymbol{l}|_{\min}}\right]^d m^{dm-|\boldsymbol{l}|_{\min}} |u|^2_{\mathcal{K}_{\boldsymbol{\alpha}}^m\left(\mathbb{R}_+^d\right)} \\ &\cdot \begin{cases} m^{\frac{(d-1)(\gamma m - |\boldsymbol{l}|_1)}{1-\gamma}} N^{|\boldsymbol{l}|_1 - m}, & \text{if } 0 < \gamma \le \dfrac{|\boldsymbol{l}|_1}{m} \\ m^{\frac{d-1}{d-\gamma}(dm - |\boldsymbol{l}|_1)} N^{-\frac{1-\gamma}{d-\gamma}(dm-|\boldsymbol{l}|_1)}, & \text{if } \dfrac{|\boldsymbol{l}|_1}{m} \le \gamma < 1 \end{cases}.\end{aligned} \tag{3.30}$$

Proof. As argued in the proof of Theorem 3.2, we arrive at

$$\left|\left|\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{l}}\left(P_{N,\gamma}^{\boldsymbol{\alpha}} u - u\right)\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{l}},\mathbb{R}_+^d} \le \max_{\boldsymbol{n} \in \boldsymbol{\Omega}_{N,OHC,\gamma}^c} \left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n}-\boldsymbol{l},\boldsymbol{\alpha}+\boldsymbol{l}}}{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{\alpha}}}\right\} \sum_{\boldsymbol{n} \in \boldsymbol{\Omega}_{N,OHC,\gamma}^c} \boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{\alpha}} \left|\hat{u}_{\boldsymbol{n}}\right|^2, \tag{3.31}$$

where $\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{\alpha}}$ is defined as in (3.19). To estimate the maximum in (3.31), we recall that

$$\frac{\boldsymbol{\rho}_{\boldsymbol{n}-\boldsymbol{l},\boldsymbol{\alpha}+\boldsymbol{l}}}{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{\alpha}}} \overset{(3.25)}{=} \prod_{j \in \mathcal{N}^c} n_j^{l_j - m} \prod_{j \in \mathcal{N}^c} \left(1 - \frac{l_j}{n_j}\right)^{-1} \cdots \left(1 - \frac{m-1}{n_j}\right)^{-1} \prod_{j \in \mathcal{N}^c} (\alpha_j + l_j + 1)_{m - l_j}, \tag{3.32}$$

where $\mathcal{N}$ and $\mathcal{N}^c$ are defined in (3.18). The maximum of the last two terms in the product of the above equality can be estimated as (3.27) and (3.22) in the proof of Theorem 3.2. It is only the first term to be estimated. It is easily verified that

$$\prod_{j \in \mathcal{N}^c} n_j^{l_j - m} \le \left(\prod_{j \in \mathcal{N}^c} |\tilde{\boldsymbol{n}}|_\infty^{l_j}\right)\left(\prod_{j \in \mathcal{N}^c} n_j\right)^{-m} = |\tilde{\boldsymbol{n}}|_\infty^{|\tilde{\boldsymbol{l}}|_1} |\tilde{\boldsymbol{n}}|_{\text{mix}}^{-m} \le |\tilde{\boldsymbol{n}}|_\infty^{|\boldsymbol{l}|_1} |\tilde{\boldsymbol{n}}|_{\text{mix}}^{-m}, \tag{3.33}$$

where $\tilde{\boldsymbol{n}}$ is defined below

$$\tilde{\boldsymbol{n}} = (n_1, \cdots, n_d) = \begin{cases} n_j, & \text{if } j \in \mathcal{N}^c \\ 0, & \text{otherwise} \end{cases}. \tag{3.34}$$

Notice that for any $\boldsymbol{n} \in \boldsymbol{\Omega}_{N,OHC,\gamma}^c$, we have

$$N^{1-\gamma} < |\boldsymbol{n}|_{\text{mix}} |\boldsymbol{n}|_\infty^{-\gamma} \le m^{d-1} |\tilde{\boldsymbol{n}}|_{\text{mix}} |\tilde{\boldsymbol{n}}|_\infty^{-\gamma} \Rightarrow \left(\frac{|\tilde{\boldsymbol{n}}|_\infty^\gamma}{|\tilde{\boldsymbol{n}}|_{\text{mix}}}\right)^{\frac{1}{1-\gamma}} < m^{\frac{d-1}{1-\gamma}} N^{-1}, \tag{3.35}$$

$$\frac{|\tilde{\boldsymbol{n}}|_\infty}{|\tilde{\boldsymbol{n}}|_{\text{mix}}} = \frac{\bar{n}_{j_0}}{\bar{n}_{j_0} \prod_{j \in \mathcal{N}^c, j \ne j_0} \bar{n}_j} \le 1, \tag{3.36}$$

since there exists $\boldsymbol{n} \in \boldsymbol{\Omega}^c_{N,OHC,\gamma}$ such that $\mathrm{card}(\mathcal{N}^c) = 1$, i.e. there is only j_0 such that $|\tilde{n}|_\infty = \bar{n}_{j_0} > N$ and all the other js belong to $\mathcal{N}$. Moreover, we have

$$N^{1-\gamma} \overset{(3.35)}{<} m^{d-1}|\tilde{\boldsymbol{n}}|_{\mathrm{mix}}|\tilde{\boldsymbol{n}}|_\infty^{-\gamma} \leq m^{d-1}|\tilde{\boldsymbol{n}}|_\infty^{d-\gamma} \Rightarrow |\tilde{\boldsymbol{n}}|_\infty > \left(\frac{N^{1-\gamma}}{m^{d-1}}\right)^{\frac{1}{d-\gamma}}. \tag{3.37}$$

Now, we are ready to estimate the maximum of the first term on the right-hand side of (3.32). If $0 < \gamma \leq \frac{|\boldsymbol{l}|_1}{m} < 1$, then

$$\begin{aligned}\max_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,OHC,\gamma}} \left\{\prod_{j\in\mathcal{N}^c} n_j^{l_j-m}\right\} &\overset{(3.33)}{<} \max_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,OHC,\gamma}} \left\{\left(\frac{|\tilde{\boldsymbol{n}}|_\infty^\gamma}{|\tilde{\boldsymbol{n}}|_{\mathrm{mix}}}\right)^{\frac{m-|\boldsymbol{l}|_1}{1-\gamma}} \left(\frac{|\tilde{\boldsymbol{n}}|_\infty}{|\tilde{\boldsymbol{n}}|_{\mathrm{mix}}}\right)^{\frac{|\boldsymbol{l}|_1-\gamma m}{1-\gamma}}\right\} \\ &\overset{(3.35),(3.36)}{\leq} m^{\frac{(d-1)(\gamma m-|\boldsymbol{l}|_1)}{1-\gamma}} N^{|\boldsymbol{l}|_1-m}.\end{aligned} \tag{3.38}$$

Otherwise, if $\frac{|\boldsymbol{l}|_1}{m} \leq \gamma < 1$, then

$$\begin{aligned}\max_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,OHC,\gamma}} \left\{\prod_{j\in\mathcal{N}^c} n_j^{l_j-m}\right\} &\overset{(3.33)}{<} \max_{\boldsymbol{n}\in\boldsymbol{\Omega}^c_{N,OHC,\gamma}} \left\{\left(\frac{|\tilde{\boldsymbol{n}}|_\infty^\gamma}{|\tilde{\boldsymbol{n}}|_{\mathrm{mix}}}\right)^{m} |\tilde{\boldsymbol{n}}|_\infty^{|\boldsymbol{l}|_1-\gamma m}\right\} \\ &\overset{(3.35),(3.37)}{\leq} m^{\frac{d-1}{d-\gamma}(dm-|\boldsymbol{l}|_1)} N^{-\frac{1-\gamma}{d-\gamma}(dm-|\boldsymbol{l}|_1)}.\end{aligned} \tag{3.39}$$

The desired result follows immediately from (3.31)-(3.39), (3.27), and (3.22)-(3.23). □

It is clear to see that

$$\left|P_{N,\gamma}^{\boldsymbol{\alpha},\boldsymbol{\beta}}u - u\right|_{\mathcal{K}^l_{\boldsymbol{\alpha}}(\mathbb{R}^d_+)} \lesssim \mathrm{card}(\boldsymbol{\Omega}_{N,OHC,\gamma})^{\frac{l-m}{2}} |u|_{\mathcal{K}^m_{\boldsymbol{\alpha}}(\mathbb{R}^d_+)},$$

where $\mathrm{card}(\boldsymbol{\Omega}_{N,OHC,\gamma}) = \mathcal{O}(N) \leq CN$. The convergence rate does not deteriorate with respect to d anymore. The effect of the dimension goes into the constant in front.

Remark 3.3. *If N in the index set $\boldsymbol{\Omega}_N$ is the same in both RHC and OHC approximation, then the convergence rate of RHC is better than that of OHC, for $d \geq 2$. If the cardinality of $\boldsymbol{\Omega}_N$ is the same in both cases, then OHC presents a faster convergence, for $d \geq 2$.*

3.2. Approximation by using Charlier polynomials

In this subsection, we shall give the error analysis of approximations by Charlier polynomials, discrete orthogonal polynomials of Askey scheme. The derivative in the continuous version is replaced by the forward difference operator. Analogous Sobolev and Koborov norms and seminorms are properly defined in (3.42) and (3.45), respectively.

Let us denote the orthogonal projection operator $P_N^{\boldsymbol{a}} : l^2_{\boldsymbol{\omega}_{\boldsymbol{a}}}\left(\mathbb{N}_0^d\right) \to X_N^{\boldsymbol{a}}$, where

$$X_N^{\boldsymbol{a}} = \mathrm{span}\left\{\boldsymbol{C}_{\boldsymbol{n}}(\boldsymbol{x};\boldsymbol{a}) : \ \boldsymbol{n} \in \boldsymbol{\Omega}_N\right\}, \tag{3.40}$$

for certain index set $\boldsymbol{\Omega}_{N,\mathrm{tensor}}$, $\boldsymbol{\Omega}_{N,RHC}$ or $\boldsymbol{\Omega}_{N,OHC,\gamma}$ defined in (3.1).

Theorem 3.4 (tensor product with Charlier polynomials). *Assume that $\boldsymbol{\Omega}_N = \boldsymbol{\Omega}_{N,\mathrm{tensor}}$ in $X_N^{\boldsymbol{a}}$. Given $u \in \mathcal{W}^m_{\boldsymbol{a}}\left(\mathbb{N}_0^d\right)$, we have for any $0 \leq l < m$,*

$$|P_N^{\boldsymbol{a}}u - u|_{\mathcal{W}^l_{\boldsymbol{a}}(\mathbb{N}_0^d)} \leq \left(|\boldsymbol{a}|_\infty^{m-l} + d\frac{|\boldsymbol{a}|_\infty^m}{|\boldsymbol{a}|^l_{\mathrm{min}}}\right)^{\frac{1}{2}} (N-m+1)^{\frac{l-m}{2}} |u|_{\mathcal{W}^m_{\boldsymbol{a}}(\mathbb{N}_0^d)}, \tag{3.41}$$

for $N \gg 1$, where

$$|u|^2_{\mathcal{W}^m_{\boldsymbol{a}}(\mathbb{N}_0^d)} = \sum_{j=1}^d \left|\left|\triangle^m_{x_j} u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{a}},\mathbb{N}_0^d}. \tag{3.42}$$

If $u \in \mathcal{K}_{\boldsymbol{a}}^m\left(\mathbb{N}_0^d\right)$, *then for any* $0 \le \boldsymbol{l} < m$, *we have*

$$\left|\left|\triangle_{\boldsymbol{x}}^{\boldsymbol{l}}\left(P_N^{\boldsymbol{a}}u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{a}},\mathbb{N}_0^d} \le C_{\boldsymbol{a}} N^{\frac{|\boldsymbol{l}|_\infty - m}{2}} |u|_{\mathcal{K}_{\boldsymbol{a}}^m\left(\mathbb{N}_0^d\right)}, \tag{3.43}$$

where

$$C_{\boldsymbol{a}} = \begin{cases} |\boldsymbol{a}|_\infty^{\frac{dm-|\boldsymbol{l}|_{\min}}{2}}, & \text{if } |\boldsymbol{a}|_\infty \ge 1 \\ |\boldsymbol{a}|_\infty^{\frac{m-|\boldsymbol{l}|_1}{2}}, & \text{if } |\boldsymbol{a}|_\infty < 1 \end{cases}, \tag{3.44}$$

for $N \gg 1$, *and*

$$|u|^2_{\mathcal{K}_{\boldsymbol{a}}^r\left(\mathbb{N}_0^d\right)} = \sum_{|\boldsymbol{k}|_\infty = r} \left|\left|\triangle_{\boldsymbol{x}}^{\boldsymbol{k}} u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{a}},\mathbb{N}_0^d}. \tag{3.45}$$

Proof. Let us show the estimate (3.41) in Sobolev space first. Let us look at

$$|P_N^{\boldsymbol{a}}u - u|^2_{\mathcal{W}_{\boldsymbol{a}}^l\left(\mathbb{N}_0^d\right)} = \sum_{j=1}^d \left(\sum_{\boldsymbol{n}\in\Lambda_N^{1,j}} + \sum_{\boldsymbol{n}\in\Lambda_N^{2,j}}\right) \boldsymbol{\rho}_{\boldsymbol{n},l\boldsymbol{e}_j,\boldsymbol{a}} \left|\hat{u}_{\boldsymbol{n}}\right|^2 := \sum_{j=1}^d (I_1 + I_2), \tag{3.46}$$

where $\Lambda_N^{1,j}$ and $\Lambda_N^{2,j}$ are defined as (3.11) in the proof of Theorem 3.1, and

$$\boldsymbol{\rho}_{\boldsymbol{n},l\boldsymbol{e}_j,\boldsymbol{a}} \overset{\text{(B.21)}}{=} \prod_{i=1,i\neq j}^d \frac{n_i!}{a_i^{n_i}} \cdot \frac{\Gamma(n_j+1)^2 (n_j-l)!}{\Gamma(n_j-l+1)^2 a_j^{n_j+l}}.$$

Let us estimate the right-hand side of (3.46) term by term:

$$I_1 \le \max_{\boldsymbol{n}\in\Lambda_N^{1,j}} \left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n},l\boldsymbol{e}_j,\boldsymbol{a}}}{\boldsymbol{\rho}_{\boldsymbol{n},m\boldsymbol{e}_j,\boldsymbol{a}}}\right\} \sum_{\boldsymbol{n}\in\Lambda_N^{1,j}} \boldsymbol{\rho}_{\boldsymbol{n},m\boldsymbol{e}_j,\boldsymbol{a}} \left|\hat{u}_{\boldsymbol{n}}\right|^2, \tag{3.47}$$

with

$$\max_{\boldsymbol{n}\in\Lambda_N^{1,j}} \left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n},l\boldsymbol{e}_j,\boldsymbol{a}}}{\boldsymbol{\rho}_{\boldsymbol{n},m\boldsymbol{e}_j,\boldsymbol{a}}}\right\} = a_j^{m-l} \max_{\boldsymbol{n}\in\Lambda_N^{1,j}} \left\{\frac{1}{(n_j-m+1)\cdots(n_j-l)}\right\} \le |\boldsymbol{a}|_\infty^{m-l}(N-m+1)^{l-m}, \tag{3.48}$$

and

$$I_2 \le \max_{\boldsymbol{n}\in\Lambda_N^{2,j}} \left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n},l\boldsymbol{e}_j,\boldsymbol{a}}}{\boldsymbol{\rho}_{\boldsymbol{n},m\boldsymbol{e}_k,\boldsymbol{a}}}\right\} \sum_{\boldsymbol{n}\in\Lambda_N^{2,j}} \boldsymbol{\rho}_{\boldsymbol{n},m\boldsymbol{e}_k,\boldsymbol{a}} \left|\hat{u}_{\boldsymbol{n}}\right|^2, \tag{3.49}$$

with

$$\begin{aligned} \max_{\boldsymbol{n}\in\Lambda_N^{2,j}} \left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n},l\boldsymbol{e}_j,\boldsymbol{a}}}{\boldsymbol{\rho}_{\boldsymbol{n},m\boldsymbol{e}_k,\boldsymbol{a}}}\right\} &= \frac{a_k^m}{a_j^l} \max_{\boldsymbol{n}\in\Lambda_N^{2,j}} \left\{\frac{(n_j-l+1)\cdots n_j}{(n_k+m-1)\cdots n_k}\right\} \le \frac{a_k^m}{a_j^l}\frac{(N-l+1)\cdots N}{(N-m+1)\cdots N} \\ &\le \frac{|\boldsymbol{a}|_\infty^m}{|\boldsymbol{a}|_{\min}^l}(N-m+1)^{l-m}. \end{aligned} \tag{3.50}$$

The result follows from (3.46)-(3.50), (3.14) and (3.17).

To show the estimate (3.43) in Koborov space. Let us denote $\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{a}}$ similarly as in (3.19):

$$\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{a}} := \left(\prod_{j\in\mathcal{N}} \rho_{n_j,l_j,a_j}\right)\left(\prod_{i\in\mathcal{N}^c} \rho_{n_i,m,a_i}\right),$$

then

$$\left|\left|\triangle_{\boldsymbol{x}}^{\boldsymbol{l}}(P_N^{\boldsymbol{a}}u - u)\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{a}},\mathbb{N}_0^d} \le \max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^c} \left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},\boldsymbol{a}}}{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{a}}}\right\} \sum_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^c} \boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{a}} \left|\hat{u}_{\boldsymbol{n}}\right|^2, \tag{3.51}$$

due to the fact that

$$\max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^c}\left\{\frac{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},\boldsymbol{a}}}{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{a}}}\right\}=\max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^c}\left\{\prod_{j\in\mathcal{N}^c}a_j^{m-l_j}\prod_{j\in\mathcal{N}^c}\frac{1}{(n_j-m+1)\cdots(n_j-l_j)}\right\}$$
$$\overset{(3.21)}{\leq}\max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^c}\left\{\prod_{j\in\mathcal{N}^c}a_j^{m-l_j}\right\}(N-m+1)^{|\boldsymbol{l}|_\infty-m},$$

and

$$\max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^c}\left\{\prod_{j\in\mathcal{N}^c}a_j^{m-l_j}\right\}\leq\max_{\boldsymbol{n}\in\boldsymbol{\Omega}_{N,\text{tensor}}^c}\left\{\prod_{j\in\mathcal{N}^c}|\boldsymbol{a}|_\infty^{m-l_j}\right\}\leq\begin{cases}|\boldsymbol{a}|_\infty^{dm-|\boldsymbol{l}|_{\min}}, & \text{if } |\boldsymbol{a}|_\infty\geq 1\\ |\boldsymbol{a}|_\infty^{m-|\boldsymbol{l}|_1}, & \text{if } |\boldsymbol{a}|_\infty<1\end{cases}=:C_{\boldsymbol{a}}^2. \tag{3.52}$$

The result (3.43) follows immediately from (3.51)-(3.52) and (3.23). ☐

Theorem 3.5 (RHC with Charlier polynomials). *Assume the index set is $\boldsymbol{\Omega}_N=\boldsymbol{\Omega}_{N,RHC}$. Given $u\in\mathcal{K}_{\boldsymbol{a}}^m\left(\mathbb{N}_0^d\right)$, for $0\leq\boldsymbol{l}<m$, we have*

$$\left|\left|\triangle_{\boldsymbol{x}}^{\boldsymbol{l}}\left(P_N^{\boldsymbol{a}}u-u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{a}},\mathbb{N}_0^d}\leq C_{\boldsymbol{a}}m^{\frac{d(2m-|\boldsymbol{l}|_\infty)-|\boldsymbol{l}|_{\min}}{2}}N^{\frac{|\boldsymbol{l}|_\infty-m}{2}}|u|_{\mathcal{K}_{\boldsymbol{a}}^m\left(\mathbb{N}_0^d\right)},$$

where $C_{\boldsymbol{a}}$ is defined in (3.44).

Proof. We proceed the proof as that of Theorem 3.2. According to (B.21), we have

$$\frac{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},\boldsymbol{a}}}{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{l},m,\boldsymbol{a}}}=\prod_{j\in\mathcal{N}^c}a_j^{m-l_j}\prod_{j\in\mathcal{N}^c}n_j^{l_j-m}\prod_{j\in\mathcal{N}^c}\left(1-\frac{m-1}{n_j}\right)^{-1}\cdots\left(1-\frac{l_j}{n_j}\right)^{-1}. \tag{3.53}$$

The maximum of the three terms on the right-hand side of (3.53) have been obtained in (3.26), (3.27) and (3.52), respectively. The conclusion follows immediately from (3.51), (3.53) and (3.23).☐

Theorem 3.6 (OHC with Charlier polynomials). *Assume the index set is $\boldsymbol{\Omega}_N=\boldsymbol{\Omega}_{N,OHC,\gamma}$, for $\gamma\in(0,1)$. For any $u\in\mathcal{K}_{\boldsymbol{a}}^m\left(\mathbb{N}_0^d\right)$, $d\geq 2$, and $0\leq|\boldsymbol{l}|_1<m$,*

$$\begin{aligned}&\left|\left|\triangle_{\boldsymbol{x}}^{\boldsymbol{l}}\left(P_{N,\gamma}^{\boldsymbol{a}}u-u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{a}},\mathbb{N}_0^d}\\ \leq& C_{\boldsymbol{a}}m^{\frac{dm-|\boldsymbol{l}|_{\min}}{2}}|u|_{\mathcal{K}_{\boldsymbol{a}}^m\left(\mathbb{N}_0^d\right)}\cdot\begin{cases}m^{\frac{(d-1)(\gamma m-|\boldsymbol{l}|_1)}{2(1-\gamma)}}N^{\frac{|\boldsymbol{l}|_1-m}{2}}, & \text{if } 0<\gamma\leq\dfrac{|\boldsymbol{l}|_1}{m}\\ m^{\frac{d-1}{2(d-\gamma)}(dm-|\boldsymbol{l}|_1)}N^{-\frac{1-\gamma}{2(d-\gamma)}(dm-|\boldsymbol{l}|_1)}, & \text{if } \dfrac{|\boldsymbol{l}|_1}{m}\leq\gamma<1\end{cases},\end{aligned}$$

where $C_{\boldsymbol{a}}$ is defined in (3.44).

Proof. We can proceed as in the proof of Theorem 3.3. The maximum of the three terms on the right-hand side of (3.53) can be obtained by (3.52), (3.27) and (3.38)-(3.39), respectively. Thus, the result follows immediately. ☐

4. Results of the approximations by other orthogonal polynomials of Askey scheme

The error analysis for Laguerre polynomials and Charlier polynomials can be applied to various orthogonal polynomials of Askey scheme. In this section, we only state the results without proofs. All the Sobolev and Koborov norms and seminorms in the following theorems are defined similarly as in (3.4) and (3.6), respectively, with appropriate weight functions.

4.1. Hermite polynomials

Theorem 4.1 (tensor product with Hermite polynomials). *Assume that $X_N := \operatorname{span}\{\boldsymbol{H_n} :\ \boldsymbol{n} \in \boldsymbol{\Omega}_{N,\text{tensor}}\}$. Given $u \in \mathcal{W}^m\left(\mathbb{R}^d\right)$, we have for any $0 \le l \le m$,*

$$|P_N u - u|_{\mathcal{W}^l(\mathbb{R}^d)} \le (1+d)^{\frac{1}{2}} (N-m+1)^{\frac{l-m}{2}} |u|_{\mathcal{W}^m(\mathbb{R}^d)}, \tag{4.1}$$

for $N \gg 1$, where

$$|u|^2_{\mathcal{W}^m(\mathbb{R}^d)} = \sum_{j=1}^{d} \left|\left|\partial^m_{x_j} u\right|\right|^2_{\boldsymbol{\omega},\mathbb{R}^d}.$$

If $u \in \mathcal{K}^m\left(\mathbb{R}^d\right)$, then for $0 \le \boldsymbol{l} < m$, we have

$$\left|\left|\boldsymbol{\partial}^{\boldsymbol{l}}_{\boldsymbol{x}} \left(P_N u - u\right)\right|\right|_{\boldsymbol{\omega},\mathbb{R}^d} \le (N-m+1)^{\frac{|\boldsymbol{l}|_\infty - m}{2}} |u|_{\mathcal{K}^m(\mathbb{R}^d)}, \tag{4.2}$$

where

$$|u|^2_{\mathcal{K}^r(\mathbb{R}^d)} = \sum_{|\boldsymbol{k}|_\infty = r} \left|\left|\boldsymbol{\partial}^{\boldsymbol{k}}_{\boldsymbol{x}} u\right|\right|^2_{\boldsymbol{\omega},\mathbb{R}^d}.$$

Theorem 4.2 (RHC with Hermite polynomials). *Assume the index set is $\boldsymbol{\Omega}_N = \boldsymbol{\Omega}_{N,RHC}$. Given $u \in \mathcal{K}^m\left(\mathbb{R}^d\right)$, for $0 \le \boldsymbol{l} < m$, we have*

$$\left|\left|\boldsymbol{\partial}^{\boldsymbol{l}}_{\boldsymbol{x}} \left(P_N u - u\right)\right|\right|_{\boldsymbol{\omega},\mathbb{R}^d} \le m^{\frac{d(2m - |\boldsymbol{l}|_\infty) - |\boldsymbol{l}|_{\min}}{2}} N^{\frac{|\boldsymbol{l}|_\infty - m}{2}} |u|_{\mathcal{K}^m(\mathbb{R}^d)},$$

for $N \gg 1$.

Theorem 4.3 (OHC with Hermite polynomials). *Assume the index set is $\boldsymbol{\Omega}_N := \boldsymbol{\Omega}_{N,OHC,\gamma}$, for $\gamma \in (0,1)$. For any $u \in \mathcal{K}^m\left(\mathbb{R}^d\right)$, $d \ge 2$, and $0 \le |\boldsymbol{l}|_1 < m$,*

$$\begin{aligned}\left|\left|\boldsymbol{\partial}^{\boldsymbol{l}}_{\boldsymbol{x}} \left(P_{N,\gamma} u - u\right)\right|\right|_{\boldsymbol{\omega},\mathbb{R}^d} \le & m^{\frac{dm - |\boldsymbol{l}|_{\min}}{2}} |u|_{\mathcal{K}^m(\mathbb{R}^d)} \\ & \cdot \begin{cases} m^{\frac{(d-1)(\gamma m - |\boldsymbol{l}|_1)}{2(1-\gamma)}} N^{\frac{|\boldsymbol{l}|_1 - m}{2}}, & \text{if } 0 < \gamma \le \dfrac{|\boldsymbol{l}|_1}{m} \\ m^{\frac{d-1}{2(d-\gamma)}(dm - |\boldsymbol{l}|_1)} N^{\frac{1-\gamma}{2(d-\gamma)}(|\boldsymbol{l}|_1 - dm)}, & \text{if } \dfrac{|\boldsymbol{l}|_1}{m} \le \gamma < 1, \end{cases}.\end{aligned}$$

4.2. Jacobi polynomials

Theorem 4.4 (tensor product with Jacobi polynomials). *Assume that $X^{\boldsymbol{\alpha},\boldsymbol{\beta}}_N := \operatorname{span}\left\{\boldsymbol{P}^{(\boldsymbol{\alpha},\boldsymbol{\beta})}_{\boldsymbol{n}} :\ \boldsymbol{n} \in \boldsymbol{\Omega}_{N,\text{tensor}}\right\}$. Given $u \in \mathcal{W}^m_{\boldsymbol{\alpha},\boldsymbol{\beta}}\left(I^d\right)$, we have for any $0 \le l < m$,*

$$\left|\mathcal{P}^{\boldsymbol{\alpha},\boldsymbol{\beta}}_N u - u\right|_{\mathcal{W}^l_{\boldsymbol{\alpha},\boldsymbol{\beta}}(I^d)} \le (1+d)^{\frac{1}{2}} 2^{m-l} (N-m)^{l-m} ||u||_{\mathcal{W}^m_{\boldsymbol{\alpha},\boldsymbol{\beta}}(I^d)}, \tag{4.3}$$

for $N \gg 1$, where $\mathcal{P}^{\boldsymbol{\alpha},\boldsymbol{\beta}}_N$ denotes the orthogonal projection operator, and the Sobolev norm is defined as

$$||u||^2_{\mathcal{W}^m_{\boldsymbol{\alpha},\boldsymbol{\beta}}(I^d)} = \sum_{0 \le |\boldsymbol{k}|_1 \le m} \left|\left|\boldsymbol{\partial}^{\boldsymbol{k}}_{\boldsymbol{x}} u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k},\boldsymbol{\beta}+\boldsymbol{k}},I^d}.$$

If $u \in \mathcal{K}^m_{\boldsymbol{\alpha},\boldsymbol{\beta}}\left(I^d\right)$, then for any $0 \le \boldsymbol{l} < m$, we have

$$\left|\left|\boldsymbol{\partial}^{\boldsymbol{l}}_{\boldsymbol{x}} \left(\mathcal{P}^{\boldsymbol{\alpha},\boldsymbol{\beta}}_N u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{l},\boldsymbol{\beta}+\boldsymbol{l}},I^d} \le 2^{dm - |\boldsymbol{l}|_{\min}} (N-m)^{|\boldsymbol{l}|_\infty - m} |u|_{\mathcal{K}^m_{\boldsymbol{\alpha}+m,\boldsymbol{\beta}+m,I^d}}, \tag{4.4}$$

where

$$|u|^2_{\mathcal{K}^r_{\boldsymbol{\alpha},\boldsymbol{\beta}}(I^d)} = \sum_{|\boldsymbol{k}|_\infty = r} \left|\left|\boldsymbol{\partial}^{\boldsymbol{k}}_{\boldsymbol{x}} u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k},\boldsymbol{\beta}+\boldsymbol{k}},I^d}.$$

Theorem 4.5 (RHC with Jacobi polynomials). *Assume the index set is $\mathbf{\Omega}_N = \mathbf{\Omega}_{N,RHC}$. Given $u \in \mathcal{K}^m_{\boldsymbol{\alpha},\boldsymbol{\beta}}\left(I^d\right)$, for $0 \le \boldsymbol{l} < m$, we have*

$$\left|\left|\boldsymbol{\partial}^{\boldsymbol{l}}_{\boldsymbol{x}}\left(\mathcal{P}^{\boldsymbol{\alpha},\boldsymbol{\beta}}_N u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{l},\boldsymbol{\beta}+\boldsymbol{l}},I^d} \le C_{\boldsymbol{\alpha},\boldsymbol{\beta}} 2^{dm-|\boldsymbol{l}|_{\min}} m^{d(2m-|\boldsymbol{l}|_\infty)-|\boldsymbol{l}|_{\min}} N^{|\boldsymbol{l}|_\infty - m} |u|_{\mathcal{K}^m_{\boldsymbol{\alpha}+m,\boldsymbol{\beta}+m,I^d}},$$

where

$$C_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \begin{cases} (|\boldsymbol{\alpha}+\boldsymbol{\beta}|_{\min}+1)^{\frac{|\boldsymbol{l}|_1-m}{2}}, & \text{if } |\boldsymbol{\alpha}+\boldsymbol{\beta}|_{\min} \ge 0 \\ (|\boldsymbol{\alpha}+\boldsymbol{\beta}|_{\min}+1)^{\frac{|\boldsymbol{l}|_{\min}-dm}{2}}, & \text{if } |\boldsymbol{\alpha}+\boldsymbol{\beta}|_{\min} < 0 \end{cases}. \tag{4.5}$$

Theorem 4.6 (OHC with Jacobi polynomials). *Assume the index set is $\mathbf{\Omega}_N := \mathbf{\Omega}_{N,OHC,\gamma}$, for $\gamma \in (0,1)$. For any $u \in \mathcal{K}^m_{\boldsymbol{\alpha}+m,\boldsymbol{\beta}+m}\left(I^d\right)$, $d \ge 2$, and $0 \le |\boldsymbol{l}|_1 < m$,*

$$\begin{aligned}\left|\left|\boldsymbol{\partial}^{\boldsymbol{l}}_{\boldsymbol{x}}\left(\mathcal{P}^{\boldsymbol{\alpha},\boldsymbol{\beta}}_{N,\gamma} u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{l},\boldsymbol{\beta}+\boldsymbol{l}},I^d} \le & C_{\boldsymbol{\alpha},\boldsymbol{\beta}} 2^{dm-|\boldsymbol{l}|_{\min}} m^{d(2m-|\boldsymbol{l}|_\infty)-|\boldsymbol{l}|_{\min}} |u|_{\mathcal{K}^m_{\boldsymbol{\alpha}+m,\boldsymbol{\beta}+m,I^d}} \\ & \cdot \begin{cases} m^{\frac{(d-1)(\gamma m-|\boldsymbol{l}|_1)}{(1-\gamma)}} N^{|\boldsymbol{l}|_1-m}, & \text{if } 0 < \gamma \le \dfrac{|\boldsymbol{l}|_1}{m} \\ m^{\frac{d-1}{(d-\gamma)}(dm-|\boldsymbol{l}|_1)} N^{-\frac{1-\gamma}{(d-\gamma)}(dm-|\boldsymbol{l}|_1)}, & \text{if } \dfrac{|\boldsymbol{l}|_1}{m} \le \gamma < 1 \end{cases},\end{aligned}$$

where $C_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ is defined in (4.5).

Remark 4.4. *Compared with Laguerre-chaos and Hermite-chaos, the Jacobi-chaos has twice faster convergence rate. It is generally believed that the slower convergence rate of Laguerre-chaos and Hermite-chaos is due to their unbounded nature.*

4.3. Krawtchouk polynomials

Theorem 4.7 (tensor product with Krawtchouk polynomials). *Assume that $X^{\boldsymbol{p},\boldsymbol{N}}_M = \operatorname{span}\{\boldsymbol{K}_{\boldsymbol{n}}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N}) : \boldsymbol{n} \in \mathbf{\Omega}_{M,\text{tensor}}\}$, for some $\boldsymbol{N} \gg 1$ and $M \le \boldsymbol{N}$. Given $u \in \mathcal{W}^m_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})$, we have for any $0 \le l < m$,*

$$\begin{aligned}&\left|P^{\boldsymbol{p},\boldsymbol{N}}_N u - u\right|_{\mathcal{W}^l_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})} \\ \le & (1+d)^{\frac{1}{2}} 2^{-m} \left[|\boldsymbol{p}|_{\min}(1-|\boldsymbol{p}|_\infty)\right]^{-\frac{l}{2}} |\boldsymbol{N}|_\infty^{\frac{m}{2}} (|\boldsymbol{N}|_{\min} - l + 1)^{-\frac{l}{2}} (M-m+1)^{\frac{l-m}{2}} |u|_{\mathcal{W}^m_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})},\end{aligned}$$

for $M \gg 1$ and $M \le \boldsymbol{N}$, where

$$|u|^2_{\mathcal{W}^m_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})} = \sum_{j=1}^{d} \left|\left|\triangle^m_{x_j} u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{p},\boldsymbol{N}},\mathbb{N}_{\boldsymbol{N}}}.$$

Given $u \in \mathcal{K}^m_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})$, for $0 \le \boldsymbol{l} < m$, we have

$$\left|\left|\triangle^{\boldsymbol{l}}_{\boldsymbol{x}}\left(P^{\boldsymbol{p},\boldsymbol{N}}_N u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{p},\boldsymbol{N}},\mathbb{N}_{\boldsymbol{N}}} \le 2^{|\boldsymbol{l}|_1 - m} |\boldsymbol{N}|_\infty^{\frac{dm-|\boldsymbol{l}|_{\min}}{2}} (M-m+1)^{\frac{|\boldsymbol{l}|_\infty - m}{2}} |u|_{\mathcal{K}^m_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})},$$

for $M \gg 1$, where

$$|u|^2_{\mathcal{K}^r_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})} = \sum_{|\boldsymbol{k}|_\infty = r} \left|\left|\triangle^{\boldsymbol{k}}_{\boldsymbol{x}} u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{p},\boldsymbol{N}},\mathbb{N}_{\boldsymbol{N}}}.$$

Theorem 4.8 (RHC with Krawtchouk polynomials). *Assume the index set is $\mathbf{\Omega}_M = \mathbf{\Omega}_{M,RHC}$, for some $\boldsymbol{N} \gg 1$ and $M < |\boldsymbol{N}|_{\text{mix}}$. Given $u \in \mathcal{K}^m_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})$, for $0 \le \boldsymbol{l} < m$, we have*

$$\left|\left|\triangle^{\boldsymbol{l}}_{\boldsymbol{x}}\left(P^{\boldsymbol{p},\boldsymbol{N}}_N u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{p},\boldsymbol{N}},\mathbb{N}_{\boldsymbol{N}}} \le 2^{|\boldsymbol{l}|_1 - m} |\boldsymbol{N}|_\infty^{\frac{dm-|\boldsymbol{l}|_{\min}}{2}} m^{\frac{d(2m-|\boldsymbol{l}|_\infty)-|\boldsymbol{l}|_{\min}}{2}} M^{\frac{|\boldsymbol{l}|_\infty - m}{2}} |u|_{\mathcal{K}^m_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})},$$

for $M \gg 1$.

Theorem 4.9 (OHC with Krawtchouk polynomials). *Assume the index set is* $\mathbf{\Omega}_M := \mathbf{\Omega}_{M,OHC,\gamma}$, *for* $\gamma \in (0,1)$, $\boldsymbol{N} \gg 1$ *and* $M < |\boldsymbol{N}|_\infty^{\frac{d-\gamma}{1-\gamma}}$. *For any* $u \in \mathcal{K}^m_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})$, $d \geq 2$, *and* $0 \leq |\boldsymbol{l}|_1 < m$,

$$
\begin{aligned}
&\left|\left|\triangle_{\boldsymbol{x}}^{\boldsymbol{l}}\left(P_{N,\gamma}^{\boldsymbol{p},\boldsymbol{N}}u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{p},\boldsymbol{N}},\mathbb{N}_{\boldsymbol{N}}} \\
\leq& 2^{|\boldsymbol{l}|_1-m}(m|\boldsymbol{N}|_\infty)^{\frac{dm-|\boldsymbol{l}|_{\min}}{2}}|u|_{\mathcal{K}^m_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})}\begin{cases} m^{\frac{(d-1)[(\gamma+1)m-2|\boldsymbol{l}|_1]}{2(1-\gamma)}}M^{\frac{|\boldsymbol{l}|_1-m}{2}}, & \text{if } 0<\gamma\leq\dfrac{|\boldsymbol{l}|_1}{m} \\ m^{\frac{d-1}{2(d-\gamma)}(dm-|\boldsymbol{l}|_1)}M^{-\frac{1-\gamma}{2(d-\gamma)}(dm-|\boldsymbol{l}|_1)}, & \text{if } \dfrac{|\boldsymbol{l}|_1}{m}\leq\gamma<1 \end{cases}.
\end{aligned}
$$

4.4. Meixner polynomials

Theorem 4.10 (tensor product with Meixner polynomials). *Assume that* $X_N^{\boldsymbol{\beta},\boldsymbol{c}} := \operatorname{span}\{\boldsymbol{M}_{\boldsymbol{n}}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c}) : \boldsymbol{n} \in \mathbf{\Omega}_N\}$. *Given* $u \in \mathcal{W}^m_{\boldsymbol{\beta},\boldsymbol{c}}(\mathbb{N}_0^d)$, *we have for any* $0 \leq l < m$,

$$
\begin{aligned}
\left|P_N^{\boldsymbol{\beta},\boldsymbol{c}}u - u\right|_{\mathcal{W}^l_{\boldsymbol{\beta},\boldsymbol{c}}(\mathbb{N}_0^d)} \leq& \left\{\left[\frac{(|\boldsymbol{c}|_\infty-1)^2}{|\boldsymbol{c}|_\infty}\right]^l(|\boldsymbol{\beta}|_\infty+m-1)^{-l} + d\left[\frac{(|\boldsymbol{c}|_{\min}-1)^2}{|\boldsymbol{c}|_{\min}}\right]^l|\boldsymbol{\beta}|_{\min}^{-l}\right\}^{\frac{1}{2}} \\
&\cdot\left[\frac{(|\boldsymbol{c}|_\infty-1)^2}{|\boldsymbol{c}|_\infty}\right]^{-\frac{m}{2}}(|\boldsymbol{\beta}|_\infty+m-1)^{\frac{m}{2}}(N-m+1)^{\frac{l-m}{2}}|u|_{\mathcal{W}^m_{\boldsymbol{\beta},\boldsymbol{c}}(\mathbb{N}_0^d)},
\end{aligned}
$$

for $N \gg 1$, *where*

$$
|u|^2_{\mathcal{W}^m_{\boldsymbol{\beta},\boldsymbol{c}}(\mathbb{N}_0^d)} = \sum_{j=1}^d\left|\left|\triangle^m_{x_j}u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\beta},\boldsymbol{c}},\mathbb{N}_0^d}.
$$

Given $u \in \mathcal{K}^m_{\boldsymbol{\beta},\boldsymbol{c}}(\mathbb{N}_0^d)$, *for* $0 \leq \boldsymbol{l} < m$, *we have*

$$
\left|\left|\triangle_{\boldsymbol{x}}^{\boldsymbol{l}}\left(P_N^{\boldsymbol{\beta},\boldsymbol{c}}u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{\beta},\boldsymbol{c}},\mathbb{N}_0^d} \leq C_{\boldsymbol{c}}C_{\boldsymbol{\beta}}N^{\frac{|\boldsymbol{l}|_\infty-m}{2}}|u|_{\mathcal{K}^m_{\boldsymbol{\beta},\boldsymbol{c}}(\mathbb{N}_0^d)},
$$

for $N \gg 1$, *where*

$$
C_{\boldsymbol{c}}^2 := \begin{cases} \left[\dfrac{(|\boldsymbol{c}|_\infty-1)^2}{|\boldsymbol{c}|_\infty}\right]^{|\boldsymbol{l}|_{\min}-dm}, & \text{if } \left[\dfrac{(|\boldsymbol{c}|_\infty-1)^2}{|\boldsymbol{c}|_\infty}\right]\leq 1 \\ \left[\dfrac{(|\boldsymbol{c}|_\infty-1)^2}{|\boldsymbol{c}|_\infty}\right]^{m-|\boldsymbol{l}|_1}, & \text{if } \left[\dfrac{(|\boldsymbol{c}|_\infty-1)^2}{|\boldsymbol{c}|_\infty}\right]<1 \end{cases}, \tag{4.6}
$$

$$
C_{\boldsymbol{\beta}}^2 := \begin{cases} (|\boldsymbol{\beta}|_\infty+m-1)^{dm-|\boldsymbol{l}|_{\min}}, & \text{if } |\boldsymbol{\beta}|_\infty+m\geq 2 \\ (|\boldsymbol{\beta}|_\infty+m-1)^{m-|\boldsymbol{l}|_1}, & \text{if } |\boldsymbol{\beta}|_\infty+m<2 \end{cases}, \tag{4.7}
$$

and

$$
|u|^2_{\mathcal{K}^r_{\boldsymbol{\beta},\boldsymbol{c}}(\mathbb{N}_0^d)} = \sum_{|\boldsymbol{k}|_\infty=r}\left|\left|\triangle_{\boldsymbol{x}}^{\boldsymbol{k}}u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\beta},\boldsymbol{c}},\mathbb{N}_0^d}.
$$

Theorem 4.11 (RHC with Meixner polynomials). *Assume the index set is* $\mathbf{\Omega}_N = \mathbf{\Omega}_{N,RHC}$. *Given* $u \in \mathcal{K}^m_{\boldsymbol{\beta},\boldsymbol{c}}(\mathbb{N}_0^d)$, *for* $0 \leq \boldsymbol{l} < m$, *we have*

$$
\left|\left|\triangle_{\boldsymbol{x}}^{\boldsymbol{l}}\left(P_N^{\boldsymbol{\beta},\boldsymbol{c}}u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{\beta},\boldsymbol{c}},\mathbb{N}_0^d} \leq C_{\boldsymbol{c}}C_{\boldsymbol{\beta}}m^{\frac{d(2m-|\boldsymbol{l}|_\infty)-|\boldsymbol{l}|_{\min}}{2}}N^{\frac{|\boldsymbol{l}|_\infty-m}{2}}|u|_{\mathcal{K}^m_{\boldsymbol{\beta},\boldsymbol{c}}(\mathbb{N}_0^d)},
$$

for $N \gg 1$, *where* $C_{\boldsymbol{c}}$ *and* $C_{\boldsymbol{\beta}}$ *are defined in* (4.6) *and* (4.7), *respectively.*

Theorem 4.12 (OHC with Meixner polynomials). *Assume the index set is* $\mathbf{\Omega}_N := \mathbf{\Omega}_{N,OHC,\gamma}$, *for* $\gamma \in (0,1)$. *For any* $u \in \mathcal{K}^m_{\boldsymbol{\beta},\boldsymbol{c}}(\mathbb{N}_0^d)$, $d \geq 2$, *and* $0 \leq |\boldsymbol{l}|_1 < m$,

$$\begin{aligned}
&\left|\left|\triangle^{\boldsymbol{l}}_{\boldsymbol{x}}\left(P^{\boldsymbol{\beta},\boldsymbol{c}}_{N,\gamma}u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{\beta},\boldsymbol{c}},\mathbb{N}_0^d} \\
\leq & C_{\boldsymbol{c}} C_{\boldsymbol{\beta}} m^{\frac{dm-|\boldsymbol{l}|_{\min}}{2}} |u|_{\mathcal{K}^m_{\boldsymbol{p},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})} \begin{cases} m^{\frac{(d-1)(\gamma m-|\boldsymbol{l}|_1)}{2(1-\gamma)}} M^{\frac{|\boldsymbol{l}|_1-m}{2}}, & \text{if } 0 < \gamma \leq \dfrac{|\boldsymbol{l}|_1}{m} \\ m^{\frac{d-1}{2(d-\gamma)}(dm-|\boldsymbol{l}|_1)} M^{-\frac{1-\gamma}{2(d-\gamma)}(dm-|\boldsymbol{l}|_1)}, & \text{if } \dfrac{|\boldsymbol{l}|_1}{m} \leq \gamma < 1 \end{cases},
\end{aligned}$$

where $C_{\boldsymbol{c}}$ *and* $C_{\boldsymbol{\beta}}$ *are defined in* (4.6) *and* (4.7)*, respectively.*

4.5. Hahn polynomials

Theorem 4.13 (tensor product with Hahn polynomials). *Assume that* $X^{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}_M = \mathrm{span}\left\{\boldsymbol{Q}_{\boldsymbol{n}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}) :\ \boldsymbol{n} \in \mathbf{\Omega}_M\right\}$, *for some* $\boldsymbol{N} \gg 1$ *and* $M \leq \boldsymbol{N}$. *Given* $u \in \mathcal{W}^m_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})$, *we have for any* $0 \leq l < m$,

$$\left|P^{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}_N u - u\right|_{\mathcal{W}^l_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})} \leq (1+d)^{\frac{1}{2}} \left[-\max\{|\boldsymbol{\alpha}|_{\min}, |\boldsymbol{\beta}|_{\min}\} - 1\right]^{\frac{m-l}{2}} (M-m+1)^{\frac{l-m}{2}} ||u||_{\mathcal{W}^m_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})},$$

for $M \gg 1$ *and* $M \leq \boldsymbol{N}$, *where*

$$|u|^2_{\mathcal{W}^m_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})} = \sum_{j=1}^d \left|\left|\triangle^m_{x_j} u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}},\mathbb{N}_{\boldsymbol{N}}}, \quad ||u||^2_{\mathcal{W}^m_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})} = \sum_{0\leq|\boldsymbol{k}|_1\leq m} \left|\left|\triangle^{\boldsymbol{k}}_{\boldsymbol{x}} u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}},\mathbb{N}_{\boldsymbol{N}}}.$$

Given $u \in \mathcal{K}^m_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})$, *for* $0 \leq \boldsymbol{l} < m$, *we have*

$$\begin{aligned}
&\left|\left|\triangle^{\boldsymbol{l}}_{\boldsymbol{x}}\left(P^{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}_N u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}},\mathbb{N}_{\boldsymbol{N}}} \\
\leq & \left[-\max\{|\boldsymbol{\alpha}|_{\min}, |\boldsymbol{\beta}|_{\min}\} - |\boldsymbol{l}|_{\min} - 1\right]^{\frac{d(m-l)}{2}} (M-m+1)^{\frac{|\boldsymbol{l}|_\infty - m}{2}} |u|_{\mathcal{K}^m_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})},
\end{aligned}$$

for $M \gg 1$, *where*

$$|u|^2_{\mathcal{K}^r_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})} = \sum_{|\boldsymbol{k}|_\infty = r} \left|\left|\triangle^{\boldsymbol{k}}_{\boldsymbol{x}} u\right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}},\mathbb{N}_{\boldsymbol{N}}}.$$

Theorem 4.14 (RHC with Hahn polynomials). *Assume the index set is* $\mathbf{\Omega}_M = \mathbf{\Omega}_{M,RHC}$, *for some* $\boldsymbol{N} \gg 1$ *and* $M < |\boldsymbol{N}|_{\mathrm{mix}}$. *Given* $u \in \mathcal{K}^m_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})$, *for* $0 \leq \boldsymbol{l} < m$, *we have*

$$\begin{aligned}
&\left|\left|\triangle^{\boldsymbol{l}}_{\boldsymbol{x}}\left(P^{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}_N u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}},\mathbb{N}_{\boldsymbol{N}}} \\
\leq & \left[-\max\{|\boldsymbol{\alpha}|_{\min}, |\boldsymbol{\beta}|_{\min}\} - |\boldsymbol{l}|_{\min} - 1\right]^{\frac{dm-|\boldsymbol{l}|_{\min}}{2}} m^{\frac{d(2m-|\boldsymbol{l}|_\infty)-|\boldsymbol{l}|_{\min}}{2}} M^{\frac{|\boldsymbol{l}|_\infty - m}{2}} |u|_{\mathcal{K}^m_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})},
\end{aligned}$$

for $M \gg 1$.

Theorem 4.15 (OHC with Hahn polynomials). *Assume the index set is* $\mathbf{\Omega}_M := \mathbf{\Omega}_{M,OHC,\gamma}$, *for* $\gamma \in (0,1)$, $\boldsymbol{N} \gg 1$ *and* $M < |\boldsymbol{N}|_\infty^{\frac{d-\gamma}{1-\gamma}}$. *For any* $u \in \mathcal{K}^m_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})$, $d \geq 2$, *and* $0 \leq |\boldsymbol{l}|_1 < m$,

$$\begin{aligned}
\left|\left|\triangle^{\boldsymbol{l}}_{\boldsymbol{x}}\left(P^{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}_{M,\gamma} u - u\right)\right|\right|_{\boldsymbol{\omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}},\mathbb{N}_{\boldsymbol{N}}} \leq & \left\{\left[-\max\{|\boldsymbol{\alpha}|_{\min}, |\boldsymbol{\beta}|_{\min}\} - |\boldsymbol{l}|_{\min} - 1\right] m\right\}^{\frac{dm-|\boldsymbol{l}|_{\min}}{2}} |u|_{\mathcal{K}^m_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}(\mathbb{N}_{\boldsymbol{N}})} \\
& \cdot \begin{cases} m^{\frac{(d-1)(\gamma m-|\boldsymbol{l}|_1)}{2(1-\gamma)}} M^{\frac{|\boldsymbol{l}|_1-m}{2}}, & \text{if } 0 < \gamma \leq \dfrac{|\boldsymbol{l}|_1}{m} \\ m^{\frac{d-1}{2(d-\gamma)}(dm-|\boldsymbol{l}|_1)} M^{-\frac{1-\gamma}{2(d-\gamma)}(dm-|\boldsymbol{l}|_1)}, & \text{if } \dfrac{|\boldsymbol{l}|_1}{m} \leq \gamma < 1 \end{cases}.
\end{aligned}$$

5. Application to Galerkin method of differential equations with random inputs

Let us recall the Galerkin method to solve differential equations with random inputs:

$$\mathcal{L}u(\boldsymbol{x}, t, \theta) = f(\boldsymbol{x}, t, \theta), \tag{5.8}$$

where $\mathcal{L}$ is differential operators in time/space, $u(\boldsymbol{x}, t, \theta)$ is the solution and $f(\boldsymbol{x}, t, \theta)$ is the source term. θ is the random parameter to describe the uncertainty of the system, which may be introduced via initial condition, boundary condition etc. The solution can be viewed as a random process with the Wiener-Askey polynomial chaos

$$u(\boldsymbol{x}, t, \theta) = \sum_{\boldsymbol{i} \in \boldsymbol{\Omega}_N} u_{\boldsymbol{i}}(\boldsymbol{x}, t) \boldsymbol{\Phi}_{\boldsymbol{i}}(\boldsymbol{Z}(\theta)), \tag{5.9}$$

where $\boldsymbol{Z} = (Z_1, \cdots, Z_d) \in \mathbb{R}^d$ are d independent random variables, $\boldsymbol{\Phi}_{\boldsymbol{i}} = \prod_{j=1}^d \Phi_{i_j}$ are the multi-variable Askey polynomials, and the summation is over certain index set $\boldsymbol{\Omega}_N$. Thus, it is clear that the total number of expansions in (5.9) depends on the cardinality of $\boldsymbol{\Omega}_N$.

Substituting the expansion (5.9) into (5.8), we obtain

$$\mathcal{L}\left(\sum_{\boldsymbol{i} \in \boldsymbol{\Omega}_N} u_{\boldsymbol{i}}(\boldsymbol{x}, t) \boldsymbol{\Phi}_{\boldsymbol{i}}(\boldsymbol{Z}(\theta)) \right) = f(\boldsymbol{x}, t, \theta).$$

The Galerkin spectral method is to project the above equation onto the linear subspace spanned by $\{\boldsymbol{\Phi}_{\boldsymbol{i}}\}_{\boldsymbol{i} \in \boldsymbol{\Omega}_N}$:

$$\left\langle \mathcal{L}\left(\sum_{\boldsymbol{i} \in \boldsymbol{\Omega}_N} u_{\boldsymbol{i}}(\boldsymbol{x}, t) \boldsymbol{\Phi}_{\boldsymbol{i}}(\boldsymbol{Z}(\theta)) \right), \boldsymbol{\Phi}_{\boldsymbol{k}} \right\rangle_{\boldsymbol{\omega}} = \langle f(\boldsymbol{x}, t, \theta), \boldsymbol{\Phi}_{\boldsymbol{k}} \rangle_{\boldsymbol{\omega}}, \tag{5.10}$$

for all $\boldsymbol{k} \in \boldsymbol{\Omega}_N$. According to the orthogonality of the polynomials, we shall arrive at a set of $\mathrm{card}(\boldsymbol{\Omega}_N)$ possibly coupled equations for $u_{\boldsymbol{i}}$s, $\boldsymbol{i} \in \boldsymbol{\Omega}_N$. It is easy to notice that the govening equation of $u_{\boldsymbol{i}}$ are deterministic. And all sorts of deterministic numerical schemes are applicable.

In the sequel, we shall consider the analogue ordinary differential equations investigated in [24] with higher-dimensional random inputs, i.e. $\boldsymbol{Z} \in \mathbb{R}^d$ with $d \geq 2$:

$$\frac{dy(t, \theta)}{dt} = -|\boldsymbol{Z}(\theta)|_1 y(t), \tag{5.11}$$

with the deterministic initial condition $y(0)$. It is easy to see that the solution to this ordinary differential equations with random inputs is

$$y(t) = y(0) e^{-|\boldsymbol{Z}|_1 t}.$$

Suppose $\boldsymbol{Z}$ are continuous random variables and the probability density of $\boldsymbol{Z}$ is known to be $f(\boldsymbol{z})$, then the mean of the stochastic solution is

$$\mathbb{E}[y](t) = y(0) \int_S e^{-|\boldsymbol{z}|_1 t} f(\boldsymbol{z}) d\boldsymbol{z},$$

where S is the support of the density function $f(\boldsymbol{z})$. If $\boldsymbol{Z}$ are discrete and the probability distribution $\mathbb{P}(\boldsymbol{Z} = \boldsymbol{z}_j) = p_j$, where $\boldsymbol{z}_j \in \mathbb{R}^d$, then

$$\mathbb{E}[y](t) = y(0) \sum_j e^{-|\boldsymbol{z}_j|_1 t} p_j.$$

where j sums over the support of the distribution. The Askey-chaos expansion is written as

$$y_N(t) = \sum_{\boldsymbol{i} \in \boldsymbol{\Omega}_N} y_{\boldsymbol{i}}(t) \boldsymbol{\Phi}_{\boldsymbol{i}}(\boldsymbol{Z}), \tag{5.12}$$

then testing with $\boldsymbol{\Phi}_{\boldsymbol{k}}$, $\boldsymbol{k} \in \boldsymbol{\Omega}_N$, we obtain

$$y'_{\boldsymbol{k}}(t) = -\left\langle \sum_{\boldsymbol{i}\in\boldsymbol{\Omega}_N} y_{\boldsymbol{i}}(t)|\boldsymbol{Z}|_1 \boldsymbol{\Phi}_{\boldsymbol{i}}(\boldsymbol{Z}), \boldsymbol{\Phi}_{\boldsymbol{k}}(\boldsymbol{Z}) \right\rangle_{\boldsymbol{\omega}} = \sum_{\boldsymbol{i}\in\boldsymbol{\Omega}_N} A_{\boldsymbol{k},\boldsymbol{i}} y_{\boldsymbol{i}}(t), \tag{5.13}$$

for all $\boldsymbol{k} \in \boldsymbol{\Omega}_N$, where $A_{\boldsymbol{k},\boldsymbol{i}} \in \mathbb{R}^{\mathrm{card}(\boldsymbol{\Omega}_N)}$ are from the orthogonality of the polynomials. In the following experiments, (5.13) is solved by *ode45* in Matlab. We define the error:

$$\epsilon_{\max} = \frac{\max_{t\in[0,T]} |y_0(t) - \mathbb{E}[y](t)|}{\max_{t\in[0,T]} |\mathbb{E}[y](t)|}, \quad \epsilon_{L^2} = \frac{||y_0(t) - \mathbb{E}[y](t)||_{L^2([0,T])}}{||\mathbb{E}[y](t)||_{L^2([0,T])}}. \tag{5.14}$$

Due to the similarity of Askey polynomials, we shall only solve (5.11) with Laguerre-chaos under the assumption of gamma distribution, Charlier-chaos with Possion distribution, and Hermite-chaos with Gaussian distribution.

5.1. Gamma distribution and Laguerre-chaos

Assume that $\boldsymbol{Z} = (Z_1, \cdots, Z_d)$ obeys i.i.d. gamma distributions with the parameter $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_d)$. The expectation of the stochastic solution is

$$\mathbb{E}[y](t) = \frac{y(0)}{(t\boldsymbol{\beta} + 1)^{\boldsymbol{\alpha}+1}}.$$

The Laguerre-chaos is naturally employed, i.e. $\boldsymbol{\Phi}_{\boldsymbol{i}} = \boldsymbol{L}_{\boldsymbol{i}}^{(\boldsymbol{\alpha})}$. Let us derive the matrix $A_{\boldsymbol{k},\boldsymbol{i}}$ in (5.13). According to (5.11), we have

$$\begin{aligned}
y'(t) &= \sum_{\boldsymbol{i}\in\boldsymbol{\Omega}_N} y'_{\boldsymbol{i}}(t)\boldsymbol{L}_{\boldsymbol{i}}^{(\boldsymbol{\alpha})}(\boldsymbol{Z}) = -\sum_{\boldsymbol{i}\in\boldsymbol{\Omega}_N} y_{\boldsymbol{i}}(t)\left[|\boldsymbol{Z}|_1 \boldsymbol{L}_{\boldsymbol{i}}^{(\boldsymbol{\alpha})}(\boldsymbol{Z})\right] \\
&= -\sum_{\boldsymbol{i}\in\boldsymbol{\Omega}_N} y_{\boldsymbol{i}}(t) \sum_{j=1}^{d} \left[L_{i_1}^{(\alpha_1)}(Z_1)\cdots\left(Z_j L_{i_j}^{(\alpha_j)}(Z_j)\right)\cdots L_{i_d}^{(\alpha_d)}(Z_d)\right] \\
&= -\sum_{\boldsymbol{i}\in\boldsymbol{\Omega}_N} y_{\boldsymbol{i}}(t) \sum_{j=1}^{d} \Big[L_{i_1}^{(\alpha_1)}(Z_1)\cdots\Big(-(i_j+1)L_{i_j+1}^{(\alpha_j)}(Z_j) + (2i_j+\alpha_j+1)L_{i_j}^{(\alpha_j)}(Z_j) - (i_j+\alpha_j)L_{i_j-1}^{(\alpha_j)}(Z_j)\Big) \\
&\qquad\qquad\qquad \cdots L_{i_d}^{(\alpha_d)}(Z_d)\Big] \\
&= -\sum_{\boldsymbol{i}\in\boldsymbol{\Omega}_N} y_{\boldsymbol{i}}(t) \sum_{j=1}^{d} \left[-(i_j+1)\boldsymbol{L}_{\boldsymbol{i}+\boldsymbol{e}_j}^{(\boldsymbol{\alpha})}(\boldsymbol{Z}) + (2i_j+\alpha_j+1)\boldsymbol{L}_{\boldsymbol{i}}^{(\boldsymbol{\alpha})}(\boldsymbol{Z}) - (i_j+\alpha_j)\boldsymbol{L}_{\boldsymbol{i}-\boldsymbol{e}_j}^{(\boldsymbol{\alpha})}(\boldsymbol{Z})\right].
\end{aligned} \tag{5.15}$$

Testing $\boldsymbol{L}_{\boldsymbol{k}}^{(\boldsymbol{\alpha})}(\boldsymbol{Z})$ on both sides, we obtain that

$$y'_{\boldsymbol{k}}(t) = \sum_{j=1}^{d} \left[k_j y_{\boldsymbol{k}-\boldsymbol{e}_j}(t) - (2k_j+\alpha_j+1)y_{\boldsymbol{k}}(t) + (k_j+\alpha_j+1)y_{\boldsymbol{k}+\boldsymbol{e}_j}\right] =: \sum_{\boldsymbol{i}\in\boldsymbol{\Omega}_N} A_{\boldsymbol{k},\boldsymbol{i}} y_{\boldsymbol{i}}(t). \tag{5.16}$$

If we write $y_{\boldsymbol{k}}$, $\boldsymbol{k} \in \boldsymbol{\Omega}_N$, in the vector form, then the above equation can be writtern in matrix form, i.e. $\vec{y}'(t) = A\vec{y}$. The matrix A is sparse. Figure 1 displays N v.s. the errors of solving (5.11) by Laguerre-chaos in different dimensions from 2 to 5. The *ode45* in Matlab has been used to numerically solve (5.16) with initial condition $y(0) = 1$. The time step is around 10^{-2}. Thus, one can't expect more accurate than 10^{-8}, due to the time marching error. Figure 1 clearly shows that the log of the error is almost linear with respect to N in OHC approximation. Although the number of basis of OHC with $\gamma = 0.5$ is significantly fewer than that of RHC (see Table 1 for Hermite-chaos), but the convergence rate of the errors are slower with RHC.

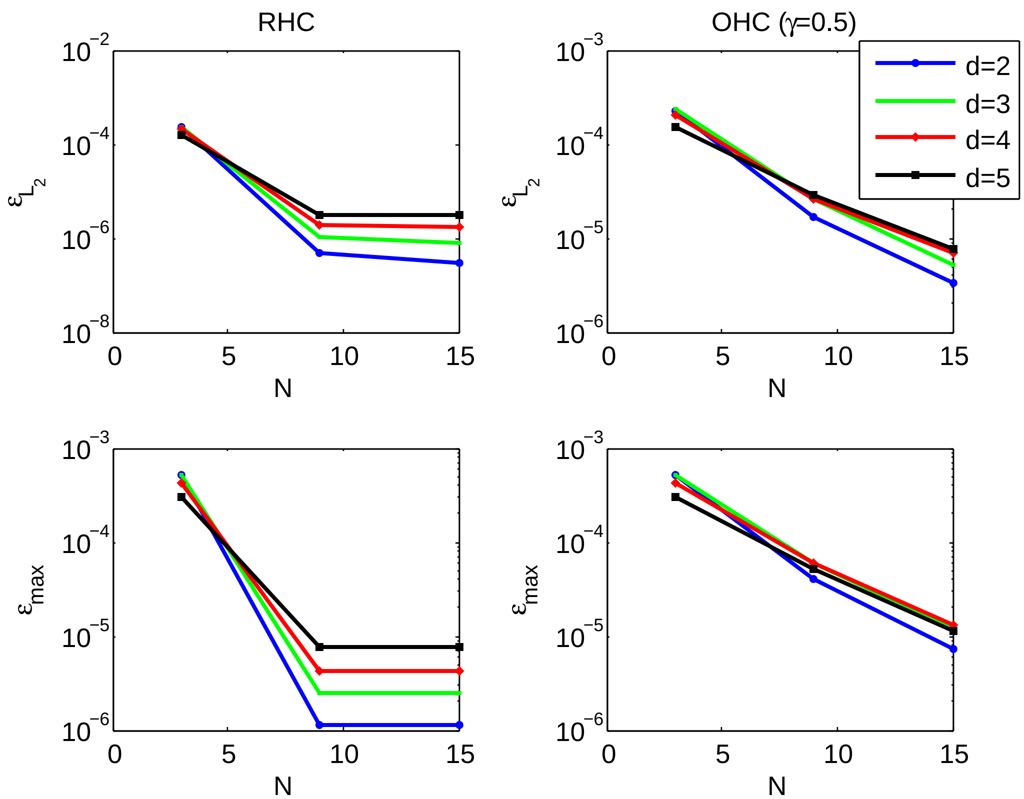

Figure 1: The $N = 3, 9, 15$ in $\mathbf{\Omega}_N$ v.s. the errors of Laguerre-chaos (including $\epsilon_{\max}$ and ϵ_{L^2} defined in (5.14)) are displayed for different dimensions. The plots corresponding to $\mathbf{\Omega}_{N,RHC}$ are in the left column, while those to $\mathbf{\Omega}_{N,OHC,0.5}$ are in the right column.

5.2. Poisson distribution and Charlier-chaos

We assume that $\boldsymbol{Z} = (Z_1, \cdots, Z_d)$, where Z_is obey i.i.d. Poisson distribution $\pi(a_i)$, $i = 1, \cdots, d$, i.e. the probability distribution is $f(\boldsymbol{z}) = \sum_{\boldsymbol{i}=0}^{\infty} e^{-|\boldsymbol{a}|_1} \frac{\boldsymbol{a}^{\boldsymbol{z}}}{\boldsymbol{z}!}$, where $\boldsymbol{a} = (a_i)_{i=1}^d$ and $\boldsymbol{z} = (z_i)_{i=1}^d$. The mean of the stochastic solution is

$$\mathbb{E}[y](t) = y(0) \exp\left[-(1 - e^{-t})|\boldsymbol{a}|_1\right].$$

The Charlier-chaos is naturally employed, i.e. $\mathbf{\Phi}_{\boldsymbol{i}} = \boldsymbol{C}_{\boldsymbol{i}}(\cdot; \boldsymbol{a})$. With similar argument in (5.15), $A_{\boldsymbol{k},\boldsymbol{i}}$ with Charlier polynomials is given by

$$y'_{\boldsymbol{k}}(t) = \sum_{j=1}^{d} \left[a_j y_{\boldsymbol{k}-\boldsymbol{e}_j}(t) - (k_j + a_j) y_{\boldsymbol{k}}(t) + (k_j + 1) y_{\boldsymbol{k}+\boldsymbol{e}_j}(t)\right], \tag{5.17}$$

for $\boldsymbol{k} \in \mathbf{\Omega}_N$. Figure 2 displays N v.s. the errors of solving (5.11) by Charlier-chaos in different dimensions from 2 to 5. Similar conclusions as those from Figure 1 can also be drawn from Figure 2.

5.3. Gaussian distribution and Hermite-chaos

We assume that $\boldsymbol{Z} = (Z_1, \cdots, Z_d)$, where Z_is obey i.i.d. Gaussian distribution $\mathcal{N}(0, 1)$. The mean of the stochastic solution is

$$\mathbb{E}[y](t) = y(0) \exp\left(-\frac{t^2 d}{2}\right).$$

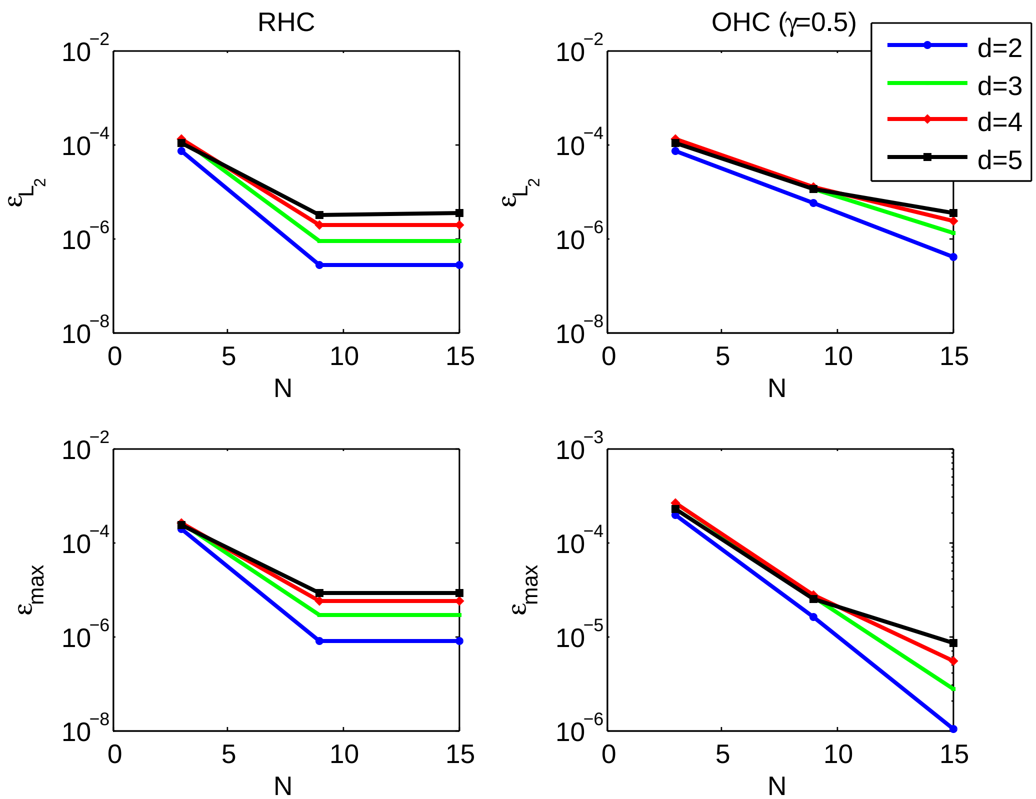

Figure 2: The $N = 3, 9, 15$ in $\boldsymbol{\Omega}_N$ v.s. the errors of Charlier-chaos (including $\epsilon_{\max}$ and ϵ_{L^2} defined in (5.14)) are displayed for different dimensions. The plots corresponding to $\boldsymbol{\Omega}_{N,RHC}$ are in the left column, while those to $\boldsymbol{\Omega}_{N,OHC,0.5}$ are in the right column.

With similar argument in (5.15), $A_{\boldsymbol{k},\boldsymbol{i}}$ with Hermite polynomials is given by

$$y'_{\boldsymbol{k}}(t) = \sum_{j=1}^{d} \left[-y_{\boldsymbol{k}-\boldsymbol{e}_j}(t) - (k_j+1) y_{\boldsymbol{k}+\boldsymbol{e}_j} \right] =: \sum_{\boldsymbol{i} \in \boldsymbol{\Omega}_N} A_{\boldsymbol{k},\boldsymbol{i}} y_{\boldsymbol{i}}(t). \tag{5.18}$$

The numbers of basis functions $H_{\boldsymbol{i}}$, $\boldsymbol{i} \in \boldsymbol{\Omega}_N$, and the numbers of nonzero elements in the matrix A are displayed in Table 1. It is clear to see that OHC approximations are with less basis functions and nonzeros elements.

dimension		2	3	4
RHC	# of basis	172	700	2453
	nonzero elements	564	3168	14024
OHC with $\gamma = 0.5$	# of basis	132	428	1232
	nonzero elements	404	1776	6432

Table 1: The numbers of basis functions $H_{\boldsymbol{i}}$, $\boldsymbol{i} \in \boldsymbol{\Omega}_N$, with $N = 30$, from dimension 2 to 4 are displayed, so do the numbers of nonzeros elements in the matrix A.

In Table 2, we experiment Hermite-chaos with RHC approximation in higher dimensions, say $d = 6, 8$ and 12 with $N = 3$. It is observed that the error only slightly grows with respect to the dimension.

dimension	CPU time	♯ of basis	ϵ_{max}	ϵ_{L^2}
6	20.6636s	448	3.3226×10^{-2}	2.4975×10^{-2}
8	498.6522s	2304	5.7095×10^{-2}	4.5479×10^{-2}
12	4747.61s*	53248	1.167×10^{-1}	9.7×10^{-2}

Table 2: The CPU times, the number of basis and the errors, with $N = 3$, of dimension 6, 8 and 12 are displayed.
* This code has been parallelled and run by 24 CPU workers.

6. Conclusion

In this paper, we simplified the error analysis in [17, 12] and applied it to gPC with the HC approximations. The error analyses of the projection onto the linear subspace spanned by all sorts of Askey polynomials have been obtained. The theorems reveal that the convergence rate of Jacobi-chaos is twice faster than any other polynomials. We believe that it is due to its continuity and the boundedness of its support. It is illustrated by the numerical experiments that solving the ordinary differential equations with random inputs using RHC approximation generally converges faster than the OHC with respect to N, while the number of the nonzeros in the stiff matrix and the number of polynomial basis of RHC grows faster than that of OHC with respect to the dimension.

Acknowledgments

The author thanks the support of National Natural Science Foundation of China (grant No. 11501023, 11471184).

Appendix A. Connections between Galerkin approximation and the standard ANOVA decomposition

It is known that there are approaches proposed to deal with the difficulties caused by high dimensional random inputs, say the ANOVA decomposition [26] and references therein. In this appendix, we shall discuss the connection between the standard ANOVA decomposition and the Galerkin approximation (2.3). Moreover, the HC approximation can be naturally combined with the ANOVA decomposition to eliminate the effect of curse of dimensionality in theory.

Recall that the Galerkin approximation of a function $f(\boldsymbol{x})$, $\boldsymbol{x} = (x_1, \cdots, x_d) \in I^d$, is

$$f_{Gal,\nu}(\boldsymbol{x}) = \sum_{\boldsymbol{i} \in \boldsymbol{\Omega}_\nu} \hat{f}_{\boldsymbol{i}} \boldsymbol{\Phi}_{\boldsymbol{i}}(\boldsymbol{x}), \tag{A.1}$$

where I is the support of the polynomial basis $\Phi_i(x)$, $\boldsymbol{\Omega}_\nu$ is a proper subset of index set $\mathbb{N}^d$ with some parameter ν and $\boldsymbol{\Phi}_{\boldsymbol{i}}(\boldsymbol{x}) = \prod_{j=1}^d \Phi_{i_j}(x_j)$ are the polynomial basis. In section 3 and 4, the error analyses have been performed with $\boldsymbol{\Omega}_\nu$ chosen to be tensor product $\boldsymbol{\Omega}_{N,tensor}$, the RHC approximation $\boldsymbol{\Omega}_{N,RHC}$ and the OHC approximation with $\gamma \in [-\infty, 1)$ $\boldsymbol{\Omega}_{N,OHC,\gamma}$ (3.1) with various Askey polynomial basis.

Recall also that the standard ANOVA decomposition represents a function $f(\boldsymbol{x})$, $\boldsymbol{x} \in I^d$, as

$$f(\boldsymbol{x}) = f_0 + \sum_{s=1}^d \sum_{j_1 < \cdots < j_s} f_{j_1, \cdots, j_s}(x_{j_1}, \cdots, x_{j_s}),$$

or equivalently,

$$f(\boldsymbol{x}) = f_0 + \sum_{1 \le j_1 \le d} f_{j_1}(x_{j_1}) + \sum_{1 \le j_1 < j_2 \le d} f_{j_1,j_2}(x_{j_1}, x_{j_2}) + \cdots + f_{1,\cdots,d}(x_1, \cdots, x_d),$$

if

$$f_0 = \int_{I^d} f(\boldsymbol{x}) d\boldsymbol{\mu}(\boldsymbol{x}) \tag{A.2}$$

and

$$\int_I f_{j_1,\cdots,j_s}(x_{j_1},\cdots,x_{j_s})d\mu(x_{j_k}) = 0, \tag{A.3}$$

for $1 \le k \le s$.

The terms in the ANOVA decomposition are computed recursively

$$f_S = \int_{I^{\sharp(S^c)}} f(\boldsymbol{x})d\boldsymbol{\mu}(\boldsymbol{x}_{S^c}) - \sum_{T\subset S} f_T(\boldsymbol{x}_T), \tag{A.4}$$

where $S = \{j_1,\cdots,j_s\}$, $\sharp(S^c)$ is the number of elements in the complement set of S, $T = \{j_1,\cdots,j_t\}$ is a proper subset of S, $f_T = f_{j_1,\cdots,j_t}$ and $\boldsymbol{x}_T = (x_{j_1},\cdots,x_{j_t})$. We call

$$f_{ANOVA,\nu}(\boldsymbol{x}) = f_0 + \sum_{1\le j_1\le d} f_{j_1}(x_{j_1}) + \cdots + \sum_{1\le j_1<\cdots<j_\nu\le d} f_{j_1,\cdots,j_\nu}(x_{j_1},\cdots,x_{j_\nu}) \tag{A.5}$$

the νth degree ANOVA approximation of $f(\boldsymbol{x})$, for some $0 \le \nu < d$. In the following proposition, we shall show that $f_{ANOVA,\nu} = f_{Gal,\nu}$ with properly chosen measure $\boldsymbol{\mu}$ and the index set $\boldsymbol{\Omega}_\nu$.

Proposition A.1. *The νth degree ANOVA approximation $f_{ANOVA,\nu}$* (A.5) *with the measure*

$$d\boldsymbol{\mu}(\boldsymbol{x}) = \frac{1}{\boldsymbol{c_0}}\boldsymbol{\Phi_0}(\boldsymbol{x})\boldsymbol{\omega}(\boldsymbol{x})d\boldsymbol{x}$$

is exactly the Galerkin approximation $f_{Gal,\nu}$ (A.1) *with the index set $\boldsymbol{\Omega}_{\le\nu} = \{\boldsymbol{n}\in\mathbb{N}_0^d : |\boldsymbol{n}|_0 \le \nu\}$, where $\boldsymbol{c_0}$ is the normalization constant, i.e.,*

$$\int_{I^d} \boldsymbol{\Phi_i}(\boldsymbol{x})\boldsymbol{\Phi_k}(\boldsymbol{x})\boldsymbol{\omega}(\boldsymbol{x})d\boldsymbol{x} = \boldsymbol{c_i}\boldsymbol{\delta_{ik}},$$

and $|\boldsymbol{n}|_0$ is the 0-norm of a vector $\boldsymbol{n}$ defined in (2.1).

Proof. It is easy to see that

$$f_0 = \int_{I^d} f(\boldsymbol{x})d\boldsymbol{\mu}(\boldsymbol{x}) = \frac{1}{\boldsymbol{c_0}}\int_{I^d}\left(\sum_{\boldsymbol{n}\in\mathbb{N}_0^d}\hat{f}_{\boldsymbol{n}}\boldsymbol{\Phi_n}(\boldsymbol{x})\right)\boldsymbol{\Phi_0}(\boldsymbol{x})\boldsymbol{\omega}(\boldsymbol{x})d\boldsymbol{x} = \hat{f}_{\boldsymbol{0}}. \tag{A.6}$$

We claim that for any $1 \le l < d$, let $S_l = \{j_1,\cdots,j_l\}$, we have

$$f_{S_l}(\boldsymbol{x}_{S_l}) = \sum_{n_{j_1},\cdots,n_{j_l}=1}^{\infty} \hat{f}_{n_{j_1}\boldsymbol{e}_{j_1}+\cdots+n_{j_l}\boldsymbol{e}_{j_l}}\phi_{n_{j_1}}(x_{j_1})\cdots\phi_{n_{j_l}}(x_{j_l}), \tag{A.7}$$

where $\boldsymbol{e}_i$ is the ith unit vector in $\mathbb{R}^d$.

In fact, by induction, for $l=1$, let $S_1 = \{j_1\}$, it is easy to check that

$$\begin{aligned}
f_{j_1}(x_{j_1}) &\overset{(A.4)}{=} \int_{I^{d-1}} f(\boldsymbol{x})\frac{1}{\boldsymbol{c}_{\boldsymbol{0},S_1^c}}\boldsymbol{\Phi}_{\boldsymbol{0},S_1^c}(\boldsymbol{x}_{S_1^c})\boldsymbol{\omega}_{S_1^c}(\boldsymbol{x}_{S_1^c})d(\boldsymbol{x}_{S_1^c}) - f_0 \\
&= \int_{I^{d-1}}\left(\sum_{\boldsymbol{n}\in\mathbb{N}_0^d}\hat{f}_{\boldsymbol{n}}\boldsymbol{\Phi_n}(\boldsymbol{x})\right)\frac{1}{\boldsymbol{c}_{\boldsymbol{0},S_1^c}}\boldsymbol{\Phi}_{\boldsymbol{0},S_1^c}(\boldsymbol{x}_{S_1^c})\boldsymbol{\omega}_{S_1^c}(\boldsymbol{x}_{S_1^c})d(\boldsymbol{x}_{S_1^c}) - f_0 \\
&\overset{(A.6)}{=} \sum_{n_{j_1}=0}^{\infty}\hat{f}_{n_{j_1}\boldsymbol{e}_{j_1}}\Phi_{n_{j_1}}(x_{j_1}) - \hat{f}_{\boldsymbol{0}} = \sum_{n_{j_1}=1}^{\infty}\hat{f}_{n_{j_1}\boldsymbol{e}_{j_1}}\Phi_{n_{j_1}}(x_{j_1}),
\end{aligned}$$

where $S_1^c = \{1,\cdots,d\}/\{j_1\}$, $\boldsymbol{c}_{\boldsymbol{0},R} = c_{r_1}\cdots c_{r_\nu}$, $\boldsymbol{\Phi}_{\boldsymbol{n},R}(\boldsymbol{x}_R) = \phi_{n_1}(x_{r_1})\cdots\phi_{n_\nu}(x_{r_\nu})$ and $\boldsymbol{\omega}_R(\boldsymbol{x}_R) = \omega_{r_1}(x_{r_1})\cdots\omega_{r_\nu}(x_{r_\nu})$, if the index set $R = \{r_1,\cdots,r_\nu\}$. Next, we assume that (A.7) holds for all

S_m with $\sharp(S_m) = m$, with $m \le l-1$. We need to show that it is also true for $S_l = \{j_1, \cdots, j_l\}$. According to (A.4), we have

$$
\begin{aligned}
f_{S_l}(\boldsymbol{x}_{S_l}) =& \int_{I^{d-l}} \left(\sum_{\boldsymbol{n}\in\mathbb{N}_0^d} \hat{f}_{\boldsymbol{n}} \boldsymbol{\Phi}_{\boldsymbol{n}}(\boldsymbol{x}) \right) \frac{1}{\boldsymbol{c}_{\boldsymbol{0},S_l^c}} \boldsymbol{\Phi}_{\boldsymbol{0},S_l^c}(\boldsymbol{x}_{S_1^c}) \boldsymbol{\omega}_{S_l^c}(\boldsymbol{x}_{S_l^c}) d(\boldsymbol{x}_{S_l^c}) - \sum_{T\subset S_l} f_T(\boldsymbol{x}_T) \\
=& \sum_{n_{j_1},\cdots,n_{j_l}=0}^{\infty} \hat{f}_{n_{j_1}\boldsymbol{e}_{j_1}+\cdots+n_{j_l}\boldsymbol{e}_{j_l}} \phi_{n_{j_1}(x_{j_1})} \cdots \phi_{n_{j_l}}(x_{j_l}) - \sum_{S_{l-1}\subset\{1,\cdots,d\}} f_{S_{l-1}}(\boldsymbol{x}_{S_{l-1}}) - \cdots - f_0 \\
\overset{(A.7),(A.6)}{=}& \sum_{n_{j_1},\cdots,n_{j_l}=0}^{\infty} \hat{f}_{n_{j_1}\boldsymbol{e}_{j_1}+\cdots+n_{j_l}\boldsymbol{e}_{j_l}} \phi_{n_{j_1}(x_{j_1})} \cdots \phi_{n_{j_l}}(x_{j_l}) \\
&- \sum_{\{j_1,\cdots,j_{l-1}\}\subset\{1,\cdots,d\}} \sum_{n_{j_1},\cdots,n_{j_{l-1}}=1}^{\infty} \hat{f}_{n_{j_1}\boldsymbol{e}_{j_1}+\cdots+n_{j_{l-1}}\boldsymbol{e}_{j_{l-1}}} \phi_{n_{j_1}}(x_{j_1}) \cdots \phi_{n_{j_{l-1}}}(x_{j_{l-1}}) - \hat{f}_{\boldsymbol{0}} \\
=& \sum_{n_{j_1},\cdots,n_{j_l}=1}^{\infty} \hat{f}_{n_{j_1}\boldsymbol{e}_{j_1}+\cdots+n_{j_l}\boldsymbol{e}_{j_l}} \phi_{n_{j_1}(x_{j_1})} \cdots \phi_{n_{j_l}}(x_{j_l}).
\end{aligned}
$$

With this claim, for $0 \le \nu < d$, the νth degree of ANOVA approximation can be written as

$$
\begin{aligned}
f_{ANOVA,\nu} =& \sum_{l=0}^{\nu} \sum_{S_l\subset\{1,\cdots,d\},\sharp S_l = l} f_{S_l}(\boldsymbol{x}_{S_l}) \\
=& \sum_{l=0}^{\nu} \sum_{\{j_1,\cdots,j_l\}\subset\{1,\cdots,d\}} \sum_{n_{j_1},\cdots,n_{j_l}=1}^{\infty} \hat{f}_{n_{j_1}\boldsymbol{e}_{j_1}+\cdots+n_{j_l}\boldsymbol{e}_{j_l}} \phi_{n_{j_1}(x_{j_1})} \cdots \phi_{n_{j_l}}(x_{j_l}) \\
=& \sum_{l=0}^{\nu} \sum_{\boldsymbol{n}\in\boldsymbol{\Omega}_l} \hat{f}_{\boldsymbol{n}} \boldsymbol{\Phi}_{\boldsymbol{n}}(\boldsymbol{x}) = \sum_{\boldsymbol{n}\in\boldsymbol{\Omega}_{\le\nu}} \hat{f}_{\boldsymbol{n}} \boldsymbol{\Phi}_{\boldsymbol{n}}(\boldsymbol{x})
\end{aligned}
$$

where $\boldsymbol{\Omega}_l = \left\{\boldsymbol{n}\in\mathbb{N}_0^d : \ |\boldsymbol{n}|_0 = l\right\}$ and $\boldsymbol{\Omega}_{\le\nu} = \left\{\boldsymbol{n}\in\mathbb{N}_0^d : \ |\boldsymbol{n}|_0 \le \nu\right\}$. $\square$

Remark 1.5. *From* (A.6)*, the* 0*th degree ANOVA approximation is exactly the same as the* 0*th order Galerkin approximation with the fact that* $\boldsymbol{\Phi}_{\boldsymbol{0}} = 1$.

Remark 1.6. *ANOVA approximation can be naturally combined with HC approximation in the following way. It is clear to see that* $\sharp(\boldsymbol{\Omega}_{\le\nu}) = \infty$ *if* $\nu \ne 0$. *Certain truncation needs to be used in the* ν*th degree ANOVA approximation, say* $\boldsymbol{\Omega}_{\le\nu,N,tensor} = \{\boldsymbol{n}\in\mathbb{N}_0^d : \ |\boldsymbol{n}|_0 \le \nu, \ |\boldsymbol{n}|_\infty \le N\}$, $\boldsymbol{\Omega}_{\le\nu,N,RHC} = \{\boldsymbol{n}\in\mathbb{N}_0^d : \ |\boldsymbol{n}|_0 \le \nu, \ |\boldsymbol{n}|_{\mathrm{mix}} \le N\}$*, etc.*

Appendix B. Orthogonal polynomials of Askey scheme

Appendix B.1. Hermite polynomials $H_n(x)$ and Gaussian distribution

The three-term recurrence of the probabilist's Hermite polynomials are given by

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x), \tag{B.1}$$

for $n = 0, 1, 2, \cdots$, with $H_{-1}(x) = 0$ and $H_0(x) = 1$. The $\{H_n\}_{n\in\mathbb{N}_0}$ forms an orthogonal basis of $L^2_\omega(\mathbb{R})$ with the weight $\omega(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$:

$$\int_{\mathbb{R}} H_n(x) H_m(x) \omega(x) dx = n!\delta_{nm}, \tag{B.2}$$

where δ_{nm} is the Kronecker function. The derivative of $H_n(x)$ is explicitly expressed, namely

$$H_n'(x) = nH_{n-1}(x). \tag{B.3}$$

Furthermore, we have

$$\frac{d^k}{dx^k}H_n(x) = \frac{n!}{(n-k)!}H_{n-k}(x) =: \mu_{n,k}H_{n-k}, \tag{B.4}$$

if $n \geq k \geq 0$.

Now we define the d-dimensional tensorial Hermite polynomial as

$$\boldsymbol{H}_{\boldsymbol{n}}(\boldsymbol{x}) = \prod_{j=1}^{d} H_{n_j}(x_j),$$

for $\boldsymbol{x} \in \mathbb{R}^d$. It verifies readily that

$$\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}}\boldsymbol{H}_{\boldsymbol{n}} = \boldsymbol{\mu}_{\boldsymbol{n},\boldsymbol{k}}\boldsymbol{H}_{\boldsymbol{n}-\boldsymbol{k}}, \tag{B.5}$$

and

$$\int_{\mathbb{R}^d} \boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}}\boldsymbol{H}_{\boldsymbol{n}}(\boldsymbol{x})\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}}\boldsymbol{H}_{\boldsymbol{m}}(\boldsymbol{x})\boldsymbol{\omega}(\boldsymbol{x})d\boldsymbol{x} = \boldsymbol{\mu}_{\boldsymbol{n},\boldsymbol{k}}^2(\boldsymbol{n}-\boldsymbol{k})!\boldsymbol{\delta}_{\boldsymbol{n}\boldsymbol{m}} =: \boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{k}}\boldsymbol{\delta}_{\boldsymbol{n}\boldsymbol{m}}, \tag{B.6}$$

where $\boldsymbol{\omega}(\boldsymbol{x}) = \prod_{j=1}^{d}\omega(x_j)$, $\boldsymbol{\mu}_{\boldsymbol{n},\boldsymbol{k}} = \prod_{j=1}^{d}\mu_{n_j,k_j}$ and $\boldsymbol{\delta}_{\boldsymbol{n}\boldsymbol{m}} = \prod_{j=1}^{d}\delta_{n_j m_j}$. Here, $\boldsymbol{\delta}_{\boldsymbol{n}\boldsymbol{m}}$ is the tensorial Kronecker function. It is clear to see that the weight $\boldsymbol{\omega}$ is the density function of the standard Gaussian distribution.

The Hermite polynomials $\{\boldsymbol{H}_{\boldsymbol{n}}(\boldsymbol{x})\}_{\boldsymbol{n}\in\mathbb{N}_0^d}$ form an orthogonal basis of $L^2_{\boldsymbol{\omega}}\left(\mathbb{R}^d\right)$. That is, for any function $u \in L^2_{\boldsymbol{\omega}}\left(\mathbb{R}^d\right)$, it can be written in the form

$$u(\boldsymbol{x}) = \sum_{\boldsymbol{n}\geq 0}\hat{u}_{\boldsymbol{n}}\boldsymbol{H}_{\boldsymbol{n}}(\boldsymbol{x}), \tag{B.7}$$

with $\hat{u}_{\boldsymbol{n}} = \frac{1}{\boldsymbol{n}!}\int_{\mathbb{R}^d} u(\boldsymbol{x})\boldsymbol{H}_{\boldsymbol{n}}(\boldsymbol{x})\boldsymbol{\omega}(\boldsymbol{x})d\boldsymbol{x}$. Hence, we have $\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}}u(\boldsymbol{x}) = \sum_{\boldsymbol{n}\geq\boldsymbol{k}}\hat{u}_{\boldsymbol{n}}\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}}\boldsymbol{H}_{\boldsymbol{n}}(\boldsymbol{x})$. Furthermore,

$$\left|\left|\boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}}u\right|\right|^2_{\boldsymbol{\omega},\mathbb{R}^d} = \sum_{\boldsymbol{n}\geq\boldsymbol{k}}\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{k}}\left|\hat{u}_{\boldsymbol{n}}\right|^2 = \sum_{\boldsymbol{n}\in\mathbb{N}_0^d}\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{k}}\left|\hat{u}_{\boldsymbol{n}}\right|^2, \tag{B.8}$$

if we define conventionally $\mu_{n,k} = 0$, for $0 \leq n < k$.

Appendix B.2. Laguerre polynomial $L_n^{(\alpha)}(x)$ and gamma distribution

The Laguerre polynomial is given by the three-term recurrence relation:

$$(n+1)L_{n+1}^{(\alpha)}(x) - (2n+\alpha+1-x)L_n^{(\alpha)}(x) + (n+\alpha)L_{n-1}^{(\alpha)}(x) = 0,$$

with $L_{-1}^{(\alpha)}(x) = 0$ and $L_0^{(\alpha)}(x) = 1$, for any $\alpha > -1$, $x \in \mathbb{R}_+$. The orthogonality of $L_n^{(\alpha)}(x)$ with respect to the weight $\omega_\alpha(x) = \frac{x^\alpha e^{-x}}{\Gamma(\alpha+1)}$ is

$$\int_{\mathbb{R}_+} L_m^{(\alpha)}(x)L_n^{(\alpha)}(x)\omega_\alpha(x)dx = \frac{(\alpha+1)_n}{n!}\delta_{mn} =: \rho_{n,\alpha}\delta_{mn}, \tag{B.9}$$

with the Pochhammer symbol $(a)_n$ defined in (2.2).

Recall that the gamma distribution has the probability density function

$$f(x) = \frac{x^\alpha e^{-x/\beta}}{\beta^{\alpha+1}\Gamma(\alpha+1)},$$

for $\alpha > -1$, $\beta > 0$. The weight function of Laguerre polynomial is the same as that of the gamma distribution with $\beta = 1$. The derivative of Laguerre polynomial is

$$\frac{d^k}{dx^k}L_n^{(\alpha)}(x) = (-1)^k L_{n-k}^{(\alpha+k)}(x), \tag{B.10}$$

if $n \geq k \geq 0$. The d-dimensional tensorial Laguerre polynomial is readily defined as

$$\boldsymbol{L}_{\boldsymbol{n}}^{(\boldsymbol{\alpha})}(\boldsymbol{x}) = \prod_{j=1}^{d} L_{n_j}^{(\alpha_j)}(x_j),$$

for $\boldsymbol{x} \in \mathbb{R}_+^d$, $\boldsymbol{\alpha} > -\mathbf{1}$. The orthogonality of the tensorial Laguerre polynomial follows immediately from (B.9):

$$\int_{\mathbb{R}_+^d} \boldsymbol{L}_{\boldsymbol{n}}^{(\boldsymbol{\alpha})}(\boldsymbol{x}) \boldsymbol{L}_{\boldsymbol{m}}^{(\boldsymbol{\alpha})}(\boldsymbol{x}) \boldsymbol{\omega}_{\boldsymbol{\alpha}}(\boldsymbol{x}) d\boldsymbol{x} = \boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{\alpha}} \boldsymbol{\delta}_{\boldsymbol{mn}}, \tag{B.11}$$

where $\boldsymbol{\omega}_{\boldsymbol{\alpha}}(\boldsymbol{x}) = \prod_{j=1}^{d} \omega_{\alpha_j}(x_j)$, $\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{\alpha}} = \prod_{j=1}^{d} \rho_{n_j,\alpha_j}$ and $\boldsymbol{\delta}_{\boldsymbol{mn}} = \prod_{j=1}^{d} \delta_{m_j n_j}$. Furthermore, it is easy to deduce that

$$\int_{\mathbb{R}_+^d} \boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}} \boldsymbol{L}_{\boldsymbol{n}}^{(\boldsymbol{\alpha})}(\boldsymbol{x}) \boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}} \boldsymbol{L}_{\boldsymbol{m}}^{(\boldsymbol{\alpha})}(\boldsymbol{x}) \boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k}}(\boldsymbol{x}) d\boldsymbol{x} \overset{\text{(B.10)},\text{(B.11)}}{=} \boldsymbol{\rho}_{\boldsymbol{n}-\boldsymbol{k},\boldsymbol{\alpha}+\boldsymbol{k}} \boldsymbol{\delta}_{\boldsymbol{mn}}.$$

Any $u(\boldsymbol{x}) \in L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}}}(\mathbb{R}_+^d)$ can be written as

$$u(\boldsymbol{x}) = \sum_{\boldsymbol{n} \in \mathbb{N}_0^d} \hat{u}_{\boldsymbol{n}} \boldsymbol{L}_{\boldsymbol{n}}^{(\boldsymbol{\alpha})}(\boldsymbol{x}),$$

with

$$\hat{u}_{\boldsymbol{n}} = \frac{1}{\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{\alpha}}} \int_{\mathbb{R}_+^d} u(\boldsymbol{x}) \boldsymbol{L}_{\boldsymbol{n}}^{(\boldsymbol{\alpha})}(\boldsymbol{x}) \boldsymbol{\omega}_{\boldsymbol{\alpha}}(\boldsymbol{x}) d\boldsymbol{x}.$$

Hence, we have

$$\left|\left| \boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}} u(\boldsymbol{x}) \right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k}},\mathbb{R}_+^d} = \sum_{\boldsymbol{n} \geq \boldsymbol{k}} \boldsymbol{\rho}_{\boldsymbol{n}-\boldsymbol{k},\boldsymbol{\alpha}+\boldsymbol{k}} \left| \hat{u}_{\boldsymbol{n}} \right|^2 = \sum_{\boldsymbol{n} \in \mathbb{N}_0^d} \boldsymbol{\rho}_{\boldsymbol{n}-\boldsymbol{k},\boldsymbol{\alpha}+\boldsymbol{k}} \left| \hat{u}_{\boldsymbol{n}} \right|^2, \tag{B.12}$$

if we let $\rho_{n-k,\alpha+k} = 0$ when $0 \leq n < k$.

Appendix B.3. Jacobi polynomial $P_n^{(\alpha,\beta)}(x)$ and beta distribution

The Jacobi polynomial is given by the three-term recurrence relation:

$$\begin{aligned} xP_n^{(\alpha,\beta)}(x) =& \frac{2(n+1)(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} P_{n+1}^{(\alpha,\beta)}(x) + \frac{\beta^2-\alpha^2}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)} P_n^{(\alpha,\beta)}(x) \\ &+ \frac{2(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)} P_{n-1}^{(\alpha,\beta)}(x), \end{aligned}$$

for $n = 2, 3, \cdots$, with $P_{-1}^{(\alpha,\beta)}(x) = 0$ and $P_0^{(\alpha,\beta)}(x) = 1$, for any $\alpha, \beta > -1$, $x \in (-1,1) =: I$. The orthogonality of $P_n^{(\alpha,\beta)}(x)$ with respect to the weight

$$\omega_{\alpha,\beta}(x) = \frac{\Gamma(\alpha+\beta+2)}{2^{\alpha+\beta}\Gamma(\alpha+1)\Gamma(\beta+1)} (1-x)^{\alpha}(1+x)^{\beta}$$

is

$$\int_I P_m^{(\alpha,\beta)}(x) P_n^{(\alpha,\beta)}(x) \omega_{\alpha,\beta}(x) dx = h_n^2 \delta_{mn}, \tag{B.13}$$

for $\alpha, \beta, \alpha+\beta > -1$, where $\Gamma(\circ)$ denotes the gamma function and

$$h_n^2 = \frac{2\Gamma(\alpha+\beta+2)\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{n!(2n+\alpha+\beta+1)\Gamma(\alpha+1)\Gamma(\beta+1)\Gamma(n+\alpha+\beta+1)}. \tag{B.14}$$

It is easy to verify that letting $\bar{\alpha} = \alpha + 1$, $\bar{\beta} = \beta + 1$ and $\bar{x} = \frac{1-x}{2}$, the weight function $\omega_{\alpha,\beta}$ is exactly the density function of beta distribution, i.e.

$$f(\bar{x}; \bar{\alpha}, \bar{\beta}) = \frac{\Gamma(\bar{\alpha} + \bar{\beta})}{\Gamma(\bar{a})\Gamma(\bar{\beta})} \bar{x}^{\bar{\alpha}-1}(1-\bar{x})^{\bar{\beta}-1},$$

for $\bar{\alpha}, \bar{\beta} > 0$, $\bar{x} \in (0, 1)$. The derivative of Jacobi polynomial is

$$\frac{d^k}{dx^k} P_n^{(\alpha,\beta)}(x) = \frac{\Gamma(\alpha + \beta + n + 1 + k)}{2^k \Gamma(\alpha + \beta + n + 1)} P_{n-k}^{(\alpha+k,\beta+k)}(x), \tag{B.15}$$

if $n \geq k \geq 0$. Now we define the d-dimensional tensorial Jacobi polynomial as

$$\boldsymbol{P}_{\boldsymbol{n}}^{(\boldsymbol{\alpha},\boldsymbol{\beta})}(\boldsymbol{x}) = \prod_{j=1}^{d} P_{n_j}^{(\alpha_j,\beta_j)}(x_j),$$

for $\boldsymbol{x} \in I^d$, $\boldsymbol{\alpha}, \boldsymbol{\beta} > -\mathbf{1}$. The orthogonality of the tensorial Jacobi polynomial follows immediately from (B.13):

$$\int_{I^d} \boldsymbol{P}_{\boldsymbol{n}}^{(\boldsymbol{\alpha},\boldsymbol{\beta})}(\boldsymbol{x}) \boldsymbol{P}_{\boldsymbol{m}}^{(\boldsymbol{\alpha},\boldsymbol{\beta})}(\boldsymbol{x}) \boldsymbol{\omega}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{x}) d\boldsymbol{x} = \boldsymbol{h}_{\boldsymbol{n}}^2 \boldsymbol{\delta}_{\boldsymbol{mn}}, \tag{B.16}$$

for $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\alpha} + \boldsymbol{\beta} > -1$, where $\boldsymbol{\omega}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{x}) = \prod_{j=1}^{d} \omega_{\alpha_j,\beta_j}(x_j)$, $\boldsymbol{\delta}_{\boldsymbol{mn}} = \prod_{j=1}^{d} \delta_{m_j n_j}$ and $\boldsymbol{h}_{\boldsymbol{n}} = \prod_{j=1}^{d} h_{n_j}$. Furthermore, it is easy to deduce that

$$\int_{I^d} \boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}} \boldsymbol{P}_{\boldsymbol{n}}^{(\boldsymbol{\alpha},\boldsymbol{\beta})}(\boldsymbol{x}) \boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}} \boldsymbol{P}_{\boldsymbol{m}}^{(\boldsymbol{\alpha},\boldsymbol{\beta})}(\boldsymbol{x}) \boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k},\boldsymbol{\beta}+\boldsymbol{k}}(\boldsymbol{x}) d\boldsymbol{x} \overset{\text{(B.15),(B.16)}}{=} \boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{k},\boldsymbol{\alpha},\boldsymbol{\beta}} \boldsymbol{\delta}_{\boldsymbol{mn}}, \tag{B.17}$$

where $\boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{k},\boldsymbol{\alpha},\boldsymbol{\beta}} = \prod_{j=1}^{d} \rho_{n_j,k_j,\alpha_j,\beta_j}$ and

$$\rho_{n_j,k_j,\alpha_j,\beta_j} = \left[\frac{\Gamma(\alpha_j + \beta_j + n_j + k_j + 1)}{2^{k_j} \Gamma(\alpha_j + \beta_j + n_j + 1)} \right]^2 h_{n_j - k_j}^2. \tag{B.18}$$

Any $u(\boldsymbol{x}) \in L^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha},\boldsymbol{\beta}}}\left(I^d\right)$ can be written as

$$u(\boldsymbol{x}) = \sum_{\boldsymbol{n} \in \mathbb{N}_0^d} \hat{u}_{\boldsymbol{n}} \boldsymbol{P}_{\boldsymbol{n}}^{(\boldsymbol{\alpha},\boldsymbol{\beta})}(\boldsymbol{x}),$$

with

$$\hat{u}_{\boldsymbol{n}} = \frac{1}{\boldsymbol{h}_{\boldsymbol{n}}^2} \int_{I^d} u(\boldsymbol{x}) \boldsymbol{P}_{\boldsymbol{n}}^{(\boldsymbol{\alpha},\boldsymbol{\beta})}(\boldsymbol{x}) \boldsymbol{\omega}_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{x}) d\boldsymbol{x}.$$

Hence, we have

$$\left|\left| \boldsymbol{\partial}_{\boldsymbol{x}}^{\boldsymbol{k}} u(\boldsymbol{x}) \right|\right|^2_{\boldsymbol{\omega}_{\boldsymbol{\alpha}+\boldsymbol{k},\boldsymbol{\beta}+\boldsymbol{k}}, I^d} = \sum_{\boldsymbol{n} \geq \boldsymbol{k}} \boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{k},\boldsymbol{\alpha},\boldsymbol{\beta}} \left| \hat{u}_{\boldsymbol{n}} \right|^2 = \sum_{\boldsymbol{n} \in \mathbb{N}_0^d} \boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{k},\boldsymbol{\alpha},\boldsymbol{\beta}} \left| \hat{u}_{\boldsymbol{n}} \right|^2, \tag{B.19}$$

if we let $\rho_{n,k,\alpha,\beta} = 0$ when $0 \leq n < k$.

Appendix B.4. Charlier polynomial $C_n(x; a)$ and Poisson distribution

Charlier polynomial $C_n(x; a)$ is given by the recurrence relation:

$$-xC_n(x; a) = aC_{n+1}(x; a) - (n + a)C_n(x; a) + nC_{n-1}(x; a), \quad a > 0,$$

for $n \geq 1$, $x \in \mathbb{N}_0$, with $C_{-1}(x; a) = 0$ and $C_0(x; a) = 1$. The orthogonality of Charlier polynomials with respect to the weight $w(x; a) = e^{-a} \frac{a^x}{x!}$ is

$$\sum_{x=0}^{\infty} C_n(x; a) C_m(x; a) w(x; a) = a^{-n} n! \delta_{mn}. \tag{B.20}$$

The probability function of Poisson distribution is exactly the weight $w(x;a)$. The forward difference of Charlier polynomial is

$$\triangle_x^k C_n(x;a) = (-1)^k \frac{\Gamma(n+1)}{\Gamma(n-k+1)a^k} C_{n-k}(x;a),$$

if $n \geq k \geq 0$. Now we define the d-dimensional tensorial Charlier polynomial

$$\boldsymbol{C_n}(\boldsymbol{x};\boldsymbol{a}) = \prod_{j=1}^{d} C_{n_j}(x_j;a_j),$$

for $\boldsymbol{a} > 0$ and $\boldsymbol{x} \in \mathbb{N}_0^d$. The orthogonality of the tensorial Charlier polynomial follows immediately from (B.20):

$$\sum_{\boldsymbol{x}\in\mathbb{N}_0^d} \boldsymbol{C_n}(\boldsymbol{x};\boldsymbol{a})\boldsymbol{C_m}(\boldsymbol{x};\boldsymbol{a})\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{a}) = \boldsymbol{a}^{-\boldsymbol{n}}\boldsymbol{n}!\boldsymbol{\delta_{nm}}.$$

Furthermore, it is easy to deduce that

$$\sum_{\boldsymbol{x}\in\mathbb{N}_0^d} \triangle_{\boldsymbol{x}}^{\boldsymbol{k}}\boldsymbol{C_n}(\boldsymbol{x};\boldsymbol{a})\triangle_{\boldsymbol{x}}^{\boldsymbol{k}}\boldsymbol{C_m}(\boldsymbol{x};\boldsymbol{a})\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{a}) = \frac{\Gamma(\boldsymbol{n}+1)^2(\boldsymbol{n}-\boldsymbol{k})!}{\Gamma(\boldsymbol{n}-\boldsymbol{k}+1)^2\boldsymbol{a}^{\boldsymbol{n}+\boldsymbol{k}}}\boldsymbol{\delta_{nm}} =: \boldsymbol{\rho_{n,k,a}}\boldsymbol{\delta_{nm}}, \tag{B.21}$$

if $\boldsymbol{n} \geq \boldsymbol{k} \geq 0$, where $\Gamma(\boldsymbol{n}) := \prod_{j=1}^{d}\Gamma(n_j)$. Any $u(\boldsymbol{x}) \in l^2_{\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{a})}\left(\mathbb{N}_0^d\right)$ (in section 3.2 we denote $\boldsymbol{\omega_a} = \boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{a})$ for short) can be written as

$$u(\boldsymbol{x}) = \sum_{\boldsymbol{n}\in\mathbb{N}_0^d} \hat{u}_{\boldsymbol{n}}\boldsymbol{C_n}(\boldsymbol{x};\boldsymbol{a}),$$

with

$$\hat{u}_{\boldsymbol{n}} = \frac{\boldsymbol{a}^{\boldsymbol{n}}}{\boldsymbol{n}!}\sum_{\boldsymbol{x}\in\mathbb{N}_0^d} u(\boldsymbol{x})\boldsymbol{C_n}(\boldsymbol{x};\boldsymbol{a})\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{a}).$$

Appendix B.5. Krawtchouk polynomial $K_n(x;p,N)$ and binomial distribution

Krawtchouk polynomial $K_n(x;p,N)$ is given by the recurrence relation:

$$-xK_n(x;p,N) = p(N-n)K_{n+1}(x;p,N) - [p(N-n)+n(1-p)]K_n(x;p,N) + n(1-p)K_{n-1}(x;p,N),$$

for $0 < p < 1$, $x \in \mathbb{N}_N$, $n \in \mathbb{N}_N$, $N \in \mathbb{N}$, where $\mathbb{N}_N = \{0,1,\cdots,N\}$, with $K_{-1}(x;p,N) = 0$ and $K_0(x;p,N) = 1$. The orthogonality of Krawtchouk polynomial with respect to the weight $\omega(x;p,N) = \binom{N}{x}p^x(1-p)^{N-x}$ is

$$\sum_{x=0}^{N} K_m(x;p,N)K_n(x;p,N)\omega(x;p,N) = \left(\frac{1-p}{p}\right)^n / \binom{N}{n}\delta_{mn}, \quad 0 < p < 1. \tag{B.22}$$

The weight function is the probability density function of binomial distribution. The forward difference of Krawtchouk polynomial is

$$\triangle_x^k K_n(x;p,N) = \frac{(-1)^k}{p^k}\frac{\Gamma(n+1)\Gamma(N-k+1)}{\Gamma(n-k+1)\Gamma(N+1)}K_{n-k}(x;p,N-k),$$

if $n \geq k \geq 0$. Now we define the d-dimensional tensorial Krawtchouk polynomial

$$\boldsymbol{K_n}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N}) = \prod_{j=1}^{d} K_{n_j}(x_j;p_j,N_j),$$

for $0 < \boldsymbol{p} < 1$ and $\boldsymbol{N} \in \mathbb{N}_0^d$. The orthogonality of the tensorial Krawtchouk polynomial follows immediately from (B.22):

$$\sum_{\boldsymbol{x}\in\mathbb{N}_{\boldsymbol{N}}} \boldsymbol{K_m}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N})\boldsymbol{K_n}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N})\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N}) = \left(\frac{1-\boldsymbol{p}}{\boldsymbol{p}}\right)^{\boldsymbol{n}} / \binom{\boldsymbol{N}}{\boldsymbol{n}} \boldsymbol{\delta_{mn}},$$

where $\mathbb{N}_{\boldsymbol{N}} = \otimes_{j=1}^d \mathbb{N}_{N_j}$ and $\binom{\boldsymbol{N}}{\boldsymbol{n}} = \prod_{j=1}^d \binom{N_j}{n_j}$. Furthermore, it is easy to deduce that

$$\begin{aligned}\sum_{\boldsymbol{x}\in\mathbb{N}_{\boldsymbol{N}-\boldsymbol{k}}} &\triangle_{\boldsymbol{x}}^{\boldsymbol{k}}\boldsymbol{K_m}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N})\triangle_{\boldsymbol{x}}^{\boldsymbol{k}}\boldsymbol{K_n}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N})\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N}-\boldsymbol{k}) \\ =&\frac{(1-\boldsymbol{p})^{\boldsymbol{n}-\boldsymbol{k}}}{\boldsymbol{p}^{\boldsymbol{n}+\boldsymbol{k}}}\left(\frac{\Gamma(\boldsymbol{n}+1)\Gamma(\boldsymbol{N}-\boldsymbol{k}+1)}{\Gamma(\boldsymbol{n}-\boldsymbol{k}+1)\Gamma(\boldsymbol{N}+1)}\right)^2 / \binom{\boldsymbol{N}-\boldsymbol{k}}{\boldsymbol{n}-\boldsymbol{k}}\boldsymbol{\delta_{mn}} =: \boldsymbol{\rho_{n,k,p,N}}\boldsymbol{\delta_{mn}},\end{aligned} \tag{B.23}$$

if $\boldsymbol{n} \geq \boldsymbol{k} \geq 0$. Any $u(\boldsymbol{x}) \in l^2_{\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N})}(\mathbb{N}_{\boldsymbol{N}})$ (in section 4.3 $\boldsymbol{\omega_{p,N}} = \boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N})$ for short) can be written as

$$u(\boldsymbol{x}) = \sum_{\boldsymbol{n}\in\mathbb{N}_{\boldsymbol{N}}} \hat{u}_{\boldsymbol{n}}\boldsymbol{K_n}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N}),$$

with

$$\hat{u}_{\boldsymbol{n}} = \binom{\boldsymbol{N}}{\boldsymbol{n}} / \left(\frac{1-\boldsymbol{p}}{\boldsymbol{p}}\right)^{\boldsymbol{n}} \sum_{\boldsymbol{x}\in\mathbb{N}_{\boldsymbol{N}}} u(\boldsymbol{x})\boldsymbol{K_n}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N})\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{p},\boldsymbol{N}).$$

Appendix B.6. Meixner polynomial $M_n(x;\beta,c)$ and negative binomial distribution

Meixner polynomial $M_n(x;\beta,c)$ is given by the recurrence relation:

$$(c-1)xM_n(x;\beta,c) = c(n+\beta)M_{n+1}(x;\beta,c) - [n+(n+\beta)c]M_n(x;\beta,c) + nM_{n-1}(x;\beta,c),$$

for $\beta > 0$, $0 < c < 1$, $x \in \mathbb{N}_0$ and $n \in \mathbb{N}_0$, with $M_{-1}(x;\beta,c) = 0$ and $M_0(x;\beta,c) = 1$. The orthogonality of Meixner polynomial with respect to the weight $\omega(x;\beta,c) = \frac{(\beta)_x}{x!}c^x(1-c)^\beta$ is

$$\sum_{x=0}^{\infty} M_m(x;\beta,c)M_n(x;\beta,c)\omega(x;\beta,c) = \frac{c^{-n}n!}{(\beta)_n}\delta_{mn}, \tag{B.24}$$

where $(\beta)_x$ is the Pochhammer notation defined in (2.2). The weight function is the probability density function of negative binomial distribution. In the case where β is an integer, it is often called Pascal distribution. The forward difference of Meixner polynomial is

$$\triangle_x^k M_n(x;\beta,c) = \left(\frac{c-1}{c}\right)^k \frac{\Gamma(n+1)\Gamma(\beta)}{\Gamma(n-k+1)\Gamma(\beta+k)} M_{n-k}(x;\beta+k,c),$$

if $n \geq k \geq 0$. Now we define the d-dimensional tensorial Meixner polynomial

$$\boldsymbol{M_n}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c}) = \prod_{j=1}^d M_{n_j}(x_j;\beta_j,c_j),$$

for $0 < \boldsymbol{\beta}$ and $0 < \boldsymbol{c} < 1$. The orthogonality of the tensorial Meixner polynomial follows immediately from (B.24):

$$\sum_{\boldsymbol{x}\in\mathbb{N}_0^d} \boldsymbol{M_m}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c})\boldsymbol{M_n}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c})\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c}) = \frac{\boldsymbol{c}^{-\boldsymbol{n}}\boldsymbol{n}!}{(\boldsymbol{\beta})_{\boldsymbol{n}}}\boldsymbol{\delta_{mn}}.$$

Furthermore, it is easy to deduce that

$$\begin{aligned}\sum_{\boldsymbol{x}\in\mathbb{N}_0^d} &\triangle_{\boldsymbol{x}}^{\boldsymbol{k}}\boldsymbol{M_m}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c})\triangle_{\boldsymbol{x}}^{\boldsymbol{k}}\boldsymbol{M_n}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c})\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{\beta}+\boldsymbol{k},\boldsymbol{c}) \\ =&\frac{(\boldsymbol{c}-1)^{2\boldsymbol{k}}}{\boldsymbol{c}^{\boldsymbol{n}+\boldsymbol{k}}}\frac{\Gamma(\boldsymbol{n}+1)^2\Gamma(\boldsymbol{\beta})^2}{\Gamma(\boldsymbol{n}-\boldsymbol{k}+1)\Gamma(\boldsymbol{\beta}+\boldsymbol{k})\Gamma(\boldsymbol{\beta}+\boldsymbol{n})}\boldsymbol{\delta_{mn}} =: \boldsymbol{\rho_{n,k,\beta,c}}\boldsymbol{\delta_{mn}},\end{aligned} \tag{B.25}$$

if $\boldsymbol{n} \geq \boldsymbol{k} \geq 0$. Any $u(\boldsymbol{x}) \in l^2_{\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c})}\left(\mathbb{N}_0^d\right)$ (in section 4.4$\boldsymbol{\omega_{\beta,c}} = \boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c})$ for short) can be written as

$$u(\boldsymbol{x}) = \sum_{\boldsymbol{n}\in\mathbb{N}_0^d} \hat{u}_{\boldsymbol{n}} \boldsymbol{M}_{\boldsymbol{n}}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c}),$$

with

$$\hat{u}_{\boldsymbol{n}} = \frac{\boldsymbol{c}^{\boldsymbol{n}}(\boldsymbol{\beta})_{\boldsymbol{n}}}{\boldsymbol{n}!} \sum_{\boldsymbol{x}\in\mathbb{N}_0^d} u(\boldsymbol{x}) \boldsymbol{M}_{\boldsymbol{n}}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c}) \boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{\beta},\boldsymbol{c}).$$

Appendix B.7. Hahn polynomial $Q_n(x;\alpha,\beta,N)$ and hypergeometric distribution

Hahn polynomial $Q_n(x;\alpha,\beta,N)$ is given by the recurrence relation:

$$-xQ_n(x;\alpha,\beta,N) = A_n Q_{n+1}(x;\alpha,\beta,N) - (A_n + C_n)Q_n(x;\alpha,\beta,N) + C_n Q_{n-1}(x;\alpha,\beta,N),$$

where

$$\begin{cases} A_n = \dfrac{(n+\alpha+\beta+1)(n+\alpha+1)(N-n)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} \\ C_n = \dfrac{n(n+\alpha+\beta+N+1)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)} \end{cases},$$

for $\alpha, \beta > -1$ or $\alpha, \beta < -N$, $n = 0, 1, \cdots, N$ and $x \in \mathbb{N}_N$. The orthogonality of Hahn polynomial with respect to the weight $\omega(x;\alpha,\beta,N) = \frac{(\alpha+1)_x(\beta+1)_{N-x}}{x!(N-x)!}$ is

$$\sum_{x=0}^{\infty} Q_m(x;\alpha,\beta,N)Q_n(x;\alpha,\beta,N)\omega(x;\alpha,\beta,N) = h_n^2(\alpha,\beta,N)\delta_{mn}, \tag{B.26}$$

where

$$h_n^2(\alpha,\beta,N) = \frac{(-1)^n(n+\alpha+\beta+1)_{N+1}(\beta+1)_n n!}{(2n+\alpha+\beta+1)(\alpha+1)_n(-N)_n N!}.$$

where $(\beta)_x$ is the Pochhammer notation defined in (2.2). It is easy to verify that when $\alpha, \beta < -N$, $(-1)^N\omega(x;\alpha,\beta,N) > 0$ and $(-1)^N h_n^2(\alpha,\beta,N) > 0$. Also when $\alpha, \beta < -N$, if we set $\alpha = -\tilde{\alpha} - 1$ and $\beta = -\tilde{\beta} - 1$ in the weight function, we obtain

$$\tilde{w} = \frac{1}{\binom{N}{N-\tilde{\alpha}-\tilde{\beta}-1}} \frac{\binom{\tilde{\alpha}}{x}\binom{\tilde{\beta}}{N-x}}{\binom{\tilde{\alpha}+\tilde{\beta}}{N}},$$

which is exactly the hypergeometric distribution, apart from the constant $1/\binom{N}{N-\tilde{\alpha}-\tilde{\beta}-1}$ in front. We shall restrict ourselves to the case $\alpha, \beta < -N$, due to the close connection to the hypergeometric distribution. The forward difference of Hahn polynomial is

$$\triangle_x^k Q_n(x;\alpha,\beta,N) = (-1)^k \frac{(n-k+1)_k(n+\alpha+\beta+1)_k}{(\alpha+1)_k(N-k+1)_k} Q_{n-k}(x;\alpha+k,\beta+k,N-k),$$

if $n \geq k \geq 0$. Now we define the d-dimensional tensorial Hahn polynomial

$$\boldsymbol{Q}_{\boldsymbol{n}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}) = \prod_{j=1}^{d} Q_{n_j}(x_j;\alpha_j,\beta_j,N_j),$$

for $\boldsymbol{\alpha}, \boldsymbol{\beta} > -1$ or $\boldsymbol{\alpha}, \boldsymbol{\beta} < -\boldsymbol{N}$. The orthogonality of the tensorial Hahn polynomial follows immediately from (B.26):

$$\sum_{\boldsymbol{x}\in\mathbb{N}_{\boldsymbol{N}}} \boldsymbol{Q}_{\boldsymbol{m}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N})\boldsymbol{Q}_{\boldsymbol{n}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N})\bar{\boldsymbol{\omega}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}) = \bar{\boldsymbol{h}}_{\boldsymbol{n}}^2(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N})\boldsymbol{\delta}_{\boldsymbol{mn}},$$

where $\bar{\boldsymbol{\omega}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}) = (-1)^{\boldsymbol{N}}\boldsymbol{\omega}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}) > 0$ and $\bar{\boldsymbol{h}}^2_{\boldsymbol{n}}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}) = \prod_{j=1}^{d}(-1)^{N_j}h^2_{n_j}(\alpha_j,\beta_j,N_j) > 0$. Furthermore, it is easy to deduce that

$$\begin{aligned}&\sum_{\boldsymbol{x}\in\mathbb{N}_{\boldsymbol{N}}} \triangle^{\boldsymbol{k}}_{\boldsymbol{x}}\boldsymbol{Q}_{\boldsymbol{m}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N})\triangle^{\boldsymbol{k}}_{\boldsymbol{x}}\boldsymbol{Q}_{\boldsymbol{n}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N})\bar{\boldsymbol{\omega}}(\boldsymbol{x};\boldsymbol{\alpha}+\boldsymbol{k},\boldsymbol{\beta}+\boldsymbol{k},\boldsymbol{N}-\boldsymbol{k})\\ &\quad=\left[\frac{(\boldsymbol{n}-\boldsymbol{k}+1)_{\boldsymbol{k}}(\boldsymbol{n}+\boldsymbol{\alpha}+\boldsymbol{\beta}+1)_{\boldsymbol{k}}}{(\boldsymbol{\alpha}+1)_{\boldsymbol{k}}(\boldsymbol{N}-\boldsymbol{k}+1)_{\boldsymbol{k}}}\right]^2\bar{\boldsymbol{h}}^2_{\boldsymbol{n}-\boldsymbol{k}}(\boldsymbol{\alpha}+\boldsymbol{k},\boldsymbol{\beta}+\boldsymbol{k},\boldsymbol{N}-\boldsymbol{k})\boldsymbol{\delta}_{\boldsymbol{m}\boldsymbol{n}} =: \boldsymbol{\rho}_{\boldsymbol{n},\boldsymbol{k},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}}\boldsymbol{\delta}_{\boldsymbol{m}\boldsymbol{n}},\end{aligned} \tag{B.27}$$

if $\boldsymbol{n}\ge\boldsymbol{k}\ge 0$. Any $u(\boldsymbol{x})\in l^2_{\bar{\boldsymbol{\omega}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N})}(\mathbb{N}_{\boldsymbol{N}})$ (in section 4.5 $\boldsymbol{\omega}_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}} = \bar{\boldsymbol{\omega}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N})$ for short) can be written as

$$u(\boldsymbol{x}) = \sum_{\boldsymbol{n}\in\mathbb{N}_{\boldsymbol{N}}}\hat{u}_{\boldsymbol{n}}\boldsymbol{Q}_{\boldsymbol{n}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}),$$

with

$$\hat{u}_{\boldsymbol{n}} = \frac{1}{\bar{\boldsymbol{h}}^2_{\boldsymbol{n}}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N})}\sum_{\boldsymbol{x}\in\mathbb{N}_{\boldsymbol{N}}}u(\boldsymbol{x})\boldsymbol{Q}_{\boldsymbol{n}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N})\bar{\boldsymbol{\omega}}(\boldsymbol{x};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{N}).$$

References

[1] R. Askey, J. Wilson, *Some basic hyperbolic polynomials that generalize Jacobi polynomials*, Mem. Amer. Math. Soc. 319, AMS, Providence, RI, 1983.

[2] I. M. Babuška, R. Tempone, G. E. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM J. Numer. Anal., 435:1005-1034, 2007.

[3] V. Barthelmann, E. Novak, K. Ritter, *High dimensional polynomial interpolation on sparse grids*, Adv. Comput. Math., 12:273-288, 2000.

[4] H.-J. Burgartz, M. Griebel, *Sparse grids*, Acta Numer., 13:147-269, 2004.

[5] R. Cameron, W. Martin, *The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals*, Ann. of Math. (2), 48:385-392, 1947.

[6] T. Gerstner, M. Griebel, *Numerical integration using sparse grids*, Numer. Algorithms, 18:209-232, 1998.

[7] R. G. Ghanem, P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 1991.

[8] M. Griebel, J. Hamakekers, *Sparse grid for the Schrödinger equation*, M2AN Math. Model. Numer. Anal., 41:215-247, 2007.

[9] K. Karhunen, *Uber lineare methoden in der wahrscheinlichkeitsrechnung*, Amer. Acad. Sci. Fennicade, Ser. A, I, 37:3-79, 1947; (Translation: RAND Corporation, Santa Monica, California, Rep. T-131, Aug. 1960).

[10] S. Knapek, *Hyperbolic cross approximation of integral operators with smooth kernel*, Tech. Report 665, SFB 256, Univ. Bonn, 2000.

[11] M. Loeve, *Fonctions aleatoires du second ordre*, supplement to P. Levy, Processus Stochastic et Mouvement Brownien, Paris, Gauthier Villars, 1948.

[12] X. Luo, S. S.-T. Yau, *Hermite spectral method with hyperbolic cross approximations to high-dimensional parabolic PDEs*, SIAM J. Numer. Anal., 51(6):3186-3212, 2013.

[13] X. Ma, N. Zabaras, *An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations*, J. Comput. Phys., 228:3084-3113, 2009.

[14] O. Le Maitre, O. Knio, H. Najm, R. Ghanem, *Uncertainty propagation using Wiener-Haar expansions*, J. Comput. Phys., 197: 28-57, 2004.

[15] F. Nobile, R. Tempone, C. Webster, *A sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal., 46(5):2309-2345, 2008.

[16] W. Schoutens, *Stochastic Processes and Orthogonal Polynomials*, Springer-Verlag, New York, 2000.

[17] J. Shen, L.-L. Wang, *Sparse spectral approximations of high-dimensional problems based on hyperbolic cross*, SIAM J. Numer. Anal., 48(3):1087-1109, 2010.

[18] S. A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR, 4:240-243, 1963.

[19] G. Szegö, *Orthogonal Polynomials*, American Mathematical Society, Providence, RI, 1939.

[20] X. Wan and G. E. Karniadakis, *Multi-element generalized polynomial chaos for arbitrary probability measures*, SIAM J. Sci. Comput., 28:901-928, 2006.

[21] N. Wiener, *The homogeneous chaos*, Amer. J. Math., 60:897-936, 1938.

[22] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, 2010.

[23] D. Xiu and J. S. Hesthaven, *High-order collocation methods for differential equations with random inputs*, SIAM J. Sci. Comput., 27:1118-1139, 2005.

[24] D. Xiu and G. E. Karniadakis, *The Wiener-Askey polynomial chaos for stochastic differential equations*, SIAM J. Sci. Comput., 24(2):619-644, 2002.

[25] D. Xiu and G. E. Karniadakis, *Modeling uncertainty in flow situations via generalized polynomial chaos*, J. Comput. Phys., 187:137-167, 2003.

[26] X. Yang, M. Choi, G. Lin and G. Karniadakis, *Adaptive ANOVA decomposition of stochastic incompressible and compressible flows*, J. Comp. Physics, 231:1587-1614, 2012.

[27] C. Zenger, *Sparse grids*, in Parallel Algorithms for Partial Differential Equations, Notes Numer. Fluid Mech. 31, W. Hackbusch, ed., Vieweg, Braunschweig, 241-251, 1991.