



**HAL**  
open science

# A Generalized Pell's equation for a class of multivariate orthogonal polynomials

Jean-Bernard Lasserre, Yuan Xu

► **To cite this version:**

Jean-Bernard Lasserre, Yuan Xu. A Generalized Pell's equation for a class of multivariate orthogonal polynomials. 2023. hal-04163153v1

**HAL Id: hal-04163153**

**<https://laas.hal.science/hal-04163153v1>**

Preprint submitted on 17 Jul 2023 (v1), last revised 31 Mar 2024 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# A GENERALIZED PELL'S EQUATION FOR A CLASS OF MULTIVARIATE ORTHOGONAL POLYNOMIALS

JEAN B. LASSERRE AND YUAN XU

ABSTRACT. We extend the polynomial Pell's equation satisfied by univariate Chebyshev polynomials on  $[-1, 1]$  from one variable to several variables, using orthogonal polynomials on regular domains that include cubes, balls, and simplexes of arbitrary dimension. Moreover, we show that such an equation is strongly connected (i) to a certificate of positivity (from real algebraic geometry) on the domain, as well as (ii) to the Christoffel functions of the equilibrium measure on the domain. In addition, the solution to Pell's equation reflects an extremal property of orthonormal polynomials associated with an entropy-like criterion.

MSC: 44A60 14Q30 42C05 90C25 90C46 32U15 14P10

## 1. INTRODUCTION

The starting point of our investigation is polynomial Pell's equation

$$(1.1) \quad T_n(x)^2 + (1 - x^2)U_{n-1}(x)^2 = 1, \quad \forall x \in \mathbb{R}, \quad n \in \mathbb{N},$$

where  $T_n \in \mathbb{Z}[x]$  and  $U_n \in \mathbb{Z}[x]$  are the Chebyshev polynomial of the first and second kinds, respectively. We aim at extending this identity to multivariate polynomials in the form of

$$(1.2) \quad P_n^\mu(\mathbf{x}) + \sum_{i=1}^r \phi_i(\mathbf{x}) P_{n,i}^{\phi_i \cdot \mu}(\mathbf{x}) = 1, \quad \forall \mathbf{x} \in \mathbb{R}^d, \quad n \in \mathbb{N},$$

where  $\phi_i$ 's are some products of generators of a domain  $\Omega \subset \mathbb{R}^d$ ,  $P_n^\mu$  (resp.  $P_{n,i}^{\phi_i \cdot \mu}$ ) is the sum of the squares of all orthonormal polynomials of degree  $n$  (resp.  $n - \deg(\phi_i)$ ) with respect to the equilibrium measure  $\mu$  on  $\Omega$  (resp. the measure  $\phi_i \cdot \mu = \phi_i d\mu$  with the same support  $\Omega$ ).

A theoretical study of this problem has been carried out recently in [8], where it is pointed out that the problem is connected to several other topics, which we describe below as they motivate our study. We first define the Christoffel function. Let  $\mu$  be a measure supported on  $\Omega \in \mathbb{R}^d$ . For  $n \in \mathbb{N}$ , let  $\mathbb{R}[\mathbf{x}]_n$  be the space of polynomials of degree  $n$  in  $\mathbf{x} \in \mathbb{R}^d$ . The Christoffel function associated with  $\mu$ , denoted by  $\Lambda_n^\mu$ , is defined by

$$(1.3) \quad \Lambda_n^\mu(\boldsymbol{\xi}) := \inf_{p \in \mathbb{R}[\mathbf{x}]_n, p(\boldsymbol{\xi})=1} \int p(\mathbf{y})^2 d\mu(\mathbf{y}), \quad \forall \boldsymbol{\xi} \in \mathbb{R}^d.$$

---

The first author's work is supported by the AI Interdisciplinary Institute ANITI funding through the French program "Investing for the Future PIA3" under grant agreement number ANR-19-PI3A-0004 and the research is also part of the program DesCartes supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) program. The second author's work is partially supported by Simons Foundation Grant # 849676.

Let  $x \mapsto g(x) := 1 - x^2$ ,  $s(n) := \binom{1+n}{n}$ , and let  $\mu$  be the Chebyshev (equilibrium) measure  $dx/\pi\sqrt{1-x^2}$  of  $[-1, 1]$ , whereas  $g \cdot \mu := \sqrt{1-x^2}dx/\pi$ . After scaling, a summation of (1.1) leads to the identity

$$(1.4) \quad \Lambda_n^\mu(x)^{-1} + (1-x^2)\Lambda_{n-1}^{g \cdot \mu}(x)^{-1} = s(n) + s(n-1),$$

where  $\Lambda_n^\mu$  (resp.  $\Lambda_n^{g \cdot \mu}$ ) is the Christoffel function of degree  $n$  associated with  $\mu$  (resp.  $g \cdot \mu$ ). The Christoffel function can also be formulated in terms of the moments of  $\mu$ ; see the next section. For the Chebyshev measure, the formulation is of the form

$$\begin{aligned} \Lambda_n^\mu(x)^{-1} &= \mathbf{v}_n(x)^T \mathbf{M}_n(\mu)^{-1} \mathbf{v}_n(x), \quad \forall x \in \mathbb{R}, \\ \Lambda_{n-1}^{g \cdot \mu}(x)^{-1} &= \mathbf{v}_{n-1}(x)^T \mathbf{M}_{n-1}(g \cdot \mu)^{-1} \mathbf{v}_{n-1}(x), \quad \forall x \in \mathbb{R}, \end{aligned}$$

where  $\mathbf{v}_n(x) = (1, x, x^2, \dots, x^n)$ ,  $\mathbf{M}_n(\mu)$  is the matrix of moments up to degree  $2n$  of  $\mu$ , and  $\mathbf{M}_{n-1}(g \cdot \mu)$  is the matrix of moments up to degree  $2n-2$  of  $g \cdot \mu$ .

While polynomial Pell's equation (1.1) originates from Pell's equation in algebraic number theory, it also can be regarded from a different angle. In [8] it was observed that (1.1) (resp. (1.4)) is the *Markov-Lukács certificate* (resp. *Putinar certificate*) that the constant polynomial  $x \mapsto p(x) = 1$  for all  $x$ , is positive on  $[-1, 1]$ . Indeed while the former states that a polynomial  $p$  of degree  $2n$ , nonnegative on  $[-1, 1]$ , can be written in the form

$$p(x) = q_0(x)^2 + (1-x^2)q_1(x)^2, \quad \forall x \in \mathbb{R},$$

for two single squares of polynomials  $q_0$  of degree  $n$  and  $q_1$  of degree  $n-1$ , the latter states that

$$(1.5) \quad p(x) = \sigma_0(x) + (1-x^2)\sigma_1(x), \quad \forall x \in \mathbb{R},$$

for two polynomials  $\sigma_0$  of degree  $2n$  and  $\sigma_1$  of degree  $2n-2$  that are *sum-of-squares*. The latter, importantly, has a multivariate generalization [12] to compact basic semi-algebraic sets of  $\mathbb{R}^d$  whereas the former is specific to the univariate case. Furthermore, the reciprocals of the Christoffel functions  $\Lambda_n^\mu$  and  $\Lambda_n^{g \cdot \mu}$  satisfy an extremal property: Namely the couple  $(\mathbf{M}_n(\mu)^{-1}, \mathbf{M}_{n-1}(g \cdot \mu)^{-1})$  of their respective Gram matrices is the unique optimal solution for the optimization problem

$$(1.6) \quad \begin{aligned} &\max_{\mathbf{A}, \mathbf{B} \succ 0} \{ \log \det(\mathbf{A}) + \log \det(\mathbf{B}) : \\ &\text{s.t. } s(n) + s(n-1) = \underbrace{\mathbf{v}_n(x)^T \mathbf{A} \mathbf{v}_n(x)}_{\sigma_0(x)} + \underbrace{\mathbf{v}_{n-1}(x)^T \mathbf{B} \mathbf{v}_{n-1}(x)}_{\sigma_1(x)} (1-x^2) \\ &\quad \forall x \in \mathbb{R} \}, \end{aligned}$$

where the ‘sup’ is over all possible positive semidefinite Gram matrices  $\mathbf{A}$  and  $\mathbf{B}$  of the sum-of-squares polynomials  $\sigma_0$  and  $\sigma_1$  respectively. Interestingly, this is a direct consequence and interpretation of a result by Nesterov [11] on a one-to-one correspondence between the respective interiors of the convex cone of polynomials of the form (1.5) and its dual; see [8, 11].

**Contribution.** In [8], the identity (1.1) is extended to (1.2) in several bivariate cases, namely for set  $\Omega \subset \mathbb{R}^2$  being the Euclidean ball, the triangle, and the cube, but only for  $n = 1, 2, 3$ . In the present paper:

(i) We first show that these extensions hold for all  $n \geq 1$  and, furthermore, for all dimensions, not just for  $d = 2$ . Together, our results consist of a significant extension of (1.1) to the multivariate setting and provide the first set of examples for (1.2).

In particular, it shows that (1.4), as a partition of unity for  $[-1, 1]$ , has a natural multivariate analog for the regular domains in our list.

The proof of our extensions comes from a special property of multivariate orthogonal polynomials. More specifically, the property is an analog of the *addition formula* for spherical harmonics, which are homogeneous harmonic polynomials restricted on the unit sphere  $\mathbb{S}^d$  of  $\mathbb{R}^{d+1}$ . Indeed, let  $\mathcal{H}_n^{d+1}$  be the space of spherical harmonics of degree  $n$  and let  $\{Y_\ell^n : 1 \leq \ell \leq \dim \mathcal{H}_n^{d+1}\}$  be an orthonormal basis of  $\mathcal{H}_n^d$ ; then the addition formula states that

$$(1.7) \quad \sum_{\ell=1}^{\dim \mathcal{H}_n^{d+1}} Y_\ell^n(\boldsymbol{\xi}) Y_\ell^n(\boldsymbol{\eta}) = \frac{2n+d-1}{d-1} C_n^{\frac{d-1}{2}}(\langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle), \quad \boldsymbol{\xi}, \boldsymbol{\eta} \in \mathbb{S}^d,$$

where  $C_n^\lambda$  denotes the usual Gegenbauer polynomial of degree  $n$  and  $\langle \cdot, \cdot \rangle$  is the ordinary inner product in  $\mathbb{R}^d$ . It shows, in particular, that

$$\sum_{\ell=1}^{\dim \mathcal{H}_n^{d+1}} Y_\ell^n(\boldsymbol{\xi})^2 = \frac{2n+d-1}{d-1} C_n^{\frac{d-1}{2}}(1), \quad \forall \boldsymbol{\xi} \in \mathbb{S}^d,$$

which is a sum of square formula or an analog of (1.2). For the Chebyshev weight function on each of the regular domains in our list, we have an analog of the addition formula, which however is often equal to a sum on its right-hand side, instead of one term as in (1.7). Nevertheless, in each case, we shall show that an appropriate combination with associated orthogonal polynomials for a family of weight functions, involving generators of the set  $\Omega$ , leads to an identity that can be used to obtain (1.2). The identity is of interest in its own as it links all orthonormal polynomials of degree  $n$  for  $\mu$  with all orthonormal polynomials of degree  $n - \deg(\phi_i)$  for  $\phi_i \cdot \mu$ ,  $i = 1, \dots, r$ , where  $(\phi_i)_{i=1}^r$  is a set of generators of the domain  $\Omega$ .

Moreover, by summing over  $n = 0, 1, \dots, t$ , we obtain an identity satisfied by reciprocals of the Christoffel functions associated with measures  $\mu$  and  $\phi \cdot \mu$ . As an analogy with (1.4), it is fair to state that these Christoffel functions are solutions of a generalized polynomial Pell's equation for every fixed degree  $n \in \mathbb{N}$ .

(ii) In addition, the inverses of the associated moment matrices for these measures satisfy an extremal property since they also form the unique optimal solution of a multivariate version of the optimization problem (1.6).

(iii) Next, in the general case of a compact set  $\Omega$  with nonempty interior, and whose set of generators is given and fixed, we show that to its associated equilibrium measure  $\mu$  is associated a sequence of polynomials  $(p_n)_{n \in \mathbb{N}}$  of increasing degree, positive on the domain. We show that the sequence  $(p_n)$  is related to the constant polynomial  $\mathbf{1}$ :

- in a weak form when  $p_n$  is seen as a density with respect to  $\mu$ : the resulting sequence of probability measures  $(p_n \mu)_{n \in \mathbb{N}}$  converges weakly to  $\mu$ .
- in a stronger form under additional assumptions: the sequence  $(p_n)_{n \in \mathbb{N}}$  converges to  $\mathbf{1}$ , uniformly on compact subsets of  $\Omega$ .

Altogether, our generalization of (1.4) to certain multivariate settings reveals unexpected links with seemingly disconnected fields (orthogonal polynomials, equilibrium measures, certificates of positivity, a certain conic duality in optimization), which we hope will be of interest to researchers working in those fields.

The paper is organized as follows. After a brief section on notations and preliminaries, we discuss the background and motivation in the third section, then describe our

results in relatively simple terms of the Christoffel functions of appropriate measures on the family of semi-algebraic sets  $\Omega \subset \mathbb{R}^d$  that we consider in the fourth section. The proof will be given in the fifth section, which contains the full strength of the identity derived from the addition formula. The nature of the proof means it is more involved and technique, for which we will introduce additional notations and restate our results for easier access and clarity of exposition.

## 2. NOTATION AND PRELIMINARIES

**Notation and definition.** Let  $\mathbb{R}[\mathbf{x}]$  denote the ring of real polynomials in the variables  $\mathbf{x} = (x_1, \dots, x_d)$  and  $\mathbb{R}[\mathbf{x}]_n \subset \mathbb{R}[\mathbf{x}]$  be the subset of polynomials of total degree at most  $n$ . Denote by  $\mathbf{1} \in \mathbb{R}[\mathbf{x}]$  the polynomial equal to 1 for all  $\mathbf{x}$ . Let  $\mathbb{N}_n^d := \{\boldsymbol{\alpha} \in \mathbb{N}^d : |\boldsymbol{\alpha}| \leq n\}$ , where  $|\boldsymbol{\alpha}| = \sum_i \alpha_i$ , which has the cardinal  $s(n) := \binom{d+n}{d}$ . Let  $\mathbf{v}_n(\mathbf{x}) = (\mathbf{x}^\alpha)_{\alpha \in \mathbb{N}_n^d}$  be the vector of monomials up to degree  $n$ , where the monomials are listed in, say, the graded lexicographical order, and let  $\Sigma[\mathbf{x}] \subset \mathbb{R}[\mathbf{x}]$  (resp.  $\Sigma[\mathbf{x}]_n \subset \mathbb{R}[\mathbf{x}]_{2n}$ ) be the convex cone of polynomials (resp. of polynomials of total degree at most  $2n$ ) that are sum-of-squares (SOS in short). For every  $p \in \mathbb{R}[\mathbf{x}]_n$  write

$$\mathbf{x} \mapsto p(\mathbf{x}) = \langle \mathbf{p}, \mathbf{v}_n(\mathbf{x}) \rangle, \quad \forall \mathbf{x} \in \mathbb{R}^d,$$

where  $\mathbf{p} \in \mathbb{R}^{s(n)}$  is the vector of coefficients of  $p$  in the monomial basis  $(\mathbf{x}^\alpha)_{\alpha \in \mathbb{N}_n^d}$ . For a real symmetric matrix  $\mathbf{A} = \mathbf{A}^T$  the notation  $\mathbf{A} \succeq 0$  (resp.  $\mathbf{A} \succ 0$ ) stands for  $\mathbf{A}$  is positive semidefinite (p.s.d.) (resp. positive definite (p.d.)).

The support of a Borel measure  $\mu$  on  $\mathbb{R}^d$  is the smallest closed set  $A$  such that  $\mu(\mathbb{R}^d \setminus A) = 0$ , and such a set  $A$  is unique. With  $\Omega \subset \mathbb{R}^d$  compact, denote by  $\mathcal{C}(\Omega)$  the Banach space of real continuous functions on  $\Omega$  equipped with the sup-norm. Its topological dual  $\mathcal{C}(\Omega)^*$  is the Banach space  $\mathcal{M}(\Omega)$  of finite signed Borel measures on  $\Omega$ , equipped with the total-variation norm.

**Moment matrix.** Associated with  $n \in \mathbb{N}$  and a real vector  $\boldsymbol{\phi} = (\phi_\alpha)_{\alpha \in \mathbb{N}_{2n}^d}$ , are the (Riesz) linear functional  $\phi \in \mathbb{R}[\mathbf{x}]_{2n}^*$  defined by:

$$p \mapsto \phi(p) := \langle \mathbf{p}, \boldsymbol{\phi} \rangle = \sum_{\alpha \in \mathbb{N}_{2n}^d} p_\alpha \phi_\alpha, \quad \forall p \in \mathbb{R}[\mathbf{x}]_{2n},$$

and the real ‘‘moment’’ matrix  $\mathbf{M}_n(\boldsymbol{\phi})$  (or  $\mathbf{M}_n(\phi)$ ) with rows and columns indexed by  $\boldsymbol{\alpha} \in \mathbb{N}_n^d$  and with entries

$$\mathbf{M}_n(\boldsymbol{\phi})(\boldsymbol{\alpha}, \boldsymbol{\beta}) := \phi(\mathbf{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}}) = \phi_{\boldsymbol{\alpha}+\boldsymbol{\beta}} \quad \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_n^d.$$

If  $\mu$  is a finite Borel measure on  $\mathbb{R}^d$  with all moments  $\boldsymbol{\mu} = (\mu_\alpha)_{\alpha \in \mathbb{N}^d}$  assumed to be finite, then

$$\mathbf{M}_n(\mu)(\boldsymbol{\alpha}, \boldsymbol{\beta}) := \int \mathbf{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}} d\mu = \mu_{\boldsymbol{\alpha}+\boldsymbol{\beta}}, \quad \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_n^d,$$

and obviously,  $\mathbf{M}_n(\mu) \succeq 0$  for all  $n$  since

$$\langle \mathbf{p}, \mathbf{M}_n(\mu) \mathbf{p} \rangle = \int p^2 d\mu \geq 0, \quad \forall p \in \mathbb{R}[\mathbf{x}]_n.$$

On the other hand, given a real vector  $\boldsymbol{\phi} = (\phi_\alpha)_{\alpha \in \mathbb{N}_{2n}^d}$ ,  $\mathbf{M}_n(\boldsymbol{\phi}) \succeq 0$  is only a necessary condition for the associated linear functional  $\phi \in \mathbb{R}[\mathbf{x}]_{2n}^*$  to have a representing measure, i.e.,

$$\phi_\alpha = \phi(\mathbf{x}^\alpha) = \int \mathbf{x}^\alpha d\varphi, \quad \forall \alpha \in \mathbb{N}_{2n}^d,$$

for some Borel measure  $\varphi$  on  $\mathbb{R}^d$ .

**Localizing matrix.** Given a real sequence  $\phi \in \mathbb{N}_{2n}^d$  and a polynomial  $g \in \mathbb{R}[\mathbf{x}]$ , one may define the new real sequence  $g \cdot \phi \in \mathbb{R}^{s(2n) - \deg(g)}$  (and associated Riesz linear functional  $g \cdot \phi \in \mathbb{R}[\mathbf{x}]_{2n - \deg(g)}^*$ ) by:

$$(g \cdot \phi)_\alpha = g \cdot \phi(\mathbf{x}^\alpha) := \phi(g \mathbf{x}^\alpha) = \sum_{\beta \in \mathbb{N}_n^d} g_\beta \phi_{\alpha+\beta}, \quad \forall \alpha \in \mathbb{N}_{2n - \deg(g)}^d.$$

Then, setting  $t_g := \lceil \deg(g)/2 \rceil (= \lceil t/2 \rceil)$ , the *localizing* matrix  $\mathbf{M}_{n-t_g}(g \cdot \phi)$  ( $n \geq t_g$ ) associated with  $g$  and  $\phi$  is just the moment matrix  $\mathbf{M}_{n-t_g}(g \cdot \phi)$  associated with the (pseudo) moment vector  $g \cdot \phi \in \mathbb{R}^{s(2n) - \deg(g)}$ .

**Orthogonal polynomials and their kernels.** Here we assume  $\mathbf{M}_n(\mu) \succ 0$  for all  $n \in \mathbb{N}$ , and therefore the inverse  $\mathbf{M}_n(\mu)^{-1}$  is well-defined for all  $n \in \mathbb{N}$ . In particular, this is true in our case of interest, i.e., when the support  $\Omega \subset \mathbb{R}^d$  of  $\mu$  is compact with a nonempty interior and  $\mu$  has a density with respect to the Lebesgue measure on  $\Omega$ . Under the assumption, a sequence of orthogonal polynomials  $(P_\alpha)_{\alpha \in \mathbb{N}^d} \subset \mathbb{R}[\mathbf{x}]$  exists in  $L^2(\mu)$ .

Let  $(P_\alpha)_{\alpha \in \mathbb{N}^d} \subset \mathbb{R}[\mathbf{x}]$  be a family of polynomials, with the degree of  $P_\alpha$  being  $|\alpha|$ , that are orthonormal with respect to  $\mu$ ; that is,

$$\int_{\mathbb{R}^d} P_\alpha(\mathbf{x}) P_\beta(\mathbf{x}) d\mu(\mathbf{x}) = \delta_{\alpha,\beta}, \quad \alpha, \beta \in \mathbb{N}^d.$$

Let  $\mathcal{V}_m^d$  be the space of orthogonal polynomials of degree exactly  $m$ . Let  $\{P_\alpha : |\alpha| = m\}$  be an orthonormal basis of  $\mathcal{V}_m^d$ . Then the reproducing kernel of the space  $\mathcal{V}_m^d$  in  $L^2(\mu)$  is defined by

$$(2.1) \quad (\mathbf{x}, \mathbf{y}) \mapsto P_m^\mu(\mathbf{x}, \mathbf{y}) := \sum_{|\alpha|=m} P_\alpha(\mathbf{x}) P_\alpha(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y}, \quad m \in \mathbb{N},$$

which is independent of the choice of bases. Summing over  $0 \leq m \leq n$  gives the reproducing kernel of  $\mathbb{R}[\mathbf{x}]_n$  in the space  $L^2(\mu)$ ,

$$(2.2) \quad K_n^\mu(\mathbf{x}, \mathbf{y}) = \sum_{j=0}^n P_j^\mu(\mathbf{x}, \mathbf{y}) = \sum_{|\alpha| \leq n} P_\alpha(\mathbf{x}) P_\alpha(\mathbf{y}),$$

which is sometimes called the Christoffel-Darboux kernel.

**Christoffel function.** The Christoffel function  $\Lambda_n^\mu : \mathbb{R} \rightarrow \mathbb{R}_+$  of degree  $n$ , associated with  $\mu$ , is defined by (1.3) and it can also be defined in terms of the moment matrix by

$$\mathbf{x} \mapsto \Lambda_n^\mu(\mathbf{x})^{-1} = \mathbf{v}_n(\mathbf{x})^T \mathbf{M}_n(\mu)^{-1} \mathbf{v}_n(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

Alternatively, in terms of the kernels  $K_n^\mu(\cdot, \cdot)$ , the Christoffel function satisfies

$$(2.3) \quad \Lambda_n^\mu(\mathbf{x})^{-1} = K_n^\mu(\mathbf{x}, \mathbf{x}) = \sum_{\alpha \in \mathbb{N}_n^d} P_\alpha(\mathbf{x})^2, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

It should be emphasized that the above identity holds for the nonnegative measure  $\mu$ , which is not necessarily a probability measure.

In its variational characterization (1.3), the Christoffel function is the optimal value of a quadratic convex optimization problem that can be solved efficiently. Its unique optimal solution  $p^* \in \mathbb{R}[\mathbf{x}]_n$  reads:

$$\mathbf{x} \mapsto p^*(\mathbf{x}) := \frac{\sum_{\alpha \in \mathbb{N}_n^d} P_\alpha(\boldsymbol{\xi}) P_\alpha(\mathbf{x})}{\sum_{\alpha \in \mathbb{N}_n^d} P_\alpha(\boldsymbol{\xi})^2} = \frac{K_n^\mu(\boldsymbol{\xi}, \mathbf{x})}{K_n^\mu(\boldsymbol{\xi}, \boldsymbol{\xi})}, \quad \mathbf{x} \in \mathbb{R}^d.$$

*Remark 2.1.* As  $\Lambda_n^\mu$  and  $K_n^\mu$  depend only on moments  $\boldsymbol{\mu} = (\mu_\alpha)_{\alpha \in \mathbb{N}_{2n}^d}$  of  $\mu$ , one may also define exactly in the same manner, the Christoffel function  $\Lambda_n^\phi$  and the CD-kernel  $K_n^\phi$  associated with a real (pseudo)-moment vector  $\phi \in \mathbb{R}^{s(2n)}$  such that  $\mathbf{M}_n(\phi) \succ 0$ .

**Equilibrium measure.** The notion of equilibrium measure associated with a given set originates from logarithmic potential theory (working in  $\mathbb{C}$  in the univariate case) to minimize some energy functional. For instance, the equilibrium (Chebyshev) measure  $d\phi := dx/\pi\sqrt{1-x^2}$  minimizes the Riesz  $s$ -energy functional

$$\int \int \frac{1}{|x-y|^s} d\mu(x) d\mu(y)$$

with  $s = 2$ , among all measures  $\mu$  equivalent to  $\phi$ . Some generalizations have been obtained in the multivariate case via pluripotential theory in  $\mathbb{C}^n$ . In particular, if  $\Omega \subset \mathbb{R}^n \subset \mathbb{C}^n$  is compact then the equilibrium measure (let us denote it by  $\mu$ ) is equivalent to the Lebesgue measure on compact subsets of  $\text{int}(\Omega)$ . It has an even explicit expression if  $\Omega$  is convex and symmetric about the origin; see e.g. Bedford and Taylor [2, Theorems 1.1 and 1.2]. Several examples of sets  $\Omega$  with its equilibrium measure given in explicit form can be found in Baran [1]. A Borel measure  $\phi$  supported on a compact set  $\Omega \subset \mathbb{R}^d$  is said to satisfy the Bernstein-Markov property if there exists a sequence of positive numbers  $(M_n)_{n \in \mathbb{N}}$  such that for all  $n$  and  $p \in \mathbb{R}[\mathbf{x}]_n$ ,

$$\sup_{\mathbf{x} \in \Omega} |p(\mathbf{x})| \leq M_n \cdot \left( \int_{\Omega} p^2 d\phi \right)^{1/2} \quad \text{and} \quad \lim_{n \rightarrow \infty} \log(M_n)/n = 0;$$

see e.g. [9, Section 4.3.3]. If a Borel measure  $\phi$  on  $\Omega$  has the Bernstein-Markov property, then the sequence of measures  $d\nu_n = \frac{d\phi(\mathbf{x})}{s(n)\Lambda_n^\phi(\mathbf{x})}$ ,  $n \in \mathbb{N}$ , converges to  $\mu$  for the weak- $\star$  topology and, in particular,

$$(2.4) \quad \lim_{n \rightarrow \infty} \int_{\Omega} \mathbf{x}^\alpha d\nu_n = \lim_{n \rightarrow \infty} \int_{\Omega} \frac{\mathbf{x}^\alpha d\phi(\mathbf{x})}{s(n)\Lambda_n^\phi(\mathbf{x})} = \int_{\Omega} \mathbf{x}^\alpha d\mu, \quad \forall \alpha \in \mathbb{N}^d$$

(see e.g. [9, Theorem 4.4.4]). In addition, if a compact  $\Omega \subset \mathbb{R}^d$  is regular then  $(\Omega, \mu)$  has the Bernstein-Markov property; see [9, p. 59]. For a brief account of equilibrium measures, see [1, 2] and the discussion in [9, Section 4-5, pp. 56–60], while more detailed expositions can be found in some of the references indicated there.

### 3. BACKGROUND AND MOTIVATION

In this section, we discuss a duality result and its connection with the convex optimization problems, which motivates our study.

Let  $g_0 := \mathbf{1}$  be the constant polynomial equal to 1, and let  $G \subset \mathbb{R}[\mathbf{x}]$  be a finite set of polynomials that contains  $g_0$ . For every  $g \in \mathbb{R}[\mathbf{x}]$ , let  $t_g := \lceil \deg(g)/2 \rceil$  and let  $Q_n(g)$  be the convex cone associated with  $G$ , defined by:

$$(3.1) \quad Q_n(g) := \left\{ \sum_{g \in G} \sigma_g g : \sigma_g \in \Sigma[\mathbf{x}]_{n-t_g} \right\}, \quad n \in \mathbb{N}.$$

Clearly, every polynomial  $p \in Q_n(G)$  is nonnegative on the set

$$(3.2) \quad \Omega := \{ \mathbf{x} \in \mathbb{R}^d : g(\mathbf{x}) \geq 0, \forall g \in G \}.$$

This is the reason that one may say that the sum-of-squares weights  $(\sigma_g^*)_{g \in G}$ , in the representation  $p = \sum_{g \in G} \sigma_g^* g$  of  $p$ , provide  $p$  with an algebraic certificate of its positivity on  $\Omega$ . Such a representation of  $p$ , however, is not unique in general and, consequently, neither is such a positivity certificate of  $p$  unique.

The associated dual convex cone  $Q_n(G)^* \subset \mathbb{R}[\mathbf{x}]_{2n}^*$  of  $Q_n(G)$  is defined by

$$(3.3) \quad Q_n(G)^* := \left\{ \phi \in \mathbb{R}^{s(2n)} : \mathbf{M}_{n-t_g}(g \cdot \phi) \succeq 0, \quad g \in G \right\}.$$

**Lemma 3.1** ([8, 11]). *If  $p \in \text{int}(Q_n(G))$  then there is some  $\phi \in \text{int}(Q_n(G)^*)$  such that*

$$(3.4) \quad p(\mathbf{x}) = \sum_{g \in G} \mathbf{v}_{n-t_g}(\mathbf{x})^T \mathbf{M}_{n-t_g}(g \cdot \phi)^{-1} \mathbf{v}_{n-t_g}(\mathbf{x}) g(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d$$

$$(3.5) \quad = \sum_{g \in G} \Lambda_{n-t_g}^{g \cdot \phi}(\mathbf{x})^{-1} g(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

Eq. (3.4) is due to Nesterov [11] while (3.5) is its interpretation in terms of Christoffel functions of appropriate linear functionals; see [8] for more details. Notice that (3.4) provides  $p$  with a *distinguished* certificate of its positivity on the set  $\{ \mathbf{x} \in \mathbb{R}^d : g(\mathbf{x}) \geq 0, \forall g \in G \}$ , stated in terms of a specific linear functional  $\phi_{2n}^* \in \text{int}(Q_n(G)^*)$ . An immediately question arises: *What is the relation between  $\phi_{2n}^*$  and  $p$ ?* We will provide an answer for the case when  $p = \mathbf{1}$  for certain sets  $\Omega \subset \mathbb{R}^d$  with specific geometry.

Next, we consider the following two convex optimization problems. For every  $n \in \mathbb{N}$  and a subset  $G_n \subseteq G$ , the first problem states

$$(3.6) \quad \rho_n = \inf_{\phi \in \mathbb{R}^{s(2n)}} \left\{ - \sum_{g \in G_n} \log \det(\mathbf{M}_{n-t_g}(g \cdot \phi)) : \right. \\ \left. \phi(\mathbf{1}) = 1, \mathbf{M}_{n-t_g}(g \cdot \phi) \succeq 0, \forall g \in G_n \right\},$$

and the second problem states

$$(3.7) \quad \rho_n^* = \sup_{\mathbf{Q}_g} \left\{ \sum_{g \in G_n} \log \det(\mathbf{Q}_g) : \mathbf{Q}_g \succeq 0, \forall g \in G_n, \right. \\ \left. \sum_{g \in G_n} s(n-t_g) = \sum_{g \in G_n} g(\mathbf{x}) \cdot \mathbf{v}_{n-t_g}(\mathbf{x})^T \mathbf{Q}_g \mathbf{v}_{n-t_g}(\mathbf{x}), \forall \mathbf{x} \in \mathbb{R}^d \right\}$$

where the supremum is taken over real symmetric matrices  $(\mathbf{Q}_g)_{g \in G_n}$  of respective sizes  $s(n-t_g)$ . These two optimization problems are closely related.

**Theorem 3.2.** *With  $n \in \mathbb{N}$  and  $G_n \subseteq G$  fixed, Problems (3.6) and (3.7) have the same finite optimal value  $\rho_n = \rho_n^*$  if and only if  $\mathbf{1} \in \text{int}(Q_n(G_n))$ . Moreover, both have a unique optimal solution  $\phi_{2n}^* \in \mathbb{R}^{s(2n)}$  and  $(\mathbf{Q}_g^*)_{g \in G_n}$  respectively, which satisfy*



$\mathbf{Q}_g^* = \mathbf{M}_{n-t_g}(g \cdot \phi_{2n}^*)^{-1}$  for all  $g \in G_n$ . And, as a consequence,

$$(3.8) \quad \begin{aligned} 1 &= \frac{1}{\sum_{g \in G_n} s(n-t_g)} \sum_{g \in G_n} g(\mathbf{x}) \mathbf{v}_{n-t_g}(\mathbf{x})^T \mathbf{M}_{n-t_g}(g \cdot \phi_{2n}^*)^{-1} \mathbf{v}_{n-t_g}(\mathbf{x}) \\ &= \frac{1}{\sum_{g \in G_n} s(n-t_g)} \sum_{g \in G_n} g(\mathbf{x}) \Lambda_{n-t_g}^{g \cdot \phi_{2n}^*}(\mathbf{x})^{-1}, \quad \forall \mathbf{x} \in \mathbb{R}^d. \end{aligned}$$

Once again, one may interpret (3.8) as providing the constant polynomial  $\mathbf{1}$  with a specific algebraic certificate of its positivity on the set  $\Omega := \{\mathbf{x} \in \mathbb{R}^d : g(\mathbf{x}) \geq 0, \forall g \in G_n\}$ . It also provides  $\Omega$  with a polynomial *partition of unity* in terms of its generators  $g \in G_n$ . Finally, it is worth noticing that both (3.6) and (3.7) are specific convex optimization problems with a ‘‘log det’’ criterion, which can be solved with off-the-shelf software such as CVX [6] or Julia.

An interesting and very specific case is when the (unique) optimal solution  $\phi_{2n}^*$  in Theorem 3.2, for some  $n_0 \in \mathbb{N}$  and every  $n \geq n_0$ , is the vector of moments up to degree  $2n$  of a same unique measure  $\mu$  on a set  $\Omega := \{\mathbf{x} \in \mathbb{R}^d : g(\mathbf{x}) \geq 0, \forall g \in G\}$ . In fact,  $\Omega$  can be defined by a smaller set  $\tilde{G}$  of generators while  $G$  is made of some products of generators in  $\tilde{G}$ . This is because (3.8) may not hold with  $\tilde{G}$  while it sometimes holds with a subset  $G_n \subseteq G$ ; however, for sufficiently large  $n$ ,  $G_n = G$ . It then turns out that  $\mu$  is necessarily the equilibrium measure of  $\Omega$ .

**Assumption 3.3.** *The set  $\Omega$  in (3.2) is compact with nonempty interior. Moreover, there exists  $R > 0$  such that the quadratic polynomial  $\mathbf{x} \mapsto \theta(\mathbf{x}) := R - \|\mathbf{x}\|^2$  is an element of  $Q_1(G)$ . In other words,  $h \in Q_1(G)$  is an ‘‘algebraic certificate’’ that  $\Omega$  is compact.*

**Theorem 3.4.** *With  $\Omega$  as in (3.2), let Assumption 3.3 hold. Let  $\phi = (\phi_\alpha)_{\alpha \in \mathbb{N}^n}$  (with  $\phi_0 = 1$ ) be such that  $\mathbf{M}_n(g \cdot \phi) \succ 0$  for all  $n \in \mathbb{N}$  and all  $g \in G$ , so that the Christoffel functions  $\Lambda_n^{g \cdot \phi}$  are all well defined (recall that  $\phi \in \mathbb{R}[\mathbf{x}]^*$  is the Riesz linear functional associated with the moment sequence  $\phi$ ). In addition, suppose that there exists  $n_0 \in \mathbb{N}$  such that*

$$(3.9) \quad 1 = \frac{1}{\sum_{g \in G_n} s(n-t_g)} \sum_{g \in G_n} g \cdot (\Lambda_{n-t_g}^{g \cdot \phi})^{-1}, \quad \forall n \geq n_0,$$

for some subset  $G_n \subseteq G$  for all  $n \geq n_0$  (with  $G_n = G$  for  $n$  sufficiently large). Then  $\phi$  is a Borel measure on  $\Omega$  and it is the unique representing measure of  $\phi$ . Moreover, if  $(\Omega, g \cdot \phi)$  satisfies the Bernstein-Markov property for every  $g \in G$ , then  $\phi = \mu$  (the equilibrium measure of  $\Omega$ ) and therefore the Christoffel polynomials  $(\Lambda_n^{g \cdot \mu})_{g \in G_n}^{-1}$  satisfy the generalized Pell’s equations

$$(3.10) \quad 1 = \frac{1}{\sum_{g \in G_n} s(n-t_g)} \sum_{g \in G_n} g \cdot (\Lambda_{n-t_g}^{g \cdot \mu})^{-1}, \quad \forall n \geq n_0.$$

The prototype example of Theorem 3.4 is the interval  $\Omega := [-1, 1]$  with Chebyshev (equilibrium) measure  $\mu = dx/\pi\sqrt{1-x^2}$ . Our main result in this paper is to identify three cases of sets  $\Omega \subset \mathbb{R}^d$  for which this is precisely the case for *all*  $n$  and in *any* dimension  $d$ , as suggested in [8] where they are verified only for  $d = 2$  and  $n = 1, 2, 3$ .

*Remark 3.1.* The equilibrium measure is a probability measure. In the identity (3.9), however, the measure  $g \cdot \mu$  is not normalized as a probability measure. This is important

since the value of the Christoffel function depends on the normalization. Indeed, if  $d\mu$  is a measure and  $\gamma$  is a positive constant, then it is easy to verify that

$$\Lambda_n^{\gamma \cdot \mu}(\mathbf{x}) = \gamma \Lambda_n^\mu(\mathbf{x}).$$

#### 4. MAIN RESULT

In this section, we state our main results on (1.2) for several compact sets  $\Omega$  with associated equilibrium measure denoted  $\mu$ . That is, we identify three cases of sets  $\Omega \subset \mathbb{R}^d$  for which Theorem 3.4 holds in any dimension  $d$ . Since the properties of the orthogonal structure remain valid under the affine transformation, we shall state our results only for regular domains. More precisely, we consider

**Unit ball:**  $\Omega = \mathbb{B}^d = \{\mathbf{x} \in \mathbb{R}^d : 1 - \|\mathbf{x}\|^2 \geq 0\}$  and  $\mu$  is proportional to

$$(4.1) \quad \frac{d\mathbf{x}}{\sqrt{1 - \|\mathbf{x}\|^2}}.$$

**Simplex:**  $\Omega = \Delta^d = \{\mathbf{x} \in \mathbb{R}_+^d : 1 - \sum_{j=1}^d x_j \geq 0\}$  and  $\mu$  is proportional to

$$(4.2) \quad \frac{d\mathbf{x}}{\sqrt{x_1} \cdots \sqrt{x_d} \sqrt{1 - \sum_{j=1}^d x_j}}.$$

**Unit cube:**  $\Omega = \square^d := [-1, 1]^d$  and  $\mu$  is proportional to

$$(4.3) \quad \frac{d\mathbf{x}}{\sqrt{1 - x_1^2} \cdots \sqrt{1 - x_d^2}}.$$

Below we state and discuss our main result for each case in a subsection. The proof is postponed to Section 5.

**4.1. On the unit ball.** Let  $\Omega = \mathbb{B}^d$  with associated equilibrium measure  $\mu$  as in (4.1), and let  $\mathbf{x} \mapsto g(\mathbf{x}) := 1 - \|\mathbf{x}\|^2$ , for all  $\mathbf{x} \in \mathbb{R}^d$ .

**Theorem 4.1.** *Let  $\mu$  be the equilibrium measure of  $\mathbb{B}^d$ , normalized to be a probability measure. Then, for every  $m \in \mathbb{N}$ ,*

$$(4.4) \quad P_m^\mu(\mathbf{x}, \mathbf{x}) + g(\mathbf{x}) P_{m-1}^{g \cdot \mu}(\mathbf{x}, \mathbf{x}) = \binom{d+m-1}{d-1} + \binom{d+m-2}{d-1}, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

And as a consequence, for every  $n \in \mathbb{N}$ :

$$(4.5) \quad \Lambda_n^\mu(\mathbf{x})^{-1} + g(\mathbf{x}) \Lambda_{n-1}^{g \cdot \mu}(\mathbf{x})^{-1} = \binom{d+n}{d} + \binom{d+n-1}{d}, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

As noted in Remark 3.1, the weight function  $g \cdot \mu$  is not normalized.

We note that the identity (4.4) is the exact multivariate analog for the unit ball  $\mathbb{B}^d$  of the univariate polynomial Pell's equation (1.1) for  $[-1, 1]$ . It is now a property of all polynomials of degree  $m$  orthonormal with respect to  $\mu$  and all polynomials of degree  $m-1$  orthonormal with respect to  $g \cdot \mu$ . Moreover, the identity (4.5) is a generalized Pell's equation satisfied by the Christoffel functions  $\Lambda_n^\mu$  and  $\Lambda_n^{g \cdot \mu}$ . In other words, we have proved that in any dimension  $d$ , the unit ball is an instance of a set  $\Omega \subset \mathbb{R}^d$  for which Theorem 3.4 holds. In this case  $G_n = G = \{g\}$  for all  $n$ , with  $\mathbf{x} \mapsto g(\mathbf{x}) = 1 - \|\mathbf{x}\|^2$ .

**Corollary 4.2.** *Let  $\mu$  be the equilibrium measure of  $\mathbb{B}^d$  normalized to be a probability measure. Then, for every  $n$ ,*

$$(4.6) \quad \inf_{\mathbf{x} \in \mathbb{B}^d} \Lambda_n^\mu(\mathbf{x}) = \min_{\mathbf{x} \in \mathbb{B}^d} \Lambda_n^\mu(\mathbf{x}) = \left[ \binom{d+n}{d} + \binom{d+n-1}{d} \right]^{-1} =: \frac{1}{\gamma_n},$$

and the minimum is attained at all points of the boundary  $\mathbb{S}^{d-1}$  of  $\mathbb{B}^d$ . Moreover

$$\mathbb{B}^d \subset \{\mathbf{x} : \Lambda_n^\mu(\mathbf{x}) \geq 1/\gamma_n\}, \quad \forall n \in \mathbb{N}.$$

*Proof.* By (4.5),

$$\binom{d+n}{d} + \binom{d+n-1}{d} - \Lambda_n^\mu(\mathbf{x})^{-1} = g(\mathbf{x}) \Lambda_{n-1}^{g \cdot \mu}(\mathbf{x})^{-1} \geq 0, \quad \forall \mathbf{x} \in \mathbb{B}^d,$$

and the right-hand side vanishes on  $\mathbb{S}^{d-1}$  because  $g(\mathbf{x}) = 0$  for all  $\mathbf{x} \in \mathbb{S}^{d-1}$ .  $\square$

The above corollary states the polynomial sublevel set  $\{\mathbf{x} : \Lambda_n^\mu(\mathbf{x})^{-1} \leq \gamma_n\}$  is an outer-approximation of  $\mathbb{B}^d$  for all  $n \in \mathbb{N}$ . It also states that the polynomial  $\mathbf{x} \mapsto \gamma_n - \Lambda_n^\mu(\mathbf{x})^{-1}$  is in the ideal  $\langle g \rangle$  generated by  $g$ .

**4.2. On the Simplex.** Let  $\Delta^d = \{\mathbf{x} \in \mathbb{R}_+^d : \sum_{i=1}^d x_i \leq 1\}$ , with associated equilibrium  $\mu$  as in (4.2). With  $\varepsilon \in \{0, 1\}^{d+1}$  and  $|\mathbf{x}| = \sum_{i=1}^d x_i$ , define

$$(4.7) \quad \mathbf{x} \mapsto g_\varepsilon(\mathbf{x}) := x_1^{\varepsilon_1} \cdots x_d^{\varepsilon_d} \cdot (1 - |\mathbf{x}|)^{\varepsilon_{d+1}}, \quad \forall \mathbf{x} \in \mathbb{R}^d,$$

and introduce the vector  $\mathbf{1} := (1, 1, \dots, 1) \in \mathbb{N}^{d+1}$ .

**Theorem 4.3.** *Let  $\mu$  be the equilibrium measure of  $\Delta^d$ , normalized to be a probability measure. Then for every  $m \in 2\mathbb{N}$  and  $\mathbf{x} \in \mathbb{R}^d$ ,*

$$(4.8) \quad \sum_{\varepsilon \in \{0,1\}^{d+1}; |\varepsilon| \in 2\mathbb{N}; |\varepsilon| \leq 2m} g_\varepsilon(\mathbf{x}) P_{m-|\varepsilon|/2}^{g_\varepsilon \cdot \mu}(\mathbf{x}, \mathbf{x}) = \binom{2n+d-1}{d-1} + \binom{2n+d-2}{d-1}.$$

And as a consequence, for every  $n \in \mathbb{N}$  and  $\mathbf{x} \in \mathbb{R}^d$ ,

$$(4.9) \quad \sum_{\varepsilon \in \{0,1\}^{d+1}; |\varepsilon| \in 2\mathbb{N}; |\varepsilon| \leq n} g_\varepsilon(\mathbf{x}) \Lambda_{n-|\varepsilon|/2}^{g_\varepsilon \cdot \mu}(\mathbf{x})^{-1} = \binom{2n+d}{d},$$

where  $(a)_n = a(a+1) \cdots (a+n-1)$  is the usual Pochhammer symbol.

Again we have proved that in any dimension  $d$ , the simplex is an instance of a set  $\Omega \subset \mathbb{R}^d$  for which Theorem 3.4 holds. Here  $G = \{g_\varepsilon : \varepsilon \in \{0, 1\}^d\}$  with  $g_\varepsilon$  as in (4.7), and  $G_n = \{g_\varepsilon \in G : \varepsilon \in 2\mathbb{N}; |\varepsilon| \leq n\}$  so that  $G_n = G$  for sufficiently large  $n$ .

As an example, the identity (4.9) for  $d = 2$  is given by

$$\begin{aligned} \Lambda_n^\mu(\mathbf{x})^{-1} + x_1 x_2 \Lambda_{n-1}^{x_1 x_2 \cdot \mu}(\mathbf{x})^{-1} + x_1 x_3 \Lambda_{n-1}^{x_1 x_3 \cdot \mu}(\mathbf{x})^{-1} \\ + x_2 x_3 \Lambda_{n-1}^{x_2 x_3 \cdot \mu}(\mathbf{x})^{-1} = (n+1)(2n+1), \end{aligned}$$

where  $\mathbf{x} = (x_1, x_2)$  and  $x_3 = 1 - x_1 - x_2$ . In this example, as well as in the Theorem 4.3, we should keep in mind that  $g_\varepsilon \cdot \mu$  is not normalized, as shown in Remark 3.1.

We also have an analog of Corollary 4.2 that follows from a similar proof.

**Corollary 4.4.** *Let  $\mu$  be the equilibrium measure of  $\Delta^d$ , normalized to be a probability measure. Then, for every  $n$ ,*

$$(4.10) \quad \inf_{\mathbf{x} \in \Delta^d} \Lambda_n^\mu(\mathbf{x}) = \min_{\mathbf{x} \in \Delta^d} \Lambda_n^\mu(\mathbf{x}) = \left[ \binom{2n+d}{d} \right]^{-1} =: \frac{1}{\gamma_n},$$

and the minimum is attained at all points of the boundary  $\partial\Delta^d$  of  $\Delta^d$ . Moreover

$$\Delta^d \subset \{\mathbf{x} : \Lambda_n^\mu(\mathbf{x}) \geq 1/\gamma_n\}, \quad \forall n \in \mathbb{N}.$$

Again, Corollary 4.4 states that the polynomial sublevel set  $\{\mathbf{x} : \Lambda_n^\mu(\mathbf{x})^{-1} \leq \gamma_n\}$  is an outer-approximation of the simplex  $\Delta^d$  for all  $n \in \mathbb{N}$ . It also states that the polynomial  $\mathbf{x} \mapsto \gamma_n - \Lambda_n^\mu(\mathbf{x})^{-1}$  is in the ideal generated by all the polynomials  $g_\varepsilon$  in (4.9).

**4.3. The unit cube.** Let  $\square^d = [-1, 1]^d$  with associated equilibrium measure  $\mu$  as in (4.3), and for every  $\varepsilon \in \{0, 1\}^d$ , define

$$(4.11) \quad \mathbf{x} \mapsto g_\varepsilon(\mathbf{x}) := \prod_{j=1}^d (1 - x_j^2)^{\varepsilon_j}, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

**Theorem 4.5.** *Let  $\mu$  be the equilibrium measure of  $\square^d$ , normalized to be a probability measure. Then, for every  $m \in \mathbb{N}$  and  $\mathbf{x} \in \mathbb{R}^d$ ,*

$$(4.12) \quad \sum_{\varepsilon \in \{0,1\}^d; |\varepsilon| \leq m} g_\varepsilon(\mathbf{x}) P_{m-|\varepsilon|}^{g_\varepsilon \cdot \mu}(\mathbf{x}, \mathbf{x}) = \sum_{j=0}^d \binom{d}{j} \left[ \binom{d+n-j}{d} - \binom{d+n-1-j}{d} \right]$$

and, as a consequence, for every  $n \in \mathbb{N}$ :

$$(4.13) \quad \sum_{\varepsilon \in \{0,1\}^d; |\varepsilon| \leq n} g_\varepsilon(\mathbf{x}) \Lambda_{n-|\varepsilon|}^{g_\varepsilon \cdot \mu}(\mathbf{x})^{-1} = \sum_{j=0}^d \binom{d}{j} \binom{d+n-j}{d}, \quad \forall \mathbf{x} \in \mathbb{R}^d.$$

As in the case for the unit ball, the identity (4.12) is satisfied by all polynomials of degree  $m - \varepsilon$ , orthonormal with respect to the measures  $g_\varepsilon \cdot \mu$ ,  $\varepsilon \in \{0, 1\}^d$  and  $|\varepsilon| \leq m$ . Moreover, the identity (4.13) is a generalized Pell's equation satisfied by the Christoffel functions  $(\Lambda_{n-|\varepsilon|}^{g_\varepsilon \cdot \mu})_{\varepsilon \in \{0,1\}^d; |\varepsilon| \leq n}$ .

Thus, we have proved that in any dimension  $d$ , the cube is also an instance of a set  $\Omega \subset \mathbb{R}^d$  for which Theorem 3.4 holds. Here  $G = \{g_\varepsilon : \varepsilon \in \{0, 1\}^d\}$  with  $g_\varepsilon$  as in (4.11), and  $G_n = \{g_\varepsilon \in G : |\varepsilon| \leq n\}$  so that  $G_n = G$  as soon as  $n \geq d$ .

As an example, (4.13) with  $d = 2$  reads:

$$\begin{aligned} \Lambda_n^\mu(\mathbf{x})^{-1} &+ (1 - x_1^2) \Lambda_{n-1}^{(1-x_1^2) \cdot \mu}(\mathbf{x})^{-1} + (1 - x_2^2) \Lambda_{n-1}^{(1-x_2^2) \cdot \mu}(\mathbf{x})^{-1} \\ &+ (1 - x_1^2)(1 - x_2^2) \Lambda_{n-2}^{(1-x_1^2)(1-x_2^2) \cdot \mu}(\mathbf{x})^{-1} = 1 + 2n(n+1). \end{aligned}$$

Thus, we have established the results initiated in [8] for *all degrees*  $n$  and *all dimensions*  $d$ , and we have demonstrated in the three cases listed above that, for every  $n$ , the Christoffel functions  $(\Lambda_n^{g \cdot \mu})_{g \in G}$  satisfy the generalized Pell's identity

$$(4.14) \quad \sum_{g \in G_n} s(n - t_g) = \sum_{g \in G_n} g(\mathbf{x}) \Lambda_n^{g \cdot \mu}(\mathbf{x})^{-1}, \quad \forall \mathbf{x} \in \mathbb{R}^d, \quad \forall n \in \mathbb{N}.$$

It is important to emphasize that the set  $G$  of generators is crucial. Indeed, in the case of the simplex for instance, if we take  $G = \{1, x_1, \dots, x_d, 1 - \sum_j x_j\}$  then (3.10) cannot hold. A similar conclusion holds for the box with  $G = \{1, 1 - x_1^2, \dots, 1 - x_d^2\}$  as products of such generators are needed.

**4.4. Further examples.** As shown in the next section, our proof relies on analog to addition formulas on the domain, which provides a powerful tool whenever it exists. Analogs of the addition formula are known to exist for orthogonal polynomials on hyperbolic surfaces and hyperboloids, but only for the subclass of orthogonal polynomials that are even in the variable on the axis of rotation [15]. With an appropriate extension of our setup to the subclass of orthogonal polynomials, it is possible to establish analogs to our main results on the hyperbolic surfaces and hyperboloids. Since our main goal in this paper is to demonstrate, via examples, the remarkable multivariate analog and generalization of Pell's polynomial equation, we decide not to include these further extensions as they are more technical in nature.

**4.5. An extremal property of Christoffel functions.** We have established the results initiated in [8] for *all degrees*  $n$  and *all dimensions*  $d$ , and we have demonstrated in the three cases listed above that, for every  $n$ , the Christoffel functions  $(\Lambda_n^{g \cdot \mu})_{g \in G}$  satisfy the generalized Pell's identity (4.14). Again it is important to emphasize that the set  $G$  of generators is crucial. Indeed, recall that in the case of the box for instance, if we take  $G = \{1, 1 - x_1^2, \dots, 1 - x_d^2\}$  then even though a unique optimal solution  $\phi_{2n}^*$  of (3.6) exists, it cannot be the vector of moments up to degree  $2n$  of the equilibrium measure  $\mu$  of  $\Omega$ .

As the property (4.14) of Christoffel functions is quite strong, it is likely to hold only for sets  $\Omega \subset \mathbb{R}^d$  with a very specific geometry and a very specific representation (3.2) since the choice  $G$  of generators is also crucial. What about different sets  $G$  of generators, or more general sets  $\Omega \subset \mathbb{R}^d$ ?

In the rest of this section, we suppose that  $G \subset \mathbb{R}[\mathbf{x}]$  is given and  $\Omega \subset \mathbb{R}^d$  in (3.2) is compact. By Theorem 3.2, if  $\mathbf{1} \in Q_n(G_n)$  for every  $n \geq n_0$  and  $G_n = G$  for sufficiently large  $n$ , there exists  $\phi_{2n}^* \in \mathbb{R}^{s(2n)}$  which is an optimal solution of (3.6), and

$$(4.15) \quad \sum_{g \in G_n} s(n - t_g) = \sum_{g \in G_n} g(\mathbf{x}) \Lambda_{n-t_g}^{g \cdot \phi_{2n}^*}(\mathbf{x})^{-1}, \quad \forall \mathbf{x} \in \mathbb{R}^d, \quad \forall n \geq n_0.$$

The crucial difference with (4.14) is that now  $\phi_{2n}^*$  depends on  $n$ , whereas in (4.14) the linear functionals  $(g \cdot \mu)_{g \in G_n}$  do *not* change. It may also happen that  $\phi_{2n}^*$  does not have a representing measure. Notice, however, that (4.15) still states an identity satisfied by Christoffel functions  $\Lambda_{n-t_g}^{g \cdot \phi_{2n}^*}$ ,  $g \in G_n$ , associated with the linear functionals  $g \cdot \phi_{2n}^*$ ,  $g \in G_n$ . Finally, we note that (4.15) also provides a partition of unity for the set  $\Omega$ .

Interestingly, the convex optimization problem (3.6) provides us with a tool to check whether  $\phi_{2n}^*$  is the moment vector of the equilibrium measure  $\mu$ . Indeed if  $\phi_{2n}^*$  is the restriction to moments up to degree  $2n$  of  $\phi_{2n+2}^*$  then indeed,  $\phi_{2n}^*$  may be the vector of moments of  $\mu$ , up to degree  $2n$ . This is very interesting because apart from sets  $\Omega$  with special geometry (like in this paper), there is no simple characterization of  $\mu$  (let alone numerical characterization).

Regarding the convex optimization problem (3.6), we mention the following assumption, in which  $Q_1(G)$  is the cone defined in (3.1).

**Assumption 4.6.** *The set  $\Omega \subset \mathbb{R}^d$  in (3.2) is compact with a nonempty interior and the quadratic polynomial  $\mathbf{x} \mapsto 1 - \|\mathbf{x}\|^2$  belongs to  $Q_1(G)$ .*

It has been shown in [8] that under Assumption 4.6, (3.6) has always an optimal solution  $\phi_{2n}^* \in \mathbb{R}^{s(2n)}$  and, in addition, the sequence  $(\phi_{2n}^*)_{n \in \mathbb{N}}$  has accumulation points

with associated converging subsequences. For each such subsequence  $(n_k)_{k \in \mathbb{N}}$ ,

$$\lim_{k \rightarrow \infty} (\phi_{2(n_k)}^*)_{\alpha} = \phi_{\alpha}^*, \quad \forall \alpha \in \mathbb{N}^d,$$

for some measure  $\phi^*$  on  $\Omega$ . A natural question arises: *is the measure  $\phi^*$  unique and, if so, is  $\phi^*$  related to the equilibrium measure of  $\Omega$ ?*

Conversely, let  $\mu$  be the equilibrium measure of the set  $\Omega$ . Then  $\mathbf{M}_n(\mu) \succ 0$  and  $\mathbf{M}_{n-t_g}(g \cdot \mu) \succ 0$  for every  $g \in G$  and every  $n \in \mathbb{N}$ ,  $n \geq n_0 := \min_{g \in G} t_g$ . Therefore, for every  $n \geq n_0$ , the polynomial

$$(4.16) \quad \mathbf{x} \mapsto p_n(\mathbf{x}) := \frac{1}{\sum_{g \in G} s(n-t_g)} \sum_{g \in G} g(\mathbf{x}) \Lambda_{n-t_g}^{g \cdot \mu}(\mathbf{x})^{-1}, \quad \forall \mathbf{x} \in \mathbb{R}^d,$$

is well-defined and belongs to  $\text{int}(Q_n(G))$ . Let  $d\mu_n := p_n d\mu$ , for every  $n \geq n_0$ . The measure  $\mu_n$  is a probability measure on  $\Omega$  and it is proved in [8] that

$$\lim_{n \rightarrow \infty} \int f d\mu_n = \int f d\mu, \quad \forall f \in \mathcal{C}(\Omega),$$

that is, the sequence of probability measures  $(\mu_n)_{n \in \mathbb{N}}$  converges weakly to  $\mu$ , denoted  $\mu_n \Rightarrow \mu$  as  $n \rightarrow \infty$ . In other words, the density  $p_n$  behaves like the constant polynomial  $\mathbf{1}$  when integrating continuous functions against  $\mu_n$  (but the right-hand-side of (4.16) is not equal to  $\mathbf{1}$ ). So it is fair to say that

$$p_n \mu \Rightarrow \mu \quad (\text{with } p_n \text{ as in (4.16)})$$

is a *weak* form of (4.15) (and also provides a weak form of Theorem 3.4).

Next, under some additional assumption, we can indeed relate the linear functional  $\phi_{2n}^*$  which is the unique optimal solution of (3.6) and the equilibrium measure  $\mu$  of  $\Omega$ .

**Corollary 4.7.** *Let  $\Omega \subset \mathbb{R}^d$  be compact with associated equilibrium measure  $\mu = f_E(\mathbf{x})d\mathbf{x}$  and suppose that  $\lim_{n \rightarrow \infty} s(n)\Lambda_n^\mu(\mathbf{x}) = 1$  uniformly on compact subsets of  $\text{int}(\Omega)$ . Assume that  $g(\mathbf{x}) > 0$  for all  $\mathbf{x} \in \text{int}(\Omega)$  and all  $g \in G$ , and let  $p_n$  be as in (4.16). Then*

$$(4.17) \quad \lim_{n \rightarrow \infty} p_n(\mathbf{x}) = 1, \quad \text{uniformly on compact subsets of } \text{int}(\Omega).$$

*In addition, assume that  $\mathbf{1} \in Q_n(G_n)$  for all  $n \geq$  and let  $\phi_{2n}^*$  be the unique optimal solution of (3.6). Then, uniformly on compact subsets of  $\Omega$ ,*

$$(4.18) \quad \lim_{n \rightarrow \infty} g(\mathbf{x}) \frac{\Lambda^{g \cdot \phi_{2n}^*}(\mathbf{x})^{-1}}{s(n-t_g)} = \lim_{n \rightarrow \infty} g(\mathbf{x}) \frac{\Lambda^{g \cdot \mu}(\mathbf{x})^{-1}}{s(n-t_g)} = 1, \quad \forall g \in G.$$

*Proof.* As  $g > 0$  on  $\text{int}(\Omega)$ , by a result of Kroó and Lubinsky [7] (see also [9, Theorem 4.4.1, p. 52]), and using our assumption on  $\mu$ ,

$$\lim_{n \rightarrow \infty} s(n) \Lambda_n^{g \cdot \mu}(\mathbf{x}) = g(\mathbf{x})$$

uniformly on compact subsets of  $\text{int}(\Omega)$ . Rewriting (4.16) as

$$p_n(\mathbf{x}) = \sum_{g \in G} \frac{g(\mathbf{x})}{s(n-t_g) \Lambda_{n-t_g}^{g \cdot \mu}(\mathbf{x})} \cdot \frac{s(n-t_g)}{\sum_g s(n-t_g)}, \quad \forall \mathbf{x} \in \mathbb{R}^d,$$

and taking limit as  $n$  grows when  $\mathbf{x} \in \text{int}(\Omega)$ , we obtain (4.17).

Next, let  $\phi_{2n}^*$  be an optimal solution of (3.6). Let  $\Delta \subset \text{int}(\Omega)$  be an arbitrary compact subset, and with  $\varepsilon > 0$  fixed, let  $n$  be large enough so that  $G_n = G$  and

- $(\sum_{g \in G} s(n - t_g))/s(n - t_g) \leq 1 + \varepsilon$  for all  $g \in G$ ,
- $1 - \varepsilon \leq g(\mathbf{x}) \frac{\Lambda_{n-t_g}^{g \cdot \mu}(\mathbf{x})^{-1}}{\sum_{g \in G} s(n-t_g)} \leq 1 + \varepsilon$ .

Then by (4.15),  $0 \leq g(\mathbf{x}) \Lambda_n^{g \cdot \phi_{2n}^*}(\mathbf{x})^{-1}/s(n - t_g) \leq (1 + \varepsilon)$ , for all  $g \in G$ , and all  $\mathbf{x} \in \Delta$ , so that for every  $\mathbf{x} \in \Delta$  and every  $g \in G$ ,

$$-2\varepsilon \leq \frac{g(\mathbf{x}) (\Lambda_{n-t_g}^{g \cdot \phi_{2n}^*}(\mathbf{x})^{-1} - \Lambda_{2n}^{g \cdot \mu}(\mathbf{x})^{-1})}{s(n - t_g)} \leq 2\varepsilon.$$

As  $\varepsilon > 0$  was arbitrary, one obtains

$$\lim_{n \rightarrow \infty} g(\mathbf{x}) \frac{\Lambda_{n-t_g}^{g \cdot \phi_{2n}^*}(\mathbf{x})^{-1}}{s(n - t_g)} = \lim_{n \rightarrow \infty} g(\mathbf{x}) \frac{\Lambda_{n-t_g}^{g \cdot \mu}(\mathbf{x})^{-1}}{s(n - t_g)} = 1, \quad \forall \mathbf{x} \in \Delta, \forall g \in G.$$

□

So, under the assumptions in Corollary 4.7, the linear functional  $\phi_{2n}^*$ , unique optimal solution of (3.6), behaves asymptotically like the equilibrium measure of  $\Omega$ , as  $n$  grows. Thus, for instance, for the unit cube with the set of generators  $G = \{1, 1 - x_1^2, \dots, 1 - x_d^2\}$ ,  $\phi_{2n}^*$  cannot be the vector of moments up to degree  $2n$  of the equilibrium measure  $\mu$ . However:

$$\lim_{n \rightarrow \infty} \frac{\Lambda_n^{\phi_{2n}^*}(\mathbf{x})^{-1}}{s(n)} = \lim_{n \rightarrow \infty} \frac{\Lambda_n^\mu(\mathbf{x})^{-1}}{s(n)} = 1, \quad \forall \mathbf{x} \in \text{int}([-1, 1]^d),$$

and for every  $i = 1, \dots, d$ ,

$$\lim_{n \rightarrow \infty} (1 - x_i^2) \frac{\Lambda_{n-1}^{(1-x_i^2) \cdot \phi_{2n}^*}(\mathbf{x})^{-1}}{s(n-1)} = \lim_{n \rightarrow \infty} (1 - x_i^2) \frac{\Lambda_{n-1}^{(1-x_i^2) \cdot \mu}(\mathbf{x})^{-1}}{s(n-1)} = 1,$$

for all  $\mathbf{x} \in \text{int}([-1, 1]^d)$ .

## 5. PROOFS

Our proof depends heavily on orthogonal polynomials on a family of weight functions associated with the Chebyshev weight. To avoid confusion, we shall adopt notations that make it transparent how kernels depend on weight functions. More precisely, let  $W$  be a weight function on a domain  $\Omega$ ; we consider orthogonal polynomials with respect to the inner product

$$\langle f, g \rangle = \int_{\Omega} f(\mathbf{x})g(\mathbf{x})W(\mathbf{x}) \, d\mathbf{x}.$$

We then denote the space  $\mathcal{V}_n^d$  of orthogonal polynomials of degree  $n$  by  $\mathcal{V}_n(W; \Omega)$ . Then

$$\dim \Pi_n^d = \binom{n+d}{n} \quad \text{and} \quad \dim \mathcal{V}_n(W; \Omega) = \binom{n+d-1}{n}.$$

Let  $P_n(W; \cdot, \cdot)$  be the reproducing kernel of  $\mathcal{V}_n(W, \Omega)$ . If  $\{P_\alpha^n : |\alpha| = n, \alpha \in \mathbb{N}_0^d\}$  is an orthonormal basis of  $\mathcal{V}_n(W, \Omega)$ , then the reproducing kernel of  $\mathcal{V}_n(W, \Omega)$  is

$$P_n(W; \mathbf{x}, \mathbf{y}) = \sum_{|\alpha|=n} P_\alpha^n(\mathbf{x})P_\alpha^n(\mathbf{y}).$$

Moreover, the reproducing kernel of  $\Pi_n^d$  in  $L^2(W, \Omega)$  is denoted by

$$K_n(W; \mathbf{x}, \mathbf{y}) = \sum_{k=0}^n P_k(W; \mathbf{x}, \mathbf{y}).$$

**5.1. Unit ball.** The proof of Theorem 4.1 relies on the addition formula (1.7) for the spherical harmonics mentioned in the introduction. Spherical harmonics are the restrictions of harmonic polynomials on the unit sphere and they are known to be orthogonal on the unit sphere. Let  $\mathcal{H}_n^{d+1}$  be the space of spherical harmonics of degree  $n$  in  $d+1$  variables. For  $n \in \mathbb{N}_0$ , let  $\{Y_\ell^n : 1 \leq \ell \leq \dim \mathcal{H}_n^{d+1}\}$  be an orthonormal basis of  $\mathcal{H}_n^{d+1}$ , so that

$$\frac{1}{\sigma_d} \int_{\mathbb{S}^d} Y_\ell^n(\boldsymbol{\xi}) Y_{\ell'}^m(\boldsymbol{\xi}) d\sigma(\boldsymbol{\xi}) = \delta_{\ell, \ell'} \delta_{n, m},$$

where  $\sigma_d$  denote the surface area of  $\mathbb{S}^d$  and  $d\sigma$  denote the Lebesgue measure on  $\mathbb{S}^d$ . We denote by  $P_n(d\sigma; \cdot, \cdot)$  the reproducing kernel of  $\mathcal{H}_n^{d+1}$ , which then satisfies the *addition formula*

$$(5.1) \quad Y_n(d\sigma; \mathbf{x}, \mathbf{y}) = \sum_{\ell=1}^{\dim \mathcal{H}_n^{d+1}} Y_\ell^n(\mathbf{x}) Y_\ell^n(\mathbf{y}) = Z_n^{\frac{d-1}{2}}(\langle \mathbf{x}, \mathbf{y} \rangle), \quad \mathbf{x}, \mathbf{y} \in \mathbb{S}^d,$$

where  $Z_n^\lambda$  is a multiple of the Gegenbauer (ultraspherical) polynomial,

$$Z_n^\lambda(t) = \frac{n + \lambda}{\lambda} C_n^\lambda(t).$$

In particular, setting  $\mathbf{y} = \mathbf{x}$  in the addition formula, we obtain

$$(5.2) \quad \sum_{\ell} |Y_\ell^n(\mathbf{x})|^2 = \frac{2n + d - 1}{d - 1} C_n^{\frac{d-1}{2}}(1) = \binom{n + d - 1}{d - 1} + \binom{n + d - 2}{d - 1},$$

which follows from  $C_n^\lambda(1) = \Gamma(n + 2\lambda) / (n! \Gamma(2\lambda))$  and a simple verification.

Taking a cue of (5.1), for any domain  $\Omega$  with weight function  $W$ , an addition formula for the reproducing kernel of  $\mathcal{V}_n(W; \Omega)$  is a closed-form formula for  $P_n(W; \cdot, \cdot)$ .

We now turn our attention to the unit ball  $\mathbb{B}^d$ . The classical orthogonal polynomials on  $\mathbb{B}^d$  are associated with the weight function  $(1 - \|\mathbf{x}\|^2)^\alpha$ ,  $\alpha > -1$ . The normalization constant  $c_\alpha$  of this weight function, so that  $c_\alpha(1 - \|\mathbf{x}\|^2)^\alpha$  is a probability measure, is given by

$$c_\alpha = \frac{\Gamma(\alpha + \frac{d+2}{2})}{\pi^{\frac{d}{2}} \Gamma(\alpha + 1)}.$$

In the following, we denote by

$$W_{-\frac{1}{2}}(\mathbf{x}) = c_{-\frac{1}{2}}(1 - \|\mathbf{x}\|^2)^{-\frac{1}{2}} \quad \text{and} \quad W_{\frac{1}{2}}(\mathbf{x}) = c_{-\frac{1}{2}}(1 - \|\mathbf{x}\|^2)^{\frac{1}{2}}.$$

Notice that  $W_{-\frac{1}{2}}$  is a probability measure on  $\mathbb{B}^d$ , but  $W_{\frac{1}{2}}$  is not. The spherical harmonics are closely related to orthogonal polynomials associated with  $W_{-\frac{1}{2}}$  and  $W_{\frac{1}{2}}$  on the unit ball. In particular, we have the following addition formula:

**Proposition 5.1.** *Let  $X = (\mathbf{x}, \sqrt{1 - \|\mathbf{x}\|^2})$ ,  $\mathbf{x} \in \mathbb{B}^d$ , and  $\mathbf{x} \in \mathbb{S}^{d+1}$ . Then*

$$(5.3) \quad Y_n(d\sigma; X, Y) = P_n\left(W_{-\frac{1}{2}}; \mathbf{x}, \mathbf{y}\right) + \sqrt{1 - \|\mathbf{x}\|^2} \sqrt{1 - \|\mathbf{y}\|^2} P_{n-1}\left(W_{\frac{1}{2}}; \mathbf{x}, \mathbf{y}\right).$$



*Proof.* Let  $\{P_\nu^n(W_\mu; \mathbf{x}) : |\nu| = n, \nu \in \mathbb{N}_0^d\}$  denote an orthogonal basis of  $\mathcal{V}_n^d(W_\mu)$ . By [5, Theorem 4.2.4], it follows that an orthogonal basis for the space  $\mathcal{H}_n^{d+1}$  of spherical harmonics consists of

$$\begin{aligned} Y_\nu^{(1)}(\mathbf{x}, x_{d+1}) &= P_\nu^n(W_{-\frac{1}{2}}, \mathbf{x}), \quad |\nu| = n, \\ Y_\nu^{(2)}(\mathbf{x}, x_{d+1}) &= x_{d+1} P_\nu^{n-1}(W_{\frac{1}{2}}, \mathbf{x}), \quad |\nu| = n-1, \end{aligned}$$

where  $\mathbf{x} \in \mathbb{B}^d$ ,  $(\mathbf{x}, x_{d+1}) \in \mathbb{S}^d$ . Using the integral identity [5, Lemma 4.2.3],

$$\int_{\mathbb{S}^d} f(\mathbf{y}) d\sigma(\mathbf{y}) = \int_{\mathbb{B}^d} \left[ f(\mathbf{x}, \sqrt{1-\|\mathbf{x}\|^2}) + f(\mathbf{x}, -\sqrt{1-\|\mathbf{x}\|^2}) \right] \frac{d\mathbf{x}}{\sqrt{1-\|\mathbf{x}\|^2}},$$

it follows readily that  $c_{-\frac{1}{2}}/2 = 1/\omega_d$  is the surface area of  $\mathbb{S}^d$  and, consequently,

$$\|Y_\nu^{(1)}\|_{L^2(d\sigma)}^2 = \|P_\nu^n(W_{-\frac{1}{2}})\|_{L^2(W_{-\frac{1}{2}})}^2$$

Moreover, we also obtain

$$\begin{aligned} \|Y_\nu^{(2)}\|_{L^2(d\sigma)}^2 &= \frac{1}{\omega_d} \int_{\mathbb{S}^d} \left| x_{d+1} P_\nu^{n-1}(W_{\frac{1}{2}}, \mathbf{x}) \right|^2 d\sigma \\ &= c_{-\frac{1}{2}} \int_{\mathbb{B}^d} \left| P_\nu^{n-1}(W_{\frac{1}{2}}, \mathbf{x}) \right|^2 \sqrt{1-\|\mathbf{x}\|^2} d\mathbf{x} = \|P_\nu^{n-1}(W_{\frac{1}{2}})\|_{L^2(W_{\frac{1}{2}})}^2. \end{aligned}$$

Consequently, since the reproducing kernel can be written as the sum of products of orthonormal polynomials, it follows that  $Y_n(d\sigma; X, Y)$  is equal to

$$\begin{aligned} &\sum_{|\nu|=n} \frac{P_\nu^n(W_{-\frac{1}{2}}, \mathbf{x}) P_\nu^n(W_{-\frac{1}{2}}, \mathbf{y})}{\|P_\nu^n(W_{-\frac{1}{2}})\|_{L^2(W_{-\frac{1}{2}})}^2} + \sum_{|\nu|=n-1} \frac{x_{d+1} y_{d+1} P_\nu^{n-1}(W_{\frac{1}{2}}, \mathbf{x}) P_\nu^{n-1}(W_{\frac{1}{2}}, \mathbf{y})}{\|P_\nu^{n-1}(W_{\frac{1}{2}})\|_{L^2(W_{\frac{1}{2}})}^2} \\ &= P_n\left(W_{-\frac{1}{2}}; \mathbf{x}, \mathbf{y}\right) + \sqrt{1-\|\mathbf{x}\|^2} \sqrt{1-\|\mathbf{y}\|^2} P_{n-1}\left(W_{\frac{1}{2}}; \mathbf{x}, \mathbf{y}\right). \end{aligned}$$

This completes the proof.  $\square$

*Remark 5.1.* As an illustration for the Remark 3.1, we note that if we replace  $W_{\frac{1}{2}}$  by the probability measure  $\widehat{W}_{\frac{1}{2}}(x) = c_{\frac{1}{2}}(1-\|x\|^2)^{\frac{1}{2}}$ , then

$$\begin{aligned} \|Y_\nu^{(2)}\|_{L^2(d\sigma)}^2 &= \frac{1}{\omega_d} \int_{\mathbb{S}^d} \left| x_{d+1} P_\nu^{n-1}(\widehat{W}_{\frac{1}{2}}, \mathbf{x}) \right|^2 d\sigma \\ &= \frac{c_{-\frac{1}{2}}}{c_{\frac{1}{2}}} \int_{\mathbb{B}^d} \left| P_\nu^{n-1}(\widehat{W}_{\frac{1}{2}}, \mathbf{x}) \right|^2 \widehat{W}_{\frac{1}{2}}(x) d\mathbf{x} = \frac{1}{d+1} \|P_\nu^{n-1}(\widehat{W}_{\frac{1}{2}})\|_{L^2(\widehat{W}_{\frac{1}{2}})}^2. \end{aligned}$$

Hence, following the proof of the above proposition, we obtain the identity

$$Y_n(d\sigma; X, Y) = P_n\left(W_{-\frac{1}{2}}; \mathbf{x}, \mathbf{y}\right) + (d+1) \sqrt{1-\|\mathbf{x}\|^2} \sqrt{1-\|\mathbf{y}\|^2} P_{n-1}\left(\widehat{W}_{\frac{1}{2}}; \mathbf{x}, \mathbf{y}\right).$$

Notice the additional  $(d+1)$  in this identity in comparison with that of (5.3)

Setting  $Y = X$  in the formula that we just proved and applying (5.2) leads to the following Pell identity on the unit ball.

**Corollary 5.2.** For  $d \geq 1$ ,  $n = 0, 1, 2, \dots$  and  $\mathbf{x} \in \mathbb{B}^d$ ,

$$(5.4) \quad P_n\left(W_{-\frac{1}{2}}; \mathbf{x}, \mathbf{x}\right) + (1-\|\mathbf{x}\|^2) P_{n-1}\left(W_{\frac{1}{2}}; \mathbf{x}, \mathbf{x}\right) = \binom{n+d-1}{d-1} + \binom{n+d-2}{d-1}$$

and, summing up the identity,

$$(5.5) \quad K_n \left( W_{-\frac{1}{2}}; \mathbf{x}, \mathbf{x} \right) + (1 - \|\mathbf{x}\|^2) K_{n-1} \left( W_{\frac{1}{2}}; \mathbf{x}, \mathbf{x} \right) = \binom{n+d}{d} + \binom{n+d-1}{d}.$$

The identity (5.4) is exactly the same as (4.4) in Theorem 4.1, whereas (5.5) is equivalent to (4.5) by (2.3).

**5.2. Simplex.** The classical orthogonal polynomials on  $\Delta^d$  are associated with the weight function

$$W_{\boldsymbol{\kappa}}(\mathbf{x}) = x_1^{\kappa_1} \cdots x_d^{\kappa_d} (1 - |\mathbf{x}|)^{\kappa_{d+1}}, \quad \mathbf{x} \in \Delta^d, \quad \kappa_i > -1.$$

The normalization constant  $c_{\boldsymbol{\kappa}}$  that makes this weight function a probability measure is given by

$$c_{\boldsymbol{\kappa}} = \frac{\Gamma(|\boldsymbol{\kappa}| + d + 1)}{\Gamma(\kappa_1 + 1) \cdots \Gamma(\kappa_{d+1} + 1)}.$$

For  $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^{d+1}$ , we define the modified weight function

$$W_{\boldsymbol{\kappa}}^{\Delta}(x) = c_{-\frac{1}{2}} W_{\boldsymbol{\kappa}}(x).$$

By this definition, the weight function  $W_{-\mathbf{1}/2}$  is normalized to have the unit integral on  $\Delta^d$  but  $W_{\boldsymbol{\kappa}}$  for  $\boldsymbol{\varepsilon} \neq -\mathbf{1}/2$  is not.

The classical orthogonal polynomials for  $W_{\boldsymbol{\kappa}}^{\Delta}$  and those on the ball are closely related.

**Lemma 5.3.** *Let  $\{P_{\boldsymbol{\nu}}^{\boldsymbol{\kappa}} : |\boldsymbol{\nu}| = n\}$  denote an orthogonal basis for  $\mathcal{V}_n(W_{\boldsymbol{\kappa}}^{\Delta}; \Delta^d)$ . Let  $\boldsymbol{\varepsilon} \in \{0, 1\}^d$ , and define, with  $\boldsymbol{\kappa} = (\boldsymbol{\kappa}', \kappa_{d+1})$  and  $\boldsymbol{\kappa}' \in \mathbb{R}^d$ ,*

$$\mathbf{x} \mapsto Q_{\boldsymbol{\nu}, \boldsymbol{\varepsilon}}^{\boldsymbol{\kappa}}(\mathbf{x}) := \mathbf{x}^{\boldsymbol{\varepsilon}} P_{\boldsymbol{\nu}}^{\boldsymbol{\kappa}' + \boldsymbol{\varepsilon}, \kappa_{d+1}}(x_1^2, \dots, x_d^2), \quad \boldsymbol{\nu} \in \mathbb{N}_0^d.$$

*Let  $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^d$ . Then  $\{Q_{\boldsymbol{\nu}, \boldsymbol{\varepsilon}}^{(-\mathbf{1}/2, \alpha)} : \boldsymbol{\varepsilon} \in \{0, 1\}^d, |\boldsymbol{\nu}| = \frac{n - |\boldsymbol{\varepsilon}|}{2} \in \mathbb{N}_0, \boldsymbol{\nu} \in \mathbb{N}_0^d\}$  is an orthogonal basis for  $\mathcal{V}_n(W_{\alpha}; \mathbb{B}^d)$  and, moreover,*

$$\left\| Q_{\boldsymbol{\nu}, \boldsymbol{\varepsilon}}^{(-\mathbf{1}/2, \alpha)} \right\|_{L^2(W_{\alpha}^{\mathbb{B}})}^2 = \left\| P_{\boldsymbol{\nu}}^{(-\mathbf{1}/2 + \boldsymbol{\varepsilon}, \alpha)} \right\|_{L^2(W_{-\mathbf{1}/2 + \boldsymbol{\varepsilon}, \alpha}^{\Delta})}^2.$$

*Proof.* For  $\boldsymbol{\varepsilon} \neq \boldsymbol{\varepsilon}'$ , the orthogonality of  $Q_{\boldsymbol{\nu}, \boldsymbol{\varepsilon}}^{(-\mathbf{1}/2, \alpha)}$  and  $Q_{\boldsymbol{\nu}', \boldsymbol{\varepsilon}' }^{(-\mathbf{1}/2, \alpha)}$  follows from the parity and the invariance of  $W_{\alpha}$  under sign changes. For fixed  $\boldsymbol{\varepsilon}$ , the orthogonality of  $\{Q_{\boldsymbol{\nu}, \boldsymbol{\varepsilon}}^{(-\mathbf{1}/2, \alpha)} : \boldsymbol{\nu} \in \mathbb{N}_0^d\}$  follows from [5, Theorem 4.4.4]. Furthermore, it is not difficult to verify that the cardinality of  $\{Q_{\boldsymbol{\nu}, \boldsymbol{\varepsilon}}^{(-\mathbf{1}/2, \alpha)} : \boldsymbol{\varepsilon} \in \{0, 1\}^d, |\boldsymbol{\nu}| = \frac{n - |\boldsymbol{\varepsilon}|}{2} \in \mathbb{N}_0, \boldsymbol{\nu} \in \mathbb{N}_0^d\}$  is equal to  $\dim \mathcal{V}_n(W_{\alpha}; \mathbb{B}^d)$ , so that the set is an orthogonal basis. Moreover, using the integral identity ([5, Lemma 4.4.1])

$$\int_{\mathbb{B}^d} f(y_1^2, \dots, y_d^2) dy = \int_{\Delta^d} f(x_1, \dots, x_d) \frac{dx}{\sqrt{x_1 \cdots x_d}},$$

it is easy to verify that  $c_{-\mathbf{1}/2, \alpha}$  is the normalization constant of  $W_{\alpha}^{\mathbb{B}}$  on the unit ball. Hence,

$$\begin{aligned} \left\| Q_{\boldsymbol{\nu}, \boldsymbol{\varepsilon}}^{(-\mathbf{1}/2, \alpha)} \right\|_{L^2(W_{\alpha}^{\mathbb{B}})}^2 &= c_{-\mathbf{1}/2, \alpha} \int_{\mathbb{B}^d} \left| \mathbf{x}^{\boldsymbol{\varepsilon}} P_{\boldsymbol{\nu}}^{(-\mathbf{1}/2 + \boldsymbol{\varepsilon}, \alpha)}(x_1^2, \dots, x_d^2) \right|^2 W_{\alpha}^{\mathbb{B}}(x) dx \\ &= c_{-\mathbf{1}/2, \alpha} \int_{\mathbb{T}^d} \left| P_{\boldsymbol{\nu}}^{(-\mathbf{1}/2 + \boldsymbol{\varepsilon}, \alpha)}(x_1, \dots, x_d) \right|^2 W_{-\mathbf{1}/2 + \boldsymbol{\varepsilon}, \alpha}^{\Delta}(x) dx \\ &= \left\| P_{\boldsymbol{\nu}}^{(-\mathbf{1}/2 + \boldsymbol{\varepsilon}, \alpha)} \right\|_{L^2(W_{-\mathbf{1}/2 + \boldsymbol{\varepsilon}, \alpha}^{\Delta})}^2, \end{aligned}$$

where the last step follows from our definition of  $W_{\kappa}^{\Delta}$ . This completes the proof.  $\square$

**Proposition 5.4.** For  $\kappa_1, \dots, \kappa_{d+1} > -1$ ,  $\mathbf{x} \in \Delta^d$ ,

$$(5.6) \quad \sum_{\varepsilon \in \{0,1\}^{d+1}, |\varepsilon| \in 2\mathbb{N}_0} P_{n-\frac{|\varepsilon|}{2}} \left( W_{-1/2+\varepsilon}^{\Delta}; \mathbf{x}, \mathbf{x} \right) = \binom{2n+d-1}{d-1} + \binom{2n+d-2}{d-1}.$$

Moreover, summing up the identity,

$$(5.7) \quad \sum_{\varepsilon \in \{0,1\}^{d+1}, |\varepsilon| \in 2\mathbb{N}_0} \mathbf{x}^{\varepsilon} K_{n-\frac{|\varepsilon|}{2}} \left( W_{-1/2+\varepsilon}^{\Delta}; \mathbf{x}, \mathbf{x} \right) = \binom{2n+d}{2n}.$$

*Proof.* Let  $\mathbf{x}^2 = (x_1^2, \dots, x_d^2)$ . As a consequence of the lemma, we obtain immediately

$$\begin{aligned} P_n(W_{\alpha}^{\mathbb{B}}; \mathbf{x}, \mathbf{y}) &= \sum_{\substack{\varepsilon \in \{0,1\}^d \\ \frac{n-|\varepsilon|}{2} \in \mathbb{N}_0}} \sum_{|\nu| = \frac{n-|\varepsilon|}{2}} \frac{Q_{\nu, \varepsilon}^{(-1/2, \alpha)}(\mathbf{x}) Q_{\nu, \varepsilon}^{(-1/2, \alpha)}(\mathbf{y})}{\|Q_{\nu, \varepsilon}^{(-1/2, \alpha)}\|_{L^2(W_{\alpha}^{\mathbb{B}})}^2} \\ &= \sum_{\substack{\varepsilon \in \{0,1\}^d \\ \frac{n-|\varepsilon|}{2} \in \mathbb{N}_0}} \mathbf{x}^{\varepsilon} \mathbf{y}^{\varepsilon} \sum_{|\nu| = \frac{n-|\varepsilon|}{2}} \frac{P_{\nu}^{(-1/2+\varepsilon, \alpha)}(\mathbf{x}^2) P_{\nu}^{(-1/2+\varepsilon, \alpha)}(\mathbf{y}^2)}{\|P_{\nu}^{(-1/2+\varepsilon, \alpha)}\|_{L^2(W_{-1/2+\varepsilon, \alpha}^{\Delta})}^2} \\ &= \sum_{\substack{\varepsilon \in \{0,1\}^d \\ \frac{n-|\varepsilon|}{2} \in \mathbb{N}_0}} \mathbf{x}^{\varepsilon} \mathbf{y}^{\varepsilon} P_{\frac{n-|\varepsilon|}{2}} \left( W_{-1/2+\varepsilon, \alpha}^{\Delta}; \mathbf{x}^2, \mathbf{y}^2 \right), \end{aligned}$$

where the sum is over those  $\varepsilon \in \{0,1\}^d$  such that  $\frac{n-|\varepsilon|}{2} \in \mathbb{N}_0$ , which implies, in particular,

$$P_n(W_{\alpha}^{\mathbb{B}}; \mathbf{x}, \mathbf{x}) = \sum_{\substack{\varepsilon \in \{0,1\}^d \\ \frac{n-|\varepsilon|}{2} \in \mathbb{N}_0}} \mathbf{x}^{2\varepsilon} P_{\frac{n-|\varepsilon|}{2}} \left( W_{-1/2+\varepsilon, \alpha}^{\Delta}; \mathbf{x}^2, \mathbf{x}^2 \right).$$

Choosing  $\alpha = \pm \frac{1}{2}$  and applying the relation (5.4), we obtain

$$\begin{aligned} Z_n^{\frac{d-1}{2}}(1) &= \sum_{\substack{\varepsilon \in \{0,1\}^d \\ \frac{n-|\varepsilon|}{2} \in \mathbb{N}_0}} \mathbf{x}^{2\varepsilon} P_{\frac{n-|\varepsilon|}{2}} \left( W_{-1/2+\varepsilon, -\frac{1}{2}}^{\Delta}; \mathbf{x}^2, \mathbf{x}^2 \right) \\ &+ (1 - \|\mathbf{x}\|^2) \sum_{\substack{\varepsilon \in \{0,1\}^d \\ \frac{n-1-|\varepsilon|}{2} \in \mathbb{N}_0}} \mathbf{x}^{2\varepsilon} P_{\frac{n-|\varepsilon|-1}{2}} \left( W_{-1/2+\varepsilon, \frac{1}{2}}^{\Delta}; \mathbf{x}^2, \mathbf{x}^2 \right). \end{aligned}$$

The two sums on the right-hand side can be combined by setting  $\varepsilon \in \{0,1\}^{d+1}$ . Thus, changing variables  $\mathbf{x} \mapsto (\sqrt{x_1}, \dots, \sqrt{x_d})$  and setting  $x_{d+1} = 1 - |\mathbf{x}|$ , we obtain

$$\sum_{\substack{\varepsilon \in \{0,1\}^{d+1} \\ \frac{n-|\varepsilon|}{2} \in \mathbb{N}_0}} P_{\frac{n-|\varepsilon|}{2}} \left( W_{-1/2+\varepsilon}^{\Delta}; \mathbf{x}, \mathbf{x} \right) = Z_n^{\frac{d-1}{2}}(1)$$

by (5.2). In particular, setting  $n \mapsto 2n$ , we have proved (5.6). Summing up this identity gives (5.7), as can be seen by verifying the sum over binomial coefficients.  $\square$

Like the case for the unit ball, (5.6) is the same as (4.8), whereas (5.7) is equivalent to (4.9) by (2.3).

5.3. **Cube.** The classical orthogonal polynomials on  $[-1, 1]^d$  are associated with the weight function

$$W_{\alpha}^{\square}(\mathbf{x}) = \prod_{i=1}^d (1 - x_i^2)^{\alpha_i}, \quad \alpha_i > -1,$$

and they are products of univariate Gegenbauer polynomials. More precisely, an orthonormal basis for  $\mathcal{V}_n(\alpha^{\square}; [-1, 1]^d)$  is given by

$$P_{\mathbf{k}}\left(W_{\alpha}^{\square}; \mathbf{x}\right) = \prod_{j=1}^d \widehat{C}_{k_j}^{\alpha_j + \frac{1}{2}}(x_j), \quad |\mathbf{k}| = n, \quad \mathbf{k} \in \mathbb{N}_0^d,$$

where  $\widehat{C}_n^{\lambda}$  denotes the orthonormal Gegenbauer polynomial of degree  $n$ . It is worth mentioning that orthonormality is defined with respect to the probability measure. In particular, normalized Chebyshev weight functions are  $\frac{1}{\pi}(1 - x^2)^{-\frac{1}{2}}$  and  $\frac{1}{2\pi}(1 - x^2)^{\frac{1}{2}}$ , so that the normalized Chebyshev polynomials  $\widehat{T}_n$  of the first kind are

$$\widehat{T}_n(x) = \sqrt{\delta(k)} T_n(x), \quad n \geq 0, \quad \text{where } \delta(k) = \begin{cases} 1 & k = 0, \\ 2 & k > 0. \end{cases}$$

whereas the Chebyshev polynomials  $U_n$  of the second kind are already orthonormal.

A closed-form formula for the reproducing kernel for  $\mathcal{V}_n(W_{-1/2}^{\square}, [-1, 1]^d)$  was derived in [14] and [3], and further explored in [4] recently. It shows, in particular, that if  $\theta = (\theta_1, \dots, \theta_d) \in [0, \pi]^d$  and  $\mathbf{k} = (k_1, \dots, k_d)$ , then

$$(5.8) \quad \sum_{|\mathbf{k}|=n, \mathbf{k} \in \mathbb{Z}^d} e^{i\mathbf{k} \cdot \theta} = [\cos \theta_1, \cos \theta_2, \dots, \cos \theta_d] H_{n,d},$$

where  $[x_1, \dots, x_d]f$  denotes the divided difference of  $f$  with knots  $x_1, \dots, x_d$  and

$$H_{n,d}(\cos \theta) = 2(-1)^{\lfloor \frac{d-1}{2} \rfloor} (\sin \theta)^{d-1} \times \begin{cases} -\sin(n\theta) & \text{for } d \text{ even,} \\ \cos(n\theta) & \text{for } d \text{ odd} \end{cases}$$

for  $n \geq 1$  and

$$H_{0,d}(\cos \theta) = 2(-1)^{\lfloor \frac{d-1}{2} \rfloor} (\sin \theta)^{d-1} \cos \frac{\theta}{2} \times \begin{cases} \cos \frac{\theta}{2} & \text{for } d \text{ even,} \\ \sin \frac{\theta}{2} & \text{for } d \text{ odd.} \end{cases}$$

Recall that the divided difference is defined inductively by

$$[x]f = f(x) \quad \text{and} \quad [x_0, \dots, x_m]f = \frac{[x_0, \dots, x_{m-1}]f - [x_1, \dots, x_m]f}{x_0 - x_m}.$$

It is a symmetric function of the knots  $x_0, \dots, x_m$ , which may coalesce. In particular, if all knots coalesce and if  $f$  is sufficiently differentiable, then the divided difference collapses to

$$[x_0, \dots, x_m]f = \frac{f^{(m)}(x_0)}{m!} \quad \text{if } x_0 = x_1 = \dots = x_m.$$

it follows that the cardinality of  $\{\mathbf{k} \in \mathbb{N}_0^d : |\mathbf{k}| = n\}$  is given by

$$(5.9) \quad M_{n,d}^{\square} = \#\{\mathbf{k} \in \mathbb{Z}^d : |\mathbf{k}| = n\} = \frac{H_{n,d}^{(d-1)}(1)}{(d-1)!} =: \frac{h_{n,d}}{(d-1)!}.$$

for  $n \geq 1$  and  $M_{0,d}^{\square} = 1$ . The value of  $h_{n,d}$  is given by [4, Lemma 2.3]. We shall give another formula for  $h_{n,d}$  after the proof of the following proposition.

**Proposition 5.5.** *For  $\mathbf{x} = (\cos \theta_1, \dots, \cos \theta_d)$  and  $\mathbf{y} = (\cos \phi_1, \dots, \cos \phi_d)$ , then*

$$\sum_{|\mathbf{k}|=n, \mathbf{k} \in \mathbb{Z}^d} e^{i\mathbf{k} \cdot (\theta - \phi)} = \sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} 2^{|\boldsymbol{\varepsilon}|} \left[ \prod_{j=1}^d (1 - x_j^2)^{\boldsymbol{\varepsilon}_j/2} (1 - y_j^2)^{\boldsymbol{\varepsilon}_j/2} \right] P_{n-|\boldsymbol{\varepsilon}|} \left( W_{-1/2+\boldsymbol{\varepsilon}}^\square; \mathbf{x}, \mathbf{y} \right).$$

*Proof.* For  $\mathbf{k} \in \mathbb{N}_0^d$ , we write  $\sigma(\mathbf{k}) = \{\mathbf{j} \in \mathbb{Z}^d : \mathbf{j} = (\pm k_1, \dots, \pm k_d)\}$  be the orbit of  $\mathbf{k}$  under  $\mathbb{Z}_2^d$ . For  $\mathbf{k} \in \mathbb{N}_0^d$  with  $k_j \neq 0$ ,  $1 \leq j \leq d$ , it is easy to see by symmetry that

$$\begin{aligned} \sum_{\mathbf{j} \in \sigma(\mathbf{k})} e^{i\mathbf{j} \cdot (\theta - \sigma)} &= 2^d \sum_{k_1, \dots, k_d} \prod_{j=1}^{\ell} \cos(k_{i_j} s_{i_j}) \cos(k_{i_j} t_{i_j}) \prod_{j=\ell+1}^d \sin(k_{i_j} s_{i_j}) \sin(k_{i_j} t_{i_j}) \\ &= \sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} 2^{|\boldsymbol{\varepsilon}|} \prod_{i=1}^d (1 - x_i^2)^{\frac{\boldsymbol{\varepsilon}_i}{2}} (1 - y_i^2)^{\frac{\boldsymbol{\varepsilon}_i}{2}} P_{\mathbf{k}}(W_{-1/2+\boldsymbol{\varepsilon}}; \mathbf{x}) P_{\mathbf{k}}(W_{-1/2+\boldsymbol{\varepsilon}}; \mathbf{y}), \end{aligned}$$

where the sum in the left-hand side of the first equation is over all possible distribution of  $\mathbf{k} = (k_1, \dots, k_d)$ , and we have used the fact that  $P_{\mathbf{k}}(W_{-1/2+\boldsymbol{\varepsilon}}; \mathbf{x})$  contains  $d - |\boldsymbol{\varepsilon}|$  many  $\widehat{T}_{k_j}(x_j) = \sqrt{2} \cos(k_j \theta_j)$  in the second identity. The sum over  $\sigma(\mathbf{k})$  with, say, exactly  $m$  elements of  $\mathbf{k}$  non-zero is the above identity with  $d$  replaced by  $m$ . Recall that the orthonormal polynomial of degree 0 is equal to 1 by our normalization, we see that the identity remains true for  $\mathbf{k} \in \mathbb{N}_0^d$  that has some zero components. Since summing over  $\{\mathbf{m} \in \mathbb{Z}^d : |\mathbf{m}| = n\}$  is the same as summing over  $\{\mathbf{k} \in \mathbb{N}_0^d : \mathbf{m} \in \sigma(\mathbf{k}), |\mathbf{m}| = n\}$ , summing the identity over  $\mathbf{k} \in \mathbb{N}_0^d$  and  $|\mathbf{k}| = n$  proves the stated identity.  $\square$

Setting  $\mathbf{y} = \mathbf{x}$  and applying (5.8) in the identity of the above proposition, we obtain the extended Pell equations for the cube:

**Corollary 5.6.** *For  $d \geq 1$ ,  $n = 0, 1, 2, \dots$  and  $\mathbf{x} \in [-1, 1]^d$ ,*

$$\begin{aligned} (5.10) \quad \sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} 2^{|\boldsymbol{\varepsilon}|} \left[ \prod_{j=1}^d (1 - x_j^2)^{\boldsymbol{\varepsilon}_j} \right] P_{n-|\boldsymbol{\varepsilon}|} \left( W_{-1/2+\boldsymbol{\varepsilon}}^\square; \mathbf{x}, \mathbf{x} \right) \\ = \sum_{j=0}^d \binom{d}{j} \left[ \binom{d+n-j}{d} - \binom{d+n-1-j}{d} \right] \end{aligned}$$

and, summing up the identity,

$$(5.11) \quad \sum_{\boldsymbol{\varepsilon} \in \{0,1\}^d} 2^{|\boldsymbol{\varepsilon}|} \left[ \prod_{j=1}^d (1 - x_j^2)^{\boldsymbol{\varepsilon}_j} \right] K_{n-|\boldsymbol{\varepsilon}|} \left( W_{-1/2+\boldsymbol{\varepsilon}}^\square; \mathbf{x}, \mathbf{x} \right) = \sum_{j=0}^d \binom{d}{j} \binom{d+n-j}{d}.$$

*Proof.* Setting  $\mathbf{y} = \mathbf{x}$  as indicated gives the identity (5.10) with the right-hand side equal to  $M_{n,d}^\square$ . Thus, it is sufficient to establish the identity:

$$(5.12) \quad M_{n,d}^\square = \frac{h_{n,d}}{(d-1)!} = \sum_{j=0}^d \binom{d}{j} \left[ \binom{d+n-j}{d} - \binom{d+n-1-j}{d} \right].$$

For this we use the following identity established in [4, Lemma 2.4],

$$\sum_{n=0}^{\infty} \frac{h_{n,d}(u)}{(d-1)!} r^n = \frac{(1-r^2)^d}{(1-2ru+r^2)^d}, \quad 0 \leq r < 1,$$

which implies, in particular, that the following generating function for  $h_{n,d}(1)$ ,

$$\sum_{n=0}^{\infty} \frac{h_{n,d}(1)}{(d-1)!} r^n = \frac{(1+r)^d}{(1-r)^d}, \quad 0 \leq r < 1,$$

We now show that the right-hand side of (5.12) satisfies the same generating function. Indeed, a quick computation shows that

$$\begin{aligned} & \sum_{n=0}^{\infty} \sum_{j=0}^d \binom{d}{j} \left[ \binom{d+n-j}{d} - \binom{d+n-1-j}{d} \right] r^n \\ &= (1-r) \sum_{j=0}^d \binom{d}{j} \sum_{n=0}^{\infty} \binom{d+n-j}{d} r^n \\ &= (1-r) \sum_{j=0}^d \binom{d}{j} \sum_{n=0}^{\infty} \binom{d+n}{d} r^{n+j} \\ &= (1-r)(1+r)^d (1-r)^{-d-1} = \frac{(1+r)^d}{(1-r)^d}. \end{aligned}$$

Since both sides of (5.12) satisfy the same generating function, they are equal. This proves (5.10). Finally, summing over  $m$  of (5.10) proves (5.11).  $\square$

As in the two previous cases, (5.10) is the same as (4.12), whereas (5.11) is equivalent to (4.13) by (2.3).

## 6. CONCLUSION

We have considered a set  $\Omega \subset \mathbb{R}^d$  that is the unit ball, the simplex, or the unit cube in  $\mathbb{R}^d$  in this paper. In each case, we proved a remarkable property of the orthonormal polynomials (of all degrees) with respect to the equilibrium measure  $\mu$  of  $\Omega$  and related measures  $g \cdot \mu$  on  $\Omega$ , absolutely continuous with respect to  $\mu$ , and whose density  $g$  is some product of polynomials that define the boundary of  $\Omega$ . This property is a multivariate analog and generalization of Pell's polynomial equation satisfied by the Chebyshev polynomial of the first and second kind for  $\Omega = [-1, 1] \subset \mathbb{R}$ . The property yields a similar and remarkable property for the Christoffel functions associated with those measures, already revealed in [8] for  $\Omega = [-1, 1]$  and only partially for the multivariate cases. Moreover, this (identity) property can be interpreted as a distinguished algebraic certificate *à la Putinar* that the constant polynomial  $\mathbf{1}$  is positive on the set  $\Omega$ . This distinguished representation of  $\mathbf{1}$  is a result of a duality between a pair of convex cones, due to Nesterov. We hope that our result will stimulate further investigation on the links of the Christoffel functions with those seemingly unrelated fields.

**Acknowledgement.** This work was initiated when both authors were hosted by the American Institute of Mathematics (AIM) in San José, CA, during a research week that was part of the SQuaRE program of AIM, and so the authors gratefully acknowledge support from AIM.

## REFERENCES

- [1] M. Baran. Complex equilibrium measure and Bernstein type theorems for compact sets in  $\mathbb{R}^n$ , Proc. Amer. Math. Soc. **123**(2), pp. 485–494, 1995.
- [2] E. Bedford, B.A. Taylor. The complex equilibrium measure of a symmetric convex set in  $\mathbb{R}^n$ , Trans. Amer. Math. Soc. **294**(2), pp. 705–717, 1986.

- [3] H. Berens and Y. Xu, Fejér means for multivariate Fourier series, *Math. Z.* **221** (1996), 449–465.
- [4] M. Buhmann, J. Jager, and Y. Xu,  $\ell^1$ -summability and Fourier series of B-splines with respect to their knots. arXiv:2208.02167.
- [5] C. Dunkl and Y. Xu, *Orthogonal polynomials of several variables*. 2nd ed. Encyclopedia of Mathematics and its Applications, vol. **155**, Cambridge Univ. Press, 2014.
- [6] M. Grant, S. Boyd. CVX: Matlab Software for Disciplined Convex Programming, version 2.1, <http://cvxr.com/cvx>, March 2014.
- [7] A. Kroó, D. S. Lubinsky. Christoffel functions and universality in the bulk for multivariate polynomials, *Canad. J. Math.* **65**(3), pp. 600–620, 2013.
- [8] J. B. Lasserre. Pell’s equation, sum-of-squares, and equilibrium of compact sets, *Comptes Rendus Mathématique* (2023). To appear. arXiv:2210.07608
- [9] J. B. Lasserre, E. Pauwels, M. Putinar. *The Christoffel-Darboux Kernel for Data Analysis*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, UK, 2022.
- [10] J. Mc Laughlin. Multivariable-polynomial solutions to Pell’s equation and fundamental units in real quadratic fields, *Pacific J. Math.* **210**(2), pp. 335–348, 2002.
- [11] Y. Nesterov. Squared functional systems and optimization problems, In *High Performance Optimization*, Applied Optimization series, vol. 33, H. Frenk, K. Roos, T. Terlaky, and Shuzong Zhang Eds., Springer, Boston, MA, 2000, pp. 405–440.
- [12] M. Putinar. Positive polynomials on compact semi-algebraic sets, *Indiana Univ. Math. J.* **42**(3), pp. 969–984, 1993.
- [13] W. A. Webb, Hisashi Yokota. Polynomial Pell’s equation, *Proc. Amer. Math. Soc.* **131**(4), pp. 993–1006, 2002.
- [14] Y. Xu, Christoffel functions and Fourier Series for multivariate orthogonal polynomials, *J. Approx. Theory*, **82** (1995), 205–239.
- [15] Y. Xu, Orthogonal structure and orthogonal series in and on a double cone or a hyperboloid. *Trans. Amer. Math. Soc.* **374** (2021), 3603–3657.

LAAS-CNRS AND INSTITUTE OF MATHEMATICS, UNIVERSITY OF TOULOUSE, LAAS, 7 AVENUE DU COLONEL ROCHE, 31077 TOULOUSE CÉDEX 4, FRANCE

*Email address:* `lasserre@laas.fr`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OREGON, EUGENE, OR 97403-1222, U.S.A.

*Email address:* `yuan@uoregon.edu`