

UC Santa Barbara

UC Santa Barbara Previously Published Works

Title

Perturbations of Christoffel–Darboux Kernels: Detection of Outliers

Permalink

<https://escholarship.org/uc/item/3c27b1c1>

Journal

Foundations of Computational Mathematics, 21(1)

ISSN

1615-3375

Authors

Beckermann, Bernhard

Putinar, Mihai

Saff, Edward B

et al.

Publication Date

2021-02-01

DOI

10.1007/s10208-020-09458-9

Peer reviewed

PERTURBATIONS OF CHRISTOFFEL-DARBOUX KERNELS. I: DETECTION OF OUTLIERS

BERNHARD BECKERMANN, MIHAI PUTINAR, EDWARD B. SAFF,
AND NIKOS STYLIANOPOULOS

ABSTRACT. Two central objects in constructive approximation, the Christoffel-Darboux kernel and the Christoffel function, are encoding ample information about the associated moment data and ultimately about the possible generating measures. We develop a multivariate theory of the Christoffel-Darboux kernel in \mathbb{C}^d , with emphasis on the perturbation of Christoffel functions and their level sets with respect to perturbations of small norm or low rank. The statistical notion of leverage score provides a quantitative criterion for the detection of outliers in large data. Using the refined theory of Bergman orthogonal polynomials, we illustrate the main results, including some numerical simulations, in the case of finite atomic perturbations of area measure of a 2D region. Methods of function theory of a complex variable and (pluri)potential theory are widely used in the derivation of our perturbation formulas.

CONTENTS

| | |
|--|----|
| 1. Introduction | 2 |
| 1.1. The Christoffel-Darboux kernel and Christoffel function | 2 |
| 1.2. Detecting outliers and anomalies in statistical data | 3 |
| 1.3. Outline | 6 |
| 2. Univariate and multivariate Christoffel-Darboux kernel | 6 |
| 2.1. Definition and basic properties in the univariate case | 7 |
| 2.2. Definition of the Christoffel-Darboux kernel in the multivariate case | 8 |
| 2.3. Basic properties of multivariate Christoffel-Darboux kernels | 12 |
| 2.4. Leverage scores, the Mahalanobis distance and Christoffel-Darboux kernels | 16 |
| 3. Approximation of Christoffel-Darboux kernels | 18 |
| 3.1. Comparing two measures | 18 |

Date: December 18, 2018.

2010 Mathematics Subject Classification. 34E10, 42C05, 46E22, 30H20, 30E05,

Key words and phrases. orthogonal polynomial, Christoffel-Darboux kernel, Green function, Siciak function, Bergman space, outlier, leverage score.

Acknowledgements. The first author was supported in part by the Labex CEMPI (ANR- 11-LABX-0007-01). The third author was partially supported by the U.S. National Science Foundation grant DMS-1516400. The fourth author was supported by the University of Cyprus grant 3/311-21027.

| | |
|---|----|
| 3.2. Small perturbations | 22 |
| 4. Additive perturbation of positive measures | 25 |
| 5. Cosine asymptotics in the univariate case | 28 |
| 6. Asymptotics in Bergman space | 33 |
| 7. Examples in the multivariate case | 41 |
| 7.1. Examples of plurisubharmonic Green functions with pole at infinity | 41 |
| 7.2. The real square with tensor product of Chebyshev weights | 43 |
| 7.3. Christoffel-Darboux kernel associated to Reinhardt domains in \mathbb{C}^d | 46 |
| 8. Index of notation | 49 |
| References | 49 |

1. INTRODUCTION

The wide scope of this first article in a series is a systematic study of perturbations of multivariate Christoffel-Darboux kernels. We depart from a Bourbaki style by treating in parallel a specific application, namely the detection of outliers in statistical data. Both themes have a rich and glorious history which cannot be condensed in a single research article. We will merely provide glimpses with precise references into the past or recent achievements as our story unfolds, continuously interlacing the two themes.

The reproducing kernel $K_n^\mu(z, w)$ on the space of univariate or multivariate polynomials of degree less than or equal to n , in the presence of an inner product derived from the Lebesgue space $L^2(\mu)$ associated with a positive measure μ in \mathbb{C}^d , bears the name of *Christoffel-Darboux kernel*. It is given in the univariate case by

$$K_n^\mu(z, w) = \sum_{j=0}^n p_j^\mu(z) \overline{p_j^\mu(w)},$$

where the p_j^μ are orthonormal polynomials of respective degrees j . For more than a century this object has repeatedly appeared (mostly in the one complex variable setting) as a central figure in interpolation, approximation, and asymptotic expansion problems.

1.1. The Christoffel-Darboux kernel and Christoffel function. Consider for instance the wide class of inverse problems leading in the end to the classical moment problem on the real line. There one has a grasp on $K_n(z, w) = K_n^\mu(z, w)$, $n \geq 1$, before knowing the representing measure(s) μ . This gives for complex and non-real values $z \in \mathbb{C}$ the radius of Weyl's circle of order n :

$$r_n(z) = \frac{1}{|z - \bar{z}| K_n(z, z)}.$$

The vanishing of the limit $r_\infty(z) = \lim_n r_n(z)$ for a single non-real z resolves without ambiguity the delicate question whether the moment problem is determinate (has a unique solution). For a real value $x \in \mathbb{R}$ and a fixed n the reciprocal $\frac{1}{K_n(x,x)}$, also known as the *Christoffel function*, gives a tight upper bound for the maximal mass a finite atomic measure with prescribed moments up to degree n can charge at the point x . An authoritative account of the central role of the reproducing kernel method in moment problems and the theory of orthogonal polynomials on the line or on the circle is contained in Freud's monograph [13]. For an enthusiastic and informative eulogy of Freud's predilection for the kernel method see [29].

Much happened meanwhile. On the theoretical side, Totik [41] has settled a long standing open problem concerning the limiting values of $\frac{n}{K_n(x,x)}$ for a positive measure on the line, culminating a long series of particular cases originating a century ago in the pioneering work of Szegő, see also the survey [37]. Roughly speaking, the limit of $\frac{1}{K_n(x,x)}$ detects the point mass charge at the point x in the support of the original measure μ , while $\frac{n}{K_n(x,x)}$ (seen as a function of x) gives the Radon-Nikodym derivative of μ with respect to the equilibrium measure of the support of orthogonality. One step further, Lubinsky [24] has studied a universality principle for the asymptotics of the kernel $K_n^\mu(x, y)$, highly relevant for the theory of random matrices. The weak-* limits of the counting measures of the eigenvalues of truncations of (unitary or self-adjoint) random matrices were recently unveiled via similar kernel methods by Simon [38].

All the above results refer to a single real variable. There is, however, notable progress in the study of the asymptotics of the Christoffel-Darboux kernel in the multivariate case cf. [7, 22, 23, 45]. These studies cover specific measures, in general invariant under a large symmetry group, the main theorems offering kernel asymptotics at points belonging to the support of the original measure. The behavior of the multivariate Christoffel-Darboux kernel outside the support of the generating measure is decided in most cases by a classical extremal problem in pluripotential theory. We shall provide details on the latter in the second section of this article and we will record a few illustrative cases in the last section.

1.2. Detecting outliers and anomalies in statistical data. But that's not all. Scientists outside the strict boundaries of the mathematics community continue to rediscover the wonders of Christoffel-Darboux type kernels. For a long time statisticians used the kernel method to estimate the probability density of a sequence of identically distributed random variables [31], or to separate observational data by the level set of a positive definite kernel, a technique originally developed for pattern recognition [44]. The last couple of years have witnessed the entry of Christoffel-Darboux kernel through the front door of the very dynamical fields of Machine Learning and Artificial Intelligence [26, 27, 19, 20]. Notice that in these applications μ is typically

an atomic (empirical) measure with finite support in \mathbb{C}^d representing a point cloud, which makes analysis like kernel asymptotics more involved.

In the case $d = 1$, Malyshkin [26] suggested the use of modified moments, that is, computing the Christoffel-Darboux kernel of a measure μ from orthogonal polynomials of a similar but classical measure (Jacobi, Laguerre, Hermite), without going through monomials, an idea also used in the modified Chebyshev algorithm of Sack and Donovan [15]. Some numerical evidence is provided for numerical stability of this approach, even for high degrees n . The approximation of the support of the original measure and detection of outliers became thus possible through the investigation of the level sets of the Christoffel function [19, 20]. Not unrelated, three of the authors of this article have successfully used the Christoffel function for reconstruction, from indirect measurements in 2D, of the boundaries of an archipelago of islands each carrying the same constant density against area measure [17].

The aim of this first essay in a series is to study the effect of various perturbations on Christoffel-Darboux kernels associated to positive measures compactly supported by \mathbb{C}^d . An important thesis one can draw from classical constructive function theory is that even purely real questions ask for the domain of the reproducing kernel, potential functions, spectra of associated operators to be extended to complex values. We therefore suggest to consider multivariate orthogonal polynomials in \mathbb{C}^d , where in applications real point clouds could be imbedded in \mathbb{C}^d by using the isomorphism with \mathbb{R}^{2d} . As we will recall in (6) for $d = 1$ and in Lemma 2.9 for $d > 1$, the Christoffel-Darboux kernel $K_n^\mu(z, z)$ typically has exponential growth in Ω (the unbounded connected component of $\mathbb{C}^d \setminus \text{supp}(\mu)$) but not in $\text{supp}(\mu)$. This confirms the claim of [26, 27, 19, 20] that level sets of the Christoffel-Darboux kernel for sufficiently large n allow us to detect *outliers*, that is, points of $\text{supp}(\mu)$ which are isolated in Ω .

However, this claim is certainly wrong for atomic measures $\tilde{\nu}$ (also referred to as discrete measures)

$$\tilde{\nu} = \sum_{j=1}^N t_j \delta_{z^{(j)}}$$

for $t_j > 0$ and distinct $z^{(j)} \in \mathbb{C}^d$ with support describing a cloud of points since, following the above terminology, every element of $\text{supp}(\tilde{\nu})$ would be an outlier, and, even worse, there is only a finite number of orthogonal polynomials making kernel asymptotics impossible. Here typically the point cloud splits into two parts, namely most of the points approximating well some continuous set $S \subseteq \mathbb{C}^d$, and some outliers sufficiently far from S . To formalize this, we suppose that $N \gg n$, and that the atomic measure $\tilde{\nu}$ can be split into two atomic measures

$$\tilde{\nu} = \tilde{\mu} + \sigma, \quad \tilde{\mu} \approx \mu, \quad \text{supp}(\sigma) \subseteq \Omega, \quad (1)$$

with μ and Ω as before and n large enough such that we have sufficient information about $K_n^\mu(z, z)$. However, we should insist that in applications we only know the (empirical) measure $\tilde{\nu}$ and possibly its moments, and we want to learn from level lines of $K_n^{\tilde{\nu}}(z, z)$ the set of outliers $\text{supp}(\sigma)$, and, if possible, $\text{supp}(\mu)$ or even μ itself.

The notion $\tilde{\mu} \approx \mu$ in (1) means that the moments of both measures up to order n are sufficiently close, measured in §3.2 in terms of modified moments, to insure that $K_n^\mu(z, z)$ is sufficiently close to $K_n^{\tilde{\mu}}(z, z)$ for all $z \in \mathbb{C}^d$. In order to deduce lower and upper bounds for $K_n^{\tilde{\nu}}(z, z)$ and actually show that level lines of this Christoffel-Darboux kernel are indeed useful for outlier detection, we still have to examine how a Christoffel-Darboux kernel changes after adding a finite number of point masses represented by σ , that is, provide lower and upper bounds for

$$K_n^{\mu+\sigma}(z, z)/K_n^\mu(z, z) \quad \text{and} \quad K_n^{\tilde{\mu}+\sigma}(z, z)/K_n^{\tilde{\mu}}(z, z),$$

which is the scope of §4. For the special case $d = 1$, we establish in §5 asymptotics for $K_n^{\mu+\sigma}(z, z)/K_n^\mu(z, z)$ in the special case where we have ratio asymptotics for the underlying μ -orthogonal polynomials.

Roughly speaking, as long as $N \gg n$, by combining the above results we find two different regimes, namely for $z \in \text{supp}(\sigma) \cup \text{supp}(\mu)$ where $K_n^{\tilde{\nu}}(z, z)$ essentially remains bounded, and on compact subsets of the complement $\Omega \setminus \text{supp}(\sigma)$ where $K_n^{\tilde{\nu}}(z, z)$ grows exponentially like $K_n^\mu(z, z)$. We may conclude that level lines for $K_n^{\tilde{\nu}}(z, z)$ for sufficiently large parameters will separate outliers from $\text{supp}(\mu)$, and refer the readers to [19, 20] for a practical choice of the parameter of critical level lines.

We expect the required critical level lines to encircle once either $\text{supp}(\mu)$ or exactly one of the outliers, but not the others. In practical terms, it is non-trivial to draw on a computer certified level lines because of the stiff gradients of $K_n^{\tilde{\nu}}(z, z)$, and hence this approach might be quite costly.

Therefore, we suggest an alternate strategy with an overall cost which scales linearly with N (and at most cubic in n): Inspired by related results in statistics and data analysis, we are considering in Section 2.4 a leverage score attached to each mass point $z^{(j)}$, namely

$$t_j K_n^{\tilde{\nu}}(z^{(j)}, z^{(j)}) \in (0, 1],$$

which we expect to be close to 1 iff $z^{(j)}$ is an outlier. For the particular case of $d = 1$ and μ being area measure on a simply connected and bounded subset of the complex plane, the above-mentioned results on the different Christoffel-Darboux kernels allow us to fully justify this claim, see Corollary 6.5 and the following discussion and numerical experiments.

In this paper we do not discuss and exploit for the moment any recurrence relation between multi-variate orthogonal polynomials and thus the d underlying commuting Hessenberg operators representing the multiplication with the independent variables z_1, z_2, \dots, z_d . This will be considered in a sequel publication.

Adding point masses to a given measure and recording the change of orthogonal polynomials is not a new theme, see for instance the historical notes in [36, p. 684]. The latter addresses measures on the unit circle, but the same applies to orthogonality on the real line. In case of several variables, quite a few authors have considered classical measures μ with explicit formulas for the corresponding orthogonal polynomials like Jacobi-type measures on the unit ball, and σ being one [9] or several mass points [10], or the Lebesgue measure on the sphere [28, Theorem 3.3].

1.3. Outline. §2 is devoted to the construction of multivariate orthogonal polynomials and the corresponding Christoffel-Darboux kernel, pointing out in §2.2 the inherent complications of the multivariate setting: choice of ordering of orthogonal polynomials, degeneracy due to algebraic constraints in the support, lack of positivity characterization of moment matrices, and so on. We also explore in §2.3 to what extent some elementary properties of univariate Christoffel-Darboux kernels recalled in §2.1 (like kernel asymptotics) generalize to the multivariate setting.

In §3.2 we give sufficient conditions insuring that, for fixed n , the Christoffel-Darboux kernels for two different measures behave similarly. This requires us to first address in §3.1 the question of how to compute a family of orthogonal polynomials p'_n from p_n^μ for two measures μ, ν , giving rise to so-called modified moment matrices.

Our first main result can be found in §4 where we discuss new upper and lower bounds for the change of the Christoffel-Darboux kernel after adding a finite number of point masses. A second main result given in §5 deals with the special case $d = 1$ of univariate polynomials and describes asymptotics for the change of the Christoffel-Darboux kernel after adding a finite number of point masses. This second result is illustrated in §6 by analyzing Bergman orthogonal polynomials. Combining these findings we get a complete analysis of outlier detection in \mathbb{C} in situation (1) under some smoothness assumptions on μ . We finally illustrate by examples in §7 that the situation is more involved in the multivariate setting, especially because of the lack of knowledge on asymptotics of some normalized Christoffel-Darboux kernel $C_n^\mu(z, w)$. For the convenience of the reader, an index of notation is included in §8.

Acknowledgements. The authors are indebted to the Mathematical Research Institute at Oberwolfach, Germany, which provided exceptional working conditions for a Research in Pairs collaboration in 2016 during which time the main ideas of this manuscript were developed.

2. UNIVARIATE AND MULTIVARIATE CHRISTOFFEL-DARBOUX KERNEL

The role of reproducing kernels in the development of the classical theories of moments, orthogonal polynomials, statistical mechanics and in general function theory cannot be overestimated. Numerous studies, old and new,

recognize the asymptotics of the Christoffel functions as the key technical ingredient for a variety of interpolation and approximation questions.

Our aim is to study the stability of the Christoffel-Darboux kernels under small perturbations or additive perturbations of the underlying measure. Quite a long time ago it was recognized that even purely real questions of, say approximation, have to be treated with methods of function theory of a complex variable and potential theory. For this reason, we advocate below the complex variable setting. For the advantages and richness of the theory of orthogonal polynomials in a single complex variable see [33, 39].

2.1. Definition and basic properties in the univariate case. Consider the Lebesgue space $L^2(\mu)$ with scalar product and norm

$$\langle f, g \rangle_{2,\mu} = \int f(z) \overline{g(z)} d\mu(z), \quad \|f\|_{2,\mu} = \left(\langle f, f \rangle_{2,\mu} \right)^{1/2}, \quad (2)$$

where μ is a positive Borel measure in the complex plane. Assume that the measure μ has compact and infinite support $\text{supp}(\mu)$, so that one can orthogonalize without ambiguity the sequence of monomials. Then the Gram-Schmidt process gives us the associated sequence of orthonormal polynomials, p_j^μ of degree j , with positive leading coefficient, together with the Christoffel-Darboux kernel

$$K_n^\mu(z, w) = \sum_{j=0}^n p_j^\mu(z) \overline{p_j^\mu(w)} \quad (3)$$

for the subspace of polynomials of degree at most n . The case $z = w$ will be referred to as the *diagonal kernel*, whereas for $z \neq w$ we speak of a *polarized kernel*.

For any $z \in \mathbb{C}$ and any polynomial p of degree at most n we have the reproducing property

$$\langle p, K_n^\mu(\cdot, z) \rangle_{2,\mu} = \int K_n^\mu(z, w) p(w) d\mu(w) = p(z); \quad (4)$$

in other words, the function $w \mapsto K_n^\mu(w, z)$ represents the linear functional of point evaluation $p \mapsto p(z)$ in the algebra of polynomials. Simple Hilbert space arguments show that $K_n^\mu(z, w) = \langle K_n^\mu(\cdot, w), K_n^\mu(\cdot, z) \rangle_{2,\mu}$, and that, for all $z \in \mathbb{C}$,

$$\frac{1}{\sqrt{K_n^\mu(z, z)}} = \min\{\|p\|_{2,\mu} : p \text{ polynomial of degree } \leq n, p(z) = 1\}, \quad (5)$$

with the unique extremal polynomial being given by $x \mapsto K_n^\mu(x, z)/K_n^\mu(z, z)$.

The reciprocal $\lambda_n(z) = 1/K_n^\mu(z, z)$ naturally appears in a variety of approximation questions in the complex plane, and is traditionally called the *Christoffel* or *Christoffel-Darboux* function. For instance, it is shown by Totik [42, Theorem 1.2] that, provided that μ is sufficiently smooth and supported on a system of arcs, the quantity $n/K_n(z, z)$ tends to the Radon-Nykodym derivative of μ on open subarcs. If z belongs to the

(two-dimensional) interior of $\text{supp}(\mu)$, a similar result is established for $n^2/K_n^\mu(z, z)$ in [42, Theorem 1.4]. Under weaker assumptions we may also describe n th root limits of the Christoffel-Darboux kernel: Denote by Ω the unbounded connected component of $\mathbb{C} \setminus \text{supp}(\mu)$, with $g_\Omega(\cdot)$ its Green function¹ with pole at ∞ . Using the inequality

$$\|p\|_{2,\mu}^2 \leq \mu(\mathbb{C}) \|p\|_{L^\infty(\text{supp}(\mu))}^2$$

in (5) and, e.g., the extremal property [35, Theorem III.1.3] of Fekete points, we get at least exponential growth for the Christoffel-Darboux kernel in Ω , namely

$$\liminf_{n \rightarrow \infty} K_n^\mu(z, z)^{1/n} \geq e^{2g_\Omega(z)} \quad (6)$$

uniformly on compact subsets of Ω . Moreover, if Ω is regular with respect to the Dirichlet problem, then for sufficiently “dense” measures, namely the class **Reg** introduced in [39, §3], it is known by [39, Theorem 3.2.3] that

$$\lim_{n \rightarrow \infty} K_n^\mu(z, z)^{1/n} = e^{2g_\Omega(z)} \quad (7)$$

uniformly on compact subsets of \mathbb{C} . In particular, there is no exponential growth on the support. Roughly saying, this is the motivation of [19, 20] for considering level sets of the Christoffel-Darboux kernel to help detect outliers, that is, points of the support of μ which are isolated in Ω .

2.2. Definition of the Christoffel-Darboux kernel in the multivariate case. The theory of orthogonal polynomials of several (real or complex) variables is much more involved and thus less developed, we refer the reader for instance to [11] or the recent summary [12] for the case of several real variables. Briefly speaking, we have to face at least two basic problems which we will address in the next two subsections. The first problem is that there is no natural ordering of multivariate monomials. Second, the multivariate monomials might not be linearly independent in $L^2(\mu)$, or, in other words, the null space

$$\mathcal{N}(\mu) = \{p \in \mathbb{C}[z] : \int |p|^2 d\mu = 0\}$$

might be nontrivial. This means in particular that we have to revisit the reproducing property (4), and the extremal problems (5) and (6), the aim of §2.3.

Henceforth $\mathbb{C}[z]$ stands for the algebra of polynomials with complex coefficients in the indeterminates $z = (z_1, z_2, \dots, z_d)$. Notice that every function continuous in \mathbb{C}^d and vanishing μ -almost everywhere has to vanish on $\text{supp}(\mu)$, which actually shows that $\mathcal{N}(\mu)$ does only depend on $\text{supp}(\mu)$:

$$\mathcal{N}(\mu) = \{p \in \mathbb{C}[z] : p \text{ vanishes on } \text{supp}(\mu)\}.$$

It immediately follows that $\mathcal{N}(\mu)$ is a (radical) ideal in the algebra $\mathbb{C}[z]$.

¹Recall that $g_\Omega(z) > 0$ for $z \in \Omega$, and equal to infinity if $\text{supp}(\mu)$ is polar, that is, of logarithmic capacity zero.

Hereafter, μ denotes a positive Borel measure with compact support in \mathbb{C}^d , with corresponding scalar product as in (2). Depending on the context, the hermitian inner product in \mathbb{C}^d is denoted as follows:

$$\langle z, w \rangle = z \cdot \bar{w} = w^* z = z_1 \bar{w}_1 + z_2 \bar{w}_2 + \cdots + z_d \bar{w}_d.$$

In a few instances we will also encounter the bilinear form

$$z \cdot w = z_1 w_1 + z_2 w_2 + \cdots + z_d w_d.$$

The standard multi-index notation

$$z^\alpha = z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_d^{\alpha_d}, \quad \alpha = (\alpha_1, \alpha_2, \dots, \alpha_d) \in \mathbb{N}^d,$$

is adopted. The graded lexicographical order on the set of indices is denoted by “ \leq_ℓ ”. Specifically

$$\alpha \leq_\ell \beta$$

if either $\alpha = \beta$ or $\alpha < \beta$; i.e.,

$$|\alpha| = \alpha_1 + \cdots + \alpha_d < \beta_1 + \cdots + \beta_d = |\beta|,$$

or

$$|\alpha| = |\beta| \quad \text{and} \quad \alpha_1 = \beta_1, \dots, \alpha_k = \beta_k, \alpha_{k+1} < \beta_{k+1},$$

for some $k < d$.

In what follows we suppose that a fixed enumeration $\alpha(0), \alpha(1), \alpha(2), \dots$ of all multi-indices is given with $\alpha(0) = 0$, for instance (but not necessarily) a graded lexicographical ordering. Henceforth we assume that multiplication by any variable increases the order;² that is, for all j, n that $z_j z^\alpha = z^\beta$ for $\beta > \alpha$. We write $\deg p = k$ if p is a linear combination of the monomials $z^{\alpha(0)}, z^{\alpha(1)}, \dots, z^{\alpha(k)}$, with the coefficient of $z^{\alpha(k)}$ being non-trivial. Note that, for $d > 1$, our notion of degree differs from the so-called total degree $\text{tdeg } p$ being the largest of the orders $|\alpha|$ of monomials z^α occurring in p .

Example 2.1. Consider the positive measure μ on \mathbb{C}^2 given by

$$\int \overline{p(z_1, z_2)} q(z_1, z_2) d\mu(z) = \int_{|u| \leq 1} \overline{p(u^3, u^2)} q(u^3, u^2) dA(u),$$

where $p, q \in \mathbb{C}[z]$ and dA stands for the area measure in \mathbb{C} . Notice that $\text{supp}(\mu)$ is compact and strictly contained in the complex affine variety

$$\{(z_1, z_2) \in \mathbb{C} : P(z_1, z_2) = 0\}; \quad P(z_1, z_2) := z_1^2 - z_2^3,$$

in particular, the ideal $\mathcal{N}(\mu)$ is generated by P . The graded lexicographical ordering for bivariate monomials is given by

$$1, z_1, z_2, z_1^2, z_1 z_2, z_2^2, z_1^3, z_1^2 z_2, z_1 z_2^2, z_2^3, z_1^4, z_1^3 z_2, z_1^2 z_2^2, z_1 z_2^3, z_2^4, \dots$$

and, for instance, $\alpha(8) = (2, 1)$. Notice that, seen as functions restricted to the above affine variety, the two monomials z_2^3 and z_1^2 cannot be distinguished, and the same is true for $z_1 z_2^3$ and z_1^4 , and for z_2^4 and $z_1^3 z_2$, and so

² This is equivalent to requiring that any subset of multi-indices $\{\alpha(0), \alpha(1), \alpha(2), \dots, \alpha(n)\}$ is downward close, compare with, e.g., [8].

on. We thus have to go to the quotient space $\mathbb{C}[z_1, z_2]/\mathcal{N}(\mu)$, with a basis given by

$$1, z_1, z_2, z_1^2, z_1 z_2, z_2^2, z_1^3, z_1^2 z_2, z_1 z_2^2, z_1^4, \dots, z_1^n, z_1^{n-1} z_2, z_1^{n-2} z_2^2, z_1^{n+1}, \dots,$$

written in the same order. We will see in Lemma 2.3 below, that exactly this maximal set of monomials being linearly independent in $L^2(\mu)$ will allow us to construct corresponding multivariate orthogonal polynomials.

Definition 2.2. *Let $k \in \mathbb{N}$. The tautological vector of ordered monomials is the row vector with $(k + 1)$ components, given by*

$$v_k(z) = (z^{\alpha(j)})_{j=0,1,\dots,k}.$$

The corresponding matrix of moments is defined by

$$\widetilde{M}_k(\mu) = \left(\langle z^{\alpha(\ell)}, z^{\alpha(j)} \rangle_{2,\mu} \right)_{j,\ell=0}^k = \int v_k(z)^* v_k(z) d\mu(z).$$

Notice that $\widetilde{M}_k(\mu)$ is a Gram matrix and hence positive semi-definite. However, these moment matrices are not necessarily of full rank $k + 1$. This may happen already in the case $d = 1$ of univariate orthogonal polynomials, where $z^{\alpha(j)} = z^j$ and the support of μ is finite. Indeed, for $d = 1$, assume that $\det \widetilde{M}_n(\mu) = 0$ and n is the first index with this property. That means there exists a non-trivial combination of the rows of the matrix $\widetilde{M}_n(\mu)$, or in other terms, there exists a monic polynomial $h(z) = z^n + \dots$ of exact degree n , such that $h \in \mathcal{N}(\mu)$. That is, μ is an atomic measure with exactly n distinct points in its support, namely the roots of h . Consequently, $h(z)z^j$ is also annihilated by μ for all $j \geq 0$, showing that $\mathcal{N}(\mu)$ is the ideal generated by h , and

$$\text{rank} \widetilde{M}_{n+j}(\mu) = \text{rank} \widetilde{M}_n(\mu) = n, \quad j \geq 0.$$

In case of several variables, the ideal $\mathcal{N}(\mu)$ is not necessarily a principal ideal, but it is finitely generated, according to Hilbert's basis theorem. However, in general the above rank stabilization phenomenon of moment matrices ceases to hold for $d > 1$.

In the case $d \geq 1$, we may still generate orthogonal polynomials p_0^μ, p_1^μ, \dots by the Gram-Schmidt procedure starting with the family of functions $z^{\alpha(0)}, z^{\alpha(1)}, \dots$; however, it may happen that not every monomial $z^{\alpha(k)}$ gives raise to a new orthogonal polynomial. Indeed, since we have fixed an ordering of the monomials, we may and will construct a maximal set of monomials $z^{\alpha(k_n^\mu)}$, with $0 = k_0^\mu < k_1^\mu < \dots$ being linearly independent in $L^2(\mu)$, and an orthogonal basis p_0^μ, p_1^μ, \dots of the set spanned by these monomials. As we will see in Lemma 2.6 below, this set together with $\mathcal{N}(\mu)$ will allow us to write the set of multivariate polynomials as an orthogonal sum.

The next result, necessarily technical, formalizes the construction of orthogonal polynomials in the presence of an ordering of the monomials and possible linear dependence of the associated moments. The interested reader might first have a look at the special case $\mathcal{N}(\mu) = \{0\}$ where the moment

matrices $\widetilde{M}_k(\mu)$ have full rank $k + 1$ for all $k \in \mathbb{N}$, and thus orthogonal polynomials exist for each degree $n = k_n^\mu$.

Lemma 2.3. *We may construct multivariate orthogonal polynomials p_0^μ, p_1^μ, \dots with*

$$\deg p_n^\mu = k_n^\mu := \min\{k \geq 0 : \text{rank} \widetilde{M}_k(\mu) = n + 1\}; \quad (8)$$

in particular, $k_0^\mu = 0$. Moreover, for $k = k_n^\mu$, there exists an invertible and upper triangular matrix $\widetilde{R}_k(\mu)$ of order $k + 1$ such that

$$\widetilde{R}_k(\mu)^* \widetilde{M}_k(\mu) \widetilde{R}_k(\mu) = E_k(\mu) E_k(\mu)^*, \quad E_k(\mu) := (\delta_{j,k_\ell^\mu})_{j=0, \dots, k, \ell=0, \dots, n}. \quad (9)$$

Furthermore, $M_n(\mu) = E_k(\mu)^ \widetilde{M}_k(\mu) E_k(\mu)$ is an maximal invertible submatrix of $\widetilde{M}_k(\mu)$, and*

$$R_n(\mu)^* M_n(\mu) R_n(\mu) = I, \quad v_n^\mu := (p_0^\mu, \dots, p_n^\mu) = v_k E_k(\mu) R_n(\mu), \quad (10)$$

with the upper triangular and invertible matrix $R_n(\mu) = E_k(\mu)^ \widetilde{R}_k(\mu) E_k(\mu)$.*

Proof. We construct orthogonal polynomials with (8) by recurrence. Since $\|1\|_{2,\mu}^2 = \mu(\mathbb{C}^d) \neq 0$, we may define $q_0^\mu(z) = 1$ and $p_0^\mu(z) = 1/\sqrt{\mu(\mathbb{C}^d)}$. At step $k \geq 1$, with $k_{n-1}^\mu < k \leq k_n^\mu$, we compute $q_k^\mu = v_k \xi$ by orthogonalizing $z^{\alpha(k)}$ against $p_0^\mu, \dots, p_{n-1}^\mu$. By recurrence hypothesis (8) and the assumption $k_{n-1}^\mu < k$, the degree of any of the polynomials $p_0^\mu, \dots, p_{n-1}^\mu$ is strictly less than k . Thus $\deg q_k^\mu = k$; that is, the last component of ξ is nontrivial. If $k = k_n^\mu$, then by definition of k_n^μ the last column of $\widetilde{M}_k(\mu)$ is linearly independent of the others, and hence $\widetilde{M}_k(\mu)\xi \neq 0$. Using the fact that $\widetilde{M}_k(\mu)$ is Hermitian and positive semi-definite, we infer

$$\|q_k^\mu\|_{2,\mu}^2 = \xi^* \widetilde{M}_k(\mu) \xi \neq 0.$$

Consequently one can define $p_n^\mu = q_k^\mu / \|q_k^\mu\|_{2,\mu}$ of degree k_n^μ , with positive leading coefficient. If however $k < k_n^\mu$, we find that $\widetilde{M}_k(\mu)\xi = 0$ and thus $\|q_k^\mu\|_{2,\mu}^2 = 0$, in other words, $q_k^\mu \in \mathcal{N}(\mu)$. The same argument shows that there is no polynomial of degree exactly k and of norm 1 that is orthogonal to $p_0^\mu, \dots, p_{n-1}^\mu$. This proves (8).

Redefining $q_k^\mu(z) = p_n^\mu(z)$ for all $k = k_n^\mu$, we have thus constructed a basis $\{q_0^\mu, \dots, q_k^\mu\}$ of the space of polynomials of degree at most k ; in other words, there exists an upper triangular and invertible matrix $\widetilde{R}_k(\mu)$ with

$$(q_0^\mu(z), \dots, q_k^\mu(z)) = v_k(z) \widetilde{R}_k(\mu),$$

and (9) follows by orthonormality.

By construction, p_n^μ contains only the powers

$$z^{\alpha(k_0^\mu)}, \dots, z^{\alpha(k_n^\mu)};$$

that is, the columns of the coefficient vectors of our orthogonal polynomials satisfy

$$\widetilde{R}_k(\mu) E_n(\mu) = E_n(\mu) E_n(\mu)^* \widetilde{R}_k(\mu) E_n(\mu) = E_n(\mu) R_n(\mu),$$

which implies (10). \square

Notice that (9) gives a full rank decomposition $\widetilde{M}_k(\mu) = XX^*$, with the factor $X^* = E_n(\mu)^* \widetilde{R}_n(\mu)^{-1}$ being of upper echelon form, the pivot in the j -th column lying in row k_j^μ . We also learn from (9) that a basis of the nullspace of $\widetilde{M}_k(\mu)$ is given by the columns of $\widetilde{R}_k(\mu)$ with indices different from k_0^μ, k_1^μ, \dots . Finally, from (9), $R_n(\mu)^{-1}$ is the Cholesky factor of the invertible matrix $M_n(\mu)$.

With the vector of orthogonal polynomials v_n^μ as in Lemma 2.3, we are now prepared to define the *multivariate Christoffel-Darboux kernel* of order n as

$$K_n^\mu(z, w) = v_n^\mu(z) v_n^\mu(w)^* = \sum_{j=0}^n p_j^\mu(z) \overline{p_j^\mu(w)},$$

where we notice that, according to (10),

$$K_n^\mu(z, w) = v_{k_n^\mu}(z) E_n(\mu) M_n(\mu)^{-1} E_n(\mu)^* v_{k_n^\mu}(w)^*.$$

For later use we mention the affine invariance of orthogonal polynomials and Christoffel functions : consider the change of variables $\tilde{z} = Bz + b$ for a diagonal and invertible $B \in \mathbb{C}^{d \times d}$ and $b \in \mathbb{C}^d$, and let $c > 0$. If we define $\tilde{\mu}$ by $\tilde{\mu}(S) = c\mu(BS + b)$ for any Borel set $S \subseteq \mathbb{C}^d$, then

$$K_n^\mu(z, w) = c K_n^{\tilde{\mu}}(Bz + b, Bw + b), \quad (11)$$

compare with [19, Lemma 1] where the authors allow also for general invertible matrices B but then require graded lexicographical ordering and the additional assumption $\mathcal{N}(\mu) = \{0\}$.

2.3. Basic properties of multivariate Christoffel-Darboux kernels.

Definition 2.4. Consider for $k = k_n^\mu$ the following two subsets of multivariate polynomials

$$\begin{aligned} \mathcal{N}_n(\mu) &:= \{v_k \xi : \widetilde{M}_k(\mu) \xi = 0\} = \{p \in \mathcal{N}(\mu) : \deg p \leq k\}, \\ \mathcal{L}_n(\mu) &:= \text{span}\{p_j^\mu : j = 0, 1, \dots, n\} = \text{span}\{z^{\alpha(k_j^\mu)} : j = 0, 1, \dots, n\}, \end{aligned}$$

and the affine variety

$$\mathcal{S}_n(\mu) := \bigcap_{p \in \mathcal{N}_n(\mu)} \{z \in \mathbb{C}^d : p(z) = 0\}.$$

By definition of $\mathcal{N}_n(\mu)$, we get $\text{supp}(\mu) \subseteq \mathcal{S}_n(\mu)$. Notice that $\mathcal{N}_n(\mu)$ is increasing in n and hence $\mathcal{S}_n(\mu)$ decreases in n . Moreover, recalling that $\mathcal{N}(\mu) = \bigcup_{n \geq 0} \mathcal{N}_n(\mu)$ is a finitely generated ideal, we see that, for n being larger than or equal to the largest of the degrees of the generators, the set $\mathcal{S}_n(\mu)$ does no longer depend on n , and we will use the notation $\mathcal{S}(\mu)$, being the *Zarisky closure* of $\text{supp}(\mu)$.

Notice that $\mathcal{N}(\mu) = \{0\}$ implies that $\mathcal{S}(\mu) = \mathbb{C}^d$. In case of an atomic measure μ with a finite number of point masses one can show that $\mathcal{S}(\mu) =$

$\text{supp}(\mu)$, but in general $\mathcal{S}(\mu)$ is larger than $\text{supp}(\mu)$; see for instance Example 2.1.

Examples 2.5. Consider two measures μ, ν , with $\text{supp}(\mu) \subseteq \text{supp}(\nu)$, subject for instance to the condition $\nu - \mu \geq 0$. Then obviously $\mathcal{N}(\nu) \subseteq \mathcal{N}(\mu)$, with equality if and only if $\text{supp}(\nu) \subseteq \mathcal{S}(\mu)$. The Lebesgue measure μ on the real or complex unit ball in \mathbb{C}^d has been extensively studied; in particular $\mathcal{N}(\mu) = \{0\}$, see §7. We conclude that, if $\text{supp}(\nu)$ has a non-empty interior in \mathbb{C}^d (or $\text{supp}(\mu) \subseteq \mathbb{R}^d$ has a non-empty interior in \mathbb{R}^d), then $\mathcal{N}(\nu) = \{0\}$ and $\mathcal{S}(\nu) = \mathbb{C}^d$.

It is easy to check that we keep the reproducing property (4) for $p \in \mathcal{L}_n(\mu)$, but it fails for nontrivial elements of $\mathcal{N}_n(\mu)$. The following result shows that the set of multivariate polynomials of degree at most $k = k_n^\mu$ can be written as an orthogonal sum $\mathcal{L}_n(\mu) \oplus \mathcal{N}_n(\mu)$. Thus the multivariate Christoffel-Darboux kernel is indeed a projection kernel onto $\mathcal{L}_n(\mu)$, and only a reproducing kernel in the case $\mathcal{N}_n(\mu) = \{0\}$.

Lemma 2.6. *Let $k, n \in \mathbb{N}$ with $k_n^\mu \leq k < k_{n+1}^\mu$, and p be a polynomial of degree $\leq k$. Then the polynomial defined by*

$$q(z) := \int K_n^\mu(z, w)p(w)d\mu(w)$$

satisfies $q \in \mathcal{L}_n(\mu)$ and $p - q \in \mathcal{N}_{n+1}(\mu)$. In particular, for $k = k_n^\mu = \deg p_n^\mu$ we have that

$$\mathcal{N}_n(\mu) = \{p \in \mathbb{C}[z] : \deg p \leq \deg p_n^\mu, p \perp_\mu p_0^\mu, \dots, p_n^\mu\}. \quad (12)$$

Proof. By construction, $q \in \mathcal{L}_n(\mu)$. Also, $\deg(p - q) \leq k < \deg p_{n+1}^\mu$, and, by orthogonality,

$$\int |p - q|^2 d\mu = \|p\|_{2,\mu}^2 - \|q\|_{2,\mu}^2 = \|p\|_{2,\mu}^2 - \sum_{j=0}^n |\langle p, p_j^\mu \rangle_{2,\mu}|^2 = 0,$$

showing that $p - q \in \mathcal{N}_{n+1}(\mu)$, as claimed in the first assertion. Notice that in the special case $k = k_n^\mu$, the same reasoning gives the slightly more precise information $p - q \in \mathcal{N}_n(\mu)$, and $p \in \mathcal{N}_n(\mu)$ iff $q = 0$ iff $p \perp_\mu p_0^\mu, \dots, p_n^\mu$, leading to the second part. \square

Let us now investigate to what extent the extremal property (5) remains true in the multivariate setting. If $z \notin \mathcal{S}_n(\mu)$, then by definition there exists $p \in \mathcal{N}_n(\mu)$ with $p(z) = 1$, and hence $\inf\{\|p\|_{2,\mu} : \deg p \leq k_n^\mu, p(z) = 1\} = 0$, showing that (5) fails. There are two ways to specialize this extremal problem in order to recover (5).

Lemma 2.7. *For all $n \in \mathbb{N}$ and $z \in \mathbb{C}^d$*

$$\frac{1}{\sqrt{K_n^\mu(z, z)}} = \min\{\|p\|_{2,\mu} : p \in \mathcal{L}_n(\mu), p(z) = 1\}, \quad (13)$$

and for all $k, n \in \mathbb{N}$ with $k_n^\mu \leq k < k_{n+1}^\mu$ and $z \in \mathcal{S}_{n+1}(\mu)$

$$\frac{1}{\sqrt{K_n^\mu(z, z)}} = \min\{\|p\|_{2,\mu} : \deg p \leq k, p(z) = 1\}. \quad (14)$$

Proof. For a proof of (13), we use exactly the same argument as in the univariate case, namely the Cauchy-Schwarz inequality and its sharpness. Let now be $z \in \mathcal{S}_{n+1}(\mu)$. Then, with p and $q \in \mathcal{L}_n(\mu)$ as in Lemma 2.6, we have that $\|p\|_{2,\mu} = \|q\|_{2,\mu}$, and $p(z) - q(z) = 0$ by definition of $\mathcal{S}_{n+1}(\mu)$. This shows that in (14) we may restrict ourselves to $p \in \mathcal{L}_n(\mu)$, and (14) follows from (13). \square

We learn from (14) that, for $z \in \mathcal{S}(\mu)$, the Christoffel-Darboux kernel $K_n^\mu(z, z)$ does not depend on the particular ordering of monomials as long as the set $\{\alpha(0), \alpha(1), \dots, \alpha(k)\}$ remains the same.

Of related interest will be the *cosine function*

$$C_n^\mu(z, w) := \frac{K_n^\mu(z, w)}{\sqrt{K_n^\mu(z, z)}\sqrt{K_n^\mu(w, w)}}. \quad (15)$$

To the best of our knowledge, there are hardly any results in the literature concerning asymptotics for $n \rightarrow \infty$ for such a normalized and polarized Christoffel-Darboux kernel. Multipoint matricial analogs of these kernels will be used later on.

Definition 2.8. Let $z_1, z_2, \dots, z_\ell, w_1, w_2, \dots, w_\ell$ be two ℓ -tuples of arbitrary points in \mathbb{C}^d , and define the $\ell \times \ell$ matrices

$$\begin{aligned} K_n^\mu(z_1, z_2, \dots, z_\ell; w_1, w_2, \dots, w_\ell) &:= (K_n^\mu(z_j, w_k))_{j,k=1}^\ell, \\ C_n^\mu(z_1, z_2, \dots, z_\ell; w_1, w_2, \dots, w_\ell) &:= (C_n^\mu(z_j, w_k))_{j,k=1}^\ell. \end{aligned}$$

If $z_1, \dots, z_n \in \mathcal{S}_n(\mu)$ are distinct and $K_n^\mu(z_1, z_2, \dots, z_\ell; z_1, z_2, \dots, z_\ell)$ is invertible (compare with Lemma 4.1 below), then this matrix also occurs in variational problems similar to the single point case. Specifically

$$\begin{aligned} &\min\{\|p\|_{2,\mu}^2 : \deg p \leq k_n^\mu, p(z_j) = c_j \text{ for } j = 1, \dots, \ell\} \\ &= c^* K_n^\mu(z_1, z_2, \dots, z_\ell; z_1, z_2, \dots, z_\ell)^{-1} c, \quad c^T = (c_1, \dots, c_\ell). \end{aligned}$$

This follows along the same lines as the proof of (14) by observing that it is sufficient to examine candidates of the form $p(z) = a_1 K_n^\mu(z, z_1) + \dots + a_\ell K_n^\mu(z, z_\ell)$ with $p(z_j) = c_j$ for $j = 1, \dots, \ell$.

We want finally to extend property (6) to the multivariate setting. As for many other asymptotic results in the literature, we have to restrict ourselves to total degree, and therefore assume that, for all $m \in \mathbb{N}$, there exists $k = \binom{m+d}{d} - 1$ such that

$$\{\alpha(0), \dots, \alpha(k)\} = \{\alpha \in \mathbb{N}^d : |\alpha| \leq m\}, \quad (16)$$

as it is, for instance, true for the graded lexicographical ordering. In the following result we will only consider Christoffel-Darboux kernels with indices

$$n =: n_{tot}(m) \text{ s.t. } k_n^\mu \leq \binom{m+d}{d} - 1 < k_{n+1}^\mu. \quad (17)$$

Also, in order to insure that we have an infinite number of orthogonal polynomials, we require that $\text{supp}(\mu)$ is not finite. Further assumptions are stated below.

Lemma 2.9. *Consider an ordering with (16), and suppose that $\text{supp}(\mu)$ is infinite. Denote by Ω be the unbounded connected component of $\mathbb{C}^d \setminus \text{supp}(\mu)$, and let $g_\Omega : \mathbb{C}^d \mapsto \mathbb{R}$ be the plurisubharmonic Green function with pole at ∞ which is strictly positive on Ω ; see the proof below. Then, for $z \in \mathcal{S}(\mu) \cap \Omega$,*

$$\liminf_{m \rightarrow \infty} K_{n_{tot}(m)}^\mu(z, z)^{1/m} \geq e^{2g_\Omega(z)}. \quad (18)$$

*If, in addition, Ω is L -regular in the sense of [21, p.186 and §5.3] and μ belongs to the multivariate generalization of the class **Reg** introduced in [22, Eqn. (1.6)], then the limit*

$$\lim_{m \rightarrow \infty} K_{n_{tot}(m)}^\mu(z, z)^{1/m} = e^{2g_\Omega(z)} \quad (19)$$

exists for all $z \in \mathcal{S}(\mu)$, and equals zero for $z \in \mathcal{S}(\mu) \setminus \Omega$.

Proof. Our proof of Lemma 2.9 is based on Siciak's function Φ in \mathbb{C}^d for the compact set $\text{supp}(\mu)$:

$$\begin{aligned} \Phi_{\text{supp}(\mu)}(z) &:= \sup \left\{ \left(\frac{|p(z)|}{\|p\|_{L^\infty(\text{supp}(\mu))}} \right)^{1/\text{tdeg}(p)} : p \neq 0 \text{ a polynomial} \right\} \\ &= \lim_{m \rightarrow \infty} \sup \left\{ \left(\frac{|p(z)|}{\|p\|_{L^\infty(\text{supp}(\mu))}} \right)^{1/m} : p \neq 0, \text{tdeg}(p) \leq m \right\}, \end{aligned}$$

together with pluripotential theory; for a summary we refer the reader to [5, 32] or [35, Appendix B] as well as the monograph [21]. (Recall that $\text{tdeg}(p)$ is the largest of the orders of monomials occurring in p .) The choice $p = 1$ shows us that $\Phi_{\text{supp}(\mu)}(z) = 1$ for $z \in \text{supp}(\mu)$.

Consider $V_{\text{supp}(\mu)} = \log \Phi_{\text{supp}(\mu)}$ and its upper semi-continuous regularization

$$V_{\text{supp}(\mu)}^*(z) = \limsup_{w \rightarrow z} V_{\text{supp}(\mu)}(w).$$

By [33, Definition B.1.8 and Theorem B.2.8], the function $V_{\text{supp}(\mu)}^*$ is plurisubharmonic whereas, in general, $V_{\text{supp}(\mu)}$ is not. We write

$$g_\Omega(z) := V_{\text{supp}(\mu)}^*(z)$$

since the latter function shares many of the properties of a univariate Green function, such as being ≥ 0 in \mathbb{C}^d and strictly positive in Ω , having a logarithmic singularity at ∞ , and $g_\Omega(z) = 0$ on $\mathbb{C}^d \setminus \Omega$ up to some pluripolar set; see the enumeration after [35, Definition B.1.8], or [21, §5.1].³

We now return to a proof of (18) and (19). The inequality

$$\|p\|_{2,\mu}^2 \leq \mu(\mathbb{C}) \|p\|_{L^\infty(\text{supp}(\mu))}^2,$$

for $p \in \mathbb{C}[z]$ allows us to relate for $z \in \mathcal{S}(\mu) \subseteq \mathcal{S}_n(\mu)$ the extremal problem (14) to Siciak's extremal problem. Observing that Siciak's extremal function is continuous at $z \in \Omega$, we obtain the lower bound $|\Phi_{\text{supp}(\mu)}(z)|^2 = \exp(2g_\Omega(z))$, as claimed in (18). For (19) we have the additional assumption of Ω being L -regular, or $V_{\text{supp}(\mu)}^* = V_{\text{supp}(\mu)}$ by [21, p.186] and thus in particular $g_\Omega(z) = 0$ for $\mathbb{C}^d \setminus \Omega$ by the maximum principle for plurisubharmonic functions.⁴ Since condition **Reg** gives us the possibility to compare the n -th roots of $\|p\|_{2,\mu}$ and $\|p\|_{L^\infty(\text{supp}(\mu))}$, the claim (19) is an immediate consequence. \square

Since for $d = 1$ the plurisubharmonic Green function becomes the classical Green function, and $n_{\text{tot}}(m) = m$, we see that Lemma 2.9 gives the (pointwise) multivariate equivalent of (6), (7), that is, we have exponential growth of the Christoffel function in $\Omega \cap \mathcal{S}(\mu)$, but (at least for sufficiently smooth μ) not in $\text{supp}(\mu)$. In particular, we should mention that (18) and (19) are wrong in general for outliers $z \in \text{supp}(\mu)$ being isolated in Ω , since the plurisubharmonic Green function remains the same after adding pluripolar sets [21, Corollary 5.2.5]. We will recall in §7.1 below explicit formulas for the plurisubharmonic Green function for a certain number of domains.

Also, we should mention that in the recent paper [6] the authors discuss Siciak functions for more general degree constraints based on convex bodies, together with its pluripotential interpretation.

2.4. Leverage scores, the Mahalanobis distance and Christoffel-Darboux kernels. The aim of this section is to recall some links between the Christoffel-Darboux kernel and several notions from statistics. For a recent related work see [30]. In this subsection we will denote elements of \mathbb{C}^d as row vectors. The mean $m(\mu)$ and positive definite covariance matrix $\mathcal{C}(\mu)$ of a random variable with values in \mathbb{C}^d with law described by some probability measure μ have the entries

$$m(\mu)_j = \int z_j d\mu(z), \quad \mathcal{C}(\mu)_{j,k} = \int \overline{(z_j - m(\mu)_j)}(z_k - m(\mu)_k) d\mu(z).$$

³In contrast to our terminology, sometimes $V_{\text{supp}(\mu)}$ is called a pluricomplex Green function, see [21, 35].

⁴Conversely, by [21, Corollary 5.1.4], if g_Ω vanishes on $\text{supp}(\mu)$ then Ω is L -regular.

The Mahalanobis distance between a point $z \in \mathbb{C}^d$ and the probability measure μ is given by

$$\Delta(z, \mu) = \sqrt{(z - m(\mu))\mathcal{C}(\mu)^{-1}(z - m(\mu))^*}.$$

The reader easily checks in terms of the moment matrices introduced in Lemma 2.3 that, for $v_d(z) = (1, z_1, \dots, z_d)$, we have that

$$V^* \widetilde{M}_d(\mu) V = \begin{bmatrix} \langle 1, 1 \rangle_{2, \mu} & 0 \\ 0 & \mathcal{C}(\mu) \end{bmatrix}, \quad V := \begin{bmatrix} 1 & -m(\mu) \\ 0 & I \end{bmatrix},$$

where we recall that $\langle 1, 1 \rangle_{2, \mu} = \mu(\mathbb{C}^d) = 1$ and hence $p_0^\mu(z) = K_0^\mu(z, z) = 1$ for all $z \in \mathbb{C}^d$. We conclude from the previous identity that, with $\mathcal{C}(\mu)$, also $\widetilde{M}_d(\mu) = M_d(\mu)$ is invertible. Then

$$\begin{aligned} K_d^\mu(z, z) &= v_d(z) M_d(\mu)^{-1} v_d(z)^* \\ &= v_d(z) V \begin{bmatrix} 1 & 0 \\ 0 & \mathcal{C}(\mu)^{-1} \end{bmatrix} V^* v_d(z)^* = 1 + \Delta(z, \mu)^2. \end{aligned} \tag{20}$$

Hence, up to some additive constant, $\sqrt{K_n^\mu(z, z)}$ can be considered as a natural generalization of the Mahalanobis distance between a point $z \in \mathbb{C}^d$ and the probability measure μ .

In the next statement we will introduce what will be referred to as our leverage score for each element of the support of a discrete measure.

Lemma 2.10. *Consider the discrete probability measure*

$$\tilde{\nu} = \sum_{j=1}^N t_j \delta_{z^{(j)}}$$

for $t_j > 0$ and distinct $z^{(j)} \in \mathbb{C}^d$. Then, for all $j = 1, 2, \dots, N$ and for all n such that $p_n^{\tilde{\nu}}$ exists, our leverage score $t_j K_n^{\tilde{\nu}}(z^{(j)}, z^{(j)})$ satisfies

$$0 \leq t_j K_n^{\tilde{\nu}}(z^{(j)}, z^{(j)}) \in (0, 1].$$

Proof. This is an immediate consequence of (14) and of the fact that $z^{(j)} \in \text{supp}(\tilde{\nu}) \subseteq \mathcal{S}(\tilde{\nu})$. \square

For a discrete probability measure $\tilde{\nu}$ as in Lemma 2.10, it is classical in data analysis to consider the data matrix $X \in \mathbb{C}^{N \times d}$ with centered and scaled rows $\sqrt{t_j} (z^{(j)} - m(\tilde{\nu}))$, and thus $\mathcal{C}(\tilde{\nu}) = X^* X$. In the literature on data analysis, the quantity

$$(X(X^* X)^{-1} X^*)_{j,j} = t_j \Delta(z^{(j)}, \tilde{\nu})^2$$

for $j = 1, \dots, N$ is known to be an element of $[0, 1)$, and referred to as a leverage score: datum $z^{(j)}$ is considered to be an outlier if $(X(X^* X)^{-1} X^*)_{j,j}$ is “close” to 1. Similar techniques have been applied for eliminating outlier

data in multiple linear regression; see for instance Cook's distance. To make the link with our leverage score for $n = d$, recall from (20) that

$$t_j \left(K_d^{\tilde{\nu}}(z^{(j)}, z^{(j)}) - 1 \right) = (X(X^*X)^{-1}X^*)_{j,j} \in [0, 1 - t_j],$$

the last inclusion following from $1 = K_0^{\tilde{\nu}}(z^{(j)}, z^{(j)}) \leq K_d^{\tilde{\nu}}(z^{(j)}, z^{(j)})$, and Lemma 2.10. As a consequence, our leverage score of Lemma 2.10 for $n = d$ can be closer to 1, and thus can be considered as slightly more precise. Choosing an index $n > d$ corresponds to considering in the regression not only random variables X_j but also their powers X^α describing mixed effects.

It remains to show that $t_j K_n^{\tilde{\nu}}(z^{(j)})$ is close to 1 (and for large n "very" close) iff $z^{(j)}$ is an outlier. Of course this necessitates giving a more precise meaning to what is called an outlier. These two questions are studied in the next sections; see, e.g., Corollary 6.4 and Corollary 6.5 for $d = 1$.

3. APPROXIMATION OF CHRISTOFFEL-DARBOUX KERNELS

The aim of the present section is to give sufficient conditions on positive measures μ, ν with compact support in \mathbb{C}^d insuring that the corresponding Christoffel-Darboux kernels $K_n^\mu(z, z)$ and $K_n^\nu(z, z)$ are close for a fixed n and for all $z \in \mathbb{C}^d$. To this aim the following matrices are essential.

Definition 3.1. For $k = k_m^\mu = k_n^\nu$, we define the matrix of mixed moments as

$$R_n(\nu, \mu) := \int v_n^\nu(z)^* v_m^\mu(z) d\nu(z) \in \mathbb{C}^{(n+1) \times (m+1)},$$

and the matrix of modified moments

$$M_m(\nu, \mu) := \int v_m^\mu(z)^* v_m^\mu(z) d\nu(z) \in \mathbb{C}^{(m+1) \times (m+1)}.$$

3.1. Comparing two measures. In this subsection we study conditions allowing one family of orthogonal polynomials to be linear combinations of the other.

Lemma 3.2. Assume that two positive measures μ, ν satisfy $k = k_m^\mu = k_n^\nu$, for some positive integers n, m . The following statements are equivalent:

- (a) $\mathcal{N}_n(\nu) \subseteq \mathcal{N}_m(\mu)$;
- (b) $\mathcal{S}_m(\mu) \subseteq \mathcal{S}_n(\nu)$;
- (c) $\{k_0^\mu, k_1^\mu, k_2^\mu, \dots, k_m^\mu\} \subseteq \{k_0^\nu, k_1^\nu, k_2^\nu, \dots, k_n^\nu\}$;
- (d) The orthogonal polynomials p_0^μ, \dots, p_m^μ can be written as a linear combination of $p_0^\nu, p_1^\nu, \dots, p_n^\nu$;
- (e) $v_m^\mu = v_n^\nu R_n(\nu, \mu)$;
- (f) There exists a constant c_k such that, for all $p \in \mathbb{C}[z]$ of degree at most k ,

$$\|p\|_{2,\mu} \leq c_k \|p\|_{2,\nu}.$$

In addition, the smallest such constant is given by $c_k = \|R_m(\mu, \nu)\|$.

Proof. We know from Lemma 2.6 that we may write the set of multivariate polynomials of degree at most $k = k_m^\mu = k_n^\nu$ as a direct sum $\mathcal{L}_m(\mu) \oplus \mathcal{N}_m(\mu) = \mathcal{L}_n(\nu) \oplus \mathcal{N}_n(\nu)$. Thus the equivalence between (a), (b), (c) and (d) follows immediately from Definition 2.4. The implication (d) \implies (e) follows by ν -orthogonality, and the implication (f) \implies (a) is trivial. Hence it only remains to show that (e) implies (f).

Suppose that (e) holds, implying $\mathcal{L}_m(\mu) \subseteq \mathcal{L}_n(\nu)$. Let $p \in \mathbb{C}[z]$ of degree at most k . Applying Lemma 2.6 we may write

$$p = q + p_3, \quad q = v_n^\nu \xi \in \mathcal{L}_n(\nu), \quad p_3 \in \mathcal{N}_n(\nu) \subseteq \mathcal{N}_m(\mu)$$

for some $\xi \in \mathbb{C}^{n+1}$. Then, applying again Lemma 2.6,

$$q = p_1 + p_2, \quad p_1 \in \mathcal{L}_m(\mu), \quad p_2 \in \mathcal{N}_m(\mu) \cap \mathcal{L}_n(\nu),$$

where

$$\begin{aligned} p_1(z) &= \int K_n^\mu(z, w) q(w) d\mu(w) \\ &= \int v_n^\mu(z) v_n^\mu(w)^* v_n^\nu(w) \xi d\mu(w) = v_n^\mu(z) R_m(\mu, \nu) \xi. \end{aligned}$$

Hence

$$\begin{aligned} \|p\|_{2,\mu} &= \|p_1\|_{2,\mu} = \|R_m(\mu, \nu) \xi\| \leq \|R_m(\mu, \nu)\| \|\xi\| \\ &= \|R_m(\mu, \nu)\| \|q\|_{2,\nu} = \|R_m(\mu, \nu)\| \|p\|_{2,\nu}, \end{aligned}$$

showing that (e) holds with $c_k = \|R_m(\mu, \nu)\|$. Taking the supremum for $\xi \in \mathbb{C}^{n+1} \setminus \{0\}$ shows that c_k cannot be smaller. \square

Under the assumptions of Lemma 3.2(e), we can make a link between $R_n(\nu, \mu)$ and the matrices defined in Lemma 2.3, namely

$$R_n(\nu, \mu) = R_n(\nu)^{-1} E_k(\nu)^T E_k(\mu) R_m(\mu). \quad (21)$$

To see this, recall from (10) that

$$v_m^\mu = v_k E_k(\mu) R_m(\mu), \quad v_n^\nu = v_k E_k(\nu) R_n(\nu),$$

where Lemma 3.2(c) tells us that $E_k(\mu) = E_k(\nu) E_k(\nu)^T E_k(\mu)$, and thus

$$v_m^\mu = v_n^\nu R_n(\nu)^{-1} E_k(\nu)^T E_k(\mu) R_m(\mu) = v_n^\nu R_n(\nu, \mu),$$

implying (21). In this case, it is not difficult to check that $R_n(\nu, \mu)$ is of full column rank $m + 1$ and of upper echelon form, and that we have full rank decomposition

$$M_m(\nu, \mu) = R_n(\nu, \mu)^* R_n(\nu, \mu). \quad (22)$$

In the remainder of this section we will always suppose that $k = k_m^\mu = k_n^\nu$ with $\mathcal{N}_n(\nu) = \mathcal{N}_m(\mu)$. Thus we may apply Lemma 3.2 twice, implying that $m = n$ and $k_j^\mu = k_j^\nu$ for $j = 0, 1, \dots, n$ by Lemma 3.2(c). In particular, we find that $E_k(\nu)^T E_k(\mu)$ is the identity of order $n + 1$. We conclude using (21) that $R_n(\nu, \mu)$ is upper triangular and invertible, with inverse given by

$$R_n(\nu, \mu)^{-1} = R_n(\mu)^{-1} R_n(\nu) = R_n(\mu, \nu). \quad (23)$$

We summarize these findings and their consequence for the Christoffel-Darboux kernel in the following statement.

Corollary 3.3. *Let $n \geq 0$ be some integer⁵ with $k_n^\mu = k_n^\nu$ and $\mathcal{N}_n(\mu) = \mathcal{N}_n(\nu)$. Then*

$$v_n^\mu(z) = v_n^\nu(z)R_n(\nu, \mu), \quad v_n^\nu(z) = v_n^\mu(z)R_n(\mu, \nu), \quad (24)$$

with the upper triangular and invertible matrices $R_n(\nu, \mu) = R_n(\nu)^{-1}R_n(\mu)$ and $R_n(\mu, \nu) = R_n(\nu, \mu)^{-1}$, allowing for the Cholesky decompositions

$$M_n(\nu, \mu) = R_n(\nu, \mu)^*R_n(\nu, \mu), \quad M_n(\mu, \nu) = R_n(\mu, \nu)^*R_n(\mu, \nu). \quad (25)$$

In particular,

$$K_n^\nu(z, w) = v_n^\mu(z)M_n(\nu, \mu)^{-1}v_n^\mu(w)^*. \quad (26)$$

Examples 3.4. (a) Let $\mu = \mu_0$ be the normalized Lebesgue measure on $(\partial\mathbb{D})^d$ making monomials orthonormal. Here $\mathcal{N}(\mu_0) = \{0\}$, and thus $k_n^{\mu_0} = n$ for all n . As a consequence, Lemma 3.2(d),(e) reduces to Lemma 2.3.

(b) Consider the case $\nu = \mu + \sigma$ for some positive measure σ with compact support; for instance, the case of adding a finite number of point masses. Since $\nu - \mu \geq 0$, clearly $\mathcal{N}(\nu) \subseteq \mathcal{N}(\mu)$, and from Lemma 3.2 we know that we may express the orthogonal polynomials p_j^μ in terms of the p_j^ν . Moreover, $\|M_n(\mu, \nu)\| = \|R_n(\mu, \nu)\|^2 \leq 1$ by Lemma 3.2(f). If, in addition, $\text{supp}(\sigma) \subseteq \mathcal{S}(\mu)$, then $\mathcal{N}(\mu) = \mathcal{N}(\mu + \sigma)$ by Example 2.5. Thus, in this case, we may also express the orthogonal polynomials p_j^ν in terms of the p_j^μ via (24), $k_j^\mu = k_j^\nu$ for all $j \geq 0$, and, by (25) and (26),

$$\forall z \in \mathbb{C}^d, \forall n \geq 0: \quad K_n^{\mu+\sigma}(z, z) \leq K_n^\mu(z, z).$$

Remarks 3.5. Lasserre and Pauwels in [20, Theorem 3.8 and Theorem 3.10] consider the graded lexicographical ordering (16) and give an explicit sequence of numbers κ_m such that the Hausdorff distance $d_H(S, S_m)$ between the sets

$$S = \text{supp}(\mu), \quad \text{and} \quad S_m = \{x \in \mathbb{R}^n : K_{n_{\text{tot}}(m)}^\mu(z, z) \leq \kappa_m\}$$

tends to 0. For this they assume that $S = \text{Clos}(\text{Int}(S))$, and that there is some constant $w_{\min} > 0$ such that $d\mu \geq w_{\min}dA|_S$, where dA denotes (unnormalized) Lebesgue measure in \mathbb{R}^d .

We summarize their reasoning, by using our notation. All measures ν considered in this context are supported in \mathbb{R}^d , with $\text{supp}(\nu)$ having non-empty interior (in \mathbb{R}^d), such that $\mathcal{N}(\nu) = \{0\}$ and $\mathcal{S}(\nu) = \mathbb{C}^d$ by Example 2.5. Notice that $d_H(S, S_m) \leq \delta$ if

$$\forall z \in \text{supp}(\mu) \text{ with } \text{dist}(z, \partial \text{supp}(\mu)) \geq \delta: \quad K_{n_{\text{tot}}(m)}^\mu(z, z) \leq \kappa_m, \quad (27)$$

$$\forall z \notin \text{supp}(\mu) \text{ with } \text{dist}(z, \partial \text{supp}(\mu)) \geq \delta: \quad K_{n_{\text{tot}}(m)}^\mu(z, z) \geq \kappa_m. \quad (28)$$

⁵The interested reader might want to check that this condition is true for all n if $\mathcal{N}(\mu) = \mathcal{N}(\nu)$.

For z as in (27) we denote by B the real unit ball in \mathbb{R}^n , and observe that $z + \delta B \subseteq S$. Hence $w_{\min} A|_{z+\delta B} \leq w_{\min} A|_S \leq \mu$. Using the fact that that an explicit upper bound κ'_m is known for $K_{n_{\text{tot}}(m)}^{A|_B}(0, 0)$ since the 90-ies, we obtain by Example 3.4(b) and (11)

$$K_{n_{\text{tot}}(m)}^\mu(z, z) \leq \frac{K_{n_{\text{tot}}(m)}^{A|_{z+\delta B}}(z, z)}{w_{\min}} = \frac{A(B)}{A(z + \delta B)} \frac{K_{n_{\text{tot}}(m)}^{A|_B}(0, 0)}{w_{\min}} = \frac{\kappa'_m}{\delta^d w_{\min}},$$

compare with [20, Lemma 6.1]. With z as in (28), they consider the annulus $S' = \{x \in \mathbb{R}^d : \delta \leq \|x\| \leq r\}$, where $r = \delta + \text{diam}(S)$ such that $S \subseteq S'$. Following [22, 23], they then consider the real *needle polynomial*

$$p(x) = T_{\lfloor m/2 \rfloor} \left(\frac{r^2 + \delta^2 - 2\|x - z\|^2}{r^2 - \delta^2} \right), \quad \text{with} \quad \|p\|_{L^\infty(S)} \leq \|p\|_{L^\infty(S')} = 1,$$

$\text{tdeg } p \leq m$ and thus $\text{deg } p \leq n_{\text{tot}}(m)$, and

$$|p(z)| = T_{\lfloor m/2 \rfloor} \left(\frac{r^2 + \delta^2}{r^2 - \delta^2} \right) \geq \frac{1}{2} e^{\lfloor m/2 \rfloor g_{\mathbb{C} \setminus [-1, 1]} \left(\frac{r^2 + \delta^2}{r^2 - \delta^2} \right)} \geq \frac{1}{2} \left(\frac{r + \delta}{r - \delta} \right)^{(m-1)/2};$$

compare with [20, Lemma 6.3]. It follows from (14) that

$$K_{n_{\text{tot}}(m)}^\mu(z, z) \geq \frac{|p(z)|^2}{\|p\|_{2, \mu}^2} \geq \frac{1}{2\mu(\mathbb{C})} \left(1 + \frac{2\delta}{\text{diam}(S)} \right)^{m-1}.$$

In view of (27), (28), it thus only remains to find δ as a function of m (possibly decreasing) and κ_m such that

$$\frac{\kappa'_m}{\delta_m^d w_{\min}} \leq \kappa_m \leq \frac{1}{2\mu(\mathbb{C})} \left(1 + \frac{2\delta_m}{\text{diam}(S)} \right)^{m-1}.$$

For instance, for $\delta_m = 1/m^\alpha$ with $\alpha \in (0, 1)$, the left-hand term grows like a power of m , and the right-hand side exponentially in m , so that the above inequalities are true for sufficiently large m . Enclosed we list some possible improvements.

- (1) The assumption $S = \text{Clos}(\text{Int}(S))$ for $S = \text{supp}(\mu)$ is strongly used, which does not allow outliers.
- (2) The assumption $S = \text{Clos}(\text{Int}(S))$ is strongly used also to give upper bounds for $K_{n_{\text{tot}}(m)}^\mu(z, z)$ in $\text{supp}(\mu)$ at some distance from the boundary. This condition could be replaced by stepping to the relative interior of $\text{supp}(\mu)$ in $\mathcal{S}(\mu) \cap \mathbb{R}^d$ (in the real setting) or $\mathcal{S}(\mu)$ in the complex setting. This may possibly result in an increase in the power of m .
- (3) The lower bound for $K_{n_{\text{tot}}(m)}^\mu(z, z)$ is very rough, and it should be possible to write something down in terms of the plurisubharmonic Green function of $\text{supp}(\mu)$. One difficulty arises in this direction: one needs bounds uniformly in z for all $z \in \text{supp}(\mu)$ with distance to $\text{supp}(\mu)$ decaying like a fractional negative power of m .

Remarks 3.6. Modified moment matrices and mixed moment matrices have been shown to be useful for explicitly computing orthogonal polynomials of one or several complex variables; see for instance the modified Chebyshev algorithm for orthogonality on the real line [15]. Here the basic idea is that one knows μ and its orthogonal polynomials p_j^μ , and computes both $R_n(\mu, \nu)$ and $v_n^\nu(z) = v_n^\mu(z)R_n(\mu, \nu)$ through ν -orthogonality, using, e.g., the Gram-Schmidt method. In case of finite precision arithmetic, we can only insure small errors if $R_n(\mu, \nu)$ has a modest condition number given by $\|R_n(\mu, \nu)\| \|R_n(\mu, \nu)^{-1}\|$. In case of one complex variable it has been shown in [3, Lemma 3.4] that this condition number grows exponentially in n unless the complements of $\text{supp}(\mu)$ and $\text{supp}(\nu)$ have the same unbounded connected component, and, according to Lemma 2.9, a similar result is expected to be true in the case of several complex variables. Thus the choice of μ is essential.

In particular, it is not always a good idea to compute p_0^ν, p_1^ν, \dots from monomials, even after scaling (11); compare with Example 3.4(a). But the modified moment matrix $M_n(\mu, \nu)$ and hence $R_n(\mu, \nu)$ might be much better conditioned for μ close to ν , or, in other words, the p_j^μ are “nearly” ν -orthogonal for $j = 0, 1, \dots, n$.

3.2. Small perturbations. In this subsection we compare for fixed n the Christoffel-Darboux kernels for two measures μ and ν that are close in a precise sense. In view of Lemma 3.2(f), the proximity of the two measures is imposed by the following condition: there exists a sufficiently small $\epsilon \in (0, 1)$ such that

$$(1 - \epsilon) \|p\|_{2,\mu}^2 \leq \|p\|_{2,\nu}^2 \leq (1 + \epsilon) \|p\|_{2,\mu}^2 \quad (29)$$

for all polynomials p of degree at most k_n^μ . As a consequence of (29), the quantity $\|p\|_{2,\mu}^2$ vanishes if and only if this is true for $\|p\|_{2,\nu}^2$; that is, $k_n^\mu = k_n^\nu$ and $\mathcal{N}_n(\mu) = \mathcal{N}_n(\nu)$. This allows us to apply Corollary 3.3, in particular $v_n^\mu = (p_0^\mu, \dots, p_n^\mu) = v_n^\nu R_n(\nu, \mu) = (p_0^\nu, \dots, p_n^\nu) R_n(\nu, \mu)$ with the Hermitian and positive definite modified moment matrix $M_n(\nu, \mu) = R_n(\nu, \mu)^* R_n(\nu, \mu)$ being similar to $M_n(\mu, \nu)^{-1}$.

This modified moment matrix allows us to restate assumption (29): for any $\xi \in \mathbb{C}^{n+1}$, by considering $p(z) = v_n^\mu(z)\xi = v_n^\nu(z)R_n(\nu, \mu)\xi$ in (29), we find

$$(1 - \epsilon) \|\xi\|^2 \leq \|R_n(\nu, \mu)\xi\|^2 = \xi^* M_n(\nu, \mu)\xi \leq (1 + \epsilon) \|\xi\|^2; \quad (30)$$

in other words, ν is so close to μ that all eigenvalues of the Hermitian matrix $M_n(\nu, \mu) - I$ lie in the interval $[-\epsilon, \epsilon]$, or $\|M_n(\nu, \mu) - I\| \leq \epsilon$. Using the Froebenius matrix norm, we therefore get the sufficient condition

$$\sum_{j,k=0}^n \left| \langle p_j^\mu, p_k^\mu \rangle_{2,\nu} - \delta_{j,k} \right|^2 = \|M_n(\nu, \mu) - I\|_F^2 \leq \epsilon^2 \quad (31)$$

which indicates how p_0^μ, \dots, p_n^μ are “nearly” ν -orthogonal.

For two Hermitian matrices X_1, X_2 we write that $X_1 \leq X_2$ if $X_2 - X_1$ is positive semi-definite. We make use of this notation in the proofs of the following results.

Proposition 3.7. *Under the assumption (29) there holds for all $z, w \in \mathbb{C}^d$*

$$(1 - \epsilon) K_n^\nu(z, z) \leq K_n^\mu(z, z) \leq (1 + \epsilon) K_n^\nu(z, z), \quad (32)$$

$$|K_n^\mu(z, w) - K_n^\nu(z, w)| \leq \epsilon \sqrt{K_n^\nu(z, z)} \sqrt{K_n^\nu(w, w)}, \quad (33)$$

$$|C_n^\mu(z, w) - C_n^\nu(z, w)| \leq 2\epsilon. \quad (34)$$

Proof. Recall from (26) that

$$K_n^\mu(z, w) = v_n^\nu(z) M_n(\mu, \nu)^{-1} v_n^\nu(w)^*.$$

Since $M_n(\mu, \nu)^{-1}$ is similar to $M_n(\nu, \mu)$, we get from (30) that

$$(1 - \epsilon)I \leq M_n(\mu, \nu)^{-1} \leq (1 + \epsilon)I, \quad (35)$$

and a combination with (26) for $w = z$ gives (32). Moreover, it follows from (35) that $\|M_n(\mu, \nu)^{-1} - I\| \leq \epsilon$, and thus, again by (26),

$$|K_n^\mu(z, w) - K_n^\nu(z, w)| \leq \epsilon \|v_n^\nu(z)\| \|v_n^\nu(w)\|,$$

implying (33). Finally, from (32) we infer

$$\frac{K_n^\mu(z, z) K_n^\mu(w, w)}{K_n^\nu(z, z) K_n^\nu(w, w)} \in \left[(1 - \epsilon)^2, (1 + \epsilon)^2 \right]$$

and thus, by definition (15) and (33),

$$\begin{aligned} & |C_n^\mu(z, w) - C_n^\nu(z, w)| \\ & \leq \left| \frac{K_n^\mu(z, w) - K_n^\nu(z, w)}{\sqrt{K_n^\nu(z, z) K_n^\nu(w, w)}} \right| + |C_n^\mu(z, w)| \left| 1 - \sqrt{\frac{K_n^\mu(z, z) K_n^\mu(w, w)}{K_n^\nu(z, z) K_n^\nu(w, w)}} \right| \\ & \leq \epsilon + |C_n^\mu(z, w)| \epsilon \leq 2\epsilon, \end{aligned}$$

as claimed in (34). \square

Remark 3.8. If assumption (29) holds for a pair of measures (μ, ν) , then it also holds for $(\mu + \sigma, \nu + \sigma)$ for any measure $\sigma \geq 0$ supported on a subset of $\mathcal{S}(\mu)$, and hence

$$(1 - \epsilon) K_n^{\nu+\sigma}(z, z) \leq K_n^{\mu+\sigma}(z, z) \leq (1 + \epsilon) K_n^{\nu+\sigma}(z, z)$$

by Proposition 3.7. Indeed, from Example 3.4(b) we know that $k_n^\mu = k_n^{\mu+\sigma}$, and so, for any polynomial p , of degree at most $k_n^{\mu+\sigma}$

$$\| \|p\|_{2, \mu+\sigma}^2 - \|p\|_{2, \nu+\sigma}^2 \| = \| \|p\|_{2, \mu}^2 - \|p\|_{2, \nu}^2 \| \leq \epsilon \|p\|_{2, \mu}^2 \leq \epsilon \|p\|_{2, \mu+\sigma}^2,$$

which implies assumption (29) for $(\mu + \sigma, \nu + \sigma)$.

Remark 3.9. Multi-point analogues of Proposition 3.7 are also available. We include them for completeness. Suppose that assumption (29) holds, and let $z_1, \dots, z_\ell \in \mathbb{C}^d$. Then

$$\begin{aligned} (1 - \epsilon) K_n^\nu(z_1, \dots, z_\ell; z_1, \dots, z_\ell) &\leq K_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell) \\ &\leq (1 + \epsilon) K_n^\nu(z_1, \dots, z_\ell; z_1, \dots, z_\ell), \end{aligned}$$

and, for the Froebenius matrix norm $\|\cdot\|_F$,

$$\|C_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell) - C_n^\nu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)\|_F \leq 2\sqrt{\ell(\ell - 1)}\epsilon.$$

For a proof, consider the matrix

$$V_n^\nu(z_1, \dots, z_\ell) := \begin{bmatrix} v_n^\nu(z_1) \\ \vdots \\ v_n^\nu(z_\ell) \end{bmatrix}.$$

From the factorization

$$K_n^\nu(z_1, \dots, z_\ell; z_1, \dots, z_\ell) = V_n^\nu(z_1, \dots, z_\ell) V_n^\nu(z_1, \dots, z_\ell)^*,$$

the identity for the polarized kernel follows by multiplying (35) on the left by $\xi^* V_n^\nu(z_1, \dots, z_\ell)$, and on the right by its adjoint. For the cosine identity, we apply (34) and obtain

$$\begin{aligned} &\|C_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell) - C_n^\nu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)\|_F^2 \\ &= \sum_{j,k=1, j \neq k}^{\ell} |C_n^\mu(z_j, z_k) - C_n^\nu(z_j, z_k)|^2 \leq 4\ell(\ell - 1)\epsilon^2, \end{aligned}$$

as required for the above claim.

Remarks 3.10. In a series of papers, Migliorati and his co-authors considered a fixed general probability measure μ compactly supported in \mathbb{R}^d (probably their results still remain valid in \mathbb{C}^d) and the discrete measure

$$\nu = \frac{1}{N} \sum_{j=1}^N w(z^{(j)}) \delta_{z^{(j)}},$$

with the N masspoints $z^{(j)} \in \text{supp}(\mu)$ being independent and identically distributed random variables with law given by some sampling probability measure $\tilde{\mu}$ with $\text{supp}(\tilde{\mu}) = \text{supp}(\mu)$ (e.g., $\mu = \tilde{\mu}$), and the density $w(z) = \frac{d\mu}{d\tilde{\mu}}(z)$ is assumed to be strictly positive in $\text{supp}(\mu)$. We now refer to [8] where, as in the present paper, the authors allow for general degree sequences and general measures μ . In particular, in [8, Theorem 2.1(i)], the authors choose N large enough such that

$$\max_{z \in \text{supp}(\mu)} w(z) K_n^\mu(z, z) \leq \frac{1 - \log 2}{2 + 2r} \frac{N}{\log(N)}$$

for some $r > 0$, and show that the probability that $\|M_n(\nu, \mu) - I\| > 1/2$ is less than $2N^{-r}$. In other words, condition (30) holds for $\epsilon = 1/2$ with a

probability $1 - 2N^{-r}$ close to 1. By checking the proof, it can be seen that a similar result is true for any $\epsilon \in (0, 1/2)$.

4. ADDITIVE PERTURBATION OF POSITIVE MEASURES

The aim of this section is to give upper and lower bounds for the ratio of Christoffel-Darboux kernels $K_n^{\mu+\sigma}(z, z)/K_n^\mu(z, z)$ for $z \notin \text{supp}(\mu)$, but possibly $z \in \text{supp}(\sigma)$. Several of the examples studied in later sections require that the cosine function $C_n^\mu(z, w)$ (see (15)) has a limit, at least along subsequences, for $z, w \notin \text{supp}(\mu)$. Hence our bounds will be formulated in terms of this cosine.

Let $\text{supp}(\mu)$ be an infinite set such that there exists an infinite number of orthogonal polynomials p_j^μ , and consider the case of adding ℓ disjoint point masses in the Zariski closure $\mathcal{S}(\mu)$ of the support of the original measure μ (and later outside $\text{supp}(\mu)$). That is,

$$\sigma = \sum_{j=1}^{\ell} t_j \delta_{z_j}, \quad t_j > 0, \quad z_1, \dots, z_\ell \in \mathcal{S}(\mu) \text{ disjoint.} \quad (36)$$

As in Remark 3.9, we depart from the canonical notation and allow z_1, \dots, z_ℓ to be elements in \mathbb{C}^d rather than the numerical coordinates of a single point z . From Example 3.4(b) we know that $\mathcal{N}(\mu) = \mathcal{N}(\mu + \sigma)$, and thus we may apply Corollary 3.3 and in particular property (26) for $\nu = \mu + \sigma$.

Lemma 4.1. *If $z_1, \dots, z_\ell \in \mathcal{S}(\mu)$ are distinct, then there exists an N such that the matrix $K_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)$ is invertible for all $n \geq N$.*

Proof. There exist multivariate Lagrange polynomials p_1, \dots, p_ℓ , each of them of minimal degree, such that $p_j(z_k) = \delta_{j,k}$ for $k = 1, \dots, \ell$. The relation $z_1, \dots, z_\ell \in \mathcal{S}(\mu)$ implies that for each $p \in \mathcal{N}(\mu)$ there holds $p(z_j) = 0$ for $j = 1, 2, \dots, \ell$. For each p_j there exists $q_j \in \text{span}\{p_0^\mu, p_1^\mu, \dots\}$ with $\deg q_j \leq \deg p_j$ and $p_j - q_j \in \mathcal{N}(\mu)$. Then $q_j(z_k) = p_j(z_k) = \delta_{j,k}$, and thus also q_1, \dots, q_ℓ are Lagrange polynomials. In particular, $\deg q_j = \deg p_j$ by minimality of $\deg p_j$. We conclude that there exists an integer N with $k_N^\mu = \max\{\deg q_j^\mu : j = 1, \dots, \ell\}$, and the rows $\{v_n^\mu(z_j) : j = 1, 2, \dots, \ell\}$ are linearly independent for $n \geq N$. This is false for $n < N$ by minimality of the degrees. Thus $K_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)$ is invertible for all $n \geq N$, but not for $n < N$. \square

Notice that Lemma 4.1 is trivial for the case $d = 1$ of univariate polynomials, where $z^{\alpha(j)} = z^j$ and hence $N = \ell - 1$.

Recalling Definition 2.8 we see that the matrices $K_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)$ and $C_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)$ are invertible for distinct points z_j . This enables us to establish the following comparison theorem consisting of a sequence of alternating inequalities.

Theorem 4.2. *Let $z_1, \dots, z_\ell \in \mathcal{S}(\mu)$ be distinct, $n \geq N$ as in Lemma 4.1, and consider the following matrices (depending on n)*

$$\begin{aligned} C &:= C_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell), \\ \tilde{C} &:= C_n^\mu(z_1, \dots, z_\ell, z; z_1, \dots, z_\ell, z) = \begin{bmatrix} C & b^* \\ b & 1 \end{bmatrix}, \quad b \in \mathbb{C}^{1 \times \ell}, \\ D &:= \text{diag} \left(\frac{1}{\sqrt{t_j K_n^\mu(z_j, z_j)}} \right)_{j=1, \dots, \ell}, \end{aligned}$$

and the constants

$$\Sigma_m := 1 - \sum_{j=0}^{m-1} (-1)^j b C^{-1} (D^2 C^{-1})^j b^*, \quad m = 1, 2, \dots$$

Then, for all $z \in \mathbb{C}^d$,

$$\frac{K_n^{\mu+\sigma}(z, z)}{K_n^\mu(z, z)} = 1 - b(D^2 + C)^{-1} b^*, \quad (37)$$

and

$$\Sigma_1 \leq \Sigma_3 \leq \Sigma_5 \leq \dots \leq \frac{K_n^{\mu+\sigma}(z, z)}{K_n^\mu(z, z)} \leq \dots \leq \Sigma_4 \leq \Sigma_2 \leq \Sigma_0. \quad (38)$$

Before presenting the proof of this theorem, we make some important observations and provide two examples where the result is applied.

Notice that, in the case of outlier detection, z_j lies in $\Omega \cap \mathcal{S}(\mu)$, with Ω the unbounded connected component of $\mathbb{C} \setminus \text{supp}(\mu)$. Hence (6) in case $d = 1$ and (18) in case $d > 1$ tell us that $\|D\| \rightarrow 0$ exponentially fast for $n \rightarrow \infty$. Also, $\|b\| \leq \sqrt{\ell}$ since $|C_n^\mu(z, z_j)| \leq 1$. Hence, as long as C^{-1} is bounded uniformly for n sufficiently large, we see that (38) provides quite sharp lower and upper bounds, even for modest values of m . This assumption on C is verified in many special cases discussed below where we show that (a unitary scaled counterpart of) C has a finite limit as $n \rightarrow \infty$.

In the next two examples, we make this last assertion a bit more explicit.

Example 4.3. We begin with the simple special case $\sigma = t_1 \delta_{z_1}$ of adding one point mass. Here, with the notation of Theorem 4.2,

$$C = 1, \quad b = C_n^\mu(z, z_1), \quad D^2 = \frac{1}{t_1 K_n^\mu(z_1, z_1)}.$$

With these data we find from (37)

$$\frac{K_n^{\mu+\sigma}(z, z)}{K_n^\mu(z, z)} = 1 - b(1 + D^2)^{-1} b^* = 1 - \frac{|C_n^\mu(z, z_1)|^2}{1 + \frac{1}{t_1 K_n^\mu(z_1, z_1)}},$$

and, from (38) for $m = 2$,

$$\begin{aligned} & -\frac{1}{(t_1 K_n^\mu(z_1, z_1))^2} \leq \Sigma_3 - \Sigma_2 \leq \frac{K_n^{\mu+\sigma}(z, z)}{K_n^\mu(z, z)} - \Sigma_2 \\ & = \frac{K_n^{\mu+\sigma}(z, z)}{K_n^\mu(z, z)} - 1 + |C_n^\mu(z, z_1)|^2 \left(1 - \frac{1}{t_1 K_n^\mu(z_1, z_1)}\right) \leq 0. \end{aligned}$$

□

Example 4.4. For σ consisting of $\ell \geq 2$ point masses as in (36), Theorem 4.2 provides the following useful estimates when taking $m \in \{0, 1, 2\}$. From (38) for $m = 0, 1$ we observe that

$$\Sigma_1 = \frac{\det C_n^\mu(z_1, \dots, z_\ell, z; z_1, \dots, z_\ell, z)}{\det C_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)} \leq \frac{K_n^{\mu+\sigma}(z, z)}{K_n^\mu(z, z)} \leq \Sigma_0 = 1. \quad (39)$$

where, in the equality on the left, we have used Schur complement techniques. Since Σ_1 vanishes when z is one of the point masses at z_j , we go one term further:

$$-\left(\Sigma_2 - \Sigma_1\right) \leq \frac{K_n^{\mu+\sigma}(z, z)}{K_n^\mu(z, z)} - \Sigma_2 \leq 0, \quad (40)$$

with

$$\begin{aligned} \Sigma_2 - \Sigma_1 & = bC^{-1}D^2C^{-1}b^* = \sum_{j=1}^{\ell} \frac{|bC^{-1}e_j|^2}{t_j K_n^\mu(z_j, z_j)} \\ & = \sum_{j=1}^{\ell} \frac{|\det C_n^\mu(z_1, \dots, z_{j-1}z, z_{j+1}, \dots, z_\ell; z_1, \dots, z_\ell)|^2}{|\det C_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)|^2} \frac{1}{t_j K_n^\mu(z_j, z_j)}, \end{aligned} \quad (41)$$

where in the last equality we have applied Cramer's rule. □

Proof of Theorem 4.2. We start by recalling that, by (26) for $\nu = \mu + \sigma$,

$$K_n^{\mu+\sigma}(z, z) = v_n^\mu(z)M_n(\mu + \sigma, \mu)^{-1}v_n^\mu(z)^*.$$

By introducing the matrices

$$V := \begin{bmatrix} v_n^\mu(z_1)/\sqrt{K_n^\mu(z_1, z_1)} \\ \vdots \\ v_n^\mu(z_\ell)/\sqrt{K_n^\mu(z_\ell, z_\ell)} \end{bmatrix} \in \mathbb{C}^{\ell \times (n+1)},$$

and D, C, b as in the statement of Theorem 4.2, it is not difficult to check that

$$M_n(\mu + \sigma, \mu) = I + V^*D^{-2}V,$$

and thus, by the Sherman-Morrison formula [16, §2.1.4],

$$M_n(\mu + \sigma, \mu)^{-1} = I - V^*D^{-1}(I + D^{-1}VV^*D^{-1})^{-1}D^{-1}V.$$

Observing that

$$C = VV^*, \quad b = \frac{v_n^\mu(z)V_n^*}{\sqrt{K_n^\mu(z, z)}},$$

we obtain (37).

If we neglect D , which usually is assumed to have small entries, we obtain in (37) that the right-hand side is the Schur complement $\Sigma_1 = \tilde{C}/C$. For $D \neq 0$ we can give lower and upper bounds: by assumption, C is Hermitian positive definite, and hence has a square root $C^{1/2}$ with inverse $C^{-1/2}$. Introducing the Hermitian positive definite matrix $E = C^{-1/2}D^2C^{-1/2}$, we have for all integers $m \geq 0$ that

$$\begin{aligned} (-1)^m \left((I + E)^{-1} - \sum_{j=0}^{m-1} (-1)^j E^j \right) &= E^m (I + E)^{-1} \\ &= ((E^{1/2})^m)^* (I + E)^{-1} (E^{1/2})^m \geq 0, \end{aligned}$$

where again we write $A \leq B$ for two Hermitian matrices if $B - A$ is positive definite. Substituting gives

$$(-1)^m \left((D + C)^{-1} - \sum_{j=0}^{m-1} C^{-1} (D^2 C^{-1})^j \right) \geq 0,$$

which together with (37) implies (38). \square

5. COSINE ASYMPTOTICS IN THE UNIVARIATE CASE

In this section we consider orthogonal polynomials of a single complex variable $z \in \mathbb{C}$. For obtaining asymptotics it is natural to assume throughout the section that $\text{supp}(\mu)$ is infinite. Thus we have the nullspace $\mathcal{N}(\mu) = 0$, $k_n^\mu = n$, and $\mathcal{S}(\mu) = \mathbb{C}$, compare with §2.2.

In the case of a single complex variable, we want to show that ratio asymptotics is sufficient to derive asymptotics of the cosine $C_n^\mu(z, w)$ and thus of $K_n^{\mu+\sigma}(z, z)/K_n^\mu(z, z)$. Our main result is the following.

Theorem 5.1. *Let Ω denote a subdomain of the unbounded component of $\mathbb{C} \setminus \text{supp}(\mu)$, and suppose that there is a function g analytic and different from zero in Ω such that*

$$\lim_{n \rightarrow \infty} \frac{p_n^\mu(z)}{p_{n+1}^\mu(z)} = g(z) \quad (42)$$

uniformly on any compact subset of Ω . Let F be a compact subset of Ω . Then $1/K_n^\mu(z, z) \rightarrow 0$ uniformly for $z \in F$, and

$$\lim_{n \rightarrow \infty} C_n^\mu(z, w) \frac{|p_n^\mu(z)|}{|p_n^\mu(z)|} \frac{|p_n^\mu(w)|}{|p_n^\mu(w)|} = \frac{\sqrt{(1 - |g(z)|^2)(1 - |g(w)|^2)}}{1 - g(z)g(w)} \quad (43)$$

uniformly for $z, w \in F$.

Furthermore, let $z_1, \dots, z_\ell, w_1, \dots, w_\ell \in \Omega$ with distinct $g(z_1), \dots, g(z_\ell)$ and distinct $g(w_1), \dots, g(w_\ell)$. Then, uniformly for $z, w \in F$, there holds

$$\begin{aligned} & \lim_{n \rightarrow \infty} \left| \frac{\det C_n^\mu(z_1, \dots, z_\ell, z; w_1, \dots, w_\ell, w)}{\det C_n^\mu(z_1, \dots, z_\ell; w_1, \dots, w_\ell)} \right| \\ &= \frac{\sqrt{(1 - |g(z)|^2)(1 - |g(w)|^2)}}{|1 - g(z)\overline{g(w)}|} \prod_{j=1}^{\ell} \left| \frac{g(w) - g(w_j)}{1 - g(w_j)\overline{g(w)}} \frac{g(z) - g(z_j)}{1 - g(z_j)\overline{g(z)}} \right|. \end{aligned} \quad (44)$$

Before giving the proof of this theorem we mention a sufficient condition for which the hypotheses hold (see [34, Proposition 3.4]) .

Lemma 5.2. *If there exists a function $G(z)$ analytic and non-zero at infinity such that*

$$\lim_{n \rightarrow \infty} \frac{z p_n^\mu(z)}{p_{n+1}^\mu(z)} = G(z),$$

uniformly in some neighborhood of infinity, then

$$\lim_{n \rightarrow \infty} \frac{p_n^\mu(z)}{p_{n+1}^\mu(z)} = g(z) := \frac{G(z)}{z}, \quad (45)$$

uniformly on every closed subset of $\Omega := \mathbb{C} \setminus \text{Co}(\text{supp}(\mu))$, where $\text{Co}(S)$ denotes the convex hull of the set S . Moreover, $0 < |g(z)| < 1$ for $z \in \Omega$.

Remark 5.3. While it is possible for the limit (45) to hold everywhere in $\mathbb{C} \setminus \text{supp}(\mu)$, it cannot hold locally uniformly with a limit of modulus less than one in any domain containing a point of $\text{supp}(\mu)$.

Proof of Theorem 5.1. We may rewrite condition (42) in the following form which will be used in what follows in the proof: for all $\epsilon > 0$ there exists an $N = N(\epsilon, F)$ such that, for all $n \geq N - 1$ and $z \in F$,

$$|p_n^\mu(z) - g(z)p_{n+1}^\mu(z)| \leq \epsilon \sqrt{|p_n^\mu(z)|^2 + |p_{n+1}^\mu(z)|^2}. \quad (46)$$

Thus, for $z, w \in F$ and $n \geq N - 1$

$$\begin{aligned} & |p_n^\mu(z)\overline{p_n^\mu(w)} - g(z)p_{n+1}^\mu(z)\overline{g(w)p_{n+1}^\mu(w)}| \\ & \leq |(p_n^\mu(z) - g(z)p_{n+1}^\mu(z))\overline{p_n^\mu(w)}| \\ & \quad + |\overline{(p_n^\mu(w) - g(w)p_{n+1}^\mu(w))}(p_n^\mu(z) - g(z)p_{n+1}^\mu(z))| \\ & \quad + |\overline{(p_n^\mu(w) - g(w)p_{n+1}^\mu(w))}p_n^\mu(z)| \\ & \leq (2\epsilon + \epsilon^2) \sqrt{|p_n^\mu(z)|^2 + |p_{n+1}^\mu(z)|^2} \sqrt{|p_n^\mu(w)|^2 + |p_{n+1}^\mu(w)|^2}. \end{aligned} \quad (47)$$

For $z = w$ we deduce for sufficiently small ϵ that

$$\frac{|g(z)|^2 - 2\epsilon - \epsilon^2}{1 + 2\epsilon + \epsilon^2} |p_{n+1}^\mu(z)|^2 \leq |p_n^\mu(z)|^2 \leq \frac{|g(z)|^2 + 2\epsilon + \epsilon^2}{1 - 2\epsilon - \epsilon^2} |p_{n+1}^\mu(z)|^2;$$

that is, by assumption on g there are constants $\tilde{q}, q > 0$ such that, for all $n \geq N$ and $z \in F$,

$$\tilde{q} |p_{n+1}^\mu(z)| \leq |p_n^\mu(z)| \leq q |p_{n+1}^\mu(z)|,$$

implying that

$$\tilde{q}^{n-N} |p_n^\mu(z)| \leq |p_N^\mu(z)| \leq q^{n-N} |p_n^\mu(z)|. \quad (48)$$

By the assumptions on Ω and F we know from [1, Theorem 1] that

$$\limsup_{n \rightarrow \infty} |p_n^\mu(z)|^{1/n} > 1,$$

which together with (48) and $\tilde{q} \neq 0$ implies that $p_N^\mu(z) \neq 0$ for $z \in F$, and hence

$$c_1 := \min_{z \in F} |p_N^\mu(z)|^2 > 0, \quad c_2 := \max_{z \in F} K_N^\mu(z, z) < \infty.$$

In addition, taking n th roots in (48), we find that $q < 1$. (The same argument implies that $|g(z)| < 1$ for $z \in \Omega$.) In particular,

$$K_n^\mu(z, z) \geq |p_n^\mu(z)|^2 \geq q^{2(N-n)} |p_N^\mu(z)|^2 \geq c_1 q^{2(N-n)},$$

showing that $1/K_n^\mu(z, z) \rightarrow 0$ uniformly on F , as claimed in the theorem. Conversely,

$$\begin{aligned} K_n^\mu(z, z) &= K_N^\mu(z, z) + \sum_{j=N+1}^n |p_j^\mu(z)|^2 \\ &\leq c_2 + |p_n^\mu(z)|^2 \sum_{j=0}^n q^{n-j} \leq c_2 + \frac{|p_n^\mu(z)|^2}{1-q}. \end{aligned}$$

As a consequence,

$$|p_N^\mu(z)| \leq \sqrt{K_N^\mu(z, z)} = o(|p_n^\mu(z)|)_{n \rightarrow \infty}, \quad \sqrt{K_n^\mu(z, z)} = \mathcal{O}(|p_n^\mu(z)|)_{n \rightarrow \infty},$$

uniformly on F .

Summing the inequality (47) for $j = N, N+1, \dots, n-1$ we obtain with help of the Cauchy-Schwarz inequality that

$$\begin{aligned} &\left| (1 - g(z)\overline{g(w)}) \left(K_n^\mu(z, w) - K_N^\mu(z, w) \right) - p_n^\mu(z)\overline{p_n^\mu(w)} + p_N^\mu(z)\overline{p_N^\mu(w)} \right| \\ &\leq \sum_{j=N}^{n-1} |p_j^\mu(z)\overline{p_j^\mu(w)} - g(z)p_{j+1}^\mu(z)\overline{g(w)p_{j+1}^\mu(w)}| \\ &\leq (2\epsilon + \epsilon^2) \sum_{j=N}^{n-1} \sqrt{|p_j^\mu(z)|^2 + |p_{j+1}^\mu(z)|^2} \sqrt{|p_j^\mu(w)|^2 + |p_{j+1}^\mu(w)|^2} \\ &\leq 2(2\epsilon + \epsilon^2) \sqrt{K_n^\mu(z, z)K_n^\mu(w, w)}. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we conclude that

$$(1 - g(z)\overline{g(w)})K_n^\mu(z, w) = p_n^\mu(z)\overline{p_n^\mu(w)}(1 + o(1)_{n \rightarrow \infty}), \quad (49)$$

$$(1 - |g(z)|^2)K_n^\mu(z, z) = |p_n^\mu(z)|^2(1 + o(1)_{n \rightarrow \infty}), \quad (50)$$

uniformly on F . This implies (43).

For a proof of (44), we may suppose that $z_1, \dots, z_\ell, w_1, \dots, w_\ell \in F$, and neglect the factors of modulus 1 on the left-hand side of (43). Since the

square roots on the right-hand side of (43) can be factored by linearity of the determinant, for establishing (44) it is sufficient to recall the well-known expression for the determinant of a Pick matrix, namely

$$\det\left(\frac{1}{1 - a_j \bar{b}_k}\right)_{j,k=1,\dots,\ell} = \frac{\prod_{j,k=1, j < k}^{\ell} (a_k - a_j)(\bar{b}_k - \bar{b}_j)}{\prod_{j,k=1}^{\ell} (1 - a_j \bar{b}_k)}. \quad (51)$$

This last formula follows by considering the left-hand determinant as a function of a_ℓ , having simple roots at $a_1, \dots, a_{\ell-1}$ and simple poles at $1/\bar{b}_1, \dots, 1/\bar{b}_\ell$ (and similarly as a function of \bar{b}_ℓ). Hence

$$\det\left(\frac{1}{1 - a_j \bar{b}_k}\right)_{j,k=1,\dots,\ell} = \frac{C}{1 - a_\ell \bar{b}_\ell} \prod_{j=1}^{\ell-1} \frac{a_\ell - a_j}{1 - a_\ell \bar{b}_j} \frac{\bar{b}_\ell - \bar{b}_j}{1 - a_j \bar{b}_\ell},$$

with a constant C independent of a_ℓ . Since the residual at $a_\ell = 1/\bar{b}_\ell$ of the right-hand product equals 1 we obtain, by recurrence on ℓ , formula (51). \square

Combining the asymptotic findings of Theorem 5.1 with our upper and lower bounds (39) and (40) for $K_n^{\mu+\sigma}(z, z)/K_n^\mu(z, z)$ we prove the following result for adding ℓ distinct point masses from Ω ; that is, $\sigma = \sum_{j=1}^{\ell} t_j \delta_{z_j}$.

Corollary 5.4. *Under the assumptions of Theorem 5.1, and in particular $z_1, \dots, z_\ell \in \Omega$ with distinct $g(z_1), \dots, g(z_\ell)$, we have uniformly on compact subsets of Ω*

$$\lim_{n \rightarrow \infty} \frac{K_n^{\mu+\sigma}(z, z)}{K_n^\mu(z, z)} = \left| \prod_{j=1}^{\ell} \frac{g(z) - g(z_j)}{1 - g(z)g(z_j)} \right|^2, \quad (52)$$

and, at a point mass $z = z_m$,

$$K_n^{\mu+\sigma}(z_m, z_m) = \frac{1}{t_m} \left(1 + \mathcal{O}\left(\frac{1}{K_n^\mu(z_m, z_m)}\right) \right)_{n \rightarrow \infty}. \quad (53)$$

Proof. We appeal to a special case of Theorem 4.2 that asserts

$$\Sigma_1 \leq \Sigma_3 \leq \frac{K_n^{\mu+\sigma}(z, z)}{K_n^\mu(z, z)} \leq \Sigma_2,$$

where we recall that tacitly all quantities depend on n and z . We will also use the abbreviations

$$B(z) := \prod_{j=1}^{\ell} \frac{g(z) - g(z_j)}{1 - g(z)g(z_j)}, \quad B_j(z) := \prod_{k=1, k \neq j}^{\ell} \frac{g(z) - g(z_k)}{1 - g(z)g(z_k)}.$$

The matrix $C = C_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)$ is positive semi-definite; hence by (39)

$$\begin{aligned} \Sigma_1 &= \frac{\det C_n^\mu(z_1, \dots, z_\ell, z; z_1, \dots, z_\ell, z)}{\det C_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)} \\ &= \left| \frac{\det C_n^\mu(z_1, \dots, z_\ell, z; z_1, \dots, z_\ell, z)}{\det C_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)} \right|, \end{aligned}$$

which, by (44) for $w_j = z_j$ and $z = w$, tends to $|B(z)|^2$ as $n \rightarrow \infty$ for $z \in \Omega$. It remains to examine $\Sigma_2 - \Sigma_1$: we recall from (41) that

$$\Sigma_2 - \Sigma_1 = \sum_{j=1}^{\ell} \frac{|\det C_n^\mu(z_1, \dots, z_{j-1}, z, z_{j+1}, \dots, z_\ell; z_1, \dots, z_\ell)|^2}{|\det C_n^\mu(z_1, \dots, z_\ell; z_1, \dots, z_\ell)|^2} \frac{1}{t_j K_n^\mu(z_j, z_j)},$$

which according to (44) behaves like

$$\sum_{j=1}^{\ell} \frac{1}{t_j K_n^\mu(z_j, z_j)} \left(1 - \left| \frac{g(z) - g(z_j)}{1 - g(z)g(z_j)} \right|^2\right) \left| \frac{B_j(z)B_j(z_j)}{B_j(z_j)B_j(z_j)} \right|^2,$$

where $B_j(z_j) \neq 0$ by assumption on $g(z_1), \dots, g(z_\ell)$. From Theorem 5.1 and its proof we know that $1/K_n^\mu(z_j, z_j)$ exponentially decays to 0 for all j as $n \rightarrow \infty$, implying that (52) holds.

Notice that (52) is not very useful at a point mass $z = z_m$ since then $B(z_m) = 0$ and even $\Sigma_1 = 0$, implying that $b^* = Ce_m$, the m th column of C . Here it is more helpful to return to the definition of Σ_j and observe that

$$\begin{aligned} \Sigma_2 - \Sigma_1 &= bC^{-1}D^2C^{-1}b^* = e_m^*D^2e_m = \frac{1}{t_m K_n^\mu(z_m, z_m)}, \\ \Sigma_2 - \Sigma_3 &= bC^{-1}D^2C^{-1}D^2C^{-1}b^* = \frac{e_m^*C^{-1}e_m}{\left(t_m K_n^\mu(z_m, z_m)\right)^2}. \end{aligned}$$

From Cramer's rule and (44) we deduce that $e_m^*C^{-1}e_m$ has a limit different from 0 as $n \rightarrow \infty$. Hence (53) holds. \square

Remark 5.5. We have precise asymptotics with error terms in (53), but such error terms are missing in (52) as well as in (43) and (44).

Under the assumption of Theorem 5.1, a slightly more careful error analysis shows that the maximal error for $z \in F$ in (49), (50) and thus in (43) is bounded by a constant times

$$\max_{z \in F} \left| \frac{p_n^\mu(z) - g(z)p_{n+1}^\mu(z)}{\sqrt{|p_n^\mu(z)|^2 + |p_{n+1}^\mu(z)|^2}} \right|$$

plus an exponentially decreasing term. Since $|C_n^\mu(\cdot, \cdot)| \leq 1$, the same is true for (44) by linearity of the determinant, and thus also for (52).

Remark 5.6. In the examples presented in the next section, $\text{supp}(\mu)$ is compact with smooth boundary, Ω is the unbounded connected component of $\mathbb{C} \setminus \text{supp}(\mu)$ being supposed to be simply connected, and (42) holds with $g(z) = 1/\Phi(z)$, with Φ the Riemann conformal map from Ω onto the exterior of the closed unit disk. In particular, g is injective and, with z_1, \dots, z_ℓ , also $g(z_1), \dots, g(z_\ell)$ are distinct. Moreover, the limit in (52) is different from 0 for $z \in \Omega$ different from a point mass.

Also, in this case, Ω is known to be regular with respect to the Dirichlet problem, and the measures μ of §6 are all of class **Reg**, which together with

(7) allows us to specify the rate of geometric convergence mentioned in the preceding Remark.

6. ASYMPTOTICS IN BERGMAN SPACE

In this section we discuss a special and well known case of orthogonality in one complex variable where the precise asymptotic behavior of the Christoffel-Darboux kernels, both in (a small neighborhood of) the support of the measure of orthogonality and far enough from the support is precisely known; Theorem 6.2 below contains precise ratio asymptotics, relevant for our study. The findings of the previous sections are then put to work, yielding the performance of the leverage score, see Corollary 6.5.

Throughout this section G denotes a bounded open subset of the complex plane with simply connected complement, with boundary Γ ; μ stands for the area measure on $\text{Clos}(G) = G \cup \Gamma$. We denote by Φ the Riemann outer conformal map from $\mathbb{C} \setminus \text{Clos}(G)$ onto $\mathbb{C} \setminus \text{Clos}(\mathbb{D})$ fixing the point at infinity. As usual, for $r > 1$, the compact level sets are defined by their complement,

$$\mathbb{C} \setminus G_r := \{z \in \mathbb{C} \setminus \text{Clos}(G) : |\Phi(z)| > r\}.$$

Henceforth c_j is used to denote some absolute, strictly positive constants neither depending on z nor on n .

Bergman orthogonal polynomials and their kernels have been used quite successfully as building blocks of conformal maps; the long history of their asymptotics is recorded in the monographs [14], [40]; the recent article [4] deals with the case of piecewise analytic boundaries. To provide some comparison basis, we start by the simplest case of the unit disk.

Example 6.1. To be more explicit, set $\mu = \mu_{\mathbb{D}}$, the area measure on the unit disk. Since $p_n^{\mu_{\mathbb{D}}}(w) = \sqrt{(n+1)/\pi} w^n$, explicit formulas for the Christoffel kernel are at hand; in particular,

$$\max_{w \in \text{supp}(\mu_{\mathbb{D}})} K_n^{\mu_{\mathbb{D}}}(w, w) = \frac{(n+1)(n+2)}{2\pi}, \quad (54)$$

$$K_n^{\mu_{\mathbb{D}}}(w, w) = \frac{n+1}{\pi} \frac{|w|^{2n+2}}{|w|^2 - 1} (1 + \mathcal{O}(1/n)_{n \rightarrow \infty}), \quad (55)$$

the second relation being true uniformly on compact subsets of $\mathbb{C} \setminus \text{supp}(\mu_{\mathbb{D}})$.

The following theorem collects all relevant estimates and asymptotics.

Theorem 6.2. *Let μ be area measure on G , and suppose that Γ is a Jordan curve which is either piecewise analytic without cusps, or otherwise possesses an arc length parametrization with a derivative being 1/2-Hölder continuous.*

(a) *For all $n \geq 0$ and $z \in G$:*

$$K_n^{\mu}(z, z) \leq \frac{1}{\pi \text{dist}(z, \Gamma)^2}.$$

(b) With $r(n) := \sqrt{1 + \frac{1}{n+1}}$ the estimates

$$\begin{aligned} \frac{1}{e} \max_{z \in G_{r(n)}} K_n^\mu(z, z) &\leq \max_{z \in \text{supp}(\mu)} K_n^\mu(z, z) \leq \max_{z \in G_{r(n)}} K_n^\mu(z, z) \\ &\leq \gamma_n := \frac{c_1}{\text{dist}(\Gamma, \partial G_{r(n)})^2}. \end{aligned}$$

hold for all $n \geq 0$.

(c) We have

$$K_n^\mu(z, z) = \frac{n+1}{\pi} \frac{|\Phi'(z)|^2}{|\Phi(z)|^2 - 1} |\Phi(z)|^{2n+2} (1 + \mathcal{O}(\frac{1}{n})_{n \rightarrow \infty})$$

uniformly on compact subsets of $\mathbb{C} \setminus \text{supp}(\mu)$.

(d) The asymptotics

$$\frac{p_n^\mu(z)}{p_{n+1}^\mu(z)} = \frac{1}{\Phi(z)} (1 + \mathcal{O}((1/n)_{n \rightarrow \infty}))$$

is valid uniformly on compact subsets of $\mathbb{C} \setminus \text{supp}(\mu)$.

Proof. Part (a) is classical (being valid for any domain G); it follows, e.g., from [14, Lemma 1]. For a proof of part (b), we claim that the first two inequalities are simple consequences of the maximum principle and (5). Indeed, for any $r \geq 1$,

$$\max_{z \in \text{supp}(\mu)} K_n^\mu(z, z) \leq \max_{z \in G_r} K_n^\mu(z) \leq \max_{z \in G_r} \max_{\deg P \leq n} \frac{|P(z)|^2}{\|P\|_{2,\mu}^2}.$$

Using the maximum principle for $P(z)$ on G_r we find the upper bound

$$\max_{z \in G_r} K_n^\mu(z) \leq \max_{\deg P \leq n} \max_{z \in \partial G_r} \frac{|P(z)|^2}{\|P\|_{2,\mu}^2} \leq \max_{z \in \partial G_r} K_n^\mu(z, z),$$

and thus the maximum of $K_n^\mu(z, z)$ is attained at the boundary. Similarly, from the maximum principle for $P(z)/\Phi(z)^n$ on $\mathbb{G} \setminus \text{supp}(\mu)$ we infer that

$$\max_{z \in \partial G_{r(n)}} K_n^\mu(z, z) \leq r(n)^{2n} \max_{\deg P \leq n} \max_{z \in \text{supp}(\mu)} \frac{|P(z)|^2}{\|P\|_{2,\mu}^2} \leq e \max_{z \in \text{supp}(\mu)} K_n^\mu(z, z),$$

showing that the first two inequalities of part (b) are true. In order to show the last inequality, we closely follow [4], and consider the row vectors

$$\begin{aligned} P &:= (\psi'(w) p_j^\mu(z))_{j=0, \dots, n}, & Q &:= \left(\sqrt{\frac{j+1}{\pi}} w^j \right)_{j=0, \dots, n}, \\ F &:= \left(\psi'(w) \frac{F'_{j+1}(z)}{\sqrt{\pi(j+1)}} \right)_{j=0, \dots, n}, \end{aligned}$$

where $w = \Phi(z)$ is of modulus larger than 1, $z = \psi(w)$, and F_n is the n th Faber polynomial associated to G . Notice that $\|P\|^2 = K_n^\mu(z, z)/|\Phi'(z)|^2$, whereas

$$\|Q\|^2 = K_n^{\mu^\mathbb{D}}(w, w) \leq \frac{n+1}{\pi} \frac{|w|^{2n+2}}{|w|^2 - 1}$$

gives the n th Bergman Christoffel-Darboux kernel for the unit disk discussed in Example 6.1. From [4, Eqn. (1.5)] we know that the k th component of $F - Q$ is given by the k th component of $(\sqrt{(j+1)/\pi}w^{-j-2})_{j=0,1,\dots}C$, with C the infinite Grunsky matrix having norm strictly less than 1, and hence

$$\forall |w| > 1: \quad \|F - Q\|^2 \leq \frac{1}{\pi(|w|^2 - 1)^2}.$$

From [4, Eqn. (2.1) and Corollary 2.3] we know that there exists an upper triangular matrix R_n with $\|R_n\| \leq 1$ and $\|R_n^{-1}\| \leq 1/\sqrt{1 - \|C\|^2}$ such that $F = PR_n$. Thus

$$\forall |w| > 1: \quad \|F\|^2 \leq \|P\|^2 \leq \frac{\|F\|^2}{1 - \|C\|^2}.$$

Combining these findings gives for $z \in \partial G_{r(n)}$ and hence $|w|^2 - 1 = 1/(n+1)$

$$\begin{aligned} K_n^\mu(z, z) = |\Phi'(z)|^2 \|P\|^2 &\leq |\Phi'(z)|^2 \frac{(\|Q\| + \|F - Q\|)^2}{1 - \|C\|^2} \\ &\leq \frac{2.5}{1 - \|C\|^2} \frac{|\Phi'(z)|^2}{(|\Phi(z)|^2 - 1)^2}. \end{aligned}$$

Recalling from [43, Theorem 3.1] that

$$\frac{1/2}{\text{dist}(z, \Gamma)} \leq \frac{|\Phi'(z)|}{|\Phi(z)| - 1} \leq \frac{2}{\text{dist}(z, \Gamma)},$$

we have established the last inequality of part (b), with $c_1 \leq \frac{2.5}{1 - \|C\|^2}$ depending only on the geometry of G .

For a proof of parts (c) and (d), we recall from [4, Eqn. (1.12)] that $\varepsilon_n = \|Ce_n\|^2 = \mathcal{O}(1/n)$ by our assumptions on Γ , and hence by [4, Theorem 1.1], uniformly for z in a compact subset of $\mathbb{C} \setminus \text{supp}(\mu)$,

$$p_n^\mu(z) = \Phi'(z)\sqrt{n+1}\pi\Phi(z)^n \left(1 + \mathcal{O}\left(\frac{1}{n+1}\right)_{n \rightarrow \infty}\right). \quad (56)$$

Then part (c) follows by taking squares in (56), summing, and finally applying (55) with $w = \Phi(z)$. Also, part (d) on ratio asymptotics is an immediate consequence of (56). \square

Remarks 6.3. (a) We claim without proof that the estimates of the previous statement and its proof could be improved if we are willing to add additional smoothness assumptions on the boundary Γ . E.g., for a \mathcal{C}^2 boundary Γ , $\|F - Q\|/\|Q\|$ can be shown to tend to zero for $n \rightarrow \infty$ uniformly for $|\Phi(z)|^2 - 1 \geq 1/(n+1)$, leading to lower and upper bounds for $K_n^\mu(z, z)$ that are sharp for $n \rightarrow \infty$ up to a constant.

(b) For general Γ as in Theorem 6.2, the quantity γ_n of Theorem 6.2(b) can be shown to behave like $(n+1)^2$ in the case when Γ has no corner of outer angle $> \pi$. Since this is true in particular for a \mathcal{C}^2 boundary, our Theorem 6.2(b) is in accordance with a recent result of Totik [42, Theorem 1.3] who showed under this additional assumption that the limit

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \max_{z \in \partial \text{supp}(\mu)} K_n^\mu(z, z) \text{ is finite and } \neq 0.$$

In the case when there is a maximal outer angle $= \alpha\pi$ with $\alpha > 1$, γ_n can be shown to behave like $(n+1)^{2\alpha}$, which we believe to be also the rate of growth of $\max_{z \in \partial \text{supp}(\mu)} K_n^\mu(z, z)$.

(c) The $\mathcal{O}(1/n)$ error term in Theorem 6.2(c) can be improved for smoother boundaries Γ , but not in Theorem 6.2(d), see Example 6.1. \square

In the next statement we improve a statement of Lasserre and Pauwels [20, Theorem 3.8 and Theorem 3.10]; compare with Remark 3.5. Here we allow for a finite number of fixed outliers in the case of one complex variable.

Corollary 6.4. *Let μ be as in Theorem 6.2, and*

$$\nu = \mu + \sigma, \quad \sigma = \sum_{j=1}^{\ell} t_j \delta_{z_j}, \quad t_j > 0, \quad z_1, \dots, z_\ell \in \mathbb{C} \setminus \text{supp}(\mu) \text{ distinct.}$$

Then the Hausdorff distance between $\text{supp}(\nu)$ and the level set $S_n := \{z \in \mathbb{C} : K_n^\nu(z, z) \leq \gamma_n\}$ with γ_n as in Theorem 6.2(b) tends to zero as $n \rightarrow \infty$.

Proof. We first recall from Theorem 4.2 that $K_n^\nu(z, z) \leq K_n^\mu(z, z)$, and hence $\text{supp}(\mu) \subseteq S_n$ by Theorem 6.2(b). Also, setting $\nu_j := \nu - t_j \delta_{z_j}$, we know from Theorem 4.2 that

$$K_n^\nu(z_j, z_j) = \frac{1/t_j}{1 + \frac{1}{t_j K_n^{\nu_j}(z_j, z_j)}}.$$

The quantity $K_n^{\nu_j}(z_j, z_j)/K_n^\mu(z_j, z_j)$ has a non-zero limit according to Corollary 5.4 and Theorem 6.2(d). Also, Theorem 6.2(c) together with Remark 6.3(b) imply that $K_n^\mu(z_j, z_j)/\gamma_n$ grows geometrically large. Hence $K_n^\nu(z_j, z_j) \rightarrow 1/t_j$ with a geometric rate, implying that $\text{supp}(\nu) \subseteq S_n$ for sufficiently large n .

For z outside of a neighborhood U of $\text{supp}(\nu)$ we write

$$\frac{K_n^\nu(z, z)}{\gamma_n} = \frac{K_n^\nu(z, z)}{K_n^\mu(z, z)} \frac{K_n^\mu(z, z)}{\gamma_n}.$$

The first factor on the right-hand side has a non-zero limit according to Corollary 5.4 and Theorem 6.2(d) uniformly for $z \notin U$. Also, again by Theorem 6.2(c) and Remark 6.3(b), $K_n^\mu(z, z)/\gamma_n$ grows at least as a constant times $|\Phi(z)|^{2n}/n^\beta$ uniformly for $z \notin U$ for some constant $\beta > 0$, and we conclude that $S_n \subseteq U$ for sufficiently large n , showing the convergence of the Hausdorff distance. \square

We are now prepared to describe a situation where the efficiency of our leverage score for detecting outliers can be tested.

Corollary 6.5. *Let μ, ν, σ be as in Corollary 6.4, and n be sufficiently large. Furthermore, consider the discrete measures*

$$\tilde{\nu} = \tilde{\mu} + \sigma, \quad \tilde{\mu} = \sum_{j=\ell+1}^N t_j \delta_{z_j}, \quad t_j > 0, \quad z_{\ell+1}, \dots, z_N \in \text{supp}(\mu) \text{ distinct,}$$

where we assume that $\tilde{\mu}$ is sufficiently close to μ such that $\|M_n(\tilde{\mu}, \mu) - I\| \leq 1/2$. Then there exist $c > 0$ and $q \in (0, 1)$ depending only on μ, σ but not on $\tilde{\nu}$ such that, for the outliers,

$$j = 1, 2, \dots, \ell : \quad 1 - t_j K_n^{\tilde{\nu}}(z_j, z_j) \leq cq^n,$$

whereas for the other mass points of $\tilde{\nu}$

$$j = \ell + 1, \dots, N : \quad t_j K_n^{\tilde{\nu}}(z_j, z_j) \leq \frac{3}{2} \min \left\{ t_j \gamma_n, \frac{t_j}{\pi \text{dist}(z_j, \Gamma)^2} \right\},$$

where γ_n is as in Theorem 6.2(b).

Proof. We first consider the case $j \in \{1, \dots, \ell\}$ of outliers. Let $\nu_j = \nu - t_j \delta_{z_j}$, $\tilde{\nu}_j = \tilde{\nu} - t_j \delta_{z_j}$. According to Remark 3.8 with $\epsilon = 1/2$ we have $K_n^{\tilde{\nu}_j}(z_j, z_j) \geq K_n^{\nu_j}(z_j, z_j)/2$, and hence

$$\begin{aligned} 1 - t_j K_n^{\tilde{\nu}}(z_j, z_j) &= \frac{1}{1 + t_j K_n^{\tilde{\nu}_j}(z_j, z_j)} \leq \frac{2}{t_j K_n^{\nu_j}(z_j, z_j)} \\ &= \frac{2}{t_j} \frac{K_n^{\mu}(z_j, z_j)}{K_n^{\nu_j}(z_j, z_j)} \frac{1}{K_n^{\mu}(z_j, z_j)}, \end{aligned}$$

and we conclude as in the previous proof using Theorem 6.2(c),(d) and Corollary 5.4 that our first assertion holds.

Now we consider the case $j \in \{\ell + 1, \dots, N\}$ of mass points $z_j \in \text{supp}(\mu)$. Applying first Theorem 4.2 and then Proposition 3.7 with $\epsilon = 1/2$, we arrive at

$$t_j K_n^{\tilde{\nu}}(z_j, z_j) \leq t_j K_n^{\tilde{\mu}}(z_j, z_j) \leq \frac{3}{2} t_j K_n^{\mu}(z_j, z_j).$$

Thus our second assertion is a consequence of Theorem 6.2(a),(b). \square

By comparing the upper left entry, our assumption $\|M_n(\tilde{\mu}, \mu) - I\| \leq 1/2$ also implies that

$$\frac{\tilde{\mu}(\mathbb{C})}{\mu(\mathbb{C})} = \frac{1}{\mu(\mathbb{C})} \sum_{j=\ell+1}^N t_j \in [1/2, 3/2],$$

and thus a typical weight of $\tilde{\mu}$ is of order $1/N$. Hence, roughly speaking, our leverage score $t_j K_n^{\tilde{\nu}}(z_j, z_j)$ is very close to 1 for the case $j \in \{1, \dots, \ell\}$ of outliers, and at least bounded for z_j lying in the interior of $\text{supp}(\mu)$, and more precisely, having a weight of typical size and z_j being at least of distance

$1/\sqrt{N}$ to the boundary Γ . Moreover, for equal weights $t_{\ell+1} = \dots = t_N$ (and thus of order $1/N$), as long as

$$\ell \text{ is fixed, } n \text{ is sufficiently large, and } N \geq \gamma_n \quad (57)$$

Corollary 6.5 tells us that we are able to identify correctly the outliers as those elements z_j of the support where our leverage score $t_j K_n^{\tilde{\nu}}(z_j, z_j)$ is close to 1.

In order to illustrate this claim, we conclude this section by reporting some numerical simulations. Here we discuss for different clouds of distinct points $z_1, \dots, z_N \in \mathbb{C}$, and equal weights $t_1 = \dots = t_N = 1/N$, with the corresponding discrete measure ν . The points z_j of the clouds are drawn following a color code given by our leverage score $t_j K_n^{\mu}(z_j, z_j)$, with red corresponding to values close to 0, and blue values close to 1. For each cloud, we draw in the upper row (without theoretical justification) the level scores for bivariate orthogonal polynomials with lexicographical ordering for total degrees 1, 4, 8 (and hence $n = 2, 14, 44$) and in the lower row for univariate Bergman orthogonal polynomials for $n = 1, 9, 44$. The first column essentially corresponds to the classical leverage score known from data analysis and mentioned in Section 2.4.

The vector of values of the leverage score has been computed in finite precision arithmetic as follows: in the univariate case, we first compute by the full Arnoldi method in complexity $\mathcal{O}(n^2 N)$ the Hessenberg matrix allowing to represent $z v_n^{\tilde{\nu}}(z)$ in terms of $v_{n+1}^{\tilde{\nu}}(z)$, and then use a link between values of the Christoffel-Darboux kernel and GMRES for the shifted Hessenberg matrix, with a total complexity of $\mathcal{O}(n^3 N)$. A generalization of this approach has been used for bivariate orthogonal polynomials. There exist more efficient approaches, but many of them suffer from loss of orthogonality. For each of our simulations we give an indicator showing that our approach does not have this drawback. Also, and this is probably *the most important message for large data sets, the complexity and memory requirements scale linearly with N* .

In the numerical experiments displayed in Figures 1, 2, and 3, we observe that the classical leverage scores in the first column do not detect any of our outliers, probably due to the small weights $t_j = 1/N$. In the experiences of the other two columns, we clearly detect outliers, but the color separation is more pronounced for one complex variable, in particular for outliers which are closer to G . In the right column for $n = 44$, the color code for the outliers is best, but in fact also corners of the domain and some other points of the support of $\tilde{\nu}$ have an outlier color code. This seems to be partly a consequence of the sampling procedure, but clearly also a consequence of the fact that we do not respect (57) since N is not large enough.

Similar sharp estimates are known for measures supported on Jordan curves, or arcs in the complex plane, but we do not detail them here.

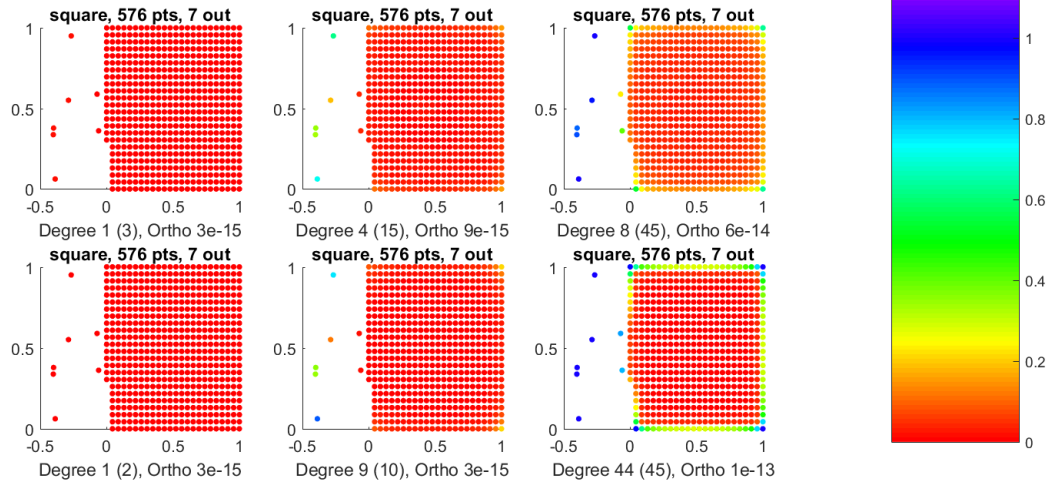


FIGURE 1. A cloud of $N = 576$ points, with 7 random outliers, the others obtained by discretizing normalized Lebesgue measure on the unit square by a regular grid.

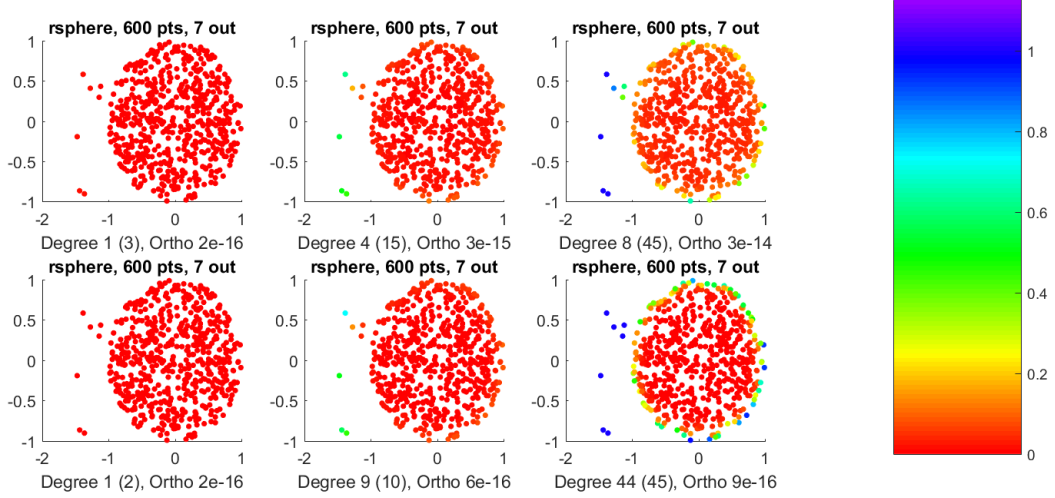


FIGURE 2. A cloud of $N = 600$ points, with $\ell = 7$ random outliers. The other points are random samplings of the unit disk.

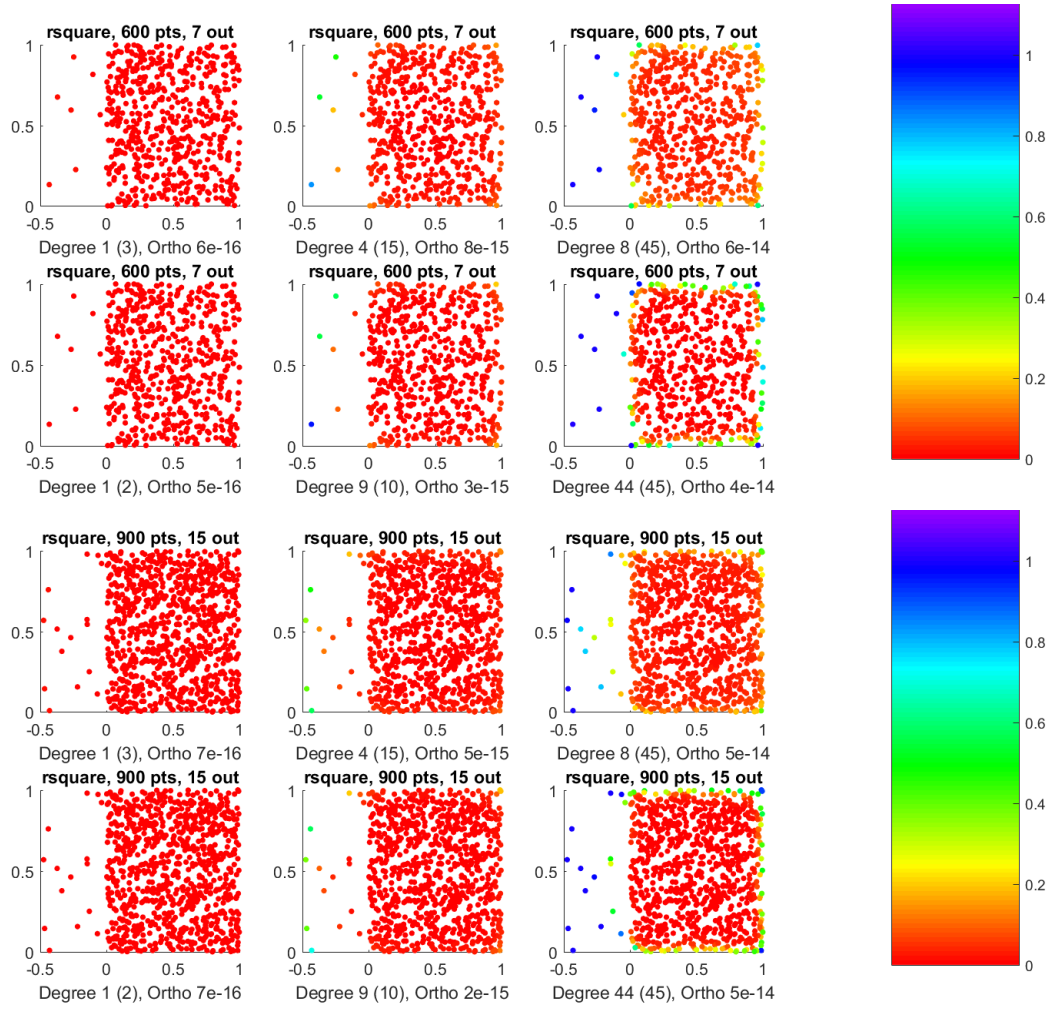


FIGURE 3. A first cloud of $N = 600$ points, with $\ell = 7$ random outliers, and a second one with $N = 900$ and 15 outliers. The other points are random samplings.

7. EXAMPLES IN THE MULTIVARIATE CASE

The present section collects a series of mostly known facts of pluripotential theory which are relevant for our study. In § 7.1 we recall several well-known formulas for the plurisubharmonic Green function $g_\Omega(z)$. In § 7.2 we consider the particular case of tensor measures on the unit square allowing to get far more precise results on asymptotics. We also include a simple numerical experiment endorsing the thesis that working in the complex domain has clear benefits even for problems formulated solely in real variables. Finally, in § 7.3 we discuss two other cases of several complex variables where explicit formulas for the Christoffel-Darboux kernel are available.

7.1. Examples of plurisubharmonic Green functions with pole at infinity. With respect to and support of Lemma 2.9, we are fortunate to have closed form expressions for the plurisubharmonic Green function $g_\Omega(z)$ with pole at infinity for a few classes of subsets of \mathbb{C}^d . All these examples are obtained via computing Siciak's extremal function, see [21]. Notice also the affine invariance not only of the Christoffel-Darboux kernel (11), but also of the plurisubharmonic Green function [21, §5.3], and thus a property being for instance valid for the real unit ball also translates to any scaled and shifted real ball. While most of results were originally stated for the total degree, and the underlying lexicographical order, a new trend of considering degrees defined by the Minkowski functional of a convex subset of the multi-indices is becoming popular [2, 6].

7.1.1. *Complex ball.* Consider for instance a complex norm $[\cdot]$ in \mathbb{C}^d and the closed ball

$$B_a(r) = \{z \in \mathbb{C}^d; [z - a] \leq r\}.$$

By a complex norm we mean a norm with the homogeneity property

$$[\lambda z] = |\lambda|[z], \quad \lambda \in \mathbb{C}, \quad z \in \mathbb{C}^d.$$

Let Ω denote the complement of $B_a(r)$ in \mathbb{C}^d . Then the set $B_a(r)$ turns out to be regular and

$$g_\Omega(z) = \log^+ \frac{[z - a]}{r}. \quad (58)$$

For a proof, see Example 5.1.1 in [21]. In particular, for the unit ball \mathbb{B}^d with respect to the standard Hermitian norm $\|\cdot\|$ we obtain

$$g_{\mathbb{C}^d \setminus \mathbb{B}^d}(z) = \log^+ \|z\|.$$

7.1.2. *Polynomial polyhedra.* Let p_1, p_2, \dots, p_d be complex polynomials in d variables, with highest homogeneous parts $\hat{p}_1, \hat{p}_2, \dots, \hat{p}_d$ respectively. Assume that

$$\sum_{j=1}^d |\hat{p}_j(z)| = 0 \quad \text{iff} \quad z = 0;$$

that is, the map $(p_1, p_2, \dots, p_d) : \mathbb{C}^d \rightarrow \mathbb{C}^d$ has finite fibres. Consider the analytic polyhedron

$$K = \{z \in \mathbb{C}^d; |p_j(z)| \leq 1, \quad 1 \leq j \leq d\},$$

and its complement $\Omega = \mathbb{C}^d \setminus K$. Then (see Corollary 5.3.2 in [21])

$$g_\Omega(z) = \max_{1 \leq j \leq d} \frac{\log^+ |p_j(z)|}{\deg p_j}. \quad (59)$$

In particular, for the polydisk

$$\mathbb{D}^d = \{z \in \mathbb{C}^d; |z_j| \leq 1, \quad 1 \leq j \leq d\}$$

we get

$$g_{\mathbb{C}^d \setminus \mathbb{D}^d}(z) = \max_{1 \leq j \leq d} \log^+ |z_j|. \quad (60)$$

7.1.3. Product domains. A theorem due to Siciak provides the naturality of the Green function on product sets. More specifically, let $K \subseteq \mathbb{C}^d$ and $L \subseteq \mathbb{C}^e$ be compact subsets. Then

$$g_{\mathbb{C}^{d+e} \setminus (K \times L)}(z, w) = \max(g_{\mathbb{C}^d \setminus K}(z), g_{\mathbb{C}^e \setminus L}(w)), \quad z \in \mathbb{C}^d, \quad w \in \mathbb{C}^e. \quad (61)$$

For details, see Theorem 5.1.8 in [21].

7.1.4. Real subsets of complex space. Although a great deal of examples live in \mathbb{R}^d , it is necessary to treat them as subsets of \mathbb{C}^d . Almost all closed form known formulas are obtained by pull-back from the Green function with pole at infinity of the real interval $[-1, 1]$, given by

$$g_{\mathbb{C} \setminus [-1, 1]}(z) = \log |z + \sqrt{z^2 - 1}|, \quad z \in \mathbb{C} \setminus [-1, 1],$$

and $g_{\mathbb{C} \setminus [-1, 1]}(z) = 0$ in for $z \in [-1, 1]$. Here the square root is chosen so that $g_{\mathbb{C} \setminus [-1, 1]} \geq 0$.

Consider a compact subset $E \subseteq \mathbb{R}^d$ that is convex, has non-empty interior and is symmetric with respect to the origin: $x \in E$ implies $-x \in E$. Following Lundin [25] one can represent E as follows

$$E = \{z \in \mathbb{C}^d : \forall \omega \in S^{d-1}, a(\omega)\omega^*z \in [-1, 1]\},$$

where S^{d-1} denotes the unit sphere in \mathbb{R}^d , and a is continuous on S^{d-1} . We can choose $a(\omega)$ to be equal to the inverse of half the width of E in the direction ω . If the boundary of E is smooth, then there is no ambiguity in defining $a(\omega)$. The main result of [25] then states that

$$g_{\mathbb{C}^d \setminus E}(z) = \max_{\omega \in S^{d-1}} g_{\mathbb{C} \setminus [-1, 1]}(a(\omega)\omega^*z).$$

A few particular cases are relevant and we list them under separate subsections.

7.1.5. *Real ball.* On the unit ball $B^d \subseteq \mathbb{R}^d$ we remark $a(\omega) = 1$ for all values of $\omega \in S^{d-1}$ and therefore

$$g_{\mathbb{C}^d \setminus B^d}(z) = \frac{1}{2} g_{\mathbb{C} \setminus [-1,1]}(\|z\|^2 + |z \cdot z - 1|), \quad z \notin B^d. \quad (62)$$

This result is due to Siciak, see Theorem 5.4.6 in [21]. In particular, for a real vector $x \in \mathbb{R}^d$ we infer:

$$g_{\mathbb{C}^d \setminus B}(x) = g_{\mathbb{C} \setminus [-1,1]}(\|x\|) = \log(\|x\| + \sqrt{\|x\|^2 - 1}) = \frac{1}{2} g_{\mathbb{C} \setminus [-1,1]}(2\|x\|^2 - 1).$$

7.1.6. *Real cube.* According to the product formula (61) we find

$$g_{\mathbb{C}^d \setminus [-1,1]^d}(z) = \max_{1 \leq j \leq d} |g_{\mathbb{C} \setminus [-1,1]}(z_j)|. \quad (63)$$

7.1.7. *Simplex.* Denote by Δ^d the standard simplex in \mathbb{R}^d ; that is, the convex hull of the vectors $(0, e_1, \dots, e_d)$, where (e_j) is the standard orthonormal basis. A base change of Lundin's formula via the map $(z_1, z_2, \dots, z_d) \mapsto (z_1^2, z_2^2, \dots, z_d^2)$ leads to the following closed form expression, discovered by Baran:

$$g_{\mathbb{C}^d \setminus \Delta^d}(z) = g_{\mathbb{C} \setminus [-1,1]}(|z_1| + |z_2| + \dots + |z_d| + |z_1 + \dots + z_d - 1|). \quad (64)$$

For details, see Example 5.4.7 in [21].

7.2. The real square with tensor product of Chebyshev weights.

Let us consider for $x = (x_1, x_2)^T \in [-1, 1]^2$ the tensor measure

$$d\mu(x) = d\omega(x_1)d\omega(x_2), \quad d\omega(x_1) = \frac{dx}{\pi\sqrt{1-x_1^2}}.$$

The orthogonal polynomials for the equilibrium measure ω on $[-1, 1]$ are explicitly known; namely, $p_0^\omega(z) = 1$ and $p_n^\omega(z) = \sqrt{2}T_n(z)$ for $n \geq 1$, where T_n are Chebyshev polynomials of the first kind. In particular, if $\alpha(n) = (j, k)$ then $p_n^\mu(x) = p_j^\omega(x_1)p_k^\omega(x_2)$, making it possible to get more explicit information about the underlying Christoffel-Darboux kernel.

Theorem 7.1. *Consider the enumeration of the multi-indices consistent with the tensor structure; that is, for all integers $n \geq 0$ with $N = n_{\max}(n) = (n+1)^2 - 1$*

$$\{\alpha(0), \dots, \alpha(n_{\max}(N))\} = \left\{ \begin{bmatrix} j \\ \ell \end{bmatrix}, 0 \leq j, \ell \leq n \right\},$$

and set

$$g(z) = g_{\mathbb{C} \setminus [-1,1]}(z_1)g_{\mathbb{C} \setminus [-1,1]}(z_2). \quad (65)$$

(a) For $z' = (1, 1)$ and $N = (n+1)^2 - 1$,

$$\max_{z \in \text{supp}(\mu)} K_N^\mu(z, z) = K_N(z', z') = (2n+1)^2 \leq 4N+1.$$

(b) For all $z \in \mathbb{R}^2 \setminus \text{supp}(\mu)$ and $N = (n+1)^2 - 1 \rightarrow \infty$,

$$K_N^\mu(z, z) \sim \begin{cases} e^{2ng(z)} & \text{in the "corner case" } |z_1| > 1 \text{ and } |z_2| > 1, \\ (n+1)e^{2ng(z)} & \text{else,} \end{cases},$$

where \sim means that the quotient has a finite and non-zero limit for $n \rightarrow \infty$.

(c) For all (generic) $z, w \in \mathbb{R}^2 \setminus \text{supp}(\mu)$ with $z_1 \neq w_1$ and $z_2 \neq w_2$, and $N = (n+1)^2 - 1 \rightarrow \infty$, n even, we have $C_N^\mu(z, w)$ tending to some finite non-zero limit in the corner case for both z, w , and else tending to 0.

Proof. We first mention that, for $N = (n+1)^2 - 1$,

$$K_N^\mu(z, w) = K_n^\omega(z_1, w_1)K_n^\omega(z_2, w_2), \quad C_N^\mu(z, w) = C_n^\omega(z_1, w_1)C_n^\omega(z_2, w_2),$$

and hence for the assertion of the Theorem it is sufficient to consider the univariate case. For a proof of part (a), it is sufficient to recall that, for $z_1 \in [-1, 1]$, $T_k(z_1)^2 \leq 1 = T_k(1)^2$, and hence $K_n^\omega(z_1, z_1) \leq K_n^\omega(1, 1) = 2n+1$.

For a proof of part (b) and (c), we write

$$z_j = \frac{1}{2}\left(u_j + \frac{1}{u_j}\right), \quad w_j = \frac{1}{2}\left(v_j + \frac{1}{v_j}\right), \quad |u_j| \geq 1, \quad |v_j| \geq 1,$$

and $\log |u_1| = g_{\mathbb{C} \setminus [-1, 1]}(z_1)$. Then by using geometric sums it is quite easy to check that

$$K_n^\omega(z_1, z_1) \sim \begin{cases} \frac{|u_1|^{2n}}{2(1-1/|u_1|^2)} & \text{if } z_1 \notin [-1, 1], \text{ that is, } |u_1| > 1, \\ (n+1) & \text{if } z_1 = \cos(t) \in [-1, 1], \text{ that is, } u_1 = e^{it}, \end{cases}$$

implying part (b).

For a proof of part (c), we recall that, for distinct real $w_1, z_1 \notin [-1, 1]$ with $|z_1| > 1, |w_1| > 1$ we obviously have ratio asymptotics $p_n^\omega(z_1)/p_{n+1}^\omega(z_1) \rightarrow 1/u_1$ for $n \rightarrow \infty$, and hence by Theorem 5.1

$$\lim_{n \rightarrow \infty} C_n^\omega(z_1, w_1) \left(\frac{z_1 w_1}{|z_1 w_1|} \right)^n = \frac{\sqrt{(u_1^2 - 1)(v_1^2 - 1)}}{u_1 v_1 - 1},$$

which simplifies for even n . If however $z_1 \in [-1, 1]$, then the quotient $|K_n^\omega(z_1, w_1)|/\sqrt{K_n^\omega(w_1, w_1)}$ is shown to be bounded in both cases $w_1 \notin [-1, 1]$, and $w_1 \in [-1, 1] \setminus \{z_1\}$, and hence $C_n^\omega(z_1, w_1)$ tends to 0 by part (b). \square

If $z \in \mathbb{R}^2 \setminus \text{supp}(\mu)$ and the masspoints of σ have distinct coordinates, we conclude as in the proof of Corollary 5.4 that $K_N^{\mu+\sigma}(z, z)/K_N^\mu(z, z)$ has a limit for $N = (n+1)^2 - 1 \rightarrow \infty$, which is different from one only if the corner case holds for z and at least one mass point. Hence, as in Corollaries 6.4 and 6.5, we verify that our bivariate leverage score works in this setting.

In order to compare the preceding theorem with Lemma 2.9, we notice that, in case of graded lexicographical ordering and $N = n_{\text{tot}}(n) = (n+1)(n+2)/2 - 1$, we still have that $\max_{z \in \text{supp}(\mu)} K_N^\mu(z, z) \leq 4N+1$. Hence, with the plurisubharmonic Green function of $[-1, 1]^2$

$$\tilde{g}(z) = \max\{g_{\mathbb{C} \setminus [-1, 1]}(z_1), g_{\mathbb{C} \setminus [-1, 1]}(z_2)\}, \quad (66)$$

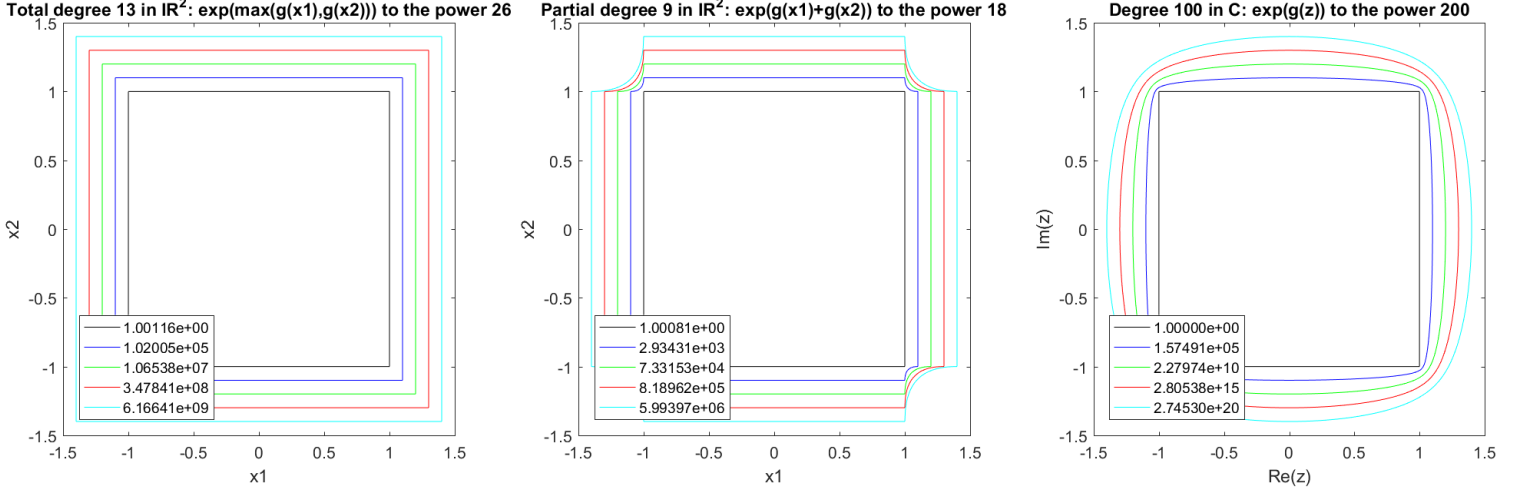


FIGURE 4. Comparison of level lines of Green functions for the unit square and three different families of orthogonal polynomials: for the left-hand image we utilized the plurisubharmonic Green function (66) of $[-1, 1]^2 \subseteq \mathbb{R}^2$, which corresponds to bivariate OP with graded lexicographical ordering (i.e., total degree); the middle image is generated by using the tensor Green function (65) corresponding to bivariate OP and partial degree is used; and for the right-most image, the complex Green function corresponding to Bergman OP of one complex variable is used.

one may show that

$$\frac{1}{2}e^{2n\tilde{g}(z)} \leq K_N^\mu(z, z) \leq e^{2n\tilde{g}(z)} \max_{\deg P \leq N} \max_{x \in \text{supp}(\mu)} \frac{|p(x)|^2}{\|P\|_{2,\mu}^2} \leq (4N + 1)e^{2n\tilde{g}(z)},$$

the left-hand side being obtained by (14) for $p(z_1, z_2) \in \{p_n^\omega(z_1), p_n^\omega(z_2)\}$. All these findings are less precise than for the case of tensor ordering, and it seems to be non-trivial to get cosine asymptotics.

In Figure 4 we compare three approaches to detect outliers outside the unit square $[-1, 1]^2$, and in particular address the question whether one should prefer an analysis in \mathbb{R}^2 or the complex plane \mathbb{C} . Though not fully justified for the case of two real variables, we expect to be able to detect successfully outliers at $z \notin [-1, 1]^2$ for a parameter N provided that $K_N^\mu(z, z)$ is large. The left-hand plot corresponds to graded lexicographical ordering (total degree) discussed in the previous paragraph, where we draw some level lines of $\exp(2n\tilde{g}(z))$ with the plurisubharmonic Green function (66) for $z \in \mathbb{R}^2 \setminus [-1, 1]^2$ instead of those of $K_N^\mu(z, z)$, $N + 1 = (n + 1)(n + 2)/2$. In the middle plot corresponding to the tensor case (partial degree) discussed in Theorem 7.1 we draw some level lines of $\exp(2ng(z))$ with the tensor Green

function (65) for $z \in \mathbb{R}^2 \setminus [-1, 1]^2$ instead of those of $K_N^\mu(z, z)$, $N + 1 = (n + 1)^2$. Finally, on the right corresponding to the case of one complex variable z discussed in §6 one finds some level lines of $\exp(2ng_{\mathbb{C} \setminus [-1, 1]^2}(z))$ for $z \in \mathbb{C} \setminus [-1, 1]^2$ instead of those of $K_N^\mu(z, z)$, $N + 1 = n + 1$, compare with Theorem 6.2(c). Notice that there is no easy closed form expression for this complex Green function $g_{\mathbb{C} \setminus [-1, 1]^2}$, we have used the Schwarz-Christoffel toolbox for Matlab of Toby Driscoll. In order to make comparison fair, we should use about the same number $N + 1$ of orthogonal polynomials: we have chosen from the left to the right $n = 13$, $n = 9$ and $n = 100$, corresponding to $N + 1 = 105$, $N + 1 = 100$ and $N + 1 = 101$, respectively.

Some observations are especially noteworthy. We start by comparing the two bivariate kernels on the left-hand side and in the middle: from the behavior of the level lines at the corners of $[-1, 1]^2$ it is apparent that one should prefer partial degree to total degree for detecting outliers with all components outside $[-1, 1]$ (called the corner case in Theorem 7.1). Comparing the values of the corresponding level lines, the opposite conclusion seems to be true for other outliers. However, much more striking, the parameters of the level lines on the right-hand side are increasing much faster, and also here the level curves do fine around the corners. This confirms our claim that outlier detection of $(2d)$ -dimensional data should be done in \mathbb{C}^d and not in \mathbb{R}^{2d} , at least for $d = 1$.

7.3. Christoffel-Darboux kernel associated to Reinhardt domains in \mathbb{C}^d . Natural domains of convergence for power series in several complex variables are quite diverse, and in general they are not bi-holomorphically equivalent. These are known as (logarithmically convex) Reinhardt domains and offer already a vast area of research for function theory and complex geometry. The recent monograph [18] gives a comprehensive account of the theory of Reinhardt domains. For our immediate purpose, this class of domains D is important because the monomials are orthogonal (but not orthonormal) with respect to Lebesgue measure μ^D (and thus $k_n^{\mu^D} = \deg p_n^{\mu^D} = n$ in Lemma 2.3); moreover in most cases the monomials are complete in the associated Bergman space.

By way of illustration we consider the two most important, and nonequivalent, Reinhardt domains: the ball and the polydisk. In the sequel of this subsection we will work in \mathbb{C}^d for some fixed $d \geq 1$, and enumerate the multivariate monomials z^α in graded lexicographical ordering. Let $N = N(n)$ be the number of monomials of total degree $\leq n$.

7.3.1. The complex ball. We first consider the Lebesgue (or volume) measure $\mu^{\mathbb{B}}$ on the unit ball

$$\mathbb{B} = \mathbb{B}^d = \{z \in \mathbb{C}^d; \|z\| < 1\},$$

having as boundary the odd dimensional sphere S^{2d-1} . Here with help of polar coordinates one gets

$$\int z^\alpha \bar{z}^\beta d\mu^{\mathbb{B}}(z) = \frac{\pi^d \alpha!}{(|\alpha| + d)!} \quad \text{for } \alpha = \beta, \text{ and zero else.}$$

Thus, up to normalization, the orthonormal polynomials are monomials, and we get for the Christoffel-Darboux kernel using the multinomial formula

$$K_N^{\mu^{\mathbb{B}}}(z, w) = \frac{1}{\pi^d} \sum_{k=0}^n \sum_{|\alpha|=k} \frac{(|\alpha| + d)!}{\alpha!} z^\alpha \bar{w}^\alpha = \frac{1}{\pi^d} \sum_{k=0}^n (k+1)_d (w^* z)^k,$$

where we use the abbreviation $w^* z = \bar{w}_1 z_1 + \dots + \bar{w}_d z_d$, together with the Pochhammer symbol $(a)_k = a(a+1)\dots(a+k-1)$. Again, to simplify notation, above and henceforth in this section $N = n_{tot}(n)$.

We already know from Siciak, see formula (58), that

$$\lim_{n \rightarrow \infty} K_N^{\mu^{\mathbb{B}}}(z, z)^{1/n} = \|z\|^2, \quad \|z\| > 1.$$

The explicit form of the polarized kernel allows a closer look at the asymptotic behavior for large degree.

Provided that $|w^* z| < 1$ (and thus in particular for $w, z \in \mathbb{B}$), the limit for $N \rightarrow \infty$ and thus for $n \rightarrow \infty$ exists, and is given by the Bergman reproducing kernel

$$K^{\mu^{\mathbb{B}}}(z, w) = \frac{1}{\pi^d} \sum_{k=0}^{\infty} (k+1)_d (w^* z)^k = \frac{d!}{\pi^d} (1 - w^* z)^{-d-1}.$$

Note that the unbounded region defined in $\mathbb{C}^d \times \mathbb{C}^d$ by the inequality $|w^* z| < 1$ may well project on one or the other coordinate in the exterior of the ball.

In the case $w^* z = 1$ we obtain

$$K_N^{\mu^{\mathbb{B}}}(z, w) = \frac{d!}{\pi^d} \sum_{k=0}^n \binom{k+d}{d} = \frac{d!}{\pi^d} \binom{n+1+d}{d+1} = \frac{(n+1)_{d+1}}{\pi^d (d+1)},$$

which is also an upper bound for $K_N^{\mu^{\mathbb{B}}}(z, w)$ in the case $|w^* z| = 1$. In particular we learn that $K_N^{\mu^{\mathbb{B}}}(z, z)$ grows at most like $n^{d+1} = \mathcal{O}(N^2)$ for z in the support of $\mu^{\mathbb{B}}$. Finally, in the case $|w^* z| > 1$,

$$K_N^{\mu^{\mathbb{B}}}(z, w) = \frac{(n+1)_d}{\pi^d} (w^* z)^n \sum_{j=0}^n \frac{(n-j+1)_d}{(n+1)_d} (w^* z)^{-j} \sim \frac{(n+1)_d}{\pi^d} \frac{(w^* z)^{n+1}}{w^* z - 1}$$

where in the last step we have used the fact that, for $0 \leq j \leq n$,

$$0 \leq 1 - \frac{(n-j+1)_d}{(n+1)_d} = \sum_{\ell=0}^{j-1} \frac{(n-\ell+1)_d - (n-\ell)_d}{(n+1)_d} \leq \frac{jd}{n+d}.$$

If we exclude the (unlikely for outlier analysis) case that z is a complex multiple of w , we conclude that, for $z, w \notin \text{supp}(\mu^{\mathbb{B}})$,

$$K_N^{\mu^{\mathbb{B}}}(z, z) \sim \frac{(n+1)_d \|z\|^{2n+2}}{\pi^d (\|z\|^2 - 1)}, \quad \lim_{N \rightarrow \infty} C_N(z, w) = 0. \quad (67)$$

7.3.2. *The polydisk.* In complete analogy, consider the Lebesgue measure $\mu^{\mathbb{P}}$ on the polydisk $\mathbb{P} = \mathbb{D}^d$. Here again multivariate monomials are orthogonal, with the normalization constant

$$\|z^\alpha\|_{2, \mu^{\mathbb{P}}}^2 = \frac{\pi^d}{\prod_{j=1}^d (\alpha_j + 1)},$$

and hence

$$K_N^{\mu^{\mathbb{P}}}(z, w) = \frac{1}{\pi^d} \sum_{k=0}^n \sum_{|\alpha|=k} z^\alpha \bar{w}^\alpha \prod_{j=1}^d (\alpha_j + 1).$$

It is not difficult to check that, provided that $|w_j z_j| < 1$ for all j ,

$$\lim_{N \rightarrow \infty} K_N^{\mu^{\mathbb{P}}}(z, w) = \frac{1}{\pi^d} \prod_{j=1}^d \frac{1}{(1 - z_j \bar{w}_j)^2},$$

which is the well known Bergman space reproducing kernel for $\mathbb{P} = \mathbb{D}^d$. Also, the reader may check that, again, $K_N^{\mu^{\mathbb{P}}}(z, z)$ grows at most polynomially in N for $z \in \text{supp}(\mu^{\mathbb{D}^d})$. As for the exterior behavior, again the extremal plurisubharmonic function approach gives

$$\lim_{n \rightarrow \infty} K_N^{\mu^{\mathbb{P}}}(z, z)^{1/n} = \max_{1 \leq j \leq d} |z_j|^2, \quad z \notin \mathbb{P},$$

cf. formula (61).

Since the further analysis is a bit involved, we will restrict ourselves to the special case $d = 2$ and $z, w \notin \mathbb{P}$; that is, $\max(|z_1|, |z_2|) > 1$ and $\max(|w_1|, |w_2|) > 1$. If $|\bar{w}_1 z_1| < |\bar{w}_2 z_2|$ then

$$\begin{aligned} \pi^2 K_N^{\mu^{\mathbb{P}}}(z, w) &= \sum_{k=0}^n (\bar{w}_2 z_2)^k \sum_{j=0}^k (j+1)(k+1-j) \left(\frac{\bar{w}_1 z_1}{\bar{w}_2 z_2} \right)^j \\ &\sim \sum_{k=0}^n (k+1) (\bar{w}_2 z_2)^k \frac{(\bar{w}_2 z_2)^2}{(\bar{w}_2 z_2 - \bar{w}_1 z_1)^2} \\ &\sim \frac{n+1}{\bar{w}_2 z_2 - 1} \frac{(\bar{w}_2 z_2)^{n+3}}{(\bar{w}_2 z_2 - \bar{w}_1 z_1)^2}. \end{aligned}$$

The case $\bar{w}_1 z_1 = \bar{w}_2 z_2$ can be similarly treated. We conclude that $K_N^{\mu^{\mathbb{P}}}(z, z)$ grows outside the support at least like $\max\{|z_1|^{2n}, |z_2|^{2n}\}$ times some polynomial in n . An asymptotic analysis for $C_N^{\mu^{\mathbb{P}}}(z, w)$ is possible but quite involved and we omit the details.

8. INDEX OF NOTATION

We list here the meanings of many of the symbols that are frequently used throughout the paper.

- $K_n^\mu(z, w)$ the Christoffel-Darboux kernel consisting of $(n + 1)$ summands;
- $\text{tdeg } p$ the total degree of a multivariate polynomial p ;
- $\text{deg } p$ the degree of a multivariate orthogonal polynomial, indicating the position $n = \text{deg } p$ of p in a prescribed linear ordering of independent monomials;
- $\mathcal{S}(\mu)$ the Zariski closure of the support of a positive measure μ ;
- $\mathcal{N}(\mu)$ the ideal of polynomials vanishing on $\mathcal{S}(\mu)$;
- $v_n(z)$ the tautologic vector of monomials of degree less than or equal to n ;
- $\widetilde{v}_n^\mu(z)$ the tautologic vector of μ -orthogonal polynomials;
- $M_n(\mu)$ the matrix of moments of degree less than or equal to n , associated to a positive measure μ ;
- $\widetilde{M}_n(\mu)$ the reduced moment matrix, a maximal invertible submatrix of $M_n(\mu)$;
- $\Delta(z, \mu)$ the Mahalanobis distance between a point z and a measure μ ;
- $t_j K_n^{\widetilde{\nu}}(z^{(j)}, z^{(j)})$ the leverage score of a masspoint $z^{(j)}$ with mass t_j of a discrete measure $\widetilde{\nu}$;
- $\mathcal{L}_n(\mu)$ the linear span of orthogonal polynomials of degree less than or equal to n ;
- $C_n^\mu(z, w)$ the cosine function associated to the Christoffel-Darboux kernel;
- $\mathcal{C}(\mu)$ the covariance matrix of a random variable with law μ ;
- g_Ω the plurisubharmonic Green function of the unbounded open set Ω , with pole at infinity;
- Φ the normalized conformal mapping of the unbounded component of the complement of a Jordan curve in \mathbb{C} onto the exterior of the unit disk;
- T_n Chebyshev polynomial of the first kind.

REFERENCES

- [1] A. Ambroladze, On Exceptional Sets or Asymptotic relations for General Orthogonal Polynomials; *J. Approx. Th.* 82 (1995) 257-273.
- [2] T. Bayraktar: Zero distribution of random sparse polynomials, *Mich. Math. J.* 66(2017), 389-419.
- [3] B. Beckermann, On the numerical condition of polynomial bases: Estimates for the Condition Number of Vandermonde, Krylov and Hankel matrices, *Habilitationsschrift*, Universität Hannover (1996).
- [4] B. Beckermann and N. Stylianopoulos, Bergman orthogonal polynomials and the Grunsky matrix, *Constr. Approx.* 47(2018) 211-235.
- [5] T. Bloom, N. Levenberg, Weighted pluripotential theory in \mathbb{C}^d , *American Journal of Mathematics* 125(2003), 57-103.
- [6] L. Bos, N. Levenberg, Bernstein-Walsh theory associated to convex bodies and applications to multivariate approximation theory, *Comp. Meth. Funct. Theory* 18(2018), 361-388. (2017).
- [7] L. Bos, B. Della Vecchia and G. Mastroianni, On the asymptotics of Christoffel functions for centrally symmetric weights functions on the ball in \mathbb{R}^n , *Rendiconti del Circolo Matematico di Palermo* 52 (1998), 277-290.

- [8] A. Cohen, G. Migliorati, Optimal weighted least-squares methods. *SMAI J. Comput. Math.* 3 (2017), 181-203.
- [9] A.M. Delgado, L. Fernández, T.E. Pérez, M.A. Piñar, On the Uvarov modification of two variable orthogonal polynomials on the disk. *Complex Anal. Oper. Theory* 6 (3) (2012), 665-676.
- [10] A. Delgado, L. Fernandez, T. Perez, M. Pinar, Multivariate Orthogonal Polynomials and Modified Moment Functionals, arXiv:1601.07194 (2017).
- [11] C. F. Dunkl, Y. Xu, *Orthogonal Polynomials of Several Variables* (second edition), Cambridge Univ. Press, Cambridge, 2014.
- [12] Y. Xu, Orthogonal polynomials of several variables, arXiv:1701.02709.
- [13] G. Freud, *Orthogonale Polynome*, Birkhäuser, Basel, 1969.
- [14] D. Gaier, *Lectures on Complex Approximation*, Birkhäuser Boston Inc., Boston, MA, 1987.
- [15] W. Gautschi, *Orthogonal polynomials: computation and approximation*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2004.
- [16] G.H. Golub, Ch.F. Loan, *Matrix computations*, 4th edition, Johns Hopkins University Press (2013).
- [17] B. Gustafsson, M. Putinar, E. Saff, and N. Stylianopoulos, *Bergman polynomials on an archipelago: Estimates, zeros and shape reconstruction*, *Advances in Math.* **222** (2009), 1405–1460.
- [18] M. Jarnicki, P. Pflug, *First Steps in Several Complex Variables: Reinhardt Domains*, Europ. Math. Soc., Zürich, 2008.
- [19] J.B. Lasserre, E. Pauwels, Sorting out typicality with the inverse moment matrix SOS polynomial, Proceedings of the 30-th NIPS Conference, 2016.
- [20] J. B. Lasserre, E. Pauwels, The empirical Christoffel function in Statistics and Machine Learning, ArXiv 1701.02886v1,
- [21] M. Klimek, *Pluripotential Theory*, Clarendon Press, Oxford, 1991.
- [22] A. Kroó and D. S. Lubinsky, Christoffel functions and universality in the bulk for multivariate orthogonal polynomials, *Canadian J. Math.* 65(2013), 600-620.
- [23] A. Kroó and D. S. Lubinsky, Christoffel functions and universality on the boundary of the ball, *Acta Math. Hungarica* 140(2013), 117-133.
- [24] D. S. Lubinsky, A new approach to universality limits involving orthogonal polynomials, *Ann. of Math.* 170(2009), 915-939.
- [25] M. Lundin, The extremal psh for the complement of convex, symmetric subsets of \mathbb{R}^N , *Michigan Math. J.* 32(1985), 197- 201.
- [26] V. G. Malyskin, Multiple Instance Learning: Christoffel Function Approach to Distribution Regression Problem, ArXiv 1511.07085v1
- [27] V. G. Malyskin, Mathematical Foundations of Realtime Equity Trading. Liquidity Deficit and Market Dynamics. Automated Trading Machines, ArXiv 1510.05510v5
- [28] C. Martinez and M.A. Pinar, Orthogonal Polynomials on the Unit Ball and Fourth-Order Partial Differential Equations, *SIGMA* **12** (2016) 1-11.
- [29] P. Nevai, Géza Freud, orthogonal polynomials and Christoffel functions. A case study, *J. Approx. Theory*, 48 (1986), pp. 3-167.
- [30] E. Pauwels, F. Bach, J-P. Vert, Relating Leverage Scores and Density using Regularized Christoffel Functions, arXiv:1805.07943
- [31] E. Parzen, On the estimation of a probability density function and mode, *Annals of Mathematical Statistics* 33(1962), 1065-1076.
- [32] W. Plesniak, Siciak's extremal function in complex and real analysis, *Annales Polonici Mathematici* **80**(2003) 37-46.
- [33] E. B. Saff, *Orthogonal polynomials from a complex perspective*, Orthogonal polynomials (Columbus, OH, 1989), Kluwer Acad. Publ., Dordrecht, 1990, pp. 363–393.

- [34] ———, *Remarks on relative asymptotics for general orthogonal polynomials*, Recent trends in orthogonal polynomials and approximation theory, Contemp. Math., vol. 507, Amer. Math. Soc., Providence, RI, 2010, pp. 233–239.
- [35] E.B. Saff, V. Totik, *Logarithmic Potentials with External Fields*, Grundlehren der mathematischen Wissenschaften, vol. 316, Springer, Heidelberg (1997).
- [36] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral theory*, AMS (2009).
- [37] B. Simon, The Christoffel-Darboux kernel, in vol. "Perspectives in Partial Differential Equations, Harmonic Analysis and Applications: A Volume in Honor of Vladimir G. Mazya's 70th Birthday", D. Mitrea, M. Mitrea, eds., Proc. Symp. Pure Math. Amer. Math. Soc., Providence, R. I., 2008, pp. 314-355.
- [38] B. Simon, Weak convergence of CD kernels and applications, Duke Math. J. 146(2009), 305-330.
- [39] H. Stahl, V. Totik, *General Orthogonal Polynomials*, Cambridge Univ. Press, Cambridge, 1992.
- [40] P. K. Suetin, *Polynomials orthogonal over a region and Bieberbach polynomials*, Proceedings of the Steklov Institute of Mathematics, **100** (1971), Amer. Math. Soc., Providence, R.I., 1974.
- [41] V. Totik, *Asymptotics for Christoffel functions for general measures on the real line*, J. Anal. Math. **81** (2000), 283-303.
- [42] V. Totik, *Christoffel functions on curves and domains*, Trans. Amer. Math. Soc. **362** (2010), no. 4, 2053–2087.
- [43] K. C. Toh and L. N. Trefethen, The Kreiss matrix theorem on a general complex domain, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 145165.
- [44] V. N. Vapnik, *Statistical Learning Theory*, Wiley Interscience, New York, 1998.
- [45] Y. Xu, Asymptotics for orthogonal polynomials and Christoffel functions on a ball, Methods Appl. Analysis 3(1996), 257-272.

LABORATOIRE PAINLEVÉ UMR 8524, DEPARTMENT OF MATHEMATICS, UNIVERSITÉ DE LILLE, 59655 VILLENEUVE D'ASCQ, FRANCE

E-mail address: `Bernhard.Beckermann@univ-lille.fr`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT SANTA BARBARA, SANTA BARBARA, CALIFORNIA, 93106-3080, and SCHOOL OF MATHEMATICS AND STATISTICS, NEWCASTLE UNIVERSITY, NEWCASTLE UPON TYNE, NE1 7RU, UK

E-mail address: `mputinar@math.ucsb.edu`, `mihai.putinar@ncl.ac.uk`

CENTER FOR CONSTRUCTIVE APPROXIMATION, DEPARTMENT OF MATHEMATICS, VANDERBILT UNIVERSITY, 1326 STEVENSON CENTER, 37240 NASHVILLE, TN, USA

E-mail address: `edward.b.saff@vanderbilt.edu`

URL: <http://my.vanderbilt.edu/edsaff/>

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF CYPRUS, P.O. BOX 20537, 1678 NICOSIA, CYPRUS

E-mail address: `nikos@ucy.ac.cy`

URL: <http://ucy.ac.cy/~nikos>