Sampling of real multivariate polynomials and pluripotential theory

Robert J. Berman, Joaquim Ortega-Cerdà

American Journal of Mathematics, Volume 140, Number 3, June 2018, pp. 789-820 (Article)

Published by Johns Hopkins University Press
DOI: *https://doi.org/10.1353/ajm.2018.0019*

SAMPLING OF REAL MULTIVARIATE POLYNOMIALS AND PLURIPOTENTIAL THEORY

By ROBERT J. BERMAN and JOAQUIM ORTEGA-CERDÀ

Abstract. We consider the problem of stable sampling of multivariate real polynomials of large degree in a general framework where the polynomials are defined on an affine real algebraic variety M, equipped with a weighted measure. In particular, this framework contains the well-known setting of trigonometric polynomials (when M is a torus equipped with its invariant measure), where the limit of large degree corresponds to a high frequency limit, as well as the classical setting of one-variable orthogonal algebraic polynomials (when M is the real line equipped with a suitable measure), where the sampling nodes can be seen as generalizations of the zeros of the corresponding orthogonal polynomials. It is shown that a necessary condition for sampling, in the general setting, is that the asymptotic density of the sampling points is greater than the density of the corresponding weighted equilibrium measure of M, as defined in pluripotential theory. This result thus generalizes the well-known Landau type results for sampling on the torus, where the corresponding critical density corresponds to the Nyqvist rate, as well as the classical result saying that the zeros of orthogonal polynomials become equidistributed with respect to the logarithmic equilibrium measure, as the degree tends to infinity.

1. Introduction.

1.1. Background.

By the classical Whittaker-Shannon-Kotelnikov sampling theorem a band-limited signal f on the real line $\mathbb{R}$, normalized so that its frequency is in $[-1,1]$ may be recovered from its values at the points $t_j = j\pi$ where j ranges over the integers and

$$\int_{\mathbb{R}} \left|f(t)\right|^2 dt = \pi \sum_j \left|f(t_j)\right|^2.$$

In mathematical terms, f is in the Paley-Wiener space $PW_1(\mathbb{R})$ consisting of all functions in $L^2(\mathbb{R})$ whose Fourier transform is supported in $[-1,1]$. More generally, in the theory of *non-regular sampling* a sequence $\Lambda := \{\lambda\}_{\lambda\in\Lambda}$ of points on the real line $\mathbb{R}$ is said to be *sampling* for $PW_1(\mathbb{R})$ if there exists a constant C such that the following sampling inequality holds

$$\frac{1}{C}\int_{\mathbb{R}} \left|f(t)\right|^2 dt \le \sum_{\lambda\in\Lambda} \left|f(\lambda)\right|^2 \le C\int_{\mathbb{R}} \left|f(t)\right|^2 dt$$

Manuscript received November 14, 2015; revised November 2, 2016.

Research of the first author supported by grants from the European and Swedish Research Council and the Knut and Alice Wallenberg foundation; research of the second author supported by the Spanish Ministerio de Economía y Competividad (grants MTM2014-51834-P, MTM2017-83499-P, and MDM-2014-0445).

American Journal of Mathematics 140 (2018), 789–820.

for any $f \in PW_1(\mathbb{R})$, ensuring that the reconstruction of f is stable in the L^2-sense. Corresponding results also hold in the higher dimensional setting where $\mathbb{R}$ is replaced with $\mathbb{R}^n$ and the band $[-1,1]$ with the unit-cube $[-1,1]^n$ (or more generally any fixed convex body of volume one). By the seminal result of Landau [17], a necessary condition for a set Λ to be sampling is that the corresponding asymptotic density of points in $\mathbb{R}^n$ (in the sense of Beurling) is at least equal to the Nyqvist rate $1/\pi^n$, i.e.,

$$\liminf_{R\to\infty} \frac{\#\{\Lambda \cap R\Omega\}}{R^n} \geq \int_\Omega \frac{1}{\pi^n} dt$$

(uniformly over translations) for any smooth domain $\Omega \subset \mathbb{R}^n$ assuming a uniform separation lower bound on the points in Λ. In one-dimension the reversed strict inequality is also a sufficient condition for sampling, but not in higher dimensions.

By a rescaling, Landau's density results may also be formulated in terms of the high frequency limit which appears when the frequency domain $[-1,1]^n$ is replaced with $[-k,k]^n$ for k large, i.e., $PW_1(\mathbb{R}^n)$ is replaced with the corresponding Paley-Wiener space $PW_k(\mathbb{R}^n)$. In this context a sequence $\Lambda_k := \{\lambda^{(k)}\}$ of sets of points on $\mathbb{R}^n$ is said to be sampling for $PW_k(\mathbb{R}^n)$ if

$$\frac{1}{C}\int_{\mathbb{R}^n} \left|f(t)\right|^2 dt \leq \frac{1}{k^n} \sum_{\lambda^{(k)}\in\Lambda_k} \left|f\left(\lambda^{(k)}\right)\right|^2 \leq C\int_{\mathbb{R}^n} \left|f(t)\right|^2 dt$$

for any $f \in PW_k(\mathbb{R}^n)$ with the constant C independent of k. For the sake of simplicity if it is clear from the context we will omit the superindex k in $\lambda^{(k)}$ and write simply $\lambda \in \Lambda_k$. The corresponding necessary density condition on the sampling points may then be reformulated as

$$\liminf_{k\to\infty} \frac{\#\{\Lambda_k \cap \Omega\}}{k^n} \geq \int_\Omega \frac{1}{\pi^n} dt$$

uniformly over translations for any domain $\Omega \subset \mathbb{R}^n$ with $|\partial\Omega| = 0$. (Landau's setting corresponds to the case when Λ_k is of the form $k^{-1}\Lambda$, but his arguments extend to this high-frequency setting).

There is also a natural *compact* analogue of the Paley-Wiener setting on $\mathbb{R}^n$ which is the one which is most relevant for the present paper, where $\mathbb{R}$ (or $\mathbb{R}^n$) is replaced with the circle $S^1 := \mathbb{R}/2\pi$ (or the n-dimensional torus). Then the role of $PW_k(\mathbb{R})$ is played by the space $H_k(S^1)$ of all finite Fourier series on $[0,2\pi]$ with frequencies in $[-k,k]$, i.e., the space of all trigonometric polynomials of degree at most k. A sampling sequence of finite sets of points $\Lambda_k \subset S^1$ in this setting is also called a Marcinkiewicz-Zygmund family [24].

From an abstract point of view the previous settings fit into a general Hilbert space framework where $H_k(M)$ is a given sequence of Hilbert spaces of functions on a set M with reproducing kernels $K_k(x,y)$. Then a sequence Λ_k of sets of points on M is said to be *sampling for* $H_k(M)$ if the family of normalized functions $\kappa_\lambda := K_k(\cdot,\lambda)/\|K_k(\cdot,\lambda)\|$ for $\lambda \in \Lambda_k$, form a *frame* in the Hilbert space $H_k(M)$, in the sense of Duffin-Schaeffer [13], i.e.:

$$\frac{1}{C}\|f\|^2 \le \sum_{\lambda\in\Lambda_k} \left|\langle f, \kappa_\lambda\rangle\right|^2 \le C\|f\|^2, \quad \forall f \in H_k(M),$$

which is equivalent to the sampling inequalities:

$$\frac{1}{C}\|f\|^2 \le \sum_{\lambda\in\Lambda_k} \frac{\left|f(\lambda)\right|^2}{K_k(\lambda,\lambda)} \le C\|f\|^2, \quad \forall f \in H_k(M), \tag{1.1}$$

where we will assume that C can be taken to be independent of k.

1.2. The present setting. The main aim of the present paper is to generalize the Landau type necessary density conditions for sampling on S^1 to a general setting where the Hilbert space $H_k(M)$ consist of polynomials of degree at most k on an affine real algebraic variety M equipped with a weighted measure. We are not dealing with the very interesting problem of finding sufficient conditions for sampling multivariate polynomials. This and its numerical implementation is a very basic question in signal analysis, see for instance [14] and the references therein for the one-variable numerical sampling.

Our setting is the following: by definition M is the variety cut out by a finite numbers of polynomials on $\mathbb{R}^m$ and $H_k(M)$ is the space of polynomials of total degree at most k restricted to M and equipped with the L^2 norm

$$\left\|p_k\right\|^2_{L^2(e^{-k\phi}\mu)} := \int_M \left|p_k\right|^2 e^{-k\phi} d\mu$$

defined by a compactly supported measure μ on M and a continuous function ϕ on M (referred to as the weight function). Following [2] we will refer to the pair (μ,ϕ) as a "weighted measure". In order that the latter norm be non-degenerate some regularity assumption has to be made on μ. The affine case, i.e., when $M = \mathbb{R}^m$ is thus the classical setting for multivariate orthogonal polynomials. We will assume two regularity conditions: the Bernstein-Markov property and moderate growth, see Section 2.1 for the precise definitions.

Our first main result in this general setting is:

THEOREM 1. *Let M be an affine real algebraic variety equipped with a non-degenerate measure μ and a weight function ϕ. Assume that the pair (μ, ϕ) satisfies the Bernstein-Markov property* (2.1) *and it is of moderate growth* (2.2). *Then a necessary condition for a sequence Λ_k of sets of points in M to be sampling for the space $H_k(M)$ of polynomials of degree at most k, with respect to the weight*

$k\phi$ and measure μ, is that

$$\liminf_{k\to\infty} \frac{1}{N_k} \sum_{\lambda\in\Lambda_k} \delta_\lambda \geq \mu_{eq} \tag{1.2}$$

in the weak topology on the measures on M, where μ_{eq} denotes the normalized equilibrium measure of the weighted measure (μ,ϕ) and $N_k = \dim(H_k(M))$.

1.3. Sampling on compact real algebraic varieties equipped with a volume form. One disadvantage of the definition of the sampling inequalities (1.1) in this general setting is that it is of a rather abstract nature as it involves the reproducing kernel $K_k(x,x)$ which in general is impossible to compute explicitly. On the other hand, only the asymptotic behavior of $K_k(x,x)$ as $k\to\infty$ is needed and these asymptotics can often be estimated. Also, if μ_{eq} is absolutely continuous with respect to Lebesgue measure than the condition (1.2) above may be written as

$$\liminf_{k\to\infty} \frac{\#\{\Lambda_k\cap\Omega\}}{\# N_k} \geq \frac{\mu_{eq}(\Omega)}{\mu_{eq}(M)} \tag{1.3}$$

for any smooth domain Ω. We will refer to the latter condition as the "pluripotential Nyqvist bound". One particularly interesting case where Theorem 1 can be made explicit is the following:

THEOREM 2. *Let M be an n-dimensional affine real algebraic variety, which is non-singular and compact, let μ be a volume form on M and let $\phi = 0$. Then there exists a positive constant C such that the reproducing kernel for $(H_k(M),\mu)$ satisfies*

$$\frac{1}{C}k^n \leq K_k(x,x) \leq Ck^n \tag{1.4}$$

and thus (μ,ϕ) is non-degenerate, it satisfies the Bernstein-Markov property and it is of moderate growth. Moreover, a necessary condition for a sequence Λ_k of sets of points on X to be sampling for $H_k(M)$ is that the density of sampling points is at least equal to the density of the equilibrium measure μ_{eq} of M, as $k\to\infty$, i.e., the pluripotential Nyquist bound (1.3) *holds.*

The definition of the equilibrium measure of M and more generally the equilibrium measure attached to a weighted measure will be recalled in Section 2.2. As pointed out above this result thus generalizes the results in [24] concerning the case when M is the unit-circle. Moreover, the case when M is the unit-sphere corresponds to the case studied in [20], where the signals in questions are spherical harmonics. While in all these special cases the equilibrium measure μ_{eq} is explicitly given by the Haar measure (since the corresponding Riemannian manifolds are homogeneous) the equilibrium measure of a general real affine algebraic variety appears to be of a highly non-explicit nature. Another generalization of the homogeneous cases was considered in [23], where the signals are "band-limited" sums of eigenfunctions of the Laplacian on a given compact Riemannian manifolds (M,g) and then the role of the equilibrium measure is played by the Riemannian volume form. In another direction sampling sequences have previously been studied in the complex geometric setting of a positively curved line bundle over a projective complex algebraic manifold in [5] (where the corresponding equilibrium measure is explicitly given by the volume form induced by the curvature form). Our setting can thus be viewed as a real analog of the latter setting. As will be explained below the main new challenge that appears in our setting comes from the absence of point-wise and precise decay estimates for the corresponding Bergman kernels.

It is not evident a priori that there are sampling sequences at all. This is assured with the following Bernstein type theorem:

THEOREM 3. *Given a smooth compact real manifold $M \subset \mathbb{R}^m$ of dimension n, the following are equivalent:*

- *M is algebraic.*
- *M satisfies a Bernstein inequality, i.e., for some $q \geq 1$ (or for all $q \geq 1$):*

$$\left\|\nabla_t p\right\|_{L^q(M)} \leq C_q \deg(p)\|p\|_{L^q(M)}.$$

- *There is a uniformly separated Λ_k such that for some (all) $q \geq 1$*

$$\int_M |p|^q dV_M \lesssim \frac{1}{k^n}\sum_{\lambda\in\Lambda_k}\left|p(\lambda)\right|^q \lesssim \int_M |p|^q dV_M, \quad \forall p \in \mathcal{P}_k\left(\mathbb{R}^m\right).$$

This generalizes the main result of [9] where the case $q = \infty$ was considered.

1.4. Sampling of multivariate real polynomials on convex domains. Another instance where Theorem 1 can be made more precise is the case were $M = \mathbb{R}^n$, μ is the Lebesgue measure restricted to a smooth bounded convex domain Ω and $\phi = 0$. In this case the equilibrium measure is very well understood, see [1, 9]. It behaves roughly as $d\mu_{eq} \simeq 1/\sqrt{d(x,\partial\Omega)}dV$, (this will also follow from the asymptotics (1.5) below).

THEOREM 4. *Let Ω be a smoothly bounded convex domain in $\mathbb{R}^n$. Then the reproducing kernel for $(H_k(\Omega), dV)$ satisfies*

$$B_k(x) = K_k(x,x) \simeq \min\left(\frac{k^n}{\sqrt{d(x)}}, k^{n+1}\right) \quad \forall x \in \Omega, \tag{1.5}$$

where $d(x)$ denotes the distance of $x \in \Omega$ to the boundary of Ω. Thus it satisfies the Bernstein-Markov property (2.1) *and it is of moderate growth* (2.2). *Moreover, a necessary condition for the sequence Λ_k of sets of points on Ω to be sampling* *for $H_k(M)$ is that the density of sampling points is at least equal to the density of the equilibrium measure μ_{eq} of Ω, as $k \to \infty$, i.e., the pluripotential Nyquist bound* (1.3) *holds.*

1.5. Interpolating sequences. A natural companion problem to that of sampling sequences are the interpolating sequences. In the same abstract point of view that we considered for sampling sequences we consider a sequence $H_k(M)$ of Hilbert spaces of functions on a set M with reproducing kernels $K_k(x,y)$ and instead of a frames we consider Riesz sequences of normalized reproducing kernels:

Definition 1. A sequence Λ_k of sets of points on M is said to be *interpolating* for $H_k(M)$ if the family of normalized reproducing kernels

$$\kappa_\lambda := K_k(\cdot,\lambda)/\|K_k(\cdot,\lambda)\|$$

for $\lambda \in \Lambda_k$, is a Riesz sequence in the Hilbert space $H_k(M)$, i.e.:

$$\frac{1}{C}\sum_{\lambda\in\Lambda_k}|c_\lambda|^2 \le \Big\|\sum_{\lambda\in\Lambda_k} c_\lambda \kappa_\lambda\Big\|^2 \le C\sum_{\lambda\in\Lambda_k}|c_\lambda|^2, \quad \forall\{c_\lambda\}_{\lambda\in\Lambda_k}\in \ell_2$$

where we will assume that C can be taken independent of k.

Landau in [17] studied also these sequences in the Paley-Wiener space, and his observation was that locally if a sequence Λ is interpolating then its density should be smaller than the local density of the space. We can again use the ideas inspired in [22] to deal with the case of polynomials in real algebraic varieties.

Our main result to this problem is:

THEOREM 5. *Let M be an affine real algebraic variety equipped with the Lebesgue measure. Then a necessary condition for a sequence Λ_k of points on M to be interpolating for the space $H_k(M)$ of polynomials of degree at most k, is that*

$$\limsup_{k\to\infty}\frac{1}{N_k}\sum_{\lambda\in\Lambda_k}\delta_\lambda \le \mu_{eq} \tag{1.6}$$

in the weak topology on M, where μ_{eq} denotes the normalized equilibrium measure of M and $N_k = \dim(H_k(M))$, i.e., the following reversed pluripotential Nyqvist bound holds:

$$\limsup_{k\to\infty}\frac{\#\{\Lambda_k\cap\Omega\}}{\# N_k} \le \frac{\mu_{eq}(\Omega)}{\mu_{eq}(M)}$$

for any given smooth domain Ω in M.

1.6. Discussion of the proof of Theorem 2. Let us make some brief comments on the circle of ideas involved in the proof of the previous theorems. First of all, since the sampling points uniquely determine a polynomial p_k on M the total number $\#\Lambda_k$ of sampling points at level k is of course at least equal to the dimension N_k of $H_k(M)$. As a well-known guiding principle the necessary conditions for sampling should come from a localized version of this argument saying the asymptotic lower density of sampling points should at least be given by the "local dimension" of the Hilbert space $(H_k(M), \|\cdot\|_{L^2(d\mu_k)})$, were $\mu_k := e^{-k\phi} d\mu$, i.e., by the leading asymptotics of N_k^{-1} times the function

$$B_k(x) := \sum_{i=1}^{N_k} \left| p_i^{(k)}(x) \right|^2 e^{-k\phi(x)}$$

where $\{p_i^{(k)}(x)\}$ is any orthonormal base in the Hilbert $(H_k(M), \|\cdot\|_{L^2(d\mu_k)})$ (note that integrating $B_k(x)$ with respect to dV indeed gives the dimension N_k of $H_k(M)$). The independence of the choice of base follows from the following extremal representation of $B_k(x)$:

$$B_k(x) := \sup_{p_k \in H_k(M)} \frac{\left| p_k(x) \right|^2 e^{-k\phi(x)}}{\int_M \left| p_k \right|^2 d\mu_k}. \tag{1.7}$$

From the point of view of general Hilbert space theory $B_k(x)$ may be written as $B_k(x) = K_k(x,x)$ where $K_k(x,y)$ is the reproducing kernel of the Hilbert space $(H_k(M), \|\cdot\|_{L^2(d\mu_k)})$, i.e., the kernel of the orthogonal projection from $L^2(d\mu_k)$ to $H_k(M)$. By the general results in [2]

$$N_k^{-1} B_k(x) d\mu_k \longrightarrow \mu_{eq}/\mu_{eq}(M) \tag{1.8}$$

weakly on M as $k \to \infty$, which in the view of the guiding principle above thus gives a strong motivation for the previous theorems. However, this guiding principle does not seem to hold in all generality and it has to be complemented with some further asymptotic information of the full reproducing kernel $K_k(x,y)$. This is already clear from Landau's classical proof in the Paley-Wiener setting on $\mathbb{R}^n$ [17], where the decay asymptotics of $K_k(x,y)$, away from the diagonal are needed in order to construct functions f_k which are well localized on a given domain Ω (using suitable Toeplitz operators). Moreover, Landau's approach also relies on certain submean inequalities for f_k which pose difficulties in our general setting. Instead we use a new approach to proving necessary conditions for sampling, which is inspired by [18, 22], where we reduce the problem to establishing two asymptotic properties of $K_k(x,y)$ (given the convergence of the Bergman function to the equilibrium measure):

- A growth property of B_k.
- A weak decay property of $|K_k(x,y)|$ away from the diagonal.

The interpolation theorem 5 is proved in a similar way, but by replacing the moderate growth property with a Bernstein type inequality, see Theorem 9.

One interesting feature is that although the statement of the problems studied are purely real, all the proofs rely on the process of complexification with one important exception: the off-diagonal estimate on the Bergman kernel (Theorem 17) which exploits, in an essential way, the real structure. Somewhat remarkably, the off-diagonal estimate for the Bergman kernel that we get (which is sharp in terms of powers of the distance function) matches precisely the decay needed in our new approach to sampling (and interpolation).

1.7. Further relations to previous results. As explained above one important ingredient in the proof of Theorem 1 is the asymptotics for $B_k(x)$ in formula (1.8) established in [2], which holds generally under the Bernstein-Markov assumption on (μ, ϕ). In turn, the latter result can be seen as a consequence of a very general result in [2] giving the convergence towards the equilibrium measure of (μ, ϕ) for the normalized Dirac measure associated to a sequence of N_k points under the condition that the points are asymptotic Fekete points for the weighted set $(\mathrm{Supp}(\mu), \phi)$. It may thus be tempting to try to deduce Theorem 1 in the present paper directly from the general convergence results in [2] by the following tentative procedure: one removes points from a sampling sequence Λ_k until one arrives at N_k points, while keeping the sampling property. But as shown by a counter example in [24, Example 2] such a procedure is doomed to fail, already in the homogenouous case of the circle.

The point-wise asymptotics for $B_k(x)$ in Theorem 2 (formula (1.4)) can be seen as an improvement—in the special case of a real algebraic manifold—of a classical result for regular compact subsets complex space going back to Siciak and Zaharyuta giving that $B_k(x)^{1/k} \to 1$ point-wise on M (this latter classical result was given a $\bar{\partial}-$proof by Demailly, [12] which can seen as a precursor to our proof). In view of the weak asymptotics (1.8) and the bounds in formula (2.3) below, it seems natural to conjecture that $B_k(x)/k^n$ in fact converges pointwise (in the almost everywhere sense) to the L^1-density of the equilibrium measure of M (and similarly in the setting of a convex domain; as in the one dimensional setting, see [28]). This would be a real analog of the point-wise asymptotics for $B_k(x)$ in [4], where the role of M is played by a complex projective variety endowed with a hermitian holomorphic line bundle (the case of positive curvature is a fundamental result in complex geometry, due to Bouche [7] and Tian [27]).

On the other hand, in the line bundle setting, the asymptotics for $|K_k(x,y)|^2$ in Theorem 17 are only known in the case of a line bundle with positive curvature, but it seems natural to expect that they hold in general (see [4] for some results in this direction in connection to the study of fluctuations of linear statistics of determinantal point processes).

Finally, we recall that in the one-dimensional case there is a vast literature on various asymptotic results for orthogonal polynomials, in particular in connection to random matrix theory. For example, the asymptotics of the scaled reproducing kernel $k^{-n}K_k(x+\frac{a}{k}, x+\frac{b}{k})$ have been established, in connection to the question of universality, under very general condition on a given measure μ on the real line, when x is a fixed point in the "bulk" of the support of μ; see for example [19] (and similar scaling result holds at the "edge" of the support). However, there seem to be very few results in the higher dimensional setting (but see [29, 16], where the case of the ball and the simplex in is settled). It would be interesting to extend the asymptotics in the present paper to study similar universality questions in higher dimensions and we leave this as a challenging problem for the future.

Acknowledgments. The authors are grateful to Ahmed Zeriahi for very useful discussions regarding the Bernstein inequality.

2. Pluripotential theory and asymptotics of real orthogonal polynomials.

2.1. Setup. Let M be an n-dimensional affine real algebraic variety, which is non-singular and compact. In particular, M is the common zero-locus of a collection of real polynomials $p_1, \ldots, p_r$ in $\mathbb{R}^m$. We denote by $H_k(M)$ the real vector space consisting of the functions on M which are restrictions of real polynomials in $\mathbb{R}^m$ of total degree at most k. We will also consider the "complexifications" X and $H_k(X)$ of the real variety M and the real vector space $H_k(M)$, respectively. More precisely, X is the complex algebraic variety in $\mathbb{C}^m$ defined by the common complex zeros of the ideal defining M and $H_k(X)$ is the complex vector space consisting of restrictions to X of polynomials in $\mathbb{C}^m$ of total degree at most k. Then M is indeed the real part of X in the sense that it consists of all points in z in X such that $\bar{z} = z$ and real vector space $H_k(M)$ is the real part of $H_k(X)$ in the sense that it consists of all p_k in $H_k(X)$ such that $\overline{p_k} = p_k$ (restricted to M).

We will denote by $K_k(x,y)$ the Bergman reproducing kernel of H_k equipped with the L^2-norm induced by a given weighed measure (μ, ϕ) and moreover we will use the notation $B_k(x) = K_k(x,x)e^{-k\phi(x)}$.

Example 6. Let M be the unit-circle realized as the zero-set in $\mathbb{R}^2$ of $p(x,y) = x^2 + y^2 - 1$. Setting $x = \cos\theta$ for $\theta \in [0, 2\pi]$ we may identify $H_k(M)$ with the space $H_k([0,2\pi])$ of all Fourier series on $[0,2\pi]$ "band-limited" to $[-k,k]$, i.e., spanned by 1 and $\cos m\theta$ and $\sin m\theta$ for $m \in [1,k] \cap \mathbb{Z}$. More precisely, $H_k([0,2\pi])$ is the pull-back of $H_k(M)$ under the corresponding map from $[0,2\pi]$ to M. To see the relation to the more standard setting where M corresponds to the unit-circle S^1 in $\mathbb{C}$ with complex coordinate τ (equal to $e^{i\theta}$ on S^1) we note that the embedding F of $\mathbb{C}^*$ in $\mathbb{C}^2$ given by

$$z := \left(\tau + \tau^{-1}\right)/2 \quad w = \left(\tau - \tau^{-1}\right)/2i$$

maps $\mathbb{C}^*$ to the complex quadric X cut out by $p(z,w) := z^2 + w^2 - 1$ and the unit-circle S^1 in $\mathbb{C}^*$ is mapped to the real part M of X. Indeed, $\tau \in S^1$ if and only if

$\bar{\tau} = \tau^{-1}$ iff $z(\tau) = \Re\tau (= \cos\theta)$ and $w(\tau) = \Im\tau (= \sin\theta)$. The pull-back of $H_k(X)$ under F is the space of Laurent polynomials on $\mathbb{C}$ spanned by the monomials τ^m for $m \in [-k, k] \cap \mathbb{Z}$. Hence, the real part of $F^* H_k(M)$ is indeed spanned by 1 and $\cos m\theta (= \Re\tau^m)$ and $\sin m\theta (= \Im\tau^m)$ for $m \in [1, k] \cap \mathbb{Z}$.

Definition 2. (Bernstein-Markov) The standard assumption on the pair (μ, ϕ) is that it satisfies the Bernstein-Markov property (with respect to the support of μ), which may be formulated as the property that the reproducing kernel $K_k(x, y)$ have sub-exponential growth on the diagonal, i.e., for any $\epsilon > 0$ there exists a positive constant C_ϵ such that

$$B_k(x) \leq C_\epsilon e^{\epsilon k} \tag{2.1}$$

uniformly on the support of μ.

For our general results to hold we will also need the following technical regularity of the growth assumption on the reproducing kernel along the diagonal:

Definition 3. (Moderate growth) We say that H_k has a reproducing kernel with moderate growth if

$$K_{k+1}(x, x) \leq C K_k(x, x) \tag{2.2}$$

on the support of μ. More precisely, for our purposes the constant C may be replaced by any sequence with growth of the order $o(k)$.

Anyway, in all examples that we are aware of the constant C in (2.2) may actually be replaced by a sequence tending to one as $k \to \infty$ (which, by iteration, actually implies the Bernstein-Markov property (2.1)). All the measures μ supported in M that we will consider are *non-degenerate*, in the sense that $\|f\| \neq 0$ if $0 \neq f \in H_k(M)$. Otherwise if the support of μ is contained in the zeros of a non-vanishing polynomial in $H_k(M)$ one may replace M with a subvariety.

While Definition 2 is standard (see [2] and references therein), Definition 3 appears to be new. We expect it to hold in great generality and we will establish it in the situations relevant to the present paper.

2.2. The extremal function attached to a real affine variety. Recall that the *Lelong class* $\mathcal{L}(\mathbb{C}^m)$ is the convex space of all plurisubharmonic (psh, for short) functions ϕ on $\mathbb{C}^m$ with logarithmic growth, in the sense that $\phi \leq \log(1 + |z|^2) + C$. The restriction of this space to X will be denoted by $\mathcal{L}(X)$ and it may be identified with the space of all (singular) metrics on the line bundle $\mathcal{O}(1)_{\bar{X}} \to \bar{X}$ with positive curvature current, see [2] and references therein for further background. The *Siciak extremal function* (sometimes called the equilibrium potential) of a compact and non-pluripolar subset K of X and a weight $\psi \in \mathbb{C}(K)$ is the function in $\mathcal{L}(X)$

defined as the upper semi-continuous regularization v_K^* of the envelope

$$v_{K,\psi}(x) = \sup_{\phi \in \mathcal{L}(X)} \left\{ \phi(x) : \phi \le \psi \text{ on } K \right\}$$

and the weighted set (K, ψ) is called *regular* if $v_{K,\psi}^* = v_{K,\psi}$. The Monge-Ampère measure

$$\mu_{K,\psi} := MA\big(v_{K,\psi}\big)$$

is called the *(pluripotential) equilibrium measure* of (K, ψ) and it is supported on K (when there is no risk of confusion we will write $\mu_{K,\psi} := \mu_{eq}$).

In the following we will take $K := M$ as above, which is thus embedded in the complex affine variety X in such a way that $M = X \cap \{y = 0\}$, where y denotes the imaginary part of $z \in \mathbb{C}^n$. We will also take $\psi = 0$.

PROPOSITION 7. *Let M be an n-dimensional affine real algebraic variety, which is non-singular and compact and denote by v_M its extremal function, defined on the complexification X of M. Then M is non-pluripolar and regular. Moreover, there exists a constant C such that*

$$\frac{1}{C}|y| \le v_M \le C|y| \tag{2.3}$$

in a neighborhood of M in X. In particular, the equilibrium measure μ_M is absolutely continuous with respect to the Lebesgue measure dV_M on M and its density is bounded from above and below by positive constants:

$$\frac{1}{D} dV_M \le \mu_M \le D dV_M \tag{2.4}$$

on M.

Proof. The lower bound on v_K follows from a simple max construction. Indeed, we may after a scaling assume that $|z| < 1$ on M and then set

$$\psi := \begin{cases} \max\left\{|y|/C, \log|z|^2\right\} & \text{if } |z| \le 2 \\ \log|z|^2 & \text{if } |z| > 2 \end{cases}$$

with C sufficiently large to ensure that ψ is continuous along $|z| = 2$. Since, $\psi = |y|/C$ close to M the function ψ is a contender for the sup defining v_M and hence $\psi \le v_M$, which proves the lower bound in (2.3). The proof of the upper bound is more involved: We know that $v_M \in L^\infty_{\text{loc}}(X)$ because M is algebraic, see [25]. Moreover there is a distance d such that $g(r) = \sup_{z \in X : d(z,M) = r} v_M(z)$ is convex in r in $[0, \delta]$, see the proof of Theorem 10 for details. Therefore $g(r) \le Cr$ and the upper bound in (2.3) follows.

Anyway, for the proof of the lower bound in (1.4) we will only need the lower bound in (2.3). Note also that combining (1.4) and the asymptotics (1.8) immediately gives the inequalities (2.4) which also follow from the inequalities (2.3) by the comparison principle for the Monge-Ampère measure, see [1, Lemma 2.1]. □

2.3. The proof of the lower bound on B_k in (1.4). Denote by B_{kv_M} the Bergman function on X defined by the L^2-norm on $H_k(X)$ induced by the measure $e^{-kv_M}dV_X$. The idea of the proof is to first show that

$$B_{kv_M} \geq Ck^{2n} \quad \text{on } M, \tag{i}$$

and then that

$$B_k \geq Ck^{-n}B_{kv_M} \quad \text{on } M. \tag{ii}$$

This would clearly imply the result in question. However, for technical reasons we will only show a slightly weaker version of these inequalities (needed for (i)) where v_M is replaced by

$$v_M^\epsilon := v_M(1-\epsilon)+\epsilon\psi$$

where ψ is a continuous function in $\mathcal{L}(X)$ such that $\psi = v_M^2/C$ in a neighborhood of M. Here ϵ is a sufficiently small positive number which is fixed once and for all. To see that such a function ψ exists we may after scaling assume that $|z| < 1$ on M and then simply set $\psi := \max\{v_M^2/C, \log|z|^2\}$ when $|z| < 2$ and $\phi := \log|z|^2$ when $|z| > 2$. The constant C is taken sufficiently large to ensure that ψ is continuous at $|z| = 2$.

Let us start with the proof of (i). To simplify the notation we will assume that $n = 1$ (but the general proof is essentially the same). To this end fix a point in M and introduce local holomorphic coordinates z on U in X, centered a the fixed point, such that $M = \{y = 0\}$ locally, i.e., on U (not to be confused with the global coordinates on $\mathbb{C}^m$ and $\mathbb{R}^n$, respectively). The idea is to first construct a local function f_k, holomorphic on U such that

$$\frac{|f_k(0)|^2}{\int_U |f_k|^2 e^{-kv_M} dV_X} \geq k^2/C \tag{2.5}$$

and then perturb f_k slightly to become a polynomial p_k by solving a global $\bar{\partial}$-equation on $\bar{X}$ with an L^2-estimate.

There is no loss of generality assuming that $f_k(0) = 1$. Working in a local coordinates and reescaling (2.5) it is enough to prove that there is a function $f \in \mathcal{H}(\mathbb{C})$ such that $f(0) = 1$ and

$$\int_{\mathbb{C}} |f(z)|^2 e^{-C|\Im z|} < \infty,$$

then $f_k(z) = f(kz)$ will satisfy (2.5). The function $f(z) = \operatorname{sinc}^2(Cz/2)$ has the desired properties.

2.3.1. Modification and globalization. Let now χ be a smooth cut-off function supported on U (say equal to one on $U/2$). In view of standard globalization arguments the problem with the present setting is that $\bar{\partial}$ of the global function χf_k on X does not have a small weighted L^2-norm (compared to the weighted norm of f_k). The reason is that $e^{-k2|y|}$ is only well localized in the y-direction. To bypass this difficulty we will instead replace v_M with v_M^ϵ and modify f_k accordingly as follows. First observe that, by definition,

$$\int_U |g_k|^2 e^{-kv_M^\epsilon} dV_X \leq \int_U |g_k|^2 e^{-kv_M(1-\epsilon)} e^{-k\epsilon 4|y|^2} dV_X$$

for any g_k. We next observe that $|e^{-2z^2}|e^{-4|y|^2} = e^{-2|z|^2}$ and hence setting $g_k := f_k e^{-k\epsilon z^2}$ gives

$$|g_k|^2 e^{-k4\epsilon|y|^2} = |f_k|^2 e^{-2k\epsilon|z|^2} \leq |f_k|^2$$

and $g_k(0) = f_k(0)$. In particular,

$$\frac{|g_k(0)|^2}{\int_U |g_k|^2 e^{-kv_M^\epsilon} dV_X} \geq \frac{|f_k(0)|^2}{\int_U |f_k|^2 e^{-k(1-\epsilon)v_M} dV_X} \geq k^2/C_\epsilon.$$

Here the optimal constant C_ϵ is slightly smaller than the previous optimal C, but on the other hand we have gained a Gaussian factor that we will next exploit. The point is that $\bar{\partial}(\chi g_k) = \bar{\partial}\chi g_k$ is supported where $|z| > 1/4$ and hence

$$\begin{aligned}\int_U |\bar{\partial}(\chi g_k)|^2 e^{-kv_M^\epsilon} dV_X &\leq C \int_{1/4\leq|z|\leq 2} |g_k|^2 e^{-kv_M(1-\epsilon)} e^{-k\epsilon 4|y|^2} dV_X \\ &= C \int_{1/4\leq|z|\leq 2} |f_k|^2 e^{-kv_M(1-\epsilon)} e^{-k\epsilon 2|z|^2} dV_X.\end{aligned}$$

Estimating the Gaussian factor $e^{-k\epsilon 2|z|^2}$ with its sup, i.e., with $e^{-k2\epsilon/4^2}$ thus gives the bound

$$\int_U |\bar{\partial}(\chi g_k)|^2 e^{-kv_M^\epsilon} dV_X \leq O(e^{-\delta k}) \int_U |f_k|^2 e^{-kv_M(1-\epsilon)} dV_X.$$

Here and henceforth $O(e^{-\delta k})$ denotes a term which is exponentially small in k (recall that ϵ is a small number which is fixed once and for all).

With this local estimate in place we can now apply a standard globalization argument: using L^2-estimates for $\bar{\partial}$ on the line bundle $\mathcal{O}(1)_X$ over $\bar{X}$, or more precisely (if the latter variety is singular) on its pull-back to a smooth resolution of

$\bar{X}$ there exists a smooth function u_k on such that $p_k := g_k - u_k$ is in $H_k(X)$ and

$$\bar{\partial} u_k = \bar{\partial}(\chi g_k), \quad \int_X |u_k|^2 e^{-kv_M^\epsilon} dV_X \leq C \int_U \left|\bar{\partial}(\chi g_k)\right|^2 e^{-kv_M^\epsilon} dV_X$$

(strictly speaking to apply L^2-estimates we have to slightly modify the weight v_M^ϵ with a k-independent term to ensure that the corresponding metric on the line bundle $k\mathcal{O}(1)_X$ has a sufficiently large uniform lower bound on its curvature form, but this only changes the L^2-estimates with an overall multiplicative constant, which is harmless). This is a standard procedure; for a precise statement which also applies in the singular setting see, for example [3, Section 2].

By the previous estimate this means that

$$\int_X \left|u_k\right|^2 e^{-kv_M^\epsilon} dV_X \leq O\left(e^{-\delta k}\right) \int_U \left|f_k\right|^2 e^{-kv_M(1-\epsilon)} dV_X.$$

Moreover, applying the mean value property for holomorphic functions on a small coordinate ball then gives

$$\left|u_k(0)\right|^2 \leq Ck^2 \int_X \left|u_k\right|^2 e^{-kv_M^\epsilon} dV_X \leq O\left(e^{-\delta k}\right) \int_U \left|f_k\right|^2 e^{-kv_M(1-\epsilon)} dV_X.$$

Hence,

$$\begin{aligned} \frac{\left|p_k(0)\right|^2}{\int_X \left|p_k\right|^2 e^{-kv_M^\epsilon} dV_X} &= \frac{\left|g_k(0) - u_k(0)\right|^2}{\int_X \left|\chi g_k - u_k\right|^2 e^{-kv_M^\epsilon} dV_X} \\ &\geq \frac{\left|g_k(0)\right|^2 - O\left(e^{-\delta k}\right) \int_U \left|f_k\right|^2 e^{-kv_M(1-\epsilon)}}{\int_U \left|\chi g_k\right|^2 e^{-kv_M^\epsilon} dV_X + O\left(e^{-\delta k}\right) \int_U \left|f_k\right|^2 e^{-kv_M(1-\epsilon)} dV_X}. \end{aligned}$$

But $|g_k(0)|^2 = |f_k(0)|^2$ and $|g_k|^2 e^{-kv_M^\epsilon} \leq |f_k|^2 e^{-kv_M(1-\epsilon)}$. Moreover, as explained above $\int_U |f_k|^2 e^{-kv_M|(1-\epsilon)} = O(k^2)$ and hence we get just as above

$$\frac{\left|p_k(0)\right|^2}{\int_X \left|p_k\right|^2 e^{-kv_M^\epsilon} dV_X} \geq C_\epsilon k^2 \left(1 + O\left(e^{-\delta k}\right)\right)$$

which concludes the proof of the bound (i)

$$B_{kv_M^\epsilon} \geq Ck^{2n} \quad \text{on } M.$$

2.3.2. The inequality between B_k and B_{kv_M}. First observe that it is enough to prove the following lemma where now y denotes the imaginary part of $z \in \mathbb{C}^n$ (so that $X \cap \{y = 0\} = M$):

LEMMA 8. *Let U_k be the set of all points in X such that $|y| \le 1/k$ (which defines a neighborhood of M in X). Then there exists a constant C such that*

$$\int_M |p_k|^2 dV_M \le C \frac{1}{\mathrm{Vol}(U_k)} \int_{U_k} |p_k|^2 \, dV_X$$

for any polynomial of total degree at most k.

Indeed, since the function v_M on X is comparable to $|y|$ close to M (by Theorem 7) and in particular kv_M^ϵ is uniformly bounded on U_k, we then get that

$$\int_M |p_k|^2 dV_M \le C' \frac{1}{k^n} \int_X |p_k|^2 e^{-kv_M^\epsilon} dV_X.$$

It follows immediately that

$$B_k \ge k^n / C' B_{kv_M^\epsilon}$$

on M, which combined with the inequality (i) thus concludes the proof of the lower bound in (1.4), given Lemma 8, to whose proof we next turn.

For any $x \in M$ there are constants C and r_0 such that for any $r < r_0$,

$$|f(x)|^2 \le \frac{C}{r^{2n}} \int_{X \cap B(x,r)} |f(y)|^2 \, dV_X(y)$$

for any f holomorphic in X. In particular if we integrate over $x \in X$ a polynomial of degree k taking $r = 1/k$ we get

$$\begin{aligned} \int_M |p_k|^2 dV_M &\le Ck^{2n} \int_{U_k} |p_k(y)|^2 V_M\big(B(y,1/k) \cap M\big) \, dV_X(y) \\ &\le Ck^n \int_{U_k} |p_k|^2 dV_X \le \frac{C}{\mathrm{Vol}(U_k)} \int_{U_k} |p_k|^2 dV_X. \end{aligned} \qquad \square$$

2.4. The L^q Bernstein inequality. Let M be a smooth compact algebraic variety in $\mathbb{R}^m$ of dimension n.

Given a polynomial $p \in \mathcal{P}_k(\mathbb{R}^m)$ and $x \in M$ we denote by $\nabla_t p(x)$ the tangential gradient of p along the manifold M. The following Bernstein type inequality holds:

THEOREM 9. *Let $q \in [1, \infty]$, then there is a constant C_q such that*

$$\|\nabla_t p\|_{L^q(M)} \le C_q \deg(p) \|p\|_{L^q(M)}.$$

The case $q = \infty$ was proved in [6]. We prove now the case $q = 1$ and the others follow by interpolation.

Let X be a complexification of M, i.e., an algebraic variety in $\mathbb{C}^m$ such that $M = X \cap \mathbb{R}^m$. We denote by $U_r \subset X$ the neighborhood of M defined as $U_r = \{x \in X : d(x,M) < r\}$. By the Cauchy inequalities we have that for any $x \in M$ and any $f \in \mathcal{H}(X)$:

$$\left|\nabla_t f(x)\right| \lesssim \frac{1}{r^{2n+1}} \int_{B(x,r)} \left|f(y)\right| dV_X(y).$$

and integrating over M we have

$$\begin{aligned}\int_M \left|\nabla_t f(x)\right| dV_M &\lesssim \frac{1}{r^{2n+1}} \int_M \int_{B(x,r)} \left|f(y)\right| dV_X(y) dV_M(x)\\ &\lesssim \frac{1}{r^{n+1}} \int_{U_r} \left|f(y)\right| dV_X(y).\end{aligned}$$

Therefore Theorem 9 follows from the following result:

THEOREM 10. *There is $C > 0$ such that for all polynomials p_k of degree k, the following inequality holds:*

$$\int_{U_{1/k}} \left|p_k\right| dV_X \leq C k^{-n} \int_M \left|p_k\right| dV_M.$$

Proof. In order to estimate the integral over $U_{1/k}$ we will integrate along surfaces surrounding M. These surfaces will be level sets of plurisubharmonic functions with Monge-Ampère 0. In this setting there is a generalization of Hadamard three circles theorem due to Demailly that will be used, see [11]. In order to use this technique we need that the psh-function that defines the level sets is smooth out of M, and its square must be smooth. We can use the function provided by Guillemin and Stenzel in [15] in their study of Grauert tubular neighborhoods around real analytic manifolds. We present the setting:

Take ψ a plurisubharmonic function in a neighborhood U of M in X defined as

$$\psi(z) = d(z,M),$$

where the distance d is given by a metric provided in a Grauert tubular neighborhood U as in [15]. The function ψ satisfies $(dd^c\psi)^n = 0$ in $U \setminus M$, ψ^2 is a real analytic Kähler potential in U and $(dd^c(\psi^2))^n$ is comparable to the volume form in X in a neighborhood of M.

We use the same notation as in [11]. Consider the pseudospheres $S_r = \{z \in U;\ \psi(z) = r\}$, $r > 0$ and the positive measures μ_r supported on S_r that are defined as

$$\mu_r(h) := \int_{S_r} h\big(dd^c\psi\big)^{n-1} \wedge d^c\psi$$

for any $h \in C(U)$. When $r = 0$, then we define

$$\bar{\mu}_0(h) := \int_M h\big(dd^c\psi\big)^n.$$

We have that $\mu_r(h)$ is continuous for $r > 0$ with $\mu_r(h) \to \bar{\mu}_0(h)$ as $r \to 0^+$ see [11, Theorem 3.2].

Moreover $(dd^c\psi)^n$ which is supported on M is comparable to the volume form in M. This is so because at any point $z \in M$ we can write local holomorphic coordinates such that M corresponds to $z \in \mathbb{C}^n : \Im z = 0$. In this coordinates $\psi(z)$ is comparable to $|\Im z|$ and since $(dd^c|\Im z|)^n$ is the Lebesgue measure on $\mathbb{R}^n$ then by the comparison principle for the Monge-Ampère measure, see [1, Lemma 2.1], the measure $\bar{\mu}_0$ is locally comparable to the volume form and M being compact it is globally comparable.

Take the psh function $V = \log|p_k|$, then [11, Corollary 6.6(a)] says that the function

$$u(r) = \log \mu_r\big(e^V\big), \quad r > 0, \quad u(0) = \log \bar{\mu}_0\big(e^V\big)$$

is convex and increasing in r. We fix $R > 0$ such that S_R belongs to the tubular neighborhood U. The convexity of u implies that for any $r > 0$

$$u(r) \le u(0)\frac{R-r}{R} + u(R)\frac{r}{R}. \tag{2.6}$$

We have that

$$u(0) = \log \int_M \big|p_k\big|\, d\bar{\mu}_0.$$

We are going fix R such and estimate $u(R)$. Since p_k is a polynomial of degree k we have that by the Bernstein-Walsh estimate

$$\sup_{S_R} \big|p_k\big| \le \sup_{S_R} e^{k\phi_M(z)} \sup_M |p|,$$

where ϕ_M is the Siciak extremal function defined in Section 2.2

It is a well-known theorem of Sadullaev, see [25] that if X is algebraic then $\phi_K \in L^\infty_{\text{loc}}(X)$ for any non-pluripolar compact set K relative to X. Certainly M is non-pluripolar relative to X since it is totally real. Therefore:

$$\sup_{S_R} \big|p_k\big| \le C^k \sup_M |p|.$$

Moreover, we need the following Bernstein-Markov type inequality:

$$\sup_M |p| \le C_M^k \int_M |p|. \tag{2.7}$$

This is easier than the standard Bernstein-Markov property since we are not requiring that C_M is close to 1. In our case (2.7) is a special case of [8, Theorem 4.1].

Finally $\sup_{S_R} |p_k| \leq C^k \int_M |p|$, and we have that $u(R) \leq (C + \log \|p\|_{L^1(M)})k$. Therefore if $r = 1/k$ and using the convexity (2.6) we deduce that

$$u(1/k) \leq u(0) + C.$$

Since $u(r)$ is increasing we have that for any $r < 1/k$

$$\int_{S_r} |p_k| d\mu_r \leq C \int_M |p_k| d\bar{\mu}_0.$$

But the measures μ_r disintegrate the form $(dd^c\psi)^{n-1} \wedge d\psi \wedge d^c\psi$, see [11, Proposition 3.9], and we have that

$$\int_0^{1/k} r^{n-1} \int_{S(r)} |p_k| d\mu_r = \int_{\psi<1/k} |p_k(z)| \psi(z)^{n-1} \big(dd^c\psi\big)^{n-1} \wedge d\psi \wedge d^c\psi.$$

Moreover

$$\psi(z)^{n-1} \big(dd^c\psi\big)^{n-1} \wedge d\psi \wedge d^c\psi = \big(dd^c\big(\psi^2\big)\big)^n.$$

But ψ^2 is a real analytic Kähler potential in X, see [15]. Thus $(dd^c(\psi^2))^n$ is equivalent to the original volume form V_X in a neighborhood of M in X:

$$\int_{\psi<1/k} |p_k| dV_X \simeq \int_0^{1/k} r^{n-1} \int_{S(r)} |p_k| d\mu_r. \qquad \square$$

Remark. With the same proof, for any $1 \leq q < \infty$,

$$\int_{U_{1/k}} |p_k|^q dV_X \lesssim k^{-n} \int_M |p_k|^q dV_M.$$

It is also true that

$$\sup_{U_{1/k}} |p_k| \lesssim \sup_M |p_k|.$$

The proof is the same, but instead of [11, Corollary 6.6(a)] one uses that

$$u(r) = \sup_{S_r} \log |p_k|,$$

is a convex function of r, see [11, Corollary 6.6(b)].

Remark. For any $x \in M$ we consider a ball $B_X(x, 1/k)$ in the complexified manifold X. By the submean value property we have that

$$|p_k(x)|^2 \lesssim k^{2n} \int_{B_X(x,1/k)} |p_k(y)|^2 \, dV_X(y) \lesssim k^n \int_M |p_k|^2 \, dV_M.$$

Therefore $K_k(x,x) \lesssim k^n$ and we have proved the upper inequality in (1.4). We include the argument for completeness but this upper bound is well known and it follows from the arguments in [30].

It is also possible to prove a converse result to Theorem 9:

THEOREM 11. *Let M be a smooth compact submanifold in $\mathbb{R}^m$. If there is a constant $C > 0$ such that for some $q \in [1,\infty]$,*

$$\left\|\nabla_t p\right\|_{L^q(M)} \le C \deg(p)\|p\|_{L^q(M)},$$

for all polynomials $p \in \mathcal{P}(\mathbb{R}^m)$, then M is algebraic.

Proof. We will need a definition

Definition 4. A sequence of finite sets $\{\Lambda_k\}$ is an ϵ-net if it is uniformly separated and $1 \le \sum_{\lambda\in\Lambda_k} \chi_{B(\lambda,\varepsilon/k)}(x) \le C$, for all $x \in M$ and $k > 0$.

By an application of the Vitali covering lemma it is possible to construct ϵ-nets for arbitrarily small ε where the constant $C = C_M$ depends on the dimension of M but not on ε.

Given an ϵ-net $\Lambda = \Lambda(\varepsilon)$ we denote by $l_k = \#\Lambda_k$. We may define: $T_k : \mathcal{P}_k(M) \to \mathbb{R}^{l_k}$ as

$$T_k(p)(\lambda) = p_{B_M(\lambda,\varepsilon/k)} := \fint_{B_M(\lambda,\varepsilon/k)} p\, dV_M \quad \forall\lambda \in \Lambda_k.$$

We will prove now that if ε is small enough then

$$\int_M |p_k|^q \lesssim \frac{1}{k^n} \sum_{\lambda\in\Lambda_k} |T_k(p)(\lambda)|^q. \tag{2.8}$$

If this is the case, then T_k is one to one and $\dim(\mathcal{P}_k(M)) \le l_k \simeq k^n$ where $n = \dim(M)$ which is much smaller than k^m. Thus M is algebraic.

Let us prove (2.8).

$$\int_M |p_k|^q\, dV_M \lesssim \frac{1}{k^n} \sum_{\lambda\in\Lambda_k} |T_k(p)(\lambda)|^q + \sum_{\Lambda_k} \int_{B_M(\lambda,\varepsilon/k)} \left|p(x) - p_{B_M(\lambda,\varepsilon/k)}\right|^q dV_M.$$

By the Poincaré inequality:

$$\int_{B_M(\lambda,\varepsilon/k)} \left|p(x) - p_{B_M(\lambda,\varepsilon/k)}\right|^q dV_M \lesssim \frac{\varepsilon^q}{k^q} \int_{B_M(\lambda,\varepsilon/k)} |\nabla_t p|^q\, dV_M.$$

By Theorem 9

$$\int_M |p_k|^q \, dV_M \lesssim \frac{1}{k^n} \sum_{\lambda \in \Lambda_k} |T_k(p)(\lambda)|^q + \varepsilon \int_M |p_k|^q \, dV_M,$$

and if ε is small enough then (2.8) follows. □

2.5. Applications. We will use now the Bernstein inequality to get some more information on sampling and interpolation sequences of finite sets.

Definition 5. A sequence of measures $\{\mu_k\}_k$ is said to be a uniformly sequence of Carleson measures if there is a $C > 0$ such that

$$\int_M |p_k|^2 \, d\mu_k \le C \|p_k\|^2, \quad \forall p_k \in \mathcal{P}_k. \tag{2.9}$$

In M we consider the balls defined by any Riemannian metric.

PROPOSITION 12. *A sequence $\{\mu_k\}_k$ is a uniformly sequence of Carleson measures if and only if there is a $C > 0$ such that*

$$\mu_k\big(B(x, 1/k)\big) < C/k^n \quad \textit{for all } x \in M,\ k \in \mathbb{N}.$$

Proof. The necessity follows from testing (2.9) against normalized reproducing kernels. For any $x \in M$, let $\kappa_{x,k}(y) = K_k(x,y)/\sqrt{K_k(x,x)}$. Then it is clear that for all $y \in M$:

$$|\kappa_{x,k}(y)| \le |\langle K_k(x,\cdot), K_k(y,\cdot)\rangle| / \sqrt{K_k(x,x)} \le \sqrt{K_k(y,y)} \simeq k^{n/2},$$

and $|\kappa_{x,k}(x)| = \sqrt{K_k(x,x)} \simeq k^{n/2}$. On the other hand since M is algebraic we have the classical Bernstein inequality, see [6]:

$$\sup_M \|\nabla_t \kappa_{x,k}\| \lesssim k \sup_M |\kappa_{x,k}| \simeq k^{n/2+1}.$$

This means that there is a $\delta > 0$ such that

$$|\kappa_{x,k}(y)| \gtrsim \kappa_{x,k}(x) \simeq k^{n/2}, \quad \forall y \in B_M(x, \delta/k).$$

Therefore if we test (2.9) with $\kappa_{x,k}$ we get that $\mu_k(B_M(x,\delta/k)) \lesssim 1/k^n$.

In the other direction for any $x \in M$ we consider a ball $B_X(x,r)$ the ball in the complexified manifold X. By the submean property of the holomorphic functions $f \in \mathcal{H}(X)$, we have that for $r \le r_0$, $|f(x)| \lesssim \fint_{B_X(x,r)} |f(y)| \, dV_X(y)$. Thus

$$|p_k(x)|^2 \lesssim \fint_{B_X(x,1/k)} |p_k(y)|^2 \, dV_X(y) \simeq k^{2n} \int_{B_X(x,1/k)} |p_k(y)|^2 \, dV_X(y).$$

Finally, thanks to Theorem 10

$$\begin{aligned}\int_X |p_k(x)|^2\, d\mu_k(x) &\lesssim \int_{x\in X} k^{2n} \int_{B_X(x,1/k)} |p_k(y)|^2\, dV_X(y), d\mu_k(x)\\ &\lesssim \int_{y\in X, d(y,M)<1/k} k^{2n} |p_k(y)|^2 \mu_k(B_X(y,1/k)\cap X) dV_X(y)\\ &\lesssim \int_{y\in X, d(y,M)<1/k} k^{n} |p_k(y)|^2 dV_X(y)\\ &\lesssim \int_X |p_k(x)|^2\, dV_M(x). \qquad \square\end{aligned}$$

An immediate corollary is the description of the sequences that satisfy the left-hand side inequality of the sampling sequences, that is a Plancherel-Polya type inequality.

We say that a sequence of finite sets Λ_k is *uniformly separated* if and only if there is an $\varepsilon > 0$ such that $d(\lambda, \lambda') \geq \varepsilon/k$ for all $\lambda \neq \lambda'$, $\lambda, \lambda' \in \Lambda_k$.

COROLLARY 13. (Plancherel-Polya type inequality) *The sequence of finite sets Λ_k is a finite union of uniformly separated sequences if and only if there is a constant $C > 0$ such that*

$$\frac{1}{k^n} \sum_{\lambda\in\Lambda_k} |p(\lambda)|^2 \leq C \int_M |p|^2, \quad \forall p \in \mathcal{P}_k.$$

Proof. Apply Proposition 12 to the measures $\mu_k = \frac{1}{k^n}\sum_{\lambda\in\Lambda_k} \delta_\lambda$. This implies that the Plancherel-Polya type inequality holds if and only if $\#\{\Lambda_k \cap B(x, 1/k)\} \leq C$ uniformly in x and k. That is, the sequence Λ_k is a finite union of uniformly separated sequences of sets. $\square$

Once we have a Bernstein type inequality the following Proposition is standard, see [26, Proposition 5, p. 47] and it allows to reduce our considerations to uniformly separated sequences.

PROPOSITION 14. *If Λ_k is a sampling sequence then there is a uniformly separated sequence of subsets $\Lambda'_k \subset \Lambda_k$ such that Λ'_k is still a sampling sequence.*

It is also completely standard that:

PROPOSITION 15. *If Λ_k is an interpolating sequence then it is uniformly separated.*

Theorem 10 can also be used to provide a sufficient condition that assures the existence of sampling sequences. More precisely, we say that the sequence Λ_k is an ϵ-net if it is uniformly separated and $1 \leq \sum_{\lambda\in\Lambda_k} \chi_{B(\lambda,\varepsilon/k)}(x) \leq C_M$, for all $x \in M$ and $k > 0$, the constant C_M depends on M but not on ε. By an application of the Vitali covering lemma it is possible to construct ϵ-nets for arbitrarily small ε.

PROPOSITION 16. *There is an ϵ_0 such that any sequence Λ_k that is an ϵ-net with $\epsilon < \epsilon_0$ is a sampling sequence.*

Proof. Take one ϵ-net $\Lambda_k = \Lambda_k(\varepsilon)$. Then

$$\begin{aligned}\int_M |p_k|\, dV_M &\le \sum_{\lambda\in\Lambda_k} \int_{B_M(\lambda,\varepsilon/k)} |p_k|\, dV_M \\ &\le \sum_{\lambda\in\Lambda_k} |p_k(\lambda)|\big|B_M(\lambda,\varepsilon/k)\big| + \sum_{\lambda\in\Lambda_k} \int_{B_M(\lambda,\varepsilon/k)} \big|p_k(x)-p_k(\lambda)\big|\, dV_M \\ &\lesssim \sum_{\lambda\in\Lambda_k} \frac{\varepsilon^n}{k^n}|p_k(\lambda)| + \sum_{\lambda\in\Lambda_k} \frac{\varepsilon}{k}\big|\nabla_t p_k(\zeta_\lambda)\big|\big|B_M(\lambda,\varepsilon/k)\big|,\end{aligned}$$

where $\zeta_\lambda \in \overline{B_M(\lambda,\varepsilon/k)}$ is such that $|\nabla_t p_k(\zeta_\lambda)| = \sup_{x\in B_M(\lambda,\varepsilon/k)} |\nabla_t p_k(x)|$. By the Cauchy inequality, if we take a ball $B_X(\lambda,1/k)$ in the complexification X of M, we have

$$\big|\nabla_t p(\zeta_\lambda)\big| \lesssim k \fint_{B_X(\lambda,1/k)} |p_k|\, dV_X.$$

Therefore,

$$\int_M |p_k| dV_M \lesssim \frac{\varepsilon^n}{k^n} \sum_{\lambda\in\Lambda_k} |p_k(\lambda)| + \sum_{\lambda\in\Lambda_k} \varepsilon^{n+1} k^n \int_{B_X(\lambda,1/k)} |p_k|\, dV_X.$$

Since Λ_k is an ε-net there are at most $C\varepsilon^{-n}$ points of Λ_k in any given ball $B_X(x,1/k)$ of center $x \in U_{1/k}$. Thus,

$$\int_M |p_k| dV_M \lesssim \frac{\varepsilon^n}{k^n} \sum_{\lambda\in\Lambda_k} |p_k(\lambda)| + \varepsilon k^n \int_{U_{1/k}} |p_k|\, dV_X.$$

We use now Theorem 10 to control the right-hand side integral. If we take ε small enough we can absorb the integral in the left-hand side and we get

$$\int_M |p_k| dV_M \lesssim \frac{1}{k^n} \sum_{\lambda\in\Lambda_k} |p_k(\lambda)|.$$

The L^∞ version: $\sup_M |p_k| \lesssim \sup_{\Lambda_k} |p_k|$ follows immediately by the Bernstein inequality proved in [6], if ε is small enough. By interpolation we get that for any $q \in [1,\infty)$

$$\int_M |p_k|^q dV_M \lesssim \frac{1}{k^n} \sum_{\lambda\in\Lambda_k} |p_k(\lambda)|^q.$$

The reverse inequality

$$\frac{1}{k^n}\sum_{\lambda\in\Lambda_k}|p_k(\lambda)|^q \lesssim \int_M |p_k|^q dV_M,$$

follows from Corollary 13 since Λ_k is uniformly separated. □

We can finish now the proof of Theorem 3

Proof. We have already proved that the algebracity of M is equivalent to the Bernstein inequality, this is Theorem 9 and 11. Moreover Proposition 16 proves that compact algebraic manifolds have uniformly separated sampling sequences. So we only need to check that if there are such sequences then M is algebraic. This is proved in a similar way to Theorem 11. We denote by $l_k = \#\Lambda_k$ as before. Define: $R_k : \mathcal{P}_k(M) \to \mathbb{R}^{l_k}$ as

$$R_k(p)(\lambda) = p(\lambda) \quad \forall\lambda \in \Lambda_k.$$

Clearly, since we have the sampling property, R_k is one-to-one. Therefore

$$\dim(P_k(M)) \leq l_k.$$

Moreover since Λ_k is uniformly separated, then $l_k \leq k^n$. This implies that M is algebraic. □

2.6. A general off-diagonal estimate on the reproducing kernel.

THEOREM 17. *Let M be an n-dimensional affine real algebraic variety (possibly singular), μ_k a sequence of non-degenerate finite measures on M with support contained in a compact of M and denote by $K_k(x,y)$ the reproducing kernel for the space $H_k(M)$, viewed as a subspace of $L^2(M,\mu_k)$. Then there exists a positive constant C such that*

$$\int_{M\times M}\frac{1}{k^n}|K_k(x,y)|^2 d\mu_k(x)\otimes d\mu_k(y)|x-y|^2 \leq C/k.$$

Remark. Observe that if we pick $\mu_k = e^{-k\phi}\mu$ the theorem covers the weighted setting as well.

Proof. Given a bounded function f on M we denote by T_f be the Toeplitz operator on $H_k(M)\cap L^2(M,\mu_k)$ with symbol f, i.e., $T_f := \Pi_k \circ f\cdot$ where Π_k denotes the orthogonal projection from $L^2(M,\mu_k)$ to $H_k(M)$, i.e., T_f is the Hermitian operator on $H_k(M)$ determined by

$$\langle T_f p_k, p_k\rangle_{L^2(M,\mu_k)} = \langle f p_k, p_k\rangle_{L^2(M,\mu_k)}$$

for any $p_k \in H_k(M)$. The following is essentially a well-known formula

$$\operatorname{Tr} T_f^2 - \operatorname{Tr} T_{f^2} = \frac{1}{2}\int_{M\times M} \left|K_k(x,y)\right|^2 d\mu_k(x)\otimes d\mu_k(y)\big(f(x)-f(y)\big)^2.$$

We provide nevertheless a proof for convenience of the reader:

CLAIM. *Let H be a reproducing kernel Hilbert space with kernel K, then for any bounded symbol f we have*

$$\iint \left|f(x)-f(y)\right|^2 \left|K(x,y)\right|^2 = \operatorname{Tr}\big(2T_{|f|^2} - T_f \circ T_{\bar f} - T_{\bar f}\circ T_f\big).$$

Proof. $K(x,y) = \sum_n f_n(x)\overline{f_n(y)}$ and

$$T_f(g)(x) = \int K(x,y) f(y) g(y).$$

We compute the traces of $T_{|f|^2}$ and of $T_f \circ T_{\bar f}$.

$$\begin{aligned}\operatorname{Tr}\big(T_{|f|^2}\big) &= \sum_n \big\langle f_n, T_{|f|^2}\big(f_n\big)\big\rangle \\ &= \sum_n \int_x f_n(x)\overline{\int_y K(x,y)|f|^2(y) f_n(y)} = \iint \left|K(x,y)\right|^2 \overline{|f|^2(y)}.\end{aligned}$$

Thus

$$\operatorname{Tr}\big(T_{|f|^2}\big) = \iint \left|K(x,y)\right|^2 |f(x)|^2 = \iint \left|K(x,y)\right|^2 \left|f(y)\right|^2.$$

Now

$$\begin{aligned}\operatorname{Tr}(T_f\circ T_{\bar f}) &= \sum_n \int_x f_n(x)\overline{\int_y K(x,y) f(y) T_{\bar f}(f_n)(y)} \\ &= \sum_n \int_x\int_y f_n(x)\overline{K(x,y)f(y)}\,\overline{\int_w K(y,w)\overline{f(w)} f_n(w)} \\ &= \iiint K(x,w)\overline{K(y,w)K(y,w)f(y)} f(w) \\ &= \iint |K(y,w)|^2 \overline{f(y)} f(w).\end{aligned}$$

Similarly

$$\operatorname{Tr}\big(T_{\bar f}\circ T_f\big) = \iint \left|K(y,w)\right|^2 f(y)\overline{f(w)}. \qquad \square$$

Now, setting $f := x_i$ for a fixed index $i \in \{1,\dots,m\}$ we note that there exists a vector subspace V_k in $H_k(M)$ such that $\dim V_k = H_k(M) - O(k^{n-1})$ such that $T_f = f$ and $T_f^2 = f^2$. We can take V_k to be the space spanned by the restrictions to

M of all polynomials of total degree at most $k-1$, i.e., H_{k-1}. The dimension of $N_k = \dim(H_k)$ is the Hilbert polynomial of degree for $k \geq k_0$. Thus $N_k = dk^n + O(k^{n-1})$ where d is the degree of the variety M and n is the dimension. In particular, denoting by W_k the orthogonal complement of V_k in $H_k(M) \cap L^2(M, \mu_k)$ then $\dim(W_k) = O(k^{n-1})$. Setting $A_k := T_f^2 - T_{f^2}$ gives $A_k = 0$ on V_k and hence

$$\operatorname{Tr} T_f^2 - \operatorname{Tr} T_{f^2} = 0 + \operatorname{Tr} A_{k|W_k} \leq C k^{n-1}$$

using that $\langle T_f p_k, p_k \rangle_{L^2(M,\mu_k)} \leq \sup |f|_M \langle p_k, p_k \rangle_{L^2(M,\mu_k)}$ and $\dim W_k = O(k^{n-1})$. $\square$

3. Sampling and interpolation of real orthogonal polynomials.

3.1. Sampling polynomials in a real variety.

Proof of Theorem 1. We equip M with the distance function d induced by the Euclidean distance in $\mathbb{R}^m$, i.e., $d(x,y) := |x-y|$. We recall that the corresponding Wasserstein L^1-distance on the space $\mathcal{P}(M)$ of all probability measures on M is defined as

$$W(\mu, \sigma) = \inf_{\rho} \iint_{M \times M} d(x,y)\, d\rho(x,y),$$

where the infimum is taken among all probability measures such that the first marginal of ρ is μ and the second σ. The Wasserstein distance metrizes the weak-$*$ convergence.

We rely on the fact that

$$\frac{1}{N_k} B_k(x)\, d\mu_k(x) \longrightarrow \mu_{eq}(x),$$

where the convergence is in the weak-$*$ topology, see [2]. Thus the way to prove the inequality of the theorem is by proving that there are constants $\{c_\lambda\}_{\lambda \in S_k}$, $0 \leq c_\lambda < 1$ such that

$$W\big(\sigma_k, \beta_k\big) \longrightarrow 0$$

where $\sigma_k = \frac{1}{N_k} \sum_{\lambda \in \Lambda_k} c_\lambda \delta_\lambda$, $\beta_k = \frac{1}{N_k} B_k(x)$. Instead of the standard Wasserstein distance we will use an alternative expression more convenient for our purpose that it is equivalent to it, see [18]:

$$W(\mu, \sigma) = \inf_{\rho} \iint_{M \times M} d(x,y)\, \big|d\rho(x,y)\big|,$$

where the inf is taken among all complex measures ρ such that the first marginal of f is μ and the second σ. The difference is that ρ is not necessarily positive and even if we don't require that σ and μ are probability measures it still metrizes the weak-$*$ convergence. Any candidate ρ with the right marginals is called a transport plan.

The transport plan ρ_k that is convenient to estimate is:

$$\rho_k(x,y) = \frac{1}{N_k}\sum_{\lambda\in\Lambda_k}\delta_\lambda(y)\times g_\lambda(x)\frac{K_k(\lambda,x)}{\sqrt{B_k(\lambda)}}\,d\mu_k(x),$$

where $K_k(\lambda,x)$ is the reproducing kernel for λ in the space H_k and $\{g_\lambda\}_{\lambda\in S_k}$ is the canonical dual frame (see [10]) to $\{\frac{K_k(\lambda,x)}{\sqrt{B_k(\lambda)}}\}_{\lambda\in\Lambda_k}$ in H_k. The latter is a frame because Λ_k is sampling.

If we compute the marginals of ρ_k we get on one hand:

$$\sigma_k(y) = \frac{1}{N_k}\sum_{\lambda\in\Lambda_k}\frac{g_\lambda(\lambda)}{\sqrt{B_k(\lambda)}}\delta_\lambda(y),$$

and the other marginal is given by

$$d\beta_k(y) = \frac{1}{N_k}\sum_{\lambda} g_\lambda(x)\frac{K_k(\lambda,x)}{\sqrt{B_k(\lambda)}}\,d\mu_k(x) = \frac{1}{N_k}K_k(x,x)\,d\mu_k(x).$$

In the last equality we have used that g_λ is a dual frame of the normalized reproducing kernels.

The fact that $\{g_\lambda\}$ it is the canonical dual frame to the normalized reproducing kernels allows us to conclude that $\frac{g_\lambda(\lambda)}{\sqrt{B_k(\lambda)}} = \langle g_\lambda(x), \frac{K_k(\lambda,x)}{\sqrt{B_k(\lambda)}}\rangle$ is positive and smaller than one. This follows from the following well-known fact:

CLAIM. *If $\{x_n\}_n$ is a frame in a Hilbert space H and $\{y_n\}_n$ is the dual frame then $\langle x_n, y_n\rangle \in [0,1]$.*

Proof. Let T be the frame operator, i.e.: $T(x) = \sum\langle x, x_n\rangle x_n$. Since $\{x_n\}_n$ is a frame then T is bounded, self-adjoint and invertible. The definition of the dual frame is $T(y_n) = x_n$. For any vector $v\in H$ we have

$$v = T\big(T^{-1}v\big) = \sum_n \big\langle T^{-1}v, x_n\big\rangle x_n.$$

In particular

$$x_k = \sum_n \big\langle y_k, x_n\big\rangle x_n,$$

and multiplying by y_k at both sides we get

$$\big\langle x_k, y_k\big\rangle = \sum_n \big|\big\langle y_k, x_n\big\rangle\big|^2.$$

Therefore $\langle x_k, y_k\rangle \geq 0$ and $\langle x_k, y_k\rangle > 0$ unless $x_k = 0$. Moreover,

$$\langle x_k, y_k\rangle - \left|\langle y_k, x_k\rangle\right|^2 = \sum_{n\neq k} \left|\langle y_k, x_n\rangle\right|^2 \geq 0.$$

Thus

$$\langle x_k, y_k\rangle \big(1 - \langle x_k, y_k\rangle\big) \geq 0,$$

therefore $\langle x_k, y_k\rangle \leq 1$ too. □

Finally we need to estimate

$$I = \iint_{M\times M} |x-y|\left|d\rho_k\right| \leq \frac{1}{N_k} \sum_{\lambda\in\Lambda_k} \int_M |\lambda - x|\left|K_k(\lambda,x)\right| \frac{g_\lambda(x)}{\sqrt{B_k(\lambda)}}\, d\mu_k(x).$$

Since $\|g_\lambda\|_2 \simeq 1$ we can estimate

$$I^2 \lesssim \frac{1}{N_k} \sum_{\lambda\in\Lambda_k} \int_M |\lambda - x|^2 \frac{|K_k(\lambda,x)|^2}{B_k(\lambda)}.$$

We would like to use the sampling inequality (1.1), and obtain that

$$\begin{aligned} &\frac{1}{N_k} \sum_{\lambda\in\Lambda_k} \int_M |\lambda - x|^2 \frac{\left|K_k(\lambda,x)\right|^2}{B_k(\lambda)}\, d\mu_k(x) \\ &\quad\leq \frac{1}{N_k} \iint_{M\times M} |y-x|^2 \left|K_k(y,x)\right|^2 d\mu_k(x) d\mu_k(y). \end{aligned} \tag{3.1}$$

This we cannot do immediately because the polynomial $(x-y)K_k(y,x)$ (in the variable y) is of degree $k+1$ instead of k as required in (1.1).

But we are assuming that (μ,ϕ) define spaces with reproducing kernels of moderate growth. Thus $B_{k+1} \simeq B_k$ in M. Therefore if $\{\Lambda_k\}_k$ is sampling for $H_k(M)$ then $\{\Lambda_{k+1}\}_k$ is sampling for $H_k(M)$. Thus, it is harmless to assume that Λ_k is sampling both for H_k and for H_{k+1} and we have established (3.1). Then, using Theorem 17, we obtain

$$W\big(\sigma_k, \beta_k\big) = O(1/\sqrt{k}),$$

as desired. □

3.2. Interpolating polynomials in a real variety. The property that a sequence of sets of points Λ_k is an interpolating family as in Definition 1 is equivalent to the two following properties. First, the Plancherel-Polya inequality:

$$\sum_{\lambda\in\Lambda_k} \frac{\left|f(\lambda)\right|^2}{K_k(\lambda,\lambda)} \leq C\|f\|^2, \quad \forall f \in H_k(M) \tag{3.2}$$

and, second, the *interpolation property*: for any sequence of sets of values $\{c_\lambda^{(k)}\}_{\lambda\in\Lambda_k}$ there are functions $f_k \in H_k$ such that $f_k(\lambda^{(k)}) = c_\lambda^{(k)}$ with

$$\|f_k\|^2 \le C \sum_{\lambda\in\Lambda_k} \frac{|c_\lambda|^2}{K_k(\lambda,\lambda)}, \tag{3.3}$$

and again the constant C should not depend on k.

The property that the collection $\{\kappa_\lambda\}_{\lambda\in\Lambda_k}$ is a frame in $H_k(M)$ is a quantitative version of the fact that the normalized reproducing kernels span the whole space and the property that they are a Riesz sequence quantifies the fact that they are linearly independent.

Proof of Theorem 5. Let $F_k \subset H_k$ be the subspace spanned by

$$\kappa_\lambda(x) = K_k(\lambda,x)/\sqrt{K_k(\lambda,\lambda)} \quad \forall \lambda \in \Lambda_k.$$

Denote by g_λ the dual (biorthogonal) basis to κ_λ in F_k. We have clearly that:

- We can span any function in F_k in terms of κ_λ, thus:

$$\sum_{\lambda\in\Lambda_k} \kappa_\lambda(x) g_\lambda(x) = \mathcal{K}_k(x,x),$$

where $\mathcal{K}_k(x,y)$ is the reproducing kernel of the subspace F_k.

- The norm of g_λ is uniformly bounded since κ_λ was a uniform Riesz sequence.
- $g_\lambda(\lambda) = \sqrt{K_k(\lambda,\lambda)}$. This is due to the biorthogonality and the reproducing property.

We are going to prove that the measure $\sigma_k = \frac{1}{N_k}\sum_{\lambda\in\Lambda_k}\delta_\lambda$, and the measure $\beta_k = \frac{1}{N_k}\mathcal{K}_k(x,x)d\mu(x)$ are very close to each other: $W(\sigma_k,\beta_k)\to 0$. In this case then since $\mathcal{K}_k(x,x) \le K_k(x,x)$ and $\frac{1}{N_k}B_k(x)\,d\mu \to d\mu_{eq}$, where μ_{eq} is the normalized equilibrium measure on M, then $\limsup_k \sigma_k \le \mu_{eq}$.

In order to prove that $W_k(\sigma_k,\beta_k)\to 0$ we use the transport plan:

$$\rho_k(x,y) = \frac{1}{N_k}\sum_{\lambda\in\Lambda_k} \delta_\lambda(y)\times g_\lambda(x)\kappa_\lambda(x)\,d\mu(x).$$

It has the right marginals, σ_k and β_k and we can estimate the integral in the same way as in the proof of Theorem 1

$$W\big(\sigma_k,\beta_k\big) \le \iint_{M\times M} |x-y|\big|d\rho_k\big| = O(1/\sqrt{k}).$$

The only point that merits a clarification is that we need an inequality similar to (3.1), i.e.:

$$\begin{aligned}\frac{1}{N_k}\sum_{\lambda\in\Lambda_k}\int_M|\lambda-x|^2\frac{\left|K_k(\lambda,x)\right|^2}{K_k(x,x)}\,d\mu(x)\\ \le\frac{1}{N_k}\iint_{M\times M}|y-x|^2\left|K_k(y,x)\right|^2d\mu(x)d\mu(y).\end{aligned}$$

This time this is true because Λ_k is a uniformly separated sequence by Proposition 15 and therefore, it is a Plancherel-Polya sequence, see Proposition 13. □

3.3. Sampling in convex domains. We proceed with the proof of Theorem 4. The only part that we need to proof are the estimates for the reproducing kernel (1.5). If these are proved, then it follows that the measure has the Bernstein-Markov property (2.1) and the kernel is of moderate growth (2.2), thus we can apply Theorem 1 in the particular case where $M=\mathbb{R}^n$ and $\mu=\chi_\Omega dm(x)$ where dm is the Lebesgue measure in $\mathbb{R}^n$ and the weight $\phi=0$.

Proof. We start by the case when $\Omega=\mathbb{B}$ is the unit ball. We denote by $B_\Omega(x)$ the Bergman function which is the reproducing kernel $K_k(x,x)$ evaluated at the diagonal of the space of polynomials of total degree k endowed with the L^2 norm with respect to the standard volume form. To get a lower bound for B_Ω we consider the cube Q such that the ball is inside it and tangent to its faces. Clearly by the comparison principle of the Bergman functions $B_\mathbb{B}\ge B_Q$ and $B_Q(x)\simeq B_I(x_1)\cdots B_I(x_n)$ where B_I is the one dimensional Bergman kernel associated to the interval. This is known to be, see [21, p. 108]:

$$B_I(x)\simeq\min\left(\frac{k}{\sqrt{d(x)}},k^2\right).$$

This implies that for points x in the interval that joins the origin with the center of one of the faces of the cube Q we have

$$B_Q(x)\simeq\min\left(\frac{k^n}{\sqrt{d(x)}},k^{n+1}\right).$$

Thus we have the lower bound for $B_\mathbb{B}$ that we wanted. To get the upper bound we will work in dimension $n=2$ for simplicity but a similar argument works in any dimension. Observe that the space of polynomials of degree smaller or equal than k is spanned by the functions $\{\rho^j\cos^j(t),\rho^j\sin^j(t)\}_{j=0,\dots,k}$ in polar coordinates in the interval $[0,1]\times[0,2\pi]$ with the measure $\rho d\rho$ in the first interval and dt in the second. Consider now the space of functions $\tilde{H}_k$ in the product interval such that it is spanned by $\{\rho^j\cos^m(x),\rho^j\sin^m(x)\}_{j=0,\dots,k\ m=0,\dots,k}$. The space $\tilde{H}_k$ is bigger than the space of polynomials thus the Bergman function at the diagonal $B_{\tilde{H}_k}(x)\ge$

$B_{\mathbb{B},k}(x)$. But $B_{\tilde{H}_k}$ is easier to analyze because it is a product space of two one-dimensional spaces: The space of one dimensional polynomials of degree smaller k with the norm $\rho\, d\rho$ in the interval $[0,1]$ and the space of trigonometric polynomials $\{\sin^j(x), \cos^j(x)\}_{j=0,\dots k}$ with the measure dx in $[0,2\pi]$. The Bergman function of $\tilde{H}_k$ is the product of the one-dimensional Bergman functions. The Bergman function corresponding to the trigonometric polynomials is constant by invariance under rotations and by dimensionality it must be $2k+1$. The space of polynomials in ρ are a space of Jacobi polynomials and its Bergman function has been estimated, see [21, p. 108]:

$$B_J(x) \simeq \min\left(\frac{k}{\sqrt{d(x)}}, k^2\right) \quad \forall x > 1/2.$$

Thus finally when $n = 2$ we get

$$B_{\mathbb{B}} \lesssim \min\left(\frac{k^2}{\sqrt{d(x)}}, k^3\right).$$

Similarly in higher dimension we get

$$B_{\mathbb{B}} \lesssim \min\left(\frac{k^n}{\sqrt{d(x)}}, k^{n+1}\right).$$

Now for an arbitrary convex domain there is an $r > 0$ (small) and an $R > 0$ (big) that depend only on the domain such that for any point x in the boundary of the domain, there is a ball $B(y,r)$ inside the domain tangent at x and with center y in the normal direction to the boundary of the domain at x and a cube $Q(R)$ tangent to the domain at x in the middle of a face of the cube and such that the domain is contained in the cube. Again by the comparison principle of the Bergman function we get

$$\begin{aligned}\min\left(\frac{k^n}{\sqrt{d(x)}}, k^{n+1}\right) &\lesssim B_{Q(R)}(x) \lesssim B_\Omega(x)\\ &\lesssim B_{B(y,r)}(x) \lesssim \min\left(\frac{k^n}{\sqrt{d(x)}}, k^{n+1}\right). \qquad \square\end{aligned}$$

3.4. Existence of simultaneously interpolating and sampling sequences in the one-dimensional setting. We conclude the paper by recalling some classical facts which are special for the one dimensional setting.

Let μ be a finite measure on $\mathbb{R}$ with compact set K and assume that μ has the Bernstein-Markov property with respect to K. By the classical Christoffel-Darboux formula there exists constants a_{k+1} such that

$$K_k(x,y) = a_{k+1} \frac{q_{k+1}(x)q_k(y) - q_k(x)q_{k+1}(y)}{x-y}$$

where q_{k+1} is the kth orthogonal polynomial (with respect to μ). Let $\Lambda_k := \{x_j^{(k)}\}$ be the set of $k+1$ zeros of q_k (which by classical results are indeed all distinct and contained in the support K of μ). Then $K_k(x_i^{(k)}, x_j^{(k)}) = 0$ if $i \neq j$, as follows immediately from the Christoffel-Darboux formula. Hence, normalizing $K_k(\cdot, x_j^{(k)})$ yields an orthonormal base in $H_k(M,\mu)$ and as a consequence the following "sampling equality" holds for any $p_k \in H_k(M)$:

$$\int_M |p_k|^2 d\mu = \sum_i \frac{1}{B_k\left(x_i^{(k)}\right)} \left|p_k\left(x_i^{(k)}\right)\right|^2,$$

and in particular the sequence Λ_k is both sampling and interpolation. Finally, recall that by classical results the normalized Dirac measure δ_k on the zeros Λ_k has the same weak limit points as $B_k/(k+1)\mu$. In particular, if μ has the Bernstein-Markov property, then $\frac{1}{k}\sum \delta_k \to \mu_{eq}$, which is thus consistent with Theorem 1 and Theorem 5.

DEPARTMENT OF MATHEMATICAL SCIENCES, CHALMERS UNIVERSITY OF TECHNOLOGY AND UNIVERSITY OF GOTHENBURG, 412 96 GÖTEBORG, SWEDEN
E-mail: robertb@chalmers.se

DEPT. MATEMÀTICA APLICADA I ANÀLISI, UNIVERSITAT DE BARCELONA AND BGS-MATH, GRAN VIA 585, 08007 BARCELONA, SPAIN
E-mail: jortega@ub.edu

REFERENCES

[1] E. Bedford and B. A. Taylor, The complex equilibrium measure of a symmetric convex set in $\mathbf{R}^n$, *Trans. Amer. Math. Soc.* **294** (1986), no. 2, 705–717.

[2] R. Berman, S. Boucksom, and D. Witt Nyström, Fekete points and convergence towards equilibrium measures on complex manifolds, *Acta Math.* **207** (2011), no. 1, 1–27.

[3] R. J. Berman, Bergman kernels and equilibrium measures for line bundles over projective manifolds, *Amer. J. Math.* **131** (2009), no. 5, 1485–1524.

[4] ———, Determinantal point processes and fermions on complex manifolds: Bulk universality, preprint, https://arxiv.org/abs/0811.3341.

[5] B. Berndtsson, Bergman kernels related to Hermitian line bundles over compact complex manifolds, *Explorations in Complex and Riemannian Geometry*, *Contemp. Math.*, vol. 332, Amer. Math. Soc., Providence, RI, 2003, pp. 1–17.

[6] L. Bos, N. Levenberg, P. Milman, and B. A. Taylor, Tangential Markov inequalities characterize algebraic submanifolds of $\mathbf{R}^N$, *Indiana Univ. Math. J.* **44** (1995), no. 1, 115–138.

[7] T. Bouche, Convergence de la métrique de Fubini-Study d'un fibré linéaire positif, *Ann. Inst. Fourier (Grenoble)* **40** (1990), no. 1, 117–130.

[8] A. Brudnyi, Local inequalities for plurisubharmonic functions, *Ann. of Math. (2)* **149** (1999), no. 2, 511–533.

[9] D. Burns, N. Levenberg, S. Ma'u, and Sz. Révész, Monge-Ampère measures for convex bodies and Bernstein-Markov type inequalities, *Trans. Amer. Math. Soc.* **362** (2010), no. 12, 6325–6340.

[10] I. Daubechies, *Ten Lectures on Wavelets*, *CBMS-NSF Regional Conf. Ser. in Appl. Math.*, vol. 61, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[11] J.-P. Demailly, Mesures de Monge-Ampère et caractérisation géométrique des variétés algébriques affines, *Mém. Soc. Math. France (N.S.)* (1985), no. 19, 124.

[12] ———, Potential theory in several complex variables, lecture notes, 1989.

[13] R. J. Duffin and A. C. Schaeffer, A class of nonharmonic Fourier series, *Trans. Amer. Math. Soc.* **72** (1952), 341–366.

[14] H. G. Feichtinger, K. Gröchenig, and T. Strohmer, Efficient numerical methods in non-uniform sampling theory, *Numer. Math.* **69** (1995), no. 4, 423–440.

[15] V. Guillemin and M. Stenzel, Grauert tubes and the homogeneous Monge-Ampère equation, *J. Differential Geom.* **34** (1991), no. 2, 561–570.

[16] A. Kroó and D. S. Lubinsky, Christoffel functions and universality in the bulk for multivariate orthogonal polynomials, *Canad. J. Math.* **65** (2013), no. 3, 600–620.

[17] H. J. Landau, Necessary density conditions for sampling and interpolation of certain entire functions, *Acta Math.* **117** (1967), 37–52.

[18] N. Lev and J. Ortega-Cerdà, Equidistribution estimates for Fekete points on complex manifolds, *J. Eur. Math. Soc. (JEMS)* **18** (2016), no. 2, 425–464.

[19] D. S. Lubinsky, Bulk universality holds in measure for compactly supported measures, *J. Anal. Math.* **116** (2012), 219–253.

[20] J. Marzo, Marcinkiewicz-Zygmund inequalities and interpolation by spherical harmonics, *J. Funct. Anal.* **250** (2007), no. 2, 559–587.

[21] P. G. Nevai, Orthogonal polynomials, *Mem. Amer. Math. Soc.* **18** (1979), no. 213, v+185.

[22] S. Nitzan and A. Olevskii, Revisiting Landau's density theorems for Paley-Wiener spaces, *C. R. Math. Acad. Sci. Paris* **350** (2012), no. 9-10, 509–512.

[23] J. Ortega-Cerdà and B. Pridhnani, Beurling-Landau's density on compact manifolds, *J. Funct. Anal.* **263** (2012), no. 7, 2102–2140.

[24] J. Ortega-Cerdà and J. Saludes, Marcinkiewicz-Zygmund inequalities, *J. Approx. Theory* **145** (2007), no. 2, 237–252.

[25] A. Sadullaev, An estimate for polynomials on analytic sets, *Math. USSR Izv.* **20** (1983), 493–502.

[26] K. Seip, *Interpolation and Sampling in Spaces of Analytic Functions*, *Univ. Lecture Ser.*, vol. 33, American Mathematical Society, Providence, RI, 2004.

[27] G. Tian, On a set of polarized Kähler metrics on algebraic manifolds, *J. Differential Geom.* **32** (1990), no. 1, 99–130.

[28] V. Totik, Asymptotics for Christoffel functions for general measures on the real line, *J. Anal. Math.* **81** (2000), 283–303.

[29] Y. Xu, On multivariate orthogonal polynomials, *SIAM J. Math. Anal.* **24** (1993), no. 3, 783–794.

[30] A. Zériahi, Inegalités de Markov et développement en série de polynômes orthogonaux des fonctions C^∞ et A^∞, *Several Complex Variables (Stockholm, 1987/1988)*, *Math. Notes*, vol. 38, Princeton Univ. Press, Princeton, NJ, 1993, pp. 683–701.