



ELSEVIER

Contents lists available at SciVerse ScienceDirect

Journal of Complexity

journal homepage: www.elsevier.com/locate/jco

Marcinkiewicz–Zygmund measures on manifolds

F. Filbir^a, H.N. Mhaskar^{b,*}^a Institute of Biomathematics and Biometry, Helmholtz Center Munich, 85764 Neuherberg, Germany^b Department of Mathematics, California State University, Los Angeles, CA, 90032, USA

ARTICLE INFO

Article history:

Received 24 August 2010

Accepted 3 March 2011

Available online 10 March 2011

Keywords:

Data defined manifolds

Marcinkiewicz–Zygmund inequalities

Quadrature formulas

Discretization inequalities

ABSTRACT

Let \mathbb{X} be a compact, connected, Riemannian manifold (without boundary), ρ be the geodesic distance on \mathbb{X} , μ be a probability measure on \mathbb{X} , and $\{\phi_k\}$ be an orthonormal (with respect to μ) system of continuous functions, $\phi_0(x) = 1$ for all $x \in \mathbb{X}$, $\{\ell_k\}_{k=0}^{\infty}$ be a nondecreasing sequence of real numbers with $\ell_0 = 1$, $\ell_k \uparrow \infty$ as $k \rightarrow \infty$, $\Pi_L := \text{span}\{\phi_j : \ell_j \leq L\}$, $L \geq 0$. We describe conditions to ensure an equivalence between the L^p norms of elements of Π_L with their suitably discretized versions. We also give intrinsic criteria to determine if any system of weights and nodes allows such inequalities. The results are stated in a very general form, applicable for example, when the discretization of the integrals is based on weighted averages of the elements of Π_L on geodesic balls rather than point evaluations.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

To avoid complicating our notations unnecessarily, the notations used in the introduction and the next section might have a different meaning from the rest of the paper.

The classical Marcinkiewicz–Zygmund (MZ) inequality states the following [31, Chapter X, Theorem (7.5)]: Let $n \geq 1$ be an integer, and S be a trigonometric polynomial of order at most n (i.e., an expression of the form $\sum_{|k| \leq n} c_k e^{ikx}$). If $1 < p < \infty$, then

$$\int_0^{2\pi} |S(x)|^p dx \leq \frac{A_p}{2n+1} \sum_{k=0}^{2n} \left| S\left(\frac{2k\pi}{2n+1}\right) \right|^p \leq AA_p \int_0^{2\pi} |S(x)|^p dx, \quad (1.1)$$

where A is an absolute, positive constant, and A_p is a positive constant depending only on p . We observe that the number of points in the summation is the same as the dimension of the space of

* Corresponding author.

E-mail addresses: filbir@helmholtz-muenchen.de (F. Filbir), hmhaska@gmail.com (H.N. Mhaskar).

all trigonometric polynomials of order at most n . The second inequality in (1.1) is valid for $p = 1, \infty$ as well. The first inequality holds for $p = 1, \infty$ if one allows more points in the summation than the dimension $2n + 1$ [31, Chapter X, Theorem (7.28)]. These inequalities are also known as large sieve inequalities or network inequalities. Inequalities of this form have many applications in approximation theory, number theory, signal processing, etc. Therefore, several analogs of these inequalities have been studied in the literature, for example, in the setting where S is a univariate algebraic polynomial, the integrals in (1.1) are replaced by weighted or Lebesgue–Stieltjes integrals on real intervals, and the weights and sampling nodes in the sum in (1.1) are chosen judiciously. A survey of many of the classical results in this direction and their applications can be found in the paper [25] by Lubinsky.

Many modern applications require an analysis of huge, unstructured, high dimensional data sets, which are not dense on any cube, unlike classical approximation theory scenarios. Coifman and his collaborators have recently introduced diffusion geometry techniques for this purpose; see [3] for an introduction. The basic idea is to assume that the data lies on an unknown low dimensional manifold. The current paper is motivated by a desire to study the analogs of MZ inequalities on manifolds and their role in data based approximation in further detail, and in a somewhat greater abstraction.

To explain the connection between MZ inequalities and data processing, we consider an example. Let \mathbb{X} be a compact, connected, Riemannian manifold (without boundary), and μ be the volume measure on \mathbb{X} . There is a positive semi-definite elliptic differential operator Δ^* on the manifold, called the Laplace–Beltrami operator (cf. (2.2)). Let $\{\phi_k\}$ be an orthonormal (with respect to μ) system of eigenfunctions of Δ^* , $\{\ell_k\}_{k=0}^\infty$ be a nondecreasing sequence of real numbers with $\ell_0 = 1, \ell_k \uparrow \infty$ as $k \rightarrow \infty$, such that $\Delta^* \phi_k = \ell_k^2 \phi_k, k = 0, 1, \dots$, and $\Pi_L := \text{span}\{\phi_j : \ell_j \leq L\}, L \geq 0$. For each $L \geq 1$, let \mathcal{C}_L be a finite subset of $\mathbb{X}, W_L = \{w_y\}_{y \in \mathcal{C}_L} \subset \mathbb{R}$. Let $1 \leq p \leq \infty$. For the purpose of this introduction, we will say that $\{(\mathcal{C}_L, W_L)\}$ is a MZ system of order p if for every $L \geq 1$ and $P \in \Pi_L$,

$$c_1 \int_{\mathbb{X}} |P(x)|^p d\mu(x) \leq \sum_{y \in \mathcal{C}_L} |w_y| |P(y)|^p \leq c_2 \int_{\mathbb{X}} |P(x)|^p d\mu(x), \tag{1.2}$$

where c_1, c_2 are positive constants depending only on $\mathbb{X}, \{\phi_k\}, \{\ell_k\}, \mu$, and p , but independent of L and P , and the customary interpretation is assumed when $p = \infty$. Given that our new object of study is the discrete sum in the middle of (1.2), we will refer to the (second) inequality giving an upper estimate of this sum as the upper inequality, and the (first) inequality giving a lower estimate for this sum as the lower inequality. We list four of the applications of such inequalities, which have inspired our own interest in these.

First, suppose one wants to approximate a continuous function $f : \mathbb{X} \rightarrow \mathbb{R}$. A standard way to do this is by means of an operator

$$T_L(f, x) = \int_{\mathbb{X}} f(y) \Phi_L(x, y) d\mu(y),$$

where Φ_L is a suitable kernel with the property that it is in Π_L as a function of x and as a function of y . A typical example is the Fourier projection, where $\Phi_L(x, y) = \sum_{\ell_j \leq L} \phi_j(x) \phi_j(y)$ is the reproducing kernel for Π_L . If the approximation is required using the values of f at points in \mathcal{C}_L , then it is natural to consider the discretization

$$T_L^D(f, x) = \sum_{y \in \mathcal{C}_L} w_y f(y) \Phi_L(x, y).$$

In the case when T_L is the Fourier projection, such a discretization has been called hyperinterpolation. It is easy to verify that the operator norms of T_L, T_L^D are given by

$$\sup_{x \in \mathbb{X}} \int_{\mathbb{X}} |\Phi_L(x, y)| d\mu(y), \quad \sup_{x \in \mathbb{X}} \sum_{y \in \mathcal{C}_L} |w_y| |\Phi_L(x, y)|$$

respectively. Since $\Phi_L(x, \circ) \in \Pi_L$ for each $x \in \mathbb{X}$, (1.2) with $p = 1$ implies that the operator norms of T_L, T_L^D have the same order of magnitude as functions of L , as $L \rightarrow \infty$. In this context, it is not

necessary that the weights w_y should be all nonnegative, a fact which may be useful in numerical computations. In particular, in the case when \mathbb{X} is a Euclidean sphere, Φ_L is the reproducing kernel for the space Π_L of spherical polynomials of degree at most L , then this leads to a simple proof of the estimates on the norm of the hyperinterpolation operators on the sphere, obtained by Sloan, Reimer, and others (cf. [17, Section 3.2] for a review). We are also aware of a work [5] by Damelin and Levesley on similar questions on hyperinterpolation in the context of projective spaces.

The second application illustrates the use of (1.2) when $p = \infty$. Suppose that the sampling nodes \mathcal{C}_L are chosen randomly. Using probabilistic estimates, one is sometimes able to estimate the probability that the quantity $\sum_{y \in \mathcal{C}_L} |w_y| |\Phi_L(x, y)|$ exceeds a given threshold for any fixed $x \in \mathbb{X}$. The inequality (1.2) with $p = \infty$ then enables one to estimate the probability that the operator norm of T_L^D is bounded. An example of this argument can be found in [24].

The third application concerns least squares approximation. If one wishes to obtain $Q \in \Pi_L$ so as to minimize $\sum_{y \in \mathcal{C}_L} |w_y| |f(y) - Q(y)|^2$, then one has to solve a system of linear equations with the matrix (the Gram matrix) G whose (j, k) -th entry $G_{j,k}$ is given by $\sum_{y \in \mathcal{C}_L} |w_y| \phi_j(y) \phi_k(y)$. In view of the Rayleigh–Ritz theorem [19, Theorem 4.2.2, p. 176], the lowest and highest eigenvalues of this matrix are given respectively by the infimum and supremum of the quotients

$$\frac{\sum_{j,k} a_j a_k G_{j,k}}{\sum_j a_j^2}$$

over all $a_j \in \mathbb{R}$, $j = 1, \dots, \dim(\Pi_L)$. Let $P = \sum_j a_j \phi_j$. Then the denominator expression above is equal to $\int_{\mathbb{X}} |P(x)|^2 d\mu(x)$. It is easy to check that the numerator expression is equal to $\sum_{y \in \mathcal{C}_L} |w_y| |P(y)|^2$. Thus, if the weights $|w_y|$'s are chosen so as to satisfy (1.2) with $p = 2$, then the lowest (respectively, the highest) eigenvalue is estimated from below (respectively, from above) by c_1 (respectively, c_2). In particular, the closer the ratio c_2/c_1 is to 1, the better conditioned is the matrix G .

Finally, we have demonstrated in a number of papers starting with [28] that the inequalities (1.2) with $p = 1$ and c_1, c_2 sufficiently close to 1 lead to the existence of positive quadrature formulas exact for integration of elements of Π_{cL} for some constant c .

To continue with the motivation behind the current paper, Lafon [22] has shown that certain positive definite matrices constructed from the mutual distances among the observed data points on the manifold with some tuning parameters converge to the “heat kernel” on the manifold, defined formally by

$$K_t(x, y) = \sum_{k=0}^{\infty} \exp(-\ell_k^2 t) \phi_k(x) \overline{\phi_k(y)}. \tag{1.3}$$

However, the data may not be sampled uniformly according to the volume measure on the manifold. In this case, the limiting kernel is not the classical heat kernel, but only some positive–semidefinite operator. Accordingly, in the representation above, one has to consider a more general measure μ , and the case when $\ell_k \uparrow \infty$ and the eigenfunctions $\{\phi_k\}$'s are orthonormal with respect to μ . This heat kernel has been used by Coifman and Maggioni [4] to define a metric on the data manifold, as well as to construct a multiresolution analysis on the manifold. In a more recent work [20], Jones, Maggioni, and Schul have demonstrated that the classical heat kernel can be used to construct a local coordinate chart on the unknown manifold.

Thus, even though the manifold is unknown, it is reasonable to assume for theoretical investigations that one knows a semi–group of positive definite kernels on the manifold, or equivalently (from a theoretical point of view) that one knows its infinitesimal generator, called the Laplacian on the unknown manifold. In [26], we started to develop a detailed theory of function approximation based on the eigenfunctions of the heat kernel. Our assumptions in [26] were formulated in terms of the behavior of the sums of the form $\sum_{\ell_k \leq L} |\phi_k(x)|^2$ and the so-called finite speed of wave propagation. However, since the actual manifold is unknown, and the heat kernel is the only easily computable quantity, we find it important in theoretical considerations to formulate the assumptions behind various theorems in terms of the heat kernel as far as possible. In [11], the

properties of a summability kernel which plays a critical role in this theory were formulated purely in terms of the heat kernel, and generalized to obtain Marcinkiewicz–Zygmund inequalities, Bernstein inequalities, and the existence of quadrature formulas.

This paper is devoted to a more detailed study of Marcinkiewicz–Zygmund inequalities in the case when \mathbb{X} is a smooth manifold, where the “heat kernel” based on the orthonormal system $\{\phi_j\}$ satisfies certain properties. In particular, our theory is valid when these are the eigenfunctions of the Laplace–Beltrami operator on the manifold, as well as in the case when they are eigenfunctions of certain weighted Laplace–Beltrami operators and a large class of other second order elliptic operators. We will establish conditions under which inequalities of the form (1.2) hold for all p , $1 \leq p \leq \infty$ with $c_1 = 1 - \eta$ and $c_2 = 1 + \eta$ for a prescribed η . Such inequalities were proved in [11] in the case when $p = 1, \infty$, and applied to obtain the existence of quadrature formulas. However, a straightforward application of the classical Riesz–Thorin interpolation theorem is not sufficient to prove these inequalities in such a sharp form for $1 < p < \infty$. We will prove an alternative form of this interpolation theorem, which appears to be new. We will also give intrinsic characterizations of the systems $\{(C_L, W_L)\}$ without reference to the system $\{\phi_j\}$ which are equivalent to the inequalities of the form (1.2) (without the requirement that c_1, c_2 should be arbitrarily close to 1). For example, we will show that the upper inequality in (1.2) holds if and only if for any $x \in \mathbb{X}$ and any geodesic ball $B(x, r)$ of radius $r > 0$ centered at x ,

$$\sum_{y \in B(x,r) \cap C_L} |w_y| \leq c \mu(B(x, r + 1/L)),$$

for some constant $c > 0$. Similar results will be proved for the lower inequality. Our aim is to provide such results for a very general setting, in particular including the case when the middle term in (1.2) involves weighted averages of elements of Π_L on balls rather than their values at finitely many points. In this paper, the heat kernel will play a somewhat indirect role. We will work with the very general setting of an arbitrary orthonormal system $\{\phi_k\}$ and sequence $\ell_k \uparrow \infty$. We will be using mainly the results in [11]. In turn, these are proved under the assumptions formulated in terms of the heat kernel defined formally by (1.3).

In Section 2, we will review a few basic facts regarding Riemannian manifolds in general, which will be needed in the rest of the paper. In Section 3, we discuss the various assumptions on the manifold, the measure, and the systems $\{\phi_j\}, \{\ell_j\}$ via the heat kernel. In Section 4, we introduce the abstract notions which enable us to generalize the theory from point evaluations to other measures. The main results are stated in Section 5, and proved in Section 7. These proofs require us to develop and review certain preparatory material, which is presented in Section 6.

2. Riemannian manifolds

The purpose of this section is to review some facts and terminology regarding Riemannian manifolds. We will avoid very technical details, which can be found in such standard texts as [1,8,9,29]. The material in this section is based mostly on [9], and is essentially the same as the Appendix to our paper [11].

Let $q \geq 1$ be an integer. A *differentiable manifold* of dimension q is a set \mathbb{X} and a family of injective mappings $\mathbf{x}_\alpha : U_\alpha \subset \mathbb{R}^q \rightarrow \mathbb{X}$ of open sets U_α into \mathbb{X} such that (i) $\cup_\alpha \mathbf{x}_\alpha(U_\alpha) = \mathbb{X}$, (ii) for any pair α, β , with $\mathbf{x}_\alpha(U_\alpha) \cap \mathbf{x}_\beta(U_\beta) = W$ being nonempty, the sets $\mathbf{x}_\alpha^{-1}(W)$ and $\mathbf{x}_\beta^{-1}(W)$ are open subsets of \mathbb{R}^q , and the mapping $\mathbf{x}_\beta^{-1} \circ \mathbf{x}_\alpha$ is (infinitely) differentiable on $\mathbf{x}_\alpha^{-1}(W)$, (iii) the family (atlas) $\mathcal{A}_\mathbb{X} = \{(U_\alpha, \mathbf{x}_\alpha)\}$ is maximal relative to the conditions (i) and (ii). The pair $(U_\alpha, \mathbf{x}_\alpha)$ (respectively, \mathbf{x}_α) with $x \in \mathbf{x}_\alpha(U_\alpha)$ is called a parametrization or coordinate chart (respectively, a system of coordinates) of \mathbb{X} around x , and $\mathbf{x}_\alpha(U_\alpha)$ is called a coordinate neighborhood of x . In the sequel, the term differentiable will mean infinitely many times differentiable. We assume also that \mathbb{X} is Hausdorff and has a countable basis as a topological space.

Intuitively, one thinks of a differentiable manifold as a surface in an ambient Euclidean space. The abstract definition above is intended to overcome the technical need for the ambient space. For all applications of our theory that we can imagine, and in particular, for an intuitive comprehension of

our paper, there is no loss in thinking of a manifold as a surface. Moreover, a theorem of Whitney [9, p. 30] provides a further justification of such a viewpoint.

Let \mathbb{X}, \mathbb{Y} be two differentiable manifolds. A mapping $f : \mathbb{X} \rightarrow \mathbb{Y}$ is called differentiable on an open set $W \subseteq \mathbb{X}$ if there exist for every $x \in W$ coordinate charts $(U, \mathbf{x}) \in \mathcal{A}_{\mathbb{X}}, (V, \mathbf{y}) \in \mathcal{A}_{\mathbb{Y}}$ with $x \in U, f(U) \subseteq V$ such that $\mathbf{y}^{-1} \circ f \circ \mathbf{x}$ is a C^∞ function. In particular, a curve in \mathbb{X} is a differentiable mapping from an interval in \mathbb{R} to \mathbb{X} . The restriction of a curve γ to a compact subinterval $[a, b]$ of I is called a curve segment, joining $\gamma(a)$ to $\gamma(b)$. We may define a piecewise differentiable curve on a manifold \mathbb{X} in an obvious manner.

If $x \in \mathbb{X}, \epsilon > 0$, and $\gamma : (-\epsilon, \epsilon) \rightarrow \mathbb{X}$ is a curve with $x = \gamma(0)$, then the tangent vector to γ at $\gamma(t_0)$ is defined to be the functional $\gamma'(t_0)$ acting on the class of all differentiable $f : \mathbb{X} \rightarrow \mathbb{R}$ by

$$\gamma'(t_0)f = \left. \frac{d(f \circ \gamma)}{dt} \right|_{t=t_0}.$$

The set of all such functionals $\gamma'(0)$ defines a vector space, called the tangent space of \mathbb{X} at x , denoted by $T_x\mathbb{X}$. Let (U, \mathbf{x}) be a coordinate chart such that $0 \in U$ and $x = \mathbf{x}(0)$, and for $j = 1, \dots, q, \partial_j(x)$ be the tangent vector at x to the coordinate curve $x_j \rightarrow (0, \dots, x_j, 0, \dots, 0)$. Then $\{\partial_j(x)\}$ is a basis for $T_x\mathbb{X}$. In particular, the dimension of $T_x\mathbb{X}$ is q . The set $\{(x, v) : x \in \mathbb{X}, v \in T_x\mathbb{X}\}$ is called the tangent bundle of \mathbb{X} , and can be endowed with the structure of a differentiable manifold of dimension $2q$. A vector field F on \mathbb{X} is a mapping that assigns to each $x \in \mathbb{X}$ a vector $F(x) \in T_x\mathbb{X}$ such that for every differentiable function g on \mathbb{X} , the mapping $x \mapsto F(x)g$ is differentiable. If G is another vector field, we may apply $G(x)$ to this mapping, obtaining thereby a second order vector field $G \circ F$. A derivative of higher order can be defined similarly.

A Riemannian metric on a differentiable manifold \mathbb{X} is given by a scalar product $\langle \circ, \circ \rangle_x$ on each $T_x\mathbb{X}$ which depends smoothly on the base point x , i.e. the function $\mathbb{X} \rightarrow \mathbb{R}, x \mapsto \langle X(x), Y(x) \rangle_x$ is $C^\infty(\mathbb{X})$. A manifold with a given Riemannian metric is called a Riemannian manifold. Let $g_{i,j} = \langle \partial_i(x), \partial_j(x) \rangle_x$ and denote by g the matrix $(g_{i,j})$. The entries of g^{-1} are denoted by $g^{i,j}$. The Riemannian metric on \mathbb{X} allows one to define a notion of length of a curve segment as well as the volume element (Riemannian measure) on \mathbb{X} . First, if F is a vector field on \mathbb{X} , we may define $\|F\|_x^2 := \langle F(x), F(x) \rangle_x$. The length of a differentiable curve $\gamma : [0, 1] \rightarrow \mathbb{X}$ is defined as $\mathcal{L}(\gamma) = \int_0^1 \| \gamma'(t) \|_{\gamma(t)} dt$. A differentiable curve $\gamma : [0, 1] \rightarrow \mathbb{X}$, such that the length of γ does not exceed that of any other piecewise differentiable curve joining $\gamma(0)$ to $\gamma(1)$, is called a geodesic [9, Proposition 3.6, Corollary 3.9].

The gradient of a function $f \in C^\infty(\mathbb{X})$ is a vector field defined by

$$\nabla f = \sum_{j=1}^q \sum_{i=1}^q g^{i,j} \partial_i f \partial_j.$$

For the gradient field we have

$$\langle (\nabla f)_x, F(x) \rangle_x = (Ff)(x) \tag{2.1}$$

for every vector field F . The divergence of a vector field $F = \sum_{j=1}^q F_j \partial_j$ is defined by

$$\text{div}(F) = \frac{1}{\sqrt{\det(g)}} \sum_{j=1}^q \partial_j \left(\sqrt{\det(g)} F_j \right).$$

The Laplace–Beltrami operator $\Delta^* f(x)$ is defined as the differential operator given by

$$\Delta^* f = -\text{div}(\nabla f) = \frac{-1}{\sqrt{\det(g)}} \sum_{j=1}^q \sum_{i=1}^q \partial_j \left(\sqrt{\det(g)} g^{i,j} \partial_i f \right). \tag{2.2}$$

The operator Δ^* is an elliptic operator. Therefore, in the case when \mathbb{X} is a compact connected manifold, the existence of a discrete spectrum and system of orthonormal eigenfunctions follows from the general theory of partial differential equations [30, Chapter 5.1].

We will have no further occasion to refer to the dimension of the manifold, and therefore, will use the symbol q with different meanings in the rest of this paper.

3. Assumptions

Let \mathbb{X} be a compact, connected, Riemannian manifold (without boundary), ρ be the geodesic distance on \mathbb{X} , μ be a probability measure on \mathbb{X} , and $\{\phi_k\}$ be an orthonormal system of smooth functions, $\phi_0(x) = 1$ for all $x \in \mathbb{X}$, and $\{\ell_k\}_{k=0}^\infty$ be a nondecreasing sequence of real numbers with $\ell_0 = 1, \ell_k \uparrow \infty$ as $k \rightarrow \infty$. For $L \geq 0$, the space $\text{span}\{\phi_j : \ell_j \leq L\}$ will be denoted by Π_L , and its members will be called *diffusion polynomials* of degree at most L . We will also write $\Pi_\infty = \cup_{L \geq 0} \Pi_L$.

We observe that μ is not necessarily the volume measure on \mathbb{X} . Similarly, the systems ℓ_k and ϕ_k are quite general, not necessarily the eigenvalues or eigenfunctions of some elliptic differential operator. In particular, the space Π_L will depend upon both of these systems. For example, if \mathbb{X} is the Euclidean unit sphere embedded in \mathbb{R}^3 , then the eigenvalues ℓ_k of the Laplace–Beltrami operator are of the form $\sqrt{j(j+1)}$ for some j depending the enumeration of the eigenvalues. It might be worthwhile to use j instead as the corresponding ℓ_k . This will change the space Π_L somewhat. However, estimates such as (6.12) on the summability kernels defined in (6.9) can then be deduced easily by well-known formulas connecting these with the Césaro operators.

We need to formulate some assumptions on these various quantities and their relationships with each other.

For $x \in \mathbb{X}, r > 0$, let

$$B(x, r) := \{y \in \mathbb{X} : \rho(x, y) \leq r\}, \quad \Delta(x, r) := \mathbb{X} \setminus B(x, r).$$

Assumption 1. We assume that there exist constants $\kappa_1, \alpha > 0$ such that

$$\mu(\{y \in \mathbb{X} : \rho(x, y) < r\}) = \mu(B(x, r)) \leq \kappa_1 r^\alpha, \quad x \in \mathbb{X}, r > 0. \tag{3.1}$$

In the case when μ is the volume measure on \mathbb{X} , then one may choose α to be the dimension of the manifold. In general, there may not be such a connection. In light of the first equation in (3.1), we will refer to two balls being disjoint to mean that their intersection is a μ -null set.

Next, we discuss the notion of the heat kernel, and the assumptions on the same. For $t > 0$, the *heat kernel* is defined formally by

$$K_t(x, y) = \sum_{k=0}^\infty \exp(-\ell_k^2 t) \phi_k(x) \overline{\phi_k(y)}. \tag{3.2}$$

Assumption 2. We assume the existence of constants $\kappa_2, \kappa_3 > 0$ such that

$$|K_t(x, y)| \leq \kappa_2 t^{-\alpha/2} \exp(-\kappa_3 \rho(x, y)^2/t), \quad x, y \in \mathbb{X}, t \in (0, 1], \tag{3.3}$$

and

$$\|\nabla_y K_t(x, y)\|_x \leq \kappa_2 t^{-\alpha/2-1} \exp(-\kappa_3 \rho(x, y)^2/t), \quad x, y \in \mathbb{X}, t \in (0, 1], \tag{3.4}$$

where ∇_y indicates that the gradient is taken with respect to y . We assume further that for some constant $\kappa_4 > 0$,

$$K_t(x, x) \geq \kappa_4 t^{-\alpha/2}, \quad x \in \mathbb{X}, t \in (0, 1]. \tag{3.5}$$

Kordyukov [21] has proved that each of our assumptions above hold for the heat kernels, when for each $k = 0, 1, \dots, \phi_k$ is the eigenfunction of a second order elliptic operator corresponding to the eigenvalue ℓ_k^2 . The elliptic operator in question is assumed to satisfy some very general conditions, which are satisfied by the Laplace–Beltrami operator on a Riemannian manifold with “bounded geometry” (see [21] for definitions). Estimates on the heat kernel and its gradients are well-understood in many other cases, including higher order partial differential operators on manifolds [7,13,2,10], with many other references given in [14]. Clearly, Assumption 2 will be satisfied (with different constants) by any other choice of $\{\ell'_k\}$ for the system $\{\ell_k\}$ provided $\ell'_k \sim \ell_k$.

We will need one more assumption, which we include here for the sake of organizational clarity, even though it requires some notations introduced in (4.2) and the paragraph which follows (4.2). A system $\{\psi_k\} \subset L^2(\mu)$ will be called a *Bessel system* if there exists a dense subset $\mathcal{D} = \mathcal{D}(\{\psi_k\})$ of $C(\mathbb{X})$ (with respect to the norm of this space), such that (i) for any $\epsilon > 0$, ball of the form $B(x, r)$, and $f \in C(\mathbb{X})$ supported on $B(x, r)$, there exists $g \in \mathcal{D}$ such that the support of g is contained in $B(x, 2r)$ and

$$\int_{\mathbb{X}} |f(x) - g(x)| d\mu(x) \leq \epsilon,$$

and (ii)

$$\sum_{k=0}^{\infty} |f, \psi_k|^2 \leq \mathcal{N}(f) < \infty, \quad f \in \mathcal{D}, \tag{3.6}$$

where $\mathcal{N}(f)$ is a positive number dependent on f , (\circ, \circ) and $\{\psi_k\}$. Obviously, any orthonormal system on \mathbb{X} is a Bessel system with $\mathcal{D} = C(\mathbb{X})$. Another interesting example is the following. Let \mathbb{X} be a Riemannian manifold, μ be the Riemannian volume measure on \mathbb{X} , $\{\phi_k\}$ be the eigenfunctions of the Laplace–Beltrami operator on \mathbb{X} , F be a vector field on \mathbb{X} , and ψ_k be defined by $\psi_k = F\phi_k$; i.e., $\psi_k(x) = F(x)(\phi_k)$, $x \in \mathbb{X}, k = 0, 1, \dots$. For the space \mathcal{D} we choose the class of all compactly supported, infinitely differentiable functions on \mathbb{X} . Using the Green’s formula [23, p. 383], we obtain for any $f \in \mathcal{D}$, and $k = 0, 1, \dots$,

$$\int_{\mathbb{X}} \psi_k(x) f(x) d\mu(x) = - \int_{\mathbb{X}} \phi_k(x) \operatorname{div}(fF)(x) d\mu(x). \tag{3.7}$$

Thus, $\{\psi_k\}$ is a Bessel family with $\mathcal{N}(f) = \|\operatorname{div}(fF)\|_{\mu;2}$. A similar fact obtains also when one considers $w d\mu$ instead of $d\mu$ for some smooth positive valued function w .

Assumption 3. We assume that for any vector field F , the system $\{F\phi_k\}$ is a Bessel system.

Constant convention:

In the sequel, the symbols c, c_1, \dots will denote positive constants depending only on $\mathbb{X}, \rho, \mu, \kappa_1, \dots, \kappa_4$, and other similar fixed quantities, but not on the systems $\{\phi_k\}, \{\ell_k\}$, nor any other variables not explicitly indicated. Their values may be different at different occurrences, even within a single formula. The notation $A \sim B$ will mean $c_1 A \leq B \leq c_2 A$.

We note some consequences for our assumptions. We have proved [11, Proposition 4.1], [27, Lemma 5.2] that the conditions (3.3) with $x = y$ and (3.5) are equivalent to

$$\sum_{\ell_j \leq L} |\phi_j(x)|^2 \sim L^\alpha. \tag{3.8}$$

We note that the conditions $\phi_0(x) \equiv 1$, and $\ell_0 = 1$ imply that

$$\int_{\mathbb{X}} K_t(x, y) d\mu(y) = 1, \quad x \in \mathbb{X}, t \in (0, 1]. \tag{3.9}$$

In [16], Grigor’yan has proved that (3.1), (3.9) and (3.3) together imply that

$$\mu(B(x, r)) \geq cr^\alpha, \quad 0 < r \leq 1, x \in \mathbb{X}. \tag{3.10}$$

Using (3.1) and (3.10), we obtain that μ satisfies the homogeneity condition

$$\mu(B(x, R)) \leq cc_1^\alpha (R/r)^\alpha \mu(B(x, r)), \quad x \in \mathbb{X}, 0 < r \leq c_1, R > 0, c_1 \geq 1. \tag{3.11}$$

4. MZ measures

In this section, we wish to express the ideas in the introduction in a more abstract and formal manner. First, it is cumbersome to write a sum of the form $\sum_{y \in \mathcal{C}} w_y f(y)$. To write such a sum, we

need to introduce the set \mathcal{C} and the weights $W = \{w_y\}$. The precise choice of these objects plays no role in our theory. Moreover, it makes it difficult to prescribe the dependence of various constants on the set \mathcal{C} and the weights W . For these reasons we prefer to use the Lebesgue–Stieltjes integral notation $\int_{\mathbb{X}} f(y) d\nu(y)$ to denote this sum, where ν is the measure that associates the mass w_y with the point y , $y \in \mathcal{C}$; i.e., for $B \subset \mathbb{X}$,

$$\nu(B) = \sum_{y \in B} w_y. \tag{4.1}$$

The notation has an additional advantage. When f is not continuous, but in some $L^p(\mu)$, then f cannot be defined everywhere. It is customary in such cases to consider averages of f on small balls around the points in \mathcal{C} . A weighted sum of these averages can be written in the form $\int_{\mathbb{X}} f(y) d\nu(y)$ as well, with a suitable choice of the measure ν . Further, some applications require the consideration of a weighted manifold (cf. [15]). Rather than dealing with the eigenfunctions for a weighted analog of the Laplace–Beltrami operator, one may wish to work with the eigenfunctions for the unweighted case. The MZ inequalities with the measure $d\nu = w d\mu$, where w is the weight function in question are expected to be useful in such situations. The class of all signed Borel measures on \mathbb{X} is a vector space, which will be denoted by \mathcal{M} .

If $\nu \in \mathcal{M}$, its total variation measure is defined for Borel measurable subsets $B \subset \mathbb{X}$ by

$$|\nu|(B) := \sup \sum_{i=1}^{\infty} |\nu(B_i)|,$$

where the supremum is taken over all countable partitions $\{B_i\}$ of B . For any signed measure ν , $|\nu|(\mathbb{X})$ is always a finite number. In the case when ν is the measure that associates the mass w_y with each $y \in \mathcal{C}$, one can easily deduce that $|\nu|(\mathbb{X}) = \sum_{y \in \mathcal{C}} |w_y|$. In the case when $d\nu = w d\mu$, one has $d|\nu| = |w| d\mu$. This includes the case when ν is a weighted sum of averages on balls. If $\nu \in \mathcal{M}$, the *support* of ν is defined by

$$\text{supp}(\nu) := \{x \in \mathbb{X} : |\nu|(B(x, r)) > 0 \text{ for every } r > 0\}.$$

In view of (3.10), $\text{supp}(\mu) = \mathbb{X}$.

Let ν be a signed measure on \mathbb{X} . If $B \subseteq \mathbb{X}$ is ν -measurable, and $f : B \rightarrow \mathbb{C}$ is a ν -measurable function, we will write

$$\|f\|_{\nu;p,B} := \begin{cases} \left\{ \int_B |f(x)|^p d|\nu|(x) \right\}^{1/p}, & \text{if } 0 < p < \infty, \\ |\nu| - \text{esssup}_{x \in B} |f(x)|, & \text{if } p = \infty. \end{cases} \tag{4.2}$$

The class of all f with $\|f\|_{\nu;p,B} < \infty$ will be denoted by $L^p(\nu; B)$, with the usual convention of considering two functions to be equal if they are equal $|\nu|$ -almost everywhere. If $B = \mathbb{X}$, we will omit its mention from the notations. The expressions $\|f\|_{\nu;p,B}$ are not norms if $p < 1$, but we prefer to continue using the same notation. The inner product of $L^2(\mu)$ will be denoted by $\langle \circ, \circ \rangle$. The $L^p(\mu)$ -closure of Π_{∞} will be denoted by $X^p(\mu)$. The class of all uniformly continuous and bounded functions on B , equipped with the uniform norm will be denoted by $C(B)$. If $1 < p < \infty$, the conjugate index p' is defined by $p' := p/(p - 1)$. We define $1' = \infty$ and $\infty' = 1$.

Thus, if $\mathcal{C} = \mathcal{C}_L$ is as in the introduction, and ν is the corresponding measure as defined in (4.1), then the inequalities (1.2) can be expressed in the concise form

$$c_1 \|P\|_{\mu;p} \leq \|P\|_{\nu;p} \leq c_2 \|P\|_{\mu;p}, \quad P \in \Pi_L. \tag{4.3}$$

Since Π_L is a finite dimensional space, such inequalities are always valid, but with the constants possibly depending on L . We are mostly interested in investigating the conditions under which these are independent of L , but wish to note another example where the constants depend polynomially on L . Since this provides an important rationale for considering general measures, apart from the wish to include averages over balls, we discuss this example in some detail. First, we note a property of diffusion polynomials, known as *Nikolskii inequalities*.

Proposition 4.1. Let $L > 0, 0 < p < r \leq \infty, P \in \Pi_L$. Then

$$\|P\|_{\mu;r} \leq cL^{\alpha(1/p-1/r)} \|P\|_{\mu;p}. \tag{4.4}$$

Proof. This proposition was proved in the case $p \geq 1$ in [27, Lemma 5.5]. In particular, we have

$$\|P\|_{\mu;\infty} \leq cL^\alpha \|P\|_{\mu;1}, \quad P \in \Pi_L.$$

Let $0 < p < 1, P \in \Pi_L$. Then using the above inequality, we obtain

$$\|P\|_{\mu;\infty} \leq cL^\alpha \int_{\mathbb{X}} |P(x)| d\mu(x) = cL^\alpha \int_{\mathbb{X}} |P(x)|^{1-p} |P(x)|^p d\mu(x) \leq cL^\alpha \|P\|_{\mu;\infty}^{1-p} \|P\|_{\mu;p}^p.$$

Therefore,

$$\|P\|_{\mu;\infty}^p = \|P\|_{\mu;\infty}^{p-1} \|P\|_{\mu;\infty} \leq cL^\alpha \|P\|_{\mu;p}^p.$$

This leads to (4.4) in the case when $r = \infty$ and $0 < p < 1$. If $0 < p < r < \infty$, then

$$\begin{aligned} \|P\|_{\mu;r}^r &= \int_{\mathbb{X}} |P(x)|^{r-p} |P(x)|^p d\mu(x) \leq \|P\|_{\mu;\infty}^{r-p} \|P\|_{\mu;p}^p \\ &\leq cL^{\alpha(r-p)/p} \|P\|_{\mu;p}^{r-p} \|P\|_{\mu;p}^p = cL^{\alpha r(1/p-1/r)} \|P\|_{\mu;p}^r. \end{aligned}$$

This implies (4.4) for all $r, p, 0 < p < r < \infty$. \square

Example 1. Let $w \geq 0$ μ -almost everywhere on \mathbb{X} , $1 < r < \infty$, and $w \in L^r(\mu)$. We define $dv = w d\mu$. Let $L > 0, P \in \Pi_L$, and $0 < p < \infty$. Using Hölder inequality followed by (4.4) with $pr' > p$ in place of r , we obtain

$$\begin{aligned} \|P\|_{v;p}^p &:= \int_{\mathbb{X}} |P(x)|^p w(x) d\mu \leq \|w\|_{\mu;r} \left\{ \int_{\mathbb{X}} |P(x)|^{pr'} d\mu(x) \right\}^{1/r'} \\ &= \|w\|_{\mu;r} \|P\|_{\mu;pr'}^p \leq cL^{\alpha p(1/p-1/(pr'))} \|w\|_{\mu;r} \|P\|_{\mu;p}^p = cL^{\alpha/r} \|w\|_{\mu;r} \|P\|_{\mu;p}^p. \end{aligned}$$

Thus, if $1 < r < \infty$ and $w \in L^r(\mu)$, then

$$\|P\|_{v;p} \leq c_1 L^{\alpha/(pr)} \|w\|_{\mu;r}^{1/p} \|P\|_{\mu;p}. \tag{4.5}$$

To obtain an inequality in the reverse direction, let $1 < q < \infty$, and $w^{-1} \in L^{q-1}(\mu)$. Using the Nikolskii inequality (4.4) first with p/q' in place of p and p in place of r , followed by Hölder inequality, we obtain

$$\begin{aligned} \|P\|_{\mu;p}^p &\leq c_2 L^{\alpha(q'-1)} \|P\|_{\mu;p/q'}^p = c_2 L^{\alpha/(q-1)} \left\{ \int_{\mathbb{X}} |P(x)|^{p/q'} w^{1/q'}(x) w^{-1/q'}(x) d\mu(x) \right\}^q \\ &\leq c_2 L^{\alpha/(q-1)} \left\{ \int_{\mathbb{X}} w^{-q/q'}(x) d\mu(x) \right\}^{q'/q} \left\{ \int_{\mathbb{X}} |P(x)|^p w(x) d\mu(x) \right\} \\ &= c_2 L^{\alpha/(q-1)} \|w^{-1}\|_{\mu;q-1} \|P\|_{v;p}^p. \end{aligned}$$

Thus, if $1 < q < \infty$ and $w^{-1} \in L^{q-1}(\mu)$, then

$$\|P\|_{\mu;p} \leq c_3 L^{\alpha/(pq-p)} \|w^{-1}\|_{\mu;q-1}^{1/p} \|P\|_{v;p}. \tag{4.6}$$

In particular, if $w \in L^r(\mu)$ and $w^{-1} \in L^r(\mu)$, then

$$\|P\|_{v;p} \leq c_1 L^{\alpha/(pr)} \|w\|_{\mu;r}^{1/p} \|P\|_{\mu;p}, \quad \|P\|_{\mu;p} \leq c_3 L^{\alpha/(pr)} \|w^{-1}\|_{\mu;r}^{1/p} \|P\|_{v;p}. \quad \square \tag{4.7}$$

The measure ν will be called an *MZ measure* if the constants c_1, c_2 appearing in (4.3) are independent of L .

5. Main theorems

Let $\mathcal{C} \subset K \subset \mathbb{X}$ be compact sets. We define the mesh norm $\delta(\mathcal{C}, K)$ of \mathcal{C} with respect to K and the minimal separation of \mathcal{C} by

$$\delta(\mathcal{C}, K) = \sup_{x \in K} \rho(x, \mathcal{C}), \quad q(\mathcal{C}) = \min_{x, y \in \mathcal{C}, x \neq y} \rho(x, y). \tag{5.1}$$

To keep the notation simple, we will write $\delta(\mathcal{C}) := \delta(\mathcal{C}, \mathbb{X})$. Of course, the quantity $q(\mathcal{C})$ is of interest only when \mathcal{C} is a finite set. It is easy to see that $q(\mathcal{C})/2 \leq \delta(\mathcal{C})$.

Our first theorem states the MZ inequalities in a sharp form in an apparently special case. (The proof of this and other statements in this section will be given in Section 7.) We note that part (a) of the following theorem was proved (with minor differences) in [11, Theorem 3.2] for the case $p = 1$.

Theorem 5.1. *Let $\mathcal{C} = \{x_1, \dots, x_M\}$ be a finite subset of \mathbb{X} satisfying*

$$\frac{1}{2}q(\mathcal{C}) \leq \delta(\mathcal{C}) \leq \kappa q(\mathcal{C}) \tag{5.2}$$

for some $\kappa \geq 1$. Let $1 \leq p < \infty$ and $A \geq 2$. In this theorem, all the constants may depend upon κ and A .

(a) *There exist $c_1, c_2 > 0$ such that for every $\eta > 0$, if $\delta(\mathcal{C}) \leq c_1, L \leq c_2 \eta (p\delta(\mathcal{C}))^{-1}$, and $P \in \Pi_L$, then*

$$\sum_{k=1}^M \mu(B(x_k, \delta(\mathcal{C}))) \sup_{z, y \in B(x_k, A\delta(\mathcal{C}))} \left| |P(z)|^p - |P(y)|^p \right| \leq \eta \int_{\mathbb{X}} |P(x)|^p d\mu(x). \tag{5.3}$$

(b) *Let \mathcal{C} be as in part (a), and $\{Y_k\}_{k=1}^M$ be a partition of \mathbb{X} such that $x_k \in Y_k \subseteq B(x_k, A\delta(\mathcal{C}))$ for each $k, 1 \leq k \leq M$. There exists $c_3 > 0$ such that for $L \leq c_3 \eta (p\delta(\mathcal{C}))^{-1}$, and $P \in \Pi_L$, we have*

$$\left| \int_{\mathbb{X}} |P(z)|^p d\mu(z) - \sum_{k=1}^M \mu(Y_k) |P(x_k)|^p \right| \leq \eta \int_{\mathbb{X}} |P(x)|^p d\mu(x). \tag{5.4}$$

(c) *There exists $c_4 > 0$ such that if $L \leq c_4 \eta \delta(\mathcal{C})^{-1}$ then*

$$\left| \|P\|_{\mu; \infty} - \max_{1 \leq k \leq M} |P(x_k)| \right| \leq \eta \|P\|_{\mu; \infty}. \tag{5.5}$$

We note that a variant of **Theorem 5.1** was stated in our paper [28, Theorem 3.1] in the case when \mathbb{X} is the Euclidean unit sphere, ϕ_i 's are spherical harmonics (so that Π_L is the class of all spherical polynomials of degree at most L), and μ is the Riemannian volume measure on the sphere. The theorem is correct for $p = 1, \infty$ as stated there, but the proof does not use the correct form of the Riesz–Thorin interpolation theorem which is needed for proving such inequalities. Also, in the proof of [11, Theorem 3.2], we had constructed a partition Y_k . However, it was an error on our part to assume that $x_k \in B(x_k, q(\mathcal{C})/2) \subset Y_k$. Both of these errors are corrected in **Theorem 5.1** and the proof of **Theorem 5.3**.

Next, we wish to give an analog of **Theorem 5.1** where general measures are involved. The transition from the finitely supported measures to the general case is achieved via the following theorem.

Theorem 5.2. *Let ν be a signed measure, $\delta(\text{supp}(\nu)) < d \leq 1/81$. Then there exists a finite subset $\mathcal{C} = \{x_1, \dots, x_M\} \subseteq \text{supp}(\nu)$ with the property that*

$$q(\mathcal{C})/2 \leq \delta(\mathcal{C}) \leq 81d \leq 162q(\mathcal{C}). \tag{5.6}$$

Moreover, there exists a partition $\{Y_k\}_{k=1}^M$ of \mathbb{X} and a finite subset $\tilde{\mathcal{C}}$ with $\mathcal{C} \subseteq \tilde{\mathcal{C}} \subseteq \text{supp}(\nu)$ such that for $k = 1, \dots, M, x_k \in Y_k \subseteq B(x_k, 81d), \mu(Y_k) \sim d^\alpha$, and $|\nu|(Y_k) \geq c \min_{x \in \tilde{\mathcal{C}}} |\nu|(B(x, d/4)) > 0$.

Theorem 5.2 helps us to use **Theorem 5.1** to arrive at the following statement, where general measures are involved.

Theorem 5.3. Let ν be a signed measure, $\delta(\text{supp}(\nu)) < d \leq 1/81$, $\{Y_k\}_{k=1}^M$ be as in Theorem 5.2.

(a) Let $1 \leq p < \infty$. There exist $c_1, c_2 > 0$ such that for every $\eta > 0$, if $d \leq c_1$, $L \leq c_2\eta(pd)^{-1}$, and $P \in \Pi_L$, then

$$\sum_{k=1}^M \left| \int_{Y_k} |P(y)|^p d\mu(y) - \frac{\mu(Y_k)}{|\nu|(Y_k)} \int_{Y_k} |P(z)|^p d|\nu|(z) \right| \leq \eta \|P\|_{\mu;p}^p. \tag{5.7}$$

(b) There exists $c_3 > 0$ such that if $L \leq c_3\eta d^{-1}$ and $P \in \Pi_L$, then

$$\left| \|P\|_{\nu;\infty} - \|P\|_{\mu;\infty} \right| \leq \eta \|P\|_{\mu;\infty}. \tag{5.8}$$

Since Theorem 5.3(b) settles the question of MZ inequalities in the case $p = \infty$, we will focus in the remainder of this paper on the case when $1 \leq p < \infty$. It is clear from Theorem 5.3(a) that the MZ inequalities for the measure ν will depend upon the relationship between $\nu(B(x, d))$ and $\mu(B(x, d)) \sim d^\alpha$ for $x \in \mathbb{X}$. Accordingly, we make the following definition.

Definition 5.4. Let $\nu \in \mathcal{M}$, $d > 0$.

(a) We say that ν is d -regular if

$$\nu(B(x, d)) \leq cd^\alpha, \quad x \in \mathbb{X}. \tag{5.9}$$

The infimum of all constants c which work in (5.9) will be denoted by $\|\nu\|_{R,d}$.

(b) We say that ν is d -dominant if

$$\nu(B(x, d)) \geq cd^\alpha, \quad x \in \mathbb{X}. \tag{5.10}$$

The supremum of all c which work in (5.10) will be denoted by $\|\nu\|_{D,d}^{-1}$.

We observe that (5.9) and (5.10) are very similar to (3.1) and (3.10) respectively. In particular, μ is both d -regular and d -dominant for every $d > 0$. In contrast, for a general measure $\nu \in \mathcal{M}$, the definition requires (5.9) and (5.10) to hold only for one value of d . Also, the function $\nu \rightarrow \|\nu\|_{R,d}$ is a norm on the space of all d -regular measures.

Example 2. Let \mathcal{C} be as in Theorem 5.1, $\delta(\mathcal{C}) \leq 2\kappa$, ν be the measure that associates the mass $\mu(B(x_k, \delta(\mathcal{C})))$ with each $x_k, k = 1, \dots, M$. Let $x \in \mathbb{X}, \tilde{\mathcal{C}} = B(x, 2\delta(\mathcal{C})) \cap \mathcal{C}$. Then (5.2) implies that the balls $B(y, \delta(\mathcal{C})/(2\kappa)), y \in \tilde{\mathcal{C}}$ are mutually disjoint, and clearly, their union is a subset of $B(x, 2\delta(\mathcal{C})(1 + 1/(2\kappa)))$. So, in view of (3.1) and (3.11), we obtain

$$\begin{aligned} \nu(B(x, 2\delta(\mathcal{C}))) &= \sum_{y \in \tilde{\mathcal{C}}} \mu(B(y, \delta(\mathcal{C}))) \leq c \sum_{y \in \tilde{\mathcal{C}}} \mu(B(y, \delta(\mathcal{C})/(2\kappa))) \\ &= c\mu\left(\bigcup_{y \in \tilde{\mathcal{C}}} B(y, \delta(\mathcal{C})/(2\kappa))\right) \\ &\leq c\mu(B(x, 2\delta(\mathcal{C})(1 + 1/(2\kappa)))) \leq c_1\delta(\mathcal{C})^\alpha. \end{aligned} \tag{5.11}$$

In the reverse direction, the definition of the mesh norm implies that $B(x, \delta(\mathcal{C})) \subset \bigcup_{y \in \tilde{\mathcal{C}}} B(y, \delta(\mathcal{C}))$. Therefore, we obtain using (3.10) that

$$\nu(B(x, 2\delta(\mathcal{C}))) = \sum_{y \in \tilde{\mathcal{C}}} \mu(B(y, \delta(\mathcal{C}))) \geq \mu(B(x, \delta(\mathcal{C}))) \geq c\delta(\mathcal{C})^\alpha. \tag{5.12}$$

Thus, ν is $2\delta(\mathcal{C})$ -regular as well as $2\delta(\mathcal{C})$ -dominant.

Example 3. Let \mathcal{C} be as in Theorem 5.1. For each $y \in \mathcal{C}$, let $q(\mathcal{C})/4 < r_y \leq q(\mathcal{C})/2$, χ_y be the characteristic (indicator) function of $B(y, r_y)$. Then using the same argument as above, it is easy to verify that the measure $d\nu = \left(\sum_{y \in \mathcal{C}} \chi_y\right) d\mu$ also satisfies (5.11) and (5.12) (with different constants). Thus, this ν is also $c_1\delta(\mathcal{C})$ -regular and $c_2\delta(\mathcal{C})$ -dominant.

Example 4. Let $w \in L^r(\mu)$ for some $1 < r < \infty$, $w \geq 0$ μ -almost everywhere on \mathbb{X} , and $d\nu = wd\mu$. Then $\text{supp}(\nu) = \mathbb{X}$. Let $x \in \mathbb{X}$, $d \in (0, 2/3)$, and in this example only, $M = \|w\|_{\mu,r}/d^{\alpha/r}$, $E = \{y \in B(x, d) : w(y) \geq M\}$. Then Hölder’s inequality implies that

$$M\mu(E) \leq \int_E wd\mu \leq \|w\|_{\mu,r}\mu(E)^{(r-1)/r},$$

and hence, $\mu(E) \leq (\|w\|_{\mu,r}/M)^r$. The second inequality above then yields

$$\int_E wd\mu \leq \|w\|_{\mu,r}^r M^{-(r-1)}.$$

Therefore, our choice of M implies that

$$\begin{aligned} \nu(B(x, d)) &= \int_{B(x,d)\setminus E} wd\mu + \int_E wd\mu \leq M\mu(B(x, d)) \\ &+ \|w\|_{\mu,r}^r M^{-(r-1)} \leq (\kappa_1 + 1)Md^\alpha, \end{aligned} \tag{5.13}$$

where κ_1 is defined in (3.1). To obtain an estimate analogous to (5.10), let $w^{-1} \in L^{q-1}(\mu)$ (i.e., $w^{-1} \in L^q(\nu)$) for some $q > 1$. The same argument as above shows that for any $M_1 > 0$, $\tilde{E} = \{y \in B(x, d) : w(y)^{-1} \geq M_1\}$, we have

$$\int_{\tilde{E}} w^{-1}d\nu = \mu(\tilde{E}) \leq \|w^{-1}\|_{\nu,q}^q M_1^{-(q-1)} = \left(\frac{\|w^{-1}\|_{\mu,q-1}}{M_1}\right)^{q-1}.$$

Hence, in view of (3.10),

$$cd^\alpha \leq \mu(B(x, d)) = \int_{B(x,d)\setminus\tilde{E}} w^{-1}d\nu + \int_{\tilde{E}} w^{-1}d\nu \leq M_1\nu(B(x, d)) + \left(\frac{\|w^{-1}\|_{\mu,q-1}}{M_1}\right)^{q-1}.$$

We now choose $M_1 = \|w^{-1}\|_{\mu,q-1}(cd^\alpha/2)^{-1/(q-1)}$, and conclude that

$$\nu(B(x, d)) \geq c_1 M_1^{-1} d^\alpha. \tag{5.14}$$

Theorem 5.5. Let $L \geq 2$, and $\nu \in \mathcal{M}$. In this theorem, all constants c, c_1, \dots , may depend upon p .

- (a) If ν is $1/L$ -regular, then $\|P\|_{\nu,p} \leq c_1 \|\nu\|_{R,1/L}^{1/p} \|P\|_{\mu,p}$ for all $P \in \Pi_L$ and $1 \leq p < \infty$. Conversely, if for some $A > 0$ and $1 \leq p < \infty$, $\|P\|_{\nu,p} \leq A^{1/p} \|P\|_{\mu,p}$ for all $P \in \Pi_L$, then ν is $1/L$ -regular, and $\|\nu\|_{R,1/L} \leq c_2 A$.
- (b) There exist constants c, c_4 such that if $L \geq c$ and ν is c_4/L -dominant, then $\|P\|_{\mu,p} \leq c_3 \|\nu\|_{D,c_4/L}^{1/p} \|P\|_{\nu,p}$ for all $P \in \Pi_L$, and for all $p, 1 \leq p < \infty$. Conversely, let ν be $1/L$ -regular, and $S > \alpha$ be an integer. If for some $A_1 > 0$, and $1 \leq p < \infty$, $\|P\|_{\mu,p} \leq A_1^{1/p} \|P\|_{\nu,p}$ for all $P \in \Pi_L$, then ν is $d = c_5(S)(\max(1, \|\nu\|_{R,1/L} A_1)^{1/(S-\alpha)})L^{-1}$ -dominant, and $\|\nu\|_{D,d} \leq c_6(S)A_1$.

The term d -regular has been used with different meanings in our other papers. The following proposition reconciles the different definitions.

Proposition 5.6. Let $d \in (0, 1]$, $\nu \in \mathcal{M}$.

- (a) If ν is d -regular, then for each $r > 0$ and $x \in \mathbb{X}$,

$$|\nu|(B(x, r)) \leq c \|\nu\|_{R,d} \mu(B(x, r+d)) \leq c_1 \|\nu\|_{R,d} (r+d)^\alpha. \tag{5.15}$$

Conversely, if for some $A > 0$, $|\nu|(B(x, r)) \leq A(r+d)^\alpha$ or each $r > 0$ and $x \in \mathbb{X}$, then ν is d -regular, and $\|\nu\|_{R,d} \leq 2^\alpha A$.

- (b) For each $\gamma > 1$,

$$\|\nu\|_{R,\gamma d} \leq c_1 (\gamma + 1)^\alpha \|\nu\|_{R,d} \leq c_1 (\gamma + 1)^\alpha \gamma^\alpha \|\nu\|_{R,\gamma d}, \tag{5.16}$$

where c_1 is the constant appearing in (5.15).

We end this section with a discussion about positive quadrature formulas. We will say that ν is a quadrature measure of order L if

$$\int_{\mathbb{X}} P(y) d\mu(y) = \int_{\mathbb{X}} P(y) d\nu(y), \quad P \in \Pi_L.$$

First, we prove a very general existence theorem for such formulas.

Theorem 5.7. *There exist constants $c_1, c_2 > 0$ with the following property: If ν is a signed measure, $\delta(\text{supp}(\nu)) < d < c_1$ and $0 < L < c_2 d^{-1}$, then there exists a simple function $W : \text{supp}(\nu) \rightarrow [0, \infty)$, satisfying*

$$\int_{\mathbb{X}} P(y) d\mu(y) = \int_{\mathbb{X}} P(y) W(y) d|\nu|(y), \quad P \in \Pi_L. \quad (5.17)$$

If ν is d -regular, then $W(y) \geq c \|\nu\|_{R,d}^{-1}, y \in \mathbb{X}$.

We observe that if ν is supported on a finite subset of \mathbb{X} , then this reduces to [11, Theorem 3.1(b)]. We find it remarkable that the only conditions on ν for (5.17) to hold are on $\text{supp}(\nu)$.

In many cases of interest, for example, the Euclidean sphere, the rotation group $SO(3)$ and projective spaces, if $P \in \Pi_L$, then $P^2 \in \Pi_{2L}$. In the Appendix, we will show that a similar fact holds for eigenfunctions of a fairly large class of elliptic operators. In the very general situation considered in this paper, we make the following product assumption as in [27]. To formulate this assumption, we need one further notation. If $f \in L^p(\mu)$, and $m > 0$, we denote

$$\text{dist}(p; f, \Pi_m) := \inf_{P \in \Pi_m} \|f - P\|_{\mu;p}.$$

Product assumption:

We assume that there exists a constant $A^* \geq 2$ with the following property: With

$$\epsilon_L := \sup_{\ell_j, \ell_k \leq L} \text{dist}(\infty; \phi_j \phi_k, \Pi_{A^*L}), \quad L > 0, \quad (5.18)$$

we have $L^c \epsilon_L \rightarrow 0$ as $L \rightarrow \infty$ for every $c > 0$.

We have conjectured in [27] that this assumption holds for every analytic manifold \mathbb{X} .

Theorem 5.8. *Let the product assumption hold. There exists a constant $c > 0$ such that if $L \geq c$ and τ is a positive quadrature measure of order $2A^*L$, then*

$$\|P\|_{\tau;p} \sim \|P\|_{\mu;p}, \quad P \in \Pi_L, \quad (5.19)$$

where the constants involved may depend upon p but not on τ or L .

6. Preparatory results

In this section, we summarize some results which will be needed in the proofs of the theorems in Section 5. In Section 6.1, we prove the Riesz–Thorin interpolation theorem in the form in which we need it. In Section 6.2, we summarize some of the properties of a localized kernel and diffusion polynomials [26,11], and extend these to the L^p setting using the Riesz–Thorin interpolation theorem. In Section 6.3, we prove, for the sake of completeness, a special case of Krein’s extension theorem for positive functionals, following a hint in [18, Exercise (14.27), p. 200]. This will be used in proving the existence of positive quadrature formulas in Theorem 5.7(b).

6.1. Riesz–Thorin interpolation theorem

Let \mathcal{X}, \mathcal{Y} be Banach spaces of functions defined on a measure space (Ω, τ) . We assume the existence of associated Banach spaces $\mathcal{X}', \mathcal{Y}'$ such that

$$\begin{aligned} \|f\|_{\mathcal{X}} &= \sup \left\{ \int_{\Omega} f(x)\overline{g(x)}d\tau(x) : \|g\|_{\mathcal{X}'} = 1 \right\}, \\ \|f\|_{\mathcal{Y}} &= \sup \left\{ \int_{\Omega} f(x)\overline{g(x)}d\tau(x) : \|g\|_{\mathcal{Y}'} = 1 \right\}. \end{aligned} \tag{6.1}$$

Let $\mathbf{W} = \{w_k\}_{k=1}^M \subset (0, \infty)$. For $1 \leq p \leq \infty$, and integer $M \geq 1$, we define for $\mathbf{a} = (a_1, \dots, a_M) \in \mathbb{C}^M$,

$$\|\mathbf{a}\|_{\mathbf{w}, \ell^p} = \begin{cases} \left(\sum_{k=1}^M w_k |a_k|^p \right)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \max_{1 \leq k \leq M} |a_k|, & \text{if } p = \infty. \end{cases}$$

It is elementary to check that

$$\|\mathbf{a}\|_{\mathbf{w}, \ell^p} = \sup \left\{ \sum_{k=1}^M w_k a_k \overline{b_k} : \|(b_1, \dots, b_M)\|_{\mathbf{w}, \ell^{p'}} = 1 \right\}. \tag{6.2}$$

We define the space $\mathcal{X}_{\mathbf{w}, p}$ to be the tensor product space $\otimes_{k=1}^M \mathcal{X}$ equipped with the norm

$$\|\mathbf{f}\|_{\mathcal{X}, \mathbf{w}, p} := \|(\|f_1\|_{\mathcal{X}}, \dots, \|f_M\|_{\mathcal{X}})\|_{\mathbf{w}, \ell^p}, \quad \mathbf{f} = (f_1, \dots, f_M) \in \mathcal{X}_{\mathbf{w}, p}.$$

The space $\mathcal{X}'_{\mathbf{w}, p}$ and the norm $\|\circ\|_{\mathcal{X}', \mathbf{w}, p}$ are defined similarly.

In the statement of the Riesz–Thorin interpolation theorem, we need another measure space. Not to complicate our notations, we will use (\mathbb{X}, μ) here. However, it should be understood that the only property we need is that this is a measure space, with μ being a positive measure. In this subsection we are not assuming any properties and other assumptions, including the fact that \mathbb{X} is a manifold, and μ is a probability measure.

Theorem 6.1. *Let $1 \leq p_0 \leq p_1 \leq \infty, 1 \leq r_0 \leq r_1 \leq \infty, 0 < t < 1, \mathcal{U}$ be a linear operator satisfying*

$$\|\mathcal{U}f\|_{\mathcal{X}, \mathbf{w}, p_j} \leq M_j \|f\|_{\mu; r_j}, \quad f \in L^{r_j}(\mu), \quad j = 0, 1, \tag{6.3}$$

and $1/p = (1 - t)/p_0 + t/p_1, 1/r = (1 - t)/r_0 + t/r_1$. Then

$$\|\mathcal{U}f\|_{\mathcal{X}, \mathbf{w}, p} \leq M_0^{1-t} M_1^t \|f\|_{\mu; r}, \quad f \in L^r(\mu). \tag{6.4}$$

The proof of **Theorem 6.1** mimics that of the usual Riesz–Thorin theorem. We could not find a reference where this theorem is stated in the form in which we need it. Therefore, we include a proof, following that of the usual Riesz–Thorin theorem as given in [31, Chapter XII, Theorem 1.11]. The first step is the following lemma.

Lemma 6.2.

$$\|\mathbf{f}\|_{\mathcal{X}, \mathbf{w}, p} = \sup \left| \sum_{k=1}^M w_k \overline{b_k} \int_{\Omega} f_k(x)\overline{g_k(x)}d\tau(x) \right|, \tag{6.5}$$

where the supremum is over all $\mathbf{b} = (b_1, \dots, b_M)$ with $\|\mathbf{b}\|_{\mathbf{w}, \ell^{p'}} = 1$ and $g_1, \dots, g_M \in \mathcal{X}'$ with $\|g_k\|_{\mathcal{X}'} = 1, k = 1, \dots, M$.

Proof. In view of Hölder’s inequality, it is clear that

$$\sup \sum_{k=1}^M w_k \overline{b_k} \int_{\Omega} f_k(x)\overline{g_k(x)}d\tau(x) \leq \|\mathbf{f}\|_{\mathcal{X}, \mathbf{w}, p}, \tag{6.6}$$

where the supremum is over all \mathbf{b} with $\|\mathbf{b}\|_{\mathbf{w}, \ell^{p'}} = 1$ and $g_1, \dots, g_M \in \mathcal{X}'$ with $\|g_k\|_{\mathcal{X}'} = 1, k = 1, \dots, M$. If $\mathbf{f} = 0$ then (6.5) is obvious. Let $\mathbf{f} \neq 0$, and $\epsilon > 0$. In view of (6.1) and (6.2), there exist $g_k \in \mathcal{X}'$ and $\mathbf{b} \in [0, \infty)^M$ such that $\|g_k\|_{\mathcal{X}'} = 1, k = 1, \dots, M, \|\mathbf{b}\|_{\mathbf{w}, \ell^{p'}} = 1$ and

$$\sum_{k=1}^M w_k \bar{b}_k \int_{\Omega} f_k(x) \overline{g_k(x)} d\tau(x) \geq (1 - \epsilon) \sum_{k=1}^M w_k \bar{b}_k \|f_k\|_{\mathcal{X}} \geq (1 - \epsilon)^2 \|\mathbf{f}\|_{\mathcal{X}, \mathbf{w}, p}. \quad \square$$

Next, we recall the Phragmén–Lindelöf maximum principle [31, Chapter XII, Theorem 1.3].

Proposition 6.3. *Suppose that f is continuous and bounded on the closed strip of the complex plane $0 \leq \Re z \leq 1$, and analytic in the interior of this strip. If $|f(z)| \leq M_0, \Re z = 0$, and $|f(z)| \leq M_1, \Re z = 1$, then $|f(z)| \leq M_0^{1-t} M_1^t$ for $\Re z = t$.*

We are now ready to prove Theorem 6.1.

Proof of Theorem 6.1. In this proof only, we will write $\alpha_j = 1/r_j, \beta_j = 1/p_j, j = 0, 1, \alpha(z) = (1 - z)\alpha_0 + z\alpha_1, \beta(z) = (1 - z)\beta_0 + z\beta_1$, so that $\alpha(t) = 1/r, \beta(t) = 1/p$. If $r_0 = r_1 = \infty$, then $r = \infty$ as well, and (6.4) is a simple consequence of Hölder inequality. So, we assume that $r_0 < \infty$, and hence, $r < \infty$. Since simple functions are dense in $L^r(\mu)$, it is enough to prove (6.4) when f is a simple function; i.e., $f = \sum_{j=1}^N d_j e^{iu_j} \chi_j$, where $N \geq 1$ is an integer, $d_j > 0, u_j \in (-\pi, \pi]$, and χ_j 's are the characteristic functions of pairwise disjoint sets. We define $f_z = \sum_{j=1}^N d_j^{\alpha(z)r} e^{iu_j} \chi_j$. Next, let $\mathbf{b} = (|b_1|e^{iv_1}, \dots, |b_M|e^{iv_M}) \in \mathbb{C}^M$ be an arbitrary vector satisfying $\|\mathbf{b}\|_{\mathbf{w}, \ell^{p'}} = 1$, and $g_1, \dots, g_M \in \mathcal{X}'$ be arbitrary functions satisfying $\|g_k\|_{\mathcal{X}'} = 1, k = 1, \dots, M$. We define

$$G_{z,k} = |b_k|^{(1-\beta(z))p'} e^{-iv_k}, \quad k = 1, \dots, M,$$

where it is understood that $G_{z,k} = 0$ if $b_k = 0$. Finally, we define

$$\begin{aligned} \Phi(z) &= \sum_{k=1}^M w_k G_{z,k} \int_{\Omega} (\mathcal{U}f_z)_k(x) \overline{g_k(x)} d\tau(x) \\ &= \sum_{k=1}^M \sum_{j=1}^N w_k G_{z,k} d_j^{\alpha(z)p} e^{iu_j} \int_{\Omega} (\mathcal{U}\chi_j)(x) \overline{g_k(x)} d\tau(x). \end{aligned}$$

We note that $f_t = f, G_{t,k} = \bar{b}_k$, and therefore,

$$\Phi(t) = \sum_{k=1}^M w_k \bar{b}_k \int_{\Omega} (\mathcal{U}f)_k(x) \overline{g_k(x)} d\tau(x). \tag{6.7}$$

Now, Φ is a finite linear combination of functions of the form e^{az} , and hence, is an entire function, bounded on the strip $0 \leq \Re z \leq 1$. If $\Re z = 0$ then $|G_{z,k}| = |b_k|^{p' \Re z (1-\beta(z))} = |b_k|^{p' (1-\beta_0)} = |b_k|^{p'/p'_0}$. Therefore, using Hölder's inequality, we obtain for $\Re z = 0$:

$$\begin{aligned} |\Phi(z)| &\leq \|(G_{z,1}, \dots, G_{z,M})\|_{\mathbf{w}, \ell^{p'_0}} \\ &\quad \times \left\| \left(\int_{\Omega} (\mathcal{U}f_z)_1(x) \overline{g_1(x)} d\tau(x), \dots, \int_{\Omega} (\mathcal{U}f_z)_M(x) \overline{g_M(x)} d\tau(x) \right) \right\|_{\mathbf{w}, \ell^{p_0}} \\ &\leq \|\mathbf{b}\|_{\mathbf{w}, \ell^{p'}}^{p'/p'_0} (\|(\mathcal{U}f_z)_1\|_{\mathcal{X}}, \dots, \|(\mathcal{U}f_z)_M\|_{\mathcal{X}})_{\mathbf{w}, \ell^{p_0}} = \|\mathcal{U}f_z\|_{\mathcal{X}, \mathbf{w}, p_0} \leq M_0 \|f_z\|_{\mu; r_0}. \end{aligned} \tag{6.8}$$

For $\Re z = 0, |d_j^{\alpha(z)r}| = d_j^{r/r_0}$. Also, at any point $x \in \mathbb{X}$, there is at most one j such that $\chi_j(x) \neq 0$. For this $j, f_z(x) = d_j^{\alpha(z)r} e^{iu_j} \chi_j(x)$, and $f(x) = d_j e^{iu_j} \chi_j(x)$. So, for $\Re z = 0$, and any $x \in \mathbb{X}, |f_z(x)|^{r_0} = \sum_{j=1}^N d_j^r \chi_j(x) = |f(x)|^r$. Thus, $\|f_z\|_{\mu; r_0} = \|f\|_{\mu; r}^{r/r_0}$. Hence, (6.8) shows that

$|\Phi(z)| \leq M_0 \|f\|_{\mu;r}^{r/r_0}$, $\Re z = 0$. Similarly, $|\Phi(z)| \leq M_1 \|f\|_{\mu;r}^{r/r_1}$, $\Re z = 1$. Proposition 6.3 then implies that $|\Phi(t)| \leq M_0^{1-t} M_1^t \|f\|_{\mu;r}$; i.e., in view of (6.7), we have

$$\left| \sum_{k=1}^M w_k \bar{b}_k \int_{\Omega} (\mathcal{U}f)_k(x) \overline{g_k(x)} d\tau(x) \right| \leq M_0^{1-t} M_1^t \|f\|_{\mu;r}.$$

Since \mathbf{b} and the functions g_k were arbitrary subject only to $\|\mathbf{b}\|_{\mathbf{w},p'} = 1$, $\|g_k\|_{\mathcal{X}'} = 1$, the estimate (6.4) follows from Lemma 6.2. \square

6.2. Localized polynomial operators

Let $h : \mathbb{R} \rightarrow [0, 1]$ be an even, C^∞ function, nonincreasing on $[0, \infty)$ such that $h(t) = 1$ if $|t| \leq 1/2$ and $h(t) = 0$ if $|t| \geq 1$. We will treat h to be a fixed function, so that the dependence of different constants on the choice of h will not be mentioned. We will write

$$\Phi_L(x, y) := \sum_{j=0}^{\infty} h(\ell_j/L) \phi_j(x) \phi_j(y). \tag{6.9}$$

For $f \in L^1(\mu)$, we define

$$\hat{f}(j) = \int_{\mathbb{X}} f(y) \phi_j(y) d\mu(y), \quad j = 0, 1, \dots,$$

and

$$\sigma_L(f, x) := \int_{\mathbb{X}} f(y) \Phi_L(x, y) d\mu(y) = \sum_{\ell_j \leq L} h(\ell_j/L) \hat{f}(j) \phi_j(x). \tag{6.10}$$

We have proved in [26, Theorem 4.1], [11, Theorem 2.1] the following:

Theorem 6.4. For every $L > 0$ and integer $S > \alpha$, we have

$$|\Phi_L(x, y)| \leq c \frac{L^\alpha}{\max(1, (L\rho(x, y))^S)}, \quad x, y \in \mathbb{X}, \tag{6.11}$$

and

$$\sup_{x \in \mathbb{X}} \int_{\mathbb{X}} |\Phi_L(x, y)| d\mu(y) \leq c. \tag{6.12}$$

Consequently, for $1 \leq p \leq \infty$,

$$\|\sigma_L(f)\|_{\mu;p} \leq c \|f\|_{\mu;p}, \quad f \in L^p(\mu). \tag{6.13}$$

We will also need the following two propositions. Proposition 6.5 is proved in [27, Proposition 5.1]. The definition of regular measures in [27] is different from the one in this paper, but Proposition 5.6 shows that they are equivalent.

Proposition 6.5. Let $d > 0$, $S > \alpha$ be an integer, and (3.1), (3.3) hold. Let ν satisfy $\|\nu\|_{\mathbb{R},d} < \infty$, $L > 0$, and κ_1 be as in (3.1). Let $1 \leq p \leq \infty$. In this proposition, all constants will depend upon S .

(a) If $g_1 : [0, \infty) \rightarrow [0, \infty)$ is a nonincreasing function, then for any $L > 0$, $r > 0$, $x \in \mathbb{X}$,

$$L^\alpha \int_{\Delta(x,r)} g_1(L\rho(x, y)) d|\nu|(y) \leq \frac{2^\alpha(\kappa_1 + (d/r)^\alpha)\alpha}{1 - 2^{-\alpha}} \|\nu\|_{\mathbb{R},d} \int_{rL/2}^\infty g_1(u) u^{\alpha-1} du. \tag{6.14}$$

(b) If $r \geq 1/L$, then

$$\int_{\Delta(x,r)} |\Phi_L(x, y)|d|v|(y) \leq c_1(1 + (dL)^\alpha)(rL)^{-S+\alpha} \|v\|_{R,d}. \tag{6.15}$$

(c) We have

$$\int_{\mathbb{X}} |\Phi_L(x, y)|d|v|(y) \leq c_2(1 + (dL)^\alpha) \|v\|_{R,d}, \tag{6.16}$$

$$\|\Phi_L(x, \circ)\|_{v;\mathbb{X},p} \leq c_3L^{\alpha/p'}(1 + (dL)^\alpha)^{1/p} \|v\|_{R,d}. \tag{6.17}$$

In the sequel, we will assume $S > \alpha$ to be a fixed, large integer, and will not indicate the dependence of the constants on S .

Next, we recall the following Proposition 6.6, proved essentially in [11, Eqn. (4.40), Lemma 4.5].

Proposition 6.6. Let $L \geq 1, \mathcal{C} = \{x_1, \dots, x_M\} \subset \mathbb{X}, \kappa > 1, A \geq 2, \delta(\mathcal{C}) \leq \kappa q(\mathcal{C})$. Let $X_k = B(x_k, \delta(\mathcal{C}))$, $\tilde{X}_k = B(x_k, A\delta(\mathcal{C}))$. In the following, all constants will depend upon A and κ . Then for every $P \in \Pi_L$, we have

$$\sum_{k=1}^M \mu(X_k) \|P\|_{\mu;\infty,\tilde{X}_k} \leq c\{(\delta(\mathcal{C})L)^\alpha + \min(1, (\delta(\mathcal{C})L)^{\alpha-S})\} \|P\|_{\mu,1}, \tag{6.18}$$

$$\sum_{k=1}^M \mu(X_k) \|\|\nabla P\|\|_{\mu;\infty,\tilde{X}_k} \leq cL\{(\delta(\mathcal{C})L)^\alpha + \min(1, (\delta(\mathcal{C})L)^{\alpha-S})\} \|P\|_{\mu,1}, \tag{6.19}$$

and

$$\|\|\nabla P\|\|_{\mu;\infty} \leq cL \|P\|_{\mu;\infty}. \tag{6.20}$$

In our proof of this proposition in [11], we used $A = 2$, but the same proof works in the more general case, verbatim, except for the following changes: The set \mathcal{I} defined after [11, eqn. (4.34)] should be redefined by $\mathcal{I} = \{j : \rho(x, \tilde{X}_j) \geq (2A + 1)\delta\}$, and the two displayed equations after [11, eqn. (4.34)] are changed to

$$|\rho(x, y) - \rho(x, \tilde{X}_j)| = |\rho(x, y) - \rho(x, z_j)| \leq \rho(y, z_j) \leq 2A\delta \leq (2A/(2A + 1))\rho(x, \tilde{X}_j),$$

and

$$\delta \leq (2A + 1)^{-1}\rho(x, \tilde{X}_j) \leq \rho(x, y) \leq \frac{4A + 1}{2A + 1}\rho(x, \tilde{X}_j)$$

respectively. We prefer not to reproduce the entire proof to accommodate these minor changes. We need to prove an L^p analog of the above proposition.

Lemma 6.7. Let $L \geq 1, \mathcal{C} = \{x_1, \dots, x_M\} \subset \mathbb{X}, A \geq 2, \kappa > 1, \delta(\mathcal{C}) \leq \kappa q(\mathcal{C})$. Let $X_k = B(x_k, \delta(\mathcal{C}))$, $\tilde{X}_k = B(x_k, A\delta(\mathcal{C}))$. Then for every $P \in \Pi_L$, we have

$$\left\{ \sum_{k=1}^M \mu(X_k) \|P\|_{\mu;\infty,\tilde{X}_k}^p \right\}^{1/p} \leq c\{(\delta(\mathcal{C})L)^\alpha + \min(1, (\delta(\mathcal{C})L)^{\alpha-S})\}^{1/p} \|P\|_{\mu;p}, \tag{6.21}$$

and

$$\left\{ \sum_{k=1}^M \mu(X_k) \|\|\nabla P\|\|_{\mu;\infty,\tilde{X}_k}^p \right\}^{1/p} \leq cL\{(\delta(\mathcal{C})L)^\alpha + \min(1, (\delta(\mathcal{C})L)^{\alpha-S})\}^{1/p} \|P\|_{\mu;p}. \tag{6.22}$$

Proof. We will prove (6.22). The proof of (6.21) is similar, but simpler, and is omitted. We observe first that for any differentiable $f : \mathbb{X} \rightarrow \mathbb{R}$, and $x \in \mathbb{X}$,

$$\|\nabla f\|_x = \sup \langle \nabla f(x), F(x) \rangle_x,$$

where the supremum is over all vector fields F with $\|F\|_x = 1$.

Let F be an arbitrarily fixed vector field with $\|F\|_x = 1$ for all $x \in \mathbb{X}$. We use Theorem 6.1 with $L^\infty(\mu)$ in place of \mathcal{X} , $p_0 = q_0 = 1$, $p_1 = q_1 = \infty$, $w_k = \mu(X_k)$, and

$$\mathcal{U}f(x) = (\chi_1(x)F(x)(\sigma_{2L}(f)), \dots, \chi_M(x)F(x)(\sigma_{2L}(f))),$$

where each χ_k is the characteristic function of \tilde{X}_k . Then, for $1 \leq p < \infty$,

$$\|\mathcal{U}f\|_{L^\infty(\mu), \mathbf{w}, p} = \left\{ \sum_{k=1}^M \mu(X_k) \|F(\sigma_{2L}(f))\|_{\mu; \infty, \tilde{X}_k}^p \right\}^{1/p},$$

and the formula holds also for $p = \infty$ with an obvious modification. Using (6.19), (6.20), (6.13) with $2L$ in place of L , $\sigma_{2L}(f)$ in place of P , we then see that for $p = 1, \infty, f \in L^p(\mu)$,

$$\begin{aligned} \|\mathcal{U}f\|_{L^\infty(\mu), \mathbf{w}, p} &\leq cL\{(\delta(\mathcal{C})L)^\alpha + \min(1, (\delta(\mathcal{C})L)^{\alpha-5})\}^{1/p} \|\sigma_{2L}(f)\|_{\mu; p} \\ &\leq c_1L\{(\delta(\mathcal{C})L)^\alpha + \min(1, (\delta(\mathcal{C})L)^{\alpha-5})\}^{1/p} \|f\|_{\mu; p}. \end{aligned}$$

Theorem 6.1 now implies

$$\left\{ \sum_{k=1}^M \mu(X_k) \|F(\sigma_{2L}(f))\|_{\mu; \infty, \tilde{X}_k}^p \right\}^{1/p} \leq cL\{(\delta(\mathcal{C})L)^\alpha + \min(1, (\delta(\mathcal{C})L)^{\alpha-5})\}^{1/p} \|f\|_{\mu; p}.$$

Since F is an arbitrary unit vector field, this leads to

$$\left\{ \sum_{k=1}^M \mu(X_k) \|\nabla(\sigma_{2L}(f))\|_{\mu; \infty, \tilde{X}_k}^p \right\}^{1/p} \leq cL\{(\delta(\mathcal{C})L)^\alpha + \min(1, (\delta(\mathcal{C})L)^{\alpha-5})\}^{1/p} \|f\|_{\mu; p}.$$

Since $\sigma_{2L}(P) = P$ for $P \in \Pi_L$, this implies (6.22). \square

6.3. Krein's extension theorem

The purpose of this section is to prove the following special case of the Krein extension theorem. Let \mathcal{X} be a normed linear space, \mathcal{K} be a subset of its normed dual \mathcal{X}^* , and \mathcal{V} be a linear subspace of \mathcal{X} . We say that a linear functional $x^* \in \mathcal{V}^*$ is positive on \mathcal{V} with respect to \mathcal{K} if $x^*(f) \geq 0$ for every $f \in \mathcal{V}$ with the property that $y^*(f) \geq 0$ for every $y^* \in \mathcal{K}$.

Theorem 6.8. *Let \mathcal{X} be a normed linear space, \mathcal{K} be a bounded subset of its normed dual \mathcal{X}^* , \mathcal{V} be a linear subspace of \mathcal{X} , $x^* \in \mathcal{V}^*$ be positive on \mathcal{V} with respect to \mathcal{K} . We assume further that there exists $v_0 \in \mathcal{V}$ such that $\|v_0\|_{\mathcal{X}} = 1$ and*

$$\inf_{y^* \in \mathcal{K}} y^*(v_0) = \beta^{-1} > 0. \tag{6.23}$$

Then there exists an extension $X^* \in \mathcal{X}^*$ of x^* which is positive on \mathcal{X} with respect to \mathcal{K} and satisfies

$$\|X^*\|_{\mathcal{X}^*} \leq \beta \sup_{y^* \in \mathcal{K}} \|y^*\|_{\mathcal{X}^*} x^*(v_0). \tag{6.24}$$

Proof. In this proof only, let $M = \sup_{y^* \in \mathcal{K}} \|y^*\|_{\mathcal{X}^*}$, and for $f_1, f_2 \in \mathcal{X}$, we will say that $f_1 \succeq f_2$ if $y^*(f_1) \geq y^*(f_2)$ for every $y^* \in \mathcal{K}$. In this proof only, let

$$p(f) = \inf\{x^*(P) : P \in \mathcal{V}, P \succeq f\}.$$

For any $f \in \mathcal{X}$ and $y^* \in \mathcal{K}$, we have $|y^*(f)| \leq M\|f\|_{\mathcal{X}} \leq \beta M\|f\|_{\mathcal{X}}y^*(v_0) = y^*(\beta M\|f\|_{\mathcal{X}}v_0)$. Since $\pm\beta M\|f\|_{\mathcal{X}}v_0 \in \mathcal{V}$, it follows that $p(f)$ is a finite number for $f \in \mathcal{X}$. It is not difficult to check that p is a sublinear functional; i.e.,

$$p(f_1 + f_2) \leq p(f_1) + p(f_2), \quad p(\gamma f_1) = \gamma p(f_1), \quad f_1, f_2 \in \mathcal{X}, \gamma \geq 0.$$

If $P, Q \in \mathcal{V}$, and $P \geq Q$ then the fact that x^* is positive on \mathcal{V} with respect to \mathcal{K} implies that $x^*(P) \geq x^*(Q)$. So, $p(Q) = x^*(Q)$ for all $Q \in \mathcal{V}$. The Hahn–Banach theorem [18, Theorem (14.9), p. 212] then implies that there exists an extension of x^* to a linear functional X^* on \mathcal{X} such that $X^*(f) \leq p(f)$, $f \in \mathcal{X}$. Then

$$X^*(f) = -X^*(-f) \geq -p(-f) = \sup\{x^*(-P) : P \in \mathcal{V}, P \geq -f\} = \sup\{x^*(Q) : Q \in \mathcal{V}, f \geq Q\}.$$

This implies two things. First, let $f \geq 0$. Choosing Q in the last supremum expression to be 0, we see that $X^*(f) \geq 0$. Second, as we observed earlier, $\beta M\|f\|_{\mathcal{X}}v_0 \geq f \geq -\beta M\|f\|_{\mathcal{X}}v_0$. Since $\pm\beta M\|f\|_{\mathcal{X}}v_0 \in \mathcal{V}$, we obtain that $|X^*(f)| \leq \beta Mx^*(v_0)\|f\|_{\mathcal{X}}$. This proves (6.24), and in particular, that $X^* \in \mathcal{X}^*$. \square

7. Proofs of the results in Section 5

We start with the proof of Theorem 5.1.

Proof of Theorem 5.1. We assume that $1 < p < \infty$, $\mathcal{C} = \{x_1, \dots, x_M\}$; the case $p = 1$ is simpler, and is essentially done in [11, Theorem 3.2]. We use the notation $X_k = B(x_k, \delta(\mathcal{C}))$, $\tilde{X}_k = B(x_k, A\delta(\mathcal{C}))$. Using the fact that $\nabla|P|^p = p|P|^{p-1}\text{sgn}(P)\nabla P$, we deduce that for any $k = 1, \dots, M, z, y \in \tilde{X}_k$,

$$||P(z)|^p - |P(y)|^p| \leq 2Ap\delta(\mathcal{C})\|P\|_{\mu; \infty, \tilde{X}_k}^{p-1} \|\nabla P\|_{\mu; \infty, \tilde{X}_k}.$$

We may assume that $L\delta(\mathcal{C}) \leq 1$. Hence, using Hölder’s inequality, (6.21) and (6.22), we obtain that

$$\begin{aligned} \sum_{k=1}^M \mu(X_k) \sup_{z, y \in \tilde{X}_k} ||P(z)|^p - |P(y)|^p| &\leq 2Ap\delta(\mathcal{C}) \sum_{k=1}^M \mu(X_k) \|P\|_{\mu; \infty, \tilde{X}_k}^{p-1} \|\nabla P\|_{\mu; \infty, \tilde{X}_k} \\ &\leq 2Ap\delta(\mathcal{C}) \left\{ \sum_{k=1}^M \mu(X_k) \|P\|_{\mu; \infty, \tilde{X}_k}^p \right\}^{1/p'} \left\{ \sum_{k=1}^M \mu(X_k) \|\nabla P\|_{\mu; \infty, \tilde{X}_k}^p \right\}^{1/p} \\ &\leq cAp\delta(\mathcal{C})L\|P\|_{\mu; p}^{p/p'} \|P\|_{\mu; p} = cAp\delta(\mathcal{C})L\|P\|_{\mu; p}^p. \end{aligned}$$

With $c_2 = 1/c$, this proves (5.3) if $LAp\delta(\mathcal{C}) \leq c_2\eta$.

To prove part (b), we observe that

$$\begin{aligned} \left| \int_{\mathcal{X}} |P(z)|^p d\mu(z) - \sum_{k=1}^M \mu(Y_k) |P(x_k)|^p \right| &= \left| \sum_{k=1}^M \int_{Y_k} \{|P(z)|^p - |P(x_k)|^p\} d\mu(z) \right| \\ &\leq \sum_{k=1}^M \int_{Y_k} ||P(z)|^p - |P(x_k)|^p| d\mu(z) \\ &\leq \sum_{k=1}^M \mu(Y_k) \sup_{z, y \in Y_k} ||P(z)|^p - |P(y)|^p| \\ &\leq c \sum_{k=1}^M \mu(X_k) \sup_{z, y \in \tilde{X}_k} ||P(z)|^p - |P(y)|^p|. \end{aligned}$$

Hence, (5.4) follows from (5.3).

The proof of part (c) is easier. Let $|P(z^*)| = \|P\|_{\mu; \infty}$. By definition of $\delta(\mathcal{C})$, there exists $x^* \in \mathcal{C}$ such that $\rho(z^*, x^*) \leq \delta(\mathcal{C})$. Then in view of (6.20), we have

$$|P(z^*)| - |P(x^*)| \leq \delta(\mathcal{C}) \|\nabla P\|_{\mu; \infty} \leq cL\delta(\mathcal{C})\|P\|_{\mu; \infty};$$

i.e.,

$$\max_{x \in \mathcal{C}} |P(x)| \leq \|P\|_{\mu; \infty} \leq |P(x^*)| + cL\delta(\mathcal{C})\|P\|_{\mu; \infty} \leq \max_{x \in \mathcal{C}} |P(x)| + cL\delta(\mathcal{C})\|P\|_{\mu; \infty}.$$

This leads to (5.5). \square

In the proofs of the other results in Section 5, we will often need the following observation. If $K \subseteq \mathbb{X}$ is a compact subset and $\epsilon > 0$, we will say that a subset $\mathcal{C} \subseteq K$ is ϵ -separated if $\rho(x, y) \geq \epsilon$ for every $x, y \in \mathcal{C}, x \neq y$. Since K is compact, there exists a finite, maximal ϵ -separated subset $\{x_1, \dots, x_M\}$ of K . If $x \in K \setminus \cup_{k=1}^M B(x_k, \epsilon)$, then $\{x, x_1, \dots, x_M\}$ is a strictly larger ϵ -separated subset of K . So, $K \subseteq \cup_{k=1}^M B(x_k, \epsilon)$. Moreover, the balls $B(x_k, \epsilon/2)$ are mutually disjoint.

The proof of Theorem 5.2 requires some further preparation. First, we recall a lemma [11, Lemma 4.4]. For a set Y , we denote the cardinality of Y by $|Y|$.

Lemma 7.1. *Let \mathcal{C} be a finite set for which (5.2) holds, $\delta(\mathcal{C}) \leq 2\kappa, A > 0$, and $x \in \mathbb{X}$. Then*

$$|\{y \in \mathcal{C} : x \in B(y, A\delta(\mathcal{C}))\}| \leq c_1(1 + A)^{2\alpha},$$

where $c_1 > 0$ is independent of A and $\delta(\mathcal{C})$. In particular,

$$|\{y \in \mathcal{C} : B(x, A\delta(\mathcal{C})) \cap B(y, A\delta(\mathcal{C})) \neq \emptyset\}| \leq c_1(1 + 2A)^{2\alpha}.$$

Proof. In this proof, let $\delta := \delta(\mathcal{C})$. Let $y_1, \dots, y_m \in \mathcal{C}$ and $x \in \cap_{k=1}^m B(y_k, A\delta)$. Then $B(x, \delta) \subseteq \cap_{k=1}^m B(y_k, (1 + A)\delta)$. Since $q(\mathcal{C}) \geq \delta/\kappa$, the balls $B(y_k, \delta/(2\kappa))$ are pairwise disjoint, and their union is a subset of $B(x, (1 + A)\delta)$. Therefore, (3.11) implies that

$$\begin{aligned} \mu(B(x, \delta)) &\leq \min_{1 \leq k \leq m} \mu(B(y_k, (1 + A)\delta)) \leq \frac{1}{m} \sum_{k=1}^m \mu(B(y_k, (1 + A)\delta)) \\ &\leq \frac{c(1 + A)^\alpha}{m} \sum_{k=1}^m \mu(B(y_k, \delta/(2\kappa))) \\ &= \frac{c(1 + A)^\alpha}{m} \mu(\cup_{k=1}^m B(y_k, \delta/(2\kappa))) \leq \frac{c(1 + A)^\alpha}{m} \mu(B(x, (1 + A)\delta)) \\ &\leq \frac{c_1(1 + A)^{2\alpha}}{m} \mu(B(x, \delta)). \end{aligned}$$

Thus, $m \leq c_1(1 + A)^{2\alpha}$. \square

The following lemma is needed in the construction of the partition in Theorem 5.3. The proof is based on some ideas in the book [6, Appendix 1] of David.

Lemma 7.2. *Let τ be a positive measure on \mathbb{X} , \mathcal{A} be a finite subset of \mathbb{X} satisfying*

$$q(\mathcal{A})/2 \leq \delta(\mathcal{A}) \leq \kappa q(\mathcal{A})$$

for some $\kappa > 0$, $\{Z_y\}_{y \in \mathcal{A}}$ be a partition of \mathbb{X} such that each $Z_y \subseteq B(y, \gamma\delta(\mathcal{A}))$ for some $\gamma \geq 1$. (We do not require that each Z_y be nonempty.) Then there exists a subset $\mathcal{G} \subseteq \mathcal{A}$ and a partition $\{Y_y\}_{y \in \mathcal{G}}$ such that for each $y \in \mathcal{G}, Z_y \subseteq Y_y, \tau(Y_y) \geq c \min_{z \in \mathcal{A}} \tau(B(z, \gamma\delta(\mathcal{A})))$, and $Y_y \subseteq B(y, 3\gamma\delta(\mathcal{A}))$. In particular, $\delta(\mathcal{G}) \leq 3\gamma\delta(\mathcal{A})$ and $q(\mathcal{G}) \geq q(\mathcal{A})$.

Proof. In this proof, let $\delta = \delta(\mathcal{A})$, $m = \min_{z \in \mathcal{A}} \tau(B(z, \gamma\delta))$. In view of Lemma 7.1, at most c^{-1} of the balls $B(y, \gamma\delta)$ can intersect each other, the constant c depending upon γ and κ . Let $\mathcal{G} = \{y \in \mathcal{A} : \tau(Z_y) \geq cm\}$. Now, we define a function ϕ as follows. If $z \in \mathcal{G}$, we write $\phi(z) = z$. Otherwise, let $z \in \mathcal{A} \setminus \mathcal{G}$. Since $\{Z_y\}$ is a partition of \mathbb{X} , we have

$$m \leq \tau(B(z, \gamma\delta)) = \sum_{y \in \mathcal{A}} \tau(B(z, \gamma\delta) \cap Z_y).$$

Since each $Z_y \subseteq B(y, \gamma\delta)$, it follows that at most c^{-1} of the Z_y 's have a nonempty intersection with $B(z, \gamma\delta)$. So, there must exist $y \in \mathcal{A}$ for which

$$\tau(B(z, \gamma\delta) \cap Z_y) \geq cm.$$

Clearly, each such y belongs to \mathcal{G} . We imagine an enumeration of \mathcal{A} , and among the y 's for which $\tau(B(z, \gamma\delta) \cap Z_y)$ is a maximum, pick the one with the lowest index. We then define $\phi(z)$ to be this y . Necessarily, $\phi(z) = y \in \mathcal{G}$, and $B(z, \gamma\delta) \cap Z_y \subseteq B(z, \gamma\delta) \cap B(y, \gamma\delta)$ is nonempty. So,

$$\rho(z, \phi(z)) \leq 2\gamma\delta, \quad B(z, \gamma\delta) \subseteq B(\phi(z), 3\gamma\delta), \quad \tau(B(z, \gamma\delta) \cap Z_{\phi(z)}) \geq cm. \tag{7.1}$$

Now, we define

$$Y_y = \cup\{Z_z : \phi(z) = y, z \in \mathcal{A}\}, \quad y \in \mathcal{G}.$$

For each $z \in \mathcal{A}$, $Z_z \subseteq Y_{\phi(z)}$. Since Z_z is a partition of \mathbb{X} , we have $\mathbb{X} = \cup_{y \in \mathcal{G}} Y_y$. If $x \in \mathbb{X}$, $x \in Y_y \cap Y_{y'}$ for $y, y' \in \mathcal{G}$, then $x \in Z_z$ with $\phi(z) = y$ and $x \in Z_{z'}$ with $\phi(z') = y'$. Since $\{Z_z\}$ is a partition of \mathbb{X} , it follows that $z = z'$, and hence $y = y'$. Thus, $\{Y_y\}$ is a partition of \mathbb{X} , $\tau(Y_y) \geq \tau(Z_y) \geq cm$, and

$$Y_y \subseteq \cup_{\phi(z)=y} Z_z \subseteq \cup_{\phi(z)=y} B(z, \gamma\delta) \subseteq B(y, 3\gamma\delta). \quad \square$$

Proof of Theorem 5.2. The partition Y_k is required to satisfy three goals: first, we wish to ensure that $\mu(Y_k) \sim d^\alpha$, second, we wish to be able to obtain a lower bound on $|\nu|(Y_k)$ as stated in Theorem 5.2, and finally, we wish to ensure that $x_k \in Y_k$. Accordingly, we will start with an appropriately dense subset of $\text{supp}(\nu)$, and construct a corresponding partition in a somewhat obvious manner. We will then use Lemma 7.2 three times to ensure the three goals; first with μ in place of τ , then with $|\nu|$ in place of τ with the resulting partition, and finally build a somewhat artificial measure τ supported on the finite subset obtained in the second step, and use the lemma with this measure. This ensures that each of the sets in the resulting partition contains at least one point of the set in the second step, but not necessarily the set created in the third step. However, this is easy to arrange with an increase in the mesh norm by a constant factor.

In this proof only, let $\mathcal{G}_1 = \{y_1, \dots, y_N\}$ be a maximal $d/2$ -separated subset of $\text{supp}(\nu)$. Then $\text{supp}(\nu) \subseteq \cup_{k=1}^N B(y_k, d/2)$, and $d/4 \leq q(\mathcal{G}_1)/2 \leq \delta(\mathcal{G}_1) \leq 3d/2$. We let $Z_{y_1,1} = B(y_1, \delta(\mathcal{G}_1))$, and for $k = 2, \dots, N$, $Z_{y_k,1} = B(y_k, \delta(\mathcal{G}_1)) \setminus \cup_{j=1}^{k-1} B(y_j, \delta(\mathcal{G}_1))$. Then $\{Z_{y,1}\}_{y \in \mathcal{G}_1}$ is a partition of \mathbb{X} and each $Z_{y,1} \subseteq B(y, \delta(\mathcal{G}_1))$, $y \in \mathcal{G}_1$.

We apply Lemma 7.2 first with μ in place of τ , resulting in a subset $\mathcal{G}_2 \subseteq \mathcal{G}_1$, $\delta(\mathcal{G}_2) \leq 3\delta(\mathcal{G}_1)$, and a partition $\{Z_{y,2}\}_{y \in \mathcal{G}_2}$ of \mathbb{X} such that for each $y \in \mathcal{G}_2$, $Z_{y,2} \subseteq B(y, 3\delta(\mathcal{G}_1)) \subseteq B(y, 3\delta(\mathcal{G}_2))$, and $c_1 \min_{z \in \mathcal{G}_1} \mu(B(z, \delta(\mathcal{G}_1))) \leq \mu(Z_{y,2}) \leq \mu(B(y, 3\delta(\mathcal{G}_1)))$; i.e., $\mu(Z_{y,2}) \sim (\delta(\mathcal{G}_1))^\alpha \sim d^\alpha$.

We apply Lemma 7.2 again with \mathcal{G}_2 in place of \mathcal{A} , $\{Z_{y,2}\}$ as the corresponding partition, and $|\nu|$ in place of τ . This yields a subset $\mathcal{G}_3 \subseteq \mathcal{G}_2$ and a partition $\{Z_{y,3}\}_{y \in \mathcal{G}_3}$ of \mathbb{X} with $\delta(\mathcal{G}_3) \leq 3\delta(\mathcal{G}_2)$, such that for each $y \in \mathcal{G}_3$, $Z_{y,2} \subseteq Z_{y,3} \subseteq B(y, 3\delta(\mathcal{G}_2)) \subseteq B(y, 9\delta(\mathcal{G}_1)) \cap B(y, 3\delta(\mathcal{G}_3))$, and $|\nu|(Z_{y,3}) \geq c_2 \min_{z \in \mathcal{G}_2} |\nu|(B(z, \delta(\mathcal{G}_2))) =: u$. Since $\mathcal{G}_2 \subseteq \text{supp}(\nu)$, u is a positive number. We note that $\mu(Z_{y,3}) \sim d^\alpha$ as well.

At this point, we still have not proved that $Z_{y,3} \cap \mathcal{G}_3$ is nonempty for each $y \in \mathcal{G}_3$. Towards this end, we repeat an application of Lemma 7.2 with the measure, to be denoted in this proof only by τ , that associates the mass $u > 0$ with each $y \in \mathcal{G}_3$. This gives us a subset $\mathcal{G}_4 \subseteq \mathcal{G}_3$ and a partition $\{Z_{y,4}\}_{y \in \mathcal{G}_4}$ with $\delta(\mathcal{G}_4) \leq 3\delta(\mathcal{G}_3)$, such that for each $y \in \mathcal{G}_4$, $Z_{y,2} \subseteq Z_{y,3} \subseteq Z_{y,4} \subseteq B(y, 3\delta(\mathcal{G}_3)) \subseteq B(y, 27\delta(\mathcal{G}_1))$, $\mu(Z_{y,4}) \sim d^\alpha$ and $\tau(Z_{y,4}) \geq c_3 \min_{z \in \mathcal{G}_3} \tau(B(z, \delta(\mathcal{G}_3))) \geq c_3 u > 0$. Necessarily, each $Z_{y,4}$ contains some element of \mathcal{G}_3 .

We pick one element from each $Z_{y,4} \cap \mathcal{G}_3$ to form the set $\mathcal{C} = \{x_1, \dots, x_M\}$, and rename the set $Z_{y,4}$ containing x_k to be Y_k . By construction, $\{Y_k\}_{k=1}^M$ is a partition of \mathbb{X} with each $x_k \in Y_k \subseteq B(x_k, 54\delta(\mathcal{G}_1))$, $\mu(Y_k) \sim d^\alpha$ and $|v|(Y_k) \geq c_3u$. \square

We are now in a position to prove [Theorem 5.3](#).

Proof of Theorem 5.3. We apply [Theorem 5.1](#) with the set \mathcal{C} and the partition $\{Y_k\}$. We observe that for any k ,

$$\begin{aligned} & \left| \int_{Y_k} |P(y)|^p d\mu(y) - \frac{\mu(Y_k)}{|v|(Y_k)} \int_{Y_k} |P(z)|^p d|v|(z) \right| \\ &= \left| \int_{Y_k} \left\{ |P(y)|^p - \frac{1}{|v|(Y_k)} \int_{Y_k} |P(z)|^p d|v|(z) \right\} d\mu(y) \right| \\ &= \frac{1}{|v|(Y_k)} \left| \int_{Y_k} \int_{Y_k} \{ |P(y)|^p - |P(z)|^p \} d|v|(z) d\mu(y) \right| \\ &\leq \mu(Y_k) \max_{z,y \in Y_k} \left| |P(y)|^p - |P(z)|^p \right|. \end{aligned}$$

Since $Y_k \subseteq B(x_k, 81d)$, [\(5.3\)](#) leads to [\(5.7\)](#). This completes the proof of part (a). Part (b) follows from [Theorem 5.1\(c\)](#). \square

We find it convenient to prove [Proposition 5.6](#) next, so that we may use such statements as [Proposition 6.5](#) which were proved with the definition as in [\(5.15\)](#) rather than the one which have used in this paper.

Proof of Proposition 5.6. In the proof of part (a) only, let $\lambda > \|v\|_{R,d}$, $r > 0$, $x \in \mathbb{X}$, and let $\{y_1, \dots, y_N\}$ be a maximal $2d/3$ -separated subset of $B(x, r + 2d/3)$. Then $B(x, r) \subseteq B(x, r + 2d/3) \subseteq \cup_{j=1}^N B(y_j, 2d/3)$. So,

$$|v|(B(x, r)) \leq |v|(B(x, r + 2d/3)) \leq \sum_{j=1}^N |v|(B(y_j, 2d/3)) \leq \sum_{j=1}^N |v|(B(y_j, d)) \leq \lambda Nd^\alpha.$$

The balls $B(y_j, d/3)$ are mutually disjoint, and $\cup_{j=1}^N B(y_j, d/3) \subseteq B(x, r + d)$. In view of [\(3.10\)](#), $d^\alpha \leq c\mu(B(y_j, d/3))$ for each j . So,

$$|v|(B(x, r)) \leq \lambda Nd^\alpha \leq c\lambda \sum_{j=1}^N \mu(B(y_j, d/3)) = c\lambda\mu(\cup_{j=1}^N B(y_j, d/3)) \leq c\lambda\mu(B(x, r + d)).$$

Since $\lambda > \|v\|_{R,d}$ was arbitrary, this leads to the first inequality in [\(5.15\)](#). The second inequality follows from [\(3.1\)](#). The converse statement is obvious. This completes the proof of part (a).

The second estimate in [\(5.16\)](#) is clear from the definitions. The first estimate in [\(5.16\)](#) follows by applying [\(5.15\)](#) with $r = \gamma d$. \square

In the proof of [Theorem 5.5](#), we will often need the following observation.

Lemma 7.3. *There exists a constant $\beta \in (0, 1/2)$ such that for any $L > 0$, $x \in \mathbb{X}$,*

$$|\Phi_L(x, y)| \geq (1/2)\Phi_L(x, x) \geq cL^\alpha, \quad \rho(x, y) \leq \beta/L. \tag{7.2}$$

Hence, for $1 \leq p < \infty$ and $v \in \mathcal{M}$,

$$\int_{B(x, \beta/L)} |\Phi_L(x, y)|^p d|v| \geq cL^{\alpha p} |v|(B(x, \beta/L)). \tag{7.3}$$

Proof. In this proof only, let $P(y) = \Phi_L(x, y)$, $y \in \mathbb{X}$, and $|P(y^*)| = \|P\|_{\mu; \infty}$. Then $P(x) = \Phi_L(x, x)$, and Schwarz inequality and (3.8) show that

$$\|P\|_{\mu; \infty} = |P(y^*)| \leq \Phi_L(x, x)^{1/2} \Phi_L(y^*, y^*)^{1/2} \leq c_1 L^\alpha \leq c_2 \Phi_L(x, x) = c_2 P(x) \leq c_2 \|P\|_{\mu; \infty}.$$

Thus, $\|P\|_{\mu; \infty} \sim L^\alpha$. Since $P \in \Pi_L$, we conclude from (6.20) that

$$|P(y) - P(x)| \leq c_3 L \rho(x, y) \|P\|_{\mu; \infty} = c_4 L \rho(x, y) P(x).$$

Hence, with $\beta = \min(1/2, 1/(2c_4))$, we obtain that $|P(y)| \geq (1/2)P(x) \geq cL^\alpha$ if $\rho(x, y) \leq \beta/L$. \square

We are now in a position to prove Theorem 5.5.

Proof of Theorem 5.5. Let ν be $1/L$ -regular, and without loss of generality, $\|\nu\|_{R, 1/L} = 1$. Using (6.16) with μ and ν , and the fact that both μ and ν are $1/L$ -regular, we deduce that

$$\sup_{x \in \mathbb{X}} \max \left\{ \int_{\mathbb{X}} |\Phi_{2L}(x, y)| d\mu(y), \int_{\mathbb{X}} |\Phi_{2L}(x, y)| d|\nu|(y) \right\} \leq c.$$

Therefore, using Fubini’s theorem, we conclude as in [27, Corollary 5.2] that for $p = 1, \infty$,

$$\|\sigma_{2L}(f)\|_{\nu; p} \leq c \|f\|_{\mu; p}.$$

Hence, the Riesz–Thorin interpolation theorem shows that this inequality is valid also for all p , $1 \leq p \leq \infty$. If $P \in \Pi_L$, we use this inequality with P in place of f , and recall that $\sigma_{2L}(P) = P$ to deduce that $\|P\|_{\nu; p} \leq c \|P\|_{\mu; p}$, as claimed in the first part of Theorem 5.5(a).

Conversely, suppose that for some p , $1 \leq p < \infty$,

$$\|P\|_{\nu; p}^p \leq A \|P\|_{\mu; p}^p, \quad P \in \Pi_L. \tag{7.4}$$

Let $x \in \mathbb{X}$. We apply (7.4) with $P(y) = \Phi_L(x, y)$. Using (6.17) with the $1/L$ -regular measure μ in place of ν , we see that $\|P\|_{\mu; p}^p \leq cL^{\alpha p/p'} = cL^{\alpha(p-1)}$. Therefore, (7.4) and (7.3) together imply that with β as defined in Lemma 7.3,

$$c_1 L^{\alpha p} |\nu|(B(x, \beta/L)) \leq \|P\|_{\nu; p}^p \leq cAL^{\alpha(p-1)}.$$

Thus, ν is β/L -regular, with $\|\nu\|_{R, \beta/L} \leq (c/c_1)A$. In view of (5.16) applied with $1/\beta$ in place of γ , this completes the proof of the converse statement in Theorem 5.5(a).

The proof of the first part of Theorem 5.5(b) relies upon Theorem 5.3(a). Let $1 \leq p < \infty$. With the constants c_1, c_2 as in that theorem, let $c_4 = c_2/(8p)$, $L \geq 4c_4 \max(81, c_1^{-1})$, $d = 4c_4/L$. Then $d \leq c_1$, and ν is $d/4$ -dominant. From the definition, this means that for every $x \in \mathbb{X}$, $\nu(B(x, d/4)) > 0$. In particular, each $B(x, d/4) \cap \text{supp}(\nu)$ is nonempty; i.e., $\delta(\text{supp}(\nu)) \leq d/4 < d \leq 1/81$. Therefore, all the conditions of Theorem 5.3(a) are satisfied so that (5.7) holds with $\eta = 1/2$. This shows that for every $P \in \Pi_L$,

$$(1/2) \|P\|_{\mu; p}^p \leq \sum_{k=1}^M \frac{\mu(Y_k)}{|\nu|(Y_k)} \int_{Y_k} |P(z)|^p d|\nu|(z). \tag{7.5}$$

Here, we recall that $\mu(Y_k) \sim d^\alpha$, and $|\nu|(Y_k) \geq c \min_{x \in \mathbb{X}} |\nu|(B(x, d/4))$. Since ν is $d/4$ -dominant, we have $|\nu|(Y_k) \geq c \|\nu\|_{D, d/4}^{-1} d^\alpha$. So, (7.5) leads to

$$\|P\|_{\mu; p}^p \leq c_3 \|\nu\|_{D, d/4} \sum_{k=1}^M \int_{Y_k} |P(z)|^p d|\nu|(z) = c_3 \|\nu\|_{D, c_4/L} \|P\|_{\nu; p}^p, \quad P \in \Pi_L.$$

This completes the proof of the first part of Theorem 5.5(b).

Finally, we prove the converse statement in Theorem 5.5(b). Accordingly, we assume that ν is $1/L$ -regular. In view of part (a) of this theorem and the assumption of the converse statement, we have with $A \sim \|\nu\|_{R, 1/L}^{-1/p}$,

$$A \|P\|_{\nu; p}^p \leq \|P\|_{\mu; p}^p \leq A_1 \|P\|_{\nu; p}^p, \quad P \in \Pi_L. \tag{7.6}$$

We will use Lemma 7.3 as before. Let $x \in \mathbb{X}$, and $P(y) = \Phi_L(x, y)$, and $r \geq 1$ to be chosen later. Using (7.3) with μ in place of ν , and (3.10), we obtain

$$\frac{c_7}{A_1} L^{\alpha(p-1)} \leq A_1^{-1} \|P\|_{\mu;p}^p \leq \|P\|_{\nu;p}^p = \int_{B(x,r/L)} |P(y)|^p d|\nu|(y) + \int_{\Delta(x,r/L)} |P(y)|^p d|\nu|(y). \tag{7.7}$$

Since ν is assumed to be $1/L$ -regular, we may apply (6.15) to conclude that

$$\int_{\Delta(x,r/L)} |P(y)|^p d|\nu|(y) \leq \|P\|_{\mu;\infty}^{p-1} \int_{\Delta(x,r/L)} |P(y)| d|\nu|(y) \leq c_8 L^{\alpha(p-1)} r^{\alpha-S} \|\nu\|_{R,1/L}. \tag{7.8}$$

We now choose

$$r = \max \left(1, \left(\frac{2c_8 A_1 \|\nu\|_{R,1/L}}{c_7} \right)^{1/(S-\alpha)} \right).$$

Then (7.7) and (7.8) together lead to

$$\frac{c_7}{2A_1} L^{\alpha(p-1)} \leq \int_{B(x,r/L)} |P(y)|^p d|\nu|(y) \leq c_9 L^{\alpha p} |\nu|(B(x, r/L)).$$

Thus,

$$|\nu|(B(x, r/L)) \geq c_{10} A_1^{-1} L^{-\alpha}.$$

This completes the proof of the converse statement. \square

Proof of Theorem 5.7. We find a finite subset $\mathcal{C} = \{x_1, \dots, x_M\}$ and a partition $\{Y_k\}$ as in Theorem 5.2. In view of (5.6), each $x_k \in Y_k \subseteq B(x_k, 324\delta(\mathcal{C}))$. Moreover, the conditions of Theorem 5.1 are satisfied with appropriate κ for this \mathcal{C} . In view of (3.11) and (5.3), we conclude that for a suitably chosen $c, L \leq c\delta(\mathcal{C})^{-1} \sim d^{-1}$,

$$\sum_{k=1}^M \mu(B(x_k, 324\delta(\mathcal{C}))) \max_{y,z \in B(x, 324\delta(\mathcal{C}))} |P(y) - P(z)| \leq (1/4) \|P\|_{\mu;1}. \tag{7.9}$$

In this proof only, let

$$x_k^*(P) = \frac{1}{|\nu|(Y_k)} \int_{Y_k} P(z) d|\nu|(z), \quad k = 1, \dots, M.$$

Then for $P \in \Pi_L$, and $k = 1, \dots, M$,

$$\begin{aligned} \left| \int_{Y_k} |P(y)| d\mu(y) - \mu(Y_k) |x_k^*(P)| \right| &= \left| \int_{Y_k} \{ |P(y)| - |x_k^*(P)| \} d\mu(y) \right| \\ &\leq \int_{Y_k} | |P(y)| - |x_k^*(P)| | d\mu(y) \\ &\leq \int_{Y_k} |P(y) - x_k^*(P)| d\mu(y) \\ &\leq \frac{1}{|\nu|(Y_k)} \int_{Y_k} \int_{Y_k} |P(y) - P(z)| d|\nu|(z) d\mu(y) \\ &\leq \mu(B(x, 324\delta(\mathcal{C}))) \max_{y,z \in B(x, 324\delta(\mathcal{C}))} |P(y) - P(z)|. \end{aligned} \tag{7.10}$$

Then (7.10) and (7.9) imply that

$$\left| \sum_{k=1}^M \int_{Y_k} |P(y)| d\mu(y) - \sum_{k=1}^M \mu(Y_k) |x_k^*(P)| \right| \leq (1/4) \|P\|_{\mu;1};$$

i.e.,

$$(3/4)\|P\|_{\mu;1} \leq \sum_{k=1}^M \mu(Y_k)|x_k^*(P)| \leq (5/4)\|P\|_{\mu;1}. \tag{7.11}$$

Moreover, if each $x_k^*(P) \geq 0$, then the same estimate as (7.10) with P in place of $|P|$ leads to

$$\left| \int_{\mathbb{X}} P(y)d\mu(y) - \sum_{k=1}^M \mu(Y_k)x_k^*(P) \right| \leq (1/4)\|P\|_{\mu;1} \leq (1/3) \sum_{k=1}^M \mu(Y_k)x_k^*(P).$$

Thus, if each $x_k^*(P) \geq 0$, then

$$\int_{\mathbb{X}} P(y)d\mu(y) \geq (2/3) \sum_{k=1}^M \mu(Y_k)x_k^*(P) \geq 0. \tag{7.12}$$

Now, we wish to use Theorem 6.8. We let \mathbb{X} be the space \mathbb{R}^M , equipped with the norm $\|y\| = \sum_{k=1}^M \mu(Y_k)|y_k|$, where $y = (y_1, \dots, y_k)$. For the set \mathcal{K} , we choose the set of coordinate functionals; $y_k^*(y) = y_k$. Then \mathcal{K} is clearly a compact subset of \mathbb{X}^* . We consider the operator $\mathcal{J} : \Pi_L \rightarrow \mathbb{R}^M$ given by $P \mapsto (x_1^*(P), \dots, x_M^*(P))$, and take the subspace \mathcal{V} of \mathbb{X} to be the range of \mathcal{J} . The lower estimate in (7.11) shows that \mathcal{J} is invertible on \mathcal{V} . We define the functional x^* on \mathcal{V} by

$$x^*(\mathcal{J}(P)) = \int_{\mathbb{X}} P(z)d\mu(z) - (1/3) \sum_{k=1}^M \mu(Y_k)x_k^*(P), \quad P \in \Pi_L.$$

Our observations in the previous paragraph show that x^* is positive on \mathcal{V} with respect to \mathcal{K} . The element $(1, \dots, 1) \in \mathcal{V}$ serves as v_0 in Theorem 6.8. Theorem 6.8 then implies that there exists a nonnegative functional X^* on $\mathbb{X} = \mathbb{R}^M$ that extends x^* . We may identify this functional with $(\tilde{W}_1, \dots, \tilde{W}_M) \in \mathbb{R}^M$, such that each $\tilde{W}_k \geq 0$. The fact that X^* extends x^* means that for each $P \in \Pi_L$,

$$\int_{\mathbb{X}} P(x)d\mu(x) = \sum_{k=1}^M (\tilde{W}_k + (1/3)\mu(Y_k))x_k^*(P) =: \sum_{k=1}^M \frac{W_k}{|v|(Y_k)} \int_{Y_k} P(y)d|v|(y). \tag{7.13}$$

Writing $W(y) = W_k/|v|(Y_k)$ for $y \in Y_k$, we have now proved (5.17). By construction, $W_k \geq (1/3)\mu(Y_k) \geq c_1d^\alpha$. If v is d -regular, then $|v|(Y_k) \leq |v|(B(x_k, 81d)) \leq c_2\|v\|_{R,d}d^\alpha$. Hence, $W(y) \geq c\|v\|_{R,d}^{-1}$ for all $y \in \mathbb{X}$. \square

The proof of Theorem 5.8 uses the following lemma proved in [27, Lemma 5.5]:

Lemma 7.4. *Let the product assumption hold, and $L > 0$. If v is a quadrature measure of order $2A^*L$, $|v|(\mathbb{X}) \leq c$, and $P_1, P_2 \in \Pi_{2L}$ then for any $p, r, 1 \leq p, r \leq \infty$ and any positive number $R > 0$,*

$$\left| \int_{\mathbb{X}} P_1P_2d\mu - \int_{\mathbb{X}} P_1P_2dv \right| \leq c_1L^{2\alpha}\epsilon_L\|P_1\|_{\mu;p}\|P_2\|_{\mu;r} \leq c(R)L^{-R}\|P_1\|_{\mu;p}\|P_2\|_{\mu;r}. \tag{7.14}$$

Proof of Theorem 5.8. Let $x \in \mathbb{X}$, and $P^* := \Phi_L(x, \circ) \in \Pi_L$. Taking β as in Lemma 7.3, we obtain from (7.3) applied with τ in place of v that

$$c_1L^{2\alpha}\tau(B(x, \beta/L)) \leq \int_{B(x, \beta/L)} |P^*(y)|^2d\tau(y) \leq \int_{\mathbb{X}} |P^*(y)|^2d\tau(y).$$

Since τ is a positive quadrature measure of order $2A^*L$, we now obtain from Lemma 7.4 used with $P_1 = P_2 = P^*, p = r = 1, R = 1$ that

$$c_1L^{2\alpha}\tau(B(x, \beta/L)) \leq \int_{\mathbb{X}} |P^*(y)|^2d\mu(y) + c\|P^*\|_{\mu;1}^2/L = \Phi_L(x, x) + c_2\|P^*\|_{\mu;1}^2/L.$$

In view of (6.12) and (3.8), $\|P^*\|_{\mu;1} \leq c_3$, $\Phi_L(x, x) \leq c_3L^\alpha$. We deduce that for sufficiently large L :

$$L^{2\alpha} \tau(B(x, \beta/L)) \leq c_4L^\alpha;$$

i.e., $\tau(B(x, \beta/L)) \leq c_4L^{-\alpha}$. In view of Theorem 5.5, this implies that $\|P\|_{\tau;p} \leq c_5\|P\|_{\mu;p}$ for all p with $1 \leq p < \infty$, and $P \in \Pi_L$.

It remains to prove that

$$\|P\|_{\mu;p} \leq c\|P\|_{\tau;p}, \quad P \in \Pi_L, \quad 1 \leq p < \infty. \tag{7.15}$$

Towards this end, we introduce the discretized version of the operator σ_L :

$$\sigma_L(\tau; f, x) = \int_{\mathbb{X}} \Phi_L(x, y)f(y)d\tau(y), \quad f \in L^1(\tau), \quad x \in \mathbb{X}, \quad L > 0.$$

If $f \in L^p(\tau)$ and $g \in L^{p'}(\mu)$, then it is easy to verify using Fubini's theorem that

$$\begin{aligned} \int_{\mathbb{X}} \sigma_L(\tau; f, x)g(x)d\mu(x) &= \int_{\mathbb{X}} \int_{\mathbb{X}} \Phi_L(x, y)f(y)d\tau(y)g(x)d\mu(x) \\ &= \int_{\mathbb{X}} \int_{\mathbb{X}} \Phi_L(x, y)g(x)d\mu(x)f(y)d\tau(y) \\ &= \int_{\mathbb{X}} \sigma_L(g, y)f(y)d\tau(y). \end{aligned}$$

Using the duality principle and (6.13), we conclude that

$$\begin{aligned} \|\sigma_L(\tau; f)\|_{\mu;p} &= \sup_{\|g\|_{\mu;p'}=1} \left| \int_{\mathbb{X}} \sigma_L(\tau; f, x)g(x)d\mu(x) \right| = \sup_{\|g\|_{\mu;p'}=1} \left| \int_{\mathbb{X}} \sigma_L(g, y)f(y)d\tau(y) \right| \\ &\leq \sup_{\|g\|_{\mu;p'}=1} \|\sigma_L(g)\|_{\mu;p'} \|f\|_{\tau;p} \leq c\|f\|_{\tau;p}. \end{aligned}$$

In particular,

$$\|\sigma_{2L}(\tau; f)\|_{\mu;p} \leq c\|f\|_{\tau;p}. \tag{7.16}$$

Since we do not assume that the product of polynomials in Π_{2L} is itself in Π_{2A^*L} , it does not follow from the fact that τ is a quadrature measure of order $2A^*L$ that $\sigma_{2L}(\tau; P) = P$ for $P \in \Pi_L$. Nevertheless, Lemma 7.4 allows us to conclude that $\sigma_{2L}(\tau; P)$ is close to P if $P \in \Pi_L$. Let $x \in \mathbb{X}$. Since τ is a quadrature measure of order $2A^*L$, we may use Lemma 7.4 with $R = 1$, $r = 1$, P in place of P_1 and $\Phi_{2L}(x, \circ)$ in place of P_2 to conclude that

$$\begin{aligned} |\sigma_{2L}(\tau; P, x) - P(x)| &= |\sigma_{2L}(\tau; P, x) - \sigma_{2L}(P, x)| \\ &= \left| \int_{\mathbb{X}} P(y)\Phi_{2L}(x, y)d\tau(y) - \int_{\mathbb{X}} P(y)\Phi_{2L}(x, y)d\mu(y) \right| \\ &\leq cL^{-1}\|\Phi_{2L}(x, \circ)\|_{\mu;1}\|P\|_{\mu;p}. \end{aligned}$$

In view of (6.12), and the fact that μ is a probability measure, this implies that

$$\|\sigma_{2L}(\tau; P)\|_{\mu;p} - \|P\|_{\mu;p} \leq \|\sigma_{2L}(\tau; P) - P\|_{\mu;p} \leq cL^{-1}\|P\|_{\mu;p}.$$

Thus, if $L \geq 2c$, then (7.16) implies

$$\|P\|_{\mu;p} \leq 2\|\sigma_{2L}(\tau; P)\|_{\mu;p} \leq c_4\|P\|_{\tau;p}.$$

This proves (7.15) and completes the proof of the theorem. \square

Acknowledgments

The research of the first author was partially funded by Deutsche Forschungsgemeinschaft grant FI 883/3-1.

The research of the second author was supported, in part, by grant DMS-0908037 from the National Science Foundation and grant W911NF-09-1-0465 from the US Army Research Office.

Appendix

In this Appendix, we follow an idea in the paper [12] of Geller and Pesenson to prove that the product assumption is valid when ϕ_k (respectively, ℓ_k^2) are eigenfunctions (respectively, eigenvalues) of a selfadjoint, uniformly elliptic partial differential operator of second order satisfying some technical conditions. In the Appendix, we will use the symbol q again to denote the dimension of the manifold \mathbb{X} .

First, we need to recall the notion of exponential maps on the manifold. For any $x \in \mathbb{X}$, there exists a neighborhood V of x , a number $\epsilon > 0$ and a C^∞ mapping $\gamma : (-2, 2) \times \mathcal{U} \rightarrow \mathbb{X}$, where $\mathcal{U} = \{(y, w) \in T\mathbb{X} : y \in V, w \in T_y\mathbb{X}, \|w\|_y < \epsilon\}$, such that $\gamma(\circ, y, w)$ is the unique geodesic of \mathbb{X} with $\gamma(0, y, w) = y$, and the tangent vector at y being w [9, Proposition 2.7]. In this Appendix, we will denote by $B_\epsilon(0)$ the open Euclidean ball in \mathbb{R}^q with center at 0 and radius ϵ . For any tangent space $T_x\mathbb{X}$, we may consider an appropriate coordinate chart on $T_x\mathbb{X}$, and view $B_\epsilon(0)$ as a subset of $T_x\mathbb{X}$, with 0 corresponding to x . The exponential map at x is the mapping $\exp_x : B_\epsilon(0) \subset T_x\mathbb{X} \rightarrow \mathbb{X}$ defined by $\exp_x(w) = \gamma(1, x, w)$, where γ is the mapping just described. Thus, $\exp_x(w)$ is the point on \mathbb{X} where one reaches by following the geodesic at x , with tangent vector given by $w/\|w\|$, for a length of $\|w\|$. For every $x \in \mathbb{X}$, there exists an $\epsilon > 0$ such that $\exp_x : B_\epsilon(0) \subset T_x\mathbb{X} \rightarrow \mathbb{X}$ is a diffeomorphism of $B_\epsilon(0)$ onto an open subset of \mathbb{X} . Since \mathbb{X} is compact, we may choose ϵ to be the same for all $x \in \mathbb{X}$. The largest value of such ϵ is called the injectivity radius of \mathbb{X} , to be denoted in this Appendix by ι . If $\delta > 0$ and $\overline{B_\delta(0)} \subset B_\iota(0) \subset T_x\mathbb{X}$, then $\exp_x(B_\delta(0))$ is called a normal ball of radius δ centered at x . Normal neighborhoods of x are defined in the obvious way. If $\{\partial_j\}$ is a basis for $T_x\mathbb{X}$, the normal coordinate system (with respect to $\{\partial_j\}$) at x is defined on a normal neighborhood of x by $\mathbf{x}(u_1, \dots, u_q) = \exp_x(\sum_{j=1}^q u_j \partial_j)$.

Let \mathcal{P} be a self-adjoint differential operator of second order. In terms of a normal coordinate system at a point $x \in \mathbb{X}$ the operator \mathcal{P} can be expressed in the form

$$\mathcal{P}f = \sum_{\mathbf{k} \in \mathbb{Z}^q, \mathbf{k} \geq 0, |\mathbf{k}| \leq 2} a_{\mathbf{k},x}(\mathbf{u}) \frac{\partial^{|\mathbf{k}|} f}{\partial \mathbf{u}^{\mathbf{k}}}.$$

The operator is strongly uniformly elliptic if there are constants $c_1, c_2 > 0$ (independent of x) such that

$$c_1 \|\mathbf{y}\|^2 \leq \sum_{|\mathbf{k}|=2} a_{\mathbf{k},x}(\mathbf{u}) \mathbf{y}^{\mathbf{k}} \leq c_2 \|\mathbf{y}\|^2, \quad \mathbf{u} \in B_\iota(0), x \in \mathbb{X}, \mathbf{y} \in \mathbb{R}^q.$$

We assume that there exists a constant $C > 1$ such that

$$\left| \frac{\partial^{|\mathbf{m}|} a_{\mathbf{k},x}}{\partial \mathbf{u}^{\mathbf{m}}} \right| \leq C^{|\mathbf{m}|}, \quad x \in \mathbb{X}, \mathbf{u} \in B_\iota(0), \mathbf{m} \in \mathbb{Z}^q. \tag{A.1}$$

The eigenvalues of \mathcal{P} can be enumerated in the form $\{\ell_k^2\}$ ($\ell_k \uparrow \infty$), and we let ϕ_k be the eigenfunction corresponding to ℓ_k^2 . We assume (by choosing a larger C if necessary) that

$$\sum_{\ell_k \leq L} \left| \frac{\partial^{|\mathbf{m}|} \phi_k(x)}{\partial \mathbf{u}^{\mathbf{m}}} \right|^2 \leq C^{2|\mathbf{m}|} L^{q+2|\mathbf{m}|}, \quad x \in \mathbb{X}, \mathbf{m} \in \mathbb{Z}^q, L > 0. \tag{A.2}$$

This result follows essentially from the estimates on the derivatives of the heat kernel corresponding to \mathcal{P} given by Kordyukov [21, Theorem 5.5], and the Tauberian theorem in our paper [11, Proposition 4.1], except for the dependence of the constants involved on \mathbf{m} . An immediate consequence of (A.2)

is the following. If $Q = \sum_k \hat{Q}(k)\phi_k \in \Pi_L$, then

$$\begin{aligned} \left| \frac{\partial^{|\mathbf{m}|} Q(x)}{\partial \mathbf{u}^{\mathbf{m}}} \right|^2 &= \left| \sum_k \hat{Q}(k) \frac{\partial^{|\mathbf{m}|} \phi_k(x)}{\partial \mathbf{u}^{\mathbf{m}}} \right|^2 \leq \left\{ \sum_k |\hat{Q}(k)|^2 \right\} \left\{ \sum_{\ell_k \leq L} \left| \frac{\partial^{|\mathbf{m}|} \phi_k(x)}{\partial \mathbf{u}^{\mathbf{m}}} \right|^2 \right\} \\ &\leq C^{2|\mathbf{m}|} L^{q+2|\mathbf{m}|} \|Q\|_{\mu;2}^2; \end{aligned}$$

i.e.,

$$\left| \frac{\partial^{|\mathbf{m}|} Q(x)}{\partial \mathbf{u}^{\mathbf{m}}} \right| \leq L^{q/2} (CL)^{|\mathbf{m}|} \|Q\|_{\mu;2}. \tag{A.3}$$

We are now ready to prove the product assumption, in fact, a much stronger statement:

Theorem A.1. *Let \mathcal{P} be a second order, strongly uniformly elliptic, self-adjoint, partial differential operator on a smooth, compact manifold \mathbb{X} (without boundary), the eigenvalues of \mathcal{P} be enumerated in the form $\{\ell_k^2\}$ ($\ell_k \uparrow \infty$), and we let ϕ_k be the eigenfunction corresponding to ℓ_k^2 . Assume further that (A.1) and (A.2) are satisfied. There exists $A^* \geq 2$ such that if $Q, R \in \Pi_L$, then $QR \in \Pi_{A^*L}$.*

Proof. Let $N \geq 2$ be an integer. In view of Leibniz’s formula, one can write

$$\mathcal{P}^N(QR)(x) = \sum_{\substack{\mathbf{k}, \mathbf{m} \in \mathbb{Z}^q, \mathbf{k}, \mathbf{m} \geq 0, \\ |\mathbf{k}| + |\mathbf{m}| \leq 2N}} b_{\mathbf{k}, \mathbf{m}}(x) \frac{\partial^{|\mathbf{m}|} Q(x)}{\partial \mathbf{u}^{\mathbf{m}}} \frac{\partial^{|\mathbf{k}|} R(x)}{\partial \mathbf{u}^{\mathbf{k}}},$$

where $b_{\mathbf{k}, \mathbf{m}}$ ’s are products of derivatives of the coefficients $a_{\mathbf{k}, x}$ in \mathcal{P} . In view of (A.1) and (A.2), we conclude that for some $A^* \geq 2$,

$$\|\mathcal{P}^N(QR)\|_{\mu;2} \leq \|\mathcal{P}^N(QR)\|_{\mu;\infty} \leq L^q (A^*L/2)^{2N} \|Q\|_{\mu;2} \|R\|_{\mu;2}. \tag{A.4}$$

In this proof only, let for $f \in L^2(\mu)$,

$$S_L(f) = \sum_{\ell_k \leq A^*L} \hat{f}(k)\phi_k.$$

We observe that $\|f - S_L(f)\|_{\mu;2} \leq \|f\|_{\mu;2}$. Since $\mathcal{P}\phi_k = \ell_k^2\phi_k$, it follows that $\mathcal{P}^N(S_L(QR)) = S_L(\mathcal{P}^N(QR))$. Consequently,

$$\begin{aligned} \|\mathcal{P}^N(QR - S_L(QR))\|_{\mu;2} &= \|\mathcal{P}^N(QR) - S_L(\mathcal{P}^N(QR))\|_{\mu;2} \leq \|\mathcal{P}^N(QR)\|_{\mu;2} \\ &\leq L^q (A^*L/2)^{2N} \|Q\|_{\mu;2} \|R\|_{\mu;2}. \end{aligned} \tag{A.5}$$

On the other hand, Parseval’s identity shows that

$$\begin{aligned} \|\mathcal{P}^N(QR - S_L(QR))\|_{\mu;2}^2 &= \sum_{\ell_k > A^*L} \ell_k^{4N} |\widehat{(QR)}(k)|^2 \geq (A^*L)^{4N} \sum_{\ell_k > A^*L} |\widehat{(QR)}(k)|^2 |\widehat{(QR)}|^2 \\ &\geq (A^*L)^{4N} \|QR - S_L(QR)\|_{\mu;2}^2. \end{aligned}$$

Together with (A.5), this implies that for every integer $N \geq 2$,

$$\|QR - S_L(QR)\|_{\mu;2} \leq L^q \left(\frac{A^*L}{2A^*L} \right)^{2N} \|Q\|_{\mu;2} \|R\|_{\mu;2} = L^q 4^{-N} \|Q\|_{\mu;2} \|R\|_{\mu;2}.$$

Letting $N \rightarrow \infty$, we conclude that $QR = S_L(QR) \in \Pi_{A^*L}$. This completes the proof. \square

References

[1] W.M. Boothby, An Introduction to Differentiable Manifolds and Riemannian Geometry, Academic Press, Elsevier, Amsterdam, 2003.
 [2] S.Y. Cheng, P. Li, S.-T. Yau, On the upper estimate of the heat kernel of a complete Riemannian manifold, Amer. J. Math. 103 (5) (1981) 1021–1063.

- [3] C.K. Chui, D.L. Donoho, Special issue: diffusion maps and wavelets, *Appl. Comput. Harmon. Anal.* 21 (1) (2006).
- [4] R.R. Coifman, M. Maggioni, Diffusion wavelets, *Appl. Comput. Harmon. Anal.* 21 (2006) 53–94.
- [5] S.B. Damelin, J. Levesley, Hyperinterpolation, packing and spherical designs on projective spaces, Manuscript, Private communication.
- [6] G. David, Wavelets and Singular Integrals on Curves and Surfaces, in: *Lecture Notes in Mathematics*, vol. 1465, Springer Verlag, Berlin, 1992.
- [7] E.B. Davies, L^p spectral theory of higher-order elliptic differential operators, *Bull. Lond. Math. Soc.* 29 (1997) 513–546.
- [8] M.P. do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice Hall, New Jersey, 1976.
- [9] M.P. do Carmo, *Riemannian Geometry*, Birkhäuser, Boston, 1992.
- [10] N. Dungey, Some remarks on gradient estimates for heat kernels, *Abstr. Appl. Anal.* (2006) Art. ID 73020, 10 pp.
- [11] F. Filbir, H.N. Mhaskar, A quadrature formula for diffusion polynomials corresponding to a generalized heat kernel, *J. Fourier Anal. Appl.* 16 (2010) 629–657.
- [12] D. Geller, I.Z. Pesenson, Band-limited localized Parseval frames and Besov spaces on compact homogeneous manifolds, 2010, arXiv:1002.3841v1 [math.FA].
- [13] A. Grigor'yan, Upper bounds of derivatives of the heat kernel on an arbitrary complete manifold, *J. Funct. Anal.* 127 (2) (1995) 363–389.
- [14] A. Grigor'yan, Estimates of heat kernels on Riemannian manifolds, in: E.B. Davies, Y. Safarov (Eds.), *Spectral Theory and Geometry*, Edinburgh, 1998, in: *London Math. Soc. Lecture Note Ser.*, vol. 273, Cambridge University Press, Cambridge, 1999, pp. 140–225.
- [15] A. Grigor'yan, Heat kernels on weighted manifolds and applications, *Contemp. Math.* 398 (2006) 93–191.
- [16] A. Grigor'yan, Heat kernels and function theory on metric measure spaces, *Cont. Math.* 338 (2003) 143–172.
- [17] K. Hesse, I.H. Sloan, Hyperinterpolation on the sphere, in: *Frontiers in Interpolation and Approximation*, in: *Pure Appl. Math. (Boca Raton)*, 282, Chapman & Hall/CRC, Boca Raton, FL, 2007, pp. 213–248.
- [18] E. Hewitt, K. Stromberg, *Real and Abstract Analysis*, Springer, New York, 1975.
- [19] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
- [20] P.W. Jones, M. Maggioni, R. Schul, Universal local parameterizations via heat kernels and eigenfunctions of the Laplacian, *Ann. Acad. Sci. Fenn. Math.* 35 (1) (2010) 131–174.
- [21] Yu.A. Kordyukov, L^p -theory of elliptic differential operators on manifolds of bounded geometry, *Acta Appl. Math.* 23 (3) (1991) 223–260.
- [22] S. Lafon, *Diffusion maps and geometric harmonics*, Ph.D. Thesis, Yale University, Dept of Mathematics & Applied Mathematics, 2004.
- [23] J.M. Lee, *Introduction to Smooth Manifolds*, Springer, New York, 2003.
- [24] Q.T. Le Gia, H.N. Mhaskar, Localized linear polynomial operators and quadrature formulas on the sphere, *SIAM J. Numer. Anal.* 47 (1) (2008) 440–466.
- [25] D.S. Lubinsky, Marcinkiewicz inequalities: methods and results, in: G.V. Milovanovic (Ed.), *Recent Progress in Inequalities*, Kluwer, Dordrecht, 1998, pp. 213–240.
- [26] M. Maggioni, H.N. Mhaskar, Diffusion polynomial frames on metric measure spaces, *Appl. Comput. Harmon. Anal.* 24 (3) (2008) 329–353.
- [27] H.N. Mhaskar, Eignets for function approximation on manifolds, *Appl. Comput. Harmon. Anal.* 29 (2010) 63–87.
- [28] H.N. Mhaskar, F.J. Narcowich, J.D. Ward, Spherical Marcinkiewicz–Zygmund inequalities and positive quadrature, *Math. Comp.* 70 (235) (2001) 1113–1130. Corrigendum: *Math. Comp.* 71, 2001, pp. 453–454.
- [29] B. O'Neill, *Elementary Differential Geometry*, Elsevier, Amsterdam, 2006.
- [30] M.E. Taylor, *Partial Differential Equations, Basic Theory*, Springer Verlag, New York, 1996.
- [31] A. Zygmund, *Trigonometric Series*, Cambridge University Press, Cambridge, 1977.