



# Weighted Alpert Wavelets

Rob Rahm<sup>1</sup> · Eric T. Sawyer<sup>2</sup> · Brett D. Wick<sup>3</sup>

Received: 14 June 2019 / Revised: 15 May 2020 / Published online: 23 November 2020  
© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

In this paper we construct a wavelet basis in  $L^2(\mathbb{R}^n; \mu)$  possessing vanishing moments of a fixed order for a general locally finite positive Borel measure  $\mu$ . The approach is based on a clever construction of Alpert in the case of the Lebesgue measure that is appropriately modified to handle the general measures considered here. We then use this new wavelet basis to study a two-weight inequality for a general Calderón–Zygmund operator on  $\mathbb{R}$  and conjecture that under suitable natural conditions, including a *weaker* energy condition, the operator is bounded from  $L^2(\mathbb{R}; \sigma)$  to  $L^2(\mathbb{R}; \omega)$  if certain *stronger* testing conditions hold on polynomials. An example is provided showing that this conjecture is logically different from existing results in the literature.

**Keywords** Calderon-Zygmund operators · Two-Weight inequality · Alpert basis

## 1 Introduction and Statement of Main Results

The use of weighted Haar wavelet expansions has its roots in connection with the  $Tb$  theorem in [2,3], and came to fruition in treating the two weight norm inequality for the Hilbert transform in [9,13], the two part paper [5,6,8]. The key features of the weighted Haar expansion  $\{h_I^\mu\}_{I \in \mathcal{D}}$  are threefold (below,  $\mathcal{D}$  is a dyadic lattice; explained below):

---

Communicated by Stephan Dahlke.

✉ Rob Rahm  
robrahm@math.tamu.edu  
Eric T. Sawyer  
sawyer@mcmaster.ca

<sup>1</sup> Texas A&M University, College Station, TX, USA

<sup>2</sup> McMaster University, Hamilton, Canada

<sup>3</sup> Washington University, St. Louis, USA

(1) The Haar functions  $\{h_I^\mu\}_{I \in \mathcal{D}}$  form an orthonormal basis of  $L^2(\mu)$ :

$$f = \sum_{I \in \mathcal{D}} \langle f, h_I^\mu \rangle_{L^2(\mu)} h_I^\mu \text{ both pointwise } \mu \text{-a.e. and in } L^2(\mu),$$

where  $\langle h_J^\mu, h_I^\mu \rangle_{L^2(\mu)} = \delta_I^J$ ,

(2) Telescoping identities hold:

$$\mathbf{1}_K \sum_{I \in \mathcal{D}: K \not\subseteq I \subset L} \langle f, h_I^\mu \rangle_{L^2(\mu)} h_I^\mu = E_K^\mu f - E_L^\mu f, \quad K \subsetneq L,$$

(3) Moment vanishing conditions hold:

$$\int h_I^\mu(x) d\mu(x) = 0, \quad I \in \mathcal{D}.$$

In the setting of Lebesgue measure, Alpert [1] introduced new wavelets with more vanishing moments in (3), while retaining orthonormality (1) and telescoping (2). The expense of imposing these extra moment conditions is that one requires additional functions in order to obtain the expansion. The purpose of this note is to extend existence of Alpert wavelets to arbitrary locally finite positive Borel measures in Euclidean space  $\mathbb{R}^n$ , and to investigate degeneracy and uniqueness in the one-dimensional case as well. To state the main result in this paper requires some notation.

Let  $\mu$  be a locally finite positive Borel measure on  $\mathbb{R}^n$ , and fix  $k \in \mathbb{N}$ . Let  $\mathcal{P}^n$  denote the collection of cubes with sides parallel to the coordinate axes. If  $Q$  is such a cube, its “dyadic children” or just its “children”—denoted  $\mathcal{C}(Q)$ —are the  $2^n$  subcubes whose side lengths are half the side length of  $Q$ . So, for example, if  $n = 1$  and  $Q = [0, 1)$ , then the two children are  $[0, \frac{1}{2})$  and  $[\frac{1}{2}, 1)$ . Or if  $n = 2$  and  $Q = [0, 1) \times [0, 1)$ , the four children are  $[0, \frac{1}{2}) \times [0, \frac{1}{2})$ ,  $[0, \frac{1}{2}) \times [\frac{1}{2}, 1)$ ,  $[\frac{1}{2}, 1) \times [0, \frac{1}{2})$ ,  $[\frac{1}{2}, 1) \times [\frac{1}{2}, 1)$ .

For  $Q \in \mathcal{P}^n$ , denote by  $L_{Q;k}^2(\mu)$  the finite dimensional subspace of  $L^2(\mu)$  that consists of linear combinations of the indicators of the children  $\mathcal{C}(Q)$  of  $Q$  multiplied by polynomials of degree at most  $k - 1$ , and such that the linear combinations have vanishing  $\mu$ -moments on the cube  $Q$  up to order  $k - 1$ :

$$L_{Q;k}^2(\mu) \equiv \left\{ f = \sum_{Q' \in \mathcal{C}(Q)} \mathbf{1}_{Q'} p_{Q';k}(x) : \int_Q f(x) x_i^\ell d\mu(x) = 0, \right. \\ \left. \text{for } 0 \leq \ell \leq k - 1 \text{ and } 1 \leq i \leq n \right\}, \tag{1.1}$$

where  $p_{Q';k}(x) = \sum_{\alpha \in \mathbb{Z}_+^n: |\alpha| \leq k-1} a_{Q';\alpha} x^\alpha$  is a polynomial in  $\mathbb{R}^n$  of degree  $|\alpha| = \alpha_1 + \dots + \alpha_n$  at most  $k - 1$ . Here  $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$ . Let  $d_{Q;k} \equiv \dim L_{Q;k}^2(\mu)$  be the dimension of the finite dimensional linear space  $L_{Q;k}^2(\mu)$ .

Define

$$\mathcal{F}_\infty^k(\mu) \equiv \left\{ \alpha \in \mathbb{Z}_+^n : |\alpha| \leq k - 1 : x^\alpha \in L^2(\mu) \right\}, \tag{1.2}$$

$$\text{and } \mathcal{P}_{\mathbb{R}^n}^k(\mu) \equiv \text{Span} \left\{ x^\alpha \right\}_{\alpha \in \mathcal{F}_\infty^k}. \tag{1.3}$$

Let  $\Delta_{Q;k}^\mu$  denote orthogonal projection onto the finite dimensional subspace  $L_{Q;k}^2(\mu)$ , let  $\mathbb{E}_{Q;k}^\mu$  denote orthogonal projection onto the finite dimensional subspace

$$\text{Span}\{\mathbf{1}_Q(x) x^\alpha : 0 \leq |\alpha| \leq k - 1\}, \tag{1.4}$$

and let  $\Delta_{\mathbb{R}^n;k}^\mu$  denote orthogonal projection onto  $\mathcal{P}_{\mathbb{R}^n}^k(\mu)$ . The projections  $\Delta_{Q;k}^\mu$  are often referred to as multiresolution projections. Observe that in many cases,  $\mathcal{F}_\infty^k(\mu)$  is empty and so  $\Delta_{\mathbb{R}^n;k}^\mu = 0$ .

Finally, by a dyadic lattice on  $\mathbf{R}$ , we mean the collection of intervals of the form  $[\frac{j}{2^m}, \frac{j+1}{2^m})$  where  $j, m \in \mathbf{Z}$ . A dyadic lattice on  $\mathbf{R}^n$  is the collection of cubes of the form  $I_1 \times \dots \times I_n$  where each  $I_k$  is a member of the dyadic lattice on  $\mathbf{R}$  (and, as the name ‘‘cube’’ implies, the side lengths of each factor are equal).

The first of two main results proved in this note is the following theorem, which establishes the existence of Alpert wavelets in all dimensions having the three important properties of orthogonality, telescoping and moment vanishing.

**Theorem 1** (Weighted Alpert Bases) *Let  $\mu$  be a locally finite positive Borel measure on  $\mathbb{R}^n$ , fix  $k \in \mathbb{N}$ , and fix a dyadic grid  $\mathcal{D}$  in  $\mathbb{R}^n$ .*

(1) *Then  $\left\{ \Delta_{\mathbb{R}^n;k}^\mu \right\} \cup \left\{ \Delta_{Q;k}^\mu \right\}_{Q \in \mathcal{D}}$  is a complete set of orthogonal projections in  $L_{\mathbb{R}^n}^2(\mu)$  and*

$$f = \Delta_{\mathbb{R}^n;k}^\mu f + \sum_{Q \in \mathcal{D}} \Delta_{Q;k}^\mu f, \quad f \in L_{\mathbb{R}^n}^2(\mu),$$

$$\left\langle \Delta_{\mathbb{R}^n;k}^\mu f, \Delta_{Q;k}^\mu f \right\rangle = \left\langle \Delta_{P;k}^\mu f, \Delta_{Q;k}^\mu f \right\rangle = 0 \text{ for } P \neq Q,$$

where convergence in the first line holds both in  $L_{\mathbb{R}^n}^2(\mu)$  norm and pointwise  $\mu$ -almost everywhere.

(2) *Moreover we have the telescoping identities*

$$\mathbf{1}_Q \sum_{Q \subsetneq I \subset P} \Delta_{I;k}^\mu = \mathbb{E}_{Q;k}^\mu - \mathbb{E}_{P;k}^\mu \text{ for } P, Q \in \mathcal{D} \text{ with } P \subsetneq Q, \tag{1.5}$$

(3) *and the moment conditions*

$$\int_{\mathbb{R}^n} \Delta_{Q;k}^\mu f(x) x^\alpha d\mu(x) = 0, \text{ for } Q \in \mathcal{D}, 0 \leq |\alpha| \leq k - 1. \tag{1.6}$$

In the special case of dimension  $n = 1$ , we further investigate uniqueness and degeneracy of the wavelets constructed in Theorem 1. The system of one-dimensional Alpert wavelets is underdetermined, in general having  $\binom{k}{2}$  additional degrees of freedom which can be used to impose additional moment conditions in (3). The degeneracy condition for Alpert wavelets is phrased in terms of a matrix of moments  $\mathbf{M}_{Q,k}$  and positive semi-definiteness, and can be interpreted as the degree to which  $\mu$  is a finite sum of point mass measures within a given child of a cube. Here is our second main result, which includes the main points of our investigation into uniqueness and degeneracy, but not all of them—see Sect. 3.4 for more. Let

$$\mathbf{M}_{Q,k} = \int_Q \begin{bmatrix} 1 & x & \dots & x^{k-2} & x^{k-1} \\ x & x^2 & \ddots & x^{k-3} & x^{k-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x^{k-2} & x^{k-3} & \ddots & x^{2k-4} & x^{2k-3} \\ x^{k-1} & x^{k-2} & \dots & x^{2k-3} & x^{2k-2} \end{bmatrix} d\mu(x) \tag{1.7}$$

be the *non-negative* symmetric matrix of moments of the measure  $\mu$  up to order  $k - 1$  on the interval  $Q$ .

**Theorem 2** *Let  $\mu$  be a locally finite positive Borel measure on  $\mathbb{R}$ , fix  $k \in \mathbb{N}$ , and fix a dyadic grid  $\mathcal{D}$  in  $\mathbb{R}$ . Then, in addition to parts (1) and (2) of Theorem 1 (restricted to  $n = 1$ ), we also have:*

(1) *The dimension of  $L^2_{Q;k}(\mu)$  is given by*

$$\dim L^2_{Q;k}(\mu) = \dim \left( \text{Range } \mathbf{L}_{Q,k} \cap \text{Range } \mathbf{R}_{Q,k} \right),$$

where  $\mathbf{L}_{Q,k}$  and  $\mathbf{R}_{Q,k}$  denote the matrices  $\mathbf{M}_{Q_{\text{left}},k}$  and  $\mathbf{M}_{Q_{\text{right}},k}$  respectively and  $Q_{\text{left}}/Q_{\text{right}}$  are the left and right halves of the interval  $Q$ . This shows in particular that  $L^2_{Q;k}(\mu) = k$  for all dyadic intervals  $Q$  if and only if  $\mathbf{M}_{Q,k} > 0$  for all dyadic intervals  $Q$ .

(2) *In the case when  $\mathbf{M}_{Q,k} > 0$  for all dyadic intervals  $Q$ , we can choose an orthonormal basis  $\{a^{\mu,\ell}_Q\}_{\ell=1}^k$  of  $L^2_{Q;k}(\mu)$  so that in addition to the moment conditions given in part (3) of Theorem 1, the following additional moment conditions hold:*

$$\int a^{\mu,\ell}_Q(x) x^i d\mu(x) = 0, \quad \text{for all } 2 \leq \ell \leq k \text{ and } k \leq i \leq k + \ell - 2.$$

The additional moment conditions in part (2) consume the remaining  $\binom{k}{2}$  degrees of freedom available in defining the  $2k^2$  coefficients in the functions  $\{a^{\mu,\ell}_Q\}_{\ell=1}^k$ , which we refer to as Alpert functions.

**Remark 3** There is an analogous theorem in higher dimensions  $n > 1$ , whose formulation and proof we leave for the interested reader.

When  $k = 1$ , these theorems reduce to the well-known weighted Haar basis in  $L^2(\mu)$  which is recalled in detail in Sect. 2. However, when  $k \geq 2$ , new wavelet bases are provided by this construction, and in Sect. 3, we prove our two main results regarding these new bases, Theorems 1 and 2. In the final section of this paper, we use the Alpert basis to study weighted inequalities for Calderón–Zygmund operators on the real line. A natural proof strategy is to decompose  $f \in L^2(\sigma)$  and  $g \in L^2(\omega)$  via a wavelet basis and then analyze the behavior of the matrix of  $T$  relative to this basis. One then arrives at testing conditions related to the number of moments that the basis possess, as well as a relaxation of the energy condition required to control certain terms in the matrix. With this new basis we study a  $Tp$  type result in the two weight setting. See Conjecture 17. In addition, these wavelets were first used to study integral equations numerically on unweighted spaces. Perhaps similar methods can be used to study integral equations or Toeplitz-like operators on weighted spaces. In addition, papers like [4] study the problem of reconstructing a function with knowledge of its “local moments” (i.e. unweighted averages on dyadic intervals of a fixed scale). A key tool in papers like [4] is the Alpert wavelet basis. So, it seems possible that similar reconstructions can be done using weighted local moments and these weighted Alpert bases.

## 2 Weighted Haar Bases

In this section we review the existence, uniqueness and degeneracy of the weighted Haar wavelets, beginning with the local case. Let  $\mu$  be a locally finite positive Borel measure on the unit interval  $[0, 1)$ .

**Definition 4** Set  $I_0^0 \equiv [0, 1)$ ,  $I_0^1 \equiv [0, \frac{1}{2})$ ,  $I_1^1 \equiv [\frac{1}{2}, 1)$ , and in general

$$I_j^m \equiv \left[ \frac{j}{2^m}, \frac{j+1}{2^m} \right), \quad 0 \leq j \leq 2^m - 1, \quad m \in \mathbb{Z}_+.$$

Note that the left half of the interval  $I_j^m$  is  $I_{j,\text{left}}^m = I_{2j}^{m+1}$  and that the right half of the interval  $I_j^m$  is  $I_{j,\text{right}}^m = I_{2j+1}^{m+1}$ . We begin by briefly reviewing the weighted Haar wavelet bases on the real line and in Euclidean space, and in the next section we will turn to weighted Alpert wavelet bases with more vanishing moments.

Suppose that  $\mu$  is a finite positive Borel measure on  $[0, 1)$  and define

$$L^2(\mu) \equiv \left\{ f : [0, 1) \rightarrow \mathbb{C} \text{ } \mu\text{-measurable} : \int |f|^2 d\mu < \infty \right\}.$$

Let  $|E|_\mu$  denote the  $\mu$ -measure of a Borel set  $E$ . Assume first the nondegeneracy condition that  $\mu$  charges every dyadic subinterval of  $[0, 1)$ ,

$$|J|_\mu > 0 \text{ for all } J = I_j^m, m \in \mathbb{Z}_+, 0 \leq j \leq 2^m - 1. \tag{2.1}$$

**Definition 5** For all  $m \in \mathbb{Z}_+$  and  $0 \leq j \leq 2^m - 1$ , define

$$h_{I_j^m}^\mu(x) \equiv -\sqrt{\frac{1}{|I_j^m|_\mu}} \sqrt{\frac{|I_{j,\text{right}}^m|_\mu}{|I_{j,\text{left}}^m|_\mu}} \mathbf{1}_{I_{j,\text{left}}^m} + \sqrt{\frac{1}{|I_j^m|_\mu}} \sqrt{\frac{|I_{j,\text{left}}^m|_\mu}{|I_{j,\text{right}}^m|_\mu}} \mathbf{1}_{I_{j,\text{right}}^m}. \tag{2.2}$$

Finally define

$$h_0^\mu(x) \equiv \sqrt{\frac{1}{|[0, 1)|_\mu}} \mathbf{1}_{[0, 1)}(x).$$

**Theorem 6** *The collection*

$$\mathcal{U}_{[0, 1)}^{\text{Haar}, \mu} \equiv \{h_0^\mu\} \cup \left\{ h_{I_j^m}^\mu \right\}_{m \in \mathbb{Z}_+, 0 \leq j \leq 2^m - 1}$$

is an orthonormal basis for  $L^2(\mu)$ .

**Proof** It is a straightforward computation to see that  $\mathcal{U}_{[0, 1)}^{\text{Haar}, \mu}$  is an orthonormal set in  $L^2(\mu)$ , and the dyadic Lebesgue differentiation theorem with respect to the measure  $\mu$ , together with the multi-resolution telescoping identities, show that  $\mathcal{U}_{[0, 1)}^{\text{Haar}, \mu}$  is complete in  $L^2(\mu)$ . □

### 2.1 Derivation of the Explicit Formula (2.2)

The coefficients on the functions  $\mathbf{1}_{I_{j,\text{left}}^m}$  and  $\mathbf{1}_{I_{j,\text{right}}^m}$  are derived in the following way. If we set

$$h_{I_j^m}(x) \equiv -\alpha \mathbf{1}_{I_{j,\text{left}}^m} + \beta \mathbf{1}_{I_{j,\text{right}}^m},$$

and demand that both

$$\begin{aligned} 0 &= \int_{I_j^m} h_{I_j^m} d\mu = -\alpha |I_{j,\text{left}}^m|_\mu + \beta |I_{j,\text{right}}^m|_\mu, \\ 1 &= \int_{I_j^m} |h_{I_j^m}|^2 d\mu = \alpha^2 |I_{j,\text{left}}^m|_\mu + \beta^2 |I_{j,\text{right}}^m|_\mu, \end{aligned}$$

then we must have

$$\beta = \frac{|I_{j,\text{left}}^m|_\mu}{|I_{j,\text{right}}^m|_\mu} \alpha \text{ and } 1 = \alpha^2 |I_{j,\text{left}}^m|_\mu + \left( \frac{|I_{j,\text{left}}^m|_\mu}{|I_{j,\text{right}}^m|_\mu} \right)^2 \alpha^2 |I_{j,\text{right}}^m|_\mu = \alpha^2 |I_j^m|_\mu \frac{|I_{j,\text{left}}^m|_\mu}{|I_{j,\text{right}}^m|_\mu},$$

which implies

$$\alpha = \sqrt{\frac{|I_{j,\text{right}}^m|_\mu}{|I_j^m|_\mu |I_{j,\text{left}}^m|_\mu}} \text{ and } \beta = \sqrt{\frac{|I_{j,\text{left}}^m|_\mu}{|I_j^m|_\mu |I_{j,\text{right}}^m|_\mu}},$$

where we see that  $\alpha$  and  $\beta$  are uniquely determined up to sign.

### 2.2 The Haar Degenerate Case

Here we examine what happens when the nondegeneracy condition (2.1) fails.

**Lemma 7** *If at least one of the two children of  $I_j^m$  fails to be charged by  $\mu$ , then if  $h_{I_j^m}$  is defined according to the derivation outlined above, we have that  $h_{I_j^m} \equiv 0$ .*

**Proof** Fix  $I_j^m$  with  $m \in \mathbb{Z}_+$  and  $0 \leq j \leq 2^m - 1$ . If  $|I_j^m|_\mu = 0$ , then clearly  $h_{I_j^m} = 0$  in  $L^2(\mu)$ . If just one of the children of  $I_j^m$  is not charged by  $\mu$ , say  $|I_{j,\text{left}}^m|_\mu = 0$  but  $|I_{j,\text{right}}^m|_\mu > 0$ , then the moment requirement above becomes  $0 = -\alpha |I_{j,\text{left}}^m|_\mu + \beta |I_{j,\text{right}}^m|_\mu = \beta |I_{j,\text{right}}^m|_\mu$ , which implies  $\beta = 0$ , and hence  $h_{I_j^m} = -\alpha \mathbf{1}_{I_{j,\text{left}}^m} + \beta \mathbf{1}_{I_{j,\text{right}}^m} = -\alpha \mathbf{1}_{I_{j,\text{left}}^m} \equiv 0$  in  $L^2(\mu)$ . Thus we see that  $h_{I_j^m} \equiv 0$  if at least one of its children is not charged by  $\mu$ .  $\square$

The resulting pared collection  $\mathcal{U}_{(0,1)}^{\text{Haar},\mu} \equiv \{h_0^\mu\} \cup \left\{ h_{I_j^m}^\mu \right\}_{m \in \mathbb{Z}_+ \text{ and } 0 \leq j \leq 2^m - 1}$ , where now  $h_{I_j^m}^\mu$  is removed if it vanishes identically, is still an orthonormal basis for  $L^2(\mu)$ . Indeed, this follows from the fact that the telescoping identities still hold, and hence Lebesgue’s dyadic differentiation theorem continues to show the pared collection  $\mathcal{U}_{(0,1)}^{\text{Haar},\mu}$  is complete in  $L^2(\mu)$ .

### 2.3 The Global Case

If  $\mu$  is a locally finite positive Borel measure on the real line  $\mathbb{R}$ ,  $\mathcal{D}$  is a dyadic grid in  $\mathbb{R}$ , and

$$h_I^\mu(x) \equiv \begin{cases} -\sqrt{\frac{1}{|I|_\mu}} \sqrt{\frac{|I_{\text{right}}|_\mu}{|I_{\text{left}}|_\mu}} \mathbf{1}_{I_{\text{left}}} + \sqrt{\frac{1}{|I|_\mu}} \sqrt{\frac{|I_{\text{left}}|_\mu}{|I_{\text{right}}|_\mu}} \mathbf{1}_{I_{\text{right}}} & \text{if } \min\{|I_{\text{right}}|_\mu, |I_{\text{left}}|_\mu\} > 0 \\ 0 & \text{if } \min\{|I_{\text{right}}|_\mu, |I_{\text{left}}|_\mu\} = 0 \end{cases}$$

for each dyadic interval  $I \in \mathcal{D}$ , then the collection

$$\mathcal{U}_{\mathbb{R}}^{\text{Haar}, \mu} \equiv \{h_I^\mu\}_{I \in \mathcal{D}}$$

is an orthonormal basis for  $L^2(\mu)$  (where we of course discard those  $h_I^\mu$  that vanish identically).

### 2.4 The Higher Dimensional Case

Again the local case  $L^2([0, 1]^n; \mu)$  and the global case  $L^2(\mathbb{R}^n; \mu)$  are treated similarly and we only consider the global case  $L^2(\mu) = L^2(\mathbb{R}^n; \mu)$  here. So suppose that  $\mu$  is a locally finite positive Borel measure on  $\mathbb{R}^n$  and that  $\mathcal{D}$  is a dyadic grid on  $\mathbb{R}^n$ . Given a dyadic cube  $Q \in \mathcal{D}$  with  $|Q|_\mu > 0$ , let  $\Delta_Q^\mu$  denote orthogonal projection onto the finite dimensional subspace  $L^2_{Q;1}(\mu)$  of  $L^2(\mu)$  that consists of linear combinations of the indicators of the children  $\mathfrak{C}(Q)$  of  $Q$  that have  $\mu$ -mean zero over  $Q$ :

$$\begin{aligned} L^2_{Q;1}(\mu) &\equiv \left\{ f = \sum_{Q' \in \mathfrak{C}(Q)} a_{Q'} \mathbf{1}_{Q'} : a_{Q'} \in \mathbb{R}, \int_Q f d\mu = 0 \right\} \\ &= \text{Span}\{\mathbf{1}_{Q'} : Q' \in \mathfrak{C}(Q), |Q'|_\mu > 0\} \ominus \text{Span}\{\mathbf{1}_Q\}. \end{aligned}$$

Thus, among other things, we see that  $\dim L^2_{Q;1} = \#\{Q' \in \mathfrak{C}(Q) : |Q'|_\mu > 0\} - 1$ . If  $|Q|_\mu = 0$ , set  $\Delta_Q^\mu = 0$ . Then we have the important telescoping property for dyadic cubes  $Q_1 \subset Q_2$  (below  $[Q_1, Q_2] = \{Q : Q_1 \subset Q \subset Q_2\}$ ):

$$\mathbf{1}_{Q_0}(x) \left( \sum_{Q \in [Q_1, Q_2]} \Delta_Q^\mu f(x) \right) = \mathbf{1}_{Q_0}(x) \left( \mathbb{E}_{Q_0}^\mu f - \mathbb{E}_{Q_2}^\mu f \right),$$

$$Q_0 \in \mathfrak{C}(Q_1), f \in L^2(\mu). \tag{2.3}$$

We will at times find it convenient to use a fixed orthonormal basis  $\{h_Q^{\mu,a}\}_{a \in \Gamma_n}$  of  $L^2_Q(\mu)$  where  $\Gamma_n$  is any convenient index set with cardinality equal to the dimension of  $L^2_Q(\mu)$ , i.e.



$$\#\Gamma_n = \dim L^2_Q(\mu) = \#\left\{Q' \in \mathfrak{C}(Q) : |Q'|_\mu > 0\right\} - 1,$$

where the second equality here follows from the fact that the functions in  $L^2_Q(\mu)$  have vanishing mean. Then  $\left\{h_Q^{\mu,a}\right\}_{a \in \Gamma_n \text{ and } Q \in \mathcal{D}}$  is an orthonormal basis for  $L^2(\mu)$ , with the understanding that we add the constant function  $\mathbf{1}$  if  $\mu$  is a finite measure. In particular, if  $\mu$  is an infinite measure, we have

$$f(x) = \sum_{Q \in \mathcal{D}} \Delta_Q^\mu f(x), \quad \mu - a.e. x \in \mathbb{R}^n, \tag{2.4}$$

$$\|f\|_{L^2(\mu)}^2 = \sum_{Q \in \mathcal{D}} \left\| \Delta_Q^\mu f \right\|_{L^2(\mu)}^2 = \sum_{Q \in \mathcal{D}} \sum_{a \in \Gamma_n} |\widehat{f}(Q)|^2, \tag{2.5}$$

where

$$|\widehat{f}(Q)|^2 \equiv \sum_{a \in \Gamma_n} \left| \left\langle f, h_Q^{\mu,a} \right\rangle_\mu \right|^2,$$

and the measure is suppressed in the notation. Indeed, this follows from (2.3) and Lebesgue’s differentiation theorem for cubes.

### 3 Weighted Alpert Wavelets with Higher Vanishing Moments

Let  $\mu$  be a locally finite positive Borel measure on  $\mathbb{R}^n$ , and fix  $k \in \mathbb{N}$ . In analogy with the definition of  $L^2_{Q;1}(\mu)$  above for  $Q \in \mathcal{P}^n$ , we denote by  $L^2_{Q;k}(\mu)$  the finite dimensional subspace of  $L^2(\mu)$  that consists of linear combinations of the indicators of the children  $\mathfrak{C}(Q)$  of  $Q$  multiplied by polynomials of degree at most  $k - 1$ , and such that the linear combinations have vanishing  $\mu$ -moments on the cube  $Q$  up to order  $k - 1$ :

$$L^2_{Q;k}(\mu) \equiv \left\{ f = \sum_{Q' \in \mathfrak{C}(Q)} \mathbf{1}_{Q'} p_{Q';k}(x) : \int_Q f(x) x_i^\ell d\mu(x) = 0, \right. \\ \left. \text{for } 0 \leq \ell \leq k - 1 \text{ and } 1 \leq i \leq n \right\},$$

where  $p_{Q';k}(x) = \sum_{\alpha \in \mathbb{Z}_+^n; |\alpha| \leq k-1} a_{Q';\alpha} x^\alpha$  is a polynomial in  $\mathbb{R}^n$  of degree  $|\alpha| = \alpha_1 + \dots + \alpha_n$  at most  $k - 1$ . Here  $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$ . Let  $d_{Q;k} \equiv \dim L^2_{Q;k}(\mu)$  be the dimension of the finite dimensional linear space  $L^2_{Q;k}(\mu)$ . Note that the space

$L^2_{Q;k}(\mu)$  can also be written as:

$$\text{Span}\{\mathbf{1}_{Q'}x^\alpha : Q' \in \mathfrak{C}(Q), |Q'|_\mu > 0, |\alpha| \leq k - 1\} \ominus (\oplus_{\alpha:|\alpha|\leq k-1} \mathbf{1}_Q x^\alpha). \tag{3.1}$$

We begin with the proof of Theorem 1 in the next subsection below, and we will complete the proof of Theorem 2 in the third subsection. In the final subsection we will give a complete and detailed answer to both uniqueness and degeneracy in the special case when  $n = 1$  and  $k = 2$ .

### 3.1 Proof of Theorem 1

We begin with an estimate of the dimension  $d_{Q;k}$ .

**Proposition 8** *We have:*

$$0 \leq d_{Q;k} \leq (2^n - 1) A_{k,n} = (2^n - 1) \binom{n+k-1}{n}, \tag{3.2}$$

where  $A_{k,n}$  denotes the number of non-negative integer solutions to  $\alpha_1 + \dots + \alpha_n \leq k - 1$ .

**Proof** First, if the functions  $\{\mathbf{1}_Q(x)x^\alpha : |\alpha| \leq k - 1\}$  are linearly independent, then

$$\dim \text{Span}\{\mathbf{1}_Q(x)x^\alpha : |\alpha| \leq k - 1\} = A_{k,n}.$$

Let  $B$  denote  $\dim \text{Span}\{\mathbf{1}_Q(x)x^\alpha : |\alpha| \leq k - 1\}$ . Then it follows that

$$\dim \text{Span}\{\mathbf{1}_{Q'}(x)x^\alpha : |\alpha| \leq k - 1\} \leq B, \quad \text{for all } Q' \in \mathfrak{C}(Q).$$

Thus

$$\dim L^2_{Q;k}(\mu) \leq \#\{Q' \in \mathfrak{C}Q : |Q'|_\mu > 0\}B - B \leq (2^n - 1)B \leq (2^n - 1)A_{k,n}.$$

Finally it is well-known that the number of nonnegative solutions to  $\alpha_1 + \dots + \alpha_n = j$  is  $\binom{n-1+j}{j}$  (simply choose  $n - 1$  boxes from a row of  $n + j - 1$  boxes and put a single ball in each of the unchosen boxes—then let  $\alpha_i$  be the number of balls between the  $(i - 1)$ th box and the  $i$ th box), and so  $A_{k,n} = \sum_{j=0}^{k-1} \binom{n-1+j}{j} = \binom{n-1+k}{k-1}$  by induction on  $k$ . □

We restrict attention to a fixed dyadic grid  $\mathcal{D}$  in  $\mathbb{R}^n$ . For  $P, Q \in \mathcal{D}$  dyadic cubes, the subspaces  $L^2_{P;k}(\mu)$  and  $L^2_{Q;k}(\mu)$  are orthogonal for  $P \neq Q$ , i.e.  $\langle f, g \rangle_{L^2(\mu)} = 0$  for  $f \in L^2_{P;k}(\mu)$  and  $g \in L^2_{Q;k}(\mu)$ . Indeed, the only case that needs checking is when either  $P \subsetneq Q$  or  $Q \subsetneq P$ . If  $P \subsetneq Q$ , then the restriction of  $g$  to  $P$  is a polynomial of degree less than or equal to  $k - 1$ . But by definition,  $L^2_{P;k}(\mu)$  is orthogonal to

such functions and so  $\langle f, g \rangle_{L^2(\mu)} = 0$ . Of course similar reasoning holds in the case  $Q \subsetneq P$ .

Next define

$$\mathcal{F}_\infty^k(\mu) \equiv \left\{ \alpha \in \mathbb{Z}_+^n : |\alpha| \leq k - 1 : x^\alpha \in L^2(\mu) \right\},$$

$$\text{and } \mathcal{P}_{\mathbb{R}^n}^k(\mu) \equiv \text{Span} \left\{ x^\alpha \right\}_{\alpha \in \mathcal{F}_\infty^k}.$$

We claim that  $L^2(\mu)$  is the Hilbert space direct sum of  $\mathcal{P}_{\mathbb{R}^n}^k(\mu)$  in  $L^2(\mu)$  and the finite dimensional subspaces  $\left\{ L_{Q;k}^2(\mu) \right\}_{Q \in \mathcal{D}}$ , namely

$$L^2(\mu) = \mathcal{P}_{\mathbb{R}^n}^k(\mu) \oplus \left( \bigoplus_{Q \in \mathcal{D}} L_{Q;k}^2(\mu) \right).$$

To see this, fix a large dyadic cube  $P \in \mathcal{D}$ . Set

$$\mathcal{P}_P^k \equiv \left\{ f = \mathbf{1}_P(x) p_k(x) : \alpha \in \mathbb{Z}_+^n : |\alpha| \leq k - 1 \right\}$$

to be the linear space of restrictions to  $P$  of polynomials  $p_k(x)$  having degree at most  $k - 1$ . Then for  $Q \subset P$ , we have

$$L_{Q;1}^2(\mu) = \left\{ f = \sum_{Q' \in \mathcal{C}(Q)} a_{Q'} \mathbf{1}_{Q'} : a_{Q'} \in \mathbb{R}, \int_Q f d\mu = 0 \right\}$$

$$\subset \overline{\text{Span}} \left\{ \mathcal{P}_P^k, \left\{ L_{R;k}^2(\mu) \right\}_{R \in \mathcal{D}: Q \subset R \subset P} \right\}.$$

Now let  $P$  tend to infinity to conclude that

$$L_{Q;1}^2(\mu) \subset \overline{\text{Span}} \left\{ \mathcal{P}_{\mathbb{R}^n}^k(\mu), \left\{ L_{R;k}^2(\mu) \right\}_{R \in \mathcal{D}: Q \subset R} \right\}.$$

But we already know that the Haar spaces  $\mathcal{P}_{\mathbb{R}^n}^1(\mu)$  and  $L_{Q;1}^2(\mu)$  form a direct sum decomposition of  $L^2(\mu)$ , i.e.

$$L^2(\mu) = \mathcal{P}_{\mathbb{R}^n}^1(\mu) \oplus \left( \bigoplus_{Q \in \mathcal{D}} L_{Q;1}^2(\mu) \right),$$

and hence we see that

$$L^2(\mu) = \mathcal{P}_{\mathbb{R}^n}^1(\mu) \oplus \left( \bigoplus_{Q \in \mathcal{D}} L_{Q;1}^2(\mu) \right) \subset \mathcal{P}_{\mathbb{R}^n}^k(\mu) \oplus \left( \bigoplus_{Q \in \mathcal{D}} L_{Q;k}^2(\mu) \right) \subset L^2(\mu).$$

Formula (3.1) gives the telescoping identities (1.5), and the moment conditions (1.6) are immediate from the definition of  $L^2_{Q;k}(\mu)$ . This completes the proof of Theorem 1.

### 3.2 Uniqueness and Degeneracy

In this subsection we begin to investigate the lack of uniqueness of the orthogonal projections  $\left\{ \Delta_{\mathbb{R}^n;k}^\mu \right\} \cup \left\{ \Delta_{Q;k}^\mu \right\}_{Q \in \mathcal{D}^n}$  and their degeneracy when

$$d_{Q;k} < (2^n - 1) \binom{n+k-1}{n}$$

is not maximal. We can of course use the Gram-Schmidt orthogonalization algorithm to find an orthonormal basis  $\{k_{\mathbb{R}^n}^{\mu,a}\}_{a \in \Gamma_n^k(\mathbb{R}^n)}$  of  $\mathcal{P}_{\mathbb{R}^n}^k(\mu)$  where  $\Gamma_n^k(\mathbb{R}^n)$  is any convenient index set with cardinality equal to  $\dim \mathcal{P}_{\mathbb{R}^n}^k(\mu) = \#\mathcal{F}_\infty^k(\mu)$ , and also an orthonormal basis  $\{h_Q^{\mu,a}\}_{a \in \Gamma_n^k}$  of  $L^2_{Q;k}(\mu)$  where  $\Gamma_n^k$  is any convenient index set with cardinality equal to  $d_{Q;k}$ , the dimension of  $L^2_{Q;k}(\mu)$ , i.e.

$$\#\Gamma_n^k(Q) = \dim L^2_{Q;k}(\mu).$$

Then for any  $J \in \mathcal{D} \cup \{\mathbb{R}^n\}$ , the set  $\{k_J^{\mu,a}\}_{a \in \Gamma_n^k(J)} \cup \{h_Q^{\mu,a}\}_{a \in \Gamma_n^k(Q)}$  and  $Q \in \mathcal{D}$  with  $Q \subset J$  is an orthonormal basis for  $L^2(\mathbf{1}_J \mu)$ .

In the case of dimension  $n = 1$  with  $k = 1$  vanishing moment, Proposition 8 shows that the dimension  $d_{Q;1}$  of the subspace  $L^2_{Q;1}(\mu)$  satisfies  $0 \leq d_{Q;1} \leq (2^1 - 1) \binom{1+1-1}{1-1} = 1$ . So if  $L^2_{Q;1}(\mu) \neq \{0\}$ , the orthogonal projection  $\Delta_{Q;1}^\mu$

is one-dimensional, and hence is given by  $\Delta_{Q;1}^\mu f = \frac{\langle f, h_Q^\mu \rangle}{\langle h_Q^\mu, h_Q^\mu \rangle} h_Q^\mu$  for a function  $h_Q^\mu$ ,

which has the especially simple formula given by (2.2). We next investigate to what extent one can find "nice" explicit bases of  $L^2_{Q;k}(\mu)$  in dimension  $n = 1$ , where in

this case  $0 \leq d_{Q;k} \leq (2^1 - 1) \binom{1+k-1}{k-1} = k$ .

### 3.3 Proof of Theorem 2

To prepare for the proof of Theorem 2, we first turn to the explicit construction of weighted Alpert bases in dimension  $n = 1$  when the number of vanishing moments is  $k \geq 2$ .

### 3.3.1 An Explicit Basis for $n = 1$ and $k \geq 2$

Let  $\mu$  be a locally finite positive Borel measure on  $\mathbb{R}$  that satisfies the Alpert nondegeneracy condition (recall that  $|J^{(k)}|_\mu = \int_J x^k d\mu$ )

$$M_{J,k} \equiv \begin{bmatrix} |J^{(0)}|_\mu & |J^{(1)}|_\mu & \dots & |J^{(k-2)}|_\mu & |J^{(k-1)}|_\mu \\ |J^{(1)}|_\mu & |J^{(2)}|_\mu & \ddots & |J^{(k-1)}|_\mu & |J^{(k)}|_\mu \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ |J^{(k-2)}|_\mu & |J^{(k-1)}|_\mu & \ddots & |J^{(2k-4)}|_\mu & |J^{(2k-3)}|_\mu \\ |J^{(k-1)}|_\mu & |J^{(k)}|_\mu & \dots & |J^{(2k-3)}|_\mu & |J^{(2k-2)}|_\mu \end{bmatrix} \succ 0, \tag{3.3}$$

for all  $J \in \mathcal{D}$ , and where we denote a positive definite matrix  $A$  by  $A \succ 0$ . For each  $x \in J$  we have that

$$M(x) \equiv \begin{bmatrix} 1 & x & \dots & x^{k-2} & x^{k-1} \\ x & x^2 & \ddots & x^{k-3} & x^{k-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x^{k-2} & x^{k-3} & \ddots & x^{2k-4} & x^{2k-3} \\ x^{k-1} & x^{k-2} & \dots & x^{2k-3} & x^{2k-2} \end{bmatrix} = \begin{bmatrix} 1 \\ x \\ \vdots \\ x^{k-2} \\ x^{k-1} \end{bmatrix} [1 \ x \ \dots \ x^{k-2} \ x^{k-1}]$$

is a dyad, namely a rank one nonnegative semidefinite matrix. Thus

$$M_{J,k} = \int_J \begin{bmatrix} 1 & x & \dots & x^{k-2} & x^{k-1} \\ x & x^2 & \ddots & x^{k-3} & x^{k-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x^{k-2} & x^{k-3} & \ddots & x^{2k-4} & x^{2k-3} \\ x^{k-1} & x^{k-2} & \dots & x^{2k-3} & x^{2k-2} \end{bmatrix} d\mu(x) = \int_J M(x) d\mu(x)$$

is also a nonnegative semidefinite matrix.

We claim that the matrix  $M_{J,k}$  is positive definite *if and only if* the functions

$$\{ \mathbf{1}_J(x), x\mathbf{1}_J(x), \dots, x^{k-1}\mathbf{1}_J(x) \}$$

are linearly independent in  $L^2(\mu)$ . Indeed, in the special case  $k = 2$  we have

$$\det M_{J,2} = \det \begin{bmatrix} |J^{(0)}|_\mu & |J^{(1)}|_\mu \\ |J^{(1)}|_\mu & |J^{(2)}|_\mu \end{bmatrix} > 0$$

if and only if

$$|J^{(1)}|_\mu^2 = \left( \int_J x d\mu(x) \right)^2 \leq \left( \int_J 1^2 d\mu(x) \right) \left( \int_J x^2 d\mu(x) \right) = |J^{(0)}|_\mu |J^{(2)}|_\mu,$$

with strict inequality if and only if the functions  $\mathbf{1}_J(x)$  and  $x\mathbf{1}_J(x)$  are linearly independent in  $L^2(\mu)$ . Here is the general case.

**Theorem 9** *Let  $\mu$  be a locally finite positive Borel measure on  $\mathbb{R}$ . Then:*

(1) *the  $k \times k$  matrix of moments*

$$M_{J,k} = \int \begin{bmatrix} 1 & x & \dots & x^{k-2} & x^{k-1} \\ x & x^2 & \dots & x^{k-3} & x^{k-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x^{k-2} & x^{k-3} & \dots & x^{2k-4} & x^{2k-3} \\ x^{k-1} & x^{k-2} & \dots & x^{2k-3} & x^{2k-2} \end{bmatrix} d\mu(x)$$

*is nonnegative semidefinite, and*

(2)  *$M_{J,k}$  has rank  $\ell$  if and only if the span of the functions  $1, x, x^2, \dots, x^{k-1}$  has dimension  $\ell$  in  $L^2(\mu)$ , i.e.*

$$\text{rank } M_{J,k} = \dim \text{Span} \{x^j\}_{j=0}^{k-1}.$$

**Proof** With  $V_k(x)^{\text{tr}} \equiv [1 \ x \ \dots \ x^{k-2} \ x^{k-1}]$ , we compute

$$\begin{bmatrix} 1 & x & \dots & x^{k-2} & x^{k-1} \\ x & x^2 & \dots & x^{k-3} & x^{k-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ x^{k-2} & x^{k-3} & \dots & x^{2k-4} & x^{2k-3} \\ x^{k-1} & x^{k-2} & \dots & x^{2k-3} & x^{2k-2} \end{bmatrix} = \begin{bmatrix} 1 \\ x \\ \vdots \\ x^{k-2} \\ x^{k-1} \end{bmatrix} [1 \ x \ \dots \ x^{k-2} \ x^{k-1}] = V_k(x) V_k(x)^{\text{tr}},$$

and so

$$M_{J,k} \equiv \int_{\mathbb{R}^k} V_k(x) V_k(x)^{\text{tr}} d\mu(x)$$

is a positive integral of dyads, hence nonnegative semidefinite. Note that the quadratic form

$$\xi^{\text{tr}} M_{J,k} \xi = \int_{\mathbb{R}^k} (\xi^{\text{tr}} V_k(x)) (V_k(x)^{\text{tr}} \xi) d\mu(x) = \int_{\mathbb{R}^k} (\xi \cdot V_k(x))^2 d\mu(x)$$

is strictly positive if and only if  $\text{Span} \{V_k(x) : x \in \text{supp } \mu\}$  has dimension  $k$ . Indeed,  $\dim \text{Span} \{V_k(x) : x \in \text{supp } \mu\} = k$  if and only if the function  $\xi \cdot V_k(x)$  is not trivial for every  $\xi \in \mathbb{R}^k \setminus \{0\}$ , and this in turn happens if and only if  $\xi^{\text{tr}} \mathbf{M}_{J,k} \xi > 0$  for all  $\xi \in \mathbb{R}^k \setminus \{0\}$ . Similarly we have  $\dim \text{Span} \{V_k(x) : x \in \text{supp } \mu\} = \ell$  if and only if the vector subspace of  $\mathbb{R}^k$  defined by

$$S \equiv \left\{ \xi \in \mathbb{R}^k : \xi \cdot V_k(x) \text{ is not trivial} \right\}$$

has dimension  $\ell$ , and this in turn happens if and only if  $\dim \text{Span} \{x^j\}_{j=0}^{k-1} = \ell$ .  $\square$

**Remark 10** The Vandermonde determinant formula shows that  $\dim \text{Span} \{x^j\}_{j=0}^{k-1} = \ell < k$  if and only if the measure  $\mu$  is a sum of positive multiples of  $\ell$  point masses, i.e.  $\mu = \sum_{j=1}^{\ell} c_j \delta_{x_j}$  with  $c_j > 0$  and the  $x_j$  distinct.

### 3.3.2 An Orthonormal Basis of the Multiresolution Projections

Let  $\mu$  be a locally finite positive Borel measure on  $\mathbb{R}$  that satisfies the Alpert nondegeneracy condition (3.3), i.e. the determinant of all principal  $\ell \times \ell$  submatrices  $\mathbf{M}_{J,\mathbf{i}}$  is positive for  $\mathbf{i} = (i_1, i_2, \dots, i_{\ell})$  and  $1 \leq \ell \leq k$ ,

$$\det \mathbf{M}_{J,\mathbf{i}} > 0, \quad \text{for all } 1 \leq \ell \leq k, \text{ and all } J = I_j^m, m \in \mathbb{Z}_+, 0 \leq j \leq 2^m - 1; \tag{3.4}$$

$$\text{where } \mathbf{M}_{J,\mathbf{i}} \equiv \begin{bmatrix} |J^{(2i_1)}|_{\mu} & |J^{(i_1+i_2)}|_{\mu} & \dots & |J^{(i_1+i_{\ell-1})}|_{\mu} & |J^{(i_1+i_{\ell})}|_{\mu} \\ |J^{(i_2+i_1)}|_{\mu} & |J^{(2i_2)}|_{\mu} & \ddots & |J^{(i_2+i_{\ell-1})}|_{\mu} & |J^{(i_2+i_{\ell})}|_{\mu} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ |J^{(i_{\ell-1}+i_1)}|_{\mu} & |J^{(i_{\ell-1}+i_2)}|_{\mu} & \ddots & |J^{(2i_{\ell-1})}|_{\mu} & |J^{(i_{\ell-1}+i_{\ell})}|_{\mu} \\ |J^{(i_{\ell}+i_1)}|_{\mu} & |J^{(i_{\ell}+i_2)}|_{\mu} & \dots & |J^{(i_{\ell}+i_{\ell-1})}|_{\mu} & |J^{(2i_{\ell})}|_{\mu} \end{bmatrix}. \tag{3.5}$$

If for  $\mathbf{i} = (i_1, i_2, \dots, i_{\ell})$  we set  $V_{\mathbf{i}}(x)^{\text{tr}} \equiv [x^{i_1} \ x^{i_2} \ \dots \ x^{i_{\ell-1}} \ x^{i_{\ell}}]$ , then by the above theorem,

$$\mathbf{M}_{J,\mathbf{i}} \equiv \int_J V_{\mathbf{i}}(x) V_{\mathbf{i}}(x)^{\text{tr}} d\mu(x) \tag{3.6}$$

is nonnegative semidefinite; and is positive definite if and only if the the span of the functions  $x^{i_1}, x^{i_2}, \dots, x^{i_{\ell}}$  has dimension  $\ell$  in  $L^2(\mu)$ , i.e.

$$\text{rank } \mathbf{M}_{J,\mathbf{i}} = \dim \text{Span} \left\{ x^{i_j} \right\}_{j=1}^{\ell}.$$

**Remark 11** Recall that a matrix  $M$  is nonnegative semidefinite if and only if  $\det \mathbf{M}_i \geq 0$  for all principal submatrices  $\mathbf{M}_i$  (the numbers  $\det \mathbf{M}_i$  are usually referred to as principal minors); and that a matrix  $M$  is positive definite if and only if  $\det \mathbf{M}_i > 0$  for the special subset of leading principal submatrices  $\mathbf{M}_i$  having  $\mathbf{i} = (1, 2, \dots, \ell)$  and  $1 \leq \ell \leq k$  (in which case  $\det \mathbf{M}_i > 0$  for all principal submatrices  $\mathbf{M}_i$ ).

Recall that we can use the Gram-Schmidt orthogonalization algorithm to find an orthonormal basis  $\{h_Q^{\mu,a}\}_{a \in \Gamma_1^k}$  of  $L^2_{Q;k}(\mu)$  where  $\Gamma_1^k$  is any convenient index set with cardinality equal to  $d_{Q;k}$ , the dimension of  $L^2_{Q;k}(\mu)$ , i.e.

$$\#\Gamma_1^k = \dim L^2_{Q;k}(\mu).$$

Then  $\{x^\alpha\}_{\alpha \in \mathcal{F}_\infty} \cup \{h_Q^{\mu,a}\}_{a \in \Gamma_n^k \text{ and } Q \in \mathcal{D}}$  is an orthonormal basis for  $L^2_{\mathbb{R}^n}(\mu)$

In the case of dimension  $n = 1$  with  $k \geq 2$  vanishing moments, the above proposition shows that the dimension  $d_{Q;k}$  of the subspace  $L^2_{Q;k}(\mu)$  satisfies  $0 \leq d_{Q;k} \leq (2^1 - 1) \binom{1+k-1}{k-1} = k$ , and in the case where  $\mathbf{M}_{J,k}$  is positive definite, we have  $d_{Q;k} = k$ . Next, we begin an explicit construction of an ‘Alpert’ orthonormal basis.

We require the Alpert functions

$$a_I^{\mu,\ell}(x) \equiv \left( \alpha_{k-1}^\ell x^{k-1} + \alpha_{k-2}^\ell x^{k-2} + \dots + \alpha_1^\ell x + \alpha_0^\ell \right) \mathbf{1}_{I_{\text{left}}}(x) + \left( \beta_{k-1}^\ell x^{k-1} + \beta_{k-2}^\ell x^{k-2} + \dots + \beta_1^\ell x + \beta_0^\ell \right) \mathbf{1}_{I_{\text{right}}}(x),$$

namely

$$\begin{aligned} a_I^{\mu,1}(x) &\equiv \left( \alpha_{k-1}^1 x^{k-1} + \alpha_{k-2}^1 x^{k-2} + \dots + \alpha_1^1 x + \alpha_0^1 \right) \mathbf{1}_{I_{\text{left}}}(x) \\ &\quad + \left( \beta_{k-1}^1 x^{k-1} + \beta_{k-2}^1 x^{k-2} + \dots + \beta_1^1 x + \beta_0^1 \right) \mathbf{1}_{I_{\text{right}}}(x), \\ a_I^{\mu,2}(x) &\equiv \left( \alpha_{k-1}^2 x^{k-1} + \alpha_{k-2}^2 x^{k-2} + \dots + \alpha_1^2 x + \alpha_0^2 \right) \mathbf{1}_{I_{\text{left}}}(x) \\ &\quad + \left( \beta_{k-1}^2 x^{k-1} + \beta_{k-2}^2 x^{k-2} + \dots + \beta_1^2 x + \beta_0^2 \right) \mathbf{1}_{I_{\text{right}}}(x), \\ &\quad \vdots \\ a_I^{\mu,k}(x) &\equiv \left( \alpha_{k-1}^k x^{k-1} + \alpha_{k-2}^k x^{k-2} + \dots + \alpha_1^k x + \alpha_0^k \right) \mathbf{1}_{I_{\text{left}}}(x) \\ &\quad + \left( \beta_{k-1}^k x^{k-1} + \beta_{k-2}^k x^{k-2} + \dots + \beta_1^k x + \beta_0^k \right) \mathbf{1}_{I_{\text{right}}}(x), \end{aligned}$$

to satisfy the  $k^2$  moment properties,

$$\int a_I^{\mu,\ell}(x) x^i d\mu(x) = 0, \quad \text{for all } 1 \leq \ell \leq k \text{ and } 0 \leq i \leq k + \ell - 2,$$



the  $\binom{k}{2}$  orthogonality properties,

$$\int a_I^{\mu,\ell}(x) a_I^{\mu,\ell'}(x) d\mu(x) = 0, \quad \text{for all } 1 \leq \ell, \ell' \leq k,$$

and the  $k$  normalization properties,

$$\left( \int |a_I^{\mu,\ell}(x)|^2 d\mu(x) \right)^{\frac{1}{2}} = 1, \quad \text{for all } 1 \leq \ell \leq k.$$

**Remark 12** (Dimension Count) Note that  $k^2 + \binom{k}{2} + k = 2k^2 - \binom{k}{2}$ , so that there are  $\binom{k}{2}$  degrees of freedom remaining in the choice of the  $2k^2$  coefficients  $\{\alpha_i^\ell, \beta_i^\ell\}_{1 \leq \ell \leq k \text{ and } 0 \leq i \leq k-1}$ . We will later show that we can impose  $\binom{k}{2}$  additional moment conditions.

So with the definition  $|J^{(i)}|_\mu \equiv \int_J x^i d\mu(x)$ , we first tackle the moment properties:

$$\begin{aligned} 0 &= \int a_I^{\mu,\ell}(x) x^i d\mu(x) \\ &= \int_{I_{\text{left}}} \left( \alpha_{k-1}^\ell x^{k-1} + \alpha_{k-2}^\ell x^{k-2} + \dots + \alpha_1^\ell x + \alpha_0^\ell \right) x^i d\mu(x) \end{aligned} \tag{3.7}$$

$$+ \int_{I_{\text{right}}} \left( \beta_{k-1}^\ell x^{k-1} + \beta_{k-2}^\ell x^{k-2} + \dots + \beta_1^\ell x + \beta_0^\ell \right) x^i d\mu(x) \tag{3.8}$$

$$= \alpha_{k-1}^\ell \left| I_{\text{left}}^{(i+k-1)} \right|_\mu + \alpha_{k-2}^\ell \left| I_{\text{left}}^{(i+k-2)} \right|_\mu + \dots + \alpha_0^\ell \left| I_{\text{left}}^{(i)} \right|_\mu \tag{3.9}$$

$$+ \beta_{k-1}^\ell \left| I_{\text{right}}^{(i+k-1)} \right|_\mu + \beta_{k-2}^\ell \left| I_{\text{right}}^{(i+k-2)} \right|_\mu + \dots + \beta_0^\ell \left| I_{\text{right}}^{(i)} \right|_\mu, \tag{3.10}$$

which lead to

$$\begin{aligned} &\left[ \left| I_{\text{left}}^{(i)} \right|_\mu \dots \left| I_{\text{left}}^{(i+k-1)} \right|_\mu \right] \begin{pmatrix} \alpha_0^\ell \\ \vdots \\ \alpha_{k-1}^\ell \end{pmatrix} + \left[ \left| I_{\text{right}}^{(i)} \right|_\mu \dots \left| I_{\text{right}}^{(i+k-1)} \right|_\mu \right] \begin{pmatrix} \beta_0^\ell \\ \vdots \\ \beta_{k-1}^\ell \end{pmatrix} = 0, \\ &0 \leq i \leq k-1, \end{aligned} \tag{3.11}$$

which in matrix form is,

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{bmatrix} \left| I_{\text{left}}^{(0)} \right|_{\mu} & \left| I_{\text{left}}^{(1)} \right|_{\mu} & \cdots & \left| I_{\text{left}}^{(k-2)} \right|_{\mu} & \left| I_{\text{left}}^{(k-1)} \right|_{\mu} \\ \left| I_{\text{left}}^{(1)} \right|_{\mu} & \left| I_{\text{left}}^{(2)} \right|_{\mu} & \cdots & \left| I_{\text{left}}^{(k-1)} \right|_{\mu} & \left| I_{\text{left}}^{(k)} \right|_{\mu} \\ \vdots & \vdots & & \vdots & \vdots \\ \left| I_{\text{left}}^{(k-2)} \right|_{\mu} & \left| I_{\text{left}}^{(k-1)} \right|_{\mu} & \cdots & \left| I_{\text{left}}^{(2k-4)} \right|_{\mu} & \left| I_{\text{left}}^{(2k-3)} \right|_{\mu} \\ \left| I_{\text{left}}^{(k-1)} \right|_{\mu} & \left| I_{\text{left}}^{(k)} \right|_{\mu} & \cdots & \left| I_{\text{left}}^{(2k-3)} \right|_{\mu} & \left| I_{\text{left}}^{(2k-2)} \right|_{\mu} \end{bmatrix} \begin{pmatrix} \alpha_0^{\ell} \\ \alpha_1^{\ell} \\ \vdots \\ \alpha_{k-2}^{\ell} \\ \alpha_{k-1}^{\ell} \end{pmatrix} \tag{3.12}$$

$$+ \begin{bmatrix} \left| I_{\text{right}}^{(0)} \right|_{\mu} & \left| I_{\text{right}}^{(1)} \right|_{\mu} & \cdots & \left| I_{\text{right}}^{(k-2)} \right|_{\mu} & \left| I_{\text{right}}^{(k-1)} \right|_{\mu} \\ \left| I_{\text{right}}^{(1)} \right|_{\mu} & \left| I_{\text{right}}^{(2)} \right|_{\mu} & \cdots & \left| I_{\text{right}}^{(k-1)} \right|_{\mu} & \left| I_{\text{right}}^{(k)} \right|_{\mu} \\ \vdots & \vdots & & \vdots & \vdots \\ \left| I_{\text{right}}^{(k-2)} \right|_{\mu} & \left| I_{\text{right}}^{(k-1)} \right|_{\mu} & \cdots & \left| I_{\text{right}}^{(2k-4)} \right|_{\mu} & \left| I_{\text{right}}^{(2k-3)} \right|_{\mu} \\ \left| I_{\text{right}}^{(k-1)} \right|_{\mu} & \left| I_{\text{right}}^{(k)} \right|_{\mu} & \cdots & \left| I_{\text{right}}^{(2k-3)} \right|_{\mu} & \left| I_{\text{right}}^{(2k-2)} \right|_{\mu} \end{bmatrix} \begin{pmatrix} \beta_0^{\ell} \\ \beta_1^{\ell} \\ \vdots \\ \beta_{k-2}^{\ell} \\ \beta_{k-1}^{\ell} \end{pmatrix} \tag{3.13}$$

But this says that

$$\begin{bmatrix} \left| I_{\text{left}}^{(0)} \right|_{\mu} & \cdots & \left| I_{\text{left}}^{(k-1)} \right|_{\mu} \\ \vdots & & \vdots \\ \left| I_{\text{left}}^{(k-1)} \right|_{\mu} & \cdots & \left| I_{\text{left}}^{(2k-2)} \right|_{\mu} \end{bmatrix} \begin{pmatrix} \alpha_0^{\ell} \\ \vdots \\ \alpha_{k-1}^{\ell} \end{pmatrix} = - \begin{bmatrix} \left| I_{\text{right}}^{(0)} \right|_{\mu} & \cdots & \left| I_{\text{right}}^{(k-1)} \right|_{\mu} \\ \vdots & & \vdots \\ \left| I_{\text{right}}^{(k-1)} \right|_{\mu} & \cdots & \left| I_{\text{right}}^{(2k-2)} \right|_{\mu} \end{bmatrix} \begin{pmatrix} \beta_0^{\ell} \\ \vdots \\ \beta_{k-1}^{\ell} \end{pmatrix}, \tag{3.14}$$

for  $1 \leq \ell \leq k$ , which we write as

$$L_I^k \alpha_k^{\ell} = -R_I^k \beta_k^{\ell}, \tag{3.15}$$

with

$$L_I^k \equiv \begin{bmatrix} \left| I_{\text{left}}^{(0)} \right|_{\mu} & \left| I_{\text{left}}^{(1)} \right|_{\mu} & \cdots & \left| I_{\text{left}}^{(k-2)} \right|_{\mu} & \left| I_{\text{left}}^{(k-1)} \right|_{\mu} \\ \left| I_{\text{left}}^{(1)} \right|_{\mu} & \left| I_{\text{left}}^{(2)} \right|_{\mu} & \cdots & \left| I_{\text{left}}^{(k-1)} \right|_{\mu} & \left| I_{\text{left}}^{(k)} \right|_{\mu} \\ \vdots & \vdots & & \vdots & \vdots \\ \left| I_{\text{left}}^{(k-2)} \right|_{\mu} & \left| I_{\text{left}}^{(k-1)} \right|_{\mu} & \cdots & \left| I_{\text{left}}^{(2k-4)} \right|_{\mu} & \left| I_{\text{left}}^{(2k-3)} \right|_{\mu} \\ \left| I_{\text{left}}^{(k-1)} \right|_{\mu} & \left| I_{\text{left}}^{(k)} \right|_{\mu} & \cdots & \left| I_{\text{left}}^{(2k-3)} \right|_{\mu} & \left| I_{\text{left}}^{(2k-2)} \right|_{\mu} \end{bmatrix}, \tag{3.16}$$

$$R_I^k \equiv \begin{bmatrix} \left| I_{\text{right}}^{(0)} \right|_{\mu} & \left| I_{\text{right}}^{(1)} \right|_{\mu} & \cdots & \left| I_{\text{right}}^{(k-2)} \right|_{\mu} & \left| I_{\text{right}}^{(k-1)} \right|_{\mu} \\ \left| I_{\text{right}}^{(1)} \right|_{\mu} & \left| I_{\text{right}}^{(2)} \right|_{\mu} & \cdots & \left| I_{\text{right}}^{(k-1)} \right|_{\mu} & \left| I_{\text{right}}^{(k)} \right|_{\mu} \\ \vdots & \vdots & & & \vdots \\ \left| I_{\text{right}}^{(k-2)} \right|_{\mu} & \left| I_{\text{right}}^{(k-1)} \right|_{\mu} & \cdots & \left| I_{\text{right}}^{(2k-4)} \right|_{\mu} & \left| I_{\text{right}}^{(2k-3)} \right|_{\mu} \\ \left| I_{\text{right}}^{(k-1)} \right|_{\mu} & \left| I_{\text{right}}^{(k)} \right|_{\mu} & \cdots & \left| I_{\text{right}}^{(2k-3)} \right|_{\mu} & \left| I_{\text{right}}^{(2k-2)} \right|_{\mu} \end{bmatrix}, \tag{3.17}$$

$$\alpha_k^\ell \equiv \begin{pmatrix} \alpha_0^\ell \\ \vdots \\ \alpha_{k-1}^\ell \end{pmatrix}, \quad \beta_k^\ell = \begin{pmatrix} \beta_0^\ell \\ \vdots \\ \beta_{k-1}^\ell \end{pmatrix} \tag{3.18}$$

and by our determinant assumption we can then solve for  $\alpha_k^\ell$  in terms of  $\beta_k^\ell$ ;

$$\alpha_k^\ell = - \left[ L_I^k \right]^{-1} R_I^k \beta_k^\ell, \quad 1 \leq \ell \leq k. \tag{3.19}$$

Thus using the Gram-Schmidt orthogonalization algorithm, there is an orthonormal basis  $\left\{ a_I^{\mu,\ell}(x) \right\}_{\ell=1}^k$  of  $L^2_{Q;k}(\mu)$  consisting of Alpert functions if and only if at least one of the matrices  $L_I^k$  and  $R_I^k$  is nonsingular (equivalently positive definite, since these symmetric matrices are always nonnegative semidefinite). For example, if we assume that  $L_I^k$  is invertible, then just as in the case  $k = 2$  discussed above, a function  $a_I^{\mu,\ell'}(x)$  is orthogonal to  $a_I^{\mu,\ell}(x)$  in  $L^2(\mu)$ , i.e.

$$\begin{aligned} 0 &= \int a_I^{\mu,\ell}(x) a_I^{\mu,\ell'}(x) d\mu(x) \\ &= \int \left\{ \left( \alpha_{k-1}^\ell x^{k-1} + \alpha_{k-2}^\ell x^{k-2} + \cdots + \alpha_1^\ell x + \alpha_0^\ell \right) \left( \alpha_{k-1}^{\ell'} x^{k-1} \right. \right. \\ &\quad \left. \left. + \alpha_{k-2}^{\ell'} x^{k-2} + \cdots + \alpha_1^{\ell'} x + \alpha_0^{\ell'} \right) \mathbf{1}_{I_{\text{left}}}(x) \right. \\ &\quad \left. + \left( \beta_{k-1}^\ell x^{k-1} + \beta_{k-2}^\ell x^{k-2} + \cdots + \beta_1^\ell x + \beta_0^\ell \right) \left( \beta_{k-1}^{\ell'} x^{k-1} \right. \right. \\ &\quad \left. \left. + \beta_{k-2}^{\ell'} x^{k-2} + \cdots + \beta_1^{\ell'} x + \beta_0^{\ell'} \right) \mathbf{1}_{I_{\text{right}}}(x) \right\} d\mu(x) \\ &= \left( \alpha_k^\ell \right)^{\text{tr}} L_I \alpha_k^{\ell'} + \left( \beta_k^\ell \right)^{\text{tr}} R_I \beta_k^{\ell'} \\ &= \left( \beta_k^\ell \right)^{\text{tr}} \left[ L_I^{-1} R_I \right]^{\text{tr}} L_I \left[ L_I^{-1} R_I \right] \beta_k^{\ell'} + \left( \beta_k^\ell \right)^{\text{tr}} R_I \beta_k^{\ell'} \end{aligned}$$

equivalently

$$0 = \left( \beta_k^\ell \right)^{\text{tr}} \left\{ R_I L_I^{-1} L_I L_I^{-1} R_I + R_I \right\} \beta_k^{\ell'}$$

$$= (\beta_k^\ell)^\text{tr} \mathbf{R}_I \mathbf{L}_I^{-1} \{ \mathbf{R}_I + \mathbf{L}_I \} \beta_k^{\ell'} = (\beta_k^\ell)^\text{tr} X_I \beta_k^{\ell'}$$

where  $X_I \equiv \mathbf{L}_I \mathbf{R}_I \mathbf{L}_I^{-1} \{ \mathbf{R}_I + \mathbf{L}_I \}$  is invertible since  $\mathbf{R}_I + \mathbf{L}_I = \mathbf{M}_{I,k} \succ 0$  by (3.25). Thus the orthogonality can be achieved simply by choosing  $\beta_k^{\ell'}$  perpendicular to the vector  $(\beta_k^\ell)^\text{tr} X_I$ .

Finally we note that the system of equations that we have solved to obtain an orthonormal basis of Alpert functions is underdetermined since we have not yet solved for the *additional* moment conditions,

$$\int a_I^{\mu,\ell}(x) x^i d\mu(x) = 0, \quad \text{for all } 2 \leq \ell \leq k \text{ and } k \leq i \leq k + \ell - 2,$$

and we now explicitly compute the additional equations under which these additional moment conditions can also be achieved for the orthonormal basis of Alpert functions. As above have

$$0 = \alpha_{k-1}^\ell \left| I_{\text{left}}^{(i+k-1)} \right|_\mu + \alpha_{k-2}^\ell \left| I_{\text{left}}^{(i+k-2)} \right|_\mu + \dots + \alpha_0^\ell \left| I_{\text{left}}^{(i)} \right|_\mu + \beta_{k-1}^\ell \left| I_{\text{right}}^{(i+k-1)} \right|_\mu + \beta_{k-2}^\ell \left| I_{\text{right}}^{(i+k-2)} \right|_\mu + \dots + \beta_0^\ell \left| I_{\text{right}}^{(i)} \right|_\mu,$$

but now for all  $2 \leq \ell \leq k$  and  $k \leq i \leq k + \ell - 2$ , which leads to

$$\left[ \left| I_{\text{left}}^{(i)} \right|_\mu \dots \left| I_{\text{left}}^{(i+k-1)} \right|_\mu \right] \begin{pmatrix} \alpha_0^\ell \\ \vdots \\ \alpha_{k-1}^\ell \end{pmatrix} + \left[ \left| I_{\text{right}}^{(i)} \right|_\mu \dots \left| I_{\text{right}}^{(i+k-1)} \right|_\mu \right] \begin{pmatrix} \beta_0^\ell \\ \vdots \\ \beta_{k-1}^\ell \end{pmatrix} = 0,$$

for all  $2 \leq \ell \leq k$  and  $k \leq i \leq k + \ell - 2$ ,

and can be written as

$$\mathbf{v}_{\text{left}}^{(i+k-1,i)} \cdot \alpha_k^\ell + \mathbf{v}_{\text{right}}^{(i+k-1,i)} \cdot \beta_k^\ell = 0, \quad 2 \leq \ell \leq k \text{ and } k \leq i \leq k + \ell - 2, \tag{3.20}$$

where

$$\mathbf{v}_{\text{left}}^{(i+k-1,i)} \equiv \left[ \left| I_{\text{left}}^{(i)} \right|_\mu \dots \left| I_{\text{left}}^{(i+k-1)} \right|_\mu \right], \quad \mathbf{v}_{\text{right}}^{(i+k-1,i)} \equiv \left[ \left| I_{\text{right}}^{(i)} \right|_\mu \dots \left| I_{\text{right}}^{(i+k-1)} \right|_\mu \right].$$

We consider the  $\binom{k}{2} = \frac{k(k-1)}{2}$  orthogonality conditions among the  $k$  Alpert functions  $\{ a_I^{\mu,\ell} \}_{\ell=1}^k$  and the  $\binom{k}{2}$  additional moment conditions arising from the

index choices  $(\ell, i)$  in (3.20), which for convenience in visualizing we arrange in triangular form as

$$\left\{ \begin{array}{cccccccc} (k, k) & (k, k + 1) & (k, k + 2) & \cdots & (k, 2k - 4) & (k, 2k - 3) & (k, 2k - 2) \\ (k - 1, k) & (k - 1, k + 1) & (k - 1, k + 2) & \cdots & (k - 1, 2k - 4) & (k - 1, 2k - 3) \\ (k - 2, k) & (k - 2, k + 1) & (k - 2, k + 2) & \cdots & (k - 2, 2k - 4) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (4, k) & (4, k + 1) & (4, k + 2) & \cdots & & & \\ (3, k) & (3, k + 1) & & \cdots & & & \\ (2, k) & & & \cdots & & & \end{array} \right\}. \tag{3.21}$$

Using  $\alpha_k^\ell = -[L_I^k]^{-1} R_I^k \beta_k^\ell$ , the corresponding additional moment conditions are given by

$$0 = \mathbf{v}_{\text{left}}^{(i+k-1,i)} \cdot \alpha_k^\ell + \mathbf{v}_{\text{right}}^{(i+k-1,i)} \cdot \beta_k^\ell = \left\{ -\mathbf{v}_{\text{left}}^{(i+k-1,i)} \cdot \mathbf{M}_I^k + \mathbf{v}_{\text{right}}^{(i+k-1,i)} \right\} \cdot \beta_k^\ell = \mathbf{w}_k^i \cdot \beta_k^\ell$$

for  $2 \leq \ell \leq k$  and  $k \leq i \leq k + \ell - 2$ ,

where  $\mathbf{M}_I^k \equiv -[L_I^k]^{-1} R_I^k \beta_k^\ell$ , and

$$\mathbf{v}_{\text{left}}^{(i+k-1,i)} \equiv \left[ \left| I_{\text{left}}^{(i)} \right|_\mu \cdots \left| I_{\text{left}}^{(i+k-1)} \right|_\mu \right], \quad \mathbf{v}_{\text{right}}^{(i+k-1,i)} \equiv \left[ \left| I_{\text{right}}^{(i)} \right|_\mu \cdots \left| I_{\text{right}}^{(i+k-1)} \right|_\mu \right],$$

and

$$\mathbf{w}_k^i \equiv -\mathbf{v}_{\text{left}}^{(i+k-1,i)} \cdot \mathbf{M}_I^k + \mathbf{v}_{\text{right}}^{(i+k-1,i)}.$$

Thus the additional moment conditions are  $0 = \mathbf{w}_k^i \cdot \beta_k^\ell$  for all  $2 \leq \ell \leq k$  and  $k \leq i \leq k + \ell - 2$ , and written in the form of triangular matrix (3.21) are

$$\begin{aligned} 0 &= \mathbf{w}_k^k \cdot \beta_k^k = \mathbf{w}_k^{k+1} \cdot \beta_k^k = \cdots = \mathbf{w}_k^{2k-2} \cdot \beta_k^k, \\ 0 &= \mathbf{w}_k^k \cdot \beta_k^{k-1} = \mathbf{w}_k^{k+1} \cdot \beta_k^{k-1} = \cdots = \mathbf{w}_k^{2k-3} \cdot \beta_k^{k-1}, \\ 0 &= \mathbf{w}_k^k \cdot \beta_k^{k-2} = \mathbf{w}_k^{k+1} \cdot \beta_k^{k-2} = \cdots = \mathbf{w}_k^{2k-4} \cdot \beta_k^{k-2}, \\ &\vdots \\ 0 &= \mathbf{w}_k^k \cdot \beta_k^4 = \mathbf{w}_k^{k+1} \cdot \beta_k^4 = \mathbf{w}_k^{k+2} \cdot \beta_k^4, \\ 0 &= \mathbf{w}_k^k \cdot \beta_k^3 = \mathbf{w}_k^{k+1} \cdot \beta_k^3, \\ 0 &= \mathbf{w}_k^k \cdot \beta_k^2. \end{aligned}$$

Recall that mutual orthogonality of the  $k$  Alpert functions  $\left\{ a_I^{\mu,\ell} \right\}_{\ell=1}^k$  requires choosing  $\beta_k^{\ell'}$  perpendicular to the vector  $(\beta_k^\ell)^\text{tr} X_I$  for  $\ell \neq \ell'$ . In addition we must have

from (3.21) that

$$\begin{aligned}
 \mathbf{w}_k^k, \mathbf{w}_k^{k+1}, \dots, \mathbf{w}_k^{2k-2} &\in (\boldsymbol{\beta}_k^k)^\perp, \\
 \mathbf{w}_k^k, \mathbf{w}_k^{k+1}, \dots, \mathbf{w}_k^{2k-3} &\in (\boldsymbol{\beta}_k^{k-1})^\perp, \\
 \mathbf{w}_k^k, \mathbf{w}_k^{k+1}, \dots, \mathbf{w}_k^{2k-4} &\in (\boldsymbol{\beta}_k^{k-2})^\perp, \\
 &\vdots \\
 \mathbf{w}_k^k, \mathbf{w}_k^{k+1}, \mathbf{w}_k^{k+2} &\in (\boldsymbol{\beta}_k^4)^\perp, \\
 \mathbf{w}_k^k, \mathbf{w}_k^{k+1} &\in (\boldsymbol{\beta}_k^3)^\perp, \\
 \mathbf{w}_k^k &\in (\boldsymbol{\beta}_k^2)^\perp,
 \end{aligned}$$

where  $\boldsymbol{\beta}^\perp \equiv \{\mathbf{u} \in \mathbb{R}^k : \mathbf{u} \cdot \boldsymbol{\beta} = 0\}$ . Now each of the matrices  $\mathbf{R}_I \mathbf{L}_I^{-1} \mathbf{R}_I$  and  $\mathbf{R}_I$  is symmetric and positive definite, and so then is their sum  $\mathbf{X}_I = \mathbf{R}_I \mathbf{L}_I^{-1} \mathbf{R}_I + \mathbf{R}_I$ . Thus we can consider the inner product defined by

$$\langle \mathbf{v}, \mathbf{w} \rangle_I \equiv \mathbf{v}^{\text{tr}} \mathbf{X}_I \mathbf{w}.$$

This gives us a  $k$ -dimensional inner product space which we denote by  $\mathcal{X}_I$ . The condition  $\mathbf{w}_k^i \cdot \boldsymbol{\beta}_k^\ell = 0$  can be written as

$$\langle \mathbf{u}_k^i, \boldsymbol{\beta}_k^\ell \rangle_I = (\mathbf{u}_k^i)^{\text{tr}} \mathbf{X}_I \boldsymbol{\beta}_k^\ell = \mathbf{w}_k^i \cdot \boldsymbol{\beta}_k^\ell = 0 \tag{3.22}$$

if we define  $\mathbf{u}_k^i = \mathbf{X}_I^{-1} (\mathbf{w}_k^i)^{\text{tr}}$ . Then our conditions become

$$\begin{aligned}
 \langle \boldsymbol{\beta}_k^\ell, \boldsymbol{\beta}_k^{\ell'} \rangle_I &= 0 \text{ for } 1 \leq \ell < \ell' \leq k, \\
 \langle \mathbf{u}_k^i, \boldsymbol{\beta}_k^\ell \rangle_I &= 0 \text{ for } 2 \leq \ell \leq k \text{ and } k \leq i \leq k + \ell - 2,
 \end{aligned} \tag{3.23}$$

and the second line above can be written out in triangular form as

$$\begin{aligned}
 \mathbf{u}_k^k, \mathbf{u}_k^{k+1}, \dots, \mathbf{u}_k^{2k-2} &\in (\boldsymbol{\beta}_k^k)^{\perp_I}, \\
 \mathbf{u}_k^k, \mathbf{u}_k^{k+1}, \dots, \mathbf{u}_k^{2k-3} &\in (\boldsymbol{\beta}_k^{k-1})^{\perp_I}, \\
 \mathbf{u}_k^k, \mathbf{u}_k^{k+1}, \dots, \mathbf{u}_k^{2k-4} &\in (\boldsymbol{\beta}_k^{k-2})^{\perp_I}, \\
 &\vdots
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{u}_k^k, \mathbf{u}_k^{k+1}, \mathbf{u}_k^{k+2} &\in \left(\boldsymbol{\beta}_k^4\right)^{\perp_I}, \\
 \mathbf{u}_k^k, \mathbf{u}_k^{k+1} &\in \left(\boldsymbol{\beta}_k^3\right)^{\perp_I}, \\
 \mathbf{u}_k^k &\in \left(\boldsymbol{\beta}_k^2\right)^{\perp_I},
 \end{aligned}
 \tag{3.24}$$

where  $\boldsymbol{\beta}^{\perp_I} \equiv \{\mathbf{u} \in \mathcal{X}_I : \langle \mathbf{u}, \boldsymbol{\beta} \rangle_I = 0\}$ .

We can satisfy conditions (3.23) by first choosing a unit vector  $\boldsymbol{\beta}_k^k \in \mathcal{X}_I$  so that the first line in (3.24) holds. Then we choose a unit vector  $\boldsymbol{\beta}_k^{k-1} \in \left(\boldsymbol{\beta}_k^k\right)^{\perp_I}$  so that the second line in (3.24) holds. Then we choose a unit vector  $\boldsymbol{\beta}_k^{k-2} \in \left(\text{Span}\left\{\boldsymbol{\beta}_k^{k-1}, \boldsymbol{\beta}_k^k\right\}\right)^{\perp_I}$  so that the third line in (3.24) holds. Continuing in this way, we find unit vectors  $\{\boldsymbol{\beta}_k^\ell\}_{\ell=1}^k$  in the inner product space  $\mathcal{X}_I$  so that our conditions (3.24) and (3.23) hold. For the generic choice of vectors  $\mathbf{v}_{\text{left}}^{(i+k-1,i)}$  and  $\mathbf{v}_{\text{right}}^{(i+k-1,i)}$ , the choice of  $\{\boldsymbol{\beta}_k^\ell\}_{\ell=1}^k$  will be unique up to sign. This completes the proof of part (2) of Theorem 2.

We have adopted the additional moment conditions introduced by Alpert in the setting of Lebesgue measure in [1], but one can in fact replace these conditions by an essentially arbitrary collection of the correct number of moment conditions. We leave the nondegenerate case for the reader in the following exercise, and the degenerate case is not treated here at all, except for the simple case when  $k = 2$  and  $n = 1$  solved below.

**Remark 13** If  $M_{I_{\text{left}},k} > 0$ ,  $M_{I_{\text{right}},k} > 0$  and  $\mathcal{M} \subset \{(\ell, i) : 1 \leq \ell \leq k \text{ and } i \geq k\}$  has cardinality  $\binom{k}{2}$ , then we can choose an Alpert basis  $\{a_I^{\mu,\ell}\}_{\ell=1}^k$  satisfying the additional moment conditions  $\mathbf{w}_k^i \cdot \boldsymbol{\beta}_k^\ell = 0$  for all  $(\ell, i) \in \mathcal{M}$ .

### 3.3.3 The Alpert Degenerate Case

In the event that one or more of the matrices  $L_I^k$  and/or  $R_I^k$  is singular, then it is easy to see from (3.14) that the maximum number of independent Alpert functions equals the dimension of the intersection of the ranges of  $L_I^k$  and  $R_I^k$  in  $\mathbb{R}^k$ , i.e.

$$\dim L_{Q;k}^2(\mu) = \dim \left( \text{Range } L_I^k \cap \text{Range } R_I^k \right).$$

In the special case when  $k = 2$ —when there are only two Alpert functions  $a_I^{\mu,1}$ ,  $a_I^{\mu,2}$  and just one additional moment condition for  $a_I^{\mu,2}$ —we will show in the next subsection that it is not always possible to arrange for this additional moment condition to hold. In fact we will show there that it holds if and only if

$$\left( \begin{array}{c} \left| I_{\text{right}}^{(2)} \right. \\ \left. \mu \right. \\ \left| I_{\text{right}}^{(3)} \right. \\ \left. \mu \right. \end{array} \right) \in \text{Range } R_I^2.$$

Similar results hold for larger  $k$ , but we will not pursue these here.

### 3.4 The Special Case $n = 1$ and $k = 2$

First we quickly review and set notation for the nondegenerate case when  $k = 2$ , and later proceed to the degenerate case. Let  $\mu$  be a locally finite positive Borel measure on  $\mathbb{R}$  that satisfies the Haar nondegeneracy condition  $|J|_\mu > 0$  for all  $J \in \mathcal{D}$  and in addition satisfies the Alpert nondegeneracy condition

$$\det \begin{bmatrix} |J^{(0)}|_\mu & |J^{(1)}|_\mu \\ |J^{(1)}|_\mu & |J^{(2)}|_\mu \end{bmatrix} > 0, \quad \text{for all } J \in \mathcal{D}. \tag{3.25}$$

Recall that the determinant in (3.25) is nonnegative by the Cauchy-Schwarz inequality, and is positive *if and only if* the functions  $\mathbf{1}_J(x)$  and  $x\mathbf{1}_J(x)$  are linearly independent on  $J$ :

$$|J^{(1)}|_\mu^2 = \left( \int_J x d\mu(x) \right)^2 \leq \left( \int_J 1^2 d\mu(x) \right) \left( \int_J x^2 d\mu(x) \right) = |J^{(0)}|_\mu |J^{(2)}|_\mu$$

with equality if and only if the functions  $\mathbf{1}_J(x)$  and  $x\mathbf{1}_J(x)$  are linearly dependent on  $J$ .

We require the functions

$$\begin{aligned} a_I^{\mu,1}(x) &\equiv (\alpha_1^1 x + \alpha_0^1) \mathbf{1}_{I_{\text{left}}}(x) + (\beta_1^1 x + \beta_0^1) \mathbf{1}_{I_{\text{right}}}(x), \\ a_I^{\mu,2}(x) &\equiv (\alpha_1^2 x + \alpha_0^2) \mathbf{1}_{I_{\text{left}}}(x) + (\beta_1^2 x + \beta_0^2) \mathbf{1}_{I_{\text{right}}}(x), \end{aligned}$$

to satisfy the moment properties,

$$\begin{aligned} (A) \quad &\int a_I^{\mu,1}(x) d\mu(x) = 0 \text{ and } \int a_I^{\mu,1}(x) x d\mu(x) = 0, \\ (B) \quad &\int a_I^{\mu,2}(x) d\mu(x) = 0 \text{ and } \int a_I^{\mu,2}(x) x d\mu(x) = 0, \\ (C) \quad &\int a_I^{\mu,2}(x) x^2 d\mu(x) = 0, \end{aligned}$$

the orthogonality property,

$$(D) \quad \int a_I^{\mu,1}(x) a_I^{\mu,2}(x) d\mu(x) = 0,$$

and the normalization properties,

$$\left( \int |a_I^{\mu,1}(x)|^2 d\mu(x) \right)^{\frac{1}{2}} = 1,$$



$$\left( \int |a_I^{\mu,2}(x)|^2 d\mu(x) \right)^{\frac{1}{2}} = 1.$$

Theorem 1 above gives the following conclusion.

**Theorem 14** *The collection*

$$\mathcal{U}^{\text{Alpert},\mu} \equiv \left\{ a_I^{\mu,1}, a_I^{\mu,2} \right\}_{I \in \mathcal{D}}$$

is an orthonormal basis for  $L^2(\mu)$ .

If we wish to include the additional moment condition (C), i.e.  $\int a_I^{\mu,2}(x) x^2 d\mu(x) = 0$ , then we must in addition solve

$$\begin{aligned} 0 &= \int \left\{ (\alpha_1^2 x + \alpha_0^2) \mathbf{1}_{I_{\text{left}}}(x) + (\beta_1^2 x + \beta_0^2) \mathbf{1}_{I_{\text{right}}}(x) \right\} x^2 d\mu(x) \\ &= \alpha_0^2 |I_{\text{left}}^{(2)}|_{\mu} + \alpha_1^2 |I_{\text{left}}^{(3)}|_{\mu} + \beta_0^2 |I_{\text{right}}^{(2)}|_{\mu} + \beta_1^2 |I_{\text{right}}^{(3)}|_{\mu} \\ &= \left( |I_{\text{left}}^{(2)}|_{\mu} \left| I_{\text{left}}^{(3)} \right|_{\mu} \right) \begin{pmatrix} \alpha_0^2 \\ \alpha_1^2 \end{pmatrix} + \left( |I_{\text{right}}^{(2)}|_{\mu} \left| I_{\text{right}}^{(3)} \right|_{\mu} \right) \begin{pmatrix} \beta_0^2 \\ \beta_1^2 \end{pmatrix} \\ &= \left\{ - \left( |I_{\text{left}}^{(2)}|_{\mu} \left| I_{\text{left}}^{(3)} \right|_{\mu} \right) L_I^{-1} R_I + \left( |I_{\text{right}}^{(2)}|_{\mu} \left| I_{\text{right}}^{(3)} \right|_{\mu} \right) \right\} \begin{pmatrix} \beta_0^2 \\ \beta_1^2 \end{pmatrix} \end{aligned}$$

which means we must choose  $\begin{pmatrix} \beta_0^2 \\ \beta_1^2 \end{pmatrix}$  perpendicular to the vector

$$- \left( |I_{\text{left}}^{(2)}|_{\mu} \left| I_{\text{left}}^{(3)} \right|_{\mu} \right) L_I^{-1} R_I + \left( |I_{\text{right}}^{(2)}|_{\mu} \left| I_{\text{right}}^{(3)} \right|_{\mu} \right).$$

Thus it must be the case that the two vectors

$$\begin{pmatrix} \beta_0^1 \\ \beta_1^1 \end{pmatrix} R_I L_I^{-1} \{R_I + L_I\}$$

and

$$- \left( |I_{\text{left}}^{(2)}|_{\mu} \left| I_{\text{left}}^{(3)} \right|_{\mu} \right) L_I^{-1} R_I + \left( |I_{\text{right}}^{(2)}|_{\mu} \left| I_{\text{right}}^{(3)} \right|_{\mu} \right)$$

are parallel. But since  $R_I L_I^{-1} \{R_I + L_I\}$  is invertible, this can clearly be achieved by choosing  $\begin{pmatrix} \beta_0^1 \\ \beta_1^1 \end{pmatrix}$  appropriately, thereby using up our last degree of freedom in the case  $k = 2$ .

Finally, we examine what happens when one or more of the nondegeneracy conditions (2.1) and (3.25) fails. Note that for a given interval  $J$ , we have that (2.1) holds and (3.25) fails if and only if  $\mu 1_J$  is a point mass. Indeed,  $\mathbf{1}_J(x)$  and  $x \mathbf{1}_J(x)$  are

dependent if and only if  $\mu|_J$  is a point mass, which we locate at  $x_J \in J$ . In this case we set

$$\mu|_J = |J|_\mu \delta_{x_J}.$$

We also recall

$$M_J = M_{J,2} \equiv \left[ \begin{array}{c|c} \left| J_{\text{left}}^{(0)} \right|_\mu & \left| J_{\text{left}}^{(1)} \right|_\mu \\ \left| J_{\text{left}}^{(1)} \right|_\mu & \left| J_{\text{left}}^{(2)} \right|_\mu \end{array} \right] \text{ for } J \in \mathcal{D}.$$

**Lemma 15** *Let  $I = I_{\text{left}} \dot{\cup} I_{\text{right}}$  be the decomposition of  $I$  into its two children.*

- (1) *If (2.1) holds for both  $I_{\text{left}}$  and  $I_{\text{right}}$ , and (3.25) holds for  $I_{\text{left}}$  but fails for  $I_{\text{right}}$ , then  $a_I^{\mu,1} = a_I^{\mu,2} \neq 0$  in  $L^2(\mu)$ . This conclusion persists in the opposite situation where (2.1) holds for both  $I_{\text{left}}$  and  $I_{\text{right}}$ , but (3.25) fails for  $I_{\text{left}}$  and holds for  $I_{\text{right}}$ .*
- (2) *In all other degenerate cases, where at least one of the matrices*

$$\left[ \begin{array}{c|c} \left| (I_{\text{left}})^{(0)} \right|_\mu & \left| (I_{\text{left}})^{(1)} \right|_\mu \\ \left| (I_{\text{left}})^{(1)} \right|_\mu & \left| (I_{\text{left}})^{(2)} \right|_\mu \end{array} \right], \left[ \begin{array}{c|c} \left| (I_{\text{right}})^{(0)} \right|_\mu & \left| (I_{\text{right}})^{(1)} \right|_\mu \\ \left| (I_{\text{right}})^{(1)} \right|_\mu & \left| (I_{\text{right}})^{(2)} \right|_\mu \end{array} \right] \text{ fails to be positive def-}$$

*inite, we have  $a_I^{\mu,1} = a_I^{\mu,2} = 0$  in  $L^2(\mu)$ .*

**Proof** Fix  $I \in \mathcal{D}$ .

Assertion (1): We have

$$\det \left[ \begin{array}{c|c} \left| (I_{\text{left}})^{(0)} \right|_\mu & \left| (I_{\text{left}})^{(1)} \right|_\mu \\ \left| (I_{\text{left}})^{(1)} \right|_\mu & \left| (I_{\text{left}})^{(2)} \right|_\mu \end{array} \right] > 0 \text{ and } \det \left[ \begin{array}{c|c} \left| (I_{\text{right}})^{(0)} \right|_\mu & \left| (I_{\text{right}})^{(1)} \right|_\mu \\ \left| (I_{\text{right}})^{(1)} \right|_\mu & \left| (I_{\text{right}})^{(2)} \right|_\mu \end{array} \right] = 0.$$

Thus  $\mathbf{1}_{I_{j,\text{right}}^m} \mu$  is a point mass located at  $x_{I_{j,\text{right}}^m}$ . Now (3.19) shows that  $\begin{pmatrix} \alpha_0^\ell \\ \alpha_1^\ell \end{pmatrix}$

is uniquely determined by  $\begin{pmatrix} \beta_0^\ell \\ \beta_1^\ell \end{pmatrix}$  for  $1 \leq \ell \leq 2$ , and we know that the

range of  $\left[ \begin{array}{c|c} \left| (I_{j,\text{right}}^m)^{(2)} \right|_\mu & \left| (I_{j,\text{right}}^m)^{(1)} \right|_\mu \\ \left| (I_{j,\text{right}}^m)^{(1)} \right|_\mu & \left| (I_{j,\text{right}}^m)^{(0)} \right|_\mu \end{array} \right]$  is just one-dimensional, since  $\mathbf{1}_{I_{j,\text{right}}^m} \mu$

is a point mass located at  $x_{I_{j,\text{right}}^m}$ . It follows that both  $a_I^{\mu,1}$  and  $a_I^{\mu,2}$  are constant on the interval  $I_{j,\text{right}}^m$  in the space  $L^2(\mu)$ , and since the null space of

$\left[ \begin{array}{c|c} \left| (I_{j,\text{right}}^m)^{(0)} \right|_\mu & \left| (I_{j,\text{right}}^m)^{(1)} \right|_\mu \\ \left| (I_{j,\text{right}}^m)^{(1)} \right|_\mu & \left| (I_{j,\text{right}}^m)^{(2)} \right|_\mu \end{array} \right]$  is one-dimensional, namely  $\text{Span} \begin{pmatrix} -x_{I_{j,\text{right}}^m} \\ 1 \end{pmatrix}$ , we

see that the functions  $a_I^{\mu,1}$  and  $a_I^{\mu,2}$  are linearly dependent, i.e. there is only one Alpert function in this case, i.e.  $a_{I_j^m}^{\mu,1} = a_{I_j^m}^{\mu,2}$ .

Assertion (2): From above we have  $\dim L^2_{Q;2}(\mu) = \dim (\text{Range } L_I^2 \cap \text{Range } \mathbf{R}_I^2)$ . Thus if one of the ranges is  $\{0\}$  we are done. So we are left with the case where (2.1) holds for both  $I_{j,\text{left}}^m$  and  $I_{j,\text{right}}^m$ , and (3.25) fails for both  $I_{j,\text{left}}^m$  and  $I_{j,\text{right}}^m$ . We then have

$$\begin{bmatrix} \left| \begin{pmatrix} I_{j,\text{left}}^m \\ I_{j,\text{left}}^m \end{pmatrix} \right|_{\mu}^{(0)} \left| \begin{pmatrix} I_{j,\text{left}}^m \\ I_{j,\text{left}}^m \end{pmatrix} \right|_{\mu}^{(1)} \\ \left| \begin{pmatrix} I_{j,\text{left}}^m \\ I_{j,\text{left}}^m \end{pmatrix} \right|_{\mu}^{(1)} \left| \begin{pmatrix} I_{j,\text{left}}^m \\ I_{j,\text{left}}^m \end{pmatrix} \right|_{\mu}^{(2)} \end{bmatrix} \begin{pmatrix} \alpha_0^1 \\ \alpha_1^1 \end{pmatrix} = - \begin{bmatrix} \left| \begin{pmatrix} I_{j,\text{right}}^m \\ I_{j,\text{right}}^m \end{pmatrix} \right|_{\mu}^{(0)} \left| \begin{pmatrix} I_{j,\text{right}}^m \\ I_{j,\text{right}}^m \end{pmatrix} \right|_{\mu}^{(1)} \\ \left| \begin{pmatrix} I_{j,\text{right}}^m \\ I_{j,\text{right}}^m \end{pmatrix} \right|_{\mu}^{(1)} \left| \begin{pmatrix} I_{j,\text{right}}^m \\ I_{j,\text{right}}^m \end{pmatrix} \right|_{\mu}^{(2)} \end{bmatrix} \begin{pmatrix} \beta_0^1 \\ \beta_1^1 \end{pmatrix}.$$

The point mass  $\mathbf{1}_{I_{\text{right}}^m} \mu$  is located at  $x_{I_{j,\text{right}}^m}$ , and the point mass  $\mathbf{1}_{I_{\text{left}}^m} \mu$  is located at

$x_{I_{j,\text{left}}^m}$ , and thus the one-dimensional ranges of  $\begin{bmatrix} \left| \begin{pmatrix} I_{j,\text{left}}^m \\ I_{j,\text{left}}^m \end{pmatrix} \right|_{\mu}^{(0)} \left| \begin{pmatrix} I_{j,\text{left}}^m \\ I_{j,\text{left}}^m \end{pmatrix} \right|_{\mu}^{(1)} \\ \left| \begin{pmatrix} I_{j,\text{left}}^m \\ I_{j,\text{left}}^m \end{pmatrix} \right|_{\mu}^{(1)} \left| \begin{pmatrix} I_{j,\text{left}}^m \\ I_{j,\text{left}}^m \end{pmatrix} \right|_{\mu}^{(2)} \end{bmatrix}$  and

$\begin{bmatrix} \left| \begin{pmatrix} I_{j,\text{right}}^m \\ I_{j,\text{right}}^m \end{pmatrix} \right|_{\mu}^{(0)} \left| \begin{pmatrix} I_{j,\text{right}}^m \\ I_{j,\text{right}}^m \end{pmatrix} \right|_{\mu}^{(1)} \\ \left| \begin{pmatrix} I_{j,\text{right}}^m \\ I_{j,\text{right}}^m \end{pmatrix} \right|_{\mu}^{(1)} \left| \begin{pmatrix} I_{j,\text{right}}^m \\ I_{j,\text{right}}^m \end{pmatrix} \right|_{\mu}^{(2)} \end{bmatrix}$  are respectively spanned by the vectors  $\begin{pmatrix} x_{I_{j,\text{left}}^m} \\ 1 \end{pmatrix}$

and  $\begin{pmatrix} x_{I_{j,\text{right}}^m} \\ 1 \end{pmatrix}$ , which are independent since  $I_{j,\text{left}}^m \cap I_{j,\text{right}}^m = \emptyset$ . Thus only the trivial

solution  $\begin{pmatrix} \alpha_0^1 \\ \alpha_1^1 \end{pmatrix} = \begin{pmatrix} \beta_0^1 \\ \beta_1^1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  exists, and it follows that both  $a_{I_j^m}^{\mu,1}$  and  $a_{I_j^m}^{\mu,2}$  vanish in  $L^2(\mu)$ . □

The resulting pared collection

$$\mathcal{U}^{\text{Alpert},\mu} \equiv \left\{ a_I^{\mu,1}, a_I^{\mu,2} \right\}_{I \in \mathcal{D}},$$

where  $a_I^{\mu,1}$  or  $a_I^{\mu,2}$  or both are removed according to the lemma above, is an orthonormal basis for  $L^2(\mu)$ .

Finally, we consider the additional moment condition (C) in the case where assertion (1) of the above lemma holds, namely when (2.1) holds for both  $I_{j,\text{left}}^m$  and  $I_{j,\text{right}}^m$ , and (3.25) holds for  $I_{j,\text{left}}^m$  but fails for  $I_{j,\text{right}}^m$ . In this case  $a_{I_j^m}^{\mu} = b_{I_j^m}^{\mu} \neq 0$  in  $L^2(\mu)$ , and this is the only case in which there is just one Alpert function (apart from the symmetric case when (3.25) holds for  $I_{j,\text{right}}^m$  but fails for  $I_{j,\text{left}}^m$ ).

From the calculations above we must have that the two vectors

$$\begin{pmatrix} \beta_0^1 \\ \beta_1^1 \end{pmatrix} \mathbf{R}_I \mathbf{L}_I^{-1} \{ \mathbf{R}_I + \mathbf{L}_I \}$$

and

$$- \left( \left| I_{\text{left}}^{(2)} \right|_{\mu} \left| I_{\text{left}}^{(3)} \right|_{\mu} \right) L_I^{-1} R_I + \left( \left| I_{\text{right}}^{(2)} \right|_{\mu} \left| I_{\text{right}}^{(3)} \right|_{\mu} \right)$$

are parallel. But since  $R_I L_I^{-1} \{R_I + L_I\}$  is no longer invertible - indeed,  $R_I L_I^{-1} \{R_I + L_I\}$  has rank 1 since  $R_I$  does - we cannot necessarily solve

$$(\beta_0^1 \ \beta_1^1) \left\{ R_I L_I^{-1} R_I + R_I \right\} = \lambda \left\{ \left( \left| I_{\text{right}}^{(2)} \right|_{\mu} \left| I_{\text{right}}^{(3)} \right|_{\mu} \right) - \left( \left| I_{\text{left}}^{(2)} \right|_{\mu} \left| I_{\text{left}}^{(3)} \right|_{\mu} \right) L_I^{-1} R_I \right\}.$$

In fact, there is a solution if and only if the transpose of the vector

$$\left( \left| I_{\text{right}}^{(2)} \right|_{\mu} \left| I_{\text{right}}^{(3)} \right|_{\mu} \right) - \left( \left| I_{\text{left}}^{(2)} \right|_{\mu} \left| I_{\text{left}}^{(3)} \right|_{\mu} \right) L_I^{-1} R_I$$

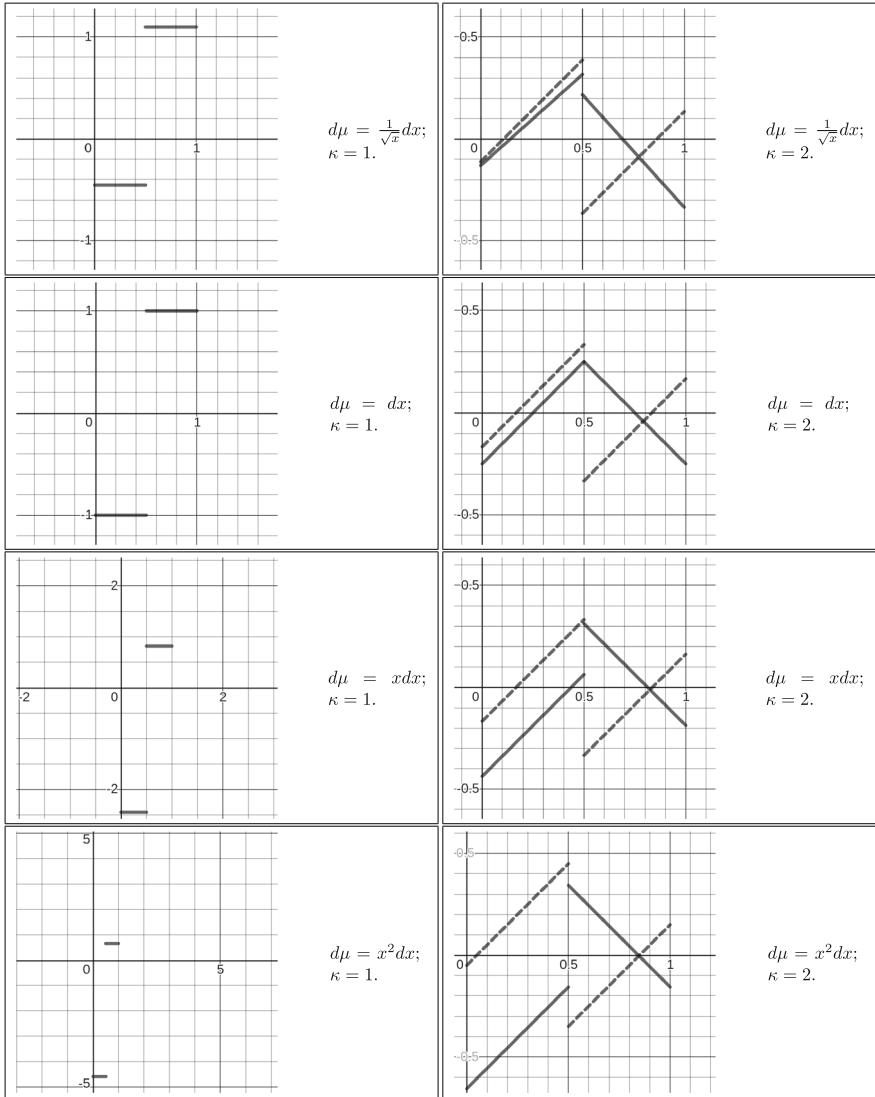
lies in the one-dimensional range of the matrix  $R_I$ , which in turn holds if and only if

$$\left( \begin{array}{c} \left| I_{\text{right}}^{(2)} \right|_{\mu} \\ \left| I_{\text{right}}^{(3)} \right|_{\mu} \end{array} \right) \in \text{Range } R_I.$$

**Remark 16** The authors thank Fletcher Gates for showing them that in the case under consideration this latter condition is always satisfied.

### 3.5 Some Examples of Wavelets in the Special Case

Here are some sample graphs (graphed on desmos.com) for measures  $x^{-\frac{1}{2}}dx$ ,  $dx$ ,  $x dx$ , and  $x^2 dx$ . Since these are all line segments, it is easy (though tedious) to verify that they satisfy the conditions claimed (except they don't satisfy normalization since this would make them too difficult to graph). However, in the unweighted case, many properties can be verified by appealing to properties of odd and even functions. The wavelets graphed with dashed lines are “odd” and the ones graphed with solid lines are “even”.



### 4 Application: A Two Weight $Tp$ Conjecture

Using the weighted Alpert wavelet bases constructed in Theorem 1, we can formulate an associated  $Tp$ -type theorem in dimension  $n = 1$  where testing over indicators is replaced by testing over indicators times polynomials of degree at most  $k - 1$ , and the energy condition is replaced by an associated  $k$ -energy condition. Unfortunately, at this point in time, we cannot demonstrate that Conjecture 17 produces new interesting weighted inequalities, despite that fact that we provide an example to show that the  $k$ -energy condition is strictly weaker than the usual energy condition, even in the

presence of the Muckenhoupt condition. But see Sect. 4.4 for a demonstration that Conjecture 17 differs ‘logically’ from existing  $T1$ -type theorems in the literature.

Let  $0 \leq \alpha < 1, k \in \mathbb{N}$  and  $0 < \delta < 1$ . We define a  $(k + \delta)$ -smooth  $\alpha$ -fractional CZ kernel  $K^\alpha(x, y)$  to be a real-valued function defined on  $\mathbb{R} \times \mathbb{R}$  satisfying the following fractional size and smoothness conditions of order  $1 + \delta$ : For  $x \neq y$ ,

$$|K^\alpha(x, y)| \leq C_{CZ} |x - y|^{\alpha-1} \text{ and } |\nabla^j K^\alpha(x, y)| \leq C_{CZ} |x - y|^{\alpha-1-j}, \quad 1 \leq j \leq k,$$

$$|\nabla^k K^\alpha(x, y) - \nabla^k K^\alpha(x', y)| \leq C_{CZ} \left( \frac{|x - x'|}{|x - y|} \right)^\delta |x - y|^{\alpha-1-k}, \quad \frac{|x - x'|}{|x - y|} \leq \frac{1}{2},$$

(4.1)

and the last inequality also holds for the adjoint kernel in which  $x$  and  $y$  are interchanged. We associate a corresponding Calderón–Zygmund operator  $T^\alpha$  is the usual way (see e.g. [7] or [10]).

The following conjectured  $Tp$  theorem with an energy side condition differs from the corresponding  $T1$  theorem with an energy side condition in two ways:

- (1) Because of the stronger moment conditions satisfied Alpert wavelets, the usual energy condition assumption from the  $T1$  theorem can be weakened in the  $Tp$  theorem.
- (2) Due to the weaker telescoping identities satisfied by Alpert wavelets, the  $T1$  testing conditions must be strengthened to testing polynomials times indicators.

**Conjecture 17** *Let  $0 \leq \alpha < 1$  and  $0 < \delta < 1$ . Let  $T^\alpha$  be a  $(k + \delta)$ -smooth  $\alpha$ -fractional Calderón–Zygmund operator on the real line. Suppose  $\sigma$  and  $\omega$  are locally finite positive Borel measures on  $\mathbb{R}$  that satisfy the  $k$ -energy condition*

$$(\mathcal{E}_{2,k}^\alpha)^2 \equiv \sup_{I=\cup_{r=1}^\infty I_r} \frac{1}{|I|_\sigma} \sum_{r=1}^\infty \left( \frac{P_k^\alpha(I_r, \mathbf{1}_I \sigma)}{|I_r|^k} \right)^2 \left\| (x - m_{I_r}^k)^k \right\|_{L^2(\mathbf{1}_I, \omega)}^2, \quad (4.2)$$

as well as the dual  $k$ -energy condition obtained by interchanging the measures  $\sigma$  and  $\omega$ . Then the operator  $T_\sigma^\alpha f \equiv T^\alpha(f\sigma)$  is bounded from  $L^2(\sigma)$  to  $L^2(\omega)$  (in the sense that tangent line truncations are uniformly bounded by a constant  $\mathfrak{A}_{T^\alpha}(\sigma, \omega)$ ) if

- (1) there is a positive constant  $\mathfrak{T}_{T^\alpha}^k(\sigma, \omega)$  such that

$$\int_Q |T_\sigma^\alpha \mathbf{1}_Q p|^2 d\omega \leq \left( \mathfrak{T}_{T^\alpha}^k(\sigma, \omega) \right)^2 \int_Q |p|^2 d\sigma,$$

for all intervals  $Q$  and polynomials  $p(x) = c_0 + c_1x + \dots + c_{k-1}x^{k-1}$  of degree at most  $k - 1$ , as well as the dual testing conditions obtained by interchanging the measures  $\sigma$  and  $\omega$  and replacing  $T_\sigma^\alpha$  with its dual  $T_\omega^\alpha$ ; and

- (2) the fractional Muckenhoupt condition is finite:  $\mathfrak{A}_2^\alpha(\sigma, \omega) < \infty$ .

A proof of this conjecture would follow the somewhat standard lines of proof for  $T1$ -type theorems already in the literature ([5,6,8,9]), namely an inner product  $\langle T_\sigma^\alpha f, g \rangle_{L^2(\omega)}$  is expanded in Alpert projections as

$$\langle T_\sigma^\alpha f, g \rangle_{L^2(\omega)} = \sum_{\substack{I \in \mathcal{D} \\ 1 \leq \ell \leq k}} \sum_{\substack{J \in \mathcal{G} \\ 1 \leq \ell' \leq k'}} \langle T_\sigma^\alpha \Delta_I^{\sigma, \ell} f, \Delta_J^{\omega, \ell'} g \rangle_{L^2(\omega)}, \tag{4.3}$$

and then decomposed into many separate infinite sums according to the relative sizes, locations and goodness of the intervals  $I$  and  $J$ , which are then all controlled differently. We first highlight the two main points of departure in controlling these different sums, followed by a brief description of the sums themselves and how they are handled in the  $Tp$  situation, as well as pointing to an obstruction to the proof.

- The estimate for norms of Alpert projections  $\|\Delta_J^\omega T^\alpha \mu\|_{L^2(\omega)}^2$ , called the Monotonicity Lemma below and in [7,10], is improved by the extra vanishing moments of Alpert wavelets to

$$\|\Delta_J^\omega T^\alpha \mu\|_{L^2(\omega)}^2 \lesssim \left( \frac{P_k^\alpha(J, \mu)}{|J|^k} \right)^2 \left\| (x - m_J^k)^k \right\|_{L^2(\mathbf{1}_J \omega)}^2,$$

which in turn can then be controlled by the  $k$ -energy condition (4.2) above, **weaker** than the usual energy condition with  $k = 1$ ;

- The telescoping identities (1.5) reduce sums of consecutive Alpert projections  $\Delta_{I;k}^\mu$  to projections  $\mathbb{E}_{Q;k}^\mu$  onto spaces of polynomials of degree at most  $k - 1$ , thus requiring the use of **stronger** testing conditions, taken locally over polynomials of degree at most  $k - 1$ , in order to bound the consecutive sums of Alpert projections that arise in the paraproduct and stopping forms in [7,10].

With these two changes in mind we can now review the main steps in the standard proof strategy for the interested reader, whom we alert to the fact that we are using here the formulation of the Lacey–Wick monotonicity lemma with an error term in [7,10], as opposed to the stronger formulation used in [6,8] that exploited special properties of the Haar basis to hide the error term. As a consequence, the reader can follow the broad outline of the one-dimensional proof in [6,8], but handling the error terms as in [7] and/or [10] (see also [11,12]). We ignore the case of common point masses, and refer the reader instead to [5,7,12].

- Step 1 Using the random grids of Nazarov, Treil and Volberg, a reduction is made to good functions  $f$  and  $g$ , i.e. those whose wavelet expansions involve only cubes from one grid that are *good* with respect to the other grid. The orthogonality of Alpert projections plays a key role here.
- Step 2 Using the testing conditions, one further restricts the supports of  $f$  and  $g$  to a finite union of large cubes.
- Step 3 Then one can implement corona constructions with Calderón–Zygmund stopping times on the averages of  $f$ , and with  $k$ -energy stopping times in place of the familiar energy stopping times.

- Step 4 The sum of inner products in (4.3) is then grouped into coronas relative to these stopping times and further decomposed into global, local and error pieces.
- Step 5 The error pieces are handled by NTV methods from [9].
- Step 6 The global inner products are controlled by the  $k$ -Poisson operator as in [7] and/or [10], which in turn has its norm inequality controlled by  $k$ -Poisson testing conditions. This latter result is proved in the same way as is done for the familiar Poisson operator.
- Step 7 The local terms are handled by Lacey’s bottom/up stopping time and recursion as in [6], with error terms from the monotonicity lemma handled as in [7] and/or [10]. The difficulty lies in using the Nazarov, Treil, Volberg method connecting back to the appropriate paraproduct terms; at the moment we are unable to control those terms and so do not have a proof of Conjecture 17.

**Remark 18** There is an analogous conjecture in higher dimensions following the arguments in [7] and/or [10], but we will not pursue this.

### 4.1 The Monotonicity Lemma

For  $0 \leq \alpha < 1$  and  $m \in \mathbb{R}_+$ , we recall the  $m$ -weighted fractional Poisson integral

$$P_m^\alpha(J, \mu) \equiv \int_{\mathbb{R}} \frac{|J|^m}{(|J| + |y - c_J|)^{m+1-\alpha}} d\mu(y),$$

where  $P_1^\alpha(J, \mu) = P^\alpha(J, \mu)$  is the standard Poisson integral.

**Lemma 19 (Monotonicity)** *Suppose that  $I$  and  $J$  are cubes in  $\mathbb{R}$  such that  $J \subset 2J \subset I$ , and that  $\mu$  is a signed measure on  $\mathbb{R}$  supported outside  $I$ . Finally suppose that  $T^\alpha$  is a standard fractional singular integral on  $\mathbb{R}$  with kernel  $K^\alpha(x, y) = K_y^\alpha(x)$ ,  $0 < \alpha < 1$ . Then there is a positive constant  $C_\alpha$  such that*

$$\Phi^\alpha(J, \mu)^2 - C_\alpha \Psi^\alpha(J, |\mu|)^2 \leq \|\Delta_J^\omega T^\alpha \mu\|_{L^2(\omega)}^2 \lesssim \Phi^\alpha(J, \mu)^2 + C_\alpha \Psi^\alpha(J, |\mu|)^2, \tag{4.4}$$

where for a measure  $\nu$ ,

$$\begin{aligned} \Phi^\alpha(J, \nu)^2 &\equiv \left| \frac{1}{k!} \int (K_y^\alpha)^{(k)}(m_J) d\mu(y) \right|^2 \|\Delta_J^\omega x^k\|_{L^2(\omega)}^2, \\ \Psi^\alpha(J, \nu)^2 &\equiv \left( \frac{P_{k+\delta}^\alpha(J, \nu)}{|J|^k} \right)^2 \left\| (x - m_J^k)^k \right\|_{L^2(\mathbf{1}_J \omega)}^2, \\ \text{where } m_J^k \in J \text{ satisfies } &\left\| (x - m_J^k)^k \right\|_{L^2(\mathbf{1}_J \omega)}^2 = \inf_{c \in J} \left\| (x - c)^k \right\|_{L^2(\mathbf{1}_J \omega)}^2. \end{aligned}$$



and where if  $\nu$  is a positive measure, then there are positive constants  $c, C$  such that

$$c\Phi^\alpha(J, \nu)^2 \leq \left(\frac{P_k^\alpha(J, \nu)}{|J|^k}\right)^2 \|\Delta_J^\omega x^k\|_{L^2(\omega)}^2 \leq C\Phi^\alpha(J, \nu)^2.$$

**Remark 20** The right hand side of (4.4) is what determines the definition of the  $k$ -energy condition (4.2) used in the stopping time arguments adapted from [7] and/or [10].

**Proof** (Proof of Lemma 19) The proof is an easy adaptation of the proofs in [7,10] restricted to dimension  $n = 1$ , but using an order  $k$  Taylor expansion instead of an order 1 expansion on the kernel  $(K_y^\alpha)(x) = K^\alpha(x, y)$ . Due to the importance of this lemma, as explained in the above remark, we give the short argument.

Let  $\{h_J^{\omega,a}\}_{a \in \Gamma}$  be an orthonormal basis of  $L^2_{J;k}(\mu)$  consisting of Alpert functions as above. Next we use the  $(k + \delta)$ -smooth Calderón–Zygmund smoothness estimate (4.1), together with Taylor’s formula

$$\begin{aligned} K_y^\alpha(x) &= T(K_y^\alpha)(x, c) + \frac{1}{k!} (K_y^\alpha)^{(k)}(\theta(x, c))(x - c)^k; \\ \text{Tay}(K_y^\alpha)(x, c) &\equiv K_y^\alpha(c) + (K_y^\alpha)'(c)(x - c) + \dots \\ &\quad + \frac{1}{(k - 1)!} (K_y^\alpha)^{(k-1)}(c)(x - c)^{k-1}, \end{aligned}$$

and the vanishing means of the Alpert functions  $h_J^{\omega,a}$ , to obtain

$$\begin{aligned} \langle T^\alpha \mu, h_J^{\omega,a} \rangle_\omega &= \int \left\{ \int K^\alpha(x, y) h_J^{\omega,a}(x) d\omega(x) \right\} d\mu(y) = \int \langle K_y^\alpha, h_J^{\omega,a} \rangle_\omega d\mu(y) \\ &= \int \langle K_y^\alpha(x) - \text{Tay}(K_y^\alpha)(x, m_J^k), h_J^{\omega,a}(x) \rangle_\omega d\mu(y) \\ &= \int \left\langle \frac{1}{k!} (K_y^\alpha)^{(k)}(\theta(x, m_J^k))(x - m_J^k)^k, h_J^{\omega,a}(x) \right\rangle_\omega d\mu(y) \\ &= \left\langle \left[ \int \frac{1}{k!} (K_y^\alpha)^{(k)}(m_J) d\mu(y) \right] (x - m_J^k)^k, h_J^{\omega,a} \right\rangle_\omega \\ &\quad + \left\langle \left[ \int \frac{1}{k!} \left[ (K_y^\alpha)^{(k)}(\theta(x, m_J^k)) - (K_y^\alpha)^{(k)}(m_J^k) \right] d\mu(y) \right] \right. \\ &\quad \left. (x - m_J^k)^k, h_J^{\omega,a} \right\rangle_\omega \\ &= \left[ \frac{1}{k!} \int (K_y^\alpha)^{(k)}(m_J) d\mu(y) \right] \langle x^k, h_J^{\omega,a} \rangle_\omega \\ &\quad + \left\langle \left[ \int \frac{1}{k!} \left[ (K_y^\alpha)^{(k)}(\theta(x, m_J^k)) - (K_y^\alpha)^{(k)}(m_J^k) \right] d\mu(y) \right] \right. \\ &\quad \left. (x - m_J^k)^k, h_J^{\omega,a} \right\rangle_\omega \end{aligned}$$

and hence

$$\begin{aligned} & \left| \langle T^\alpha \mu, h_J^{\omega, a} \rangle_\omega - \left[ \frac{1}{k!} \int (K_y^\alpha)^{(k)}(m_J) d\mu(y) \right] \langle x^k, h_J^{\omega, a} \rangle_\omega \right| \\ & \leq \frac{1}{k!} \left| \left\langle \left[ \int \sup_{\theta \in J} \left| (K_y^\alpha)^{(k)}(\theta) - (K_y^\alpha)^{(k)}(m_J^k) \right| d\mu(y) \right] |x - m_J^k|^k, |h_J^{\omega, a}\rangle_\omega \right| \right| \\ & \lesssim C_{CZ} \frac{P_{k+\delta}^\alpha(J, |\mu|)}{|J|^k} \left\| (x - m_J^k)^k \right\|_{L^2(\mathbf{1}_J \omega)} \end{aligned}$$

where in the last line we have used

$$\begin{aligned} & \int \sup_{\theta \in J} \left| (K_y^\alpha)^{(k)}(\theta) - (K_y^\alpha)^{(k)}(m_J^k) \right| d\mu(y) \\ & \lesssim C_{CZ} \int \left( \frac{|J|}{|y - c_J|} \right)^\delta \frac{d|\mu|(y)}{|y - c_J|^{k+1-\alpha}} = C_{CZ} \frac{P_{k+\delta}^\alpha(J, \mu)}{|J|^k}. \end{aligned}$$

Thus we have

$$\begin{aligned} \|\Delta_J^\omega T^\alpha \mu\|_{L^2(\omega)}^2 &= \sum_{a \in \Gamma(J)} |\langle T^\alpha \mu, h_J^{\omega, a} \rangle_\omega|^2 \\ &= \left| \frac{1}{k!} \int (K_y^\alpha)^{(k)}(m_J) d\mu(y) \right|^2 \sum_{a \in \Gamma(J)} |\langle x^k, h_J^{\omega, a} \rangle_\omega|^2 \\ &\quad + O\left( \frac{P_{k+\delta}^\alpha(J, \mu)}{|J|^k} \right)^2 \left\| (x - m_J^k)^k \right\|_{L^2(\mathbf{1}_J \omega)}^2, \end{aligned}$$

and hence

$$\begin{aligned} & c_1 \left( \frac{1}{k!} \int (K_y^\alpha)^{(k)}(m_J) d\mu(y) \right)^2 \|\Delta_J^\omega x^k\|_{L^2(\omega)}^2 - C_2 \left( \frac{P_{k+\delta}^\alpha(J, |\mu|)}{|J|^k} \right)^2 \left\| (x - m_J^k)^k \right\|_{L^2(\mathbf{1}_J \omega)}^2 \\ & \leq \|\Delta_J^\omega T^\alpha \mu\|_{L^2(\omega)}^2 \\ & \leq C_1 \left( \frac{1}{k!} \int (K_y^\alpha)^{(k)}(m_J) d\mu(y) \right)^2 \|\Delta_J^\omega x^k\|_{L^2(\omega)}^2 \\ & \quad + C_2 \left( \frac{P_{k+\delta}^\alpha(J, |\mu|)}{|J|^k} \right)^2 \left\| (x - m_J^k)^k \right\|_{L^2(\mathbf{1}_J \omega)}^2, \end{aligned}$$

where

$$\left| \frac{1}{k!} \int (K_y^\alpha)^{(k)}(m_J) d\mu(y) \right| \approx \frac{P_k^\alpha(J, \mu)}{|J|^k}.$$

□

### 4.2 Comparison of $k$ -Energy and the Usual 1-Energy

We can write

$$\left(\frac{P_k^\alpha(J, \mathbf{1}_I\sigma)}{|J|^k}\right)^2 \left\| (x - m_J^k)^k \right\|_{L^2(\mathbf{1}_J\omega)}^2 = P_k^\alpha(J, \mathbf{1}_I\sigma)^2 \left\| \left(\frac{x - m_J^k}{|J|}\right)^k \right\|_{L^2(\mathbf{1}_J\omega)}^2,$$

and clearly we have the inequalities

$$\begin{aligned} P_k^\alpha(J, \mathbf{1}_I\sigma) &= \int_{\mathbb{R}} \frac{|J|^k}{(|J| + |y - c_J|)^{k+1-\alpha}} d\sigma(y) \\ &= \int_{\mathbb{R}} \left(\frac{|J|}{(|J| + |y - c_J|)}\right)^{k-\ell} \frac{|J|^\ell}{(|J| + |y - c_J|)^{\ell+1-\alpha}} d\sigma(y) \\ &\leq \int_{\mathbb{R}} \frac{|J|^\ell}{(|J| + |y - c_J|)^{\ell+1-\alpha}} d\sigma(y) = P_\ell^\alpha(J, \mathbf{1}_I\sigma), \end{aligned}$$

and

$$\begin{aligned} \left\| \left(\frac{x - m_J^k}{|J|}\right)^k \right\|_{L^2(\mathbf{1}_J\omega)}^2 &\leq \left\| \left(\frac{x - m_J^\ell}{|J|}\right)^k \right\|_{L^2(\mathbf{1}_J\omega)}^2 = \int_J \left(\frac{x - m_J^\ell}{|J|}\right)^{2k} d\omega(x) \\ &\leq \int_J \left(\frac{x - m_J^\ell}{|J|}\right)^{2\ell} d\omega(x) = \left\| \left(\frac{x - m_J^\ell}{|J|}\right)^\ell \right\|_{L^2(\mathbf{1}_J\omega)}^2, \end{aligned}$$

for  $1 \leq \ell \leq k$  since  $\left|\frac{x - m_J^\ell}{|J|}\right| \leq 1$ , and as a consequence we obtain

$$\mathcal{E}_{2,k}^\alpha \leq \mathcal{E}_{2,\ell}^\alpha, \quad \text{for } 1 \leq \ell \leq k.$$

### 4.3 An Example with $\mathcal{A}_2^0 < \infty$ , $\mathcal{E}_{2,k}^0 < \infty$ and $\mathcal{E}_{2,1}^0 = \infty$

Define intervals

$$I_j \equiv \left[ \frac{1}{4^j} - \frac{1}{\sqrt{j}4^j}, \frac{1}{4^j} \right], \quad j = 2, 3, 4, \dots$$

and for  $\varepsilon > 0$  define measures

$$d\sigma(y) = \delta_0 + \sum_{j=2}^\infty \frac{4^j}{j^{\frac{1}{2}+\varepsilon}} \mathbf{1}_{I_j}(y) dy,$$

$$d\omega(x) = \sum_{j=2}^{\infty} \frac{j^{\frac{1}{2}}}{4^j} \mathbf{1}_{I_j}(x) dx,$$

so that

$$\begin{aligned} |I_j| &= \frac{1}{\sqrt{j}4^j}, \\ |I_j|_{\sigma} &= \frac{4^j}{j^{\frac{1}{2}+\varepsilon}} \frac{1}{\sqrt{j}4^j} = \frac{1}{j^{1+\varepsilon}}, \\ |I_j|_{\omega} &= \frac{j^{\frac{1}{2}}}{4^j} \frac{1}{\sqrt{j}4^j} = \frac{1}{4^{2j}}. \end{aligned}$$

Then we compute that

$$\begin{aligned} \frac{|I_j|_{\sigma} |I_j|_{\omega}}{|I_j|^2} &= \frac{\frac{1}{j^{1+\varepsilon}} \frac{1}{4^{2j}}}{\left(\frac{1}{\sqrt{j}4^j}\right)^2} = \frac{1}{j^{\varepsilon}}, \\ \frac{\left|[0, \frac{1}{4^j}]\right|_{\sigma} \left|[0, \frac{1}{4^j}]\right|_{\omega}}{\left|[0, \frac{1}{4^j}]\right|^2} &= \frac{\left(1 + \sum_{i=j}^{\infty} |I_i|_{\sigma}\right) \left(\sum_{i=j}^{\infty} \frac{1}{4^{2i}}\right)}{\left(\frac{1}{4^j}\right)^2} \approx 1, \end{aligned}$$

and in fact it can be verified that  $\mathcal{A}_2(\sigma, \omega) \approx 1$ . We also have

$$\left\| \frac{x - m_{I_j}^1}{|I_j|} \right\|_{L^2(\mathbf{1}_{I_j}\omega)}^2 = \int_{I_j} \left( \frac{x - m_{I_j}^1}{|I_j|} \right)^2 d\omega(x) \approx |I_j|_{\omega} = \frac{1}{4^{2j}}$$

and writing  $P = P_I^0$  with  $I = [0, 1]$ , we have

$$\begin{aligned} P(I_j, \mathbf{1}_I\sigma) &= \int_I \frac{|I_j|}{[|I_j| + |y - c_{I_j}|]^2} d\sigma(y) \approx \frac{|I_j|}{[|I_j| + |0 - c_{I_j}|]^2} \\ &\quad + \int_{I_j} \frac{|I_j|}{[|I_j| + |y - c_{I_j}|]^2} d\sigma(y) \\ &\approx \frac{1}{\left[\frac{1}{4^j}\right]^2} + \frac{|I_j|_{\sigma}}{|I_j|} = \frac{4^j}{\sqrt{j}} + \frac{1}{\sqrt{j}4^j} = \frac{4^j}{j^{\frac{1}{2}}} + \frac{4^j}{j^{\frac{1}{2}+\varepsilon}} \approx \frac{4^j}{j^{\frac{1}{2}}}, \end{aligned}$$

where we note that the delta mass  $\delta_0$  in  $\sigma$  contributes the dominant term. From this we compute

$$\mathcal{E}_{2,1}^\alpha \geq \frac{1}{|I|_\sigma} \sum_{j=2}^\infty P(I_j, \mathbf{1}_I \sigma)^2 \left\| \frac{x - m_{I_j}^1}{|I_j|} \right\|_{L^2(\mathbf{1}_{I_j} \omega)}^2 \approx \sum_{j=2}^\infty \left( \frac{4j}{j^{\frac{1}{2}}} \right)^2 \frac{1}{4^{2j}} = \sum_{j=2}^\infty \frac{1}{j} = \infty.$$

On the other hand, writing  $P_k = P_k^0$  we have

$$\begin{aligned} P_k(I_j, \mathbf{1}_I \sigma) &= \int_I \left( \frac{|I_j|}{(|I_j| + |y - c_{I_j}|)} \right)^{k-1} \frac{|I_j|}{[|I_j| + |y - c_{I_j}|]^2} d\sigma(y) \\ &\approx \left( \frac{|I_j|}{(|I_j| + |0 - c_{I_j}|)} \right)^{k-1} \frac{|I_j|}{[|I_j| + |0 - c_{I_j}|]^2} \\ &\quad + \int_{I_j} \frac{|I_j|}{[|I_j| + d(y, I_j)]^2} d\sigma(y) \\ &\approx \left( \frac{\frac{1}{\sqrt{j}4^j}}{\frac{1}{4^j}} \right)^{k-1} \frac{\frac{1}{\sqrt{j}4^j}}{\left[\frac{1}{4^j}\right]^2} + \frac{|I_j|_\sigma}{|I_j|} \approx \frac{4^j}{j^{\frac{k}{2}}} + \frac{1}{j^{1+\varepsilon}} \approx \frac{4^j}{j^{\frac{1}{2}+\varepsilon}}, \end{aligned}$$

where when  $k \geq 2$ , the dominant term arises from that part of  $\sigma$  supported on  $I_j$ , and hence is significantly smaller than  $P(I_j, \mathbf{1}_I \sigma)$ . Then for  $k \geq 2$  we compute

$$\begin{aligned} \frac{1}{|I|_\sigma} \sum_{j=2}^\infty P_k(I_j, \mathbf{1}_I \sigma)^2 \left\| \left( \frac{x - m_{I_j}^k}{|I_j|} \right)^k \right\|_{L^2(\mathbf{1}_{I_j} \omega)}^2 &\lesssim \sum_{j=2}^\infty \left( \frac{4^j}{j^{\frac{1}{2}+\varepsilon}} \right)^2 \frac{1}{4^{2j}} \\ &= \sum_{j=2}^\infty \frac{1}{j^{1+2\varepsilon}} < \infty, \end{aligned}$$

and it can be shown that in fact  $\mathcal{E}_{2,k}^\alpha < \infty$  by considering arbitrary decompositions  $I = \dot{\cup}_{r=1}^\infty I_r$ .

#### 4.4 A Calderón–Zygmund Operator Satisfying Testing Conditions

We do not have an example of a weight pair  $(\sigma, \omega)$ , and a familiar Calderón–Zygmund operator  $T$ , to which Conjecture 17 applies, and to which the known  $T1$ -type theorems fail to apply. Our purpose here is to instead construct a rather artificial example to demonstrate that Conjecture 17 is at least ‘logically’ different from the known  $T1$ -type theorems in [7,10].

For this we first construct a dyadic operator  $T_{(\sigma,\omega)}^{\text{dy}}$  that always satisfies testing conditions for a given weight pair  $(\sigma, \omega)$ . For an arbitrary weight pair  $(\sigma, \omega)$  define

$$T_{(\sigma,\omega)}^{\text{dy}} f \equiv \sum_{I \in \mathcal{D}} \sum_{1 \leq i, j \leq 2} \left\langle f, a_I^{\sigma,i} \right\rangle_{L^2(\sigma)} a_I^{\omega,j},$$

where  $\{a_I^{\sigma,1}, a_I^{\sigma,2}\}_{I \in \mathcal{D}}$  and  $\{a_I^{\omega,1}, a_I^{\omega,2}\}_{I \in \mathcal{D}}$  are Alpert bases for  $L^2(\sigma)$  and  $L^2(\omega)$  respectively. Then for any interval  $Q \in \mathcal{D}$  and polynomial  $p$  of degree at most 1 we have

$$\begin{aligned} \int_Q \left| T_{(\sigma,\omega)}^{\text{dy}} \mathbf{1}_Q p \right|^2 d\omega &= \int_Q \left| \sum_{I \in \mathcal{D}} \sum_{1 \leq i, j \leq 2} \left\langle \mathbf{1}_Q p, a_I^{\sigma,i} \right\rangle_{L^2(\sigma)} a_I^{\omega,j} \right|^2 d\omega \\ &= \sum_{I \in \mathcal{D}} \sum_{1 \leq j \leq 2} \int_Q \left| \sum_{1 \leq i \leq 2} \left\langle \mathbf{1}_Q p, a_I^{\sigma,i} \right\rangle_{L^2(\sigma)} \right|^2 |a_I^{\omega,j}|^2 d\omega \\ &= \sum_{I \in \mathcal{D}: I \not\subseteq Q} \sum_{1 \leq j \leq 2} \int_Q \left| \widehat{\mathbf{1}_Q p}(I) \right|^2 |a_I^{\omega,j}|^2 d\omega \\ &= \sum_{I \in \mathcal{D}: I \not\subseteq Q} \left| \widehat{\mathbf{1}_Q p}(I) \right|^2 \leq \int_Q |p|^2 d\sigma. \end{aligned}$$

Similarly we have the dual testing condition.

To construct a Calderón–Zygmund operator  $T$  that satisfies the testing conditions for  $(\sigma, \omega)$  and whose kernel is  $(k + \delta)$ -smooth as in (4.1) for any  $k + \delta > 1$ , we choose a subgrid  $\mathcal{D}'$  of  $\mathcal{D}$  satisfying:

$$\text{if } I, J \in \mathcal{D}' \text{ satisfy } |I| \leq |J| \text{ and } \frac{11}{9}I \cap \frac{11}{9}J \neq \emptyset, \text{ then } \frac{10}{9}I \subset J, \quad (4.5)$$

and define functions  $a_I^\sigma$  and  $a_I^\omega$  to satisfy the following conditions with  $\mu$  equal  $\sigma$  and  $\omega$ :

$$\begin{aligned} \text{Supp} a_I^\mu &\subset \frac{11}{9}I, \\ 0 &\leq a_I^\mu(x) \leq c |I|^{-\frac{1}{2}} \text{ for } x \in \mathbb{R}, \\ a_I^\mu(x) &= c \text{ for } x \in \frac{10}{9}I, \\ \int |a_I^\mu|^2 d\mu &= 1, \\ \int_{\mathbb{R}} x a_I^\mu(x) d\mu(x) &= \int_{\mathbb{R}} a_I^\mu(x) d\mu(x) = 0, \\ \left| \frac{d^\ell}{dx^\ell} a_I^\mu(x) \right| &\leq C_\ell |I|^{-\frac{1}{2}-\ell}, \quad \ell \geq 0. \end{aligned}$$

Note that these functions are *not* Alpert functions, but that they *do* form an orthonormal collection in  $L^2(\mu)$ . Then we set

$$Tf = \sum_{I \in \mathcal{D}'} \langle f, a_I^\sigma \rangle_{L^2(\sigma)} a_I^\omega$$

and note that  $T$  is a  $(k + \delta)$ -smooth Calderón–Zygmund operator (see below) that satisfies the testing conditions for the weight pair  $(\sigma, \omega)$  by the same argument above that established the testing conditions for the dyadic operator  $T_{(\sigma, \omega)}^{\text{dy}}$ . If we take the weight pair  $(\sigma, \omega)$  constructed above, then Conjecture 17 gives the boundedness of  $T$  from  $L^2(\sigma)$  to  $L^2(\omega)$ .

**Lemma 21** *The operator  $T$  defined above is a  $(k + \delta)$ -smooth 0-fractional standard Calderón–Zygmund operator for all  $k + \delta > 1$ .*

**Proof** Note that the kernel is  $K(x, y) = \sum_{I \in \mathcal{D}'} a_I^\omega(x) a_I^\sigma(y)$ . Assume that  $|x - y| \simeq \frac{11}{9} 2^{-\ell}$  (i.e.  $\frac{11}{9} 2^{-\ell} \lesssim |x - y| \lesssim \frac{11}{9} 2^{-\ell+1}$ ). The sum in the expression of  $K(x, y)$  only contains those terms which correspond to the intervals in  $\mathcal{D}'$  for which  $x, y \in \frac{11}{9} I$ . The smallest such interval has length  $2^{-\ell}$ . Also, for  $m = 0, 1, \dots, \ell - 3$  there are at most two intervals of length  $2^{-\ell+m}$  for which  $a_I^\omega(x) a_I^\sigma(y) \neq 0$ . Thus there are at most  $2(\ell - 2)$  terms in the sum for  $K(x, y)$ .

Also, using the properties of  $a_I^\omega$  and  $a_I^\sigma$  there holds:

$$\begin{aligned} |a_I^\omega(x) a_I^\sigma(y)| &\lesssim |I|^{-\frac{1}{2}} |I|^{-\frac{1}{2}} = |I|^{-1}, \\ \left| \frac{\partial}{\partial x} a_I^\omega(x) a_I^\sigma(y) \right| &\lesssim |I|^{-\frac{1}{2}-1} |I|^{-\frac{1}{2}} = |I|^{-2}, \\ \left| \frac{\partial}{\partial y} a_I^\omega(x) a_I^\sigma(y) \right| &\lesssim |I|^{-\frac{1}{2}} |I|^{-\frac{1}{2}-1} = |I|^{-2}, \\ \left| \nabla^k a_I^\omega(x) a_I^\sigma(y) \right| &\leq C_k |I|^{-1-k}, \end{aligned}$$

for all  $k \geq 1$ , and where  $\nabla = \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$ .

We use these estimates to show that the kernel  $K(x, y)$  is a  $(k + \delta)$ -smooth 0-fractional CZ kernel for all  $k + \delta > 1$ . In fact it is well known (using the mean value theorem) that it suffices to show

$$|\nabla^k K(x, y)| \leq C_k \frac{1}{|x - y|^{k+1}}, \quad \text{for all } k \geq 0.$$

Recalling that  $|x - y| \simeq \frac{11}{9} 2^{-\ell}$  we have for  $k \geq 0$ ,

$$\left| \nabla^k K(x, y) \right| \leq C_k \sum_{I \in \mathcal{D}'} \sum_{i+j=k} \left| \left( \nabla_x^i a_I^\omega(x) \right) \left( \nabla_y^j a_I^\sigma(y) \right) \right| \leq C_k \sum_{I \in \mathcal{D}'} |I|^{-k-1}$$

$$\leq C_k \sum_{m=0}^{\ell-3} 2^{(k+1)(\ell-m)} \approx C_k 2^{\ell(k+1)} \approx C_k \frac{1}{|x-y|^{k+1}}.$$

□

**Remark 22** However, we can write this measure  $\sigma$  as  $\delta_0 + \tilde{\sigma}$  and note that the weight pair  $(\tilde{\sigma}, \omega)$  satisfies the usual energy condition (the 1-energy condition), and so  $T$  is bounded from  $L^2(\tilde{\sigma})$  to  $L^2(\omega)$  by results in either [7] or [10], and hence

$$\begin{aligned} \int |Tf\sigma|^2 d\omega &\lesssim \int |Tf\tilde{\sigma}|^2 d\omega + \int |Tf\delta_0|^2 d\omega \\ &\lesssim \int |f|^2 d\tilde{\sigma} + |f(0)|^2 \int \left| \sum_{I \in \mathcal{D}'} a_I^\sigma(0) a_I^\omega \right|^2 d\omega. \end{aligned}$$

Since  $|f(0)|^2 = \int |f|^2 d\delta_0 \leq \int |f|^2 d\sigma$ , it now remains only to prove that

$$\int \left| \sum_{I \in \mathcal{D}'} a_I^\sigma(0) a_I^\omega \right|^2 d\omega = \sum_{I \in \mathcal{D}'} |a_I^\sigma(0)|^2 \lesssim 1,$$

which in turn holds simply because there is at most one  $I \in \mathcal{D}'$  containing the origin, and for such an interval we have that  $|a_I^\sigma(0)| \lesssim 1$  due to the presence of the unit point mass at the origin, a much better bound than the general bound  $c|I|^{-\frac{1}{2}}$ . Indeed, the only intervals  $I \in \mathcal{D}$  that contain the origin are the intervals  $K_j = [0, 4^{-j})$  for some integer  $j$ , and by (4.5) there is at most one of these in  $\mathcal{D}'$  containing the origin. Thus we see that the Calderón–Zygmund operator  $T$  constructed above is very artificial. Nonetheless, it does show that Conjecture 17 contains boundedness results not included in [7,10], and potentially some not so trivial as that above as well.

**Acknowledgements** E. T. Sawyer research supported in part by a grant from the National Science and Engineering Research Council of Canada. B. D. Wick research supported in part by National Science Foundation DMS Grants # 1560955 and 1800057.

## References

1. Alpert, B.K.: A class of bases in  $L^2$  for the sparse representation of integral operators. *SIAM J. Math. Anal.* **1**, 246–262 (1993)
2. Coifman, R.R., Jones, P.W., Semmes, S.: Two elementary proofs of the  $L^2$  boundedness of Cauchy integrals on Lipschitz curves. *J. A. M. S.* **2**, 553–564 (1989)
3. David, G., Journé, J.-L., Semmes, S.: Opérateurs de Calderón–Zygmund, fonctions para-accrétives et interpolation. *Rev. Mat. Iberoam.* **1**, 1–56 (1985)
4. Donoho, D.L., Dyn, N., Levin, D., Yu, T.P.Y.: Smooth multiwavelet duals of Alpert bases by moment-interpolating refinement. *Appl. Comput. Harmon. Anal.* **9**, 166–203 (2000)
5. Hytönen, T.: The two weight inequality for the Hilbert transform with general measures. *Proc. Lond. Math. Soc.* **3**(3), 483–526 (2018)
6. Lacey, M.T.: Two weight inequality for the Hilbert transform: a real variable characterization, II. *Duke Math. J.* **163**(15), 2821–2840 (2014)



7. Lacey, M.T., Wick, B.D.: Two weight inequalities for Riesz transforms: uniformly full dimension weights. Preprint (2016). [arXiv:1312.6163](https://arxiv.org/abs/1312.6163)
8. Lacey, M.T., Sawyer, E.T., Shen, C.-Y., Uriarte-Tuero, I.: Two weight inequality for the Hilbert transform: a real variable characterization I. *Duke Math. J* **163**(15), 2795–2820 (2014)
9. Nazarov, F., Treil, S., Volberg, A.: Two weight estimate for the Hilbert transform and corona decomposition for non-doubling measures. Preprint (2004). [arXiv:1003.1596](https://arxiv.org/abs/1003.1596)
10. Sawyer, Eric T., Shen, C.-Y., Uriarte-Tuero, I.: A two weight theorem for  $\alpha$ -fractional singular integrals with an energy side condition. *Rev. Mat. Iberoam.* **32**(1), 79–174 (2016)
11. Sawyer, E.T., Shen, C.-Y., Uriarte-Tuero, I.: A two weight fractional singular integral theorem with side conditions, energy and  $k$ -energy dispersed. In: *Harmonic Analysis, Partial Differential Equations, Complex Analysis, Banach Spaces, and Operator Theory*, vol. 2 (Celebrating Cora Sadosky's Life). Springer, Cham (2017) (see also [arXiv:1603.04332v2](https://arxiv.org/abs/1603.04332v2))
12. Sawyer, E.T., Shen, C.-Y., Uriarte-Tuero, I.: A good- $\lambda$  lemma, two weight  $T1$  theorems without weak boundedness, and a two weight accretive global  $Tb$  theorem. In: *Harmonic Analysis, Partial Differential Equations and Applications (In Honor of Richard L. Wheeden)*. Birkhäuser, Boston (2017) (see also [arXiv:1609.08125v2](https://arxiv.org/abs/1609.08125v2))
13. Volberg, A.: Calderón–Zygmund capacities and operators on nonhomogeneous spaces. In: *CBMS Regional Conference Series in Mathematics* (2003). MR2019058 (2005c:42015)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.