Nearly tight framelet systems from nested Marcinkiewicz–Zygmund measures on compact Riemannian manifolds

Hao-Ning Wu† Xiaosheng Zhuang‡

August 21, 2026

Abstract

Tight framelet systems on manifolds provide exact energy preservation and one-pass reconstruction, but their standard semi-discretization relies on polynomial-exact quadrature rules, which are difficult to reconcile with scattered and progressively refined data. We develop nested Marcinkiewicz–Zygmund (MZ) measures on compact Riemannian manifolds as a quantitative alternative. Using dyadic quasi-uniform nested point sets and local partition weights, we establish a sequence of deterministic MZ inequalities for diffusion polynomial spaces. The result has two complementary forms: a fixed-tolerance version, in which the polynomial bandwidth grows dyadically with the level, and a fixed-bandwidth refinement version, in which the MZ tolerance improves as more nested nodes are added. When these measures are used to discretize continuous tight framelets, the MZ tolerance transfers directly to the deviation of the frame bounds from one. Thus quadrature exactness is relaxed with a controlled loss of tightness, and near-tightness improves along the same nested hierarchy. For the fully discrete setting, we develop filter-bank analysis and synthesis procedures, characterize one-pass and canonical reconstruction, and show that the MZ tolerance also controls the conditioning of the frame-operator equation. Fast implementation of filter-bank transforms is also presented. Numerical experiments on the sphere and the flat torus illustrate the resulting multiscale decompositions and reconstruction behavior.

Keywords. Marcinkiewicz–Zygmund; nested measures; nearly tight frames; framelets; filter banks.
AMS subject classifications. 42C15, 42C40, 41A55, 41A10, 41A17, 58C40, 94A12.

1 Introduction

Multiresolution analysis was originally developed largely for data in Euclidean spaces, such as one-dimensional signals, two-dimensional images, and three-dimensional volumetric or video data. Multiscale representation systems that capture sparsity in $\mathbb{R}^d$, including wavelets, framelets, curvelets, and shearlets, have therefore been extensively developed and widely used; see, e.g., [6, 10, 13, 14, 18, 34, 35, 41]. Many modern data sets, however, are more naturally modeled on non-Euclidean domains, such as spherical images, directional data, signals on surfaces, point clouds, and data graphs. In such settings, multiscale methods can separate coarse geometric features from fine local details while respecting the geometry of the underlying domain. This perspective underlies diffusion maps and diffusion wavelets [15, 16], spectral graph wavelets [27], wavelet or frame constructions on compact manifolds and spheres [24, 37, 45], and so on. In this paper, we work on compact Riemannian manifolds, which provide a common spectral setting for these constructions.

When such manifold systems are discretized, the available samples are not always placed on a prescribed quadrature grid. In many practical situations, they may be scattered or added progressively. A multiscale method that can keep the old samples and add new ones at finer levels is then attractive. Previously acquired samples do not have to be discarded, and the same hierarchy of point sets can serve both sampling and filter-bank implementation.

To see where the sampling problem enters, it is useful to recall the continuous framelet system before discretization. On a compact manifold $\mathcal{M}$, let $\{u_\ell, \lambda_\ell\}_{\ell\geq 0}$ be the Laplace–Beltrami eigen-pairs.

†School of Mathematical Sciences, Xiamen University, Xiamen, Fujian, 361005, China (hnwu@connect.hku.hk)
‡Department of Mathematics, City University of Hong Kong, Hong Kong, China (xzhuang7@cityu.edu.hk)

Let $J_0 \in \mathbb{N}_0$ be an integer, and let $\Psi = \{\alpha; \beta^1, \dots, \beta^r\}$ with $r \geq 1$ be a finite collection of framelet generators. From Ψ, one forms the continuous kernels

$$\varphi_{J,y}(x) = \sum_{\ell \geq 0} \widehat{\alpha}\left(\frac{\lambda_\ell}{2^J}\right) \overline{u_\ell(y)} u_\ell(x), \qquad \psi^k_{j,y}(x) = \sum_{\ell \geq 0} \widehat{\beta^k}\left(\frac{\lambda_\ell}{2^j}\right) \overline{u_\ell(y)} u_\ell(x),$$

where $J \geq J_0$ is the coarsest scaling level, $j \geq J$ indexes the finer detail levels with $k = 1, \dots, r$, and $y \in \mathcal{M}$ is the center variable. The continuous framelet system is then given by

$$\mathrm{CFS}_J(\Psi) = \{\varphi_{J,y} : y \in \mathcal{M}\} \cup \{\psi^k_{j,y} : y \in \mathcal{M},\ j \geq J,\ k = 1, \dots, r\}. \tag{1.1}$$

Under the usual spectral partition conditions, $\mathrm{CFS}_J(\Psi)$ is a continuous tight frame. Tightness is particularly desirable. The framelet coefficients preserve the L^2-energy exactly, and reconstruction uses the same system without inverting a frame operator. To obtain a computable system, however, one must replace the continuum of centers $y \in \mathcal{M}$ by finitely many weighted nodes at each scale. The passage from the continuous system to a semi-discrete one therefore raises the central question of this paper: how much of the tight-frame property can be retained when the center variable is sampled on scattered and progressively refined node sets?

1.1 Quadrature rules and nested sampling

In the tight framelet constructions, see, e.g., [45, 49], this discretization of the center variable y is achieved by polynomial-exact quadrature rules. In a typical form, for a finite set $\boldsymbol{X} \subset \mathcal{M}$ and positive weights $\{w_x\}_{x \in \boldsymbol{X}}$, one asks

$$\int_{\mathcal{M}} P(y)\, d\mu(y) = \sum_{x \in \boldsymbol{X}_j} w_{j,x} P(x), \qquad P \in \mathbb{P}_{m_j}, \qquad m_j \asymp 2^j, \tag{1.2}$$

where $\mathbb{P}_m$ denotes the diffusion polynomial space of bandwidth m. In the standard quadrature-based discretization of manifold framelets, the exactness condition (1.2) on the relevant product spaces transfers the continuous tight-frame identity to the semi-discrete system. In particular, it integrates the squared coefficient functions exactly and therefore preserves the Parseval energy identity. On the sphere, for example, a canonical equal-weight realization of such exact rules is given by spherical t-designs [5, 20], which have also been used in spherical framelet constructions [49].

Although powerful, such exact rules are highly structured, whereas data nodes in applications are often scattered. For scattered data, it is natural to replace exact integration by stable discretization of polynomial norms. This leads to Marcinkiewicz–Zygmund inequalities on compact manifolds [23, 36]. An L^p Marcinkiewicz–Zygmund inequality for $\mathbb{P}_n$ has the form

$$A\|P\|^p_{L^p(\mu)} \leq \sum_{x \in \boldsymbol{X}} w_x |P(x)|^p \leq B\|P\|^p_{L^p(\mu)}, \qquad P \in \mathbb{P}_n, \tag{1.3}$$

with constants $0 < A \leq B < \infty$. The Marcinkiewicz–Zygmund inequality is not merely a qualitative stability substitute for exact quadrature. When $A = 1 - \eta$ and $B = 1 + \eta$, it preserves the continuous polynomial norm up to the relative error η. Applied to the center-variable framelet coefficients and combined with the continuous tight-frame identity, this gives a semi-discrete frame whose bounds are $1 - \eta$ and $1 + \eta$. Thus the loss of exactness leads to a quantitatively controlled, rather than arbitrary, loss of tightness. Marcinkiewicz–Zygmund inequalities have long been used in stable sampling for positive quadrature and approximation [7, 17, 19, 22, 25, 32], hyperinterpolation [1, 2, 3, 39, 46], and numerical methods for partial differential equations and integral equations [4, 47, 48]. For a broader discussion of the role and limitations of exactness in the design of quadrature rules, we refer the reader to [43]. Related relaxations of spherical needlet discretizations have also been studied in [8] recently, where the analysis relies on quasi-Monte Carlo designs [9, 40] rather than on Marcinkiewicz–Zygmund inequalities.

The Marcinkiewicz–Zygmund point of view provides a controlled relaxation of exact quadrature rules. For $p = 2$, exact integration of $|P|^2$ corresponds to the special case $A = B = 1$, whereas sufficiently fine Marcinkiewicz–Zygmund measures allow the constants to approach one while accommodating geometrically controlled scattered nodes. Such node sets can be characterized geometrically through their mesh norm and separation in the sense that existing nodes with suitable geometric quality may be retained and assigned appropriate positive weights, rather than replaced by a specially constructed exact quadrature rule.

For multiscale constructions, there is a second requirement: the sampling sets should be compatible across levels. Ideally they are *nested* in the sense of

$$\boldsymbol{X}_j \subset \boldsymbol{X}_{j+1}, \qquad j \geq 0.$$

Nestedness is the spatial analogue of dyadic refinement: a finer level keeps the old centers and adds new ones. This is automatic for some classical one-dimensional rules, such as dyadically refined Chebyshev–Lobatto points [42], but it is much less automatic for discretization rules on manifolds. For spherical t-designs, this issue has recently been addressed in [50, 51] by showing that prescribed point sets can be enlarged to spherical designs of higher degree. These works show that nesting is a genuine and useful constraint in polynomial discretization on manifolds. Our aim is therefore to construct sampling hierarchies that combine three properties: geometrically controlled scattered nodes, nestedness across scales, and Marcinkiewicz–Zygmund constants close to one. The first two properties make the systems compatible with progressively acquired data, while the third retains the tight-frame identity up to a prescribed error. Figure 1 illustrates the kind of nested scattered sampling hierarchy.

(a) $\boldsymbol{X}_2$ with $|\boldsymbol{X}_2| = 75$ **(b)** $\boldsymbol{X}_3$ with $|\boldsymbol{X}_3| = 243$ **(c)** $\boldsymbol{X}_4$ with $|\boldsymbol{X}_4| = 867$

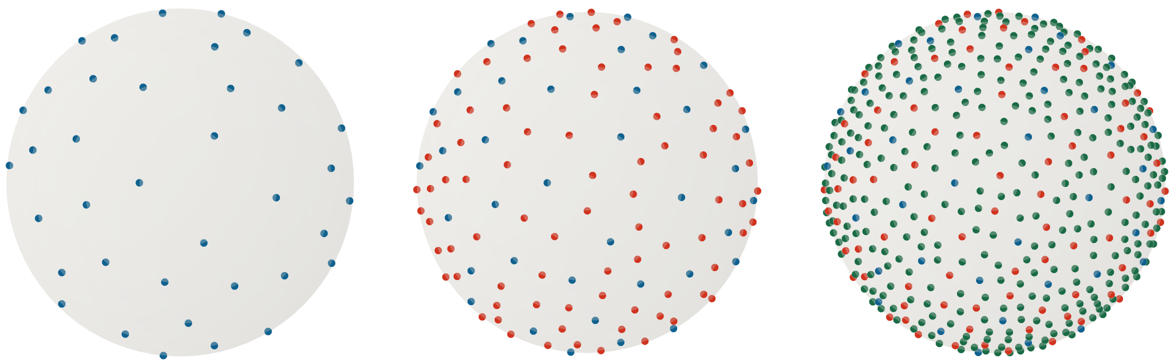

Figure 1: An illustrative nested scattered sampling family on the sphere. Blue points form $\boldsymbol{X}_2$, red points are the new points in $\boldsymbol{X}_3 \setminus \boldsymbol{X}_2$, and green points are the new points in $\boldsymbol{X}_4 \setminus \boldsymbol{X}_3$.

Accordingly, we introduce nested Marcinkiewicz–Zygmund measures. In the dyadic setting, the basic goal is to construct nested sets $\boldsymbol{X}_0 \subset \boldsymbol{X}_1 \subset \cdots$ and positive level-dependent weights such that

$$(1-\eta)\|P\|_{L^p(\mu)}^p \leq \sum_{x \in \boldsymbol{X}_j} w_{j,x}|P(x)|^p \leq (1+\eta)\|P\|_{L^p(\mu)}^p, \qquad P \in \mathbb{P}_{n_j}, \qquad n_j \asymp 2^j, \tag{1.4}$$

with a fixed η independent of j. This is the fixed-tolerance version in the sense that the Marcinkiewicz–Zygmund constants are kept uniformly controlled while the degree increases with the scale. For $p = 2$, this fixed Marcinkiewicz–Zygmund tolerance becomes exactly the near-tightness parameter of the resulting framelet system. We also prove a complementary fixed-degree version. If the degree n is fixed, then the same construction gives

$$(1-\eta_j)\|P\|_{L^p(\mu)}^p \leq \sum_{x \in \boldsymbol{X}_j} w_{j,x}|P(x)|^p \leq (1+\eta_j)\|P\|_{L^p(\mu)}^p, \qquad P \in \mathbb{P}_n, \tag{1.5}$$

where $\eta_j \to 0$ as the nested sets are refined. These two forms reflect two common uses of multiscale sampling: either one lets the approximation space grow with the dyadic scale, or one oversamples a fixed space more and more accurately.

1.2 Nearly tight framelet systems and fully discrete transforms

The second purpose of the paper is to use nested Marcinkiewicz–Zygmund measures satisfying (1.4) or (1.5) in framelet constructions. At the semi-discrete level, the measures discretize the center variable y in the continuous framelet system (1.1). With the same notation as above, the semi-discrete system is

$$\mathrm{FS}_J(\Psi; \boldsymbol{X}, w) = \{w_{J,x}^{1/2}\varphi_{J,x} : x \in \boldsymbol{X}_J\} \cup \{w_{j+1,x}^{1/2}\psi_{j,x}^k : x \in \boldsymbol{X}_{j+1},\ j \geq J,\ k = 1, \ldots, r\},$$

where $J \geq J_0$ is the coarsest scaling level, and $j \geq J$ indexes the finer detail levels with $k = 1, \ldots, r$. In the quadrature-based setting, exactness is what preserves the tight energy identity after discretization. With Marcinkiewicz–Zygmund sampling, exact tightness is replaced by quantitatively controlled near-tightness. A level-independent Marcinkiewicz–Zygmund tolerance η produces frame bounds $1-\eta$ and $1+\eta$. Moreover, sampling the same coefficient functions on s additional nested refinement levels reduces the deviation to $C_{\mathcal{M}} 2^{-s}$. Hence the framelet systems become asymptotically tight along the nested hierarchy.

At the fully discrete level, the same nested measures store the scaling and detail coefficient sequences. A filter bank

$$\Theta = \{a; b^1, \ldots, b^r\}$$

with low-pass filter a and high-pass filters $b^1, \ldots, b^r$ is linked to the generators Ψ by the refinement equations

$$\widehat{\alpha}(2\xi) = \widehat{a}(\xi)\widehat{\alpha}(\xi), \qquad \widehat{\beta^k}(2\xi) = \widehat{b^k}(\xi)\widehat{\alpha}(\xi).$$

On a finite range $J \leq j \leq K$, after the finest weighted scaling sequence $\mathbf{v}_K$ has been computed on $\boldsymbol{X}_K$, the analysis transform is obtained by repeated one-level filter-bank steps

$$\mathbf{v}_{j-1} = (\mathbf{v}_j *_j a^*) \downarrow_j, \qquad \mathbf{w}_{j-1}^k = \mathbf{v}_j *_j (b^k)^*, \quad k = 1, \ldots, r.$$

Here $*_j$ denotes the level-j spectral convolution and $\downarrow_j$ is the downsampling operator, implemented by evaluating the resulting bandlimited function on the coarser weighted node set. The corresponding direct synthesis, which we call one-pass synthesis, applies the adjoint filter-bank steps to the computed framelet coefficients without solving an additional equation. For a tight system, the frame operator is the identity, so one-pass synthesis recovers the original data exactly. For a Marcinkiewicz–Zygmund-based system, however, the frame bounds generally differ, and one-pass synthesis instead applies the finite-level frame operator. Near-tightness quantifies this departure from exact reconstruction: the frame operator remains close to the identity, while the associated frame-operator equation remains well conditioned. Canonical reconstruction is obtained by inverting this operator, which we implement numerically using conjugate gradients.

1.3 Contributions

The main contributions of the paper can be summarized as follows.

- **(i)** We establish nested Marcinkiewicz–Zygmund measures on compact manifolds and give a deterministic dyadic construction of such measures. The construction produces nested point sets with controlled mesh norm, separation radius, and optimal cardinality order. Combined with the discretization theorem of Filbir and Mhaskar [23], it yields nested deterministic L^p Marcinkiewicz–Zygmund inequalities for $1 \leq p < \infty$, in two regimes: a fixed-tolerance regime in which $n_j \asymp 2^j$, and a fixed-degree refinement regime in which the Marcinkiewicz–Zygmund tolerance tends to zero as the nested sets are refined.
- **(ii)** We use nested Marcinkiewicz–Zygmund measures to construct nearly tight semi-discrete framelet systems from scattered sampling nodes. For any prescribed $0 < \eta < 1$, the basic construction has frame bounds $1-\eta$ and $1+\eta$. For a fixed nested hierarchy, an oversampling shift s improves the bounds to $1 - C_{\mathcal{M}} 2^{-s}$ and $1 + C_{\mathcal{M}} 2^{-s}$. Thus near-tightness can be improved without relocating or discarding previously used nodes.
- **(iii)** We formulate the fully discrete filter-bank transform on the same nested measures. The Marcinkiewicz–Zygmund bounds control both the deviation of one-pass synthesis from the identity and the condition number of the canonical frame-operator equation. We derive reconstruction and approximation estimates, develop the corresponding conjugate-gradient reconstruction, and record fast implementation and complexity estimates.

1.4 Organization

The rest of the paper is organized as follows. Section 2 introduces the notation and the nested Marcinkiewicz–Zygmund definitions. Section 3 gives the deterministic construction and proves the corresponding L^p inequalities. Section 4 constructs framelets from nested Marcinkiewicz–Zygmund measures and proves the semi-discrete frame bounds. Section 5 develops the fully discrete filter-bank transform, coefficient-space reconstruction, and fast implementation. Section 6 derives the corresponding polynomial reconstruction and approximation estimates. Section 7 presents numerical experiments on the sphere and the torus, and Section 8 concludes the paper.

2 Notation and definitions

For nonnegative quantities A and B, we write $A \lesssim B$ if $A \leq CB$ for some constant $C > 0$, $A \gtrsim B$ if $B \lesssim A$, and $A \asymp B$ if both hold. Implicit constants are independent of the varying parameters but may depend on the fixed manifold and other fixed parameters.

Let $(\mathcal{M}, g)$ be a compact, connected, d-dimensional smooth Riemannian manifold without boundary, with Riemannian metric g. Let μ be the normalized Riemannian volume measure on $\mathcal{M}$ with $\mu(\mathcal{M}) = 1$. We denote by dist the geodesic distance on $\mathcal{M}$, that is,

$$\operatorname{dist}(x, y) := \inf_{\omega(0)=x,\ \omega(1)=y} \int_0^1 |\dot{\omega}(t)|_g \, dt, \qquad x, y \in \mathcal{M},$$

where the infimum is taken over all piecewise smooth curves $\omega : [0, 1] \to \mathcal{M}$ joining x and y. For $R > 0$, let $B(x, R) := \{y \in \mathcal{M} : \operatorname{dist}(x, y) \leq R\}$ be the closed geodesic ball.

We work with diffusion polynomials associated with the Laplace–Beltrami operator $\Delta_{\mathcal{M}}$ on $\mathcal{M}$. Let $\{u_\ell\}_{\ell \geq 0}$ be an orthonormal basis of $L^2(\mu)$ consisting of eigenfunctions of $-\Delta_{\mathcal{M}}$ in the sense of

$$-\Delta_{\mathcal{M}} u_\ell = \lambda_\ell^2 u_\ell.$$

The nondecreasing spectral parameters satisfy $0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots$ with $\lambda_\ell \to \infty$. For $f \in L^2(\mu)$, its generalized Fourier coefficients are

$$\widehat{f}_\ell := \langle f, u_\ell \rangle_{L^2(\mu)} = \int_{\mathcal{M}} f(x) \overline{u_\ell(x)} \, d\mu(x), \qquad \ell \geq 0.$$

Then

$$f = \sum_{\ell=0}^{\infty} \widehat{f}_\ell u_\ell \quad \text{in } L^2(\mu), \qquad \|f\|_{L^2(\mu)}^2 = \sum_{\ell=0}^{\infty} |\widehat{f}_\ell|^2.$$

For $n \geq 1$, the diffusion polynomial space of bandwidth n is

$$\mathbb{P}_n := \operatorname{span}\{u_\ell : \lambda_\ell \leq n\}.$$

On the sphere this is equivalent, up to the usual change of bandwidth parameter, to the spherical polynomial spaces generated by spherical harmonics of bounded degree. Indeed, if ℓ denotes the spherical harmonic degree, then the Laplace–Beltrami spectral parameter is $\sqrt{\ell(\ell + d - 1)}$. Thus statements formulated with the harmonic degree or with λ_ℓ differ only by harmless constants in the bandwidth. In this compact manifold setting the standard Weyl-type growth gives $\dim \mathbb{P}_n \asymp n^d$.

We begin with the single-level notion. A Marcinkiewicz–Zygmund measure for $\mathbb{P}_n$ is a measure whose L^p norm on $\mathbb{P}_n$ is equivalent to the ambient manifold L^p norm.

Definition 2.1 (Marcinkiewicz–Zygmund measure)**.** *Let $1 \leq p < \infty$, $n \in \mathbb{N}$, and let ν be a positive finite Borel measure on $\mathcal{M}$. We say that ν is an L^p* Marcinkiewicz–Zygmund measure *for $\mathbb{P}_n$ with constants $A, B > 0$ if*

$$A\|P\|_{L^p(\mu)}^p \leq \int_{\mathcal{M}} |P(x)|^p \, d\nu(x) \leq B\|P\|_{L^p(\mu)}^p, \qquad P \in \mathbb{P}_n.$$

If ν is discrete, say $\nu = \sum_{x \in \boldsymbol{X}} w_x \delta_x$, then the condition becomes

$$A\|P\|_{L^p(\mu)}^p \leq \sum_{x \in \boldsymbol{X}} w_x |P(x)|^p \leq B\|P\|_{L^p(\mu)}^p, \qquad P \in \mathbb{P}_n.$$

The nested object is obtained by requiring such inequalities at a sequence of polynomial degrees, while also requiring the supporting node sets to be nested. The constants may either be uniform in the level or improve along a refinement sequence, depending on the intended application.

Definition 2.2 (Nested Marcinkiewicz–Zygmund measures). *Let $\{n_j\}_{j\geq 0}$ be an increasing sequence of positive integers. A sequence of discrete measures $\nu_j = \sum_{x\in \boldsymbol{X}_j} w_{j,x}\delta_x$ with $j \geq 0$ is called* nested L^p Marcinkiewicz–Zygmund measures *for $\{\mathbb{P}_{n_j}\}_{j\geq 0}$ if*

$$\boldsymbol{X}_j \subset \boldsymbol{X}_{j+1}, \qquad j \geq 0,$$

and there exist constants $A_j, B_j > 0$ such that

$$A_j\|P\|^p_{L^p(\mu)} \leq \sum_{x\in \boldsymbol{X}_j} w_{j,x}|P(x)|^p \leq B_j\|P\|^p_{L^p(\mu)}, \qquad P \in \mathbb{P}_{n_j}, \quad j \geq 0.$$

We shall use standard geometric parameters of finite point sets on $\mathcal{M}$. These parameters provide a convenient deterministic route to Marcinkiewicz–Zygmund inequalities. For a finite set $\boldsymbol{X} \subset \mathcal{M}$, we define its *mesh norm* and *separation distance* by

$$h_{\boldsymbol{X}} := \sup_{y\in\mathcal{M}} \min_{x\in\boldsymbol{X}} \operatorname{dist}(y,x) \qquad \text{and} \qquad \operatorname{sep}(\boldsymbol{X}) := \min_{x,y\in\boldsymbol{X},\ x\neq y} \operatorname{dist}(x,y),$$

respectively. The mesh norm measures how well $\boldsymbol{X}$ covers the manifold: every point of $\mathcal{M}$ lies within distance $h_{\boldsymbol{X}}$ of some node of $\boldsymbol{X}$. The separation distance measures how far distinct nodes are from each other, and rules out excessive clustering when it is bounded from below.

For multiscale applications the most relevant spatial scale is dyadic. We therefore emphasize the dyadic geometry of the nodes. These node sets are nested, cover $\mathcal{M}$ at dyadic resolution, and remain uniformly separated at the same scale.

Definition 2.3 (Dyadic nested sampling family). *A sequence of finite point sets $\{\boldsymbol{X}_j\}_{j\geq 0}$ on $\mathcal{M}$ is called a* dyadic nested sampling family *if*

$$\boldsymbol{X}_j \subset \boldsymbol{X}_{j+1}, \qquad j \geq 0,$$

and there exist constants $c, C > 0$, independent of j, such that

$$h_{\boldsymbol{X}_j} \leq C2^{-j}, \qquad \operatorname{sep}(\boldsymbol{X}_j) \geq c2^{-j}, \qquad |\boldsymbol{X}_j| \asymp 2^{jd}.$$

Since $\mathcal{M}$ is compact, its injectivity radius is positive and its curvature is bounded. Hence there exist constants $R_{\mathcal{M}} > 0$ and $c_1, c_2 > 0$, depending only on $\mathcal{M}$, such that

$$c_1R^d \leq \mu(B(x,R)) \leq c_2R^d, \qquad x \in \mathcal{M}, \quad 0 < R \leq R_{\mathcal{M}}. \tag{2.1}$$

This standard local volume estimate on compact Riemannian manifolds can be found in, e.g., [11, Chapter III, Sections 3–4].

3 Deterministic dyadic nested sampling families

In this section, we construct dyadic nested point sets with controlled geometry and then establish nested Marcinkiewicz–Zygmund measures.

3.1 Dyadic nested quasi-uniform sets

Let $0 < \gamma < R_{\mathcal{M}}$, where $R_{\mathcal{M}}$ is the constant in the estimate (2.1), and set $n_j := 2^j$ and $R_j := \gamma n_j^{-1}$ for $j \geq 0$.

Lemma 3.1 (Nested quasi-uniform sets). *There exists a sequence of finite subsets $\{\boldsymbol{X}_j\}_{j\geq 0}$ of $\mathcal{M}$ such that, for any $j \geq 0$,*

$$\boldsymbol{X}_j \subset \boldsymbol{X}_{j+1},$$

and

$$h_{\boldsymbol{X}_j} \leq R_j, \qquad \operatorname{sep}(\boldsymbol{X}_j) \geq R_j, \qquad |\boldsymbol{X}_j| \asymp R_j^{-d} \asymp \gamma^{-d}2^{jd}.$$

The constants in $\asymp$ depend only on $\mathcal{M}$.

Proof. We construct the sets inductively. A point set $\boldsymbol{Y} \subset \mathcal{M}$ is called maximal R-separated if it is R-separated and no point of $\mathcal{M} \setminus \boldsymbol{Y}$ can be added while preserving R-separation. Choose $\boldsymbol{X}_0$ to be a maximal R_0-separated subset of $\mathcal{M}$. Then $\operatorname{sep}(\boldsymbol{X}_0) \geq R_0$. Moreover, maximality implies the covering property: if some $y \in \mathcal{M}$ satisfied $\operatorname{dist}(y, x) > R_0$ for all $x \in \boldsymbol{X}_0$, then $\boldsymbol{X}_0 \cup \{y\}$ would still be R_0-separated, a contradiction. Hence $h_{\boldsymbol{X}_0} \leq R_0$. Now assume that $\boldsymbol{X}_j$ has been constructed. Since $R_{j+1} = R_j/2$ and $\operatorname{sep}(\boldsymbol{X}_j) \geq R_j$, the old set $\boldsymbol{X}_j$ is R_{j+1}-separated. Enlarge $\boldsymbol{X}_j$, without removing any old point, to a maximal R_{j+1}-separated set and call the result $\boldsymbol{X}_{j+1}$. Compactness ensures that this enlargement is finite. By construction,

$$\boldsymbol{X}_j \subset \boldsymbol{X}_{j+1}, \qquad \operatorname{sep}(\boldsymbol{X}_{j+1}) \geq R_{j+1}.$$

The same maximality argument gives $h_{\boldsymbol{X}_{j+1}} \leq R_{j+1}$. This proves the nestedness, mesh norm, and separation estimates.

It remains only to estimate the cardinality. Since $R_j = \gamma 2^{-j} \leq \gamma < R_{\mathcal{M}}$, the local volume estimate (2.1) gives $\mu(B(x, R)) \asymp R^d$ for $0 < R \leq R_{\mathcal{M}}$, with constants depending only on $\mathcal{M}$. Since $\boldsymbol{X}_j$ is R_j-separated, the balls $B(x, R_j/3)$, $x \in \boldsymbol{X}_j$, are pairwise disjoint. Therefore

$$|\boldsymbol{X}_j| R_j^d \lesssim \sum_{x \in \boldsymbol{X}_j} \mu(B(x, R_j/3)) \leq 1,$$

and so $|\boldsymbol{X}_j| \lesssim R_j^{-d}$. Conversely, since $h_{\boldsymbol{X}_j} \leq R_j$, the balls $B(x, R_j)$, $x \in \boldsymbol{X}_j$, cover $\mathcal{M}$. Thus

$$1 \leq \sum_{x \in \boldsymbol{X}_j} \mu(B(x, R_j)) \lesssim |\boldsymbol{X}_j| R_j^d,$$

which gives $|\boldsymbol{X}_j| \gtrsim R_j^{-d}$. Combining the two bounds yields $|\boldsymbol{X}_j| \asymp R_j^{-d} \asymp \gamma^{-d} 2^{jd}$. $\square$

3.2 Dyadic partitions and weights

Geodesic Voronoi diagrams on Riemannian manifolds are standard; see, e.g., [33]. For the weighted discretization below, we use a measurable version obtained by assigning points on Voronoi boundaries according to a fixed ordering of the nodes.

Definition 3.1 (Partition weights)**.** *Let $\boldsymbol{X}_j \subset \mathcal{M}$ be a finite point set. A measurable partition $\{Y_{j,x} : x \in \boldsymbol{X}_j\}$ of $\mathcal{M}$ is called a* nearest-neighbor partition associated with $\boldsymbol{X}_j$ *if*

$$\mathcal{M} = \bigcup_{x \in \boldsymbol{X}_j} Y_{j,x}, \qquad x \in Y_{j,x},$$

and

$$Y_{j,x} \subset \left\{ y \in \mathcal{M} : \operatorname{dist}(y, x) = \min_{z \in \boldsymbol{X}_j} \operatorname{dist}(y, z) \right\}, \qquad Y_{j,x} \subset B(x, h_{\boldsymbol{X}_j}),$$

where points tied between several nearest nodes are assigned to the first such node in a fixed ordering of $\boldsymbol{X}_j$. This produces a Borel partition because the geodesic distance is continuous and $\boldsymbol{X}_j$ is finite. The corresponding partition weights *are*

$$w_{j,x} := \mu(Y_{j,x}), \qquad x \in \boldsymbol{X}_j.$$

The next lemma shows that, for the dyadic nested sampling family constructed above, these geometrically defined weights have the expected volume scale at every level.

Lemma 3.2 (Dyadic weights)**.** *Let $\{\boldsymbol{X}_j\}_{j \geq 0}$ be the dyadic nested sampling family constructed in Lemma 3.1, and equip each $\boldsymbol{X}_j$ with the partition weights from Definition 3.1. Then*

$$w_{j,x} \asymp R_j^d \asymp \gamma^d 2^{-jd}, \qquad x \in \boldsymbol{X}_j, \quad j \geq 0.$$

The constants depend only on $\mathcal{M}$.

Proof. Fix $j \geq 0$ and $x \in \boldsymbol{X}_j$. The upper bound uses only the mesh norm. By construction of the nearest-neighbor partition, $Y_{j,x} \subset B(x, h_{\boldsymbol{X}_j})$. Lemma 3.1 gives $h_{\boldsymbol{X}_j} \leq R_j$, and hence $Y_{j,x} \subset B(x, R_j)$. The local volume estimate (2.1) therefore gives

$$w_{j,x} = \mu(Y_{j,x}) \leq \mu(B(x, R_j)) \lesssim R_j^d.$$

For the lower bound we use the separation. Let $y \in \mathcal{M}$ satisfy $\operatorname{dist}(y, x) < R_j/2$, and let $z \in \boldsymbol{X}_j$ with $z \neq x$. Since $\operatorname{sep}(\boldsymbol{X}_j) \geq R_j$, we have $\operatorname{dist}(x, z) \geq R_j$. By the triangle inequality,

$$\operatorname{dist}(y, z) \geq \operatorname{dist}(x, z) - \operatorname{dist}(y, x) > R_j - R_j/2 = R_j/2.$$

On the other hand, $\operatorname{dist}(y, x) < R_j/2$. Thus x is the unique nearest node to every point in the open ball $B^\circ(x, R_j/2) := \{y : \operatorname{dist}(y, x) < R_j/2\}$. Hence

$$w_{j,x} \geq \mu(B^\circ(x, R_j/2)) \geq \mu(B(x, R_j/3)) \gtrsim R_j^d,$$

where the last inequality follows from (2.1). Combining the two estimates gives $w_{j,x} \asymp R_j^d$. Since $R_j = \gamma 2^{-j}$, this is the same as $w_{j,x} \asymp \gamma^d 2^{-jd}$. □

3.3 Deterministic Marcinkiewicz–Zygmund inequalities

We now recall the form of the Filbir–Mhaskar discretization theorem [23, Theorem 5.1(b)] on compact manifolds. The statement is specialized to the nearest-neighbor partitions used in this paper.

Proposition 3.1 (Filbir–Mhaskar [23])**.** *Let $1 \leq p < \infty$, $0 < \eta < 1$, $\kappa \geq 1$, and $n \geq 1$. There exists a constant $c = c(\mathcal{M}, \kappa, p) > 0$ with the following property. Let $\boldsymbol{X} \subset \mathcal{M}$ be a finite set satisfying the quasi-uniformity condition $h_{\boldsymbol{X}} \leq \kappa \operatorname{sep}(\boldsymbol{X})$, and $\{Y_x\}_{x \in \boldsymbol{X}}$ a measurable partition of $\mathcal{M}$ such that*

$$x \in Y_x, \qquad Y_x \subset B(x, h_{\boldsymbol{X}}), \qquad x \in \boldsymbol{X}.$$

If $n h_{\boldsymbol{X}} \leq c\,\eta/p$, then, for every real- or complex-valued $P \in \mathbb{P}_n$,

$$\left| \int_{\mathcal{M}} |P(y)|^p \, d\mu(y) - \sum_{x \in \boldsymbol{X}} \mu(Y_x) |P(x)|^p \right| \leq \eta \int_{\mathcal{M}} |P(y)|^p \, d\mu(y).$$

The condition in Proposition 3.1 involves three quantities: the polynomial degree n, the mesh norm $h_{\boldsymbol{X}}$, and the Marcinkiewicz–Zygmund tolerance η. In a nested construction, the mesh norm changes from level to level. This gives two natural ways to read the result. The first one fixes a tolerance η and chooses the degree at level j so that $n_j h_{\boldsymbol{X}_j}$ stays uniformly small. The second one fixes a degree n and lets the tolerance decrease as the nested sets become finer.

Theorem 3.1 (Fixed-tolerance dyadic nested Marcinkiewicz–Zygmund measures)**.** *Let $1 \leq p < \infty$ and $0 < \eta < 1$. There exist $\gamma_{\eta,p,\mathcal{M}} > 0$, dyadic nested point sets $\boldsymbol{X}_j \subset \boldsymbol{X}_{j+1}$, $n_j = 2^j$, and positive weights $\{w_{j,x}\}_{x \in \boldsymbol{X}_j}$ such that*

$$|\boldsymbol{X}_j| \asymp \gamma_{\eta,p,\mathcal{M}}^{-d} n_j^d, \qquad w_{j,x} \asymp \gamma_{\eta,p,\mathcal{M}}^d n_j^{-d}, \qquad x \in \boldsymbol{X}_j, \quad j \geq 0,$$

and

$$(1 - \eta) \|P\|_{L^p(\mu)}^p \leq \sum_{x \in \boldsymbol{X}_j} w_{j,x} |P(x)|^p \leq (1 + \eta) \|P\|_{L^p(\mu)}^p$$

for every $P \in \mathbb{P}_{n_j}$ and every $j \geq 0$. The parameter $\gamma_{\eta,p,\mathcal{M}}$ may be chosen so that

$$0 < \gamma_{\eta,p,\mathcal{M}} < \min\{R_{\mathcal{M}}, c\,\eta/p\}, \tag{3.1}$$

where c is the constant in Proposition 3.1 for $\kappa = 1$. The implicit constants in the cardinality and weight estimates depend only on $\mathcal{M}$; the dependence on η and p is carried by the explicit factor $\gamma_{\eta,p,\mathcal{M}}$.

Proof. Let c be the constant in Proposition 3.1 with $\kappa = 1$. Choose $\gamma_{\eta,p,\mathcal{M}} > 0$ so small that (3.1) holds. Construct $\boldsymbol{X}_j$ with $R_j = \gamma_{\eta,p,\mathcal{M}} n_j^{-1}$ and define $w_{j,x} = \mu(Y_{j,x})$ as above. For each j, Lemma 3.1 gives $h_{\boldsymbol{X}_j} \leq R_j$ and $\operatorname{sep}(\boldsymbol{X}_j) \geq R_j$. Hence $h_{\boldsymbol{X}_j} \leq \operatorname{sep}(\boldsymbol{X}_j)$, so the quasi-uniformity condition of Proposition 3.1 holds with $\kappa = 1$. Moreover, $h_{\boldsymbol{X}_j} \leq R_j = \gamma_{\eta,p,\mathcal{M}} n_j^{-1}$, and therefore $n_j h_{\boldsymbol{X}_j} \leq \gamma_{\eta,p,\mathcal{M}} \leq c\eta/p$. Proposition 3.1 therefore gives, for every $P \in \mathbb{P}_{n_j}$,

$$\left| \int_{\mathcal{M}} |P(y)|^p \, d\mu(y) - \sum_{x \in \boldsymbol{X}_j} \mu(Y_{j,x}) |P(x)|^p \right| \leq \eta \int_{\mathcal{M}} |P(y)|^p \, d\mu(y).$$

This is exactly the stated Marcinkiewicz–Zygmund inequality. The cardinality and weight estimates follow from Lemmas 3.1 and 3.2, since $R_j = \gamma_{\eta,p,\mathcal{M}} n_j^{-1}$. □

Theorem 3.2 (Fixed-degree dyadic nested Marcinkiewicz–Zygmund measures). *Let $1 \le p < \infty$ and let $n \ge 1$ be fixed. There exist dyadic nested point sets $\boldsymbol{X}_j \subset \boldsymbol{X}_{j+1}$ and positive weights $\{w_{j,x}\}_{x \in \boldsymbol{X}_j}$, constructed with a fixed parameter $\gamma = \gamma(\mathcal{M})$, such that*

$$|\boldsymbol{X}_j| \asymp \gamma^{-d} 2^{jd}, \qquad w_{j,x} \asymp \gamma^d 2^{-jd}, \qquad x \in \boldsymbol{X}_j, \quad j \ge 0,$$

and, for all sufficiently large j,

$$(1-\eta_j)\|P\|^p_{L^p(\mu)} \le \sum_{x \in \boldsymbol{X}_j} w_{j,x} |P(x)|^p \le (1+\eta_j)\|P\|^p_{L^p(\mu)}, \qquad P \in \mathbb{P}_n,$$

where $\eta_j \le C_{\mathcal{M},p}\, n\, 2^{-j}$. In particular, the Marcinkiewicz–Zygmund constants tend to 1 as $j \to \infty$.

Proof. Let c be the constant in Proposition 3.1 with $\kappa = 1$. Fix once and for all a number $0 < \gamma < R_{\mathcal{M}}$, for instance $\gamma = R_{\mathcal{M}}/2$. Construct the dyadic nested sets with $R_j = \gamma 2^{-j}$, and define $w_{j,x} = \mu(Y_{j,x})$ as in Lemma 3.2. For each j, Lemma 3.1 gives $h_{\boldsymbol{X}_j} \le R_j = \gamma 2^{-j}$ and $h_{\boldsymbol{X}_j} \le \operatorname{sep}(\boldsymbol{X}_j)$. Define

$$\eta_j := \frac{p}{c}\, n\, h_{\boldsymbol{X}_j}.$$

Since $h_{\boldsymbol{X}_j} \le \gamma 2^{-j}$, we have

$$\eta_j \le \frac{p\gamma}{c}\, n\, 2^{-j}.$$

Because γ depends only on $\mathcal{M}$, the constant $p\gamma/c$ may be absorbed into $C_{\mathcal{M},p}$. This also shows that $\eta_j \to 0$ as $j \to \infty$, and hence $\eta_j < 1$ for all sufficiently large j. For such j, $n h_{\boldsymbol{X}_j} = c\,\eta_j/p$, so Proposition 3.1 applies to $\mathbb{P}_n$ with tolerance η_j and gives the desired Marcinkiewicz–Zygmund inequality. The cardinality and weight estimates follow from Lemmas 3.1 and 3.2, since $R_j = \gamma 2^{-j}$. □

4 Nearly tight framelet systems from nested Marcinkiewicz–Zygmund measures

We now turn the nested sampling inequalities into multiscale systems. The continuous framelet construction provides spectral building blocks at dyadic scales, with a center variable y ranging over the whole manifold $\mathcal{M}$. To obtain a usable semi-discrete system, this center variable must be sampled at each scale.

4.1 Continuous framelets on compact manifolds

We first recall the continuous framelet construction on a compact manifold, following the spectral approach in [45]. Let $\Psi := \{\alpha; \beta^1, \dots, \beta^r\}$ be a finite collection of *framelet generators* in $L_1(\mathbb{R})$ whose Fourier transforms are compactly supported on the nonnegative frequency axis. Throughout this section we use the 2π-Fourier convention $\widehat{h}(\xi) := \int_{\mathbb{R}} h(t) e^{-2\pi i t\xi}\, dt$. We fix an arbitrary starting scale $J \in \mathbb{N}_0$, and define, for $j \ge J$ and $y \in \mathcal{M}$, the low- and high-pass continuous framelet kernels by

$$\varphi_{j,y}(x) := \sum_{\ell=0}^{\infty} \widehat{\alpha}\left(\frac{\lambda_\ell}{2^j}\right) \overline{u_\ell(y)}\, u_\ell(x), \qquad x \in \mathcal{M}, \tag{4.1}$$

and

$$\psi^k_{j,y}(x) := \sum_{\ell=0}^{\infty} \widehat{\beta^k}\left(\frac{\lambda_\ell}{2^j}\right) \overline{u_\ell(y)}\, u_\ell(x), \qquad x \in \mathcal{M}, \quad k = 1, \dots, r, \tag{4.2}$$

respectively. The variable $y \in \mathcal{M}$ plays the role of the translation parameter.

Definition 4.1 (Continuous framelet system). *Let $J_0 \in \mathbb{N}_0$ be an integer. For any $J \ge J_0$, the family*

$$\operatorname{CFS}_J(\Psi) := \{\varphi_{J,y} : y \in \mathcal{M}\} \cup \{\psi^k_{j,y} : y \in \mathcal{M},\ j \ge J,\ k = 1, \dots, r\}$$

is called the continuous framelet system *generated by Ψ from scale J.*

The system $\mathrm{CFS}_J(\Psi)$ is called a continuous *tight* frame for $L^2(\mathcal{M},\mu)$ if its elements belong to $L^2(\mathcal{M},\mu)$ and if, for every $f \in L^2(\mathcal{M},\mu)$,

$$\|f\|^2_{L^2(\mu)} = \int_{\mathcal{M}} |\langle f, \varphi_{J,y}\rangle|^2 \, d\mu(y) + \sum_{j=J}^{\infty} \sum_{k=1}^{r} \int_{\mathcal{M}} |\langle f, \psi^k_{j,y}\rangle|^2 \, d\mu(y). \tag{4.3}$$

The tightness condition (4.3) can be verified in a spectral manner. Assume that, for every $\ell \geq 0$,

$$\lim_{j\to\infty} \left| \widehat{\alpha}\left(\frac{\lambda_\ell}{2^j}\right)\right| = 1, \tag{4.4}$$

and that, for every $j \geq J$ and every $\ell \geq 0$,

$$\left| \widehat{\alpha}\left(\frac{\lambda_\ell}{2^{j+1}}\right)\right|^2 = \left| \widehat{\alpha}\left(\frac{\lambda_\ell}{2^j}\right)\right|^2 + \sum_{k=1}^{r} \left| \widehat{\beta^k}\left(\frac{\lambda_\ell}{2^j}\right)\right|^2. \tag{4.5}$$

Then the identities (4.5) telescope to

$$\left| \widehat{\alpha}\left(\frac{\lambda_\ell}{2^J}\right)\right|^2 + \sum_{j=J}^{\infty} \sum_{k=1}^{r} \left| \widehat{\beta^k}\left(\frac{\lambda_\ell}{2^j}\right)\right|^2 = 1, \qquad \ell \geq 0. \tag{4.6}$$

Consequently, $\mathrm{CFS}_J(\Psi)$ is a continuous tight frame. To see this, first let f be a finite spectral expansion, $f = \sum_{\ell\geq 0} \widehat{f}_\ell u_\ell$. Then

$$\langle f, \varphi_{j,y}\rangle = \sum_{\ell\geq 0} \widehat{f}_\ell \, \overline{\widehat{\alpha}\left(\frac{\lambda_\ell}{2^j}\right)} u_\ell(y), \qquad \langle f, \psi^k_{j,y}\rangle = \sum_{\ell\geq 0} \widehat{f}_\ell \, \overline{\widehat{\beta^k}\left(\frac{\lambda_\ell}{2^j}\right)} u_\ell(y). \tag{4.7}$$

Parseval's identity in the y variable gives

$$\int_{\mathcal{M}} |\langle f, \varphi_{J,y}\rangle|^2 \, d\mu(y) = \sum_{\ell\geq 0} |\widehat{f}_\ell|^2 \left| \widehat{\alpha}\left(\frac{\lambda_\ell}{2^J}\right)\right|^2$$

and the analogous identity for each high-pass term. Summing over the high-pass levels and using (4.6), the continuous framelet energy is exactly $\sum_{\ell\geq 0} |\widehat{f}_\ell|^2$. For a general $f \in L^2(\mathcal{M},\mu)$, the same identity follows by approximating f by finite spectral sums and using Tonelli's theorem for the nonnegative energy terms. Further details and equivalent conditions can be found in [45, Theorem 2.1].

4.2 Nested Marcinkiewicz–Zygmund sampling and semi-discrete framelets

The continuous system $\mathrm{CFS}_J(\Psi)$ uses discrete scales and continuous centers. Semi-discretization further replaces the center continuum at each scale by finitely many weighted nodes while preserving $L^2(\mathcal{M},\mu)$ frame bounds. Exact-quadrature constructions in [45, 49] preserve tightness but require structured rules, such as spherical t-designs. Here we instead use the nested measures from Section 3, whose Marcinkiewicz–Zygmund stability also allows geometrically controlled scattered and nested nodes.

For a fixed function f, the quantities to be discretized are the center-variable coefficient functions

$$P^{\varphi}_J(y) := \langle f, \varphi_{J,y}\rangle, \qquad P^{\psi^k}_j(y) := \langle f, \psi^k_{j,y}\rangle. \tag{4.8}$$

To use the Marcinkiewicz–Zygmund measures, these functions should lie in finite-dimensional diffusion polynomial spaces. Similar to the exact quadrature-based discretization [45, Corollary 2.6], we ensure this by imposing a compact spectral support condition on the generators. Namely, we assume that Ψ satisfies (4.4) and (4.5), and

$$(\operatorname{supp} \widehat{\alpha} \cap [0,\infty)) \cup \bigcup_{k=1}^{r} \Big(\operatorname{supp} \widehat{\beta^k} \cap [0,\infty)\Big) \subset [0,1]. \tag{4.9}$$

If a multiplier at scale j is supported in $[0,1]$ after the normalization $\lambda_\ell \mapsto \lambda_\ell/2^j$, then only eigenvalues $\lambda_\ell \leq 2^j$ can occur. The next lemma records the resulting bandwidth of the coefficient functions (4.8).

Lemma 4.1 (Bandwidth of coefficient functions)**.** *Assume that Ψ satisfies* (4.9)*. For every $f \in L^2(\mathcal{M}, \mu)$,*

$$P_J^{\varphi} \in \mathbb{P}_{2^J}, \qquad P_j^{\psi^k} \in \mathbb{P}_{2^j}, \qquad j \geq J, \quad k = 1, \dots, r. \tag{4.10}$$

Proof. By the coefficient expansion (4.7),

$$P_J^{\varphi}(y) = \sum_{\ell \geq 0} \widehat{f}_\ell \, \overline{\widehat{\alpha}\left(\frac{\lambda_\ell}{2^J}\right)} u_\ell(y).$$

The support condition (4.9) implies that the coefficient multiplying u_ℓ is zero unless $\lambda_\ell / 2^J \leq 1$, or equivalently $\lambda_\ell \leq 2^J$. Therefore $P_J^{\varphi} \in \mathbb{P}_{2^J}$. The proof for $P_j^{\psi^k}$ is similar. □

Remark 4.1 (On products of diffusion polynomials)**.** *Since the product of two diffusion polynomials may not necessarily be a diffusion polynomial, exact-quadrature discretizations often require a product assumption, such as $|P|^2 \in \mathbb{P}_{Cn}$ for $P \in \mathbb{P}_n$; see, e.g., [45]. Filbir and Mhaskar proved the existence of a manifold-dependent constant $C_{\mathrm{prod}} \geq 2$ such that*

$$P\overline{Q} \in \mathbb{P}_{C_{\mathrm{prod}} 2^j} \qquad \textit{for all } P, Q \in \mathbb{P}_{2^j};$$

see [23, Theorem A.1]. Consequently, exactness through degree $C_{\mathrm{prod}} 2^j$ makes the level-j weighted Fourier map an isometry. On the sphere and the flat torus one may take $C_{\mathrm{prod}} = 2$, so exactness through degree 2^{j+1} is enough. On a general compact manifold, however, C_{prod} need not equal 2, and the quadrature degree must be enlarged accordingly. Indeed, the product inclusion and the corresponding exactness rule give $\sum_{x \in \boldsymbol{X}_j} w_{j,x} P(x) \overline{Q(x)} = \langle P, Q \rangle_{L^2(\mu)}$. Thus, exactness alone is not the relevant condition, and it must be paired with a product-bandwidth inclusion of this form. The Marcinkiewicz–Zygmund approach in the current paper, however, needs no such hypothesis: Proposition 3.1 directly provides the required two-sided estimate for $|P|^2$.

Since the Marcinkiewicz–Zygmund inequalities are weighted, we absorb $w_x^{1/2}$ into each sampled atom. The resulting unweighted ℓ^2-sum of framelet coefficients then coincides with the corresponding weighted Marcinkiewicz–Zygmund sum. We first define this semi-discrete system and then establish its frame property.

Definition 4.2 (Semi-discrete framelets from nested Marcinkiewicz–Zygmund measures)**.** *Let $J_0 \in \mathbb{N}_0$ and $J \geq J_0$. For each $j \geq 0$, let $\boldsymbol{X}_j$ be a finite subset of $\mathcal{M}$ and let $\{w_{j,x}\}_{x \in \boldsymbol{X}_j}$ be positive weights. Assume that the sampling family is nested, that is, $\boldsymbol{X}_j \subset \boldsymbol{X}_{j+1}$ for all $j \geq 0$. The family*

$$\mathrm{FS}_J(\Psi; \boldsymbol{X}, w) = \{w_{J,x}^{1/2} \varphi_{J,x} : x \in \boldsymbol{X}_J\} \cup \{w_{j+1,x}^{1/2} \psi_{j,x}^k : x \in \boldsymbol{X}_{j+1}, \ j \geq J, \ k = 1, \dots, r\} \tag{4.11}$$

is called the semi-discrete framelet system from nested Marcinkiewicz–Zygmund measures*, associated with the continuous framelet system* $\mathrm{CFS}_J(\Psi)$.

Recall that a semi-discrete framelet system FS_J is a *frame* for $L^2(\mathcal{M}, \mu)$ with frame bounds $0 < A \leq B < \infty$ if its framelet coefficients satisfy

$$A \|f\|_{L^2(\mu)}^2 \leq \sum_{g \in \mathrm{FS}_J} |\langle f, g \rangle|^2 \leq B \|f\|_{L^2(\mu)}^2$$

for every $f \in L^2(\mathcal{M}, \mu)$. If $A = B = 1$, the system is *tight*. With the normalization inherited from the continuous tight frame, we call the system ε-*nearly tight* if

$$(1 - \varepsilon) \|f\|_{L^2(\mu)}^2 \leq \sum_{g \in \mathrm{FS}_J} |\langle f, g \rangle|^2 \leq (1 + \varepsilon) \|f\|_{L^2(\mu)}^2, \qquad 0 \leq \varepsilon < 1.$$

Equivalently, its frame operator S satisfies $\|S - I\| \leq \varepsilon$.

Theorem 4.1 (Nearly tight framelet systems from nested Marcinkiewicz–Zygmund measures)**.** *Let $J_0 \in \mathbb{N}_0$ and $J \geq J_0$. Assume that Ψ satisfies* (4.4)*,* (4.5)*, and* (4.9)*. Let $0 < \eta < 1$, and form the nested semi-discrete framelet system* (4.11) *from the fixed-tolerance nested Marcinkiewicz–Zygmund measures in Theorem 3.1, applied with $p = 2$ and constructed with the parameter $\gamma_{\eta,2,\mathcal{M}}$. Then $\mathrm{FS}_J(\Psi; \boldsymbol{X}, w)$ is an η-nearly tight frame for $L^2(\mathcal{M}, \mu)$, with frame bounds $1 - \eta$ and $1 + \eta$.*

Proof. For fixed $f \in L^2(\mathcal{M},\mu)$, by Lemma 4.1, $P_J^{\varphi} \in \mathbb{P}_{2^J}$ and $P_j^{\psi^k} \in \mathbb{P}_{2^j}$. Since the level-$q$ rule in Theorem 3.1 is a Marcinkiewicz–Zygmund rule for $\mathbb{P}_{2^q}$, we apply it at level $q = J$ to the low-pass term. For a high-pass term at scale j, we use the level-$(j+1)$ rule; this is valid because $\mathbb{P}_{2^j} \subset \mathbb{P}_{2^{j+1}}$. The low-pass estimate is

$$(1-\eta)\int_{\mathcal{M}} |P_J^{\varphi}(y)|^2\, d\mu(y) \le \sum_{x\in \boldsymbol{X}_J} w_{J,x}|P_J^{\varphi}(x)|^2 \le (1+\eta)\int_{\mathcal{M}} |P_J^{\varphi}(y)|^2\, d\mu(y),$$

which is the same as

$$(1-\eta)\|P_J^{\varphi}\|_{L^2(\mu)}^2 \le \sum_{x\in \boldsymbol{X}_J} |\langle f, w_{J,x}^{1/2}\varphi_{J,x}\rangle|^2 \le (1+\eta)\|P_J^{\varphi}\|_{L^2(\mu)}^2. \tag{4.12}$$

Similarly, for every $j \ge J$ and $k = 1, \dots, r$,

$$(1-\eta)\int_{\mathcal{M}} |P_j^{\psi^k}(y)|^2\, d\mu(y) \le \sum_{x\in \boldsymbol{X}_{j+1}} |\langle f, w_{j+1,x}^{1/2}\psi_{j,x}^k\rangle|^2 \le (1+\eta)\int_{\mathcal{M}} |P_j^{\psi^k}(y)|^2\, d\mu(y). \tag{4.13}$$

Summing the low-pass estimate (4.12) and the high-pass estimates (4.13) over $J \le j \le L$ gives the same two-sided inequality for the corresponding finite partial coefficient energy. Since all terms are nonnegative, we may let $L \to \infty$ and use monotone convergence on the discrete sums. On the continuous side, the corresponding partial energies converge to $\|f\|_{L^2(\mu)}^2$ by the tight identity (4.3). This proves the frame bounds $1-\eta$ and $1+\eta$ for $\mathrm{FS}_J(\Psi; \boldsymbol{X}, w)$. □

Remark 4.2 (Sampling cost of near-tightness)**.** *The tolerance η in Theorem 4.1 is the same at every level, but it can be chosen when constructing the measures. Choosing a smaller η gives tighter frame bounds $1 \pm \eta$. The price is a denser sampling family. Indeed, by Theorem 3.1 with $p = 2$ and $n_j = 2^j$, the construction gives*

$$|\boldsymbol{X}_j| \asymp \gamma_{\eta,2,\mathcal{M}}^{-d} 2^{jd}, \qquad w_{j,x} \asymp \gamma_{\eta,2,\mathcal{M}}^{d} 2^{-jd}, \qquad x \in \boldsymbol{X}_j, \quad j \ge 0.$$

Here the construction parameter is chosen subject to $0 < \gamma_{\eta,2,\mathcal{M}} < \min\{R_{\mathcal{M}}, c\eta/2\}$, where c is the constant in Proposition 3.1 with $p = 2$ and $\kappa = 1$. Thus improving the frame bounds by decreasing η generally amplifies the number of sampling centers through the factor $\gamma_{\eta,2,\mathcal{M}}^{-d}$, which is of order η^{-d} when $\gamma_{\eta,2,\mathcal{M}}$ is chosen proportional to η. A high-pass term at scale j uses $\boldsymbol{X}_{j+1}$, and hence has $|\boldsymbol{X}_{j+1}| \asymp \gamma_{\eta,2,\mathcal{M}}^{-d} 2^{(j+1)d}$ sampling centers.

There is another way to improve the frame bounds once a nested family has been constructed. Instead of rebuilding the whole family with a smaller value of $\gamma_{\eta,2,\mathcal{M}}$, one may keep the same nested sets and add s refinement steps beyond the basic sampling levels. Thus the low-pass term at scale J is sampled on $\boldsymbol{X}_{J+s}$, whereas a high-pass coefficient function at scale j, whose basic center set is $\boldsymbol{X}_{j+1}$, is sampled on $\boldsymbol{X}_{j+1+s}$. This is the sense in which the next construction is oversampled: all sampling levels in the basic semi-discrete system are shifted upward by s.

Definition 4.3 (Oversampled semi-discrete framelets from nested Marcinkiewicz–Zygmund measures)**.** *Let $J_0 \in \mathbb{N}_0$ and $J \ge J_0$. Assume that Ψ satisfies (4.4), (4.5), and (4.9). Let $\{\boldsymbol{X}_j, w_{j,x}\}_{j\ge 0}$ be a nested weighted sampling family, and let $s \ge 0$ be an integer. The shifted sampling family remains nested in the sense of $\boldsymbol{X}_{q+s} \subset \boldsymbol{X}_{q+1+s}$, $q \ge 0$. Define*

$$\varphi_{J,x}^{(s)} := w_{J+s,x}^{1/2}\varphi_{J,x}, \qquad x \in \boldsymbol{X}_{J+s},$$

and

$$\psi_{j,x}^{k,(s)} := w_{j+1+s,x}^{1/2}\psi_{j,x}^k, \qquad x \in \boldsymbol{X}_{j+1+s}, \quad j \ge J, \quad k = 1, \dots, r.$$

The family

$$\mathrm{FS}_J^{(s)} := \{\varphi_{J,x}^{(s)} : x \in \boldsymbol{X}_{J+s}\} \cup \{\psi_{j,x}^{k,(s)} : x \in \boldsymbol{X}_{j+1+s},\ j \ge J,\ k = 1, \dots, r\} \tag{4.14}$$

is called the oversampled semi-discrete framelet system from nested Marcinkiewicz–Zygmund measures *with* oversampling shift *s.*

Theorem 4.2 (Near-tightness under nested oversampling)**.** *Let $J_0 \in \mathbb{N}_0$ and $J \geq J_0$. Assume that Ψ satisfies (4.4), (4.5), and (4.9). Let $\{\boldsymbol{X}_j, w_{j,x}\}_{j\geq 0}$ be the nested weighted family constructed in Lemmas 3.1 and 3.2 with a fixed parameter $0 < \gamma < R_{\mathcal{M}}$, and form the oversampled nested semi-discrete framelet system (4.14) from this family. Let c be the constant in Proposition 3.1 with $p = 2$ and $\kappa = 1$, and set $C_{\mathcal{M}} := 2\gamma/c$. Let $s \geq 0$ be an integer such that $\rho_s := C_{\mathcal{M}} 2^{-s} < 1$. Then $\mathrm{FS}_J^{(s)}$ is a ρ_s-nearly tight frame, where $\rho_s = C_{\mathcal{M}} 2^{-s}$. In particular, these systems become asymptotically tight as $s \to \infty$.*

Proof. We prove the estimate for a high-pass coefficient function; the low-pass term is treated similarly. By Lemma 4.1, $P_j^{\psi^k} \in \mathbb{P}_{2^j}$. It is sampled on level $j + 1 + s$. For the dyadic construction in Lemma 3.1, the mesh norm at this sampling level satisfies $h_{\boldsymbol{X}_{j+1+s}} \leq \gamma 2^{-(j+1+s)}$. Therefore, with $N_j := 2^j$, we have $N_j h_{\boldsymbol{X}_{j+1+s}} \leq (\gamma/2)2^{-s}$. Set $\eta_{j,s}^{\psi} := (2/c)N_j h_{\boldsymbol{X}_{j+1+s}}$. Proposition 3.1 with $p = 2$ and $\kappa = 1$, together with $\eta_{j,s}^{\psi} \leq (C_{\mathcal{M}}/2)2^{-s} \leq \rho_s < 1$, gives a Marcinkiewicz–Zygmund inequality for $P_j^{\psi^k}$ with tolerance $\eta_{j,s}^{\psi}$. The constant is independent of the framelet scale j, the high-pass channel k, and the starting scale J. By the definition of $\psi_{j,x}^{k,(s)}$, this Marcinkiewicz–Zygmund inequality becomes

$$
\begin{aligned}
(1-\rho_s)\|P_j^{\psi^k}\|_{L^2(\mu)}^2 \leq & (1-\eta_{j,s}^{\psi})\|P_j^{\psi^k}\|_{L^2(\mu)}^2 \leq \sum_{x\in \boldsymbol{X}_{j+1+s}} |\langle f, \psi_{j,x}^{k,(s)}\rangle|^2 \leq (1+\eta_{j,s}^{\psi})\|P_j^{\psi^k}\|_{L^2(\mu)}^2 \\
\leq & (1+\rho_s)\|P_j^{\psi^k}\|_{L^2(\mu)}^2.
\end{aligned}
$$

For $P_J^{\varphi} \in \mathbb{P}_{2^J}$, sampled on $\boldsymbol{X}_{J+s}$, we have $2^J h_{\boldsymbol{X}_{J+s}} \leq \gamma 2^{-s}$; hence the same Marcinkiewicz–Zygmund argument gives tolerance at most $C_{\mathcal{M}}2^{-s} = \rho_s$. Summing the low-pass estimate and the high-pass estimates over $J \leq j \leq L$ gives the desired two-sided estimate for the finite partial coefficient energy. Letting $L \to \infty$, monotone convergence applies to the discrete sums, while the continuous partial energies converge to $\|f\|_{L^2(\mu)}^2$ by (4.3). This proves the claimed frame bounds. $\square$

Remark 4.3 (Oversampling cost)**.** *At high-pass scale j, for example, the construction uses*

$$
|\boldsymbol{X}_{j+1+s}| \asymp \gamma^{-d}2^{(j+1+s)d}, \qquad w_{j+1+s,x} \asymp \gamma^{d}2^{-(j+1+s)d}.
$$

Thus the improvement from $1 \pm O(1)$ to $1 \pm O(2^{-s})$ is obtained by using about 2^{sd} more centers than the basic high-pass set $\boldsymbol{X}_{j+1}$.

5 Fully discrete framelet transform and fast implementation

We now pass from semi-discrete framelets to a fully discrete transform. We use the frame-theoretic terms *analysis* and *synthesis*. In the filter-bank language, analysis is implemented by *decomposition*, whereas synthesis gives *reconstruction* directly only in the perfect-reconstruction case. The same nested weighted node sets as in the construction of semi-discrete framelets are used throughout.

5.1 Coefficients at the finest level

Fix a coarsest level J and a finest scaling level $K > J$. Since the generators are compactly supported on the nonnegative spectral axis, we define the finite active index set $\Lambda_j^{\alpha} := \{\ell \geq 0 : \widehat{\alpha}(\lambda_\ell/2^j) \neq 0\}$, $j \geq J$. The support assumption (4.9) implies $\Lambda_j^{\alpha} \subset \{\ell : \lambda_\ell \leq 2^j\}$. We assume, as is the case for the filters used below, that $\Lambda_{j-1}^{\alpha} \subset \Lambda_j^{\alpha}$.

For $\widehat{\mathbf{z}} = (z_\ell)_{\ell\in\Lambda_j^{\alpha}} \in \mathbb{C}^{\Lambda_j^{\alpha}}$, the weighted discrete Fourier transform $\mathbf{F}_j : \mathbb{C}^{\Lambda_j^{\alpha}} \longrightarrow \mathbb{C}^{\boldsymbol{X}_j}$ associated with $\boldsymbol{X}_j$ is defined by

$$
(\mathbf{F}_j\widehat{\mathbf{z}})(x) := \sum_{\ell\in\Lambda_j^{\alpha}} z_\ell w_{j,x}^{1/2} u_\ell(x), \qquad x \in \boldsymbol{X}_j. \tag{5.1}
$$

Its adjoint discrete Fourier transform $\mathbf{F}_j^* : \mathbb{C}^{\boldsymbol{X}_j} \longrightarrow \mathbb{C}^{\Lambda_j^{\alpha}}$ is then given by

$$
(\mathbf{F}_j^*\mathbf{v})_\ell := \sum_{x\in\boldsymbol{X}_j} v_x w_{j,x}^{1/2}\overline{u_\ell(x)}, \qquad \ell \in \Lambda_j^{\alpha}. \tag{5.2}
$$

For the oversampled construction we use instead $\mathbf{F}_j^{(s)} : \mathbb{C}^{\Lambda_j^{\alpha}} \to \mathbb{C}^{\boldsymbol{X}_{j+s}}$, defined by replacing $w_{j,x}$ and $\boldsymbol{X}_j$ in (5.1) by $w_{j+s,x}$ and $\boldsymbol{X}_{j+s}$. For the Marcinkiewicz–Zygmund measures considered below, Proposition 5.1 shows that $\mathbf{F}_j$ is injective and well conditioned, although $\mathbf{F}_j^*\mathbf{F}_j$ need not be the identity.

Proposition 5.1 (Gram bounds for the weighted Fourier map)**.** *Suppose the level-j sampling rule has L^2 Marcinkiewicz–Zygmund constants $1-\varepsilon_j$ and $1+\varepsilon_j$, where $0 \le \varepsilon_j < 1$. Then on $\mathbb{C}^{\Lambda_j^\alpha}$,*

$$(1-\varepsilon_j)I \le \mathbf{F}_j^*\mathbf{F}_j \le (1+\varepsilon_j)I. \tag{5.3}$$

Proof. For $\widehat{\mathbf{z}} \in \mathbb{C}^{\Lambda_j^\alpha}$, set $P_{\widehat{\mathbf{z}}} := \sum_{\ell\in\Lambda_j^\alpha} z_\ell u_\ell$. Since $\Lambda_j^\alpha \subseteq \{\ell : \lambda_\ell \le 2^j\}$, we have $P_{\widehat{\mathbf{z}}} \in \mathbb{P}_{2^j}$. Orthonormality gives $\|P_{\widehat{\mathbf{z}}}\|_{L^2(\mu)}^2 = \|\widehat{\mathbf{z}}\|_2^2$, while $\|\mathbf{F}_j\widehat{\mathbf{z}}\|_2^2 = \sum_{x\in\boldsymbol{X}_j} w_{j,x}|P_{\widehat{\mathbf{z}}}(x)|^2$. The Marcinkiewicz–Zygmund inequality therefore yields

$$(1-\varepsilon_j)\|\widehat{\mathbf{z}}\|_2^2 \le \|\mathbf{F}_j\widehat{\mathbf{z}}\|_2^2 \le (1+\varepsilon_j)\|\widehat{\mathbf{z}}\|_2^2.$$

Since $\|\mathbf{F}_j\widehat{\mathbf{z}}\|_2^2 = \langle \mathbf{F}_j^*\mathbf{F}_j\widehat{\mathbf{z}}, \widehat{\mathbf{z}}\rangle$, this is equivalent to (5.3). □

For $f \in L^2(\mathcal{M},\mu)$, to start the finite transform, we need the spectral vector

$$\widehat{\mathbf{v}}_K = \left(\widehat{f}_\ell\, \overline{\widehat{\alpha}(\lambda_\ell/2^K)}\right)_{\ell\in\Lambda_K^\alpha}$$

of the finest scaling coefficient function $v_K(y) := \langle f, \varphi_{K,y}\rangle$. If only the weighted nodal values $\mathbf{v}_K$ of f are available, we compute a spectral fit by

$$\widehat{\mathbf{v}}_K = \underset{\mathbf{z}\in\mathbb{C}^{\Lambda_K^\alpha}}{\operatorname{argmin}} \|\mathbf{F}_K\mathbf{z} - \mathbf{v}_K\|_2^2,$$

equivalently, $\mathbf{F}_K^*\mathbf{F}_K\widehat{\mathbf{v}}_K = \mathbf{F}_K^*\mathbf{v}_K$. Algorithm 1 solves it by a conjugate-gradient iteration. Here and below, ϵ_{mach} denotes machine precision.

Algorithm 1: Finest-level scaling coefficient computation

Input: A weighted finest-level scaling sequence $\mathbf{v}_K$, operators $\mathbf{F}_K$ and $\mathbf{F}_K^*$, tolerance $\tau > 0$, and maximum iteration number $M_{\max}$

Output: Approximation to $\widehat{\mathbf{v}}_K \in \mathbb{C}^{\Lambda_K^\alpha}$

$\mathbf{b} \leftarrow \mathbf{F}_K^*\mathbf{v}_K$, $\mathbf{z} \leftarrow 0$;
$\mathbf{r} \leftarrow \mathbf{b}$, $\mathbf{p} \leftarrow \mathbf{r}$;
for $m = 0, 1, \dots, M_{\max}-1$ **do**
  if $\|\mathbf{r}\|/\max\{\|\mathbf{b}\|, \epsilon_{\mathrm{mach}}\} \le \tau$ **then**
    return $\widehat{\mathbf{v}}_K \leftarrow \mathbf{z}$
  end if
  $\mathbf{q} \leftarrow \mathbf{F}_K^*\mathbf{F}_K\mathbf{p}$;
  $\theta_m \leftarrow \langle \mathbf{r},\mathbf{r}\rangle / \langle \mathbf{p},\mathbf{q}\rangle$;
  $\mathbf{z} \leftarrow \mathbf{z} + \theta_m\mathbf{p}$;
  $\mathbf{r}_{\mathrm{new}} \leftarrow \mathbf{r} - \theta_m\mathbf{q}$;
  $\delta_m \leftarrow \langle \mathbf{r}_{\mathrm{new}},\mathbf{r}_{\mathrm{new}}\rangle / \langle \mathbf{r},\mathbf{r}\rangle$;
  $\mathbf{p} \leftarrow \mathbf{r}_{\mathrm{new}} + \delta_m\mathbf{p}$, $\mathbf{r} \leftarrow \mathbf{r}_{\mathrm{new}}$;
end for
$\widehat{\mathbf{v}}_K \leftarrow \mathbf{z}$;

Remark 5.1. *Let C_{prod} be a product-bandwidth constant as in Remark 4.1. Suppose that the level-K sampling rule is exact for polynomials of degree up to $C_{\mathrm{prod}}2^K$, and that the corresponding product condition holds at level K. For $\widehat{\mathbf{z}}, \widehat{\mathbf{y}} \in \mathbb{C}^{\Lambda_K^\alpha}$, set $P_{\widehat{\mathbf{z}}} := \sum_{\ell\in\Lambda_K^\alpha} z_\ell u_\ell$ and $P_{\widehat{\mathbf{y}}} := \sum_{\ell\in\Lambda_K^\alpha} y_\ell u_\ell$. These functions belong to $\mathbb{P}_{2^K}$. Hence exactness applied to $P_{\widehat{\mathbf{z}}}\overline{P_{\widehat{\mathbf{y}}}}$ gives*

$$\langle \mathbf{F}_K\widehat{\mathbf{z}}, \mathbf{F}_K\widehat{\mathbf{y}}\rangle = \sum_{x\in\boldsymbol{X}_K} w_{K,x} P_{\widehat{\mathbf{z}}}(x)\overline{P_{\widehat{\mathbf{y}}}(x)} = \langle \widehat{\mathbf{z}}, \widehat{\mathbf{y}}\rangle.$$

Thus $\mathbf{F}_K^*\mathbf{F}_K = I$. Consequently, Algorithm 1 is unnecessary, and the finest-level coefficients can be computed directly as $\widehat{\mathbf{v}}_K = \mathbf{F}_K^*\mathbf{v}_K$ using the adjoint discrete Fourier transform. This recovers the setting considered in [45].

5.2 Filter banks

The continuous framelet system $\mathrm{CFS}_J(\Psi)$ was described in terms of generators $\Psi = \{\alpha; \beta^1, \ldots, \beta^r\}$. For an actual multilevel algorithm, one also needs a filter bank

$$\Theta = \{a; b^1, \ldots, b^r\},$$

where $a, b^1, \ldots, b^r \in \ell_1(\mathbb{Z})$. Here a is the low-pass filter and $b^1, \ldots, b^r$ are high-pass filters. We write their symbols, using the same 2π-Fourier convention as above, as $\widehat{a}(\xi) = \sum_{m\in\mathbb{Z}} a_m e^{-2\pi i m\xi}$ and $\widehat{b^k}(\xi) = \sum_{m\in\mathbb{Z}} b^k_m e^{-2\pi i m\xi}$. The role of the filter bank is to connect adjacent scales. The generators and the filters are related by the refinement relations

$$\widehat{\alpha}(2\xi) = \widehat{a}(\xi)\widehat{\alpha}(\xi), \qquad \widehat{\beta^k}(2\xi) = \widehat{b^k}(\xi)\widehat{\alpha}(\xi), \quad k = 1, \ldots, r. \tag{5.4}$$

Thus α and β^k describe the framelet atoms at absolute scales, while a and b^k describe the transition from one scale to the next. For example,

$$\widehat{\alpha}\left(\frac{\lambda_\ell}{2^{j-1}}\right) = \widehat{a}\left(\frac{\lambda_\ell}{2^j}\right)\widehat{\alpha}\left(\frac{\lambda_\ell}{2^j}\right), \qquad \widehat{\beta^k}\left(\frac{\lambda_\ell}{2^{j-1}}\right) = \widehat{b^k}\left(\frac{\lambda_\ell}{2^j}\right)\widehat{\alpha}\left(\frac{\lambda_\ell}{2^j}\right). \tag{5.5}$$

We assume that the filters split the energy of every active coefficient:

$$\left|\widehat{a}\left(\frac{\lambda_\ell}{2^j}\right)\right|^2 + \sum_{k=1}^{r}\left|\widehat{b^k}\left(\frac{\lambda_\ell}{2^j}\right)\right|^2 = 1, \qquad \ell \in \Lambda_j^\alpha, \quad j > J. \tag{5.6}$$

Together, these assumptions provide the algebraic basis of the fully discrete transform.

Proposition 5.2 (Generator conditions from a filter bank)**.** *Let $\Psi = \{\alpha; \beta^1, \ldots, \beta^r\} \subset L_1(\mathbb{R})$ and the filter bank $\Theta = \{a; b^1, \ldots, b^r\} \subset \ell_1(\mathbb{Z})$. Suppose that the refinement equations (5.4) hold for all $\xi \in \mathbb{R}$, that the active spectral partition (5.6) holds, and that*

$$|\widehat{\alpha}(0)| = 1, \qquad \operatorname{supp}\widehat{\alpha} \cap [0, \infty) \subset [0, 1/2].$$

Then Ψ satisfies the low-pass limit (4.4), the identity (4.5), and the support condition (4.9).

Proof. Since $\alpha \in L_1(\mathbb{R})$, its Fourier transform is continuous. For each fixed $\ell \geq 0$, $\lambda_\ell/2^j \to 0$ as $j \to \infty$. Hence

$$\lim_{j\to\infty}\left|\widehat{\alpha}\left(\frac{\lambda_\ell}{2^j}\right)\right| = |\widehat{\alpha}(0)| = 1,$$

which is (4.4). Next fix $j \geq J$ and $\ell \geq 0$, and set $\xi = \lambda_\ell/2^{j+1}$. The refinement equations give

$$\widehat{\alpha}\left(\frac{\lambda_\ell}{2^j}\right) = \widehat{a}(\xi)\widehat{\alpha}\left(\frac{\lambda_\ell}{2^{j+1}}\right), \qquad \widehat{\beta^k}\left(\frac{\lambda_\ell}{2^j}\right) = \widehat{b^k}(\xi)\widehat{\alpha}\left(\frac{\lambda_\ell}{2^{j+1}}\right).$$

If $\widehat{\alpha}(\lambda_\ell/2^{j+1}) = 0$, then all terms in (4.5) are zero. Otherwise $\ell \in \Lambda_{j+1}^\alpha$, and (5.6), applied at level $j+1$, yields

$$\left|\widehat{\alpha}\left(\frac{\lambda_\ell}{2^j}\right)\right|^2 + \sum_{k=1}^{r}\left|\widehat{\beta^k}\left(\frac{\lambda_\ell}{2^j}\right)\right|^2 = \left(|\widehat{a}(\xi)|^2 + \sum_{k=1}^{r}|\widehat{b^k}(\xi)|^2\right)\left|\widehat{\alpha}\left(\frac{\lambda_\ell}{2^{j+1}}\right)\right|^2 = \left|\widehat{\alpha}\left(\frac{\lambda_\ell}{2^{j+1}}\right)\right|^2.$$

This proves (4.5). Finally, the assumed support of $\widehat{\alpha}$ gives $\operatorname{supp}\widehat{\alpha} \cap [0, \infty) \subset [0, 1]$. For each k, the second refinement equation implies $\widehat{\beta^k}(\eta) = \widehat{b^k}(\eta/2)\widehat{\alpha}(\eta/2), \eta \in \mathbb{R}$. Thus, on the nonnegative axis, $\widehat{\beta^k}(\eta) \neq 0$ can occur only when $\eta/2 \in \operatorname{supp}\widehat{\alpha} \cap [0, \infty) \subset [0, 1/2]$, that is, only when $\eta \in [0, 1]$. Hence $\operatorname{supp}\widehat{\beta^k} \cap [0, \infty) \subset [0, 1]$ for every k, and (4.9) follows. □

5.3 Analysis

We next define the level-transition maps used by the filter bank. Let $\mathbf{v}_j = \mathbf{F}_j\widehat{\mathbf{v}}_j$ be a weighted nodal sequence on $\boldsymbol{X}_j$, where $\widehat{\mathbf{v}}_j = (\widehat{v}_{j,\ell})_{\ell\in\Lambda_j^\alpha}$. For a filter $h \in \ell_1(\mathbb{Z})$, we define the discrete convolution of $\mathbf{v}_j$ with h by

$$(\mathbf{v}_j *_j h)_x := \sum_{\ell\in\Lambda_j^\alpha} \widehat{v}_{j,\ell}\widehat{h}\left(\frac{\lambda_\ell}{2^j}\right) w_{j,x}^{1/2} u_\ell(x), \quad x \in \boldsymbol{X}_j.$$

The downsampling operator is defined in the same spectral language:

$$(\mathbf{v}_j \downarrow_j)_x := \sum_{\ell \in \Lambda_j^\alpha} \widehat{v}_{j,\ell} w_{j-1,x}^{1/2} u_\ell(x), \quad x \in \boldsymbol{X}_{j-1}.$$

Thus $\downarrow_j$ is not a deletion of entries; it is a change of weighted sampling nodes from $\boldsymbol{X}_j$ to $\boldsymbol{X}_{j-1}$. Conversely, the upsampling operator from level $j-1$ to level j is defined by

$$(\mathbf{v}_{j-1} \uparrow_j)_x := \sum_{\ell \in \Lambda_{j-1}^\alpha} \widehat{v}_{j-1,\ell} w_{j,x}^{1/2} u_\ell(x), \quad x \in \boldsymbol{X}_j.$$

The next result explains why the analysis step is implemented by the conjugate filters. For a filter h, let h^* be characterized by $\widehat{h^*}(\xi) = \overline{\widehat{h}(\xi)}$ on real ξ.

Theorem 5.1 (One-step coefficient decomposition)**.** *Suppose that the generators and filters satisfy the conditions in Proposition 5.2. Fix $f \in L^2(\mathcal{M}, \mu)$ and a level $j > J$. Define the coefficient functions*

$$v_j(y) := \langle f, \varphi_{j,y} \rangle, \quad v_{j-1}(y) := \langle f, \varphi_{j-1,y} \rangle, \qquad w_{j-1}^k(y) := \langle f, \psi_{j-1,y}^k \rangle, \quad k = 1, \dots, r.$$

Let $\mathbf{v}_j$ and $\mathbf{v}_{j-1}$ be the weighted nodal sequences of v_j and v_{j-1} on $\boldsymbol{X}_j$ and $\boldsymbol{X}_{j-1}$, respectively, and let $\mathbf{w}_{j-1}^k$ be the weighted nodal sequence of w_{j-1}^k on $\boldsymbol{X}_j$. Then

$$\mathbf{v}_{j-1} = (\mathbf{v}_j *_j a^*) \downarrow_j, \qquad \mathbf{w}_{j-1}^k = \mathbf{v}_j *_j \left(b^k\right)^*, \quad k = 1, \dots, r.$$

Proof. By the coefficient expansion (4.7), the spectral coefficients of v_j are $\widehat{v}_{j,\ell} = \widehat{f}_\ell \, \overline{\widehat{\alpha}\,(\lambda_\ell/2^j)}$, $\ell \in \Lambda_j^\alpha$. The scaled refinement equations (5.5) therefore give, for every $\ell \in \Lambda_j^\alpha$,

$$\widehat{v}_{j,\ell} \, \overline{\widehat{a}\left(\frac{\lambda_\ell}{2^j}\right)} = \widehat{f}_\ell \, \overline{\widehat{\alpha}\left(\frac{\lambda_\ell}{2^{j-1}}\right)}, \qquad \widehat{v}_{j,\ell} \, \overline{\widehat{b^k}\left(\frac{\lambda_\ell}{2^j}\right)} = \widehat{f}_\ell \, \overline{\widehat{\beta^k}\left(\frac{\lambda_\ell}{2^{j-1}}\right)}, \qquad k = 1, \dots, r.$$

For $\ell \notin \Lambda_j^\alpha$, the scaled refinement equations show that both coarse- and high-pass coefficients on the right vanish as well. Thus these identities determine all the spectral coefficients of the corresponding coefficient functions $v_{j-1}(y) = \langle f, \varphi_{j-1,y} \rangle$ and $w_{j-1}^k(y) = \langle f, \psi_{j-1,y}^k \rangle$. Evaluating the first one on $\boldsymbol{X}_{j-1}$ with the weights $w_{j-1,x}$ yields $(\mathbf{v}_j *_j a^*) \downarrow_j = \mathbf{v}_{j-1}$, while evaluating the second identity on $\boldsymbol{X}_j$ with the weights $w_{j,x}$ yields $\mathbf{v}_j *_j (b^k)^* = \mathbf{w}_{j-1}^k$ for $k = 1, \dots, r$. ☐

Starting from the finest-level vector $\widehat{\mathbf{v}}_K$ computed by Algorithm 1, we repeat the one-step decomposition for $j = K, K-1, \dots, J+1$. The restriction $\mathcal{R}_j^{j-1} : \mathbb{C}^{\Lambda_j^\alpha} \to \mathbb{C}^{\Lambda_{j-1}^\alpha}$ is taken onto Λ_{j-1}^α, that is,

$$\left(\mathcal{R}_j^{j-1} \widehat{\mathbf{z}}\right)_\ell = z_\ell, \qquad \ell \in \Lambda_{j-1}^\alpha.$$

Algorithm 2 first performs the spectral filtering and then stores each branch on its prescribed weighted node set.

Algorithm 2: Multi-level coefficient analysis

Input: Finest-level scaling vector $\widehat{\mathbf{v}}_K \in \mathbb{C}^{\Lambda_K^\alpha}$, filters $a, b^1, \dots, b^r$, and levels $J < K$
Output: Coarse sequence $\mathbf{v}_J$ and detail sequences $\mathbf{w}_{j-1}^k$, $j = J+1, \dots, K$, $k = 1, \dots, r$

for $j = K, K-1, \dots, J+1$ **do**
  for $k = 1, \dots, r$ **do**
    $\widehat{\mathbf{w}}_{j-1}^k \leftarrow \widehat{\mathbf{v}}_j \overline{\widehat{b^k}(2^{-j}\lambda)}$;
    $\mathbf{w}_{j-1}^k \leftarrow \mathbf{F}_j \widehat{\mathbf{w}}_{j-1}^k$;
  end for
  $\widehat{\mathbf{v}}_{j-1} \leftarrow \mathcal{R}_j^{j-1}\left(\widehat{\mathbf{v}}_j \overline{\widehat{a}(2^{-j}\lambda)}\right)$;
end for
$\mathbf{v}_J \leftarrow \mathbf{F}_J \widehat{\mathbf{v}}_J$;

Let $\mathcal{A}_{J,K} : \mathbb{C}^{\Lambda_K^\alpha} \to \mathbb{C}^{\boldsymbol{X}_J} \times \prod_{j=J+1}^K \left(\mathbb{C}^{\boldsymbol{X}_j}\right)^r$ be the linear map implemented by Algorithm 2, namely

$$\mathcal{A}_{J,K} \widehat{\mathbf{v}}_K = \left(\mathbf{v}_J, \{\mathbf{w}_{j-1}^k\}\right).$$

Equip the domain and codomain with their standard Euclidean inner products. The adjoint $\mathcal{A}_{J,K}^*$ is therefore the unique linear map satisfying

$$\langle \mathcal{A}_{J,K}\widehat{\mathbf{v}}_K, \mathbf{c}\rangle = \langle \widehat{\mathbf{v}}_K, \mathcal{A}_{J,K}^*\mathbf{c}\rangle$$

for every $\widehat{\mathbf{v}}_K$ and coefficient collection $\mathbf{c}$. For an oversampling shift $s \geq 0$, the operator $\mathcal{A}_{J,K}^{(s)}$ is defined by the same spectral filtering steps, with $\mathbf{F}_J$ and $\mathbf{F}_j$ replaced by $\mathbf{F}_J^{(s)}$ and $\mathbf{F}_j^{(s)}$, respectively.

5.4 Synthesis

We now synthesize a finest-level scaling vector from a coefficient collection

$$\mathbf{c} = \left(\mathbf{c}_J, \{\mathbf{c}_{j-1}^k : j = J+1, \ldots, K,\ k = 1, \ldots, r\}\right),$$

using the adjoint filter bank. Define the zero-extension map $\mathcal{E}_{j-1}^j : \mathbb{C}^{\Lambda_{j-1}^\alpha} \to \mathbb{C}^{\Lambda_j^\alpha}$ by

$$\left(\mathcal{E}_{j-1}^j\widehat{\mathbf{z}}\right)_\ell = \begin{cases} z_\ell, & \ell \in \Lambda_{j-1}^\alpha, \\ 0, & \ell \in \Lambda_j^\alpha \setminus \Lambda_{j-1}^\alpha. \end{cases}$$

It is the Euclidean adjoint of the restriction operator $\mathcal{R}_j^{j-1}$ in the analysis step. The recursion for the Euclidean adjoint $\mathcal{A}_{J,K}^*$ is

$$\widehat{\mathbf{z}}_J := \mathbf{F}_J^*\mathbf{c}_J, \qquad \widehat{\mathbf{z}}_j := \mathcal{E}_{j-1}^j\widehat{\mathbf{z}}_{j-1}\widehat{a}(2^{-j}\lambda.) + \sum_{k=1}^r \left(\mathbf{F}_j^*\mathbf{c}_{j-1}^k\right)\widehat{b^k}(2^{-j}\lambda.), \quad j = J+1, \ldots, K.$$

This recursion is obtained by taking the Euclidean adjoint of each step of Algorithm 2: evaluation by $\mathbf{F}_j$ becomes $\mathbf{F}_j^*$, while multiplication by the conjugate filters becomes multiplication by $\widehat{a}$ and $\widehat{b^k}$. It is described in the following Algorithm 3, computing $\mathcal{A}_{J,K}^*\mathbf{c}$.

Algorithm 3: Multi-level coefficient one-pass synthesis

Input: Coarse sequence $\mathbf{c}_J$, detail sequences $\mathbf{c}_{j-1}^k$, filters $a, b^1, \ldots, b^r$, and levels $J < K$
Output: One-pass finest-level spectral vector $\widehat{\mathbf{z}}_K$

$\widehat{\mathbf{z}}_J \leftarrow \mathbf{F}_J^*\mathbf{c}_J$;
for $j = J+1, J+2, \ldots, K$ **do**
  $\widehat{\mathbf{z}}_j \leftarrow \mathcal{E}_{j-1}^j\widehat{\mathbf{z}}_{j-1}\widehat{a}(2^{-j}\lambda.)$;
  for $k = 1, \ldots, r$ **do**
    $\widehat{\mathbf{e}}_{j-1}^k \leftarrow \mathbf{F}_j^*\mathbf{c}_{j-1}^k$;
    $\widehat{\mathbf{z}}_j \leftarrow \widehat{\mathbf{z}}_j + \widehat{\mathbf{e}}_{j-1}^k\widehat{b^k}(2^{-j}\lambda.)$;
  end for
end for

When the sampling rules have sufficient algebraic exactness, the adjoint filter-bank pass is already an exact inverse; no iterative correction is needed. More precisely, we have the following result.

Proposition 5.3 (Exact one-pass synthesis)**.** *Suppose that the refinement relations* (5.4) *and the filter-bank partition* (5.6) *hold. Let* C_{prod} *be a product-bandwidth constant as in Remark 4.1. Assume, for every* $j = J, \ldots, K$*, that the level-*j *sampling rule is exact for polynomials of degree up to* $C_{\mathrm{prod}}2^j$*, and that the corresponding product condition holds. Then*

$$\mathcal{A}_{J,K}^*\mathcal{A}_{J,K} = I \quad \text{on } \mathbb{C}^{\Lambda_K^\alpha}. \tag{5.7}$$

Consequently, if $\mathbf{c} = \mathcal{A}_{J,K}\widehat{\mathbf{v}}_K$*, then Algorithm 3 returns* $\widehat{\mathbf{v}}_K$ *after one pass.*

Proof. Recall that in this case, $\mathbf{F}_j^*\mathbf{F}_j = I$ on $\mathbb{C}^{\Lambda_j^\alpha}$ for every j, including the coarsest branch at level J; see the argument in Remark 5.1. We consider one spectral filter-bank step. The low-pass branch has coefficients $\widehat{\mathbf{v}}_j\overline{\widehat{a}}$, restricted to Λ^{α}_{j-1}, and the kth detail branch has coefficients $\widehat{\mathbf{v}}_j\overline{\widehat{b^k}}$. The refinement relation makes the low-pass product vanish outside Λ^{α}_{j-1}. Therefore the partition identity gives

$$\|\widehat{\mathbf{v}}_j\|_2^2 = \|\widehat{\mathbf{v}}_{j-1}\|_2^2 + \sum_{k=1}^{r} \|\widehat{\mathbf{w}}^k_{j-1}\|_2^2.$$

Iterating this identity shows that the spectral filter-bank map

$$\mathcal{W}_{J,K} : \mathbb{C}^{\Lambda^{\alpha}_K} \longrightarrow \mathbb{C}^{\Lambda^{\alpha}_J} \oplus \bigoplus_{j=J+1}^{K} \left(\mathbb{C}^{\Lambda^{\alpha}_j}\right)^r, \qquad \widehat{\mathbf{v}}_K \longmapsto \left(\widehat{\mathbf{v}}_J, \{\widehat{\mathbf{w}}^k_{j-1}\}\right),$$

is an isometry, so $\mathcal{W}^*_{J,K}\mathcal{W}_{J,K} = I$. Let

$$\mathcal{D}_{J,K} : \mathbb{C}^{\Lambda^{\alpha}_J} \oplus \bigoplus_{j=J+1}^{K} \left(\mathbb{C}^{\Lambda^{\alpha}_j}\right)^r \longrightarrow \mathbb{C}^{\boldsymbol{X}_J} \oplus \bigoplus_{j=J+1}^{K} \left(\mathbb{C}^{\boldsymbol{X}_j}\right)^r$$

be the block-diagonal map that applies $\mathbf{F}_J$ to the coarse branch and $\mathbf{F}_j$ to each level-j detail branch. The fact $\mathbf{F}_j^*\mathbf{F}_j = I$ shows that every block is an isometry, hence $\mathcal{D}^*_{J,K}\mathcal{D}_{J,K} = I$.

Since the analysis operator implemented by Algorithm 2 factors as $\mathcal{A}_{J,K} = \mathcal{D}_{J,K}\mathcal{W}_{J,K}$, we obtain

$$\mathcal{A}^*_{J,K}\mathcal{A}_{J,K} = \mathcal{W}^*_{J,K}\mathcal{D}^*_{J,K}\mathcal{D}_{J,K}\mathcal{W}_{J,K} = \mathcal{W}^*_{J,K}\mathcal{W}_{J,K} = I \quad \text{on } \mathbb{C}^{\Lambda^{\alpha}_K}.$$

The conclusion follows because the one-pass algorithm computes $\mathcal{A}^*_{J,K}\mathbf{c}$. □

For Marcinkiewicz–Zygmund sampling, the weighted Fourier maps are well conditioned but generally not isometries. Hence one-pass synthesis is still the adjoint operation, but $\mathcal{A}^*_{J,K}\mathcal{A}_{J,K} \neq I$ in general. To recover the canonical coefficient vector, we instead solve

$$\widehat{\mathbf{u}}_K = \operatorname*{argmin}_{\widehat{\mathbf{z}} \in \mathbb{C}^{\Lambda^{\alpha}_K}} \frac{1}{2}\|\mathcal{A}_{J,K}\widehat{\mathbf{z}} - \mathbf{c}\|_2^2.$$

Its normal equation, involving the coefficient frame operator $\mathbf{S}_{J,K} := \mathcal{A}^*_{J,K}\mathcal{A}_{J,K}$, is given by

$$\mathbf{S}_{J,K}\widehat{\mathbf{u}}_K = \mathcal{A}^*_{J,K}\mathbf{c}.$$

Algorithm 4 applies conjugate gradients to this normal equation on the full coefficient space.

Algorithm 4: Coefficient-space frame-operator CG reconstruction

Input: Coefficient collection $\mathbf{c}$, tolerance $\tau > 0$, and maximum iteration number $M_{\max}$
Output: Canonical finest-level coefficient vector $\widehat{\mathbf{u}}_K \in \mathbb{C}^{\Lambda^{\alpha}_K}$

$\mathbf{b} \leftarrow \mathcal{A}^*_{J,K}\mathbf{c}$, $\widehat{\mathbf{u}}_K \leftarrow 0$;
$\mathbf{r} \leftarrow \mathbf{b}$, $\mathbf{p} \leftarrow \mathbf{r}$;
for $m = 0, 1, \ldots, M_{\max} - 1$ **do**
  if $\|\mathbf{r}\| / \max\{\|\mathbf{b}\|, \epsilon_{\mathrm{mach}}\} \leq \tau$ **then**
    return $\widehat{\mathbf{u}}_K$
  end if
  $\mathbf{q} \leftarrow \mathcal{A}^*_{J,K}\mathcal{A}_{J,K}\mathbf{p}$;
  $\theta_m \leftarrow \langle \mathbf{r}, \mathbf{r} \rangle / \langle \mathbf{p}, \mathbf{q} \rangle$;
  $\widehat{\mathbf{u}}_K \leftarrow \widehat{\mathbf{u}}_K + \theta_m \mathbf{p}$;
  $\mathbf{r}_{\mathrm{new}} \leftarrow \mathbf{r} - \theta_m \mathbf{q}$;
  $\delta_m \leftarrow \langle \mathbf{r}_{\mathrm{new}}, \mathbf{r}_{\mathrm{new}} \rangle / \langle \mathbf{r}, \mathbf{r} \rangle$;
  $\mathbf{p} \leftarrow \mathbf{r}_{\mathrm{new}} + \delta_m \mathbf{p}$, $\mathbf{r} \leftarrow \mathbf{r}_{\mathrm{new}}$;
end for

The Marcinkiewicz–Zygmund bounds do more than ensure that the least-squares problem is well posed. They quantify the conditioning of its normal equation and therefore the convergence of conjugate gradients.

Proposition 5.4 (Finite-level near-tightness and CG reconstruction)**.** *Assume that the refinement relations* (5.4) *hold and that* (5.6) *holds for* $j = J+1, \ldots, K$.

(i) *For the fixed-tolerance nested Marcinkiewicz–Zygmund measures of Theorem 3.1, with* $p = 2$ *and tolerance* $0 < \eta < 1$,

$$(1-\eta)\|\widehat{\mathbf{z}}\|_2^2 \le \|\mathcal{A}_{J,K}\widehat{\mathbf{z}}\|_2^2 \le (1+\eta)\|\widehat{\mathbf{z}}\|_2^2 \quad \textit{for all } \widehat{\mathbf{z}} \in \mathbb{C}^{\Lambda_K^\alpha}.$$

(ii) *In the oversampled setting of Theorem 4.2, with* $\rho_s = C_{\mathcal{M}} 2^{-s} < 1$*, the corresponding operator* $\mathcal{A}_{J,K}^{(s)}$ *satisfies the same inequality with* η *replaced by* ρ_s.

Both conclusions remain valid after restriction to a spectral subspace. In either case, let $\mathcal{T} = \mathcal{A}_{J,K}$ *in* (i) *and* $\mathcal{T} = \mathcal{A}_{J,K}^{(s)}$ *in* (ii)*, and write* $\varepsilon = \eta$ *or* $\varepsilon = \rho_s$*, respectively. Then* $\mathbf{S}_{\mathcal{T}} := \mathcal{T}^*\mathcal{T}$ *is a positive-definite self-adjoint operator on* $\mathbb{C}^{\Lambda_K^\alpha}$*. Let* $\lambda_{\min}(\mathbf{S}_{\mathcal{T}})$ *and* $\lambda_{\max}(\mathbf{S}_{\mathcal{T}})$ *denote its smallest and largest eigenvalues, respectively. Its spectral condition number is*

$$\kappa(\mathbf{S}_{\mathcal{T}}) := \frac{\lambda_{\max}(\mathbf{S}_{\mathcal{T}})}{\lambda_{\min}(\mathbf{S}_{\mathcal{T}})} \le \frac{1+\varepsilon}{1-\varepsilon}. \tag{5.8}$$

Denote by $\|\widehat{\mathbf{z}}\|_{\mathbf{S}_{\mathcal{T}}} := \langle \mathbf{S}_{\mathcal{T}}\widehat{\mathbf{z}}, \widehat{\mathbf{z}}\rangle^{1/2}$*. If* $\widehat{\mathbf{u}}_K$ *solves* $\mathbf{S}_{\mathcal{T}}\widehat{\mathbf{u}}_K = \mathcal{T}^*\mathbf{c}$ *and* $\widehat{\mathbf{u}}_K^{(m)}$ *is the mth CG iterate (using* $\mathcal{T}$ *in place of* $\mathcal{A}_{J,K}$ *in Algorithm 4), then*

$$\|\widehat{\mathbf{u}}_K^{(m)} - \widehat{\mathbf{u}}_K\|_{\mathbf{S}_{\mathcal{T}}} \le 2\left(\frac{\sqrt{\kappa(\mathbf{S}_{\mathcal{T}})}-1}{\sqrt{\kappa(\mathbf{S}_{\mathcal{T}})}+1}\right)^m \|\widehat{\mathbf{u}}_K^{(0)} - \widehat{\mathbf{u}}_K\|_{\mathbf{S}_{\mathcal{T}}}. \tag{5.9}$$

Proof. For $\widehat{\mathbf{z}} \in \mathbb{C}^{\Lambda_K^\alpha}$, iterating the one-level filter-bank energy identity gives

$$\|\widehat{\mathbf{z}}\|_2^2 = \|\widehat{\mathbf{v}}_J\|_2^2 + \sum_{j=J+1}^{K}\sum_{k=1}^{r} \|\widehat{\mathbf{w}}_{j-1}^k\|_2^2.$$

The polynomial represented by $\widehat{\mathbf{v}}_J$ belongs to $\mathbb{P}_{2^J}$, while the polynomial represented by $\widehat{\mathbf{w}}_{j-1}^k$ belongs to $\mathbb{P}_{2^j}$. In case (i), the levelwise Marcinkiewicz–Zygmund inequalities therefore give

$$(1-\eta)\|\widehat{\mathbf{z}}\|_2^2 \le \|\mathcal{A}_{J,K}\widehat{\mathbf{z}}\|_2^2 \le (1+\eta)\|\widehat{\mathbf{z}}\|_2^2$$

after summing over the coarse and detail branches. In case (ii), the same argument samples a degree-2^j branch on $\boldsymbol{X}_{j+s}$. Theorem 3.2 gives the uniform error $\rho_s = C_{\mathcal{M}}2^{-s}$, and hence the asserted bounds for $\mathcal{A}_{J,K}^{(s)}$. Restricting either operator to a spectral subspace preserves these inequalities. The resulting frame bounds imply $(1-\varepsilon)I \le \mathbf{S}_{\mathcal{T}} \le (1+\varepsilon)I$, which proves positive definiteness and (5.8). Finally, (5.9) is the standard conjugate-gradient estimate for a positive-definite linear system; see, e.g., [44, Lecture 38]. □

5.5 Fast algorithms and computational cost

Fast algorithms for Fourier maps and their adjoints exist on many standard manifolds, including fast spherical harmonic transforms on the sphere and nonequispaced Fourier transforms on the torus; see, e.g., [21, 26, 28, 30, 31, 38]. Let $\mathsf{N}_j := |\boldsymbol{X}_j|$, $M_j := |\Lambda_j^\alpha|$, and let $\mathfrak{C}_j$ be an upper bound for the cost of one application of either $\mathbf{F}_j$ or $\mathbf{F}_j^*$. The estimates below assume a fixed fast-transform accuracy and exclude one-time planning and precomputation costs. Diagonal filter multiplication costs $O(M_j)$.

If I_{fit} iterations are used in Algorithm 1, finest-level coefficient preparation costs

$$O\big((2I_{\text{fit}}+1)\mathfrak{C}_K + I_{\text{fit}}M_K\big).$$

A coefficient analysis pass costs

$$O\left(r\sum_{j=J+1}^{K}\mathfrak{C}_j + \mathfrak{C}_J + (r+1)\sum_{j=J+1}^{K} M_j\right),$$

and an adjoint synthesis pass has the same order. If I_{rec} coefficient-space CG iterations are used, canonical reconstruction costs one initial adjoint synthesis plus I_{rec} analysis–synthesis pairs and

$O(I_{\mathrm{rec}} M_K)$ additional vector operations. Projection of a spectral coefficient vector onto a subspace such as $\mathbb{P}_{2^L}$ costs $O(M_K)$. For fixed-tolerance dyadic measures, the constants may depend on the fixed tolerance η, but $\mathsf{N}_j \asymp M_j \asymp 2^{jd}$. Here the estimate for N_j follows from the nested construction, while that for M_j follows from Weyl's law: continuity and $\widehat{\alpha}(0) \neq 0$ make $\widehat{\alpha}$ nonzero near the origin, and its compact support gives the matching upper bound.

For fixed dimension and transform accuracy, the torus nonequispaced fast Fourier transforms (NFFT) and its adjoint satisfy $\mathfrak{C}_j = O(M_j \log M_j + \mathsf{N}_j)$ [31]. On $\mathbb{S}^2$, letting $M_j \asymp n_j^2$, the nonequispaced fast spherical Fourier transform (NFSFT) and its adjoint satisfy $\mathfrak{C}_j = O\big(n_j^2 \log^2 n_j + \mathsf{N}_j\big) = O\big(M_j \log^2 M_j + \mathsf{N}_j\big)$ [30]. The diagonal weight scaling costs $O(\mathsf{N}_j)$ and does not change these estimates. Thus both cases are covered by

$$\mathfrak{C}_j = O(M_j \log^\sigma M_j + \mathsf{N}_j)\,, \qquad \sigma = \begin{cases} 1, & \mathcal{M} = \mathbb{T}^d, \\ 2, & \mathcal{M} = \mathbb{S}^2. \end{cases}$$

Since $M_j \asymp \mathsf{N}_j$ and both grow geometrically, one analysis or adjoint synthesis pass costs

$$O((r+1)\mathsf{N}_K \log^\sigma \mathsf{N}_K)\,.$$

Hence I_{rec} coefficient-space CG iterations cost this quantity times I_{rec}, up to the initial adjoint synthesis, and the coefficient storage is $O((r+1)\mathsf{N}_K)$. For an oversampling shift s, $\mathsf{N}_{j+s} \asymp 2^{sd}\mathsf{N}_j$, while M_j is unchanged. Therefore

$$\mathfrak{C}_j^{(s)} = O(M_j \log^\sigma M_j + \mathsf{N}_{j+s})\,,$$

and one oversampled analysis or adjoint synthesis pass costs

$$O\big((r+1)\big(M_K \log^\sigma M_K + \mathsf{N}_{K+s}\big)\big)\,.$$

The corresponding CG cost is this quantity times I_{rec}, up to the initial adjoint synthesis, and the storage is $O((r+1)\mathsf{N}_{K+s})$. Thus oversampling increases the node-dependent cost and storage by a factor of order 2^{sd}.

6 Approximation error

The fully discrete transform reconstructs an arbitrary vector in the terminal coefficient space. For function approximation, we instead prescribe a target polynomial space $\mathbb{P}_{2^L}$ and restrict the reconstruction to that space. Let

$$\tau_\alpha := \max\{\tau > 0 : \widehat{\alpha}(\xi) = 1 \text{ for } 0 \le \xi \le \tau\} > 0$$

and

$$q_\alpha := \min\{q \in \mathbb{N}_0 : 2^{-q} \le \tau_\alpha\}.$$

We analyze $\mathbb{P}_{2^L}$ from the terminal level $K_\alpha := L + q_\alpha$, where we assume $K_\alpha > J$. Then

$$\widehat{\alpha}\left(\lambda_\ell / 2^{K_\alpha}\right) = 1, \qquad \lambda_\ell \le 2^L. \tag{6.1}$$

Consequently, the terminal scaling coefficients of $P \in \mathbb{P}_{2^L}$ are its ordinary Fourier coefficients, extended by zero to $\Lambda^\alpha_{K_\alpha}$. We denote this coefficient vector by $\widehat{\mathbf{P}}^{K_\alpha}$, and define the restricted analysis map

$$\mathcal{B}_{J,L} P := \mathcal{A}_{J,K_\alpha} \widehat{\mathbf{P}}^{K_\alpha}, \qquad P \in \mathbb{P}_{2^L}. \tag{6.2}$$

Its adjoint is the one-pass synthesis followed by spectral truncation:

$$\mathcal{B}^*_{J,L}\mathbf{c} = \sum_{\lambda_\ell \le 2^L} \big(\mathcal{A}^*_{J,K_\alpha}\mathbf{c}\big)_\ell u_\ell.$$

The canonical reconstruction can be computed by Algorithm 4 with $\mathcal{B}_{J,L}$ in place of $\mathcal{A}_{J,K_\alpha}$. In each application of $\mathcal{B}^*_{J,L}$, the coefficients with $\lambda_\ell > 2^L$ are discarded. This projection is part of the restricted adjoint and ensures that the CG iteration remains in $\mathbb{P}_{2^L}$.

Theorem 6.1 (Polynomial reconstruction and approximation)**.** *Let $J_0 \in \mathbb{N}_0$ be an integer and $J \geq J_0$. Let $K_\alpha = L + q_\alpha > J$, and suppose that $0 \leq \varepsilon < 1$ and*

$$(1-\varepsilon)\|P\|_{L^2(\mu)}^2 \leq \|\mathcal{B}_{J,L}P\|_2^2 \leq (1+\varepsilon)\|P\|_{L^2(\mu)}^2, \qquad P \in \mathbb{P}_{2^L}.$$

For $f \in L^2(\mathcal{M},\mu)$, let

$$f_L := \sum_{\lambda_\ell \leq 2^L} \widehat{f}_\ell u_\ell$$

be its $L^2(\mu)$-orthogonal projection onto $\mathbb{P}_{2^L}$, and consider the data model

$$\mathbf{c} = \mathcal{B}_{J,L} f_L + \mathbf{e}, \tag{6.3}$$

where $\mathbf{e}$ encodes potential processing of the framelet coefficients. Define the one-pass reconstruction by

$$\widetilde{f}_L := \mathcal{B}_{J,L}^* \mathbf{c}$$

and the canonical reconstruction by

$$u_L := \operatorname*{argmin}_{P \in \mathbb{P}_{2^L}} \|\mathcal{B}_{J,L}P - \mathbf{c}\|_2^2.$$

Then

$$\|\widetilde{f}_L - f\|_{L^2(\mu)} \leq \|f - f_L\|_{L^2(\mu)} + \varepsilon\|f_L\|_{L^2(\mu)} + (1+\varepsilon)^{1/2}\|\mathbf{e}\|_2, \tag{6.4}$$

and

$$\|u_L - f\|_{L^2(\mu)} \leq \|f - f_L\|_{L^2(\mu)} + \frac{1}{\sqrt{1-\varepsilon}}\|\mathbf{e}\|_2. \tag{6.5}$$

Proof. Set $\mathcal{S}_L := \mathcal{B}_{J,L}^*\mathcal{B}_{J,L}$. The assumed bounds imply

$$(1-\varepsilon)I \leq \mathcal{S}_L \leq (1+\varepsilon)I.$$

Hence $\|\mathcal{S}_L - I\| \leq \varepsilon$ and $\|\mathcal{B}_{J,L}^*\| = \|\mathcal{B}_{J,L}\| \leq (1+\varepsilon)^{1/2}$. For the one-pass reconstruction, the data model (6.3) gives

$$\widetilde{f}_L = \mathcal{B}_{J,L}^*\mathbf{c} = \mathcal{S}_L f_L + \mathcal{B}_{J,L}^*\mathbf{e}.$$

Consequently, $\widetilde{f}_L - f_L = (\mathcal{S}_L - I)f_L + \mathcal{B}_{J,L}^*\mathbf{e}$, and therefore

$$\|\widetilde{f}_L - f_L\|_{L^2(\mu)} \leq \varepsilon\|f_L\|_{L^2(\mu)} + (1+\varepsilon)^{1/2}\|\mathbf{e}\|_2.$$

Combining this estimate with $\|\widetilde{f}_L - f\|_{L^2(\mu)} \leq \|\widetilde{f}_L - f_L\|_{L^2(\mu)} + \|f_L - f\|_{L^2(\mu)}$ proves (6.4). For the canonical reconstruction, its normal equation $(\mathcal{B}_{J,L}^*\mathcal{B}_{J,L})u_L = \mathcal{B}_{J,L}^*\mathbf{c}$ and the data model (6.3) yield $\mathcal{S}_L u_L = \mathcal{B}_{J,L}^*\mathbf{c} = \mathcal{S}_L f_L + \mathcal{B}_{J,L}^*\mathbf{e}$, hence

$$\mathcal{S}_L(u_L - f_L) = \mathcal{B}_{J,L}^*\mathbf{c} - \mathcal{B}_{J,L}^*\mathcal{B}_{J,L}f_L = \mathcal{B}_{J,L}^*\mathbf{e}.$$

Since $\mathcal{S}_L$ is positive definite on $\mathbb{P}_{2^L}$ and hence invertible, it follows

$$u_L - f_L = \mathcal{S}_L^{-1}\mathcal{B}_{J,L}^*\mathbf{e}.$$

Moreover,

$$\big(\mathcal{S}_L^{-1}\mathcal{B}_{J,L}^*\big)\big(\mathcal{S}_L^{-1}\mathcal{B}_{J,L}^*\big)^* = \mathcal{S}_L^{-1}\mathcal{B}_{J,L}^*\mathcal{B}_{J,L}\mathcal{S}_L^{-1} = \mathcal{S}_L^{-1}\mathcal{S}_L\mathcal{S}_L^{-1} = \mathcal{S}_L^{-1}.$$

Hence $\|\mathcal{S}_L^{-1}\mathcal{B}_{J,L}^*\| \leq (1-\varepsilon)^{-1/2}$, and therefore $\|u_L - f_L\|_{L^2(\mu)} \leq \|\mathbf{e}\|_2/\sqrt{1-\varepsilon}$. The triangle inequality with $f - f_L$ now proves (6.5). ☐

Remark 6.1. *If $\mathbf{e} = 0$, meaning that the decomposed coefficients are not perturbed or otherwise modified, then $u_L = f_L$. In particular, every $f \in \mathbb{P}_{2^L}$ is reconstructed exactly in this case.*

Remark 6.2 (Finite-iteration CG error)**.** *The function u_L above is the exact solution of the restricted normal equation. In computation, a conjugate-gradients method produces an iterate $u_L^{(m)}$ with residual*

$$\mathbf{r}_m := \mathcal{B}_{J,L}^*\mathbf{c} - \mathcal{S}_L u_L^{(m)} = \mathcal{S}_L(u_L - u_L^{(m)}).$$

Therefore, $u_L^{(m)} - u_L = -\mathcal{S}_L^{-1}\mathbf{r}_m$. Using $\|\mathcal{S}_L^{-1}\| \le (1-\varepsilon)^{-1}$,

$$\|u_L^{(m)} - u_L\|_{L^2(\mu)} \le \frac{\|\mathbf{r}_m\|_2}{1-\varepsilon}.$$

Combining this estimate with (6.5) gives the total error bound for the finite-iteration reconstruction

$$\|u_L^{(m)} - f\|_{L^2(\mu)} \le \|f - f_L\|_{L^2(\mu)} + \frac{\|\mathbf{e}\|_2}{\sqrt{1-\varepsilon}} + \frac{\|\mathbf{r}_m\|_2}{1-\varepsilon}.$$

Thus the total error consists of the polynomial approximation error, the coefficient perturbation error, and the finite-iteration CG error. Moreover, the spectral bounds imply $\kappa(\mathcal{S}_L) \le (1+\varepsilon)/(1-\varepsilon)$, so a smaller frame-bound deviation also improves the conditioning of the CG system.

Corollary 6.1 (Polynomial reconstruction for nested Marcinkiewicz–Zygmund measures)**.** *Let $K_\alpha = L + q_\alpha > J$, and use the notation of Theorem 6.1.*

(i) *For the basic system in Definition 4.2, with $p = 2$ and tolerance η, Proposition 5.4*(i) *gives $\varepsilon = \eta$. Hence*

$$\|u_L - f\|_{L^2(\mu)} \le \|f - f_L\|_{L^2(\mu)} + \frac{\|\mathbf{e}\|_2}{\sqrt{1-\eta}}.$$

(ii) *For the oversampled system in Definition 4.3, let $\mathcal{B}_{J,L}^{(s)}$ denote the corresponding restricted analysis map, and suppose that $\rho_s = C_{\mathcal{M}} 2^{-s} < 1$. By Proposition 5.4*(ii)*, this map has $\varepsilon = \rho_s$. If $\mathbf{c} = \mathcal{B}_{J,L}^{(s)} f_L + \mathbf{e}$, its canonical reconstruction $u_L^{(s)}$ satisfies*

$$\|u_L^{(s)} - f\|_{L^2(\mu)} \le \|f - f_L\|_{L^2(\mu)} + \frac{\|\mathbf{e}\|_2}{\sqrt{1-\rho_s}}.$$

The error term $\|f - f_L\|_{L^2(\mu)}$ in these estimate can be refined if f has extra smoothness.

Remark 6.3 (Sobolev approximation [29])**.** *If $t > 0$ and f belongs to the spectral Sobolev space*

$$H^t(\mathcal{M}) := \left\{ f \in L^2(\mathcal{M}) : \sum_{\ell \ge 0} (1 + \lambda_\ell^2)^t |\widehat{f}_\ell|^2 < \infty \right\},$$

then $\|f - f_L\|_{L^2(\mu)} \le 2^{-Lt}\|f\|_{H^t(\mathcal{M})}$.

7 Multiscale data analysis

We then illustrate the nested framelet transform on the sphere and on the torus, focusing on oversampling, multiscale decomposition, and reconstruction.

7.1 Common numerical setup

We consider the two-high-pass filter bank of Wang and Zhuang in [45], built from the smooth cutoff used in [18, Chapter 4]. Let

$$\vartheta(t) = \begin{cases} 0, & t < 0, \\ t^4(35 - 84t + 70t^2 - 20t^3), & 0 \le t \le 1, \\ 1, & t > 1. \end{cases}$$

On the frequency interval $0 \le |\xi| \le 1/2$, the filters are defined by

$$
\begin{aligned}
\widehat{a}(\xi) &= \begin{cases} 1, & |\xi| < 1/8, \\ \cos\left(\frac{\pi}{2}\vartheta(8|\xi|-1)\right), & 1/8 \le |\xi| \le 1/4, \\ 0, & 1/4 < |\xi| \le 1/2, \end{cases} \\
\widehat{b^1}(\xi) &= \begin{cases} 0, & |\xi| < 1/8, \\ \sin\left(\frac{\pi}{2}\vartheta(8|\xi|-1)\right), & 1/8 \le |\xi| \le 1/4, \\ \cos\left(\frac{\pi}{2}\vartheta(4|\xi|-1)\right), & 1/4 < |\xi| \le 1/2, \end{cases} \\
\widehat{b^2}(\xi) &= \begin{cases} 0, & |\xi| < 1/4, \\ \sin\left(\frac{\pi}{2}\vartheta(4|\xi|-1)\right), & 1/4 \le |\xi| \le 1/2. \end{cases}
\end{aligned}
\tag{7.1}
$$

The associated generators obtained from the refinement equations (5.4) are

$$
\begin{aligned}
\widehat{\alpha}(\xi) &= \begin{cases} 1, & |\xi| < 1/4, \\ \cos\left(\frac{\pi}{2}\vartheta(4|\xi|-1)\right), & 1/4 \le |\xi| \le 1/2, \\ 0, & \text{otherwise}, \end{cases} \\
\widehat{\beta^1}(\xi) &= \begin{cases} \sin\left(\frac{\pi}{2}\vartheta(4|\xi|-1)\right), & 1/4 \le |\xi| < 1/2, \\ \cos^2\left(\frac{\pi}{2}\vartheta(2|\xi|-1)\right), & 1/2 \le |\xi| \le 1, \\ 0, & \text{otherwise}, \end{cases} \\
\widehat{\beta^2}(\xi) &= \begin{cases} \cos\left(\frac{\pi}{2}\vartheta(2|\xi|-1)\right)\sin\left(\frac{\pi}{2}\vartheta(2|\xi|-1)\right), & 1/2 \le |\xi| \le 1, \\ 0, & \text{otherwise}. \end{cases}
\end{aligned}
\tag{7.2}
$$

Figure 2 displays the filter symbols and the associated generator functions.

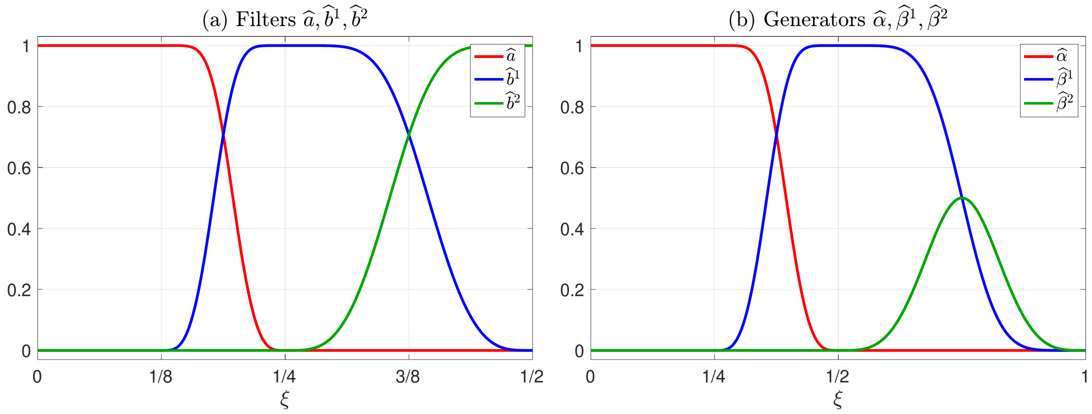

Figure 2: A two-high-pass filter bank and the associated framelet generators.

The assumptions used in the preceding sections are readily checked for this choice. The filters in (7.1) satisfy

$$
|\widehat{a}(\xi)|^2 + |\widehat{b^1}(\xi)|^2 + |\widehat{b^2}(\xi)|^2 = 1, \qquad 0 \le |\xi| \le 1/2. \tag{7.3}
$$

This follows piecewise from $\cos^2\theta + \sin^2\theta = 1$, and, together with the refinement relations (5.4), gives (4.5). The support condition (4.9) is immediate from (7.2) since $\operatorname{supp}\widehat{\alpha} \subset [-1/2, 1/2]$ and $\operatorname{supp}\widehat{\beta^1} \cup \operatorname{supp}\widehat{\beta^2} \subset [-1, 1]$. Moreover, $\widehat{\alpha}(\xi) = 1$ for $0 \le |\xi| \le 1/4$. Hence the low-pass limit condition (4.4) holds and $\tau_\alpha = 1/4$. In the coefficient formulation of Section 5 this gives $q_\alpha = 2$: a target polynomial space $\mathbb{P}_{2^L}$ is analyzed from finest scaling level $K = L + 2$. At that level the finest-level scaling multiplier is identically one on every frequency $\lambda_\ell \le 2^L$, as stated in (6.1). This finest-level offset is used in the polynomial reconstruction and approximation results of Section 6. The finite numerical experiments below instead keep the prescribed finest bandwidth fixed: they first compute a weighted least-squares projection in that space and then apply the coefficient filter bank directly, without adding two finest levels.

Both CG routines start from zero and terminate when the relative residual of the corresponding normal equation satisfies

$$\frac{\|\mathbf{r}_m\|_2}{\max\{\|\mathbf{b}\|_2, \epsilon_{\text{mach}}\}} \leq \tau. \tag{7.4}$$

or when the prescribed iteration cap is reached. Here $H\mathbf{z} = \mathbf{b}$ is either the finest-level least-squares normal equation, with $H = \mathbf{F}^*\mathbf{F}$, or the canonical frame-operator equation, with $H = \mathcal{A}^*\mathcal{A}$. We use $\tau_{\text{fit}} = 10^{-10}$ for every finest-level least-squares projection. The canonical reconstruction tolerance is $\tau_{\text{rec}} = 10^{-9}$ in the following Examples 1–3 and $\tau_{\text{rec}} = 10^{-8}$ in Example 4. The corresponding iteration caps $(M_{\text{fit}}, M_{\text{rec}})$ are $(140, 50)$, $(120, 40)$, $(80, 40)$, and $(80, 25)$, respectively.

All numerical experiments were performed in MATLAB R2024a on a MacBook Pro (16-inch, 2019) with a 2.6 GHz 6-core Intel Core i7 processor and 16 GB RAM. The fast transforms were evaluated with the NFFT 3.5.3 MATLAB package [31] via its macOS MEX interfaces; the spherical experiments use its NFSFT routines when available, and the torus experiments use its two-dimensional NFFT routines.

7.2 On the sphere

For the spherical experiments, at a dyadic level j we take $n_j = 2^j$ as the spherical harmonic degree cutoff and choose $|\boldsymbol{X}_j| = \lceil m(2^j + 1)^2 \rceil = \lceil m(n_j + 1)^2 \rceil$, where $m \geq 1$ is the oversampling factor. The nodes are generated from a large deterministic Fibonacci candidate set by a greedy farthest-point ordering with chordal distance; the sets $\boldsymbol{X}_j$ are nested prefixes of this ordering with the cardinalities specified above. Given the nested node set $\boldsymbol{X}_j$, we use the surface-area version of the nearest-neighbor partition weights in Definition 3.1. Numerically, a dense Fibonacci proxy set $\{z_q\}_{q=1}^{Q}$ is assigned to the nearest sample points and we set

$$w_{j,x} = \frac{4\pi}{Q} \#\{q : x \text{ is nearest to } z_q\}.$$

On the unit sphere, nearest neighbors in chordal and geodesic distance coincide, so these weights approximate the areas of the corresponding spherical nearest-neighbor cells. Here we use the standard surface-area normalization $\sum_{x \in \boldsymbol{X}_j} w_{j,x} = 4\pi$; this differs from the probability normalization in the theoretical sections only by a constant factor. These farthest-point prefixes with partition-type weights are not spherical designs or exact quadrature rules; they are used here as simple quasi-uniform nested Marcinkiewicz–Zygmund-type sampling rules.

Example 1: Oversampling for a Wendland test function. We first test the effect of oversampling on reconstruction. The signal is a sum of Wendland-type bumps [12] on $\mathbb{S}^2$:

$$f(x) = \sum_{c \in \{\pm e_1, \pm e_2, \pm e_3\}} \phi(\|x - c\|_2),$$

where

$$\phi(r) = \frac{(1-r)_+^{10} \left(429r^4 + 450r^3 + 210r^2 + 50r + 5\right)}{5\tau_4}, \qquad \tau_4 = \frac{15\,\Gamma(9/2)}{2\,\Gamma(5)}.$$

Here e_1, e_2, e_3 are the coordinate vectors in $\mathbb{R}^3$, and $(t)_+ = \max\{t, 0\}$. We use two finest polynomial degrees, $n = 16$ and $n = 32$, corresponding respectively to the dyadic levels $j = 4, 3, 2$ and $j = 5, 4, 3$. For each fixed degree, we vary the oversampling factor over the twelve values $m_i = 1.2 + 0.1i(i+1)$ for $i = 1, \ldots, 12$.

Figure 3 shows the Wendland signal, the one-pass reconstruction and its pointwise error, and the canonical CG reconstruction and its pointwise error, for the case $n = 32$ with oversampling factor $m = 1.4$. The pointwise errors are measured against the bandlimited projection on an independent Fibonacci test grid.

Figure 4 reports the corresponding relative errors as the number of finest-level nodes increases. The one-pass synthesis error decreases as the sampling density increases, while the canonical CG reconstruction error is several orders of magnitude smaller. Its eventual scale is compatible with the reconstruction residual tolerance $\tau_{\text{rec}} = 10^{-9}$, after allowing for the conditioning and the fast-transform error.

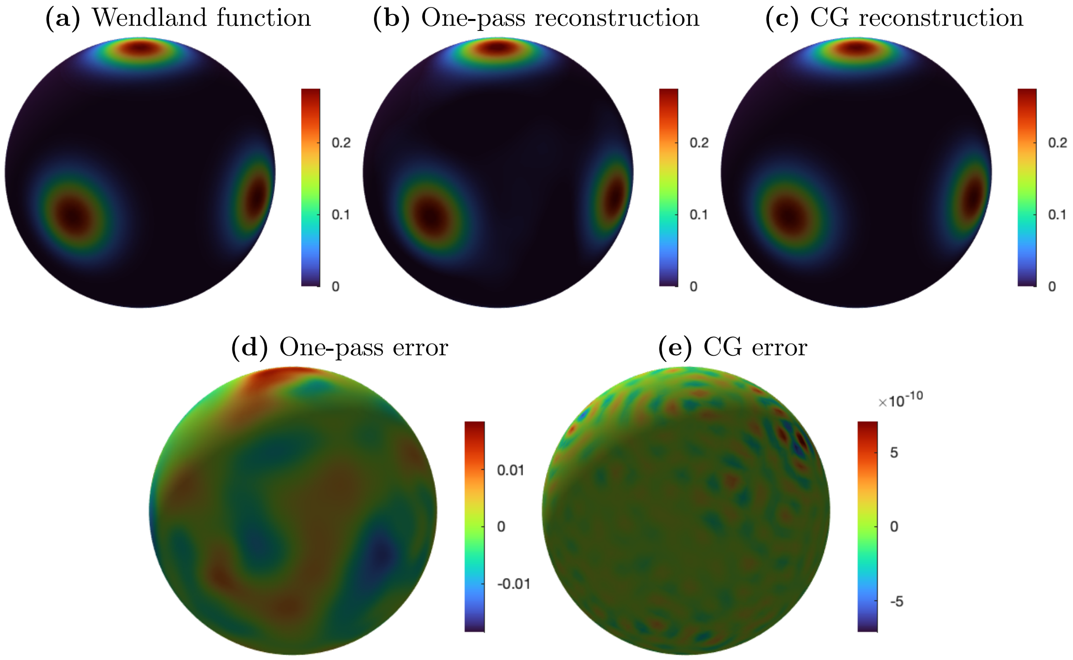

Figure 3: Wendland-type signal reconstructions on $\mathbb{S}^2$, for $n = 32$ and oversampling factor $m = 1.4$.

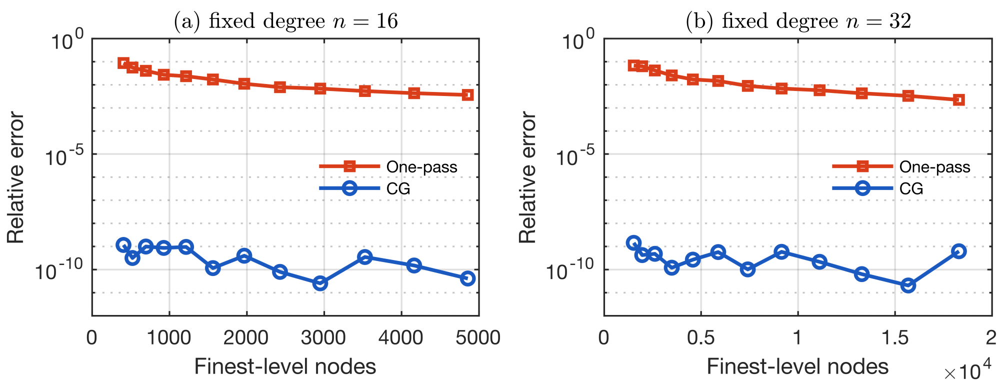

Figure 4: Relative reconstruction error for the Wendland-type test function as the number of finest-level nodes increases.

Example 2: Multiscale analysis of an Earth texture signal. We next apply the same transform to an Earth texture signal. The input is obtained from a Natural Earth texture map* by converting the image to grayscale and sampling it at the finest nested nodes. The experiment therefore treats image intensity as a scalar function on $\mathbb{S}^2$, not as physical elevation data. We use levels $j = 7, 6, 5$, with spherical harmonic degree cutoffs $n_7 = 128$, $n_6 = 64$, and $n_5 = 32$. With oversampling factor $m = 2.5$, the nested farthest-point prefixes have cardinalities $|\boldsymbol{X}_7| = 41603$, $|\boldsymbol{X}_6| = 10563$, and $|\boldsymbol{X}_5| = 2723$. The ordering is selected from 249618 Fibonacci candidates; a separate Fibonacci proxy set with 499236 points is used only to approximate the nearest-neighbor cell weights.

Let $\mathbf{v}$ denote the sampled texture data on $\boldsymbol{X}_7$. We compute its weighted least-squares projection $\mathbf{v}_7$ onto spherical harmonics of degree at most 128, using conjugate gradients with tolerance 10^{-10} and at most 120 iterations. The residual is $\mathbf{w}_7 = \mathbf{v} - \mathbf{v}_7$. Starting from the finest spectral coefficient vector, two filter-bank analysis steps produce $\mathbf{v}_6, \mathbf{w}_6^1, \mathbf{w}_6^2$ and then $\mathbf{v}_5, \mathbf{w}_5^1, \mathbf{w}_5^2$. The projection relative error is 1.5885×10^{-1}. Relative to the projected signal $\mathbf{v}_7$, the one-pass synthesis error is 3.4203×10^{-2}, while the canonical coefficient-space CG reconstruction reduces this error to 6.4390×10^{-10}. The corresponding relative error against the original image data remains 1.5885×10^{-1}, with SNR 15.888 dB. These values are consistent with (7.4): the canonical error relative to the projection is at the algebraic solver scale, whereas the much larger error relative to the image is the bandlimited projection error and is unaffected by further CG iterations. These quantities are shown in Figure 5: the approximation terms retain the large-scale texture, while the detail terms capture oscillatory features at the corresponding dyadic resolutions.

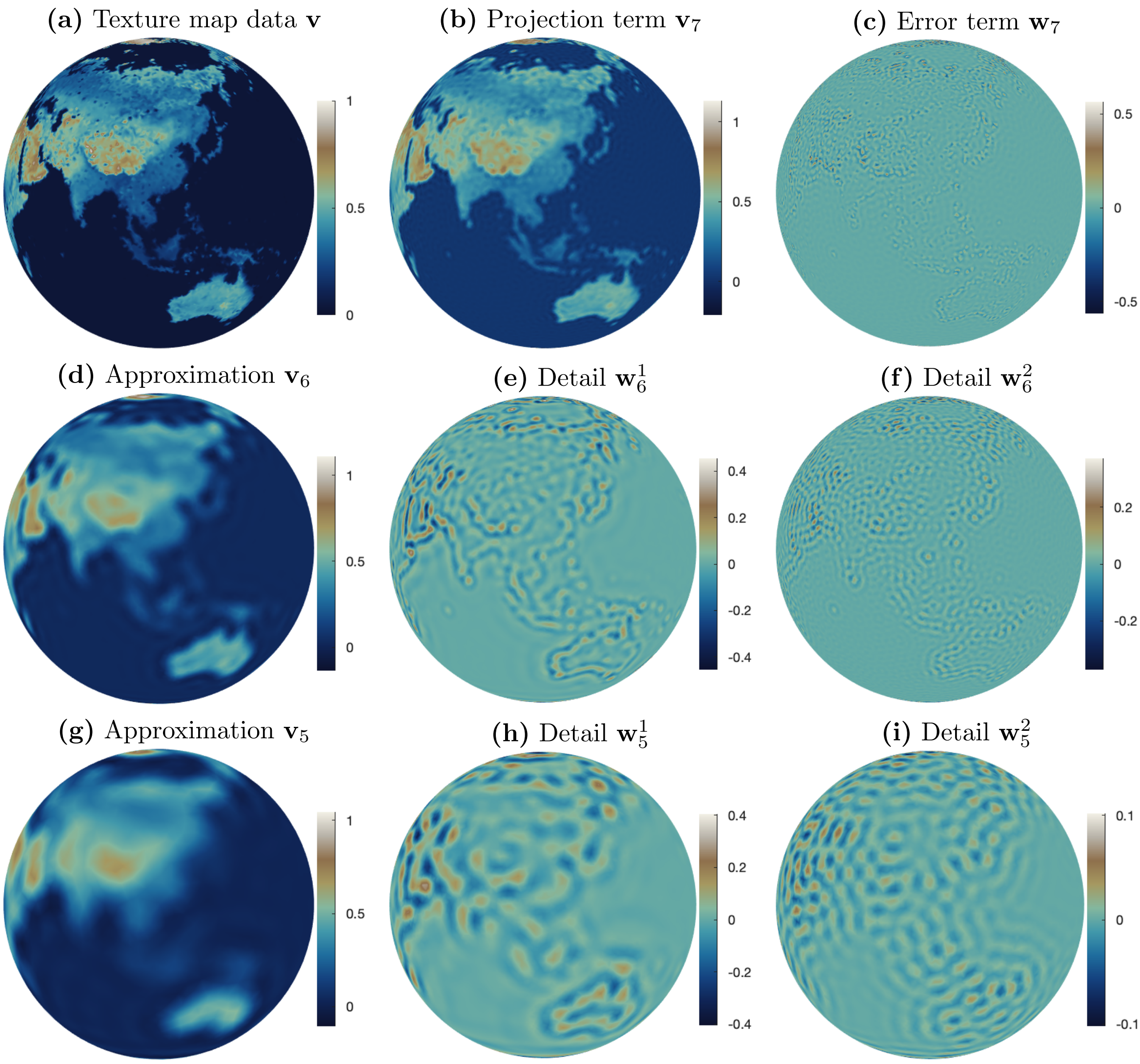

Figure 5: Multiscale decomposition of an Earth texture signal on nested spherical sampling sets.

*Adapted from Natural Earth III by Tom Patterson (`https://www.shadedrelief.com`); public domain.

7.3 On the torus

We also consider the flat torus $\mathbb{T}^2 = [0,1)^2$, with periodic boundary conditions and normalized Lebesgue measure. With the 2π-Fourier convention, $u_k(x) = e^{2\pi i k\cdot x}$, $k \in \mathbb{Z}^2$, we have $-\Delta u_k = (2\pi)^2|k|^2 u_k = \lambda_k^2 u_k$ with $\lambda_k = 2\pi|k|$, and we write $\mathbb{P}_n(\mathbb{T}^2) = \operatorname{span}\{e^{2\pi i k\cdot x} : |k| \le n\}$, where $D_n = \#\{k \in \mathbb{Z}^2 : |k| \le n\}$. For dyadic degrees $n_j = 2^j$, we measure oversampling relative to D_{n_j}. Given $m \ge 1$, set $q_J = \lceil \sqrt{mD_{n_J}} \rceil$ and $q_j = 2^{j-J} q_J$, so that $|\boldsymbol{X}_j| = q_j^2$ has the same order as mD_{n_j}. The set $\boldsymbol{X}_j$ is a $q_j \times q_j$ periodic tensor-cell rule with one jittered point per cell: each grid center is randomly displaced by a fixed fraction of the cell width and then reduced modulo one. The finest jittered array is generated first, and coarser levels are dyadic subarrays, so $\boldsymbol{X}_J \subset \cdots \subset \boldsymbol{X}_L$. We use the cell-area weights $w_{j,x} = q_j^{-2}$. These are not exact tensor-product quadrature rules, but simple quasi-uniform nested Marcinkiewicz–Zygmund-type sampling rules. For visualization, the flat torus data are displayed on the standard embedded torus in $\mathbb{R}^3$.

Example 3: Multiscale analysis on an oscillatory signal. We use dyadic levels $j = 6, 5, 4$, with $n_6 = 64$, $n_5 = 32$, and $n_4 = 16$. The oversampling factor is $m = 2.5$, the jitter parameter is 0.25, and the random seed is fixed at 5. This gives $q_6 = 180$, $q_5 = 90$, and $q_4 = 45$, and hence $|\boldsymbol{X}_6| = 32400$, $|\boldsymbol{X}_5| = 8100$, and $|\boldsymbol{X}_4| = 2025$. The test signal is a synthetic periodic function consisting of low-frequency trigonometric modes, localized periodic Gaussian bumps, and two high-frequency oscillatory modes. More precisely, before normalization to the range $[0,1]$, we use

$$
\begin{aligned}
f(u,v) = 0.32\cos(2\pi u) + 0.22\sin(2\pi v) + 0.18\cos(2\pi(u-v)) + \sum_{\ell=1}^{4} A_\ell \exp\left(-\frac{d_{\mathbb{T}^2}((u,v),c_\ell)^2}{2\sigma_\ell^2}\right) \\
+ 0.16\sin(2\pi(24u+16v)) + 0.12\cos(2\pi(31u-11v)),
\end{aligned}
$$

where $(A_1, A_2, A_3, A_4) = (1.00, 0.70, 0.85, 0.55)$, $c_1 = (0.22, 0.28)$, $c_2 = (0.62, 0.36)$, $c_3 = (0.48, 0.76)$, $c_4 = (0.82, 0.68)$, and $(\sigma_1, \sigma_2, \sigma_3, \sigma_4) = (0.055, 0.075, 0.060, 0.045)$. The finest weighted least-squares projection uses conjugate gradients with tolerance 10^{-10}. The resulting spectral coefficient vector is used as the input to the multilevel analysis transform. In this experiment the projection relative error is 4.8773×10^{-11}. The one-pass synthesis differs from the projected signal by 1.2697×10^{-2}, while the canonical coefficient-space CG reconstruction reduces this error to 2.7717×10^{-10}. The latter is below the prescribed residual tolerance 10^{-9}, which is possible because the stopping criterion is an upper threshold on the normal-equation residual rather than an identity for the reconstructed function error. These numbers describe the observed behavior of the implemented jittered Marcinkiewicz–Zygmund-type rule and should not be interpreted as computed theoretical frame bounds. Figure 6 shows the resulting multiscale decomposition. The approximation terms retain the low-frequency trend and localized bump structure, whereas the high-pass terms mainly capture the imposed oscillatory components at the corresponding dyadic scales. The flat torus data are displayed on the standard embedded torus in $\mathbb{R}^3$.

Example 4: Denoising on an oscillatory signal. We next use the same jittered nested torus rule to test a simple framelet thresholding procedure. The levels, oversampling factor, jitter parameter, and NFFT implementation are the same as in Example 3. For this experiment the clean signal is the smoother version of the periodic synthetic function, obtained by retaining the low-frequency trigonometric modes and localized periodic Gaussian bumps and replacing the high-frequency oscillatory terms by $0.025\sin(2\pi(5u+3v))$. We add Gaussian noise $g = f + \sigma\,\mathrm{std}(f)\,\epsilon$, where the entries of ϵ are independent standard normal random variables. After computing the finest weighted least-squares projection, we apply the multilevel analysis transform to the recovered finest spectral coefficient vector, use soft thresholding with threshold $\eta_{\mathrm{thr}} = 0.2\,\sigma\,\mathrm{std}(f)$, and compare two reconstruction procedures: direct one-pass synthesis and canonical coefficient-space CG reconstruction. Errors in Table 1 are dominated by the added noise and coefficient thresholding; the reconstruction tolerance $\tau_{\mathrm{rec}} = 10^{-8}$ is several orders of magnitude smaller and has no visible effect at the reported precision.

Figure 7 shows the denoising result for $\sigma = 0.20$. For this noise level, one-pass synthesis already reduces the error of the noisy observation, but the canonical coefficient-space CG reconstruction gives a smaller relative error and a higher SNR. This is consistent with the fact that one-pass synthesis also contains the finite coefficient-frame discretization error, whereas the canonical reconstruction compensates for it.

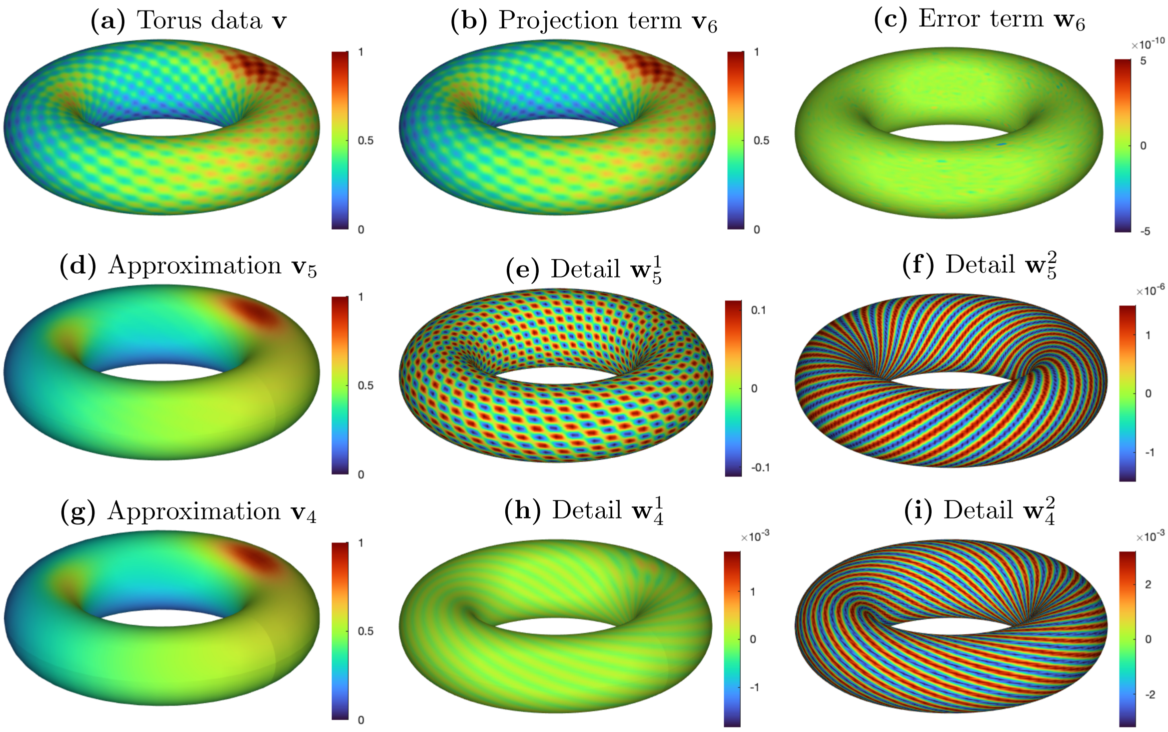

Figure 6: Multiscale decomposition of a synthetic signal on jittered nested sampling sets on $\mathbb{T}^2$.

σ	noisy rel. err.	one-pass rel. err.	canonical rel. err.	noisy SNR	one-pass SNR	CG SNR
0.10	4.5288×10^{-2}	2.0487×10^{-2}	1.7673×10^{-2}	26.880	33.770	35.054
0.20	9.0461×10^{-2}	3.7380×10^{-2}	3.5799×10^{-2}	20.871	28.547	28.922
0.30	1.3592×10^{-1}	5.4773×10^{-2}	5.3544×10^{-2}	17.334	25.229	25.426

Table 1: Denoising results on the jittered torus rule. SNR is reported in dB.

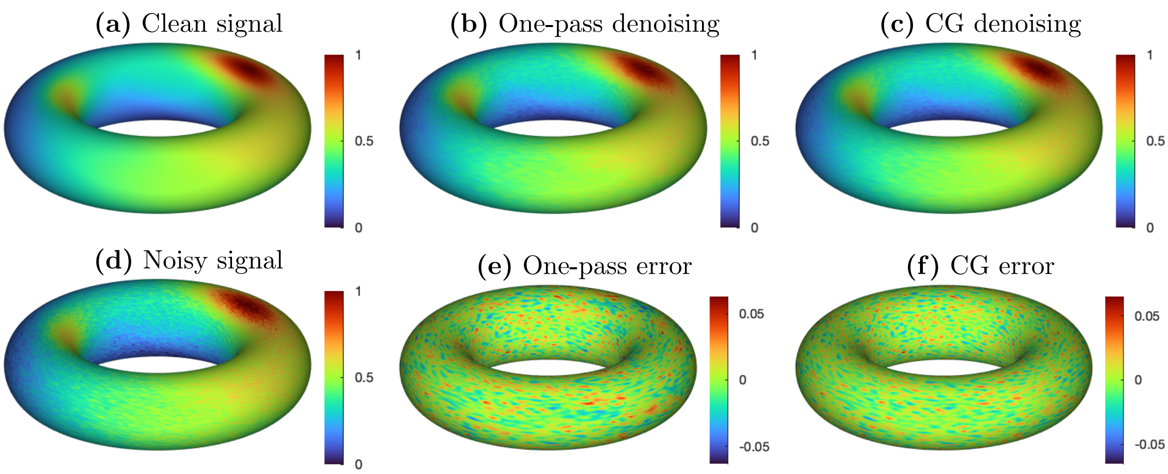

Figure 7: Framelet denoising on the jittered torus rule for noise level $\sigma = 0.20$.

8 Concluding remarks

Exact quadrature plays two logically distinct roles in many framelet constructions on manifolds. It provides a stable discretization of bandlimited functions, but it also preserves the relevant inner products exactly, leading to tightness and one-pass perfect reconstruction. The present work separates these two roles. Marcinkiewicz–Zygmund inequalities retain the stable discretization property but without requiring polynomial exactness. This relaxation allows us to work with geometrically controlled scattered nodes and, in particular, to construct nested measures that retain all previous nodes as the resolution is refined.

Applying the nested Marcinkiewicz–Zygmund measures to continuous tight framelets produces nearly tight semi-discrete systems with explicitly controlled frame operators. For a prescribed tolerance η, the frame operator lies within η of the identity. Alternatively, nested oversampling reduces this deviation at the rate $O(2^{-s})$. Therefore, relaxing polynomial exactness does not lead to an uncontrolled loss of tightness. It makes scattered and progressively refined sampling possible while retaining energy preservation and reconstruction stability to a quantifiable degree.

At the fully discrete level, the same near-tightness parameter controls both the error of one-pass synthesis and the conditioning of canonical CG reconstruction. The resulting framework thus connects the geometry of nested scattered samples, Marcinkiewicz–Zygmund discretization, nearly tight framelets, and practical multiscale transforms on manifolds.

References

[1] C. An, J. Ran, and H.-N. Wu, *The path of hyperinterpolation: A survey*, Dolomites Res. Notes Approx., 18 (2025), pp. 135–145, https://doi.org/10.25430/pupj-DRNA-2025-1-11.

[2] C. An and H.-N. Wu, *On the quadrature exactness in hyperinterpolation*, BIT, 62 (2022), pp. 1899–1919, https://doi.org/10.1007/s10543-022-00935-x.

[3] C. An and H.-N. Wu, *Bypassing the quadrature exactness assumption of hyperinterpolation on the sphere*, J. Complexity, 80 (2024), p. 101789, https://doi.org/10.1016/j.jco.2023.101789.

[4] C. An and H.-N. Wu, *Spherical configurations and quadrature methods for integral equations of the second kind*, SIAM J. Numer. Anal., 64 (2026), pp. 148–169, https://doi.org/10.1137/24M1688370.

[5] C. An and X. Zhuang, *A survey on spherical designs: Existence, numerical constructions, and applications*, arXiv preprint arXiv:2601.11963, (2026).

[6] B. G. Bodmann, G. Kutyniok, and X. Zhuang, *Gabor shearlets*, Appl. Comput. Harmon. Anal., 38 (2015), pp. 87–114, https://doi.org/10.1016/j.acha.2014.03.006.

[7] A. Böttcher, S. Kunis, and D. Potts, *Probabilistic spherical Marcinkiewicz–Zygmund inequalities*, J. Approx. Theory, 157 (2009), pp. 113–126, https://doi.org/10.1016/j.jat.2008.07.006.

[8] J. S. Brauchart, P. J. Grabner, I. H. Sloan, and R. S. Womersley, *Needlets liberated*, Appl. Comput. Harmon. Anal., 73 (2024), pp. Paper No. 101693, 16, https://doi.org/10.1016/j.acha.2024.101693.

[9] J. S. Brauchart, E. B. Saff, I. H. Sloan, and R. S. Womersley, *QMC designs: optimal order quasi Monte Carlo integration schemes on the sphere*, Math. Comp., 83 (2014), pp. 2821–2851, https://doi.org/10.1090/S0025-5718-2014-02839-1.

[10] E. Candès, L. Demanet, D. Donoho, and L. Ying, *Fast discrete curvelet transforms*, Multiscale Model. Simul., 5 (2006), pp. 861–899, https://doi.org/10.1137/05064182X.

[11] I. Chavel, *Riemannian Geometry: A Modern Introduction*, Cambridge University Press, Cambridge, 2 ed., 2006.

[12] A. Chernih, I. H. Sloan, and R. S. Womersley, *Wendland functions with increasing smoothness converge to a Gaussian*, Adv. Comput. Math., 40 (2014), pp. 185–200, https://doi.org/10.1007/s10444-013-9304-5.

[13] C. K. Chui, *An Introduction to Wavelets*, vol. 1 of Wavelet Analysis and its Applications, Academic Press, Inc., Boston, MA, 1992.

[14] A. Cohen, I. Daubechies, and P. Vial, *Wavelets on the interval and fast wavelet transforms*, Appl. Comput. Harmon. Anal., 1 (1993), pp. 54–81, https://doi.org/10.1006/acha.1993.1005.

[15] R. R. Coifman and S. Lafon, *Diffusion maps*, Appl. Comput. Harmon. Anal., 21 (2006), pp. 5–30, https://doi.org/10.1016/j.acha.2006.04.006.

[16] R. R. Coifman and M. Maggioni, *Diffusion wavelets*, Appl. Comput. Harmon. Anal., 21 (2006), pp. 53–94, https://doi.org/10.1016/j.acha.2006.04.004.

[17] F. Dai and H. Wang, *Positive cubature formulas and Marcinkiewicz-Zygmund inequalities on spherical caps*, Constr. Approx., 31 (2010), pp. 1–36, https://doi.org/10.1007/s00365-009-9041-7.

[18] I. Daubechies, *Ten Lectures on Wavelets*, vol. 61 of CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, PA, 1992, https://doi.org/10.1137/1.9781611970104.

[19] S. De Marchi and A. Kroó, *Marcinkiewicz-Zygmund type results in multivariate domains*, Acta Math. Hungar., 154 (2018), pp. 69–89, https://doi.org/10.1007/s10474-017-0769-4.

[20] P. Delsarte, J.-M. Goethals, and J. J. Seidel, *Spherical codes and designs*, Geom. Dedicata, 6 (1977), pp. 363–388, https://doi.org/10.1007/bf03187604.

[21] J. R. Driscoll and D. M. Healy, Jr., *Computing Fourier transforms and convolutions on the 2-sphere*, Adv. in Appl. Math., 15 (1994), pp. 202–250, https://doi.org/10.1006/aama.1994.1008.

[22] M. Ehler and K. Gröchenig, *An abstract approach to Marcinkiewicz-Zygmund inequalities for approximation and quadrature in modulation spaces*, Math. Comp., 93 (2024), pp. 2885–2919, https://doi.org/10.1090/mcom/3930.

[23] F. Filbir and H. N. Mhaskar, *Marcinkiewicz–Zygmund measures on manifolds*, J. Complexity, 27 (2011), pp. 568–596, https://doi.org/10.1016/j.jco.2011.03.002.

[24] D. Geller and A. Mayeli, *Nearly tight frames and space-frequency analysis on compact manifolds*, Math. Z., 263 (2009), pp. 235–264, https://doi.org/10.1007/s00209-008-0406-6.

[25] K. Gröchenig, *Sampling, Marcinkiewicz-Zygmund inequalities, approximation, and quadrature rules*, J. Approx. Theory, 257 (2020), p. No. 105455, https://doi.org/10.1016/j.jat.2020.105455.

[26] N. Hale and A. Townsend, *A fast, simple, and stable Chebyshev-Legendre transform using an asymptotic formula*, SIAM J. Sci. Comput., 36 (2014), pp. A148–A167, https://doi.org/10.1137/130932223.

[27] D. K. Hammond, P. Vandergheynst, and R. Gribonval, *Wavelets on graphs via spectral graph theory*, Appl. Comput. Harmon. Anal., 30 (2011), pp. 129–150, https://doi.org/10.1016/j.acha.2010.04.005.

[28] D. M. Healy, Jr., D. N. Rockmore, P. J. Kostelec, and S. Moore, *FFTs for the 2-sphere-improvements and variations*, J. Fourier Anal. Appl., 9 (2003), pp. 341–385, https://doi.org/10.1007/s00041-003-0018-9.

[29] K. Hesse and I. H. Sloan, *Hyperinterpolation on the sphere*, in Frontiers in Interpolation and Approximation, vol. 282 of Pure Appl. Math. (Boca Raton), Chapman & Hall/CRC, Boca Raton, 2007, pp. 213–248.

[30] J. Keiner, S. Kunis, and D. Potts, *Efficient reconstruction of functions on the sphere from scattered data*, J. Fourier Anal. Appl., 13 (2007), pp. 435–458, https://doi.org/10.1007/s00041-006-6915-y.

[31] J. Keiner, S. Kunis, and D. Potts, *Using NFFT 3—a software library for various nonequispaced fast Fourier transforms*, ACM Trans. Math. Software, 36 (2009), Article 19, pp. 1–30, https://doi.org/10.1145/1555386.1555388.

[32] Q. T. Le Gia and H. N. Mhaskar, *Localized linear polynomial operators and quadrature formulas on the sphere*, SIAM J. Numer. Anal., 47 (2009), pp. 440–466, https://doi.org/10.1137/060678555.

[33] G. Leibon and D. Letscher, *Delaunay triangulations and voronoi diagrams for riemannian manifolds*, in Proceedings of the Sixteenth Annual Symposium on Computational Geometry, New York, 2000, ACM, pp. 341–349, https://doi.org/10.1145/336154.336221.

[34] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, Inc., San Diego, CA, 1998.

[35] Y. Meyer, *Ondelettes et opérateurs. I*, Actualités Mathématiques. [Current Mathematical Topics], Hermann, Paris, 1990.

[36] H. N. Mhaskar, F. J. Narcowich, and J. D. Ward, *Spherical Marcinkiewicz–Zygmund inequalities and positive quadrature*, Math. Comp., 70 (2001), pp. 1113–1130, https://doi.org/10.1090/S0025-5718-00-01240-0.

[37] F. J. Narcowich, P. Petrushev, and J. D. Ward, *Localized tight frames on spheres*, SIAM J. Math. Anal., 38 (2006), pp. 574–594, https://doi.org/10.1137/040614359.

[38] V. Rokhlin and M. Tygert, *Fast algorithms for spherical harmonic expansions*, SIAM J. Sci. Comput., 27 (2006), pp. 1903–1928, https://doi.org/10.1137/050623073.

[39] I. H. Sloan, *Polynomial interpolation and hyperinterpolation over general regions*, J. Approx. Theory, 83 (1995), pp. 238–254, https://doi.org/10.1006/jath.1995.1119.

[40] I. H. Sloan and R. S. Womersley, *QMC designs—cubature on the sphere without polynomial exactness*, Constructive Approximation, (2026), https://doi.org/10.1007/s00365-026-09760-9.

[41] H. F. Smith, *A parametrix construction for wave equations with $C^{1,1}$ coefficients*, Ann. Inst. Fourier, 48 (1998), pp. 797–835, https://doi.org/10.5802/aif.1640.

[42] L. N. Trefethen, *Approximation Theory and Approximation Practice*, SIAM, Philadelphia, PA, 2013.

[43] L. N. Trefethen, *Exactness of quadrature formulas*, SIAM Rev., 64 (2022), pp. 132–150, https://doi.org/10.1137/20M1389522.

[44] L. N. Trefethen and D. Bau, III, *Numerical linear algebra*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997, https://doi.org/10.1137/1.9780898719574.

[45] Y. G. Wang and X. Zhuang, *Tight framelets and fast framelet filter bank transforms on manifolds*, Appl. Comput. Harmon. Anal., 48 (2020), pp. 64–95, https://doi.org/10.1016/j.acha.2018.04.006.

[46] H.-N. Wu, *Hyperinterpolation beyond exact cubature: a spectral multiplier approach*, arXiv preprint arXiv:2605.17739, (2026).

[47] H.-N. Wu, *Marcinkiewicz–Zygmund stability of quadrature-based Galerkin schemes*, Appl. Math. Lett., 182 (2026), p. 110033, https://doi.org/10.1016/j.aml.2026.110033.

[48] H.-N. Wu and X. Yuan, *Breaking quadrature exactness: A spectral method for the Allen–Cahn equation on spheres*, arXiv preprint arXiv:2305.04820, (2023).

[49] Y. Xiao and X. Zhuang, *Spherical framelets from spherical designs*, SIAM J. Imaging Sci., 16 (2023), pp. 2072–2104, https://doi.org/10.1137/22M1542362.

[50] R. Zheng and X. Zhuang, *On the existence and estimates of nested spherical designs*, Appl. Comput. Harmon. Anal., 73 (2024), p. 101672, https://doi.org/10.1016/j.acha.2024.101672.

[51] R. Zheng and X. Zhuang, *On the probability estimates of quadrature rules from uniformly sampled points on spheres*, Anal. Appl., (2026), https://doi.org/10.1142/S0219530526500119.