DIFFERENTIABLE AND ACCELERATED DIRECTIONAL WAVELET TRANSFORMS ON THE SPHERE AND BALL

M. A. Price [*,1], A. Polanska [1], J. Whitney [1], and J. D. McEwen [†1,2]

[1] Mullard Space Science Laboratory (MSSL), University College London (UCL), Holmbury St Mary, Dorking, Surrey RH5 6NT, UK and
[2] Alan Turing Institute, Euston Road, London NW1 2DB, UK

ABSTRACT

Directional wavelet dictionaries are hierarchical representations which efficiently capture and segment information across scale, location and orientation. Such representations demonstrate a particular affinity to physical signals, which often exhibit highly anisotropic, localised multiscale structure. Many physically important signals are observed over spherical domains, such as the celestial sky in cosmology. Leveraging recent advances in computational harmonic analysis, we design new highly distributable and automatically differentiable directional wavelet transforms on the 2-dimensional sphere $\mathbb{S}^2$ and 3-dimensional ball $\mathbb{B}^3 = \mathbb{R}^+ \times \mathbb{S}^2$ (the space formed by augmenting the sphere with the radial half-line). We observe up to a 300-fold and 21800-fold acceleration for signals on the sphere and ball, respectively, compared to existing software, whilst maintaining 64-bit machine precision. Not only do these algorithms dramatically accelerate existing spherical wavelet transforms, the gradient information afforded by automatic differentiation unlocks many data-driven analysis techniques previously not possible for these spaces. We publicly release both S2WAV ⬡ and S2BALL ⬡, open-sourced JAX libraries for our transforms that are automatically differentiable and readily deployable both on and over clusters of hardware accelerators (*e.g.* GPUs & TPUs).

Subject headings: Astronomical instrumentation, methods and techniques – Methods: data analysis – Methods: numerical – Techniques: image processing

1. INTRODUCTION

Many fields of research fundamentally depend on the distillation of scientifically pertinent information from data that lives on spherical manifolds; that is data which lives on the 2-dimensional sphere $\mathbb{S}^2$. In many cases such data is radially distributed, and therefore lives on the 3-dimensional ball $\mathbb{B}^3 = \mathbb{R}^+ \times \mathbb{S}^2$ (the space formed by augmenting the sphere with the non-negative radial half-line). The diversity of such fields is remarkable, ranging from: quantum chemistry (Ritchie & Kemp 1999; Choi et al. 1999), to molecular modelling and protein prediction (Boomsma & Frellsen 2017; Jumper et al. 2021), to biomedical imaging (Tuch 2004; Goodwin-Allcock et al. 2022), to geophysics and planetary science (Audet 2011; Simons et al. 2011; Marignier et al. 2020) to atmospheric and climate physics (Racah et al. 2017; Weyn et al. 2020; Ravuri et al. 2021), and to the wider cosmos (McEwen et al. 2008a,b; Price et al. 2021b; Wallis et al. 2021; Loureiro et al. 2022). Increasingly often spherical data is encountered in modern computer vision tasks, *e.g.* monocular depth estimation and semantic segmentation of 360° images (Jiang et al. 2019; Zhang et al. 2019; Ocampo et al. 2023), or 3D object retrieval (Kondor et al. 2018; Esteves et al. 2020; Cobb et al. 2021).

In recent years machine learning techniques have become increasingly widespread within these fields (see *e.g.* Ntampaka et al. 2019; Bronstein et al. 2021; Huertas-Company & Lanusse 2023). Despite numerous advantages afforded by machine learning approaches, their efficacy is typically predicated on an abundance of training

data. In many fields such data is not available, and simulation of such data is simply not feasible; astrophysics being a classic example. When data is limited, existing tools can be reimagined and employed to create hybrid methods, at great effect. One such class of tools is that of the wavelet transform and its derivatives (see *e.g.* Bruna & Mallat 2013; McEwen et al. 2022; Pedersen et al. 2023).

Wavelet theory on the sphere is a mature field of study, with a plethora of associated wavelet representations, including spherical; wavelets (Schröder & Sweldens 1995; Antoine & Vandergheynst 1998, 1999), needlets (Geller et al. 2008; Baldi et al. 2009), curvelets (Starck et al. 2006, 2009; Chan et al. 2017), and ridgelets (McEwen & Price 2019) each tailored towards specific applications. Of particular interest are scale-discretised spherical wavelets (Leistedt et al. 2013; McEwen et al. 2015b), which satisfy important quasi-exponential localisation and asymptotic uncorrelation properties (McEwen et al. 2018), and can be composed to form isometrically invariant representations which are stable to diffeomorphisms, *i.e.* effective representations for learning on the sphere (McEwen et al. 2022). Slepian wavelet representations on the sphere have also been developed (Roddy & McEwen 2021, 2023) which are particularly well suited to the masked sky, often encountered in astrophysics.

A subset of these wavelet representations have been lifted to the ball, to facilitate the analysis of spherically symmetric and radially distributed data. The exact properties of such wavelets is, in part, determined by the choice of radial discretisation. Radial needlets (Durastanti et al. 2014) adopt exponential radial basis functions whereas isotropic wavelets (Lanusse et al. 2012) adopt spherical Bessel functions, and are conse-

[*] E-mail: m.price.17@ucl.ac.uk
[†] E-mail: jason.mcewen@ucl.ac.uk

quently built on the Fourier-Bessel transform – which is encountered throughout cosmology (Abramo et al. 2010; Rassat & Refregier 2012). As the spherical Bessel function does not admit exact quadrature (Lemoine 1994), these isotropic wavelets are not exact. Instead, Fourier-Laguerre wavelets (Leistedt & McEwen 2012) select Laguerre polynomials with which to tile the radial half-line, and leverage Gauss-Laguerre quadrature to provide theoretically exact transforms. Interestingly, the spherical Bessel transform may be evaluated exactly by first projecting onto Laguerre polynomials, for signals bandlimited in the Fourier-Laguerre sense (Leistedt & McEwen 2012). Slepian wavelets have also been extended to the ball (Khalid et al. 2016) which are built upon a Fourier-Laguerre spectral decomposition.

Classically, wavelets representations have been effectively applied as compressive sensing regularisers when solving inverse problems (*e.g.* Carrillo et al. 2014; Pratley et al. 2018; Price et al. 2021a). Wavelets may be composed to form expressive scattering representations, which are both effective summary statistics and statistical generative models for highly non-Gaussian textures (Mallat 2012; Mallat et al. 2020; Allys et al. 2019, 2020; Cheng et al. 2020; Zhang & Mallat 2021; Eickenberg et al. 2022; McEwen et al. 2022; Price et al. 2023). Perhaps less well known are their applications for data compression (Balan et al. 2009; McEwen et al. 2011) and the synthesis of multifractal fields (Robitaille et al. 2020). Recently, researchers have achieved state-of-the-art performance by embedding wavelet filters directly within machine learning models (Huang et al. 2017; Liu et al. 2019), multiscale conditioning of diffusion-based generative models (Guth et al. 2022), and by solving partial differential equations with wavelet neural operators (Gupta et al. 2021; Tripura & Chakraborty 2022).

Many of these technological advances are fundamentally reliant on readily accessible gradient information, which is necessary for back-propagation during model training. A further advantage of many modern technologies is their affinity for high throughput evaluation; hence, they benefit greatly from deployment on hardware accelerators (*e.g.* GPUs and TPUs). Existing software packages which provide wavelet transforms on the sphere, S2LET ⟳, and ball, FLAGLET ⟳, are not engineered with this in mind, and provide neither differentiability nor acceleration. Consequently, it has not been possible to integrate these wavelet transforms with modern machine learning techniques on the sphere. To harness the potential of wavelets for next-generation spherical data-analysis techniques, new algorithms and software are needed.

In this work we design spherical wavelet transforms which overcome these fundamental limitations. Building upon recently released GPU accelerated spherical harmonic and Wigner transforms, provided by the S2FFT ⟳ package (Price & McEwen 2023), we develop highly parallel algorithms for the scale-discretised wavelet transform on the sphere and ball. We implement these spherical wavelet transforms in JAX, a differentiable Python library developed by Google for high-performance deep learning research (Bradbury et al. 2018). Our transforms are engineered so as to provide efficient automatic differentiation and be highly distributable both on and over hardware accelerators; hence, facilitating the integration

the future integration of wavelet techniques with modern machine learning technologies over spherically symmetric spaces. To maximise accessibility our algorithms are designed to be sampling agnostic, providing support for McEwen-Wiaux (McEwen & Wiaux 2011), Driscoll-Healy (Driscoll & Healy 1994), and HEALPix (Górski et al. 2005) sampling of the sphere at launch. Differentiability notwithstanding, the acceleration afforded by our algorithms alone facilitates previously infeasible analysis techniques, *e.g.* sampling methods and machine learning, with myriad applications from molecular modelling to the study of the cosmos.

The remainder of this article is structured as follows. In Section 2 we review mathematical background for harmonic analysis, with wavelet analysis provided in Section 3. In Section 4 we develop the directional wavelet transform on the sphere and outline our associated software package S2WAV ⟳. Subsequently, in Section 5 we abstract to the ball and outline our associated software package S2BALL ⟳. Finally, in Section 6 we draw conclusions and make closing remarks.

2. HARMONIC ANALYSIS ON THE SPHERE, ROTATION GROUP, AND BALL

In this section we review harmonic analysis on the two-sphere $\mathbb{S}^2$, rotation group SO(3), three-dimensional ball $\mathbb{B}^3 = \mathbb{R}^+ \times \mathbb{S}^2$, and rotational ball $\mathbb{H}^4 = \mathbb{R}^+ \times$ SO(3). We provide a very brief summary of necessary theory relating to spherical harmonics, Wigner functions, and Fourier-Laguerre polynomials. Our discussion covers both axisymmetric and directional convolutions on the aforementioned manifolds, for functions of arbitrary spin. We constrain this discussion to continuous transforms so as to remain sampling agnostic. In later sections we connect to corresponding transforms over discrete spaces, bridging the gap to practical applications.

2.1. *Functions on the sphere*

Spin-s functions on the sphere $_sf \in \mathrm{L}^2[\mathbb{S}^2]$ are characterised by an additional $\mathbb{U}(1)$ symmetry such that they transform by (Newman & Penrose 1966; Goldberg et al. 1967)

$$_sf' = e^{-is\chi}{}_sf, \qquad (1)$$

under right-handed rotations $\chi \in [0, 2\pi)$ in the tangent plane centered at $\omega = (\vartheta, \varphi)$, for longitude $\varphi \in [0, 2\pi)$ and colatitude $\vartheta \in [0, \pi]$. The spin-weighted spherical harmonics $_sY_{\ell m}(\omega) : \mathbb{S}^2 \to \mathbb{C}$ form the canonical orthogonal basis for square integrable functions $\mathrm{L}^2[\mathbb{S}^2]$ on the sphere for natural $\ell \in \mathbb{N}$ and integers $m, s \in \mathbb{Z}$ such that $|m|, |s| \le \ell$. Any spin-s function $_sf \in \mathrm{L}^2[\mathbb{S}^2]$ may be decomposed into this representation such that

$$_s\hat{f} = \langle\, _sf,\, _sY_{\ell m} \rangle = \int_{\mathbb{S}^2} \mathrm{d}\Omega(\omega)\, _sf(\omega)\, _sY_{\ell m}^*(\omega), \qquad (2)$$

where $\Omega(\omega) = \sin\vartheta\, \mathrm{d}\vartheta\, \mathrm{d}\varphi$ is the standard invariant measure on the sphere. By orthogonality and completeness of the spin-weighted spherical harmonics, $_sf$ may be reconstructed exactly by

$$_sf = \sum_{\ell \in \mathbb{N}} \sum_{|m| \le \ell} {}_s\hat{f}_{\ell m}\, _sY_{\ell m}. \qquad (3)$$

For practical calculation this infinite summation is truncated at a maximum harmonic degree L such that $_s\hat{f}_{\ell m} = 0$ for all $\ell \geq L$. The terminology is then such that we say the function $_s f$ is bandlimited at L. Efficient GPU algorithms have been developed to perform discretised versions of the transforms given by Equations 2 and 3 (see S2FFT ○, Price & McEwen 2023).

2.2. Functions on the rotation group

As they will play a pivotal role in the construction of directional wavelet transforms on the sphere we introduce functions on the rotation group, again very briefly. The Wigner D-functions $D_{mn}^{\ell}(\rho) : SO(3) \to \mathbb{C}$ for natural $\ell \in \mathbb{N}$, integers $m, n \in \mathbb{Z}$ such that $|m|, |n| \leq \ell$, and zyz Euler angles $\rho = (\alpha, \beta, \gamma)$ form an irreducible unitary representation of the rotation group $SO(3)$ in three dimensions (Varshalovich et al. 1989).

Any square integrable function $f \in L^2[SO(3)]$ on the rotation group may, as before, be decomposed into this representation such that

$$\hat{f} = \langle f, D_{mn}^{\ell} \rangle = \int_{SO(3)} d\Omega(\rho) \, f(\rho) \, D_{mn}^{\ell}(\rho), \quad (4)$$

where $\Omega(\rho) = \sin\beta \, d\alpha \, d\beta \, d\gamma$ is the usual invariant measure on the rotation group. By the orthogonality and completeness of the Wigner D-functions, f may be reconstructed exactly by

$$f = \sum_{\ell \in \mathbb{N}} \frac{2\ell + 1}{8\pi^2} \sum_{|m| \leq \ell} \sum_{|n| \leq \ell} \hat{f}_{mn}^{\ell} D_{mn}^{\ell*}. \quad (5)$$

As for the spin-weighted spherical harmonic transforms, the infinite summation is bandlimited at L. Efficient GPU algorithms have been developed to perform discretised version of the Wigner transform, which we leverage (see S2FFT ○, Price & McEwen 2023).

2.3. Functions on the ball

To extend harmonic analysis radially one must introduce basis functions along the positive half-line $\mathbb{R}^+$. As the canonical choice of atlas on spherical spaces are the spherical polar co-ordinates, which are separable into angular and radial components, we are free to straightforwardly adopt either spin-weighted spherical harmonics or Wigner D-functions for the angular components and independent basis functions along the radial half-line.

The canonical choice of radial basis functions are the Bessel functions which, when combined with the angular basis functions, produce the spherical Bessel functions. Unfortunately, these functions do not afford exact transforms and can be numerically unstable in practical settings. Fortunately, alternate basis functions with desirable qualities are available. The Laguerre basis functions $K_p(r) : \mathbb{R}^+ \to \mathbb{R}^+$, which are orthogonal by Gram-Schmidt and are straightforwardly complete, are defined to be

$$K_p(r) = \sqrt{\frac{p!}{(p+2)!}} \frac{e^{\frac{-r}{2\tau}}}{\sqrt{\tau^3}} L_p^{(2)}\left(\frac{r}{\tau}\right), \quad (6)$$

where $L_p^{(2)}$ are the pth-associated 2nd-order Laguerre polynomials, for natural $p \in \mathbb{N}$, and where $\tau \in \mathbb{R}^+$ is a scaling parameter. Any square integrable function

$_s f \in L^2[\mathbb{R}^+]$ may be decomposed into this representation such that

$$_s\hat{f} = \langle _s f, K_p \rangle = \int_{\mathbb{R}^+} dr \, r^2 \, _s f(r) \, K_p(r), \quad (7)$$

and for the sampling theorem presented by Leistedt & McEwen (2012) may be recovered exactly by

$$_s f(r) = \sum_{p \in \mathbb{N}} {}_s\hat{f}_p \, K_p(r). \quad (8)$$

Most real-world functions are, to a good approximation, radially bandlimited which is to say that $_s\hat{f}_p = 0$ for all $p > P$. This is the same as bandlimiting in the angular components, where instead we say a function f is radially bandlimited at P. Adopting Gauss-Laguerre quadrature to compute the integral of Equation 7 discrete Laguerre transforms can be evaluated exactly (Leistedt & McEwen 2012; McEwen & Leistedt 2013).

One can then straightforwardly compose radial and angular basis functions to form a set of basis functions on the ball. Suppose for the angular components we adopt the spin-weighted spherical harmonics, then our overall basis functions are defined to be

$$_s Z_{\ell mp}(b) = K_p(r) \, _s Y_{\ell m}(\omega), \quad (9)$$

for $b = (r, \omega) \in \mathbb{B}^3 = \mathbb{R}^+ \times \mathbb{S}^2$. These basis functions inherit the characteristics of their constituents, and are therefore trivially orthogonal and complete. Consequently, any square integrable spin-s function $_s f \in L^2[\mathbb{B}^3]$ may be decomposed into their spherical-Laguerre representation such that

$$_s\hat{f} = \langle _s f, \, _s Z \rangle = \int_{\mathbb{B}^3} d\Omega(b) \, _s f(b) \, _s Z_{\ell mp}^*(b), \quad (10)$$

where $d\Omega(b) = d\Omega(r) d\Omega(\omega) = r^2 \sin\vartheta \, dr \, d\vartheta \, d\varphi$ is the standard invariant measure on the ball. It then follows that $_s f$ may be reconstructed exactly by

$$_s f = \sum_{p, \ell \in \mathbb{N}} \sum_{|m| \leq \ell} {}_s\hat{f}_{\ell mp} \, _s Z_{\ell mp}, \quad (11)$$

where in practice infinite summations over ℓ, p are angularly and radially bandlimited by L, P respectively. Building on previous work (Leistedt & McEwen 2012; McEwen & Leistedt 2013; Leistedt et al. 2015), in this work we develop differentiable and highly accelerated GPU algorithms to evaluate this spherical-Laguerre transform, which we extend to the directional setting.

2.4. Directional convolutions on the sphere

Consider a rotation operator $\mathcal{R}_\rho$ with action

$$(\mathcal{R}_\rho \, _s f)(\omega) = e^{-is\vartheta} \, _s f(\mathcal{R}_\rho^{-1}\omega), \quad (12)$$

when applied to a spin-s square integrable function on the sphere. The exponential term here comes from the additional $\mathbb{U}(1)$ symmetry discussed in Equation 1 (McEwen et al. 2015b). By noting the additive property of the Wigner D-functions, a rotated function permits a harmonic representation

$$(\mathcal{R}_\rho \, _s f)_{\ell m} = \sum_{|n| \leq \ell} D_{mn}^{\ell} \, _s f_{\ell n}. \quad (13)$$

Suppose one would like to convolve two spin-s square integrable functions $_sf, _sg \in \mathrm{L}^2[\mathbb{S}^2]$. Conceptually, this is given as the product between $_sf$ and $\mathcal{R}_\rho \, _sg$ for all possible Euler angles $\rho \in \mathrm{SO}(3)$. Mathematically, this reads

$$(_sf \circledast _sg)(\rho) = \langle \, _sf, \, \mathcal{R}_\rho \, _sg \, \rangle$$
$$= \int_{\mathbb{S}^2} \mathrm{d}\Omega(\omega) \, _sf(\omega) \, (\mathcal{R}_\rho \, _sg)^*(\omega), \quad (14)$$

where $\circledast$ is the directional convolution. As this is a function on the rotation group, we may consider the Wigner representation for simplicity

$$(_sf \circledast _sg)^\ell_{mn} = \frac{8\pi^2}{2\ell+1} \, _sf_{\ell m} \, _sg^*_{\ell n}, \quad (15)$$

from which the convolved function can be recovered exactly by (McEwen et al. 2007, 2015b; Cobb et al. 2021)

$$(_sf \circledast _sg)(\rho) = \sum_{\ell \in \mathbb{N}} \sum_{|m,n| \le \ell} \, _sf_{\ell m} \, _sg^*_{\ell n} \, D^{\ell*}_{mn}. \quad (16)$$

When we restrict the possible rotations to the sphere, that is when $\rho = (\alpha, \beta, \gamma) \rightarrow (\alpha, \beta, 0) = (\vartheta, \varphi) = \omega$, and when $_sg$ is axisymmetric $_sg_{\ell m} = \, _sg_{\ell 0}\delta_{m0}$, this convolution becomes the axisymmetric convolution on the sphere $\mathbb{S}^2$.

2.5. *Convolutions on the radial half-line*

Further consider a generalised radial translation[1] $\mathcal{T}_r$ for $r \in \mathbb{R}^+$ defined by its application to the spherical-Laguerre basis functions

$$(\mathcal{T}_{r'} \, K_p)(r) = K_p(r') \, K_p(r), \quad (17)$$

presented in McEwen & Leistedt (2013) with action on square integrable functions $f \in \mathbb{R}^+$ given by

$$(\mathcal{T}_{r'} \, f)(r) = \sum_{p \in \mathbb{N}} \hat{f}_p \, K_p(r') \, K_p(r)$$
$$\Rightarrow (\mathcal{T}_{r'} \, f)_p = K_p(r') \, \hat{f}_p. \quad (18)$$

Using this definition of the translation operator we can define the convolution between two functions $f, g \in \mathrm{L}^2[\mathbb{R}^+]$ to be given as

$$(f \star g)(r) = \langle f, \mathcal{T}_r g \rangle$$
$$= \int_{\mathbb{R}^+} \mathrm{d}r r^2 \, f(r) \, (\mathcal{T}_{r'} \, g)(r), \quad (19)$$

which straightforwardly leads to

$$(f \star g)_p = f_p \, g_p. \quad (20)$$

2.6. *Directional convolutions on the ball*

Finally, consider the combined 4-dimensional symmetry transformation $\mathcal{L}_h = \mathcal{T}_r \mathcal{R}_\rho$ for $h = (r, \rho) \in \mathbb{H}^4 = \mathbb{R}^+ \times \mathrm{SO}(3)$, which describes the complete space of rotations and translations present in our system. The total

---

[1] From our construction of the spherical-Laguerre basis functions this is closer to a translation across Laguerre polynomials, however we will refer to this as radial translation throughout this article.

convolution of two spin-s functions $f, g \in \mathrm{L}^2[\mathbb{B}^3]$ is defined analogously to before by (Price & McEwen 2021)

$$(_sf \circledast _sg)(h) = \langle \, _sf, \, \mathcal{L}_h \, _sg \, \rangle$$
$$= \int_{\mathbb{B}^3} \mathrm{d}\Omega(b) \, _sf(b) \, (\mathcal{L}_h \, _sg)^*(b), \quad (21)$$

which is the product between f and g over the composite space of 4-dimensional translations and rotations. Comparing against Equations 15 and 19 one finds

$$(_sf \circledast _sg)^\ell_{mnp} = \frac{8\pi^2}{2\ell+1} \, _sf^\ell_{mp} \, _sg^{\ell*}_{np}, \quad (22)$$

from which the convolved function can be recovered in pixel-space by

$$(_sf \circledast _sg)(h) = \sum_{p,\ell \in \mathbb{N}} \sum_{|m,n| \le \ell} \frac{8\pi^2}{2\ell+1} \, _sf^\ell_{mp} \, _sg^{\ell*}_{np} \, Q^{\ell*}_{mnp}(h), \quad (23)$$

where we have overloaded $\circledast$ to denote the convolution under the general symmetry transformation $\mathcal{L}_h$.

Here we have implicitly defined the Wigner-Laguerre basis functions Q^ℓ_{mnp} which are orthogonal and complete on $\mathbb{H}^4$. By construction the Wigner-Laguerre functions are straightforwardly separable, hence the decomposition of square integrable functions $f \in \mathrm{L}^2[\mathbb{H}^4]$ is given by

$$\hat{f} = \langle f, Q^\ell_{mnp} \rangle = \int_{\mathbb{H}^4} \mathrm{d}\Omega(h) \, f(h) \, Q^{\ell*}_{mnp}(h), \quad (24)$$

where $\Omega(h) = r^2 \sin\beta \, \mathrm{d}r \, \mathrm{d}\alpha \, \mathrm{d}\beta \, \mathrm{d}\gamma$ is the invariant measure on $\mathbb{H}^4$. As in previous settings, the original function may be exactly synthesised by

$$f = \sum_{p,\ell \in \mathbb{N}} \frac{2\ell+1}{8\pi^2} \sum_{|m,n| \le \ell} \hat{f}^\ell_{mnp} \, Q^{\ell*}_{mnp}, \quad (25)$$

which is practically extremely expensive to evaluate. Previously, Leistedt & McEwen (2012); Leistedt et al. (2015) developed algorithms to perform this transform. In this work, we redesign these algorithms to leverage recent advances in GPU accelerated and differentiable harmonic analysis (see S2FFT ⟳, Price & McEwen 2023).

3. WAVELET TRANSFORMS ON THE SPHERE AND BALL

In this section we discuss the wavelet analysis of signals on the sphere $\mathbb{S}^2$ and ball $\mathbb{B}^3 = \mathbb{R}^+ \times \mathbb{S}^2$. We formally define the directional wavelet transform on both the sphere and ball, leveraging much of the mathematics provided in Section 2. Again we remain in the continuous setting.

3.1. *Directional wavelet transform on the sphere*

Spherical wavelet filters are square integrable bandlimited functions $_s\Psi^{(j)} \in \mathrm{L}^2[\mathbb{S}^2]$ constructed to exhibit strong localisation properties in both the harmonic and spatial domain (McEwen et al. 2018), as illustrated in Figure 1. Note that here $j \in \mathbb{N} \le J$ denotes wavelet scale, which determines the harmonic degrees over which the wavelet has support (see Appendix A). The wavelet functions form a dictionary into which spin-s functions
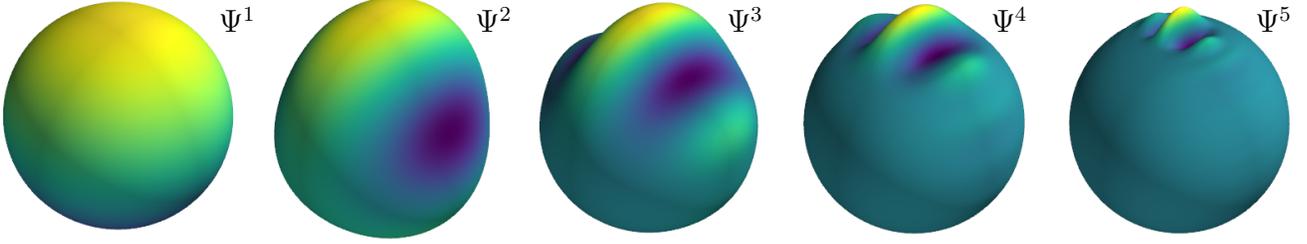
FIG. 1.—. Directional ($L = 256, N = 5$) scale-discretised wavelet filters on the sphere, for wavelet scales $j \in \{1, \dots, 5\}$ from left to right. As the wavelet scale increases the filter becomes more highly localised, both in pixel and frequency space (McEwen et al. 2018). Notice the clearly elongated structure which gives rise to the desired directional sensitivity (McEwen et al. 2015b). With an azimuthal bandlimit of $N = 5$ this filter bank is augmented by 9 rotated filters, which are not shown here.

$_s f \in \mathrm{L}^2[\mathbb{S}^2]$ may be decomposed through the directional convolution

$$W^{s \Psi^j}(\rho) = (_s f \circledast {}_s \Psi^j)(\rho) = \langle {}_s f, \mathcal{R}_\rho {}_s \Psi^j \rangle, \qquad (26)$$

which from Equation 14 can be given in Wigner space by the expression

$$(W^{s \Psi^j})^\ell_{mn} = \frac{8\pi^2}{2\ell + 1} {}_s \hat{f}_{\ell m} {}_s \Psi^{j*}_{\ell n}. \qquad (27)$$

To capture low frequency information an axisymmetric scaling function $_s \Phi \in \mathrm{L}^2[\mathbb{S}^2]$ is introduced into which f may be decomposed as

$$W^{s \Phi}(\omega) = (_s f \odot {}_s \Phi)(\omega) = \langle {}_s f, \mathcal{R}_\omega {}_s \Phi \rangle, \qquad (28)$$

where $\odot$ denotes the axisymmetric convolution. Again, this can be given in harmonic space by the expression

$$(W^{s \Phi})_{\ell m} = \sqrt{\frac{4\pi}{2\ell + 1}} {}_s f_{\ell m} {}_s \Phi^*_{\ell 0}. \qquad (29)$$

Provided a wavelet dictionary which satisfies the admissibility condition

$$\frac{4\pi}{2\ell + 1} |_s \Phi_{\ell 0}|^2 + \frac{8\pi^2}{2\ell + 1} \sum_{j=0}^{J} \sum_{|m| \leq \ell} |_s \Psi^j_{\ell m}|^2 = 1 \quad \forall \ell, \quad (30)$$

of which there are many (see e.g. Leistedt et al. 2013; Chan et al. 2017; McEwen & Price 2019), one may exactly reconstruct $_s f$ by

$$_s f = \int_{\mathbb{S}^2} \mathrm{d}\Omega(\omega) W^{s \Phi}(\omega)(\mathcal{R}_\omega \Phi)(\omega')$$
$$+ \sum_{j=0}^{J} \int_{\mathrm{SO}(3)} \mathrm{d}\Omega(\rho) W^{s \Psi^j}(\rho)(\mathcal{R}_\rho W^{s \Psi^j})(\omega'). \qquad (31)$$

In general this transform can be expensive to compute, motivating the development of efficient algorithms. In this article we adopt scale-discretised wavelets (Wiaux et al. 2008; Leistedt et al. 2013; McEwen et al. 2015b). These wavelets exhibit good harmonic and spatial localisation (McEwen et al. 2018), and permit exact synthesis, at least in such a case that a sampling theorem on the sphere and ball is provided (see e.g. Driscoll & Healy 1994; McEwen & Wiaux 2011; McEwen & Leistedt 2013).

3.2. Directional wavelet transform on the ball

Ball wavelet filters are square integrable functions $_s \Psi^{jj'} \in \mathrm{L}^2[\mathbb{B}^3]$ with angular and radial bandlimits L and P respectively. They are designed to exhibit strong localisation properties in both spherical-Laguerre and spatial domain (McEwen et al. 2018), as illustrated in Figure 2. Here we further introduce $j' \in \mathbb{N} \leq J'$ to denote radial wavelet scale (see Appendix A for further details). As in the spherical case, these wavelet functions form an overcomplete dictionary into which spin-s functions $_s f \in \mathrm{L}^2[\mathbb{B}^3]$ may be decomposed through the overloaded directional convolution

$$W^{s \Psi^{jj'}}(h) = (_s f \circledast {}_s \Psi^{jj'})(h) = \langle {}_s f, \mathcal{L}_h {}_s \Psi^{jj'} \rangle, \quad (32)$$

which from Equation 22 is represented in Wigner-Laguerre space by the expression

$$(W^{s \Psi^{jj'}})^\ell_{mnp} = \frac{8\pi^2}{2\ell + 1} {}_s \hat{f}_{\ell m p} {}_s \Psi^{jj'*}_{\ell n p}. \qquad (33)$$

To capture low frequency information an axisymmetric scaling function $_s \Phi \in \mathrm{L}^2[\mathbb{B}^3]$ is introduced

$$W^{s \Phi}(b) = (_s f \odot {}_s \Phi)(b) = \langle {}_s f, \mathcal{L}_b {}_s \Phi \rangle, \qquad (34)$$

where $\mathcal{L}_b$ is the axisymmetric simplification of $\mathcal{L}_h$. The spherical-Laguerre representation of the scaling coefficients is given as

$$(W^{s \Phi})_{\ell m p} = \sqrt{\frac{4\pi}{2\ell + 1}} {}_s f_{\ell m p} {}_s \Phi^*_{\ell 0 p}. \qquad (35)$$

On the ball, the wavelet admissibility condition as presented in Leistedt et al. (2015) reads

$$\frac{4\pi}{2\ell + 1} |_s \Phi_{\ell 0 p}|^2 + \frac{8\pi^2}{2\ell + 1} \sum_{mjj'} |_s \Psi^{jj'}_{\ell m p}|^2 = 1 \quad \forall \ell, p, \quad (36)$$

which, if satisfied, permits exact synthesis by

$$_s f = \int_{\mathbb{B}^3} \mathrm{d}\Omega(b) W^{s \Phi}(b)(\mathcal{L}_b \Phi)(b)$$
$$+ \sum_{j,j'=0}^{JJ'} \int_{\mathbb{H}^4} \mathrm{d}\Omega(h) W^{s \Psi^{jj'}}(h)(\mathcal{L}_h W^{s \Psi^{jj'}})(b). \quad (37)$$

This transform can be extremely expensive to evaluate, however noticing that each wavelet scale jj' has compact support can dramatically reduce the total number of calculations required. Such an acceleration is exploited in what is referred to as a multiresolution algorithm.
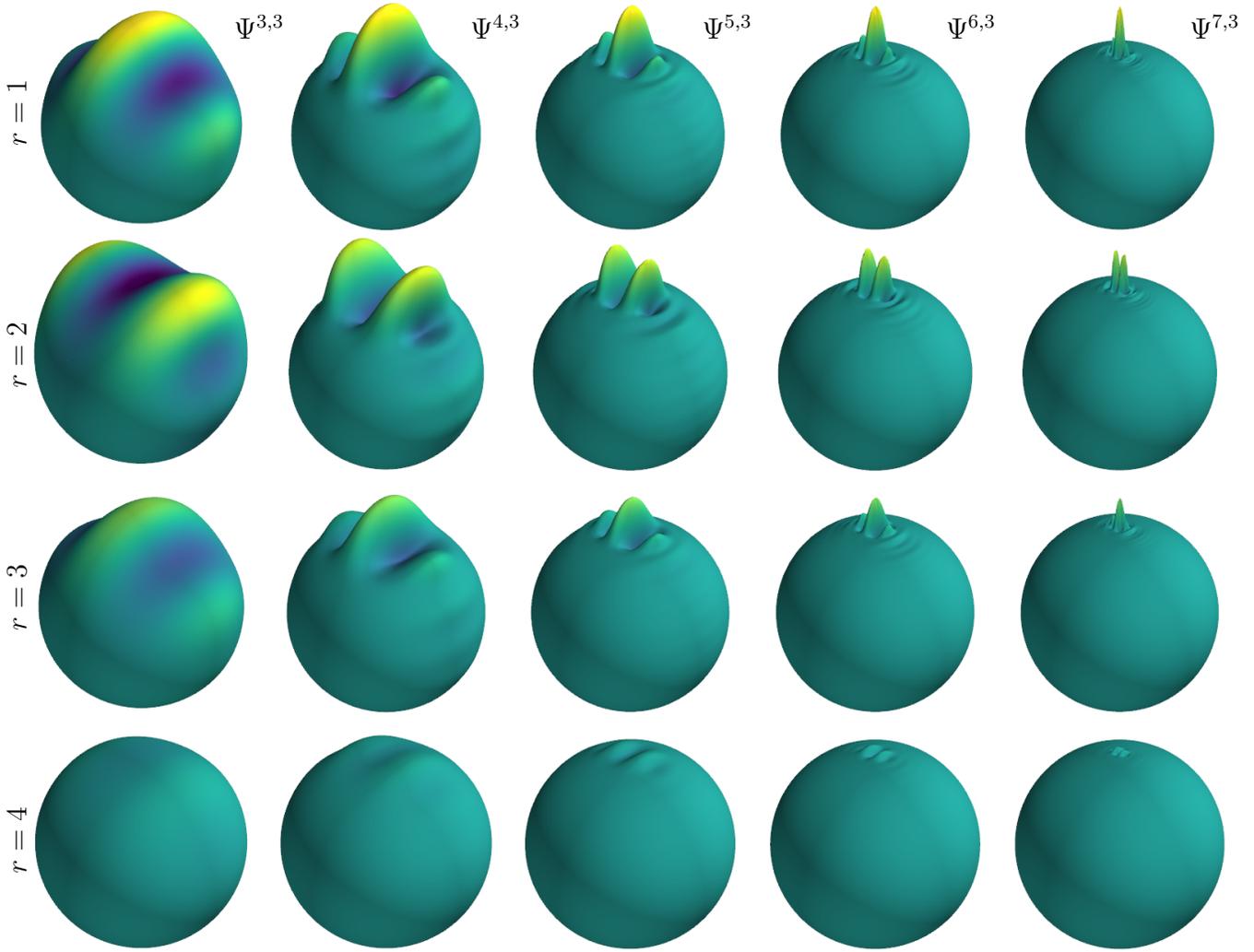
FIG. 2.—. Directional ($N = 5$) scale-discretised wavelet filters on the ball ($P = L = 256$), for angular wavelet scales $j \in [3, 7]$ from left to right, and radial wavelet scale $j' = 3$ for radial nodes $r \in [1, 4]$. As the angular wavelet scale increases the filter becomes more highly localised, both in pixel and frequency space (McEwen et al. 2018). Equally, as the radial scale increases the localisation along the radial half-line increases. For a given radial scale (as shown here) the energy of a given filter $\Psi(r)$ decays exponentially with r as expected. Notice the clearly elongated structure which gives rise to the desired directional sensitivity.
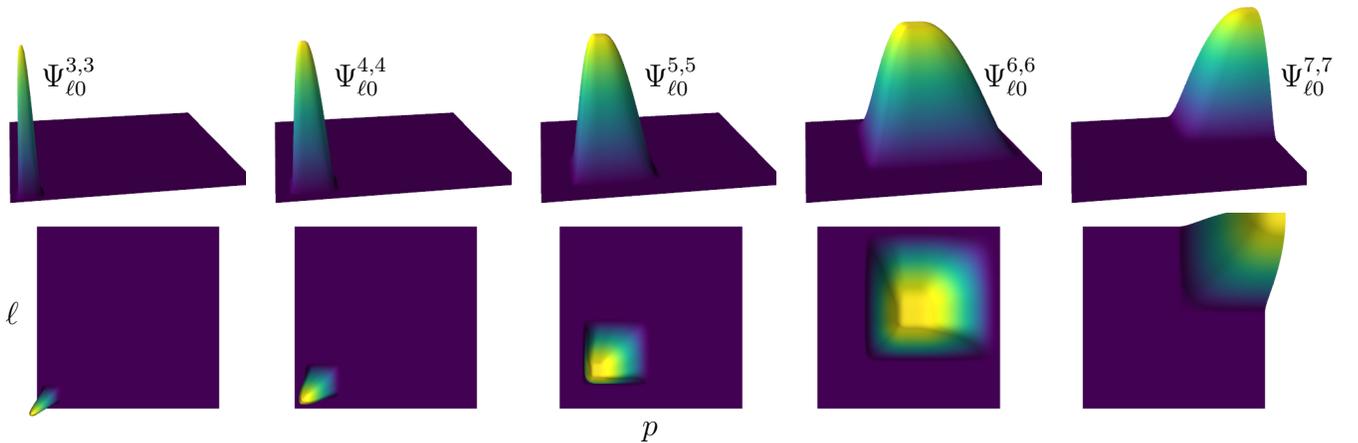

FIG. 3.—. Surface visualisation of a subset of our tiling of Fourier-Laguerre space with infinitely differentiable Cauchy-Schwartz functions (see Appendix A.3). Here we present the axisymmetric ($m = 0$) components of our wavelet filters for $j = j'$. Notice that if one were to consider all $j < j'$ and *vice versa* this filtering scheme would span the entire domain.

3.3. *Multiresolution Algorithms*

For the development of multiresolution algorithms it is critical to first note that a given wavelet scale has strictly compact support over finitely many harmonic degrees ℓ, and polynomials p, as shown in Figure 3. Therefore, both the expressions for the forward and inverse wavelet transforms, on both the sphere and ball, may be performed at varying resolutions, without loss of information.

Specifically, the wavelet $_s\Psi^j$ at scale j is only non-zero over the finite interval of harmonic degrees

$$\ell \in \left[\lfloor \lambda^{j-1} \rfloor, \lceil \lambda^{j+1} \rceil \right], \tag{38}$$

where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor and ceiling functions respectively, and $\lambda \in \mathbb{N} > 1$ is the dilation parameter of the wavelets outlined in Appendix A. As the computational time of a given harmonic degree ℓ scales as $\mathcal{O}(\ell^3)$, the overall complexity of the directional wavelet transform is dominated by the highest degrees. It then follows that only the highest two wavelet scales have non-zero support near the harmonic bandlimit $\ell \sim L$, hence the overall complexity is effectively that of this single scale.

Precisely the same argument may be applied in the radial direction, wherein a given radial scale j' has compact support over the finite interval

$$p \in \left[\lfloor \nu^{j'-1} \rfloor, \lceil \nu^{j'+1} \rceil \right], \tag{39}$$

Consequently, multiresolution algorithms designed to evaluate the aforementioned wavelet transforms can be accelerated by a factor of J and $J \times J'$ over the sphere and ball respectively, without loss of information.

4. THE S2WAV LIBRARY

Wavelet transforms have demonstrated utility in various emergent technologies, whether this be embedding wavelet filters directly within machine learning models (Huang et al. 2017; Liu et al. 2019) or through multiscale conditioning (Guth et al. 2022). In any case, to encorporate such techniques within modern machine learning technologies requires that the wavelet transforms in question are differentiable, so as to facilitate the back-propagation of gradient information. Furthermore, a primary advantage of such technologies is their high throughput, potentially with real-time evaluation. Therefore the ability to deploy such technologies on hardware accelerators (*e.g.* GPUs/TPUs) is almost mandatory. Wavelet transforms have been developed for Euclidean applications which satisfy the aforementioned remit, however no such transforms exist on the sphere, limiting progress in this area.

In this section we develop and release S2WAV ⟳, a professionally developed open-source JAX library (Bradbury et al. 2018), which provides support for the directional wavelet transform on the sphere. Specifically, by leveraging the recently released S2FFT ⟳ software package (Price & McEwen 2023) we provide GPU accelerated and automatically differentiable implementations of the directional scale-discretised wavelet transform outlined in Section 3.1. Building upon novel Wigner d-function recursions, S2FFT is designed to be extremely parallelisable, and asymptotically recovers linear scaling across multiple accelerators. Interestingly this results in S2FFT

demonstrating an effective linear compute scaling with bandlimit L, which is unprecedented. We inherit both this computational scaling and functionality, with S2WAV transforms being efficiently distributable across multiple hardware devices. In what follows we drop spin subscripts for notational brevity.

4.1. *Mathematical Overview*

Explicitly, we are primarily concerned with the efficient evaluation of two transforms corresponding to the forward (analysis) and inverse (synthesis) wavelet transform over $\mathbb{S}^2$. Consider the Wigner space representation of the wavelet coefficients given in Equation 27. Introducing $\mathbf{D}^{-1}$ as an operator which applies the inverse Wigner transform in Equation 5 to each scale j and $\mathbf{Y}$ as the forward spin spherical harmonic transform in Equation 2 the wavelet transform of a function $f \in L^2[\mathbb{S}^2]$ may be written as

$$W^{\Psi^j} = \mathbf{D}^{-1} \, \mathbf{N} \, \mathbf{\Psi}^j \, \mathbf{Y} f, \tag{40}$$

where $\mathbf{\Psi}^j$, with normalised entries given by $\Psi_{\ell n}^{j*}$, is an operator which applies the tensor outer product of $\hat{f}$ with each scale j of a given wavelet dictionary $\{\Psi^j\}$, *i.e.* the directional convolution on the sphere defined in Equation 27, and $\mathbf{N}$ applies the normalisation $8\pi^2/(2\ell+1)$. From Equation 29 the scaling coefficients are straightforwardly given by

$$W^{\Phi} = \mathbf{Y}^{-1} \, \mathbf{\Phi} \, \mathbf{Y} f, \tag{41}$$

where $\mathbf{\Phi}$, with entries given by $\sqrt{4\pi/(2\ell+1)}\Phi_{\ell 0}^*$, is an operator which denotes the inner product of $\hat{f}$ with a scaling dictionary Φ, *i.e.* the axisymmetric convolution.

Adopting analogous operators, the wavelet synthesis transform presented in Equation 31 may be written as

$$f = \mathbf{Y}^{-1}\left(\mathbf{\Phi}\mathbf{Y}W^{\Phi} + \sum_j \mathbf{\Psi}^{j\dagger} \, \mathbf{D} \, W^{\Psi^j}\right), \tag{42}$$

where the summation is a tensor contraction over each scale j. Notice that $\mathbf{\Phi}$ remains the same for both transforms as $\Phi_{\ell 0}^* = \Phi_{\ell 0}$ and the normalisation prefactor is the usual harmonic normalisation in both directions (contrast this with the different normalisations for the Wigner transform, which arise due to the normalisation factor appearing in Equation 5).

As the spherical harmonic and Wigner transforms are provided by S2FFT, we need only efficiently compute the necessary tensor operations whilst integrating the multiscale acceleration method outlined in Section 3.3.

4.2. *Precomputed Components*

From Equations 40 and 42 it is apparent that the wavelet and scaling filters, Ψ and Φ respectively, may be calculated and cached for future use rather than evaluated on the fly. In a multiresolution framework, the memory complexity associated with each of these arrays is $\mathcal{O}(NL)$ and $\mathcal{O}(L)$ respectively, which is extremely small. Though perhaps somewhat clunky, this avoids unnecessary compute and potential memory issues due to sequential reallocation of memory during *e.g.* optimisation or training.

Additionally, we configure our transforms to support the precompute functionality within S2FFT. Specifically,

Algorithm 1 Directional wavelet transform on $\mathbb{S}^2$

```
import s2fft                        ▷ JAX Spherical harmonic transforms
from s2fft import wigner            ▷ JAX Wigner transforms
from jax.numpy import einsum        ▷ JAX Tensor operations
procedure ANALYSIS WAVELET TRANSFORM(f ∈ 𝕊²):
    f̂ ← Yf                                    ▷ s2fft.forward(f,L)
    Ŵ^Φ ← Φf̂                                  ▷ einsum("lm,l->lm",f̂,Φ)
    W^Φ ← Y⁻¹Ŵ^Φ                              ▷ s2fft.inverse(Ŵ^Φ,L)
    for j ∈ [0, J] do
        Ŵ^Ψj ← NΨj f̂                          ▷ einsum("lm,ln->nlm",f̂,NΨ^j*)
        W^Ψj ← D⁻¹Ŵ^Ψj                        ▷ wigner.inverse(Ŵ^Ψj,L,N)
    return {W^Ψ, W^Φ}
procedure SYNTHESIS WAVELET TRANSFORM({W^Ψ, W^Φ}):
    Ŵ^Φ ← YW^Φ                                ▷ s2fft.forward(W^Φ,L)
    f̂ ← ΦŴ^Φ                                  ▷ einsum("lm,l->lm",Ŵ^Φ,Φ)
    for j ∈ [0, J] do
        Ŵ^Ψj ← DW^Ψj                          ▷ wigner.forward(W^Ψj,L,N)
        f̂ +← Ψ†Ŵ^Ψj                           ▷ einsum("nlm,ln->lm",Ŵ^Ψj,Ψ^j)
    f ← Y⁻¹f̂                                   ▷ s2fft.inverse(f̂,L)
    return f
```

this optional acceleration precomputes the real polar-d functions necessary to evaluate forward spin spherical harmonic, and by extension Wigner, transforms. Though this can provide extremely fast transforms, it comes with an $\mathcal{O}(NL^3)$ memory overhead, which limits the resolution at which these transforms may be applied (*e.g.* Cobb et al. 2021).

4.3. *JAX Tensor Operations*

Given precomputed wavelet and scaling filters, implementing highly efficient algorithms to evaluate the tensor operations $\mathbf{\Psi}$ and $\mathbf{\Phi}$ is straightforward. By design, JAX in fact provides support for all such operations through einsum, which converts understandable symbolic notation into linear algebraic array operations. This transform is outlined in Algorithm 1 which includes a sketch of the associated code.

As both the spherial harmonic and Wigner transforms provided by S2FFT and the einsum primitives are natively differentiable, the directional wavelet transforms provided by S2WAV also provide automatic differentiation. Our transforms can therefore be straightforwardly integrated within existing frameworks to extend, *e.g.* multi-scale conditioned generative models (Huang et al. 2017; Liu et al. 2019; Guth et al. 2022) or scattering covariances (Allys et al. 2020; Mallat et al. 2020) to the spherical setting (Mousset *et al* in prep).

Moreover, we design S2WAV to utilise the single program multiple data (SPMD) functionality of S2FFT to distribute compute across hardware accelerators. Given that the complexity of Equations 40 and 42 is dominated by the Wigner transforms, and noting the discussion of Price & McEwen (2023), with sufficient compute our wavelet transforms asymptotically recover an effective linear time complexity. In the case where a small number of GPU devices are available one should expect to asymptotically recover a further acceleration by the number of devices.

4.4. *Numerical Validation*

We benchmark our directional wavelet transforms against the existing C alternative S2LET ⟳ (Leistedt et al. 2013; McEwen et al. 2015b). Our protocol is straightforward: generate a random bandlimited signal $\hat{f}$; map this to a random bandlimited function on the sphere $f \leftarrow \mathbf{Y}^{-1}\hat{f}$; apply the forward wavelet transform followed by the inverse transform to recover f'; map this function onto its harmonic coefficients $\hat{f}' \leftarrow \mathbf{Y}f'$; and evaluate both the round-trip time and relative error $\mathbb{E}(|\hat{f} - \hat{f}'|)$.

S2LET transforms were executed on a multithreaded Xeon(R) E5-2650L v3 dedicated CPU and S2WAV transforms were executed on both a single and three NVIDIA A100 GPUs. Further testing over many more GPUs has been left for future development due to computational constraints. In addition, we provide benchmark results for both operating modalities discussed in Section 4.2; wherein real polar d-functions are evaluated on-the-fly or simply precomputed.

The results of this benchmarking are presented for recursive and precompute transforms in Table 1. As expected, when distributing across three GPUs we asymptotically recover an additional factor of three acceleration, reaching as high as $100\times$ and $300\times$ faster than existing S2LET transforms, for recursive and precompute respectively. In all cases our transforms are exact to 64-bit machine precision. Note that here we adopt McEwen-Wiaux sampling which affords a sampling theorem (McEwen & Wiaux 2011; McEwen et al. 2015a) and is theoretically exact. We also provide support for HEALPix sampling (Górski et al. 2005) which does not support a sampling theorem, and therefore produces approximate transforms.

5. THE S2BALL LIBRARY

As discussed in Section 4, it has been demonstrated that various modern machine learning technologies may be enhanced by the incorporation of wavelet representations; with the added caveat that the associated wavelet transforms must be differentiable and, ideally, deployable on hardware accelerators. Hardware acceleration is even more critical when considering functions on the ball. In short, this is because wavelet transforms over this space are prohibitively expensive. In addition to the aforementioned contributions, extending spherical wavelet dictionaries radially opens up many salient scientific applications, particularly in the study of geophysics (Simons et al. 2011; Marignier et al. 2020) and molecular modelling (Boomsma & Frellsen 2017; Jumper et al. 2021).

To this end, we develop and release S2BALL ⟳, a highly optimised open-source JAX library which provides support for the directional wavelet transform on the ball. Specifically, we provide GPU accelerated and automatically differentiable implementations of the Wigner-Laguerre wavelet transform outlined in Section 3.2. In this case, we do not integrate the S2FFT ⟳ package for spin spherical harmonic and Wigner transforms, instead providing bespoke implementations for computational efficiency, as will become apparent.

5.1. *Mathematical Overview*

Explicitly, this package is designed to efficiently evaluate the forward (analysis) and inverse (synthesis) direc-

		On-the-fly transform				Precompute transform	
L	Time `S2LET`	Time `S2WAV`	Speed-Up $1 \times$ GPU	Speed-Up $3 \times$ GPU	$3 \times$ GPU Ratio	Time `S2WAV`	Speed-Up $1 \times$ GPU
8	1.8×10^{-1}	–	–	–	–	1.4×10^{-1}	1.3
16	1.3×10^{0}	–	–	–	–	1.6×10^{-1}	8.1
32	9.5×10^{0}	–	–	–	–	2.2×10^{-1}	43.2
64	3.3×10^{1}	7.9×10^{1}	0.4	0.7	1.8	4.6×10^{-1}	71.7
128	2.1×10^{2}	1.5×10^{2}	1.4	2.8	2.0	1.5×10^{0}	140
256	1.8×10^{3}	3.0×10^{2}	6.0	13	2.2	6.0×10^{0}	300
512	1.3×10^{4}	9.4×10^{2}	14	32	2.3	–	–
1024	9.8×10^{4}	4.3×10^{3}	23	58	2.5	–	–
2048	1.1×10^{6}	3.0×10^{4}	37	92	2.5	–	–

TABLE 1. Numerical validation of the recursive (on-the-fly) and precompute `S2WAV` directional wavelet transform on the sphere against the existing `S2LET` package. Note that all timings are quoted in milliseconds. The precompute mode of `S2WAV` caches elements of the real polar d-functions necessary to evaluate the spin spherical harmonics transforms. Therefore, though this greatly increases throughput, the peak memory overhead scales as $\mathcal{O}(L^3)$ which limits the maximum operational bandlimit. Here we consider a single fixed azimuthal bandlimit $N = 5$ which corresponds to 9 rotations of our wavelet filters within each local tangent plane; for the majority of applications this is more than sufficient to recover almost all directional structure. `S2LET` functions were executed on a multithreaded Xeon(R) E5-2650L v3 CPU and `S2WAV` transforms were evaluated on a single and collection of three NVIDIA A100 GPUs. For low bandlimits L communication and GPU memory allocation costs slow down our transform, however in such cases compute is rarely an issue. Furthermore at low to moderate resolutions a precompute approach may be adopted, which can provide up to $\sim 300\times$ faster. In higher resolution cases our transforms become up to $\sim 100\times$ faster, whilst retaining 64-bit machine precision.

tional wavelet transform of L, P bandlimited functions $f \in \mathrm{L}^2[\mathbb{B}^3]$. First, consider the Wigner-Laguerre space representation of the wavelet coefficeints given in Equation 33. Recalling our definition for the inverse Wigner-Laguerre given in Equation 25 it is clear that the pixel-space representation of scale $\{j, j'\}$ is given by

$$W^{\Psi^{jj'}}(h) = \sum_{p, \ell \in \mathbb{N}} \frac{2\ell + 1}{8\pi^2} \sum_{|m,n| \le \ell} (W^{\Psi^{jj'}})^{\ell}_{mnp} \, Q^{\ell*}_{mnp}(h). \tag{43}$$

Introducing the operators $\mathbf{Z}$ to denote the forward spherical-Laguerre transforms defined in Equation 10, and introducing $\mathbf{Q}^{-1}$ to denote the inverse Wigner-Laguerre transform defined in Equation 25, one may recast Equation 43 to read

$$W^{\Psi^{jj'}} = \mathbf{Q}^{-1} \, \mathbf{N} \Psi^{jj'} \, \mathbf{Z} f, \tag{44}$$

where $\Psi^{jj'}$ is an operator which applies a scale-wise tensor outer product of $\hat{f}$ with each scale $\{j, j'\}$ of a given wavelet dictionary $\{\Psi^{j,j'}\}$, *i.e.* the directional convolution on the ball defined in Equation 33. Adopting similar operators, from Equation 35 the scaling coefficients are given by

$$W^{\Phi} = \mathbf{Z}^{-1} \, \Phi \, \mathbf{Z} f, \tag{45}$$

where Φ is an operator which applies the tensor inner product with scaling dictionary Φ, *i.e.* the axisymmetric convolution. Finally, the continuous synthesis transform presented in Equation 37 may be written as

$$f = \mathbf{Z}^{-1}\big(\Phi \mathbf{Z} W^{\Phi} + \sum_{jj'} \Psi^{jj'\dagger} \mathbf{Q} \, W^{\Psi^{jj'}}\big), \tag{46}$$

where the summation performs a tensor contraction over each scale $\{j, j'\}$.

5.2. *Precomputed Components*

Naïve matrix representations of $\mathbf{Q}$ and $\mathbf{Z}$ can unsurpsingly become rather large. Noting that the memory complexity of a function $f \in \mathrm{L}^2[\mathbb{H}^4]$ scales as $\mathcal{O}(NPL^2)$, and that P typically weakly scales as L, we cannot hope for better than $\sim \mathcal{O}(NL^3)$ memory requirements. With some care, the memory complexity for matrix representations of $\mathbf{Q}$ and $\mathbf{Z}$ can be reduced to $\mathcal{O}(NL^3)$.

In the same way that the Wigner-Laguerre basis functions can be decomposed into the orthogonal Laguerre and Wigner basis functions, the Wigner-Laguerre transform can be expanded as $\mathbf{Q} = \mathbf{KD}$ where $\mathbf{K}$ represents the forward Laguerre transform defined in Equation 7. Furthermore, by decomposing the Wigner-D functions in terms of real polar d-functions

$$D^{\ell}_{mn}(\alpha, \beta, \gamma) = d^{\ell}_{mn}(\beta)e^{-i(n\gamma + m\alpha)} \tag{47}$$

the Wigner transform defined in Equation 4 reduces to

$$\hat{f} = \int_{\beta} d^{\ell}_{mn}(\beta) \underbrace{\int_{\alpha, \gamma} \mathrm{d}\Omega(\rho) \, f(\rho) \, e^{-i(n\gamma + m\alpha)}}_{\text{Fourier Transform}}. \tag{48}$$

In the discrete setting one may evaluate this double integral as a 2-dimensional fast Fourier transform (Cooley & Tukey 1965), which require minimal memory and are famously extremely efficient with $\mathcal{O}(N^2 \log N)$ complexity. The integral over β is effectively a projection onto the real polar d-functions, which are often computed recursively (Varshalovich et al. 1989). However, to store all necessary $d^{\ell}_{mn}(\beta)$ requires $\mathcal{O}(NL^3)$ memory which is equivalent to the memory required to store a single signal on the ball, as outlined above. Therefore, for our purposes we suffer no additional memory complexity from simply precomputing and caching these matrices, with the enormous upside of theoretically optimal speed at runtime.

For completeness, consider the Laguerre transform $\mathbf{K}$ defined in Equation 7. This is simply a projection onto

the Laguerre polynomials $K_p(r)$ with memory complexity $\sim \mathcal{O}(P^2) \leq \mathcal{O}(L^2)$ and therefore negligible. Hence, we may also precompute and store all $K_p(r)$ for efficiency. Finally, we must consider the memory required to store the wavelet Ψ and scaling Φ filters. The scaling filter Φ comes with memory complexity $\mathcal{O}(PL) \leq \mathcal{O}(L^2)$ which is negligible. In contrast, the wavelet filters Ψ exhibit memory complexity $\mathcal{O}(JJ'NPL)$ which can nominally be rather large. Adopting the multiresolution scheme discussed in Section 3.3 only the highest scale $\{J, J'\}$ need be stored at full resolution, and thus the complexity becomes $\mathcal{O}(NPL) \leq \mathcal{O}(L^3)$. We are therefore free to compute and cache all necessary tensors offline with at most $\mathcal{O}(NL^3)$ memory, which we subsequently use to execute transforms extremely efficiently at runtime.

Identical arguments hold for the spherical-Laguerre transform and their corresponding inverse transforms. In S2BALL we default to this operational modality, which has the added benefit of automatically providing the adjoint wavelet transformations needed for optimisation based algorithms.

5.3. *JAX Tensor Operations*

As discussed above, we can precompute effectively all intermediate tensors. Consequently, implementing extremely distributed operations is very straightforward. The entire ball wavelet transform can be written using the aforementioned `einsum` and the `JAX` fast Fourier transform API. This transform is outlined in Algorithm 2 which includes a sketch of the associated code.

In this case providing cross GPU distribution through SPMD is not necessary. At low resolutions the additional communication overhead outweighs the potential acceleration. At high resolutions this balance may tip in the favour of distribution across multiple GPUs, however long before this point the memory overhead becomes untenable. In theory, one could shard precomputed tensors across multiple (potentially many) GPU devices, however we leave this for future work.

5.4. *Numerical Validation*

We benchmark our directional wavelet transform against the existing C alternative FLAGLET ⟳ (Leistedt & McEwen 2012; McEwen & Leistedt 2013). Our protocol is the same as outlined in Section 4.4, with spherical harmonic transforms $\mathbf{Y}$ interchanged with spherical-Laguerre transforms $\mathbf{Z}$. FLAGLET transforms were executed on a multithreaded Xeon(R) E5-2650L v3 dedicated CPU and S2BALL transforms were executed on a single NVIDIA A100 GPU.

During these tests we restrict ourselves to azimuthal bandlimit $N = 1$ for simplicity, and maintain the most general setting of $P = L$ radially. It should be noted that in practice one may work with different radial and angular resolutions, which are often dictated by the data at hand. The results of this benchmark are presented in Table 2, which includes both round-trip precision and timings. In every case we recover 64-bit machine precision. For low bandlimits $L \sim 8$ our ball wavelet transform is an order of magnitude faster than their C counterparts. At higher resolutions this acceleration increases dramatically, peaking at 21800 times faster for $L = P = 256$. This substantial acceleration opens up the possibility of

---

Algorithm 2 Directional wavelet transform on $\mathbb{B}^3$

```
from baller import laguerre
from baller import wigner_laguerre as wlaguerre
from jax.numpy import einsum
```
procedure ANALYSIS WAVELET TRANSFORM($f \in \mathbb{B}^3$):

$\quad \hat{f} \leftarrow \mathbf{Z}f$ ▷ `laguerre.forward(`f,L,P`)`

$\quad \hat{W}^\Phi \leftarrow \mathbf{\Phi}\hat{f}$ ▷ `einsum("plm,pl->plm",`$\hat{f},\Phi$`)`

$\quad W^\Phi \leftarrow \mathbf{Z}^{-1}\hat{W}^\Phi$ ▷ `laguerre.inverse(`$\hat{W}^\Phi,L,P$`)`

$\quad$ **for** $j \in [0, J]$ and $j' \in [0, J']$ **do**

$\quad\quad \hat{W}^{\Psi^{jj'}} \leftarrow \mathbf{N}\Psi^{jj'}\hat{f}$ ▷ `einsum("plm,pln->pnlm",`$\hat{f},\mathbf{N}\Psi^{jj'*}$`)`

$\quad\quad W^{\Psi^{jj'}} \leftarrow \mathbf{Q}^{-1}\hat{W}^{\Psi^{jj'}}$ ▷ `wlaguerre.inverse(`$\hat{W}^{\Psi^{jj'}},L,P,N$`)`

$\quad$ **return** $\{W^\Psi, W^\Phi\}$

procedure SYNTHESIS WAVELET TRANSFORM($\{W^\Psi, W^\Phi\}$):

$\quad \hat{W}^\Phi \leftarrow \mathbf{Z}W^\Phi$ ▷ `laguerre.forward(`W^Φ,L,P`)`

$\quad \hat{f} \leftarrow \mathbf{\Phi}\hat{W}^\Phi$ ▷ `einsum("plm,pl->plm",`$\hat{W}^\Phi,\Phi$`)`

$\quad$ **for** $j \in [0, J]$ and $j' \in [0, J']$ **do**

$\quad\quad \hat{W}^{\Psi^{jj'}} \leftarrow \mathbf{Q}W^{\Psi^{jj'}}$ ▷ `wlaguerre.forward(`$W^{\Psi^{jj'}},L,P,N$`)`

$\quad\quad \hat{f} \stackrel{+}{\leftarrow} \mathbf{\Psi}^\dagger \hat{W}^{\Psi^{jj'}}$ ▷ `einsum("pnlm,pln->plm",`$\hat{W}^{\Psi^{jj'}},\Psi^{jj'}$`)`

$\quad f \leftarrow \mathbf{Z}^{-1}\hat{f}$ ▷ `laguerre.inverse(`$\hat{f},L,P$`)`

$\quad$ **return** f

---

computationally expensive statistical methods such as Bayesian sampling methods, with direct applications in *e.g.* geophysical imaging (Marignier et al. 2021).

An additional factor worth highlighting is somewhat hidden within the wavelet transform. The spherical-and Wigner-Laguerre transforms we implement and use to construct the S2BALL wavelet transforms are, in and of themselves, interesting and of potential use. In much the same way that spherical harmonic and Wigner transforms can be seen as generalised Fourier transforms (GFFTs) over $\mathbb{S}^2$ and $\mathrm{SO}(3)$ respectively, the spherical- and Wigner-Laguerre transforms can be seen to be GFFTs on $\mathbb{B}^3$ and $\mathbb{H}^4$ respectively. In addition to wavelets, S2BALL provides JAX functions with which one may readily evaluate these transforms.

Efficient, differentiable, and exact Fourier based convolution algorithms on $\mathbb{B}^3$ and $\mathbb{H}^4$ present an exciting opportunity to develop rotational and radially translational equivariant networks on 3-dimensional spaces; a natural progression of equivariant learning on the sphere (see *e.g.* Cohen et al. 2018; Esteves et al. 2020; Bronstein et al. 2021; Cobb et al. 2021; Ocampo et al. 2023). Not only do such networks provide state-of-the-art performance but they are much more data-efficient, which is a key consideration for modern astrophysics.

6. CONCLUSIONS

In this article we develop novel efficient algorithms for the directional scale-discretised wavelet transforms on the sphere and ball. By construction, our transforms are automatically differentiable and highly distributable both on and across hardware accelerators, *e.g.* GPUs/TPUs. To increase accesibility, and in the spirit of reproducible science, we provide professionally developed open-sourced JAX libraries in both settings.

For the multiscale analysis of functions on the sphere, we publicly release S2WAV ⟳. This package provides JAX algorithms for scalable and automatically differ-

Validation of directional ball wavelet ($P = L$)			
L	Time FLAGLET	Time S2BALL	Ratio
8	7.0×10^0	7.5×10^{-1}	13
16	1.0×10^2	9.5×10^{-1}	116
32	8.3×10^2	1.3×10^0	698
64	1.0×10^4	3.0×10^0	3690
128	1.0×10^5	1.7×10^1	6400
256	3.2×10^6	1.6×10^2	21800

TABLE 2. Numerical validation of our Wigner-Laguerre wavelet transform against the previous implementation FLAGLET. In each case we consider $N = 1$ and maintain a highly general setting where $L = P$. All CPU operations were evaluated on a multithreaded Xeon(R) E5-2650L v3 CPU whereas our transforms were evaluated on a single NVIDIA A100 GPU. The overall complexity scaling of the Wigner-Laguerre transform is $\sim \mathcal{O}(NL^4)$, with an associated memory overhead of $\mathcal{O}(NL^3)$ *i.e.* the memory required to store the highest scale wavelet coefficients a single function $f \in \mathbb{H}^4$. In all cases we are at least an order of magnitude faster, rising as high as $\sim 22,000$ times faster, whilst being exact to machine precision.

entiable directional wavelet transforms on the sphere. These transforms are sampling agnostic, and may be distributed across many GPUs for extreme parallelisation. Through benchmarking we find that our transforms are several orders of magnitude faster than their C counterparts, whilst retaining 64-bit floating point precision.

For the multiscale analysis of functions on the ball, we publicly release S2BALL ⭕. This package provides JAX algorithms for extremely fast but memory intensive directional wavelet transforms on the ball. As part of S2BALL, we provide APIs to evaluate differentaible fast Fourier transforms on the ball and its SO(3) anologue. Our transforms are built on a well-known Laguerre discretisation of the radial half-line, however unlike previous methods our angular components are sampling agnostic. All transforms are automatically differentiable. Through benchmarking we find that our transforms are ~ 22 thousand times faster than their C counterparts, whilst retaining 64-bit floating point precision.

The algorithms and software we provide are critical for the fusion of wavelet theory and machine learning on the sphere and ball, unlocking the advantages such a union entails. Wavelet enhanced technologies over Euclidean spaces are already producing state-of-the-art results, *e.g.* as embedded representations (Huang et al.

2017), for multiscale conditioning during diffusion-based generative models (Guth et al. 2022), or for equivariant machine learning (McEwen et al. 2022; Ocampo et al. 2023). This work provides the tools by which anologous results may be realised on the sphere and ball.

Concomitantly with this work, we are currently developing spherical scattering covariances (an extremely compact wavelet-based representation) which effectively encodes complex non-Gaussian structure and from which realistic cosmological fields may readily be generated (Mousset et al. in prep), which relies directly on this current work. In future work we will encorporate the results of this work to enhance diffusion models on the sphere, with a plethora of applications including weather and climate prediction tasks. As differentiable programming and hardware acceleration grows in popularity throughout the scientific community, it is pertinent that our software tools are modernised. In this work we do just this, infusing the previous generation of multiscale analysis tools with characteristics necessary for integration to truly next generation technologies.

CONTRIBUTION STATEMENT

Author contributions are specified below, following the Contributor Roles Taxonomy (CRediT[2]).
Matthew A. Price: Conceptualisation, Methodology, Software, Investigation, Validation, Writing (Original Draft, Review & Editing, Supervision) **Jason D. McEwen:** Conceptualisation, Methodology, Writing (Review & Editing, Supervision). **Alicja Polanska & Jessica Whitney:** Software, Investigation, Validation.

REFERENCES

Abramo L.R., Reimberg P.H., Xavier H.S., 2010, Physical Review D, 82, 4, 043510
Allys E., Levrier F., Zhang S., Colling C., Regaldo-Saint Blancard B., Boulanger F., Hennebelle P., Mallat S., 2019, Astronomy & Astrophysics, 629, A115
Allys E., Marchand T., Cardoso J.F., Villaescusa-Navarro F., Ho S., Mallat S., 2020, Physical Review D, 102, 10, 103506
Antoine J.P., Vandergheynst P., 1998, 39, 8, 3987
Antoine J.P., Vandergheynst P., 1999, 7, 1
Audet P., 2011, 116, E1
Balan R., Bodmann B.G., Casazza P.G., Edidin D., 2009, Journal of Fourier Analysis and Applications, 15, 4, 488
Baldi P., Kerkyacharian G., Marinucci D., Picard D., 2009, 37 No.3, 1150, arXiv:math/0606599
Boomsma W., Frellsen J., 2017, Advances in neural information processing systems, 30

Bradbury J., et al., 2018, JAX: composable transformations of Python+NumPy programs
Bronstein M.M., Bruna J., Cohen T., Veličković P., 2021, arXiv preprint arXiv:2104.13478
Bruna J., Mallat S., 2013, IEEE transactions on pattern analysis and machine intelligence, 35, 8, 1872
Carrillo R.E., McEwen J.D., Wiaux Y., 2014, in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), invited contribution*, arXiv:1406.0359
Chan J.Y.H., Leistedt B., Kitching T.D., McEwen J.D., 2017, IEEE Trans. Sig. Proc., 1, 5, arXiv:1511.05578
Cheng S., Ting Y.S., Ménard B., Bruna J., 2020, Monthly Notices of the Royal Astronomical Society, 499, 4, 5902
Choi C.H., Ivanic J., Gordon M.S., Ruedenberg K., 1999, J. Chem. Phys., 111, 19, 8825
Cobb O.J., Wallis C.G.R., Mavor-Parker A.N., Marignier A., Price M., d'Avezac M., McEwen J.D., 2021, in *International Conference on Learning Representations*, arXiv:2010.11661
Cohen T.S., Geiger M., Köhler J., Welling M., 2018, arXiv preprint arXiv:1801.10130
Cooley J.W., Tukey J.W., 1965, 19, 297

Driscoll J.R., Healy D.M.J., 1994, Adv. Appl. Math., 15, 202
Durastanti C., Fantaye Y., Hansen F., Marinucci D., Pesenson
    I.Z., 2014, 90, 10, 103532, arXiv:1408.1095
Eickenberg M., et al., 2022, arXiv preprint arXiv:2204.07646
Esteves C., Makadia A., Daniilidis K., 2020, Advances in Neural
    Information Processing Systems, 33, 8614
Geller D., Hansen F.K., Marinucci D., Kerkyacharian G., Picard
    D., 2008, 78, 12, 123533, arXiv:0811.2881
Goldberg J.N., Macfarlane A.J., Newman E.T., Rohrlich F.,
    Sudarshan E.C.G., 1967, 8, 11, 2155
Goodwin-Allcock T., McEwen J.D., Gray R., Nachev P., Zhang
    H., 2022, in CDMRI 2022, Springer, 101–112
Górski K.M., Hivon E., Banday A.J., Wandelt B.D., Hansen F.K.,
    Reinecke M., Bartelmann M., 2005, 622, 759, astro-ph/0409513
Gupta G., Xiao X., Bogdan P., 2021, Advances in neural
    information processing systems, 34, 24048
Guth F., Coste S., De Bortoli V., Mallat S., 2022, Advances in
    Neural Information Processing Systems, 35, 478
Huang H., He R., Sun Z., Tan T., 2017, in Proceedings of the
    IEEE international conference on computer vision, 1689–1697
Huertas-Company M., Lanusse F., 2023, Publications of the
    Astronomical Society of Australia, 40, e001
Jiang C., Huang J., Kashinath K., Marcus P., Niessner M., et al.,
    2019, arXiv preprint arXiv:1901.02039
Jumper J., et al., 2021, Nature, 596, 7873, 583
Khalid Z., Kennedy R.A., McEwen J.D., 2016, Applied Comput.
    Harm. Anal., 40, 3, 470, arXiv:1403.5553
Kondor R., Lin Z., Trivedi S., 2018, Advances in Neural
    Information Processing Systems, 31
Lanusse F., Rassat A., Starck J.L., 2012, 540, A92,
    arXiv:1112.0561
Leistedt B., McEwen J.D., 2012, IEEE Trans. Sig. Proc., 60, 12,
    6257, arXiv:1205.0792
Leistedt B., McEwen J.D., Kitching T.D., Peiris H.V., 2015,
    Phys. Rev. D., 92, 123010, arXiv:1509.06750
Leistedt B., McEwen J.D., Vandergheynst P., Wiaux Y., 2013,
    Astron. & Astrophys., 558, A128, 1, arXiv:1211.1680
Lemoine D., 1994, The Journal of chemical physics, 101, 5, 3936
Liu P., Zhang H., Lian W., Zuo W., 2019, IEEE Access, 7, 74973
Loureiro A., Whiteway L., Sellentin E., Lafaurie J.S., Jaffe A.H.,
    Heavens A.F., 2022, arXiv preprint arXiv:2210.13260
Mallat S., 2012, Communications on Pure and Applied
    Mathematics, 65, 10, 1331, arXiv:1101.2286
Mallat S., Zhang S., Rochette G., 2020, Information and
    Inference: A Journal of the IMA, 9, 3, 721
Marignier A., Ferreira A.M., Kitching T., 2020, Geochemistry,
    Geophysics, Geosystems, 21, 9, e2020GC009276
Marignier A., McEwen J.D., Ferreira A.M.G., Kitching T.D.,
    2021, IEEE Trans. Image Proc., submitted, arXiv:2107.06500
McEwen J.D., Büttner M., Leistedt B., Peiris H.V., Wiaux Y.,
    2015a, IEEE Sig. Proc. Let., 22, 12, 2425, arXiv:1508.03101
McEwen J.D., Durastanti C., Wiaux Y., 2018, Applied Comput.
    Harm. Anal., 44, 1, 59, arXiv:1509.06767
McEwen J.D., Hobson M.P., Lasenby A.N., Mortlock D.J., 2008a,
    Mon. Not. Roy. Astron. Soc., 388, 2, 659, arXiv:0803.2157
McEwen J.D., Hobson M.P., Mortlock D.J., Lasenby A.N., 2007,
    IEEE transactions on Signal Processing, 55, 2, 520
McEwen J.D., Leistedt B., 2013, in 10th International Conference
    on Sampling Theory and Applications (SampTA), invited
    contribution, 329–333, arXiv:1307.1307
McEwen J.D., Leistedt B., Büttner M., Peiris H.V., Wiaux Y.,
    2015b, IEEE Trans. Sig. Proc., submitted, arXiv:1509.06749
McEwen J.D., Price M.A., 2019, in 27th European Signal
    Processing Conference (EUSIPCO), arXiv:1510.01595
McEwen J.D., Wallis C.G.R., Mavor-Parker A.N., 2022, in
    International Conference on Learning Representations, in
    press, arXiv:2102.02828
McEwen J.D., Wiaux Y., 2011, IEEE Trans. Sig. Proc., 59, 12,
    5876, arXiv:1110.6298

McEwen J.D., Wiaux Y., Eyers D.M., 2011, Astron. &
    Astrophys., 531, A98, 1, arXiv:1108.3900
McEwen J.D., Wiaux Y., Hobson M.P., Vandergheynst P.,
    Lasenby A.N., 2008b, Mon. Not. Roy. Astron. Soc., 384, 4,
    1289, arXiv:0704.0626
Mousset L., Price M.A., Allys E., McEwen J.D., ????, Astronomy
    & Astrophysics, in prep
Newman E.T., Penrose R., 1966, 7, 5, 863
Ntampaka M., et al., 2019, arXiv preprint arXiv:1902.10159
Ocampo J., Price M.A., McEwen J.D., 2023, in The Eleventh
    International Conference on Learning Representations
Pedersen C., Eickenberg M., Ho S., 2023, arXiv preprint
    arXiv:2307.14362
Pratley L., McEwen J.D., d'Avezac M., Carrillo R.E., Onose A.,
    Wiaux Y., 2018, Mon. Not. Roy. Astron. Soc., 473, 1, 1038,
    arXiv:1610.02400
Price M.A., Mars M., Docherty M.M., Mancini A.S., Marignier
    A., McEwen J., 2023, The Open Journal of Astrophysics,
    arXiv:2307.04798
Price M.A., McEwen J.D., 2021, IEEE Sig. Proc. Let., submitted,
    arXiv:2105.05518
Price M.A., McEwen J.D., 2023, Journal of Computational
    Physics, submitted, arXiv:2311.14670
Price M.A., McEwen J.D., Cai X., Kitching T.D., Wallis C.G.R.,
    2021a, Mon. Not. Roy. Astron. Soc., in press, arXiv:1812.04014
Price M.A., McEwen J.D., Pratley L., Kitching T.D., 2021b,
    Mon. Not. Roy. Astron. Soc., 500, 4, 5436, arXiv:2004.07855
Racah E., Beckham C., Maharaj T., Ebrahimi Kahou S., Prabhat
    M., Pal C., 2017, Advances in neural information processing
    systems, 30
Rassat A., Refregier A., 2012, Astronomy & Astrophysics, 540,
    A115
Ravuri S., et al., 2021, Nature, 597, 7878, 672
Ritchie D.W., Kemp G.J.L., 1999, J. Comput. Chem., 20, 4, 383
Robitaille J.F., Abdeldayem A., Joncour I., Moraux E., Motte F.,
    Lesaffre P., Khalil A., 2020, Astronomy & Astrophysics, 641,
    A138
Roddy P.J., McEwen J.D., 2021, IEEE Trans. Sig. Proc.,
    submitted, arXiv:2106.02023
Roddy P.J., McEwen J.D., 2023, arXiv preprint arXiv:2302.06006
Schröder P., Sweldens W., 1995, in Computer Graphics
    Proceedings (SIGGRAPH '95), 161–172
Simons F.J., et al., 2011, 187, 969, 1104.3151
Starck J., Moudden Y., Bobin J., 2009, 497, 931,
    arXiv:0902.0574
Starck J.L., Moudden Y., Abrial P., Nguyen M., 2006, 446, 1191,
    astro-ph/0509883
Tripura T., Chakraborty S., 2022, arXiv preprint
    arXiv:2205.02191
Tuch D.S., 2004, 52, 6, 1358, ISSN 1522-2594
Varshalovich D.A., Moskalev A.N., Khersonskii V.K., 1989,
    Quantum theory of angular momentum, World Scientific,
    Singapore
Wallis C.G.R., Price M.A., McEwen J.D., Kitching T.D., Leistedt
    B., Plouviez A., 2021, Mon. Not. Roy. Astron. Soc., in press,
    arXiv:1703.09233
Weyn J.A., Durran D.R., Caruana R., 2020, Journal of Advances
    in Modeling Earth Systems, 12, 9, e2020MS002109
Wiaux Y., McEwen J.D., Vandergheynst P., Blanc O., 2008, Mon.
    Not. Roy. Astron. Soc., 388, 2, 770, arXiv:0712.3519
Zhang C., Liwicki S., Smith W., Cipolla R., 2019, in Proceedings
    of the IEEE/CVF International Conference on Computer
    Vision, 3533–3541
Zhang S., Mallat S., 2021, Applied and Computational Harmonic
    Analysis, 53, 199

APPENDIX

A. WAVELET CONSTRUCTION

Scale-discretised wavelets on the sphere and ball $_s\Psi^{jj'}$ provide a natural dictionary in which many physical fields are sparsely distributed. For exact decomposition and synthesis into and from a wavelet dictionary respectively $_s\Psi^{jj'}$ must satisfy the admissibility condition

$$\frac{4\pi}{2\ell+1}\,|_s\phi_{\ell 0,p}|^2 + \frac{8\pi^2}{2\ell+1}\sum_{ijm}|_s\Psi^{jj'}_{\ell m,p}|^2 = 1, \tag{A1}$$

for all ℓ and p. When considering functions with no radial extent one may simply ignore both p and radial scales j'. A key factor which increases the efficacy of wavelets is their localisation properties, and so careful consideration of how $_s\Psi^{jj'}$ are constructed is critically important. In this section we consider the scale-discretised wavelets which are separable in the harmonic ℓ, azimuthal m, and radial p components and satisfy the admissibility condition in Equation A1. In the following sections we discuss one such choice of these tiling functions which has proved to be particularly effective.

A.1. *Tiling of the harmonic line*

To capture information along the harmonic line ℓ we define positive real functions $\kappa^{(j)}(t)$ for $t \in \mathbb{N}^+$ such that $t \leq L$, where j denotes angular wavelet scale. The construction of this function is given by the following. First, consider the infinitely differentiable C^∞ Schwartz function

$$s_\lambda(t) = s\left(\frac{2\lambda}{\lambda-1}(t - \frac{1}{\lambda}) - 1\right), \tag{A2}$$

with compact support $t \in [\lambda^{-1}, 1]$ and for dilation parameter $\lambda \in \mathbb{R}^+_{>1}$. Note that the most common choice of dilation parameter is $\lambda = 2$ in which case we recover dyadic wavelets. The function $s(t)$ is given by

$$s(t) = \exp\left(-\frac{1}{1-t^2}\right), \tag{A3}$$

for $t \in [-1, 1]$ and 0 elsewhere. From these functions with compact support we can define the function

$$k_\lambda(t) = \frac{\int_t^1 \frac{\mathrm{d}t'}{t'} s_\lambda^2(t')}{\int_{\lambda^{-1}}^1 \frac{\mathrm{d}t'}{t'} s_\lambda^2(t')}, \tag{A4}$$

which decays smoothly from unity at $t < \lambda^{-1}$ to zero at $t > 1$. We can now focus the localisation in t by defining

$$\kappa_\lambda(t) = \sqrt{k_\lambda(\lambda^{-1}t) - k_\lambda(t)}, \tag{A5}$$

which is the generating function for the harmonic component of the wavelet construction, with which the scale-discretised wavelet harmonic kernel for scale j is defined by $\kappa^{(j)}(\ell) = \kappa_\lambda(\lambda^j L^{-1}\ell)$, which is compact on $\ell \in \left[\lfloor \lambda^{-(1+j)}L \rfloor, \lceil \lambda^{1-j}L \rceil \right]$. With this construction, it is clear that each wavelet scale need only be evaluated between some lower and upper harmonic degrees, which we will denote hereon out as $L_{j\pm}$ respectively. To capture low frequency information one must also introduce a scaling tiling function $\eta_\lambda(t) = \sqrt{k_\lambda(t)}$.

A.2. *Tiling of the azimuthal line*

To capture the azimuthal (directional) information content of functions on the sphere we define the spin-s square integrable function $_s\zeta \in \mathrm{L}^2[\mathbb{S}^2]$ with harmonic representation $\langle _s\zeta, _sY_{\ell m}\rangle$, following Equation 2. An azimuthal bandlimited N is introduced such that $_s\zeta_{\ell m} = 0$ for all ℓ, m where $|m| \geq N$, in much the same way as was done for bandlimited functions on the sphere. The azimuthal tiling function is defined as

$$_s\zeta_{\ell m} = \xi\mu\sqrt{\frac{1}{2^q}\binom{q}{(q-m)/2}}, \tag{A6}$$

which is a specific form of the directional auto-correlation with $\xi = 1$ for even $N - 1$ and $\xi = i$ else, and $\mu = [1 - (-1)^{N+m}]/2$ and $q = \min\{N - 1, \ell - [1 + (-1)^{N+\ell}]/2\}$. The full details of this derivation are extremely involved and can be found in related work.

A.3. *Tiling of the ball*

The wavelet generating function on the radial half-line is in fact identical to that of the harmonic line, the only difference being that here we consider tiling over the associated 2nd-order Laguerre polynomials indexed by p. Therefore identical logic to that of Section A.1 may be applied to derive the radial tiling function $\kappa^{(j')}(p)$ and scaling function $\eta(p)$, where we define a radial dilation parameter $\lambda \to \nu$.

Combining this radial tiling and the harmonic tiling discussed in Sections A.1 one may construct a hybrid generating function for scale-discretised wavelets on the ball, given simply by

$$\kappa^{(jj')}(\ell, p) = \kappa^{(j)}(\ell)\,\kappa^{(j')}(p). \tag{A7}$$

We finally define the hybrid scaling function generator

$$\eta_{\lambda\nu}(t, t') = \sqrt{k_\lambda\left(\frac{t}{\lambda}\right)k_\nu(t') + k_\lambda(t)k_\nu\left(\frac{t'}{\nu}\right) - k_\lambda(t)k_\nu(t')}, \tag{A8}$$

which both captures low frequency information and ensures the constructed dictionary on the ball satisfies the wavelet admissibility condition.

A.4. *Scale-discretised Wavelet Dictionary*

Combining the results from this section we can explicitly define our directional spin-s wavelet functions in their harmonic representation

$$_s\Psi_{\ell m p}^{j j'} = \sqrt{\frac{2\ell + 1}{8\pi^2}}\, \kappa^j(\ell)\, \kappa^{j'}(p)\, _s\zeta_{\ell m}. \tag{A9}$$

The corresponding scaling functions $_s\Phi$ are given by

$$_s\Phi_{\ell 0 p} = \begin{cases} \sqrt{\frac{2\ell+1}{4\pi}}\eta_\nu\left(\frac{p}{\nu^{J_0}}\right) & \text{if } \ell > \lambda^{J_0'}, p \le \nu^{J_0} \\ \sqrt{\frac{2\ell+1}{4\pi}}\eta_\lambda\left(\frac{\ell}{\lambda^{J_0'}}\right) & \text{if } \ell \le \lambda^{J_0'}, p > \nu^{J_0} \\ \sqrt{\frac{2\ell+1}{4\pi}}\eta_{\lambda\nu}\left(\frac{\ell}{\lambda^{J_0'}}, \frac{p}{\nu^{J_0}}\right) & \text{if } \ell < \lambda^{J_0'}, p < \nu^{J_0}, \end{cases} \tag{A10}$$

and zero elsewhere. Here J_0 and J_0' denote the minimum wavelet scales jj', below which the scaling function captures any remaining information.