arXiv:2611.04389v1 [cs.CV] 3 Nov 2026

Spin Spherical Fourier Neural Operators: Efficient Equivariant Learning for Vector and Spin Fields on the Sphere

Kevin Mulder [1] **Matthew A. Price** [1] **Eirini Leftaki** [1] **Jason D. McEwen** [1 2]

Abstract

Many fields of physical interest on the sphere, such as wind, the polarisation of the cosmic microwave background and the shear of weak gravitational lensing, are spin fields, whose values change phase when the local frame rotates. We extend spherical Fourier neural operators (SFNOs) to fields of any spin, through lifted spherical convolutions that are exactly rotationally equivariant. Lifting spin fields to Mackey functions on the rotation group provides the conceptual framework, while computation is performed exactly and efficiently through their harmonic representation and by appealing to sampling theorems on the sphere and rotation group. We construct spin SFNOs (S^2FNOs) from these building blocks. We demonstrate empirically that rotational equivariance holds to machine precision for these layers. We achieve state-of-the-art performance on existing benchmarks, and introduce new benchmarks on which spin-aware models outperform their scalar counterparts.

1. Introduction

In many fields of science data are measured on the sphere, from the relic radiation of the Big Bang and the shapes of distant galaxies in cosmology, to the winds of the atmosphere and the currents of the ocean. Many of these data sets are not scalar. The horizontal wind and the ocean currents are vector, or spin-1, fields, while the polarisation of the cosmic microwave background (CMB) and the shear induced by weak gravitational lensing are spin-2 fields (Zaldarriaga & Seljak, 1997; Castro et al., 2005; Wallis et al., 2022). The value of a spin-s field at a point is defined relative to a frame in the tangent plane and acquires the phase $e^{-\mathrm{i}s\chi}$ when that frame rotates by χ.

[1]Department of Space & Climate Physics, University College London, London, UK [2]The Alan Turing Institute, London, UK. Correspondence to: Jason D. McEwen <jason.mcewen@ucl.ac.uk>.

Preprint. November 4, 2026.

On the sphere, rotations play the role that translations play in the plane. Spherical CNNs consequently build on $\mathrm{SO}(3)$-equivariant convolutions (Cohen et al., 2018; Esteves et al., 2018; Kondor et al., 2018; Cobb et al., 2021; Ocampo et al., 2023), and spherical Fourier neural operators (SFNOs) (Bonev et al., 2023) bring the operator-learning perspective of Fourier neural operators (Li et al., 2021; Kovachki et al., 2023) to the sphere. When applied to vector or spin fields, however, scalar spherical CNNs and SFNOs typically consider the components of spin fields (*i.e.* real and imaginary components) as unrelated scalar channels. This is the standard approach in data-driven weather models for wind (Pathak et al., 2022; Bi et al., 2023; Lam et al., 2023; Esteves et al., 2023), which does not respect general rotations.

Equivariant networks for spin fields exist (Esteves et al., 2020; Xu et al., 2022; Ballerin et al., 2025), but at a price. Those that convolve on the full rotation group are costly in computation and parameters. The group CNN of Ballerin et al. (2025), for example, has about 2 million parameters on spherical MNIST (whereas we achieve similar performance with 50 thousand parameters). The equivariance of these networks, moreover, is asserted or illustrated rather than quantified, and benchmarks for spin fields are scarce. Our layers build on McEwen (2026), who makes the lift–convolve–project construction of Cohen et al. (2019, Eq. 14) and Xu et al. (2022) explicit on the sphere and shows that it unifies the spherical convolutions of the literature, so that our layers can learn filters of each of these forms.

Our contributions are fourfold. First, we extend spherical Fourier neural operators to fields of any spin (Sec. 4), building spin SFNOs (S^2FNOs) from two blocks, the spin spherical convolution and the convolution on the rotation group, so that rotational equivariance is exact for every convolutional layer and holds in the continuum for every activation. Second, we describe the layers by lifting spin fields to Mackey functions on the rotation group, which provides their conceptual framework and proofs (McEwen, 2026), but compute them in harmonic space to ensure S^2FNOs are efficient in representation, computation and parameters (Sec. 5). Third, we show that the convolutional layers of S^2FNOs are exactly equivariant on spin fields, with errors of order 10^{-15}, that whole S^2FNO blocks and UNets with the activations

and the normalisation removed are equivariant to round-off, and we discuss the equivariance errors introduced by the other layers (Sec. 6). Fourth, we introduce three new public benchmarks for spin fields spanning classification, weather and cosmology (Sec. 7), on which our networks achieve state-of-the-art performance. Our layers, networks and benchmarks will be released with the `s2ai` code (`https://github.com/astro-informatics/s2ai`).

2. Background and related work

2.1. Spin fields and their harmonic analysis

A spin-s field ${}_sf$ is a complex function on the sphere whose values are defined relative to a local tangent frame and transform as

$$ {}_sf'(x) = e^{-\mathrm{i}s\chi}\ {}_sf(x) \tag{1}$$

under a right-handed rotation of the frame by χ (Newman & Penrose, 1966; Goldberg et al., 1967). We write ${}_sf_{\ell m} = \langle {}_sf, {}_sY_{\ell m}\rangle$, with $\ell \geq |s|$, for its spin spherical harmonic coefficients and $F^{\ell}_{mn} = \langle F, D^{\ell *}_{mn}\rangle$ for the Wigner coefficients of $F \in L^2(\mathrm{SO}(3))$, with the Wigner functions D^{ℓ}_{mn} in zyz Euler angles (α, β, γ) and the unnormalised Haar measure. The coefficients of every spin rotate by the same Wigner matrices under the spin rotation operator $\mathcal{R}_\rho$ of McEwen et al. (2015b) (Appendix A.1). Fields are bandlimited at L if their coefficients vanish for $\ell \geq L$, and functions on $\mathrm{SO}(3)$ azimuthally bandlimited at N if also $F^{\ell}_{mn} = 0$ for $|n| \geq N$. By appealing to the sampling theorems of McEwen & Wiaux (2011) and McEwen et al. (2015a), the `s2fft` code (Price & McEwen, 2024) transforms such signals exactly, at $\mathcal{O}(L^3)$ for spin spherical harmonic transforms and $\mathcal{O}(NL^3)$ for Wigner transforms.

2.2. Neural operators and equivariant networks on the sphere

The spherical Fourier neural operator (SFNO) of Bonev et al. (2023) learns one weight per spherical harmonic degree in the block structure of the Fourier neural operator (FNO) of Li et al. (2021). Like the FNO, it has a continuous representation through harmonic space, since the harmonic coefficients of bandlimited signals, computed exactly by appealing to sampling theorems, give access to the underlying continuous operator, as in the continuous approaches to spherical convolution described by Ocampo et al. (2023). Its layer is the non-localised zonal filter of Esteves et al. (2018), the isotropic scalar member of the generalised convolutions of Cobb et al. (2021). Spherical CNNs that learn underlying continuous representations (Esteves et al., 2018; Cobb et al., 2021) are thus SFNOs, neural operators in the sense of Kovachki et al. (2023), although they did not carry the name (Appendix B). This point is also made for the discrete-continuous (DISCO) convolutions of Ocampo et al. (2023), whose filters are continuous, in recent work that adopts them as local neural-operator layers (Liu-Schiaffini et al., 2024; Bonev et al., 2025).

On the rotation group, spherical CNNs apply pointwise non-linearities (Cohen et al., 2018) or exact tensor products (Kondor et al., 2018); Cobb et al. (2021) combine both and develop more efficient tensor-product activations. For spin fields, Esteves et al. (2020) learn spin-changing harmonic filters, with non-linearities on each field, while Xu et al. (2022) lift to the rotation group only for a pointwise non-linearity before a fixed projection, and Ballerin et al. (2025) apply full group convolutions and pointwise non-linearities. In Secs. 4 to 7 we combine these ingredients in a neural operator for spin fields with the block structure of the SFNO, form the lift only where it is needed, for directional filters and the non-linear coupling of spins, measure equivariance on spin fields, and introduce benchmarks that include spin 2, whereas the experiments of the networks above use spins of at most 1.

3. Equivariant spin convolutions by lifting

Our layers rest on the generalised lifted convolution of McEwen (2026), which we recall concisely here, with the results that we use stated in Appendix A.

3.1. Lifting spin fields to the rotation group

Spin-s fields are the sections of the line bundle associated by the character $e^{\mathrm{i}s\gamma}$ to the rotation group, the principal $\mathrm{SO}(2)$-bundle of frames over the sphere (Cohen et al., 2019; Weiler et al., 2021; McEwen, 2026), and correspond to the Mackey functions of type s on the rotation group, those with $F(gR_z(\gamma)) = e^{-\mathrm{i}s\gamma}F(g)$. A spin-$s$ field is lifted to its Mackey function by Λ_s and returned by its adjoint, the projection $\Lambda_s^\dagger$, which integrates over the fibre against $e^{\mathrm{i}s\gamma}$:

$$ (\Lambda_s f)(\alpha, \beta, \gamma) = \frac{1}{\sqrt{2\pi}}\, e^{-\mathrm{i}s\gamma}\, f(\beta, \alpha); \tag{2}$$

$$ (\Lambda_s^\dagger F)(\beta, \alpha) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{\mathrm{i}s\gamma}\, F(\alpha, \beta, \gamma)\, \mathrm{d}\gamma. \tag{3}$$

The lift is an isometry with $\Lambda_s^\dagger \Lambda_s = \mathrm{Id}$, occupies the single Wigner column $n = -s$, $(\Lambda_s f)^{\ell}_{mn} = (-1)^s \sqrt{8\pi^2/(2\ell+1)}\ {}_sf_{\ell m}\, \delta_{n,-s}$, and is equivariant, intertwining the regular representation $\pi_{\mathrm{SO}(3)}$ with the spin rotation $\pi_s(R_\rho) = \mathcal{R}_\rho$. Conversely, a function on $\mathrm{SO}(3)$ is azimuthally bandlimited at N if and only if it is a sum of at most $2N-1$ lifted spin fields, one for each $|s| < N$.

3.2. Lifted convolutions in harmonic space

We adopt the group convolution, which for $\Psi, F \in L^2(\mathrm{SO}(3))$ reads:

$$(\Psi \star F)(\rho) = \int \Psi^*(\rho^{-1}\rho')\, F(\rho')\, \mathrm{d}\mu(\rho'); \tag{4}$$

$$(\Psi \star F)^\ell_{mn} = \sum\nolimits_{m'} F^\ell_{mm'}\, \Psi^{\ell *}_{nm'}. \tag{5}$$

The generalised lifted convolution lifts a spin field to the rotation group, convolves it there with a kernel Ψ and projects the result back to a spin field on the sphere. It resolves to four cases depending on whether a lift is applied to the input and a projection to the output:

$$\begin{aligned}
\mathcal{C}_{s\to s'}[\Psi] &= \Lambda^\dagger_{s'}(\Psi \star \Lambda_s\, \cdot\,), \\
\mathcal{C}_{s\to \mathrm{SO}(3)}[\Psi] &= \Psi \star \Lambda_s\, \cdot\,, \\
\mathcal{C}_{\mathrm{SO}(3)\to s'}[\Psi] &= \Lambda^\dagger_{s'}(\Psi \star \cdot\,), \\
\mathcal{C}_{\mathrm{SO}(3)\to \mathrm{SO}(3)}[\Psi] &= \Psi \star \cdot\,.
\end{aligned} \tag{6}$$

All four are equivariant for every square-integrable kernel, and by the single-column lift only one kernel entry per degree reaches a spin-s' output (McEwen, 2026),

$$\big(\mathcal{C}_{s\to s'}[\Psi]\ {}_s f\big)_{\ell m} = (-1)^{s+s'}\ {}_s f_{\ell m}\ \Psi^{\ell *}_{-s',-s}, \tag{7}$$

so that the spin spherical convolution $\mathcal{C}_{s\to s'}$ maps spin s to spin s' by one multiplier per degree, without forming a signal on the rotation group. Specialising the equivariant kernel theorems (Kondor & Trivedi, 2018; Kondor et al., 2018; Cohen et al., 2019; Xu et al., 2022) to the sphere shows that these multipliers are complete, in the sense that every bounded equivariant complex-linear map T from spin-s to spin-s' fields acts as

$$(T\ {}_s f)_{\ell m} = c_\ell\ {}_s f_{\ell m} \tag{8}$$

for a unique bounded sequence $(c_\ell)_{\ell \geq \max(|s|,|s'|)}$, and that every such sequence defines one (McEwen, 2026).

A directional layer $\mathcal{C}_{s\to \mathrm{SO}(3)}[\Psi]$ whose kernel column $m' = -s$ is supported on $|n| < N$ is exactly a tuple of at most $2N-1$ spin-changing layers, which relates the directional layers of Cobb et al. (2021) to the multi-spin layers of Esteves et al. (2020).

A key result of McEwen (2026, Corollary V.7) is that directional kernels, a spherical output of one spin and rotational equivariance cannot be achieved simultaneously. Between two fixed spins the kernel of an equivariant map is radial apart from the frame phases, so that a layer whose output is a field of one spin cannot be both directional and equivariant. Directional information must therefore be kept on the rotation group or, equivalently, in a tuple of spin fields.

Since the generalised lifted convolution reduces to the spherical convolutions of the literature, each recovered by the support of its kernel, the choice of spins and whether the lift or the projection is omitted (McEwen, 2026), layers built on it can learn filters of each of these forms rather than fix one by design. The spin spherical convolution contains every bounded equivariant complex-linear map between two spins, among them the axisymmetric convolutions of scalar spherical CNNs, the left convolution of Driscoll & Healy (1994), the spin-weighted convolutions of Esteves et al. (2020) and the spherical Kaiser–Squires relation of weak lensing (Kaiser & Squires, 1993; Wallis et al., 2022), and the convolution on the rotation group contains the directional, steerable and group convolutions (Wiaux et al., 2006; McEwen et al., 2015b; Cohen et al., 2018; Cobb et al., 2021). In our networks the filters are restricted by their parameterisation, through interpolated profiles in ℓ and, on the rotation group, a small internal spin set and a kernel diagonal on it.

4. Spin spherical Fourier neural operators

We build S^{2}FNOs from two blocks with exactly equivariant linear layers, joined by the lift Λ_s and its adjoint $\Lambda^\dagger_s$.

4.1. Convolutional building blocks

The spin spherical convolution (Fig. 1a) maps a tuple of spin fields to another by one multiplier (7) per degree for each pair of spins and channels, without explicitly forming the lift. It provides complete linear spin mixing (Sec. 3.2) but no non-linear coupling of spins, since its non-linearity acts on each field separately.

The convolution on the rotation group (Fig. 1b) acts on the Wigner coefficients of a function on $\mathrm{SO}(3)$ by (5), and its non-linearity acts pointwise on samples on the group, with $2N-1$ samples in γ. In our networks its kernel is diagonal on an internal spin set $\mathcal{S}$, so that the linear part filters each spin separately and the spins are coupled by the non-linearity on the group, which is much less constrained than one on the sphere (Sec. 4.2). The projection $\Lambda^\dagger_s$ then reads the Wigner column $n = -s$ of each spin $s \in \mathcal{S}$, and our implementation computes only these columns, at $\mathcal{O}(|\mathcal{S}|(L^3 + NL^2))$ rather than the $\mathcal{O}(NL^3)$ of full Wigner transforms (Sec. 5.2).

4.2. Equivariant non-linearities

Activations for spin fields may be applied in two settings, to the spin fields themselves on the sphere or pointwise to their lift on the rotation group; each setting brings its own constraint for equivariance. On the sphere a rotation turns the local frame and multiplies a spin-s value by the phase of (1), so that for $s \neq 0$ a pointwise activation is equivariant only if it commutes with that phase, $\zeta(e^{\mathrm{i}\phi} z) = e^{\mathrm{i}\phi}\zeta(z)$, and may therefore act on the modulus alone, as magnitude,

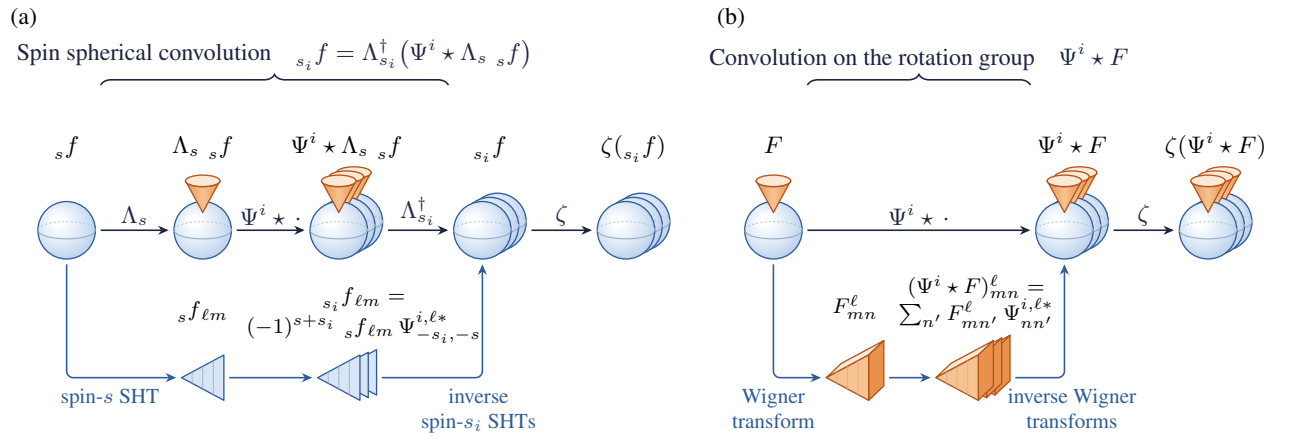


Figure 1. The spin spherical convolution (a), one multiplier per degree between spin spherical harmonic transforms (SHTs) without forming the lift, and the convolution on the rotation group (b), each followed by ζ and computed in harmonic space below. Spheres denote fields on S^2, spheres with a cone denote functions lifted to $\mathrm{SO}(3)$, triangles denote spin spherical harmonic coefficients and prisms Wigner coefficients, and i indexes output spins s_i in (a) and kernels in (b).

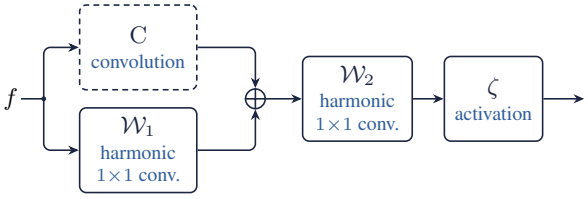


Figure 2. The S^{2}FNO block (10), with C the spin spherical convolution or the convolution on the rotation group, $\mathcal{W}_1$ and $\mathcal{W}_2$ harmonic 1×1 convolutions and ζ the activation (in detail in Fig. 10).

norm and gated non-linearities do (Worrall et al., 2017; Weiler et al., 2018; Esteves et al., 2020). On $\mathrm{SO}(3)$ rotations act on the argument alone, so that every pointwise map is equivariant (Cohen et al., 2018; Cobb et al., 2021; Ballerin et al., 2025), with a much weaker constraint. A tuple of spin fields may therefore be lifted and summed, passed through any pointwise map and returned (Appendix C).

On the sphere, for non-zero spin fields, we adopt the modulus non-linearity

$$\zeta(z) = \mathrm{ReLU}\big(a|z| + b\big)\,\frac{z}{|z| + \epsilon}, \tag{9}$$

with a learned real gain a and offset b, and $\epsilon = 10^{-8}$, which keeps the phase and is therefore equivariant for every spin. For scalar fields on the sphere ($s = 0$) we adopt standard activation functions. On $\mathrm{SO}(3)$ we apply a pointwise activation function to samples of the lifted tuple, such as the modulus non-linearity, and return to the sphere by the fixed projection $\Lambda_s^\dagger$ followed by a learned filter for each spin.

4.3. S^{2}FNO blocks

The S^{2}FNO block (Fig. 2), acting on a tuple f of spin fields,

$$f \mapsto \zeta\big(\mathcal{W}_2(\mathrm{C}\, f + \mathcal{W}_1 f)\big), \tag{10}$$

with C the convolution of either building block, follows the SFNO block of Bonev et al. (2023) in summing a harmonic-space convolution and a pointwise branch, but replaces the MLP that follows the sum by $\mathcal{W}_2$ and has no identity residual connection. The operators $\mathcal{W}_1$, in parallel with the convolution, and $\mathcal{W}_2$, after the sum, are harmonic 1×1 convolutions. Each transforms its input to harmonic space, truncates or pads the coefficients to the output bandlimit, mixes the channels with one matrix W and the spins with one matrix M, both constant in ℓ, and transforms back; on $\mathrm{SO}(3)$ it reads the columns $n = -s$, mixes them and lifts again. Between equal spins this is exactly the pointwise 1×1 convolution of the bypass of the FNO. Between different spins it keeps the harmonic coefficients and changes their spin, which in pixel space is a normalised power of the spin raising or lowering operator (McEwen, 2026); since this is not local, it is computed in harmonic space. The role of the harmonic 1×1 convolutions is to enrich the channel and spin mixing of the block.

4.4. Networks

We compare two constructions, each with Conv blocks, a convolution and its non-linearity, or with S^{2}FNO blocks, and name the four models by block and construction, Conv-S^2, Conv-$\mathrm{SO}(3)$, S^{2}FNO-S^2 and S^{2}FNO-$\mathrm{SO}(3)$. The spherical construction never lifts explicitly, whereas the rotation-group construction lifts after a first spin spherical convolution, so that its filters act on the rotation group. We construct classifier and dense prediction networks based on these ar-

chitectural building blocks. The classifiers stack six blocks and pool rotation-invariant features, and we construct UNets with encoders and decoders built from the same blocks, with skip connections that concatenate features at equal bandlimit and spin set (Fig. 11). The blocks include a normalisation, a group normalisation in the UNets and a batch normalisation in the classifiers, and the spin UNets for wind and weak lensing are Conv-SO(3) networks (Appendix D).

5. Efficient representation, computation and parameterisation

Mackey functions and the lift give the conceptual description of the layers and their proofs (McEwen, 2026), whereas the harmonic representation provides exact and efficient computation. We form the lift explicitly only where it provides expressivity.

5.1. Representation and spin mixing

In harmonic space the lift need not be formed. The lift of a spin-s field occupies a single Wigner column and so is held in the $\mathcal{O}(L^2)$ spin spherical harmonic coefficients of the field rather than the $\mathcal{O}(L^3)$ of a general function on $\mathrm{SO}(3)$, and a tuple of fields with spins in $\mathcal{S}$ in $\mathcal{O}(|\mathcal{S}|L^2)$ coefficients. Only an explicit lift, its synthesis at $2N-1$ samples in γ for a pointwise non-linearity, requires $\mathcal{O}(NL^2)$.

Linear spin mixing needs no lift. Every bounded equivariant complex-linear map between two spins is a spin spherical convolution, and a tuple of spin fields, one for each spin $|s| < N$, is exactly a function on $\mathrm{SO}(3)$ of azimuthal bandlimit N (McEwen et al., 2015a). A group convolution mapping between such functions is therefore an array of $(2N-1)^2$ spin spherical convolutions, one for each pair of spins, so that, as classes of linear maps, the spherical construction realises every convolution on $\mathrm{SO}(3)$ at that azimuthal bandlimit without forming the lift. The convolution on the rotation group in our networks uses only the diagonal of this array, so that its spins are coupled by the non-linearity alone, and the rotation-group construction mixes spins linearly in its first layer. A pointwise complex-linear map cannot mix spins equivariantly, since under a frame rotation (1) fields of different spins acquire different phases, so that an equivariant 1×1 convolution on the samples mixes channels only within a spin and the pointwise bypass of the FNO, extended to spin fields, must mix spins in harmonic space.

General non-linear spin coupling is what the explicit lift achieves. On the sphere it needs maps whose outputs have a definite spin, such as the modulus $z \mapsto |z|$, which takes every spin to spin 0 and is the phase collapse of spherical scattering networks (McEwen et al., 2022; Mousset et al., 2024), or products of fields, whose spins add. The modulus non-linearity (9) keeps the phase, and hence the spin, of each field, so that the non-linearities of our spherical construction couple no spins. On $\mathrm{SO}(3)$ every pointwise map is equivariant, and one non-linearity after the lift couples all spins at once. In the rotation-group construction spins are thus coupled non-linearly through the lift.

Representing a vector or spin field as a spin field, rather than as scalar channels, also matters under reflections of the sphere. Networks on scalar fields built from isotropic multipliers and pointwise maps, without positional inputs, are $\mathrm{O}(3)$-equivariant (Weiler et al., 2021) and cannot represent maps that anticommute with the antipodal map P, such as the advection of a vorticity treated as a scalar. Spin fields escape this restriction, since P reverses the orientation of the tangent frame and so acts on them through conjugation as well as pull-back (Appendix C.4). A multiplier between spins commutes with P if and only if its coefficients are real, so that with real profiles our classifiers would be exactly invariant under reflections, and they therefore use complex profiles. Real profiles, whose multipliers commute with P, could instead be useful for fields with antipodal symmetry, such as those of diffusion magnetic resonance imaging (Descoteaux et al., 2007).

5.2. Computation

We leverage sampling theorems and fast harmonic transforms on S^2 and $\mathrm{SO}(3)$ (McEwen & Wiaux, 2011; McEwen et al., 2015a; Price & McEwen, 2024). Since the sampling theorems make the transforms of bandlimited spin fields and functions on $\mathrm{SO}(3)$ exact, every linear layer acting on bandlimited input is computed exactly, consistent with the machine-precision equivariance errors of Sec. 6. A spin spherical convolution costs $\mathcal{O}(L^3)$ for each field, and so $\mathcal{O}(|\mathcal{S}|L^3)$ for a tuple, against $\mathcal{O}(NL^3)$ for a convolution on $\mathrm{SO}(3)$ at azimuthal bandlimit N, whose output fills all $2N-1$ spins, and $\mathcal{O}(L^4)$ for a full convolution on $\mathrm{SO}(3)$, with $N = L$. These are upper bounds for the algorithms in harmonic space (McEwen, 2026). Against convolutions that fill every spin, the saving of the spherical construction lies in a spin set smaller than $2N-1$ and in storing no samples on the rotation group, which the rotation-group construction samples only for its non-linearity. This non-linearity, which couples all spins, costs $\mathcal{O}(NL^3)$ with full Wigner transforms and $\mathcal{O}(|\mathcal{S}|(L^3 + NL^2))$ in our implementation, which computes only the Wigner columns that are read.

5.3. Parameterisation

We parameterise every linear layer in harmonic space. Our multipliers are interpolated in ℓ from ten knots placed at fixed fractions of the bandlimit, so that, as with the anchor points of Cobb et al. (2021), filters at high resolution need few parameters. The knots reduce the $\mathcal{O}(L)$ multipliers of a

spin spherical convolution for each pair of spins and channels to ten, against the $\mathcal{O}(NL)$ values of a kernel column of azimuthal bandlimit N and the $\mathcal{O}(L^3)$ of a full kernel on $\mathrm{SO}(3)$ (McEwen, 2026). In the UNets the profiles are real, with a channel matrix for each output spin, and in the classifiers complex (Appendix D). Each harmonic 1×1 convolution has one channel matrix and one spin matrix, and the activations learn an offset for each channel and, in the UNets, a gain as well. Parameters therefore attach, through the knots, to fractions of the bandlimit rather than to samples, so that their number depends neither on the bandlimit nor on the sampling grid, and a trained layer applies on any grid that supports exact transforms at its bandlimit. Our layers are thus neural operators in the sense of Kovachki et al. (2023).

Our layers are built on the `s2fft` code (Price & McEwen, 2024) and will be released with the `s2ai` code.

6. Equivariance

The convolutional layers of S^{2}FNOs are exactly equivariant to rotations of spin fields, which we confirm numerically, and we discuss the equivariance errors introduced by the other layers, such as the activations and the normalisation. Equivariance errors have been measured before for spherical CNNs on scalar signals (Cohen et al., 2018; Esteves et al., 2018; Cobb et al., 2021; Ocampo et al., 2023), but not for vector and spin fields on the sphere under general rotations.

6.1. Sources of equivariance error

As for spherical CNNs on scalar signals, convolutional layers are computed exactly in harmonic space. Any error therefore enters only through the non-linear pointwise activations and the normalisation. The output $\zeta(F)$ of an activation is not bandlimited, so that analysing its samples on the same grid aliases its content above the bandlimit, and on $\mathrm{SO}(3)$ beyond the azimuthal bandlimit. Oversampling reduces this aliasing (Cobb et al., 2021). The group normalisation of the UNets aliases for every spin, because it rescales the modulus by $|x - \mu|/|x|$, up to constants for each group (Appendix D.1), and on the sphere it adds, for non-zero spin, a structural error, since the mean that it subtracts is not rotation invariant, of the kind that pointwise spin mixing would cause (Sec. 5.1). A normalisation with statistics computed in harmonic space and no mean subtracted for $s \neq 0$ would be exactly equivariant, which we leave to future work.

6.2. Equivariance tests

We measure the mean relative equivariance error (MREE), a quadrature-weighted relative L^2 error, over 100 random bandlimited signals and 100 random rotations at $L = 16$ in double precision (Appendix E). Linear layers attain errors of order 10^{-15}, the activations alias with errors of order 0.1 to 1, apart from the complex ReLU, measured at initialisation where it is the identity up to ϵ, and the group normalisation gives errors of order 10^{-2} on the sphere and 10^{-3} on the rotation group (Table 1). With the activations and the normalisation removed, whole S^{2}FNO blocks and UNets, with spin and channel mixing, are equivariant to the round-off floor of the transforms, of order 10^{-14} at the larger bandlimit $L = 32$ (Tables 1 and 7), which confirms that the errors of our blocks and networks arise only from their activations and normalisation.

Table 1. Equivariance errors in double precision on the sphere (S^2) and the rotation group ($\mathrm{SO}(3)$), measured by the mean relative equivariance error (MREE) between rotating the input and rotating the output (Appendix E). Conv-S^2 and Conv-$\mathrm{SO}(3)$ are the spin spherical convolution $\mathcal{C}_{s\to s'}$ and the directional convolution $\mathcal{C}_{s\to\mathrm{SO}(3)}$, ReLU and Swish are applied directly to complex data, whereas the complex activations act on the modulus with the phase kept (9), and the complex ReLU is measured at initialisation, where it is the identity up to ϵ. The S^{2}FNO blocks and the UNets of the spin spherical MNIST architecture have no activation or normalisation, are at $L = 32$, with $N = 8$ on $\mathrm{SO}(3)$, and reach the round-off floor of the transforms. All other rows are at $L = 16$, with $N = 5$ on $\mathrm{SO}(3)$.

		Equivariance error	
Operation	Spin	S^2	$\mathrm{SO}(3)$
Conv-S^2	0	5.9×10^{-15}	–
Conv-S^2	1	5.2×10^{-15}	–
Conv-$\mathrm{SO}(3)$	2	–	6.0×10^{-15}
Harmonic 1×1 conv.		5.4×10^{-15}	5.4×10^{-15}
S^{2}FNO block	$\{0, \pm 1\}$	2.1×10^{-14}	2.4×10^{-14}
S^{2}FNO UNet	$0 \to 1$	2.0×10^{-14}	1.8×10^{-14}
S^{2}FNO UNet	$1 \to 0$	2.0×10^{-14}	1.6×10^{-14}
ReLU (complex data)		0.453	0.592
Swish (complex data)		0.581	1.047
Complex ReLU, init.		–	2.2×10^{-9}
Complex Swish		–	0.0885
Group norm.		0.013	0.0033

7. Spin benchmarks and experiments

In this section we first introduce three new benchmarks for spin fields, spin spherical MNIST, global wind forecasting and weak gravitational lensing, and then compare our networks on each of them with scalar counterparts and published results (Appendix G).

7.1. New benchmarks

Public benchmarks for learning on spin fields are lacking. The benchmarks of spherical CNNs, such as spherical MNIST (Cohen et al., 2018), are scalar, and the vector-field MNIST of Esteves et al. (2020), built from the Sobel gradients of the images and also used by Xu et al. (2022), is

limited to spin 1. We introduce three new benchmarks, a contribution of this article in their own right, whose inputs and targets have a definite spin, of any spin for spin spherical MNIST, and which span classification and regression on synthetic, atmospheric and cosmological fields. Their generation code will be released with `s2ai`, and Appendix F describes them in full.

Spin spherical MNIST. MNIST digits (LeCun et al., 1998) are projected onto the sphere (Cohen et al., 2018) at $L = 32$. The Sobel-gradient construction of Esteves et al. (2020), the only one available so far, gives a spin-1 field only. We instead synthesise spin fields of any spin from the same harmonic coefficients in the spin-s basis, so that rotations act exactly on the bandlimited fields, and use the harmonic spin-1 and spin-2 fields alongside the Sobel spin-1 field. We consider the classification of the digit from one of these fields, and dense prediction, which maps the spin-0 image to its spin-1 field or back and so changes spin. Both follow the NR/R protocol of training on unrotated and testing on randomly rotated data (Cohen et al., 2018; Cobb et al., 2021; Ocampo et al., 2023), which tests the invariance of a classifier and the equivariance of a dense prediction to rotations not seen in training (Appendix F.1).

Global wind forecasting. The task is to predict the horizontal wind at 50 hPa at 18:00 UTC on a given day from the wind at 00:00, 06:00 and 12:00 UTC of the same day, with the wind represented by the spin-1 field $u + \mathrm{i}v$ of its zonal and meridional components at $L = 128$. The data are the ERA5 reanalysis (Hersbach et al., 2020) of WeatherBench 2 (Rasp et al., 2024) from 1980 to 2022, from which eight days drawn at random are held out in every year, so that the benchmark tests interpolation within the same years rather than generalisation to later years (Appendix F.2).

Weak gravitational lensing. The task is to reconstruct the spin-0 convergence κ on the full sky from the noisy spin-2 shear $\gamma = \gamma_1 + \mathrm{i}\gamma_2$. Matter along the line of sight, most of it dark matter, distorts the images of distant galaxies by the shear, which on the full sky is a spin $0 \to 2$ multiplier of the convergence, a projection of the matter density (Bartelmann & Schneider, 2001; Kilbinger, 2015). The spherical Kaiser–Squires estimator inverts this multiplier as a spin $2 \to 0$ spin spherical convolution (Kaiser & Squires, 1993; Wallis et al., 2022), but practical considerations make the recovery of κ ill-posed and call for priors learned from simulations, which on the full sky have so far taken the two shear components as scalar channels (Wang & Yu, 2026). We simulate $10{,}000$ pairs of convergence and noisy shear maps for training and validation and 600 for evaluation at $L = 256$ with GLASS (Tessore et al., 2023) and the shape noise of a Stage IV survey (Laureijs et al., 2011) (Appendix F.3).

Table 2. Spin spherical MNIST classification: NR/R accuracy (%) of our spherical (S^2) and rotation-group ($\mathrm{SO}(3)$) networks (best accuracy and fewest parameters in bold).

		Sobel	Harmonic	
Model	Params	spin 1	spin 1	spin 2
Conv-S^2	**50k**	99.15	99.19	99.17
S^{2}FNO-S^2	182k	98.82	99.16	**99.34**
Conv-$\mathrm{SO}(3)$	134k	99.01	99.10	99.11
S^{2}FNO-$\mathrm{SO}(3)$	136k	99.17	**99.21**	99.28
Esteves et al.	58k	98.2	–	–
Ballerin et al., sCNN	2.2M	99.20	–	–
Ballerin et al., GsCNN	2.0M	**99.32**	–	–

7.2. Spin spherical MNIST

We train four models as classifiers and as UNets for dense prediction (Appendix D). The classifiers are selected on a non-rotated validation split and the prediction UNets on the rotated held-out set, which is also their evaluation set. The Conv-S^2 classifier on the harmonic spin-1 field is trained with three seeds, whose mean accuracy we report, and the architecture of the classifiers was developed with reference to their test accuracies (Appendix G.1).

The four classifiers generalise to rotations not seen in training (Table 2). On the Sobel field our best classifier comes within 0.15% of the published accuracy of the general spherical CNN (GsCNN) of Ballerin et al. (2025), and on the harmonic spin-1 and spin-2 fields, for which no other results are published, all four exceed 99%. The classifiers use complex harmonic filters, since with real profiles they would be exactly invariant under reflections of the sphere (Sec. 5.1).

In dense prediction, a spin-changing task, each rotation-group UNet has fewer parameters and a lower error than its spherical counterpart in all four mappings (Table 3 and Fig. 3). Moreover, the Conv-$\mathrm{SO}(3)$ UNet gives the lowest errors of the models compared from images to fields, and the S^{2}FNO-$\mathrm{SO}(3)$ UNet from fields to images. The published values, for another construction and given for context only, lie above those of the Conv-$\mathrm{SO}(3)$ UNet, which has about eight times fewer parameters than that of Esteves et al. (2020).

7.3. Global wind forecasting

We compare a spin UNet, a Conv-$\mathrm{SO}(3)$ network that takes the spin field as input, with a scalar spherical UNet that takes its real and imaginary parts as scalar channels, and with persistence, which forecasts the 12:00 UTC field for 18:00 UTC (Appendix D.3). Checkpoints are selected on the held-out days, which are also the evaluation set. The spin UNet reduces the MSE by 44.9% with 34.9% fewer parameters and improves every other metric (Table 4 and

Table 3. Spin-changing prediction on spin spherical MNIST: NR/R MSE from the spin-0 image to its spin-1 field ($0 \to 1$) and back ($1 \to 0$), with the best model and the fewest parameters in bold.

		MSE $\times 10^3 \downarrow$			
		Sobel field		Harmonic field	
Model	Params	$0 \to 1$	$1 \to 0$	$0 \to 1$	$1 \to 0$
Conv-S^2	23k	3.681	3.212	5.580	3.229
S^{2}FNO-S^2	29k	4.253	4.544	3.944	4.468
Conv-SO(3)	**14k**	**3.016**	3.144	**2.876**	3.101
S^{2}FNO-SO(3)	20k	3.866	**2.948**	3.063	**3.005**
Esteves et al.	≈112k	4.3	4.0	–	–
Xu et al.	–	3.8	–	–	–

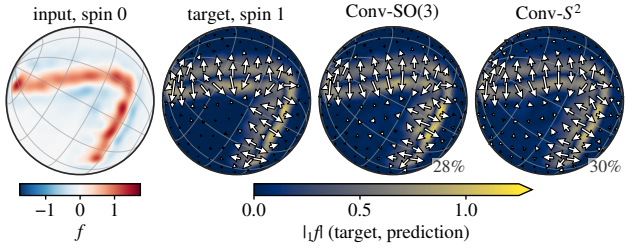


Figure 3. Prediction of the spin-1 Sobel field from the spin-0 image of a rotated test digit, chosen in advance, by the Conv-SO(3) and Conv-S^2 UNets, with their relative errors. Both predictions follow the rotated frame of the target (more in Fig. 13).

Fig. 4). It is the only model that improves on persistence in every metric, whereas the scalar UNet is worse than persistence in every metric except the correlation. These results support the practical value of the spin representation.

7.4. Weak gravitational lensing

Table 4 and Fig. 5 compare a spin UNet, which lifts the spin-2 shear and returns the spin-0 convergence, with a scalar UNet of DISCO convolutions (Ocampo et al., 2023) and residual blocks (He et al., 2016) that takes γ_1 and γ_2 as scalar channels, both five-level UNets (Ronneberger et al., 2015) (Appendix D.3). The spin UNet reduces the MSE by 37.9% with 13.4% fewer parameters and improves every other metric.

8. Conclusions

We have developed S^{2}FNOs (spin spherical Fourier neural operators), which extend spherical Fourier neural operators to fields of any spin, with rotational equivariance exact for every convolutional layer and holding in the continuum for every activation. Described by Mackey functions but computed in harmonic space, S^{2}FNOs are efficient in representation, computation and parameters, with far fewer parameters than networks that convolve spin fields on the full rotation group. On spin fields their convolutional layers

Table 4. Wind forecasting at 50 hPa on the held-out days and mass mapping from the noisy spin-2 shear of weak lensing at $L = 256$, averaged over 600 maps (best results and fewest parameters in bold; all metrics in Tables 8 and 9). All models are spherical, and the scalar UNet takes the spin field as two scalar channels. Persistence forecasts the 12:00 UTC field for 18:00 UTC.

Model	Params	MSE ↓	SNR ↑ (dB)	PSNR ↑ (dB)	PCC ↑
Global wind forecasting (MSE in m^2 s^{-2})					
Spin UNet	**48k**	**9.52**	**14.02**	**25.06**	**0.979**
Scalar UNet	74k	17.27	11.43	22.64	0.969
Persistence	–	14.55	12.17	23.19	0.962
Weak gravitational lensing (MSE $\times 10^6$)					
Spin UNet	**81k**	**1.28**	**7.53**	**23.47**	**0.907**
Scalar UNet	93k	2.06	5.45	21.40	0.846

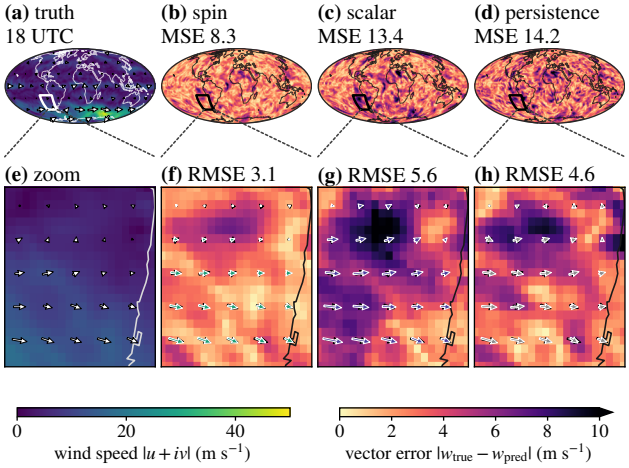


Figure 4. Wind forecast for one held-out day: (a) the target speed, (b–d) the vector errors of the spin and scalar UNets and persistence, and (e–h) the same in the outlined region, with true (white) and predicted (coloured) vectors. The scalar UNet leaves larger, coherent patches of error (more in Fig. 14).

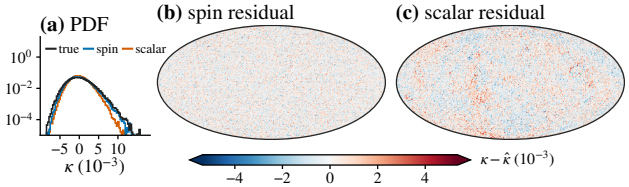


Figure 5. Mass mapping for one evaluation map at $L = 256$: (a) PDFs of the true and reconstructed convergence, and residuals $\kappa - \hat{\kappa}$ of (b) the spin and (c) the scalar UNet. The scalar UNet misses the tails and leaves large-scale structure in its residual, whereas the spin residual is noise-like (maps in Fig. 15).

are equivariant to 10^{-15}, and whole blocks and UNets without activations or normalisation to round-off. On three new benchmarks our networks achieve state-of-the-art results.

Impact Statement

This paper presents work whose goal is to advance the field of machine learning, and in particular the modelling of geophysical and astrophysical fields on the sphere. Improved treatment of spin fields such as wind and polarisation may benefit weather, climate and cosmological applications. There are many potential societal consequences of our work, none of which we feel must be specifically highlighted here.

Declaration on the use of large language models

LLMs were used in the writing of this manuscript, but the theory was worked out manually. We take full responsibility for the accuracy of the material presented.

References

Bacon, D. J., Goldberg, D. M., Rowe, B. T. P., and Taylor, A. N. Weak gravitational flexion. *Monthly Notices of the Royal Astronomical Society*, 365(2):414–428, 2006. doi: 10.1111/j.1365-2966.2005.09624.x. arXiv:astro-ph/0504478.

Ballerin, F., Blaser, N., and Grong, E. SO(3)-equivariant neural networks for learning from scalar and vector fields on spheres. arXiv:2503.09456, 2025.

Bartelmann, M. and Schneider, P. Weak gravitational lensing. *Physics Reports*, 340(4–5):291–472, 2001. doi: 10.1016/S0370-1573(00)00082-X.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., and Tian, Q. Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, 619(7970):533–538, 2023. doi: 10.1038/s41586-023-06185-3. arXiv:2211.02556.

Bonev, B., Kurth, T., Hundt, C., Pathak, J., Baust, M., Kashinath, K., and Anandkumar, A. Spherical Fourier neural operators: Learning stable dynamics on the sphere. In *International Conference on Machine Learning (ICML)*, 2023. arXiv:2306.03838.

Bonev, B., Kurth, T., Mahesh, A., Bisson, M., Kossaifi, J., Kashinath, K., Anandkumar, A., Collins, W. D., Pritchard, M. S., and Keller, A. FourCastNet 3: A geometric approach to probabilistic machine-learning weather forecasting at scale. arXiv:2507.12144, 2025.

Castro, P. G., Heavens, A. F., and Kitching, T. D. Weak lensing analysis in three dimensions. *Physical Review D*, 72(2):023516, 2005. doi: 10.1103/PhysRevD.72.023516. arXiv:astro-ph/0503479.

Cobb, O. J., Wallis, C. G. R., Mavor-Parker, A. N., Marignier, A., Price, M. A., d'Avezac, M., and McEwen, J. D. Efficient generalized spherical CNNs. In *International Conference on Learning Representations (ICLR)*, 2021. arXiv:2010.11661.

Cohen, T. S., Geiger, M., Köhler, J., and Welling, M. Spherical CNNs. In *International Conference on Learning Representations (ICLR)*, 2018. arXiv:1801.10130.

Cohen, T. S., Geiger, M., and Weiler, M. A general theory of equivariant CNNs on homogeneous spaces. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 9142–9153, 2019. arXiv:1811.02017.

Descoteaux, M., Angelino, E., Fitzgibbons, S., and Deriche, R. Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine*, 58(3):497–510, 2007. doi: 10.1002/mrm.21277.

Driscoll, J. R. and Healy, Jr., D. M. Computing Fourier transforms and convolutions on the 2-sphere. *Advances in Applied Mathematics*, 15(2):202–250, 1994. doi: 10.1006/aama.1994.1008.

Esteves, C., Allen-Blanchette, C., Makadia, A., and Daniilidis, K. Learning SO(3) equivariant representations with spherical CNNs. In *European Conference on Computer Vision (ECCV)*, pp. 54–70, 2018. arXiv:1711.06721.

Esteves, C., Makadia, A., and Daniilidis, K. Spin-weighted spherical CNNs. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2020. arXiv:2006.10731.

Esteves, C., Slotine, J.-J., and Makadia, A. Scaling spherical CNNs. In *International Conference on Machine Learning (ICML)*, 2023. arXiv:2306.05420.

Goldberg, J. N., Macfarlane, A. J., Newman, E. T., Rohrlich, F., and Sudarshan, E. C. G. Spin-s spherical harmonics and $\eth$. *Journal of Mathematical Physics*, 8(11):2155–2161, 1967. doi: 10.1063/1.1705135.

Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., and Bartelmann, M. HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622(2):759–771, 2005. doi: 10.1086/427976.

He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016. doi: 10.1109/CVPR.2016.90.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

Kaiser, N. and Squires, G. Mapping the dark matter with weak gravitational lensing. *The Astrophysical Journal*, 404:441–450, 1993. doi: 10.1086/172297.

Kilbinger, M. Cosmology with cosmic shear observations: a review. *Reports on Progress in Physics*, 78(8):086901, 2015. doi: 10.1088/0034-4885/78/8/086901.

Kingma, D. P. and Ba, J. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

Kondor, R. and Trivedi, S. On the generalization of equivariance and convolution in neural networks to the action of compact groups. In *International Conference on Machine Learning (ICML)*, pp. 2747–2755, 2018. arXiv:1802.03690.

Kondor, R., Lin, Z., and Trivedi, S. Clebsch–Gordan nets: A fully Fourier space spherical convolutional neural network. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018. arXiv:1806.09231.

Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A., and Anandkumar, A. Neural operator: Learning maps between function spaces with applications to PDEs. *Journal of Machine Learning Research*, 24(89):1–97, 2023. arXiv:2108.08481.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Vinyals, O., Stott, J., Pritzel, A., Mohamed, S., and Battaglia, P. Learning skillful medium-range global weather forecasting. *Science*, 382(6677):1416–1421, 2023. doi: 10.1126/science.adi2336. arXiv:2212.12794.

Laureijs, R. et al. Euclid definition study report, 2011. arXiv:1110.3193.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

Lewis, A., Challinor, A., and Lasenby, A. Efficient computation of cosmic microwave background anisotropies in closed Friedmann-Robertson-Walker models. *The Astrophysical Journal*, 538(2):473–476, 2000. doi: 10.1086/309179.

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., and Anandkumar, A. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations (ICLR)*, 2021. arXiv:2010.08895.

Liang, C., Lin, Y., Kryvenko, A., Yu, W., Li, C., Xie, J., Qian, X., and Ji, S. Spin-weighted spherical harmonics enable complete and scalable E(3)-equivariant networks. arXiv:2607.01408, 2026.

Liu-Schiaffini, M., Berner, J., Bonev, B., Kurth, T., Azizzadenesheli, K., and Anandkumar, A. Neural operators with localized integral and differential kernels. In *International Conference on Machine Learning (ICML)*, 2024. arXiv:2402.16845.

McEwen, J. D. Unifying spherical convolutions through lifting. arXiv:2610.13892, 2026.

McEwen, J. D. and Wiaux, Y. A novel sampling theorem on the sphere. *IEEE Transactions on Signal Processing*, 59(12):5876–5887, 2011.

McEwen, J. D., Büttner, M., Leistedt, B., Peiris, H. V., and Wiaux, Y. A novel sampling theorem on the rotation group. *IEEE Signal Processing Letters*, 22(12):2425–2429, 2015a. doi: 10.1109/LSP.2015.2490676. arXiv:1508.03101.

McEwen, J. D., Leistedt, B., Büttner, M., Peiris, H. V., and Wiaux, Y. Directional spin wavelets on the sphere. arXiv:1509.06749, 2015b.

McEwen, J. D., Wallis, C. G. R., and Mavor-Parker, A. N. Scattering networks on the sphere for scalable and rotationally equivariant spherical CNNs. In *International Conference on Learning Representations (ICLR)*, 2022. arXiv:2102.02828.

Mousset, L., Allys, E., Price, M. A., Aumont, J., Delouis, J.-M., Montier, L., and McEwen, J. D. Generative models of astrophysical fields with scattering transforms on the sphere. *Astronomy & Astrophysics*, 691:A269, 2024. doi: 10.1051/0004-6361/202451396. arXiv:2407.07007.

Newman, E. T. and Penrose, R. Note on the Bondi–Metzner–Sachs group. *Journal of Mathematical Physics*, 7(5):863–870, 1966. doi: 10.1063/1.1931221.

Ocampo, J., Price, M. A., and McEwen, J. D. Scalable and equivariant spherical CNNs by discrete-continuous (DISCO) convolutions. In *International Conference on Learning Representations (ICLR)*, 2023. arXiv:2209.13603.

Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., Kurth, T., Hall, D., Li, Z., Azizzadenesheli, K., Hassanzadeh, P., Kashinath, K., and Anandkumar, A. FourCastNet: A global data-driven high-resolution weather model using adaptive Fourier neural operators. arXiv:2202.11214, 2022.

Planck Collaboration, Aghanim, N., et al. Planck 2018 results. VI. cosmological parameters. *Astronomy & Astrophysics*, 641:A6, 2020. doi: 10.1051/0004-6361/201833910.

Price, M. A. and McEwen, J. D. Differentiable and accelerated spherical harmonic and Wigner transforms. *Journal of Computational Physics*, 510:113109, 2024. arXiv:2311.14670.

Ramachandran, P., Zoph, B., and Le, Q. V. Searching for activation functions, 2017. arXiv:1710.05941.

Rasp, S., Hoyer, S., Merose, A., Langmore, I., Battaglia, P., Russell, T., Sanchez-Gonzalez, A., Yang, V., Carver, R., Agrawal, S., et al. WeatherBench 2: A benchmark for the next generation of data-driven global weather models. *Journal of Advances in Modeling Earth Systems*, 16(6): e2023MS004019, 2024.

Ronneberger, O., Fischer, P., and Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, volume 9351 of *Lecture Notes in Computer Science*, pp. 234–241, 2015. doi: 10.1007/978-3-319-24574-4_28.

Tessore, N., Loureiro, A., Joachimi, B., von Wietersheim-Kramsta, M., and Jeffrey, N. GLASS: Generator for large scale structure. *The Open Journal of Astrophysics*, 6:11, 2023. doi: 10.21105/astro.2302.01942.

Varshalovich, D. A., Moskalev, A. N., and Khersonskii, V. K. *Quantum Theory of Angular Momentum*. World Scientific, Singapore, 1988. doi: 10.1142/0270.

Wallis, C. G. R., Price, M. A., McEwen, J. D., Kitching, T. D., Leistedt, B., and Plouviez, A. Mapping dark matter on the celestial sphere with weak gravitational lensing. *Monthly Notices of the Royal Astronomical Society*, 509(3):4480–4497, 2022. doi: 10.1093/mnras/stab3235. arXiv:1703.09233.

Wang, Y. and Yu, Y. Advancing weak lensing mass mapping with a mask-aware HEALPix transformer. *Physical Review D*, 113(4):043553, 2026. doi: 10.1103/kc9z-jllp. arXiv:2603.25471.

Weiler, M. and Cesa, G. General E(2)-equivariant steerable CNNs. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019. arXiv:1911.08251.

Weiler, M., Geiger, M., Welling, M., Boomsma, W., and Cohen, T. S. 3D steerable CNNs: Learning rotationally equivariant features in volumetric data. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2018. arXiv:1807.02547.

Weiler, M., Forré, P., Verlinde, E., and Welling, M. Coordinate independent convolutional networks – isometry and gauge equivariant convolutions on Riemannian manifolds. arXiv:2106.06020, 2021.

Wiaux, Y., Jacques, L., Vielva, P., and Vandergheynst, P. Fast directional correlation on the sphere with steerable filters. *The Astrophysical Journal*, 652(1):820–832, 2006. doi: 10.1086/507692. arXiv:astro-ph/0508516.

Worrall, D. E., Garbin, S. J., Turmukhambetov, D., and Brostow, G. J. Harmonic networks: Deep translation and rotation equivariance. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7168–7177, 2017. doi: 10.1109/CVPR.2017.758. arXiv:1612.04642.

Xie, Y., Daigavane, A., Kotak, M., and Smidt, T. The price of freedom: Exploring expressivity and runtime tradeoffs in equivariant tensor products. In *International Conference on Machine Learning (ICML)*, 2025. arXiv:2506.13523.

Xie, Y., Daigavane, A., Kotak, M., and Smidt, T. Asymptotically fast Clebsch–Gordan tensor products with vector spherical harmonics. arXiv:2602.21466, 2026.

Xu, Y., Lei, J., Dobriban, E., and Daniilidis, K. Unified Fourier-based kernel and nonlinearity design for equivariant networks on homogeneous spaces. In *International Conference on Machine Learning (ICML)*, 2022. arXiv:2206.08362.

Zaldarriaga, M. and Seljak, U. All-sky analysis of polarization in the microwave background. *Physical Review D*, 55 (4):1830–1840, 1997. doi: 10.1103/PhysRevD.55.1830. arXiv:astro-ph/9609170.

A. Background: harmonic analysis, lifting and lifted convolutions

In this appendix we collect the conventions and results of McEwen (2026) on which our layers rest, each cited in its heading by its number there, where it is proved. Definition 1, Lemma 2, Corollary 3, Theorem 4 and Proposition 6 are stated in the same words, and Theorem 5 and Corollary 7 in shortened form. Only the equivariance of the lifted non-linearity is proved in this article (Appendix C).

A.1. Notation and conventions

We adopt the conventions of McEwen (2026, Sec. II and App. A). Points of the sphere S^2 are written $\omega = (\theta, \varphi)$, with colatitude $\theta \in [0, \pi]$ and longitude $\varphi \in [0, 2\pi)$, and $\eta = (0, 0)$ denotes the north pole. We adopt the measure $\mathrm{d}\mu(\omega) = \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi$ and the inner product $\langle f, g \rangle = \int_{S^2} f \, g^* \, \mathrm{d}\mu$. Rotations are parameterised by zyz Euler angles $\rho = (\alpha, \beta, \gamma)$ in the active convention, $R(\alpha, \beta, \gamma) = R_z(\alpha) R_y(\beta) R_z(\gamma)$, with the unnormalised Haar measure $\mathrm{d}\mu(\rho) = \sin\beta \, \mathrm{d}\alpha \, \mathrm{d}\beta \, \mathrm{d}\gamma$ of total mass $8\pi^2$. The map $p(\rho) = R_\rho \eta$ sends $R(\alpha, \beta, \gamma)$ to the point (β, α), so that $S^2 \cong \mathrm{SO}(3)/\mathrm{SO}(2)$, with $\mathrm{SO}(2) = \{R_z(\gamma)\}$ the stabiliser of the north pole.

The spherical harmonics $Y_{\ell m}$ carry the Condon–Shortley phase, $Y^*_{\ell m} = (-1)^m Y_{\ell,-m}$, and the spin raising and lowering operators $\eth$ and $\bar{\eth}$ (Newman & Penrose, 1966; Goldberg et al., 1967) generate the spin spherical harmonics from them by

$$\eth \, {}_sY_{\ell m} = \sqrt{(\ell - s)(\ell + s + 1)} \; {}_{s+1}Y_{\ell m}, \qquad \bar{\eth} \, {}_sY_{\ell m} = -\sqrt{(\ell + s)(\ell - s + 1)} \; {}_{s-1}Y_{\ell m}, \qquad {}_0Y_{\ell m} = Y_{\ell m}, \tag{11}$$

so that ${}_sY_{\ell m}$ exists only for $\ell \geq |s|$ and, for fixed s, these form an orthonormal basis of the space $L^2_s(S^2)$ of square integrable spin-s functions on the sphere.

The Wigner D-functions, the matrix elements of the irreducible unitary representations of $\mathrm{SO}(3)$ (Varshalovich et al., 1988), may be written

$$D^\ell_{mn}(\alpha, \beta, \gamma) = e^{-\mathrm{i}m\alpha} \, d^\ell_{mn}(\beta) \, e^{-\mathrm{i}n\gamma}, \tag{12}$$

where the Wigner d-functions d^ℓ_{mn} are real. They satisfy $D^\ell(\rho\rho') = D^\ell(\rho) D^\ell(\rho')$ and $D^\ell(\rho^{-1}) = D^\ell(\rho)^\dagger$ and the orthogonality relation $\langle D^\ell_{mn}, D^{\ell'}_{m'n'} \rangle = \frac{8\pi^2}{2\ell+1} \delta_{\ell\ell'} \delta_{mm'} \delta_{nn'}$, and by the Peter–Weyl theorem every $F \in L^2(\mathrm{SO}(3))$ admits the Wigner expansion

$$F(\rho) = \sum_{\ell=0}^{\infty} \frac{2\ell + 1}{8\pi^2} \sum_{m,n=-\ell}^{\ell} F^\ell_{mn} \, D^{\ell *}_{mn}(\rho), \qquad F^\ell_{mn} = \int_{\mathrm{SO}(3)} F(\rho) \, D^\ell_{mn}(\rho) \, \mathrm{d}\mu(\rho). \tag{13}$$

The identity tying the sphere to the rotation group is

$$e^{-\mathrm{i}s\gamma} \; {}_sY_{\ell m}(\beta, \alpha) = (-1)^s \sqrt{\frac{2\ell + 1}{4\pi}} \; D^{\ell *}_{m,-s}(\alpha, \beta, \gamma). \tag{14}$$

Since the lift of a spin-s function thus occupies the Wigner column $n = -s$ and intertwines the spin rotation with left translation (Lemma 2(c) and (e)), and left translation acts on the first Wigner index alone, the harmonic coefficients of a function of any spin rotate by the same matrices under the spin rotation operator $\mathcal{R}_\rho$ of McEwen et al. (2015b),

$$(\mathcal{R}_\rho \; {}_sf)_{\ell m} = \sum\nolimits_n D^\ell_{mn}(\rho) \; {}_sf_{\ell n}. \tag{15}$$

Figs. 6 and 7 show spin at a point and over the sphere.

A.2. Spin fields, Mackey functions and the lift

The rotation group is a principal $\mathrm{SO}(2)$-bundle over the sphere with projection p, whose fibre over a point is the circle of unit tangent vectors there (Fig. 8). We adopt the section

$$s(\theta, \varphi) = R(\varphi, \theta, 0), \qquad \textit{i.e.} \qquad (\beta, \alpha) \mapsto (\alpha, \beta, 0) \tag{16}$$

for $0 < \theta < \pi$, with $s(\eta) = e$ and $s(-\eta) = R_y(\pi)$ at the poles. It selects at each point the frame aligned with the meridian direction e_θ and is measurable, which suffices since all constructions are in L^2. The irreducible representations of $\mathrm{SO}(2)$

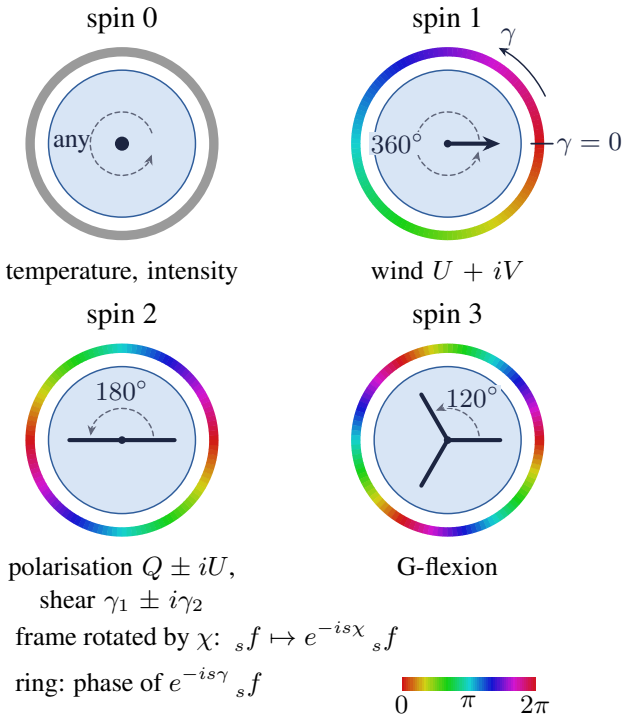


Figure 6. What spin means at a point. Each panel is the tangent plane at x, with e_θ to the right. A spin-s value is drawn as s equal spokes, and rotating the frame by χ multiplies it by $e^{-is\chi}$, as in (1). The ring, the fibre of $\mathrm{SO}(3)$ over x (Sec. 3.1), is coloured by the phase of the lift of the field, which winds s times around it.

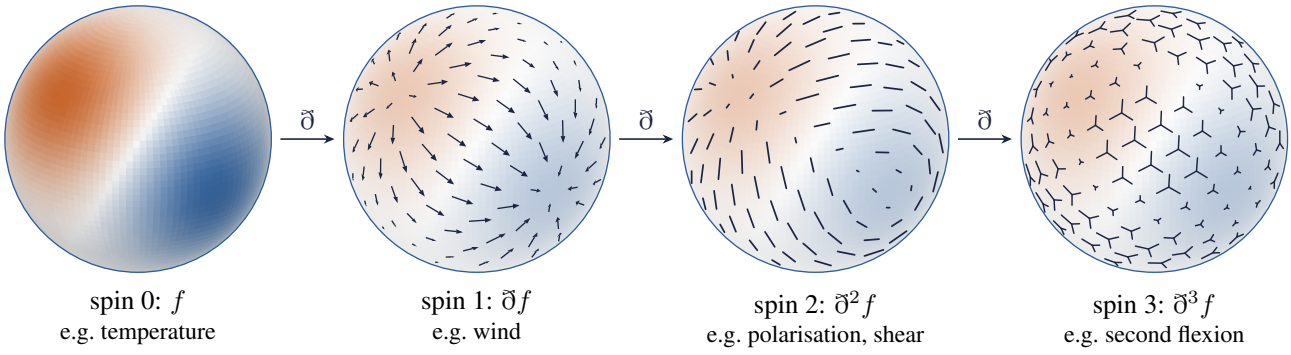


Figure 7. Spin fields on the sphere: a scalar field f and the fields $\eth f$, $\eth^2 f$ and $\eth^3 f$ obtained from it by the spin raising operator $\eth$ of (11), drawn with glyphs whose symmetry is their spin (Fig. 6). Wind has spin 1, polarisation and shear spin 2, and the second flexion of weak lensing spin 3 (Bacon et al., 2006).

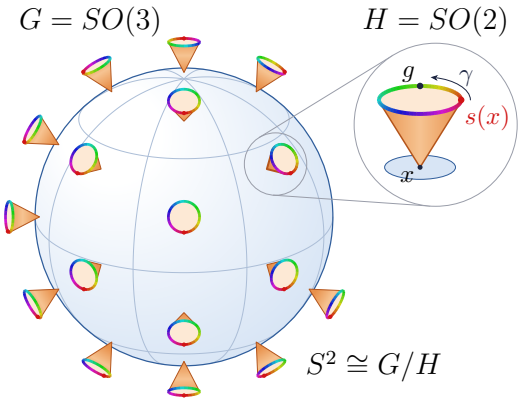


Figure 8. The rotation group as a principal $\mathrm{SO}(2)$-bundle over the sphere. The fibre over x, the rotations $R(\alpha,\beta,\gamma)$ with $R(\alpha,\beta,\gamma)\eta = x$, is drawn as the rim of a cone at x, coloured by the fibre coordinate γ, with the section $s(x) = R(\alpha,\beta,0)$ as a red dot and a general element $g = s(x)R_z(\gamma)$ in the inset.

are the characters $\rho^{(s)}(R_z(\gamma)) = e^{\mathrm{i}s\gamma}$, $s \in \mathbb{Z}$, and the Mackey functions of type s (McEwen, 2026, Definition IV.4) are the elements of

$$\mathcal{M}_s = \{F \in L^2(\mathrm{SO}(3)) : F(gR_z(\gamma)) = e^{-\mathrm{i}s\gamma}F(g)\}. \tag{17}$$

Changing the section to $s(x)R_z(\chi(x))$ multiplies the coordinate expression $\sqrt{2\pi}\,F(s(x))$ of a Mackey function by $e^{-\mathrm{i}s\chi(x)}$, which is the frame law (1) of a spin-s field (McEwen, 2026, eq. 36).

Definition 1 (Lift and projection; McEwen, 2026, Definition IV.6)**.** For $s \in \mathbb{Z}$ define the lift $\Lambda_s : L^2_s(S^2) \to L^2(\mathrm{SO}(3))$ and its adjoint, the projection $\Lambda_s^\dagger : L^2(\mathrm{SO}(3)) \to L^2_s(S^2)$, by

$$(\Lambda_s f)(\alpha,\beta,\gamma) = \frac{1}{\sqrt{2\pi}}\, e^{-\mathrm{i}s\gamma}\, f(\beta,\alpha), \tag{18}$$

$$(\Lambda_s^\dagger F)(\beta,\alpha) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{\mathrm{i}s\gamma}\, F(\alpha,\beta,\gamma)\,\mathrm{d}\gamma. \tag{19}$$

Lemma 2 (Structure of the lift; McEwen, 2026, Lemma IV.7)**.** *Let $s \in \mathbb{Z}$.*

(a) Λ_s is an isometry of $L^2_s(S^2)$ onto $\mathcal{M}_s$; $\Lambda_s^\dagger\Lambda_s = \mathrm{Id}$ and $\Lambda_s\Lambda_s^\dagger$ is the orthogonal projector onto $\mathcal{M}_s$. The subspaces $\mathcal{M}_s$ are mutually orthogonal and $L^2(\mathrm{SO}(3)) = \widehat{\bigoplus}_{s\in\mathbb{Z}}\mathcal{M}_s$.

(b) On $\mathcal{M}_s$ the projection is evaluation on the section: $\Lambda_s^\dagger F = \sqrt{2\pi}\, F \circ s$.

(c) In harmonic space,

$$(\Lambda_s f)^\ell_{mn} = (-1)^s \sqrt{\frac{8\pi^2}{2\ell+1}}\; {}_sf_{\ell m}\,\delta_{n,-s}, \tag{20}$$

$$(\Lambda_s^\dagger F)_{\ell m} = (-1)^s \sqrt{\frac{2\ell+1}{8\pi^2}}\; F^\ell_{m,-s}, \tag{21}$$

where the coefficients on the sphere are taken with respect to the spin-s harmonics.

(d) $\Lambda_s\ {}_sY_{\ell m} = (-1)^s\sqrt{\frac{2\ell+1}{8\pi^2}}\; D^{\ell *}_{m,-s}$: lifted spin harmonics are unit-norm Wigner functions.

(e) $\pi_G(g)\Lambda_s = \Lambda_s\pi_s(g)$ and $\Lambda_s^\dagger\pi_G(g) = \pi_s(g)\Lambda_s^\dagger$ for all $g \in \mathrm{SO}(3)$, where π_G is the left regular representation of $\mathrm{SO}(3)$ on $L^2(\mathrm{SO}(3))$, $(\pi_G(g)F)(g') = F(g^{-1}g')$, and π_s is the representation of $\mathrm{SO}(3)$ on spin-s fields induced from $\rho^{(s)}$; both are developed in McEwen (2026, Sec. IV.D), where (e) is proved as Proposition IV.10.

By McEwen (2026, Proposition IV.10), the induced representation $\pi_s(g) = \Lambda_s^\dagger\pi_G(g)\Lambda_s$ is the spin rotation operator $\mathcal{R}_\rho$ of McEwen et al. (2015b).

Corollary 3 (Functions on $\mathrm{SO}(3)$ as tuples of spin fields; McEwen, 2026, Corollary IV.12)**.** *Every $F \in L^2(\mathrm{SO}(3))$ decomposes uniquely as $F = \sum_{s\in\mathbb{Z}} \Lambda_s \Lambda_s^\dagger F$, its Fourier series in γ converging in L^2, with $(\Lambda_s \Lambda_s^\dagger F)^\ell_{mn} = F^\ell_{mn}\,\delta_{n,-s}$; the component $\Lambda_s^\dagger F$ is a spin-s field whose harmonic coefficients are the Wigner column $F^\ell_{m,-s}$ up to the constant in* (20)*. In particular F is azimuthally band-limited at N, $F^\ell_{mn} = 0$ for $|n| \geq N$, if and only if $F = \sum_{|s|<N} \Lambda_s f_s$ for spin fields $f_s = \Lambda_s^\dagger F$.*

A.3. Lifted convolutions

The statement summarised in Sec. 3.2, for the convolutions (4) and (6) (McEwen, 2026, Definitions III.4 and V.1), is the following.

Theorem 4 (Harmonic form; McEwen, 2026, Theorem V.2)**.** *Let $\Psi \in L^2(\mathrm{SO}(3))$, $s, s' \in \mathbb{Z}$, and let $\widetilde{\Psi}(g) = \Psi^*(g^{-1})$. For $|m|, |n| \leq \ell$, with $\ell \geq |s|$ whenever the input is a spin-s field and $\ell \geq |s'|$ whenever the output is a spin-s' field,*

$$\big(\mathcal{C}_{s\to G}[\Psi]\ {}_sf\big)^\ell_{mn} = (-1)^s \sqrt{\tfrac{8\pi^2}{2\ell+1}}\ {}_sf_{\ell m}\ \Psi^{\ell *}_{n,-s}, \tag{22}$$

$$\big(\mathcal{C}_{s\to s'}[\Psi]\ {}_sf\big)_{\ell m} = (-1)^{s+s'}\ {}_sf_{\ell m}\ \Psi^{\ell *}_{-s',-s}, \tag{23}$$

$$\big(\mathcal{C}_{G\to s'}[\Psi]\,F\big)_{\ell m} = (-1)^{s'} \sqrt{\tfrac{2\ell+1}{8\pi^2}} \sum_{m'=-\ell}^{\ell} F^\ell_{mm'}\ \Psi^{\ell *}_{-s',m'}, \tag{24}$$

$$\big(\mathcal{C}_{G\to G}[\Psi]\,F\big)^\ell_{mn} = \sum_{m'=-\ell}^{\ell} F^\ell_{mm'}\ \Psi^{\ell *}_{nm'}, \tag{25}$$

and the output of $\mathcal{C}_{s\to s'}[\Psi]$ has no component of degree $\ell < \max(|s|, |s'|)$. All four operators are bounded and $\mathrm{SO}(3)$-equivariant for every Ψ: $\mathcal{C}_{s\to s'}[\Psi]\,\pi_s(g) = \pi_{s'}(g)\,\mathcal{C}_{s\to s'}[\Psi]$ and correspondingly with π_G in place of π_s or $\pi_{s'}$. Their adjoints are $\big(\mathcal{C}_{s\to G}[\Psi]\big)^\dagger = \mathcal{C}_{G\to s}[\widetilde{\Psi}]$, $\big(\mathcal{C}_{s\to s'}[\Psi]\big)^\dagger = \mathcal{C}_{s'\to s}[\widetilde{\Psi}]$ and $\big(\mathcal{C}_{G\to G}[\Psi]\big)^\dagger = \mathcal{C}_{G\to G}[\widetilde{\Psi}]$.

In outline, the lift of ${}_sf$ occupies the Wigner column $n = -s$ (Lemma 2(c)), and the convolution (5) multiplies each block of degree ℓ on the right by $(\Psi^\ell)^\dagger$, so that only the kernel column $m' = -s$ is read, which gives (22). The projection (21) reads the column $n = -s'$, so that only $\Psi^\ell_{-s',-s}$ reaches the output and the ℓ-dependent constants cancel, which gives (23). Equivariance follows from Lemma 2(e) and the fact that the group convolution commutes with left translation, and the full proof is in McEwen (2026, Theorem V.2).

By (15) the spaces $L^2_s(S^2)$ and $L^2_{s'}(S^2)$ carry one copy of the irreducible representation D^ℓ for each $\ell \geq \max(|s|, |s'|)$, so that completeness follows from Schur's lemma.

Theorem 5 (Completeness; McEwen, 2026, Theorem V.5)**.** *Let $T : L^2_s(S^2) \to L^2_{s'}(S^2)$ be a bounded complex-linear map with $T\pi_s(g) = \pi_{s'}(g)T$ for all $g \in \mathrm{SO}(3)$. Then there is a unique bounded sequence $(c_\ell)_{\ell\geq\max(|s|,|s'|)}$ such that*

$$(T\ {}_sf)_{\ell m} = c_\ell\ {}_sf_{\ell m}, \tag{26}$$

with input degrees absent from the target annihilated. Conversely every bounded sequence defines such a T, with $\|T\| = \sup_\ell |c_\ell|$.

Proposition 6 (Azimuthally band-limited and steerable kernels; McEwen, 2026, Proposition VI.6)**.** *If the column $m' = -s$ of Ψ is supported on $|n| < N$ then*

$$\mathcal{C}_{s\to G}[\Psi] = \sum_{|s'|<N} \Lambda_{s'}\,\mathcal{C}_{s\to s'}[\Psi], \tag{27}$$

i.e. the output is represented by at most $2N-1$ spin-changing components $\mathcal{C}_{s\to s'}[\Psi]$, $|s'| < N$, with $2\min(\ell, N-1)+1$ admissible output slots at degree ℓ. Conversely, any finite tuple of bounded equivariant spin-changing maps from the same input spin assembles into a bounded map $\mathcal{C}_{s\to G}$ with a generally distributional kernel. A square integrable representative exists exactly when the sum of their Hilbert–Schmidt norms squared is finite. At finite band-limit the assembled kernel is always square integrable.

Corollary 7 (Incompatibility of directional kernels, spherical output and equivariance; McEwen, 2026, Corollary V.7)**.** *For a bounded complex-linear equivariant map $T : L^2_s(S^2) \to L^2_{s'}(S^2)$ between one fixed input spin and one fixed output spin,*

the two-point kernel is

$$K(x,y) = (-1)^{s+s'} \sum_{\ell} c_\ell \frac{2\ell+1}{4\pi} D^{\ell}_{-s',-s}\big(s(x)^{-1}s(y)\big), \tag{28}$$

in the distributional sense in general. It is radial apart from the frame phases prescribed by the two spins. Directional kernels, output on the sphere and $\mathrm{SO}(3)$*-equivariance therefore cannot be achieved simultaneously beyond those prescribed phases for this input/output pair.*

B. Spherical layers as neural operators

Table 5 summarises the spherical layers of Sec. 2.2 as neural operators, with the year of the first public version of each work (the FNO and the generalised spherical CNN appeared in the same month), although resolution-independent parameters imply neither exact sampled non-linear equivariance nor learned weights at unseen degrees.

Table 5. Spherical convolutional layers as neural operators. A layer is a neural operator in the sense of Kovachki et al. (2023) when its parameters do not depend on the sampling of the sphere. Representative costs at bandlimit L (azimuthal bandlimit $N \le L$, n planar grid points) assume fixed channel and retained spin counts and, for DISCO, bounded local neighbourhood size.

Method	Year	Signals	Linear layer	Cost
Spherical CNN (Esteves et al., 2018)	2018	S^2, scalar	harmonic zonal filter, one weight per ℓ	L^3
Spherical CNN (Cohen et al., 2018)	2018	$\mathrm{SO}(3)$, scalar	grid filter via generalised FFT	L^4
Clebsch–Gordan net (Kondor et al., 2018)	2018	$\mathrm{SO}(3)$, harmonic	per-ℓ maps, CG products	L^5
Spin-weighted CNN (Esteves et al., 2020)	2020	S^2, spins	harmonic filters interpolated over ℓ	L^3
Generalised spherical CNN (Cobb et al., 2021)	2020	S^2, $\mathrm{SO}(3)$	per-ℓ maps, azimuthal bandlimit N	NL^3
FNO (Li et al., 2021)	2020	plane	one multiplier per mode	$n \log n$
DISCO (Ocampo et al., 2023)	2022	S^2, scalar	continuous filter, quadrature	L^2
SFNO (Bonev et al., 2023)	2023	S^2, scalar	one weight per ℓ	L^3
S^2FNO (this work)	2026	S^2, $\mathrm{SO}(3)$, spins	real profile in ℓ per spin pair (10 knots)	L^3

C. Equivariant non-linearities

We consider the two settings of Sec. 4.2 in turn, the sphere and the rotation group.

C.1. Non-linearities on the sphere

A rotation moves a spin-s field and multiplies its value at each point by a phase $e^{-\mathrm{i}s\chi}$ of the form (1), where χ is the angle by which the local frame turns there. A pointwise map $\zeta : \mathbb{C} \to \mathbb{C}$ commutes with the move and is therefore equivariant if and only if it commutes with every such phase, which for $s \neq 0$ reads $\zeta(e^{\mathrm{i}\phi}z) = e^{\mathrm{i}\phi}\zeta(z)$ for all ϕ and z, since the rotations about each point turn its frame by every angle. The modulus non-linearity (9) multiplies z by a real function of $|z|$, so that

$$\zeta(e^{\mathrm{i}\phi}z) = \mathrm{ReLU}\big(a|z| + b\big)\,\frac{e^{\mathrm{i}\phi}z}{|z|+\epsilon} = e^{\mathrm{i}\phi}\,\zeta(z), \tag{29}$$

and it is therefore equivariant for every spin. For scalar fields ($s = 0$) every pointwise map is equivariant and standard activation functions apply.

C.2. Non-linearities on the rotation group

On $\mathrm{SO}(3)$ a rotation only translates the argument of a function, so that every pointwise map commutes with it. The lift-and-sum $\overline{\Lambda}$ places a tuple of spin fields, one for each distinct spin within a channel, in one function on $\mathrm{SO}(3)$, each in its own Wigner column by Lemma 2(c), a pointwise map ζ acts on that function, and a return $\mathcal{P}$ of projecting convolutions maps the result back to a tuple of spin fields. Definition 8 states this construction and Theorem 9 its equivariance.

Definition 8 (Lifted non-linearity)**.** For finite input and output spin sets, indexed by (s_i) and (s'_j), kernels $\Psi^j \in L^2(\mathrm{SO}(3))$ and a measurable pointwise map $\zeta : \mathbb{C} \to \mathbb{C}$ satisfying $|\zeta(z)| \le A + B|z|$ for finite nonnegative constants A, B, let

$$\sigma = \mathcal{P} \circ \zeta \circ \overline{\Lambda}, \qquad \mathcal{P} = \bigoplus_j \mathcal{C}_{G \to s'_j}[\Psi^j], \qquad \overline{\Lambda}\Big(\bigoplus_i {}_{s_i} f\Big) = \sum_i \Lambda_{s_i\ s_i} f. \tag{30}$$

Theorem 9 (Equivariance of the lifted non-linearity; Xu et al., 2022)**.** *Under the hypotheses of Definition 8, σ is a well-defined map between the corresponding finite sums of spin L^2 spaces and*

$$\sigma \circ \bigoplus_i \pi_{s_i}(g) = \bigoplus_j \pi_{s'_j}(g) \circ \sigma \qquad \text{for all } g \in \mathrm{SO}(3). \tag{31}$$

Theorem 9 is the result of Xu et al. (2022, Sec. 3.4), who show that their lift-and-sum, Eq. (3), followed by an elementwise non-linearity and a projection satisfies their condition, Eq. (2), proved in the same way. The growth condition is ours.

Proof of Theorem 9. The finite sum $\overline{\Lambda}$ is bounded into $L^2(\mathrm{SO}(3))$. The growth hypothesis gives $\|\zeta(F)\|_2 \le A\sqrt{8\pi^2} + B\|F\|_2$, and each component of $\mathcal{P}$ is bounded, so the composition is well defined. With the notation of Definition 8,

$$\begin{aligned}
\sigma \circ \bigoplus_i \pi_{s_i}(g) &= \mathcal{P}\,\zeta\,\overline{\Lambda} \bigoplus_i \pi_{s_i}(g) \\
&= \mathcal{P}\,\zeta\,\pi_G(g)\,\overline{\Lambda} && \text{since } \overline{\Lambda}\textstyle\bigoplus_i \pi_{s_i}(g) = \sum_i \Lambda_{s_i}\pi_{s_i}(g) = \sum_i \pi_G(g)\Lambda_{s_i} = \pi_G(g)\overline{\Lambda} \text{ by Lemma 2(e)} \\
&= \mathcal{P}\,\pi_G(g)\,\zeta\,\overline{\Lambda} && \text{since } \zeta \text{ acts pointwise and } \pi_G(g) \text{ only translates the argument} \\
&= \bigoplus_j \pi_{s'_j}(g)\,\mathcal{P}\,\zeta\,\overline{\Lambda} && \text{by Theorem 4 applied to each } \mathcal{C}_{G\to s'_j}[\Psi^j] \\
&= \bigoplus_j \pi_{s'_j}(g) \circ \sigma.
\end{aligned} \tag{32}$$

□

In our networks each component of the return is the fixed projection $\Lambda^\dagger_{s'}$, read as the Fourier column $n = -s'$ of the sampled group signal, followed by a learned filter for each channel and spin (Appendix D). By (21) this projection is the projecting convolution (24) with the kernel row $\Psi^{\ell *}_{-s',m'} = \delta_{m',-s'}$, a distributional kernel that lies in $L^2(\mathrm{SO}(3))$ once truncated at the bandlimit. The filter multiplies this row by one value for each degree and channel, and the constants of the implemented lift-and-sum and Fourier read can be absorbed into ζ and the filter, so that Theorem 9 covers the non-linearity of our networks in the continuum, and the group normalisation between ζ and the return uses invariant statistics and so preserves equivariance (Sec. 6.1). Xu's fixed fibre integral is $\sqrt{2\pi}\Lambda^\dagger_{s'}$ with their unnormalised stabiliser measure (Xu et al., 2022), and the orthogonal group projection of Ballerin et al. (2025) is $\Lambda_{s'}\Lambda^\dagger_{s'}$.

C.3. Discretisation

Theorem 9 concerns continuum functions, and $\zeta(F)$ is generally not bandlimited even when F is. The exact truncation $Q_{L,N}$ to the Wigner coefficients with $\ell < L$ and $|n| < N$ commutes with rotations, which act only on m, whereas computing these coefficients from samples aliases the content of $\zeta(F)$ beyond the bandlimits into them, which oversampling reduces (Cobb et al., 2021). Apart from the Conv classifiers, which oversample by a factor of two (Appendix D), our networks evaluate ζ without oversampling.

C.4. Parity

The antipodal map $P : x \mapsto -x$ commutes with rotations and multiplies scalar harmonic coefficients by $(-1)^\ell$. Networks on scalar fields built from isotropic multipliers and pointwise maps are therefore $\mathrm{O}(3)$-equivariant (Weiler et al., 2021) and cannot represent a map that anticommutes with P, $T(f \circ P) = -T(f) \circ P$, such as the advection $J(\psi, \zeta) = \hat{r} \cdot (\nabla\psi \times \nabla\zeta)$ of a vorticity $\zeta = \nabla^2\psi$ treated as a scalar. This is the counterpart, for such networks, of the selection rule of Gaunt tensor products (Xie et al., 2025). Spin-weighted and vector spherical signals restore the couplings of odd total degree (Liang et al., 2026; Xie et al., 2026), and spin fields escape the restriction, since P acts on a spin-s field by $f \mapsto (-1)^s f^* \circ P$. The vorticity is carried by the imaginary part of $\bar{\eth}(v_\theta + \mathrm{i}v_\varphi)$, and one spin-1 stage represents its advection through $(\eth\psi)^*\eth\zeta = \nabla\psi \cdot \nabla\zeta + \mathrm{i}J(\psi, \zeta)$ with an appropriate pointwise map on $\mathrm{SO}(3)$ that is not odd, unlike the modulus non-linearity. A multiplier between spins commutes with P, equivalently with the mirror $f \mapsto f^*(\theta, -\varphi)$, if and only if its coefficients are real, the spin counterpart of the fixed phase of reflection-equivariant planar kernels (Weiler & Cesa, 2019). The opposite parities of the E and B modes (Zaldarriaga & Seljak, 1997) are an instance of this criterion, which a mirror test of our classifiers with real profiles confirms. Completeness among real-linear maps requires the conjugate field f^*, as the conjugate frequencies do for planar harmonic networks (Weiler & Cesa, 2019).

D. Network architectures

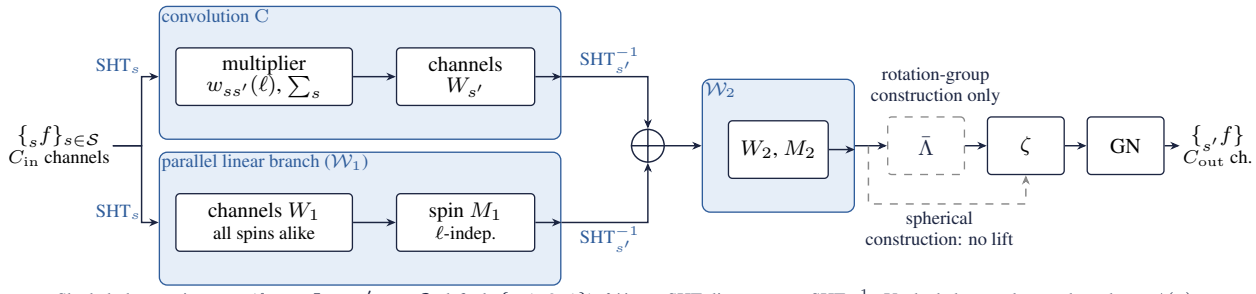
We describe the layers and networks of `s2ai` used for the reported results, with the notation and names of Sec. 4.4.

D.1. Layers

Spin spherical convolution. The spherical layer computes the forward spin-s transform of every input channel and spin, truncates or zero-pads the coefficients to the output bandlimit L_{out}, and forms for each input channel c and output spin s'

$$ {}_{s'}g^{c}_{\ell m} = \sqrt{\tfrac{4\pi}{2\ell+1}} \sum_{s \in \mathcal{S}_{\text{in}}} w^{c}_{ss'}(\ell)\ {}_{s}f^{c}_{\ell m}, \tag{33}$$

followed by a real channel-mixing matrix $W_{s'}$ for each output spin and inverse spin-s' transforms, so that the multipliers of the layer form a real and parameter-efficient subfamily of those of Theorem 5, interpolated in ℓ. Each real profile $w^{c}_{ss'}(\ell)$, independent of the output channel, is the piecewise-linear interpolation of 10 learned values at knots evenly spaced on $[0, L_{\text{out}}]$. The profile values are initialised Glorot-uniformly with fans $C_{\text{in}}|\mathcal{S}_{\text{in}}||\mathcal{S}_{\text{out}}|$ and 10, and the channel matrices $W_{s'}$, of shape $[C_{\text{in}}, C_{\text{out}}, |\mathcal{S}_{\text{out}}|]$, Glorot-uniformly with the default fan axes of JAX. Neither has a bias. The factor $\sqrt{4\pi/(2\ell+1)}$ is that of evaluation on the section, and the factor $(-1)^{s+s'}$ and the conjugation of (7) are absorbed into the real weights. Fig. 9 shows an S^{2}FNO block as implemented.



Spaces. Shaded: harmonic space ($\ell < L$; $s, s' \in \mathcal{S}$, default $\{-1, 0, 1\}$); $\mathcal{W}_{1,2}$: SHT, linear maps, SHT^{-1}. Unshaded: samples on the sphere. $\zeta(z) = \text{ReLU}(a|z| + b)\, z/(|z| + \varepsilon)$ and GN keep the phase. $\bar{\Lambda}$ (fixed): $F(\gamma) = \sum_s e^{-\text{i}s\gamma}\ {}_sf$ on $2N-1$ samples; the next block reads the columns $n = -s$.

Figure 9. An S^{2}FNO block as implemented. The convolution C, here the spin spherical convolution (33), and a harmonic 1×1 convolution ($\mathcal{W}_1$) with one channel matrix and one spin matrix shared across degrees act in parallel, and their sum passes through a second such convolution ($\mathcal{W}_2$), the modulus non-linearity ζ and group normalisation (GN). Only the rotation-group construction forms the lift-and-sum $\bar{\Lambda}$.

Lift-and-sum and projecting convolution. The lift-and-sum, which has no parameters, maps a tuple $({}_sf)_{s\in\mathcal{S}}$ to the samples $F(\gamma_k, x) = \sum_{s\in\mathcal{S}} e^{-\text{i}s\gamma_k}\ {}_sf(x)$ at $\gamma_k = 2\pi k/(2N-1)$, $k = 0, \dots, 2N-2$, without the factor $(2\pi)^{-1/2}$ of Definition 1. The projecting layer reads the columns $n = -s$, $s \in \mathcal{S}$, of a group signal by the discrete Fourier sum $(2N-1)^{-1}\sum_k F(\gamma_k, x)\, e^{\text{i}s\gamma_k}$, takes the forward spin-s transform of each, truncates or pads, multiplies by $\sqrt{4\pi/(2\ell+1)}\, w^{c}_{s}(\ell)$ with one real profile per channel and spin, mixes channels within each spin and applies the inverse spin-s transform. Spin s is thus mapped to spin s only, by the fixed adjoint up to a constant followed by a learned spin-preserving filter.

Harmonic 1×1 convolutions. The harmonic 1×1 convolutions $\mathcal{W}_1$ and $\mathcal{W}_2$ of the S^{2}FNO blocks (Fig. 10) mix channels by one real matrix of shape $[C_{\text{in}}, C_{\text{out}}]$ shared by all spins, and spins by one real matrix of shape $[|\mathcal{S}_{\text{in}}|, |\mathcal{S}_{\text{out}}|]$ shared by all degrees, orders and channels, both initialised Glorot-uniformly and without bias. The spin matrix gives spin-changing multipliers constant in ℓ, so that the stage is equivariant.

Activation and normalisation. The activation (9) has a real gain and offset for each channel, initialised to 1 and 0, so that it is the identity up to $\epsilon = 10^{-8}$ at initialisation. Group normalisation splits the channels into groups and, for each sample and group, computes the quadrature-weighted mean μ of the complex field and the mean squared modulus σ^2 of $x - \mu$. It returns $\big(a|x-\mu|/\sqrt{\sigma^2+\epsilon} + b\big)\, x/(|x| + \epsilon)$, with real a and b for each group initialised to 1 and 0, so that only the modulus is normalised and the phase of x is kept. Its statistics are computed for each spin separately on the sphere and pooled over the γ samples on the group. The pooled mean on the group is the $n = 0$ column, a spin-0 quantity, and hence an invariant constant, whereas the mean on the sphere of a field of non-zero spin is not rotation invariant, so that the mean subtraction on

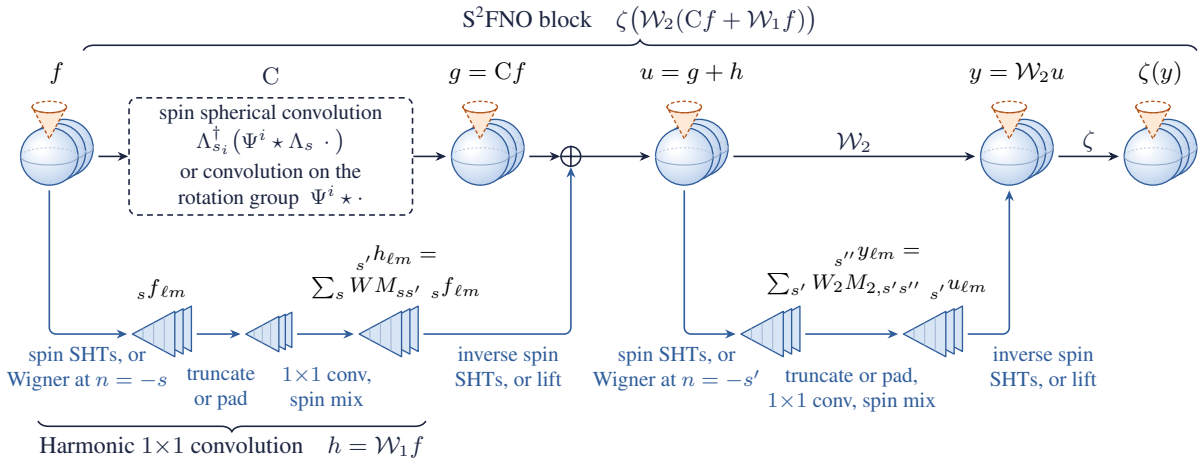


Figure 10. The generic S²FNO block (10) in detail. Stacked spheres with a dashed cone denote a tuple of spin fields or its lift to $\mathrm{SO}(3)$, and the normalisation is not shown.

the sphere is not exactly equivariant for $s \neq 0$. A normalisation with statistics computed in harmonic space and no mean subtracted for $s \neq 0$ would be exactly equivariant (Sec. 6.1), but the group normalisation already gives reasonable, if not exact, rotational equivariance, with errors of order 10^{-2} on the sphere and 10^{-3} on the rotation group (Table 1).

D.2. Blocks and spin spherical MNIST networks

Blocks. Writing C for the spin spherical convolution, C_G for the projecting convolution, $\overline{\Lambda}$ for the lift-and-sum, $\mathcal{W}_1$ and $\mathcal{W}_2$ for the harmonic 1×1 convolutions and N for the normalisation, the blocks of the four families are

$$\text{Conv-}S^2\text{:}\quad x \mapsto \mathrm{N}\,\zeta\,\mathrm{C}\,x, \qquad \text{Conv-SO(3):}\quad x \mapsto \mathrm{N}\,\zeta\,\overline{\Lambda}\,\mathrm{C}_G\,x,$$
$$\text{S}^2\text{FNO-}S^2\text{:}\quad x \mapsto \mathrm{N}\,\zeta\,\mathcal{W}_2\,(\mathrm{C}\,x + \mathcal{W}_1 x), \qquad \text{S}^2\text{FNO-SO(3):}\quad x \mapsto \mathrm{N}\,\zeta\,\mathcal{W}_2\,(\overline{\Lambda}\,\mathrm{C}_G\,x + \mathcal{W}_1 x),$$

and no block has an identity residual connection. The first layer of a Conv-$\mathrm{SO}(3)$ network is a spin spherical convolution followed by $\overline{\Lambda}$, ζ and N, and the first block of a S²FNO-$\mathrm{SO}(3)$ network is a S²FNO-S^2 block without activation and normalisation, followed by $\overline{\Lambda}$, ζ and N. In the UNets the input layer maps the data spin, 0 for images and 1 for vector fields, to $\mathcal{S} = \{-1, 0, 1\}$, with $N = 8$ in the rotation-group construction.

Classifiers. The classifiers of Table 2 follow Esteves et al. (2020) in their normalisation, downsampling and learning-rate schedule. Each stacks six blocks of a convolution, a normalisation and an activation, with $1 \to 4 \to 4 \to 8 \to 8 \to 16 \to 16$ channels for Conv-S^2, $1 \to 8 \to 8 \to 16 \to 16 \to 32 \to 32$ for Conv-$\mathrm{SO}(3)$ and S²FNO-S^2 and $1 \to 4 \to 8 \to 16 \to 16 \to 32 \to 32$ for S²FNO-$\mathrm{SO}(3)$, internal spins $\{0, 1\}$ in the spherical construction and azimuthal bandlimit $N = 4$ in the rotation-group construction. The bandlimit falls from 32 to 16 and 8 over successive pairs of blocks, and the Conv classifiers evaluate their activations on grids oversampled by a factor of two against aliasing. In the Conv classifiers each pair of channels has its own complex profile, and in the S²FNO classifiers each input channel has a complex profile for each pair of spins, followed by channel matrices. The batch normalisation takes statistics over the batch and the quadrature on the sphere (and γ on the group) and subtracts the mean only from spin-0 fields and on the group, so that it is equivariant. The activation is $\mathrm{ReLU}(\mathrm{Re}\,x + b)$, with a bias b for each channel, on spin-0 fields and on the samples on the group, and the modulus non-linearity $\mathrm{ReLU}(|x| + b)\,x/|x|$, with b initialised to -0.5, on fields of non-zero spin. The pooling returns only rotation-invariant features, the real and imaginary parts of the mean of the spin-0 field, or of the function on the group, and the mean modulus of every field. A linear layer to 128 units, a Swish activation (Ramachandran et al., 2017) and a linear layer then give the 10 class scores.

UNets. The MNIST UNets (Fig. 11) have an encoder of four levels with 8, 16, 32 and 32 channels and a decoder of two levels with 16 and 8 channels, each decoder level concatenating the previous output with the normalised encoder features of equal bandlimit and spin set. In the Conv UNets the bandlimit is 32 at the first two levels and 16 at the last two, and is restored to 32 in the decoder, by truncating and zero-padding the harmonic coefficients inside the convolutions, whereas the

S^{2}FNO UNets keep $L = 32$ throughout. The final layer, which maps the concatenation of $8 + 8$ channels to one channel of the target spin, is a spin spherical convolution without activation in Conv-S^2, a projecting convolution in the rotation-group construction and a S^{2}FNO-S^2 block, with its activation and normalisation, in S^{2}FNO-S^2.

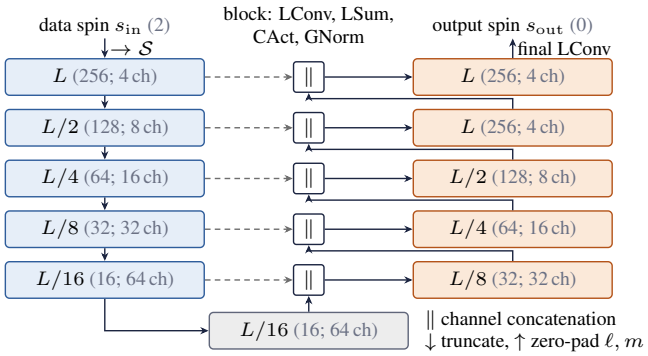


Figure 11. Composition of blocks into the UNets used for dense prediction. Every box is a block of a lifting convolution (LConv), the lift-and-sum (LSum), the activation (CAct) and group normalisation (GNorm). The first block maps the data spin onto the internal spin set $\mathcal{S}$ and the last layer returns the output spin. Brackets give the five-level spin UNet of the mass-mapping benchmark (bandlimit; channels), whereas the MNIST UNets are shallower and the S^{2}FNO UNets keep $L = 32$ at every level.

D.3. Networks for wind forecasting and weak lensing

For wind forecasting the spin UNet, built from Conv-SO(3) blocks like that of mass mapping, has a lifting input layer that maps the three complex input fields onto the spins $\{-1, 0, 1\}$, and the scalar UNet takes the real and imaginary parts of each input field as separate channels, six input channels in place of three.

The two mass-mapping UNets (Fig. 12) have five levels that halve the bandlimit from $L = 256$ to $L = 16$, with encoders of 4, 8, 16, 32 and 64 channels. In the spin UNet the first layer lifts the spin-2 shear to the internal spin set, every subsequent block acts on the lifted features, and the final layer returns the spin-0 convergence, whereas the scalar UNet combines DISCO convolutions (Ocampo et al., 2023) with residual layers (He et al., 2016), except in its last block, a residual layer alone, and its bottleneck adds a convolution from 128 to 64 channels.

D.4. Parameter counts

Table 6 counts trainable parameters only, and its MNIST counts are equal for spin-0 and spin-1 inputs.

Table 6. Trainable parameter counts of the architectures evaluated in Sec. 7.

Task	Architecture	Trainable parameters
Classification	Conv-S^2	49,674
Classification	Conv-SO(3)	133,562
Classification	S^{2}FNO-S^2	182,016
Classification	S^{2}FNO-SO(3)	136,088
Prediction	Conv-S^2	23,330
Prediction	Conv-SO(3)	13,890
Prediction	S^{2}FNO-S^2	29,097
Prediction	S^{2}FNO-SO(3)	20,034
Wind forecasting	Spin UNet	48,446
Wind forecasting	Scalar UNet	74,398
Mass mapping	Spin UNet	80,702
Mass mapping	Scalar UNet	93,198

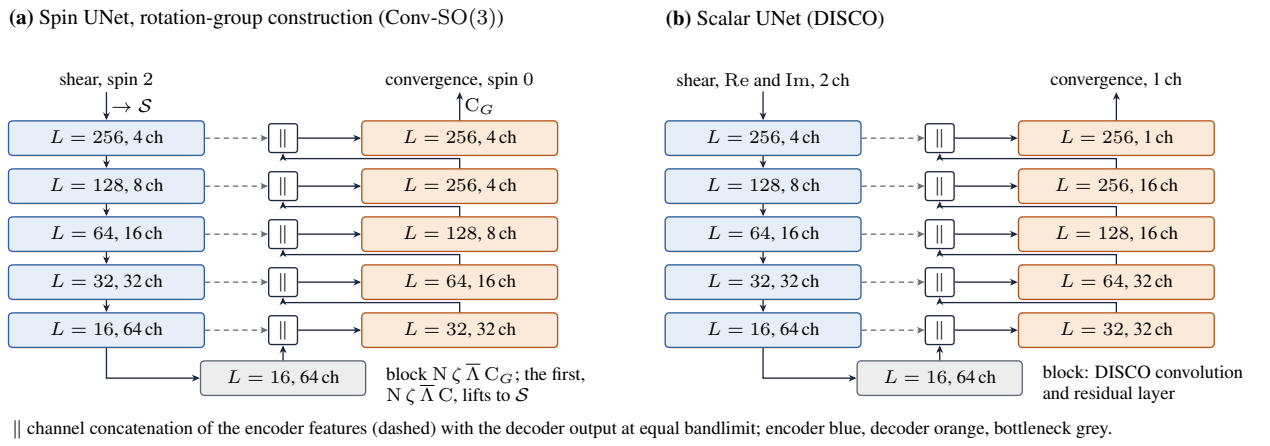


Figure 12. The two mass-mapping UNets, with five levels from $L = 256$ to $L = 16$. Each box gives the bandlimit and the number of channels of a block, and the dashed lines carry the encoder features, which are concatenated ($\|$) with the decoder output at equal bandlimit. (a) The spin UNet of the rotation-group construction, whose first block lifts the spin-2 shear to the internal spin set $\mathcal{S}$ and whose last layer, a projecting convolution C_G, returns the spin-0 convergence. (b) The scalar UNet, of DISCO convolutions and residual layers, which takes the real and imaginary parts of the shear as two channels and returns the convergence as its single channel.

E. Equivariance tests

We measure equivariance by the round-trip test of Cobb et al. (2021, Appendix D), also adopted by Ocampo et al. (2023). We consider $N_{\text{sig}} = 100$ test signals $\{f_i\}_{i=1}^{N_{\text{sig}}}$ (spin-s signals on the sphere or signals on the rotation group) with complex Gaussian random harmonic coefficients bandlimited at L, and $N_{\text{rot}} = 100$ rotations $\{\rho_j\}_{j=1}^{N_{\text{rot}}}$ drawn from the Haar distribution on $\mathrm{SO}(3)$. For each pair (f_i, ρ_j) we rotate a copy of the coefficients exactly in harmonic space by (15), synthesise both copies on the sampling grid of McEwen & Wiaux (2011), without oversampling, and pass them through the operator under test. We then transform both outputs to harmonic space, rotate the output from the unrotated input by the same rotation, transform both back to real space and compute the mean relative equivariance error

$$\text{MREE} = \frac{1}{N_{\text{sig}}} \frac{1}{N_{\text{rot}}} \sum_{i=1}^{N_{\text{sig}}} \sum_{j=1}^{N_{\text{rot}}} \frac{\|f_{ij} - g_{ij}\|_2}{\|g_{ij}\|_2}, \tag{34}$$

where f_{ij} denotes the rotated output obtained from the unrotated input, g_{ij} the output obtained from the rotated input, and $\|\cdot\|_2$ is the L^2 norm, as in Cobb et al. (2021), computed over samples, channels and spins or, for group signals, also the γ samples, with the colatitude quadrature weights of the sampling theorem. All errors are computed in double precision, at $L = 16$ with $N = 5$ on the rotation group, except for the linear models of Table 7, at $L = 32$.

The non-linear rows of Table 1 are tested on spin-0 signals on the sphere and on group signals, for which every pointwise map commutes with rotations in the continuum, so that the errors of the activations arise from aliasing. The rows ReLU and Swish (complex data) apply these functions directly to complex values, the ReLU keeping z when $\mathrm{Re}\, z > 0$ and setting it to zero otherwise. The complex activations instead act on the modulus and keep the phase. The complex ReLU is the modulus non-linearity (9) at its initialisation, $a = 1$ and $b = 0$, where it is the identity up to $\epsilon = 10^{-8}$, so that its error of order 10^{-9} is that of the round trip, and it would grow once a and b are learned. Complex Swish is the same non-linearity with the Swish $x/(1 + e^{-x})$ (Ramachandran et al., 2017) in place of the ReLU, with the same initial gain and offset. Equivariance of the trained networks, with their learned activations and normalisations, is not measured.

We also test blocks and whole UNets with several spins and channels, with every activation and normalisation removed so that every model is linear (Table 7). The UNets with $c = 8$ and $N = 8$ have the architecture of our spin spherical MNIST UNets and are tested in both directions, $0 \to 1$ and $1 \to 0$. The identity maps give the round-off floor of the transforms, of order 10^{-14} on the sphere and 10^{-13} on the rotation group. The positive control, a layer that adds one to the samples in the northern hemisphere, gives an error of order 10^{-1}, which shows that the test detects broken equivariance. Every linear block and UNet gives errors of order 10^{-14}, the floor on the sphere, as expected.

Model	N	Spins	Channels	Params	Equivariance error (MREE)
Identity on S^2		1	1	0	2.20×10^{-14}
Identity on $\mathrm{SO}(3)$	8		1	0	1.99×10^{-13}
Positive control		1	1	0	8.03×10^{-2}
S^{2}FNO-S^2 block		$0 \to 0 \to 0$	$1, 1, 1$	15	2.27×10^{-14}
S^{2}FNO-S^2 block		$1 \to 1 \to 1$	$1, 1, 1$	15	2.28×10^{-14}
S^{2}FNO-S^2 block‡		$0 \to \{-1, 0, 1\} \to \{-1, 0, 1\}$	$1, 8, 8$	138	2.00×10^{-14}
S^{2}FNO-S^2 block		$\{-1, 0, 1\}$	$8, 16, 16$	1,506	2.14×10^{-14}
S^{2}FNO-S^2 block		$\{-2, \dots, 2\}$	$4, 4, 4$	1,162	2.07×10^{-14}
S^{2}FNO-$\mathrm{SO}(3)$ block	5	$\{2\}$	$1, 1, 1$	15	2.38×10^{-14}
S^{2}FNO-$\mathrm{SO}(3)$ block	8	$\{-1, 0, 1\}$	$8, 16, 16$	1,026	2.35×10^{-14}
S^{2}FNO-$\mathrm{SO}(3)$ block	8	$\{-2, \dots, 2\}$	$4, 4, 4$	362	2.28×10^{-14}
S^{2}FNO-S^2 UNet		$0 \to \{-1, 0, 1\} \to 1$	$c = 8$	28,825	2.02×10^{-14}
S^{2}FNO-S^2 UNet		$1 \to \{-1, 0, 1\} \to 0$	$c = 8$	28,825	1.98×10^{-14}
S^{2}FNO-S^2 UNet		$0 \to \{-2, \dots, 2\} \to 1$	$c = 4$	24,865	2.52×10^{-14}
S^{2}FNO-$\mathrm{SO}(3)$ UNet	8	$0 \to \{-1, 0, 1\} \to 1$	$c = 8$	19,762	1.76×10^{-14}
S^{2}FNO-$\mathrm{SO}(3)$ UNet	8	$1 \to \{-1, 0, 1\} \to 0$	$c = 8$	19,762	1.61×10^{-14}
S^{2}FNO-$\mathrm{SO}(3)$ UNet	4	$0 \to \{-1, 0, 1\} \to 1$	$c = 4$	6,346	2.29×10^{-14}

Table 7. Equivariance errors of linear models at $L = 32$, measured by the MREE (34) over the 10^4 pairs of signal and rotation. Spins and channels are given as input, hidden and output, or as one set of spins used throughout, and c is the base width of a UNet, with encoder widths c, $2c$, $4c$ and $4c$. ‡The first block of the S^{2}FNO-S^2 UNet.

F. Datasets

We describe the three benchmarks of Sec. 7.1 in turn.

F.1. Spin spherical MNIST

We project planar digits onto the sphere following Cohen et al. (2018) and retain their scalar harmonic coefficients up to bandlimit L. We also construct a Sobel variant, motivated by Esteves et al. (2020); Xu et al. (2022), in which the Sobel derivatives of the projected image along the colatitude and longitude axes of the sampling grid, with their means subtracted, are combined into $X_\theta + \mathrm{i}X_\varphi$, whose modulus is rescaled to $[0, 1]$ for each sample with the phase kept, and the result is transformed as a spin-1 field before any rotation.

The harmonic variants are computed with `s2fft` (Price & McEwen, 2024), by a forward spherical harmonic transform of the scalar image at bandlimit L, a rotation in harmonic space by (15), exact at that bandlimit, and an inverse spin spherical harmonic transform at the desired spin, which synthesises ${}_sf = \sum_{\ell=|s|}^{L-1} \sum_m f_{\ell m}\, {}_sY_{\ell m}$. The synthesis keeps the scalar coefficients of the degrees $\ell \geq |s|$, without the degree-dependent factors of differential spin raising, and commutes with rotations by (15). A spin-1 field so constructed has no monopole, whereas the scalar fields keep a small one, since the mean subtracted from each image is the unweighted mean over the grid rather than the quadrature-weighted spherical mean.

For the prediction experiments we generate $60{,}000$ samples at $L = 32$ on the sampling grid of McEwen & Wiaux (2011) with 33×64 samples, from the first $51{,}000$ images of the MNIST training set and the first $9{,}000$ of the MNIST test set, which form an $85/15$ split into training and held-out sets. Rotations of the held-out samples use Euler angles (α, β, γ) drawn uniformly from $[0, 2\pi)^3$, which is not the Haar distribution on $\mathrm{SO}(3)$ (that would draw $\cos\beta$ uniformly), so that the rotated test set over-represents some rotation axes. The generation code and the exact split indices will be released with `s2ai`.

Splits and protocol. The classification experiments use a regenerated split of the same construction, in the image, Sobel spin-1, harmonic spin-1 and harmonic spin-2 variants. The first $55{,}000$ images of the MNIST training set are used for training and the remaining $5{,}000$, unrotated, for validation, and the $10{,}000$ images of the MNIST test set for testing, with Haar-distributed rotations drawn once with fixed seeds. Following Cohen et al. (2018), models are trained on unrotated samples and evaluated on unrotated (NR/NR) and randomly rotated (NR/R) test samples, and all results reported in Sec. 7.2 are NR/R.

F.2. Global wind forecasting

The training set is a subset of the WeatherBench 2 1.5° archive of ERA5 reanalysis data (Rasp et al., 2024; Hersbach et al., 2020) at the 50 hPa pressure level, covering four six-hourly observations per day between 1980 and 2022. The observations at 00:00, 06:00 and 12:00 UTC form the input and the observation at 18:00 UTC the target, with the zonal and meridional components combined into the spin-1 field $u + \mathrm{i}v$ (Sec. 7.1), which is the field $v_\theta + \mathrm{i}v_\varphi$ of the meridian frame up to the constant factor $-\mathrm{i}$. Data are converted from the equiangular latitude–longitude grid to the sampling of McEwen & Wiaux (2011) at $L = 128$, with 129×256 samples, and standardised by a real scale, after subtracting the mean of the training set. For evaluation, eight indices are drawn at random in the range $(0, 365)$, ignoring leap years, and the corresponding days, which cover all seasons, are held out from every year. The preprocessing code and the held-out indices will be released with `s2ai`.

F.3. Weak gravitational lensing

Weak lensing is described by two fields derived from the lensing potential ψ, namely the spin-0 convergence $\kappa = \frac{1}{2}\bar{\eth}\bar{\eth}\psi$, which traces the projected matter density, and the spin-2 shear $\gamma = \gamma_1 + \mathrm{i}\gamma_2 = \frac{1}{2}\eth\eth\psi$. Since both derive from the same potential, they are related in harmonic space by the spherical Kaiser–Squires relation (Kaiser & Squires, 1993; Wallis et al., 2022), which in our conventions reads

$$ {}_2\gamma_{\ell m} = \mathcal{D}_\ell \, \kappa_{\ell m}, \qquad \mathcal{D}_\ell = -\sqrt{\frac{(\ell+2)(\ell-1)}{\ell(\ell+1)}}, \quad \ell \geq 2. \tag{35} $$

The forward map (35) is therefore the spin $0 \to 2$ layer of Theorem 5 with $c_\ell = \mathcal{D}_\ell$, and spherical Kaiser–Squires is the spin $2 \to 0$ layer with $c_\ell = \mathcal{D}_\ell^{-1}$, where $|\mathcal{D}_\ell^{-1}| \leq \sqrt{3/2}$, on the degrees $\ell \geq 2$ that the shear determines. Practical considerations make the recovery of κ ill-posed and call for priors learned from simulations.

The convergence maps are simulated with GLASS (Tessore et al., 2023) and CAMB (Lewis et al., 2000) for a Planck 2018 cosmology (Planck Collaboration et al., 2020), with lognormal matter shells of 200 Mpc from redshift $z = 0$ to 1, one Gaussian source bin at $z = 0.5$, HEALPix resolution $N_{\mathrm{side}} = 512$ (Górski et al., 2005) and $\ell_{\max} = 1024$. The shear maps are forward modelled with `DarkMax`[1] with the noise of a Stage IV survey (Laureijs et al., 2011), a galaxy number density $n_{\mathrm{gal}} = 30$ arcmin^{-2} and an ellipticity dispersion $\sigma_e = 0.37$, with a different noise seed for each map. Maps are used at $L = 256$ on the sampling grid of McEwen & Wiaux (2011), with 257×512 samples. The training set holds $10{,}000$ shear–convergence pairs, split $80/20$ into training and validation sets, and the evaluation set 600 further pairs from the same pipeline.

G. Training details and additional results

Each model is trained once, except where stated, so that our results carry no estimate of their variation between runs.

G.1. Spin spherical MNIST

Training. The classifiers are trained in single precision on a shared NVIDIA GB10 with Adam (Kingma & Ba, 2015) ($\beta_1 = 0.9$) at a learning rate of 10^{-3}, multiplied by 0.2 at 50% and again at 83% of the run, in batches of 32 for 12 epochs (5 or 6 for Conv-SO(3), Table 2), minimising the categorical cross-entropy, and the reported checkpoint is the epoch of best accuracy on the non-rotated validation split. The prediction UNets are trained in double precision with Adam ($\beta_1 = 0.9$) at a learning rate of 5×10^{-4} for up to 2000 epochs in batches of 32, minimising the squared modulus of the difference between prediction and target, integrated with the quadrature weights of the sampling theorem, on NVIDIA GH200 and A100 GPUs with identical code. The reported prediction model is the checkpoint with the lowest loss on the rotated held-out set, on which the NR/R results are also evaluated, so that, with no separate validation set, the reported prediction errors are those of the best checkpoint on the evaluation data.

Prediction examples. The samples of Fig. 13 are fixed in advance, not chosen by their errors.

[1] Will be made public at `https://github.com/astro-informatics/darkmax`.

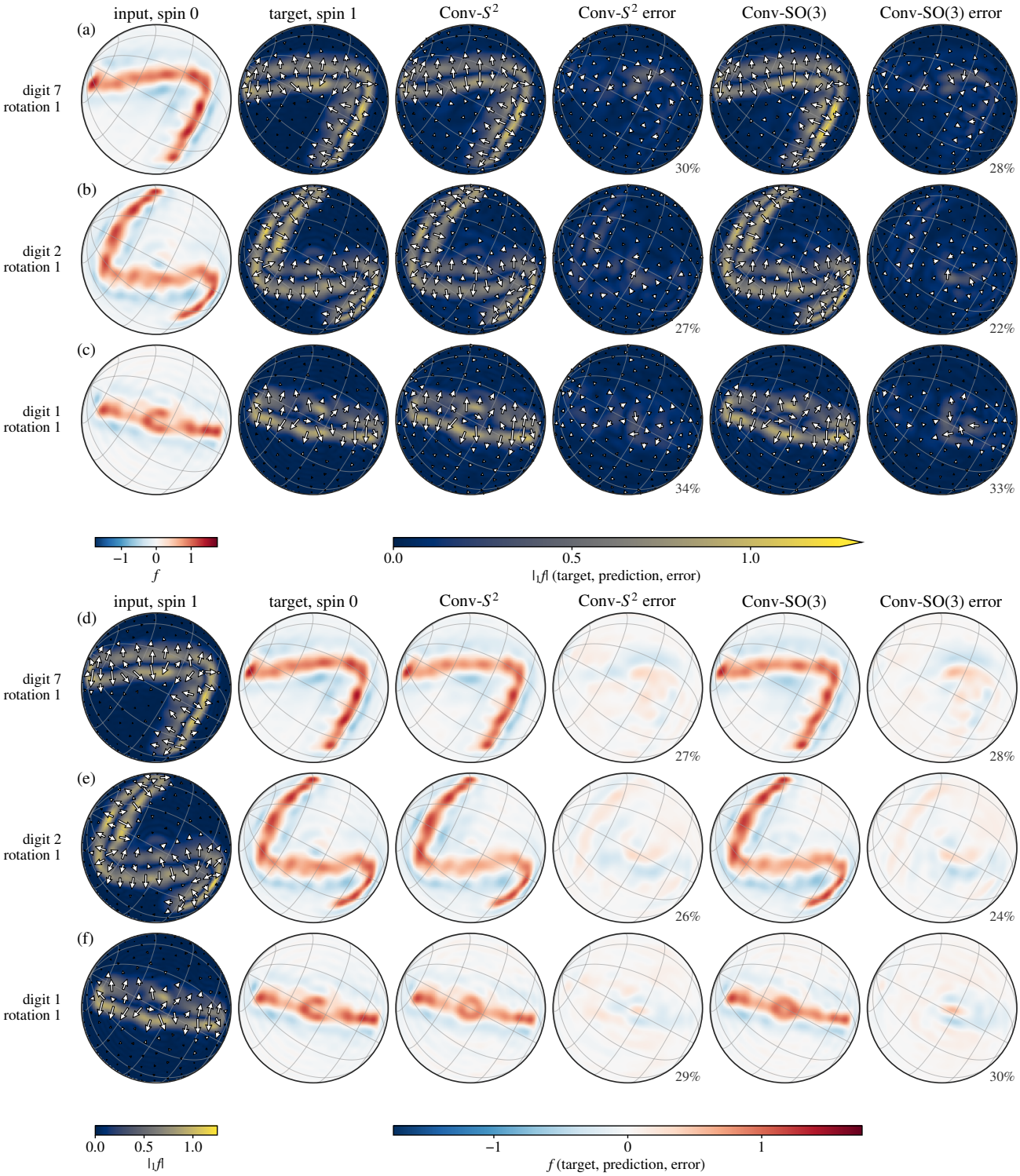


Figure 13. Prediction by the Conv-S^2 and Conv-SO(3) UNets at $L = 32$ from the spin-0 image to the spin-1 Sobel field (a–c) and back (d–f), for the first three test digits under one rotation. Errors (prediction minus target) are on the scales of the target, spin-1 fields are drawn as arrows over their modulus, and percentages give the relative error $\sqrt{\mathrm{MSE}/\overline{|f|^2}}$.

G.2. Wind forecasting

Training. Both UNets of Appendix D.3 are trained for 200 epochs with Adam (Kingma & Ba, 2015) at its default settings and a constant learning rate, minimising the mean squared error computed on the sphere, with the learning rates and batch sizes of Table 8, chosen with ablation studies. The reported model of each kind is the checkpoint with the lowest loss on the held-out days.

Table 8. Wind forecasting at 50 hPa and $L = 128$: configurations and all metrics on the held-out days, in physical units (MSE in $\mathrm{m}^2\,\mathrm{s}^{-2}$, RMSE and MAE in $\mathrm{m\,s}^{-1}$, SNR and PSNR in dB). Input shapes are batch $\times\ \theta \times\ \varphi \times$ channels, and best results are in bold.

Model	Params	LR	Input shape	MSE ↓	RMSE ↓	MAE ↓	SNR ↑	PSNR ↑	PCC ↑
Spin UNet	48,446	9×10^{-4}	20×129×256×3	**9.52**	**3.09**	**2.67**	**14.02**	**25.06**	**0.979**
Scalar UNet	74,398	8×10^{-4}	50×129×256×6	17.27	4.16	3.56	11.43	22.64	0.969
Persistence	–	–	–	14.55	3.82	3.31	12.17	23.19	0.962

G.3. Weak gravitational lensing

Training. Both UNets of Appendix D.3 are trained for 200 epochs on the $10{,}000$ shear–convergence pairs of Appendix F.3, split $80/20$ into training and validation sets, minimising a spherically weighted mean squared error with Adam (Kingma & Ba, 2015), with the inputs, batch sizes and learning rates of Table 9. Evaluation is on the 600 held-out maps.

Table 9. Weak lensing mass mapping at $L = 256$ with shape noise: configurations and all metrics, averaged over the 600 evaluation maps. Input shapes are batch $\times\ \theta \times\ \varphi \times$ channels, and best results are in bold.

Model	Params	LR	Input shape	MSE ↓ $\times 10^6$	RMSE ↓ $\times 10^3$	MAE ↓ $\times 10^3$	SNR ↑ (dB)	PSNR ↑ (dB)	PCC ↑
Spin UNet	80,702	5×10^{-4}	1×257×512×1	**1.28**	**1.13**	**0.90**	**7.53**	**23.47**	**0.907**
Scalar UNet	93,198	6×10^{-4}	20×257×512×2	2.06	1.43	1.14	5.45	21.40	0.846

Results. The residual of the scalar UNet retains high-amplitude, large-scale structure (Fig. 15), and the reconstruction of the spin UNet follows the tails of the one-point distribution of the target more closely (Fig. 15d).

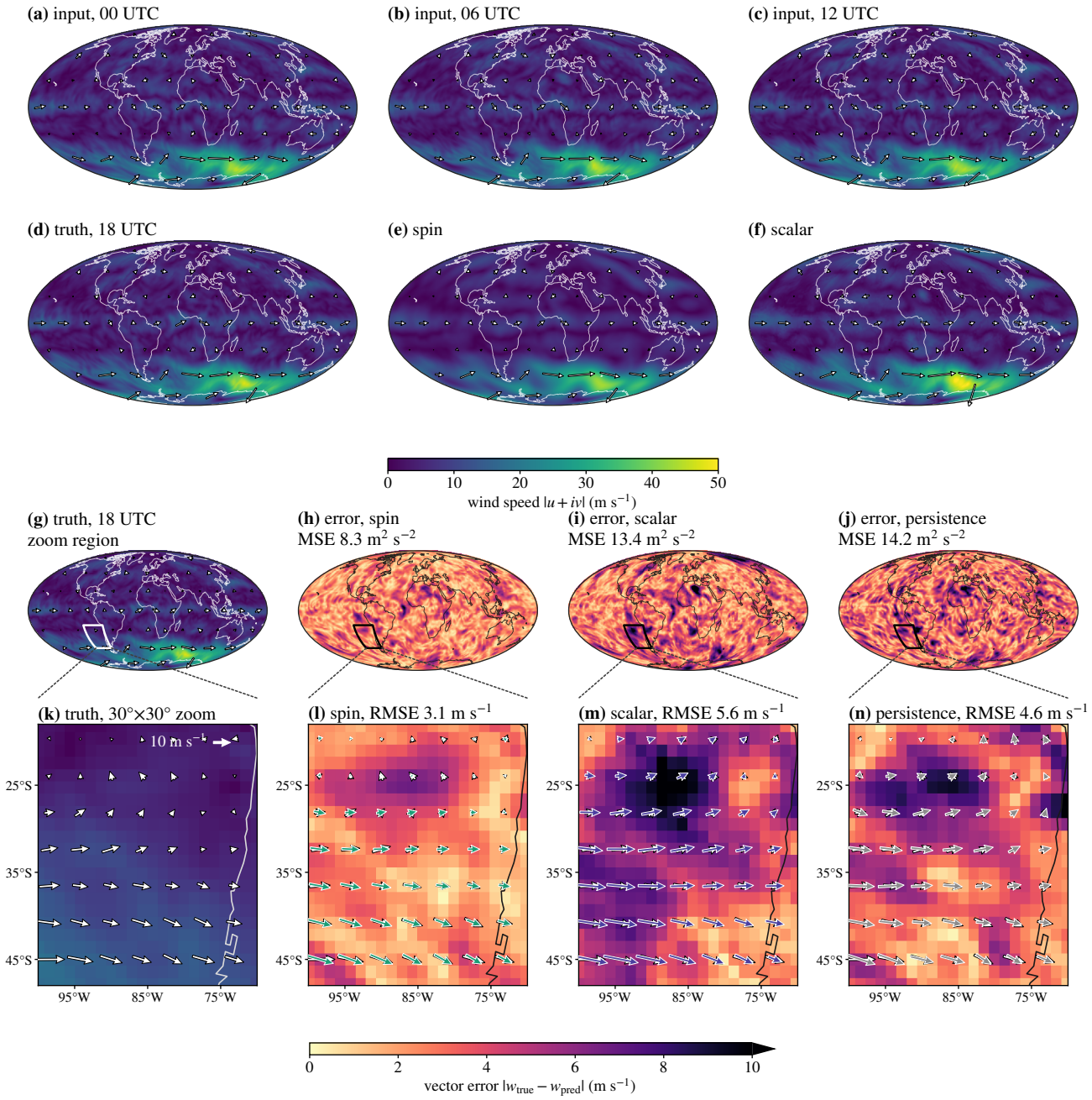


Figure 14. Wind forecast for one held-out day at $L = 128$: the speed $|u + iv|$ of the inputs, the target and the forecasts (a–f), the vector error $|w_{\text{true}} - w_{\text{pred}}|$ of the spin UNet, the scalar UNet and persistence with the area-weighted MSE of this day (h–j), and a zoom into the region outlined in (g), with true (white) and predicted (coloured) vectors and the RMSE over the region (k–n).

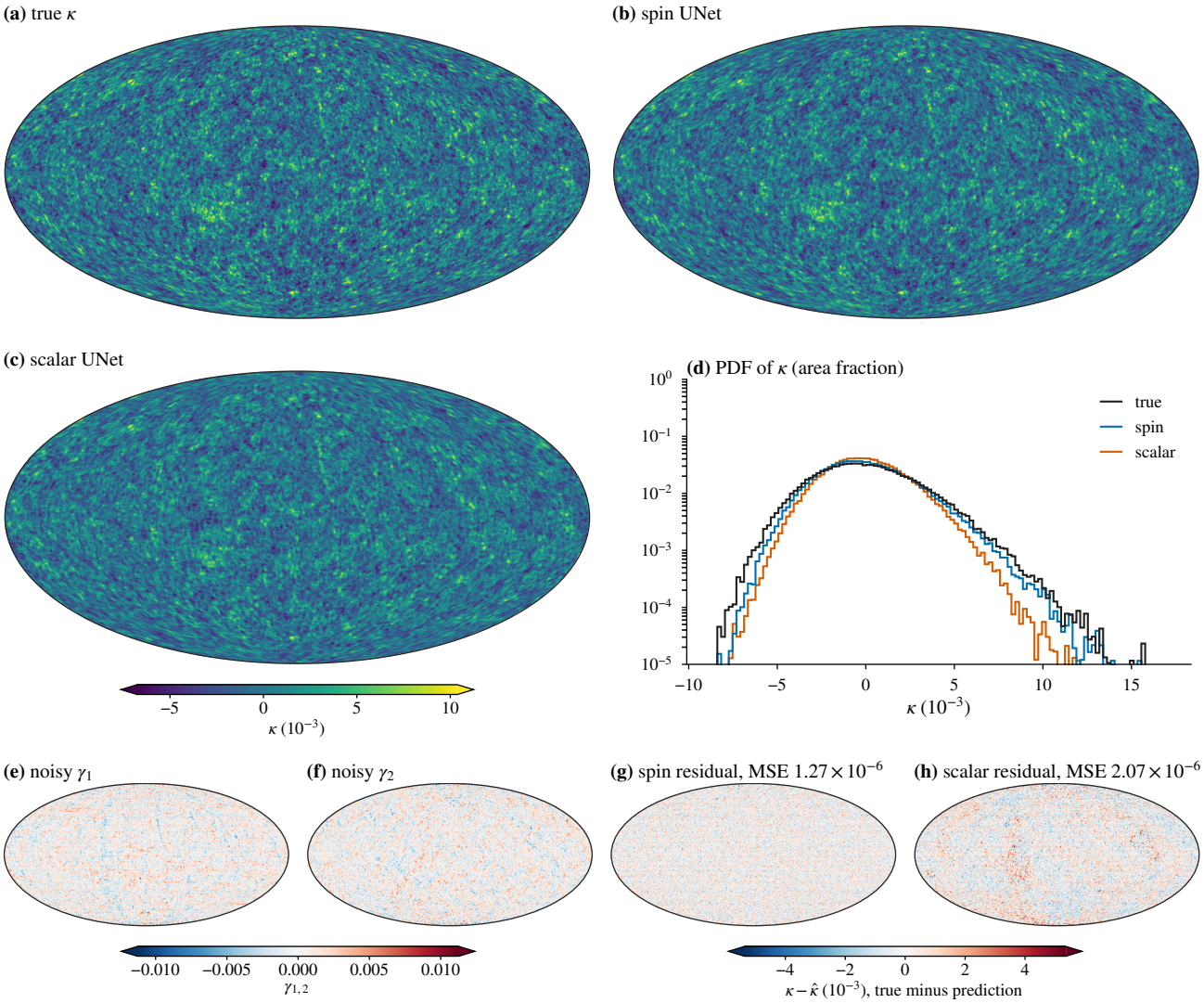


Figure 15. Mass mapping for one evaluation map at $L = 256$, with the true convergence (a), its reconstructions by the spin and scalar UNets (b, c) and their one-point distributions with that of the target (d), the noisy input shear (e, f) and the residuals $\kappa - \hat{\kappa}$ of the two UNets with their MSE on this map (g, h). The scalar residual retains large-scale structure that the spin residual does not.