APPROXIMATE EQUIVARIANCE SO(3) NEEDLET CONVOLUTION

Kai Yi[*]
Data Science Hub
& School of Mathematics and Statistics
University of New South Wales
kai.yi@unsw.edu.au

Jialin Chen[*] **& Yu Guang Wang**[†]
Institute of Natural Sciences
& School of Mathematical Sciences
Shanghai Jiao Tong University
{sjtuchenjl,yuguang.wang}@sjtu.edu.cn

Bingxin Zhou
Zhangjiang Institute for Advanced Study
Shanghai Jiao Tong University
University of Sydney Business School
The University of Sydney
bzho3923@uni.sydney.edu.au

Pietro Liò
Department of Computer Science and Technology
& Cambridge Centre for AI in Medicine
University of Cambridge
pl219@cam.ac.uk

Yanan Fan
Data Science Hub
& School of Mathematics and Statistics
University of New South Wales
y.fan@unsw.edu.au

Jan Hamann
School of Physics
University of New South Wales
jan.hamann@unsw.edu.au

ABSTRACT

This paper develops a rotation-invariant needlet convolution for rotation group SO(3) to distill multiscale information of spherical signals. The spherical needlet transform is generalized from $\mathbb{S}^2$ onto the SO(3) group, which decomposes a spherical signal to approximate and detailed spectral coefficients by a set of tight framelet operators. The spherical signal during the decomposition and reconstruction achieves rotation invariance. Based on needlet transforms, we form a Needlet approximate Equivariance Spherical CNN (NES) with multiple SO(3) needlet convolutional layers. The network establishes a powerful tool to extract geometric-invariant features of spherical signals. The model allows sufficient network scalability with multi-resolution representation. A robust signal embedding is learned with wavelet shrinkage activation function, which filters out redundant high-pass representation while maintaining approximate rotation invariance. The NES achieves state-of-the-art performance for quantum chemistry regression and Cosmic Microwave Background (CMB) delensing reconstruction, which shows great potential for solving scientific challenges with high-resolution and multi-scale spherical signal representation. The official code implementation is https://github.com/ykiiiiii/NES.

1 INTRODUCTION

Many data types in the real world can be modeled as spherical data, such as omnidirectional images Coors et al. (2018), molecules Boomsma & Frellsen (2017), and cosmic microwave background Akrami et al. (2020). Such spherical signals contain abundant topological features. Unfortunately, existing research Caldeira et al. (2019); Yi et al. (2020) usually maps spherical signals to $\mathbb{R}^2$ for convenient modeling with convolutional neural networks (CNNs), which results in distorted signals and ineffective shift equivariance Marinucci et al. (2008).

[*]Equal Contribution
[†]Corresponding Author

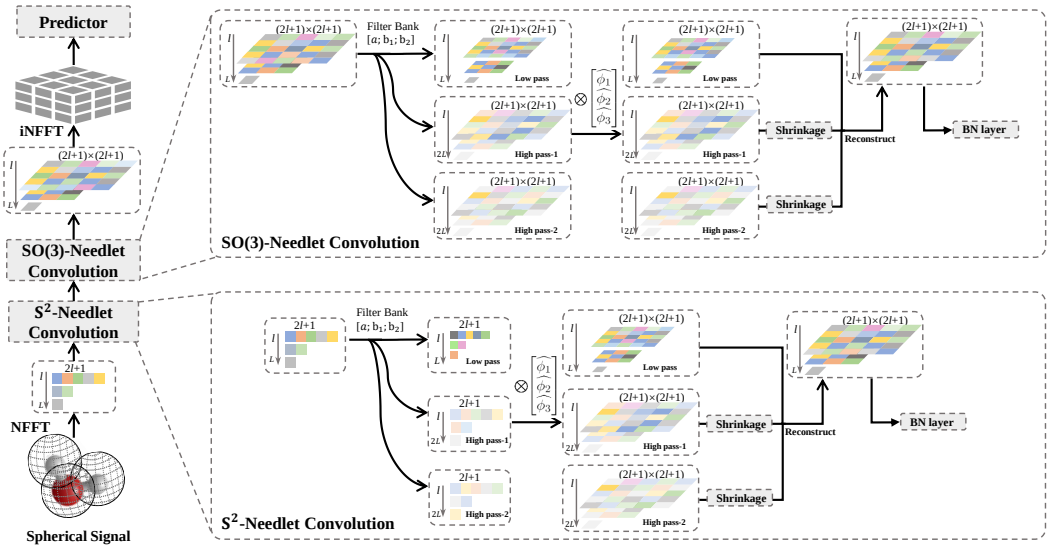

Figure 1: This figure shows the framework of our NES. As the left column shows, we first carry on a non-equispaced fast Fourier transform (NFFT) with predefined weights on the spherical signal. The following are an $\mathbb{S}^2$-Needlet Convolution and $\mathrm{SO}(3)$-Needlet Convolutions, which can be used to decompose the signal in multi scales. Then, we use the inverse NFFT (iNFFT) over the output of the $\mathrm{SO}(3)$-Needlet Convolution and feed the reconstructed signal into the downstream predictor.

Alternatively, geometric deep learning Bronstein et al. (2017; 2021) provides a universal blueprint for learning stable representation of high-dimensional data in different domains to build equivariant or invariant neural network layers that respect exact or approximate data symmetries, such as translation, rotation, and permutation. As a fundamental requirement for many applications, it has been proven critical to preserving the symmetry property in deep learning algorithm design Baek et al. (2021); Davies et al. (2021); Méndez-Lucio et al. (2021).

Equivariance is a significant property of geometric deep learning models as required by many physical sciences, such as chemistry Atz et al. (2021) and biology Townshend et al. (2021). This paper develops a scalable geometric deep learning model for spherical signal processing and learning with theoretically guaranteed rotation equivariance. Our model is based on needlet convolution on $\mathbb{S}^2$ and rotation group SO(3). The former describes the data representation on spherical point locations, while the latter records three-dimensional rotation angles of the signal. The input data features are embedded in each spherical point.

The main convolution computational unit is based on spherical needlets, which define a wavelet-like system on the two-dimensional sphere $\mathbb{S}^2$ that forms a tight frame on the sphere Narcowich et al. (2006a;b); Wang et al. (2017). A needlet is characterized by a highly-localized spherical radial polynomial, which covers a large scale but captures detailed features in local regions.

The *needlet convolution* on SO(3) decomposes spherical signals into low-pass and high-pass needlet coefficients. By separately storing and processing approximate and detailed information of the input, the network establishes hidden embeddings with enhanced scalability. In addition, the wavelet shrinkage operation Donoho (1995); Baldi et al. (2009) gains robust representations by filtering out redundant high-pass information in the framelet domain. The exact multiscale embeddings by SO(3)-needlet convolutions are invariant to rotation. Such convolutions can construct a deep neural network that distills the geometric invariant features of a spherical signal. We name it **N**eedlet approximate **E**quivariance **S**pherical CNN (**NES**). Inside the network, we utilize the convolution over the rotation group in multi scales to guarantee rotation equivariance.

The NES with shrinkage activation gains provably approximate equivariance, where the equivariance error diminishes at sufficiently high scales. Moreover, the needlet convolution is implemented efficiently with fast Fourier transforms (FFTs) on the sphere and rotation group. We validate the proposed NES on different real-world scientific problems with high-resolution and multi-scale

spherical signal inputs including regressing quantum chemistry molecules and reconstructing lensing Cosmic Microwave Background, for which our method achieves state-of-the-art performance.

2 SPHERICAL NEEDLET FRAMEWORK

Needlets are a type of framelets Wang & Zhuang (2020); Han (2017) that enjoys good localization properties in both spatial and harmonic spaces. We formulate a spherical needlet transform, which projects the given spherical signal to a set of multi-scale needlet representations in the framelet domain. The new representations can be uniquely decomposed. They are easy to compute, and divide approximate and detailed information into different scale levels, as traditional wavelets.

2.1 CHARACTERIZATION OF MULTI-SCALE SPHERICAL NEEDLETS

Needlets are defined on a Riemannian manifold $\mathcal{M}$. This paper considers a special case of $\mathcal{M}$, i.e., on $\mathbb{S}^2$ or $\mathrm{SO}(3)$. We define the spherical needlets with a *filter bank* $\eta := \left\{a; b^1, \ldots, b^r\right\} \subset l_1(\mathbb{Z}) := \left\{h = \{h_k\}_{k\in\mathbb{Z}} \subset \mathbb{C} : \sum_{k\in\mathbb{Z}} |h_k| < \infty\right\}$ and a set of associated *generating functions* $\Psi = \{\alpha; \beta^1, \cdots, \beta^r\} \subset L_1(\mathbb{R})$. We name filter a the *low-pass filter*, and filters $\{b^1, \cdots, b^r\}$ the *high-pass filters*. The former distills approximate information from the input signal, and the latter reserves more detailed information and together with noise. The associated generating functions and filter bank satisfy the relationship

$$\widehat{\alpha}(2\xi) = \widehat{a}(\xi)\widehat{\alpha}(\xi), \quad \widehat{\beta^n}(2\xi) = \widehat{b^n}(\xi)\widehat{\alpha}(\xi), \tag{1}$$

where $n = 1, \ldots, r$, and $\xi \in \mathbb{R}$.

To discretize the continuous needlets with zero numerical error, we utilize *Polynomial-exact Quadrature Rule* Wang et al. (2017) that are generated by the tensor product of the Gauss-Legendre nodes on the interval $[-1, 1]$ and equi-spaced nodes in the dimension with non-equal weights, such as longitude on sphere. Let $v_{j,k}$ represent low-pass coefficients, and $w^n_{j,k}$ represent high-pass coefficients of the signal function f, where $k = 0, \ldots, N_{j+1}$ and $j \geq J$, N_j is the number of sampling points at scale j. The low-pass (or high-pass) coefficients are defined by the inner products of low-pass (or high-pass) needlets and f. In practice, we calculate the coefficients in the Fourier domain for fast computation by

$$\widehat{v}_{j,\ell} = \widehat{f}_\ell \overline{\widehat{\alpha}\left(\frac{\lambda_\ell}{2^j}\right)}, \quad \widehat{w}^n_{j-1,\ell} = \widehat{f}_\ell \overline{\widehat{\beta^n}\left(\frac{\lambda_\ell}{2^{j-1}}\right)}. \tag{2}$$

We denote $\widehat{f}_\ell$ as the generalized Fourier coefficients of f at degree ℓ. More details about the filter bank and construction of needlets on $\mathbb{S}^2$ and $\mathrm{SO}(3)$ are given in Appendix **??**.

2.2 SPHERICAL NEEDLET CONVOLUTION

The *spherical needlet convolution* on $\mathcal{M}$ is defined by

$$[\phi \star f](R) = \langle L_R\phi, f\rangle = \int_{\mathcal{M}} \phi(R^{-1}x) f(x) dx, \tag{3}$$

where f is a signal, ϕ is a learnable locally supported filter, $L_R\phi(x) = \phi(R^{-1}x)$, and $\mathcal{M}$ represents $\mathbb{S}^2$ or $\mathrm{SO}(3)$. The constructed needlet convolution is rotation equivariant. Formally, a neural network (i.e., a function on $\mathcal{M}$) Φ is said *rotation equivariant* if for an arbitrary rotation operator L_R, there exists an operator T_R such that $\Phi \circ L_R = T_R \circ \Phi$. A rotation equivariant neural network provides more efficient and accurate prediction with theoretical support, which properties are desired for rotatable signals. It is provable that the convolution in (3) satisfies the Fourier theorem, i.e., $\widehat{[\phi \star f]}_\ell = \widehat{f}_\ell \cdot \widehat{\phi}^\dagger_\ell$, where $\dagger$ denotes the conjugate transpose and ℓ is the degree parameter. The operation $\cdot$ is matrix multiplication for $\mathrm{SO}(3)$ and outer product for $\mathbb{S}^2$.

The formulation in (3) has been adopted by Spherical CNN (Cohen et al., 2018), which induces convolution on Fourier coefficients. We define the convolution using needlet coefficients of a spherical signal. We construct the needlet coefficients with the needlet system defined in Section 2.1. We take $n = 2$ and get $\{\widehat{v}_{1,\ell}\}_{\ell=1}^{\Lambda_{J_0}}$, $\{\widehat{w}^1_{1,\ell}\}_{\ell=1}^{\Lambda_{J_1}}$ and $\{\widehat{w}^2_{1,\ell}\}_{\ell=1}^{\Lambda_{J_1}}$ for a low-pass and two high-pass needlet

coefficients, where Λ_j denotes sequence length of Fourier series of quadrature rule sampling points at scale j, and J_0, J_1 are the scale of low pass and high pass respectively.

These needlet coefficients can be used to reconstruct the Fourier coefficients $\hat{f}$ of signal f of degree ℓ. We denote this relation as $\left[\widehat{v}_{1,\ell}, \widehat{w}^1_{1,\ell}, \widehat{w}^2_{1,\ell}\right]^\top \asymp \widehat{f}_\ell$, where $\asymp$ means formal equivalence. We hereby establish formally an equivalent expression of $[\widehat{\phi \star f}]_\ell$ with multi-scale information and rotation equivariance:

$$\begin{bmatrix}\widehat{\phi_{1\ell}}\\ \widehat{\phi_{2\ell}}\\ \widehat{\phi_{3\ell}}\end{bmatrix} \odot \widehat{f}_\ell \asymp \begin{bmatrix}\widehat{\phi_{1\ell}}\\ \widehat{\phi_{2\ell}}\\ \widehat{\phi_{3\ell}}\end{bmatrix} \odot \begin{bmatrix}\widehat{v}_{1,\ell}\\ \widehat{w}^1_{1,\ell}\\ \widehat{w}^2_{1,\ell}\end{bmatrix} = \begin{bmatrix}\widehat{\phi_{1\ell}} \cdot \widehat{f}_\ell \overline{\widehat{\alpha}\left(\frac{\lambda_\ell}{2^{J_0}}\right)}\\ \widehat{\phi_{2\ell}} \cdot \widehat{f}_\ell \overline{\widehat{\beta^1}\left(\frac{\lambda_\ell}{2^{J_0}}\right)}\\ \widehat{\phi}_{3,\ell} \cdot \widehat{f}_\ell \overline{\widehat{\beta^2}\left(\frac{\lambda_\ell}{2^{J_0}}\right)}\end{bmatrix} = \begin{bmatrix}[\widehat{\phi_1 \star f}]_\ell \overline{\widehat{\alpha}\left(\frac{\lambda_\ell}{2^{J_0}}\right)}\\ [\widehat{\phi_2 \star f}]_\ell \overline{\widehat{\beta^1}\left(\frac{\lambda_\ell}{2^{J_0}}\right)}\\ [\widehat{\phi_3 \star f}]_\ell \overline{\widehat{\beta^2}\left(\frac{\lambda_\ell}{2^{J_0}}\right)}\end{bmatrix} \asymp [\widehat{\phi \star f}]_\ell.$$

Here $\widehat{\phi}_{i\ell}$ $(i = 1, 2, 3)$ are three learnable filters defined in the frequency domain and $\odot$ is the Hadamard product.

2.3 Rotation equivariance Error

Shrinkage Function One potential drawback of the spherical CNNs comes from the non-linear activation in each layer. The Fourier transforms introduce redundancy to feature representation in the frequency domain, which results in a heavy computational cost. To best preserve rotation equivariance at a reduced computational complexity, we employ a non-linear activation directly in the frequency domain with a small rotation equivariance error. Similar to UFGConv (Zheng et al., 2021), we cut off the high-pass coefficients x in the frequency domain by a shrinkage thresholding, i.e.,

$$\text{Shrinkage}(x) = \text{sgn}(x)(|x| - \lambda)_+, \forall x \in \mathbb{R},$$

with the threshold value $\lambda = \sigma\sqrt{2\text{log}(N)}/\sqrt{N}$ for N coefficients. The hyperparameter σ is an analogue to the noise level of the denoising model. Note that we do not cut off the low-pass framelet coefficients, as they distill important approximate information of input data. It is critical to offering an approximate rotation equivariance for the shrinkaged needlet convolution, as we discuss as follows.

Theorem 2.1. *Let $\mathbb{W}^s_p(\mathbb{S}^2)$ with $s > 2/p$ and $1 \leq p \leq \infty$ be a Sobolev space embedded in $\mathbb{L}_p(\mathbb{S}^2)$. For $f \in \mathbb{W}^s_p(\mathbb{S}^2)$, ϕ is a filter, then the rotation equivariance error due to using the shrinkage function is defined as the maximum of the following over all $R \in \text{SO}(3)$,*

$$\mathcal{E}(f) = \max_{R\in\text{SO}(3)} \sum_{\ell=0}^{B} \left\| \text{Shr}(\widehat{L_R f \star \phi})^{(\text{H})}_\ell - D^\ell(R)\text{Shr}(\widehat{f \star \phi}^{(\text{H})}_\ell) \right\|^2,$$

where B is the bandwidth for spherical signal embedding depending on the specific quadrature rule used, $\text{Shr}(\cdot)$ *represents shrinkage function, superscript* (H) *indicates the high-pass coefficients. Then, the approximate equivariance error for f is*

$$\mathcal{E}(f) \leq C 2^{-(J_0+1)s},$$

where J_0 is the scale of the low pass, C is a constant depending only on s, ϕ and the Sobolev norm of f.

The shrinkage mechanism thus introduces a stable rotation equivariance error. The condition in Theorem 2.1 $s > 2/p$ indicates that each function of $\mathbb{W}^s_p(\mathbb{S}^2)$ has a representation in the continuous function space on $\mathbb{S}^2$. Then, the numerical computation for f makes sense.

Pooling Operator We also establish a spectral pooling in the frequency domain to circumvent repeated Fourier transforms. A spectral pooling removes coefficients with degree larger than $\ell/2$ for the spectral representation $\hat{f} = [\hat{f}_0, \hat{f}_1, \cdots, \hat{f}_\ell]$. We prove that the spectral pooling operator is rotation equivariant

Network Architecture The framework is scalable to any application scenarios that can be represented by spherical signals. We illustrate the overall workflow of our proposed Needlet Spherical CNNs in Figure 1 with application scenario for bio-molecular prediction, where the input is a set

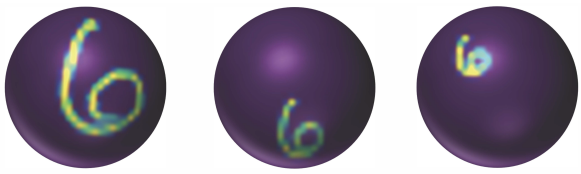

Figure 2: Illustration of a projected MNIST digit onto the sphere with 3 different downscale ratios (10%, 50%, 90%, left to right). The higher downscale indicates that the size of the digit is smaller on the sphere, which increases the difficulty of the model feature extraction where more detailed information needs captured.

Downscale Ratio	10%	30%	50%	70%	90%
Spherical CNN	94.99	92.17	86.92	83.73	78.71
NES	**97.84**	**97.30**	**96.74**	**95.21**	**92.66**

Table 1: Test accuracy on spherical MNIST with varying scales.

of spherical signals centered at atoms of a molecule. The spherical data is sampled on the points of a polynomial-exact quadrature rule. Based on the rule, we implement non-equispaced fast Fourier transforms (NFFTs) with predefined weights. The Fourier representations are sent through an $\mathbb{S}^2$-needlet convolution to $\mathrm{SO}(3)$. A number of rotation equivariant $\mathrm{SO}(3)$-needlet convolutions are repeatedly conducted. Inside each needlet convolution, we use a wavelet shrinkage to threshold small high-pass coefficients, following a pooling operator to downsize the representation. The final output of $\mathrm{SO}(3)$-needlet convolution is handled by the inverse NFFT (iNFFT) to feed into a downstream predictor.

3 Experiments

The main advantage of our model is the property of equivariance to $\mathrm{SO}(3)$ transforms with multi-scale representation for complicated real-world application. This section validates the model with three experiments. Our models are trained on 24G NVIDIA GeForce RTX 3090 Ti GPUs. The hyperparameters are obtained by grid search. Adam Kingma & Ba (2014) is used as our optimizer.

3.1 Local MNIST Classification

Dataset The first experiment evaluates the effectiveness of the needlet convolution neural network in capturing high-frequency information. We follow Cohen et al. (2018) and use a modified spherical MNIST classification dataset, where the images are projected onto a sphere to establish rotated training and test sets. Here the samples of the training set are all rotated by the same rotation while those in the test set are rotated by another rotation. We downscale the original MNIST images into five different resolutions and then project them onto a scalable area of the sphere.

Setup Our model is compared with Spherical CNN Cohen et al. (2018). We adopt the same architecture $\mathbb{S}^2$conv-ReLU-$\mathrm{SO}(3)$conv-ReLU-FC-softmax, with bandwidth L = 30, 10, 6 and k = 20, 40, 10 channels: when it comes to our model, we replace $\mathbb{S}^2$conv and $\mathrm{SO}(3)$conv with $\mathbb{S}^2$-needlet convolution and $\mathrm{SO}(3)$-needlet convolution, bandwidth L = 30, 10, 6 and k = 20, 40, 10 channels. We select the batch size of 64 and learning rate 0.001.

Results The test accuracy for spherical MNIST is presented in Table 1. To test the rotation equivariance of the models, we rotate the training dataset and test dataset with two different rotations. That is, the input training data are all rotated with a same rotation in SO(3), and all test data are rotated by another rotation. We also test on downscaled datasets with various scales: the higher the scale, the less size of the spherical digit is on the sphere, and the signal is more localized. Table 1 indicates that both models keep high test accuracy with both training and test data rotated. We can observe that our model consistently achieves high accuracy on datasets for different downscale ratios.

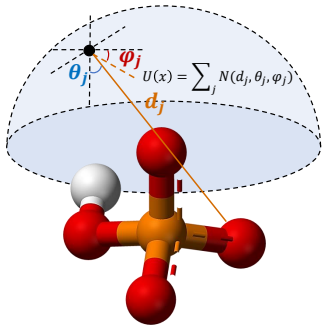

Figure 3: An illustration for computing the spherical signal of a molecule. We aggregate the information of each atom in the molecule with relative distance, polar angle and azimuthal angle.

Method	RMSE	Params
MLP/SORTED CM	16.06	-
MLP/RANDOM CM	5.96	-
GCN	7.32 ± 0.23	0.8M
SPHERICAL CNN†	8.47	1.4M
CLEBSCH–GORDAN†	7.97	$\geq$1.1M
NES† (Ours)	**7.21** $\pm$ 0.46	0.9M

Table 2: Test RMSE of atomic energy on **QM7**. † indicates the method is rotation equivariant.

Especially for the high ratio, the digit is concentrated at a small region, and the model is required to capture more details of the spherical data. In contrast, Spherical CNN has poorer performance with higher downscale. It demonstrates a reliable performance of our model in effectively distilling detailed and local features while maintaining rotation equivariance of the needlet convolutional layer.

3.2 MOLECULAR PROPERTY PREDICTION

Molecule	sGDML	SchNet	DimeNet	SphereNet	NES
Aspirin	29.5	58.5	21.6	18.6	**15.2**
Ethanol	14.3	16.9	10.0	**9.0**	9.2
Malonaldehyde	17.8	28.6	16.6	14.7	**13.6**
Naphthalene	4.8	25.2	9.3	7.7	**3.5**
Salicylic	**12.1**	36.9	16.2	15.6	14.2
Toluene	**6.1**	24.7	9.4	6.7	**6.1**
Uracil	**10.4**	24.3	13.1	11.6	10.8

Table 3: Test MAE of forces in $\mathrm{meV}/\mathring{A}$ on **MD17**.

Datasets The second experiment predicts molecular property over two widely used datasets (**QM7** and **MD17** Chmiela et al. (2017)) to evaluate the model's expressivity to bio-molecular simulation. **QM7** contains $7,165$ molecules. Each molecule contains at most $N = 23$ atoms of $T = 5$ types (H, C, N, O, S), which is to regress over the atomic energy of molecules given the corresponding position p_i and charges z_i of each atom i. **MD17** predicts the energies and forces at the atomic level for several organic molecules with up to 21 atoms and four chemical elements, using the molecular dynamics trajectories.

We follow Rupp et al. (2012) to generate spherical signals for every molecule. We define a sphere S_i centered at p_i for each atom i and define the potential functions as $U_z(x) = \sum_{j \neq i, z_j = z} \frac{z_i \cdot z}{\|x - p_i\|}$, where z is the charge of the atom, and x is taken from $\mathbb{S}^2$. For every molecule, N spherical signals

are produced in T channels. We use the Gauss-Legendre rule to discretize the continuous functions on the sphere with $L = 20$ and create a sparse $N \times T \times (2L + 1) \times (L + 1)$ tensor as the input signal representation. For **QM7**, we generalize the *Coulomb matrix* ($C \in \mathbb{R}^{N\times N}$) and obtain 23 spherical signals for every molecule. For **MD17**, we create N spherical signals that are centered at the positions of each atom for every sample, where N is the number of atoms in the molecule. For the atom i, we define a corresponding spherical signal $U_i(x)$, where x is taken from the sphere by the Gauss-Legendre sampling method. The relative position of each atom to x is calculated with the absolute Cartesian coordinates of atoms provided by **MD17**. The spherical signal U_i is defined as $U_i(x) = \sum_{j=1}^{N} \mathcal{N}(d_j, \theta_j, \varphi_j)$, where $(d_j, \theta_j, \varphi_j)$ is the position of atom j relative to x. The d_j, θ_j, and φ_j denote the radial distance, polar angle, and the azimuthal angle respectively (see Figure 3). Different with **QM7**, **MD17** does not have a *Coulomb matrix*. The number of spherical signals N can thus be taken from a neural network or a mathematical operator to extract effective features with the relative positions. Here we choose the first approach of neural networks to adaptively learn feature. We fine-tune the hyperparameters individually for every type of molecules on the validation sets with $1,000$ samples for each type.

Setup The bandwidth L is from $20, 20, 10, 10, 5$ to 5 in the final block and the feature dimension is from $5, 5, 8, 16, 32$ to 64. The hyperparameter σ is taken as 0.001 for shrinkage. We run 10 epochs for **QM7** with a batch size of 32 and a learning rate of $5e - 4$. For **MD17**, we choose a batch size of 32 and a learning rate of $2e - 4$. We run the model for $1,000$ epochs.

Results We report the experimental results of **QM7** and **MD17** respectively in Tables 2-3. For **QM7**, we compare the root mean square error (RMSE) of our proposed NES with MLP/Random CM, MLP/Sorted CM Montavon et al. (2012), GCN Kipf & Welling (2017), Spherical CNN and Clebsch-Gordan Net Kondor et al. (2018). The scores are averaged over 10 trials with standard deviation. Our model uses approximately 0.9 million parameters to achieve the lowest RMSE at 7.21 ± 0.46 among all rotation equivariant models. Our model enjoys the advantages of both a smaller number of parameters and a lower prediction error, owing to the incorporation of efficient computation and multiscale analysis architecture. In **MD17** task, we focus on atomic forces and measure the mean absolute error (MAE) averaged over all samples and atoms. SchNet Schütt et al. (2017) and DimeNet Chmiela et al. (2018) are 3D graph models that incorporate relative distance information. SphereNet Liu et al. (2021) is a 3D graph model with physically-based representations of geometric information. Most of previous state-of-the-art models are graph-based models with hand-engineered features or expert knowledge. Instead, our model utilizes the adaptive learning of input features and incorporates multiscale analysis to improve the representation ability. Results show that the proposed model outperforms baseline models with strong performance and better generalization in molecular simulation, due to the rotation equivariance. NES achieves better performance on four types of molecules. Compared to NES, sGDML Chmiela et al. (2018) is a kind of kernel method, which relies on human expertise and extra annotation, thus suffering from poor generalization to a new type of molecule.

3.3 Delensing Cosmic Microwave Background

The existence of a stochastic Primordial Gravitational Wave Background (PGWB) is a common prediction in the majority of inflationary models. It is formed when microscopic quantum fluctuations of the metric were stretched up to super-horizon scales by the sudden expansion of space-time that occurred during inflation Caprini & Figueroa (2018). Since it has been able to free-stream from time as early as (possibly) the Planck time, PGWB has the potential of becoming one of the most powerful cosmological probes. The information about phase transitions and particle creation/annihilation may have taken place in the early universe, which allows new independent measurements of cosmological parameters. In order to discover PGWB, we need to constrain some parameters, such as the ratio between tensor and scalar perturbations $r = \mathcal{P}_t(k)/\mathcal{P}_s(k)$. Such a parameter relies on a high signal-to-noise ratio (SNR) reconstruction of the lensing potential, i.e., the projected weighted gravitational potential along the line-of-sight between us and the CMB. Photons in the CMB are deflected by the intervening mass distributions when they travel to us. The lensing effect distorts the recombination of the CMB and interferes with our ability to constrain early universe physics. Therefore, removing the lensing effect from observed data is critical to decoding early-universe physics. In this experiment, we use NES convolution to reconstruct the unlensed B-mode

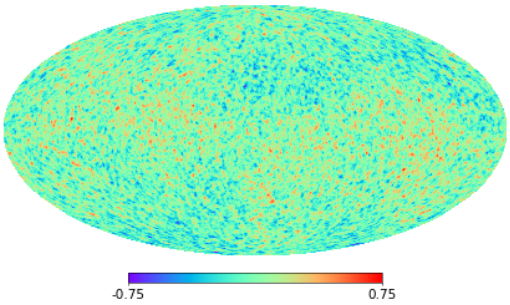

Figure 4: B-CMB multipoles unlensed map with tensor-to-scalar ratio $r = 0.2$, which is one of the main constraints in detecting the Primordial Gravitational Wave Background. We color the map with the intensity values to predict.

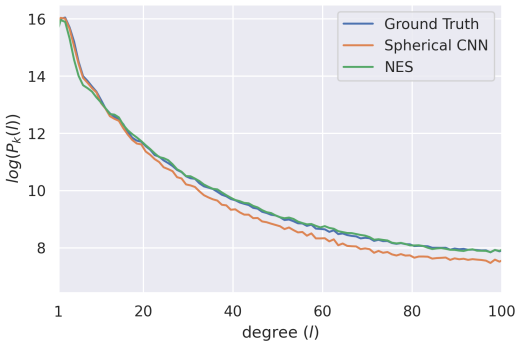

Figure 5: The power spectrum of unlensed B-map predicted by NES compared with Spherical CNNs and the ground truth.

(Figure 4) component of the CMB polarization from the lensed Q, U maps that are orthonormal bases corresponding to Stokes parameters.

Dataset Spherical CNN and NES are used to reconstruct the unlensed B map from the lensed Q, U maps. We simulate $10,000$ lensed Q, U maps and B-CMB multipoles unlensed map with $N_{\text{Side}} = 64$. Then, transforming the original sample rules from HEALPix to Gauss-Legendre tensor product rules with the bandwidth $L = 128$ by taking the average of the four nearest HEALPix points in Gauss-Legendre coordinates. We split the whole dataset into 80% training, 10% validation, and 10% test sets.

Setup We follow the U-Net architecture from Caldeira et al. (2019) and replace the standard image convolution with Spherical CNN and NES. In the encoder, the bandwidth is $128, 64, 32, 16$ and the feature dimensions are $2, 16, 32, 64$ respectively in each block. In the decoder, the bandwidth increases from $16, 32, 64$ to 128. The encoder layers are skip-connected to decoder layers, which is consistent with the standard architecture of U-Net. We sum the mean squared error (MSE) in the pixel domain and in power spectrum as the loss function. We use a batch size of 16, a learning rate of $5e-5$, and weight decay of $3e-4$ to train the model.

Result Figure 5 compares the power spectrum of two models' predicted B-unlensed map with ground truth. We can see Spherical CNNs always underestimates the ground truth when degree l is larger than 20. NES can capture more high-frequency information of data in each block during training. The estimated power spectrum of the model is consistent with the ground truth even at a large degree of $l \geq 100$.

4 Conclusion

We develop a Needlet approximate Equivariance Spherical CNN using multiscale representation systems on the sphere and rotation group. The needlet convolution inherits the multiresolution analysis ability from needlet transforms and allows rotation invariance in network propagation. Wavelet shrinkage is used as a network activation to filter out the high-pass redundancy, which helps improve the robustness of the network. The shrinkage brings controllable equivariance error for the needlet CNN, which is small when the scale is high. Empirical study shows the proposed model can achieve excellent performance on real scientific problems.

ACKNOWLEDGMENTS

YW acknowledges support from the Shanghai Municipal Science and Technology Major Project (2021SHZDZX0102) and Huawei-SJTU ExploreX Funding (SD6040004/034). We are also grateful to the anonymous reviewers for their feedback.

REFERENCES

Yashar Akrami, Frederico Arroja, M Ashdown, J Aumont, Carlo Baccigalupi, M Ballardini, Anthony J Banday, RB Barreiro, Nicola Bartolo, S Basak, et al. Planck 2018 results-x. constraints on inflation. *Astronomy & Astrophysics*, 641:A10, 2020.

Kenneth Atz, Francesca Grisoni, and Gisbert Schneider. Geometric deep learning on molecular representations. *Nature Machine Intelligence*, pp. 1–10, 2021.

Minkyung Baek, Frank DiMaio, Ivan Anishchenko, Justas Dauparas, Sergey Ovchinnikov, Gyu Rie Lee, Jue Wang, Qian Cong, Lisa N Kinch, R Dustin Schaeffer, et al. Accurate prediction of protein structures and interactions using a 3-track network. *Science*, 373:871–876, 2021.

Paolo Baldi, Gérard Kerkyacharian, Domenico Marinucci, Dominique Picard, et al. Asymptotics for spherical needlets. *The Annals of Statistics*, 37(3):1150–1171, 2009.

Wouter Boomsma and Jes Frellsen. Spherical convolutions and their application in molecular modelling. *Advances in neural information processing systems*, 30, 2017.

Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, Groups, Graphs, Geodesics, and Gauges. *arXiv preprint arXiv:2104.13478*, 2021.

João Caldeira, WL Kimmy Wu, Brian Nord, Camille Avestruz, Shubhendu Trivedi, and Kyle T Story. Deepcmb: Lensing reconstruction of the cosmic microwave background with deep neural networks. *Astronomy and Computing*, 28:100307, 2019.

Chiara Caprini and Daniel G Figueroa. Cosmological backgrounds of gravitational waves. *Classical and Quantum Gravity*, 35(16):163001, 2018.

Stefan Chmiela, Alexandre Tkatchenko, Huziel E Sauceda, Igor Poltavsky, Kristof T Schütt, and Klaus-Robert Müller. Machine learning of accurate energy-conserving molecular force fields. *Science Advances*, 3(5):e1603015, 2017.

Stefan Chmiela, Huziel E Sauceda, Klaus-Robert Müller, and Alexandre Tkatchenko. Towards exact molecular dynamics simulations with machine-learned force fields. *Nature Communications*, 9(1): 1–10, 2018.

Taco S Cohen, Mario Geiger, Jonas Köhler, and Max Welling. Spherical CNNs. In *ICLR*, 2018.

Benjamin Coors, Alexandru Paul Condurache, and Andreas Geiger. Spherenet: Learning spherical representations for detection and classification in omnidirectional images. In *Proceedings of the European conference on computer vision (ECCV)*, pp. 518–533, 2018.

Alex Davies, Petar Veličković, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomašev, Richard Tanburn, Peter Battaglia, Charles Blundell, András Juhász, et al. Advancing mathematics by guiding human intuition with AI. *Nature*, 600(7887):70–74, 2021.

David L Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3): 613–627, 1995.

Bin Han. *Framelets and wavelets – Algorithms, Analysis, and Applications*. Applied and Numerical Harmonic Analysis. Birkhäuser XXXIII Cham. Springer, 2017.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *ICLR*, 2014.

Thomas N. Kipf and Max Welling. Semi-Supervised Classification with Graph Convolutional Networks. In *ICLR*, 2017.

Risi Kondor, Zhen Lin, and Shubhendu Trivedi. Clebsch–Gordan Nets: a fully Fourier space spherical convolutional neural network. In *NeurIPS*, volume 31, pp. 10117–10126, 2018.

Yi Liu, Limei Wang, Meng Liu, Xuan Zhang, Bora Oztekin, and Shuiwang Ji. Spherical message passing for 3D graph networks. *arXiv preprint arXiv:2102.05013*, 2021.

D Marinucci, D Pietrobon, A Balbi, P Baldi, P Cabella, G Kerkyacharian, Paolo Natoli, D Picard, and N Vittorio. Spherical needlets for cosmic microwave background data analysis. *Monthly Notices of the Royal Astronomical Society*, 383(2):539–545, 2008.

Oscar Méndez-Lucio, Mazen Ahmad, Ehecatl Antonio del Rio-Chanona, and Jörg Kurt Wegner. A geometric deep learning approach to predict binding conformations of bioactive molecules. *Nature Machine Intelligence*, 3(12):1033–1039, 2021.

Grégoire Montavon, Katja Hansen, Siamac Fazli, Matthias Rupp, Franziska Biegler, Andreas Ziehe, Alexandre Tkatchenko, Anatole Lilienfeld, and Klaus-Robert Müller. Learning invariant representations of molecules for atomization energy prediction. *NeurIPS*, 25:440–448, 2012.

F Narcowich, P Petrushev, and J Ward. Decomposition of Besov and Triebel–Lizorkin spaces on the sphere. *Journal of Functional Analysis*, 238(2):530–564, 2006a.

Francis J Narcowich, Pencho Petrushev, and Joseph D Ward. Localized tight frames on spheres. *SIAM Journal on Mathematical Analysis*, 38(2):574–594, 2006b.

Matthias Rupp, Alexandre Tkatchenko, Klaus-Robert Müller, and O Anatole Von Lilienfeld. Fast and accurate modeling of molecular atomization energies with machine learning. *Physical Review Letters*, 108(5):058301, 2012.

Kristof T Schütt, Pieter-Jan Kindermans, Huziel E Sauceda, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. Schnet: A continuous-filter convolutional neural network for modeling quantum interactions. In *NeurIPS*, 2017.

Raphael JL Townshend, Stephan Eismann, Andrew M Watkins, Ramya Rangan, Maria Karelina, Rhiju Das, and Ron O Dror. Geometric deep learning of RNA structure. *Science*, 373(6558): 1047–1051, 2021.

Yu Guang Wang and Xiaosheng Zhuang. Tight framelets and fast framelet filter bank transforms on manifolds. *Applied and Computational Harmonic Analysis*, 48(1):64–95, 2020.

Yu Guang Wang, Quoc T Le Gia, Ian H Sloan, and Robert S Womersley. Fully discrete needlet approximation on the sphere. *Applied and Computational Harmonic Analysis*, 43(2):292–316, 2017.

Kai Yi, Yi Guo, Yanan Fan, Jan Hamann, and Yu Guang Wang. CosmoVAE: Variational autoencoder for cmb image inpainting. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7. IEEE, 2020.

Xuebin Zheng, Bingxin Zhou, Junbin Gao, Yu Guang Wang, Pietro Liò, Ming Li, and Guido Montúfar. How framelets enhance graph neural networks. In *ICML*, 2021.