Fiber Orientation Distribution Function Estimation by Spherical Needlets

Jie Peng

Department of Statistics, University of California, Davis

February 2016
Diffusion MRI

- Diffusion MRI is a magnetic resonance imaging technology which measures *diffusion* of water molecules along a set of (predetermined) directions.
  - In vivo, non-invasive, no radiation.
- Diffusion MRI uses water diffusion as a proxy to probe the anatomy of biological tissues.
- Raw data from a D-MRI experiment:
  - Multiple grey scale images corresponding to multiple gradient directions plus a few images with no diffusion weighting.
  - Each image consists of intensities for pixels on a 3D grid (e.g., \( \sim 256 \times 256 \times 59 \) for a human brain).
  - Image resolution: \( 1 \sim 3 \text{mm} \).
D-MRI Provides Information on Brain Connectivity

- Neuron axons with similar destinations form big bundles called *white matter fiber tracts*.
- When applied to human brains, diffusion MRI reveals detailed anatomy of white matter tracts such as their location, size, shape and how they are connected to each other.
  - Human connectome project
Tractographic Reconstruction of Neural Connections


- With diffusion information at each voxel, fiber tracts can be reconstructed using computer-aided 3D tracking techniques called tractography.
Clinical Applications of Diffusion MRI

- Detect brain abnormality in white matter regions such as specific axonal loss, deformation in brain tumors.
- Differentiate types of tumor and growth orientation.
- Measure anatomy of immature brains.
- Monitor status of specific white matter tracts.
Water Diffusion in Biological Tissues

*Anisotropic* due to the presence of fibers with coherent orientations.

- Water tends to diffuse faster along fibers.
- Information on water diffusion may be used to probe tissue structure.
Diffusion in Brain

- **White matter.**
  - Astronomical number of connections: “cables” of the brain.
  - Presence of axonal bundles at image resolution \(\Rightarrow\) diffusion appears anisotropic.

- **Grey matter.**
  - \(\sim\) 100 billion neurons: “CPU” of the brain.
  - Lack of coherent fiber organization at image resolution (\(\sim\)2mm) \(\Rightarrow\) diffusion appears isotropic.
Sensitize MRI Signal by Water Diffusion

Mori and Zhang, 2006, Neuron 51
Diffusion Weighted Signals

At voxel $s$, along direction $q$, diffusion weighted signal:

$$S_0(s) \int_{R^3} p_{s,\Delta t}(r) \exp(i\gamma\delta q \cdot r)dr.$$ 

- $S_0(s)$: signal intensity without diffusion weighting at voxel $s$.
- $p_{s,\Delta t}(\cdot)$: p.d.f. of water displacement in time duration $\Delta t$ at voxel $s$.
- $\Delta t$: time between “dephasing” and “rephasing”.
- $\delta$: duration of dephasing/rephasing.
- $\gamma$: gyromagnetic ratio.

Diffusion weighted signal is the inverse Fourier transform of the diffusion probability density function.
Gaussian Diffusion and Single Tensor Model

\[ p_{s,\Delta t}(r) = \frac{1}{(2\pi)^{3/2}} |D(s)\Delta t|^{-\frac{1}{2}} \exp \left( -\frac{r^\top D(s)^{-1}r}{2\Delta t} \right), \quad r \in \mathbb{R}^3. \]

- **D(s): diffusion tensor**, a 3 \times 3 p.d. matrix. Its principal eigenvector captures the fiber orientation within the voxel.
- **DWI signal along gradient direction u:**

\[ S(u) = S_0 \exp(-bu^\top Du), \quad b = \frac{\gamma^2 \delta^2 \|q\|^2 \Delta}{2}. \]

- **D(s) can be recovered with as few as 6 distinct gradient directions.**
Crossing Fibers

More than 30% voxels have multiple fiber bundles with distinct orientations (under D-MRI image resolution).
Limitations of Single Tensor Model

- Single tensor model can not resolve multiple fiber orientations within a voxel since a tensor only has one principal direction. It may incorrectly lead to:
  - Oblate ($\lambda_1 \approx \lambda_2 >> \lambda_3$) tensor estimation.
  - Low anisotropy and random diffusion directions.
  - Consequently, early termination of fiber tracking or bias/switching of fiber tracking.

- Tensor Mixture Model.

- **Nonparametric methods using HARDI data.**
HARDI Techniques

High angular resolution diffusion imaging (HARDI) techniques enable the detection of multi-modal diffusion signals.

- Q-ball imaging: Gradients are sampled from a single spherical shell of a particular radius (a single bvalue).
  - Diffusion orientation distribution function (ODF) (Tuch, 2004, Descoteaux et al., 2007).
  - Fiber orientation density (FOD) function (Tournier et al., 2004, 2007).
Fiber Orientation Density Function

FOD is a symmetric p.d.f. on $\mathbb{S}^2$ which describes the distribution of fiber orientations (corresponding to coherently oriented fiber bundles) at a voxel.

- Example. $K$ distinct fiber bundles:

\[
F(\theta, \phi) = \sum_{k=1}^{K} w_k \delta_{\theta_k, \phi_k}(\theta, \phi), \quad \theta \in [0, \pi], \quad \phi \in [0, 2\pi),
\]

where $w_k > 0$, $\sum_{k=1}^{K} w_k = 1$ are the volume fractions and $\theta_k$ (polar angle) and $\phi_k$ (azimuthal angle) are the spherical coordinates of the $k$-th fiber direction.
Assumptions

- DWI signals are the summation of signals originated from distinct fiber bundles.
  - No water exchange between distinct fiber bundles.
  - No water exchange between orientationally distinct segments of the same fiber bundle.
- Diffusion characteristics along all fiber bundles are (i) identical no matter the direction or abundance of the fiber bundle, and (ii) axially symmetric around the fiber direction.
  - DWI signal from a single coherently oriented fiber bundle can be represented by an axially symmetric response function.
  - The response function is identical across fiber bundles.
Response Function

- An axially symmetric kernel

\[ R : [-1, 1] \rightarrow \mathbb{R} \]

which describes DWI signal resulting from water diffusion along a single fiber bundle aligned with the z-axis.

- Estimation of the response function. Assume response function is identical across voxels.
  - Fit the single tensor model to every voxel.
  - Identify voxels with high FA values and find their principal eigenvectors.
  - For each such voxel, rotate the DWI signals such that the principal eigenvectors are aligned with the z-axis.
  - Average the rotated DWI profiles across these voxels.
• Example. Gaussian diffusion with $\lambda_1 = \lambda_2 < \lambda_3$:

$$R(\cos(\theta)) = S_0 \exp^{-b \mathbf{q}(\theta, \phi)^T \Lambda \mathbf{q}(\theta, \phi)} = S_0 \exp^{-b(\lambda_1 \sin^2 \theta + \lambda_3 \cos^2 \theta)},$$

where $\theta \in [0, \pi]$, $\mathbf{q}(\theta, \phi) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)^T$ and $\Lambda = \text{Diag}(\lambda_1, \lambda_2, \lambda_3)$.

• Since water diffuses fastest along the dominant fiber direction, the response function is attenuated the most along the z-axis.

Figure: Gaussian diffusion response function with $\lambda_1 = \lambda_2 = 20, \lambda_3 = 1000.$
Spherical Convolution Model of Diffusion Signals

- DWI signal $S(\cdot)$ is the *spherical convolution* of the response function $R(\cdot)$ with the FOD $F(\cdot)$:

$$S(x) = R \ast F(x) = \int_{S^2} R(x^T y) F(y) dy, \quad x \in S^2.$$ 

- The FOD $F(\cdot)$ can be obtained by performing the *spherical deconvolution* of the response function $R(\cdot)$ from the DWI signal function $S(\cdot)$.

- Spherical deconvolution can be achieved through *spherical harmonic representation*. 
Connection between the FOD model and the multi-tensor model.

- In the multi-tensor model, if the tensors $D_k$'s have the same set of eigenvalues satisfying $\lambda_1 = \lambda_2 < \lambda_3$, then it corresponds to the FOD model with:
  - Response function
    \[
    R(\cos(\theta)) = S_0 \exp^{-b(\lambda_1 \sin^2 \theta + \lambda_3 \cos^2 \theta)}, \quad \theta \in [0, \pi],
    \]
  - FOD
    \[
    F(\theta, \phi) = \sum_{k=1}^{K} w_k \delta_{\theta_k, \phi_k}(\theta, \phi), \quad \theta \in [0, \pi], \quad \phi \in [0, 2\pi),
    \]

where $(\theta_k, \phi_k)$ denotes the principal eigenvector of the tensor $D_k$ ($k = 1, \ldots, K$).
**Figure:** *Spherical Convolution*. Left to right: response function $R(\cdot)$, FOD $F(\cdot)$ and the DWI signal $S(\cdot)$. 
Spherical Harmonics

- Spherical harmonics, denoted by $\tilde{\Phi}_{l,m}$:

$$
\tilde{\Phi}_{l,m}(\theta, \phi) = \sqrt{\frac{2l + 1}{4\pi} \frac{(l - m)!}{(l + m)!}} P^m_l(\cos(\theta)) \exp^{im\phi}, \ \theta \in [0, \pi], \ \phi \in [0, 2\pi).
$$

- $l(\geq 0)$ denotes the harmonic order and $m (-l \leq m \leq l)$ denotes the phase factor.
- Angular frequency increases with order $l$. Harmonics with even $l$ are symmetric and those with odd $l$ are anti-symmetric.
- $P^m_l$ is an associated Legendre polynomial of order $(l, m)$. 
Real Symmetric Harmonic Basis

- For $l = 0, 2, \ldots, l_{\text{max}}$ and $m = -l, \ldots, 0, \ldots, l$

$$\Phi_{l,m} = \begin{cases} \frac{\sqrt{2}}{2}(\tilde{\Phi}_{l,m} + (-1)^m\tilde{\Phi}_{l,-m}) & \text{if } 0 < m \leq l \\ \Phi^0_l & \text{if } m = 0 \\ \frac{\sqrt{2}}{2i}((-1)^{m+1}\tilde{\Phi}_{l,m} + \tilde{\Phi}_{l,-m}) & \text{if } -l \leq m < 0 \end{cases}$$

- $\Phi_{l,m}$ form an orthonormal basis for real symmetric square-integrable functions (including $R$ and $F$) defined on $\mathbb{S}^2$.

Figure: Real symmetric spherical harmonics.
Since $S = R \star F$, so

$$s_{lm} = \sqrt{\frac{4\pi}{2l + 1}} r_l f_{lm}, \quad l = 0, 2, 4, \ldots , m = -l, \ldots , 0, \ldots l,$$

where $s_{lm} = \langle S, \Phi_{l,m} \rangle$, $r_l = \langle R, \Phi_{l,0} \rangle$ and $f_{lm} = \langle F, \Phi_{l,m} \rangle$ are the spherical harmonics (rotational harmonics) coefficients of $S$, $R$ and $F$, respectively.

- Assume that $S(\cdot), R(\cdot), F(\cdot)$ can be represented by finite-order spherical harmonics functions $\{ \Phi_{l,m} : -l \leq m \leq l \}_{l=0,2,\ldots,l_{\text{max}}}$. 
- The number of SH functions: $L = (l_{\text{max}} + 1)(l_{\text{max}} + 2)/2$. 
Regression Model For DWI Measurements

The observed D-MRI signals:

\[ y = \Phi \mathbf{R} \mathbf{f} + \epsilon. \]

- \( y = (y(\theta_1, \phi_1), \ldots, y(\theta_n, \phi_n))^T \) is the \( n \times 1 \) vector of observed DWI measurements.
- \( \Phi \) is the \( n \times L \) matrix of the SH functions evaluations at the \( n \) gradient directions \( \{(\theta_i, \phi_i)\}_{i=1}^n \).
- \( \mathbf{R} \) is an \( L \times L \) diagonal matrix with diagonal elements \( \sqrt{4\pi/(2l+1)}r_l \) (SH coefficients of the response function) in blocks of size \( 2l + 1 \) for \( l = 0, 2, \ldots, l_{\text{max}} \).
- \( \mathbf{f} \) is the \( L \times 1 \) vector of SH coefficients of the FOD \( F \).
SH-ridge Estimator of FOD

- Penalized regression:

\[
\hat{f} = \arg \min_f \| y - \Phi R f \|^2_2 + \lambda \mathbb{E}(F), \quad F = \sum_{l,m} f_{l,m} \Phi_{l,m}
\]

- Laplace-Beltrami penalty, \( \mathbb{E}(F) \), a measure of roughness.

\[
\mathbb{E}(F) := \int_{\Omega} (\Delta_b F)^2 d\Omega = \mathbf{f}^T P \mathbf{f},
\]

where \( P \) is a diagonal matrix with entries \( l^2(l+1)^2 \) in blocks of size \( 2l + 1 \).

- SH-ridge estimator:

\[
\hat{f}^{LB} = (R^T \Phi^T \Phi R + \lambda P)^{-1} R^T \Phi^T y, \quad \hat{F}^{LB} = \sum_{l,m} \hat{f}^{LB}_{l,m} \Phi_{l,m}.
\]

* Tournier et al. (2004).*
sCSD Sharpening

Suppress small values of the estimated FOD and sharpens the peak(s) of the FOD estimator.

1. Initial step: Get an initial estimator \( \hat{f}_0 \) by SH-ridge.
2. At the \( k + 1 \) updating step

\[
\hat{f}_{k+1} = \arg\min_{f} \| y - \Phi Rf \|^{2} + \lambda \| P_k f \|^{2} \tag{1}
\]

where \( P_k \) is an \( n \times L \) matrix,

\[
P_{k,i,(l,m)} := \begin{cases} 
\Phi_{i,(l,m)} & \text{if } \hat{F}_{k,i} < \tau \\
0 & \text{if } \hat{F}_{k,i} > \tau 
\end{cases} \tag{2}
\]

where \( \hat{F}_{k,i} \) is the \( i \) the element of the vector \( \hat{F}_k = \Phi \hat{f}_k \)

\[ Tournier \ et \ al.(2007) \]
Spherical Needlets

- Spherical needlets $\psi_{j,k}$s are constructed from spherical harmonics functions (Nrcowich et al., 2006).
- Needlets are spatially localized with exponential concentration with respect to the frequency index $j$.
- Needlets provide sparse representations for spherical functions with small spatial scale features.

**Figure**: Real symmetric spherical needlets.
Needlets-lasso Estimator of FOD

- Assume that $S(\cdot)$, $R(\cdot)$, $F(\cdot)$ can be represented by finite-order spherical harmonics functions.
- Then they can be represented by finite needlets functions
  $\{\psi_{k,j} : k \in \chi_j\}_{j=0,\ldots,j_{\text{max}}}$.  
- The number of symmetric needlets: $N = 2^{2j_{\text{max}}+1} - 1$.
- There is a transition matrix $C$ such that the spherical harmonics coefficients
  \[ f = C\beta, \]
  where $\beta$ is the corresponding spherical needlets coefficients which are expected to be sparse.
- The observed D-MRI signals:
  \[ y = \Phi R C \beta + \epsilon. \]
\( \ell_1 \) penalized regression with nonnegativity constraints:

\[
\hat{\beta} = \arg \min_{\beta : \tilde{\Phi}C\beta \geq 0} \| y - \Phi R C \beta \|_2^2 + \lambda \| \beta \|_1 .
\]

- \( \lambda \) is a tuning parameter which controls the degree of sparsity.
- \( \hat{F}^{NL} = \tilde{\Phi}C\beta \geq 0 \) ensures that the estimated FOD is nonnegative on the evaluation grid.
- This is a constrained convex minimization problem and can be solved by the ADMM algorithm.
Simulation Setting

- FODs with two distinct fiber bundles:

\[ F(\theta, \phi) = w_1 \delta(\theta_1, \phi_1) + w_2 \delta(\theta_2, \phi_2). \]

- \( w_1 = w_2 = 0.5. \)
- Separation angle between the two fiber bundles:

\( \theta_{sep} = 45, 60, 75, 90 \) degrees.
- Gaussian diffusion response function with \( \frac{\lambda_3}{\lambda_1} = 50. \)
- \( n = 81 \) gradient directions, sampled from an equal angle grid.
- bvalue=1000s/mm\(^2\), 3000s/mm\(^2\).
- \( SNR = \frac{S_0}{\sigma} = 20, \) where \( S_0 \) is the \( b_0 \) image intensity and \( \sigma \) is the Rician noise standard deviation.

The number of gradient directions, SNR and bvalue are typical/close to those in real D-MRI studies.
Simulation Results

- FODs are estimated by SH-ridge with $l_{\text{max}} = 8 \ (L = 45 \ \text{SH functions})$, by sCSD sharpening, and by needlets-lasso with $j_{\text{max}} = 4 \ (N = 511 \ \text{needlets functions})$.
- Tuning parameters in SH-ridge and needlets-lasso are chosen by BIC. Those of sCSD are set as the recommended values by the paper.
- The needlets-lasso estimator has much sharper peaks.
- sCSD is sensitive to the penalty parameter $\tau$ as well as $l_{\text{max}}$ and thus is hard to automate in the real data setting where there are hundreds of thousands voxels with different fiber population characteristics.
Figure: Mean plus 2-SD plots across 100 replicates. $bvalue=1000$. Top: $\theta_{sep} = 45$; Middle: $\theta_{sep} = 60$; Bottom: $\theta_{sep} = 90$. Left: SH-ridge; Middle: sCSD; Right: needlets-lasso.
**Figure:** Mean plus 2-SD plots across 100 replicates. $b_{\text{value}}=3000$. $\theta_{\text{sep}} = 45$. Left: SH-ridge; Middle: sCSD; Right: needlets-lasso.
Three fiber bundles simulation.

Figure: Mean plus 2-SD plots across 100 replicates. $b$-value=1000. Top: $\theta_{sep} = 75$; Bottom: $\theta_{sep} = 90$. Left: SH-ridge; Middle: sCSD; Right: needlets-lasso.
Discussion

- Borrow information from neighboring voxels.
- Inference based on bootstrapping.
- Feature extraction and multiscale analysis.
Sensitize MRI Signal by Water Diffusion

- **Excitation**: Apply a strong homogeneous field $\Rightarrow$ water molecules resonate at the same frequency and phase.
- **Dephasing**: Apply a linearly inhomogeneous gradient field $\Rightarrow$ water molecules resonate at different frequencies depending on their locations.
- Apply the homogeneous field $\Rightarrow$ water molecules resonate at the same frequency again, but signal is still out of phase.
- **Rephasing**: Apply an opposite gradient field.
  - If water molecules had moved along the gradient direction, then there would be a disruption of phase $\Rightarrow$ signal loss and signals are *diffusion weighted*. 
MRI can not measure the phase of individual water molecules, but it can detect imperfect rephasing through signal loss.

- During MR measurements, the amount of water molecular displacement is about $1 \sim 20 \mu m$, depending on sample, temperature, duration of experiment, etc.
- DT-MRI experiments are designed such that this amount of diffusion leads to $10 \sim 90\%$ signal loss.

*Mori and Zhang, 2006, Neuron 51*
Diffusion Weighted Signals

- “Dephasing” encodes the location information of water molecules through their signal phase.
- MR can not measure the phase of individual water molecules, but it can detect imperfect rephasing through signal loss.
  - Perfect rephasing only happens when water molecules remain stationary between the two applications of gradients.
  - If water moved between “dephasing” and “rephasing”, there will be a disruption of phase across the sample.
  - Then after rephasing, some of the molecules that moved will have different phases from the stationary molecules. This leads to an overall signal attenuation – signals are diffusion weighted.
- Water motion along directions perpendicular to the gradient direction will not cause signal loss and thus can not be detected.
  - Multiple gradient directions need to be applied if water diffuse anisotropically.
Fig. 1.9 An example of a dephase–rephase experiment by gradient application. Red, green, and blue circles indicate three water molecules located at different positions in a sample tube. Thick arrows indicate the strengths of magnetic field strength ($B_0$), and narrow arrows indicate phases of MR signals from each molecule.
Model Diffusion Signals

- Signal loss equals to the summation (across locations within a voxel) of the sinusoid waves with shifted signal phases weighted by the proton density at their corresponding location.

- Applying a gradient field $\mathbf{q}$ with duration $\delta$ introduces a phase shift in space:
  \[ \gamma \delta \mathbf{q} \cdot \mathbf{r}, \quad \mathbf{r} \in \mathbb{R}^3 \]
  which is proportional to the projected distance on $\mathbf{q}$.

  - $\mathbf{q} \in \mathbb{R}^3$: gradient field: $||\mathbf{q}||$ – field strength, $\mathbf{u} = \mathbf{q}/||\mathbf{q}||$ – gradient direction.
  - $\mathbf{r} \in \mathbb{R}^3$: displacement vector.
  - $\delta$: duration of dephasing/rephasing (assumed to be short, so ignore water movement during $\delta$).
  - $\gamma$: gyromagnetic ratio.
• After the “dephasing” stage, water molecules start to move during the time period $\Delta t (\gg \delta)$. Their final locations are distributed according to a diffusion probability density function:

$$p_{\Delta t}(r), \quad r \in \mathbb{R}^3$$

– density of protons having a displacement $r$ in time duration $\Delta t$.

• $\Delta t$: time between “dephasing” and “rephasing”. Typically, the molecular displacement is $1 \sim 20 \mu m$.

• The probability of water molecules having displacements $r$ and $-r$ is the same:

$$p_{\Delta}(r) = p_{\Delta}(-r).$$
Microscopic vs. Macroscopic

- The average distance that water molecules move during the MR measurement is $1 \sim 20\mu m$.
- Thus only barriers that have a smaller dimension may cause anisotropy in water diffusion, this includes microscopic cellular architecture ($<10\mu m$) such as protein filaments, cell membranes, myelin sheaths.
- However, the image resolution is much coarser ($\sim 2mm$). Therefore information is averaged within each voxel.
- Thus diffusion MRI provides information on “macroscopic coherent arrangement of anisotropic microscopic anatomy” (Mori, 2007). Only when both factors exist in a voxel, one can observe diffusion anisotropy.
Spherical Needelts Construction

- Two main ideas: (i) discretization of the sphere by an exact quadrature formula; (ii) Littlewood-Paley decomposition.
- The quadrature formula discretizes the sphere into cubature points and cubature weights:

\[
\text{Theorem 1. Denote } \mathcal{H}_l \text{ as the space spanned by } \{Y_{lm} : m = -l, \ldots, l \}, \text{ and let } \mathcal{K}_l = \bigoplus_{k=0}^{l} \mathcal{H}_k. \text{ For any } l \in \mathbb{N}, \text{ there exist a finite subset } \mathcal{X}_l = \{\xi_{lk} : k = 1, \ldots, n_l\} \text{ of } S^2 \text{ and positive real numbers } \{\lambda_{lk} : k = 1, \ldots, n_l\} \text{ such that}
\]

\[
\int_{S^2} f(x) \, dx = \sum_{k=1}^{n_l} \lambda_{lk} f(\xi_{lk}),
\]

for any \( f \in \mathcal{K}_l \). Here \( \xi_{lk} \) and \( \lambda_{lk} \) are called cubature points and cubature weights, respectively.
Spherical Needlets Construction (Cont’d)

- Given a frequency $j \in \mathbb{N}_0$ and cubature points $\xi_{jk}$ and weights $\lambda_{jk}$, the spherical needlets with frequency $j$ ($B > 1$):

$$\psi_{jk}(x) = \sqrt{\lambda_{jk}} \sum_{l=\lfloor B^{j-1} \rfloor}^{B^{j+1}} b\left(\frac{l}{B^j}\right) \sum_{m=-l}^{l} \tilde{\Phi}_{l,m}(\xi_{jk}) \tilde{\Phi}_{l,m}(x), \quad x \in \mathbb{S}^2.$$

- $b(\cdot)$ is a window function satisfying: (i) $\text{supp}(b) = [1/B, B]$; (ii) $\sum_{j=0}^{\infty} b^2(t/B^j) = 1$, for any $t > 0$; (iii) $b \in C^M$ for some $M \geq 1$.

- $b(\cdot)$ decomposes the frequency domain into several overlapping intervals ($B^{j-1}, B^{j+1}$).

- $\xi_{jk}$ determines the location of the needlets $\psi_{jk}$ and $\lambda_{jk}$ determines to what extent $\psi_{jk}$ is localized.

- Varying $\xi_{jk}$ and $j$ is analogous to translation and dilation in multiscale analysis.
Spherical Needlets Properties

- Needlets are real-valued spherical functions.
- They are localized in the frequency domain since the window function has compact support.
- Needlets are spatially localized with exponential concentration with respect to the frequency index $j$.
- Spherical needlets provide sparse representation of spherical functions with sharp local peaks.
- Needlets (together with the first spherical harmonics $\tilde{\Phi}_{00}$) form a tight frame on $L^2(\mathbb{S}^2)$:
  \[
  \|f\|_{L^2}^2 = \sum_{j,k} |\langle f, \psi_{jk} \rangle|^2 + a_{00}^2.
  \]
- They are almost orthogonal: for $|j - j'| \geq 2$,
  \[
  \langle \psi_{jk}, \psi_{j'k'} \rangle = 0.
  \]
- The spherical needlets $\psi_{jk}$ and $\psi_{j'k'}$ are asymptotically uncorrelated as the frequency $j$ increases and the distance between them remains fixed.