Spherical wavelets for CMB temperature and polarization data analysis

Daryl Geller
Stony Brook University

mini-Workshop on Computational AstroStatistics
Harvard-Smithsonian Center for Astrophysics

August 24, 2010

Joint work with (in various combinations)

- Domenico Marinucci
- Frode Hansen
- Azita Mayeli
- Xiaohong Lan
- S. Scodeller
- O. Rudjord

CMB temperature and isotropy

Points on unit sphere $x = (\theta, \varphi)$, $0 \leq \theta \leq \pi \quad 0 \leq \varphi < 2\pi$
CMB temperature $T(x)$ is a random field, of which we make a *single* observation.
$ET(x) = 0$ for all x (after subtracting off a constant).

CMB temperature and isotropy

Points on unit sphere $x = (\theta, \varphi)$, $0 \leq \theta \leq \pi \quad 0 \leq \varphi < 2\pi$
CMB temperature $T(x)$ is a random field, of which we make a *single* observation.
$ET(x) = 0$ for all x (after subtracting off a constant).

Second-order isotropy condition:

$$E(T(Rx)T(Ry)) = E(T(x)T(y)) < \infty$$

for all x, y in the sphere, and all rotations R.
"Physics is the same in all directions."

CMB temperature and isotropy

Points on unit sphere $x = (\theta, \varphi)$, $0 \leq \theta \leq \pi \quad 0 \leq \varphi < 2\pi$
CMB temperature $T(x)$ is a random field, of which we make a *single* observation.
$ET(x) = 0$ for all x (after subtracting off a constant).

Second-order isotropy condition:

$$E(T(Rx)T(Ry)) = E(T(x)T(y)) < \infty$$

for all x, y in the sphere, and all rotations R.
"Physics is the same in all directions."

Note: If $\{T(x) : x \in S^2\}$ is a Gaussian family, second-order isotropy is equivalent to isotropy:
for any $x_1, \ldots, x_n$ on the sphere, the joint probability distribution of $T(Rx_1), \ldots, T(Rx_n)$ is independent of the rotation R.

Standard way to determine angular power spectrum

Expand $T(x)$ in spherical harmonics (analogue of Fourier series for the sphere):

$$T(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm} Y_{lm}(x)$$

where now the A_{lm} are also random variables with mean 0.
It is crucial to note that there are $2l+1$ orthonormal spherical harmonics for each degree l.

Standard way to determine angular power spectrum

Expand $T(x)$ in spherical harmonics (analogue of Fourier series for the sphere):

$$T(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm} Y_{lm}(x)$$

where now the A_{lm} are also random variables with mean 0.
It is crucial to note that there are $2l+1$ orthonormal spherical harmonics for each degree l.

Consequence of second-order isotropy: for each l,

$$C_l := E(|A_{lm}|^2)$$

is *independent* of m. Moreover the A_{lm} are all uncorrelated.

Standard way to determine angular power spectrum

Expand $T(x)$ in spherical harmonics (analogue of Fourier series for the sphere):

$$T(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm} Y_{lm}(x)$$

where now the A_{lm} are also random variables with mean 0.
It is crucial to note that there are $2l + 1$ orthonormal spherical harmonics for each degree l.

Consequence of second-order isotropy: for each l,

$$C_l := E(|A_{lm}|^2)$$

is *independent* of m. Moreover the A_{lm} are all uncorrelated.

Law of large numbers: as $l \to \infty$,

$$C_l \sim \frac{1}{2l+1} \sum_{m=-l}^{l} |A_{lm}|^2$$

Angular power spectrum

The C_l, the *angular power spectrum*, are used to estimate various physical quantities in the early universe:

- matter density
- baryon-photon ratio
- curvature
- cosmological constant

Why wavelets?

If, again,

$$T(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm} Y_{lm}(x)$$

then

$$A_{lm} = \int T(x) \overline{Y}_{lm}(x) dS.$$

Integral is over the entire sphere, with respect to usual surface measure ($= \sin\theta d\theta d\varphi$).

Why wavelets?

If, again,

$$T(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm} Y_{lm}(x)$$

then

$$A_{lm} = \int T(x) \overline{Y}_{lm}(x) dS.$$

Integral is over the entire sphere, with respect to usual surface measure ($= \sin\theta d\theta d\varphi$).

We look for more information by considering

$$\beta_{jk} = \int T(x) \psi_{jk}(x) dS$$

where the ψ_{jk} are needlets, a type of spherical wavelet that is very well-localized in space and also in frequency with $j \sim 2^l$.

Some Applications of Needlets to Cosmology

- Handling foregrounds and masked regions
- Analogues for CMB polarization and other spin fields
- Searching for features (asymmetries) (cold spot)
- Probability density estimation for cosmic rays (for source detection)

Standard needlets – tight frame property

(Narcowich, Petrushev and Ward (2006) – definition in a moment The needlets ψ_{jk} do not form an orthonormal basis for L^2 on the sphere (as the spherical harmonics do). Rather they are a tight frame, which by definition is a countable set of functions $\{e_i\}$ such that for all L^2 functions f on the sphere,

$$\|f\|_2^2 = \sum_i |\langle f, e_i \rangle|^2$$

Standard needlets – tight frame property

(Narcowich, Petrushev and Ward (2006) – definition in a moment The needlets ψ_{jk} do not form an orthonormal basis for L^2 on the sphere (as the spherical harmonics do). Rather they are a tight frame, which by definition is a countable set of functions $\{e_i\}$ such that for all L^2 functions f on the sphere,

$$\|f\|_2^2 = \sum_i |\langle f, e_i \rangle|^2$$

From this one obtains also the reproducing formula

$$f = \sum_i \langle f, e_i \rangle e_i$$

with equality in the L^2 sense.

Standard needlets – definition

$$\psi_{jk}(x) = \sqrt{\lambda_{jk}} \sum_{\ell} b(\frac{\ell}{B^j}) \sum_{m=-\ell}^{\ell} \overline{Y}_{\ell m}(x) Y_{\ell m}(\xi_{jk});$$

Standard needlets – definition

$$\psi_{jk}(x) = \sqrt{\lambda_{jk}} \sum_{\ell} b(\frac{\ell}{B^j}) \sum_{m=-\ell}^{\ell} \overline{Y}_{\ell m}(x) Y_{\ell m}(\xi_{jk});$$

We say ψ_{jk} is centered at ξ_{jk}.

Standard needlets – definition

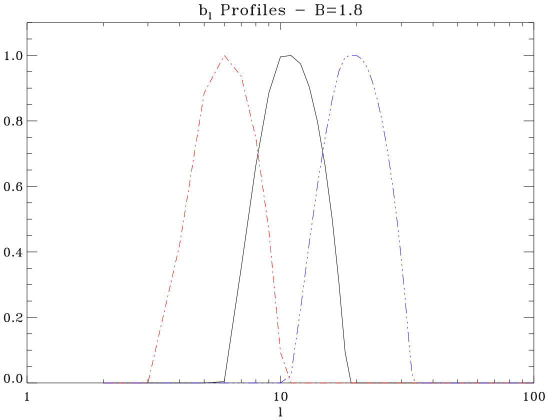
$$\psi_{jk}(x) = \sqrt{\lambda_{jk}} \sum_{\ell} b(\frac{\ell}{B^j}) \sum_{m=-\ell}^{\ell} \overline{Y}_{\ell m}(x) Y_{\ell m}(\xi_{jk});$$

We say ψ_{jk} is centered at ξ_{jk}.
Here $B > 1$, and the ξ_{jk} can be taken as HealPix points, which have the property

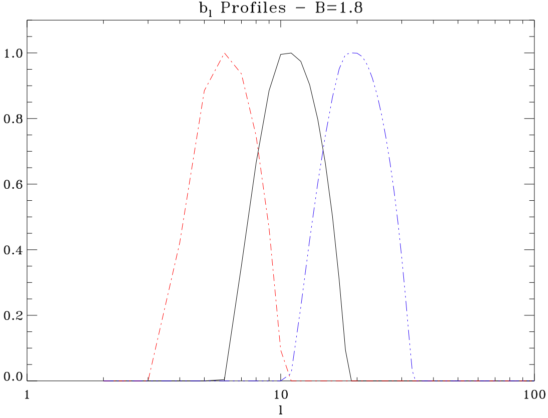
$$\int P = \sum_{k} \lambda_{jk} P(\xi_{jk})$$

for any polynomial P of degree $\leq B^{j+1}$.
Also $\lambda_{jk} \approx B^{-2j} \approx$ pixel area.
Which function b to take?

Standard $b(\frac{\cdot}{B^j})$

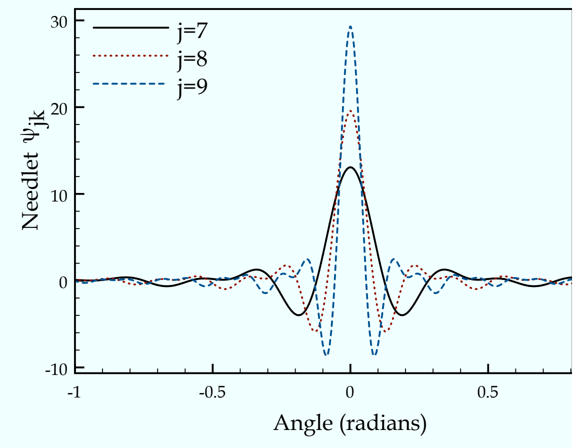


Standard $b(\frac{\cdot}{B^j})$



"Partition of unity" property – the sums of the squares of these functions is to be 1.

Needlets in Pixel Space



- These are centered at North Pole, angle $= \theta$, otherwise rotate
- well-localized (width $\approx B^{-j}$) , but a lot of oscillation

Asymptotic Uncorrelation

Under some mild regularity conditions on C_l, one has (Baldi, Kerkyacharian, Marinucci and Picard (2006)):

Asymptotic Uncorrelation

Under some mild regularity conditions on C_l, one has (Baldi, Kerkyacharian, Marinucci and Picard (2006)):

$$\left|Corr(\beta_{j,k}, \beta_{j,k'})\right| \leq \frac{c_M}{(1 + B^j d(\xi_{j,k}, \xi_{j,k'}))^M} \tag{1}$$

Asymptotic Uncorrelation

Under some mild regularity conditions on C_l, one has (Baldi, Kerkyacharian, Marinucci and Picard (2006)):

$$\left|Corr(\beta_{j,k},\beta_{j,k'})\right| \leq \frac{c_M}{(1+B^j d(\xi_{j,k},\xi_{j,k'}))^M} \tag{1}$$

where $d(\xi_{j,k},\xi_{j,k'})$ = distance between these points along a great circle $= \arccos(\langle \xi_{j,k},\xi_{j,k'} \rangle)$.

Asymptotic Uncorrelation

Under some mild regularity conditions on C_l, one has (Baldi, Kerkyacharian, Marinucci and Picard (2006)):

$$\left|Corr(\beta_{j,k},\beta_{j,k'})\right| \leq \frac{c_M}{(1+B^j d(\xi_{j,k},\xi_{j,k'}))^M} \tag{1}$$

where $d(\xi_{j,k},\xi_{j,k'})$ = distance between these points along a great circle $= \arccos(\langle\xi_{j,k},\xi_{j,k'}\rangle)$.
If the centers are different, the needlet coefficients are asymptotically uncorrelated.

Spectral Estimator

Consider the estimator

$$\Gamma_j = \frac{1}{N_j} \sum_{k=1}^{N_j} \{\widehat{\beta}_{j,k}^2\}$$

where

$$\mathrm{E}\widehat{\beta}_{j,k}^2 = \sum_{B^{j-1} \leq l \leq B^{j+1}} b^2(\tfrac{l}{B^j}) C_l \frac{2l+1}{4\pi}.$$

See Pietrobon, Balbi and Marinucci (2006), Baldi, Kerkyacharian, Marinucci and Picard (2006,2007), Fay et al. (2008), Fay and Guilloux (2008), Pietrobon et al. (2008).

Spectral Estimator

Due to localization and uncorrelation properties, the previous estimator can be evaluated on subsets of the sky, and used to search for features/anisotropies (Pietrobon et al., Phys Rev D (2008)). Statistical significance can be evaluated analytically and from simulations.

Needlet coefficients - features

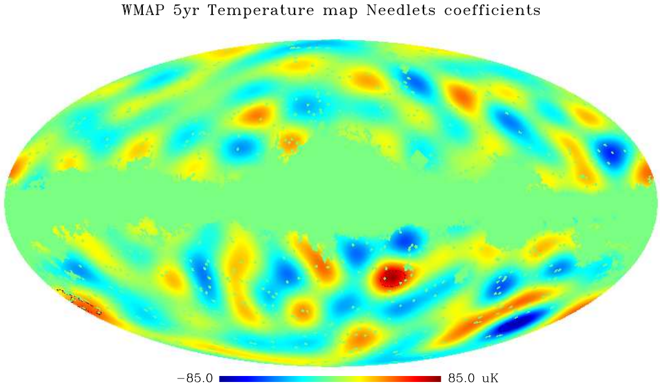
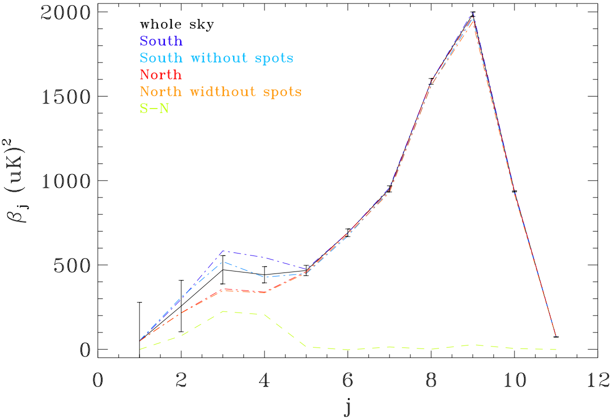


Figure: j=4, B=1.8

Asymmetries in the angular power spectrum



Angular power spectrum estimator

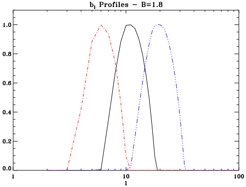
Mexican needlets

(Geller-Mayeli (2009); Scodeller, Rudjord, Hansen, Marinucci, Geller, Mayeli (2010))

Recall:

$$\psi_{jk}(x) = \sqrt{\lambda_{jk}} \sum_{\ell} b\left(\frac{\ell}{B^j}\right) \sum_{m=-\ell}^{\ell} \overline{Y}_{\ell m}(x) Y_{\ell m}(\xi_{jk});$$

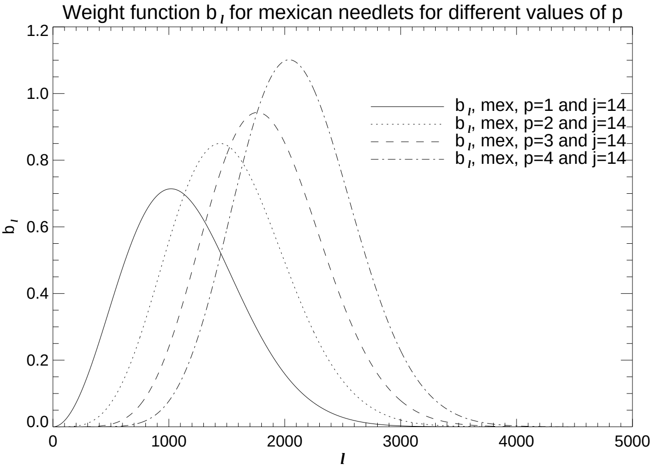
Before, we took



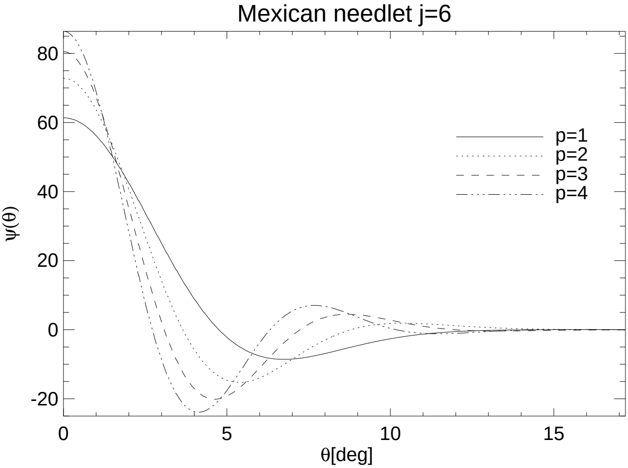
For Mexican needlets, we take instead

$$b\left(\frac{\ell}{B^j}\right) = \left(\frac{\ell}{B^j}\right)^{2p} e^{-\ell^2/B^{2j}}$$

Mexican needlets



Mexican needlets in pixel space



Advantages of Mexican needlets

- Extremely good localization in pixel space: the tails decay as $\exp(-B^{2j}x^2/4)$, as $j \to \infty$ ($x =$ angular distance)
- p adjustable if one wants more localization in frequency (and less in pixel space)
- Negligibly different from a tight frame
- Analytic expressions can be provided for their high-frequency behavior in pixel space.
- Very little oscillation in pixel space, so implementation there can be numerically stable
- Still have asymptotic uncorrelation, if $C_\ell = G(\ell)\ell^{-\alpha}$, and $\alpha < 4p + 2$ (Mayeli (2008)i, Lan and Marinucci (2008))
- For physically realistic C_ℓ, outperform standard needlets in uncorrelation properties

Spin Needlets for CMB Polarization data

Let s be an integer (for polarization, $s = 2$). Spin needlets are defined as (Geller and Marinucci (2008))

$$\psi_{jk;s}(p) = \sqrt{\lambda_{jk}} \sum_{l} b\left(\frac{l}{B^j}\right) \sum_{m=-l}^{l} Y_{l;ms}(p) \overline{Y_{l;ms}}(\xi_{jk}),$$

where the $Y_{l;ms}$ are spin spherical harmonics. More rigorously, if $e_{\ell s} = (\ell - |s|)(\ell + |s| + 1)$,

$$\psi_{jk;s}(p) = \sqrt{\lambda_{jk}} \sum_{l} b\left(\frac{\sqrt{e_{\ell s}}}{B^j}\right) \sum_{m=-l}^{l} \left\{ Y_{l;ms}(p) \otimes \overline{Y_{l;ms}}(\xi_{jk}) \right\} .$$

As before, $\{\lambda_{jk}, \xi_{jk}\}$ are cubature points and weights, $b(\cdot) \in C^{\infty}$ is nonnegative, and has compact support in $[1/B, B]$.

Spin Needlet Transform

The spin needlet transform is defined by

$$\int_{\mathbb{S}^2} f_s(p)\overline{\psi_{jk;s}}(p)dp = \beta_{jk;s} \ ,$$

and the same inversion property holds as for standard needlets, i.e.

$$f_s(p) = \sum_{jk} \beta_{jk;s}\psi_{jk;s}(p) \ ,$$

the equality holding in the L^2 sense. The coefficients of spin needlets are

$$\beta_{jk;s} = \int_{\mathbb{S}^2} f_s(p)\overline{\psi_{jk;2}}(p)dp = \sqrt{\lambda_{jk}} \sum_{l} b\left(\frac{l}{B^j}\right) \sum_{m=-l}^{l} a_{l;ms} Y_{l;ms}\left(\xi_{jk}\right) \ . \tag{2}$$

Tight frame, localization, asymptotic uncorrelation under mild hypotheses on C_ℓ

Power Spectrum Estimation

(Geller and Marinucci (2008), Geller, Lan and Marinucci (2009))

$$\Gamma_{j;s} := E \sum_k |\beta_{jk;s}|^2 = \sum_k \sum_l b^2 \left(\frac{\sqrt{e_{ls}}}{B^j} \right) C_l (2l+1)$$

$$\Gamma^*_{j;sG} := 4\pi (\sum \lambda_{jk})^{-1} \sum_k \left| \beta^*_{jk;s} \right|^2$$

where the sum is over those k with ξ_{jk} outside G^ϵ, and where

$$\beta^*_{jk;s} = \int_{G^c} P(x) \overline{\psi}_{jk,s}(x) dx.$$

(G is a masked region, and G^ϵ is a small neighborhood of it.) We have

$$\frac{\widehat{\Gamma}^*_{j;sG} - \Gamma_{j;s}}{\sqrt{Var\left\{ \widehat{\Gamma}^*_{j;sG} \right\}}} \to_d N(0,1) \text{ , as } j \to \infty \text{ .}$$

under a Gaussianity assumption.

Mixed Needlets

(Geller and Marinucci (2010))

$$\psi_{jk;s\mathcal{M}}(p) = \sqrt{\lambda_{jk}} \sum_l b\left(\frac{\sqrt{e_{\ell s}}}{B^j}\right) \sum_{m=-l}^{l} \left\{ Y_{l;ms}(p)\, \overline{Y_{lm}}(\xi_{jk}) \right\} \ .$$

Still have

- near-diagonal localization
- asymptotic uncorrelation under usual hypotheses

But mixed needlets can be used for cross-spectrum estimation, for the C_ℓ^{TE}, again under a Gaussianity assumption.

Directional Data

Baldi, Kerkyacharian, Marinucci and Picard (AoS 2009)
Assume we observe $X_1, ..., X_n \in S^2$.
We wish to estimate their density on the sphere $f(x)$. We know that

$$f(x) = \sum_{j,k} \beta_{jk}\psi_{jk}(x) \tag{3}$$

where

$$\beta_{jk} = \int_{S^2} f(x)\psi_{jk}(x)dx \tag{4}$$

The coefficients β_{jk} can be estimated by

$$\widehat{\beta}_{jk} = \frac{1}{n} \sum_{i=1}^{n} \psi_{jk}(X_i) \tag{5}$$

leading to the linear wavelet estimator

$$\widehat{f}(x) = \sum_{j,k} \widehat{\beta}_{jk} \psi_{jk}(x) \tag{6}$$

It can be shown, however, that (nearly) optimal estimates are obtained by thresholding, i.e.

$$\widehat{f}(x) = \sum_{j,k} \widehat{\beta}_{jk}^{H} \psi_{jk}(x) \tag{7}$$

$$\widehat{\beta}_{jk}^{H} = \widehat{\beta}_{jk} I(|\widehat{\beta}_{jk}| > t_n) \tag{8}$$

Intuitively, the smallest coefficients are expected to be dominated by noise and hence discarded. One takes $t_n = k_0 \sqrt{\frac{logn}{n}}$

Simulations

As an example, we try to estimate the following mixture of Gaussian density

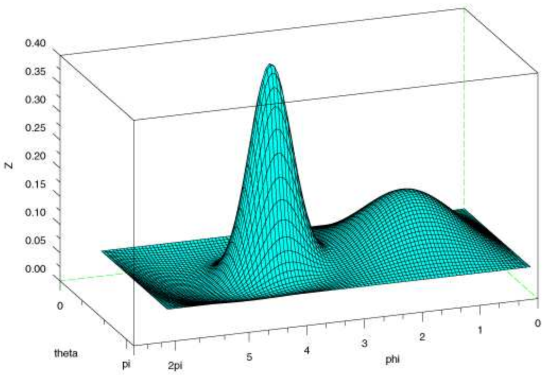
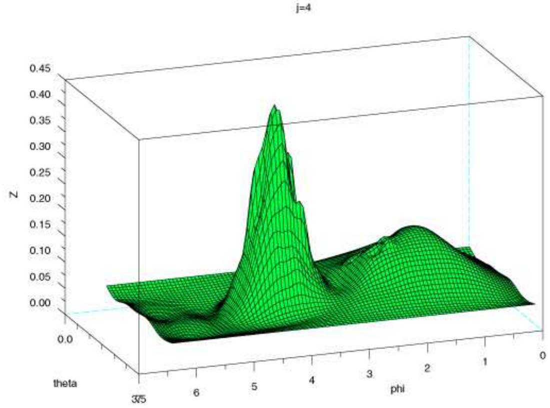
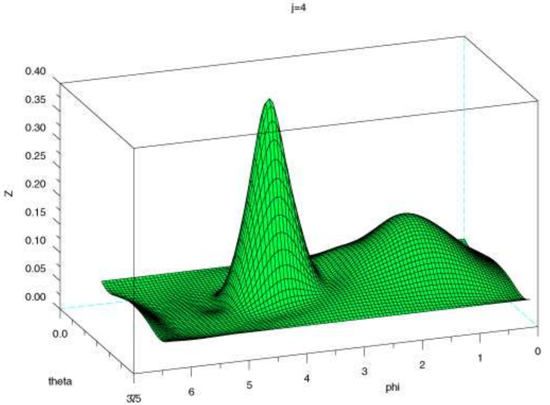


Figure: The target density

Small threshold



Medium threshold



High threshold