

Localized spectral analysis on the sphere

Mark A. Wieczorek¹ and Frederik J. Simons^{2,*}

¹Département de Géophysique Spatiale et Planétaire, Institut de Physique du Globe de Paris, France. E-mail: wieczor@ipgp.jussieu.fr

²Geosciences Department, Guyot Hall, Princeton University, Princeton, NJ 08540, USA

Accepted 2005 May 13. Received 2005 May 9; in original form 2004 October 12

SUMMARY

It is often advantageous to investigate the relationship between two geophysical data sets in the spectral domain by calculating admittance and coherence functions. While there exist powerful Cartesian windowing techniques to estimate spatially localized (cross-)spectral properties, the inherent sphericity of planetary bodies sometimes necessitates an approach based in spherical coordinates. Direct localized spectral estimates on the sphere can be obtained by tapering, or multiplying the data by a suitable windowing function, and expanding the resultant field in spherical harmonics. The localization of a window in space and its spectral bandlimitation jointly determine the quality of the spatio-spectral estimation. Two kinds of axisymmetric windows are here constructed that are ideally suited to this purpose: bandlimited functions that maximize their spatial energy within a cap of angular radius θ_0 , and spacelimited functions that maximize their spectral power within a spherical harmonic bandwidth L . Both concentration criteria yield an eigenvalue problem that is solved by an orthogonal family of data tapers, and the properties of these windows depend almost entirely upon the space–bandwidth product $N_0 = (L + 1)\theta_0/\pi$. The first $N_0 - 1$ windows are near perfectly concentrated, and the best-concentrated window approaches a lower bound imposed by a spherical uncertainty principle. In order to make robust localized estimates of the admittance and coherence spectra between two fields on the sphere, we propose a method analogous to Cartesian multitaper spectral analysis that uses our optimally concentrated data tapers. We show that the expectation of localized (cross-)power spectra calculated using our data tapers is nearly unbiased for stochastic processes when the input spectrum is white and when averages are made over all possible realizations of the random variables. In physical situations, only one realization of such a process will be available, but in this case, a weighted average of the spectra obtained using multiple data tapers well approximates the expected spectrum. While developed primarily to solve problems in planetary science, our method has applications in all areas of science that investigate spatio-spectral relationships between data fields defined on a sphere.

Key words: Slepian concentration problem, spatio-spectral localization, spectral estimation, spherical harmonics.

1 INTRODUCTION

It is common in geophysics to work with vector fields via the gradient of a scalar function. By expressing this function as a sum of spectral components, many mathematical operators such as the Laplacian and the radial derivative can be considerably simplified. It further arises that the relationship between two fields (such as gravity and topography) is linear in the spectral domain, and that the admittance or transfer function can be inverted for geophysical parameters (e.g. Turcotte *et al.* 1981). Moreover, for some geophysical observables,

such as satellite-derived gravity, the primary data are expressed in the spectral domain (e.g. Lemoine *et al.* 1998).

The spectral properties of data distributed on the sphere often vary as a function of position, and it is not uncommon that the data might also be only locally known. For these reasons, it is important to be able to obtain reliable spectral and cross-spectral estimates of data localized to specific geographic regions of interest. To this end, many powerful techniques have been developed in the Cartesian domain which generally consist of multiplying the data by a selective window, or taper, before performing the spectral expansion. The question thus naturally arises: What is the best form of this localizing window function? Slepian and coworkers (for a review, see Slepian 1983) posed and solved this problem in one dimension by finding a family of orthogonal functions that are optimally concentrated in

*Now at: Department of Earth Sciences, University College London, London WC1E 6BT, UK. E-mail: fjsimons@alum.mit.edu

the time or spectral domain. These were subsequently employed by Thomson (1982) to estimate the power spectrum of time-series data (for a comprehensive discussion, see Percival & Walden 1993). Solutions in two dimensions were given by Bronez (1988), Liu & van Veen (1992) and Hanssen (1997), and, in more general settings, by Daubechies (1988) and Daubechies & Paul (1988). Multidimensional multitaper methods have been applied in geophysics by Simons *et al.* (2000, 2003), among others.

On the Earth, the assumption of Cartesian geometry is often locally valid. However, on small bodies such as the Moon and Mars, the effects of curvature cannot be neglected and this necessitates an approach developed in spherical geometry. In this study, therefore, we solve Slepian's concentration problem on the surface of a sphere and obtain two classes of windows that are optimally concentrated: one spatially, within an angular radius θ_0 , the other spectrally, within a spherical harmonic bandwidth L . Here, we limit ourselves to axisymmetric windows and we refer to Simons *et al.* (2005) for extensions to arbitrary domains and non-zonal windows. Our results share many features with their Cartesian counterparts: the coefficients of the windows are solutions to an eigenvalue problem, their properties depend almost exclusively upon a space-bandwidth product, $N_0 = (L + 1)\theta_0/\pi$, and the number of near perfectly concentrated windows is well approximated by $N_0 - 1$.

When the input data are governed by a stationary stochastic process, a simple analytic expression exists that relates the input spectrum to the localized spectrum averaged over all possible realizations of the random variables. In the case where the power spectrum of the data is white, the localized spectral estimates using our windows have a quantifiably small bias away from the true power spectrum of the data. If, as in most physical situations, only one realization of the process is available for analysis (such as the gravity or topography field of a planet), then the true power spectrum is well approximated by a weighted average of spectra obtained from individual orthogonal data tapers. Such multitaper power spectral estimates display smaller estimation variance as N_0 and the number of employed tapers increase. What needs to be specified by the analyst is the size of the concentration region, θ_0 , and the spectral bandwidth, L , which together determine the quality of the spatio-spectral localization, and the bias and variance of the resulting estimates. From these, localized admittance and coherence spectra can be readily calculated.

Our windows solve an optimization criterion for spatio-spectral localization. As they outperform any other non-optimized window design, we expect our method to be immediately applicable for the analysis of planetary gravity and topography fields (e.g. Freedman & Windheuser 1997; Simons *et al.* 1997; McGovern *et al.* 2002; Lawrence & Phillips 2003; Smrekar *et al.* 2003; Hoogenboom *et al.* 2004), planetary magnetic fields (e.g. Voorhies *et al.* 2002), geodesy (e.g. Albertella *et al.* 1999), hydrology (e.g. Swenson & Wahr 2002; Swenson *et al.* 2003), but also cosmology (e.g. Hivon *et al.* 2002; Efstathiou 2004), medical imaging (e.g. Polyakov 2002) and numerical analysis (Ould Kaber 1996; Jakob-Chien & Alpert 1997; Swartztrauber & Spatz 2000). Since our primary motivation is to analyse geophysical data fields, we concentrate our comparisons on the widely used method of Simons *et al.* (1997) which uses truncated spherical harmonic expansions of boxcar polar caps. An extension of our method to irregularly shaped concentration regions, as well as proofs of certain asymptotic features of our zonal windows, is given by Simons *et al.* (2005). A study of Martian lithospheric properties using localized estimates of gravity-topography admittance and coherence spectra using our method is described by Belleguic *et al.* (2005).

This paper is organized as follows. First, we give an overview of spherical harmonics and describe how windowing modifies the spectral coefficients of an input data field. Next, we derive new windows that are optimally concentrated in the space and spectral domain, respectively. Following this, we discuss the properties of these windows, quantify their performance in terms of a spherical uncertainty relationship, and compare their merits with those of a spectrally truncated spherical cap. We next derive the relationship between localized power spectra and the global spectra of stationary stochastic processes, quantify the bias and variance of the spectral estimates, and describe a practical multitaper method that uses the new optimized data tapers to obtain localized admittance, coherence and power spectra. Finally, we present an example of our method by estimating the localized free-air gravity power spectrum of the Earth, we discuss the practical issues that arise when analysing low-resolution finite bandwidth data sets, and we describe the necessary steps involved in a typical localized multitaper analysis in recipe form.

2 ESTIMATING LOCAL SPECTRA ON THE SPHERE

Any square-integrable function on the unit sphere can be expressed by a linear combination of spherical harmonics as (e.g. Dahlen & Tromp 1998, Appendix B)

$$f(\Omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm} Y_{lm}(\Omega), \quad (1)$$

where Y_{lm} is a spherical harmonic of degree l and order m , f_{lm} is the corresponding spherical harmonic expansion coefficient, and $\Omega = (\theta, \phi)$ represents position on the sphere in terms of colatitude, θ , and longitude, ϕ . (See Table 1 for a description of the most often used symbols in this paper.) The wavelength λ on the surface of a sphere of radius R that is asymptotically equivalent to a spherical harmonic degree l is given by the Jeans relation $\lambda = 2\pi R/\sqrt{l(l+1)} \approx 2\pi R/(l+1/2)$. We define real spherical harmonics as

$$Y_{lm}(\Omega) = \begin{cases} \bar{P}_l(\cos\theta) \cos m\phi & \text{if } m \geq 0 \\ \bar{P}_{l|m|}(\cos\theta) \sin |m|\phi & \text{if } m < 0, \end{cases} \quad (2)$$

with

$$\bar{P}_l(x) = (-1)^m \sqrt{(2 - \delta_{0m})(2l+1)} \frac{(l-m)!}{(l+m)!} P_{lm}(x). \quad (3)$$

Here, δ_{ij} is the Kronecker delta function, and P_{lm} is an associated Legendre function (e.g. Masters & Richards-Dinger 1998; Press *et al.* 1992, pp. 246–248). These spherical harmonics are orthogonal over the sphere according to

$$\int_{\Omega} Y_{lm}(\Omega) Y_{l'm'}(\Omega) d\Omega = 4\pi \delta_{ll'} \delta_{mm'}, \quad (4)$$

where $d\Omega = \sin\theta d\theta d\phi$, and with this relationship, the individual spherical harmonic coefficients of eq. (1) can be shown to be

$$f_{lm} = \frac{1}{4\pi} \int_{\Omega} f(\Omega) Y_{lm}(\Omega) d\Omega. \quad (5)$$

It is straightforward to verify Parseval's theorem on the sphere: the total power in the space and spectral domains is related via

$$\frac{1}{4\pi} \int_{\Omega} f^2(\Omega) d\Omega = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm}^2 = \sum_{l=0}^{\infty} S_{ff}(l), \quad (6)$$

Table 1. Essential notation and mathematical symbols used in this paper.

Symbol	Definition
$Y_{lm}(\Omega)$	Real spherical harmonic of degree l and order m whose inner product over the unit sphere is 4π
Ω	The unit sphere, or the position on the unit sphere in terms of colatitude, θ , and longitude, ϕ
$P_{lm}(x)$	Associated Legendre function of degree l and order m
$\bar{P}_{lm}(x)$	Semi-normalized associated Legendre function of degree l and order m whose inner product over $[-1, 1]$ is $(4 - 2\delta_{0m})$
$f(\Omega), g(\Omega)$	Arbitrary square-integrable real-valued functions defined on the sphere
f_{lm}, g_{lm}	Spherical harmonic expansion coefficients of $f(\Omega)$ and $g(\Omega)$
$h(\theta)$	Arbitrary axisymmetric (zonal) windowing function
$h^{(k)}(\theta)$	k th axisymmetric windowing function from an orthogonal family satisfying a spatio-spectral optimization criterion
h_l	Spherical harmonic expansion coefficients of the zonal windowing function $h(\theta)$
L_h, L_f, L_g	Maximum spherical harmonic degrees of $h(\theta)$, $f(\Omega)$ and $g(\Omega)$
L	Bandwidth (for bandlimited functions), or effective bandwidth (for spacelimited functions) of the window $h(\theta)$
θ_0	Colatitude of a spherical cap centred at $\theta = 0$
N_0	Axisymmetric space–bandwidth product, or Shannon number, $(L + 1)\theta_0/\pi$
λ_k	Spatial, or spectral, concentration factor of the window $h^{(k)}(\theta)$
$\Phi(\Omega), \Gamma(\Omega)$	Windowed fields corresponding to $h(\theta)f(\Omega)$ and $h(\theta)g(\Omega)$
Φ_{lm}, Γ_{lm}	Spherical harmonic coefficients of $\Phi(\Omega)$ and $\Gamma(\Omega)$
$S_{ff}(l)$	Spectral power of the function $f(\Omega)$ for degree l
$S_{fg}(l)$	Cross-spectral power of the functions $f(\Omega)$ and $g(\Omega)$ for degree l
$S_{\Phi\Gamma}(l)$	Cross-spectral power of $\Phi(\Omega)$ and $\Gamma(\Omega)$ for degree l
$S_{\Phi\Gamma}^{(k)}(l)$	Cross-spectral power of $f(\Omega)$ and $g(\Omega)$ each windowed by $h^{(k)}(\theta)$ for degree l
$S_{\Phi\Gamma}^{(ml)}(l)$	Multitaper cross-spectral power estimate of $f(\Omega)$ and $g(\Omega)$ for degree l
$S_{\Phi\Gamma}^{(j,k)}(l, m)$	Cross-spectral power for degree l and order m of $f(\Omega)$ and $g(\Omega)$ windowed by $h^{(j)}(\theta)$ and $h^{(k)}(\theta)$, respectively
$\langle \dots \rangle, \text{cov} \{ \dots \}, \text{var} \{ \dots \}$	Expectation, covariance, and variance operators

where we define $S_{ff}(l)$ to be the power spectrum of the function f . Analogously, the total cross-power of two functions f and g can be written as

$$\frac{1}{4\pi} \int_{\Omega} f(\Omega)g(\Omega) d\Omega = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm} g_{lm} = \sum_{l=0}^{\infty} S_{fg}(l), \quad (7)$$

where $S_{fg}(l)$ is the cross-power spectrum. (For a zero-mean process, $S_{ff}(l)$ and $S_{fg}(l)$ are commonly referred to as the degree variance and covariance, respectively.) Power spectra, since they contain a sum over all orders m , are invariant under a rotation of the coordinate system (e.g. Kaula 1967; Lowes 1974). Some authors prefer to work with the average power per degree, or power spectral density, $S(l)/(2l + 1)$, as this ensures that the spectral coefficients of a spherical Dirac delta function are constant, independent of degree (e.g. Chevrot *et al.* 1998; Dahlen & Tromp 1998; Hipkin 2001; Simons *et al.* 2005).

Our 4π -normalization of the spherical harmonics in eq. (4) is commonly used in the geodesy community (e.g. Kaula 2000), and with these definitions each individual harmonic possesses unit power. Orthonormal (or unit energy) harmonics are commonly used in seismology (e.g. Dahlen & Tromp 1998), whereas Schmidt-normalized harmonics, whose power is $1/(2l + 1)$, are generally used in the geomagnetic community (e.g. Lowes 1966; Blakely 1995).

Let us now consider an axisymmetric (zonal) windowing function, $h(\theta)$, that is expanded in spherical harmonics up to a maximum degree L_h :

$$h(\theta) = \sum_{j=0}^{L_h} h_j Y_{j0}(\Omega) = \sum_{j=0}^{L_h} h_j \bar{P}_j(\cos \theta). \quad (8)$$

The total cross-power of two arbitrary functions $f(\Omega)$ and $g(\Omega)$, each multiplied by such a window, is given by

$$\begin{aligned} & \frac{1}{4\pi} \int_{\Omega} [h(\theta)f(\Omega)][h(\theta)g(\Omega)] d\Omega \\ &= \sum_{l=0}^{\infty} \sum_{m=-l}^l \Phi_{lm} \Gamma_{lm} = \sum_{l=0}^{\infty} S_{\Phi\Gamma}(l). \end{aligned} \quad (9)$$

Here, Φ_{lm} and Γ_{lm} represent the coefficients of the two windowed fields, $\Phi(\Omega) = h(\theta)f(\Omega)$ and $\Gamma(\Omega) = h(\theta)g(\Omega)$. Using eq. (5), these can be calculated according to

$$\begin{aligned} \Phi_{lm} &= \frac{1}{4\pi} \int_{\Omega} [h(\theta)f(\Omega)] Y_{lm}(\Omega) d\Omega \\ &= \sum_{j=0}^{L_h} \sum_{i=|l-j|}^{l+j} h_j f_{im} \sqrt{(2i+1)(2j+1)(2l+1)} \\ &\quad \times (-1)^m \begin{pmatrix} i & j & l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} i & j & l \\ m & 0 & -m \end{pmatrix}, \end{aligned} \quad (10)$$

where the matrix symbols in parentheses are Wigner 3- j functions, and the second equality follows from a well-known result quoted as eq. (C10) in Appendix C. We henceforth refer to the (cross-)power spectrum of the windowed field(s) in eq. (9) as a localized (cross-)power spectrum estimate. Since no assumption has been made as to the form of the global power spectrum, this localized power spectrum is a direct spectrum estimate.

We note that if the spherical harmonic coefficients of $f(\Omega)$ are unknown beyond a maximum spherical harmonic degree L_f , then only the first $L_f - L_h$ coefficients of the windowed field are reliable. This fact is a simple consequence of the limits of the second summation in eq. (10) and poses a major limitation in obtaining localized spectral estimates of finite-resolution data (Simons *et al.* 1997). Eq. (10) further demonstrates that each windowed coefficient of degree l receives contributions from the data coefficients over the range $|l - L_h| \leq l \leq l + L_h$. Thus, the windowed coefficients are spectrally smoothed versions of the data coefficients. The amount of smoothing depends on the maximum degree of the window, L_h , and on the shape of its power spectrum within this bandwidth.

Although we postpone to Section 5 a detailed analysis of the statistical relation of the spectrum of a windowed field to its global spectrum, it is obvious that a key requirement to obtaining reliable localized spectral estimates on the sphere is to apply a windowing function that maximizes its spatial concentration inside the region of interest. At the same time, as a result of eq. (10), the windowing

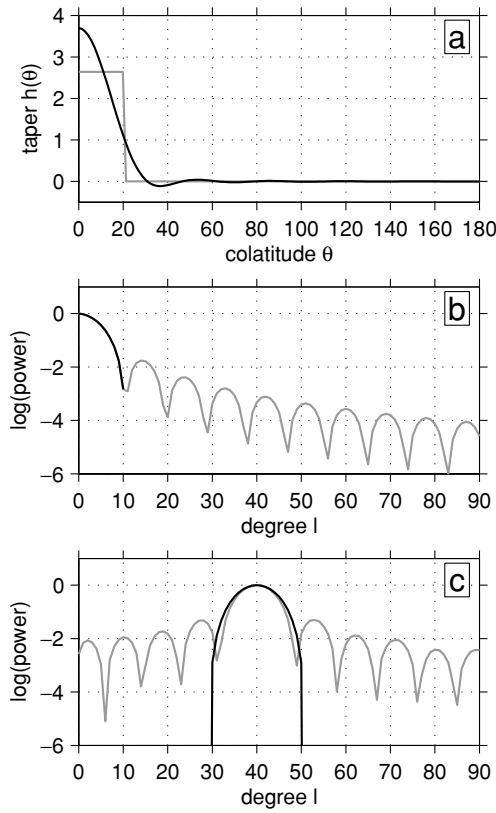


Figure 1. (a) Spatial rendition of a spacelimited axisymmetric polar cap of colatitudinal radius $\theta_0 = 20^\circ$ (grey), and the bandlimited version obtained by truncating its spectral coefficients beyond $L = 10$ (black). (b) Normalized power spectra of the perfect (grey) and truncated (black) caps in (a). (c) Localized power spectra of a spatial field defined by a single harmonic $Y_{40,0}$ obtained after tapering with the windows in (a). While the spectral leakage is confined to the interval $l \pm L$ for the spectrally truncated window (black), it is omnipresent for the perfect spherical cap window (grey).

coefficients must possess the smallest possible spectral bandwidth in order to limit the spectral smoothing.

We demonstrate the drawbacks of a poor window design by considering a spherical cap that is sharply localized in space, having a constant value for $\theta \leq 20^\circ$ and vanishing elsewhere, as shown in Fig. 1 (grey lines). The spatial (Fig. 1a) and spectral (Fig. 1b) representations (expanded to $L_h = 90$) of this naive localizing window show that perfect spatial localization is achieved at the expense of an infinite spectral bandwidth (e.g. Mallat 1998). In particular, the window possesses spectral sidelobes beyond its main lobe at $L = 10$, and the magnitude of the sidelobes decays only slowly with increasing spherical harmonic degree. The power spectrum (Fig. 1c) of a single zonal ($m = 0$) harmonic of degree $l = 40$ windowed with this function displays leakage across all adjacent degrees, as per eq. (10). Thus, by using windows that are sharply localized, each value of the windowed spectrum at a particular degree will depend on every degree of the input data. The higher the dynamic range of the input spectrum, the more significantly this contamination affects the windowed spectrum (see Percival & Walden 1993, Chap. 6). Nevertheless, with an effective bandwidth of $L = 10$, the majority of the spectral leakage in this example is confined to the degree range $40 - L \leq l \leq 40 + L$.

The undesirable qualities of the above example can be improved somewhat by reducing the magnitude of the spectral sidelobes of the windowing function, or eliminating them altogether. The latter

approach has been advocated by Simons *et al.* (1997), and forms the basis of a method of spatio-spectral concentration widely used in planetary science (e.g. Simons & Hager 1997; McGovern *et al.* 2002; Lawrence & Phillips 2003; Smrekar *et al.* 2003; Hoogenboom *et al.* 2004). In this procedure, the spherical harmonic coefficients of a sharply localized spherical cap are truncated beyond a degree close to the edge of the main lobe. This bandlimitation, here to $L = 10$, however, comes at the expense of spatial sidelobes exterior to the region of interest. While the absence of spectral sidelobes in the window confines the spectral leakage to the interval $40 - L \leq l \leq 40 + L$, the localized spectrum now contains information from data outside the target region in space.

3 THE ZONAL CONCENTRATION PROBLEM

In the preceding section, we showed that windows with sharp spatial cut-offs, or those whose spectrum has been sharply truncated, make for poor localizing windows. In this section, we use an optimization criterion to construct two classes of functions with better spatio-spectral localization properties: the first bandlimited (zero power outside of a bandwidth L , but optimally concentrated in space), the second spacelimited (zero energy outside of a radius θ_0 , but optimally concentrated spectrally). A linear combination of both would possess intermediate spatial and spectral concentration properties (Landau & Pollak 1960). For simplicity, we consider only axisymmetric polar windows (given by zonal spherical harmonics with order $m = 0$). Standard algorithms can then be used to rotate these to an arbitrary location on the unit sphere (e.g. Blanco *et al.* 1997).

Our optimization of spatio-spectral localization to a zonal polar cap is the spherical analogue to what is known as Slepian's problem in time-series analysis (e.g. Slepian 1983; Percival & Walden 1993). It has received attention previously from Grünbaum *et al.* (1982), whose results we will use to calculate efficiently the spherical harmonic coefficients of the data tapers. A related problem of finding bandlimited functions optimally concentrated within a latitudinal belt was investigated by Albertella *et al.* (1999). The extension of our procedure to concentration regions of arbitrary shape is developed by Simons *et al.* (2005).

3.1 Space concentration of a bandlimited function

Consider a bandlimited axisymmetric data taper, $h(\theta)$, expressed in spherical harmonics up to a maximum degree L (as in eq. 8, but dropping the distinction between L_h and L). We seek those functions that are optimally concentrated within a spherical cap extending over the colatitudes $0 \leq \theta \leq \theta_0$. The quality of spatial concentration will be quantified by the parameter λ , defined to be the ratio of the energy of the function within the region $0 \leq \theta \leq \theta_0$ to the energy over the entire sphere:

$$\lambda = \frac{\int_0^{2\pi} \int_0^{\theta_0} h^2(\theta) \sin \theta \, d\theta \, d\phi}{\int_0^{2\pi} \int_0^\pi h^2(\theta) \sin \theta \, d\theta \, d\phi}. \quad (11)$$

Using eq. (6), along with the definition of $h(\theta)$ from eq. (8), we may show that

$$\lambda = \frac{\sum_{l=0}^L \sum_{l'=0}^L h_l D_{ll'}(\theta_0) h_{l'}}{\sum_{l=0}^L h_l^2}, \quad (12)$$

where the elements of the square symmetric localization kernel \mathbf{D} are given by

$$D_{ll'}(\theta_0) = \frac{1}{2} \int_{\cos \theta_0}^1 \bar{P}_l(x) \bar{P}_{l'}(x) dx, \quad (13)$$

with $\bar{P}_l = \bar{P}_{l_0}$, and $x = \cos \theta$. Eq. (12) can be concisely written in matrix notation as

$$\lambda = \mathbf{h}^T \mathbf{D} \mathbf{h} / \mathbf{h}^T \mathbf{h}, \quad (14)$$

where \mathbf{h} represents an array of the $L + 1$ window coefficients h_l . The solution that maximizes λ also maximizes the matrix equation

$$\lambda \mathbf{h} = \mathbf{D} \mathbf{h}, \quad (15)$$

which is a standard eigenvalue equation that is solved by an orthogonal family of functions. The spherical harmonic coefficients of these functions are given by the $L + 1$ eigenvectors labelled $\mathbf{h}^{(k)}$, and the corresponding eigenvalues λ_k represent the quality of their spatial concentration. We will refer to k as the taper number, and $h^{(k)}$ as the k th eigentaper, ordered such that

$$1 > \lambda_1 > \lambda_2 > \dots > \lambda_{L+1} > 0. \quad (16)$$

The eigenvector $\mathbf{h}^{(1)}$ by definition maximizes λ . The eigenvectors of eq. (15) are most easily found by diagonalizing a tridiagonal matrix that commutes with \mathbf{D} (see Grünbaum *et al.* 1982, and Appendix A), and the eigenvalues can then be obtained by either eqs (11) or (14). Exact expressions for the elements of the kernel \mathbf{D} are derived in Appendix B.

The above results can be interpreted in an analogous way in the space domain. We may start with eq. (15) written in indicial notation

$$\lambda h_l = \sum_{l'=0}^L D_{ll'} h_{l'} \quad (17)$$

and then multiply both sides by \bar{P}_l and sum over all l :

$$\begin{aligned} \lambda h(\theta) &= \lambda \sum_{l=0}^L h_l \bar{P}_l(\cos \theta) \\ &= \sum_{l=0}^L \sum_{l'=0}^L \left(\frac{1}{2} \int_{\cos \theta_0}^1 \bar{P}_l(x') \bar{P}_{l'}(x') dx' \right) \bar{P}_l(x) h_{l'} \\ &= \frac{1}{2} \int_{\cos \theta_0}^1 \left(\sum_{l=0}^L \bar{P}_l(x') \bar{P}_l(x) \right) \left(\sum_{l'=0}^L h_{l'} \bar{P}_{l'}(x') \right) dx'. \end{aligned} \quad (18)$$

Defining the symmetric localization kernel

$$D_L(\theta, \theta') = \frac{1}{2} \sum_{l=0}^L \bar{P}_l(\cos \theta) \bar{P}_l(\cos \theta'), \quad (19)$$

and using the expansion formula for $h(\theta)$, eq. (18) reduces to

$$\lambda h(\theta) = \int_0^{\theta_0} D_L(\theta, \theta') h(\theta') \sin \theta' d\theta', \quad 0 \leq \theta \leq \pi. \quad (20)$$

Thus, the solutions to the space-concentration problem are also solutions of a Fredholm integral equation of the second kind (e.g. Tricomi 1970). We note that the kernel $D_L(\theta, \theta')$ can be easily computed using the Christoffel–Darboux identity (e.g. Szegő 1975; Swartrauber & Spatz 2000)

$$\begin{aligned} &\sum_{l=0}^L (2l+1) P_l(x) P_l(x') \\ &= \begin{cases} \frac{L+1}{x-x'} [P_{L+1}(x) P_L(x') - P_L(x) P_{L+1}(x')] & \text{if } x \neq x' \\ (L+1) [P'_{L+1}(x) P_L(x) - P'_L(x) P_{L+1}(x)] & \text{if } x = x', \end{cases} \end{aligned} \quad (21)$$

where P'_l represents the first derivative of the degree l Legendre polynomial with respect to x .

3.2 Spectral concentration of a spacelimited function

As an alternative to the space-concentration problem of bandlimited functions, we now seek spacelimited functions, defined to be zero at colatitudes $\theta > \theta_0$, that are optimally concentrated within an effective spherical harmonic bandwidth L . Perfectly spacelimited functions of this kind will possess an infinite number of spherical harmonic coefficients (see again Fig. 1), and we here minimize the power associated with degrees greater than L . In practice, these windows will only be calculated up to a finite degree $L_h \geq L$.

The quality of spectral concentration λ is now defined as the ratio of the power in the degree range $0 \leq l \leq L$ to the power over the entire spectral band:

$$\lambda = \sum_{l=0}^L h_l^2 / \sum_{l=0}^{\infty} h_l^2. \quad (22)$$

Using eqs (5) and (6), this can be written as

$$\lambda = \frac{\int_0^{\theta_0} \int_0^{\theta_0} h(\theta) D_L(\theta, \theta') h(\theta') \sin \theta \sin \theta' d\theta d\theta'}{\int_0^{\theta_0} h^2(\theta) \sin \theta d\theta}, \quad (23)$$

where we have used the property that $h(\theta)$ is by definition zero exterior to θ_0 , and $D_L(\theta, \theta')$ is the kernel previously defined by eq. (19). Maximizing λ is thus readily shown to be equivalent to solving the following Fredholm integral equation of the second kind:

$$\lambda h(\theta) = \int_0^{\theta_0} D_L(\theta, \theta') h(\theta') \sin \theta' d\theta', \quad 0 \leq \theta \leq \theta_0. \quad (24)$$

We remark that the integrals on the right hand side of eqs (20) and (24) are identical. The sole exception with these equations is that in the first, the function $h(\theta)$ is valid for all θ , whereas in the second, it is valid only for $0 \leq \theta \leq \theta_0$. Hence, the first $L + 1$ solutions of the space and spectral concentration problems are identical on the interval $0 \leq \theta \leq \theta_0$, and the first $L + 1$ eigenvalues of the two problems are equivalent. While eq. (24) has an infinite number of eigensolutions, only the first $L + 1$ eigenvalues are non-zero (see Simons *et al.* 2005). Eq. (24) can be solved numerically (e.g. Simons *et al.* 2005; Press *et al.* 1992, pp. 782–785), but in practice it is more convenient to solve eq. (15) and to then set the obtained functions equal to zero outside of θ_0 .

4 EIGENSOLUTIONS OF THE ZONAL CONCENTRATION PROBLEM

The eigenfunctions and eigenvalues of Slepian's Cartesian concentration operator depend exclusively on the product of the window length and its spectral bandwidth (Slepian 1983), a parameter often referred to as the Shannon number (Percival & Walden 1993). Because of the finite size of the sphere, an exact scaling of our spherical eigensolutions cannot be expected. Nonetheless, we show that the properties of these are well described by the spherical analogue of the Cartesian Shannon number,

$$N_0 = (L+1) \frac{\theta_0}{\pi}. \quad (25)$$

The subscript of N_0 is used to distinguish between our case of axisymmetry, when $m = 0$, and the more general concentration problem that involves non-zero values of m . As shown by Simons *et al.* (2005), this space–bandwidth product is asymptotically equivalent to the Cartesian Shannon number in the limit $L \rightarrow \infty$ and $\theta_0 \rightarrow 0$. The following subsections illustrate the properties of the spherical

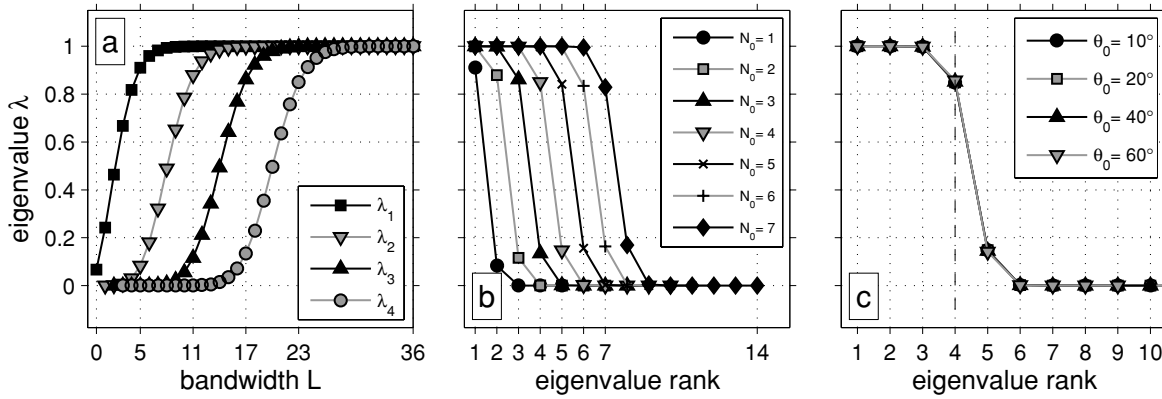


Figure 2. Eigenvalue spectra of the localization kernels. (a) Eigenvalues of the first four tapers with $\theta_0 = 30^\circ$ for increasing L . Whenever L increases by π/θ_0 , one additional eigenfunction is seen to be near perfectly concentrated. (b) Eigenvalue spectra for $\theta_0 = 30^\circ$ and varying space–bandwidth product $N_0 = (L + 1)\theta_0/\pi$. Each time N_0 increases by one, another eigenvalue attains a value of unity; $N_0 - 1$ tapers are always near perfectly concentrated. (c) Eigenvalue spectra for $N_0 = 4$ and different concentration regions θ_0 and bandwidths L . At constant N_0 , the individual spectra are nearly indistinguishable.

eigenvalues and eigenfunctions, quantify their performance using a spherical uncertainty principle, and contrast their behaviour to the spectrally truncated spherical cap windowing function developed by Simons *et al.* (1997).

4.1 Eigenvalues of the localization kernels

The eigenvalues of the kernels $D_{ll'}(\theta_0)$ and $D_L(\theta, \theta')$ are measures of the concentration of the data tapers in the space and spectral domains, respectively. Fig. 2 illustrates their dependence on L and θ_0 , and their asymptotic scaling with N_0 . In Fig. 2(a), the eigenvalues of the first four space-concentrated tapers are plotted as a function of their spectral bandwidth L , when $\theta_0 = 30^\circ$. As expected, each taper possesses a near-zero concentration factor for small window bandwidths, and as L increases, so does its corresponding concentration. The form of the $\lambda - L$ curve is similar for each taper, with each curve being offset from its neighbors by a nearly constant value of π/θ_0 . The entire eigenvalue spectrum for $\theta_0 = 30^\circ$ is plotted in Fig. 2(b) for several values of the space–bandwidth product N_0 . Each time N_0 is increased by an integer value, one additional near-unity eigenvalue is obtained. In Fig. 2(c) we plot the eigenvalue spectra for several combinations of L and θ_0 that have a constant value of $N_0 = 4$. In each case, the first $N_0 - 1$ eigenvalues are nearly unity, the subsequent two eigenvalues are intermediate in value, and the remaining eigenvalues are nearly zero. Regardless of the values chosen for L and θ_0 , the first $N_0 - 1$ eigenfunctions are always near perfectly concentrated.

The eigenvalue spectra shown in Fig. 2 display a steep transition between near-unity and near-zero values. As a result of this behaviour, the number of near perfectly concentrated eigenfunctions of the localization kernel is to a good approximation equal to the sum of its eigenvalues, which is well approximated by N_0 (see Simons *et al.* 2005). Thus, the space–bandwidth product N_0 is a practical approximation for the number of well-concentrated eigenfunctions obtained from the concentration problem.

4.2 Data tapers: the localization eigenfunctions

The spherical harmonic coefficients of the space-concentrated tapers are given by the eigenfunctions of the kernel \mathbf{D} . From these coefficients, the windows can be reconstructed in the space domain by use of eq. (1), and the spacelimited windows can be obtained simply by setting them to zero exterior to θ_0 . We normalize our windows such that they have unit power over the sphere, which guarantees

that the gain associated with the windowing process is close to unity (see Appendix C). Arbitrarily, we choose $h(\theta = 0) > 0$.

In Fig. 3, we plot the first four space-concentrated (grey) and spacelimited (black) tapers, in the space domain (top row), and in the spectral domain (bottom row). We recall that both classes of functions are identical for $\theta \leq \theta_0$ (see Section 3.2). The ringing of the space-concentrated tapers beyond θ_0 is a reflection of their eigenvalues being smaller than unity, and since the eigenvalue of each additional taper is smaller than the previous, this phenomenon becomes progressively more prominent for higher taper numbers.

By construction, the space-concentrated tapers are orthogonal over the entire sphere, whereas the spacelimited tapers are orthogonal over the region $0 \leq \theta \leq \theta_0$. Since both function classes are identical over the concentration region, it follows that the space-concentrated tapers are orthogonal over this restricted interval as well. In order to maintain orthogonality, each additional taper must have one additional zero-crossing within the interval $0 \leq \theta \leq \theta_0$.

The bottom row of Fig. 3 also shows the power spectra of the first four space-concentrated (grey) and spacelimited (black) tapers. By definition, the space-concentrated tapers are bandlimited within L , whereas for the spacelimited functions, the spectral power is maximized inside of, though not confined to, this bandwidth. The magnitude of the spectral power in the sidelobes beyond L in the latter functions is quantified by their non-unity eigenvalues. Since each additional eigenvalue is by definition smaller than the previous, the spectral sidelobes become more prominent as the taper number increases.

Finally, Fig. 3 demonstrates that the spectral power becomes more evenly distributed within the main spectral lobe as the taper number increases. Thus, even though the spectral smoothing associated with the windowing process is principally restricted to a bandwidth L , this smoothing effect will be somewhat greater for larger taper numbers. The power spectra for $l \leq L$ of both classes of functions differ only by a factor equal to their corresponding eigenvalue (Simons *et al.* 2005), and as these are near unity in this example, their spectra are nearly identical over this interval. The small differences that exist arise from our having calculated the window coefficients for the spacelimited tapers using a discrete implementation of eq. (24) (see Simons *et al.* 2005).

The asymptotic result that our windows and their associated power spectra are close to being scaled versions of each other when the space–bandwidth product N_0 is held constant (Simons *et al.* 2005) is illustrated in Fig. 4. There, the first four space-concentrated windows

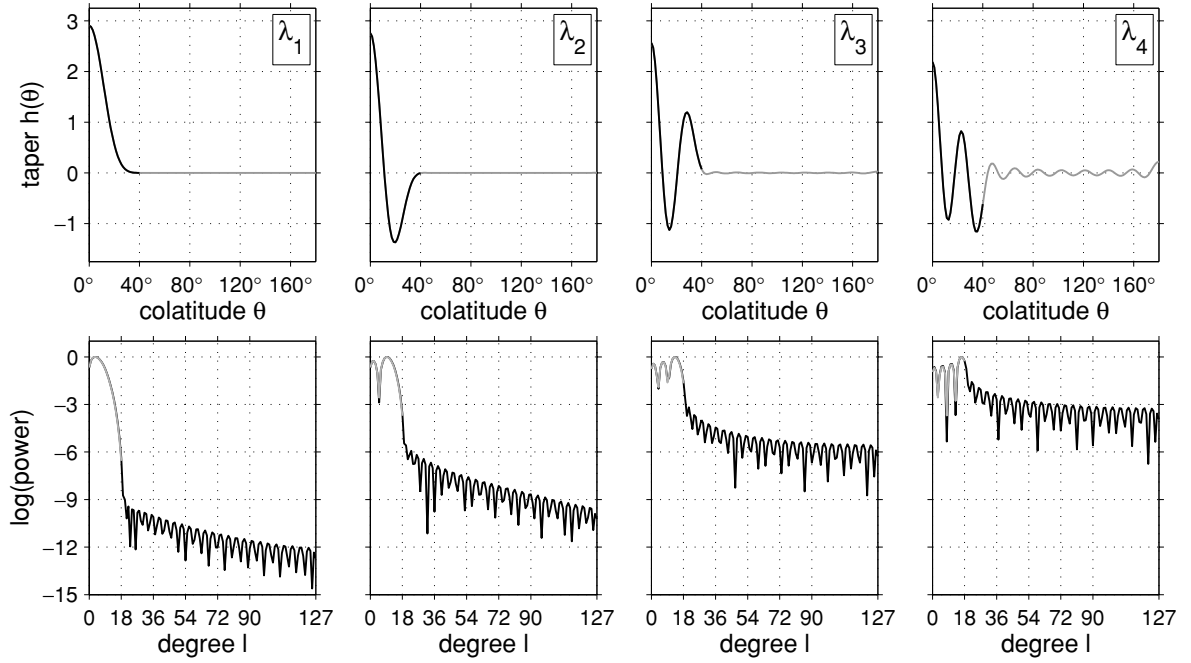


Figure 3. First four tapers in the space (top row) and spectral (bottom row) domains, for $\theta_0 = 40^\circ$, $L = 18$, and $N_0 = 4$. The spacelimited (black) and space-concentrated (grey) functions are identical for $\theta \leq \theta_0$, and their power spectra for $l \leq L$ differ only by a near-unity scaling factor. The power in the spectral sidelobes of the spacelimited functions (expanded to $L_h = 127$, but with an effective bandwidth of $L = 18$) becomes increasingly important with each additional taper, as does the spatial energy in the sidelobes of the bandlimited functions (where $L_h = L = 18$).

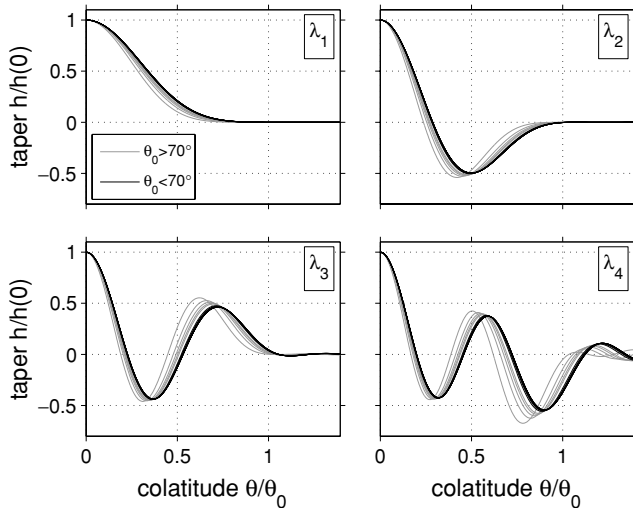


Figure 4. Scaling behaviour of the localization eigenfunctions. Plotted are the first four space-concentrated tapers as a function of scaled distance, θ/θ_0 , for a constant space-bandwidth product $N_0 = 4$. Sixteen different combinations of θ_0 and L are plotted, in grey when $\theta_0 > 70^\circ$, and in black when $\theta_0 < 70^\circ$. At constant N_0 , and as $L \rightarrow \infty$ and $\theta_0 \rightarrow 0$, the individual eigenfunctions become asymptotically indistinguishable.

are plotted as a function of scaled colatitude, θ/θ_0 , for several values of θ_0 and L corresponding to $N_0 = 4$. As is seen, when θ_0 is less than about 70° , the scaling is quite good, whereas for larger sized windows, the scaling becomes progressively worse. This should come as no surprise as the sphere is finite in size and functions cannot be scaled beyond the interval $0 \leq \theta \leq \pi$. However, when θ_0 is much smaller than π , the finite size of the sphere is of lesser importance, and the scaling is nearly perfect. A similar scaling applies to the

power spectra of our windows as a function of spherical harmonic degree scaled by the bandwidth. Our results imply that if a single scalable function were to be chosen as a data window on the sphere (e.g. Kido *et al.* 2003), the concentration properties of that window would become progressively worse upon scaling to larger and larger sizes.

4.3 A spherical uncertainty principle

In Section 3, we constructed two classes of functions that were perfectly limited in the spatial or spectral domain, and optimally concentrated in the other. It is natural to inquire whether a class of functions exists that are jointly concentrated in both domains. In the Cartesian plane, weighted Hermite polynomials possess this property, and a single parameter determines the trade-off between spatial and spectral concentration (see Simons *et al.* 2003, and references therein). While we do not derive such a class of functions here, we note that an uncertainty principle exists that relates the trade-off between a function's localization in space and its average Laplacian (Narcowich & Ward 1996). As the latter quantity is a second-order moment of the power spectrum (see below), this uncertainty principle is an alternative criterion that can be used for quantifying the quality of a window's spatio-spectral localization.

In deriving the uncertainty relationship, it is necessary to interpret $h^2(\Omega)$ as a probability density. As such, in this section, we require the windowing function $h(\theta)$ to be normalized to unit energy. All other sections in this paper employ a unit-power normalization (see Section 2). We first quote expressions for the variance of the position and average Laplacian of a function defined on the sphere (see Narcowich & Ward 1996 for further details). The average position of an axisymmetric function $h(\theta)$ is given by

$$\mu = \hat{\mathbf{z}} 2\pi \int_{-1}^1 h^2(\cos^{-1} x) x dx, \quad (26)$$

where $x = \cos \theta$, and its variance is readily shown to be

$$\sigma_\mu^2 = 1 - |\mu|^2. \quad (27)$$

The Laplace-Beltrami operator on the surface of a sphere obeys the eigenvalue equation (e.g. Dahlen & Tromp 1998, Appendix B)

$$\nabla_\Omega^2 Y_{lm}(\Omega) = -l(l+1)Y_{lm}(\Omega). \quad (28)$$

Since $h(\Omega)$ is periodic in the longitudinal coordinate ϕ , the expectation of this operator can be shown to be the zero vector, and its variance is thus

$$\sigma_L^2 = \int_\Omega h(\theta) [-\nabla_\Omega^2 h(\theta)] d\Omega = 4\pi \sum_{l=0}^{\infty} l(l+1)h_l^2. \quad (29)$$

By using an identity derived from quantum mechanics, as well as the Schwarz inequality, Narcowich & Ward (1996) show that the following inequality holds:

$$\sigma_z \sigma_L = \left(\frac{1 - |\mu|^2}{|\mu|^2} \right)^{1/2} \left(4\pi \sum_{l=0}^{\infty} l(l+1)h_l^2 \right)^{1/2} \geq 1, \quad (30)$$

where $\sigma_z = \sigma_\mu/|\mu|$. This equation differs from theirs only by a constant that is related to our use of 4π -normalized spherical harmonics.

The spherical uncertainty principle of eq. (30) states that the product of the variance of a function in the space domain and the variance of its Laplacian will always exceed a certain constant value. As shown by eq. (29), the variance of the Laplacian of a function is a second-order moment of its power spectrum. For the purposes of spatio-spectral localization, both terms in the uncertainty product should be as small as possible; the bound of unity for their product represents a theoretical lower limit.

We next use the uncertainty product $\sigma_z \sigma_L$ as a criterion to evaluate the performance of our windows. This quantity is plotted in Fig. 5 for the first five space-concentrated tapers as a function of the space-bandwidth product, N_0 , demonstrating that the uncertainty product depends almost exclusively upon the taper number k and N_0 . In particular, regardless of the values chosen for θ_0 and L ,

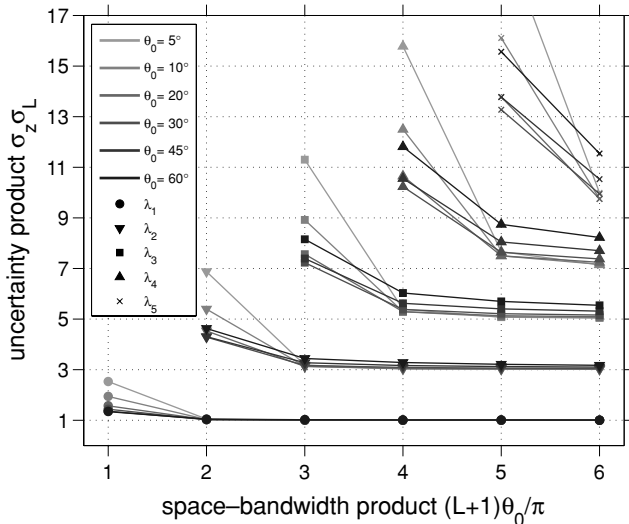


Figure 5. Spherical uncertainty product of the first five space-concentrated eigentapers of the localization problem for a variety of concentration regions θ_0 and bandwidths L , shown as a function of the space-bandwidth product, N_0 . The uncertainty product of the k th taper is seen to tend to an asymptotic value of $2k - 1$. The first taper in every sequence closely approaches the lower limit of unity.

the uncertainty product is empirically seen to rapidly approach an asymptotic value of $2k - 1$ for $N_0 \geq k + 1$. In addition, we note that only the first space-concentrated taper approaches, and nearly attains, the lower limit of unity imposed by eq. (30). The uncertainty product for the spacelimited tapers is similar to that of the space-concentrated tapers and shows the same asymptotic behaviour.

While neither the first nor the second class of eigenfunctions were designed with the uncertainty product of eq. (30) as a minimization criterion (cf. Carruthers & Nieto 1968; Bluhm *et al.* 1995; Tegmark 1995; Kowalski *et al.* 1996; Tegmark 1996; González & del Olmo 1998; Kowalski & Rembieliński 2000, 2002), it is apparent from Fig. 5 that our best concentrated windows, optimized relative to eqs (11) and (22), perform extremely well with regard to this measure. The increasingly poor performance that is seen to occur with increasing taper number is a result of the fact that the power spectrum of our orthogonal data tapers becomes increasingly uniform within the bandwidth L with increasing taper number (see Fig. 3).

4.4 Performance comparisons with other windows

We noted in Section 2 that a simple way to obtain localized direct spectral estimates on the sphere was to multiply a data set by a perfectly localized spherical cap. However, because of the sharp boundaries associated with this naive window, such an operation was shown to possess unfavourable spectral leakage characteristics. In particular, as a result of the infinite bandwidth of this window, power from every input spherical harmonic was seen to leak into all of the harmonics of the windowed field. Thus, if the power spectrum of the input field possessed a high dynamic range, the windowed spectrum could be significantly biased.

We also showed that some of the unfavourable aspects of windowing with a perfect spherical cap could be ameliorated by truncating the window coefficients beyond a certain degree. One approach developed by Simons *et al.* (1997), and subsequently used by McGovern *et al.* (2002), Lawrence & Phillips (2003), Smrekar *et al.* (2003), and Hoogenboom *et al.* (2004), is to truncate the spherical cap coefficients exterior to a degree close to the edge of the main spectral lobe. Using our definition of the space-bandwidth product, their truncated windows can be described as having $N_0 = 1$. As the first spectral lobe of a spherical cap corresponds to $N_0 \approx 1.2$, these windows are truncated within the main spectral lobe. Armed with the definition of the space-concentration factor λ of eq. (11), and with the uncertainty relationship of eq. (30), we are now in a position to formally contrast the performance of such spectrally truncated spherical-cap windows with our own space-concentrated bandlimited tapers.

In the upper portion of Fig. 6 we plot the concentration factor as a function of the space-bandwidth product for the two types of windows. As noted in Section 4.1, the first space-concentrated taper is near perfectly concentrated for all values of $N_0 \geq 2$. In contrast, the spectrally truncated spherical-cap windows are less well concentrated, and only slowly approach unity. Indeed, even for $N_0 = 5$, the concentration factor of this latter window is only ~ 0.98 . For the particular case of $N_0 = 1$, both types of windows are seen to be poorly concentrated, and only possess a concentration factor between 0.90 and 0.92. Thus, if these windows were to be used to spatially localize a geographic region on the sphere, up to 10 per cent of the signal could originate from outside the concentration region.

In the lower portion of Fig. 6 we plot the spherical uncertainty product, $\sigma_z \sigma_L$, as a function of the space-bandwidth product, N_0 , for both types of windows. As noted in Section 4.3, the first space-concentrated taper nearly attains the lower limit imposed

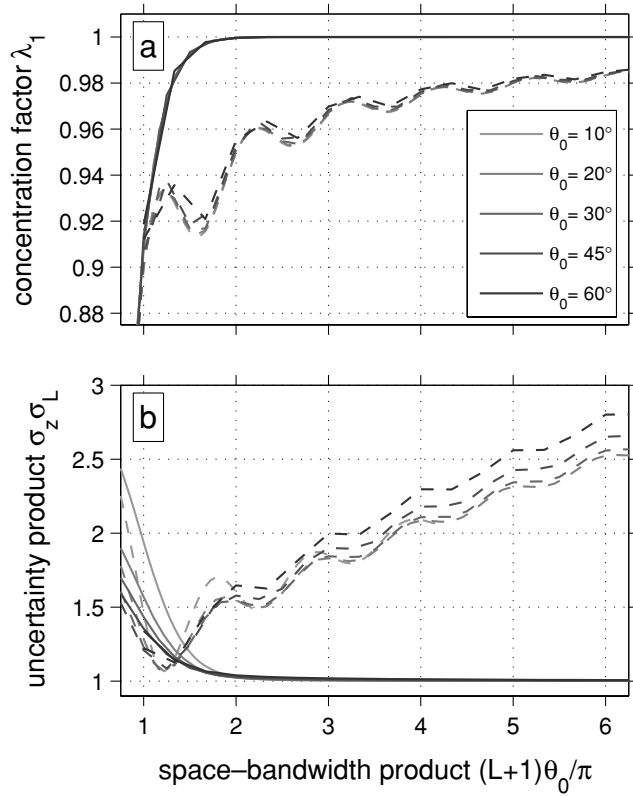


Figure 6. Comparison of the first space-concentrated tapers (solid lines) with those of a spectrally truncated spherical cap (dashed lines). (a) Concentration factor plotted as a function of the space–bandwidth product for several values of θ_0 and L . (b) Spherical uncertainty product plotted as a function of the space–bandwidth product N_0 .

by the uncertainty product for all values of $N_0 \geq 2$. In contrast, for the spectrally truncated spherical-cap windows, the uncertainty product at first decreases, but then becomes progressively worse for all values of the space–bandwidth product greater than about 1.2. While the spectrally truncated spherical cap does perform better than our windows in a small interval near this minimum value, it is always inferior in terms of its spatial concentration factor.

5 MULTITAPER (CROSS-)SPECTRAL ANALYSIS

It is well known in Cartesian analysis that the periodogram of a function, the squared Fourier coefficients of an unwindowed time-limited signal, does not yield a robust representation of its power spectrum (Percival & Walden 1993, Chap. 6). The drawbacks associated with a finite data length can be alleviated by forming a modified periodogram of the data tapered with a suitable window function. Thomson (1982) has also shown that it is often preferable to average the results obtained from several orthogonal tapers. His multitaper technique has many attractive attributes (e.g. Percival & Walden 1993, Chap. 7) and is directly applicable to our problem of localized spectral estimation on the sphere.

One benefit of an analysis with multiple orthogonal tapers is related to the fact that the combined energy of all data tapers is more evenly spread across the concentration region than that of any single window. As the energy of a single window is distributed non-uniformly over the data, the single-taper direct spectral estimates will never be completely representative of the region of interest; wherever the window energy approaches zero, the corresponding

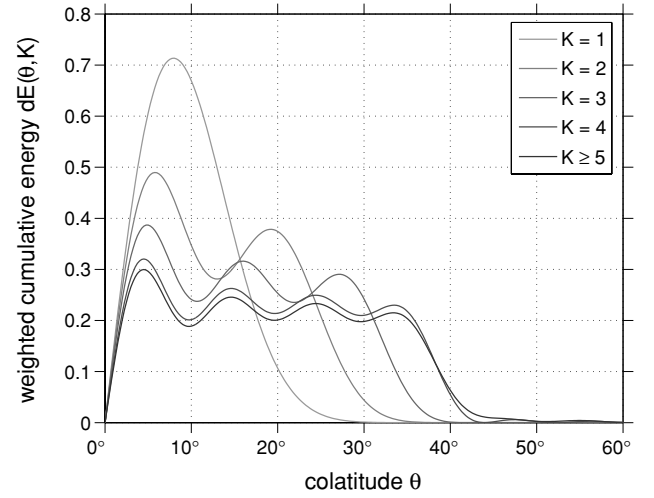


Figure 7. Cumulative energy density distribution of K band limited axisymmetric data windows with $\theta_0 = 40^\circ$, $L = 18$ and $N_0 = 4$. More tapers ensure a more uniform energy density over the concentration region.

data will be downweighted and ultimately discarded. If, however, an average is formed of several direct spectral estimates from orthogonal tapers, then the energy coverage of the data will be more uniform (Walden 1990).

We illustrate the spatial coverage of our bandlimited axisymmetric data windows in Fig. 7 by plotting the normalized cumulative energy density as a function of colatitude θ of a set of K tapers:

$$dE(\theta, K) = 2\pi \sum_{k=1}^K \lambda_k [h^{(k)}(\theta)]^2 \sin \theta \bigg/ \sum_{k=1}^K \lambda_k. \quad (31)$$

Here, the energy of each taper is weighted by its corresponding eigenvalue in order to minimize the contribution of the poorly concentrated tapers. An alternative weighting scheme would be to average only the first $N_0 - 1$ data tapers that are always near perfectly concentrated. When the first window is used alone ($K = 1$), we notice a heavily downweighted energy density near the pole and near the rim of the polar cap at θ_0 . However, as more data tapers are included ($K > 1$), the energy distribution progressively covers the concentration region more evenly (see also Simons *et al.* 2005). Thus, as in the Cartesian case, it is reasonable to expect that the average of multiple direct spectral estimates from an entire set of orthogonal tapers would be more representative of the desired concentration region than any given single taper.

In this section, we extend the Thomson (1982) technique of multitaper spectral analysis to the sphere using our newly constructed windows. First, for the case where the spectral coefficients of a global field are governed by a stationary stochastic process, we will quantify how windowed direct spectral estimates relate to the global power spectrum. Second, we show that the direct spectral estimates based on individual orthogonal tapers are approximately uncorrelated, and that the multitaper estimate of a single realization of a stochastic process approaches its theoretical value as more tapers are included (by increasing N_0). Third, we discuss ways to remove the bias present in the localized spectrum. Finally, we describe how to obtain estimates of the spectral admittance and coherence between two data sets.

5.1 Bias of a localized (cross-)spectral estimate

In Cartesian multitaper spectral analysis the true spectral power of a field is estimated by averaging windowed spectral estimates

from several orthogonal tapers (Thomson 1982). This is possible because the individual spectral estimates are asymptotically unbiased when the input signal is a smoothly varying stochastic process (e.g. Percival & Walden 1993, Chap. 6). Our understanding of spectral analysis in the Fourier domain is facilitated because the windowed power relates to the coefficients of data and windows by a convolution in the spectral domain. Here, we show that on the sphere, the power spectrum of a function windowed by our axisymmetric tapers is a nearly unbiased estimate of the global spectrum when the spectral coefficients of the data are governed by a white stochastic process. The bias can be significant when the spectrum of the data is red, but it is easily quantifiable as the effect of an operation that is reminiscent of a spectral convolution.

To show this, consider two functions on the sphere, $f(\Omega)$ and $g(\Omega)$, each multiplied by an axisymmetric window $h(\theta)$. The product of the coefficients of the windowed fields at degree l and order m is given by

$$\begin{aligned} \Phi_{lm} \Gamma_{lm} &= \frac{1}{4\pi} \int_{\Omega} [h(\theta) f(\Omega)] Y_{lm}(\Omega) d\Omega \\ &\times \frac{1}{4\pi} \int_{\Omega} [h(\theta) g(\Omega)] Y_{lm}(\Omega) d\Omega, \end{aligned} \quad (32)$$

and the cross-power spectrum of these windowed fields is, as previously defined in eq. (9), equal to

$$S_{\Phi\Gamma}(l) = \sum_{m=-l}^l \Phi_{lm} \Gamma_{lm}. \quad (33)$$

Before proceeding any further, it is useful to comment on the meaning of the above equation. By definition, $S_{\Phi\Gamma}$ is the global cross-power spectrum (expressed in global spherical harmonic basis functions) of the windowed fields $\Phi(\Omega) = h(\theta)f(\Omega)$ and $\Gamma(\Omega) = h(\theta)g(\Omega)$. However, as a result of the windowing procedure, this power spectrum is based almost exclusively on data within the concentration region. For this reason, we refer to $S_{\Phi\Gamma}$ as a localized direct power spectral estimate. When the global fields $f(\Omega)$ and $g(\Omega)$ are stationary, it is intuitive to expect that the localized power spectrum $S_{\Phi\Gamma}$ will be a function of, and approximate, the true global power spectrum S_{fg} ,

$$S_{\Phi\Gamma} \approx S_{fg}, \quad (34)$$

in a yet to be determined way. In essence, a local subset of the data is used to estimate the global (cross-)power spectrum of the stochastic process that generated them. When $f(\Omega)$ and $g(\Omega)$ are not stationary, we may think of eq. (34) as approximating the global power spectrum as if the data within the concentration region were representative of the entire sphere, thus enabling non-stationary analysis by assuming local stationarity over the concentration region.

If global data are to be regarded as governed by an underlying stochastic process, we must for the moment assume that many realizations of such a process are available for analysis. This will enable us to calculate the theoretically expected values of the localized spectra by averaging over all possible realizations of the random variables. How well a single realization of a stochastic process can be approximated by these expectation values will be illustrated in Section 5.2.

In order to obtain an analytic expression for the expected value of the localized power spectrum, we will make three simplifying assumptions. First, we will assume that $f(\Omega)$ and $g(\Omega)$ are fields generated by globally uniform and isotropic processes. This implies that their localized power spectra are invariant under a rotation of the spatial coordinates, and as such, the centring of our windows

on $\theta = 0^\circ$ is by no means restrictive. Second, we will assume that the coefficients f_{lm} are random variables with zero mean and with a variance that only depends upon the spherical harmonic degree l . Finally, we will assume that the spectral coefficients of $f(\Omega)$ and $g(\Omega)$ are linearly related by an isotropic admittance function $g_{lm} = Z_l f_{lm}$. Using these assumptions, we show in Appendix C that the expectation of the localized spectrum $S_{\Phi\Gamma}$, or its average value over all possible realizations of the random variable f_{lm} , is related to the true global spectrum S_{fg} as follows:

$$\langle S_{\Phi\Gamma}(l) \rangle = (2l+1) \sum_{j=0}^{L_h} h_j^2 \sum_{i=|l-j|}^{l+j} S_{fg}(i) \begin{pmatrix} i & j & l \\ 0 & 0 & 0 \end{pmatrix}^2. \quad (35)$$

For our spacelimited tapers, L_h is much greater than L . However, given the shape of their spectra, the sum over j can in practice be truncated at the effective bandwidth L .

Eq. (35) is valid for any arbitrary axisymmetric windowing function $h(\theta)$, and shows that the expected cross-power spectrum of the windowed fields is related to the global cross-power spectrum by an operation acting like a convolution in the spectral domain. The original spectrum S_{fg} is smoothed to yield the localized spectrum $S_{\Phi\Gamma}$, and the windowed cross-spectral estimate at degree l receives contributions from the global spectrum over the range $|l - L_h| \leq l \leq l + L_h$.

Independently from our work, the smoothing of the global power spectrum that occurs when a non-global data set is expanded in spherical harmonics has been studied by Peebles (1973) and Hauser & Peebles (1973) for astronomical purposes. Wandelt *et al.* (2001) and Hivon *et al.* (2002) have additionally given expressions analogous to our eq. (35) applicable to arbitrary non-axisymmetric windows for applications in cosmology.

We emphasize that since the global fields to be analysed were assumed to be stationary, eq. (35) cannot be used to obtain localized spectral estimates of an arbitrary non-stationary field. Nevertheless, this equation has two important uses. First, if one has a theoretical expression for the global (cross-)power spectrum that is based upon a stationary model, it can be used to obtain the theoretical windowed power spectrum. This spectrum can then be compared with localized spectral estimates of real observations. Second, given a localized (cross-)power spectrum of data on the sphere, it can be used to invert for a global power spectrum by assuming the localized region is representative of the whole sphere (see Section 5.3 for further details).

As demonstrated in the introduction of this section, the energy coverage of the concentration region becomes cumulatively more uniform when an increasing number of orthogonal data tapers is averaged. It is thus reasonable to suspect that a weighted average of spectra obtained by windowing with different tapers will be statistically more representative of the data than a localized direct spectrum obtained after windowing with a single taper. With this in mind, we define a multitaper spectral estimate as the average of $K \leq L + 1$ direct spectral estimates,

$$S_{\Phi\Gamma}^{(mt)}(l) = \sum_{k=1}^K a_k S_{\Phi\Gamma}^{(k)}(l), \quad (36)$$

where a_k is a generic weighting function, and $S^{(k)}$ is the spectrum obtained after windowing the data with the k th eigentaper $h^{(k)}$. In general, the optimal form of the weights will be dependent upon several factors including the spherical harmonic degree, the spectral bandwidth of the window, and the actual spectrum of the localized data (Thomson 1982; Percival & Walden 1993). Nevertheless, a

minimal requirement will be that poorly concentrated tapers are given near-zero weights.

In Appendix C, we derive expressions for the multitaper spectral estimates using our bandlimited tapers, when, as in eq. (31), the weights given to each individual spectrum are proportional to the eigenvalue of the eigentaper:

$$a_k = \frac{\lambda_k}{\sum_{i=1}^{L+1} \lambda_i}. \quad (37)$$

For the case of our family of bandlimited tapers, the expectation of the multitaper spectral estimate of eq. (36) is given by the expression

$$\langle S_{\Phi\Gamma}^{(mt)}(l) \rangle = \frac{(2l+1)}{\text{tr} \mathbf{D}} \sum_{j=0}^L D_{jj} \sum_{i=|l-j|}^{l+j} S_{fg}(i) \begin{pmatrix} i & j & l \\ 0 & 0 & 0 \end{pmatrix}^2, \quad (38)$$

where we have denoted the sum of the eigenvalues by the trace of the kernel \mathbf{D} (eq. 13). While eq. (38) is convenient for computational purposes and will be used throughout this subsection, a more conservative weighting scheme would be to equally weight only those eigenspectra that possess eigenvalues near unity (see Section 5.4).

Fig. 8 shows the relationship between the globally known power spectrum and the expectation of the localized individual and multitaper spectral estimates of eqs (35) and (38). We consider global

power spectra given by a power law

$$S_{ff}(l) \sim l^\beta, \quad (39)$$

and illustrate the theoretical effect of windowing the data with bandlimited tapers, $L = 17$, spatially concentrated within $\theta_0 = 30^\circ$, and thus characterized by a Shannon number $N_0 = 3$. For our first example, plotted in Fig. 8(a), we assume that this power spectrum is white ($\beta = 0$) and bandlimited ($L_f = 100$). We show the expectation of the localized estimates using each of the first three tapers, according to eq. (35), and the expectation of the multitaper estimates of eq. (38). In the wavelength range $l < L$, the windowed spectral estimates are seen to be heavily biased. Between $L \leq l \leq L_f - L$, the spectral estimates approach their globally known value as l increases, and here, the maximum spectral bias is always less than 50 per cent. Finally, for $L_f - L < l \leq L_f + L$, the spectral estimates approach zero. This is simply a result of the fact that this wavelength range receives increasing contributions from the input power spectrum beyond the bandwidth of the input data, L_f , which is here assumed to be zero. If the input data were not bandlimited to L_f , but were simply truncated at this degree, then the spectral estimates for $l > L_f - L$ would be unreliable.

Fig. 8(b) shows multitaper spectral estimates of the same white stochastic process using windows with several different values of

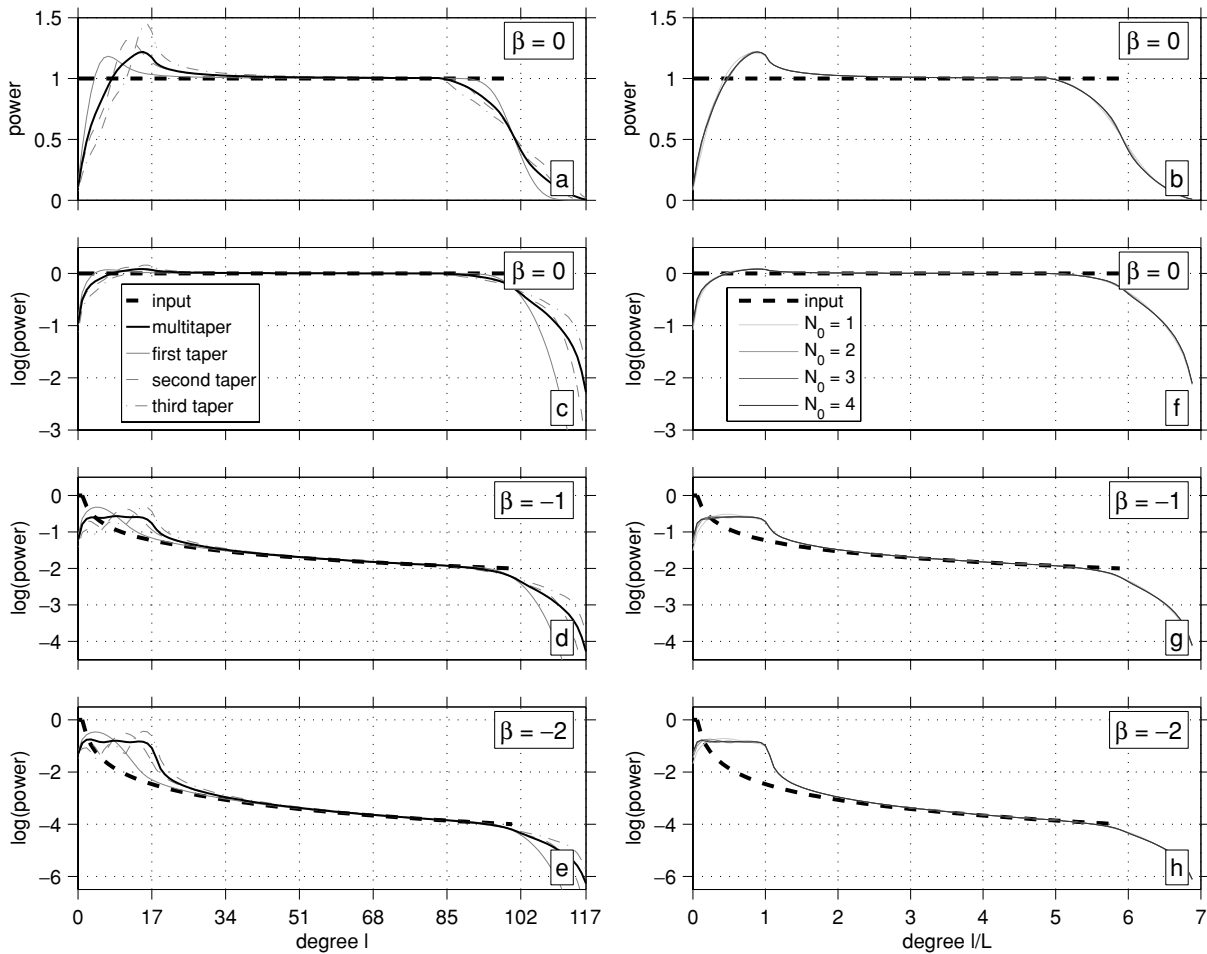


Figure 8. The expectation of individual and multitaper localized spectral estimates of power-law ($\sim l^\beta$) stationary stochastic processes. (a) Windowed spectrum using each of the first three bandlimited tapers (grey lines) and multitaper spectral estimate (solid black line), compared to the known global input spectrum (dashed black line). Here, $\beta = 0$, $\theta_0 = 30^\circ$, $L = 17$, and $N_0 = 3$. (b) Multitaper spectral estimates plotted as a function of l/L for several values of N_0 obtained by varying θ_0 with $L = 17$, for $\beta = 0$. (c–e) Same as (a) except for $\beta = 0, -1$ and -2 , on a logarithmic scale. (f–h) Same as (b), except for $\beta = 0, -1$ and -2 , on a logarithmic scale.

N_0 . Here, we varied the spatial concentration θ_0 but kept the bandwidth $L = 17$ unchanged. The results are shown as a function of l/L , the spherical harmonic degree scaled by the spectral bandwidth of the window. Regardless of the value of N_0 , all curves are nearly identical. Around the bandwidth of the window at $l = L$, the spectral bias is approximately 20 per cent, but this decreases to less than 3 per cent when $l = 2L$.

The remaining panels of Fig. 8 show that the situation is considerably different when the input spectra are red ($\beta < 0$). Figs 8(c)–(e) show plots analogous to Fig. 8(a) on a logarithmic scale for $\beta = 0, -1$ and -2 . The spectral bias can be significant for $0 < l \simeq L$, and the localized spectral estimates of non-white processes are seen to approach the global values with increasing l more slowly. Indeed, for $\beta = -2$, the spectral bias close to $l = L$ is about two orders of magnitude. This is simply a result of eq. (35) which shows that the windowed power spectrum is the result of smoothing the known global spectrum with the window coefficients: the more the input spectrum is red, the greater is the contribution of the low-degree high-power terms to the windowed spectrum. This results everywhere in a positive bias. Figs 8(f)–(h) illustrate this effect for varying values of N_0 , as in Fig. 8(b), but on a logarithmic power scale.

We must emphasize that, when using axisymmetric windowing functions, it is only possible to estimate isotropic (cross-)power spectra as a function of the degree l , but not the order m . Eq. (C13) in the Appendix can be used to show that the expected values of the windowed power as a function of both l and m are drastically different from the global values for individual harmonics. This can be interpreted as being due to the fact that while the cumulative energy density coverage of the concentration region, as defined in eq. (31), is nearly uniform as a function of θ for our axisymmetric windows (see Fig. 7), the energy per unit area within the region is not. Non-zonal windows are required to obtain a uniform energy coverage per unit area (see Simons *et al.* 2005), and with these, it should be possible to obtain spectral estimates for individual harmonics, thus allowing the estimation of (cross-)power spectral anisotropy (e.g. Simons *et al.* 2003).

5.2 Variance of a localized (cross-)spectral estimate

The results derived in Section 5.1 apply to ensemble averages: eqs (35) and (38) are applicable only to the case when the (cross-) spectral estimates are averaged over all possible realizations of the random variables f_{lm} . In most data applications, such as in planetary gravity and topography analysis, only a single realization will be available. The question thus arises as to whether the windowed spectral and cross-spectral estimates of a single realization of a stochastic process will be well approximated by the theoretical expectation values of the previous section. Here, we will first show that the spectral estimates obtained from our individual orthogonal tapers are approximately uncorrelated. By treating these individual spectral estimates as separate realizations of the data, we then show that the multitaper estimate approaches the theoretical value as the number of tapers used increases.

The variance of the multitaper estimate as a function of degree l and the number of employed tapers K can be calculated in a manner analogous to the Cartesian case (see Walden *et al.* 1994). The derivations are relegated to Appendix D, and here we only quote the result for the variance of the multitaper power-spectral estimate of a single field, $f(\Omega)$. While the equations in Appendix D are valid for any arbitrarily weighted multitaper estimate, here we focus on a multitaper estimate that is the unweighted average of $K \leq L + 1$

individual eigenspectra, that is,

$$S_{\Phi\Phi}^{(mt)}(l) = \sum_{k=1}^K a_k S_{\Phi\Phi}^{(k)}(l) \quad \text{and} \quad a_k = 1/K. \quad (40)$$

In this case, the variance of the multitaper estimate is

$$\text{var} \left\{ S_{\Phi\Phi}^{(mt)}(l) \right\} = \frac{2}{K^2} \sum_{j=1}^K \sum_{k=1}^K \sum_{m=-l}^l \langle S_{\Phi\Phi}^{(j,k)}(l, m) \rangle^2, \quad (41)$$

where

$$\begin{aligned} \langle S_{\Phi\Phi}^{(j,k)}(l, m) \rangle &= \sum_{l_1}^{L_h} h_{l_1}^{(j)} \sum_{l_2}^{L_h} h_{l_2}^{(k)} \sum_{i=i_{\min}}^{i_{\max}} S_{ff}(i) \\ &\times \frac{\sqrt{(2l_1+1)(2l_2+1)}}{2l+1} C_{i0l_1 0}^{i0} C_{i0l_2 0}^{i0} C_{iml_1 0}^{im} C_{iml_2 0}^{im}, \end{aligned} \quad (42)$$

with $i_{\min} = \max(|m|, |l - l_1|, |l - l_2|)$ and $i_{\max} = \min(l + l_1, l + l_2)$. We use Clebsch–Gordan coefficients instead of Wigner $3-j$ symbols for compactness, and made the assumption that the windowed coefficients follow a Gaussian distribution (see Appendix D for more details).

Fig. 9 shows the behaviour of the variance of the multitaper spectral estimates of a white stochastic process ($\beta = 0$). Again, this is illustrated with our bandlimited windows for $\theta_0 = 30^\circ$, $L = 29$, and $N_0 = 5$. Fig. 9(a) shows the standard deviation of the multitaper estimate as a function of the degree, l , and the number of tapers, K . It is immediately seen that the uncertainty of the estimates for $l \leq L$ is generally much larger than for $l > L$, and that the standard deviation of the spectral estimates decreases with increasing l . In particular, for $K = 1$, when $l = L$, the standard deviation of the power spectral estimate is about 46 per cent of the known global value, whereas when $l = 2L$, the uncertainty decreases to 32 per cent. This reduction relates to the fact that the number of equivalent wavelengths covering the concentration region increases with the degree l .

Cross-sections through Fig. 9(a) are plotted in Fig. 9(b) for the degree multiples $l = 29, 58$ and 87 (solid black lines), along with the standard deviations that would arise if the spectral estimates from our orthogonal tapers were completely uncorrelated, calculated by setting the cross-terms $j \neq k$ in eq. (41) to zero (dashed black lines). Both sets of curves are slightly different, implying that the spectral estimates are not perfectly uncorrelated. Nevertheless, given their similarity in magnitude and form, to a good approximation we can think of the spectral estimates obtained from our orthogonal tapers as approximately uncorrelated, as in the Cartesian case (Walden *et al.* 1994). Fig. 9(b) also shows that the decrease in the standard deviation is roughly proportional to $1/\sqrt{K}$ when the number of utilized tapers is smaller than N_0 (grey lines). The uncertainty would decrease exactly as $1/\sqrt{K}$ if the individual spectral estimates were both uncorrelated and possessed the same variance.

In Fig. 9(c), we plot the total covariance matrix of the individually tapered direct spectral estimates for one spherical harmonic degree, $l = 2L$. It is roughly diagonal, which supports our above assertion that the power spectral estimates associated with individual orthogonal tapers are approximately uncorrelated. We further observe that the variance of the power spectral estimates is largest for the best concentrated taper, at first decreases as the taper number increases, and increases again towards the worst concentrated tapers. In essence, by employing our best (worst) concentrated windows as data tapers, the associated spectral estimates are based upon only that limited portion of the data which resides within (outside of) the concentration region. In contrast, the energy distribution of

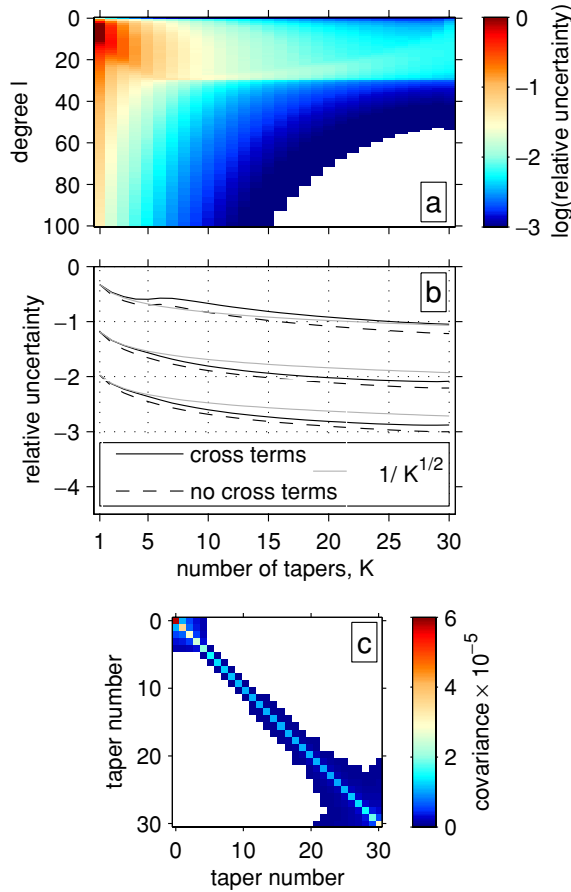


Figure 9. (a) Standard deviation at degree l of localized multitaper spectral estimates of a $\beta = 0$ power-law process, as a function of the number of tapers used, K , when $\theta_0 = 30^\circ$, $L = 29$ and $N_0 = 5$. Values less than $1/30$ of the maximum are in white. (b) Standard deviations as a function of K at the degree multiples $l = 29, 57$ and 87 (solid lines), the uncertainty that would arise if the individual estimates were uncorrelated (dashed lines), and the curve for uncorrelated estimates of constant variance, $1/\sqrt{K}$ (grey lines), scaled by the variance at $K = 1$. (c) Entire covariance matrix of tapered direct spectral estimates for $l = 2L$. Values below $1/100$ of the maximum are in white.

the intermediate tapers more evenly covers the entire sphere, which bases these estimates on a relatively larger quantity of data. If the input data were stationary, any data taper could be used for obtaining spectral estimates. However, for non-stationary data only those tapers that are near perfectly concentrated should be employed. We have confirmed that the basic properties of Fig. 9 hold when the input spectra are red ($\beta < 0$). In particular, if the standard deviation is scaled by the input power spectrum, we have confirmed that it is approximately independent of the spectral slope β when $l \gg L$.

Finally, in Fig. 10, we address the question as to whether the windowed spectrum of a single realization of a stochastic process approximates its theoretical expectation. We plot the eigenvalue-weighted multitaper (thick solid lines) and first four windowed (thin solid lines) spectrum estimates and compare these to the theory of eqs. (35) and (38) (dashed lines), for a single realization of a white Gaussian stochastic process with $\theta_0 = 30^\circ$, $L = 29$ and $N_0 = 4$. The individual spectral estimates for a single taper vary by almost ± 50 per cent about the expected value and only slightly improve as l increases, consistent with the theoretical results presented in Fig. 9. However, since we have shown that the spectra are approximately

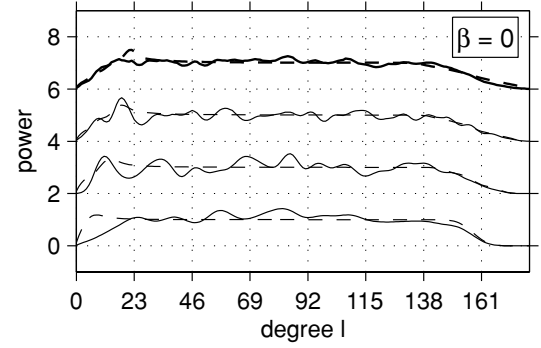


Figure 10. Individual and multitapered power spectra for a single realization of a $\beta = 0$ power-law stochastic processes (solid lines), and their theoretical expectations (dashed lines), for $\theta_0 = 30^\circ$, $L = 23$, and $N_0 = 4$. From bottom to top, spectra are for the first three tapers (thin lines) and multitaper estimate (thick lines). The input power spectrum is defined to be unity for $L_f \leq 157$, and all curves are offset for clarity.

uncorrelated, they can be thought of as each representing a different realization of the underlying stochastic process. For $l > L$ the theoretical expectation values for each taper are nearly identical, and thus we should expect the multitaper estimate of these spectra to resemble its theoretical expectation value (which is obtained by averaging over all possible realizations of the random variables) more closely than any of the individual estimates. This is indeed seen to be the case for the multitaper estimate of the example plotted in Fig. 10.

5.3 Inverting for the global power spectrum

We showed in Section 5.1 that the expectation of the windowed power spectrum of stationary data relates to the global spectrum by a simple smoothing operation. Here we address the question of whether it is possible to invert for the global spectrum from a knowledge of the smoothing operation and the expectation of a measured windowed spectrum. We emphasize that (geophysical) data distributed on a sphere are in general non-stationary. Nevertheless, in that case, the inverted global power spectrum could be interpreted as the spectrum that would arise if the localized data were representative of the entire sphere.

Since the smoothing operations of eqs (35) and (38) are linear, it will be useful to proceed using matrix notation. We define $\mathbf{S}_{\Phi\Gamma}^{(k)}$ to be a vector of the windowed cross-spectral estimates measured up to a maximum degree $L_{\Phi\Gamma}$, which thus has a dimension of $L_{\Phi\Gamma} + 1$. These windowed cross-spectral estimates depend upon the first $L_{\Phi\Gamma} + L$ degrees of the unknown global cross-power spectrum which we collect in the vector \mathbf{S}_{fg} of dimension $L_{\Phi\Gamma} + L + 1$. After relabeling indices and using the symmetry properties of the Wigner 3- j symbols, eq. (35) can be written as

$$\langle \mathbf{S}_{\Phi\Gamma}^{(k)} \rangle = \mathbf{M}^{(k)} \mathbf{S}_{fg}, \quad (43)$$

where the elements of the $(L_{\Phi\Gamma} + 1) \times (L_{\Phi\Gamma} + L + 1)$ matrix $\mathbf{M}^{(k)}$ are given by

$$M_{ij}^{(k)} = (2i + 1) \sum_{l=0}^L (h_l^{(k)})^2 \begin{pmatrix} i & j & l \\ 0 & 0 & 0 \end{pmatrix}^2, \quad (44)$$

with indices i and j starting at 0. Eq. (43) is simply the matrix version of eq. (35). Analogously, the expectation of the multitaper spectral estimate is related to the global spectrum by

$$\langle \mathbf{S}_{\Phi\Gamma}^{(mt)} \rangle = \mathbf{M}^{(mt)} \mathbf{S}_{fg}, \quad (45)$$

where

$$\mathbf{M}^{(mt)} = \sum_{k=1}^K a_k \mathbf{M}^{(k)}. \quad (46)$$

Eq. (45) is the matrix version of eq. (38). As there are always fewer windowed power spectral coefficients than global coefficients in eqs (43) and (45), it is fundamentally impossible to uniquely invert for the true global spectrum without imposing additional constraints, such as positivity or a minimum norm. Methods for solving the non-negative least squares problem can be found in Lawson & Hanson (1995, Chap. 23).

Independently from our work, in cosmology, \mathbf{M} is known as a mode coupling matrix (e.g. Wandelt *et al.* 2001; Hivon *et al.* 2002). Truncating this matrix to be square, it is possible under some conditions to obtain a deconvolved unbiased estimate of the power spectrum by simple matrix inversion (Efstathiou 2004). For a binary window masking contaminating sky sources, this approach has been followed by Hinshaw *et al.* (2003).

5.4 Localized admittance and coherence estimation

If the spectral coefficients of a function on the sphere are realizations of a white Gaussian process, we have seen that the localized multitaper spectral estimate of eq. (36) obtained by windowing with several orthogonal tapers is approximately unbiased. While the spectral estimate can be significantly biased when the global spectrum is red, this bias is easily quantifiable via eqs (35) and (38), and can potentially be removed by inversion. Here, we describe how localized spectral and cross-spectral estimates of two fields defined on the sphere can be used to estimate their localized spectral admittance and coherence functions.

In the foregoing we have generally defined the multitaper spectral estimate as a weighted sum of spectra from windowed data, the weight given by the eigenvalue of the window (eqs 36–37). The explicit assumption of stationary input fields renders the exact manner in which tapered estimates are weighted unimportant: employing a taper with poor concentration properties has little impact when the data are governed by the same process everywhere. In geophysical applications this is unlikely to ever be the case, and we thus advocate a stricter definition of the multitaper estimate in which only those tapers that are near perfectly concentrated are used ($\lambda_k \approx 1$). In particular, we suggest an average of only the first $K = N_0 - 1$ tapers:

$$S_{\Phi\Gamma}^{(mt)}(l) = \frac{1}{K} \sum_{k=1}^K S_{\Phi\Gamma}^{(k)} = \frac{1}{K} \sum_{k=1}^K \sum_{m=-l}^l \Phi_{lm}^{(k)} \Gamma_{lm}^{(k)}. \quad (47)$$

While the theoretical uncertainties associated with these estimates can be calculated via eqs (41) and (42), depending upon the chosen parameters, this can be somewhat computationally expensive. It is thus more practical to estimate the uncertainties from the multitaper sample variance (see also Thomson & Chave 1991):

$$\text{var} \left\{ S_{\Phi\Gamma}^{(mt)}(l) \right\} = \frac{1}{K-1} \sum_{k=1}^K \left(\sum_{m=-l}^l \Phi_{lm}^{(k)} \Gamma_{lm}^{(k)} - S_{\Phi\Gamma}^{(mt)}(l) \right)^2. \quad (48)$$

As was shown in Section 5.2, when the spherical harmonic degree is greater than the bandwidth L of the tapers, the variance decreases approximately as $1/K$ for $K < N_0$.

Next, consider that the spectral components of two fields are linearly related by an equation of the form

$$g_{lm} = Z(l) f_{lm} + I_{lm}, \quad (49)$$

where Z_l is an isotropic admittance function, and I_{lm} is either some form of measurement noise, or that portion of the signal that is not described by the admittance model. Multiplying both sides by f_{lm} , summing over all m , and rearranging yields the following expression for the admittance:

$$Z(l) = \frac{S_{fg}(l)}{S_{ff}(l)} - \frac{S_{If}(l)}{S_{ff}(l)}. \quad (50)$$

If the spectral coefficients of the noise, I , and the data, f are uncorrelated, which we will henceforth assume, the expectation of the last term is zero. The phase consistency between both fields is contained in the coherence function

$$\gamma(l) = \frac{S_{fg}(l)}{\sqrt{S_{ff}(l) S_{gg}(l)}}, \quad (51)$$

whose extreme values of 1 and -1 represent correlated and anti-correlated data sets, respectively. (Note that there are several naming conventions related to eq. (51). Whereas some use the term ‘coherence’ or ‘complex coherence’, others refer to the ‘coherence’ as the absolute magnitude of this equation. Alternatively, some use the term ‘correlation’ and reserve the word ‘coherence’ for the correlation squared. In any case, we emphasize that squaring or taking the absolute value of this equation discards information.)

The admittance and coherence localized to a given concentration region can be calculated by replacing the global power and cross-power spectra with their localized equivalents. These can be estimated in multitaper fashion by eq. (47), which gives rise to the following localized admittance and coherence estimates

$$Z_{\Phi\Gamma}^{(mt)}(l) = \frac{S_{\Phi\Gamma}^{(mt)}(l)}{S_{\Phi\Phi}^{(mt)}(l)}, \quad (52)$$

$$\gamma_{\Phi\Gamma}^{(mt)}(l) = \frac{S_{\Phi\Gamma}^{(mt)}(l)}{\sqrt{S_{\Phi\Phi}^{(mt)}(l) S_{\Gamma\Gamma}^{(mt)}(l)}}. \quad (53)$$

We emphasize that, since the localized spectral and cross-spectral estimates could be severely biased when the global power spectrum is red, the corresponding localized estimates of the admittance and coherence could be similarly biased as well (see Pérez-Gussinyé *et al.* 2004, for a Cartesian example of this phenomenon).

For modeling purposes, there are three ways in which the potential bias in the multitaper admittance and coherence estimates could be accounted for. First, if a simple analytic formula for the admittance function exists, synthetic coefficients of g_{lm} could be generated by use of eq. (49), and synthetic localized admittance and coherence estimates could be computed using eqs (52) and (53). Any inherent bias that might be associated with the windowing procedure would thus similarly affect the model and data. Indeed, for this case, it is irrelevant whether or not f_{lm} is a random variable as was assumed previously. If the theoretical admittance function depends upon one or more model parameters, then these could be inverted for by comparing the goodness-of-fit between the synthetic and real-data estimates. This approach has been employed in several studies (e.g. Simons *et al.* 1997; McGovern *et al.* 2002; Kido *et al.* 2003; Lawrence & Phillips 2003; Smrekar *et al.* 2003; Hoogenboom *et al.* 2004).

For certain theoretical models, however, the relationship between two fields is statistical in nature, and no deterministic admittance function exists that relates the coefficients f_{lm} and g_{lm} . For instance, in the Cartesian admittance and coherence model of Forsyth (1985), a statistical model was used to obtain analytic expressions for the expectation value of the power and cross-power spectra of the two

fields as a function of wavelength. Using these relationships, an isotropic admittance function was then calculated using the Cartesian analogues to eqs (52) and (53). If analytic expressions for the power and cross-power spectra are available in the spherical harmonic domain, then eq. (35) can be used to calculate the expected windowed spectra, and from these, the windowed admittance and coherence via eqs (52) and (53).

Finally, if windowed (cross-)power spectra have been calculated using the multitaper method, one can attempt to invert the windowed spectrum for the global spectrum, as discussed in Section 5.3. This would in principle yield unbiased (cross-)power spectral estimates which could then be compared directly to a theoretical stationary model.

6 DISCUSSION

6.1 An example

As a demonstration of the benefits of using our optimally concentrated data tapers for the purpose of localized spectral estimation, we apply a multitaper spectral analysis to the Earth's free-air gravity field. Differences and similarities between our approach and that using the single window of Simons *et al.* (1997) highlight several key factors that must be carefully considered when performing such an analysis.

The data set used here is the free-air gravity model EGM96 (Lemoine *et al.* 1998), from which we have removed the degree-2 zonal term that is dominated by the rotational flattening of the Earth. As shown in Fig. 11, we concentrate on two distinct geographic regions: Hudson Bay, which was previously analysed by

Simons & Hager (1997), and a typical locale in the South Pacific. These regions were chosen because of their contrasting gravity signatures; in the first, a large negative anomaly is present that is the result of incomplete post-glacial rebound, whereas in the second, the gravity field is relatively nondescript, lacking any apparent short- or long-wavelength structure.

For the multitaper analyses, we use our bandlimited tapers with a concentration region of $\theta_0 = 20^\circ$ and a spectral bandwidth of $L = 35$, corresponding to $N_0 = 4$. For the analyses using the single window of Simons *et al.* (1997), the same concentration region is used, but the spectral bandwidth of the window is equal to $L = 8$, corresponding to $N_0 = 1$. We plot in the upper-left panel of Fig. 11 the gravity field localized to Hudson Bay using the spectrally truncated spherical cap window of Simons *et al.* (1997), and in the upper-centre panel, the same field is plotted windowed by our first optimally concentrated taper. As the former window is concentrated only at ~ 90 per cent in the region of interest, significant spatial leakage is observed. The localized spectrum estimates of these two approaches are plotted in the upper-right panel. For the multitaper spectrum estimate, we average the first three eigenspectra and plot the corresponding sample variance (see Section 5.4). Since the variance of the spectral estimates is predicted to be large for degrees less than the bandwidth of the localizing window (see Section 5.2), a direct comparison of the two spectra is only valid for degrees greater than $L = 35$. For these degrees, the spectrum estimate based on the single window of Simons *et al.* (1997) is seen to lie generally within the error bounds of our multitaper estimate.

In contrast to the case of Hudson Bay, a significant discrepancy between the two methods is observed for the region in the South

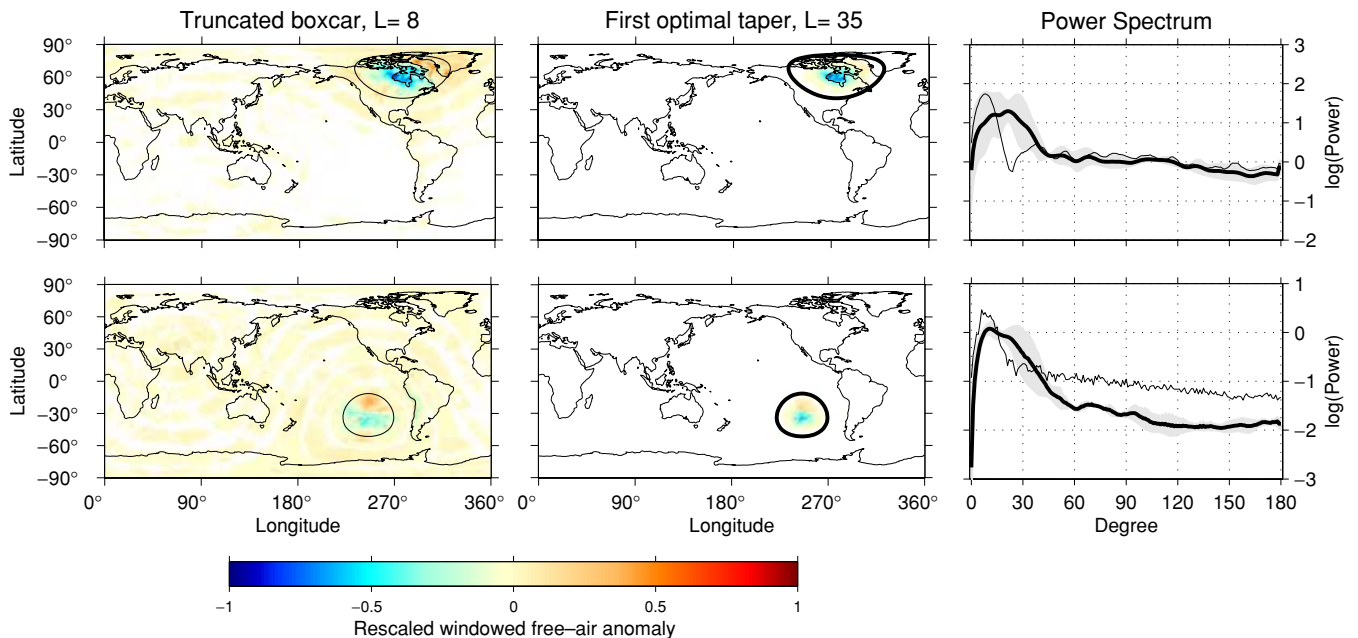


Figure 11. Localized spectral analyses of the Earth's free-air gravity field for two geographically distinct areas using the spectrally truncated spherical cap window of Simons *et al.* (1997) and the multitaper spectral analysis approach. For the spectrally truncated spherical cap (thin grey lines), $\theta_0 = 20^\circ$ and $L = 8$, corresponding to $N_0 = 1$, whereas for the multitaper analysis (thick black lines), $\theta_0 = 20^\circ$ and $L = 35$, corresponding to $N_0 = 4$. Spatial renditions of the windowed fields (scaled from -1 to 1 , and with absolute values less than $1/200$ set to zero) are shown for the spectrally truncated spherical cap (left column), and our first optimally concentrated bandlimited window (centre column). The multitaper spectrum estimate (right column) is obtained by averaging the first three eigenspectra, and the corresponding sample variance is plotted in light grey. While the windowed power spectrum based on the spectrally truncated boxcar window is similar to the multitaper estimate for the region of Hudson Bay for degrees larger than the window bandwidth of $L = 35$, it differs by almost an order of magnitude for a typical region in the South Pacific.

Pacific. We first remark that the multitaper power spectral estimates for the South Pacific region are about two orders of magnitude less at all degrees than the corresponding values for Hudson Bay. This is simply a result of the lack of any large-amplitude gravity anomalies in the concentration region. In comparing the multitaper spectrum estimate with that from the single window of Simons *et al.* (1997), it is readily apparent that the two spectra differ significantly, by almost an order of magnitude. This discrepancy is a result of both the poor localizing properties of the single window, as well as the small amplitude of the signal within the region of interest. In particular, in the example of Hudson Bay, the large signal amplitude within the concentration region partially mitigates against the poor localization properties of the single window. In contrast, for the area in the South Pacific, the signal within the concentration region is very low, and this gives greater importance to the down-weighted features exterior to the concentration region, such as the subduction-zone signature of western South America.

This example illustrates the two main advantages of our multitaper spectral analysis approach. First, as a consequence of using multiple orthogonal localizing windows, it is possible to obtain an estimate of the uncertainty of the localized spectral estimates. Second, the near perfect concentration properties of the data windows ensure that the spectral estimates are not influenced by data exterior to the concentration region. However, as is discussed in the following section, the large spectral bandwidths that are required for a multitaper analysis can sometimes place restrictive limits on the number of useful spectral estimates that can be analysed.

6.2 Working with low resolution data

To apply a localized multitaper spectral analysis on the sphere, several well-localized windows are desired. As a result of the definition of the space–bandwidth product in eq. (25), if the concentration region, θ_0 , remains fixed in size, then the spectral bandwidth of the windows must increase with the number of near perfectly localized data tapers. As an example, consider the case of $\theta_0 = 20^\circ$. If only one near perfectly concentrated taper was needed, the space–bandwidth product N_0 must be chosen to be about 2 and the corresponding spectral bandwidth would be $L = 17$. However, if three tapers were desired, N_0 must be around 4, with a corresponding spectral bandwidth of $L = 35$.

Such large bandwidths pose a fundamental problem when working with low-resolution data sets. While models of a global gravity field may only be expressed up to a maximum degree L_f , this does not imply that the gravity coefficients beyond this degree are zero—they are just not known. After windowing such a field by a taper of bandwidth L , the resultant coefficients are only reliable up to $L_f - L$ (see eq. 10), and the first L coefficients exhibit large uncertainties (see Section 5.2). In practice, the first $L + 2$ windowed coefficients of a planetary gravity and topography field will be biased by the contribution resulting from its rotational flattening (i.e. the coefficient of the $Y_{2,0}$ harmonic). Thus, using the above example of $\theta_0 = 20^\circ$ with three tapers, for an input field expressed up to $L_f = 85$, as is the case for the Martian gravity field (e.g. Yuan *et al.* 2001), only the restricted wavelength range $35 \leq l \leq 50$ could be analysed robustly. If four tapers were desired, none of the windowed coefficients would be usable.

6.3 Localized spherical spectral analysis: A recipe

Our method of estimating localized (cross-)power spectra on the sphere is straightforward and can be summarized in the following

steps that are applicable to a typical application in planetary gravity–topography analysis.

- (1) Select the class of localizing windows. If the data set is only locally known on the sphere, the spacelimited windows are most appropriate. Conversely, if the data set is globally known, the bandlimited space-concentrated tapers will minimize the effects of spectral leakage when the input spectrum is red.

- (2) For a given cap size θ_0 , choose the bandwidth, L . If only one taper is desired, this can be chosen such that a specified concentration factor is obtained (say, 99 per cent). Alternatively, if K well-concentrated tapers are desired to reduce the multitaper variance, take $L = (K + 1)\pi/\theta_0 - 1$.

- (3) Calculate the coefficients of the tridiagonal matrix \mathbf{T} in Appendix A and diagonalize it: its eigenvectors are the spherical harmonic coefficients of the bandlimited space-concentrated tapers whose concentration factors λ can be computed from eqs (11) or (14).

- (4) Rotate the window to the region of interest. This can be efficiently performed in the spherical harmonic domain by use of Wigner rotation matrices which can be calculated by standard recursion relations (e.g. Blanco *et al.* 1997).

- (5) Expand the windows to the space domain and multiply them with the data. If spacelimited tapers are desired, set the windowed fields to zero outside of the concentration region at this time.

- (6) Expand the windowed fields in spherical harmonics. This is easily performed by Fourier transformation over longitude and (Gauss–Legendre) integration over latitude (e.g. Driscoll & Healy 1994; Sneeuw 1994).

- (7) For multitaper analysis, repeat the above steps for each of the K tapers. Multitaper spectral and cross-spectral estimates are computed by simple averaging.

- (8) Calculate localized estimates of the (cross-)power spectra by eq. (33), and admittance and coherence by eqs (52) and (53). These can be compared with theoretical models in one of three ways. First, synthetic fields can be generated from the theoretical model and then windowed in the same manner as the data. Second, if a theoretical model is available only for the power and cross-power spectra of the two fields, localized admittance and coherence models can be computed by use of eq. (35). Finally, one could invert the windowed spectra for the global spectra, subject to positivity constraints (e.g. Lawson & Hanson 1995), to calculate unbiased admittance and coherence functions unaffected by the windowing procedure.

7 CONCLUSIONS

Localized spectral estimates on the sphere can be obtained by multiplying data by a suitable windowing function. The ideal windows for this purpose are concentrated within θ_0 in the spatial and within L in the spectral domain. We have obtained such windows by solving Slepian's concentration problem on the sphere. Two classes of windows exist: bandlimited functions that are optimally concentrated in space, and spacelimited functions that are optimally concentrated within a spherical harmonic bandwidth. Their spherical harmonic coefficients and corresponding concentration factors solve a simple eigenvalue equation. Each class yields a family of orthogonal windows, allowing for multitaper analysis.

The spatial and spectral concentration problems are complementary: they give rise to identical functions within the concentration domain, they possess identical eigenvalue spectra, and the spectral coefficients of the windows within the prescribed bandwidth differ

only by a multiplicative factor corresponding to their eigenvalue. The properties of the windows are almost entirely described by their space–bandwidth product, $N_0 = (L + 1)\theta_0/\pi$, and only the first $N_0 - 1$ windows are near perfectly concentrated. In addition, the first taper nearly attains a lower bound imposed by a spherical uncertainty relationship.

If the data sets to be localized are stationary and stochastic, it is possible to quantify the relationship between the global spectrum and the localized direct spectral estimates. In particular, if the input spectrum is white, the expectation of the localized spectral estimates (averaged over all realizations of the random variables) is nearly unbiased for all degrees greater than L . In contrast, when the input spectrum is red, the spectral estimates close to L can be severely biased, but this bias is easily quantifiable. Averaging the spectra obtained from several orthogonal tapers reduces the estimation variance, and the multitaper spectral estimate is more representative of the concentration region than any single individual taper as a result of its more even energy coverage.

In many situations, such as planetary gravity and topography analysis, only one realization of a stochastic process is available. As the individual spectral estimates obtained from orthogonal tapers are nearly uncorrelated, each spectral estimate can be thought of as an individual realization of the underlying stochastic process. While a single taper underperforms at estimating the localized spectrum, when several estimates from orthogonal tapers are averaged, the resulting multitaper estimate approaches the expected spectrum as the number of tapers used increases.

Finally, model parameters can be found by analysing the goodness-of-fit between a localized spectral estimate of a data set with a similarly localized theoretical model. A classic application in geophysics is the calculation of localized admittance and coherence spectra between gravity anomalies and topography, which is a primary source of information on planetary lithospheres.

ACKNOWLEDGMENTS

We thank Bruce Bills, Tony Dahlen, Francis Nimmo, Sofia Olhede, Richard Holme, and an anonymous reviewer for comments and discussions that improved the clarity of this paper. We added the references to cosmology as the paper went to press after being alerted to the parallel endeavours in this field by Luis Tenorio and Phil Stark. This work was supported by the Programme National de Planétologie and the European Community's Improving Human Potential Program under contract RTN2-2001-00414, MAGE. FJS acknowledges the financial support of IPGP. Fortran 90/95 and Matlab programs for performing the calculations in this paper are available from the authors. This is IPGP contribution 2066.

REFERENCES

- Albertella, A., Sansò, F. & Sneeuw, N., 1999. Band-limited functions on a bounded spherical domain: the Slepian problem on the sphere, *J. Geodesy*, **73**, 436–447.
- Belleguic, V., Lognonné, P. & Wiczeorek, M., 2005. Constraints on the Martian lithosphere from gravity and topography data, *J. geophys. Res.*, in press.
- Blakely, R.J., 1995. *Potential Theory in Gravity and Magnetic Applications*, Cambridge Univ. Press, New York.
- Blanco, M.A., Flórez, M. & Bermejo, M., 1997. Evaluation of the rotation matrices in the basis of real spherical harmonics, *J. Mol. Struct. (Theochem)*, **419**, 19–27.
- Bluhm, R., Kostecký, V.A. & Tudose, B., 1995. Elliptical squeezed states and Rydberg wave packets, *Phys. Rev. A*, **52**, 2234–2244.
- Bronez, T.P., 1988. Spectral estimation of irregularly sampled multidimensional processes by generalized prolate spheroidal sequences, *IEEE Trans. Acoust. Speech Signal Process.*, **36**, 1862–1873.
- Byerly, W.E., 1893. *An Elementary Treatise on Fourier's Series and Spherical, Cylindrical, and Ellipsoidal Harmonics*, Ginn & Co., Boston, Mass.
- Carruthers, P. & Nieto, M.M., 1968. Phase and angle variables in quantum mechanics, *Rev. Mod. Phys.*, **40**, 411–440.
- Chevrot, S., Montagner, J.-P. & Snieder, R.K., 1998. The spectrum of tomographic Earth models, *Geophys. J. Int.*, **133**, 783–788.
- Dahlen, F.A. & Tromp, J., 1998. *Theoretical Global Seismology*, Princeton Univ. Press, Princeton, N. J.
- Daubechies, I., 1988. Time-frequency localization operators: A geometric phase space approach, *IEEE Trans. Inform. Theory*, **34**, 605–612.
- Daubechies, I. & Paul, T., 1988. Time-frequency localisation operators—a geometric phase space approach: II. The use of dilations, *Inv. Probl.*, **4**, 661–680.
- Driscoll, J.R. & Healy, D.M., 1994. Computing Fourier transforms and convolutions on the 2-sphere, *Adv. Appl. Math.*, **15**, 202–250.
- Efstathiou, G., 2004. Myths and truths concerning estimation of power spectra: the case for a hybrid estimator, *Mon. Not. R. Astron. Soc.*, **349**, 603–626.
- Feller, W., 1971. *An Introduction to Probability Theory and Its Applications*, Vol. 2, 3 edn, Wiley, New York, NY.
- Forsyth, D.W., 1985. Subsurface loading and estimates of the flexural rigidity of continental lithosphere, *J. geophys. Res.*, **90**, 12 623–12 632.
- Freeden, W. & Windheuser, U., 1997. Combined spherical harmonic and wavelet expansion—a future concept in Earth's gravitational determination, *Appl. Comput. Harm. Anal.*, **4**, 1–37.
- González, J.A. & del Olmo, M.A., 1998. Coherent states on the circle, *J. Phys. A: Math. Gen.*, **31**, 8841–8857.
- Grünbaum, F.A., Longhi, L. & Perlstadt, M., 1982. Differential operators commuting with finite convolution integral operators: some non-abelian examples, *SIAM J. Appl. Math.*, **42**, 941–955.
- Hanssen, A., 1997. Multidimensional multitaper spectral estimation, *Signal Process.*, **58**, 327–332.
- Hauser, M.G. & Peebles, P.J.E., 1973. Statistical analysis of catalogs of extragalactic objects. II. The Abell catalog of rich clusters, *Astroph. J.*, **185**, 757–785.
- Hinshaw, G. *et al.*, 2003. First-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: the angular power spectrum, *Astroph. J. Supp. Ser.*, **148**, 135–159.
- Hipkin, R.G., 2001. The statistics of pink noise on a sphere: applications to mantle density anomalies, *Geophys. J. Int.*, **144**, 259–270.
- Hivon, E., Górski, K.M., Netterfield, C.B., Crill, B.P., Prunet, S. & Hansen, F., 2002. MASTER of the Cosmic Microwave Background anisotropy power spectrum: a fast method for statistical analysis of large and complex Cosmic Microwave Background data sets, *Astroph. J.*, **567**, 2–17.
- Homeier, H.H.H. & Steinborn, E.O., 1996. Some properties of the coupling coefficients of real spherical harmonics and their relation to Gaunt coefficients, *J. Mol. Struct. (Theochem)*, **368**, 31–37.
- Hoogenboom, T., Smrekar, S.E., Anderson, F.S. & Houseman, G., 2004. Admittance survey of type 1 coroneae on Venus, *J. geophys. Res.*, **109**, E03,002, doi:10.1029/2003JE002171.
- Jakob-Chien, R. & Alpert, B.K., 1997. A fast spherical filter with uniform resolution, *J. Comput. Phys.*, **136**, 580–584.
- Kaula, W.M., 1967. Theory of statistical analysis of data distributed over a sphere, *Rev. Geophys.*, **5**, 83–107.
- Kaula, W.M., 2000. *Theory of Satellite Geodesy: Applications of Satellites to Geodesy*, Dover, Mineola, NY.
- Kido, M., Yuen, D.A. & Vincent, A.P., 2003. Continuous wavelet-like filter for a spherical surface and its application to localized admittance function on Mars, *Phys. Earth planet. Inter.*, **135**, 1–14.
- Kowalski, K. & Rembieliński, J., 2000. Quantum mechanics on a sphere and coherent states, *J. Phys. A: Math. Gen.*, **33**, 6035–6048.
- Kowalski, K. & Rembieliński, J., 2002. On the uncertainty relations and squeezed states for the quantum mechanics on a circle, *J. Phys. A: Math. Gen.*, **35**, 1405–1414.

- Kowalski, K., Rembieliński, J. & Papaloucas, L.C., 1996. Coherent states for a quantum particle on a circle, *J. Phys. A: Math. Gen.*, **29**, 4149–4167.
- Landau, H.J. & Pollak, H.O., 1960. Prolate spheroidal wave functions, Fourier analysis and uncertainty—II, *Bell Syst. Tech. J.*, **40**, 65–84.
- Lawrence, K.P. & Phillips, R.J., 2003. Gravity/topography admittance inversion on Venus using niching genetic algorithms, *Geophys. Res. Lett.*, **30**, 1994, doi:10.1029/2003GL017515.
- Lawson, C.L. & Hanson, R.J., 1995. *Solving Least Squares Problems*, SIAM, Philadelphia.
- Lemoine, F. *et al.*, 1998. *The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96*, NASA Goddard Space Flight Cent., Greenbelt, MD.
- Liu, T.-C. & van Veen, B.D., 1992. Multiple window based minimum variance spectrum estimation for multidimensional random fields, *IEEE Trans. Signal Process.*, **40**, 578–589.
- Lowes, F.J., 1966. Mean-square values on sphere of spherical harmonic fields, *J. geophys. Res.*, **71**, 2179.
- Lowes, F.J., 1974. Spatial power spectrum of the main geomagnetic field and extrapolation to core, *Geophys. J. R. astr. Soc.*, **36**, 717–730.
- Luscombe, J.J. & Luban, M., 1998. Simplified recursive algorithm for Wigner 3j and 6j symbols, *Phys. Rev. E*, **57**, 7274–7277.
- Mallat, S., 1998. *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA.
- Masters, G. & Richards-Dinger, K., 1998. On the efficient calculation of ordinary and generalized spherical harmonics, *Geophys. J. Int.*, **135**, 307–309.
- McGovern, P.J. *et al.*, 2002. Localized gravity/topography admittance and correlation spectra on Mars: Implications for regional and global evolution, *J. geophys. Res.*, **107**, 5136, doi:10.1029/2002JE001854.
- Narcowich, F.J. & Ward, J.D., 1996. Nonstationary wavelets on the m-sphere for scattered data, *App. Comput. Harm. Anal.*, **3**, 324–336.
- Ould Kaber, S.M., 1996. A Legendre pseudospectral viscosity method, *J. Comput. Phys.*, **128**, 165–180.
- Peebles, P.J.E., 1973. Statistical analysis of catalogs of extragalactic objects. I. Theory, *Astroph. J.*, **185**, 413–440.
- Percival, D.B. & Walden, A.T., 1993. *Spectral Analysis for Physical Applications, Multitaper and Conventional Univariate Techniques*, Cambridge Univ. Press, New York.
- Pérez-Gussinyé, M., Lowry, A., Watts, A.B. & Velicogna, I., 2004. On the recovery of effective elastic thickness using spectral methods: Examples from synthetic data and from the Fennoscandian Shield, *J. geophys. Res.*, **109**(B10), 409, doi:10.1029/2003JB002788.
- Polyakov, A.S., 2002. Local basis expansions for linear inverse problem, *PhD thesis*, New York University.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P., 1992. *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed., Cambridge Univ. Press, Cambridge, UK.
- Schulten, K. & Gordon, R.G., 1975. Exact recursive evaluation of 3j-coefficients and 6j-coefficients for quantum-mechanical coupling of angular momenta, *J. Math. Phys.*, **16**, 1961–1970.
- Simons, F.J., Zuber, M.T. & Korenaga, J., 2000. Isostatic response of the Australian lithosphere: estimation of effective elastic thickness and anisotropy using multitaper spectral analysis, *J. geophys. Res.*, **105**, 19 163–19 184.
- Simons, F.J., van der Hilst, R.D. & Zuber, M.T., 2003. Spatio-spectral localization of isostatic coherence anisotropy in Australia and its relation to seismic anisotropy: Implications for lithospheric deformation, *J. geophys. Res.*, **108**, 2250, doi:10.1029/2001JB000704.
- Simons, F.J., Dahlen, F.A. & Wieczorek, M.A., 2005. Spatiospectral concentration on a sphere, *SIAM Rev.*, in press.
- Simons, M. & Hager, B.H., 1997. Localization of the gravity field and the signature of glacial rebound, *Nature*, **390**, 500–504.
- Simons, M., Solomon, S.C. & Hager, B.H., 1997. Localization of gravity and topography: constraints on the tectonics and mantle dynamics of Venus, *Geophys. J. Int.*, **131**, 24–44.
- Slepian, D., 1983. Some comments on Fourier-analysis, uncertainty and modeling, *SIAM Rev.*, **25**, 379–393.
- Smrekar, S.E., Comstock, R. & Anderson, F.S., 2003. A gravity survey of Type 2 coroneae on Venus, *J. geophys. Res.*, **108**, 5090, doi:10.1029/2002JE001935.
- Sneeuw, N., 1994. Global spherical harmonic-analysis by least-squares and numerical quadrature methods in historical perspective, *Geophys. J. Int.*, **118**, 707–716.
- Strang, G., 1986. *Introduction to Applied Mathematics*, Wellesley-Cambridge, Wellesley, MA.
- Swartztrauber, P.N., & Spatz, W.F., 2000. Generalized discrete spherical harmonic transforms, *J. Comput. Phys.*, **159**, 213–230.
- Swenson, S. & Wahr, J., 2002. Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time-variable gravity, *J. geophys. Res.*, **107**, 2193, doi:10.1029/2001JB000576.
- Swenson, S., Wahr, J. & Milly, P.C.D., 2003. Estimated accuracies of regional water storage variations inferred from the Gravity Recovery and Climate Experiment (GRACE), *Water Resources Res.*, **39**, 1223, doi:10.1029/2002WR001808.
- Szegő, G., 1975. *Orthogonal Polynomials*, 4 edn, American Mathematical Society, Providence, RI.
- Tegmark, M., 1995. A method for extracting maximum resolution power spectra from galaxy surveys, *Astroph. J.*, **455**, 429–438.
- Tegmark, M., 1996. A method for extracting maximum resolution power spectra from microwave sky maps, *Mon. Not. R. Astron. Soc.*, **280**, 299–308.
- Thomson, D.J., 1982. Spectrum estimation and harmonic analysis, *Proc. IEEE*, **70**, 1055–1096.
- Thomson, D.J. & Chave, A.D., 1991. Jackknifed error estimates for spectra, coherences, and transfer functions, in *Advances in Spectrum Analysis and Array Processing*, Vol. 1, chap. 2, pp. 58–113, edn Haykin, S., Prentice Hall, Englewood Cliffs, N. J.
- Tricomi, F.G., 1970. *Integral Equations*, 5 edn, Interscience, New York.
- Turcotte, D.L., Willemann, R.J., Haxby, W.F. & Norberry, J., 1981. Role of membrane stresses in the support of planetary topography, *J. geophys. Res.*, **86**, 3951–3959.
- Varshalovich, D.A., Moskalev, A.N. & Khersonskii, V.K., 1988. *Quantum theory of angular momentum*, World Scientific, Singapore.
- Voorhies, C.V., Sabaka, T.J. & Purucker, M., 2002. On magnetic spectra of Earth and Mars, *J. geophys. Res.*, **107**, 5034, doi:10.1029/2001JE001534.
- Walden, A.T., 1990. Improved low-frequency decay estimation using the multitaper spectral-analysis method, *Geophys. Prospect.*, **38**, 61–86.
- Walden, A.T., McCoy, E.J. & Percival, D.B., 1994. The variance of multitaper spectrum estimates for real Gaussian processes, *IEEE Trans. Signal Process.*, **2**, 479–482.
- Wandelt, B.D., Hivon, E. & Górski, K.M., 2001. Cosmic microwave background anisotropy power spectrum statistics for high precision cosmology, *Phys. Rev. D*, **64**, 083003.
- Yuan, D.N., Sjogren, W.L., Konopliv, A.S. & Kucinskis, A.B., 2001. Gravity field of Mars: A 75th degree and order model, *J. geophys. Res.*, **106**, 23 377–23 401.

APPENDIX A: DIAGONALIZATION MADE SIMPLE

Grünbaum *et al.* (1982) found a differential operator that commutes with the convolution integral of eq. (20), and a simple tridiagonal matrix, \mathbf{T} , that commutes with the kernel \mathbf{D} of eq. (13). For a concentration region θ_0 and a bandwidth L , the elements of \mathbf{T} are given by

$$T_{i,i} = \alpha_i \cos \theta_0 \quad \text{for } i = 1 \rightarrow L + 1 \quad (\text{A1})$$

$$T_{i-1,i} = T_{i,i-1} = \gamma_{i-1} \quad \text{for } i = 2 \rightarrow L + 1, \quad (\text{A2})$$

where

$$\alpha_i = i(i-1) \quad \text{and} \quad \gamma_i = \frac{i(L+1)^2 - i^3}{\sqrt{4i^2 - 1}}. \quad (\text{A3})$$

The sign of \mathbf{T} differs from that quoted in Grünbaum *et al.* (1982) to ensure that its eigenvalues have the same ordering as those of \mathbf{D} . Both \mathbf{T} and \mathbf{D} possess the same eigenfunctions, and although their eigenvalues differ, the concentration factors of the latter can be simply calculated from eq. (14). Alternatively, the eigenvalues can be computed by numerical integration of eq. (11), which avoids the construction of \mathbf{D} altogether.

In addition to computation speed, calculating the localization eigenfunctions from \mathbf{T} has the advantage that it has a monotonic and well-spaced spectrum of eigenvalues (Simons *et al.* 2005). The best concentrated eigenfunctions of \mathbf{D} , on the other hand, cannot be reliably calculated for $N_0 \geq 7$ as many of the eigenvalues would be indistinguishable from unity in double machine precision. While the worst concentrated tapers will not be needed for typical applications, they can be calculated only from \mathbf{T} as the smallest eigenvalues of \mathbf{D} would be all equal to zero to machine precision.

APPENDIX B: ANALYTICAL KERNEL ELEMENTS

In addition to by the procedure outlined in Appendix A, the solution of the space-concentration problem may be obtained by diagonalizing the kernel \mathbf{D} whose elements are (see eq. 13)

$$D_{ll'} = \frac{\sqrt{(2l+1)(2l'+1)}}{2} \int_{\cos\theta_0}^1 P_l(x)P_{l'}(x) dx. \quad (\text{B1})$$

As its integrand is a terminating polynomial of degree $l+l'$, it can be computed exactly using Gauss–Legendre quadrature (e.g. Press *et al.* 1992, pp. 140–155). Here, we develop exact expressions that are computationally faster and less affected by finite-precision round-off errors.

For the off-diagonal terms, we use an expression from Byerly (1893, p. 172) for the integral of two Legendre polynomials over the interval $[x, 1]$, namely

$$\int_x^1 P_l P_{l'} dx = \frac{(1-x^2)(P_{l'} P_l' - P_l P_{l'}')}{l(l+1) - l'(l'+1)}, \quad l \neq l', \quad (\text{B2})$$

where $P_l'(x)$ is the first derivative of the Legendre polynomial with respect to x as given by the identity

$$P_l'(x) = \frac{-lx P_l(x) + l P_{l-1}(x)}{(1-x^2)}. \quad (\text{B3})$$

Eq. (B2) is not valid for the diagonal terms, and we develop an expression for this case by expanding its integrand in a Legendre series. We first note that the integrand, $f = P_l^2$, is a polynomial of degree $2l$. It can thus be expressed as a finite sum of Legendre polynomials

$$f(x) = \sum_{j=0}^{2l} a_j^l P_j(x) \quad \text{with} \quad (\text{B4})$$

$$a_j^l = \frac{(2j+1)}{2} \int_{-1}^1 f(x) P_j(x) dx. \quad (\text{B5})$$

The coefficients a_j^l can be written in analytic form using the Wigner 3- j functions as

$$a_j^l = (2j+1) \begin{pmatrix} l & l & j \\ 0 & 0 & 0 \end{pmatrix}^2, \quad (\text{B6})$$

noting that the coefficients are zero for $j = \text{odd}$ (e.g. Dahlen & Tromp 1998, Appendix C). The diagonal terms of the kernel are

now given by

$$D_{ll}(\theta_0) = \frac{(2l+1)}{2} \sum_{j=0}^{2l} a_j^l \int_{\cos\theta_0}^1 P_j(x) dx, \quad (\text{B7})$$

where the integral can be evaluated analytically using a variant of eq. (B2),

$$\int_x^1 P_l(x') dx' = \frac{(1-x^2)P_l'(x)}{l(l+1)}, \quad l \neq 0. \quad (\text{B8})$$

Finally, using eq. (B3), the diagonal elements of $\mathbf{B1}$ have the following expression:

$$D_{ll}(\theta_0) = \frac{(2l+1)}{2} \sum_{j=0}^{2l} \frac{(2j+1)}{(j+1)} \begin{pmatrix} l & l & j \\ 0 & 0 & 0 \end{pmatrix}^2 \times [P_{j-1}(x_0) - x_0 P_j(x_0)], \quad (\text{B9})$$

where we define $P_{-1}(x)$ to be 1.

While analytical formulas exist for the Wigner 3- j symbols with arbitrary values as arguments, it is more convenient computationally to calculate these via recursion relationships. For our case, where j increases by 2, it can be shown that (e.g. Varshalovich *et al.* 1988, p. 248, 255)

$$\begin{pmatrix} l & l & 0 \\ 0 & 0 & 0 \end{pmatrix} = \frac{(-1)^l}{(2l+1)^{1/2}} \quad (\text{B10})$$

$$\begin{pmatrix} l & l & j \\ 0 & 0 & 0 \end{pmatrix} = \frac{(1-j)}{j} \begin{pmatrix} l & l & j-2 \\ 0 & 0 & 0 \end{pmatrix} \times \left(\frac{4l^2 + 4l - j^2 + 2j}{4l^2 + 4l - j^2 + 1} \right)^{1/2}. \quad (\text{B11})$$

APPENDIX C: LOCALIZATION BIAS

Consider two functions $f(\Omega)$ and $g(\Omega)$, each multiplied by an axisymmetric (zonal) data taper $h(\theta)$. We derive expressions for the expectation of the windowed cross-spectral power when the coefficients f_{lm} are governed by a stochastic process, and when the coefficients g_{lm} are linearly related to the latter by the isotropic admittance equation

$$g_{lm} = Z_l f_{lm}. \quad (\text{C1})$$

We start with the definition of the cross-spectral power of two windowed fields (see eqs 9 and 32):

$$\Phi_{lm} \Gamma_{lm} = \frac{1}{4\pi} \int_{\Omega} [h(\theta) f(\Omega)] Y_{lm}(\Omega) d\Omega \times \frac{1}{4\pi} \int_{\Omega} [h(\theta) g(\Omega)] Y_{lm}(\Omega) d\Omega. \quad (\text{C2})$$

After expanding the functions in spherical harmonics by way of eq. (5), and using the shorthand notation

$$\sum_{ij}^{\infty} = \sum_{i=0}^{\infty} \sum_{j=-i}^i \quad \text{and} \quad \sum_l^L = \sum_{l=0}^L \quad (\text{C3})$$

this can be written as

$$\Phi_{lm} \Gamma_{lm} = \frac{1}{(4\pi)^2} \sum_{l_1}^{L_h} h_{l_1} \sum_{l_2}^{L_h} h_{l_2} \sum_{ij}^{\infty} f_{ij} \sum_{i'j'}^{\infty} g_{i'j'} \times \int_{\Omega} Y_{ij} Y_{l_1 0} Y_{lm} d\Omega \int_{\Omega} Y_{i'j'} Y_{l_2 0} Y_{lm} d\Omega. \quad (\text{C4})$$

We make the assumption that the coefficients f_{ij} are random variables with zero mean, and that the coefficients g_{ij} are proportional to these via eq. (C1). The expectation of eq. (C4) is equivalent to the average over all possible realizations of f_{lm} and is given by

$$\begin{aligned} \langle \Phi_{lm} \Gamma_{lm} \rangle &= \frac{1}{(4\pi)^2} \sum_{l_1}^{L_h} h_{l_1} \sum_{l_2}^{L_h} h_{l_2} \sum_{ij}^{\infty} \sum_{i'j'}^{\infty} \langle f_{ij} g_{i'j'} \rangle \\ &\quad \times \int_{\Omega} Y_{ij} Y_{l_1 0} Y_{lm} d\Omega \int_{\Omega'} Y_{i'j'} Y_{l_2 0} Y_{lm} d\Omega', \end{aligned} \quad (C5)$$

where $\langle \dots \rangle$ is the expectation operator. Using the identity

$$\text{cov}\{X_1, X_2\} = \langle X_1 X_2 \rangle - \langle X_1 \rangle \langle X_2 \rangle, \quad (C6)$$

where $\text{cov}\{\dots\}$ is the covariance operator, we remark that

$$\langle f_{ij} g_{i'j'} \rangle = \text{cov}\{f_{ij}, g_{i'j'}\} \delta_{ii'} \delta_{jj'}, \quad (C7)$$

since the random variables are by definition uncorrelated and have zero mean. If we further assume that the covariance of f_{ij} and g_{ij} is only dependent upon the spherical harmonic degree, and not on the order, then

$$\langle f_{ij} g_{i'j'} \rangle = \frac{S_{fg}(i)}{(2i+1)} \delta_{ii'} \delta_{jj'} \quad (C8)$$

where $S_{fg}(i)$ is the cross-spectral power of the two fields for degree i (cf. eq. 7). The expectation of the cross-spectral power of the windowed field can thus be written as

$$\begin{aligned} \langle \Phi_{lm} \Gamma_{lm} \rangle &= \frac{1}{(4\pi)^2} \sum_{l_1}^{L_h} h_{l_1} \sum_{l_2}^{L_h} h_{l_2} \sum_{ij}^{\infty} \frac{S_{fg}(i)}{(2i+1)} \\ &\quad \times \int_{\Omega} Y_{ij} Y_{l_1 0} Y_{lm} d\Omega \int_{\Omega'} Y_{ij} Y_{l_2 0} Y_{lm} d\Omega'. \end{aligned} \quad (C9)$$

The next step is to express the integral of three real spherical harmonics as products of Wigner 3- j or Clebsch-Gordan symbols. When one or more of the angular orders is equal to zero, we have the equality

$$\begin{aligned} &\int_{\Omega} Y_{l_1 m_1} Y_{l_2 0} Y_{lm} d\Omega \\ &= (-1)^m 4\pi \sqrt{(2l_1+1)(2l_2+1)(2l+1)} \\ &\quad \times \begin{pmatrix} l_1 & l_2 & l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l \\ m_1 & 0 & -m \end{pmatrix} \\ &= 4\pi \left[\frac{(2l_1+1)(2l_2+1)}{2l+1} \right]^{1/2} C_{l_1 m_1 l_2 0}^{l m}, \end{aligned} \quad (C10)$$

which, in order to be non-zero, must satisfy the following selection rules (e.g. Dahlen & Tromp 1998, Appendix C):

$$\begin{aligned} m &= m_1 \\ |m| &\leq l_1, l \\ |l_1 - l_2| &\leq l \leq l_1 + l_2 \\ |l - l_1| &\leq l_2 \leq l + l_1 \\ |l_2 - l| &\leq l_1 \leq l_2 + l \\ l_1 + l_2 + l &= \text{even}. \end{aligned} \quad (C11)$$

Wigner 3- j symbols are invariant under cyclic column permutation and are related to the Clebsch-Gordan coefficients by

$$C_{l_1 m_1 l_2 m_2}^{l m} = \frac{(-1)^{l_1 - l_2 + m}}{(2l+1)^{-1/2}} \begin{pmatrix} l_1 & l_2 & l \\ m_1 & m_2 & -m \end{pmatrix}. \quad (C12)$$

These can be calculated using algorithms discussed by Schulten & Gordon (1975) and Luscombe & Luban (1998). Eq. (C10) is

most easily verified by noting that the integral of real harmonics is equivalent to the integral of complex harmonics $Y_{l_1 m_1} Y_{l_2 0} Y_{lm}^*$, for which the expansion is well known (e.g. Varshalovich *et al.* 1988, p. 148). This, however, is not generally true for non-zero orders (Homeier & Steinborn 1996).

The expectation of the cross-spectral power for a given spherical harmonic is obtained by inserting eq. (C10) into (C9)

$$\begin{aligned} \langle S_{\Phi\Gamma}(l, m) \rangle &= \langle \Phi_{lm} \Gamma_{lm} \rangle \\ &= \sum_{l_1}^{L_h} h_{l_1} \sum_{l_2}^{L_h} h_{l_2} \sum_i^{\infty} S_{fg}(i) \frac{\sqrt{(2l_1+1)(2l_2+1)}}{2l+1} \\ &\quad \times C_{i 0 l_1 0}^{l 0} C_{i 0 l_2 0}^{l 0} C_{i m l_1 0}^{l m} C_{i m l_2 0}^{l m}, \end{aligned} \quad (C13)$$

where we have used the selection rule that j must equal m . This equation is unchanged by replacing m with $-m$.

We estimate the total cross-power for a given degree from these expectation values by summing over all orders m

$$\begin{aligned} \langle S_{\Phi\Gamma}(l) \rangle &= \sum_{m=-l}^l \langle S_{\Phi\Gamma}(l, m) \rangle \\ &= \sum_{l_1}^{L_h} h_{l_1} \sum_{l_2}^{L_h} h_{l_2} \sum_i^{\infty} S_{fg}(i) \frac{\sqrt{(2l_1+1)(2l_2+1)}}{2l+1} \\ &\quad \times C_{i 0 l_1 0}^{l 0} C_{i 0 l_2 0}^{l 0} \sum_{m=-l}^l C_{i m l_1 0}^{l m} C_{i m l_2 0}^{l m}. \end{aligned} \quad (C14)$$

This expression can be considerably simplified by use of the identity (e.g. Varshalovich *et al.* 1988, p. 259)

$$\sum_{\alpha=-c}^c \sum_{\gamma=-c}^c C_{\alpha\alpha b\beta}^{c\gamma} C_{\alpha\alpha b'\beta'}^{c\gamma} = \frac{(2c+1)}{(2b+1)} \delta_{bb'} \delta_{\beta\beta'}, \quad (C15)$$

which reduces the sum over m to

$$\sum_{m=-l}^l C_{i m l_1 0}^{l m} C_{i m l_2 0}^{l m} = \sum_{m=-l}^l \sum_{j=-l}^l C_{ij l_1 0}^{l m} C_{ij l_2 0}^{l m} = \frac{(2l+1)}{(2l_1+1)} \delta_{l_1 l_2}. \quad (C16)$$

Finally, after taking into account the selection rules, and relabeling the index l_1 to j , the expectation of the total windowed power for a given degree can be written as

$$\langle S_{\Phi\Gamma}(l) \rangle = \sum_j^{L_h} h_j^2 \sum_{i=|l-j|}^{l+j} S_{fg}(i) (C_{i 0 j 0}^{l 0})^2, \quad (C17)$$

which appeared as eq. (35) in the main text. When $L_h = 0$, this reduces to the input power spectrum multiplied by h_0^2 , as $(C_{i 0 0 0}^{l 0})^2 = 1$ for all l . Thus, if the windows are normalized to unit power, windowing with a constant everywhere recovers the input spectrum without gain. The expectation operators used in this section average over all possible realizations of the random variables f_{ij} . Any single realization would of course differ from its theoretical expectation.

Eqs (C13) and (17) are valid for any axisymmetric windowing function $h(\theta)$, and we next consider windowed spectrum estimates using our family of orthogonal bandlimited tapers. In particular, we will define a localized multitaper spectrum estimate as a weighted average of all windowed cross-power spectra. When the weights correspond to the eigenvalues of the data tapers, we have

$$S_{\Phi\Gamma}^{(mt)}(l, m) = \frac{1}{\text{tr } \mathbf{D}} \sum_{k=1}^{L+1} \lambda_k S_{\Phi\Gamma}^{(k)}(l, m), \quad (C18)$$

$$S_{\Phi\Gamma}^{(mt)}(l) = \frac{1}{\text{tr } \mathbf{D}} \sum_{k=1}^{L+1} \lambda_k S_{\Phi\Gamma}^{(k)}(l), \quad (C19)$$

where the sum of the eigenvalues has been denoted by the trace of the kernel \mathbf{D} . Since the kernel \mathbf{D} can always be decomposed into a product of three matrices

$$\mathbf{D} = \mathbf{H} \mathbf{\Lambda} \mathbf{H}^T, \quad (\text{C20})$$

where \mathbf{H} contains its eigenvectors as column vectors and $\mathbf{\Lambda}$ is a diagonal matrix containing the corresponding eigenvalues (e.g. Strang 1986), the elements of the kernel can be written as

$$D_{ij} = \sum_{k=1}^{L+1} \lambda_k h_i^{(k)} h_j^{(k)}. \quad (\text{C21})$$

With this expression, the multitaper cross-spectral estimate for a single harmonic, eq. (18), weighted over all orthogonal tapers is equal to

$$\begin{aligned} \langle S_{\Phi\Gamma}^{(mt)}(l, m) \rangle &= \frac{1}{\text{tr} \mathbf{D}} \sum_{l_1}^L \sum_{l_2}^L D_{l_1 l_2} \sum_{i=i_{\min}}^{i_{\max}} S_{f_g}(i) \\ &\times \frac{\sqrt{(2l_1+1)(2l_2+1)}}{2l+1} C_{i0l_1 0}^{l_0} C_{i0l_2 0}^{l_0} C_{iml_1 0}^{lm} C_{iml_2 0}^{lm}, \end{aligned} \quad (\text{C22})$$

where $i_{\min} = \max(|m|, |l-l_1|, |l-l_2|)$ and $i_{\max} = \min(l+l_1, l+l_2)$. Similarly, the multitaper estimate of the total power for a single degree, eq. (C19), can be written as

$$\langle S_{\Phi\Gamma}^{(mt)}(l) \rangle = \frac{1}{\text{tr} \mathbf{D}} \sum_j^L D_{jj} \sum_{i=|l-j|}^{l+j} S_{f_g}(i) (C_{i0j0}^{l_0})^2, \quad (\text{C23})$$

introduced as eq. (38) in the main text.

APPENDIX D: LOCALIZATION VARIANCE

We seek to determine the variance of the multitaper power spectral estimate. We will use the following two identities:

$$\text{var} \left\{ \sum_{i=1}^N a_i X_i \right\} = \sum_{j=1}^N \sum_{k=1}^N a_j a_k \text{cov}\{X_j, X_k\} \quad (\text{D1})$$

$$\text{cov} \left\{ \sum_{i=1}^N X_i, \sum_{j=1}^M X_j \right\} = \sum_{i=1}^N \sum_{j=1}^M \text{cov}\{X_i, X_j\}, \quad (\text{D2})$$

which follow from the definition of covariance of multiple variables X_i , and Isserlis' theorem (Walden *et al.* 1994)

$$\begin{aligned} \text{cov}\{Z_1 Z_2, Z_3 Z_4\} &= \text{cov}\{Z_1, Z_3\} \text{cov}\{Z_2, Z_4\} \\ &+ \text{cov}\{Z_1, Z_4\} \text{cov}\{Z_2, Z_3\}, \end{aligned} \quad (\text{D3})$$

which is valid for random variables Z_i that are Gaussian distributed with zero mean. Using eq. (D1), the variance of the multitaper spectral estimate

$$S_{\Phi\Gamma}^{(mt)}(l) = \sum_{k=1}^K a_k S_{\Phi\Gamma}^{(k)}(l), \quad (\text{D4})$$

where a_k is a generic weight applied to the k th direct spectral estimate $S^{(k)}$, is given by

$$\text{var} \left\{ S_{\Phi\Gamma}^{(mt)}(l) \right\} = \sum_{j=1}^K \sum_{k=1}^K a_j a_k \text{cov} \left\{ S_{\Phi\Gamma}^{(j)}(l), S_{\Phi\Gamma}^{(k)}(l) \right\}. \quad (\text{D5})$$

Using eq. (D2) and the definition of the cross-power spectrum (eq. 33) the covariance of the cross-power spectral estimates can be expressed as

$$\text{cov} \left\{ S_{\Phi\Gamma}^{(j)}(l), S_{\Phi\Gamma}^{(k)}(l) \right\} = \sum_{m=-l}^l \sum_{m'=-l}^l \text{cov} \left\{ \Phi_{lm}^{(j)} \Gamma_{lm}^{(j)}, \Phi_{lm'}^{(k)} \Gamma_{lm'}^{(k)} \right\}. \quad (\text{D6})$$

At this point, Isserlis' theorem (eq. D3) can be used to simplify the above expression if the coefficients Φ_{lm} and Γ_{lm} were known to be Gaussian distributed with zero mean. It is evident from eq. (10) that since the coefficients of the global fields f_{lm} and g_{lm} have been assumed to possess a zero mean, that the windowed coefficients will as well. If we furthermore assume that each f_{lm} and g_{lm} is Gaussian distributed, then each of the $\sim L_h^2$ terms in eq. (10) will likewise be Gaussian distributed. If the input fields additionally have a white power spectrum, then it is likely as a result of the Lindeberg-Feller central limit theorem (Feller 1971, pp. 256–258) that the windowed coefficients will asymptotically approach a normal distribution with increasing L_h . As this theorem requires that the variance of each term in eq. (10) is small compared to the sum of all the terms' variances, it is probable that this asymptotic behaviour will not be attained for small degrees when the global power spectrum is red. With these caveats in mind, Isserlis' theorem can be used to express eq. (D6) as

$$\begin{aligned} \text{cov} \left\{ S_{\Phi\Gamma}^{(j)}(l), S_{\Phi\Gamma}^{(k)}(l) \right\} &= \sum_{m=-l}^l \sum_{m'=-l}^l \left(\text{cov} \left\{ \Phi_{lm}^{(j)}, \Phi_{lm'}^{(k)} \right\} \text{cov} \left\{ \Gamma_{lm}^{(j)}, \Gamma_{lm'}^{(k)} \right\} \right. \\ &\left. + \text{cov} \left\{ \Phi_{lm}^{(j)}, \Gamma_{lm'}^{(k)} \right\} \text{cov} \left\{ \Gamma_{lm}^{(j)}, \Phi_{lm'}^{(k)} \right\} \right). \end{aligned} \quad (\text{D7})$$

For zero-mean localized spectral coefficients, it is evident from eq. (C6) and a generalized form of eq. (C13) that

$$\begin{aligned} \text{cov} \left\{ \Phi_{lm}^{(j)}, \Gamma_{lm'}^{(k)} \right\} &= \langle \Phi_{lm}^{(j)} \Gamma_{lm'}^{(k)} \rangle = \langle S_{\Phi\Gamma}^{(j,k)}(l, m) \rangle \delta_{mm'} \\ &= \sum_{l_1}^{L_h} h_{l_1}^{(j)} \sum_{l_2}^{L_h} h_{l_2}^{(k)} \sum_{i=i_{\min}}^{i_{\max}} S_{f_g}(i) \frac{\sqrt{(2l_1+1)(2l_2+1)}}{2l+1} \\ &\times C_{i0l_1 0}^{l_0} C_{i0l_2 0}^{l_0} C_{iml_1 0}^{lm} C_{iml_2 0}^{lm} \delta_{mm'}, \end{aligned} \quad (\text{D8})$$

which is symmetric in j and k , with $i_{\min} = \max(|m|, |l-l_1|, |l-l_2|)$ and $i_{\max} = \min(l+l_1, l+l_2)$. Together, eqs (D5)–(D8) yield

$$\begin{aligned} \text{var} \left\{ S_{\Phi\Gamma}^{(mt)}(l) \right\} &= \sum_{j=1}^K \sum_{k=1}^K a_j a_k \sum_{m=-l}^l \left(\left\langle S_{\Phi\Phi}^{(j,k)}(l, m) \right\rangle \left\langle S_{\Gamma\Gamma}^{(j,k)}(l, m) \right\rangle \right. \\ &\left. + \left\langle S_{\Phi\Gamma}^{(j,k)}(l, m) \right\rangle \left\langle S_{\Gamma\Phi}^{(j,k)}(l, m) \right\rangle \right), \end{aligned} \quad (\text{D9})$$

and, for single-field multitaper spectral estimates,

$$\text{var} \left\{ S_{\Phi\Phi}^{(mt)}(l) \right\} = 2 \sum_{j=1}^K \sum_{k=1}^K a_j a_k \sum_{m=-l}^l \left\langle S_{\Phi\Phi}^{(j,k)}(l, m) \right\rangle^2. \quad (\text{D10})$$