

# Minimum-Variance Multitaper Spectral Estimation on the Sphere

Mark A. Wieczorek and Frederik J. Simons

Communicated by Matthias Holschneider

**ABSTRACT.** We develop a method to estimate the power spectrum of a stochastic process on the sphere from data of limited geographical coverage. Our approach can be interpreted either as estimating the global power spectrum of a stationary process when only a portion of the data are available for analysis, or estimating the power spectrum from local data under the assumption that the data are locally stationary in a specified region. Restricting a global function to a spatial subdomain — whether by necessity or by design — is a windowing operation, and an equation like a convolution in the spectral domain relates the expected value of the windowed power spectrum to the underlying global power spectrum and the known power spectrum of the localization window. The best windows for the purpose of localized spectral analysis have their energy concentrated in the region of interest while possessing the smallest effective bandwidth as possible. Solving an optimization problem in the sense of Slepian (1960) yields a family of orthogonal windows of diminishing spatio-spectral localization, the best concentrated of which we propose to use to form a weighted multitaper spectrum estimate in the sense of Thomson (1982). Such an estimate is both more representative of the target region and reduces the estimation variance when compared to estimates formed by any single bandlimited window. We describe how the weights applied to the individual spectral estimates in forming the multitaper estimate can be chosen such that the variance of the estimate is minimized.

## 1. Introduction

Spectral analysis is an indispensable tool in many branches of the physical and mathematical sciences, with common applications ranging from one-dimensional time series to two-dimensional image analysis. For many purposes it is sufficient to employ a Cartesian geometry, and for this case a plethora of sophisticated techniques have been developed, such as parametric, maximum-entropy, and multitaper spectral analysis (see [18] for a

*Math Subject Classifications.* 33C55, 34L05, 42B35, 42C10, 62M15.

*Keywords and Phrases.* Spherical harmonics, multitaper spectral analysis.

comprehensive review). However, for certain problems, especially in geophysics, it is necessary to obtain spectral estimates from data that are localized to specific regions on the surface of a sphere. While subsets of these data could be mapped to a two-dimensional plane, enabling the use of Cartesian methods, this procedure is bound to introduce some error into the obtained spectral estimates. As the size of the region approaches a significant fraction of the surface area of the sphere, these would naturally become increasingly unreliable.

Spectral analysis on the sphere is an important tool in several scientific disciplines, and two examples suffice to illustrate the range of problems that are often encountered. First, in geophysics and geodesy it is common to represent the gravity field and topography of the terrestrial planets as spherical harmonic expansions and to use their cross-spectral properties to investigate the interior structure of the body [29]. However, since the relationship between the gravity and topography coefficients depends upon the geologic history of the geographic area of interest, it is often necessary to consider only a localized subset of these data. A second example is in the field of cosmology where the power spectrum of the cosmic microwave background radiation is used to place important constraints on the structure and constitution of the universe [25]. In contrast to many geophysical problems, the measured temperature fluctuations are often assumed to be derived from a globally stationary process. Nevertheless, when estimating the power spectrum from satellite and terrestrial-based measurements, it is necessary to mask out regions that are contaminated by emissions emanating from the plane of our own galaxy [9].

On the sphere one is thus concerned generally with estimating the power spectrum of a certain process from data confined to a restricted region. As illustrated by the above examples, this can be interpreted in one of two ways. In one case, the function is assumed to be stationary, and an estimate of the global power spectrum is desired based on a localized subset of data. In the second case, the data are known to be nonstationary, and an estimate of a “localized” power spectrum is sought by assuming local stationarity of the data within the specified region.

The restriction of data to a specified region is equivalent to multiplying a globally defined function by a localization window or “data taper,” and the objective is to relate the power spectrum of this localized field to the global one. This has been investigated by using single binary masks [10], as well as families of orthogonal isotropic windows with the purpose of obtaining a “multitaper” estimate [31]. The use of multiple localization windows, as originally pioneered in the Cartesian domain by Thomson [26], possesses many advantages over that of a single window, in particular, smaller variances of the resulting spectral estimates and more uniform coverage of the localization region. Here, we extend our previous approach [31] of using zonal tapers (i.e., those with azimuthal symmetry about a polar axis) to the general case that includes nonzonal tapers [21]. The inclusion of nonzonal data tapers greatly increases the number of individual spectral estimates that make up the multitaper estimate, and this leads to a significant reduction in the variance of this estimate. We demonstrate how the weights of the individual estimates can be chosen to minimize the multitaper estimation variance.

In this article, we first describe the theory of estimating the global power spectrum of a stochastic stationary process through the use of localization windows. This includes quantifying the relationship between the global and localized power spectra, the design of windows that are optimal for this purpose, and the formation of a minimum-variance multitaper spectral estimate. The majority of the theoretical development relating to these problems is contained in four appendices. Following this, we describe the statistical properties of the multitaper estimates (such as their bias and variance) heuristically when the data are governed by a stochastic process with either a “white” or “red” power spectrum.

Next, we give an example of estimating the power spectrum from a single realization of a stochastic process. Finally, we conclude by emphasizing avenues of future research.

## 2. Theory

### 2.1 Localized Spectral Estimation

Any real square-integrable function on the unit sphere can be expressed by a linear combination of orthogonal functions as

$$f(\Omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm} Y_{lm}(\Omega), \tag{2.1}$$

where  $Y_{lm}$  is a real spherical harmonic of degree  $l$  and order  $m$ ,  $f_{lm}$  is the corresponding expansion coefficient, and  $\Omega = (\theta, \phi)$  represents position on the sphere in terms of colatitude,  $\theta$ , and longitude,  $\phi$ . The real spherical harmonics are defined in terms of a product of a Legendre function in colatitude and either a sine or cosine function in longitude,

$$Y_{lm}(\Omega) = \begin{cases} \bar{P}_{lm}(\cos \theta) \cos m\phi & \text{if } m \geq 0 \\ \bar{P}_{l|m|}(\cos \theta) \sin |m|\phi & \text{if } m < 0, \end{cases} \tag{2.2}$$

and the normalized Legendre functions used in this investigation are given by

$$\bar{P}_{lm}(\mu) = \sqrt{(2 - \delta_{0m})(2l + 1)} \sqrt{\frac{(l - m)!}{(l + m)!}} P_{lm}(\mu), \tag{2.3}$$

where  $\delta_{ij}$  is the Kronecker delta function and  $P_{lm}$  is the standard associated Legendre function,

$$P_{lm}(\mu) = \frac{1}{2^l l!} (1 - \mu^2)^{m/2} \left(\frac{d}{d\mu}\right)^{l+m} (\mu^2 - 1)^l. \tag{2.4}$$

With the above definitions, the spherical harmonics of (2.2) are orthogonal over the sphere and possess unit power,

$$\frac{1}{4\pi} \int_{\Omega} Y_{lm}(\Omega) Y_{l'm'}(\Omega) d\Omega = \delta_{ll'} \delta_{mm'}, \tag{2.5}$$

where  $d\Omega = \sin \theta d\theta d\phi$ . We note that this unit-power normalization is consistent with that used by the geodesy community, but differs from the physics and seismology communities that use orthonormal harmonics [3, 27] and the magnetics community that uses Schmidt semi-normalized harmonics [2]. Furthermore, we omit the Condon-Shortley phase factor of  $(-1)^m$  in the definition of the spherical harmonics, which is consistent with the usage of the geodesy and magnetics communities. Using (2.1) and (2.5), the total power of a real function  $f$  can be shown to be related to its spectral coefficients by a generalization of Parseval's theorem:

$$\frac{1}{4\pi} \int_{\Omega} [f(\Omega)]^2 d\Omega = \sum_{l=0}^{\infty} S_{ff}(l), \tag{2.6}$$

where

$$S_{ff}(l) = \sum_{m=-l}^l f_{lm}^2 \tag{2.7}$$

is referred to as the power spectrum of  $f$ . Similarly, the cross-power of two real functions  $f$  and  $g$  is given by

$$\frac{1}{4\pi} \int_{\Omega} f(\Omega) g(\Omega) d\Omega = \sum_{l=0}^{\infty} S_{fg}(l), \quad (2.8)$$

where the cross-power spectrum is

$$S_{fg}(l) = \sum_{m=-l}^l f_{lm} g_{lm}. \quad (2.9)$$

The power and cross-power spectra possess the property that they are unmodified by a rotation of the coordinate system [12]. Efficient and accurate algorithms for calculating the normalized Legendre functions and spherical harmonic coefficients of a function can be found in [5] and [11], respectively. The relationship between real and complex spherical harmonics, which is necessary for certain derivations in the appendices, is presented in Appendix A.

If the function  $f$  of (2.1) were known globally, it would be a trivial matter to obtain its spherical harmonic coefficients  $f_{lm}$ , and by consequence  $S_{ff}$ , its power spectrum [4]. Unfortunately, in spherical analyses it is common that the function is only known within a restricted domain on the sphere, or conversely, that it is necessary to ignore certain contaminated or unrepresentative regions of the global data set. Alternatively, it may arise that the global function is known to be nonstationary, and that a local power spectrum estimate is desired under the assumption of local stationarity for a specified region.

The localization of a global function to a given domain can be formulated as a windowing operation,

$$\Phi(\Omega) = h(\Omega) f(\Omega), \quad (2.10)$$

where  $h$  is the localization window and  $\Phi$  is the localized version of  $f$ . While a naive binary mask is sometimes used for  $h$ , indicating either the presence or absence of data, this choice will later be shown to possess undesirable properties analogous to those of the standard periodogram [18]. After having chosen  $h$ , the power spectrum of the localized function,  $S_{\Phi\Phi}$ , is easily calculated from (2.6) and (2.7). It should be clear that while  $S_{\Phi\Phi}$  might resemble  $S_{ff}$ , the two will not be equal as a result of the windowing operation. The quantity  $S_{\Phi\Phi}$  will here be referred to as both the *localized* and *windowed* power spectrum.

The relationship between  $S_{ff}$  and  $S_{\Phi\Phi}$  is a complicated one, and it is generally not possible to invert for the former given the latter. Nevertheless, if it is assumed that the function  $f$  is a stationary stochastic process, a simple relationship exists between the *expectation* of  $S_{\Phi\Phi}$  and  $S_{ff}$ , namely,

$$\langle S_{\Phi\Phi}(l) \rangle = \sum_{j=0}^L S_{hh}(j) \sum_{i=|l-j|}^{l+j} S_{ff}(i) \left( C_{j0i0}^{l0} \right)^2, \quad (2.11)$$

where  $\langle \dots \rangle$  denotes the expectation operator,  $L$  is the spherical harmonic bandwidth of  $h$ , and the symbol in parentheses is a Clebsch-Gordan coefficient. Various forms of this relationship have been previously derived independently [10, 17, 31], and in Appendix B we generalize this to the case of localized cross-power spectra. Here, it is sufficient to note that for (2.11) to hold the spherical harmonic coefficients of  $f_{lm}$  are required to be zero-mean random variables with a variance that depends only on degree  $l$ . The expectation of

the localized power spectrum is to be considered as an average over all possible realizations of the random variables  $f_{lm}$ .

Equation (2.11) shows that the expectation of the localized power spectrum is related to the power spectrum of  $f$  and the localization window  $h$  by an operation reminiscent of a convolution. In particular, it is important to note that each degree  $l$  of the localized power spectrum contains contributions from the global spectrum  $S_{ff}$  within the degree range  $l \pm L$ . Thus, the spherical harmonic bandwidth of the localization window directly controls how the global power spectrum  $S_{ff}$  is “smoothed” in determining the expectation of the localized spectrum  $S_{\Phi\Phi}$ . Given the power spectrum of a localization window, as well as an estimate of the localized power spectrum expectation, as discussed in Section 4 and [31], several techniques could be used to invert for the global spectrum. This would make the resulting “deconvolved” localized spectral estimate statistically unbiased, but would come at the cost of a higher estimation variance [4].

### 2.2 Window Design

The convolution-type operation of (2.11) that relates the localized power spectrum to the global and localization window power spectra demonstrates the importance of using a window  $h$  with as small a spherical harmonic bandwidth as possible. For instance, if a naive binary mask were used to isolate certain domains on the sphere, the bandwidth of this window would be infinite, and every localized spectral estimate  $S_{\Phi\Phi}$  would be influenced by every degree of the global spectrum  $S_{ff}$ . If the global power spectrum possessed a significant dynamic range, degrees with high power could positively bias the localized power spectrum at degrees where the global power spectrum is small. Such spectral leakage could hinder attempts to invert for the global spectrum  $S_{ff}$  given knowledge of  $S_{hh}$  and  $S_{\Phi\Phi}$ .

In order to spatially localize a function on the sphere, it is clear that the localization window should possess zero or near-zero amplitudes exterior to a specified region of interest  $R$ . Additionally, in order to limit the effects of spectral leakage, the effective spectral bandwidth of the window should be as small as possible. The problem of designing windows that are ideally localized in both the space and spectral domains was originally posed as an optimization problem by Slepian and coworkers in the Cartesian domain (see [24] for a review), and later extended to the sphere by various authors [6, 20, 21, 31].

One form of this optimization problem is to find those functions whose spatial power is concentrated within a given domain  $R$ , but yet are bandlimited to a spherical harmonic degree  $L$ ; i.e., to find those functions that maximize the ratio

$$\lambda = \int_R h^2(\Omega) d\Omega / \int_{\Omega} h^2(\Omega) d\Omega . \tag{2.12}$$

It can be shown [21, 31] that this equation reduces to an eigenvalue equation

$$\mathbf{D} \mathbf{h} = \lambda \mathbf{h} , \tag{2.13}$$

where  $\mathbf{h}$  is a vector of length  $(L + 1)^2$  containing the spherical harmonic coefficients of the window, and  $\mathbf{D}$  is a  $(L + 1)^2 \times (L + 1)^2$  localization kernel. The solution of this eigenvalue problem yields a family of orthogonal windows  $h^{(k)}$  (which we normalize to have unit power), with corresponding spatial concentration factors ordered such that

$$1 > \lambda_1 \geq \dots \lambda_k \dots \geq \lambda_{(L+1)^2} > 0 . \tag{2.14}$$

As the eigenvalue spectrum has been found empirically to transition quickly from values near unity to zero, the sum of the eigenvalues corresponds approximately to the number of

windows with good spatio-spectral localization properties, and this ‘‘Shannon number’’ has been shown [21] to be equal to

$$N = \sum_{k=1}^{(L+1)^2} \lambda_k = (L+1)^2 \frac{A}{4\pi}, \quad (2.15)$$

where  $A$  is the area spanned by the region  $R$  on the unit sphere.

An alternative criterion for designing a localization window is to instead find those functions that are perfectly contained within a region  $R$  and whose spectral power is concentrated within a spherical harmonic bandwidth  $L$ ; i.e., to maximize

$$\lambda = \frac{\sum_{l=0}^L \sum_{m=-l}^l h_{lm}^2}{\sum_{l=0}^{\infty} \sum_{m=-l}^l h_{lm}^2}, \quad (2.16)$$

where the function  $h$  is defined to be zero exterior to  $R$ . This spectral optimization problem is complementary to the spatial concentration problem [21, 31]. In particular, the functions are identical within the domain  $R$ , the spectral- and spatial-concentration eigenvalue spectra are identical up to the  $(L+1)^2$ -th eigenvalue, and the power spectra of the functions for  $l \leq L$  differ only by a factor equal to the square of the eigenvalue. Thus, if one desires localization windows that are perfectly restricted to  $R$ , yet nonetheless possess some spectral power beyond the effective bandwidth  $L$ , it is only necessary to compute the space-concentration windows by means of (2.13) and to then set these functions equal to zero exterior to  $R$ .

The geometry of the localization domain  $R$  determines the form of the localization kernel  $\mathbf{D}$ . While this matrix and its eigenvalues and eigenfunctions can be computed for any arbitrary domain [21], certain geometries greatly simplify this task. For bandlimited windows, if the domain  $R$  is a spherical cap located at the North pole ( $\theta = 0^\circ$ ),  $\mathbf{D}$  is block diagonal, allowing the eigenvalue problem of (2.13) to be solved separately for each individual angular order  $m$ . More importantly, there exists a tridiagonal matrix with analytically prescribed elements that commutes with  $\mathbf{D}$ , and hence shares the same eigenfunctions [6, 21]. Similarly, for the case of two antipodal spherical caps (or equivalently, an equatorial belt), a simple commuting tridiagonal matrix exists as well [20].

The theoretical development presented in this article is valid for any irregularly shaped concentration domain, and for either space-limited or bandlimited localization windows. In this investigation, however, we will employ exclusively bandlimited localization windows whose spatial power is optimally concentrated within a spherical cap located at  $\theta = 0^\circ$ . A spherical-cap concentration domain is a close analog to the square domain that is commonly used in 2-D Cartesian analyses and is likely to find broad applicability to many problems [1, 7, 16, 30].

For demonstration purposes, we use a concentration domain with an angular radius of  $\theta_0 = 30^\circ$  and a spectral bandwidth of  $L = 29$ , corresponding to  $N \simeq 60$ . We will further restrict ourselves to those windows that are nearly perfectly localized with  $\lambda > 0.99$ , of which there are 34. As illustrated in Figure 1, each of these windows possesses nonzero spherical harmonic coefficients for only a single angular order  $m$ . Whereas only windows with  $m \geq 0$  are shown in this figure, we note that each nonzonal window (i.e.,  $m \neq 0$ ) has an identically concentrated twin of angular order  $-m$  that differs only by an azimuthal rotation of  $90^\circ/m$ . In contrast, if we were to employ only those windows that are isotropic (i.e.,  $m = 0$ ), we would have only 4 nearly perfectly concentrated windows at our disposal. In that case, as shown by [31], the zonal Shannon number would be given by  $N_0 = (L+1)\theta_0/\pi = 5$ .

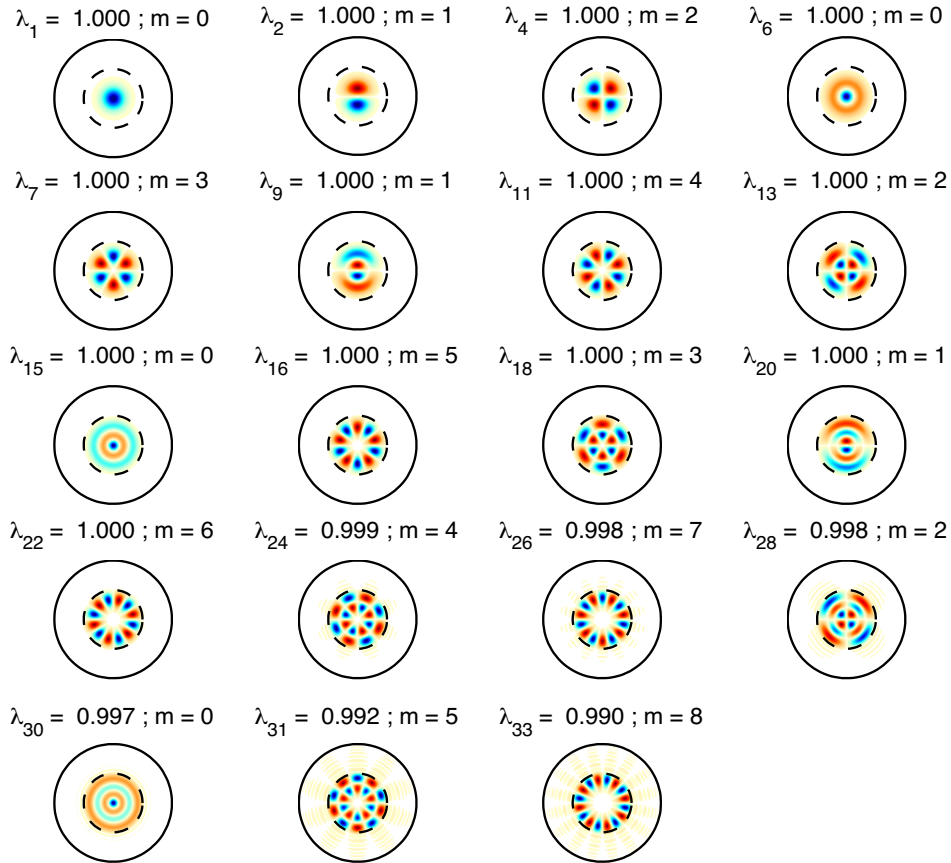


FIGURE 1 Spatial rendition of those functions with angular order  $m \geq 0$  that are nearly perfectly concentrated within a spherical cap of angular radius  $\theta_0 = 30^\circ$  and with a spherical harmonic bandwidth of  $L = 29$ . For this case,  $N \simeq 60$ , and the total number of windows with  $\lambda > 0.99$  is 34. Each nonzonal function possesses a twin with angular order  $-m$  that is rotated azimuthally by  $90^\circ/m$  (not shown).

### 2.3 Minimum-Variance Multitaper Spectral Estimation

Several studies have used single localization windows to obtain localized spectral estimates on the sphere [1, 10, 16, 22]. Nevertheless, as originally formulated by Thomson [26], the use of multiple orthogonal localization windows can have significant advantages over the use of any single window [18, 28, 31]. In particular, the energy of a single bandlimited window will always nonuniformly cover the desired concentration region, and this will result in some data being statistically over- or under-represented when forming the spectral estimate. In contrast, the cumulative energy of the orthogonal windows solving (2.13) more uniformly covers the concentration region. Second, since the spectral estimates that result from using orthogonal windows are somewhat uncorrelated, a multitaper average of these will possess a smaller estimation variance. This is especially important since most investigations are limited to analyzing a single realization of a stochastic process.

To demonstrate the first of the above advantages of using multiple orthogonal localization windows, we plot the cumulative energy (i.e., the squared amplitude) of the windows utilized in this study as a function of colatitude in Figure 2. Since windows of angular order  $m$  and  $-m$  are included, this function is independent of azimuth  $\phi$ . In particular, the

cumulative energy of the four zonal windows is shown, as well as that for the 34 zonal and nonzonal windows of Figure 1. As is readily seen, the energy of the four zonal windows is peaked near the center of the spherical cap, giving lesser importance to data located close to the cap edge. In contrast, the cumulative energy of all zonal and nonzonal windows is more evenly spread across the concentration domain. While data adjacent to the cap edge are still somewhat downweighted, this is not nearly as drastic as for the case when only zonal tapers are used. We note that if all  $(L + 1)^2$  localization windows were to be used, the cumulative energy would be  $(L + 1)^2$  everywhere [21], though in this case, the resulting multitaper estimate would not be localized to any particular spatial region of the sphere.

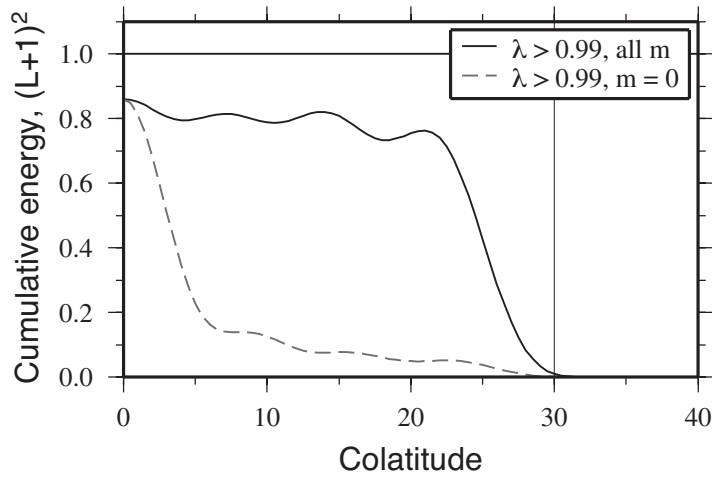


FIGURE 2 Cumulative energy of the nearly perfectly concentrated ( $\lambda > 0.99$ ) localization windows of Figure 1 ( $\theta_0 = 30^\circ$  and  $L = 29$ ). The solid curve is the cumulative energy for all 34 tapers, whereas the dashed curve is for the subset of the four zonal tapers. If all  $(L + 1)^2$  tapers were employed, the cumulative energy would be  $(L + 1)^2$  everywhere.

We define the multitaper localized power spectrum estimate of a function  $f$  as a weighted average of direct spectral estimates obtained from  $K$  orthogonal tapers:

$$S_{\Phi\Phi}^{(mt)}(l) = \sum_{k=1}^K a_k S_{\Phi\Phi}^{(k)}(l), \quad (2.17)$$

with the constraint that the sum of the weights is unity,

$$\sum_{k=1}^K a_k = 1. \quad (2.18)$$

It is clear from (2.11) and (2.17)–(2.18) that the expectation of this estimate is given by

$$\langle S_{\Phi\Phi}^{(mt)}(l) \rangle = \sum_{j=0}^L \left( \sum_{k=1}^K a_k S_{hh}^{(k)}(j) \right) \sum_{i=|l-j|}^{l+j} S_{ff}(i) \left( C_{j0i0}^{l0} \right)^2, \quad (2.19)$$



and that its bias is

$$\text{bias} \left\{ S_{\Phi\Phi}^{(mt)}(l) \right\} = \sum_{k=1}^K a_k \left( \left\langle S_{\Phi\Phi}^{(k)}(l) \right\rangle - S_{ff}(l) \right) = \sum_{k=1}^K a_k B_k, \quad (2.20)$$

where the elements of the vector  $\mathbf{B}$  depend implicitly upon the power spectrum of the global function, the power spectrum of the  $k$ -th localization window, and the spherical harmonic degree  $l$ . Using the above definitions, we show in Appendix C that the variance of the multitaper estimate can be written as

$$\text{var} \left\{ S_{\Phi\Phi}^{(mt)}(l) \right\} = \sum_{j=1}^K \sum_{k=1}^K a_j F_{jk} a_k, \quad (2.21)$$

where the symmetric covariance matrix  $\mathbf{F}$  depends both upon  $S_{ff}(l)$  and the expansion coefficients of the window, and is given by the lengthy Equations (C.12) and (C.23). In Appendix C we generalize the expression (2.21) to apply to cross-power spectra as well. Appendix D describes how to calculate the covariance of two multitaper spectral estimates at two different degrees.

We next address the question of which values to use for the weights  $a_j$  when constructing a multitaper estimate. Two cases that are in common use in the time series community are either to take weights that are all equal to  $1/K$ , or weights that are proportional to the eigenvalues (2.14) of the localization windows (this latter case helps simplify some mathematical relationships [4, 31]). An alternative approach would be to instead solve for those weights that minimize some combination of the variance and bias of the estimate. As an example, if it were important to obtain estimates that possessed both low variance and low bias, then an appropriate measure to minimize might be the mean-squared error of the spectral estimate at a particular degree  $l$ , which is simply a sum of the variance and squared bias:

$$\text{mse} = \text{var} + \text{bias}^2 = \sum_{j=1}^K \sum_{k=1}^K a_j (F_{jk} + B_j B_k) a_k. \quad (2.22)$$

Thomson [26] advocated a similar (though approximate) approach that he referred to as “adaptive weighting.” However, since his study used windows that were not bandlimited, the choice was made to consider only that portion of the bias that resulted from frequencies greater than the effective bandwidth of the window (i.e., the broadband bias). Since both the covariance matrix  $\mathbf{F}$  and bias  $\mathbf{B}$  depend upon the unknown global power spectrum, it is clear that such a minimization procedure would, in general, be iterative.

Instead of attempting to minimize the mean-squared error, a different philosophy is to solve for those weights  $a_j$  that minimize solely the variance of the multitaper spectral estimate. While the bias of such an estimate would naturally be larger, this potentially unfavorable characteristic is countered by the fact that the bias is completely *quantifiable* when the global power spectrum is known. For a large class of inverse problems, one is concerned with comparing the biased multitaper spectral estimate directly to a similarly biased theoretical model. In this situation, all that is important is how the goodness-of-fit between the two spectra varies as a function of the theoretical model parameters, and not how closely the windowed power spectra match their global equivalents. When it is easy to account for the estimation bias, the relevant quantity to minimize is naturally the variance of the windowed spectral estimates. Since many scientific problems that

use multitaper analyses are done so in the context of comparing forward models to the observations, minimum-variance multitaper spectral estimation will be emphasized in the following sections.

The numerical values of the weights  $a_j$  that minimize the mean-squared error of (2.22) will here be solved for subject to the constraint that the sum of the weights is unity. (To obtain the minimum variance solution, it is only necessary to disregard the vector  $\mathbf{B}$ .) This is easily accomplished by minimizing the objective function

$$\Psi = \sum_{i=1}^K \sum_{j=1}^K a_i (F_{ij} + B_i B_j) a_j + \lambda \left( \sum_{k=1}^K a_k - 1 \right) \quad (2.23)$$

with respect to the weights  $a_j$  and Lagrange multiplier  $\lambda$ , which yields the following set of linear equations:

$$\frac{\partial \Psi}{\partial a_n} = \sum_{i=1}^K a_i (F_{in} + B_i B_n) + \sum_{j=1}^K a_j (F_{nj} + B_n B_j) + \lambda = 0, \quad (2.24)$$

$$\frac{\partial \Psi}{\partial \lambda} = \sum_{k=1}^K a_k - 1 = 0. \quad (2.25)$$

Since  $\mathbf{F}$  is symmetric, these equations can be written in matrix notation as

$$\begin{bmatrix} 2(F_{11} + B_1 B_1) & \cdots & 2(F_{1K} + B_1 B_K) & 1 \\ 2(F_{21} + B_2 B_1) & \cdots & 2(F_{2K} + B_2 B_K) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 2(F_{K1} + B_K B_1) & \cdots & 2(F_{KK} + B_K B_K) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_K \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \quad (2.26)$$

In order to find the optimal values of the weights  $a_j$  (as well as  $\lambda$ , which is not further needed), it is only necessary to solve a simple linear equation. If one needs to calculate the covariance matrix  $\mathbf{F}$ , the weights are easily obtained with little additional computational effort.

### 3. White and Red Stochastic Processes

Many physical processes obey power-law behavior in the sense that their power spectrum varies as

$$S_{ff}(l) \sim l^\beta. \quad (3.1)$$

When the exponent  $\beta$  is equal to zero, the total power per spherical harmonic degree is constant, and we refer to the process as *white*. In contrast, when the exponent is less than zero, the power decreases with increasing spherical harmonic degree and the spectrum is *red*. We note that this terminology depends on the definition of the power spectrum (2.7), which may differ between fields of application [4]. Common examples of red spectra include planetary gravitational fields and topography. In this section, we take two representative values of  $\beta$ , namely 0 and  $-2$ , and describe in detail the properties of localized power spectrum estimates using the windows described in Section 2.2. For a discussion on Cartesian spectral analysis of power law processes, see [15].

The relationship between a global power spectrum  $S_{ff}$  and the expectation of its localized equivalent  $S_{\Phi\Phi}$  is described by (2.11). In particular, for a given degree  $l$ , the expectation of the localized spectrum depends upon the global power spectrum within the degree range  $l \pm L$ , where  $L$  is the bandwidth of the localization window. Thus, while  $S_{\Phi\Phi}$  should be expected to resemble the global spectrum, it will nevertheless be biased. In Figure 3, we plot the localized power spectra of a white (left) and red (right) power law process using the 34 windows displayed in Figure 1. As is readily seen, the localized spectra do indeed resemble the global spectra (shown by the heavy dashed lines), and appear to asymptotically approach the global values at high degrees. However, for degrees close to or less than the bandwidth of the window, the bias can be appreciable. In particular, for degrees close to zero the localized power spectrum is always biased down, whereas for degrees greater than about  $L/2$  the bias is positive.

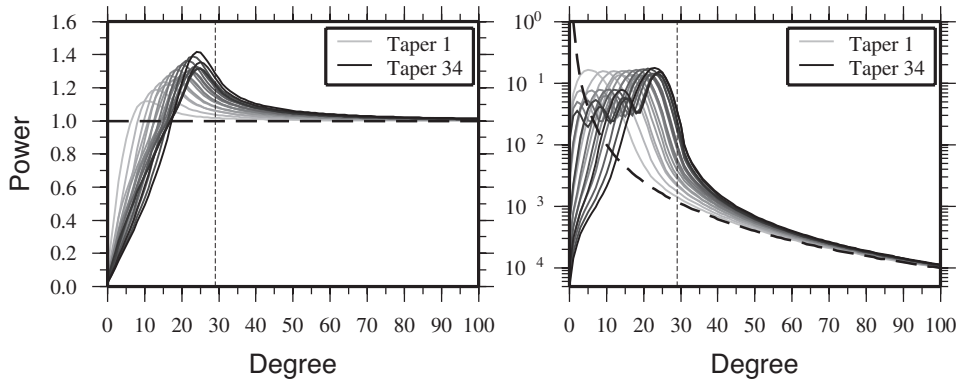


FIGURE 3 Expectations of localized power spectra for stationary stochastic global processes. Heavy dashed lines represent white (left) and red (right) global power spectra with  $\beta = 0$  and  $-2$ , respectively. The expectations of the localized spectra were generated using each of the 34 localization windows shown in Figure 1; the taper numbers increase from light gray to black. Vertical lines denote the spectral bandwidth  $L = 29$  of the tapers.

The bias is clearly worse for the red-spectrum example than for the white example. This is simply a result of the fact that for red spectra, the range of the global power spectrum within the degree range  $l \pm L$  is greater for small  $l$  than for large  $l$ . As a result, the high power at small  $l$  in the global spectrum will disproportionately influence the sum in (2.11), “leak” towards higher degrees, and bias the localized power spectrum upwards. In contrast, for large  $l$ , the global power spectrum can be considered to be approximately constant within the degree range that contributes to the localized spectral estimate, and the bias properties are similar to those in the white-spectrum example. We note that for degrees greater than the bandwidth of the window, the bias appears to increase with increasing taper number. This is simply related to the shape of the power spectrum of the localization windows [21, 31]. For the best localization windows, the power is concentrated at the lowest degrees, which acts to give the window an “effective” bandwidth that is somewhat less than  $L$ . In contrast, for higher taper numbers, the power becomes more evenly spread across the nominal bandwidth, ensuring that all degrees in the range  $l \pm L$  contribute appreciably to the sum in (2.11).

The variance properties of localized multitaper spectral estimates at degrees  $l = 30$  (left) and  $65$  (right) are displayed in Figure 4 for the case where the global power spectrum is white. The upper panel plots the covariance matrix  $\mathbf{F}$  (2.21), and as is seen, the largest

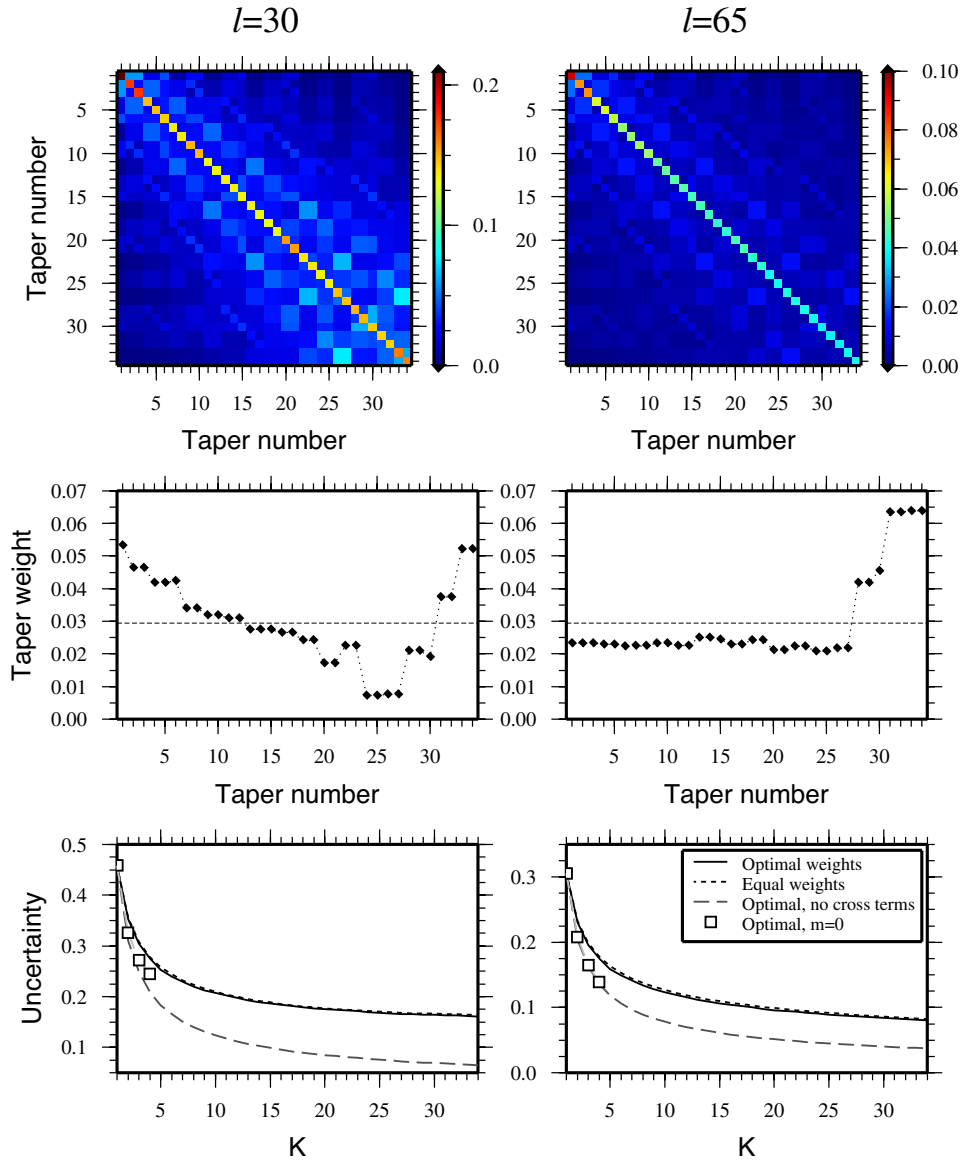


FIGURE 4 (top) Covariance matrix between the 34 best concentrated tapers of the spherical-cap concentration problem in Figure 1 for a white global spectrum at  $l = 30$  (left) and 65 (right). (center) Optimal weights,  $a_j$ , that minimize the variance of the spectral estimate at  $l = 30$  and 65 when all 34 tapers are used. The dotted horizontal line corresponds to equal weights of  $1/34$ . (bottom) Uncertainty of the spectral estimate at  $l = 30$  and 65 as a function of the number of employed tapers. Shown are the cases of using optimal weights, equal weights, optimal weights if all the off-diagonal terms of the covariance matrix were zero, and optimal weights for only the zonal tapers.

contributors are the diagonal terms, indicating that the spectral estimates from individual windows are not too highly correlated. The off-diagonal terms appear to be relatively less important for the case of  $l = 65$  in comparison to  $l = 30$ , and this trend continues with increasing degree. Thus, as quantified below, the individual spectral estimates that make

up the multitaper estimate can be considered as being somewhat statistically independent, with the level of this approximation improving with increasing degree.

The middle panel of Figure 4 shows the weights that minimize the variance of the localized multitaper spectral estimate when all 34 tapers are used, and the lower panel shows its square root, the uncertainty, as a function of  $K$ . Somewhat surprisingly, even though the optimal weights  $a_j$  are decidedly nonuniform, the uncertainty of the multitaper estimate using these weights (solid curves) differs only insignificantly from what would arise if the weights were all equal (short-dashed curves). For comparative purposes, we plot the optimal uncertainty using only the four zonal localization windows (squares). As noted by [31], the multitaper estimates using zonal windows are nearly uncorrelated, and their uncertainty decreases as  $1/\sqrt{K}$ . By including the 30 nonzonal windows, which are equally well concentrated as the zonal ones, the uncertainty has been reduced by an additional 50%. The uncertainty that would arise if all the individual spectral estimates were completely uncorrelated is also shown, calculated by setting the cross-terms of the covariance matrix equal to zero (long-dashed curves). This variance is lower in magnitude than the corresponding curve that includes the cross terms, which demonstrates that while the nonzonal windows are useful for reducing the uncertainty of multitaper spectral estimates, the individual spectral estimates are not entirely independent.

The uncertainties associated with the optimal multitaper spectral estimates are shown in the left panel of Figure 5 for spherical harmonic degrees between  $l = 30$  and 100 and as a function of the number of employed localization windows. Here, we only show results for  $l > L$  since (1) the smaller degrees are influenced by the magnitude of the degree-0 term, which is often statistically unrelated to the other degrees for many physical processes, (2) as shown in Figure 3, the windowed spectrum is highly biased for smaller degrees, and (3) it is unreasonable to expect that wavelengths larger than the size of the window would be well resolved in the localized spectra. As demonstrated in the previous figure, the uncertainty of the multitaper spectral estimate decreases with increasing number of localization windows. Furthermore, the variance decreases with increasing spherical harmonic degree, though

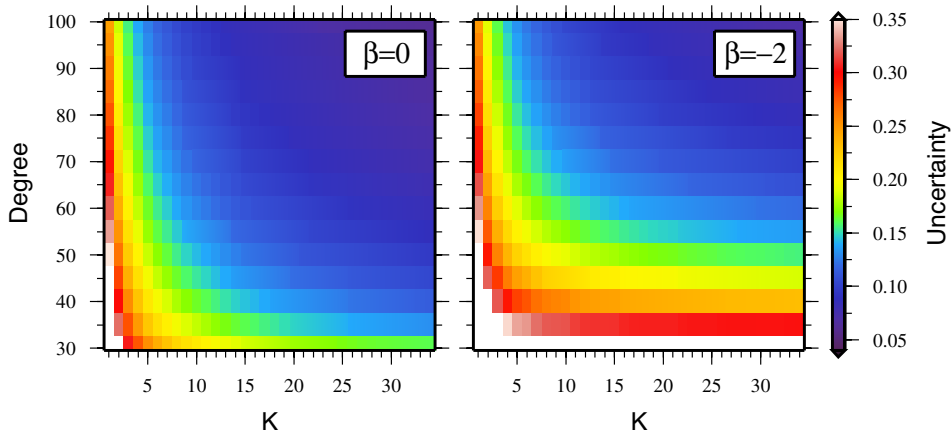


FIGURE 5 Optimal localized multitaper uncertainty of a white (left) and red (right) stochastic process as a function of spherical harmonic degree and the number of employed tapers. For the case of the red spectrum, the uncertainty is scaled by the square root of the global power spectrum. As in the previous figures, the windows were constructed using  $\theta_0 = 30^\circ$  and  $L = 29$ . As a result of the computationally intensive nature of these calculations, the uncertainty was calculated only for degrees that are multiples of 5.

somewhat more slowly. As should be readily visible, if only a few windows were used to generate a multitaper estimate, its uncertainty could be higher than 30%.

Figures 5 (right) and 6 show analogous results for the case when the underlying global power spectrum is red, with  $\beta = -2$ . We first note that for  $l = 65$  in Figure 6 (right), the covariance matrix and multitaper uncertainties are very similar to those in the white example. As mentioned previously, this is simply because the global power spectrum varies slowly at high degrees and is approximately constant within the degree range  $65 \pm L$ . In contrast, the results for  $l = 30$  are dramatically different. First, the diagonal elements of

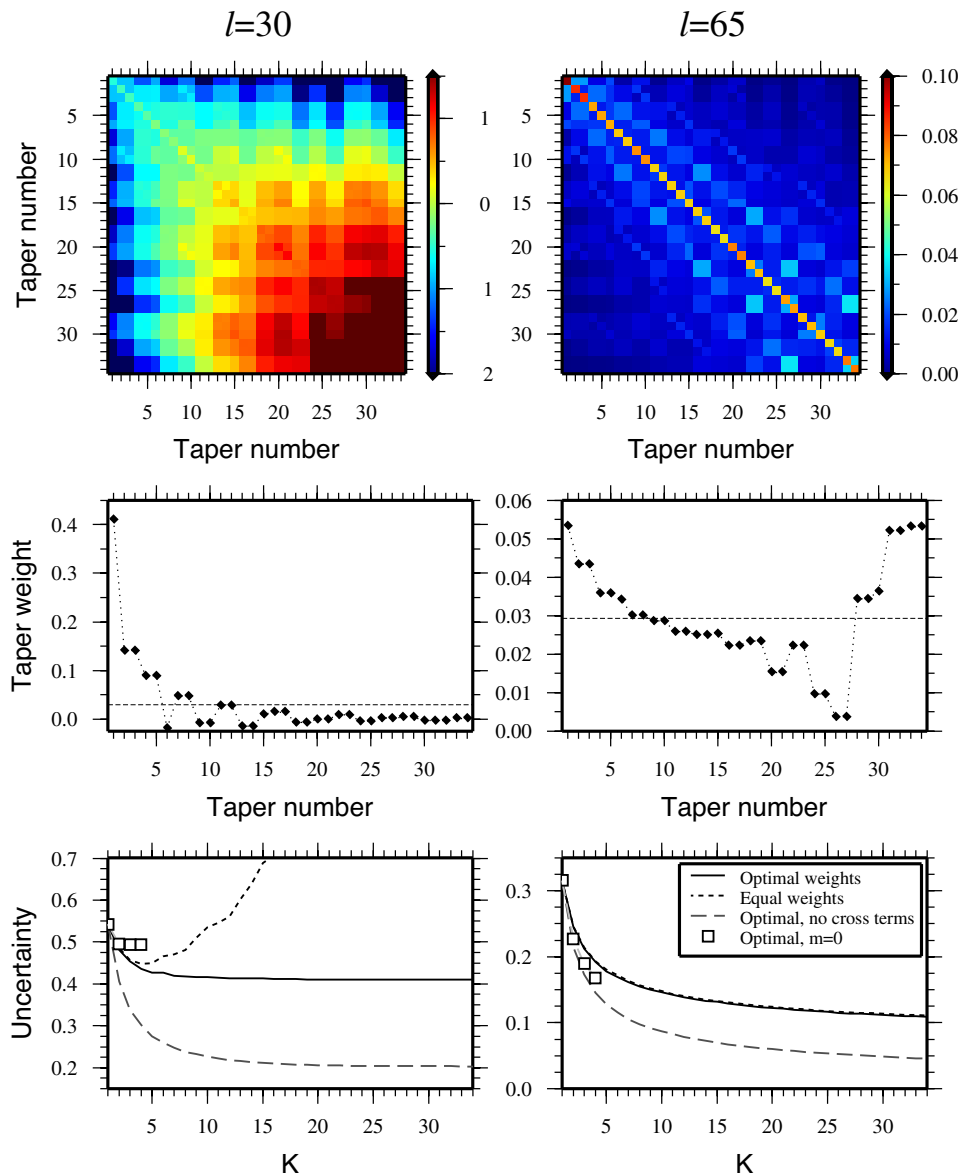


FIGURE 6 Same as Figure 4, but for a red global spectrum with  $\beta = -2$ . The covariance matrix (top) was scaled by  $[S_{ff}(l)]^2$ , and the uncertainty of the spectral estimates (bottom) by  $S_{ff}(l)$ .

the covariance matrix are seen to increase appreciably with increasing taper number. Thus, the spectral estimates obtained using higher taper numbers will possess larger uncertainties than those using smaller taper numbers. This is related to the fact that the bias for each localization window increases significantly with increasing taper number (see Figure 3). The off-diagonal terms are also seen to become increasingly important as both indices of the covariance matrix  $F_{ij}$  increase.

The lower left panel of Figure 6 shows that if one were to use equal weights for obtaining a multitaper spectral estimate at  $l = 30$ , the uncertainty of this estimate would actually increase after using more than the first five localization windows. This is simply a result of the larger uncertainties associated with higher taper numbers as demonstrated by the covariance matrix. For this degree, the use of optimal weights for minimizing the uncertainty of the multitaper estimate is critical. As shown in the middle panel, the uncertainty of the spectral estimate is minimized in this case by, in essence, only utilizing the first five localization windows. The inclusion of additional windows leads to no appreciable decrease in variance. It is further noted that given the highly structured covariance matrix for this example, the minimum obtainable variance is far from what would result if all the spectral estimates were uncorrelated.

The examples shown in Figures 4–6 demonstrate that calculating the covariance matrix  $\mathbf{F}$  and the associated optimal weights is in practice only necessary for low degrees when the underlying process is significantly red. As computation of the covariance matrix can be somewhat time consuming (depending on the bandwidth of the localization window, the number of localization windows, and the spherical harmonic degree) it would be useful to have a criterion for when the use of equal weights is adequate, and when the use of optimal weights is necessary. The simplest approach would be to use equal weights initially, and to then plot the uncertainty of the multitaper spectral estimate as a function of the number of tapers used in its construction. If the uncertainty fails to decrease, as in the lower left panel of Figure 6, using optimal weights could reduce the estimation variance significantly.

#### 4. Single Realizations of Stochastic Processes

In the preceding section we assumed that the stochastic process giving rise to the global power spectrum was known, and this allowed for the analytic computation of the expectation (2.19) and variance (2.21) of the multitaper spectral estimate. Unfortunately, for many physical processes, not only is the underlying power spectrum of the stochastic process unknown, but only a single realization is available for analysis: there is only one gravitational field of the Earth, only one cosmic microwave background, and so on. As Figures 4 and 6 demonstrate, the spectral estimates that result from orthogonal localization windows are, in general, somewhat uncorrelated. Thus, even though only a single realization of a process might be available for analysis, each windowed spectral estimate can be treated approximately as if it were derived from a separate realization. As the number of localization windows used in constructing the multitaper estimate increases, we expect the variance of this estimate to decrease accordingly.

When the underlying global power spectrum is not known *a priori*, the primary difficulty lies in how to estimate the uncertainty of the multitaper estimate. One approximate approach would be to assume that the multitaper estimate is equal to the global value, and to compute the expected variance using the expressions in Appendix C. If a more accurate estimate of the uncertainty were desired, one could attempt inverting for the global power spectrum (see below), and using this to calculate the expected multitaper uncertainty. From

the calculated covariance matrix, optimal weights that minimize the multitaper variance could be obtained, and these could be used to form a new multitaper estimate. By repeating this process, one would ultimately expect to converge on the minimum-variance multitaper spectrum estimate. This procedure, however, suffers from having to calculate the covariance matrix  $\mathbf{F}$  at each degree  $l$  several times. For many problems, this approach is unfeasible given the computationally intensive nature of the covariance matrix calculations.

As a more practical, but necessarily approximate, approach we will make the assumption that for a given degree  $l$  the individual spectral estimates  $S_{\Phi\Phi}^{(k)}$  that contribute to the multitaper spectrum estimate are statistically independent and Gaussian distributed with identical variance  $\sigma_{\Phi\Phi}^2$ . This amounts to assuming that the off-diagonal terms of the covariance matrix are zero and that the diagonal terms are all equal. In this case it is easily shown that the variance of the estimate (2.17)–(2.18) is

$$\text{var}\{S_{\Phi\Phi}^{(mt)}\} = \sigma_{\Phi\Phi}^2 \sum_{k=1}^K a_k^2 \quad \text{with} \quad \text{var}\{S_{\Phi\Phi}^{(k)}\} = \sigma_{\Phi\Phi}^2. \quad (4.1)$$

Defining the weighted sample variance of individual spectral estimates as

$$\sigma^2 = \sum_{k=1}^K a_k \left( S_{\Phi\Phi}^{(k)} - S_{\Phi\Phi}^{(mt)} \right)^2, \quad (4.2)$$

its expectation is found to be equal to

$$\langle \sigma^2 \rangle = \sigma_{\Phi\Phi}^2 \left( 1 - \sum_{k=1}^K a_k^2 \right) + \sum_{k=1}^K a_k \left( \langle S_{\Phi\Phi}^{(k)} \rangle - \langle S_{\Phi\Phi}^{(mt)} \rangle \right)^2. \quad (4.3)$$

The second term in the above equation is a measure of the variability of the expectations of the  $K$  windowed estimates  $S_{\Phi\Phi}^{(k)}$  about their weighted mean (referring to Figure 3, of the curves of the expected values of single-taper estimates around the expected value of the multitaper estimate). By assuming that the statistical spread of each windowed estimate is greater than the spread of the individual expectations, this term can be ignored. Combining (4.1) and (4.3), the following unbiased estimate for the variance of the multitaper estimate is obtained:

$$\text{var}\{S_{\Phi\Phi}^{(mt)}\} \approx \sigma^2 \left( \frac{\sum_k a_k^2}{1 - \sum_k a_k^2} \right). \quad (4.4)$$

In contrast to (2.21), which requires knowledge of the global spectrum, this estimate, though approximate, is determined from the data alone. When the weights  $a_k$  are all equal, the variance of the multitaper spectrum estimate is simply  $\sigma^2/(K-1)$ . Having ignored what is essentially the sample variance of the expectations of the windowed estimates, we are insured that the calculated uncertainty (4.4) will be an overestimate. However, we should also note that the assumption of each individually tapered spectral estimate being statistically independent will not in general be true, and this will cause the above uncertainty to underestimate the true value. We note that [31, Eq. 48] previously advocated using the sample variance  $\sigma^2$  of the individual spectral estimates as an estimate for the variance of the multitaper estimate, which, in hindsight, is unnecessarily conservative.

In order to assess the suitability of the above approximations to estimate the uncertainty via (4.4), we generated three realizations of a white stochastic process and calculated localized multitaper estimates. In the upper panel of Figure 7, the absolute value of the



difference between the multitaper estimate and its expectation based on the known input spectrum is shown as a function of degree  $l$  and number of tapers  $K$  for the three realizations. As is readily seen, the difference between the two almost everywhere decreases with increasing number of employed localization windows. Furthermore, comparison with Figure 5 shows that the difference between the two is compatible with the expected uncertainty of the multitaper estimate. The bottom panel plots the uncertainty of the multitaper estimate using (4.4), and this is seen to be generally comparable to the difference between the multitaper expectation and single realization in the upper panel. Nevertheless, it should be noted that this estimation of the uncertainty underestimates the true value as shown in Figure 5 by a small factor. This is most likely a result of the fact that the individual spectral estimates are not uncorrelated as we assumed in deriving (4.4).

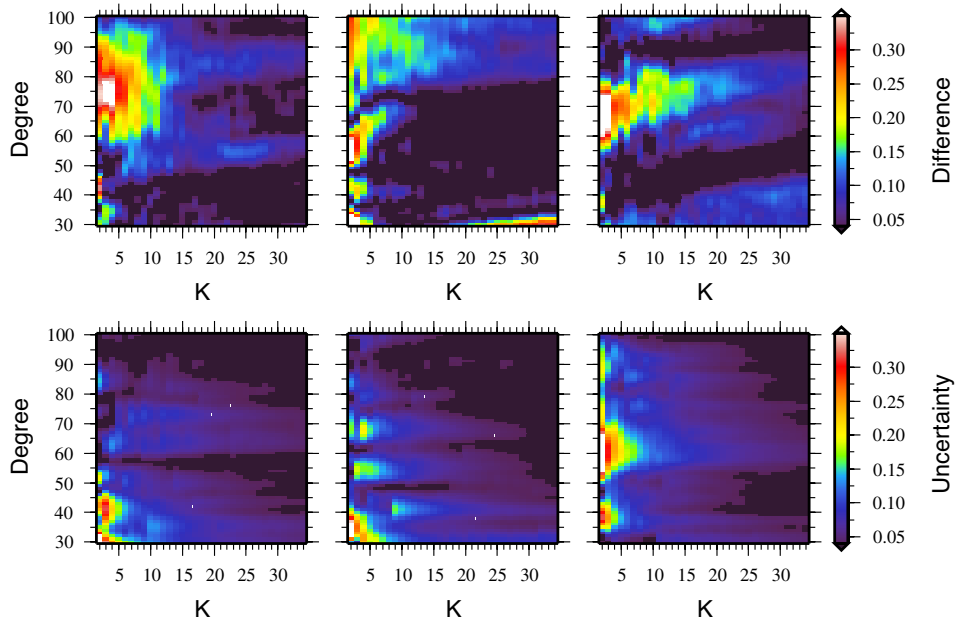


FIGURE 7 Difference and estimated uncertainty for three multitaper realizations of a white stochastic process. (top) Absolute value of the difference between the known multitaper spectrum expectation and the single realization of the multitaper estimate. (bottom) The estimated uncertainty of the multitaper spectrum estimate calculated from (4.4) using equal weights. Between 2 and 34 tapers were used in generating the multitaper estimates, the tapers were constructed using  $\theta_0 = 30^\circ$  and  $L = 29$ , and the data were localized at the North pole. For comparison purposes, the color bars possess the same range as those in Figure 5.

Finally, we note that it is possible under certain circumstances to invert for the global power spectrum using (2.19) combined with knowledge of the multitaper spectral estimates and their uncertainties. As shown in [31], this equation can be written in matrix notation as

$$\langle \mathbf{S}_{\Phi\Phi}^{(mt)} \rangle = \mathbf{M}^{(mt)} \mathbf{S}_{ff} , \tag{4.5}$$

where  $\mathbf{S}_{\Phi\Phi}^{(mt)}$  is a vector containing the  $L_{\Phi\Phi} + 1$  multitaper spectral estimates,  $\mathbf{S}_{ff}$  is a vector containing the  $L_{\Phi\Phi} + L + 1$  elements of the global power spectrum, and  $\mathbf{M}^{(mt)}$  is an  $(L_{\Phi\Phi} + 1) \times (L_{\Phi\Phi} + L + 1)$  matrix that maps the latter into the former. Assigning the

index 0 to the first row and column of  $\mathbf{M}$ , the elements are given by

$$M_{ij}^{(mt)} = \sum_{l=0}^L \sum_{k=1}^K a_k S_{hh}^{(k)}(l) \left( C_{l0j0}^{i0} \right)^2. \quad (4.6)$$

The individual linear equations of (4.5) could be weighted by the measurement uncertainties by dividing each row of  $\mathbf{S}_{\Phi\Phi}^{(mt)}$  and  $\mathbf{M}^{(mt)}$  by the multitaper standard deviation obtained from (4.4).

Inverting  $\mathbf{M}^{(mt)}$  to obtain  $\mathbf{S}_{ff}$  from  $\mathbf{S}_{\Phi\Phi}^{(mt)}$  is an underdetermined inversion problem, as the dimension of  $\mathbf{S}_{ff}$  will always be greater than that of  $\mathbf{S}_{\Phi\Phi}^{(mt)}$ , and  $\mathbf{M}^{(mt)}$  is never full-rank. Hence, most of our attempts to invert (4.5) for the global power spectrum, using a variety of linear techniques, have been unsatisfactory for this example. For instance, choosing that solution for which the norm of  $\mathbf{S}_{ff}$  is minimized yielded a power spectrum with both positive and negative values, a situation that is clearly unphysical. Truncating the matrix  $\mathbf{M}$  to be square yielded similar results. A nonnegative least squares inversion [13, Chap. 23] yielded a solution for which the majority of the  $S_{ff}(l)$  were zero. One method that did yield acceptable results was to assume that  $S_{ff}$  was constant in bins of width  $\Delta l$ . However, for the problem at hand,  $\Delta l$  was required to be greater than about 15 in order to obtain positive values with reasonable variances. More sophisticated nonlinear and Monte Carlo techniques that utilize positivity constraints and upper and lower bounds are worth investigating. Alternatively, one could parametrize the global power spectrum by a smoothly varying function (such as a power law) and invert for the parameter values that best fit the observed multitaper spectral estimates. A maximum-likelihood inversion approach is described in [4] and [8].

## 5. Concluding Remarks

In this study, we have demonstrated how the global power spectrum of a stochastic process on the sphere can be estimated from spatially limited observation domains. In particular, the act of restriction to a certain region can be formulated as a windowing operation, and the expectation of the windowed power spectrum has been shown to be related to the global power spectrum by a convolution-type equation. The best localization windows are those that are spatially concentrated in the region of interest, and yet have as small an effective spectral bandwidth as possible. Solution of a simple optimization problem yields a family of orthogonal functions, and a multitaper spectrum estimate can be constructed using those windows with good spatio-spectral localization properties. The multitaper estimate has the benefits that the data are more evenly weighted and that it has a reduced variance when compared to single-window estimates. It is straightforward to choose the weights used in constructing the multitaper estimate in order to minimize the estimation variance.

While much of the theoretical groundwork has been developed for the problem of localized spectral estimation on the sphere, several promising lines of future research could improve upon the procedure developed in this work. In particular, we have only concerned ourselves with localization windows that are solutions to the spherical-cap concentration problem [6, 21, 31]. While sufficient for many problems, such as estimating the localized spectral properties of planetary gravitational fields and topography, more complex localization windows might be desired for other applications. The techniques developed here are easily generalizable to other concentration domains [20, 21], and we do not expect the general character of our results to differ significantly from those presented here.

As a second line of future research, we note that we have restricted ourselves to using only a single localization window when calculating the localized cross-power spectrum  $S_{\Phi\Gamma}^{(k)}$ . However, as demonstrated in Appendix B, we could have localized each field by a different window to form localized cross-power spectra  $S_{\Phi\Gamma}^{(j,k)}$ . Using two different windows, up to  $K(K+1)/2$  individual cross-power spectral estimates are, in principle, available for analysis compared to  $K$  as used in this study. The covariance properties of these windows remain to be investigated, as well as whether their use would significantly decrease the multitaper estimation variance.

Of more practical concern are the computational demands to obtain the covariance matrix. Generating the left and right panels of Figure 4 took about 12 hours and 2 days of dedicated time on a modern desktop computer, respectively. The calculations for Figure 5 took considerably longer, about one month. Clearly, such computations will become increasingly infeasible as the spherical harmonic degree increases, and alternative means will eventually be necessary for computing the covariance matrix, optimal weights, and uncertainties. Asymptotic relations for the Clebsch-Gordan coefficients and/or covariance matrix may be used to simplify these numerical computations at high degrees [4].

Finally, we note that most of our discussion has emphasized the quantification of the bias and uncertainty of the multitaper spectrum estimate. However, in practice, one is often instead interested in obtaining an unbiased estimate of the global power spectrum. While one can in principle invert (2.19) for this quantity, the standard linear inversion techniques discussed in Section 4 yielded mixed results. More sophisticated nonlinear and Monte Carlo techniques utilizing bounds and positivity constraints may be worth investigating in this context.

## Appendix

### A. Real and Complex Spherical Harmonics

The derivations in this study are considerably simplified if complex spherical harmonics are used. By inserting the identities

$$\cos \beta = \frac{e^{i\beta} + e^{-i\beta}}{2} \quad \text{and} \quad \sin \beta = \frac{e^{i\beta} - e^{-i\beta}}{2i} \quad (\text{A.1})$$

into (2.1), the spherical harmonic expansion of a real function  $f$  is expressed as

$$\begin{aligned} f(\Omega) &= \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm} Y_{lm}(\Omega) \\ &= \sum_{l=0}^{\infty} \sum_{m=0}^l \left[ f_{lm} \left( \frac{e^{im\phi} + e^{-im\phi}}{2} \right) + f_{l-m} \left( \frac{e^{im\phi} - e^{-im\phi}}{2i} \right) \right] \bar{P}_{lm}(\cos \theta) \\ &= \sum_{l=0}^{\infty} \sum_{m=0}^l \left[ e^{im\phi} \left( \frac{f_{lm} - if_{l-m}}{2} \right) + e^{-im\phi} \left( \frac{f_{lm} + if_{l-m}}{2} \right) \right] \bar{P}_{lm}(\cos \theta) \\ &= \sum_{l=0}^{\infty} \sum_{m=-l}^l f_l^m Y_l^m(\Omega), \end{aligned} \quad (\text{A.2})$$

where the complex spherical harmonics,  $Y_l^m$ , are defined as

$$Y_l^m(\Omega) = \sqrt{2l+1} \sqrt{\frac{(l-m)!}{(l+m)!}} P_{lm}(\cos\theta) e^{im\phi}. \quad (\text{A.3})$$

Since

$$P_{l-m} = (-1)^m \frac{(l-m)!}{(l+m)!} P_{lm}, \quad (\text{A.4})$$

the complex harmonics satisfy the identity

$$Y_l^{m*}(\Omega) = (-1)^m Y_l^{-m}(\Omega), \quad (\text{A.5})$$

and normalization

$$\frac{1}{4\pi} \int_{\Omega} Y_l^{m*}(\Omega) Y_{l'}^{m'}(\Omega) d\Omega = \delta_{ll'} \delta_{mm'}, \quad (\text{A.6})$$

where the superscript \* indicates complex conjugation. To be consistent with the definition of the real spherical harmonic functions in Section 2.1, the complex harmonics here also do not include the Condon-Shortley phase factor of  $(-1)^m$  that generally appears in the physics and seismology communities. Regardless, we note that the inclusion or exclusion of this phase will not affect the results presented in this article. The complex coefficients are related to the real coefficients by

$$f_l^m = \begin{cases} (f_{lm} - i f_{l-m})/\sqrt{2} & \text{if } m > 0 \\ f_{l0} & \text{if } m = 0 \\ (-1)^m f_l^{-m*} & \text{if } m < 0, \end{cases} \quad (\text{A.7})$$

and by using the orthogonality properties of the spherical harmonics, these can be shown to be related to the real function  $f$  by the relation

$$f_l^m = \frac{1}{4\pi} \int_{\Omega} f(\Omega) Y_l^{m*}(\Omega) d\Omega. \quad (\text{A.8})$$

It is straightforward to show that the total power of a real function  $f$  is related to its complex spectral coefficients by a generalization of Parseval's theorem:

$$\frac{1}{4\pi} \int_{\Omega} [f(\Omega)]^2 d\Omega = \sum_{l=0}^{\infty} S_{ff}(l), \quad (\text{A.9})$$

where the power spectrum is

$$S_{ff}(l) = \sum_{m=-l}^l f_l^m f_l^{m*}. \quad (\text{A.10})$$

Similarly, the cross-power of two real functions  $f$  and  $g$  is given by

$$\frac{1}{4\pi} \int_{\Omega} f(\Omega) g(\Omega) d\Omega = \sum_{l=0}^{\infty} S_{fg}(l), \quad (\text{A.11})$$

where the cross-power spectrum is

$$S_{fg}(l) = \sum_{m=-l}^l f_l^m g_l^{m*}. \quad (\text{A.12})$$

If the functions  $f$  and  $g$  have a zero mean (i.e., their degree-0 terms are equal to zero), then  $S_{ff}(l)$  and  $S_{fg}(l)$  represent the contributions to their variance and covariance, respectively, for degree  $l$ .

### B. Bias of a Localized Spectral Estimate

In this section, expressions will be derived that relate the cross-power spectrum of two global fields to their windowed equivalents. We assume that the real spherical harmonic coefficients of a function  $f$  are zero-mean random variables, and that the power spectrum of the function is isotropic, i.e., depends only upon degree  $l$ :

$$\langle f_{lm} f_{l'm'} \rangle = \frac{S_{ff}(l)}{(2l + 1)} \delta_{ll'} \delta_{mm'} , \tag{B.1}$$

where  $\langle \dots \rangle$  signifies the expectation operator. When considering two fields, it will be assumed in a similar manner that their cross-power at a given degree is also isotropic:

$$\langle f_{lm} g_{l'm'} \rangle = \frac{S_{fg}(l)}{(2l + 1)} \delta_{ll'} \delta_{mm'} . \tag{B.2}$$

It can be verified using (A.7) that

$$\langle f_{lm} f_{l'm'} \rangle = \langle f_l^{m*} f_{l'}^{m'} \rangle = \langle f_l^{m*} f_{l'}^{m'} \rangle^* , \tag{B.3}$$

and

$$\langle f_{lm} g_{l'm'} \rangle = \langle f_l^{m*} g_{l'}^{m'} \rangle = \langle f_l^{m*} g_{l'}^{m'} \rangle^* . \tag{B.4}$$

The goal of this appendix is to find an expression for the expectation of the cross-power spectrum of the functions  $f$  and  $g$ , each localized by a different data taper  $h^{(i)}$  and  $h^{(j)}$ , respectively. This quantity will be denoted as  $\langle S_{\Phi\Gamma}^{(i,j)}(l) \rangle$ , where  $\Phi$  and  $\Gamma$  represent the localized fields  $f h^{(i)}$  and  $g h^{(j)}$ , respectively. The spectral bandwidth of each localization window will be assumed to be the same. We start with the product of two windowed coefficients for a given degree and order:

$$\begin{aligned} \Phi_l^{m(i)} \Gamma_l^{m(j)*} &= \frac{1}{4\pi} \int_{\Omega} [h^{(i)}(\Omega) f(\Omega)] Y_l^{m*}(\Omega) d\Omega \\ &\times \frac{1}{4\pi} \int_{\Omega'} [h^{(j)*}(\Omega') g^*(\Omega')] Y_l^m(\Omega') d\Omega' . \end{aligned} \tag{B.5}$$

Expanding the windowed functions in spherical harmonics, and utilizing the short-hand notation

$$\sum_{lm} = \sum_{l=0}^L \sum_{m=-l}^l \quad \text{and} \quad \sum_l = \sum_{l=0}^L , \tag{B.6}$$

yields

$$\begin{aligned} \Phi_l^{m(i)} \Gamma_l^{m(j)*} &= \frac{1}{(4\pi)^2} \sum_{l_1 m_1}^L h_{l_1}^{m_1(i)} \sum_{l_2 m_2}^{\infty} f_{l_2}^{m_2} \sum_{l_3 m_3}^L h_{l_3}^{m_3(j)*} \sum_{l_4 m_4}^{\infty} g_{l_4}^{m_4*} \\ &\times \int_{\Omega} Y_{l_1}^{m_1} Y_{l_2}^{m_2} Y_l^{m*} d\Omega \int_{\Omega'} Y_{l_3}^{m_3*} Y_{l_4}^{m_4*} Y_l^m d\Omega' . \end{aligned} \tag{B.7}$$

Averaging this equation over all possible combinations of the random variables gives its expectation:

$$\begin{aligned} \left\langle \Phi_l^{m(i)} \Gamma_l^{m(j)*} \right\rangle &= \frac{1}{(4\pi)^2} \sum_{l_1 m_1}^L h_{l_1}^{m_1(i)} \sum_{l_3 m_3}^L h_{l_3}^{m_3(j)*} \sum_{l_2 m_2}^{\infty} \frac{S_{fg}(l_2)}{(2l_2 + 1)} \\ &\times \int_{\Omega} Y_{l_1}^{m_1} Y_{l_2}^{m_2} Y_l^{m*} d\Omega \int_{\Omega'} Y_{l_3}^{m_3*} Y_{l_2}^{m_2*} Y_l^m d\Omega'. \end{aligned} \quad (\text{B.8})$$

The integral of a triple product of spherical harmonics is real and can be evaluated by a well-known relationship involving Clebsch-Gordan coefficients [27, p. 148]

$$\int_{\Omega} Y_{l_1}^{m_1} Y_{l_2}^{m_2} Y_l^{m*} d\Omega = 4\pi \sqrt{\frac{(2l_1 + 1)(2l_2 + 1)}{(2l + 1)}} C_{l_1 0 l_2 0}^{l 0} C_{l_1 m_1 l_2 m_2}^{l m}, \quad (\text{B.9})$$

which is nonzero only when the following selection rules are satisfied [27]:

$$m = m_1 + m_2 \quad (\text{B.10})$$

$$|m| \leq l; |m_1| \leq l_1; |m_2| \leq l_2 \quad (\text{B.11})$$

$$|l_1 - l_2| \leq l \leq l_1 + l_2 \quad (\text{B.12})$$

$$|l_2 - l| \leq l_1 \leq l_2 + l \quad (\text{B.13})$$

$$|l - l_1| \leq l_2 \leq l + l_1 \quad (\text{B.14})$$

$$l_1 + l_2 + l = \text{even}. \quad (\text{B.15})$$

Combining (B.8) and (B.9), and making use of (A.5), yields

$$\begin{aligned} \left\langle \Phi_l^{m(i)} \Gamma_l^{m(j)*} \right\rangle &= \sum_{l_1 m_1}^L h_{l_1}^{m_1(i)} \sum_{l_3 m_3}^L h_{l_3}^{m_3(j)*} \sum_{l_2 m_2}^{\infty} S_{fg}(l_2) \\ &\times \frac{\sqrt{(2l_1 + 1)(2l_3 + 1)}}{2l + 1} C_{l_1 0 l_2 0}^{l 0} C_{l_3 0 l_2 0}^{l 0} C_{l_1 m_1 l_2 m_2}^{l m} C_{l_3 - m_3 l_2 - m_2}^{l - m}, \end{aligned} \quad (\text{B.16})$$

where a phase factor is set equal to unity because of (B.10). We next sum this entire equation over all values of  $m$ , employ the symmetry relationship of the Clebsch-Gordan coefficients [27, p. 245]

$$C_{l_1 - m_1 l_2 - m_2}^{l - m} = (-1)^{l_1 + l_2 - l} C_{l_1 m_1 l_2 m_2}^{l m}, \quad (\text{B.17})$$

set a phase factor equal to unity because of (B.15), and rearrange the sum over  $m_2$  to obtain

$$\begin{aligned} \left\langle S_{\Phi \Gamma}^{(i,j)}(l) \right\rangle &= \sum_{l_1 m_1}^L h_{l_1}^{m_1(i)} \sum_{l_3 m_3}^L h_{l_3}^{m_3(j)*} \sum_{l_2}^{\infty} S_{fg}(l_2) \\ &\times \frac{\sqrt{(2l_1 + 1)(2l_3 + 1)}}{2l + 1} C_{l_1 0 l_2 0}^{l 0} C_{l_3 0 l_2 0}^{l 0} \sum_{m=-l}^l \sum_{m_2=-l_2}^{l_2} C_{l_1 m_1 l_2 m_2}^{l m} C_{l_3 m_3 l_2 m_2}^{l m}. \end{aligned} \quad (\text{B.18})$$

The final sum over  $m$  and  $m_2$  is greatly simplified by use of the identity [27, p. 259]

$$\sum_{\alpha} \sum_{\gamma} C_{\alpha \alpha \beta \beta}^{c \gamma} C_{\alpha \alpha \beta' \beta'}^{c \gamma} = \frac{(2c + 1)}{(2b + 1)} \delta_{bb'} \delta_{\beta \beta'}, \quad (\text{B.19})$$

where the summations are implicitly over all values which are nonzero, and the symmetry relationship [27, p. 245]

$$C_{l_1 m_1 l_2 m_2}^{lm} = (-1)^{l_1 + l_2 - l} C_{l_2 m_2 l_1 m_1}^{lm}. \quad (\text{B.20})$$

Taking into account the selection rule (B.14), (B.19) can be succinctly written as

$$\langle S_{\Phi\Gamma}^{(i,j)}(l) \rangle = \sum_{l_1 m_1}^L h_{l_1}^{m_1(i)} h_{l_1}^{m_1(j)*} \sum_{l_2=|l-l_1|}^{l+l_1} S_{fg}(l_2) \left( C_{l_1 0 l_2 0}^{l 0} \right)^2, \quad (\text{B.21})$$

where it is clear that this quantity is real. While the sum over  $m_1$  might sometimes be zero, in general, this will not be the case, even for when  $h^{(i)}$  and  $h^{(j)}$  are orthogonal. Thus, if  $K$  tapers are being employed in the spectral estimation procedure, up to  $K(K+1)/2$  cross-spectral estimates can be obtained. For the specific case where  $f$  and  $g$  are localized by the same window, the expectation of the localized cross-power spectrum is, after a change of variables,

$$\langle S_{\Phi\Gamma}(l) \rangle = \sum_{j=0}^L S_{hh}(j) \sum_{i=|l-j|}^{l+j} S_{fg}(i) \left( C_{j 0 i 0}^{l 0} \right)^2. \quad (\text{B.22})$$

For computational purposes, we note that the Wigner 3- $j$  symbols are related to the Clebsch-Gordan coefficients by the definition [27, p. 236]

$$C_{l_1 m_1 l_2 m_2}^{lm} = (-1)^{l_1 - l_2 + m} \sqrt{2l+1} \begin{pmatrix} l_1 & l_2 & l \\ m_1 & m_2 & -m \end{pmatrix}. \quad (\text{B.23})$$

Algorithms for calculating the Wigner 3- $j$  symbols are discussed by [14] and [19].

### C. Variance of a Localized Spectral Estimate

A multitaper estimate for the localized cross-power spectrum of two fields  $f$  and  $g$  will be defined as a weighted average of direct spectral estimates obtained using  $K$  orthogonal tapers:

$$S_{\Phi\Gamma}^{(mt)}(l) = \sum_{i=1}^K \sum_{j=1}^K a_{ij} S_{\Phi\Gamma}^{(i,j)}(l), \quad (\text{C.1})$$

where the elements of  $a_{ij}$  are the weights applied to the cross-spectral estimate subject to the constraint

$$\sum_{i=1}^K \sum_{j=1}^K a_{ij} = 1. \quad (\text{C.2})$$

It is easy to verify that this quantity is real when  $a_{ij}$  is symmetric. To simplify the following derivations, and to reduce the amount of time required to calculate these numerically, we will consider only the case where  $i = j$ . Further research is required to determine the benefit of employing a spectral estimate obtained using two different tapers. In this case, the multitaper spectral estimate reduces to

$$S_{\Phi\Gamma}^{(mt)}(l) = \sum_{k=1}^K a_k S_{\Phi\Gamma}^{(k)}(l), \quad (\text{C.3})$$

with the constraint

$$\sum_{k=1}^K a_k = 1. \quad (\text{C.4})$$

This appendix seeks to determine the variance of such a spectral estimate.

We start with the definitions of variance and covariance of complex variables:

$$\text{var} \left\{ \sum_{i=1}^N a_i X_i \right\} = \sum_{j=1}^N \sum_{k=1}^N a_j \text{cov}\{X_j, X_k\} a_k \quad (\text{C.5})$$

$$\text{cov}\{X_j, X_k\} = \langle X_j X_k^* \rangle - \langle X_j \rangle \langle X_k^* \rangle. \quad (\text{C.6})$$

which give the following expression for the variance of (C.3):

$$\text{var} \left\{ S_{\Phi\Gamma}^{(mt)}(l) \right\} = \sum_{j=1}^K \sum_{k=1}^K a_j \text{cov} \left\{ S_{\Phi\Gamma}^{(j)}(l), S_{\Phi\Gamma}^{(k)}(l) \right\} a_k. \quad (\text{C.7})$$

By using the definition of the cross-power spectrum (A.11)–(A.12) with the identity

$$\text{cov} \left\{ \sum_{i=1}^N X_i, \sum_{j=1}^M X_j \right\} = \sum_{i=1}^N \sum_{j=1}^M \text{cov}\{X_i, X_j\}, \quad (\text{C.8})$$

the covariance of two spectral estimates using tapers  $j$  and  $k$  can be written as

$$\text{cov} \left\{ S_{\Phi\Gamma}^{(j)}(l), S_{\Phi\Gamma}^{(k)}(l) \right\} = \sum_{m=-l}^l \sum_{m'=-l}^l \text{cov} \left\{ \Phi_l^{m(j)} \Gamma_l^{m(j)*}, \Phi_l^{m'(k)} \Gamma_l^{m'(k)*} \right\}. \quad (\text{C.9})$$

We proceed by using Isserlis' theorem [28]

$$\text{cov}\{Z_1 Z_2, Z_3 Z_4\} = \text{cov}\{Z_1, Z_3\} \text{cov}\{Z_2, Z_4\} + \text{cov}\{Z_1, Z_4\} \text{cov}\{Z_2, Z_3\}, \quad (\text{C.10})$$

which is valid for zero-mean Gaussian complex random variables  $Z_i$ . Given that we have previously assumed that the coefficients  $f_{lm}$  and  $g_{lm}$  have a zero mean, so will the spectral coefficients of the localized fields  $\Phi_{lm}$  and  $\Gamma_{lm}$ . We may then rely on a central-limit theorem [31] to subsequently assume that the localized coefficients will approach a Gaussian distribution. Alternatively, if we assume that the coefficients of the unwrapped fields were Gaussian to begin with, then it is easily shown that the windowed coefficients will be Gaussian as well. Under these conditions (C.9) can be written as

$$\begin{aligned} \text{cov} \left\{ S_{\Phi\Gamma}^{(j)}(l), S_{\Phi\Gamma}^{(k)}(l) \right\} &= \sum_{m=-l}^l \sum_{m'=-l}^l \left( \text{cov} \left\{ \Phi_l^{m(j)}, \Phi_l^{m'(k)} \right\} \text{cov} \left\{ \Gamma_l^{m(j)*}, \Gamma_l^{m'(k)*} \right\} \right. \\ &\quad \left. + \text{cov} \left\{ \Phi_l^{m(j)}, \Gamma_l^{m'(k)*} \right\} \text{cov} \left\{ \Gamma_l^{m(j)*}, \Phi_l^{m'(k)} \right\} \right), \end{aligned} \quad (\text{C.11})$$

and it is trivial to generalize (B.17) to show that

$$\begin{aligned} \text{cov} \left\{ \Phi_l^{m(j)}, \Gamma_l^{m'(k)*} \right\} &= \left\langle \Phi_l^{m(j)} \Gamma_l^{m'(k)*} \right\rangle = \sum_{l_1 m_1}^L h_{l_1}^{m_1(j)} \sum_{l_3 m_3}^L h_{l_3}^{m_3(k)*} \sum_{l_2 m_2}^{\infty} S_{fg}(l_2) \\ &\quad \times \frac{\sqrt{(2l_1+1)(2l_3+1)}}{2l+1} C_{l_1 0 l_2 0}^{l 0} C_{l_3 0 l_2 0}^{l 0} C_{l_1 m_1 l_2 m_2}^{l m} C_{l_3 m_3 l_2 m_2}^{l m'}. \end{aligned} \quad (\text{C.12})$$



For computational purposes, we note that the sums over  $l_2$  and  $m_2$  are considerably restricted as a result of the selection rules (B.10)–(B.15), which imply

$$m_2 = m - m_1, \quad (\text{C.13})$$

$$m_2 = m' - m_3, \quad (\text{C.14})$$

$$m - m_1 = m' - m_3, \quad (\text{C.15})$$

$$l_1 + l_2 + l = \text{even}, \quad (\text{C.16})$$

$$l_3 + l_2 + l = \text{even}, \quad (\text{C.17})$$

$$l_1 + l_3 = \text{even}. \quad (\text{C.18})$$

The expression for the variance is somewhat simplified when only one field,  $f$ , is considered. Noting that

$$\langle \Phi_l^{m(j)*} \Phi_l^{m'(k)*} \rangle = \langle \Phi_l^{m(j)} \Phi_l^{m'(k)} \rangle^*, \quad (\text{C.19})$$

$$\langle \Phi_l^{m(j)*} \Phi_l^{m'(k)} \rangle = \langle \Phi_l^{m(j)} \Phi_l^{m'(k)*} \rangle^*, \quad (\text{C.20})$$

the variance of the multitaper spectral estimate can be written as

$$\text{var} \left\{ S_{\Phi\Phi}^{(ml)}(l) \right\} = \sum_{j=1}^K \sum_{k=1}^K a_j F_{jk} a_k, \quad (\text{C.21})$$

where  $\mathbf{F}$  implicitly depends upon  $l$  and  $S_{ff}$  and is given by

$$F_{jk} = \sum_{m=-l}^l \sum_{m'=-l}^l \left( \left| \langle \Phi_l^{m(j)} \Phi_l^{m'(k)*} \rangle \right|^2 + \left| \langle \Phi_l^{m(j)} \Phi_l^{m'(k)} \rangle \right|^2 \right). \quad (\text{C.22})$$

Using the negative angular order symmetry relations (A.7) combined with the fact that the above sums are performed over all values of  $m$ ,  $\mathbf{F}$  further simplifies to

$$F_{jk} = 2 \sum_{m=-l}^l \sum_{m'=-l}^l \left| \langle \Phi_l^{m(j)} \Phi_l^{m'(k)*} \rangle \right|^2. \quad (\text{C.23})$$

The covariance matrix is naturally symmetric.

Finally, we note that (C.12), and hence the computation of  $\mathbf{F}$ , can be simplified for the case where the windows are solutions of the spherical-cap concentration problem [21]. For this situation, each window  $j$  and  $k$  has nonzero *real* spherical harmonic coefficients only for a single angular order  $m_j$  and  $m_k$ , respectively. When the windows are expressed in complex form, this implies that the only nonzero coefficients are for  $m_1 = \pm m_j$ ,  $m_3 = \pm m_k$ ,  $l_1 \geq |m_j|$  and  $l_3 \geq |m_k|$ .

## D. Correlation of Multitaper Spectral Estimates

It was shown in Appendix B and (2.19) that the expectation of a multitaper spectral estimate at degree  $l$  depends upon the global power spectrum within the degree range  $l \pm L$ , where  $L$  is the bandwidth of the localization windows. It is thus natural to expect that multitaper

spectral estimates separated by less than  $2L$  degrees will be partially correlated. Following the methodology presented in Appendix C, this correlation is quantified by the covariance of the two multitaper spectral estimates, which can be shown to equal

$$\text{cov} \left\{ S_{\Phi\Gamma}^{(ml)}(l), S_{\Phi\Gamma}^{(m'l')}(l') \right\} = \sum_{i=1}^K \sum_{j=1}^K a_i F_{ij}^{ll'} a_j, \quad (\text{D.1})$$

where

$$F_{ij}^{ll'} = \sum_{m=-l}^l \sum_{m'=-l'}^{l'} \left( \text{cov} \left\{ \Phi_l^{m(i)}, \Phi_{l'}^{m'(j)} \right\} \text{cov} \left\{ \Gamma_l^{m(i)*}, \Gamma_{l'}^{m'(j)*} \right\} \right. \\ \left. + \text{cov} \left\{ \Phi_l^{m(i)}, \Gamma_{l'}^{m'(j)*} \right\} \text{cov} \left\{ \Gamma_l^{m(i)*}, \Phi_{l'}^{m'(j)} \right\} \right), \quad (\text{D.2})$$

and

$$\text{cov} \left\{ \Phi_l^{m(i)}, \Gamma_{l'}^{m'(j)*} \right\} = \left\langle \Phi_l^{m(i)} \Gamma_{l'}^{m'(j)*} \right\rangle = \sum_{l_1 m_1}^L h_{l_1}^{m_1(i)} \sum_{l_3 m_3}^L h_{l_3}^{m_3(j)*} \sum_{l_2 m_2}^{\infty} S_{fg}(l_2) \\ \times \sqrt{\frac{(2l_1+1)(2l_3+1)}{(2l+1)(2l'+1)}} C_{l_1 0 l_2 0}^{l 0} C_{l_3 0 l_2 0}^{l' 0} C_{l_1 m_1 l_2 m_2}^{l m} C_{l_3 m_3 l_2 m_2}^{l' m'}. \quad (\text{D.3})$$

If only cross-power spectra of a single function are being considered, the matrix  $\mathbf{F}$  can be considerably simplified to

$$F_{ij}^{ll'} = 2 \sum_{m=-l}^l \sum_{m'=-l'}^{l'} \left| \left\langle \Phi_l^{m(i)} \Phi_{l'}^{m'(j)*} \right\rangle \right|^2. \quad (\text{D.4})$$

## Acknowledgements

We thank two anonymous reviewers for comments that helped clarify portions of this manuscript. Financial support for this work has been provided by the U. S. National Science Foundation under Grants EAR-0710860 and EAR-0105387 awarded to FJS at Princeton University, and by U. K. Natural Environmental Research Council New Investigator Award NE/D521449/1 and Nuffield Foundation Grant for Newly Appointed Lecturers NAL/01087/G to FJS at University College London. Software for performing the computations in this article can be found on the authors' web sites. This is IPGP contribution 2236.

## References

- [1] Belleguic, V., Lognonné, P., and Wieczorek, M. A. (2005). Constraints on the Martian lithosphere from gravity and topography data, *J. Geophys. Res.* **110**.
- [2] Blakely, R. J. (1995). *Potential Theory in Gravity and Magnetic Applications*, Cambridge University Press, New York.
- [3] Dahlen, F. A. and Tromp, J. (1998). *Theoretical Global Seismology*, Princeton University Press, Princeton, NJ.

- [4] Dahlen, F. A. and Simons, F. J. (2007). Spectral estimation on a sphere in geophysics and cosmology, *Geophys. J. Int.*, submitted manuscript.
- [5] Driscoll, J. R. and Healy, D. M. (1994). Computing Fourier transforms and convolutions on the 2-sphere, *Adv. Appl. Math.* **15**, 202–250.
- [6] Grünbaum, F. A., Longhi, L., and Perlstadt, M. (1982). Differential operators commuting with finite convolution integral operators: Some non-abelian examples, *SIAM, J. Appl. Math.* **42**, 941–955.
- [7] Han, S.-C. and Simons, F. J. (2007). Spatiospectral localization of global geopotential fields from GRACE reveals the coseismic gravity change due to the 2004 Sumatra-Andaman earthquake, *J. Geophys. Res.*, in press.
- [8] Hansen, F. K., Górski, K. M., and Hivon, E. (2002). Gabor transforms on the sphere with applications to CMB power spectrum estimation, *Mon. Not. R. Astron. Soc.* **336**, 1304–1328.
- [9] Hinshaw, G., Nolta, M. R., Bennett, C. L., Bean, R., Doré, O., Greason, M. R., Halpern, M., Hill, R. S., Jarosik, N., Kogut, A., Komatsu, E., Limon, M., Odegard, N., Meyer, S. S., Page, L., Peiris, H. V., Spergel, D. N., Tucker, G. S., Verde, L., Weiland, J. L., Wollack, E., and Wright, E. L. (2006). Three-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Temperature analysis, arXiv, astro-ph/0603451.
- [10] Hivon, E., Górski, K. M., Netterfield, C. B., Crill, B. P., Prunet, S., and 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.
- [11] Holmes, S. A. and Featherstone, W. E. (2002). A unified approach to the Clenshaw summation and the recursive computation of very high degree and order normalised associated Legendre functions, *J. Geod.* **76**, 279–299.
- [12] Kaula, W. M. (1967). Theory of statistical analysis of data distributed over a sphere, *Rev. Geophys.* **5**, 83–107.
- [13] Lawson, C. L. and Hanson, R. J. (1995). Solving least squares problems, *Classics Appl. Math. SIAM* **15**.
- [14] Luscombe, J. J. and Luban, M. (1998). Simplified recursive algorithm for Wigner  $3j$  and  $6j$  symbols, *Phys. Rev. E* **57**, 7274–7277.
- [15] McCoy, E. J., Walden, A. T., and Percival, D. B. (1998). Multitaper spectral estimation of power law processes, *IEEE Trans. Signal Proc.* **46**, 655–688.
- [16] McGovern, P. J., Solomon, S. C., Smith, D. E., Zuber, M. T., Simons, M., Wieczorek, M. A., Phillips, R. J., Neumann, G. A., Aharonson, O., and Head, J. W. (2002). Localized gravity/topography admittance and correlation spectra on Mars: Implications for regional and global evolution, *J. Geophys. Res.* **107**, 5136.
- [17] Peebles, P. J. E. (1973). Statistical analysis of catalogs of extragalactic objects, I. Theory, *Astroph. J.* **185**, 413–440.
- [18] Percival, D. B. and Walden, A. T. (1993). *Spectral Analysis for Physical Applications, Multitaper and Conventional Univariate Techniques*, Cambridge University Press.
- [19] Schulten, K. and 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.
- [20] Simons, F. J. and Dahlen, F. A. (2006). Spherical Slepian functions and the polar gap in geodesy, *Geophys. J. Int.* **166**, 1039–1061.
- [21] Simons, F. J., Dahlen, F. A., and Wieczorek, M. A. (2006). Spatiospectral localization on the sphere, *SIAM Rev.* **48**, 504–536.
- [22] Simons, M., Solomon, S. C., and Hager, B. H. (1997). Localization of gravity and topography: Constraints on the tectonics and mantle dynamics of Venus, *Geophys. J. Int.* **131**, 24–44.
- [23] Slepian, D. and Pollak, H. O. (1960). Prolate spheroidal wave functions, Fourier analysis and uncertainty — I, *Bell Syst. Tech. J.* **40**, 43–63.
- [24] Slepian, D. (1983). Some comments on Fourier-analysis, uncertainty and modeling, *SIAM Rev.* **25**, 379–393.
- [25] Spergel, D. N., Bean, R., Doré, O., Nolta, M. R., Bennett, C. L., Hinshaw, G., Jarosik, N., Komatsu, E., Page, L., Peiris, H. V., Verde, L., Barnes, C., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L. (2006). Wilkinson Microwave Anisotropy Probe (WMAP) three year results: Implications for cosmology, arXiv, astro-ph/0603449.
- [26] Thomson, D. J. (1982). Spectrum estimation and harmonic analysis, *Proc. IEEE* **70**, 1055–1096.
- [27] Varshalovich, D. A., Moskalev, A. N., and Khersonskii, V. K. (1988). Quantum theory of angular momentum, *World Scientific*, Singapore.

- [28] Walden, A. T., McCoy, E. J., and Percival, D. B. (1994). The variance of multitaper spectrum estimates for real Gaussian processes, *IEEE Trans. Signal Process.* **2**, 479–482.
- [29] Wieczorek, M. A. (2007). Gravity and topography of the terrestrial planets, in *Treatise on Geophysics, 10*, Eds. Spohn, T. and Schubert, T., Elsevier-Pergamon, Oxford, in press.
- [30] Wieczorek, M. A. (2007). Constraints on the composition of the Martian south polar cap from gravity and topography, *Icarus*, in press.
- [31] Wieczorek, M. A. and Simons, F. J. (2005). Localized spectral analysis on the sphere, *Geophys. J. Int.* **162**, 655–675.

---

Received September 27, 2006

Revision received May 29, 2007

Equipe d'Etudes Spatiales et Planétologie, Institut de Physique du Globe de Paris,  
94107 Saint Maur, France  
e-mail: wieczor@ipgp.jussieu.fr

Department of Earth Sciences, University College of London,  
Gower Street, London, WC1E 6BT, United Kingdom  
and

Department of Geosciences, Princeton University, Guyot Hall, Princeton, NJ 08544, USA  
e-mail: fjsimons@alum.mit.edu