## Spatiospectral localization of isostatic coherence anisotropy in Australia and its relation to seismic anisotropy: Implications for lithospheric deformation

Frederik J. Simons,<sup>1</sup> Rob D. van der Hilst,<sup>2</sup> and Maria T. Zuber Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Received 3 July 2001; revised 25 October 2002; accepted 28 January 2003; published 15 May 2003.

[1] We investigate the two-dimensional (2-D) nature of the coherence between Bouguer gravity anomalies and topography on the Australian continent. The coherence function or isostatic response is commonly assumed to be isotropic. However, the fossilized strain field recorded by gravity anomalies and their relation to topography is manifest in a degree of isostatic compensation or coherence which does depend on direction. We have developed a method that enables a robust and unbiased estimation of spatially, directionally, and wavelength-dependent coherence functions between two 2-D fields in a computationally efficient way. Our new multispectrogram method uses orthonormalized Hermite functions as data tapers, which are optimal for spectral localization of nonstationary, spatially dependent processes, and do not require solving an eigenvalue problem. We discuss the properties and advantages of this method with respect to other techniques. We identify regions on the continent marked by preferential directions of isostatic compensation in two wavelength regimes. With few exceptions, the short-wavelength coherence anisotropy is nearly perpendicular to the major trends of the suture zones between stable continental domains, supporting the geological observation that such zones are mechanically weak. Mechanical anisotropy reflects lithospheric strain accumulation, and its presence must be related to the deformational processes affecting the lithosphere integrated over time. Threedimensional models of seismic anisotropy obtained from surface wave inversions provide an independent estimate of the lithospheric fossil strain field, and simple models have been proposed to relate seismic anisotropy to continental deformation. We compare our measurements of mechanical anisotropy with our own model of the azimuthally anisotropic seismic wave speed structure of the Australian lithosphere. The correlation of isostatic anisotropy with directions of fast wave propagation gleaned from the azimuthal anisotropy of surface waves decays with depth. This may support claims that above  $\sim 200$  km, internally coherent deformation of the entire lithosphere is responsible for the anisotropy present in surface wave speeds or split shear waves. INDEX TERMS: 1236 Geodesy and Gravity: Rheology of the lithosphere and mantle (8160); 1234 Geodesy and Gravity: Regional and global gravity anomalies and Earth structure; 3299 Mathematical Geophysics: General or miscellaneous; 7218 Seismology: Lithosphere and upper mantle; 8180 Tectonophysics: Evolution of the Earth: Tomography; KEYWORDS: anisotropy, lithosphere, wavelets, coherence, isostasy, shear wave speed

**Citation:** Simons, F. J., R. D. van der Hilst, and M. T. Zuber, Spatiospectral localization of isostatic coherence anisotropy in Australia and its relation to seismic anisotropy: Implications for lithospheric deformation, *J. Geophys. Res.*, 108(B5), 2250, doi:10.1029/2001JB000704, 2003.

#### 1. Introduction

[2] The elastic properties of the Earth's lithosphere range from its local, instantaneous response to a seismic wave to

Copyright 2003 by the American Geophysical Union. 0148-0227/03/2001JB000704\$09.00

the long-term accommodation of topographic loading at various interfaces. The former is characterized by the elasticity tensor, which represents seismic wave speeds and their dependence on propagation and polarization direction, whereas the flexural or mechanical response is reflected in the admittance [*McKenzie and Bowin*, 1976] and coherence [*Forsyth*, 1985] between gravity anomalies and (surface and interface) topography. The isotropic and anisotropic seismic structure of continents [e.g., *Silver*, 1996; *Savage*, 1999] and the isotropic mechanical behavior of the lithosphere are well studied [e.g., *Burov and Diament*, 1995; *McKenzie and Fairhead*, 1997].

<sup>&</sup>lt;sup>1</sup>Now at Geosciences Department, Princeton University, Princeton, New Jersey, USA.

<sup>&</sup>lt;sup>2</sup>Now at Faculteit Aardwetenschappen, Universiteit Utrecht, Utrecht, Netherlands.

[3] Finding the relation between seismic and mechanical structure can shed light on the structure and deformation of the lithosphere [*Griot and Montagner*, 1998; *Chen and Özalaybey*, 1998; *Vauchez et al.*, 1998; *Meissner et al.*, 2002; *Simons and van der Hilst*, 2002]. One hypothesis holds that seismic anisotropy is predominantly caused by strain fossilized in the subcontinental mantle since the last major episode of tectonic activity [*Silver and Chan*, 1991] rather than current mantle deformation [*Vinnik et al.*, 1995]. If lithospheric deformation is internally coherent, the structural geologic trends observed at the surface will be predictors for the seismic anisotropy [*Silver and Chan*, 1988].

[4] The hypothesis of vertically coherent deformation was inspired by the comparison of shear wave splitting measurements to geologic deformation indicators measured at the surface. *Rümpker and Silver* [1998], *Saltzer et al.* [2000], and others have shown, however, that shear wave splitting gives an oversimplified view of continental anisotropy. Anisotropic surface wave tomography affords the vertical resolution lacking in shear wave splitting studies. That the seismic anisotropy of the Australian lithosphere is indeed complex was shown by *Debayle and Kennett* [2000] and *Simons et al.* [2002]. On the basis of the pattern and the correlation length scale of azimuthal anisotropy derived from surface wave modeling, both groups argue that a regime of "fossil" deformation overlies a zone characterized by a predominant influence of present-day mantle deformation.

[5] The objective of this paper is to quantify how good a predictor of the tectonic fossil strain field the observed seismic anisotropy in the lithosphere is and to which depth the correspondence between them might be valid. Rather than relying on surface strain indicators, we infer the dominant fossil strain field of the lithosphere from the orientation of gravity anomalies relative to topography. We will do this in the spectral domain by extracting the azimuthal variation of the gravity-topography coherence function. As a concept, such "mechanical anisotropy" is well documented in the geological literature [Vauchez et al., 1998; Tommasi and Vauchez, 2001], but although it has been the subject of some investigations by spectral methods [Stephenson and Beaumont, 1980; Stephenson and Lambeck, 1985a; Lowry and Smith, 1995; Escartin and Lin, 1998], it has only recently been studied with fully two-dimensional (2-D) spectral approaches [Simons et al., 2000].

[6] The robust and unbiased estimation of coherence functions (depending on space, direction, and wavelength) between two 2-D fields requires specialized techniques. The method we present here is aimed at satisfying three design criteria: first, the correct retrieval of coherence in the spectral domain (the wavelength dependence); second, the localization of the coherence in the space domain (the spatial dependence); and third, ensuring an isotropic response of the spectral estimator in all azimuths (to get an unbiased estimate of its directional dependence). Traditional approaches to coherence estimation fail to reach all three goals. Mirrored periodogram methods and modified periodograms using a single window yield less stable estimates than multitaper methods [see, e.g., Percival and Walden, 1993]. Phase-averaging schemes such as isotropic wave number binning techniques [Bechtel et al., 1987] obliterate the directionality in the signal. Both problems are overcome by the Thomson multitaper method [Thomson, 1982]

extended to two-dimensional fields [*Liu and van Veen*, 1992; *Hanssen*, 1997]: the average of several periodograms calculated on data windowed with different Slepian windows [*Slepian*, 1978] proves to be a good estimator of the coherence [*Scheirer et al.*, 1995; *McKenzie and Fairhead*, 1997] and its anisotropy [*Simons et al.*, 2000] for lithospheric loading problems.

[7] To enable a local comparison of gravity-topography coherence anisotropy with predictions from seismic anisotropy, we require a method that has the ability to map out the spatial variation of the coherence function in great detail. Lateral and directional resolution come naturally to seismic surface waves, since they sample the Earth in a spatially well-defined manner. Spectral properties such as coherence, on the other hand, are by definition not localized in space: pure frequencies have infinite extent. Keeping Thomson's method, some spatial resolution can be achieved by assuming local stationarity, and analyzing discrete portions of the data set. In addition to being ad hoc, however, the assumption of local stationarity can lead to strongly erroneous results [Frazer and Boashash, 1994; Bayram and Baraniuk, 1996]. More advanced time-frequency [Flandrin, 1998] and time-scale (wavelet) methods [Mallat, 1998] have been designed to optimize resolution both in the spatial and the spectral domain. Such techniques are ideally suited to analyze processes with temporally or spatially varying spectral properties.

[8] Wavelet methods are used pervasively for the univariate time-scale characterization and representation of geophysical fields [Kumar and Foufoula-Georgiou, 1997; Torrence and Compo, 1998; Bergeron et al., 2000], including their anisotropy [Kumar, 1995], but the calculation of multivariate coherence functions by wavelet transforms is hampered by the lack of suitable smoothing schemes [Liu, 1994; Meredith, 1999]. A handful of multiwavelet methods does yield adequate multivariate time and frequency resolution, but these techniques are restricted to 1-D signals (time series) [Daubechies and Paul, 1988; Lilly and Park, 1995; Olhede and Walden, 2002]. Simons et al. [1997] formulated a spherical-wavelet-based spatial localization of the admittance between gravity and topography, but their method was not designed to measure its anisotropy (though there is no a priori reason to preclude this method from being modified to do so).

[9] We compare the capability for spatiospectral localization of a space-scale multiwavelet and a space-frequency multiwindow method by an analysis of the concentration domains in the space-frequency plane that are attained by the wavelets or windows they employ. The method we prefer uses a set of orthonormal Hermite windowing functions and is capable of detecting both the anisotropy and the spatial variations of the spectral coherence function in realistic settings. In our previous study [Simons et al., 2000], optimal spectral localization of the coherence function was achieved by using Slepian data windows and the method due to Thomson [1982], which assumes spatial stationarity. Bayram and Baraniuk [2001] and Çakrak and Loughlin [2001] demonstrated that orthonormal Hermite polynomials used as data windows offer spatial and spectral localization simultaneously. The additional localization in the spatial domain, which enables the study of nonstationary processes, comes at the expense of some spectral resolution

overall [*Parks and Shenoy*, 1990], but the benefits are twofold. First, the spectral resolution is completely isotropic by construction: the spectral properties of the windowing tapers are identical in all azimuths and are hence ideally suited for the detection of possibly anisotropic components in the coherence. Second, there is a significant reduction in computational expense with respect to the Thomson-Slepian method because the need to solve an eigenvalue problem of the size of the data can be avoided.

[10] We study the mechanical anisotropy of Australia by calculating coherence functions between topography and Bouguer gravity anomalies. We discuss the mechanical properties of the Australian lithosphere in comparison with the fast directions obtained from high-resolution seismic surface wave tomography with azimuthal anisotropy [*Simons et al.*, 2002]. Identifying directions that have accumulated more than the isotropic average of gravitational anomalies for a given amount of topography, we test the simple hypothesis that seismic anisotropy is reflected in the anisotropy of fossil strain.

#### 2. Methodology

[11] Sections 2–4 of this paper are concerned with the methodology of nonstationary coherence estimation. We illustrate the performance of a method with multiple Hermite windows for the univariate as well as multivariate (cross-)spectral characterization of 1-D and 2-D processes and, in particular, for a geologically relevant loading scenario. Sections 5–7 deal with the practical application to the study of the spatial and azimuthal variability of the isostatic response of the Australian continent, its comparison to seismic anisotropy, and the implications of our findings for continental deformation.

[12] Appendix A summarizes how the mechanical properties of the lithosphere are reflected in the coherence between Bouguer gravity anomalies and topography. For two nonstationary random processes {*X*} (gravity) and {*Y*} (topography), defined on **r** in the spatial domain and on **k** in the Fourier domain, the coherence-square function relating both fields,  $\gamma_{XY}^2$ , is defined as the ratio of their cross-spectral density,  $S_{XY}$  normalized by the individual power spectral densities,  $S_{XX}$  and  $S_{YY}$  [Bendat and Piersol, 2000]:

$$\gamma_{XY}^{2}(\mathbf{r}, \mathbf{k}) = \frac{|S_{XY}(\mathbf{r}, \mathbf{k})|^{2}}{S_{XX}(\mathbf{r}, \mathbf{k})S_{YY}(\mathbf{r}, \mathbf{k})}$$
$$= \frac{|E\{\tilde{X}(\mathbf{r}, \mathbf{k})\tilde{Y}^{*}(\mathbf{r}, \mathbf{k})\}|^{2}}{E\{\tilde{X}(\mathbf{r}, \mathbf{k})\tilde{X}^{*}(\mathbf{r}, \mathbf{k})\}E\{\tilde{Y}(\mathbf{r}, \mathbf{k})\tilde{Y}^{*}(\mathbf{r}, \mathbf{k})\}}.$$
(1)

Here, *E* denotes an expectation operator, tildes refer to the Fourier-transformed signal, and the asterisk refers to the complex conjugate. The periodogram  $\tilde{X}\tilde{X}^*$  is a direct spectral estimator of *X*, although not a particularly accurate one [*Percival and Walden*, 1993].

[13] The coherence is a measure of the consistency of the phase relationship between both processes. Therefore it cannot simply be estimated from the ratios of the periodograms, as the periodogram does not record phase information. The averaging operators in equation (1) are required to avoid that  $\gamma_{XY}^2 = 1$  everywhere [*Bendat and Piersol*, 1993]. Multiwindow or multiwavelet methods (see Appendix B1)

provide a way of smoothing across different, approximately uncorrelated estimates to obtain a stable average estimate, rather than smoothing within the spectral or spatial domain as required by single-window methods [Simons et al., 2000]. Each individual direct spectral estimate is formed as a spectrogram or a scalogram (for definitions, see Appendix B1) obtained with a different window or wavelet. Windows and wavelets can be designed to afford maximum resolution while minimizing bias, variance, and spectral leakage [Percival and Walden, 1993]. Resolution is commonly defined by the properties of an operator that concentrates the estimate in a well-defined domain on the frequency axis, the time-frequency plane (in one dimension) or the space containing spatial dimensions and spatial frequencies (in two dimensions) (see Appendices B2 and B3). The multiple eigenfunctions of such operators are the windowing functions or wavelets that are used to estimate spectrum and coherence with multiwindow or multiwavelet methods.

[14] *Simons et al.* [2000] focused on the analysis of stationary processes by a multitaper method with Slepian functions (see also Appendix B2). Here we focus on two possible extensions to the nonstationary case (see also Appendix B3). We compare the space-scale concentration achieved by Slepian wavelets (section 3.1) to the concentration in the space-frequency domain of orthonormal Hermite spectral windowing functions (section 3.2).

# Spatiospectral Localization Properties Slepian Wavelets

[15] Lilly and Park [1995] extended Slepian's frequency concentration criterion to the time-scale domain and found wavelets that are eigenfunctions of a time-frequency concentration operator (see Appendix B3). To compare the properties of the Slepian functions (prolate spheroidal wave functions or pswf) for stationary analysis [see Simons et al., 2000, Figure 2] with the Slepian wavelets for nonstationary analysis, we calculated six Slepian wavelets (N samples spaced  $\Delta t$  apart) with a central frequency  $f_c = 3/N/\Delta t$  and a half bandwidth  $W/2 = 2.5/N/\Delta t$ . Figure 1 shows the wavelets in the time domain (Figures 1a-1c) and their magnitude response in the frequency domain (Figures 1d-1f). Six wavelets are plotted, grouped in pairs (solid and dashed lines) with similar magnitude response. The frequency axis is in terms of the smallest resolvable frequency  $1/N/\Delta t$  [Kav and Marple, 1981]. For every wavelet of the set the spectral windows are concentrated in two symmetric bands of width 5, centered at 3 on the duration-times-frequency scale. The frequency concentration properties of the Slepian wavelets are fairly similar to those of the pswf, and their multiplicity enables the averaging required for coherence analysis of time-varying 1-D signals. The eigenvalues, a measure of the concentration of energy in the central lobe of the spectral windows, are extremely close to unity [Lilly and Park, 1995]. However, unlike Slepian functions, Slepian wavelets do not have 2-D analogues.

#### 3.2. Hermite Windows

[16] Slepian wavelets of length  $T = N\Delta t$  and bandwidth W are localized in a rectangular domain  $[-T/2, T/2] \times [-W/2, W/2]$  [Lilly and Park, 1995] of the time-frequency



**Figure 1.** Six prolate spheroidal (Slepian) wavelets [*Lilly* and Park, 1995] in the (a–c) time and (d–f) frequency domain, grouped in three pairs (solid and dashed lines) with similar frequency response. The  $\Psi_j(t)$ ,  $j = 0 \rightarrow 5$ , are centered on a frequency equal to 3 times the fundamental frequency,  $f_c = 3/N/\Delta t$ , and have a half bandwidth  $W/2 = 2.5/N/\Delta t$ . Their concentration properties are comparable to those of the prolate spheroidal wave functions of *Slepian* [1978], but the band-pass magnitude response enables time-frequency analysis.

plane, but such a rectangular tiling of the (t, f) plane is not optimal in the sense that it does not provide maximal simultaneous concentration of time and frequency [*Parks* and Shenoy, 1990]. Circular (t, f) concentration domains allow optimal resolution in time and frequency to be achieved simultaneously, without one being at the expense of the other. In other words, the estimated spectrum has isotropic distortion in time and frequency (see Appendix B1). If the circular symmetry applies, instead of to time and frequency, to space and spatial frequency (both horizontal and vertical), the resulting spherical concentration domains allow the estimation of spatially localized directional spectral properties without anisotropic bias. The spectral resolution is identical in all possible azimuths, and no spurious anisotropy will be introduced.

[17] The eigenfunctions of operators concentrating in a disk-shaped time-frequency domain  $t^2 + (2\pi f)^2 \leq R^2$  are orthonormal Hermite functions [Daubechies, 1988; Flandrin, 1988]. Their analytic form is given by [Olhede and Walden, 2002]

$$h_j(t) = \frac{H_j(t)e^{-t^2/2}}{\pi^{1/4}\sqrt{2^j j!}},$$
(2)

i.e., they are Hermite polynomials,  $H_j$ , modulated by a Gaussian function. Starting from  $H_0(t) = 1$  and  $H_1(t) = 2t$ , Hermite polynomials can be calculated using the recurrence relation

$$H_{n+1}(t) = 2tH_n(t) - 2nH_{n-1}(t).$$
(3)

The eigenvalues are dependent on the radius of the concentration region, R, and are given by

$$\lambda_j(R) = \frac{1}{j!} \gamma \left( \frac{R^2}{2, j+1} \right),\tag{4}$$

where  $\gamma$  denotes the incomplete gamma function.

[18] The remarkable property of the Hermite functions and their eigenvalues is that the eigenfunctions of the concentration operator do not depend on the width *R* of the domain; the *R* dependence is contained in the eigenvalues. This makes the Hermite method computationally very fast. To design a set of windows with a particular timefrequency resolution, it is sufficient to multiply the  $h_j$  with the appropriate  $\lambda_j(R)$  to make their average concentration lie within a domain of radius *R*. This is not the case for the pswf or Slepian wavelets, where an eigenvalue problem needs to be solved for every choice of bandwidth and signal length or central frequency and for every different data length.

[19] In Figure 2a, five orthonormal Hermite functions are plotted (again compare with the pswf shown by *Simons et al.* [2000, Figure 2]). We let  $t: -5 \rightarrow +5$  in equation (2). The concentration region *R*, expressed in the same units, then corresponds to a time concentration within *R*/10 multiplied by the actual physical length of the window. The spectral windows  $\mathcal{H}(f; R)$  are plotted in Figure 2b for varying values of *R*. They are calculated as

$$\mathcal{H}(f;R) = \frac{1}{J} \sum_{j=0}^{J-1} \lambda_j(R) \left| \int_{-\infty}^{\infty} h_j(\tau) e^{-i2\pi f \tau} d\tau \right|^2.$$
(5)

On the dimensionless duration-times-frequency scale, the corresponding frequency concentration (around 10 dB of bias reduction) lies within the gray shaded regions. The eigenvalues  $\lambda_j$  for the three choices of *R* are plotted in



**Figure 2.** Hermite functions and their eigenvalues. (a) Orthonormal  $h_j(t)$ ,  $j = 0 \rightarrow 4$ . (b) Frequency concentration, spectral windows  $\mathcal{H}(f; R)$  (see equation (5)). (c) Eigenvalues  $\lambda_R(j)$  for three different radii of concentration. (d) Time concentration, energy U(t; R) (see equation (6)).

Figure 2c. The temporal and spectral resolution of the multispectrogram estimates is controlled by the length and overlap of the windows, and by the value of the concentration radius R. As R becomes larger and resolution degrades, J grows to reduce the estimation variance (see Appendix B4). The decreasing number of significantly contributing window functions with increased resolution (smaller R) is a manifestation of the duality of resolution and variance. In practice, we take  $J = R^2$  tapers [Daubechies, 1988].

[20] As a measure of the concentration of the windows in the time domain we define their energy as [Walden, 1990a]

$$U(t;R) = \frac{1}{J} \sum_{j=0}^{J-1} \lambda_j(R) h_j^2(t),$$
(6)

which we plotted in Figure 2d. We have normalized the time axis by the entire data length, so the energy of the windows is contained within  $0.5 \pm R/10$  (gray shaded regions). The even extraction of energy from the signal is achieved by the calculation of the spectrogram at specific time intervals corresponding to an effective overlap of the averaging window.

[21] Besides the ease with which they can be calculated and the elegance of their construction, perhaps the most attractive characteristic of the Hermite functions is the fact that the eigenfunctions of n-dimensional circularly symmetric concentration operators are outer products of the same Hermite functions [Daubechies, 1988]. The extension from time-frequency to multidimensional space-frequency analysis therefore becomes practical. Multidimensional extensions to space-scale (wavelet) approaches are conceptually more difficult [Foufoula-Georgiou and Kumar, 1994]. Concentration operator formalisms have been derived for the space-scale phase plane, but rather than the circularly symmetric concentration regions of the time-frequency case, the associated wavelets have complex, nonsymmetrical concentration domains [Daubechies and Paul, 1988; Olhede and Walden, 2002].

#### 3.3. Comparison of Concentration Domains

[22] The best way of characterizing and comparing the effect of windowing or wavelet functions for space-frequency or space-scale analysis is not by their average periodograms or energy functions (as in Figures 1 and 2) but by their average Wigner-Ville transform (WV; see Appendix B1). In Figure 3, we have plotted the alias-free WV [Jeong and Williams, 1992] of two sets of Slepian wavelets (Figures 3a and 3b) and of two sets of Hermite windows (Figures 3c and 3d). We give examples of Slepian wavelet concentrations at one center frequency for two different bandwidths and of Hermite concentration regions for two radii R. The rectangular nature of the Slepian wavelet concentration domain and the disk-shaped concentration domains of the Hermite functions are well captured by their WV. The symmetry and smoothness of the Hermite function kernels make them attractive for their use in timefrequency coherence analysis. For multidimensional properties, this symmetry applies to each dimension: for spatial data, no one direction is favored. This is particularly important for the estimation of anisotropic processes.

[23] The high concentration and symmetric qualities of the multiwindow Hermite technique are responsible for its

superior performance compared to traditional (i.e., singlewindow) spectrogram methods (an advantage it shares with the pswf) [Bronez, 1992] and to methods of nonstationary estimation which employ the pswf in a sliding-window fashion [Frazer and Boashash, 1994; Bayram and Baraniuk, 1996; Mellors et al., 1998]. Section 4 illustrates this performance advantage with physical examples.

#### **Examples of Performance** 4.

#### 4.1. Univariate 1-D: Power Spectra of Time Series

[24] We compare the Slepian multiwavelet approach to the Hermite multiwindow technique for the characterization of the synthetic signal shown in Figure 4a [see Lilly and Park, 1995, Figure 5]. The signal consists of a sinemodulated linear chirp followed by a spike. In Figure 4b, the Slepian multiwavelet scalogram is plotted (scale was converted to frequency), calculated with wavelets of center frequency  $f_c = 7/N/\Delta t$  and half width  $W/2 = 3/N/\Delta t$ . The number of wavelets used per scale was smaller or equal to the Shannon number of 2W. This determined the wavelet length and the maximum resolvable frequency. In Figure 4c, we have plotted the Hermite multispectrogram, with R = 1. The window length was 0.13 in units of normalized time, and the values were calculated at time steps corresponding to an overlap of the windows of 99%.

[25] An important difference between time-scale and time-frequency methods is that the time resolution of wavelets is inversely proportional to the frequencies at which the signal is analyzed: the high-frequency components are analyzed with much shorter wavelets than the low-frequency portions of the signal. This is evident in the typical pyramidal form generated by discontinuities in the signal. It also explains the particular efficiency of the wavelet transform for signal representation with a small number of coefficients [Strang and Nguyen, 1997]. For analysis rather

Figure 3. Concentration of Slepian wavelets and Hermite functions in the time-frequency plane by their normalized average Wigner-Ville energy distribution. Contours are drawn at [0.1:0.2:1]. (a) Slepian wavelets with  $f_c = 6$  and W/2 = 2 and (b) with  $f_c = 6$  and W/2 = 3 in dimensionless units normalized by the fundamental frequency. Concentration of Hermite functions in the time-frequency plane for radii (c) R = 2 and (d) R = 4.





**Figure 4.** Comparison of the Hermite multiple-spectrogram method with the Slepian multiwavelet method. (a) Synthetic signal [after *Lilly and Park*, 1995]. (b) Power spectrum based on the Slepian multiwavelet transform. Slepian wavelets with  $f_c = 7/N/\Delta t$  and  $W/2 = 3/N/\Delta t$  were used. (c) Power spectrum from the multispectrogram method, the concentration region R = 1, window length  $L = 32\Delta t = 0.13 N\Delta t$ , overlap 99%, number of frequency bins nfft = 256. The thin solid lines in Figures 4b and 4c correspond to the instantaneous frequency of the input chirp shown in Figure 4a.

than representation of signals, on the other hand, we prefer the results generated by the Hermite method.

#### 4.2. Bivariate 1-D: Coherence Between Time Series

[26] We take the synthetic example proposed by *Santoso* et al. [1997] to show that the Hermite function approach is capable of retrieving both the magnitude and the phase of the coherence between both signals; we can distinguish variations in time as well as frequency. Santoso et al. [1997] were unable to provide frequency resolution. The synthetic signals, plotted in Figures 5a and 5b, consist of two time segments with a sinusoid at 1/4 and 1/2 of the Nyquist frequency (Figure 5a), and another at 1/4 and 1/1 of the Nyquist (Figure 5b),  $-\pi/2$  radians out of phase with the first. Uniform random noise (with a signal-to-noise ratio of 5 dB) was added. The localization properties of the Hermite functions are excellent, both in time and in frequency. This is evident from the individual power spectral estimations of both signals, plotted in Figures 5c and 5d, and in the strong peak of unit coherence in Figure 5e. Where the coherence is unity the measurement of its phase shown in Figure 5f correctly yields the input phase difference between both segments.

#### 4.3. Bivariate 2-D: Coherence Between 2-D Fields

[27] For a 2-D test we follow the method of *Lowry and Smith* [1994] to create pairs of real synthetic fields with a known (an)isotropic coherence between them. Four pairs of coherent fields were joined together, forming a spatial input pattern of coherence shown in Figure 6. The Hermite multispectrogram method was applied on spatially overlapping tiles of the data. Figure 6 illustrates how spatially varying anisotropic coherence functions between two 2-D fields may be retrieved with confidence.

#### 4.4. A Geologically Relevant Example

[28] Finally, we show a realistic example of a lithosphere loaded by two statistically independent processes at the surface and at one subsurface interface. We assume that isostatic compensation is described by the effective elastic thickness and operates in an directionally isotropic way. In this classic loading scenario the equilibrium topography, Bouguer anomaly, and predicted coherence function can be predicted by the analytic expressions first derived by *Forsyth* [1985]. We have generalized these expressions to a multilayered case in Appendix A.

[29] The purpose of our example is twofold. First, we show how a coherence function predicted by the loading model is observed by the Hermite method for various values of the resolution-variance parameter R (see section 3.2). Second, we show that anisotropy in either gravity or topography does not introduce spurious anisotropy in the isostatic response. By construction, the coherence function represents the average cross-spectral properties of gravity and topography,



**Figure 5.** Coherence estimation with the Hermite method. (a-b) Synthetic signals X and Y [after Santoso et al., 1997] and (c-d) their time-varying power spectra,  $S_X$  and  $S_F$ Time-varying coherence-square function (e)  $\gamma^2$  and (f) its phase  $\phi$ . Where the input signals are coherent, between 0.125 and 0.450, the measured coherence equals unity and the input phase difference of  $-\pi/2$  is correctly retrieved. Concentration region R = 3, window length  $L = 256\Delta t =$ 0.12  $N\Delta t$ , overlap 99%, nfft = 128.



**Figure 6.** Retrieval of spatially varying anisotropic coherence functions. Two synthetic fields are related by a different coherence function in each of four spatial quadrants. (a) Spatial variation of the input coherence, in wave vector space, with the long wavelengths at the center of the tile. (b) Retrieval with Hermite multispectrogram method. The measured coherence functions (four in each quadrant) are centered on the location of the measurement, but measurement regions overlapped spatially by 25%. Inferred anisotropic directions (solid lines) are compared to the input directions (dashed lines). Concentration radius R = 3.

normalized by the individual power spectra of both fields, and hence no such artifacts are expected to occur.

[30] We made two synthetic data sets with power spectra identical to data observed in central Australia, but with randomly perturbed phase. The resulting fields have realistic, anisotropic power spectra. The random phase guarantees that both loading processes are statistically independent, as required for the coherence method to be valid [Forsyth, 1985]. We applied the topography shown in Figures 7a and 7b as loads to the surface and one interface, respectively, and added a third, unloaded, compensating density interface. The wavelength-dependent ratio of bottom-to-top loading [Forsyth, 1985] has an average of 0.38 and a standard deviation over all wave numbers of 0.17. The effective elastic thickness of the lithosphere is  $T_e = 25$ . The (isotropic) isostatic compensation leads to an equilibrium surface topography shown in Figure 7c, a Bouguer gravity anomaly in Figure 7d, and predicted coherence between the two in Figure 7e. Prior to analysis, we added Gaussian random noise to the equilibrium surface topography (with mean equal to 50% of the mean absolute value of the topography). Figures 7f-7h show how the predicted coherence is retrieved by analyzing the data shown in Figure 7c (topography) and Figure 7d (Bouguer anomaly) with the Hermite multiwindow approach, for different values of the resolution parameter R. As shown in section 3.2, increased values of R degrade the resolution of the estimate (the coherence peak is broadened), but at the same time, the variance of the estimate decreases. The variance of the coherence estimate is not explicitly shown in Figures 7f-7h, but its decrease is apparent from the decreasingly oscillatory nature of the observed coherence in regions where null coherence is predicted.

[31] Changing R affects the transition wavelength between high and low coherence, a measure often used in inversions for effective elastic thickness. The value of R (or NW in the Thomson-Slepian approach) combined with the physical size

of the data determine the effective bandwidth of the estimate (see section 3.1) [Walden et al., 1995]. The same dependence is implicitly present in single-window or mirrored periodogram techniques [Simons et al., 2000]. Hence no two estimates of  $T_e$  should be compared without a comparison of the resolution bandwidth. The difference in the (average, isotropic) transition wavelength between Figures 7e–7h is on the order of 100 km, which seriously jeopardizes attempts to attach absolute values of effective elastic thickness to similar measurements. For the purposes of analysis of anisotropy in the isostatic response, however, Figure 7h shows that our choice of R = 3 affords low-variance estimates of the coherence, without artificial introduction of anisotropy, even when the power spectra of the data analyzed are themselves anisotropic. Higher values of R were not required.

[32] In summary, the Hermite multispectrogram method allows for unbiased, high-resolution, low-variance coherence calculations in any number of dimensions. It compares favorably to traditional and Slepian multitaper spectrogram and multiwavelet scalogram methods of time-frequency analysis. The  $(\mathbf{r}, \mathbf{k})$  space is treated as a whole without sacrificing temporal or spatial resolution to frequency resolution. Resolution and variance are at the discretion of the analyst, who can choose the radius of the concentration region. The Hermite windows and their eigenvalues are easy to calculate recursively, and no matrix diagonalization is required.

#### 5. Gravity and Topography Data of Australia

[33] The continental gravity and topography data sets used in this study are the same as used by *Simons et al.* [2000]. Topography and bathymetry data (see Figure 8) are from the compilation by *Smith and Sandwell* [1997]; over the continent they are identical to the GTOPO30 data set [*Gesch et al.*, 1999]. Oceanic bathymetry was added to provide continuity with the continental topography. How-



**Figure 7.** Predicted and observed coherence for a realistic loading scenario. The model has three interfaces at depths  $\mathbf{z} = (0,15,35)$  km, density contrasts of  $\Delta \boldsymbol{\rho} = (2600, 300, 400)$  kg m<sup>-3</sup>, an effective elastic thickness  $T_e = 25$  km, Young's modulus  $E = 1.78 \times 10^{11}$  and Poisson's ratio  $\nu = 0.25$ . Loading topography is applied at the top two interfaces (Figures 7a and 7b), and the isotropic isostatic compensation produces an equilibrium topography, shown with noise added (see text), and Bouguer gravity anomaly (Figures 7c and 7d). The predicted coherence function, with the longest wave numbers at the center, is plotted in Figure 7e. Observed coherence, calculated with after tapering the data with Hermite windows, is shown in Figures 7f–7h for varying values of the resolution parameter *R*. The optimal value is R = 3, when the isotropic coherence is most faithfully retrieved.

ever, the intricacies involved in the generation of the bathymetry data from the integration of ship-sounding data with satellite gravity field measurements (in which an isotropic coherence function is implicitly assumed by *Smith and Sandwell* [1997]) prevent our coherence analyses from being meaningful over the purely oceanic areas.

[34] The continental Bouguer gravity anomaly map plotted in Figure 9 is from Geoscience Australia (formerly the Australian Geological Survey Organisation). A crustal density of 2670 kg m<sup>-3</sup> is assumed. A background gravity field obtained from a spherical harmonic expansion of satellitederived geoidal coefficients up to degree 10 was removed (for details, see *Simons et al.* [2000]). Over the oceans we computed Bouguer anomalies with a crustal thickness of 6 km using the method of *Parker* [1972]. Oceanic and continental data sets were adjusted to the same baseline. The data were projected onto a Cartesian grid with a sampling interval of ~5.5 km in both x and y. Figures 8 and 9 also show the location and size of the analyses boxes used to generate the results.

#### 6. Results: Coherence Analysis of Australia

[35] We applied the Hermite multispectrogram method to characterize the coherence between the data sets shown in Figures 8 and 9. The full extent of the data was subdivided into square boxes with a size of 720 km and an overlap of

30%. The data were analyzed using 81 different Hermite tapers (the outer products of 9 tapers in each dimension) and a concentration region of R = 3. We follow two approaches for extracting the directional dependence of the 2-D coherence function.

#### 6.1. Long-Wavelength Response

[36] The first method concentrates on the longest-wavelength parts of the coherence and is directly related to the traditional method of determining the elastic thickness, or  $T_e$ , of the continental plate [Watts, 2001]. The primary diagnostic to distinguish mechanically weak from strong plates (the "weak" direction being the one which has accumulated most of the deformation for a given amount of topography) is the transition wavelength at which the change occurs from isostatically compensated loads (and, thus, high coherence of the Bouguer anomaly with the topography) to uncompensated loads supported by the elastic strength of the plate (low or zero coherence) [Forsyth, 1985]. For every coherence estimate  $\gamma^2(\mathbf{k})$  we determine the wavelength  $\lambda_{1/2}(\theta)$  at which the coherence drops to half of its maximum (long-wavelength) value as a function of the azimuth  $\theta$  of the profile defined in  $(k_x, k_y)$  space. This transition wavelength  $\lambda_{1/2}(\theta)$  is well defined if the coherence falls off as predicted by simple compensation modeling of elastic plates (see Appendix A), and the function  $\lambda_{1/2}(\theta)$ should show a clear minimum if there is a dominant



**Figure 8.** Topography and bathymetry of Australia and its surrounding areas, in km. Data from *Smith and Sandwell* [1997]. Inverted triangles indicate center of analysis boxes. Box size and overlap are plotted top left. The boxes labeled a–c correspond to the regions analyzed in Figure 10.

direction of weakness in the plate. The interpretation with respect to standard curves for varying  $T_e$  can then be verified by comparing  $\gamma^2(k, \theta_s)$  (in the strongest direction  $\theta_s$ ) and  $\gamma^2(k, \theta_w)$  (in the weakest direction  $\theta_w$ ).

[37] In Figure 10, we give three examples of results obtained using this approach. From left to right, Figures 10a-10c consist of a  $720 \times 720$  km<sup>2</sup> tile of topography, h(x, y), the corresponding Bouguer gravity anomaly, B(x, y), and the coherence function,  $\gamma^2(\mathbf{k}) = \gamma^2(k_x, k_y)$ . The scale of the  $\gamma^2(\mathbf{k}) = \gamma^2(k_x, k_y)$  panel is linear in  $(k_x, k_y)$ , with the longest resolvable (also known as Rayleigh) wavelength,  $\lambda_R = 720$  km, in the center, and a wavelength of  $\pm 2\lambda_N = 20$ km, equal to twice the shortest resolvable (Nyquist) wavelength, on the sides. Next is the azimuth-dependent transition wavelength,  $\lambda_{1/2}(\theta)$ , where  $\lambda$  is in km and  $\theta$  in radians, measured anticlockwise from the equatorial line through the center of the  $\gamma^2(\mathbf{k}) = \gamma^2(k_x, k_y)$  panel. The maximum and minimum  $\lambda_{1/2}(\theta)$  define a "strong", or least deformed (solid line), and a "weak", or most deformed (dashed line) direction of the plate. Last, the coherence is shown as a function of the wave number, for the interpreted strongest ( $\gamma^2(k, \theta_s)$ , solid lines) and weakest ( $\gamma^2(k, \theta_w)$ , dashed lines) directions. Because the 2-D plots in  $(k_x, k_y)$  space are symmetric through the center, only half of the coherence needs to be plotted (here we use positive wave numbers only). Also plotted are the error bars on the coherence measurements (see Appendix B4). From the shape of the minimum of  $\lambda_{1/2}(\theta)$ , we can see that the direction corresponding to the minimum transition wavelength is not always unambiguous; a conservative estimate of its uncertainty can be as large as 45°. On the other hand, the plots of  $\gamma^2(k)$  for the strongest and the weakest direction show a significant difference in the coherence, and such coherence behavior will correspond to an interpretable difference in  $T_e$  of several km (for sample curves, see Figure 7b of *Forsyth* [1985]). We also note that the error bars were obtained for a unique azimuth without averaging in  $(k_x, k_y)$  space. They could be further reduced by considering the coherence in a small azimuthal range, as opposed to a single azimuth, but we have not found it necessary to do so.

[38] The noticable difference in transition wavelength can be interpreted as a directional dependence of the effective elastic thickness,  $T_e$ , but due to the complexity of the assumptions involved [e.g., Burov and Diament, 1995; Lowry and Smith, 1995; Banks et al., 2001], we prefer to show only the coherence data rather than calculate actual values for  $T_e$  in various directions. Perhaps the most striking difference between theory and actual observations of continental gravity/topography coherence is the failure of the coherence to reach unity at the longest wavelengths. Among the possible reasons are the influence of erosional processes on the spectral content of the data [McKenzie and Fairhead, 1997], the improper accounting of dynamic topography due to deeper mantle processes [D'Agostino and McKenzie, 1999], or the displacement of interfaces due to tectonic stresses which do not result in complete isostatic compensation even at large wavelengths (M. P. Doin, personal communication, 2002). Furthermore, the longest-wave-



**Figure 9.** Oceanic and continental Bouguer gravity anomalies, in mGal (1 mGal =  $10^{-5}$  m s<sup>-2</sup>). Continental Bouguer anomaly data from the Australian Geological Survey Organisation. Over the oceans, *Parker*'s [1972] method was used to compute the Bouguer correction with an average crustal thickness of 6 km. The boxes labeled A– C correspond to the regions analyzed in Figure 11.



**Figure 10.** (a–c) Coherence anisotropy between Bouguer gravity and topography for the regions shown in Figure 8 and (d) a synthetic example with random phase data. From left to right: topography, Bouguer anomaly (tick marks every 100 km), coherence square function and transition wavelength, in km, in function of the azimuth  $\theta$ . Minimum transition wavelength at  $\theta_w$  represents the weakest direction of the plate (dashed lines). Maximum at  $\theta_s$  indicates strongest direction (solid lines). Right panel on every row shows strong and weak coherence with their error in function of the wave number. Figure 10d indicates that if the phase relation of topography to gravity was purely random, no directional differences would be detectable.

length value of the coherence, as well as the exact position of the transition wavelength are, as we have argued in section 4.4, directly related to the size of the analysis window. With decreasing window size, the effective averaging bandwidth of the coherence estimation broadens. This results in a lowered coherence at the longest resolvable wavelength and a transition wavelength that is shifted to shorter values. Our analysis cautions against the comparison of absolute values of  $T_e$  between studies performed with different window sizes, but leaves the interpretation of relative variations and anisotropic components sound.

[39] We may assess the significance of our results by comparing them with realistic synthetic topography and gravity. Two random phase realizations of data with power spectra identical to those observed in central Australia are shown in Figure 10d. The coherence shows no particular pattern. The transition wavelengths are erratic and no unique minimum can be found. The wave number coherence is near zero for all wavelengths. Furthermore, we checked the reproducibility of our measurements by rotating circularly tapered data sets over a set of angles  $\theta$ , and verifying that the coherence pattern as well as the interpreted directions rotated over approximately the same angle. We have also checked the consistency of the results by shifting the centers of the data boxes (inverted triangles in Figures 8 and 9).

[40] The analysis described above was carried out on all topography/gravity tiles, and "good" measurements (such

as the examples of Figure 10) were separated from "fair" measurements with less well defined minima. "Bad" measurements were rejected. We summarize the results in section 7.

#### 6.2. Short-Wavelength Response

[41] The anisotropy measured by the transition wavelengths described in section 6.1. probably involves the deeper parts of the elastic lithosphere. Besides from the transition wavelengths we can extract directional variability from the shorter-wavelength portions of the coherence. Isostatic compensation can be aided by shallow, crustal processes such as faulting [*Bechtel*, 1989; *Lowry and Smith*, 1995]. If this occurs in a preferred direction the coherence in that direction will be higher than the isotropic average. Disregarding the long-wavelength parts of the coherence spectrum, or the wavelengths longer than 150 km, we attempt to find the direction in which the average coherence exceeds the value of the other directions.

[42] Figure 11 shows three examples of short-wavelength coherence anisotropy in Australia. Figures 11a-11c show actual data taken from the locations labeled in Figure 9, For comparison, Figure 11d shows a synthetic example with random phase topography unrelated to the gravity anomaly (taken from central Australia). The first two left panels show the topography and gravity fields, whereas the third panel shows the coherence-square function without its longwavelength portion. The right panel gives the average coherence as a function of azimuth with 68% and 95% confidence intervals. The direction of maximum coherence, the weaker or more easily isostatically compensated direction of the continental plate is marked by a dashed line. On the basis of the peaks in the average coherence we identify ("good" measurements), along with more ambiguous peaks ("fair"), and reject "bad" or "null" measurements. All results are summarized in section 7.

#### 7. Discussion

#### 7.1. Mechanical Anisotropy and Zones of Weakness

[43] The "weak" directions obtained from the transition wavelengths are plotted in Figure 12b. The directions where the coherence is higher than average in the short wavelength range are shown in Figure 12c. High-quality measurements are shown with thick solid lines, and fair ones with thin solid lines.

[44] The western two thirds of the Australian continent are a Precambrian amalgamation of numerous continental fragments [Rutland, 1976; Myers et al., 1996] joined by weak zones characterized by extensively reworked crust, which are often manifest as strong gradients in the gravity and magnetic anomaly maps [Wellman, 1998] (Figure 12a). Precambrian Australia comprises the Archean Western Australian craton (notably the Yilgarn and Pilbara cratons), a North Australian craton which extends into the offshore areas of the continental platform [Zielhuis and van der Hilst, 1996; Simons et al., 1999], and a South Australian craton containing the Archean-Proterozoic Arunta and Gawler units. The eastern third of Australia accreted during the Paleozoic and has been the scene of intense orogenic activity with a predominantly N-S oriented strike. The Central Australian region, which extends to the northwest

and separates the western and southern from the northern Australian mega-elements, is an extensively reworked zone with major collisional and other boundaries, which are oriented predominantly E-W in the central parts of Australia [*Wellman*, 1998; *Shaw et al.*, 1995].

[45] Figure 12a [after *Wellman*, 1998] shows the location of major geological and geophysical boundaries. Such zones represent substantial rheological heterogeneities. Their role of weak zones has been invoked to explain the relative stability of the cratonic parts they circumscribe [*Vauchez et al.*, 1998; *Lenardic et al.*, 2000; *Tommasi and Vauchez*, 2001]. The short-wavelength anisotropy (Figure 12c) correlates well with these suture zones. With some exceptions, the weak direction is indeed oriented nearly perpendicularly to the trend directions mapped in Figure 12. In addition, several long-wavelength weak directions are oriented at high angles to the ocean-continent boundary. This may be a manifestation of the rheological weakness associated with the junction of oceanic and continental crust.

#### 7.2. Mechanical Anisotropy and Lithospheric Stress

[46] The clear N-S anisotropy of the central Australian lithosphere evident in the short-wavelength coherence of Figure 12c corroborates the findings of *Simons et al.* [2000], who tentatively related the N-S direction of weakness to the presence of pervasive E-W running zones of faulting [*Lambeck et al.*, 1988] and rheological effects due to differential sediment burial rates [*Sandiford and Hand*, 1998].

[47] If the above average isostatic compensation in the N-S direction observed in central Australia is interpreted as a lowered effective elastic thickness [Simons et al., 2000], then it could be explained as an effect of the regional N-S pattern of maximum compressive stress [Lambeck et al., 1984; Lowry and Smith, 1995]. Similarly, an explanation for mechanical anisotropy elsewhere could be sought in relation to the contemporaneous intraplate stress field. However, regional stress indicators for Australia (e.g., borehole breakouts [Hillis et al., 1998, 1999] or focal mechanisms [Lambeck et al., 1984; Spassov, 1998; Spassov and Kennett, 2000]) are poorly consistent with each other, and Australia has very low seismicity. In addition to the difficulty of identifying a representative stress direction responsible for anisotropic isostatic compensation, there is no a priori causal relationship between them. Loading of an elastic plate and the creation of gravity anomalies might have reflected the stress state of the plate at the moment the loading and deformation took place, but from the preserved gravity structure it is not possible to derive the present-day stress field, unless the present-day stress and fossil strain are still related, which seems to be the case in central Australia, or in some areas of the continental United States [Lowry and Smith, 1995].

#### 7.3. Mechanical Anisotropy and Lithospheric Strain

[48] An interpretation of coherence anisotropy in terms of strain is more promising. Finite strain causes lattice preferred orientation (LPO) of anisotropic mantle minerals and, hence, seismic anisotropy [*Mainprice and Silver*, 1993]. Experimental [e.g., *Zhang and Karato*, 1995] and theoretical [e.g., *Ribe*, 1992] studies have shown that the [100] axis



**Figure 11.** Coherence anisotropy for the shortest wavelengths. (a–c) Data examples (see boxes A–C in Figure 9 for location). (d) A synthetic example with random phase topography unrelated to the gravity anomaly (taken from central Australia). Topography and gravity fields are plotted together with the short-wavelength part of their coherence function. The average coherence per azimuth is plotted with  $1\sigma$  and  $2\sigma$  error bars (dark and light shadings). "Weak" directions are found by looking for a well-identified peaks and direction of higher than average coherence (dashed lines).

of olivine becomes oriented parallel to the extensional direction, or perpendicular to the shortening direction but parallel to transpressional structures [*Nicolas and Poirier*, 1976; *Tommasi*, 1998]. If deformation has operated in a vertically coherent way, the fast polarization direction of *SKS* splitting or the direction of maximum propagation speed of azimuthally anisotropic surface waves will be parallel to the structural trends, or nearly perpendicular to the compression direction in the case of large-scale continental collision [*Silver*, 1996]. By inference, the fast axis will be perpendicular to the direction that has accumulated the most deformation per unit of topographic loading: the

"weak" direction from our coherence analysis, in which deformation is most easily accommodated by folding, faulting or buckling (Figure 13). We will compare the weak plate directions with the fast seismic axes and investigate to which extent they are indeed at right angles to each other, and to which depth.

[49] Interpreting anisotropic isostatic compensation as a function of the strain directions it records arguably provides a better handle on the dominant deformation mechanism than a surface mapping of geologic trends (strike of faults, province boundaries, stress directions, etc.) alone. Isostatic compensation involves all of the elastic lithosphere, and the



isostatic response thus represents the time- and depthintegrated dominant mode of deformation. This interpretation is not unique, however, as we need to assume that the observed topography and gravity are in static equilibrium with each other. Ideally, the flexural response and erosion should be studied as coupled processes [*Stephenson*, 1984; *Stephenson and Lambeck*, 1985b]. However, an alternative model, in which the topography is altered by erosion without affecting the subsurface loads might give rise to anisotropic coherence and cannot be discarded without further study.

[50] Both body wave [Clitheroe and van der Hilst, 1998; Girardin and Farra, 1998; Özalaybey and Chen, 1999] and surface wave [Debayle and Kennett, 2000; Simons et al., 2002] studies suggest that the seismic anisotropy of the Australian lithosphere is more complex than either endmember model, vertically coherent deformation [Silver and Chan, 1988] or present-day mantle deformation [Vinnik et al., 1995] would predict. Measurements of the birefringence of SKS phases in Australia by Clitheroe and van der Hilst [1998] and Özalaybey and Chen [1999] are at odds with each other. Whereas Clitheroe and van der Hilst [1998] attribute the splitting they observe at high frequencies (f >0.3Hz) to a small amount of azimuthal anisotropy at shallow depths, Özalaybey and Chen [1999] contend they are a result of scattering associated with heterogeneities at depth. The splitting of body phases poorly constrains the depth range of lithospheric anisotropy [Savage, 1999], but surface waves enable us to detect horizontal as well as vertical variations in anisotropy and wave speed. However, we emphasize that surface wave studies are conducted at lower frequencies than shear wave splitting studies. Hence surface wave models represent averages of seismic structure over several hundred kilometers.

[51] The two most detailed waveform tomographic models to date, by Debayle and Kennett [2000] and Simons et al. [2002] both indicate a change in anisotropic nature in the midlithosphere. Above 150-200 km depth, the highly variable nature and relatively strong amplitudes of azimuthal anisotropy are suggestive of a regime of frozen strain, whereas the smoother, weaker anisotropy below 200 km is more consistent with active mantle convective processes deforming the lowermost lithosphere and aligning the fast axes with the direction of absolute plate motion. Elsewhere [Simons et al., 2002] we present a model of the 3-D S wave heterogeneity and azimuthal anisotropy from Rayleigh waves propagating in the Australian upper mantle, placing particular emphasis on the construction of the model, its robustness and error structure, and comparing the differences between our model and that of Debayle and Kennett

**Figure 12.** (opposite) (a) Major trend directions on the Australian continent [from *Wellman*, 1998]. Measurements of anisotropy in the coherence between Bouguer anomalies and topography. (b) Long-wavelength weak directions corresponding to the azimuths of minimum transition wavelength. (c) Short-wavelength weak directions from the azimuth with higher than average coherence in the range from 20 to 150 km. "Good" measurements are indicated by thick solid lines; "fair" measurements are indicated by thin solid lines.



**Figure 13.** Vertically coherent deformation of the lithosphere [modified after *Silver*, 1996] (with permission from the Annual Review of Earth and Planetary Sciences, vol. 24, copyright 1996 by Annual Reviews www.annualreviews. org) and the relation between seismic and mechanical anisotropy.

[2000] in detail. The qualitative agreement between both models and the interpretation based on them is encouraging, but many differences remain at the small scales. Consequently, we will focus on our own seismic model for the comparison with mechanical anisotropy as a first-order exercise, and reserve the detailed comparison between seismic models for the more specialized literature.

[52] We have averaged the anisotropy model of *Simons et al.* [2002] over equally sized, similarly overlapping regions of the continent as we used in the coherence analysis. The fast axes for various depths are shown in Figure 14 (black lines scaled to the magnitude of the anisotropy). Overlaying these results, we plot the long-wavelength directions of mechanical weakness (gray lines) in Figures 14a–14d (corresponding to the thick solid lines in Figure 12b), and the short-wavelength results in Figures 14e–14h (the thick solid lines of Figure 12c). We show this comparison only for the locations of the best measurements of the coherence anisotropy.

[53] In most cases the mechanical and seismic anisotropy are at large angles. For different depths, we quantify this by calculating the percentage of weak directions that falls in the quadrant containing the perpendicular to the fast axis of seismic anisotropy. Figure 15 suggests there is indeed some agreement between the predictions based on the long-wavelength mechanical anisotropy and the seismically anisotropic directions. Although the resolution of the seismic model at 30 km depth is arguably the worst, the seismic directions are perpendicular to the directions of mechanical weakness for about 60% of the measurements, while only 10% of the weak directions are parallel to the fast axis of anisotropy. With increasing depth, these figures change to about 30% (perpendicular) and 20% (parallel) at 200 km depth. For the mechanical anisotropy derived from the highfrequency component of the coherence, the relation with seismic anisotropy is more ambiguous. As we have argued before, its origin probably lies in the shallow crust.

[54] In addition to comparing the directions of mechanical weakness with the depth-dependent directions of surface wave anisotropy, we have compared them with splitting directions from SKS, SKKS, and PKS phases [Clitheroe and van der Hilst, 1998; T. Iidaka et al., unpublished data, 2001]. Shear wave splitting times in Australia are, in general, small. Of the 89 measurements reported in the above two studies, only 13 display splitting times equal than or larger than 1 s. At only four stations a corresponding high-quality measurement of coherence anisotropy could be made, and for three out of those four cases, the mechanical anisotropy was very nearly perpendicular to the splitting direction. As shear wave splitting measurements are much influenced by shallow structures [Saltzer et al., 2000], this perpendicular relationship is in agreement with the inference made from Figure 15 that such a relationship holds predominantly for the uppermost upper mantle.

#### 8. Conclusions

[55] Continental deformation is recorded by topography and gravity anomalies, and the coherence function relating them is a wavelength-dependent measure of the amount of flexure experienced by density interfaces in the lithosphere due to loading by a unit of topography. The 2-D nature of the coherence is an expression of the elastic strength or weakness of the plate. The azimuth in which the transition from high to low coherence occurs at the shortest wavelength indicates a direction of mechanical weakness or strain concentration. We can further identify directions of preferential isostatic compensation on the basis of their anomalously high coherence with respect to the isotropic average. Used as a directional deformation indicator, the coherence can supplant the limited and ambiguous mapping of surface trend directions. We can thus infer the lithospheric deformation direction dominant over time and integrated over depth (the "fossil" strain).

[56] Multiwindow methods are necessary to study the anisotropic coherence of 2-D fields. To study variations of coherence with space, multiwavelet or multispectrogram methods are required. The spatiospectral concentration properties of windows or wavelets are evident from their average Wigner-Ville transform (WV). Using the WV, we have shown that Slepian sequences and Slepian wavelets tile the (r, k) space with rectangular concentration regions. A better way of analyzing space-varying spectral properties is with Hermite windowing functions. Their concentration regions are circular and treat the phase space as a whole, without trading spatial for spectral resolution. Our multispectrogram method of coherence analysis uses Hermite data windows and is able to retrieve spatial, azimuthal and wavelength-dependent variations of coherence, in a computationally efficient way.

[57] We have investigated the spatial variations and anisotropy of the coherence between Bouguer gravity anomalies



**Figure 14.** Seismic and mechanical anisotropy. (a–d) Long-wavelength mechanical anisotropy (weak directions from Figure 12b, in gray) compared to the fast axis of seismic anisotropy at different depths [from *Simons et al.*, 2002], proportional to its magnitude. (e–h) Short-wavelength mechanical anisotropy (weak directions from Figure 12c, in gray) compared to seismic fast axes.

and topography on the Australian continent. Some of our measurements appear to be at large angles to boundaries between stable continental cratons, which is consistent with the notion that such boundaries are mechanically weaker.

[58] We have compared the 3-D distribution of fast axes of seismic anisotropy obtained from surface wave waveform tomography to the locally dominant deformation directions from the gravity-topography coherence analysis. Down to about 200 km most mechanically weak and seismically fast directions of anisotropy are at high angles to each other, but at large depth this relationship vanishes. This observation is consistent with the notion that large-scale deformation processes have affected the lithosphere coherently to about 200 km depth. The pattern of anisotropy at 200 and 300 km 8 - 16



**Figure 15.** Relationship between mechanically weak directions and fast axes of seismic anisotropy. Percentage of long (squares, solid lines) and short-wavelength (triangles, dashed lines) weak directions that falls in the quadrant containing (a) perpendicular or (b) parallel to the seismic fast axis.

depth may reflect the approximately northward motion of the Australian plate.

### Appendix A: Flexure, Admittance, Coherence A1. Flexure of the Elastic Lithosphere

[59] In one dimension, the equation describing the deflection n(x) due to a load l(x) of a thin elastic plate overlying a fluid substrate is given by [*Turcotte and Schubert*, 1982]:

$$\left(\frac{d^4}{dx^4} + \frac{\Delta_R}{D}g\right)n(x) = -\frac{\Delta_F g}{D}l(x).$$
 (A1)

The interface being loaded is represented by the density contrast  $\Delta_{\rm F}$  across it (the driving force), and  $\Delta_{\rm R}$  is the restoring force exerted by the interfaces being flexed. The flexural rigidity is denoted by *D*, and *g* is the gravitational acceleration. The effective elastic thickness, *T<sub>e</sub>*, relates to *D* by

$$D = \frac{ET_e^3}{12(1-\nu^2)},$$
 (A2)

where *E* is Young's modulus and  $\nu$  is Poisson's ratio.

[60] For a wave number  $k = 2\pi/\lambda$ , the Fourier transform of the equilibrium load L(k) is related to the Fourier transform of the deflection N(k) by

$$N(k) = -L(k)\frac{\Delta_{\rm F}}{\frac{Dk^4}{g} + \Delta_{\rm R}}.$$
 (A3)

[61] Let us examine the driving and restoring forces in a four-layer (three-interface) density case composed of a layer of air or water (density  $\rho_w$ ), an upper ( $\rho_u$ ) and a lower crust  $(\rho_l)$ , and a half-space mantle  $(\rho_m)$ . The density jumps at the interfaces are then  $\rho_u - \rho_w = \Delta_1$ ,  $\rho_l - \rho_u = \Delta_2$  and  $\rho_m - \rho_l =$  $\Delta_3$ . We denote the topography on interface *i* by  $W_i$ , its topographic surface expression after flexure by  $_{i}H$  (see Table A1), and the resulting Bouguer gravity anomaly observed at the surface by  $_{i}G_{B}$ . For loading at the surface, the equilibrium topography is identical to the surface expression of the load. Furthermore, the deflection due to loading on interface i > 1 is assumed to be expressed equally at all interfaces  $j \neq i$ . As L(k) and N(k) represent equilibrium values, the initial applied topography I(k) is given by the difference of the equilibrium topography and the deflection [Forsyth, 1985].

[62] Neglecting finite amplitude effects, the Bouguer gravity anomaly associated with a warped density interface  $W_i$  located at depth  $z_i$  is given by  $\Delta_i 2\pi G W_i(k) e^{-kzi}$ , where G is the gravitational constant [*Turcotte and Schubert*, 1982]. To obtain the free-air anomaly, the surface contribution is taken into account by adding  $\Delta_1 2\pi G_i H(k)$ .

#### A2. Response of Gravity to Topography

[63] From equation (A3) and Table A1, surface loading of a multilayered system produces warped interfaces,  $W_{i>1}$  at  $z_i$ , related to the loading topography,  $_1H$ , as

$$W_i = -{}_1 H \frac{\Delta_1}{\left[\frac{Dk^4}{g} + \sum_{j>1} \Delta_j\right]} \quad , \quad i > 1.$$
(A4)

The Fourier transform of the Bouguer gravity anomaly due to top loading,  ${}_{1}G_{B}$ , is the sum of the individual gravity anomalies  $\Delta_{i}2\pi GW_{i}e^{-kzi}$  generated by the surface load. Hence the ratio of gravity anomaly to its surface expression, the Bouguer admittance,  $Z_{1}$ , is expressed as

$$Z_1 = -2\pi G \frac{\Delta_1}{\left[\frac{Dk^4}{g} + \sum_{i>1} \Delta_i\right]} \sum_{i>1} \Delta_i e^{-kz_i}.$$
 (A5)

Similarly, the free-air admittance is given by  $Z_1 + 2\pi G\Delta_1$ . Loading  $W_{i>1}$  on an interface different from the surface and

**Table A1.** Forcing  $(\Delta_F)$  and Restoring  $(\Delta_R)$  Density Contrasts, Interface Deflections (*N*), and Equilibrium (*L*) and Initial (*I*) Applied Topography for Loading on Interface *i* 

	_				
i	$\Delta_{\rm F}$	$\Delta_{\rm R}$	L(k)	N(k)	I(k)
1	$\Delta_1$	$\Delta_2 + \Delta_3$	$_{1}H$	$W_2 = W_3$	$_{1}H - W_{2}$
2	$\Delta_2$	$\Delta_1 + \Delta_3$	$W_2$	$_2H$	$W_2 - {}_2H$
3	$\Delta_3$	$\Delta_1 + \Delta_2$	$W_3$	3H	$W_3 - {}_3H$

the subsequent deflection of all other interfaces  $W_{j \neq i}$  (expressed identically at the surface as  $_iH$ ) are related as

$$_{i}H = -W_{i}\frac{\Delta_{i}}{\left[\frac{Dk^{4}}{g} + \sum_{j \neq i} \Delta_{j}\right]} \quad , \quad i > 1.$$
 (A6)

[64] For such subsurface loading scenarios, we get a contribution to the gravity anomaly from the loading term,  $\Delta_i 2\pi G W_i e^{-kzi}$ . In addition, there is a contribution from the deflections of the other interfaces, which, by assumption, are all identical to  $_iH$ , thus  $\Delta_j 2\pi G_i H e^{kz_j}$ . The Bouguer admittance thus becomes

$$Z_{i>1} = 2\pi G\left(\sum_{j\neq i,j>1} \Delta_j e^{-kz_j} - e^{-kz_i} \left[\frac{Dk^4}{g} + \sum_{j\neq i} \Delta_j\right]\right), \quad (A7)$$

and the free-air admittance is given by  $Z_{i>1} + 2\pi G \Delta_1$ .

#### A3. Predicting Coherence

[65] Combining equations (A4) and (A6) and Table A1, the initial applied loads,  $I_i(k)$ , relate to their surface expressions,  $_iH(k)$ , as

$${}_{1}H = \begin{pmatrix} \frac{Dk^{4}}{g} + \sum_{i>1} \Delta_{i} \\ \frac{Dk^{4}}{g} + \sum_{i} \Delta_{i} \end{pmatrix} I_{1},$$
(A8)

$$_{i}H = -\left(\frac{\Delta_{i}}{\frac{Dk^{4}}{g} + \sum_{j}\Delta_{j}}\right)I_{i}$$
,  $i > 1.$  (A9)

Note that to obtain the actual load (in Pa),  $I_i$  (which is expressed in m) needs to be multiplied by  $g\Delta_i$ . The total equilibrium topography resulting from different loading processes is the sum of the individual  $_iH$ , and the total Bouguer anomaly by the sum of  $_iG_B = _iHZ_i$ . For the example of Figure 7, Figures 7a and 7b represent two levels of loading,  $I_1$  and  $I_2$ , and the equilibrium topography of Figure 7c was obtained by adding equations (A8) and (A9). The individual Bouguer anomaly contributions were obtained from equations (A5) and (A7) and equations (A8) and (A9). If the loading processes and the noise are statistically uncorrelated [*Forsyth*, 1985], the coherence square is obtained from

$$\gamma^{2} = \frac{\left(\sum_{i} |_{i} H(\mathbf{k})|^{2} Z_{i}\right)^{2}}{\sum_{i} |_{i} H(\mathbf{k})|^{2} \sum_{i} |_{i} H(\mathbf{k})|^{2} Z_{i}^{2}}.$$
 (A10)

This paper focuses on the estimation of nonstationary  $\gamma^2(\mathbf{r}, \mathbf{k})$ , when spectral coherence varies spatially as well.

#### Appendix B: Spectral Windows and Wavelets B1. Multiwindow and Multiwavelet Methods

[66] All multiple window or wavelet methods have in common that a spectral estimate is made with each different wavelet or window. The results are averaged.

[67] The qualities desired of a set of *J* direct spectral estimators,  $\hat{S}_j$ ,  $j = 0 \rightarrow J - 1$ , of the spectral density, *S*, are unbiasedness (equation (B1a)), a low estimation variance (equation (B1b)), and the minimization of leakage from adjacent wave vectors (equation (B1c)):

$$\forall \mathbf{k} : E\{\hat{S}(\mathbf{k})\} = S(\mathbf{k}),\tag{B1a}$$

$$\lim_{J \to \infty} \operatorname{Var}\{\hat{S}(\mathbf{k})\} = 0, \tag{B1b}$$

$$\forall \mathbf{k}' \neq \mathbf{k} : \operatorname{Cov}\{\hat{S}(\mathbf{k}'), \hat{S}(\mathbf{k})\} = 0.$$
(B1c)

[68] In the following, for the clarity of the presentation we will write the relations for 1-D signals, x(t), defined in the time-frequency plane, (t, f). For stationary signals, the criteria represented by equations (B1a)–(B1c) are met by a class of quadratic spectrum estimators [*Mullis and Scharf*, 1991] consisting of an average of *J* modified (windowed) periodograms acting on data windowed by different tapers,  $h_j$ . The multitaper spectral estimate,  $\hat{S}^{\text{MT}}(f)$ , is defined as [*Thomson*, 1982]

$$\hat{S}^{\rm MT}(f) = \frac{1}{J} \sum_{j=0}^{J-1} \left| \int_{-\infty}^{\infty} h_j(\tau) x(\tau) e^{-i2\pi f \tau} d\tau \right|^2.$$
(B2)

Satisfying equation (B1a), the first moment of  $\hat{S}^{\text{MT}}$  is the true spectrum, *S*, smoothed in frequency space by a kernel,  $\mathcal{H}$ , which is the average periodogram of the windowing functions  $h_i(\tau)$  [*Percival and Walden*, 1993]:

$$E\{\hat{S}^{\mathrm{MT}}(f)\} = S(f) \otimes \mathcal{H}(f).$$
(B3)

Frequency domain convolution is denoted by  $\otimes$ . The explicit form of the kernel function is

$$\mathcal{H}(f) = \frac{1}{J} \sum_{j=0}^{J-1} \left| \int_{-\infty}^{\infty} h_j(\tau) e^{-i2\pi f \tau} d\tau \right|^2.$$
(B4)

The desired (as per equation (B1b)) reduction of estimation variance is obtained by the averaging of more than one direct spectral estimate, while the requirements contained in equation (B1c) can be met by an appropriate choice of the windowing functions  $h_i(\tau)$  (see Appendix B2).

[69] For the estimation of nonstationary spectra, we must first introduce three time-frequency distributions [*Mallat*, 1998]. They are the Wigner-Ville transform ( $P_{WV}$ ), the windowed Fourier transform ( $P_{WF}$ ), and the wavelet transform ( $P_{WT}$ ):

$$P_{WV}[x](t,f) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x\left(t - \frac{\tau}{2}\right) e^{-i2\pi f\tau} d\tau,$$
  

$$P_{WF}[x](t,f) = \int_{-\infty}^{\infty} h(\tau - t) x(\tau) e^{-i2\pi f\tau} d\tau,$$
 (B5)  

$$P_{WT}[x](t,s) = \int_{-\infty}^{\infty} x(\tau) \frac{1}{\sqrt{s}} \psi\left(\frac{\tau - t}{s}\right) d\tau.$$

Here,  $\Psi$  is a real-valued wavelet, and *s* a scaling parameter which can be thought of as a proxy for frequency [*Kumar* and Foufoula-Georgiou, 1997]. The WV is a quadratic energy distribution. It is the Fourier transform of a local autocorrelation function. The energy density function associated with the windowed Fourier transform is called the spectrogram and given by

$$|P_{WF}(t,f)|^2$$
. (B6)

The wavelet energy density function is termed the scalogram and is given by

$$\left|P_{WT}(t,s)\right|^2.\tag{B7}$$

In analogy with equation (B2), *Bayram and Baraniuk* [1996] proposed a multitaper spectrogram,  $\hat{S}^{WF}(t, f)$ :

$$\hat{S}^{\rm WF}(t,f) = \frac{1}{J} \sum_{j=0}^{J-1} \left| \int_{-\infty}^{\infty} x(\tau) h_j(\tau-t) e^{-i2\pi f \tau} d\tau \right|^2.$$
(B8)

On the other hand, *Lilly and Park* [1995] proposed a multiwavelet scalogram,  $\hat{S}^{WT}(t, s)$ , to estimate S(t, f):

$$\hat{S}^{\text{WT}}(t,s) = \frac{2}{J} \sum_{j=0}^{J-1} \left| x(t) \otimes \frac{1}{\sqrt{s}} \psi_j \left(\frac{-t}{s}\right) \right|^2.$$
(B9)

Note that we have rewritten the continuous wavelet transform as a convolution with the time-reversed wavelet. The parallelism between equations (B9) and (B3) can be extended as follows: the time-varying spectrum, S(t, f), is the expectation of the WV [*Mallat*, 1998]:

$$S(t,f) = E\{P_{WV}(t,f)\}.$$
 (B10)

Both the scalogram and the spectrogram are members of a *Cohen*'s [1989] class: they can be obtained by time-frequency or time-scale (affine) smoothing of the WV [*Rioul and Flandrin*, 1992; *Frazer and Boashash*, 1994]. For any such estimator  $\hat{S}$ ,

$$\hat{S} = P_{WV}[x] \otimes \oplus \frac{1}{J} \sum_{j=0}^{J-1} P_{WV}[h_j].$$
(B11)

Consequently, the expectation of either equation (B8) or (B9) is the true spectrum, *S*, smoothed by a kernel which is the average WV of the windowing functions  $h_j(\tau)$  or wavelets  $\psi_j(\tau)$ . Convolution in time is denoted by  $\otimes$  and frequency or scale domain convolution by  $\otimes$ . Comparing equations (B10) and (B11) with equation (B3), we see that the WV takes the role of the periodogram in stationary spectrum analysis. The properties of the WV of the wavelets (see section 3.1) or windows (see section 3.2) used determine the distortion of the estimate with respect to the true properties.

#### **B2.** Stationary Spectral Analysis

[70] *Slepian* [1978] discovered that windowing functions  $h_i(\tau)$  could be found that, notwithstanding the finite data

length, are optimally concentrated in the frequency domain. Their narrow central lobe and low sidelobe level assure low bias and leakage, while their multiplicity reduces the variance of the estimate when used as in equation (B2).

[71] It is straightforward to see how a finite data length, T, implies the action of a time projection operator **T**:

$$\mathbf{T}\{h(t)\} = h(t)R_{-T/2}^{T/2}(t),$$
(B12)

where  $R_{-T/2}^{T/2}$  denotes a boxcar function of length *T*. A low-pass projection within a bandwidth *W* is achieved by the operator L [*Flandrin*, 1988]:

$$\mathbf{L}\{h(t)\} = \int_{-W/2}^{W/2} \hat{h}(f) e^{i2\pi f t} df.$$
 (B13)

As is easily shown, the combined effect of both operators is

$$\mathbf{TL}\{h(t)\} = \int_{-T/2}^{T/2} \frac{\sin \pi W(t-\tau)}{\pi(t-\tau)} h(\tau) \, d\tau.$$
(B14)

The Shannon number, *TW*, is the number of optimally concentrated windows  $h_j(\tau)$  that can be found as the eigenfunctions of equation (B14). These  $h_j(\tau)$  are known as prolate spheroidal wave functions (pswf) or Slepian functions [*Slepian*, 1983]. In our previous paper [*Simons et al.*, 2000], we have outlined their properties, and for more information we refer to *Percival and Walden* [1993].

#### **B3.** Timescale Analysis

[72] By construction, the spectral windows  $\mathcal{H}(f)$  associated with the pswf  $h_j(\tau)$  are low-pass filters concentrated in a domain  $|f| \leq W/2$ . As shown by *Lilly and Park* [1995], wavelets are obtained when equation (B14) is rewritten to find band-pass filters centered around a frequency  $f_c$  as  $|f - f_c| \leq W/2$ . From symmetry considerations (compare with equation (B14)), they are the eigenfunctions of

$$\int_{T/2}^{T/2} \left\{ \frac{\sin 2\pi (f_c + W/2)(t - \tau)}{\pi (t - \tau)} - \frac{\sin 2\pi (f_c - W/2)(t - \tau)}{\pi (t - \tau)} \right\} \psi(\tau) \, d\tau$$
(B15)

The Shannon number is now 2*TW*. (The factor of 2 arises from using both positive and negative frequency intervals in the optimization condition; it is also present in equation (B9)). The 2*TW* solutions  $\psi_f(T, f_c, W; \tau)$  are indeed wavelets: they are asymptotically self-similar, and as their length increases their sensitivity shifts to lower frequencies and smaller bandwidths. *Lilly and Park* [1995] have called them Slepian wavelets. Examples of their usage are shown by *Bear and Pavlis* [1997, 1999].

#### **B4.** Statistics of Multiwindow Coherence Estimates

[73] The statistics of coherence square estimators have been studied extensively [e.g., *Carter*, 1987; *Thomson and Chave*, 1991]. *Munk and Cartwright* [1966] and *Bendat* [1978] derive ad hoc expressions for the bias and variance of the estimator  $\hat{\gamma}^2$ . If  $\hat{\gamma}^2$  is calculated as the average of *J* uncorrelated direct estimates (such as the expression in



**Figure B1.** Bias and variance of coherence square estimates with multiwindow methods. (a) Expectation and (b) standard deviation of the estimator  $\hat{\gamma}^2$  in function of the true  $\gamma^2$ , from distribution theory. (c) Approximate expectation and (d) standard deviation of  $\hat{\gamma}^2$  used in this paper.

equation (B8)), its first and second moments can be expressed in terms of the true coherence square function  $\gamma^2$  as

$$E\{\hat{\gamma}^2\} = \gamma^2 + \frac{(1-\gamma^2)^2}{J},$$
 (B16)

$$\sigma^{2}\left\{\hat{\gamma}^{2}\right\} = 2\gamma^{2}\frac{\left(1-\gamma^{2}\right)^{2}}{J}.$$
(B17)

[74] Munk and Cartwright [1966] and Bendat [1978] do not agree on the expression for equation (B16), but as we state it, equation (B16) constitutes the best approximation of the exact expressions of Carter et al. [1973] and Touzi et al. [1999] (see Figures B1a and B1b). Equation (B16) implies the estimate is positively biased. The bias is accentuated for low coherences based on few estimates. The effect of the overestimation of the coherence, which, with traditional windowed periodogram methods, occurs mostly in the long-wavelength range of the spectrum, motivated Simons et al. [2000] to reevaluate the single-taper Australian coherence measurements by Zuber et al. [1989] with a multiple-taper technique. The more tapers are included in the calculation of the spectrogram, the lower the variance of the estimate. However, as we have seen, the  $\sim 1/J$  decrease in the estimation variance is achieved at the expense of widening the concentration region  $R = \sqrt{J}$ , which degrades the spectral resolution. The individual estimates based on data windowed with orthonormal windows are indeed all approximately uncorrelated, as long as the frequency is not too close to 0 or the Nyquist frequency, the number of tapers is less than the Shannon number, and furthermore, if the true spectral density function is smoothly varying [Thomson, 1982; Percival and Walden, 1993; Walden et al., 1994].

[75] The approximate result stated in equation (B17) is a Cramer-Rao lower bound on the variance which is asymptotically achieved by maximum likelihood estimates [*Sey*-

mour and Cumming, 1994; Touzi et al., 1999]. Exact expressions for bias and variance are based on distribution theory [*Carter et al.*, 1973; Touzi et al., 1996]. The distribution functions for the coherence square estimator of two Gaussian processes are complicated expressions involving gamma functions and generalized hypergeometric functions. Some approximations are given by *Walden* [1990b]. The validity of these expressions has been verified experimentally in Monte Carlo experiments [*Guarnieri and Prati*, 1997].

[76] We compare the exact and the approximate (Cramer-Rao) expressions for the bias (plotted in Figures B1a and B1c) and standard deviation (Figures B1b and B1d) of the coherence-square estimate in function of the number of uncorrelated estimates, J. Note that, in two dimensions, a concentration radius of R = 3 gives rise to  $J = R^2$  tapers per dimension, which amounts to  $J = R^4$  approximately uncorrelated spectral estimates from which the coherence function is calculated.

[77] From Figure B1 we conclude the following. First, our estimates based on several approximately uncorrelated tapered estimates are nearly unbiased, and this holds for all coherence values and for all regions in the wave vector plane. There is no need for a bias correction, as needed for mirrored or single-window estimates [Munk and Cartwright, 1966; Zuber et al., 1989]. Second, when basing our estimates on nine different windowing functions in each of two dimensions (leading to 81 two-dimensional windows), we may safely quote the easily computed approximate error of equation (B17), calculated with the coherence-square estimate instead of the true, unknown, coherence-square function. As a cautionary note, we add that even when formal error estimates are available, the significance of our results has to be assessed by experiments on the data, as described in the text.

[78] Acknowledgments. We thank Oded Aharonson, Mark Behn, Don Forsyth, Felix Herrmann, Hrafnkell Kárason, Dan McKenzie, Gilbert Strang, Paul Tapponnier, and Andrew Walden for help and advice. The comments by Marie-Pierre Doin, Brian Kennett, Guust Nolet, and the anonymous Associate Editor were of great value in improving the presentation of this material. This research was supported by NSF grant EAR-0001136 and NASA grant NAG5-11113.

#### References

- Banks, R. J., S. C. Francis, and R. G. Hipkin, Effects of loads in the upper crust on estimates of the elastic thickness of the lithosphere, Geophys. J. Int., 145, 291-299, 2001.
- Bayram, M., and R. G. Baraniuk, Multiple window time-frequency analysis, in Proceedings of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, pp. 173-176, Inst. of Electr, Electron. Eng., New York, 1996.
- Bayram, M., and R. Baraniuk, Multiple window time-varying spectrum estimation, in Non-linear and Nonstationary Signal Processing, edited by R. L. Smith et al., pp. 292-316, Cambridge Univ. Press, New York, 2001.
- Bear, L. K., and G. Pavlis, Estimation of slowness vectors and their uncertainties using multi-wavelet seismic array processing, Bull. Seismol. Soc. Am., 87, 755-769, 1997.
- Bear, L. K., and G. Pavlis, Multi-channel estimation of time residuals from broadband seismic data using multi-wavelets, Bull. Seismol. Soc. Am., 89, 681-692, 1999.
- Bechtel, T. D., Mechanisms of isostatic compensation in East Africa and North America, Ph.D. thesis, Brown Univ., Providence, R. I., 1989.
- Bechtel, T. D., D. W. Forsyth, and C. J. Swain, Mechanisms of isostatic compensation in the vicinity of the East African Rift, Kenya, Geophys. J. R. Astron. Soc., 90, 445-465, 1987.
- Bendat, J. S., Statistical errors in measurement of coherence functions and input/output quantities, J. Sound Vibration, 59, 405-421, 1978.
- Bendat, J. S., and A. G. Piersol, Engineering Applications of Correlation and Spectral Analysis, 2nd ed., John Wiley, New York, 1993.
- Bendat, J. S., and A. G. Piersol, Random data: Analysis and Measurement Procedures, 3rd ed., John Wiley, New York, 2000.
- Bergeron, S. Y., D. A. Yuen, and A. P. Vincent, Capabilities of 3-D wavelet transforms to detect plume-like structures from seismic tomography, Geophys. Res. Lett., 27, 3433-3436, 2000.
- Bronez, T. P., On the performance advantage of multitaper spectral analysis, *IEEE Trans. Signal Process.*, 40, 2941–2946, 1992.
- Burov, E. B., and M. Diament, The effective elastic thickness  $(T_e)$  of continental lithosphere: What does it really mean?, J. Geophys. Res., 100, 3905-3927, 1995.
- Cakrak, F., and P. J. Loughlin, Multiple window time-varying spectral analysis, IEEE Trans. Signal Process., 49, 448-453, 2001
- Carter, G. C., Coherence and time-delay estimation, Proc. IEEE, 75, 236-255, 1987.
- Carter, G. C., C. H. Knapp, and A. H. Nuttal, Statistics of the estimate of the magnitude-coherence function, IEEE Trans. Audio Electroacoust., AU21, 388-389, 1973.
- Chen, W.-P., and S. Özalaybey, Correlation between seismic anisotropy and Bouguer anomalies in Tibet and its implications for lithospheric structures, Geophys. J. Int., 135, 93-101, 1998.
- Clitheroe, G., and R. D. van der Hilst, Complex anisotropy in the Australian lithosphere from shear-wave splitting in broad-band SKS-records, in Structure and Evolution of the Australian Continent, Geodyn. Ser., vol. 26, edited by J. Braun et al., pp. 73-78, AGU, Washington, D. C., 1998.
- Cohen, L., Time-frequency distributions-A review, Proc. IEEE, 77, 941-981, 1989.
- D'Agostino, N., and D. McKenzie, Convective support of long-wavelength topography in the Apennines (Italy), Terra Nova, 11, 234-238, 1999.
- Daubechies, I., Time-frequency localization operators: A geometric phase space approach, IEEE Trans. Inf. Theory, 34, 605-612, 1988.
- Daubechies, I., and T. Paul, Time-frequency localization operators: A geometric phase space approach: II, The use of dilations, Inverse Problems, 4. 661-680, 1988.
- Debayle, E., and B. L. N. Kennett, The Australian continental upper mantle: Structure and deformation inferred from surface waves, J. Geophys. Res., 105, 25, 423-25, 450, 2000.
- Escartin, J., and J. Lin, Tectonic modification of axial crustal structure: Evidence from spectral analyses of residual gravity and bathymetry of the Mid-Atlantic Ridge flanks, Earth Planet. Sci. Lett., 154, 279-293, 1998
- Flandrin, P., Maximum signal energy concentration in a time-frequency domain, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4, pp. 2176-2179, Inst. of Electr. and Electron. Eng., New York, 1988.
- Flandrin, P., Temps-Fréquence, 2 ed., Hermès, Paris, 1998.
- Forsyth, D. W., Subsurface loading and estimates of the flexural rigidity of continental lithosphere, J. Geophys. Res., 90, 12,623-12,632, 1985.

- Foufoula-Georgiou, E., and P. Kumar (Eds.), Wavelets in Geophysics, Academic, San Diego, Calif., 1994
- Frazer, G., and B. Boashash, Multiple window spectrogram and time-frequency distributions, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4, pp. 293-296, Inst. of Electr. and Electron. Eng., New York, 1994. Gesch, D., K. L. Verdin, and S. K. Greenlee, New land surface digital
- elevation model covers the Earth, Eos Trans. AGU, 80, 69-70, 1999.
- Girardin, N., and V. Farra, Azimuthal anisotropy in the upper mantle from observations of P-to- S converted phases: Application to southeast Australia, Geophys. J. Int., 133, 615-629, 1998.
- Griot, D.-A., and J.-P. Montagner, Confrontation of mantle seismic anisotropy with two extreme models of strain, in central Asia, Geophys. Res. Lett., 25, 1447-1450, 1998.
- Guarnieri, A. M., and C. Prati, SAR interferometry: A "quick and dirty" coherence estimator for data browsing, IEEE Trans. Geosci. Remote Sens., 35, 660-669, 1997.
- Hanssen, A., Multidimensional multitaper spectral estimation, Signal Process., 58, 327-332, 1997.
- Hillis, R. R., J. J. Meyer, and S. D. Reynolds, The Australian stress map, Explor. Geophys., 29, 420-427, 1998.
- Hillis, R. R., J. R. Enever, and S. D. Reynolds, In situ stress field of eastern Australia, Aust. J. Earth Sci., 46, 813-825, 1999.
- Jeong, J., and W. J. Williams, Alias-free generalized discrete-time timefrequency distributions, IEEE Trans. Signal Process., 40, 2757-2765, 1992
- Kay, S. M., and S. L. Marple, Spectrum analysis: A modern perspective, Proc. IEEE, 69, 1380-1419, 1981.
- Kumar, P., A wavelet-based methodology for scale-space anisotropic analysis, Geophys. Res. Lett., 22, 2777-2780, 1995
- Kumar, P., and E. Foufoula-Georgiou, Wavelet analysis for geophysical applications, Rev. Geophys., 35, 385-412, 1997.
- Lambeck, K., H. W. S. McQueen, R. A. Stephenson, and D. Denham, The state of stress within the Australian continent, Ann. Geophys., 2, 723-742, 1984.
- Lambeck, K., G. Burgess, and R. D. Shaw, Teleseismic travel-time anomalies and deep crustal structure in central Australia, Geophys. J. Int., 94, 105-124, 1988.
- Lenardic, A., L. Moresi, and H. Mühlhaus, The role of mobile belts for the longevity of deep cratonic lithosphere: The crumple zone model, Geophys. Res. Lett., 27, 1235-1238, 2000.
- Lilly, J. M., and J. Park, Multiwavelet spectral and polarization analyses of seismic records, Geophys. J. Int., 122, 1001-1021, 1995.
- Liu, P. C., Wavelet spectrum analysis and ocean wind waves, in Wavelets in Geophysics, edited by E. Foufoula-Georgiou and P. Kumar, pp. 151-166, Academic, San Diego, Calif., 1994.
- Liu, T.-C., and B. D. van Veen, Multiple window based minimum variance spectrum estimation for multidimensional random fields, IEEE Trans. Signal Process., 40, 578-589, 1992.
- Lowry, A. R., and R. B. Smith, Flexural rigidity of the Basin and Range-Colorado Plateau-Rocky Mountain transition from coherence analysis of gravity and topography, J. Geophys. Res., 99, 20,123-20,140, 1994. Lowry, A. R., and R. B. Smith, Strength and rheology of the western U.S.
- Cordillera, J. Geophys. Res., 100, 17,947-17,963, 1995.
- Mainprice, D., and P. G. Silver, Interpretation of SKS-waves using samples from the subcontinental lithosphere, Phys. Earth Planet. Inter., 78, 257-280, 1993
- Mallat, S., A Wavelet Tour of Signal Processing, Academic, San Diego, Calif., 1998.
- McKenzie, D. P., and C. Bowin, The relationship between bathymetry and gravity in the Atlantic Ocean, J. Geophys. Res., 81, 1903-1915, 1976.
- McKenzie, D. P., and J. D. Fairhead, Estimates of the effective elastic thickness of the continental lithosphere from Bouguer and free air gravity anomalies, J. Geophys. Res., 102, 27,523-27,552, 1997.
- Meissner, R., W. D. Mooney, and I. Artemieva, Seismic anisotropy and mantle creep in young orogens, Geophys. J. Int., 149, 1-14, 2002.
- Mellors, R. J., F. L. Vernon, and D. J. Thomson, Detection of dispersive signals using multitaper dual-frequency coherence, Geophys. J. Int., 135, 146-154, 1998.
- Meredith, R. W., Coherence estimation for high-frequency narrowband cw pulsed signals in shallow water, J. Acoust. Soc. Am., 106, 828-836, 1999.
- Mullis, C. T., and L. L. Scharf, Quadratic estimators of the power spectrum, in Advances in Spectrum Analysis and Array Processing, edited by S. Haykin, vol. 1, pp. 1-57, Prentice-Hall, Old Tappan, N. J., 1991.
- Munk, W. H., and D. E. Cartwright, Tidal spectroscopy and prediction, Philos. Trans. R. Soc. London, Ser. A, 259, 533-581, 1966
- Myers, J. S., R. D. Shaw, and I. M. Tyler, Tectonic evolution of Proterozoic Australia, Tectonics, 15, 1431-1446, 1996.
- Nicolas, A., and J.-P. Poirier, Crystalline Plasticity and Solid State Flow in Metamorphic Rocks, John Wiley, New York, 1976.

- Olhede, S., and A. T. Walden, Generalized morse wavelets, *IEEE Trans.* Signal Process., 50, 2661–2670, 2002.
- Özalaybey, S., and W. P. Chen, Frequency-dependent analysis of SKS-SKKS waveforms observed in Australia: Evidence for null birefringence, *Phys. Earth Planet. Inter.*, 114, 197–210, 1999.
- Parker, R. L., The rapid calculation of potential anomalies, *Geophys. J. R. Astron. Soc.*, 31, 447–455, 1972.
- Parks, T. W., and R. G. Shenoy, Time-frequency concentrated basis functions, in *Proceedings of IEEE International Conference on Acoustics*, *Speech, and Signal Processes*, vol. 5, pp. 2459–2462, Inst. of Electr, Electron. Eng., New York, 1990.
- Percival, D. B., and A. T. Walden, Spectral Analysis for Physical Applications, Multitaper and Conventional Univariate Techniques, Cambridge Univ. Press, New York, 1993.
- Ribe, N. M., On the relation between seismic anisotropy and finite strain, J. Geophys. Res., 97, 8737–8747, 1992.
- Rioul, O., and P. Flandrin, Time-scale energy distributions: A general class extending wavelet transforms, *IEEE Trans. Signal Process.*, 40, 1746– 1757, 1992.
- Rümpker, G., and P. G. Silver, Apparent shear-wave splitting parameters in the presence of vertically varying anisotropy, *Geophys. J. Int.*, 135, 790– 800, 1998.
- Rutland, R. W. R., Orogenic evolution of Australia, *Earth Sc. Rev.*, 12, 161–196, 1976.
- Saltzer, R. L., J. B. Gaherty, and T. H. Jordan, How are vertical shearwave splitting measurements affected by variations in the orientation of azimuthal anisotropy with depth?, *Geophys. J. Int.*, 141, 374–390, 2000.
- Sandiford, M., and M. Hand, Controls on the locus of intraplate deformation in central Australia, *Earth Planet. Sci. Lett.*, 162, 97–110, 1998.
- Santoso, S., E. J. Powers, R. D. Bengston, and A. Ouroua, Time-series analysis of nonstationary plasma fluctuations using wavelet transforms, *Rev. Sci. Instrum.*, 68, 898–901, 1997.
- Savage, M. K., Seismic anisotropy and mantle deformation: What have we learned from shear wave splitting?, *Rev. Geophys.*, 37, 65–106, 1999.
- Scheirer, D., D. W. Forsyth, and A. Hosford, Multitaper estimates of the lithospheric strength of the Basin and Range province (abstract), *Eos Trans. AGU*, 76(17), Spring Meet. Suppl., S282, 1995.
- Seymour, M. S., and I. G. Cumming, Maximum likelihood estimation for SAR interferometry, in *Proceedings of the International Geosciences and Remote Sensing Symposium, IGARSS '94*, vol. 4, pp. 2272–2275, Inst. of Electr, Electron. Eng., New York, 1994.
- Shaw, R. D., P. Wellman, P. Gunn, A. J. Whitaker, C. Tarlowski, and M. P. Morse, Australian crustal elements map, AGSO Res. Newslett., 23, 1–3, 1995.
- Silver, P. G., Seismic anisotropy beneath the continents: Probing the depths of geology, *Annu. Rev. Earth Planet. Sci.*, 24, 385–432, 1996.
- Silver, P. G., and W. W. Chan, Implications for continental structure and evolution from seismic anisotropy, *Nature*, 335, 34–39, 1988.
- Silver, P. G., and W. Chan, Shear wave splitting and subcontinental mantle deformation, J. Geophys. Res., 96, 16,429–16,454, 1991.
- Simons, F. J., and R. D. van der Hilst, Age-dependent seismic thickness and mechanical strength of the Australian lithosphere, *Geophys. Res. Lett.*, 29, 1529, doi:10.1029/2002GL014962, 2002.
- Simons, M., S. C. Solomon, and B. H. Hager, Localization of gravity and topography: Constraints on the tectonics and mantle dynamics of Venus, *Geophys. J. Int.*, 131, 24–44, 1997.
- Simons, F. J., A. Zielhuis, and R. D. van der Hilst, The deep structure of the Australian continent from surface-wave tomography, *Lithos*, 48, 17–43, 1999.
- Simons, F. J., M. T. Zuber, and J. Korenaga, 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, 2000.
- Simons, F. J., R. D. van der Hilst, J.-P. Montagner, and A. Zielhuis, Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle, *Geophys. J. Int.*, 151, 1–17, 2002.
- Slepian, D., Prolate spheroidal wave functions, Fourier analysis and uncertainty, V, The discrete case, *Bell Syst. Tech. J.*, 57, 1371–1429, 1978.
- Slepian, D., Some comments on Fourier-analysis, uncertainty and modeling, *SIAM Rev.*, 25, 379–393, 1983.
- Smith, W. H. F., and D. T. Sandwell, Global sea floor topography from satellite altimetry and ship depth soundings, *Science*, 277, 1956–1962, 1997.
- Spassov, E., The stress field in Australia from composite fault plane solutions of the strongest earthquakes in the continent, J. Seismol., 2, 173– 178, 1998.

- Spassov, E., and B. L. N. Kennett, Stress and faulting in southeast Australia as derived from the strongest earthquakes in the region, *J. Asian Earth Sci.*, 18, 17–23, 2000.
- Stephenson, R., Flexural models of continental lithosphere based on the long-term erosional decay of topography, *Geophys. J. R. Astron. Soc.*, 77, 385–413, 1984.
- Stephenson, R., and C. Beaumont, Small-scale convection in the upper mantle and the isostatic response of the Canadian shield, in *Mechanisms* of *Continental Drift and Plate Tectonics*, edited by P. A. Davies and S. K. Runcorn, pp. 111–122, Academic, San Diego, Calif., 1980.
- Stephenson, R., and K. Lambeck, Isostatic response of the lithosphere with in-plane stress: Application to central Australia, J. Geophys. Res., 90, 8581–8588, 1985a.
- Stephenson, R., and K. Lambeck, Erosion-isostatic rebound models for upllift: An application to south-eastern Australia, *Geophys. J. R. Astron.* Soc., 82, 31–55, 1985b.
- Strang, G., and T. Nguyen, *Wavelets and Filter Banks*, 2nd ed., Wellesley-Cambridge Press, Wellesley, Mass., 1997.
- Thomson, D. J., Spectrum estimation and harmonic analysis, *Proc. IEEE*, 70, 1055–1096, 1982.
- Thomson, D. J., and A. D. Chave, Jackknifed error estimates for spectra, coherences, and transfer functions, in *Advances in Spectrum Analysis and Array Processing*, edited by S. Haykin, vol. 1, pp. 58–113, Prentice-Hall, Old Tappan, N. J., 1991.
- Tommasi, A., Forward modeling of the development of seismic anisotropy in the uppper mantle, *Earth Planet. Sci. Lett.*, *160*, 1–13, 1998.
- Tommasi, A., and A. Vauchez, Continental rifting parallel to ancient collisional belts: An effect of the mechanical anisotropy of the lithospheric mantle, *Earth Planet. Sci. Lett.*, 185, 199–210, 2001.
- Torrence, C., and G. P. Compo, A practical guide to wavelet analysis, Bull. Am. Meteorol. Soc., 79, 61–78, 1998.
- Touzi, R., A. Lopes, and P. Vachon, Estimation of the coherence function for interferometric SAR applications, paper presented at European Conference on Synthetic Aperture Radar 1996, EUSAR, Königswinter, Germany, 1996.
- Touzi, R., A. Lopes, J. Bruniquel, and P. Vachon, Coherence estimation for SAR imagery, *IEEE Trans. Geosci. Remote Sens.*, 37, 135–149, 1999.
- Turcotte, D. L., and G. Schubert, *Geodynamics, Application of Continuum Physics to Geological Problems*, John Wiley, New York, 1982.
- Vauchez, A., A. Tommasi, and G. Barruol, Rheological heterogeneity, mechanical anisotropy and deformation of the continental lithosphere, *Tectonophysics*, 296, 61–86, 1998.
- Vinnik, L. P., R. W. E. Green, and L. O. Nicolaysen, Recent deformations of the deep continental root beneath southern Africa, *Nature*, 375, 50– 52, 1995.
- Walden, A. T., Improved low-frequency decay estimation using the multitaper spectral-analysis method, *Geophys. Prospect.*, 38, 61–86, 1990a.
- Walden, A. T., Maximum likelihood estimation of magnitude-squared multiple and ordinary coherence, *Signal Process.*, 19, 75–82, 1990b.
- Walden, A. T., E. J. McCoy, and D. B. Percival, The variance of multitaper spectrum estimates for real Gaussian processes, *IEEE Trans. Signal Pro*cess., 2, 479–482, 1994.
- Walden, A. T., E. J. McCoy, and D. B. Percival, The effective bandwidth of a multitaper spectral estimator, *Biometrika*, 82, 201–214, 1995.
- Watts, A. B., *Isostasy and Flexure of the Lithosphere*, Cambridge Univ. Press, New York, 2001.
- Wellman, P., Mapping of geophysical domains in the Australian continental crust using gravity and magnetic anomalies, in *Structure and Evolution of the Australian Continent, Geodyn. Ser.*, vol. 26, edited by J. Braun et al., pp. 59–71, AGU, Washington, D. C., 1998.
- Zhang, S. Q., and S. Karato, Lattice preferred orientation of olivine aggregates deformed in simple shear, *Nature*, 375, 774–777, 1995.
- Zielhuis, A., and R. D. van der Hilst, Upper-mantle shear velocity beneath eastern Australia from inversion of waveforms from Skippy portable arrays, *Geophys. J. Int.*, *127*, 1–16, 1996.
- Zuber, M. T., T. D. Bechtel, and D. W. Forsyth, Effective elastic thicknesses of the lithosphere and mechanisms of isostatic compensation in Australia, *J. Geophys. Res.*, 94, 9353–9367, 1989.

F. J. Simons, Geosciences Department, Princeton University, Guyot Hall 321B, Princeton, NJ 08544, USA. (fjsimons@alum.mit.edu)

R. D. van der Hilst, Faculteit Aardwetenschappen, Universiteit Utrecht, 3584 CD Utrecht, Netherlands.

M. T. Zuber, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Rm 54-517A, Cambridge, MA 02139, USA.