The spherical Slepian basis as a means to obtain spectral consistency between mean sea level and the geoid

D. C. Slobbe, F. J. Simons & R. Klees

🙋 Springer

The spherical Slepian basis as a means to obtain spectral consistency between mean sea level and the geoid

D. C. Slobbe · F. J. Simons · R. Klees

Abstract The mean dynamic topography (MDT) can be computed as the difference between the mean sea level (MSL) and a gravimetric geoid. This requires that both data sets are spectrally consistent. In practice, it is quite common that the resolution of the geoid data is less than the resolution of the MSL data, hence, the latter need to be low-pass filtered before the MDT is computed. For this purpose conventional low-pass filters are inadequate, failing in coastal regions where they run into the undefined MSL signal on the continents. In this paper, we consider the use of a bandlimited, spatially concentrated Slepian basis to obtain a low-resolution approximation of the MSL signal. We compute Slepian functions for the oceans and parts of the oceans and compare the performance of calculating the MDT via this approach with other methods, in particular the iterative spherical harmonic approach in combination with Gaussian low-pass filtering, and various modifications. Based on the numerical experiments, we conclude that none of these methods provide a low-resolution MSL approximation at the sub-decimetre level. In particular, we show that Slepian functions are not appropriate basis functions for this problem, and a Slepian representation of the low-resolution MSL signal suffers from broadband leakage. We also show that a meaningful definition of a low-resolution MSL over incomplete spherical domains involves orthogonal basis functions with additional properties that Slepian functions do not possess. A low-resolution MSL signal, spectrally consistent with a given geoid model, is obtained by a suitable truncation of the expansions of the MSL signal in terms of these orthogonal basis functions. We compute one of these sets of orthogonal basis functions using the Gram–Schmidt orthogonalization for spherical harmonics. For the oceans, we could construct an orthogonal basis only for resolutions equivalent to a spherical harmonic degree 36. The computation of a basis with a higher resolution fails due to inherent instabilities. Regularization reduces the instabilities but destroys the orthogonality and, therefore, provides unrealistic low-resolution MSL approximations. More research is needed to solve the instability problem, perhaps by finding a different orthogonal basis that avoids it altogether.

1 Introduction

Since ocean currents are nearly in geostrophic balance on time scales longer than a few days and spatial scales longer than a few tens of kilometers, the velocity of the surface geostrophic current is proportional to the gradient in the average height of the sea surface expressed relative to the geoid (Wunsch and Gaposchkin 1980). This so-called mean dynamic topography (MDT) can be computed by subtracting a gravimetric geoid from an altimetric mean sea level (MSL) model, provided that both are expressed relative to the same reference ellipsoid and in the same permanent tide system. However, especially in the open ocean, the altimetric MSL has a much higher spatial resolution than the gravimetric geoid. Consequently, the apparently straightforward computation of the MDT becomes problematic since

D. C. Slobbe (✉) · R. Klees
Delft University of Technology,
P. O. Box 5048, 2600 GA, Delft, The Netherlands
e-mail: D.C.Slobbe@tudelft.nl

F. J. Simons
Department of Geosciences, Princeton University,
Guyot Hall, Princeton, NJ 08544, USA

the high-frequency part that is lacking from the geoid cannot be subtracted from the MSL. This causes a non-random error that can be misinterpreted as a part of the MDT (Losch et al. 2002; Albertella and Rummel 2009). Before the MDT is computed by taking the difference between altimetric MSL and gravimetric geoid, the latter two have to be made "spectrally consistent": they have to cover the same spectral range. Hence, a suitable low-pass filter has to be applied to the altimetric MSL.

The problem of ensuring spectral consistency between two signals before combining them is more general in nature than the specific context in which it appeared above. Indeed, it arises whenever different data sets and/or models with various spatial or temporal resolutions are to be merged for interpretation e.g. in seismology and geomagnetic studies (Trampert and Snieder 1996; Boschi and Dziewonski 1999; Schachtschneider et al. 2010; Schott and Thébault 2011).

To derive the MDT of the North Atlantic Ocean, Jayne (2006) applied a Hamming window smoother in the spatial domain on both geoid and the MSL. However, since low-pass filtering results in the replacement of the original signal by its weighted spatial average, this operation fails in coastal regions, because the MSL is not defined on land. Alternatively, adaptive filters based on principal component analysis (PCA) over the domain of interest might be used. For example, Vianna et al. (2007) used a singular spectrum analysis (SSA) expansion to filter noise in a GRACE-based MDT for the South Atlantic region. SSA could also be used to obtain spectral consistency of altimetric MSL and gravimetric geoid.

On the other hand, we might make use of the fact that geoid models are bandlimited, i.e., they are expressed as a set of spherical harmonic expansion coefficients, complete to some degree L_g. Likewise, the altimetric MSL could be expanded in spherical harmonics complete to some degree L_r, typically with $L_r > L_g$. Spectral consistency could then be obtained by truncating the MSL expansion at degree L_g. Two problems are inherent in this approach. First, since the spherical harmonics are not an orthogonal basis over incomplete spherical domains like the oceans, the estimation of the expansion coefficients from radar altimeter data is ill-conditioned (Simons and Dahlen 2006). Second, truncation of the spherical harmonic expansion beyond degree L_g gives rise to Gibbs phenomena that can only be suppressed with appropriate spectral windows.

The first problem, ill-conditioning, is traditionally addressed by regularized least-squares or truncated singular value decomposition (SVD) approaches (e.g., Xu 1998). More recently, truncated Slepian basis representations have been proposed (Simons and Dahlen 2006), about which more is to follow below. Alternatively, as advocated by, e.g., Tapley et al. (2003), missing MSL data on land and in uncovered

ocean regions like the polar gaps can be replaced by geoid information. Rather than estimating the spherical harmonic coefficients from a signal defined over a subdomain of the sphere, the estimation is then performed using the combined whole-sphere signal. The geoid is used to extend the MSL since the two do not differ more than by about 3 m. Despite this small difference, discontinuities persist at the coastlines, introducing Gibbs phenomena. Bingham et al. (2008) suggest to further reduce the Gibbs effects in the MDT using the same geoid on land as is used to compute the MDT. In an elaboration of the method (Tapley et al. 2003) discussed by Albertella et al. (2008), the transition from land to sea is smoothed by iteratively estimating the spherical harmonic coefficients of MSL up to degree L_r from the combined land–ocean data set. In each iteration, the values over land are replaced by those obtained from a spherical harmonic synthesis of the last derived set of coefficients. This process is repeated until some pre-defined stopping criterion has been satisfied. Finally, low-pass filtering is applied to obliterate the signal above L_g and to reduce Gibbs effects. Albertella et al. (2008) and Albertella and Rummel (2009) use a Gaussian filter, but other (low-pass) filters may be used as well (for more details about the Gaussian filter and alternatives, we refer to, e.g., Jekeli 1981). Generally, however, the low-pass filtered MSL signal still contains energy above degree L_g and will thus not be spectrally consistent with the geoid to which it is compared.

In summary, we might say that no ideal approach currently exists to obtain low-resolution approximations to MSL that are spectrally consistent with the geoid. In studying a one-dimensional version of this problem, Albertella and Rummel (2009) came to the conclusion that extending MSL to the entire globe using, e.g., geoid information unavoidably results in a distortion of its spectral content, even if utmost care is taken to derive smooth transitions from ocean to land. In contrast, they conclude that representing MSL in a basis of Slepian functions, which are suitable for signals of limited bandwidth living on limited intervals (Simons 2010), holds much promise in solving both problems above. Somewhat pessimistically, though, they wrote that "it may prove difficult to apply it to the real world case with the complicated shapes of ocean basins".

Since the efficient generation of Slepian functions on domains of arbitrary geometry, whether in spherical (Simons et al. 2006) or Cartesian (Simons and Wang 2011) coordinates, presents no *intrinsic* difficulties, we are here able to consider their use in the context of the work by Albertella and Rummel (2009), on which we build. Our main goal is to design a bandlimited Slepian basis for the ocean basins in spherical geometry and to evaluate the utility of this basis in solving the problem stated above, which is to derive suitable representations of altimetric MSL while maintaining spectral consistency with the gravimetric geoid. Hwang (1991),

see also Hwang (1993) for a shorter version, used another set of orthogonal functions on the oceans to represent the sea surface topography derived from radar altimeter data. This set of orthogonal basis functions has been generated from spherical harmonics using the Gram–Schmidt process, see Golub and van Loan (1996).

In this article, we first summarize some basics about spherical Slepian functions (Sect. 2 and Appendix A). Following this, we address the problem of spectral consistency and bandwidth for functions given on a part of the spherical domain (Sect. 3). Next, we describe the setup of several numerical experiments, which were designed to investigate the performance of a Slepian basis representation of the MSL signal which is spectrally consistent with a given gravimetric geoid (Sect. 4). This includes a discussion about the choice of the optimal number of Slepian basis functions. Thereafter, we present and discuss the results of the numerical experiments (Sect. 5 and Appendix B). The main result is that all methods discussed in Sect. 5 fail to provide a low-resolution MSL signal with an adequate accuracy. Therefore, in Sect. 6, we reformulate the problem, provide a mathematical solution, and discuss its applicability to high-resolution data. Finally, we conclude by emphasizing the main findings and identifying topics for future research.

2 The spherical Slepian basis

The functions now named after David Slepian grew out of the work by Slepian and Pollak (1961) and Landau and Pollak (1961, 1962), who solved a long-standing problem in information theory, namely, that of optimally concentrating a given signal in both the time and frequency domains. Since timelimited functions cannot be simultaneously bandlimited in the frequency domain, nor vice versa, the optimally concentrated signal is considered to be the one with the least energy outside the interval of interest. The concentration problem has been extended and generalized for the purpose of signal estimation, representation and analysis on geographical domains by Albertella et al. (1999), Pail et al. (2001) and Simons and Dahlen (2006) in geodesy, and by Wieczorek and Simons (2007) and Dahlen and Simons (2008) in more general settings. The quadratic maximization of the spatial energy of bandlimited functions is one way to achieve localization in one domain while curbing leakage in the other. Other constructions may have similarly desirable properties, but even those are often judged on how closely they satisfy quadratic optimality constraints (e.g. Freeden and Michel 1999; Guilloux et al. 2009). We therefore remain faithful to the original approach of Slepian, as transformed into spherical geometry by Simons et al. (2006). Appendix A provides a review of spherical Slepian functions, limited to those aspects

of the theory which are necessary to understand the remaining sections.

3 Spectral consistency and the choice of the bandwidth

Suppose we have access to a high-resolution MSL model and a low-resolution geoid model for the oceans or a part of the oceans, e.g., an ocean basin. Then, spectral consistency between MSL and geoid is obtained when both are represented in a bandlimited, spatially concentrated Slepian basis involving the same set of Slepian basis functions. The choice of the bandwidth should be dictated by the signal that has the lowest resolution, in our case the geoid, since this is the resolution at which we need to describe the MSL to compute the MDT reliably. Geoid information is typically provided in terms of a spherical harmonic expansion complete to some maximum degree, say L_g. This maximum degree describes solely the spatial resolution of the geoid, and, therefore, is also used as a descriptor of the bandwidth.

According to Eq. (A-15), we can transform any given spherical harmonic expansion of the geoid complete to degree L_g into a bandlimited, spatially concentrated Slepian basis invoving $(L_g + 1)^2$ Slepian basis functions. Therefore, a natural choice of the bandwidth of the Slepian functions would be L_g. This is definitely correct as long as the geoid is considered as a function on the whole domain Ω. However, when the geoid signal is confined to a part of the domain such as the oceans or an ocean basin, this definition of the bandwidth of the geoid becomes meaningless as we will show below. Therefore, the choice of the correct bandwidth of the Slepian basis functions is still open and not necessarily given by the maximum degree of the spherical harmonic expansion of the geoid.

In order to investigate the choice of the bandwidth for a signal given in a region Ω_0 of the sphere Ω, we design the following experiment. We assume that a certain signal is given in terms of a spherical harmonic expansion complete to degree $L_g = 48$. We try to reconstruct this bandlimited signal from data inside various regions Ω_0 of different size using spherical harmonic expansions complete to degree $L \leq L_g$. In our experiment, the region Ω_0 is always a spherical cap; the radius of the cap varies in increments of $10°$ between $10°$ and $180°$. The latter case corresponds to the choice $\Omega_0 = \Omega$, the entire sphere. Despite its simplistic geometry, the simple spherical cap is an appropriate choice for a trial region. Its advantage is that the Slepian functions for this particular geometry can be calculated via the painless procedure devised by Grünbaum et al. (1982), as shown by Simons et al. (2006). Furthermore, for complete generality with respect to where we position the cap in the analysis we consider a bandlimited white Gaussian stationary process defined on the region Ω_0, i.e.,

$$w(\hat{\mathbf{x}}) = \begin{cases} \sum_{l=0}^{L_g} \sum_{m=-l}^{l} w_{lm} \hat{Y}_{lm}(\hat{\mathbf{x}}) & \text{if } \hat{\mathbf{x}} \in \Omega_0, \\ \text{unknown} & \text{if } \hat{\mathbf{x}} \in \Omega - \Omega_0, \end{cases} \quad (1)$$

where

$$\mathrm{E}\{w_{lm}\} = 0, \quad \mathrm{Cov}\{w_{lm} w_{l'm'}\} = \delta_{ll'} \delta_{mm'}, \ 0 \le l, l', \le L_g, \tag{2}$$

and $\mathrm{E}\{\cdot\}$ denotes expectation and $\mathrm{Cov}\{\cdot\}$ denotes covariance. The signal $w(\hat{\mathbf{x}})$ is approximated by a function $\hat{w}(\hat{\mathbf{x}})$, which is given by

$$\hat{w}(\hat{\mathbf{x}}) = \sum_{l=0}^{L} \sum_{m=-l}^{l} \hat{w}_{lm} \hat{Y}_{lm}(\hat{\mathbf{x}}), \tag{3}$$

for some maximum degree $L \le L_g$. For continuous data, and using the notation established in Eqs. (A-16)–(A-19), the least-squares estimate of the spherical harmonic coefficients $\hat{w}_{lm}$ is

$$\hat{\mathbf{w}}_L = \mathbf{D}_L^{-1} \mathbf{D}_{L_g} \mathbf{w}_{L_g}, \tag{4}$$

where $\hat{\mathbf{w}}_L = (\hat{w}_{lm} : l, m \le L)^{\mathrm{T}}$, $\mathbf{w}_{L_g} = (w_{lm} : l, m \le L_g)^{\mathrm{T}}$ and $\mathbf{D}_L$ and $\mathbf{D}_{L_g}$ are the matrices with entries $D_{lm,l'm'}$ as in Eq. (A-7) for degrees $l = 0, \dots, L$ and $l = 0, \dots, L_g$, in both dimensions or in one of each dimension: $\mathbf{D}_L$ is square, $\mathbf{D}_{L_g}$ is rectangular, respectively.

If the signal is known at M points, collected in a vector $\mathbf{w}$, we write Eq. (1) as

$$\mathbf{w} = \mathbf{Y}_{L_g}^{\mathrm{T}} \mathbf{w}_{L_g}, \tag{5}$$

using the notation established in Eqs. (A-16)–(A-19). The least-squares estimate $\hat{\mathbf{w}}_L$ is given by

$$\hat{\mathbf{w}}_L = (\mathbf{Y}_L \mathbf{Y}_L^{\mathrm{T}})^{-1} \mathbf{Y}_L \mathbf{w}, \tag{6}$$

$$= (\mathbf{Y}_L \mathbf{Y}_L^{\mathrm{T}})^{-1} (\mathbf{Y}_L \mathbf{Y}_{L_g}^{\mathrm{T}}) \mathbf{w}_{L_g}, \tag{7}$$

which is the analogue to Eq. (4) for discrete data. Irrespectively of the radius Θ of the spherical cap, when $L = L_g$, we thus in principle recover the input without bias, i.e.,

$$\hat{\mathbf{w}}_{L_g} = \mathbf{w}_{L_g}, \tag{8}$$

although the condition number of the normal matrix $\mathbf{Y}_{L_g} \mathbf{Y}_{L_g}^{\mathrm{T}}$ in that case will determine how close we can get. The numerical integration error (i.e., the number and spatial distribution of the samples and the chosen weights) determines how well $\hat{\mathbf{w}}_L$ as defined in Eq. (6) matches the elements $\hat{w}_{lm}$ as defined in Eq. (4). Assuming that the function of interest $w(\hat{\mathbf{x}})$ is densely sampled inside Ω_0, e.g. on the M nodes of a Fibonacci grid (González 2010), where $M \gg |\Omega_0| \left(\frac{L_g}{\pi}\right)^2$, and the weights are equal to $\Delta\Omega = 4\pi/M$, the normal matrix should be well approximated by (Simons 2010)

$$\mathbf{Y}_L \mathbf{Y}_L^{\mathrm{T}} \approx \Delta\Omega^{-1} \mathbf{D}_L. \tag{9}$$

We also note via Eqs. (A-1) and (A-7) that

$$\mathbf{D}_L \to \mathbf{I} \quad \text{for} \quad \Theta \to 180^\circ \quad \text{or} \quad \Omega_0 \to \Omega, \tag{10}$$

where $\mathbf{I}$ is the $(L+1)^2 \times (L+1)^2$ unit matrix. When $L < L_g$, the estimate (7) contains *broadband leakage* (Simons and Dahlen 2006): high-degree signal contributions to the estimated low-degree coefficients. The broadband leakage is $\hat{\mathbf{w}}_L - \mathbf{w}_L$, which follows from

$$\hat{\mathbf{w}}_L = (\mathbf{Y}_L \mathbf{Y}_L^{\mathrm{T}})^{-1} \mathbf{Y}_L (\mathbf{Y}_{L_g}^{\mathrm{T}} \mathbf{w}_{L_g}),$$
$$= (\mathbf{Y}_L \mathbf{Y}_L^{\mathrm{T}})^{-1} \mathbf{Y}_L (\mathbf{Y}_L^{\mathrm{T}} \mathbf{w}_L + \mathbf{Y}_{\to L_g}^{\mathrm{T}} \mathbf{w}_{\to L_g}),$$
$$= \mathbf{w}_L + (\mathbf{Y}_L \mathbf{Y}_L^{\mathrm{T}})^{-1} \mathbf{Y}_L \mathbf{Y}_{\to L_g}^{\mathrm{T}} \mathbf{w}_{\to L_g}, \tag{11}$$

where $\mathbf{Y}_{\to L_g}$ is the lower subblock of the matrix $\mathbf{Y}_{L_g}$ that complements $\mathbf{Y}_L$, and $\mathbf{w}_{\to L_g}$ is the lower subportion of $\mathbf{w}_{L_g}$, covering the degrees $L + 1 \to L_g$. Hence, the broadband leakage is

$$\hat{\mathbf{w}}_L - \mathbf{w}_L = (\mathbf{Y}_L \mathbf{Y}_L^{\mathrm{T}})^{-1} \mathbf{Y}_L \mathbf{Y}_{\to L_g}^{\mathrm{T}} \mathbf{w}_{\to L_g}. \tag{12}$$

Via Eqs. (9)–(10) we deduce from (12) that

$$\hat{\mathbf{w}}_L - \mathbf{w}_L \approx \begin{cases} \Delta\Omega \, (\mathbf{Y}_L \mathbf{Y}_{\to L_g}^{\mathrm{T}}) \mathbf{w}_{\to L_g} & \text{for } \Theta = 180^\circ, \\ \Delta\Omega \, \mathbf{D}_L^{-1} (\mathbf{Y}_L \mathbf{Y}_{\to L_g}^{\mathrm{T}}) \mathbf{w}_{\to L_g} & \text{for } \Theta < 180^\circ. \end{cases} \tag{13}$$

Since the least-squares estimate $\hat{\mathbf{w}}_L$ for the case $\Theta < 180^\circ$ depends on the inverse of the localization kernel $\mathbf{D}_L$, the broadband leakage directly depends on the size and shape of the region of missing data, as well as on the bandwidth $L < L_g$ of the estimate. In the spatial domain, the broadband leakage generates a bias, which Simons and Dahlen (2006) refer to as *broadband bias*.

In practice, it is hard to find a stable inverse of $\mathbf{D}_L$, since for $\Omega_0 \subset \Omega$, the matrix $\mathbf{D}_L$ tends to be poorly conditioned. This problem can be solved either by utilizing a regularized least-squares approach or by using a truncated Slepian basis. The latter refers to a Slepian basis with less than $(L + 1)^2$ basis functions, which improves the condition number of the normal matrix. The former may be obtained by a truncated SVD of the matrix $\mathbf{Y}_L^{\mathrm{T}}$ (Xu 1998; Simons 2010). Here we solve Eq. (6) as

$$\hat{\mathbf{w}}_L = \left(\mathbf{Y}_L^{\mathrm{T}}\right)^+ \mathbf{w}, \tag{14}$$

where

$$\left(\mathbf{Y}_L^{\mathrm{T}}\right)^+ = \mathbf{V} \mathbf{\Sigma}^+ \mathbf{U}^{\mathrm{T}}, \tag{15}$$

and the truncation is accomplished via

$$\Sigma_{ii}^+ = \begin{cases} \Sigma_{ii}^{-1} & \text{for } |\Sigma_{ii}| > \delta, \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$
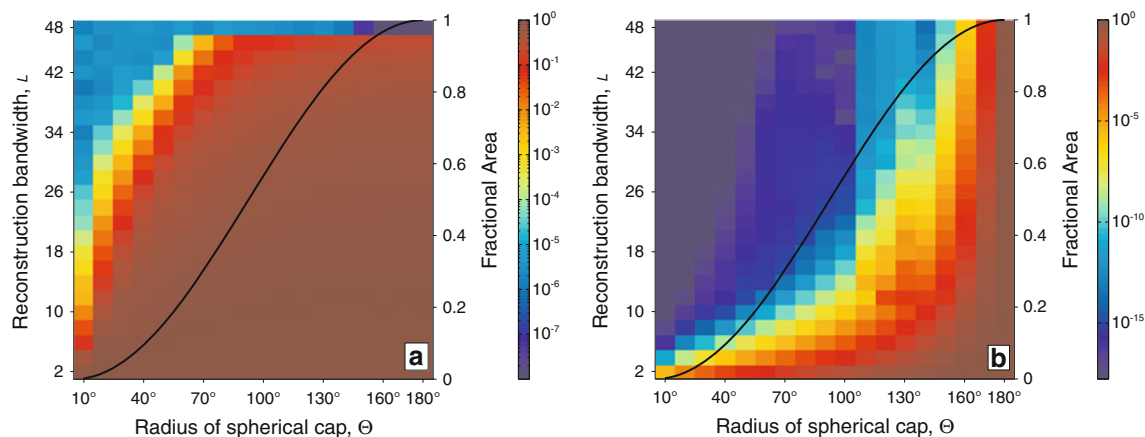
Fig. 1 Reconstruction, via Eq. (14), of a $L_g = 48$ bandlimited white signal, as defined in Eqs. (1)–(2), from a set of observations made on a Fibonacci grid inside north-polar spherical caps of various radii Θ, for various spherical harmonic reconstruction bandwidths L. **a** shows the root mean squared (rms) error normalized by the rms strength of the signal inside the cap. Note that the rms values are computed at points confined to the cap that are different from the data points. **b** shows the inverse of the condition number of the matrix $\mathbf{Y}_L$ in Eq. (6). The *black curves* and the markings on the *right axis* represent the fractional area of the spherical caps, $|\Omega_0|/(4\pi)$

The SVD of matrix $\mathbf{Y}_L^{\mathsf{T}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathsf{T}}$, $\delta = \varepsilon M \max(\Sigma_{ii})$, and ε is the machine epsilon. This is in fact implemented in Matlab's `pinv` routine (MATLAB 7.10.0.499, R2010a). Both the regularized least-squares approach and the truncated Slepian approach introduce a bias in the solution, but lower its variance if there is noise (Simons and Dahlen 2006). This so-called *truncation bias* is the third source of misfit in the approximation.

The propagation of errors in $\hat{\mathbf{w}}_L$ to the spatial domain for each combination of bandwidth L and cap size Θ is expressed in terms of the rms error (rmse), which is computed as the rms difference between the original signal and its least-squares estimate at a set of points inside Ω_0 that are different from the points used in the inversion. This metric, normalized by the rms signal, is shown as a function of Θ and L in Fig. 1a. For all spherical cap radii Θ the rmse decreases if the bandwidth L increases to meet the original L_g, to relative values on the order of 10^{-5} for radii $\Theta \ll 180°$ and 10^{-14} for radii $\Theta \approx 180°$. This decrease is in line with what we could expect, since increasing the bandwidth means that more basis functions are used to approximate the signal.

We derive insight into the behavior of the rmse shown in Fig. 1a via b which shows the inverse of the condition number (an estimate of the ratio of the largest to the smallest singular value) of $\mathbf{Y}_L$ as a function of Θ and L. For radii $\Theta \ll 180°$ the condition number is very large and strong truncation (i.e., regularization) is required to solve Eq. (6). This explains why the normalized rmse does not drop below a value of about 10^{-5} for spherical cap radii significantly smaller than $180°$ even if a bandwidth $L = 48$ is chosen. It is the truncation bias that prevents a significantly smaller rmse. Vice versa, for spherical cap radii close to $180°$, the

truncation bias is negligible, and the rmse is at the level of numerical round-off errors.

Figure 1a also reveals that the rmse as a function of the bandwidth L drops more rapidly for smaller values of the spherical cap radius Θ (i.e. for smaller regions Ω_0). For example, when $\Theta = 10°$ and $L = 22$, we achieve already a very small normalized rmse of 10^{-4}. For larger spherical cap radii, say, $\Theta > 90°$, a similar reduction of the rmse can only be observed if $L = 48$ is chosen. In general, for large regions Ω_0, the rmse decreases very little if the bandwidth is increased, with an abrupt drop at the transition to full bandwidth.

Counter-intuitive is the behavior of the rmse as function of the spherical cap radius Θ for fixed degree L (cf. Fig. 1a). The smaller the spherical cap radius, the smaller the rmse. That is, the bandlimited signal (bandwidth $L = 48$) can be approximated very well by a low-degree ($L < 48$) spherical harmonic expansion provided that the data area is sufficiently small. For fixed degree L, the rmse *increases* with increasing size of the region Ω_0. This suggests that we have a considerable degree of freedom in fitting data from a high-bandwidth model, sampled inside a small region Ω_0, with a low-bandwidth approximation. The smaller the area is, the more freedom we have. This counter-intuitive result is further illustrated in Fig. 2. We observe that the least-squares approximation complete to degree and order L is closer to the original signal (complete to degree and order L_g) than the original signal truncated at degree L.

From this experiment, we conclude that in case a spherical harmonic expansion is fitted to data given on a part of the sphere, the optimal bandwidth to carry out this procedure is no longer a measure for the spatial resolution of the data set.

Fig. 2 The original,
bandlimited signal complete to
degree 48 (*black*), the original
signal truncated at degree L
(*red*), and the least-squares
estimate complete to degree L
(*blue*) for a meridional arc
crossing a spherical cap of
radius Ψ. From *top* to *bottom*:
$L = 2$, $\Psi = 10°$, $L = 22$, $\Psi = 20°$, $L = 40$, $\Psi = 50°$, and
$L = 48$, $\Psi = 100°$. The *thin
black lines* indicate the
boundary of the spherical cap.
Note that within the spherical
cap no differences between the
original signal and the
least-squares approximation are
visible for degrees $L \geq 22$. The
labels indicate the rms error of
the differences between the
original, bandlimited signal
complete to degree 48 on the
one hand and the original signal
truncated at degree L (*left*) or
the least-squares approximation
complete to degree L (*right*) on
the other hand, normalized by
the rms strength of the signal
inside the cap
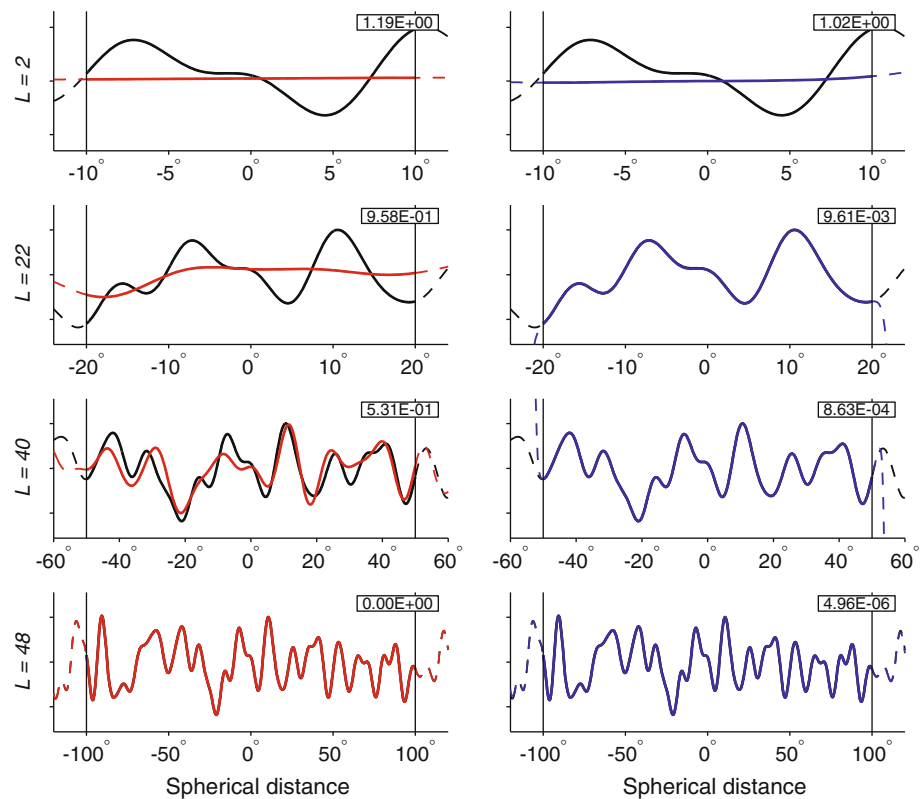
The smaller the area covered with data is, the larger the difference between the optimal bandwidth and the "true" bandwidth. In general, we can expect that the optimal bandwidth is always smaller than the "true" one.

Note that we would come to the same conclusion if we used an actual signal rather than white noise, data with much higher spatial resolution, a different data distribution rather than a Fibonacci grid or a different region rather than a spherical cap centred at the North Pole. The conclusion also does not change if we repeat the experiment on the circle using series expansions in Legendre polynomials.

The major implication for the main objective of this study (i.e., to get a MSL that is spectrally consistent with a given geoid) is that depending on how closely we want to represent the geoid within a region $\Omega_0 \subset \Omega$, a bandwidth L smaller than the nominal value L_g may be appropriate, depending on the size of the region. Hence, the bandwidth of the Slepian basis functions should be set equal to this optimal bandwidth L instead of the nominal bandwidth L_g. This is contrary to the approach of Albertella and Rummel (2009), who maintained L_g as the bandwidth. In practical applications, the choice of the optimal bandwidth L will also depend on the commission error. That is, there is no need to obtain an 'exact' representation of the given geoid in the presence of noise; any approximation will do as long as the approximation error is smaller than the commission error. This further reduces the

optimal bandwidth L to the benefit of a reduced numerical complexity for the computation of the Slepian basis functions and the least-squares estimation of the Slepian basis function coefficients.

Another implication of this experiment is that it will be very difficult if not impossible to extract a low-resolution signal from high-resolution data given on a part of the sphere. As pointed out before, if a low-resolution spherical harmonic expansion is fitted to data given on a part of the sphere by least-squares, the fit of the model to the data is optimized according to the least-squares principle. Hence, the coefficients representing the low-resolution approximation will explain as much of the signal as possible at the complete set of frequencies to minimize the residual sum of squares, and therefore will not provide a good representation of the true low-resolution signal. In Sect. 5, we will investigate whether Slepian functions offer a solution to this problem and, if not, what the alternatives are.

4 Experimental setup

In this section, we discuss the setup of the numerical experiments related to the use of a Slepian basis to extract a low-resolution signal from high-resolution data given on a part of the sphere. This comprises the choice of the concentration region

and the construction of the Slepian functions, the generation of the sampled geoid, the truncation level in the inversion, and the criteria we use to assess the quality of the results, which are to be presented in Sect. 5.

4.1 Concentration region

In the numerical experiments, we consider two concentration regions: (i) a spherical cap with radius 40° centred in the Pacific Ocean and (ii) the union of the world's ocean basins as defined next. Due to the non-polar orbits of all radar altimeter satellites, caps centered on both poles are left without data coverage and therefore are not part of the concentration region. For the TOPEX/Poseidon satellite, the radii of these polar gaps are approximately 24°. For other radar altimetry missions this radius may be smaller, but in our experiments we define Ω_0 as the part of the oceans covered by TOPEX/Poseidon, i.e:

$$\Omega_0 = \Omega - \bar{\Omega}_0, \tag{17}$$

where $\bar{\Omega}_0$ is the union of Eurasia, Africa, North and South America, Antarctica, Greenland, and Australia as defined and shown individually by Simons et al. (2006, 2009) and Simons (2010), with the subsequent addition of New Guinea, Borneo, Madagascar, Sumatra, Honshu, the United Kingdom, and the further exclusion of the two polar gaps. The fractional area of this region on the unit sphere is $|\Omega_0|/(4\pi) \approx 0.67$. Because the bandwidths that we will use to construct the "true" MSL signal and its approximation are relatively small, all islands smaller than $200,000\,\mathrm{km}^2$ are in fact neglected in the localization kernel $\mathbf{D}$. The latter is computed as the difference of the localization kernel for the latitudinal belt exclusive those continents that are partly located inside the latitudinal belt minus the localization kernels for individual continental regions completely located inside the latitudinal belt.

Spatial expansions of some of the eigenfunctions of $\mathbf{D}$ that result from this procedure, for a bandwidth $L = 36$, are shown with their eigenvalues λ in Fig. 3. These are the Slepian functions that we will use as a spatiospectrally localized basis in the forthcoming analysis.

4.2 Construction of the sampled geoid and MSL signals

Samples of the geoid and MSL signals are derived from available spherical harmonic models. They are evaluated at a set of M points of a Fibonacci grid, which provides a homogeneous sampling of the region of interest, Ω_0. In order to avoid sampling errors, M is chosen sufficiently large so that the distance between the sample points does not exceed half of the shortest wavelength contained in the signal. The global geoid signal is derived from the Earth gravitational model EGM2008 (Pavlis et al. 2008) truncated at degree L_g. The MSL signal is defined as the sum of the MDT model DOT2008A (U.S. National

Geospatial-Intelligence Agency EGM Development Team 2010) and the geoid model EGM2008, both truncated at degree L_r, where $L_r > L_g$. In the experiments of Sect. 5, we choose $L_r = 48$, $L_g = 36$, $M \approx 10,000$ (if no land data are used), and $M \approx 13,000$ (otherwise). The number of control points which are used to assess the quality of the various solutions is about 55,000 randomly distributed over the target area Ω_0. The MSL that is spectrally consistent with the geoid is referred to as the "true" low-resolution MSL. It is defined as the sum of the DOT2008A MDT and the EGM2008 geoid both truncated at degree $L_g = 36$, as displayed in Fig. 4.

4.3 Optimal truncation level

To compute the MDT as the difference between the MSL (bandwidth L_r) and the geoid (bandwidth L_g) given on Ω_0, where $L_r > L_g$, we need to find a suitable representation of the MSL, which is spectrally consistent with the geoid signal. A Slepian basis with bandwidth L_g comprising $(L_g + 1)^2$ basis functions provides such a representation. However, as mentioned in Appendix A, the number of Slepian basis functions required to obtain a faithful approximation of a given signal on a subdomain of the sphere may be smaller than $(L_g + 1)^2$. Simons and Dahlen (2006) have shown that in the presence of noise the optimal number of Slepian basis functions is determined by the signal-to-noise ratio of the data from which the expansion coefficients are derived by inversion, provided this is numerically feasible. In our experiments, however, the sampled geoid and MSL signals are assumed to be noise-free and numerical considerations do come into play. Therefore, we follow another strategy to determine the optimal truncation level J_o. We evaluate the spherical harmonic expansion of the "true" low-resolution MSL signal at the M points of the Fibonacci grid (these samples serve as data) and at a set of different control points which, however, both cover the target area Ω_0. Then, we fit a Slepian basis function representation comprising J basis functions to the data by least-squares. J is varied between the Shannon number K, defined in Eq. (A-14), and the maximum number of Slepian basis functions, $(L_g + 1)^2$. The optimal number J_o is found to be the one which minimizes the rms difference between the least-squares solution and the "true" low-resolution MSL signal evaluated at the control points. In the following, the above is referred to as "the Slepian approach", whether data are used on the entire globe Ω or on a subregion Ω_0 will be clear from the context.

4.4 Quality assessment

In order to assess the quality of the Slepian approach we first compute statistics of the differences with respect to the known "true" low-resolution MSL as defined in Sect. 4.2.
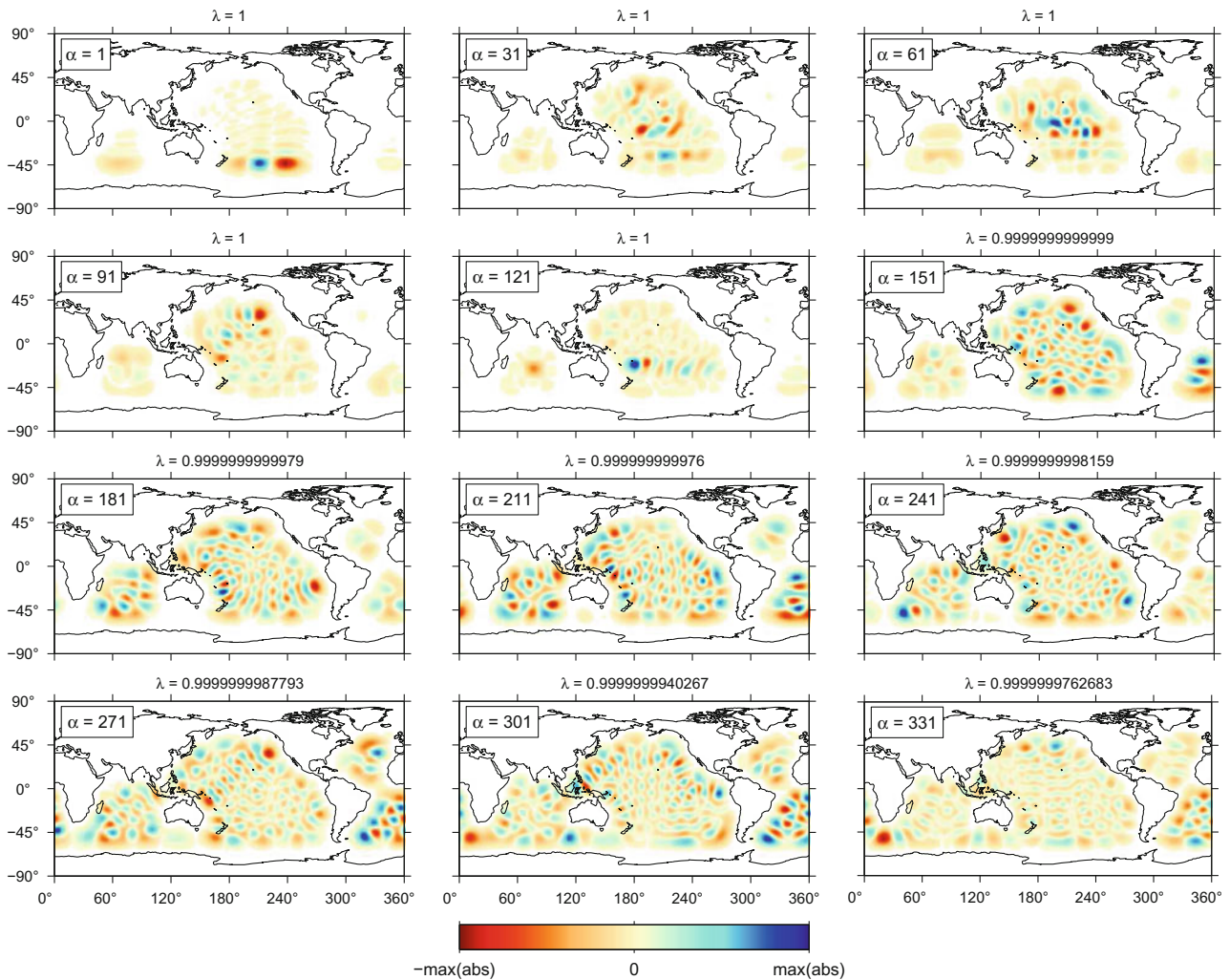
Fig. 3 Slepian functions that are bandlimited to $L = 36$ but optimally concentrated within the oceanic concentration region Ω_0 as computed via Eqs. (A-4)–(A-7). We integrated Eq. (A-7) using Gauss–Legendre quadrature over the colatitudes and analytically over longitudes. A selection of eigenfunctions g_α is shown with their concentration factors λ_α, as labeled. The rounded Shannon number, from Eq. (A-14), is $K = 918$. Every function was normalized to max(abs(function)). Values smaller than 0.01 on this scale are rendered white. The sign of an eigenfunction is arbitrary since the concentration is quadratic. Altogether the Slepian functions form a complete basis for bandlimited processes anywhere on the sphere. The first K functions provide an approximate basis for bandlimited signals concentrated in the oceanic region Ω_0, as discussed in the text

Fig. 4 The "true" low-resolution MSL signal used in the experiments to be presented in Sect. 5. It is the sum of the EGM2008 geoid and the DOT2008A mean dynamic topography both truncated at degree 36
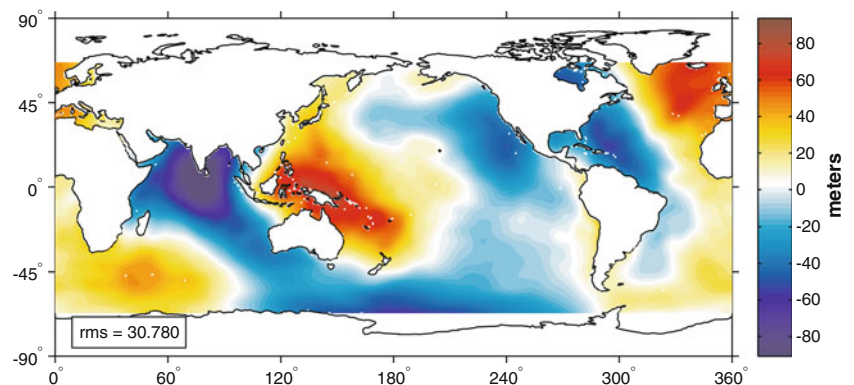
Table 1 Experiment I: statistics of the differences between "true" and approximated low-resolution MSL signal, evaluated at a set of control points in the target area Ω_0 (columns 2–5) and the rms difference between data and approximated low-resolution MSL signal in the tar-get area Ω_0 (column 6). $L_r = 48$, $L_g = 36$. The "true" low resolution MSL signal is the sum of the EGM2008 geoid and the DOT2008A mean dynamic topography both truncated at degree 36 (cf. Fig. 4). The domain Ω_0 comprises the world's oceans defined in Sect. 4.1

Method	Control points				Data points
	Rms (m)	Min (m)	Max (m)	Mean (m)	Rms difference (m)
Iterative SH, $L = L_r$, geoid data outside, Gaussian filter	0.660	−4.76	3.79	−0.017	0.499
Iterative SH, $L = L_r$, geoid data outside, ideal low-pass filter	0.071	−0.963	0.526	0.000	0.859
Direct SH, $L = L_r$, no data outside Ω_0, Gaussian filter	0.660	−4.76	3.79	−0.017	0.499
Direct SH, $L = L_r$, no data outside Ω_0, ideal low-pass filter	2.71×10^{-5}	-1.06×10^{-4}	7.27×10^{-5}	-2.27×10^{-6}	0.872
Direct SH, $L = L_g$, geoid data outside Ω_0	0.230	−1.50	2.96	0.016	0.773
Direct SH, $L = L_g$, no data outside Ω_0	0.596	−5.38	7.78	0.030	0.610
Slepian functions (bandwidth L_g), no data outside Ω_0, $J_o = 1,180$	0.508	−4.40	6.04	0.025	0.685
Slepian functions (bandwidth L_g), geoid data outside Ω_0, $J_o = 1,322$	0.230	−1.50	2.96	0.016	0.772
MSL signal in the band $37 \leq l \leq 48$	0.874	−7.15	6.46	−0.040	

We also compare the estimated Slepian representation of the low-resolution MSL with a low-resolution MSL model obtained using the iterated spherical harmonic and Gaussian smoothing approach. The latter is reported by Albertella et al. (2008) as providing the best results in their experiments. This iterative spherical harmonic approach consists of five steps:

1. A global data set is formed using samples of the high-resolution MSL signal inside the region Ω_0 and samples of the low-resolution geoid in the complementary domain $\bar{\Omega}_0$.
2. A spherical harmonic analysis using the data from step 1 is performed complete to degree L_r.
3. The solution of step 2 is synthesized at the data points in the domain $\bar{\Omega}_0$ and a new data vector is formed.
4. Steps 2 and 3 are repeated until convergence to within error.
5. A Gaussian low-pass filter with correlation length Ψ_o is applied to the final set of spherical harmonic coefficients to remove the contributions from the degrees $L_g + 1 \rightarrow L_r$ and to reduce ringing effects. In the experiments, we use an optimal correlation length, Ψ_o, of the Gaussian filter, which is empirically derived by minimizing the rms difference between the "true" low-resolution MSL signal and the smoothed MSL signal inside the domain Ω_0.

5 Experimental results and discussion

We refer to Table 1 for an overview of the statistics of the differences between the "true" and the "approximated" low-resolution MSL signal. The computations in the Slepian basis have been done using open source software provided by the second author; see http://www.frederik.net.

5.1 Experiment I

In Experiment I, the concentration region Ω_0 is identical to the oceans as defined in Sect. 4.1. Figure 5a shows the MSL in the band $37 \leq l \leq 48$, i.e. the difference between the "high-resolution" MSL signal and the "low-resolution" MSL signal. This is exactly the signal that we want to eliminate before computing a reliable MDT.

Using the iterative spherical harmonic approach (with geoid data on land in the first iteration at step 1), 10,303 iterations are required to reduce the maximum absolute residue to 0.01 m. To extract the low-resolution MSL signal, we apply a Gaussian filter with $\Psi_o = 170$ km (cf. Sect. 4.4). Evaluated at a set of random locations, the rms approximation error equals 0.660 m; pointwise errors attain extreme values of several metres (see Table 1 for more statistics). Hence, the iterative spherical harmonic approach fails to provide a representation of the low-resolution MSL signal with an accuracy of a few centimetres. Figure 5b shows a geographic map of the approximation errors. They strongly

Fig. 5 a Shows the differences between high-resolution and low-resolution mean sea level (MSL), i.e. the MSL in the spherical harmonic band $37 \leq l \leq 48$. b Shows the differences between the iterative spherical harmonic approximation to the low-resolution MSL, with geoid data over the continents and with Gaussian filtering, and the "true" low-resolution MSL shown in Fig. 4. c Shows the differences between the 1,180-term Slepian approximation of the low-resolution MSL and the "true" low-resolution MSL shown in Fig. 4. d Shows the differences between the 1,322-term Slepian approximation of the low-resolution MSL, obtained using geoid data on land, and the "true" low-resolution MSL shown in Fig. 4
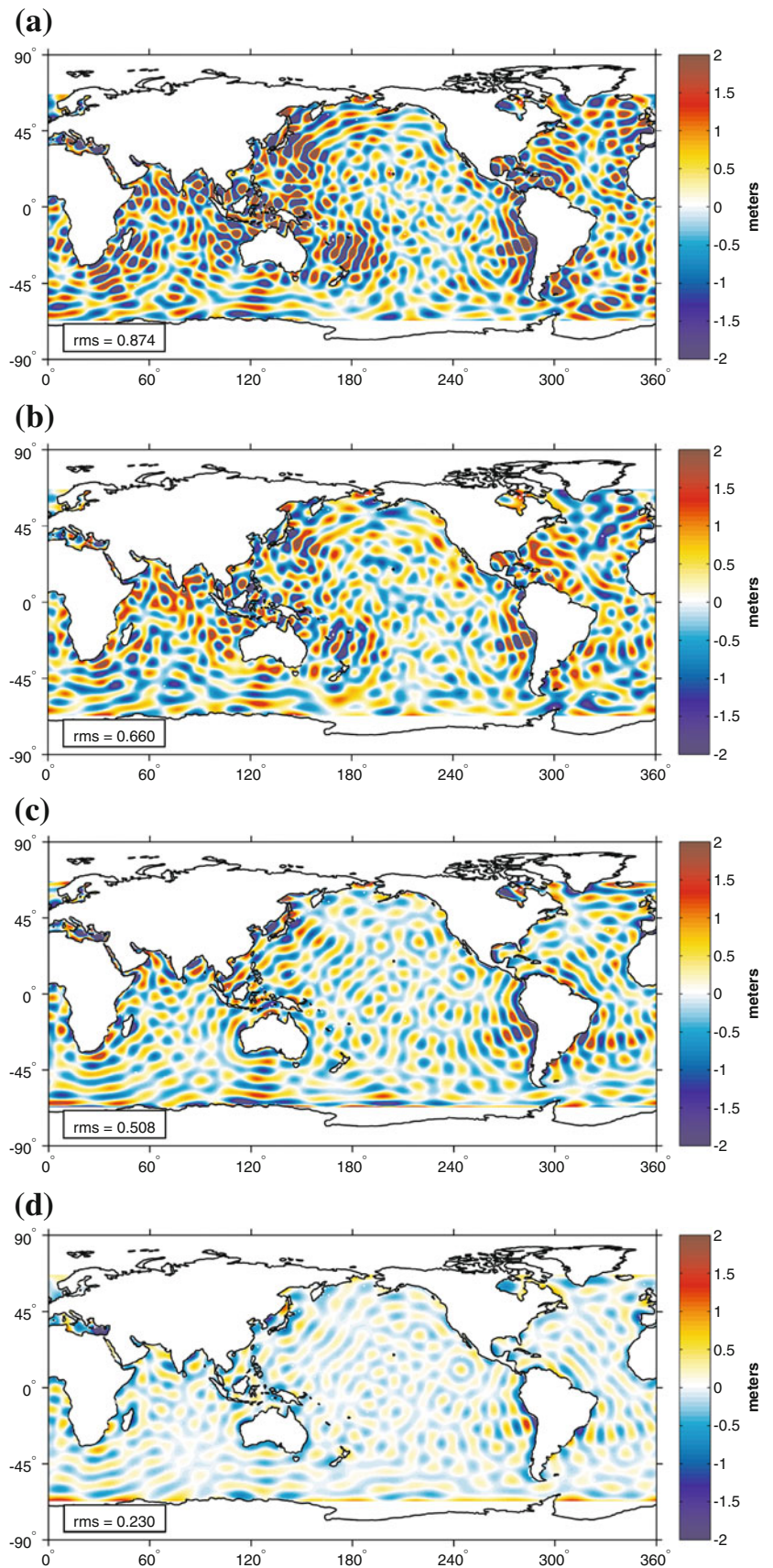
correlate with the MSL signal in the band $37 \leq l \leq 48$, Fig. 5a. This is mainly due to the weak performance of the Gaussian filter, which is unable to extract the low-resolution MSL signal from the high-resolution (bandwidth 48) least-squares solution. Too much signal from the band $37 \leq l \leq 48$ is left after filtering. This is confirmed by the rms difference between data and the approximated low-resolution MSL signal (last column of Table 1): the rms difference is 0.499 m, which is not close to the 0.874 m rms MSL signal in the band $37 \leq l \leq 48$, which indicates a leakage from the band $37 \leq l \leq 48$ into the low-resolution MSL solution.

Using the iterative spherical harmonic approach with geoid data on land but replacing the Gaussian filter with a boxcar low-pass filter in the frequency domain (i.e., hard truncation at degree $L_g = 36$) provides a much better approximation of the low-resolution MSL signal: the rms approximation error improves from 0.660 to 0.071 m (cf. Table 1). A look at the rms difference between the data and the approximated low-resolution MSL signal confirms this result: it is 0.859 m, close to the 0.87 m rms MSL signal in the band $37 \leq l \leq 48$. Hence, almost no signal in the band $37 \leq l \leq 48$ leaks into the estimated coefficients. This is explained by the fact that Ω_0 covers about 67% of the whole sphere. The main error contributors are the geoid data on land, which are used to initially allow for a global spherical harmonic analysis. Although the rms approximation error is only 0.071 m, the maximum pointwise error is about 1 m, which is far above the target accuracy of a few centimetres. Finally, we will show in Sect. 5.3 that hard truncation performs very poorly if the size of Ω_0 is a much smaller fraction of the whole sphere than the entirety of the ocean basins. Therefore, an ideal low-pass filter cannot generally be the method of choice.

For the Slepian approach that uses only MSL data over the oceans, the rms approximation error of the low-resolution MSL signal is 0.508 m. A spatial rendition of the approximation errors is shown in Fig. 5c. This figure represents the optimal solution in the sense explained in Sect. 4.3 whereby $J_o = 1,180$ Slepian basis functions. The optimal bandwidth in the sense in which it appeared in Sect. 3 for the oceans is found to be $L = 36$ (i.e., identical to the maximum degree of the global geoid model), which corresponds to 1,369 basis functions and an rmse of 0.596 m (see Table 1). We explain the fact that less than the total number of Slepian basis functions provides the smallest rms approximation error by invoking the partial cancellation of truncation error and broadband bias. At 0.596 m the rms approximation error for 1,369 Slepian basis functions is only slightly higher than the 0.508 m obtained for $J_o = 1,180$ Slepian basis functions. A positive effect of using fewer Slepian basis functions is a significant improvement of the condition number of the normal matrix: from 10^8 (with 1,369 terms) to only 10 (with 1,180 terms).

The performance of the Slepian approach is not much better than for the iterative spherical harmonic approach (rms error of 0.508 m as compared to 0.660 m). The reason for the poor performance of the Slepian approach is the presence of broadband leakage (frequency domain) and broadband bias (spatial domain). Though Slepian functions with the same bandwidth are orthogonal on Ω_0, this does not apply to Slepian functions of different bandwidths. The similarities between the error pattern shown in Fig. 5c (Slepian approach) and Fig. 5a (MSL signal in the band $37 \leq l \leq 48$) are evidence for the presence of broadband bias and leakage, as they were for the iterative spherical harmonic approach shown in Fig. 5b.

The quality of the Slepian approach improves if MSL data on the oceans are complemented by geoid data on land. Then, the solution with minimum rms approximation error is obtained with 1,322 Slepian basis functions. Figure 5d shows a geographic map of the approximation errors in that case, whose rms approximation error has been reduced from 0.508 to 0.230 m. Note, however, that this approach is almost identical to a direct least-squares fit to the global data of a spherical harmonic model complete to degree 36 (rms error 0.230 m), which can be understood from the fact that almost all Slepian basis functions are used.

A simple alternative to the approaches discussed so far is a direct least-squares fit of a spherical harmonic expansion complete to degree $L_r = 48$ to the data (no data are used outside Ω_0) followed by a hard truncation at degree $L_g = 36$. The extreme values of the low-resolution MSL approximation error are -1.06×10^{-4} and 7.27×10^{-5} m, respectively, i.e. at the sub-millimetre level and the rmse is 2.71×10^{-5} m (cf. Table 1). This approach performs by far the best for the current test setup. However, extensive simulations with different target areas Ω_0 and different bandwidths L_r (not shown here) reveal that the performance of this straightforward approach depends critically on (i) the size of the domain and (ii) the bandwidth of the signal. The smaller the domain, the larger the approximation error. Moreover, the larger the bandwidth and the smaller the target area, the larger the condition number of the normal matrix. Then, regularization is indispensable, which introduces an additional bias and reduces the quality of the solution. Therefore, this simple approach works for the oceans in combination with low-resolution MSL data as used here ($L_r = 48$), but fails for smaller target areas and/or MSL data with a higher resolution.

From the results shown in Table 1, we can also conclude that the iterative spherical harmonic approach does not depend on the data used on land provided that enough iterations are performed. This has been verified in several numerical simulations. In the limit, it is equivalent to a direct least-squares fit of a high-resolution spherical harmonic model (i.e., complete to degree $L_r = 48$) to the ocean data followed by Gaussian filtering in reaching an rmse of

0.660 m. Expressed differently, neither iteration nor information on land is needed to match the performance of the iterative spherical harmonic approach proposed by Albertella et al. (2008). From Table 1, we also conclude that the quality of the iterative spherical harmonic approach and the direct spherical harmonic approach that estimates coefficients complete to degree $L_r = 48$ is solely determined by the performance of the filter. A Gaussian filter is definitely not the preferred choice; we expect that the use of other filters may reduce the errors but will not reduce them down to the level of several centimetres.

5.2 Experiment II

Experiment I has demonstrated that the Slepian approach fails to recover the low-resolution MSL with adequate accuracy. This has been explained by broadband leakage and truncation bias. Trampert and Snieder (1996) have proposed a method to suppress leakage, which in fact uses a different cost function than the one being used in the classical least-squares solution. For details about the implementation, the setup of the numerical experiment, and the results, we refer to Appendix B. In summary, the method works perfectly for very low-resolution data, e.g., $L_r = 18$ and $L_g = 12$. However, already for moderate resolutions such as $L_r = 36$ and $L_g = 48$ and for higher resolutions, the performance is not better than the one presented in Sect. 5.1 for both the direct spherical harmonic approach with only MSL data and the Slepian approach.

The results for the direct spherical harmonic approach with only MSL data are not in line with the findings of Trampert and Snieder (1996). We explain this by the fact that Trampert and Snieder (1996) consider a non-homogeneous, but global data distribution, whereas in our experiment we lack data on the continents, which can be considered as an example of an extremely non-homogeneous global data distribution. For more details, we refer to Appendix B. We conjecture that the method of Trampert and Snieder (1996) will always fail if the diameter of the area without data is large compared to the spatial resolution we aim at. A more precise analysis of the relation between approximation error, spatial resolution, and size of the data gaps requires further research and is beyond the scope of this paper.

5.3 Experiment III

In Sect. 5.1, we found that a direct least-squares fit to data over all of the oceans of a spherical harmonic expansion complete to degree $L_r = 48$ allows an almost perfect recovery of the low-resolution (degree $L_g = 36$) MSL signal if an ideal low-pass filter in the frequency domain is applied. The question is whether this also applies to smaller areas. Furthermore, we found that for the oceans as target area, almost

all Slepian basis functions are needed to obtain a good fit to the data. Using all $(L_g + 1)^2$ Slepian functions, however, is equivalent with a spherical harmonic expansion complete to degree L_g. In that case, however, using a Slepian basis does not offer any advantage compared to spherical harmonics. Moreover, the condition number of the normal matrix increases exponentially with increasing bandwidth L_g both for Slepian functions and spherical harmonics, which makes regularization indispensable, at the cost of additional bias. If significantly less than $(L_g + 1)^2$ Slepian functions allow a good least-squares fit to the data, this would reduce the computational costs compared to spherical harmonics and would also reduce the condition number, thus making regularization superfluous. In order to answer these questions, we repeat Experiment I with a smaller target domain Ω_0, which is now a spherical cap of radius $40°$ in the Pacific Ocean centered at $210°$ longitude and $5°$ southern latitude.

Table 2 summarizes the main results. First of all, we observe that a direct least-squares fit of a spherical harmonic expansion complete to degree $L_r = 48$ followed by a truncation at degree $L_g = 36$ now has an rms of 0.369 m and thus fails to recover the low-resolution MSL signal with an accuracy of a few centimetres. This is completely different from the results of Experiment I, but in line with what could be expected based on the experiment in Sect. 3. If the size of the domain Ω_0 decreases, the distribution of the energy over the spherical harmonic coefficients is no longer preserved because the spherical harmonics are not orthogonal over Ω_0. Therefore, a hard truncation of the expansion at degree L_g does not allow to recover the original spectrum at degrees $l \leq L_g$, which explains the poor performance of this approach for domains significantly smaller than the whole sphere. The fact that this method still performs better than the iterative spherical harmonic approach and the Slepian approach is due to the fact that a spherical cap of radius $40°$ is large enough to allow for this. However the performance of this approach will further decrease if the size of Ω_0 decreases.

When using Slepian functions, we obtain the lowest rms approximation error if significantly less than the total number of basis functions, which is 1,369 for $L_g = 36$, are used. The optimal number of Slepian functions turned out to be 289 if only MSL data inside Ω_0 are used and 1,179 if the MSL data are complemented by geoid data outside Ω_0. Note that the Shannon number of the $40°$ spherical cap is $K = 161$.

The attempted recovery of a signal complete to degree $L_g = 36$ on a spherical cap of radius $40°$ from data complete to degree $L_r = 48$ is an unstable problem no matter what basis functions are used and what approach is followed. Strong regularization was necessary in all cases to obtain a solution. We always used a truncated SVD for regularization, which may not be the optimal choice due to its global character. A better choice could be a regularization scheme which constrains the MSL variance over land as used by

Table 2 Experiment III: statistics of the differences between "true" and approximated low-resolution MSL signal, evaluated at a set of control points (columns 2–5) and the rms difference between the data and the approximated low-resolution MSL signal (column 6).

$L_r = 48$, $L_g = 36$. The "true" low resolution MSL signal is the sum of the EGM2008 geoid and the DOT2008A mean dynamic topography both truncated at degree 36 (cf. Fig. 4). The domain Ω_0 is a spherical cap with radius 40°

Method	Control points				Data points
	Rms (m)	Min (m)	Max (m)	Mean (m)	Rms difference (m)
Iterative SH, $L = L_r$, geoid data outside, Gaussian filter	0.462	−1.66	2.14	0.006	0.317
Iterative SH, $L = L_r$, geoid data outside, ideal low-pass filter	0.153	−0.572	0.589	−0.001	0.525
Direct SH, $L = L_r$, no data outside Ω_0, Gaussian filter	0.462	−1.67	2.14	0.006	0.316
Direct SH, $L = L_r$, no data outside Ω_0, Ideal low-pass filter	0.369	−1.37	1.35	0.000	0.637
Direct SH, $L = L_g$, no data outside Ω_0	0.342	−3.91	1.90	0.000	0.430
Slepian functions (bandwidth L_g), no data outside Ω_0, $J_o = 289$	0.499	−3.83	4.31	0.000	0.311
Slepian functions (bandwidth L_g), geoid data outside Ω_0, $J_o = 1,179$	0.342	−3.91	1.90	0.000	0.430
MSL signal in the band $37 \leq l \leq 48$	0.874	−7.15	6.46	−0.040	

Kusche and Schrama (2005). The condition number will further increase if higher resolutions are aimed at, with the risk of a larger regularization bias, and, therefore, a reduced accuracy of the recovered low-resolution MSL signal. For completeness we want to remark that also Experiment II has been executed for the spherical cap of 40° radius. However, the main conclusions are the same as for the world's oceans.

6 Reformulation and solution of the problem

So far we defined the MSL data on the domain Ω_0 in terms of a spherical harmonic expansion complete to degree L_r. The "true" low-resolution MSL signal on the domain Ω_0 has been identified with the same spherical harmonic expansion, but now truncated at degree $L_g < L_r$, where L_g is the maximum degree of the expansion of the global geoid in spherical harmonics. Several statistics of the differences between the methods designed to recover the low-resolution MSL signal from the MSL data inside Ω_0 were used to assess the quality of the solutions. We found that more or less all methods fail in the sense that point-wise errors exceed the level of several metres, whereas in practice we would like to recover the low-resolution MSL with errors comparable to the errors in the MSL data and the geoid model, which are at the level of a few centimetres.

This definition of the "true" low-resolution MSL signal is in line with what other authors also used in similar studies (e.g. Tapley et al. 2003; Albertella et al. 2008; Bingham et al. 2008). From the viewpoint of a "globalized" MSL signal (using geoid information on land) this may be justified. If, however, the MSL is considered to be signal which is only defined on a part of the sphere, this definition needs

to be reconsidered as the results presented in Sect. 3 have shown. The reason is that on a domain smaller than the whole sphere, the degree of the spherical harmonic representation of a global signal is no longer a measure of its resolution. Usually, a spherical harmonic representation of a lower maximum degree is adequate to reproduce the data within acceptable error bounds. The smaller the size of Ω_0, the lower the maximum degree needed to represent a given signal.

Moreover, when we want to extract a low-resolution signal from high-resolution data given on a domain $\Omega_0 \subset \Omega$, we need to represent the data in a basis orthogonal on Ω_0. We also showed that not every orthogonal basis is equally well suited for this. In particular Slepian functions, which are designed to maximize the concentration of bandlimited functions in a spatial domain Ω_0, are not appropriate here, because every Slepian basis function depends on all spherical harmonics. According to Eq. (A-10), the transformation of a spherical harmonic basis (collected in a vector $\hat{\mathbf{y}}$) into the Slepian basis (collected in a vector $\mathbf{g}$) is given by

$$\mathbf{g} = \mathbf{Q}_S \, \hat{\mathbf{y}}, \tag{18}$$

where the matrix $\mathbf{Q}_S$ is a full unitary matrix. Hence, every Slepian basis function is a linear combination of *all* spherical harmonics up to the bandwidth L, and all basis functions have the same bandwidth, which is equal to the maximum degree of the spherical harmonic basis. This is why we cannot obtain a low-resolution approximation of a signal represented in a Slepian basis simply by a truncation of the Slepian representation. To obtain such an approximation requires the use of a Slepian basis with a smaller bandwidth, which involves spherical harmonics of a lower maximum degree. These low-bandwidth Slepian functions are, however, not

orthogonal to the Slepian functions representing the full signal. Moreover, the approximation of a high-resolution signal using a low-resolution (i.e., low bandwidth) Slepian basis always causes broadband leakage, which is another consequence of the fact that the Slepian transform matrix $\mathbf{Q}_S$ is a full matrix.

A suitable orthogonal basis on Ω_0 is found if the well-known Gram–Schmidt orthogonalization procedure is applied to spherical harmonics (e.g., Golub and van Loan 1996). Let L be the maximum degree of the spherical harmonic expansion. Then, the new basis functions, which are orthogonal on Ω_0, are related to the spherical harmonics as

$$q_\alpha(\hat{\mathbf{x}}) = \sum_{l=0}^{L} \sum_{m=-l}^{l} q_{lm\,\alpha} \, \hat{Y}_{lm}(\hat{\mathbf{x}}), \quad \alpha = 1, \ldots, (L+1)^2,$$
(19)

where

$$q_{lm\,\alpha} = 0 \text{ for all } l \geq \alpha.$$
(20)

In matrix-vector notation, this is written as

$$\mathbf{q} = \mathbf{Q}_{GS}\,\hat{\mathbf{y}},$$
(21)

where, due to Eq. (20), $\mathbf{Q}_{GS}$ is a lower triangular matrix. To compute the coefficients $\{q_{lm\,\alpha}\}$, we only need to compute the Cholesky decomposition of the Gram matrix of the basis functions $\{\hat{Y}_{lm}\}$, which is proportional to the matrix $\mathbf{D}$ from Eqs. (A-6)–(A-7), namely

$$\frac{1}{|\Omega_0|} \mathbf{D} = \mathbf{R}\mathbf{R}^{\mathrm{T}}.$$
(22)

Then,

$$\mathbf{Q}_{GS} = \mathbf{R}^{-1},$$
(23)

where $\mathbf{R}$ is the lower triangular Cholesky factor. A lower triangular matrix $\mathbf{Q}_{GS}$ implies that a basis function q_α depends only on spherical harmonics of degree $0 \leq l \leq \alpha - 1$. This is a property Slepian basis functions do not possess (cf. Eq. A-10). It has an important practical implication for the problem at hand. Suppose the restriction of the geoid to Ω_0 is represented in terms of spherical harmonics complete to degree L (note that commonly $L < L_g$ as shown in Sect. 3, where L_g is the maximum degree of the *global* representation of the geoid in spherical harmonics). Then,

$$s^{(g)}(\hat{\mathbf{x}}) = \sum_{l=0}^{L_g} \sum_{m=-l}^{l} s_{lm}^{(g)} \, \hat{Y}_{lm}(\hat{\mathbf{x}})$$
$$= \sum_{\alpha=1}^{N} s_\alpha^{(g)} q_\alpha(\hat{\mathbf{x}}), \hat{\mathbf{x}} \in \Omega_0, \ N = (L_g+1)^2. \quad (24)$$

In the same way, the MSL data can be written as

$$s^{(\mathrm{MSL})}(\hat{\mathbf{x}}) = \sum_{l=0}^{L_r} \sum_{m=-l}^{l} s_{lm}^{(\mathrm{MSL})} \, \hat{Y}_{lm}(\hat{\mathbf{x}})$$
$$= \sum_{\alpha=1}^{M} s_\alpha^{(\mathrm{MSL})} q_\alpha(\hat{\mathbf{x}}), \hat{\mathbf{x}} \in \Omega_0, \ M = (L_r+1)^2.$$
(25)

Due to property (20), the low-resolution MSL signal that is spectrally consistent with the restriction of the geoid to Ω_0 is

$$s_{lr}^{(\mathrm{MSL})}(\hat{\mathbf{x}}) := \sum_{\alpha=1}^{N} s_\alpha^{(\mathrm{MSL})} q_\alpha(\hat{\mathbf{x}}), \quad \hat{\mathbf{x}} \in \Omega_0.$$
(26)

Hence, the MDT can be computed as

$$s^{(\mathrm{MDT})}(\hat{\mathbf{x}}) = s_{lr}^{(\mathrm{MSL})}(\hat{\mathbf{x}}) - s^{(g)}(\hat{\mathbf{x}})$$
$$= \sum_{\alpha=1}^{N} \left(s_\alpha^{(\mathrm{MSL})} - s_\alpha^{(g)} \right) q_\alpha(\hat{\mathbf{x}}), \quad \hat{\mathbf{x}} \in \Omega_0. \quad (27)$$

Note that the coefficients $\{s_\alpha^{(\mathrm{MSL})}\}$ can be computed from the given MSL data by least-squares. Compared to a least-squares fit of a Slepian representation, the least-squares fit of a representation in the basis $\{q_\alpha\}$ does not suffer from broadband leakage, which is a direct consequence of Eq. (20).

Table 3 shows the result of Experiment I (cf. Table 1) with the difference that now the "true" low-resolution MSL is defined in the basis $\{q_\alpha\}$ with $N = 1,369$. A geographical plot of the differences between the "true" low-resolution MSL signal of Fig. 4, which is used as reference in Table 1, and the "true" low-resolution MSL signal in the orthogonal basis $\{q_\alpha\}$, which is used as reference in Table 3, is shown in Fig. 6. Note that if Ω_0 is identical to the oceans as defined in Sect. 4.1, the restriction to Ω_0 of a global geoid model complete to degree 36 requires $(36+1)^2 = 1,396$ basis functions. That is why $N = 1,369$.

For most methods, the fit to the new low-resolution MSL signal $s_{lr}^{(\mathrm{MSL})}$, Eq. (26), is not better than to the low-resolution MSL signal used in Sect. 5 (compare Table 1 with Table 3). However, a direct least-squares fit of a spherical harmonic expansion complete to degree $L_g = 36$ to the MSL data provides an almost perfect fit to $s_{lr}^{(\mathrm{MSL})}$. The same is valid for a least-squares fit of 1, 369 Slepian functions of bandwidth $L = 36$, because both solutions are equivalent. This surprising result implies that the unavoidable broadband leakage in the spherical harmonic solution and the Slepian solution, respectively, is negligible for the chosen experimental setup. That is, almost no signal from bandwidth $37 \leq l \leq 48$ leaks into the solution, though spherical harmonics are not orthogonal over Ω_0, and the Slepian solution suffers from broadband leakage as shown in Sect. 2. We suppose that this is due

Table 3 Experiment I (repeated): statistics of the differences between "true" and approximated low-resolution MSL signal, evaluated at a set of control points (columns 2–5) and the rms difference between data and approximated low-resolution MSL signal (column 6). $L_r = 48$, $L_g = 36$. The "true" low-resolution MSL is defined in Eq. (26). The domain Ω_0 comprises the world's oceans defined in Sect. 4.1

Method	Control points				Data points
	Rms (m)	Min (m)	Max (m)	Mean (m)	Rms difference (m)
Iterative SH, $L = L_r$, geoid data outside, Gaussian filter	0.537	−7.23	4.08	−0.047	0.499
Iterative SH, $L = L_r$, geoid data outside, ideal low-pass filter	0.577	−7.67	5.34	−0.030	0.859
Direct SH, $L = L_r$, no data outside Ω_0, Gaussian filter	0.537	−7.14	4.10	−0.047	0.499
Direct SH, $L = L_r$, no data outside Ω_0, Ideal low-pass filter	0.596	−7.78	5.38	−0.030	0.872
Direct SH, $L = L_g$, no data outside Ω_0	1.40×10^{-8}	-5.29×10^{-8}	5.77×10^{-8}	1.57×10^{-9}	0.610
Slepian functions (bandwidth L_g), no data outside Ω_0, $J_o = 1,369$	1.20×10^{-9}	-7.87×10^{-9}	1.19×10^{-8}	2.94×10^{-11}	0.685
Slepian functions (bandwidth L_g), geoid data outside Ω_0, $J_o = 1,322$	0.457	−6.49	4.48	−0.014	0.772
MSL signal in the band $37 \leq l \leq 48$	0.874	−7.15	6.46	−0.040	

Fig. 6 Difference between the "true" low-resolution MSL signal used in the experiments of Sect. 5 (cf. Fig. 4) and the "true" low-resolution MSL signal in the orthogonal basis $\{q_\alpha\}$
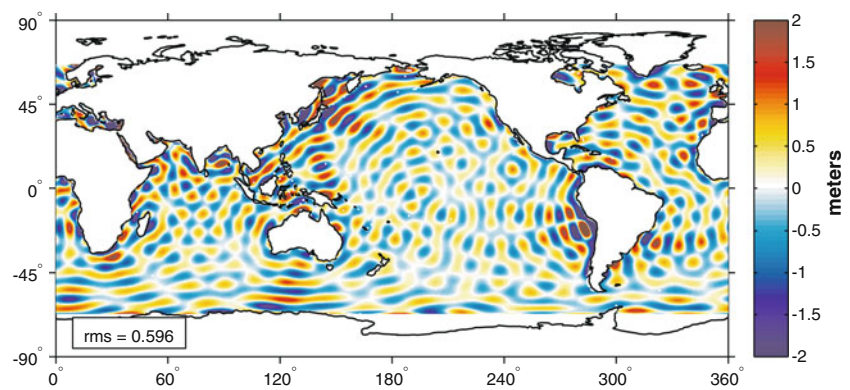
to the fact that in our experiment the "true" low-resolution MSL signal has a moderate maximum degree ($L_g = 36$) and, at the same time, the target area Ω_0 constitutes a significant part of the whole sphere. We expect, however, that broadband leakage will become significant for larger spherical harmonic degrees L_g. The latter is relevant for practical applications, because state-of-the-art global geoid models based on data of the dedicated gravity missions CHAMP, GRACE, and GOCE will be complete to degree 200–250. Therefore, we expect that the direct least-squares fit of a spherical harmonic expansion complete to degree L_g and the equivalent Slepian approach using all $(L_g + 1)^2$ basis functions do not provide an accurate enough approximation to the low-resolution MSL signal $s_{lr}^{(MSL)}$, Eq. (26), for spatial resolutions relevant in practical applications.

Instead, the correct solution to the problem at hand needs to be found as follows: suppose (i) a global geoid model is given in terms of an expansion in spherical harmonics complete to degree L_g; (ii) MSL data are given on a region Ω_0; (iii) the resolution of the MSL data is higher than that of the geoid model. Then,

1. Synthesize geoid data on Ω_0 and compute a least-squares fit of a spherical harmonic expansion complete to degree L. The corresponding spherical harmonic coefficients are denoted $s_{lm}^{(g)}$. Usually, $L < L_g$, in particular for regions with $|\Omega_0| \ll 4\pi$ as found in Sect. 3. The choice of the maximum degree L depends on how well the model should fit the data. In practical applications this depends on the target accuracy of the MDT and/or the accuracy of geoid and MSL data. If the geoid is not given in terms of a spherical harmonic expansion, but as a set of scattered or gridded data, the spherical harmonic synthesis is dropped.
2. Compute the Gram matrix for the spherical harmonics complete to degree L, Eq. (22).

3. Compute the Cholesky decomposition of the Gram matrix: $\mathbf{R}\mathbf{R}^{T}$.
4. Invert the lower triangular Cholesky factor: $\mathbf{Q}_{GS} = \mathbf{R}^{-1}$.
5. Compute the $N = (L+1)^2$ orthogonal basis functions according to Eqs. (21) and (23).
6. Compute the N coefficients $s_\alpha^{(MSL)}$ by a least-squares fit to the given MSL data or by numerical integration.

The representation of the MDT is then given as

$$s^{(MDT)}(\hat{\mathbf{x}}) = \sum_{\alpha=1}^{N} s_\alpha^{(MSL)} q_\alpha(\hat{\mathbf{x}})$$
$$- \sum_{l=0}^{L} \sum_{m=-l}^{l} s_{lm}^{(g)} \hat{Y}_{lm}(\hat{\mathbf{x}}), \quad \hat{\mathbf{x}} \in \Omega_0. \quad (28)$$

Though the correct solution to the problem at hand has been found, the question is whether it is useful in practice. The latter means that it must be possible to construct the orthogonal basis $\{q_\alpha\}$ using Gram–Schmidt for practically relevant target regions Ω_0 and spatial resolutions. Unfortunately, several test computations reveal that the construction of the orthogonal basis fails already for moderate spatial resolutions and/or medium-size target regions Ω_0. For instance, for the oceans as defined in Sect. 4.1 we were able to construct the orthogonal basis $\{q_\alpha\}$ for a maximum degree 36. However, the Cholesky factorization could not be performed anymore for a maximum degree 48. Test computations for a spherical cap of radius 40° failed already for much lower maximum degrees. These problems related to the computation of the orthogonal basis are caused by the ill-conditioning of the Gram matrix, Eq. (22). The condition number of the Gram matrix grows exponentially with (a) the maximum degree L and (b) the size of the target region Ω_0. The higher L and the smaller Ω_0, the larger the condition number. Therefore, for all practically relevant situations, we need to apply regularization to compute the Cholesky factor $\mathbf{R}$ and its inverse. We found, however, that even minimal regularization already destroys the orthogonality of the basis, and the approximation errors attain values of several metres or become even meaningless. Note that Hwang (1991) did extensive numerical experiments for the construction of an orthogonal basis for the oceans using Gram–Schmidt orthogonalization for spherical harmonics. However, successful computations of the basis functions are documented only up to a maximum degree 28.

7 Summary and conclusions

In this paper, we investigated whether bandlimited, spatially concentrated Slepian functions provide a low-resolution MSL signal that is spectrally consistent with a given geoid restricted to an incomplete part of the sphere. The recovered low-resolution MSL signal should have an accuracy that is comparable to the accuracy of the geoid and/or the MSL data, typically on the order of a few centimetres. In a number of numerical simulations, we quantified the errors of the Slepian approach and compared them with errors of alternative approaches suggested in the literature.

We showed that Slepian functions are not suited to provide a MSL signal that is spectrally consistent with a given geoid. Though they are orthogonal on the target domain, they lack the important property that the matrix, which transforms a spherical harmonic basis into the Slepian basis, is lower-triangular. Therefore, a least-squares fit of a truncated Slepian basis to given MSL data suffers from broadband leakage, hence, is unable to extract the low-resolution MSL signal with sufficient accuracy.

The iterative spherical harmonic approach proposed by Albertella et al. (2008) performs slightly worse than the Slepian approach, though the differences are not significant for practical applications. Moreover, we could show that this rather time consuming method is equivalent to a direct least-squares fit of a spherical harmonic representation to the given MSL data, which is numerically much easier to implement and less time-consuming. Several variants of the iterative spherical harmonic approach and the Slepian approach do not provide significantly better results for the world's oceans.

We also showed that the definition of the "true" low-resolution MSL signal on a domain $\Omega_0 \subset \Omega$ requires some care. A reasonable definition requires an orthogonal basis on Ω_0, which is linked to the spherical harmonic basis by a lower-triangular matrix. Then, the "true" low-resolution MSL signal is a truncated version of the series expansion of the MSL data in the orthogonal basis. This orthogonal basis can be constructed using Gram–Schmidt orthogonalization. In this study, we applied Gram–Schmidt to the spherical harmonics. We showed that the construction of the basis from spherical harmonics is a highly unstable problem in particular for small target areas and/or high-resolution geoid data. Consequently, the Gram–Schmidt orthogonalization for spherical harmonics breaks down already for geoid models of maximum degree, say, 48. The orthogonality of the constructed basis is very sensitive to any regularization, which implies that regularization is not a solution to the instability problem. Using 128-bit arithmetic may allow the Cholesky factorization for somehow higher maximum degrees, but will not be enough for the more recent high-resolution GRACE/GOCE-based geoid models. Whether the application of Gram–Schmidt to other basis functions allows the construction of an orthogonal basis for high-resolution geoid data remains to be investigated.

Appendix A: the spherical Slepian basis

Following the notation of Simons (2010) we use bold, serifed fonts (e.g., $\mathbf{f}$, $\mathbf{D}$) for vectors or matrices that are entirely composed of spectral quantities, and bold, sans-serif fonts (e.g. f, Y) for those that depend on at least one spatial variable. The colatitude of a geographical point $\hat{\mathbf{x}}$ on the surface Ω of the unit sphere is denoted by ϑ and the longitude by λ, with $0 \leq \vartheta \leq \pi$ and $0 \leq \lambda < 2\pi$. We use Ω_0 to denote a region of Ω, of area $|\Omega_0|$, within we seek to concentrate a bandlimited function of position $\hat{\mathbf{x}}$. Using orthonormalized real surface spherical harmonics $\hat{Y}_{lm}(\hat{\mathbf{x}})$, whereby

$$\int_{\Omega} \hat{Y}_{lm}(\hat{\mathbf{x}})\hat{Y}_{l'm'}(\hat{\mathbf{x}})\,d\Omega = \delta_{ll'}\delta_{mm'}, \tag{A-1}$$

we can express a square-integrable real function $f(\hat{\mathbf{x}})$ on the surface of the unit sphere as

$$f(\hat{\mathbf{x}}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} f_{lm}\hat{Y}_{lm}(\hat{\mathbf{x}}), \quad f_{lm} = \int_{\Omega} f(\hat{\mathbf{x}})\hat{Y}_{lm}(\hat{\mathbf{x}})\,d\Omega. \tag{A-2}$$

The bandlimited Slepian basis (bandwidth L) that is spatially concentrated to the region Ω_0 is the collection of functions that have no power outside the spectral interval $0 \leq l \leq L$ but as much of their power as possible concentrated within Ω_0:

$$g(\hat{\mathbf{x}}) = \sum_{l=0}^{L}\sum_{m=-l}^{l} g_{lm}\hat{Y}_{lm}(\hat{\mathbf{x}}), \quad g_{lm} = \int_{\Omega} g(\hat{\mathbf{x}})\hat{Y}_{lm}(\hat{\mathbf{x}})\,d\Omega, \tag{A-3}$$

for which the spatial concentration factor

$$\lambda = \frac{\int_{\Omega_0} g^2(\hat{\mathbf{x}})\,d\Omega}{\int_{\Omega} g^2(\hat{\mathbf{x}})\,d\Omega} = \text{maximum}. \tag{A-4}$$

Note that by convention, λ is also used to indicate the longitude. Its exact meaning will be clear from the context. Maximizing Eq. (A-4) can be achieved in the spectral domain by solving the algebraic eigenvalue problem:

$$\mathbf{D}\,\mathbf{g} = \lambda\mathbf{g}, \tag{A-5}$$

where $\mathbf{g}$ is a $(L+1)^2$-dimensional vector that represents a Slepian eigenfunction expressed in spherical harmonics, i.e. $\mathbf{g} = (g_{00}\cdots g_{lm}\cdots g_{LL})^{\mathrm{T}}$ and $\mathbf{D}$ is the $(L+1)^2 \times (L+1)^2$-dimensional spectral-basis projection operator or localization kernel:

$$\mathbf{D} = \begin{pmatrix} D_{00,00} & \ldots & D_{00,LL} \\ \vdots & & \vdots \\ D_{LL,00} & \ldots & D_{LL,LL} \end{pmatrix}, \tag{A-6}$$

whose elements $D_{lm,l'm'}$ are given by

$$D_{lm,l'm'} = \int_{\Omega_0} \hat{Y}_{lm}(\hat{\mathbf{x}})\hat{Y}_{l'm'}(\hat{\mathbf{x}})\,d\Omega. \tag{A-7}$$

As a consequence of the symmetry $\mathbf{D}^{\mathrm{T}} = \mathbf{D}$, the eigenvectors $\{\mathbf{g}_\alpha : \alpha = 1, \ldots, (L+1)^2\}$ are mutually orthogonal. When choosing them to be orthonormal, we have

$$\mathbf{g}_\alpha^{\mathrm{T}}\mathbf{g}_\beta = \delta_{\alpha\beta} \quad \text{and} \quad \mathbf{g}_\alpha^{\mathrm{T}}\mathbf{D}\,\mathbf{g}_\beta = \lambda_\alpha\delta_{\alpha\beta}, \tag{A-8}$$

where $\delta_{\alpha\beta}$ is the Kronecker delta defined as

$$\delta_{\alpha\beta} = \begin{cases} 1 & \text{if } \alpha = \beta, \\ 0 & \text{otherwise.} \end{cases} \tag{A-9}$$

The corresponding spatial Slepian functions,

$$g_\alpha(\hat{\mathbf{x}}) = \sum_{l=0}^{L}\sum_{m=-l}^{l} g_{lm\alpha}\hat{Y}_{lm}(\hat{\mathbf{x}}), \quad \alpha = 1, \ldots, (L+1)^2, \tag{A-10}$$

are orthonormal over the whole sphere Ω and orthogonal over the region Ω_0.

Rather than determining *bandlimited* functions that are concentrated in a spatial region of interest Ω_0 we may also find *spacelimited* functions $h(\hat{\mathbf{x}})$ that are concentrated in a spectral interval $0 \leq l \leq L$. Instead of Eq. (A-5), the $h(\hat{\mathbf{x}})$ satisfy the spatial-domain eigenvalue equation

$$\int_{\Omega_0} D(\hat{\mathbf{x}}, \hat{\mathbf{x}}')h(\hat{\mathbf{x}}')\,d\Omega' = \lambda h(\hat{\mathbf{x}}), \quad \text{for} \quad \hat{\mathbf{x}} \in \Omega_0, \tag{A-11}$$

whose kernel is given by the bandlimited delta function

$$D(\hat{\mathbf{x}}, \hat{\mathbf{x}}') = \sum_{l=0}^{L}\left(\frac{2l+1}{4\pi}\right)P_l(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}'), \tag{A-12}$$

and where P_l is the Legendre function of degree l. Eqs. (A-5) and (A-11) have identical eigenvalues λ. In both cases the eigenvalues are a measure of the quality of the spatiospectral concentration. In the latter case, λ expresses the spectral concentration

$$\lambda = \frac{\sum_{l=0}^{L} \sum_{m=-l}^{l} h_{lm}^2}{\sum_{l=0}^{\infty} \sum_{m=-l}^{l} h_{lm}^2} = \text{maximum}. \tag{A-13}$$

The sum of the eigenvalues, which is known as the Shannon number, K, is given by

$$K = \sum_{\alpha=1}^{(L+1)^2} \lambda_\alpha = \sum_{l=0}^{L} \sum_{m=-l}^{l} D_{lm,lm} = \int_{\Omega_0} D(\hat{\mathbf{x}}, \hat{\mathbf{x}}) \, d\Omega$$

$$= (L+1)^2 \frac{|\Omega_0|}{4\pi}. \tag{A-14}$$

From Eq. (A-4) it follows that the closer an eigenvalue λ is to 1 the better the corresponding Slepian function $g(\hat{\mathbf{x}})$ is concentrated within Ω_0. If an eigenvalue is small, i.e., $\lambda \ll 1$, the corresponding Slepian function is mostly concentrated in the domain $\bar{\Omega}_0 = \Omega - \Omega_0$. Hence, the Shannon number K, Eq. (A-14), is roughly equivalent to the *number* of well concentrated ($\lambda \approx 1$) eigenfunctions.

If the first K eigenfunctions $g_1, g_2, \ldots, g_K$ are well concentrated in the region Ω_0, the remaining eigenfunctions $g_{K+1}, g_{K+2}, \ldots, g_{(L+1)^2}$ are well concentrated in the complementary region $\bar{\Omega}_0$. From Eq. (A-14) it then follows that when the area of region Ω_0 is a small fraction of the sphere, $K \ll (L+1)^2$ holds true. By implication there are many more eigenfunctions with insignificant eigenvalues ($\lambda \approx 0$) than well-concentrated eigenfunctions with significant eigenvalues ($\lambda \approx 1$). If on the other hand the area $|\Omega_0|$ approximates the area of the sphere, $K \approx (L+1)^2$, i.e. there will be many more well concentrated than well-excluded eigenfunctions.

Bandlimited spatially concentrated Slepian functions designed for a given domain Ω_0 form a basis for bandlimited signals *anywhere* on the sphere. When ranked by decreasing eigenvalue λ, the first J members of such a Slepian basis provide efficient constructive approximations to bandlimited functions $s(\hat{\mathbf{x}})$ that are themselves spatially concentrated inside the same region Ω_0:

$$s(\hat{\mathbf{x}}) = \sum_{l=0}^{L} \sum_{m=-l}^{l} s_{lm} \hat{Y}_{lm}(\hat{\mathbf{x}}) = \sum_{\alpha=1}^{(L+1)^2} t_\alpha g_\alpha(\hat{\mathbf{x}})$$

$$\approx \sum_{\alpha=1}^{J} t_\alpha g_\alpha(\hat{\mathbf{x}}), \quad \hat{\mathbf{x}} \in \Omega_0, \tag{A-15}$$

where t_α are the Slepian expansion coefficients. Due to the characteristic flat, and then rapidly declining, shape of the eigenvalue spectrum, taking $J = K$ will provide very reasonable approximations to $s(\hat{\mathbf{x}})$ within the region Ω_0. Equality prevails globally when $J = (L+1)^2$. From an inverse modeling standpoint, the optimal (in the mean-squared error sense) truncation level J in estimating a localized signal from noisy data over the region is determined by the signal-to-noise ratio.

In the noiseless case, $J = (L+1)^2$ holds true, while in the special case of white noise contaminating a white signal, the optimal value for J is when the corresponding eigenvalue $\lambda_J = N/S$, with N the noise variance and S the signal variance. For more details, see Simons and Dahlen (2006).

In geodetic practice we are likely to have M samples of a certain signal $s(\hat{\mathbf{x}})$, which, collected in a vector $\mathbf{s}$, allows us to rewrite Eq. (A-15) as:

$$\mathbf{s} = \mathbf{Y}_L^\mathsf{T} \mathbf{s} = \mathbf{Y}_L^\mathsf{T} \mathbf{G} \mathbf{t} = \mathbf{G}^\mathsf{T} \mathbf{t}, \tag{A-16}$$

where $\mathbf{s} = (s(\hat{\mathbf{x}}_1) \cdots s(\hat{\mathbf{x}}_j) \cdots s(\hat{\mathbf{x}}_M))^\mathsf{T}$, $\mathbf{s} = (s_{00} \cdots s_{lm} \cdots s_{LL})^\mathsf{T}$, and $\mathbf{t} = (t_1 \cdots t_\alpha \cdots t_{(L+1)^2})^\mathsf{T}$, and, following Simons (2010),

$$\mathbf{Y}_L = \begin{pmatrix} \hat{Y}_{00}(\hat{\mathbf{x}}_1) & \ldots & \hat{Y}_{00}(\hat{\mathbf{x}}_j) & \ldots & \hat{Y}_{00}(\hat{\mathbf{x}}_M) \\ \vdots & & \hat{Y}_{lm}(\hat{\mathbf{x}}_j) & & \vdots \\ \hat{Y}_{LL}(\hat{\mathbf{x}}_1) & \ldots & \hat{Y}_{LL}(\hat{\mathbf{x}}_j) & \ldots & \hat{Y}_{LL}(\hat{\mathbf{x}}_M) \end{pmatrix}, \tag{A-17}$$

$$\mathbf{G} = \begin{pmatrix} g_1(\hat{\mathbf{x}}_1) & \ldots & g_1(\hat{\mathbf{x}}_j) & \ldots & g_1(\hat{\mathbf{x}}_M) \\ \vdots & & g_\alpha(\hat{\mathbf{x}}_j) & & \vdots \\ g_{(L+1)^2}(\hat{\mathbf{x}}_1) & \ldots & g_{(L+1)^2}(\hat{\mathbf{x}}_j) & \ldots & g_{(L+1)^2}(\hat{\mathbf{x}}_M) \end{pmatrix}, \tag{A-18}$$

$$\mathbf{G} = \begin{pmatrix} g_{00\,1} & \cdots & g_{00\,\alpha} & \cdots & g_{00\,(L+1)^2} \\ \vdots & & g_{lm\,\alpha} & & \vdots \\ g_{LL\,1} & \cdots & g_{LL\,\alpha} & \cdots & g_{LL\,(L+1)^2} \end{pmatrix}. \tag{A-19}$$

In this notation, Eq. (A-3) is rewritten as $\mathbf{G} = \mathbf{G}^\mathsf{T} \mathbf{Y}_L$, and the left-hand side of Eq. (A-8) becomes $\mathbf{G}^\mathsf{T} \mathbf{G} = \mathbf{I}$, the identity matrix.

Appendix B: experiment II

We start with Eq. (5), which we split into a low-degree and a high-degree portion as

$$\mathbf{w} = \mathbf{Y}_{L_g}^\mathsf{T} \mathbf{w}_{L_g} = \mathbf{Y}_L^\mathsf{T} \mathbf{w}_L + \mathbf{Y}_{\to L_g}^\mathsf{T} \mathbf{w}_{\to L_g}. \tag{B-1}$$

Instead of using the functional model

$$\mathbf{w} = \mathbf{Y}_L^\mathsf{T} \mathbf{w}_L \tag{B-2}$$

in combination with the classical least-squares cost function

$$\Phi(\mathbf{w}) = (\mathbf{w} - \mathbf{Y}_L^\mathsf{T} \mathbf{w}_L)^\mathsf{T} \mathbf{C}_w^{-1} (\mathbf{w} - \mathbf{Y}_L^\mathsf{T} \mathbf{w}_L), \tag{B-3}$$

where $\mathbf{C}_w$ is the data noise variance-covariance matrix, we use the cost function

$$\tilde{\Phi}(\mathbf{w}) = \mathbf{v}^\mathsf{T}\mathbf{C}_w{}^{-1}\mathbf{v} + \mathbf{w}_L^\mathsf{T}\mathbf{C}_L{}^{-1}\mathbf{w}_L + \mathbf{w}_{\to L_g}^\mathsf{T}\mathbf{C}_{\to L_g}{}^{-1}\mathbf{w}_{\to L_g},$$
(B-4)

where

$$\mathbf{v} := \mathbf{w} - \mathbf{Y}_L^\mathsf{T}\mathbf{w}_L - \mathbf{Y}_{\to L_g}^\mathsf{T}\mathbf{w}_{\to L_g}.$$
(B-5)

Here, $\mathbf{C}_L{}^{-1}$ and $\mathbf{C}_{\to L_g}{}^{-1}$ define weight matrices in model (spectral) space. Minimizing $\tilde{\Phi}(\mathbf{w})$ gives

$$\hat{\mathbf{w}}_L = \left(\mathbf{Y}_L\mathbf{C}^{-1}\mathbf{Y}_L^\mathsf{T} + \mathbf{C}_L{}^{-1}\right)^{-1}\mathbf{Y}_L\mathbf{C}^{-1}\mathbf{w},$$
(B-6)

with the data-space anti-leakage operator

$$\mathbf{C} = \mathbf{Y}_{\to L_g}\mathbf{C}_{\to L_g}\mathbf{Y}_{\to L_g}^\mathsf{T} + \mathbf{C}_w.$$
(B-7)

The advantage of the least-squares solution Eq. (B-6) is that it is not biased by $\mathbf{w}_{\to L_g}$ (Trampert and Snieder 1996).

In order to investigate the performance of the estimator Eq. (B-6), we use the same data as in Experiment I. We define diagonal covariance matrices in the spherical harmonic domain using Kaula's rule for degrees $l \geq 2$ and the mean variance per coefficient computed from the data for degrees 0 and 1. The diagonal covariance matrix $\mathbf{C}_L$ is propagated into the Slepian domain providing a full covariance matrix (cf. Eq. B-4). Numerical instabilities did not allow the computation of a solution for noiseless data. Therefore, zero-mean white noise with a standard deviation of 0.04 m was added to the data (0.04 is the accuracy we expect for MSL from radar altimetry, see e.g., Andersen and Knudsen 2009), and the correct noise-covariance matrix $\mathbf{C}_w$ was used.

Scaling factors for all three matrices were computed using variance component estimation (VCE). For the spherical harmonic approach, the estimated variance components are 0.99 for the data, 103.1 for the spherical harmonic coefficients complete to degree 36, and 201.6 for the spherical harmonic coefficients from degree $37 \leq l \leq 48$. For the Slepian approach, 1,346 basis functions gave the best solution. The corresponding variance components are 0.99 for the data, 2.33×10^4 for the first 1,346 Slepian functions, 3.47×10^7

for the neglected, 23 Slepian functions, and 185.3 for the spherical harmonic coefficients from degree $37 \leq l \leq 48$. Nevertheless, the estimation of the parameters turned out to be very unstable, and the approximation errors are quite large. This can be explained by the fact that we were forced to ignore the correlations among the used and neglected basis functions.

For the direct spherical harmonic approach with geoid data on land this can be explained by the Gibbs effects along the boundary of the domain Ω_0. Note that for this particular experiment the total number of observations was reduced from 13,000 to 10,000 in order to be able to solve the system of equations. For the direct spherical harmonic approach with only MSL data, there is no straightforward explanation for the poor performance. It is not caused by leakage from the signal in the band $37 \leq l \leq 48$ into the low-resolution MSL solution as shown in the last column of Table 4: the rms of the residuals is close to the MSL signal in the bandwidth $37 \leq l \leq 48$. This is not in line with the findings of Trampert and Snieder (1996). We explain the discrepancy by the different set-up of our experiment compared with the experiment considered by Trampert and Snieder (1996). The latter is based on a (moderately) non-homogeneous, but global data distribution. The lack of data on land and in the polar regions as considered in our experiment is an example of an *extremely* non-homogeneous, global data distribution. Our experiment indicates that the method of Trampert and Snieder (1996) is not well suited for this type of data distribution. We also tested different power laws to describe the signal behavior in the basis of spherical harmonics, but the quality of the results did not change significantly. We do wish to mention that when we repeated this experiment with $L_r = 18$ and $L_g = 12$, the method of Trampert and Snieder (1996) allowed us to *perfectly* recover the low-resolution MSL signal. No instabilities have been observed and we could also compute without problems a solution with noiseless data. This is in line with the results obtained with the experiment described in Sect. 3 (cf Fig. 1b): the condition number depends on the size of the

Table 4 Experiment II: statistics of the differences between "true" and approximated low-resolution MSL, evaluated at a set of control points and the rms of residuals. All solutions minimize the cost function Eq. (B-4). The "true" low resolution MSL signal is the sum of the

EGM2008 geoid and the DOT2008A mean dynamic topography both truncated at degree 36 (cf. Fig. 4). The domain Ω_0 comprises the world's oceans defined in Sect. 4.1

Method	Control points				Data points
	Rms (m)	Min (m)	Max (m)	Mean (m)	Rms of residuals (m)
Direct SH approach, cost function Eq. (B-4), noisy MSL data, No data on land	0.175	−1.26	1.05	−0.002	0.856
Direct SH approach, cost function Eq. (B-4), noisy MSL data, geoid data on land	0.080	−0.671	0.446	0.003	0.875
Slepian approach, cost function Eq. (B-4), noisy MSL data, no data on land	0.262	−2.23	3.52	0.003	0.831
MSL signal in the band $37 \leq l \leq 48$	0.874	−7.15	6.46	−0.040	

domain Ω_0 *and* on the bandwidth. For a given domain Ω_0, the condition number increases exponentially with increasing bandwidth.

References

Albertella A, Rummel R (2009) On the spectral consistency of the altimetric ocean and geoid surface: a one-dimensional example. J Geod 83:805–815

Albertella A, Sansò F, Sneeuw N (1999) Band-limited functions on a bounded spherical domain: the Slepian problem on the sphere. J Geod 73:436–447

Albertella A, Savcenko R, Bosch W, Rummel R (2008) Dynamic Ocean Topography—The Geodetic Approach. Technical Report 27, Institut für Astronomische und Physikalische Geodäsie, Forschungseinrichtung Satellitengeodäsie, München

Andersen OB, Knudsen P (2009) DNSC08 mean sea surface and mean dynamic topography models. J Geophys Res (Oceans) 114(C13):C11001. doi:10.1029/2008JC005179

Bingham RJ, Haines K, Hughes CW (2008) Calculating the ocean's mean dynamic topography from a mean sea surface and a geoid. J Atmospheric Ocean Technol 25:1808–1822. doi:10.1175/2008JTECHO568.1

Boschi L, Dziewonski AM (1999) High- and low-resolution images of the Earth's mantle: Implications of different approaches to tomographic modeling. J Geophys Res 104(B11):25567–25594. doi:10.1029/1999JB900166

Dahlen FA, Simons FJ (2008) Spectral estimation on a sphere in geophysics and cosmology. Geophys J Int 174:774–807. doi:10.1111/j.1365-246X.2008.03854.x

Freeden W, Michel V (1999) Constructive approximation and numerical methods in geodetic research today—an attempt at a categorization based on an uncertainty principle. J Geod 73(9):452–465

Golub G, Loan Cvan (1996) Matrix computations, 3rd edn. Johns Hopkins, Baltimore

González Á (2010) Measurement of areas on a sphere using Fibonacci and latitude–longitude lattices. Math Geosci 42(1):49–64

Grünbaum FA, Longhi L, Perlstadt M (1982) Differential operators commuting with finite convolution integral operators: some non-Abelian examples. SIAM J Appl Math 42(5):941–955

Guilloux F, Faÿ G, Cardoso J-F (2009) Practical wavalet design on the sphere. Appl Comput Harmon Anal 26:143–160. doi:10.1016/j.acha.2008.03.003

Hwang C (1991) Orthogonal functions over the oceans and applications to the determination of orbit error, geoid and sea surface topography from satellite altimetry. Technical Report 414, Department of Geodetic Science and Surveying, The Ohio State University, Columbus

Hwang C (1993) Spectral analysis using orthonormal functions with a case study on the sea surface topography. Geophys J Int 115:1148–1160. doi:10.1111/j.1365-246X.1993.tb01517.x

Jayne SR (2006) Circulation of the North Atlantic Ocean from altimetry and the gravity recovery and climate experiment geoid. J Geophys Res 111:C03005. doi:10.1029/2005JC003128

Jekeli C (1981) Alternative methods to smooth the Earth's gravity field. Report 327, Department of Geodetic Science and Surveying, The Ohio State University, Columbus, Ohio

Kusche J, Schrama E (2005) Surface mass redistribution inversion from global GPS deformation and gravity recovery and climate experiment (GRACE) gravity data. J Geophys Res 110. doi:10.1029/2004JB003556

Landau HJ, Pollak HO (1961) Prolate spheroidal wave functions, Fourier analysis and uncertainty-II. Bell Syst Tech J 40(1):65–84

Landau HJ, Pollak HO (1962) Prolate spheroidal wave functions, Fourier analysis and uncertainty-III: the dimension of the space of essentially time- and band-limited signals. Bell Syst Tech J 41(4):1295–1363

Losch M, Sloyan BM, Schröter J, Sneeuw N (2002) Box inverse models, altimetry and the geoid: Problems with the omission error. J Geophys Res 107:3078. doi:10.1029/2001JC000855

Pail R, Plank G, Schuh W-D (2001) Spatially restricted data distributions on the sphere: the method of orthonormalized functions and applications. J Geod 75:44–56

Pavlis N, Holmes S, Kenyon S, Factor J (2008) An earth gravitational model to degree 2160: Egm2008, 2008. http://earth-info.nga.mil/GandG/wgs84/gravitymod/egm2008/first_release.html presented at the 2008 General Assembly of the European Geosciences Union, Vienna, Austria, April 13–18

Schachtschneider R, Holschneider M, Mandea M (2010) Error distribution in regional inversion of potential field data. Geophys J Int 181:1428–1440. doi:10.1111/j.1365-246X.2010.04598.x

Schott J-J, Thébault E (2011) Modelling the earth's magnetic field from global to regional scales. In: Mandea M, Korte M (eds) Geomagnetic Observations and Models. IAGA Special Sopron Book Ser., vol. 5. Springer, Berlin

Simons FJ (2010) Slepian functions and their use in signal estimation and spectral analysis. In: Freeden W, Nashed MZ, Sonar T (eds) Handbook of Geomathematics, vol. 30. Springer, Berlin, pp 891–923. doi:10.1007/978-3-642-01546-5_30

Simons FJ, Dahlen FA (2006) Spherical Slepian functions and the polar gap in geodesy. Geophys J Int 166:1039–1061. doi:10.1111/j.1365-246X.2006.03065.x

Simons FJ, Dahlen FA, Wieczorek MA (2006) Spatiospectral concentration on a sphere. SIAM Rev 48(3):504–536. doi:10.1137/S0036144504445765

Simons FJ, Hawthorne JC, Beggan CD (2009) Efficient analysis and representation of geophysical processes using localized spherical basis functions. In: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, vol. 7446, Aug. 2009. doi:10.1117/12.825730

Simons FJ, Wang DV (2011) Spatiospectral concentration in the Cartesian plane. Int. J. Geomath. doi:10.1007/s13137-011-0016-z

Slepian D, Pollak HO (1961) Prolate spheroidal wave functions, Fourier analysis and uncertainty-I. Bell Syst Tech J 40(1):43–63

Tapley BD, Chambers DP, Bettadpur S, Ries JC (2003) Large scale ocean circulation from the GRACE GGM01 Geoid. Geophys Res Lett 30(22):2200001

Trampert J, Snieder R (1996) Model estimations biased by truncated expansions: possible artifacts in seismic tomography. Science 271(5253):1257–1260. doi:10.1126/science.271.5253.1257

US National Geospatial-Intelligence Agency EGM Development Team (2010) Egm2008-files & products for oceanographic applications. http://earth-info.nga.mil/GandG/wgs84/gravitymod/egm2008/oceano.html Accessed July, 2010

Vianna ML, Menezes VV, Chambers DP (2007) A high resolution satellite-only GRACE-based mean dynamic topography of the South Atlantic Ocean. Geophys Res Lett 34:24604. doi:10.1029/2007GL031912

Wieczorek MA, Simons FJ (2007) Minimum-variance spectral analysis on the sphere. J Fourier Anal Appl 13(6):665–692. doi:10.1007/s00041-006-6904-1

Wunsch C, Gaposchkin EM (1980) On using satellite altimetry to determine the general circulation of the oceans with application to geoid improvement. Rev Geophys 18:725–745. doi:10.1029/RG018i004p00725

Xu P (1998) Truncated SVD methods for discrete linear ill-posed problems. Geophys J Int 135(2):505–541