

# Optimized discrete wavelet transforms in the cubed sphere with the lifting scheme—implications for global finite-frequency tomography

Sébastien Chevrot,<sup>1</sup> Roland Martin<sup>2</sup> and Dimitri Komatitsch<sup>3</sup>

<sup>1</sup>IRAP, CNRS UMR 5277, Université Toulouse 3 Paul Sabatier, Observatoire Midi-Pyrénées, 14 avenue Edouard Belin, F-31400 Toulouse, France.  
E-mail: sebastien.chevrot@irap.omp.eu

<sup>2</sup>GET, CNRS UMR 5563, Université Toulouse 3 Paul Sabatier, Observatoire Midi-Pyrénées, 14 avenue Edouard Belin, F-31400 Toulouse, France

<sup>3</sup>LMA, CNRS UPR 7051, Université Aix-Marseille, Centrale Marseille, F-13402 Marseille Cedex 20, France

Accepted 2012 September 18. Received 2012 September 14; in original form 2012 June 13

## SUMMARY

Wavelets are extremely powerful to compress the information contained in finite-frequency sensitivity kernels and tomographic models. This interesting property opens the perspective of reducing the size of global tomographic inverse problems by one to two orders of magnitude. However, introducing wavelets into global tomographic problems raises the problem of computing fast wavelet transforms in spherical geometry. Using a Cartesian cubed sphere mapping, which grids the surface of the sphere with six blocks or ‘chunks’, we define a new algorithm to implement fast wavelet transforms with the lifting scheme. This algorithm is simple and flexible, and can handle any family of discrete orthogonal or bi-orthogonal wavelets. Since wavelet coefficients are local in space and scale, aliasing effects resulting from a parametrization with global functions such as spherical harmonics are avoided. The sparsity of tomographic models expanded in wavelet bases implies that it is possible to exploit the power of compressed sensing to retrieve Earth’s internal structures optimally. This approach involves minimizing a combination of a  $\ell_2$  norm for data residuals and a  $\ell_1$  norm for model wavelet coefficients, which can be achieved through relatively minor modifications of the algorithms that are currently used to solve the tomographic inverse problem.

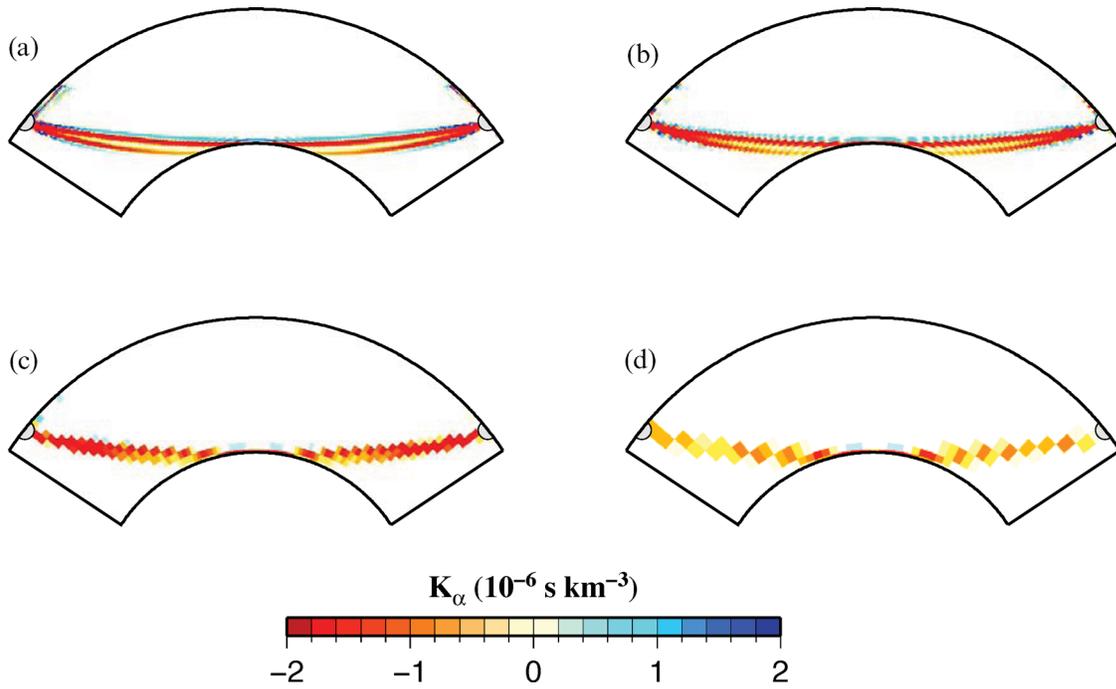
**Key words:** Inverse theory; Tomography; Seismic tomography; Computational seismology.

## 1 INTRODUCTION

In the last decades, finite-frequency effects have started to be accounted for in seismic tomography (e.g. Spetzler & Trampert 2003; Montelli *et al.* 2004a). Since Born–Fréchet kernels describe finite frequency and wavefront healing effects on traveltimes (Hung *et al.* 2000, 2001) in contrast to ray theory, their introduction in seismic tomography fuelled great hopes about improving the resolution of both regional and global tomographic models. However, these expectations were soon shown to be over-optimistic, the tomographic models obtained with finite-frequency theory being statistically similar to those obtained with asymptotic ray theory (e.g. Montelli *et al.* 2004b; Trampert & Spetzler 2006).

The reason why finite-frequency theory gave so far results similar to ray theory is simple: in order to be numerically accurate, sensitivity kernels have to be computed on a very fine grid. These are the kernels that are always shown in scientific publications. The sensitivity kernel for a traveltimes measured by cross-correlation with a synthetic seismogram is a banana-shaped region surrounding the geometrical ray (Dahlen *et al.* 2000). Fig. 1(a) shows such a finite-frequency kernel for a Pdiff wave with a dominant period of 2 s recorded at a distance of 103°. This kernel has been calculated in a fine grid, for a cell size of 0.35°, using a database

of strain Green’s functions pre-computed with the Direct Solution Method (Fuji *et al.* 2012). However, to keep the inverse problem tractable, the kernels that are effectively used in tomographic inversions are always projected on much coarser tomographic grids. The results of such projections are shown in Figs 1(b)–(d), where the same kernel as in Fig. 1(a) has been projected on coarser regular grids with block sizes of 0.7°, 1.4° and 2.8°. For blocks as small as  $\sim 3^\circ$ , which are smaller than those used classically in global tomography, the hole in the kernel and the second Fresnel zone are no longer present; the projected kernels look like fat rays. There is a good reason for this, which is also well understood: the Fermat principle tells us that when we sum the contributions of all the individual paths that contribute to an observed waveform, only those that are close to the stationary phase path are important. The others have a phase that varies rapidly and thus interfere destructively. When integrating kernels over blocks with a size comparable to the width of the Fresnel zone, we thus expect to recover the results of the stationary phase approximation, namely ray theory. Of course, there is nothing wrong with using finite-frequency theory and a coarse tomographic grid, but it is just a very inefficient way of implementing ray theory tomography. Another factor that has a detrimental effect on tomographic models comes from the regularization constraints (damping and smoothing) that are introduced



**Figure 1.** Sensitivity kernel for the traveltime of a 2 s Pdiff wave recorded at an epicentral distance of  $103^\circ$ , projected in a grid with  $0.35^\circ$  (a),  $0.70^\circ$  (b),  $1.40^\circ$  (c) and  $2.80^\circ$  (d) cells.

to stabilize the inversion. In many cases, and especially in global tomographic studies, the resulting loss in resolution is such that it is not possible to resolve structures smaller than the size of Fresnel volumes, meaning that considering 3-D sensitivity kernels in the inversion is simply pointless. This suggests that previous global finite-frequency tomographic studies found models almost identical to those obtained with ray theory because of the utilization of coarse tomographic grids and/or of an insufficient resolution potential to benefit from accounting for finite-frequency effects on seismic traveltimes.

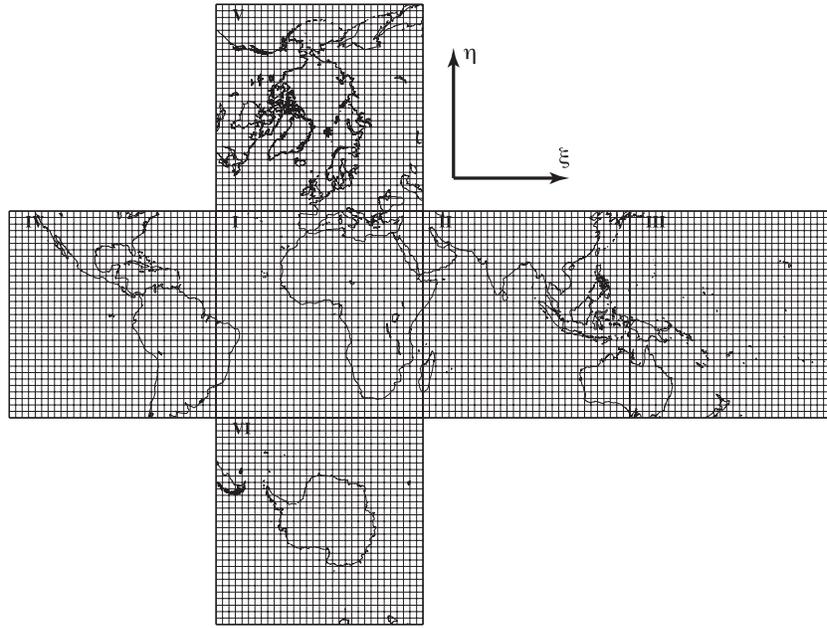
These simple considerations suggest that in order to improve the resolution in finite-frequency tomography, it is necessary to use finer parametrizations than used in the past, capable of keeping the detailed structural sensitivity to perturbations of seismic velocities. Of course, this is not a sufficient condition since resolution will also depend on ray coverage and data quality, for example. In any case, using finer tomographic grids implies considering much larger parameter spaces and consequently solving much larger tomographic inverse problems. Recognizing this problem, Chevrot & Zhao (2007) proposed to project sensitivity kernels on a basis of discrete Haar wavelets in a 3-D Cartesian grid. This allowed them to significantly compress the number of coefficients necessary to describe the information contained in 3-D finite-frequency kernels. Using a conservative compression ratio of 8, they obtained a reconstruction error around 2 per cent. In this study, we explore the potential of using discrete wavelet bases in global tomography. This raises the new problem of computing discrete wavelet transforms in spherical geometry. This problem has been addressed previously by Schröder & Sweldens (1995) using a tessellation of the sphere with triangles. In contrast, our method relies on the so-called ‘cubed sphere’ mapping (Sadourny 1972; Ronchi *et al.* 1996; Komatitsch & Tromp 2002), which leads to a much simpler algorithm and better performance. While a similar construction has been used by

Simons *et al.* (2011) to compute 2-D wavelet transform on the surface of the cubed sphere, our algorithm is different and much easier to implement because it does not require to consider overlapping ‘superchunks’, which are obtained by extending the original chunks in the cubed sphere by 50 per cent. Our algorithm builds on the power, efficiency and simplicity of the lifting scheme (Sweldens 1995) in Cartesian coordinates, which has been adapted to spherical geometry.

The paper is organized as follows: We first describe the construction of the cubed sphere, and the forward and inverse mappings between spherical coordinates and cubed-sphere coordinates in Section 2. In Section 3, we present the lifting scheme algorithm on the cubed sphere surface and the principle of compression by thresholding wavelet coefficients. We then show the compression rates that can be attained for Earth’s topography and for tomographic mantle model S40RTS (Ritsema *et al.* 2011) at 200 km depth. The algorithm is then generalized to the 3-D case in Section 4 by adding an additional wavelet transform along the radial or vertical dimension. We demonstrate that using bi-orthogonal Cohen–Daubechies–Feauveau (CDF) wavelets (Cohen *et al.* 1992), we obtain very high compression rates on both 3-D finite-frequency kernels and tomographic models. We then explore in Section 5 the potential of exploiting the compact or sparse representations of such basic ingredients of finite-frequency tomography to optimize efficiency and resolution in future global tomographic studies.

## 2 CONSTRUCTION OF THE CUBED SPHERE

The cubed sphere is constructed by inflating a cube to make it fit the shape of a sphere. Following Ronchi *et al.* (1996), we define the angular coordinates  $\xi$  and  $\eta$  that span the interval  $[-\pi/4, \pi/4]$  on each face of the cubed sphere. The Cartesian coordinates of a point



**Figure 2.** Construction of the cubed sphere and numbering of its six faces. The spherical surface is mapped to the six faces of a cube shown here in plane view. Each face is represented by an equidistant mesh with surface coordinates  $\xi$  and  $\eta$  in the interval  $[-\pi/4, \pi/4]$ . In this example, each face has  $2^{2N}$  points, with  $N = 5$ .

$(\xi, \eta, r)$ , where  $r$  is the radial coordinate, are given by

$$(x, y, z) = \begin{cases} r(1, \tan \xi, \tan \eta)/s & \text{if } k = 1, \\ r(-\tan \xi, 1, \tan \eta)/s & \text{if } k = 2, \\ r(-1, -\tan \xi, \tan \eta)/s & \text{if } k = 3, \\ r(\tan \xi, -1, \tan \eta)/s & \text{if } k = 4, \\ r(-\tan \eta, \tan \xi, 1)/s & \text{if } k = 5, \\ r(\tan \xi, \tan \eta, -1)/s & \text{if } k = 6, \end{cases} \quad (1)$$

where  $s = \sqrt{1 + \tan^2 \eta + \tan^2 \xi}$  and  $k$  is the face number. This cubed sphere transformation maps the spherical surface of the Earth on the six faces of the cube with  $2^N \times 2^N$  elements, as shown in Fig. 2. The total number of elements for the whole spherical surface is thus  $6 \times 2^{2N}$ . Note that the  $\xi$  and  $\eta$  coordinate axes are common to the six faces of the cube in the 2-D plane. However, each face has its own origin defined on its lower left corner.

Using conventions (1) and defining  $t = \max(|x|, |y|, |z|)$ , the inverse mapping is

$$(\xi, \eta, k) = \begin{cases} [\tan^{-1}(y/x), \tan^{-1}(z/x), 1] & \text{if } t = x, \\ [\tan^{-1}(-x/y), \tan^{-1}(z/y), 2] & \text{if } t = y, \\ [\tan^{-1}(y/x), \tan^{-1}(-z/x), 3] & \text{if } t = -x, \\ [\tan^{-1}(-x/y), \tan^{-1}(-z/y), 4] & \text{if } t = -y, \\ [\tan^{-1}(y/z), \tan^{-1}(-x/z), 5] & \text{if } t = z, \\ [\tan^{-1}(-y/z), \tan^{-1}(-x/z), 6] & \text{if } t = -z. \end{cases} \quad (2)$$

### 3 THE LIFTING SCHEME

The lifting scheme is an algorithm that reduces the number of operations to compute discrete wavelet transforms. It also reduces the memory required for its implementation, since all operations are performed in place, within the input array data. For a tutorial on

the lifting scheme, we refer the reader, for instance, to Sweldens & Schröder (1996) and the textbook ‘Ripples in Mathematics’ (Jensen & la Cour-Harbo 2001). Let us only give a brief outline below.

To illustrate the principle of the lifting scheme, let us consider the simple Haar wavelet transform Haar (1910). The Haar transform consists in replacing two neighbouring samples  $s_{n,2l}$  and  $s_{n,2l+1}$  of an input signal at scale  $n$  with  $2^n$  samples by their average and difference:

$$s_{n-1,l} = \frac{s_{n,2l+1} + s_{n,2l}}{2}, \quad (3)$$

$$d_{n-1,l} = s_{n,2l+1} - s_{n,2l}. \quad (4)$$

After applying these operations to the  $2^n$  samples of the input signal, we are left with  $2^{n-1}$  averages  $s_{n-1,l}$  and  $2^{n-1}$  differences  $d_{n-1,l}$ . We can think of the averages  $s_{n-1}$  as a coarser representation of the input signal  $s_n$  and of the differences  $d_{n-1}$  as the detail information needed to recover the original signal from its coarser representation. If there is some correlation in the input signal, neighbouring samples will have similar values and difference or detail coefficients will be much smaller than the original signal values. This important property implies that it is possible to obtain a sparse representation of the input signal, with a much smaller number of coefficients. This wavelet transform can be repeated iteratively at scales up to  $n$  before running out of samples, after which we obtain a representation of the signal at different scales  $j$  with  $0 \leq j \leq n-1$ , each with  $2^j$  coefficients.

The power of the lifting scheme is to perform the same wavelet transform in-place by over-writing the values at locations  $s_{n,2l+1}$  and  $s_{n,2l}$  by the average  $s_{n-1,l}$  and difference  $d_{n-1,l}$ . This can be done with an algorithm involving three steps: split, predict and update.

#### (i) Split

This stage involves splitting the signal into a set of samples  $s_{2l}$  with even indices and samples  $s_{2l+1}$  with odd indices. Each set contains half as many samples as the input signal.

(ii) Predict

Given the sets of coefficients with even indices at scale  $j$ , we predict the values of the coefficients with odd indices at scale  $j - 1$ . For the Haar wavelet, this step is

$$d_{j-1,l} = s_{j,2l+1} - s_{j,2l}. \tag{5}$$

(iii) Update

The update step imposes that the coarser signal has the same average value as the input signal:

$$s_{j-1,l} = s_{j,2l} + \frac{d_{j-1,l}}{2}. \tag{6}$$

The cost of computing a wavelet transform with the lifting scheme is proportional to the number of signal samples  $N$ . It compares favourably with the fast Fourier transform whose cost is proportional to  $N \log N$ . Since computations are performed in-place, it is also optimal in terms of memory requirements.

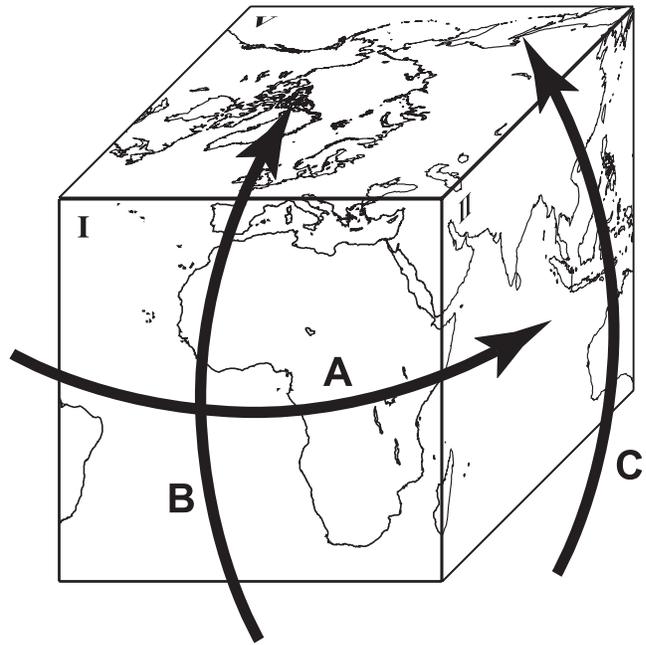
The Haar transform uses a predictor that eliminates zeroth-order correlation (i.e. the average). The update operator preserves the average in the coarser signal. It is possible to construct predict and update operators that preserve higher order moments; such a construction leads to the so-called CDF wavelets (Cohen *et al.* 1992). For example, CDF(2, 2) wavelets preserve the average and first moment, meaning that detail or wavelet coefficients will be zero if the signal is linear. These wavelets lead to a linear approximation of the signal at different scales. Similarly, CDF(4, 4) wavelets will lead to a cubic polynomial representation of the signal at different scales. To implement these wavelet transforms, one simply needs to change the lifting and filter coefficients that are used to implement the predict and update operators. It is relatively easy to compute them for any type of CDF wavelets (Fernández *et al.* 1996; Jensen & la Cour-Harbo 2001). These coefficients are given by Uytterhoeven *et al.* (1997) for CDF wavelets up to order 6.

#### 4 LIFTING SCHEME IN THE CUBED SPHERE

At this point, we have mapped a spherical surface to the six faces of a cube. While it is easy to implement a lifting scheme in the Cartesian domain associated to each chunk separately, doing so introduces spurious artefacts along the edges of the chunks, as pointed out by Simons *et al.* (2011). To overcome this problem, these authors proposed an algorithm involving partially overlapping so-called ‘superchunks’ that are obtained by extending the dimensions of the normal chunks by 50 per cent along the  $\xi$  and  $\eta$  directions. While this algorithm is a viable solution to implement discrete wavelet transforms in the cubed sphere, we believe that using ‘superchunks’ leads to unnecessary and complicated bookkeeping. In addition, the size of the ‘superchunks’ limits the maximum scale that can be handled by the wavelet transform. Let us now introduce a new algorithm to implement the lifting scheme on the cubed sphere that avoids the use of these ‘superchunks’ and yet is free of any artefact around the edges of the cube faces.

##### 4.1 Principle

The principle of our algorithm is summarized in Fig. 3. The idea is to perform 1-D wavelet transforms along three different paths, named A, B and C in Fig. 3. Each path involves a sequence of four



**Figure 3.** Basic principle of the lifting scheme in the cubed sphere. One-dimensional wavelet transforms are performed sequentially along paths A, B and C and recombined to get a 2-D wavelet transform on the cubed sphere.

contiguous faces of the cube, following the order:

$$\begin{aligned} A &\rightarrow (I, II, III, IV), \\ B &\rightarrow (I, V, III, VI), \\ C &\rightarrow (II, V, IV, VI). \end{aligned} \tag{7}$$

Using conventions (1), the directions along which the wavelet transform operates on each face are:

$$\begin{aligned} A &\rightarrow [+ \xi (I), + \xi (II), + \xi (III), + \xi (IV)], \\ B &\rightarrow [+ \eta (I), + \eta (V), - \eta (III), + \eta (VI)], \\ C &\rightarrow [+ \eta (II), - \xi (V), - \eta (IV), + \xi (VI)]. \end{aligned} \tag{8}$$

To be more specific, for each path, we construct a 2-D domain from the sequences of four contiguous faces given in (7) sampled along the directions (8). For each path, we thus obtain a 2-D domain with  $4 \times 2^N$  points along the  $X$  direction and  $2^N$  points along the  $Y$  direction. The 1-D wavelet transforms are then performed sequentially along each of the  $2^N$  rows of the 2-D domain, using periodic boundary conditions. These domains are mapped back to the cubed sphere surface after completion of each wavelet transform on a particular path. As can be seen from (8), each face of the cubed sphere is sampled twice along the  $\xi$  and  $\eta$  directions after completion of the wavelet transform along the three paths. The inverse wavelet transform is simply obtained by applying all the operations in reverse order. If all the wavelet coefficients are kept, the reconstruction is exact. However, note that some wavelet coefficients encode the difference between scaling and wavelet coefficients when the wavelet support cross some of the edges of our cubed sphere construction. Since we are not interested in the physical interpretation of these wavelet coefficients but rather in the possibility to exploit discrete wavelets to compress global fields expressed on the surface of the Earth, this is not an issue for our purpose.

## 4.2 Compression

A wavelet transform leads to a representation of a volume of data with a number of coefficients that is equal to the number of input data. However, in most cases, a large number of wavelet coefficients are very small and can thus be neglected without losing any significant amount of information. We will demonstrate in the following that both sensitivity kernels and tomographic models indeed have compact or sparse representations in the wavelet domain. This is the basic motivation for formulating the tomographic inverse problem in a discrete basis of orthogonal or bi-orthogonal wavelets.

With orthogonal wavelets such as those of Haar or Daubechies, the wavelet transform preserves energy. For an input signal  $s$  of length  $N$ , this property can be written

$$\sum_{n=1}^N s_n^2 = \sum_{n=1}^N w_n^2, \quad (9)$$

where the  $w_n$  are the coefficients of the wavelet expansion of signal  $s$  in the wavelet basis. This property is very convenient for quantizing the compression in the wavelet domain. The compression algorithm using the  $\ell_2$  norm involves three steps (Stollnitz *et al.* 1995):

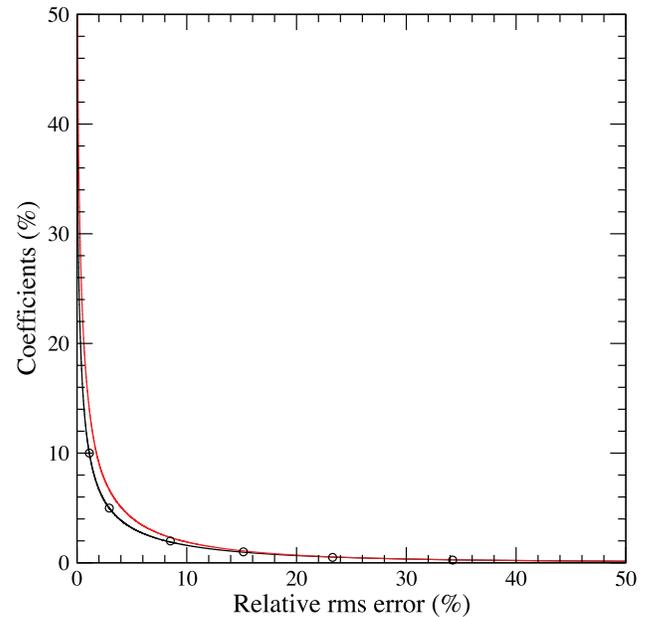
- (i) Compute the wavelet coefficients representing the input signal in the wavelet basis.
- (ii) Sort these coefficients in order of decreasing absolute values.
- (iii) Determine the error as a function of the number of wavelet coefficients kept for the reconstruction, taking the coefficients sorted by decreasing absolute values.

If  $M$  is the number of wavelet coefficients kept for the reconstruction, the reconstruction error is simply given by

$$E(M) = \sum_{n=M+1}^N w_n^2, \quad (10)$$

In contrast, bi-orthogonal wavelets such as the CDF wavelets do not preserve energy. However, energy in the CDF wavelet coefficients deviates by only a few percents so that in practice CDF wavelets can be considered as nearly orthogonal. To give an idea of the accuracy of this approximation, let us compare the exact reconstruction errors with the approximate errors given by (10) on global tomographic mantle model S40RTS (Ritsema *et al.* 2011) at 200 km depth with CDF(2, 2) wavelets. This comparison (Fig. 4) shows that the approximate errors (solid black line) are very close to the exact errors (solid circles), which demonstrates that we can indeed use (10) to compute reconstruction errors for CDF wavelets with very good accuracy.

For comparison, we have also computed the compression curve obtained with the Daubechies wavelets D4. This wavelet has the same number of vanishing moments as the CDF(2, 2) wavelets. We note that while this orthogonal wavelet transform is more complicated to implement than the bi-orthogonal CDF (2, 2) wavelet because it involves two update steps instead of one in the case of CDF(2, 2) wavelets, it clearly under-performs CDF(2, 2). Indeed, the superiority of bi-orthogonal wavelets compared to orthogonal wavelets to compress signals is now well established (e.g. Usevitch 2001), which more than compensates for their disadvantage of not preserving energy. In fact, bi-orthogonal wavelets are now the standard to compress images and are used, for instance, in the JPEG2000 encoder. In the following, we will thus use bi-orthogonal CDF wavelets and Haar wavelets.



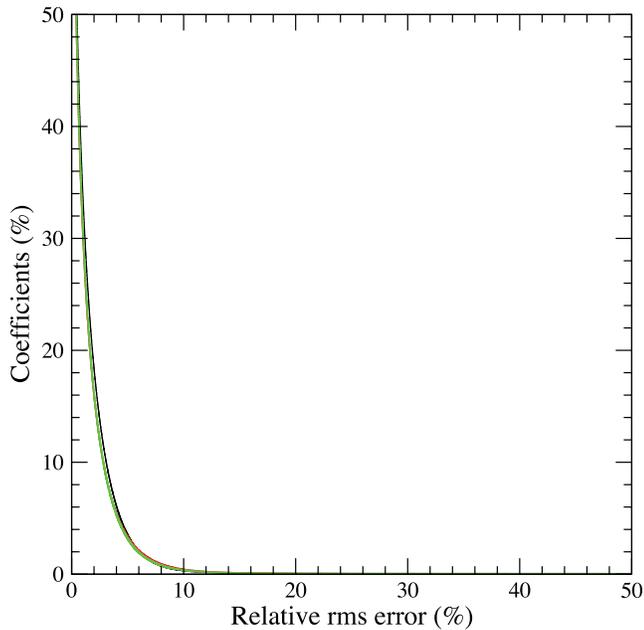
**Figure 4.** Relative rms error as a function of the percentage of thresholded wavelet coefficients used in the reconstruction of mantle model S40RTS at 200 km depth with CDF(2, 2) (black line) and D4 (red line) wavelets over four decomposition levels. The black circles show the exact rms error computed in the spatial domain for the reconstructed model.

## 4.3 Example 1: topography of the Earth

Let us now illustrate our wavelet transform on the 2-D surface of the sphere. We first consider the topography of the Earth, sampled with a resolution of 5 arc minutes, taken from model etopo1 (Amante & Eakins 2009). This data set is mapped on the cubed sphere with  $2^{10} \times 2^{10}$  elements on each face, for a total of 6 291 456 elements. We consider the Haar, CDF(2, 2) and CDF(4, 4) bases up to five scale levels. Fig. 5 shows the relative root-mean-square (rms) error as a function of the percentage of coefficients kept for the reconstruction using these three families of wavelets. The most striking feature is that the three wavelet bases give very similar compression curves. As in Schröder & Sweldens (1995), we interpret this observation by the fact that topography is not smooth. Consequently, the smoothness of the wavelets has no influence on the compression rate. However, the choice of the wavelet basis has a strong influence on the visual quality of the reconstructed topography. Reconstruction with Haar wavelets after thresholding 96.47 per cent of the wavelet coefficients (Fig. 6) is clearly contaminated by block artefacts. These artefacts are absent from the reconstruction obtained with the CDF(4, 4) wavelets (Fig. 7). While the 5 per cent rms error is very similar in both cases, the CDF(4, 4) wavelets lead to visually better reconstructions, and with slightly fewer wavelet coefficients (96.82 per cent). In any case, the three wavelets allow us to compress topographic data by a factor of  $\sim 10$  with very little loss of information, the error being around 3 per cent. The reconstructed topography is free of any visible artefact around the edges of the chunks, which suggests that our algorithm correctly solves the problem of connecting the different faces of the cubed sphere within the lifting scheme.

## 4.4 Example 2: tomographic model

We now turn to a more relevant case from the point of view of seismic tomography by considering tomographic mantle model S40RTS



**Figure 5.** Relative rms error as a function of the percentage of thresholded wavelet coefficients used in the reconstruction of Earth topography with Haar (black line), CDF(2, 2) (red line) and CDF(4, 4) (green line) over five decomposition levels.

(Ritsema *et al.* 2011) at 200 km depth. Since this model is described by spherical harmonics up to degree 40, it is straightforward to compute the velocity perturbations at the exact locations of each node in the cubed sphere. We use a grid scale  $N = 7$ , meaning that each face of the cubed sphere contains  $2^7 \times 2^7$  or  $128 \times 128$  cells, and perform the wavelet transforms up to level 4. In Fig. 8, we show the relative rms error as a function of the percentage of wavelet coefficients kept for the reconstruction with the Haar, CDF(2, 2) and CDF(4, 4) wavelets. The compression curves exhibit a very different behaviour than previously observed in the case of topographic data. We obtain increasingly higher compression rates when considering increasingly smoother wavelets. Only 5 per cent (820) of the wavelet coefficients are required to reconstruct the tomographic model with an error smaller than 2 per cent with CDF(4, 4). This strongly suggests that, contrary to the case of topography, tomographic Earth models are smooth and highly compressible when represented in bases of discrete wavelets. Obviously, the value of the compression ratio is not very significant here, since it mainly depends on the initial number of spatial nodes used to map the tomographic model. Nevertheless, this example demonstrates that discrete wavelets are competitive to represent smooth global fields on the surface of the sphere, with performance levels comparable to spherical harmonics.

## 5 WAVELET COMPRESSION IN THE 3-D CUBED SPHERE

Extending the lifting scheme to the 3-D cubed sphere is straightforward. The additional depth or radial dimension can be seen as a stack of cubed sphere surfaces similar to that shown in Fig. 2. The depth dimension thus only adds a third Cartesian dimension into the lifting scheme. Any type of wavelet can be used to compute the wavelet expansion along that depth direction, and it may differ from that used to compute the wavelet expansion on the cubed sphere surfaces. However, the radial wavelet transform is

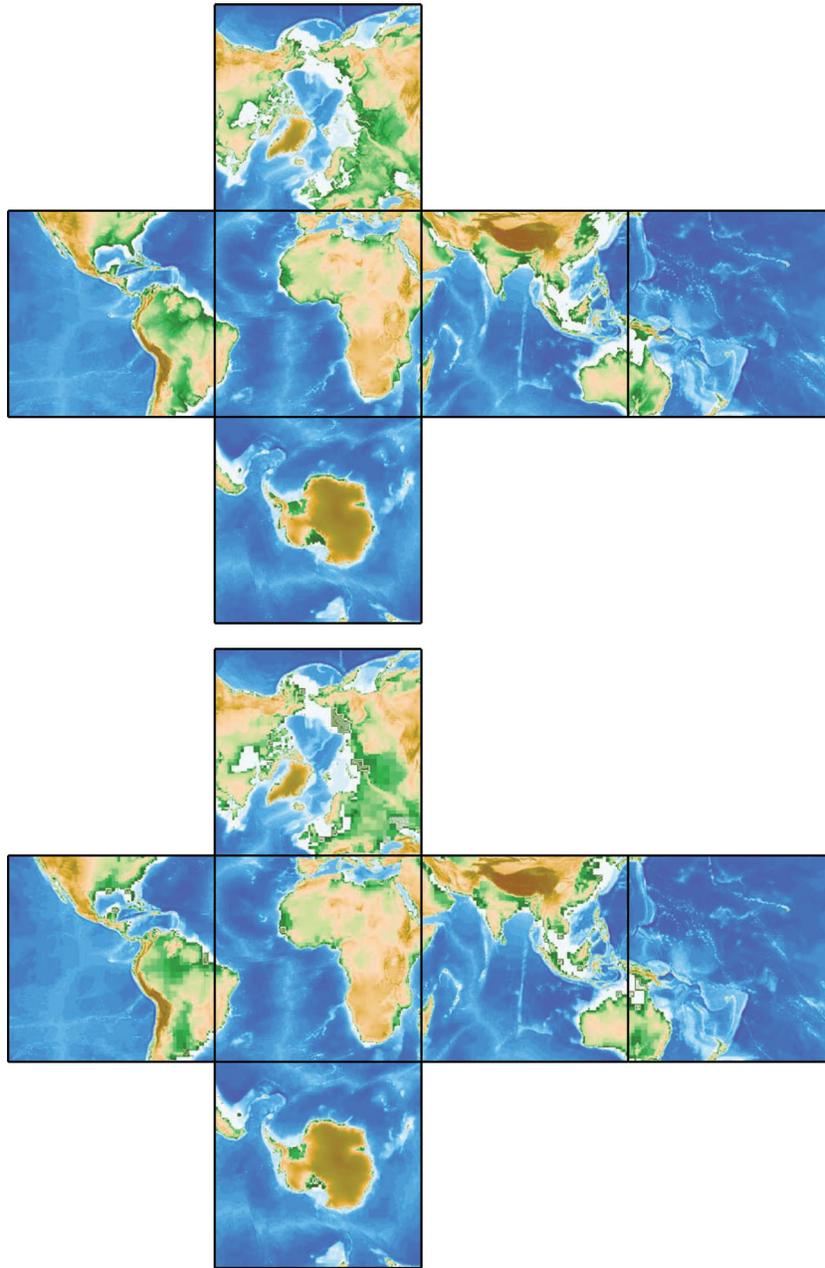
performed in a finite interval, bounded by the free surface at the top and by the core–mantle boundary at the bottom, if one considers the whole mantle and only that region. To preserve the vanishing moments of the signal, it is thus necessary to use specific filter coefficients close to the boundaries (Jensen & la Cour-Harbo 2001; Mallat 2009). While their computations can be quite involved in the general case, it turns out that they are relatively simple for CDF wavelets (Fernández *et al.* 1996), which is another good reason for preferring these wavelets over other families such as Daubechies wavelets. Handling the wavelet transform in the vicinity of depth boundaries properly is important to avoid any artefact in the reconstructed model.

### 5.1 Compression of 3-D tomographic models

Let us compute the compression curves for global mantle model S40RTS from the surface down to the core–mantle boundary. In our model extraction, we use a parametrization with 64 ( $2^6$ ) points along the vertical dimension. Note that our algorithm is general and does not require the number of depth nodes to be a power of 2. For the 2-D wavelet transforms on the cubed sphere surfaces, we use the CDF(4, 4) wavelets, which we find to give the best compression rates at 200 km depth (Fig. 8). Along the vertical direction, we compare the Haar, CDF(2, 2) and CDF(4, 4) wavelets, computed up to three scale levels. The resulting compression curves are shown in Fig. 9. The first striking observation is that 3-D compression rates are much larger than 2-D compression rates, whatever the wavelet basis used to compress along the third dimension. This result was expected because compression rates depend to first order on the dimensionality of the problem. The behaviour of the compression curves is similar to that observed in Fig. 8 which means that vertical variations of seismic velocities in tomographic models, like lateral variations, are smooth. This was expected, since model S40RTS is expanded in the basis of spherical harmonics up to order 40 only. Using the CDF(4, 4) wavelets, only 1.5 per cent of the wavelet coefficients are necessary to reconstruct that tomographic model with an error smaller than 2 per cent, a level at which no significant reconstruction artefact is visible.

### 5.2 Compression of 3-D sensitivity kernels

Let us now examine the main ingredient of finite-frequency tomography: the finite-frequency traveltimes sensitivity kernels. The spectral-element method is widely used to model wave propagation in 3-D Earth models at both regional and global scale (Komatitsch *et al.* 2005; Peter *et al.* 2011). Combined with an adjoint wavefield computation (Tromp *et al.* 2005, 2008), this approach allows us to compute 3-D sensitivity kernels. While this approach is very costly in terms of computations, it allows us to compute exact 3-D kernels in a reference 3-D model. We compute the sensitivity kernel for the direct  $P$  wave generated by the deep June 9, 1994 earthquake in Bolivia, recorded by seismic station ANMO in Albuquerque, New Mexico (USA), located at an epicentral distance of  $61.12^\circ$ . We employ hexahedral elements to mesh the whole Earth, with  $256 \times 256$  spectral elements used at the surface of each of the six chunks of the cubed sphere. As is classical in SPEC-FEM3D\_GLOBE, we use a polynomial degree  $N = 4$  to describe the wave field inside each spectral element, and thus each such element contains  $(N + 1)^3 = 125$  local and non-evenly spaced Gauß–Lobatto–Legendre (GLL) grid points. Our kernel calculation is thus accurate down to a minimum seismic period of about 17 s. Since the finite-element mesh

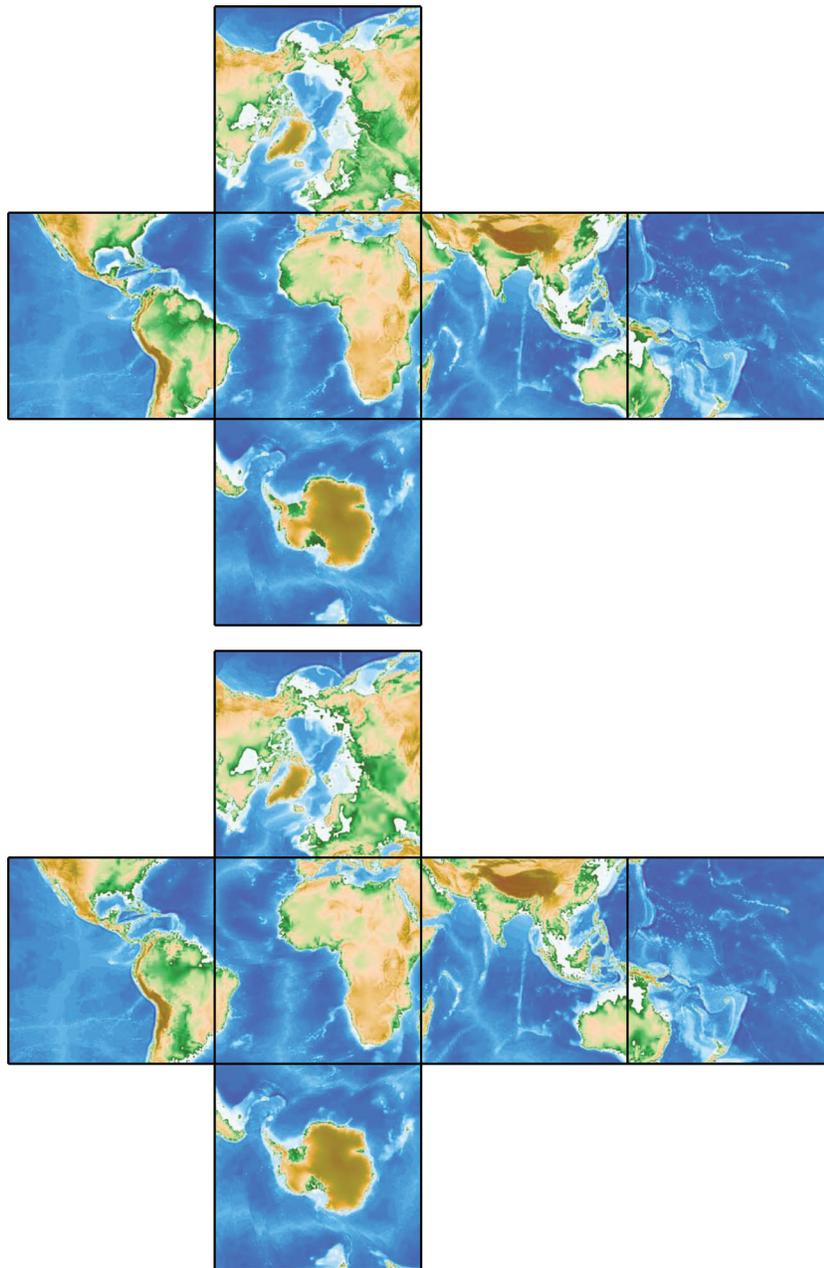


**Figure 6.** Initial (top) and reconstructed (bottom) topography of the Earth with Haar wavelets using five decomposition levels. The reconstructed topography has been obtained by thresholding 96.47 per cent of the wavelet coefficients, leading to a reconstruction error of 5 per cent.

created by SPEC3D\_GLOBE is non-structured, i.e. it does not consist of a regular grid whose topology could be described by a topologically regular grid of indices  $(i, j, k)$  and each grid point may have a valence (total number of neighbours) greater than 8, we interpolate the spectral-element kernel onto such a topologically regular grid of indices of size  $NX = 256 \times NY = 256 \times NZ = 200$ . Since the polynomial finite-element basis functions can be evaluated exactly (analytically) at any point inside a finite element and not only at its GLL grid points, this process leads to no error.

Fig. 10(a) shows a cross-section of the  $P$  kernel interpolated and projected on the 3-D cubed sphere grid. Since the sensitivity of the  $P$  wave is non-zero in a very small fraction of the cubed sphere, the compression curves have to be computed with respect to the number of significant kernel coefficients. We have arbitrarily

chosen to consider as non-zero all the coefficients that are larger than  $10^{-4}$  times the maximum absolute value of the kernel coefficients in the original cubed sphere grid. We find that the  $P$  wave kernel shown in Fig. 10(a) is described by 709 843 non-zero coefficients, which represents only about 1 per cent of the total number of cells in the cubed sphere. The compression curves obtained with the Haar, CDF(2, 2) and CDF(4, 4) wavelets up to four decomposition levels are shown in Fig. 11. The reconstructed kernels obtained with CDF(4, 4) for reconstruction errors of 1, 2, and 5 per cent, are shown in Figs 10(b)–(d), respectively. The reconstruction kernel with a tolerance error of 1 per cent is almost indistinguishable from the initial kernel. For a reconstruction error of 2 per cent, some artefacts become visible. They may be still acceptable to attain an aggressive compression level, but they are probably too strong with

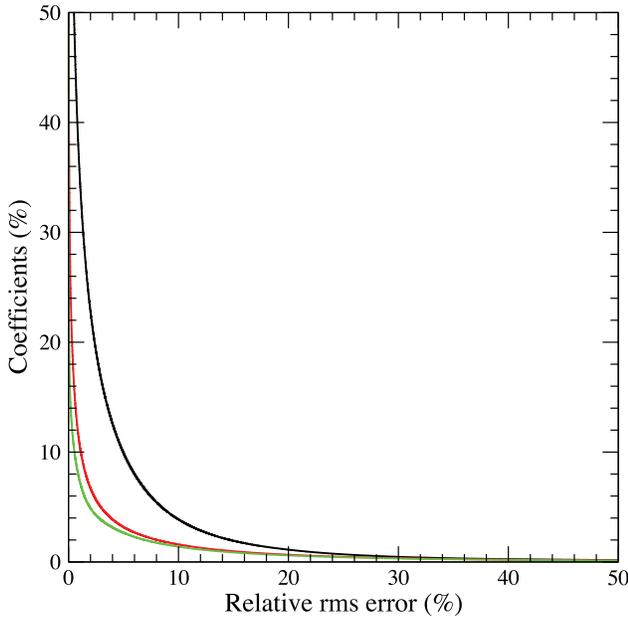


**Figure 7.** Initial (top) and reconstructed (bottom) topography of the Earth with CDF(4, 4) wavelets using five decomposition levels. The reconstructed topography has been obtained by thresholding 96.82 per cent of the wavelet coefficients, leading to a reconstruction error of 5 per cent.

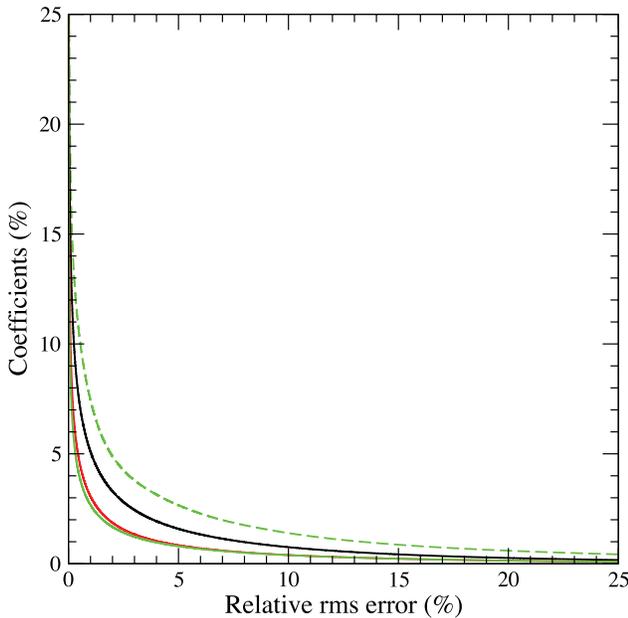
a reconstruction error of 5 per cent. The reconstructed kernel with 1 per cent error is obtained with only 11 358 wavelet coefficients, which represents about 1.6 per cent of the non-zero coefficients that are necessary to describe the initial kernel in the spatial domain, which is a very significant compression ratio. Fig. 12 shows the compression curves for the Pdiff kernel of Fig. 1(a). They show a very different behaviour. In this case, the best compression ratios are obtained with the CDF(2, 2) wavelets, rather than with the smoother CDF(4, 4) wavelets. Since this kernel corresponds to a much shorter dominant period, the sensitivity is concentrated in a very narrow tube, resulting in a far less smooth distribution of sensitivity. Using CDF(2, 2) wavelets, 74 737 wavelet coefficients are necessary to reconstruct this kernel with an error of 1 per cent. This comparison clearly demonstrates that kernels having large first

Fresnel zone because of the long dominant period of the reference wavelet are much more compressible than kernels corresponding to shorter dominant periods. The benefit of using wavelets thus increases with the dominant period of waves, when finite-frequency effects become stronger.

While efficient methods to compute accurate 3-D sensitivity kernels have been developed recently (Zhao & Chevrot 2011a,b; Fuji *et al.* 2012), their use in massive inverse tomographic problems would typically involve storing hundred of thousands of 3-D sensitivity kernels. This can represent very large volumes of data, especially if a fine tomographic grid is considered, which is necessary to describe their detailed structure as mentioned above. Wavelet compression of sensitivity kernels in the cubed sphere would drastically reduce the space required to store them, but also



**Figure 8.** Relative rms error as a function of the percentage of thresholded wavelet coefficients used in the reconstruction of mantle model S40RTS at 200 km depth with Haar (black line), CDF(2, 2) (red line) and CDF(4, 4) (green line) wavelets over four decomposition levels.



**Figure 9.** Relative rms error as a function of the percentage of thresholded wavelet coefficients used in the reconstruction of 3-D mantle model S40RTS with Haar (black line), CDF(2, 2) (red line) and CDF(4, 4) (green line) wavelet transforms over three decomposition levels along the vertical dimension. In the three cases, the CDF(4, 4) wavelets over four decomposition levels have been used for the 2-D wavelet transforms on the cubed sphere surfaces. For comparison, we also plot the compression curve obtained with the CDF(4, 4) wavelets on model S40RTS at 200 km depth (dashed green line).

simplify considerably their distribution to the scientific community in the form of pre-computed databases. This is another important practical problem, since the computation of 3-D kernels requires significant computational resources, and is typically performed on

national supercomputing resources, in a remote computing centre. It is then necessary to transfer these kernels back to the laboratory where they can be incorporated in the construction of a massive tomographic inverse problem. While this is easily done for a few kernels computed in a fine 3-D grid, the large size of the files containing these kernels may considerably complicate and slow down the process. While discrete wavelets in the cubed sphere are clearly the solution of choice to compress sensitivity kernels, it would be necessary to design an efficient encoder and define a new data format such as the JPEG2000 standard encoder for 2-D images (e.g. Taubman & Marcellin 2001). So far, we have not considered this problem, which is beyond the scope of this study.

## 6 IMPLICATIONS FOR TOMOGRAPHIC INVERSIONS IN THE WAVELET DOMAIN

A canonical tomographic problem consists in finding the vector model  $\mathbf{m}$  expanded in a spherical grid that minimizes the misfit with respect to the data vector  $\mathbf{d}$ :

$$S(\mathbf{m}) = (\mathbf{G} \cdot \mathbf{m} - \mathbf{d})^t \mathbf{C}_d^{-1} (\mathbf{G} \cdot \mathbf{m} - \mathbf{d}) + \mathbf{m}^t \mathbf{C}_m^{-1} \mathbf{m}, \quad (11)$$

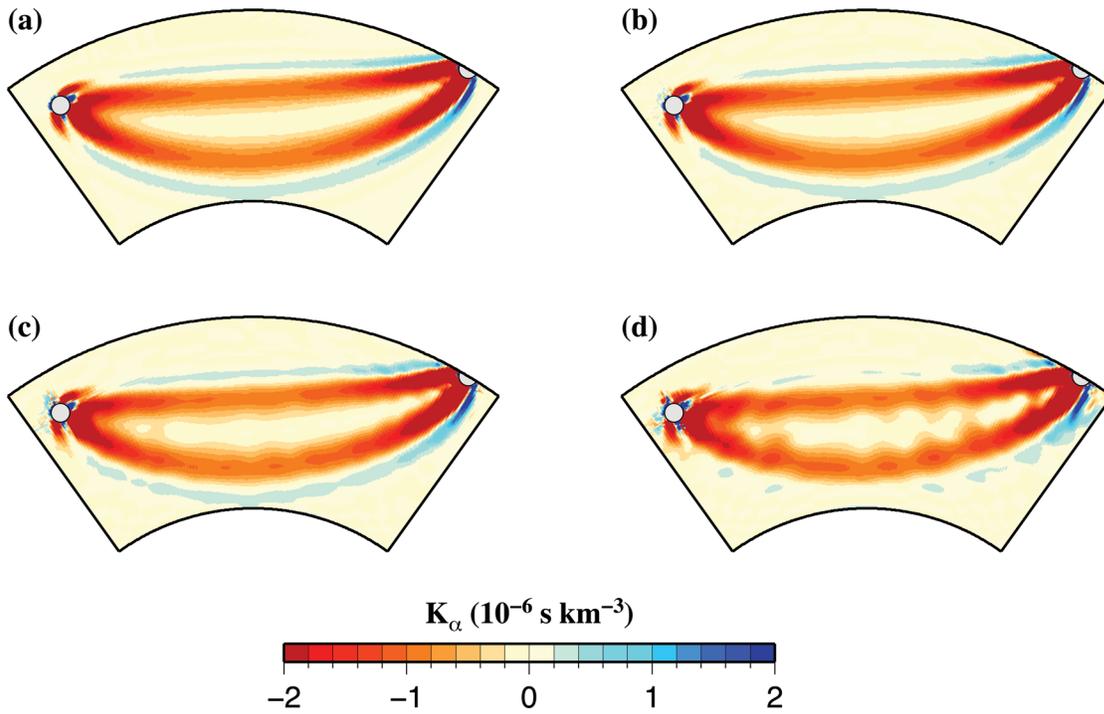
where  $\mathbf{C}_d$  and  $\mathbf{C}_m$  are, respectively, the data and model covariance matrices, and each line of the sensitivity matrix  $\mathbf{G}$  contains a 3-D sensitivity kernel. To be more specific, an element  $G_{ij}$  of the sensitivity matrix represents the partial derivative of the traveltime of path  $i$  to a velocity perturbation in grid cell  $j$ . This minimization problem can be recast into the linear inverse problem:

$$\begin{pmatrix} \mathbf{C}_d^{-1/2} \mathbf{G} \\ \mathbf{C}_m^{-1/2} \end{pmatrix} \mathbf{m} = \begin{pmatrix} \mathbf{C}_d^{-1/2} \mathbf{d} \\ \mathbf{0} \end{pmatrix}. \quad (12)$$

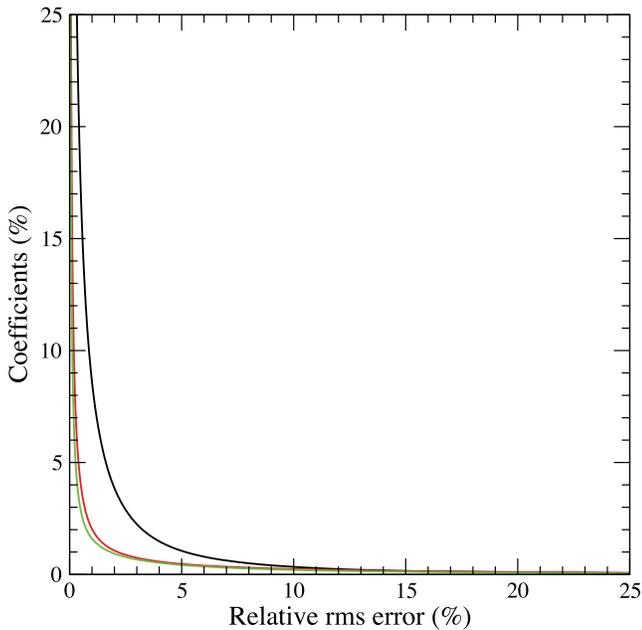
As already stated, the basic motivation for using discrete wavelet bases to parametrize tomographic models relies on the compact representation that can be achieved in the wavelet domain for both 3-D sensitivity kernels and tomographic problems. Indeed, the tests presented in the previous section demonstrate that the number of coefficients necessary to describe sensitivity kernels can be reduced by almost two orders of magnitude without losing any significant amount of information. This result is important because it will allow one to dramatically reduce the number of non-zero elements in the sensitivity matrix  $\mathbf{G}$ . Formulating (12) into the wavelet domain leads to

$$\begin{pmatrix} \mathbf{C}_d^{-1/2} \mathbf{G}_w \\ \mathbf{C}_w^{-1/2} \end{pmatrix} \mathbf{w} = \begin{pmatrix} \mathbf{C}_d^{-1/2} \mathbf{d} \\ \mathbf{0} \end{pmatrix}, \quad (13)$$

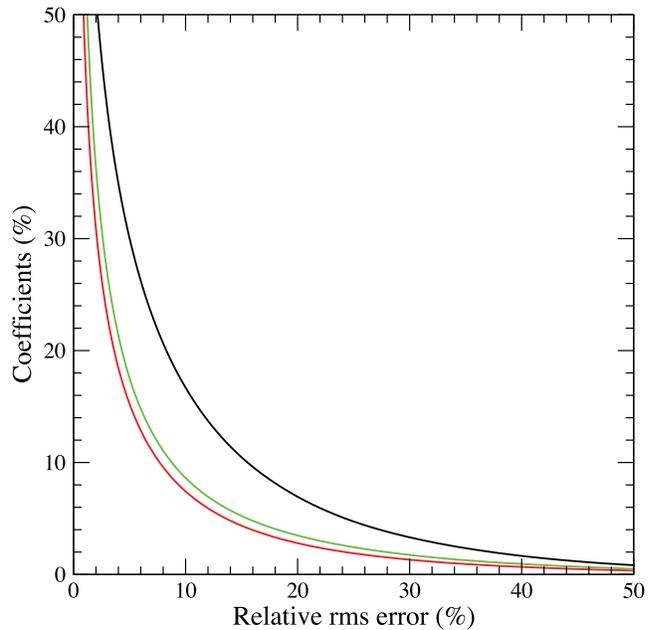
where the lines of  $\mathbf{G}_w$  now contain the wavelet coefficients of the 3-D sensitivity kernels, and the new model vector  $\mathbf{w}$  contains the wavelet coefficients of  $\mathbf{m}$ . A popular technique to solve (13) is the LSQR algorithm (Paige & Saunders 1982) which can easily be adapted to exploit the sparsity of  $\mathbf{G}_w$ . Since this algorithm only requires computing matrix–vector products involving  $\mathbf{G}_w$  or  $\mathbf{G}_w^t$ , it can easily be adapted to sparse matrix representations to perform the computations with the non-zero elements of these matrices only. Solving (13) instead of (12) can be done for a fraction of the cost in terms of both memory and CPU time. Since finite-frequency theory predicts that sensitivity is broadly distributed around the geometrical ray, it leads to much larger tomographic problems than those classically considered in the framework of ray theory. Working within the wavelet domain reduces the finite-frequency tomographic problem to a size comparable to a ray tomographic problem. Therefore, multi-scale



**Figure 10.** Sensitivity kernel for the traveltime of a 20 s *P* wave recorded at an epicentral distance of 61.12° projected in the cubed sphere with 256 × 256 elements on the surface and 200 elements along the radial direction (a). Also shown are the reconstructed kernels using CDF(4, 4) wavelets for reconstruction errors of 1 per cent (b), 2 per cent (c) and 5 per cent (d).



**Figure 11.** Relative rms error as a function of the percentage of thresholded wavelet coefficients used in the reconstruction of the *P* kernel shown in Fig. 10(a) with Haar (black line), CDF(2, 2) (red line) and CDF(4,4) (green line) wavelets over four decomposition levels.



**Figure 12.** Relative rms error as a function of the percentage of thresholded wavelet coefficients used in the reconstruction of the *Pdiff* kernel shown in Fig. 1(a) with Haar (black line), CDF(2, 2) (red line) and CDF(4, 4) (green line) wavelets over four decomposition levels.

tomography will allow one to account for finite-frequency effects at no additional cost compared to classical ray tomography.

The sparsity of tomographic models can also be exploited within the recently developed mathematical theory of compressed sensing (Donoho 2006; Tsaig & Donoho 2006; Lustig *et al.* 2007). In this approach, a  $\ell_1$  norm is used to damp the model wavelet coefficient

and we need to solve a non-linear inverse problem to find the model  $\mathbf{w}$  that minimizes the new misfit function

$$S_2(\mathbf{m}) = \|\mathbf{G}_w \cdot \mathbf{w} - \mathbf{d}\|_2^2 + \lambda \|\mathbf{w}\|_1. \tag{14}$$

The motivation to use the  $\ell_1$  norm lies in the fact that it will tend to favour solutions with a small number of large significant

coefficients, in contrast to the usual  $\ell_2$  norm that will favour a large number of small coefficients. If the model is sparse, this makes its reconstruction more robust. Different algorithms have been proposed to find the minimum of (14) (e.g. Loris & Verhoeven 2012). These algorithm requires computing the gradient of the data misfit at a number of iterations which, like in the LSQR algorithm, only involves computing matrix–vector products. Using a sparse representation of matrix  $\mathbf{G}_w$ , these algorithms can be made extremely efficient. Therefore, the  $\ell_1$  regularization can be used at almost no extra cost compared to the  $\ell_2$  norm. This new regularization scheme was first tested on simple tomographic inverse problems by Loris *et al.* (2007), and initially proved promising, but later studies on more realistic models (Loris *et al.* 2010) found little improvement over the classical  $\ell_2$  norm. Other types of non-linear regularization methods are described in Loris & Verhoeven (2012).

We should emphasize here a fundamental difference with the resolution of the inverse problem as envisioned in Simons *et al.* (2011). Indeed, Simons *et al.* (2011) solve the tomographic problem in the spatial domain, as in (12), but regularize the  $\ell_1$  norm of the model wavelet transform. Therefore, they do not exploit the sparsity of both 3-D sensitivity kernels and tomographic models in the resolution of the tomographic problem. It is interesting to note, however, that in the case of ray tomography, matrix  $\mathbf{G}$  is sparser than  $\mathbf{G}_w$ , because in ray theory, sensitivity is concentrated along the singularity of the ray, while it is widely distributed over different scaling and wavelet coefficients at different scales in  $\mathbf{G}_w$ . Working in the spatial domain is thus more efficient in the case of ray theory, but it is more efficient to work in the wavelet domain in the case of finite-frequency theory. However, we would argue that it may still be interesting to solve a ray theory tomographic problem in the wavelet domain. Indeed, redistribution of sensitivity over different scales may improve the quality of the reconstructed tomographic model. This also opens new perspectives to improve ray-theory global tomographic models that we plan to explore in future work.

## 7 CONCLUSIONS

We have introduced an extension of the lifting scheme in the cubed sphere that allows one to compute fast wavelet transforms in 2-D or 3-D spherical domains. The computational grid is obtained by a tessellation of the sphere with six chunks, within which one defines Cartesian coordinates. Mapping the radial dimension simply consists in considering stacks of such constructions. The algorithm is simple and flexible and can be implemented with any family of orthogonal or bi-orthogonal wavelets. However, our experiments suggest that CDF(4, 4) represents an excellent trade-off between efficiency, ease of implementation and capacity to obtain a sparse representation of both 3-D finite-frequency kernels and tomographic models in the wavelet domain. These wavelets are very well suited to represent the whole spectrum of seismic heterogeneities, dominated by longer wavelengths (Su & Dziewonski 1991, 1992; Chevrot *et al.* 1998; Simons *et al.* 2011) while retaining the capacity to describe the smaller scale structures that are heavily smoothed in current global tomographic models (Gudmundsson *et al.* 1990; Margerin & Nolet 2003; Garcia *et al.* 2009). Using these wavelets, the size of the sensitivity matrix can be reduced by almost two orders of magnitude without losing any significant amount of information regarding finite-frequency sensitivity to structural details.

Ultimately, resolution in tomographic models of the Earth will still be limited by the level of noise in seismic observables and by uneven spatial coverage. However, parametrizing tomographic models with wavelets keeps the inverse problem tractable even in a

very fine tomographic grid. Another advantage of parametrizing tomographic models with wavelets is that the grid can easily be refined in a sub-domain, in which the wavelet expansion can be continued towards shorter scales. It is thus straightforward to embed a regional tomographic problem into a global tomographic problem in order to mitigate the leakage of unmodelled structural heterogeneities inside the regional model. Finally, a wavelet parametrization offers versatility and flexibility. When the resolution potential of a tomographic data set allows for resolving fine structures in a given region, the quality of the inversion will not be degraded because of other regions having much lower resolution. In other words, regularization of the inverse tomographic problem will not penalize resolution in the best sampled parts of the model, which is a major drawback in current classical tomographic approaches.

## ACKNOWLEDGMENTS

This work was granted access to the high-performance computing resources of the French super computer center CCRT under allocation #2012-046351 awarded by GENCI (Grand Équipement National de Calcul Intensif).

## REFERENCES

- Amante, C. & Eakins, B.W., 2009. Etopo1 1 arc-minute global relief model: procedures, data sources and analysis, Tech. rep., NOAA Technical Memorandum NESDIS NGDC-24.
- Chevrot, S., Montagner, J.P. & Snieder, R., 1998. The spectra of tomographic models, *Geophys. J. Int.*, **133**, 783–788.
- Chevrot, S. & Zhao, L., 2007. Multi-scale finite-frequency Rayleigh wave tomography of the Kaapvaal craton, *Geophys. J. Int.*, **169**, 201–215.
- Cohen, A., Daubechies, I. & Feauveau, J.C., 1992. Biorthogonal bases of compactly supported wavelets, *Commun. Pure appl. Math.*, **45**(5), 485–560.
- Dahlen, F.A., Hung, S.H. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes–I. Theory, *Geophys. J. Int.*, **141**, 157–174.
- Donoho, D.L., 2006. Compressed sensing, *IEEE Trans. Inf. Theory*, **52**, 1289–1306.
- Fernández, G., Periaswamy, S. & Sweldens, W., 1996. LIFTPACK: a software package for wavelet transforms using lifting, in *Wavelet Applications in Signal and Image Processing IV*, Proc. SPIE 2825, pp. 396–408, eds Unser, M., Aldroubi, A. & Laine, A.F., SPIE, Denver, CO.
- Fuji, N., Chevrot, S., Zhao, L., Geller, R.J. & Kawai, K., 2012. Finite-frequency structural sensitivities of short-period compressional body waves, *Geophys. J. Int.*, **190**, 522–540.
- Garcia, R., Chevrot, S. & Calvet, M., 2009. Statistical study of seismic heterogeneities at the base of the mantle from PKP differential travel times, *Geophys. J. Int.*, **179**, 1607–1616.
- Gudmundsson, O., Davies, J.H. & Clayton, R.W., 1990. Stochastic analysis of global traveltime data: mantle heterogeneity and random errors in the ISC data, *Geophys. J. Int.*, **102**, 25–43.
- Haar, A., 1910. Zur Theorie der orthogonalen Funktionensysteme, *Math. Annalen*, **69**, 331–371.
- Hung, S.H., Dahlen, F.A. & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes II. Examples, *Geophys. J. Int.*, **141**, 175–203.
- Hung, S.H., Dahlen, F.A. & Nolet, G., 2001. Wavefront healing: a banana-doughnut perspective, *Geophys. J. Int.*, **146**, 289–312.
- Jensen, A. & la Cour-Harbo, A., 2001. *Ripples in Mathematics*, Springer-Verlag, Berlin.
- Komatitsch, D. & Tromp, J., 2002. Spectral-element simulations of global seismic wave propagation I. Validation, *Geophys. J. Int.*, **149**, 390–412.
- Komatitsch, D., Tsuboi, S. & Tromp, J., 2005. The spectral-element method in seismology, in *Seismic Earth: Array Analysis of Broadband Seismograms*, Geophys. Monogr. Ser. Vol. 157, pp. 205–227, American Geophysical Union, Washington, DC.

- Loris, I., Douma, H., Nolet, G., Daubechies, I. & Regone, C., 2010. Non-linear regularization techniques for seismic tomography, *J. Comp. Phys.*, **229**, 890–905.
- Loris, I., Nolet, G., Daubechies, I. & Dahlen, F.A., 2007. Tomographic inversion using  $\ell_1$ -norm regularization of wavelet coefficients, *Geophys. J. Int.*, **170**, 359–370.
- Loris, I. & Verhoeven, C., 2012. Iterative algorithms for total variation-like reconstructions in seismic tomography, *Int. J. Geomath.*, in press.
- Lustig, M., Donoho, D. & Pauly, J.M., 2007. Sparse MRI: the application of compressed sensing for rapid MR imaging, *Magn. Reson. Med.*, **58**, 1182–1195.
- Mallat, S., 2009. *A Wavelet Tour of Signal Processing*, 3rd edn, Academic Press, Waltham, MA.
- Margerin, L. & Nolet, G., 2003. Multiple scattering of high-frequency seismic waves in the deep Earth: PKP precursor analysis and inversion for mantle granularity, *J. geophys. Res.*, **108**, 2514, doi:10.1029/2003JB002455.
- Montelli, R., Nolet, G., Masters, G., Dahlen, F.A. & Hung, S.H., 2004a. Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, **303**, 338–343.
- Montelli, R., Nolet, G., Masters, G., Dahlen, F.A. & Hung, S.H., 2004b. Global P and PP traveltime tomography: rays versus waves, *Geophys. J. Int.*, **158**, 637–654.
- Paige, C.C. & Saunders, M.A., 1982. LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Softw.*, **8**, 43–71.
- Peter, D. *et al.*, 2011. Forward and adjoint simulations of seismic wave propagation on fully unstructured hexahedral meshes, *Geophys. J. Int.*, **186**, 721–739.
- Ritsema, J., Deuss, A., van Heijst, H.J. & Woodhouse, J.H., 2011. S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements, *Geophys. J. Int.*, **184**, 1223–1236.
- Ronchi, C., Iacono, R. & Paolucci, P.S., 1996. The ‘Cubed Sphere’: a new method for the solution of partial differential equations in spherical geometry, *J. Comp. Phys.*, **124**, 93–114.
- Sadourny, R., 1972. Conservative finite-difference approximations of the primitive equations on quasi-uniform spherical grids, *Mon. Weather Rev.*, **100**, 136–144.
- Schröder, P. & Sweldens, W., 1995. Spherical wavelets: efficiently representing functions on the sphere, in *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, ACM, New York, NY.
- Simons, F.J. *et al.*, 2011. Solving or resolving global tomographic models with spherical wavelets, and the scale and sparsity of seismic heterogeneity, *Geophys. J. Int.*, **187**, 969–988.
- Spetzler, J. & Trampert, J., 2003. Implementing spectral leakage corrections in global surface wave tomography, *Geophys. J. Int.*, **155**, 532–538.
- Stollnitz, E.J., deRose, T.D. & Salesin, D.H., 1995. Wavelets for computer graphics: a primer, part 1, *IEEE Comput. Graph. Appl.*, **15**(3), 76–86.
- Su, W.J. & Dziewonski, A.M., 1991. Predominance of long wavelength heterogeneity in the mantle, *Nature*, **352**, 121–126.
- Su, W.J. & Dziewonski, A.M., 1992. On the scale of mantle heterogeneity, *Phys. Earth planet. Inter.*, **74**, 29–54.
- Sweldens, W., 1995. The lifting scheme: a new philosophy in biorthogonal wavelet constructions, in *Wavelet Applications in Signal and Image Processing III*, Proc. SPIE 2569, pp. 68–79, eds Laine, A.F. & Unser, M., SPIE, Denver, CO.
- Sweldens, W. & Schröder, P., 1996. Building your own wavelets at home, in *Wavelets in Computer Graphics*, ACM SIGGRAPH Course Notes.
- Taubman, D.S. & Marcellin, M.W., 2001. *JPEG 2000: Image Compression Fundamentals, Standards and Practice*, Kluwer Academic Publishers, Norwell, MA.
- Trampert, J. & Spetzler, J., 2006. Surface wave tomography: finite-frequency effects lost in the null space, *Geophys. J. Int.*, **164**, 394–400.
- Tromp, J., Komatitsch, D. & Liu, Q., 2008. Spectral-element and adjoint methods in seismology, *Comput. Phys. Commun.*, **3**, 1–32.
- Tromp, J., Tape, C. & Liu, Q.Y., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**(1), 195–216.
- Tsaig, Y. & Donoho, D.L., 2006. Extensions of compressed sensing, *Signal Process.*, **86**, 549–571.
- Usevitch, B.E., 2001. A tutorial on modern lossy wavelet image compression: foundations of jpeg 2000, *IEEE Signal Process. Mag.*, **18**, 22–35.
- Uytterhoeven, G., Roose, D. & Bultheel, A., 1997. Wavelet transforms using the lifting scheme, Tech. rep., Report ITA-Wavelets—WP1.1 (Revised Version).
- Zhao, L. & Chevrot, S., 2011a. An efficient and flexible approach to the calculation of three-dimensional full-wave Fréchet kernels for seismic tomography—I. Theory, *Geophys. J. Int.*, **185**, 922–938.
- Zhao, L. & Chevrot, S., 2011b. An efficient and flexible approach to the calculation of three-dimensional full-wave Fréchet kernels for seismic tomography—II. Numerical results, *Geophys. J. Int.*, **185**, 939–954.