 doi: 10.1111/j.1365-246X.2007.03511.x

Simultaneous iterative time-domain sparse deconvolution to teleseismic receiver functions

Christian Escalante, Yu J. Gu and Mauricio Sacchi

Institute for Geophysical Research, Department of Physics, University of Alberta, Edmonton AB T6G2G7*, Canada. E-mail: cescalan@phys.ualberta.ca*

Accepted 2007 May 29. Received 2007 May 10; in original form 2006 April 8

SUMMARY

In this paper, we present a new approach to estimate high-resolution teleseismic receiver functions using a simultaneous iterative time-domain sparse deconvolution. This technique improves the deconvolution by using reweighting strategies based on a Cauchy criterion. The resulting sparse receiver functions enhance the primary converted phases and its multiples. To test its functionality and reliability, we applied this approach to synthetic experiments and to seismic data recorded at station ABU, in Japan. Our results show *Ps* conversions at approximately 4.0 s after the primary *P* onset, which are consistent with other seismological studies in this area. We demonstrate that the sparse deconvolution is a simple, efficient technique in computing receiver functions with significantly greater resolution than conventional approaches.

Key words: deconvolution, inversion, sparse regularization, teleseismic receiver functions.

1 INTRODUCTION

Crustal and mantle discontinuities play an important role in characterizing major changes in seismic velocity, rheology, and potentially chemical composition (Ringwood & Irifune 1988; Stixrude 1997; Shearer 2000). In recent years, several methods have been proposed to estimate large-scale properties of discontinuities (Shearer & Masters 1992; Lee & Grand 1996; Flanagan & Shearer 1998; Gu *et al.* 1998; Flanagan & Shearer 1999; Gu & Dziewonski 2002). These methods are complemented by high-resolution approaches such as teleseismic receiver functions (Petersen *et al.* 1993; Kind *et al.* 1995; Sheehan 1995; Bostock 1996; Shen *et al.* 1996, 1998; Vinnik *et al.* 1996; Dueker & Sheehan 1997; Gurrola & Minster 1998; Li *et al.* 2000; Zhu & Kanamori 2000) in regions well sampled by seismic stations. Receiver functions are time series representing *P*-to-*S* conversions of teleseismic waves at discontinuities in response to the earth structure beneath the receiver. Assuming nearly horizontal layers, most of the converted energy is observed on the radial and transverse components. By deconvolving the horizontal component with the vertical component, the influence of the source and near source structure can be effectively removed. In other words, the vertical component is used as an effective source time function to isolate the local response from the horizontal component (Langston 1977, 1979, 1981; Owens & Crosson 1988; Ammon *et al.* 1990; Ammon 1991).

One of the most commonly used techniques to compute receiver functions is frequency-domain deconvolution (Clayton & Wiggins 1976; Bostock 1996; Dueker & Sheehan 1997; Zhu & Kanamori 2000; Lawrence & Shearer 2006), also called water-level deconvolution. This method performs a spectral division between the horizontal component (either radial or transverse) and the vertical component (source spectrum) and uses a water-level stabilizer to avoid the amplification of small spectral components. Another approach is time-domain deconvolution (Oldenburg 1981; Abers *et al.* 1995; Gurrola *et al.* 1995; Yuan *et al.* 1997; Zhu 2004) that can be solved using the method of least squares. In this case, a trade-off parameter is used to stabilize the inversion. These two approaches generally produce similar results when the available data are broadband with a high signal-to-noise ratio (Ligorría & Ammon 1999).

In reality, seismic data often contain substantial noise, especially when the source is relatively weak and/or the structural complexity is high. Therefore, successful applications of a given method often depends on its effectiveness in removing the noise and enhancing the desirable features of the signal. For this reason, several deconvolution approaches have been proposed in seismic and ultrasonic applications (Pham & de Figueiredo 1989; Mendel 1990; Sacchi 1997; Olofsson 2004). These methods attempt to construct a high-resolution solution to the deconvolution problem by imposing sparsity on the unknown reflectivity model. In this paper, we adopt the technique proposed by Sacchi (1997) for the deconvolution of seismic events in reflection seismology. The method consists of solving the deconvolution inverse problem with a regularization strategy based on a Cauchy criterion. The latter is used to retrieve sparse reflectivity sequences. The term 'sparse' is used to refer to a solution that contain a few non-zero samples (reflectors). It is important to mention that results similar to the ones found in this paper could be obtained using the so-called minimalist receiver functions proposed by Menke (personal communication, 2007). Minimalist receiver functions are found via a non-linear optimization method where one needs to estimate the arrival times and amplitudes for an assumed number of spikes. This method produces models that consists of a minimal number of pulses needed to fit the observations

within a desired level of accuracy (see also Menke & Levin 2004 for its application to *SKS* splitting).

The main objective of this paper is to introduce an iterative time-domain sparse deconvolution method for computing teleseismic receiver functions. This method was applied to both synthetic and actual seismic data recorded at station ABU, Japan. Our results clearly demonstrate that the sparse deconvolution approach is a simple, efficient and high-resolution alternative to the existing methods.

2 SPARSE DECONVOLUTION

Receiver function determination is inherently a deconvolution problem. The solutions obtained by common techniques, such as water-level based spectral division (e.g. Ammon 1991) and time-domain deconvolution (e.g. Gurrola *et al.* 1995), can be equally effective when using sufficiently broad-band records with good signal-to-noise levels. Unfortunately, in most of the cases, the band-limitation of data leads to suboptimal deconvolution results. Hence, the choice of an appropriate deconvolution technique is vital.

A sparse deconvolution problem is usually referred to: given some observation sequence $\boldsymbol{R} = \{R_1, \ldots, R_M\}$, find the sparse signal $\boldsymbol{r} = \{r_1, \ldots, r_N\}$ such that

$$\boldsymbol{R} = \boldsymbol{Z}\boldsymbol{r} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{Z}$ is a convolution matrix and $\boldsymbol{n}$ represents the noise. The signal $\boldsymbol{r}$ is known to be sparse, that is, only a few of its samples have non-zero values.

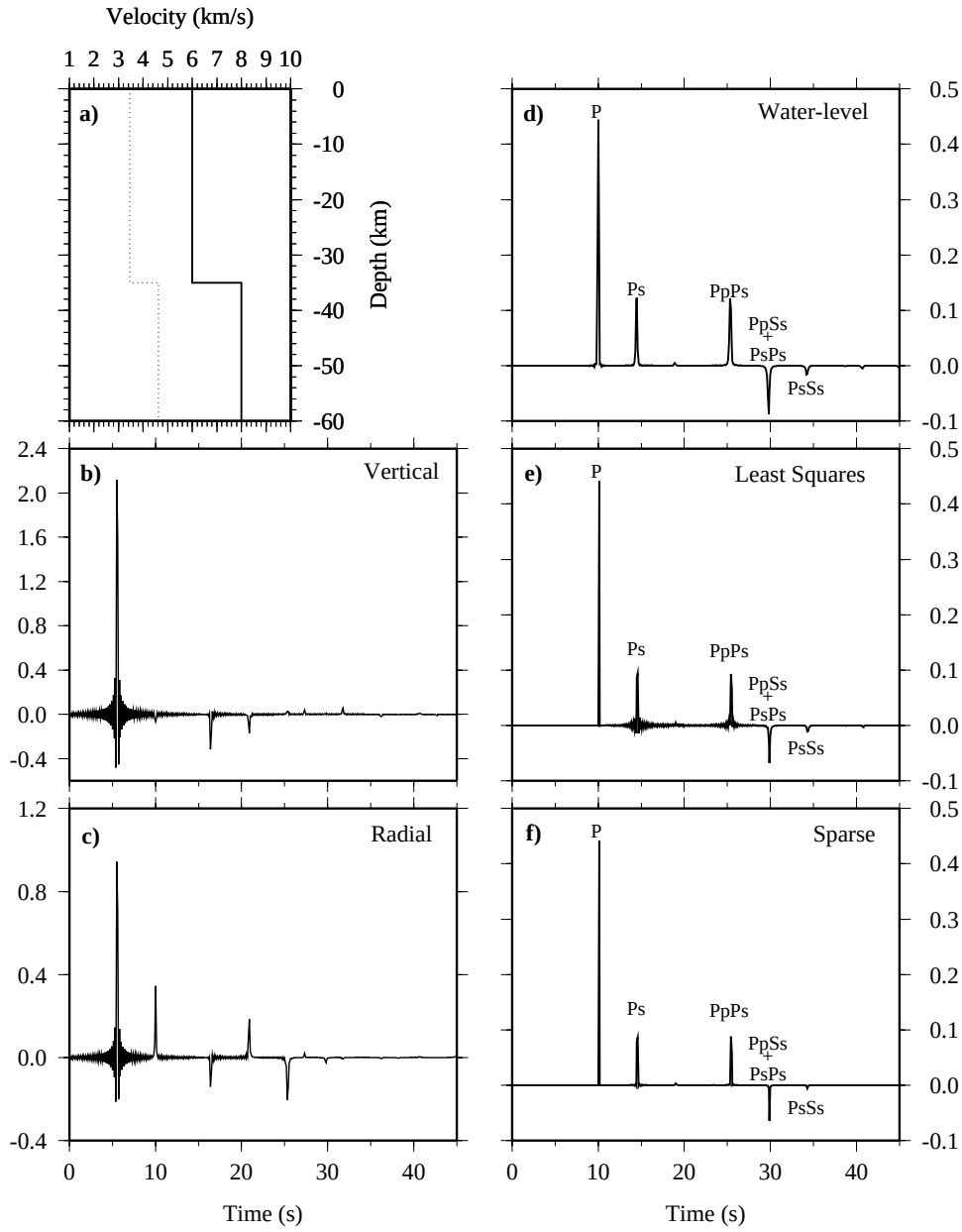

Figure 1. Estimated receiver functions for a simple two-layer velocity model: (a) velocity model (solid and dashed lines indicate *P* and *S* velocities, respectively); (b) vertical-component seismogram; (c) radial-component seismogram; (d) receiver function obtained after applying water-level deconvolution (Ammon 1991); (e) computed receiver function after applying the least-squares deconvolution (Gurrola *et al.* 1995) and (f) estimated receiver function using the sparse deconvolution (this study). *Ps* conversions and multiples are indicated.

In theory, a solution of $\boldsymbol{r}$ that minimizes the squared error $E^2 = ||\boldsymbol{R} - \boldsymbol{Zr}||_2^2$ can be easily found if $\boldsymbol{Z}$ is known. In practice, however, a stable solution cannot be uniquely determined due to the ill-conditioned character of $\boldsymbol{Z}$ (typically band-limited). A generic approach to overcome this problem is to apply regularization techniques by including some *a priori* knowledge of the solution.

Sacchi (1997) used different reweighting strategies to improve the deconvolution of seismic records and retrieved sparse reflectivity sequences by applying a Cauchy criterion. This criterion is derived by considering that the unknown reflectivity can be modeled via a Cauchy distribution. The latter is a long-tail probability distribution that is commonly used as a prior for the inversion of sparse models and signals.

In time-domain, the deconvolution can be achieved by solving the convolution equation:

$$\boldsymbol{Z(t)} * r(t) = \boldsymbol{R(t)}, \tag{2}$$

which can be solved using a least-squares approach. For an ideal case, we can rewrite this equation in matrix form as (Gurrola *et al.* 1995)

$$\boldsymbol{R} = \boldsymbol{Zr}, \tag{3}$$

where $\boldsymbol{R}$ can be either the radial or transverse component and contains M elements, $\boldsymbol{R}$ is the receiver function composed of N elements, and $\boldsymbol{Z}$ is the convolution matrix created by the vertical component ($\boldsymbol{z}$) of the seismogram of k elements. In other words, each column of $\boldsymbol{Z}$ contains a vector $\boldsymbol{z}$ properly padded with zeros. The size of $\boldsymbol{Z}$ is $N \times M$ as follows:

$$\boldsymbol{Z} = \begin{pmatrix} z_1 & 0 & \vdots & \vdots & \dots & 0 \\ \vdots & z_1 & 0 & \vdots & \vdots & \vdots \\ z_k & \vdots & z_1 & \ddots & \vdots & \vdots \\ 0 & z_k & \vdots & \ddots & \vdots & \vdots \\ \vdots & 0 & z_k & \vdots & z_1 & 0 \\ \vdots & \vdots & 0 & \ddots & \vdots & z_1 \\ \vdots & \vdots & \vdots & \vdots & z_k & \vdots \\ \vdots & \vdots & \vdots & \vdots & 0 & z_k \end{pmatrix}. \tag{4}$$

According to Gurrola *et al.* (1995), in a noise-free case we can solve eq. (3) directly for the receiver function, which in principle should produce a result similar to the water-level deconvolution. In reality $\boldsymbol{R}$ contains noise, therefore, an exact solution to (3) would introduce noise into our results due to overfitting the data. Here, we are looking for a solution in which the difference between the observations $\boldsymbol{R}$ and predicted $\boldsymbol{Zr}$ is minimized. A solution to this problem can be expressed as

$$\boldsymbol{r} = \left(\lambda \boldsymbol{I} + \boldsymbol{Z}^T \boldsymbol{Z}\right)^{-1} \boldsymbol{Z}^T \boldsymbol{R}, \tag{5}$$

where $\boldsymbol{I}$ is an $N \times N$ identity matrix, and λ is a trade-off parameter. The appropriate value for λ would be one that satisfies eq. (3) within a tolerance level.

This method, as well as frequency-domain method, are both examples of a damped least-squares deconvolution (Gurrola *et al.* 1995). In either case, we test a range of water-level values or λ to obtain an optimal deconvolution result.

The sparse constraint can be implemented by minimizing the following cost function:

$$\boldsymbol{J} = \|\boldsymbol{Zr} - \boldsymbol{R}\|_2^2 + \mu \sum_{i=1} ln\left(1 + a r_i^2\right), \tag{6}$$

where the second term on the right-hand side represents the cost function of a Cauchy distribution. In this equation, μ and a are hyperparameters needed to enforce sparseness to the model. We then obtain

$$\boldsymbol{r} = \left[\boldsymbol{Z}^T \boldsymbol{Z} + \mu \boldsymbol{Q(r)}\right]^{-1} \boldsymbol{Z}^T \boldsymbol{R}, \tag{7}$$

where $\boldsymbol{Q(r)}$ is a diagonal matrix of weights with elements given by

$$Q_{ii} = \frac{2a}{1 + a r_i^2}. \tag{8}$$

We can solve this problem iteratively by starting with an initial $\boldsymbol{r}$ such that we are able to compute $\boldsymbol{Q}$, then $\boldsymbol{r}$, update $\boldsymbol{Q}$ until reaching convergence as depicted in the following steps (modified from Sacchi 1997).

(1) Start with an initial receiver function $\boldsymbol{r}^{(0)}$.
(2) Select the hyperparameters, μ and a.
(3) Compute $\boldsymbol{Q}^{(0)}$.
(4) Iteratively solve eq. (7) using:

$$\boldsymbol{r}^{(l)} = \left[\boldsymbol{Z}^T \boldsymbol{Z} + \mu \boldsymbol{Q(r)}^{(l-1)}\right]^{-1} \boldsymbol{Z}^T \boldsymbol{R},$$

where l is the iteration number.

(5) The iterative scheme is stopped when the tolerance is satisfied:

$$2\frac{\left|\boldsymbol{J}^{(l)} - \boldsymbol{J}^{(l-1)}\right|}{\left(\left|\boldsymbol{J}^{(l)}\right| + \left|\boldsymbol{J}^{(l-1)}\right|\right)} \leq \text{tolerance}.$$

(6) Compute the data misfit. Select new hyperparameters if the misfit is not satisfactory.

In each iteration, the weighting matrix $\boldsymbol{Q(r)}$ is modified.

A simultaneous deconvolution of several teleseismic events can be achieved, as in Gurrola *et al.* (1995), by modifying eq. (3):

$$\begin{pmatrix} \boldsymbol{R_1} \\ \boldsymbol{R_2} \\ \vdots \\ \boldsymbol{R_{NR}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Z_1} \\ \boldsymbol{Z_2} \\ \vdots \\ \boldsymbol{Z_{NR}} \end{pmatrix} \boldsymbol{r}, \tag{9}$$

where each $\boldsymbol{Z}_j$, and $\boldsymbol{R}_j [j = 1, \dots, NR$ (number of seismic records)] are the same as $\boldsymbol{Z}$, and $\boldsymbol{R}$ in (3) for the jth seismogram. In the

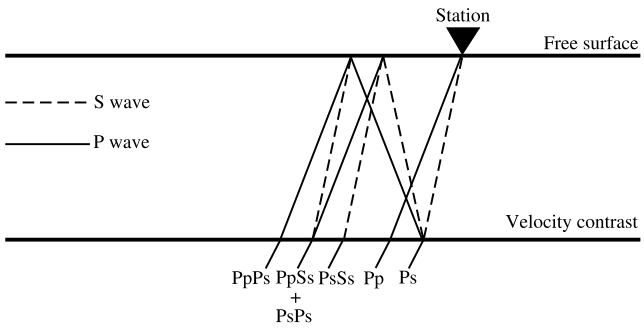

Figure 2. Simplified diagram showing the ray paths of the main conversions (*Ps*) and its multiples (*PsPs*, *PpSs* and *PpPs*). Except from the first arrival (*P*), upper-case letters indicate upgoing travel paths, and lower-case letters indicate downgoing travel paths (modified from Ammon 1991).

same manner as in (3), we can solve for the receiver function $\boldsymbol{R}$ as follows:

$$\boldsymbol{r} = \left[\sum_{j=1}^{NR} \boldsymbol{Z}_j{}^T \boldsymbol{Z}_j + \mu \boldsymbol{Q}(\boldsymbol{r}) \right]^{-1} \sum_{j=1}^{NR} \boldsymbol{Z}_j{}^T \boldsymbol{R}_j. \tag{10}$$

Therefore, a simultaneous deconvolution is no more complex than the deconvolution of a single seismogram (Gurrola *et al.* 1995).

The selection of the two hyperparameters μ and a in eq. (7) is very important; unfortunately there is no clear way to determine them *a priori* and a poor selection may lead us to an unreasonable solution. In general, we observed that μ must be small compared with a to preserve fidelity with the data and, at the same time, to obtain a sparse solution. As we increase a we enforce sparseness. In contrast, the solution approaches to a damped least-squares solution as a tends to 0.

The hyperparameters μ and a can be determined following the same procedure as in Sacchi (1997), that is, by using χ^2 statistic as a target misfit, $E[\chi^2] = N_{\text{obs}}$ (total number of observations), where the largest acceptable value is approximately $N_{\text{obs}} + 3.3\sqrt{N_{\text{obs}}}$ (Press *et al.* 1992). Using this test, we find multiple candidate pairs of parameters that produce similar misfits; however, when proper choices of μ and a are introduced, the resulting receiver functions only vary in sparseness but not in the presence and timing of the main phases.

The sparse deconvolution approach requires the inversion of a non-Toeplitz matrix n times, where n is the number of iterations. The computational cost of each inversion is proportional to M^3, where

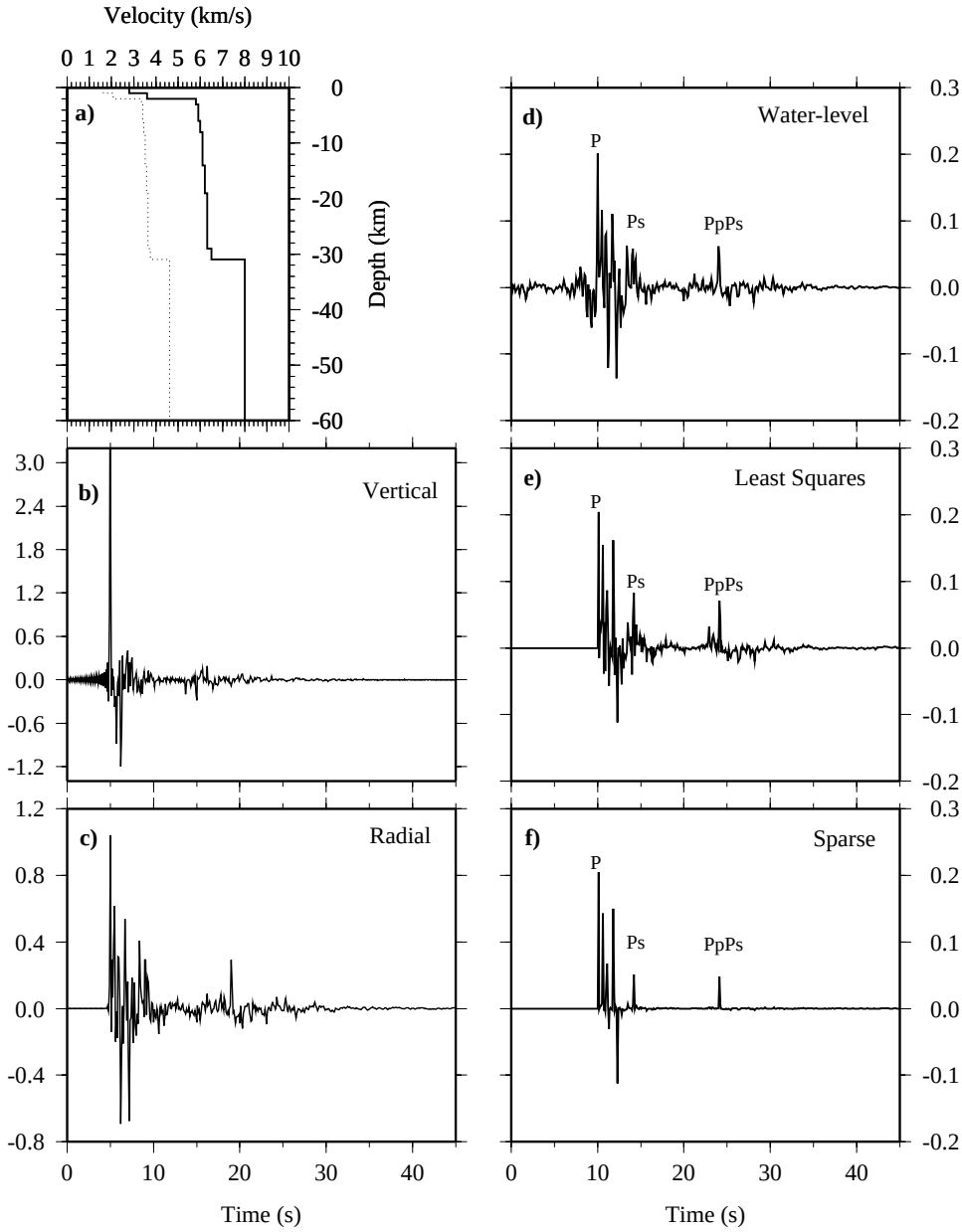

Figure 3. Estimated receiver functions for a multilayer velocity model: (a) velocity model (solid and dashed lines indicate *P* and *S* velocities, respectively); (b) vertical-component seismogram; (c) radial-component seismogram; (d) receiver function obtained from water-level deconvolution (Ammon 1991); (e) computed receiver function from least-squares deconvolution (Gurrola *et al.* 1995) and (f) estimated receiver function using the sparse deconvolution (this study).

M is the dimension of the matrix. In comparison, as shown in Eq. (5), the least-squares solution involves inverting a Toeplitz matrix that can be solved using fast methods such as Levinson recursion (Marple 1987). In this case, the number of computations needed is proportional to M^2 and, therefore, less expensive than the sparse deconvolution. However, with today's computer technology, such computational cost is insignificant. For instance, a receiver function of $N = 900$ samples and 15 iterations can be computed in 0.3 s with a 2.16 GHz Intel Core 2 Duo processor.

In the following sections, we will test the functionality and reliability of the simultaneous iterative time-domain sparse deconvolution in computing teleseismic receiver functions.

3 SYNTHETIC EXPERIMENTS

We performed three numerical experiments to test the effectiveness of the sparse deconvolution approach. We used the method of Randall (1989) to compute synthetic seismograms based on two simple velocity models: a two-layer velocity model of Ammon (1991), and a multilayer model of Ligorría & Ammon (1999). The resulting receiver functions are compared with two other approaches: water-level frequency-domain deconvolution (Ammon 1991) and time-domain least-squares deconvolution (Gurrola *et al.* 1995). In all cases, the resulting models ensure the same misfit (the rms difference between the estimated radial components and the observed radial components). As in Ligorría & Ammon (1999), we present the estimated receiver functions from water-level deconvolution using Gaussian width factors of 1.5 and 10 for the two-layer and multilayer models, respectively. In the case of time-domain techniques, we performed the χ^2 test described in the previous section.

Fig. 1 shows the computed receiver functions for the two-layer model. For this idealized case, all three methods retrieved high quality receiver functions with clear primary *Ps* conversions and multiples (Fig. 2). The arrival times of the majority of the signals agree

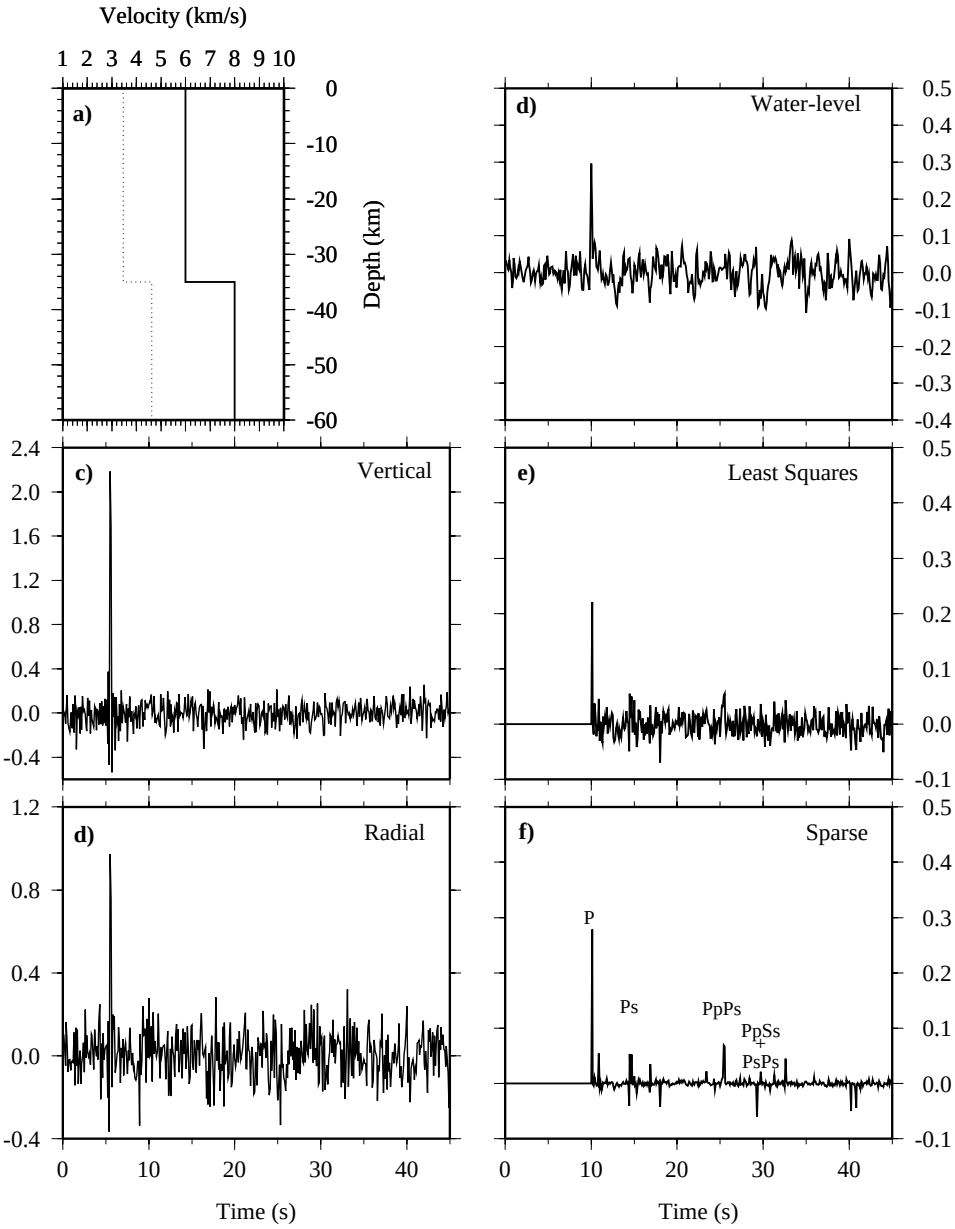

Figure 4. Same as in Fig. 1 but in this case we added noise to the vertical and radial component seismograms (signal-to-noise ratio = 0.1).

very well, though minor differences in signal amplitude are visible among the output of these three approaches. Our sparse deconvolution result appears to be more consistent with that of water-level deconvolution where little ringing, if any, is present on the receiver functions.

Consistent receiver functions can be uniformly obtained even when the input model is significantly more complex (Fig. 3) than the previous case. This is expected since the data are broadband with good signal-to-noise levels, thus the differences between these three methods are insignificant. In Fig. 3, the sparse deconvolution approach shows a pronounced zero-phase arrival at approximately 10 s. This is highly consistent with the results of the least-squares deconvolution; both highlight the zero-phase nature of the primary *P* arrival after deconvolution. In general, the sparse deconvolution slightly increase the resolution of closely spaced phases on the receiver functions. These two experiments show the functionality and reliability of the sparse deconvolution technique in determining receiver functions for relatively simple models and well-recorded data sets.

To demonstrate the distinct advantages of our sparse deconvolution approach over previous techniques. We added noise to the original vertical and radial component seismograms for the two-layer model. The computed receiver functions are presented in Fig. 4. All the results, as in the previous tests, ensure similar misfits for a fair comparison of the models. Although not as clear as in the noise-free case (Fig. 1), the sparse deconvolution approach is able to highlight the primary conversions and its multiples. Thus, we show the ability of our approach to compute receiver functions when dealing with less ideal data. In the following section, we use band-limited seismic data recorded in Japan to demonstrate the effectiveness of our technique to compute high-resolution receiver functions.

4 ESTIMATION OF RECEIVER FUNCTIONS IN SOUTHERN JAPAN

We used seismic data recorded at station ABU (Southern Japan), from the F-net broad-band seismic array, during 2003. A total of 88 teleseismic events were selected in the distance range between 30° and 90° with magnitudes m_b greater than 5.5 (Fig. 5) and the event-station pairs naturally form four azimuthal bins. The small data set meets the primary objective of our pilot study, which is to evaluate the suitability of the sparse deconvolution as means to estimate high-resolution receiver functions. We used a time window of 120 s, starting from 30 s before the *P* arrival.

We then performed both least-squares deconvolution (Fig. 6a) and sparse deconvolution (Fig. 6c) to determine the radial receiver functions for the NE azimuthal direction. We used hyperparemeters ($\lambda = 10$; $\mu = 0.014$ and $a = 10\,000$) that produce the same misfit as shown in Fig. 6(b) where the trade-off curves for the two approaches are displayed. In this figure, the intercepts between the trade-off curves and the dotted line indicate the hyperparameters used to compute the receiver functions.

Fig. 6(c) demonstrates that the sparse deconvolution approach is able to provide much greater resolution and improve the identification of important phases in the receiver functions when using seismic data (band-limited data). The overall improvement results from the sparseness constraint that creates the simplest earth model possible. This will increase the frequency content since the result

Figure 5. Azimuthal orthographic projection of teleseismic events recorded at station ABU, Southern Japan, during 2003. Only events from four azimuthal direction were used in this study. Equidistant lines at every 30° are shown by concentric circles. Solid lines represent ray paths.

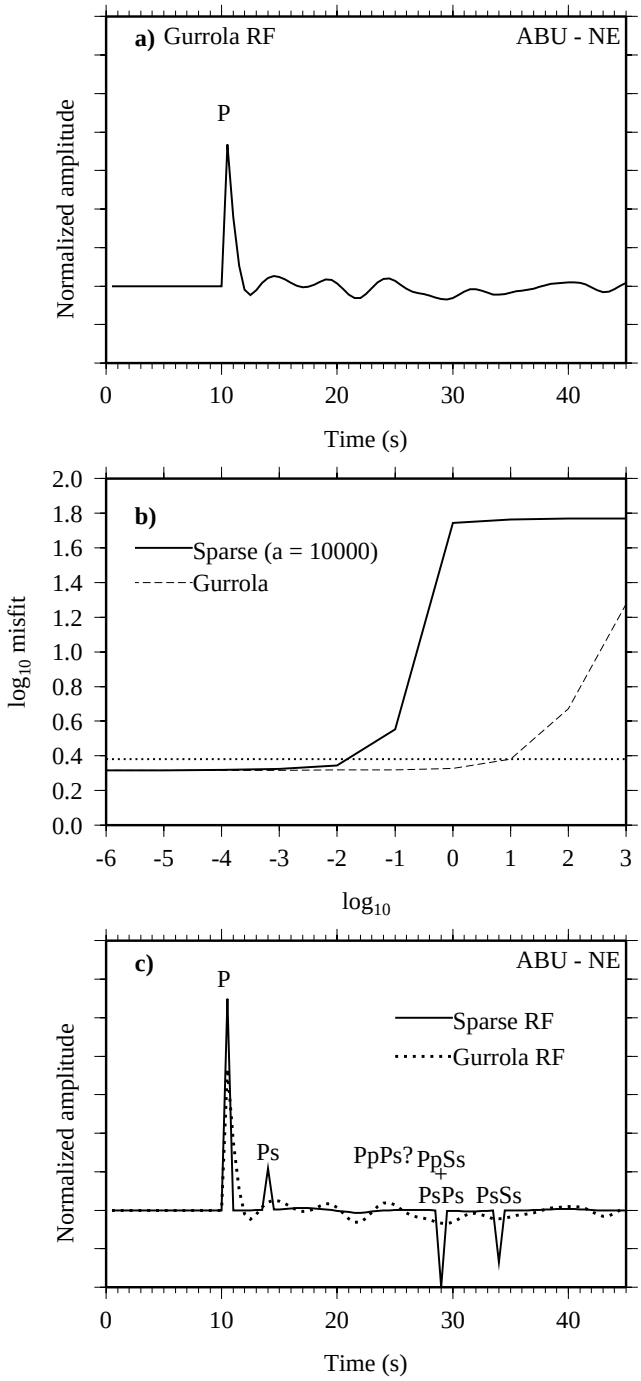

Figure 6. A comparison of the estimated receiver function for the NE azimuthal direction using least-squares deconvolution and the sparse deconvolution approach. (a) Receiver function obtained from least-squares deconvolution (Gurrola *et al.* 1995). (b) Trade-off curves for the least-squares approach (dashed line) and the sparse deconvolution approach (solid line). The dotted line marks the hyperparameters (intercepts with trade-off curves) used to estimate the receiver functions that produce the same misfit. (c) Receiver function determined using the sparse deconvolution (solid line); the result displayed in (a) is also shown (dashed line) for a better comparison.

is composed of a superposition of broad-band spikes. Therefore, by using the sparseness constraint we are able to recover the high frequency component of the data (Sacchi 1997).

To observe the effect of the selection of hyperparameters μ and a, we computed several receiver functions using different hyperparameters while preserving the misfit for a fair comparison of the results (Fig. 7). It is shown that as a tends to 0 the model approaches the least-squares solution (Fig. 6a), whereas increasing a will enhance the sparseness of the solution. In general, all the solutions are similar after reaching certain a value.

Fig. 8 shows the estimated and observed radial-component seismograms after using the sparse deconvolution approach for the NE azimuthal direction. Clearly, there is a remarkable fit between the calculated seismograms and the observations, although small differences are observed, especially at the later part of the radial components.

The computed receiver functions for all azimuthal directions (Fig. 9) show the same pattern with *Ps* conversions arriving 3.5–4.0 s after the *P* onset. Some multiples are also observed such as *PpPs*, *PpSs* + *PsPs* and *PsSs* at approximately 14.0, 19.0 and 25.0 s, respectively, after the *P* arrival. The similarity of the resulting receiver functions reflects the stability of the sparse deconvolution approach. Moreover, Li *et al.* (2000) used receiver functions to image the crustal and upper-mantle discontinuities beneath Japan with more than 5000 teleseismic records. They found *Ps* conversions at around 4 s after the primary *P* arrival in areas close to station ABU which is consistent with our results.

Slight differences can be observed among the receiver functions (Fig. 9), especially in the timing of the different phases. These differences arise due to the lack of distance and heterogeneity corrections since our main purpose is to demonstrate the functionality of this technique, rather than to provide a detailed seismic interpretation. Furthermore, the *PpPs* multiple is notably missing on the receiver function from the NE azimuthal direction. This is because the sparse deconvolution approach tends to emphasize the most prominent features and remove small/weak phases; relatively few events (4) used in the receiver function calculations (Fig. 5) may also contribute to the lack of this phase. In other words, a comparable resolution of the *PpPs* with synthetic examples shown in Fig. 1 will require data with a greater bandwidth, a denser azimuthal coverage and a high signal-to-noise ratio. It should be noted that a single receiver function using the events from all azimuthal directions can be computed, but an accurate interpretation would require proper corrections prior to the inversion procedure. Otherwise, the receiver function will be highly affected by the different traveltime moveouts.

From a broader perspective, the sparse deconvolution approach presented here nicely complement the ongoing global effort in providing a high-resolution image for seismic structure and discontinuity depths. Our approach can be equally effective in computing transverse receiver functions (Park & Levin 2000; Levin *et al.* 2002) that are highly sensitive to seismic anisotropy. The sparseness constraint can significantly improve the accuracy of phase identification and timing of important phases such as *P*-to-*S* conversions. This directly translates to improved analyses of structure, interface, and anisotropy in the Earth's crust and mantle.

5 CONCLUSION

We present the sparse deconvolution as an alternative approach to existing methods such as water-level deconvolution or least-squares deconvolution for a receiver function estimation. The functionality and reliability of this new approach are examined through the applications on both synthetic and observed seismic data. The results clearly demonstrate that the simultaneous iterative time-domain sparse deconvolution is a suitable method for the determination of high-resolution telesesimic receiver functions.

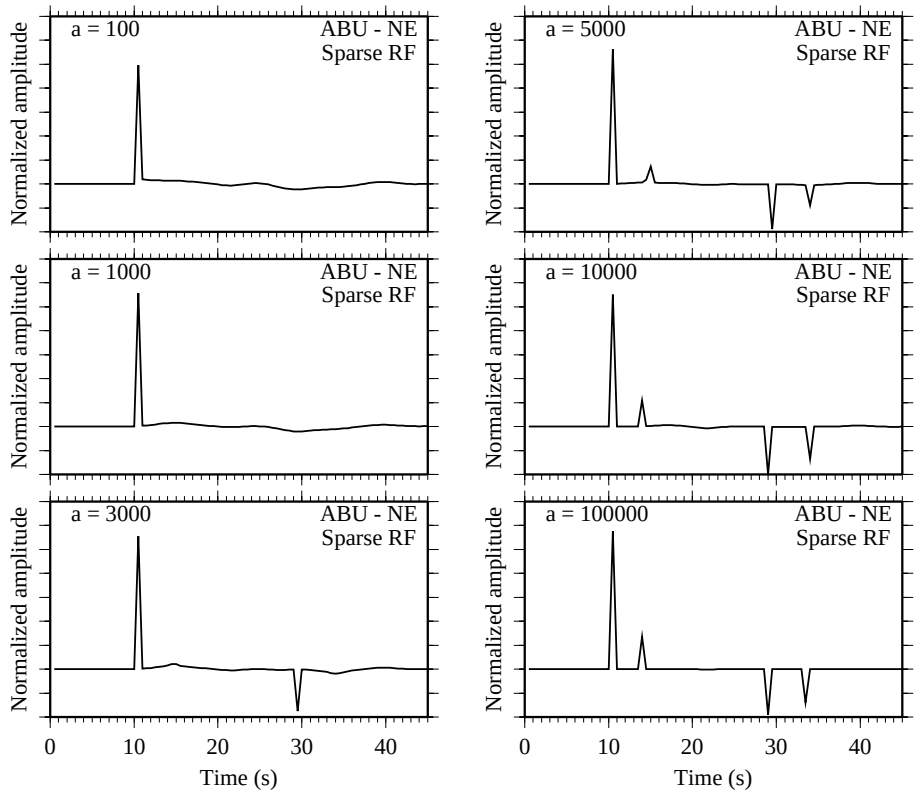

Figure 7. Receiver functions computed for different hyperparameters a and μ. All the resulting models ensure the same misfit. Higher a values enforce sparseness to the model. Conversely, the receiver function approaches the damped least-squares solution when a approaches 0.

ACKNOWLEDGMENTS

We are grateful to Andrew Curtis and two anonymous reviewers for their constructive suggestions and comments. We thank the National Research Institute for Earth Science and Disaster Prevention of Japan for providing the F-net data. This research is funded by NSERC (EQPEGQ 300040, RGPIN 288160-04), the Canadian Foundation for Innovation and Alberta Ingenuity.

REFERENCES

Abers, G.A., Hu, X. & Sykes, L.R., 1995. Source scaling of earthquakes in the Shumagin Region, Alaska: time-domain inversions of regional waveforms, *Geophys. J. Int.,* **123,** 41–58.

Ammon, C.J., 1991. The isolation of receiver effects from teleseismic *P* waveforms, *Bull. seism. Soc. Am.,* **81,** 2504–2510.

Ammon, C.J., Randall, G.E. & Zandt, G., 1990. On the nonuniqueness of receiver function inversions, *J. geophys. Res.,* **95,** 15 303–15 318.

Bostock, M.G., 1996. *Ps* conversions from the upper mantle transition zone beneath the Canadian landmass, *J. geophys. Res.,* **101,** 8393–8402.

Clayton, R.W. & Wiggins, R.A., 1976. Source shape estimation and deconvolution of teleseismic body waves, *Geophys. J. R. astr. Soc.,* **47,** 151–177.

Dueker, K.G. & Sheehan, A.F., 1997. Mantle discontinuity structure from midpoint stacks of converted *P* to *S* waves across the Yellowstone hot spot track, *J. geophys. Res.,* **102,** 8313–8327. teleseismic body waves, *Geophys. J. R. astr. Soc.*, 47, 151–177.

Flanagan, M.P. & Shearer, P.M., 1998. Global mapping of topography on transition zone velocity discontinuities by stacking SS precursors, *J. geophys. Res.,* **103,** 2673–2692.

Flanagan, M.P. & Shearer, P.M., 1999. A map of topography on the 410-km discontinuity from *PP* precursors, *Geophys. Res. Lett.,* **26,** 549–552.

Gu, Y.J. & Dziewonski, A.M., 2002. Global variability of transition zone thickness, *J. geophys. Res.,* **107,** doi:10.1029/2001JB000489.

Gu, Y.J., Dziewonski, A.M. & Agee, C.B., 1998. Global de-correlation of the topography of transition zone discontinuities, *Earth Planet. Sci. Lett.,* **157,** 57–67.

Gurrola, H. & Minster, J.B., 1998. Thickness estimates of the upper-mantle transition zone from bootstrapped velocity spectrum stacks of receiver functions, *Geophys. J. Int.,* **133,** 31–43.

Gurrola, H., Baker, G.E. & Minster, J.B., 1995. simultaneous time domain deconvolution with application to the computation of receiver functions, *Geophys. J. Int.,* **120,** 537–543.

Kind, R., Kosarev, G.L. & Petersen, N.V., 1995. Receiver functions at the stations of the German Regional Seismic Network (GRSN), *Geophys. J. Int.,* **121,** 191–202.

Langston, C.A., 1977. The effect of planar dipping structure on source and receiver responses for constant ray parameter, *Bull. seism. Soc. Am.,* **67,** 1029–1050.

Langston, C.A., 1979. Structure under Mount Rainer, Washington, inferred from teleseismic body waves, *J. geophys. Res.,* **84,** 4749–4762.

Langston, C.A., 1981. Evidence for the subducting lithosphere under southern Vancouver Island and western Oregon from teleseismic *P* wave conversions, *J. geophys. Res.,* **86,** 3857–3866.

Lawrence, J.F. & Shearer, P.M., 2006. A global study of transition zone thickness using receiver functions, *J. geophys. Res.,* **111,** doi:10.1029/2005JB003973.

Lee, D.-K. & Grand, S., 1996. Depth of upper mantle discontinuities beneath the East Pacific Rise, *Geophys. Res. Lett.,* **23,** 3369–3372.

Levin, V., Park, J., Brandon, M., Lees, J., Peyton, V., Gordeev, E. & Ozerov, A., 2002. Crust and upper mantle of Kamchatka from teleseismic receiver functions, *Tectonophys,* **358,** 233–265.

Li, X., Sobolev, S.V., Kind, R., Yuan, X. & Estabrook, Ch., 2000. A detailed receiver function image of the upper mantle discontinuities in the Japan subduction zone, *Earth Planet. Sci. Lett.,* **183,** 527–541.

Ligorría, J.P. & Ammon, C.J., 1999. Iterative deconvolution and receiver-function estimation, *Bull. seism. Soc. Am.,* **89,** 1395–1400.

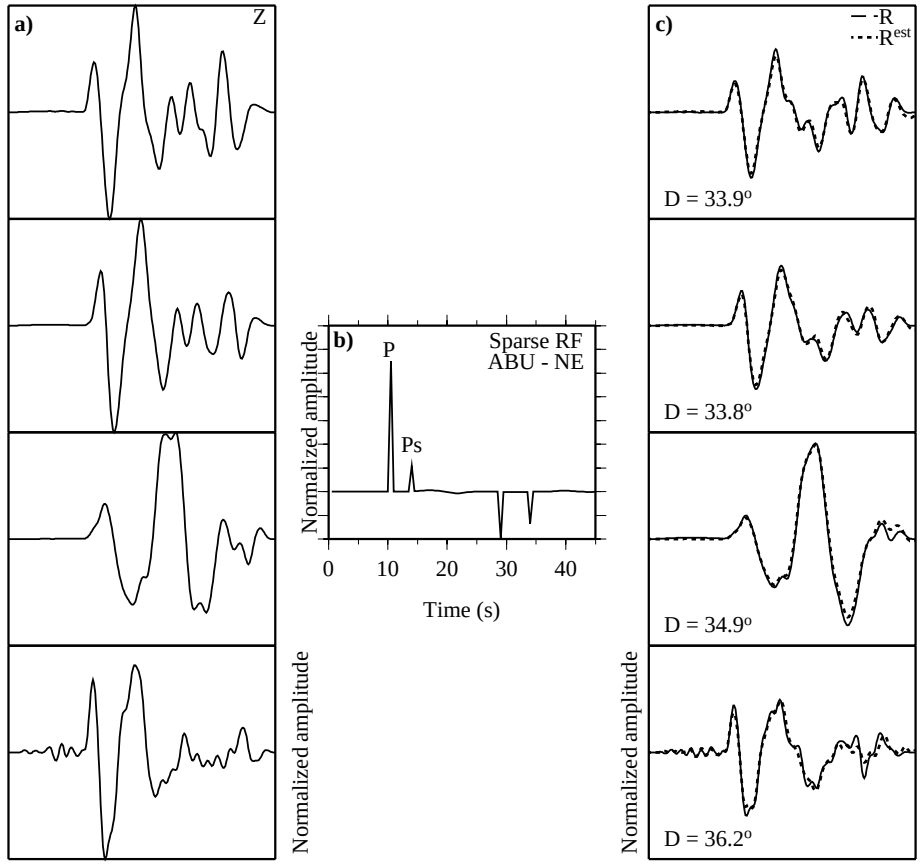

Figure 8. Estimated radial-component seismograms from events located in the NE azimuthal direction. (a) Vertical-component seismograms; (b) receiver function estimated using time-domain sparse deconvolution and (c) estimated (dashed lines) and observed (solid lines) radial-component seismograms. The epicentral distances are also indicated.

Marple, S.L., Jr., 1987. *Digital Spectral Analysis with Applications*, Prentice-Hall, New Jersey.

Mendel, J.M., 1990. *Maximum Likelihood Deconvolution*, Springer, New York.

Menke, W. & Levin, V., 2004. The cross-convolution method for interpreting *SKS* splitting observations, with application to one and two-layer anisotropic earth models, *Geophys. J. Int.*, **154,** 379–392.

Oldenburg, D.W., 1981. A comprehensive solution to the linear deconvolution problem, *Geophys. J. R. astr. Soc.*, **65,** 331–357.

Olofsson, T., 2004. Semi-sparse deconvolution robust to uncertainties in the impulse responses, *Ultrasonics*, **42,** 969–975.

Owens, T.J. & Crosson, R.S., 1988. Shallow structure effects on broadband teleseismic *P P* waveforms, *Bull. seism. Soc. Am.*, **78,** 96–108.

Park, J. & Levin, V., 2000. Receiver functions from multiple-taper spectral correlation estimates, *Bull. seism. Soc. Am.*, **90,** 1507–1520.

Pham, T. & de Figueiredo, R., 1989. Maximum likelihood estimation of a class of non-Gaussian densities with application to l_p deconvolution, *IEEE Trans. Signal Proc.*, **37,** 73–82.

Petersen, N., Vinnik, L.P., Kosarev, G., Kind, R., Oreshin, S. & Stammler, K., 1993. Sharpness of the mantle discontinuities, *Geophys. Res. Lett.*, **20,** 859–862.

Press, W., Teukolsky, S., Vettering, W. & Flannery, B., 1992. *Numerical Recipes*, Cambridge University Press, Cambridge.

Randall, G.E., 1989. Efficient calculation of differential seismograms for lithospheric receiver functions, *Geophys. J. Int.*, **99,** 469–481.

Ringwood, A.E. & Irifune, T., 1988. Nature of the 650-km seismic discontinuity - implications for mantle dynamics and differentiation, *Nature*, **331,** 131–136.

Sacchi, M.D., 1997. Reweighting strategies in seismic deconvolution, *Geophys. J. Int.*, **129,** 651–656.

Shearer, P.M., 2000. Upper mantle seismic discontinuity, in *Earth's Deep Interior: Mineral Physics and Tomography from the Atomic to the Global Scale, AGU Geophysical Monograph*, Vol. 117, 115–131. *Nature*, 355, 791–796.

Shearer, P.M. & Masters, G., 1992. Global mapping of topography on the 660-km discontinuity, *Nature*, **355,** 791–796.

Sheehan, A.F., Abers, G.A., Jones, C.H. & Lernerlam, A.L., 1995. Crustal thickness variations across the Colorado Rocky-Mountain from teleseismic receiver functions, *J. geophys. Res.*, **100,** 20 391–20 404.

Shen, Y., Solomon, S.C., Bjarnason, I.T. & Purdy, G.M., 1996. Hot mantle transition zone beneath Iceland and the adjacent Mid-Atlantic Ridge inferred from *P*-to-*S* conversions and the 410- and 660-km discontinuities, *Geophys. Res. Lett.*, **23,** 3527–3530.

Shen, Y., Sheehan, A.F., Dueker, K.G., de Groot-Hedlin, C. & Gilbert, H., 1998. Mantle discontinuity structure beneath the southern East Pacific Rise from *P*-to-*S* converted phases, *Science*, **280,** 1232–1235.

Stixrude, L., 1997. Structure and sharpness of phase transitions and mantle discontinuities, *J. geophys. Res.*, **102,** 14 835–14 852.

Vinnik, L.P., Kosarev, G. & Petersen, N., 1996. Mantle transition zone beneath Eurasia, *Geophys. Res. Lett.*, **23,** 1485–1488.

Yuan, X., Ni, J., Kind, R., Sandvol, E. & Mechie, J., 1997. Lithospheric and upper mantle structure of southern Tibet from a seismological passive source experiment, *J. geophys. Res.*, **102,** 27 491–27 500.

Zhu, L., 2004. Lateral variation of the Tibetan lithospheric structure inferred from teleseismic waveforms, in *Advancements in Seismology and Physics of the Earth Interior in China*, Seismology Press, Beijing.

Zhu, L. & Kanamori, H., 2000. Moho depth variation in Southern California from teleseismic receiver functions, *J. geophys. Res.*, **105,** 2969–2980.

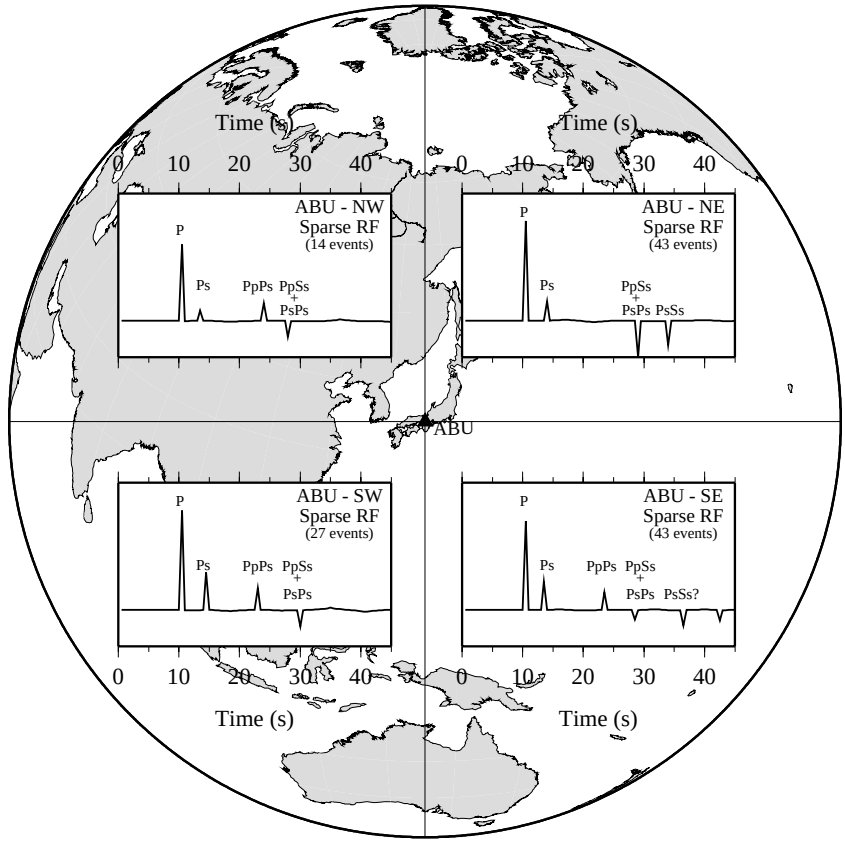

Figure 9. Receiver functions at station ABU (Southern Japan) estimated by the simultaneous iterative sparse deconvolution for all azimuthal directions. The number of events used to compute the receiver functions are also indicated. There are slight differences among the receiver functions, especially in the timing of the phases due to the lack of distance and heterogeneity corrections to the receiver functions.