

Improvement of coda phase detectability and reconstruction of global seismic data using frequency–wavenumber methods

Simon Schneider,^{1,*} Christine Thomas,¹ Ramin M.H. Dokht,² Yu Jeffrey Gu² and Yunfeng Chen²

¹*Institut für Geophysik, Westfälische Wilhelms-Universität Münster, Corrensstraße 24, D-48149 Münster, Germany.*

E-mail: simon.andreas.schneider@gmail.com

²*Department of Physics, University of Alberta, Edmonton, AB T6G2E1, Canada*

Accepted 2017 November 1. Received 2017 October 30; in original form 2017 April 5

SUMMARY

Due to uneven earthquake source and receiver distributions, our abilities to isolate weak signals from interfering phases and reconstruct missing data are fundamental to improving the resolution of seismic imaging techniques. In this study, we introduce a modified frequency–wavenumber (fk) domain based approach using a ‘Projection Onto Convex Sets’ (POCS) algorithm. POCS takes advantage of the sparsity of the dominating energies of phase arrivals in the fk domain, which enables an effective detection and reconstruction of the weak seismic signals. Moreover, our algorithm utilizes the 2-D Fourier transform to perform noise removal, interpolation and weak-phase extraction. To improve the directional resolution of the reconstructed data, we introduce a band-stop 2-D Fourier filter to remove the energy of unwanted, interfering phases in the fk domain, which significantly increases the robustness of the signal of interest. The effectiveness and benefits of this method are clearly demonstrated using both simulated and actual broadband recordings of PP precursors from an array located in Tanzania. When used properly, this method could significantly enhance the resolution of weak crust and mantle seismic phases.

Key words: Fourier analysis; Spatial analysis; Time-series analysis; Body waves; Coda waves.

INTRODUCTION

Investigations of the Earth’s internal structures often involve seismic arrays and array-based techniques that shift and stack recordings for different slowness and backazimuth values (e.g. Rost & Thomas 2002). In many applications, vespagrams (Muirhead & Datt 1976; Rost & Thomas 2002) are adopted to identify or detect small amplitude arrivals by using the slowness of a wave as the discriminating factor. Examples include arrivals from the lowermost mantle (e.g. Cobden & Thomas 2013) or P wave underside reflections off mantle transition zone discontinuities that arrive as small amplitude precursors to the PP phase (i.e. PP precursors; e.g. Rost & Weber 2001; Deuss 2009; Saki *et al.* 2015).

While signal detection is simple in the case of isolated arrivals, the problem becomes much more complex when multiple seismic phases arrive in the same time window. In the case of PP precursors, the strong coda of the diffracted P wave or the reflected wave at the inner core boundary (PKiKP) may interfere with low-amplitude precursors, thereby hampering phase identification and analysis

(e.g. Lessing *et al.* 2015). Despite high signal-to-noise ratios of the main phases, interferences on weak secondary phases such as PP or SS precursors often result in the removal of significant fractions of potentially usable data. The resultant, often mistaken, the absence of underside reflections from seismic discontinuities (e.g. 410 and 660) is conducive to misinterpretations of the temperature and/or mineralogy of the mantle transition zone. Therefore it is critical to minimize the phase interference between strong arrivals and weak signals of interest for a robust phase identification and extraction on a vespagram.

The most commonly adopted approach is stacking (delay-and-sum) due to its simplicity and effectiveness in suppressing incoherent (though not necessarily low-amplitude) arrivals. Recently, more sophisticated resolution enhancement approaches such as Radon transform (Gu & Sacchi 2009; Schultz & Gu 2013), local-slant-stack transform (Ventosa *et al.* 2012) and slant-stacklet transform (Ventosa & Romanowicz 2015) have been proposed. The local-slant-stack transform was successfully applied to long-period SS precursors (Zheng *et al.* 2015), which effectively differentiates between seismic arrivals based on moveout or epicentral distance. Other noteworthy methods, which are primarily used in small-scale seismic surveys to reconstruct gaps in data, include minimum weighted norm interpolation (Naghizadeh & Sacchi 2010),

*Now at: Department of Earth Sciences, Utrecht University, 3508 TC Utrecht, The Netherlands.

Singular Spectrum Analysis (SSA; Oropeza & Sacchi 2011; Dokht *et al.* 2017) and least-squares de-noising and interpolation (Naghizadeh 2012). These rank-reduction based approaches transform the data to frequency–space (fx) or frequency–wavenumber (fk) domains to reconstruct missing information, and have been successfully extended to multidimensional seismic analysis. For example, Kreimer *et al.* (2013) simultaneously remove the noise and reconstruct a 5-D seismic gather using tensor completion in the fx domain, and Chen *et al.* (2016) further improves the robustness of 5-D data reconstruction using the damped truncated singular value decomposition. These powerful reconstruction methods, however, have not been well-documented in global seismic studies. Only recently has SSA been utilized to simultaneously remove the random noise and reconstruct missing traces of SS precursors and P -to- S conversions from mantle seismic discontinuities (Dokht *et al.* 2017).

In this study, we introduce a novel fk based method for the reconstruction and resolution improvement of broadband seismic array data. Our method takes advantage of the 2-D Fourier transform, which enables an effective detection of sparse dominant energies corresponding to the weak seismic signals of interest in the fk domain. The following sections will provide the details of this method, which incorporates a number of useful algorithms in the fk domain for noise removal, interpolation and weak-phase extraction.

METHOD

Data regularization

One of the prerequisites for 2-D fast Fourier transformation (FFT) is uniform sampling in both temporal and spatial directions. Uniform sampling is not necessarily needed for the discrete Fourier transformation, however, many seismic applications use FFT and therefore require uniform sampling in space and time. The uniform sampling in time is generally fulfilled by digital seismic instruments via a constant data acquisition rate, the latter requirement could be more problematic in practice since seismic receivers, especially those in regional arrays, cannot always be deployed in a uniform or near-uniform pattern. As a result, data regularization, for example, interpolation and resampling, are required prior to data analysis and interpretation. In this study, we introduce two methods to produce a uniform grid where seismic traces are evenly distributed, both of which assume that the structure in the region of study is well-represented by a 1-D model (i.e. the lateral velocity variation is small). Any deviation from this assumption would introduce a travel time deviation and hence a mislocation vector to the data (Krüger & Weber 1992, Jacobeit *et al.* 2013).

One example of non-uniform spatial sampling is shown in Fig. 1(a) for a synthetic seismogram section of three coherent linear arrivals with two different slowness values. In this section, the data sampling is particularly dense at between 100° – 104° , whereas a clear gap is present at the epicentral range from 106° to 110° .

The first method, which we refer to as ‘Static Resampling’, takes the largest interstation spacing and arranges the data on a uniform grid by applying a partial stacking scheme (Fig. 1b). The resampled grid contains the same number of traces at the same aperture as the original data set. To implement this method, the existing traces are binned and repositioned to the new node location (centre of the bin; Fig. 1b), where all data will be stacked to form the new trace if data exist and otherwise the traces will be left empty.

This method assumes a linear moveout of the target phase, defined as a seismic arrival that is influenced by a coda or a different phase. If the sizes of the bins are too large (e.g. greater than 0.5° for the example in Fig. 4), then the seismic arrivals will be shifted in time to correct for the moveout before stacking. The amount of time shift is calculated from the arrival time difference between the reference phase of the original trace (e.g. the first peak of each trace in Fig. 1a) and that of the new (non-empty) trace based on PREM (Dziewonski & Anderson 1981). In our case, we utilize PP as a reference phase to correct for the moveout of PP precursors due to their similar slowness values. Similarly, phases like SS could also be used to correct for SS precursors. A similar approach has been applied to the construction of a common midpoint gather in exploration seismology (Yilmaz 2001). We will discuss problems due to nonlinear moveout below.

An alternative spatial resampling procedure is to redistribute the traces using a moving window, which we refer to as ‘Sliding Window resampling’. Similar to the ‘Static Resampling’ approach, this method partially stacks the entire data for a constant bin size (b). Additionally, a sliding window of size Δ is introduced during the partial stacking procedure to allow for the overlap between adjacent bins. This method essentially applies a moving average to the data, which slightly reduces the spatial resolution but improves the coherency of the nearby traces. Specifically, after populating the first new trace, the bin is shifted by $b - \Delta$ and stacked. In contrast to the ‘Static Resampling’, this method does not limit the number of new traces to the same as the original array, which permits a more flexible choice of the spatial sampling interval. This method assumes a linear moveout of the target phase just in each bin.

Data reconstruction

Despite the uniform spatial distribution, our resampled traces can still suffer from large data gaps at certain distances (see 106° – 108° in Figs 1b and c). We use a Projection onto Convex Sets (POCS) algorithm (Abma & Kabir 2006), a compressive sensing based technique (Candes & Romberg 2007), to interpolate the missing traces. The iterative algorithm is similar to the one suggested by Dokht *et al.* (2017):

$$\mathbf{d}_{i+1} = \mathbf{d}_i + SFT^{-1}(\alpha \mathbf{T}_i(\mathbf{D}_i)) \quad (1)$$

where \mathbf{d}_i represents the data matrix (with missing traces) in the time-space (tx) domain for the i th iteration, and \mathbf{D}_i denotes the corresponding fk spectrum. The operator \mathbf{S} represents the sampling matrix (same size as the original data), with values of 1 where data are missing and 0 elsewhere. The operator FT^{-1} represents the inverse Fourier Transform and \mathbf{T}_i is the threshold operator, which sets all values of amplitude of \mathbf{D}_i below the threshold to 0. The parameter α is a predefined factor that lowers the threshold after each iteration. The initial conditions of eq. (1) are defined by

$$\mathbf{T}_0 = \max(|\mathbf{D}_0|) \quad (2)$$

where \mathbf{T}_0 and \mathbf{D}_0 represent the initial threshold operator and Fourier spectrum of the input data \mathbf{d}_0 , respectively.

We explain this process using a synthetic data set with a missing trace (Fig. 2). The data set is transformed into the fk domain, where the maximum absolute value of the fk spectrum is picked and defines the first threshold value for the data reconstruction. The fk amplitude spectrum below this threshold will be muted by setting all values to zero (Fig. 2b). The new fk domain signal, which contains

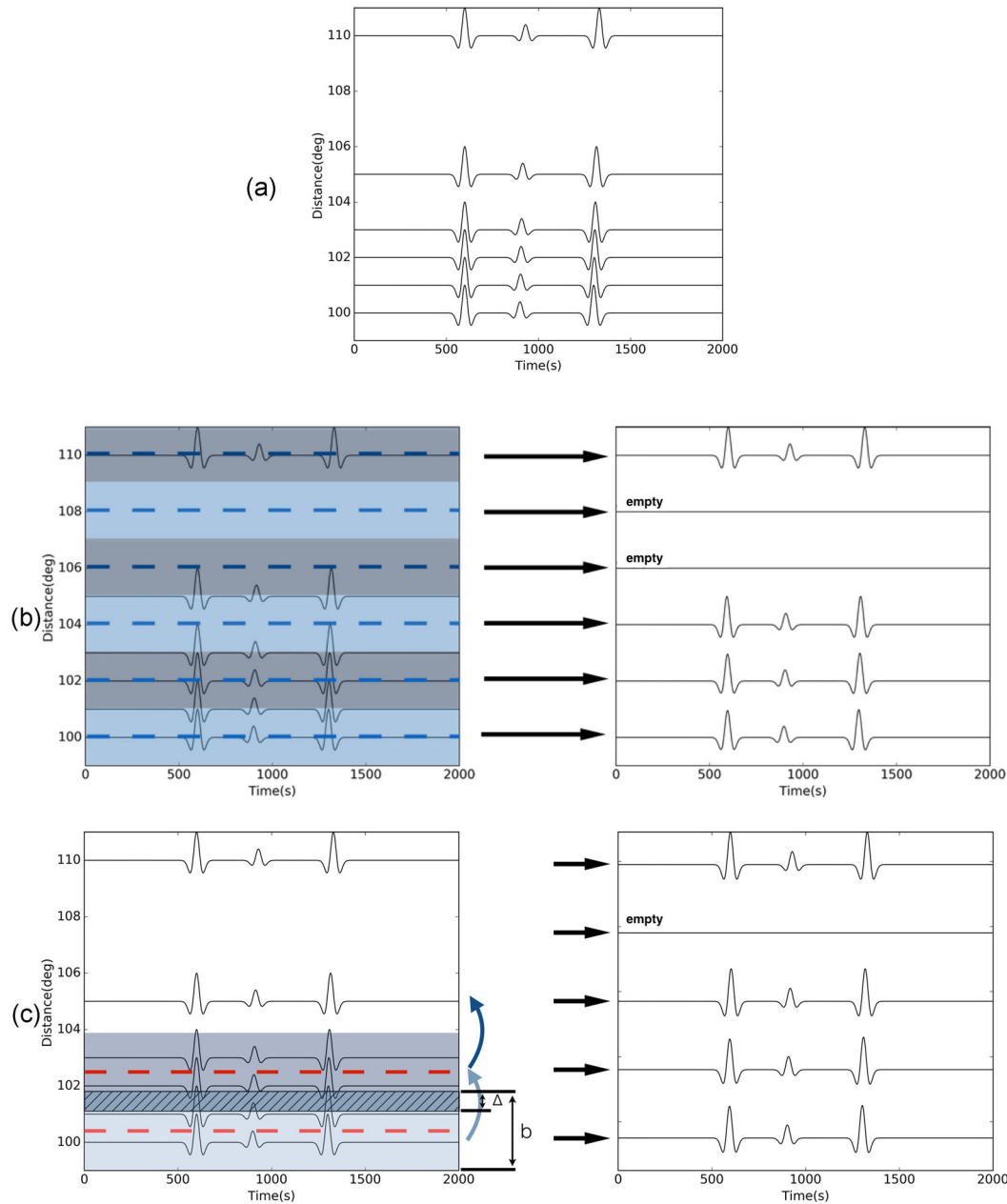


Figure 1. (a) A spatially non-uniformly distributed data set with three different coherent signals. (b) ‘Static Resampling’ scheme where the new equidistant grid is depicted with blue dashed lines. The old traces (left panel) within a bin centred at grid position are stacked to form the new trace (right panel), which is assigned empty if no trace is present in the bin (e.g. at 106° and 108° in the right panel). (c) ‘Sliding Window’ resampling method with a bin of size b (light blue area). In our example a large bin size of 3° and overlap window (Δ) of 30 per cent are used, which would be much smaller in real application. Data within the moving window are stacked and assigned to the centre of the bin (red dashed line), whereas the bin is left empty if no trace contributes to the stacking.

non-zero values only above the threshold, will be transformed back to the tx domain. This transformation procedure partially recovers the signal, which represents the dominant frequency and wavenumber of the data set, on the formerly empty trace. This new trace is subsequently reinserted into the original data set, while the rest stays untouched (Fig. 2c).

The output of the first iteration (Fig. 2c) is then used as input for the next iteration, which is subsequently transformed back to the fk domain. The threshold fk amplitude in the second iteration is lowered (i.e. α per cent of that of the first iteration), below which all signals are muted again (Fig. 2d). This procedure is repeated, with a gradually decreasing threshold after each iteration, until a

predefined number of iterations is reached or the difference between iterations ($\mathbf{d}_{i+1} - \mathbf{d}_i$) is small enough (e.g. below a predefined cut-off value). In this example the signals (i.e. amplitude and wavelength) in the tx domain remain nearly unchanged after six iterations, which is evidence for convergence. The fk spectrum is now broadened by including the signals outside the dominant slopes, which results in increased frequency and wavenumber contents, and therefore the recovery of greater details in the tx domain.

This algorithm is dependent on two variables, α as the decreasing factor of the threshold value and the number of iterations. To determine the optimal α and the number of iterations, it is useful to test various combinations of these two parameters and compare the

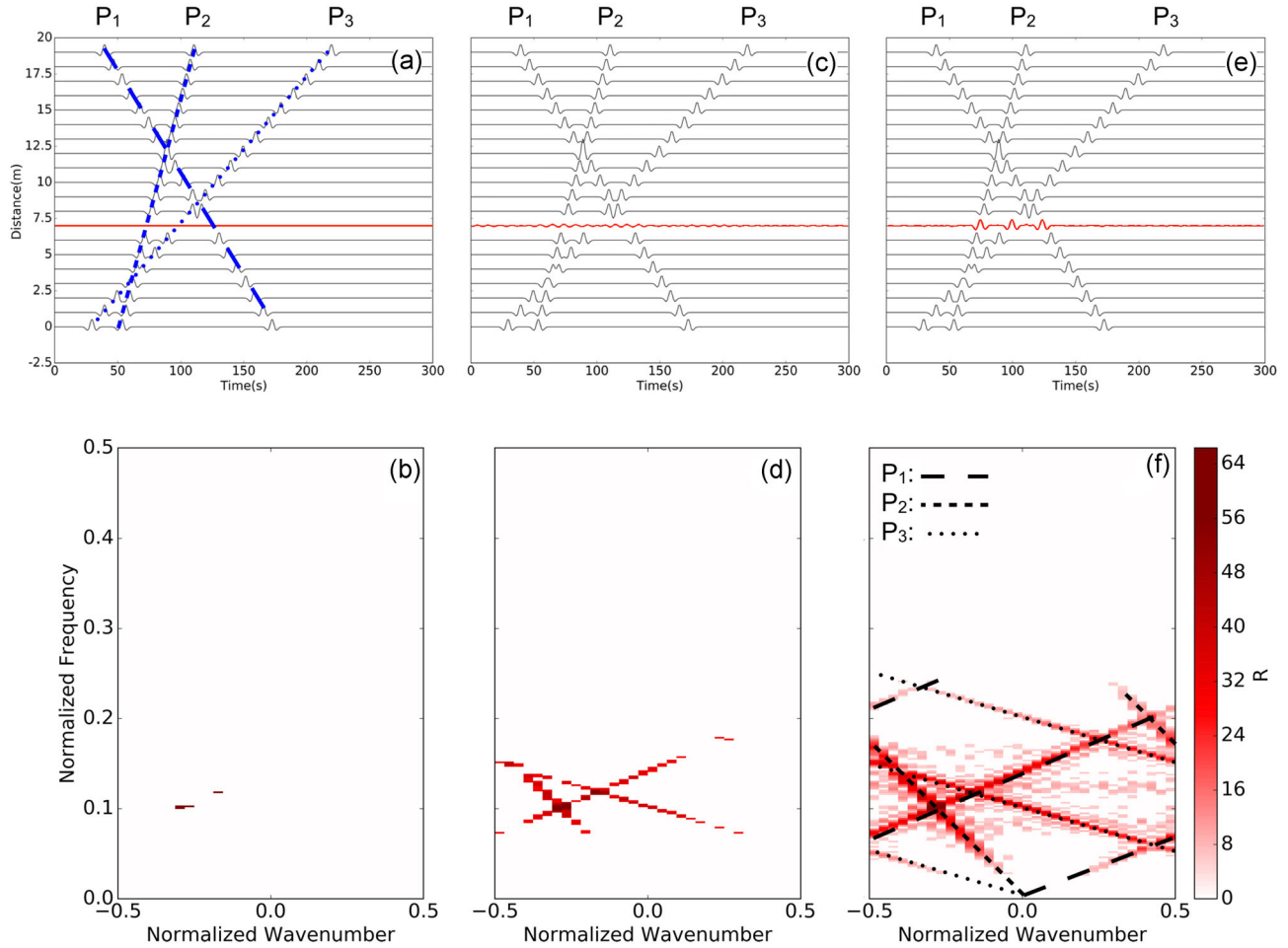


Figure 2. (a) Example data set contains three coherent events with different slowness values (P_1 , P_2 , P_3 , marked with blue dashed and dotted lines). A gap (red) exists at 7° distance due to the missing trace. (b) The fk -spectrum of the data set in (a) after applying the threshold of $\alpha = 0.56$ of the POCS. (c) Recovered data in tx domain using the spectrum in (b) after the first iteration of the POCS. The formerly empty trace now contains some data (red trace). (d) The fk -spectrum of the data in (c) after lowering the threshold with the same α . (e) Recovered tx domain signals and (f) the corresponding fk -diagram after six iterations of the POCS algorithm. The branch of each event in the fk domain is marked by the corresponding line in black. The missing trace (red) is fully recovered with minimum energy leakage in the fk domain.

results of the reconstructed data \mathbf{d}_{rec} with an ideal data set \mathbf{d}_{org} without missing traces. For a quantitative comparison a quality parameter Q is introduced (Kazemi *et al.* 2016):

$$Q = 10 \log \left(\frac{\|\mathbf{d}_{\text{org}}\|_2^2}{\|\mathbf{d}_{\text{org}} - \mathbf{d}_{\text{rec}}\|_2^2} \right). \quad (3)$$

The choices of α and the number of iterations largely determine the robustness of the reconstructed signals and the effectiveness of the interpolation process. For instance, lowering the threshold too fast will cause more noise-related energy in the fk spectrum, thereby degrading the quality of the reconstruction (Dokht *et al.* 2017). To determine the optimal combination of two parameters for our synthetic example, we calculate the quality parameter Q using a grid search (Fig. 3), which shows that convergence is achieved after 6 iterations for a relatively small (~ 0.55) factor α . Our synthetic data sets are calculated using the IRIS Synthetics Engine (IRIS DMC 2015; van Driel *et al.* 2016) that provides a fast way of generating a reference data set \mathbf{d}_{org} for a given reference model. We used the same synthetic data sets to calculate \mathbf{d}_{org} for our real data examples.

Design of an fk filter

The simplest way to modify signals in the fk spectrum is by muting or filtering a wavenumber range. A key consideration is the number of stations n , which influences the number of sampled points in the fk domain. A higher n results in a greater wavenumber resolution. Therefore, depending on n , the geometry and array aperture, different shapes of the filter may be adopted (Gubbins 2004).

Several candidate filters may be used, most simply a boxcar. Since ripples and overshoots in the signal (i.e. Gibbs oscillations) are introduced by the sharp edges of the filter (Hewitt & Hewitt 1979), it is useful to design the filter with a transfer function $F(f)$ with a certain slope (Yilmaz 2001). For instance, we could apply a linear function

$$F(f) = m(f + f_c), \quad (4)$$

with the slope m , or a Butterworth power function

$$F(f) = \frac{1}{1 + \left(\frac{f}{f_c}\right)^{2a}} \quad (5)$$

where f_c is the cut-off frequency (or cut-off wavenumber k_c in case of wavenumber filtering) and the exponent a controls the sharpness

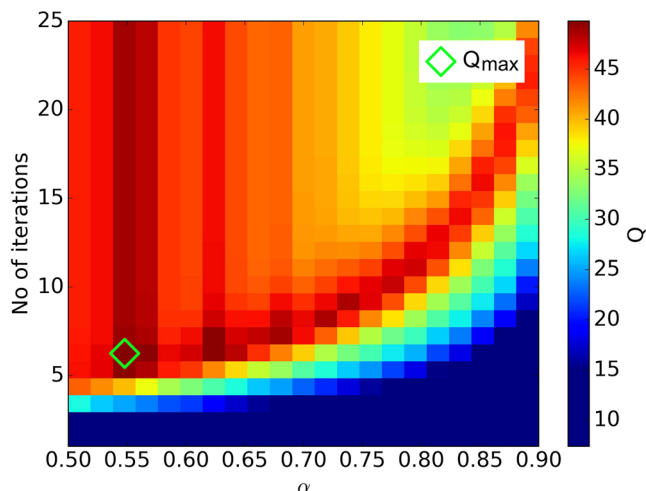


Figure 3. Colour-scale diagram of the quality parameter Q as a function of number of iterations and threshold value α for the synthetic data set shown in Fig. 2. The maximum value of Q is marked with a diamond and is taken as the best combination of iterations and α .

of the cut-off. We show possibilities of different filters tested here in Fig. S1. To create the fk filter, we calculate the filter functions for each frequency slice of the fk domain; therefore, generating a 2-D transfer function (as shown in Fig. S1d) centred around a wavenumber of 0. The corner wavenumbers are defined at half of the maximum amplitude of the transfer function.

Bootstrapping

To quantify the uncertainties we use the bootstrap algorithm (Efron & Tibshirani 1991). This is conducted by creating m new data sets, containing randomly selected traces from the original data set. The new data sets consist of the same number of traces as the original data set, but with random traces being left out and others being used repeatedly. The standard error σ is calculated as

$$\sigma = \left(\frac{\sum_{k=1}^m (\mathbf{d}_{\text{stacked}} - \mathbf{d}_k)^2}{m(m-1)} \right)^{1/2}.$$

Here $\mathbf{d}_{\text{stacked}}$ is the original data stack and \mathbf{d}_k denotes the k th randomly drawn stack.

PROCESSING STEPS

We process the data in different steps to fulfil the purpose of data filtering. First, we analyse the spatial data distribution and redistribute based on either a ‘sliding window’ or ‘static resampling’ approach. As mentioned above, the ‘static resampling’ scheme assumes the linearity of the target phase’s moveout in the entire data set, which limits the subsequent processing steps such as reconstruction or fk -filtering; the ‘sliding window’ scheme assumes the linearity of target phase in each bin only. Due to the stacking of other phases with potentially nonlinear arrival times, the signal of interest may, to varying degrees, suffer from decreased amplitudes during the reconstruction and filtering process, especially in the event of large spatial gaps in the resampled data set. In step 2, we reconstruct the missing data of the resampled data set using the POCS algorithm. Artefacts in the data such as high amplitude outliers, a shift in arrival times and differing amplitudes of the target phase can influence the result of the reconstruction, which will be discussed in detail below.

In a final step, we apply the fk filter on the codas or arrivals that interact with the target phase. Because of the nature of the fk domain, all signals with a linear moveout in the tx domain also show a linear moveout in the fk domain and share a common origin ($f = 0$ and $k = 0$). Therefore, the fk filter removes a part of the lowest frequencies of all coherent signals with a linear moveout. Because of the narrow k window, the filter is the most effective on the data in the centre of the array and least effective on data at the edges.

EXAMPLES

POCS reconstruction

Fig. 4 highlights the performance of the POCS reconstruction. In this example we adopt a synthetic data set containing two arriving phases (marked as P and PP) and test different input configurations. From the example in Figs 4(a) we randomly remove up to 30 per cent of the traces and add 20 per cent Gaussian noise to the remaining seismograms (Figs 4b, e and h). We then apply the POCS algorithm to reconstruct the missing traces, the results of which (Figs 4c, f and i) show the reconstruction of two coherent events. The algorithm also reproduces the noise onto the new traces, which is different from the approach of Dokht *et al.* (2017) where random noise is simultaneously reduced during reconstruction. The difference between reconstructed and original data sets (see Figs 4d, g and j) suggests satisfactory recovery of the main features, but data sets with large trace gaps (e.g. Fig 4h) exhibit considerable clutter after reconstruction (see Fig 4j). The maximum permissible noise levels for these configurations are 70 per cent Gaussian noise in case of two missing centre traces (Fig. S2a), 50 per cent for random missing traces (Fig. S2d) and 50 per cent in the presence of a large data gap (Fig. S2g), beyond which the quality of reconstruction is severely degraded. In all these cases the reconstructed traces show less noise and sharper coherent signals than the original data, although the best-fitted Q decreases with increasing noise levels (see Fig. 5).

We further investigate the limits of our POCS reconstruction by performing several tests using synthetic data. Distortions of the arrival time or amplitude of the target phase or high amplitude outliers in the input data set manifest themselves as erroneous signals in the reconstructed data (Fig. 6). After stacking, for example, in a vespagram, these artefacts are no longer observed (Fig. S3). The influence of an interfering signal on the same data sets as in Fig. 6 is shown in Fig. 7. The main difference is the reduced amplitudes in the reconstructed centre traces (see Fig. 7g). The influence of the high amplitude outlier is also apparent in the reconstruction (Figs 7k and l), especially in the neighbouring trace where a high amplitude signal is introduced after reconstruction. These tests show limitations of our reconstruction algorithm. After reconstructing, filtering and stacking, the results of the bootstrap shows that the time-delays introduce an uncertainty as an increase of the standard error and make it harder to identify the precursors (Figs 8a–c). Although the standard error decreases after fk -filtering in each case (Figs 8d–f), the reconstruction and filtering may introduce artefacts on neighbouring traces (Fig. 7).

fk -filtering

To assess the performance of the fk filter further (Fig. 9), we generate a synthetic data gather where a main phase (e.g. PP) and its two precursors with 10 per cent amplitudes are introduced (Fig. 9a). This example is inspired by mantle reflections from transition zone

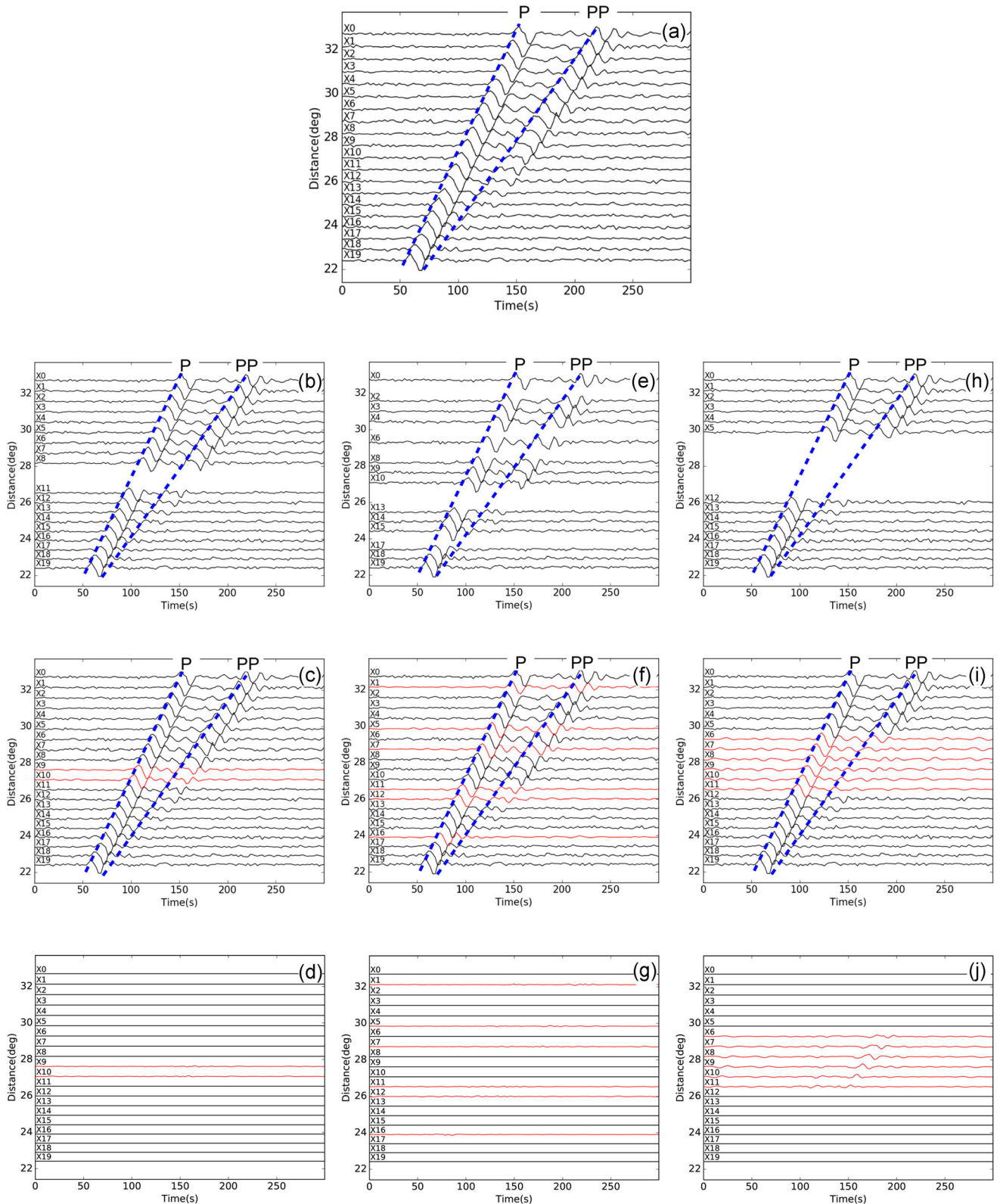


Figure 4. (a) A synthetic data set of a *P* and *PP* phase with 20 per cent added Gaussian noise. The waveforms are created with AxiSEM (Nissen Meyer *et al.* 2014) using realistic station-event geometries and an assumed earthquake source at 20 km depth. The phases of interest are marked by the blue dashed lines. (b) The data set with the two centre traces removed. (c) The reconstructed data using POCS with $\alpha = 0.59$ after five iterations. (d) The difference between the reconstructed and original data sets. (e–g) Similar to (b)–(d) but for a data set with 30 per cent of traces randomly removed. (h–j) Same as (b)–(d) but for a large data gap of around 4° .

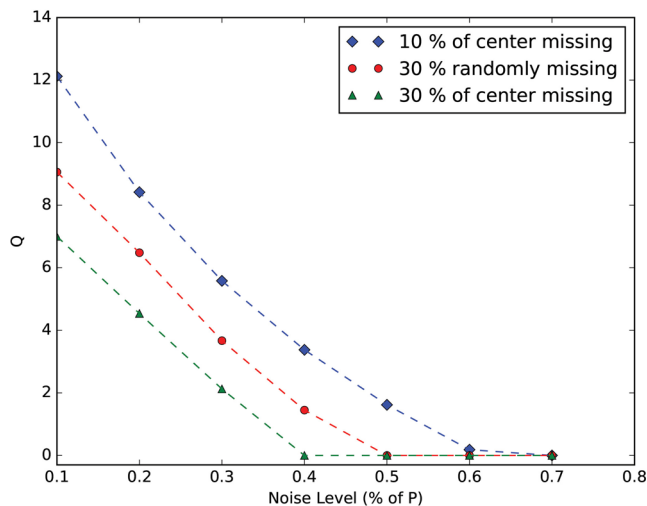


Figure 5. Q for reconstructed data shown in Fig. 4(a) with different noise levels (measured in percentage of amplitude of P). Blue diamonds show the trace distribution used in Fig. 4(b), red circles for trace distribution of Fig. 4(e) and green triangles for those of Fig. 4(h).

discontinuities at 410 and 660 km depths (e.g. Shearer & Flanagan 1999; Lessing *et al.* 2015). The corresponding vespagram shows clear energy clusters corresponding to the accurate slowness values of the three input time domain signals. Since the fk transform maps events with the same slowness to the same ω/k ratio, one can extract precursors from those data sets. To illustrate this we introduce an interfering phase (coda) with the same amplitude but different slowness as the main phase in the time window of 200–600 s (Fig. 9b). This long period phase effectively simulates a gradually attenuating signal observed in the P -wave, diffracted P (Pdiff) wave coda or PKiKP (e.g. Thomas & Billen 2009; Lessing *et al.* 2014, 2015; Saki *et al.* 2015). Our result shows strong interference from the coda wave on the first precursor, whereas the second precursor is nearly unaffected (Fig. 9b). The corresponding vespagram of the coda at a lower slowness provides further evidence for the absence of the first precursor and well preserved second precursor. A similar behaviour has been documented by Lessing *et al.* (2015), where a PP precursor strongly interferes with a core reflection.

After applying our Butterworth fk filter, the signal of the interfering synthetic coda is suppressed and both precursors are clearly visible (Fig. 9c). It should be noted, however, that the waveforms are altered due to the filtering process: side-lobes are introduced to all phases and the amplitude/waveform is also slightly modified, but the main shape as a zero phase wavelet in all cases is preserved, that is, no phase-shift is introduced. The application of the fk filter effectively removes the coda energy on all traces, though the outermost traces in our example still show residual coda energy. The slope of the fk filter and the spatial sampling of the data set are mainly responsible for this behaviour. For a larger data set with ~ 60 traces (Figs 6 and 7), the remaining energy due to the slope is confined exclusively to the outermost traces, which suggests that the spectrum is better resolved. In all cases, this residual coda energy in the outermost traces is suppressed and only the main phase and precursors are visible during the stacking process in the vespagram procedure.

Further we apply our filter to synthetic data sets, using different event depths, to investigate the application on filtering out depth phases such as pP or sS that may interfere with our target phase. We calculate synthetic data sets using source depths of 700 km

(Fig. 10) that show depth phases of the P wave that dominate the time window where the precursors arrive. After filtering we find clearer precursory arrivals, for both P410P and P660P. An additional test using a 500 km depth source shows similar improvement on the precursors after filtering process (Fig. S4).

The fk filter is equally effective on real data from an earthquake in the Philippines (1995 May 5) recorded at the Tanzania BB Experiment Array (Fig. 12). This event was used in a search for PP precursors from mantle transition zone discontinuities beneath the Indian Ocean (Reiß *et al.* 2017). The unequally spaced traces are resampled using the ‘Sliding Window’ method prior to the reconstruction. The resampled and reconstructed data (Fig. 11b) clearly show the arrivals of P and PP at a regular spatial interval. The weak arrivals in the time window that contains the precursors (e.g. between 150–350 s) are also well recovered.

Figs 12(a) and (b) shows the fk spectrum and vespagram before fk -filtering of the data used in Fig. 11. The PP phase is clearly visible in the vespagram but P shows a coda of more than 100 s. The fk spectrum displays the P and P coda energy at wavenumber 0 after aligning on its slowness. There is no evidence for a precursor near the theoretical arrival times but small, diffuse arrivals are detected in the time window surrounding PP . This recording would generally have been rejected due to difficulties in picking arrivals times (e.g. Thomas & Billen 2009; Lessing *et al.* 2014; Saki *et al.* 2015). After applying a Butterworth band stop fk filter, the region around wavenumber 0, which includes the energies of P phase and its coda, is muted for all frequencies in the fk spectrum (Fig. 12c). The corresponding vespagram (Fig. 12d) shows a clear P410P arrival at the expected slowness and travel time, even a weak P660P arrival can be identified in the data. The amplitude difference between P410P and P660P, which has been reported previously (e.g. Estabrook & Kind 1996; Deuss 2009), can also be seen here but is unrelated to filtering effects.

DISCUSSION

A plane wave approximation is commonly used in global seismology, leading to the assumption of travel time linearity in the tx domain. This assumption is used for most interpolation and reconstruction techniques, especially in active source, near surface seismic studies (Naghizadeh & Sacchi 2010; Oropeza & Sacchi 2011; Naghizadeh 2012). In this study we adopt a POCS algorithm (Stark 1987) for reconstruction, which are highly effective for both exploration and solid Earth seismic data. A known drawback of POCS is an undesirable tendency to interpolate in time direction (Abma & Kabir 2006), that is, during the removal of spatial gaps some well-sampled traces in the tx domain may be influenced. Due to the linearity of events within in relatively narrow time windows and spatial gathers, this problem could be alleviated through partial stacking based on a sliding time window prior to reconstruction. The POCS heavily relies on the distribution of missing data. For an optimal solution, gaps should be as random as possible in all dimensions, though in reality missing data often follow a linear geometry (e.g. due to complex terrain). Large gaps in data sets show clear limits to the POCS reconstruction, where the POCS tends to produce averaging effects similar to smoothing in 3-D seismic velocity tomography (Koroni & Trampert 2015). Despite these limitations, the spatial interpolation from the fk filter is generally effective, especially in view that most of the data have been pre-conditioned through resampling (i.e. the moveouts of the target phase such as PP are retained while other interfering phases are preferentially suppressed).

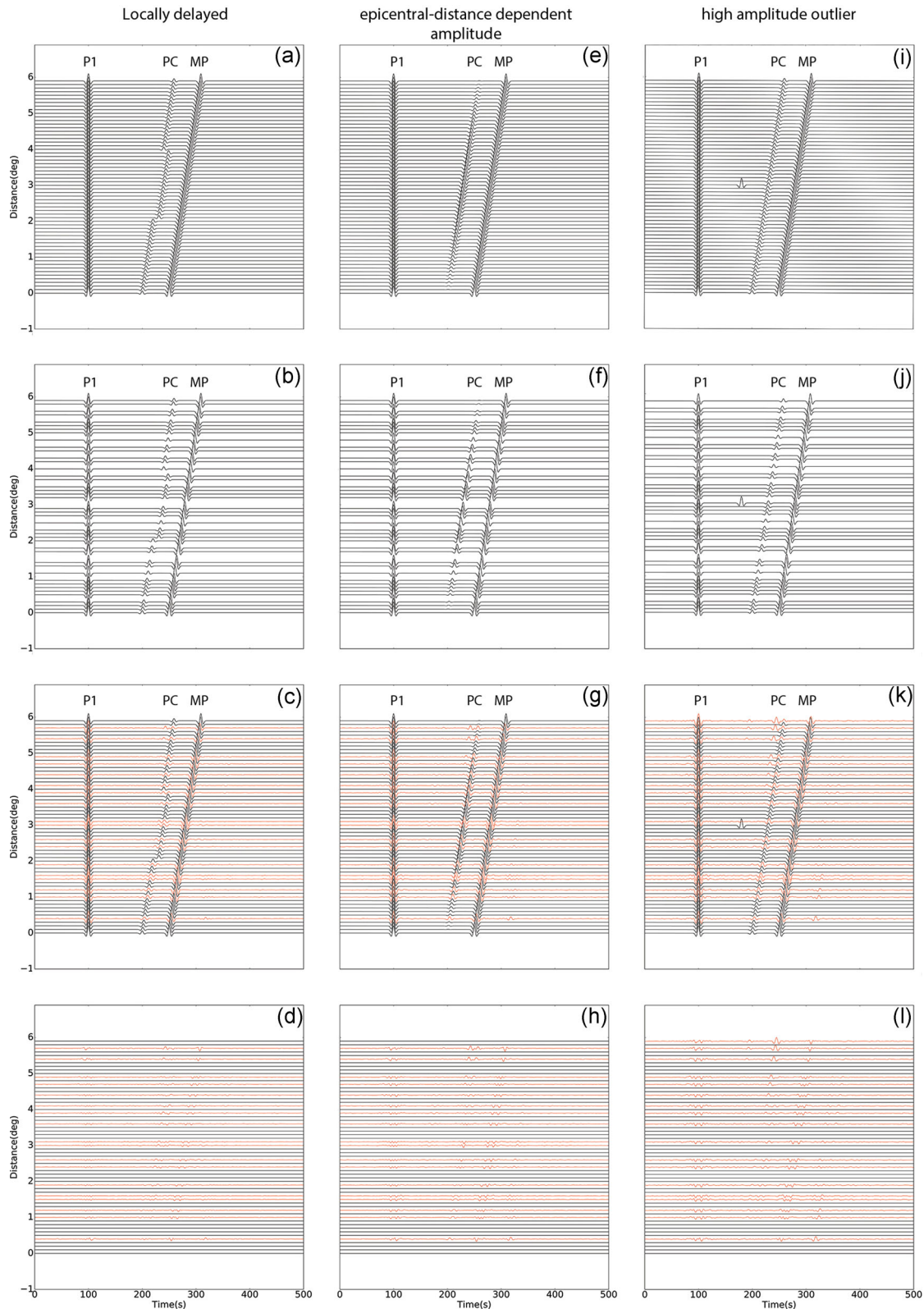


Figure 6. Results of the POCS reconstruction, when precursor (PC) arrivals are locally delayed by 4 s (a–d). The reconstructed section shows distorted or doubled waveforms at each trace with a delayed signal. (e–h) Results for a linear distance dependent PC amplitude, with its maximum at 3° . The reconstruction shows satisfying results in the centre of the array, but artefacts are clearly visible at both edges of the array. (i–l) Results for a high amplitude outlier in the data set. Again, artefacts and falsely reconstructed signals can be seen at the edge traces of the array. In all test cases, 30 per cent of the traces were randomly removed before reconstruction.

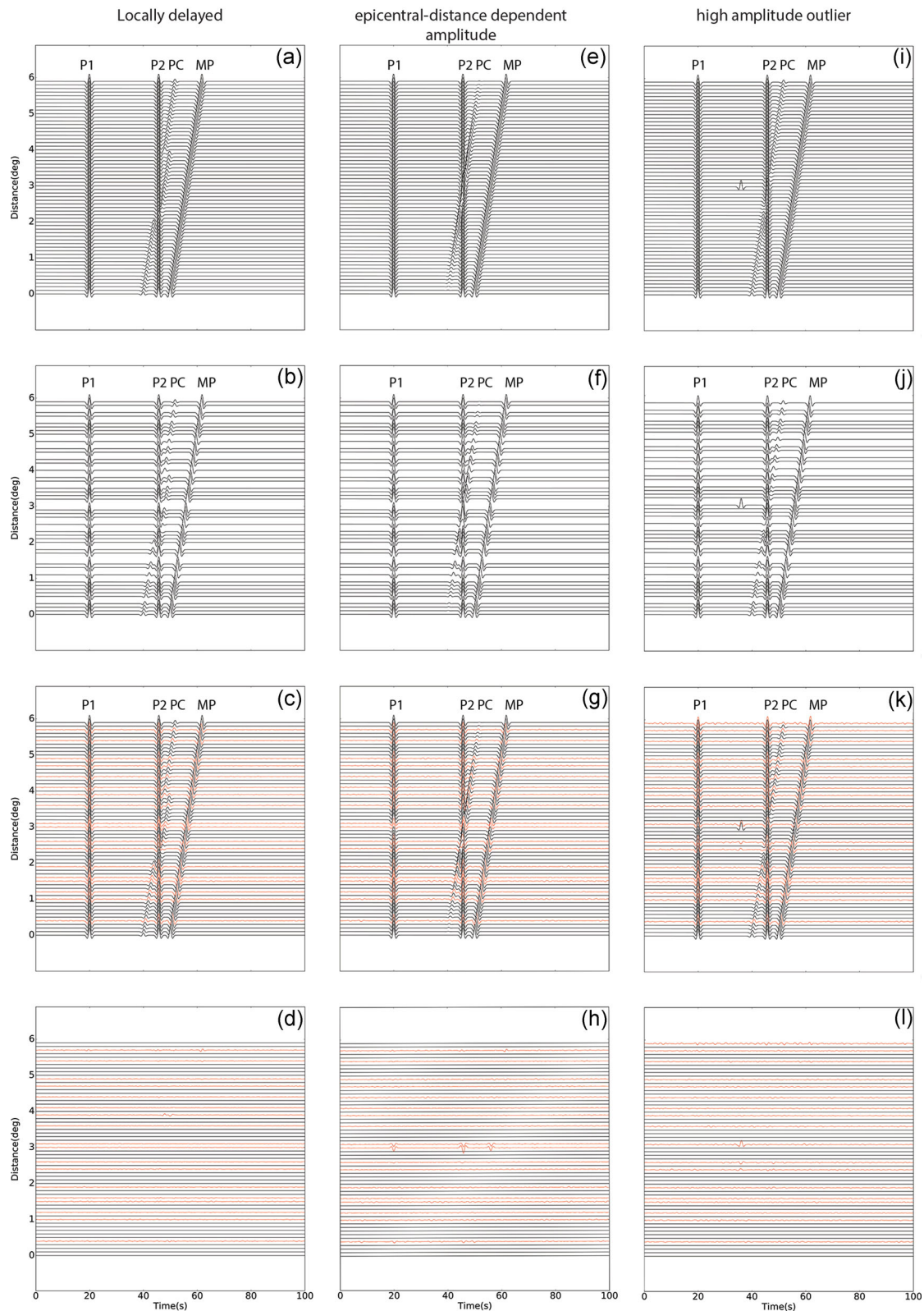


Figure 7. Same as in Fig. 6, but with an additional interfering phase P2.

The series of examples described above show that it is possible to interpolate complex data sets with a high precision. It is possible to redistribute data on a uniform grid for further processing and, because of its simplicity and efficiency, POCS seems to be a powerful tool in data pre-processing. It only depends on two pa-

rameters [number of iterations i and threshold level α (eq. 1)], and the best combination of those can be examined with a Q parameter test. Our algorithm is fast, easy to implement/execute, and it is not restricted to the 2-D used in this study. By adding a second spatial dimension provided by a seismic array, it is possible to apply it to

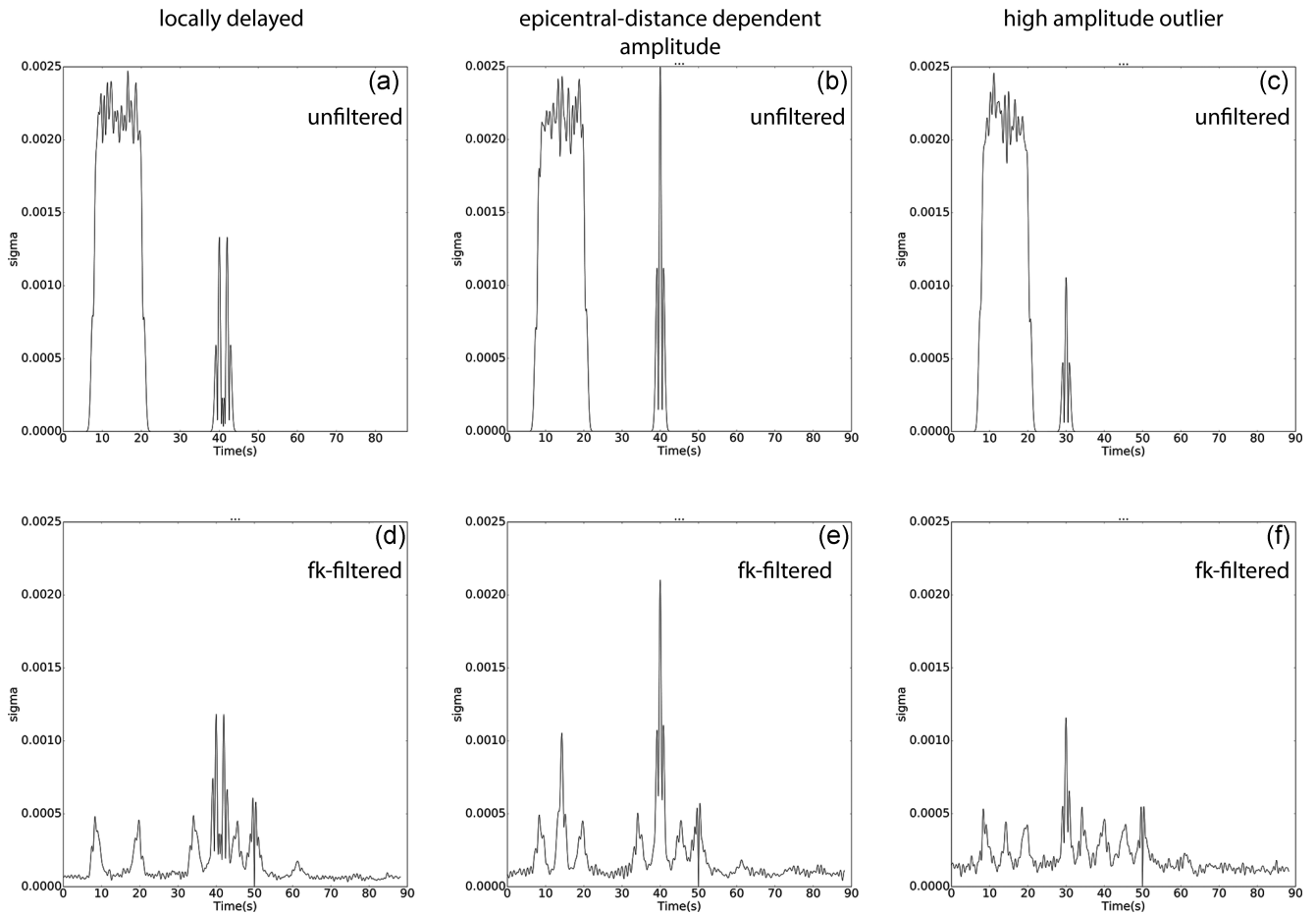


Figure 8. Bootstrap results of data shown in Fig. 6. A total of 200 trials are conducted. For each trial, traces are randomly selected and are aligned on the precursor to correct for the moveout before stacking. Bootstrap results of data with (a) a locally delayed arrival of the precursor (c), (b) distance dependent, amplitude varying precursors (Fig. 6g) and (c) a high amplitude outlier (Fig. 6k). (d–f) Same as (a)–(c) but for the fk filtered data. The stacking uncertainties (sigma) around the arrival time of the precursor are significantly decreased, but the distribution of uncertainty becomes broader.

case studies with two spatial and one temporal dimensions (Abma & Kabir 2006; Wang *et al.* 2010). Computational costs rise with the size of the data set, but not to the extent that it becomes impractical, due to its simplicity. Just as importantly, POCS has no restriction on transformation and regularly sampled data are not required as it is applicable to both linear and nonlinear transformations, for example, Radon transform (Kabir & Verschuur 1995; Schultz & Gu 2013), slant-stacklet transform (Ventosa & Romanowicz 2015), wavelet transform (Daubechies 1990) or anti-leakage Fourier transform (Xu *et al.* 2005) which could also be incorporated in this algorithm. This could potentially alleviate the smearing problem due to the assumption of linear moveout during the stacking as discussed above.

It is worth noting that the limitations of the fk filter are largely dictated by the distortions in the data sets. For example, the quality of the reconstruction varies depending on travel time delays or amplitude changes. The target signal in the seismic section, for example, strong arrivals, is usually reconstructed effectively while artefacts such as multiples are not. Locally delayed signals tend to increase the uncertainty of the reconstructed and measured signals.

A key problem in seismic data analysis is the interference of weak seismic phases with unwanted neighbouring phases or their codas. At different slownesses, these phases should be separate and distinguishable on a vespagram (e.g. Muirhead & Datt 1976; Rost & Thomas 2002). However, the masking of weak phases by interfering

signal is possible (e.g. Lessing *et al.* 2015), as are misinterpretations. We tested the filter on other events with long interfering coda of PKiKP and Pdiff and the process allows a much larger number of events to be used in the analysis. Even PP precursor phases that were apparently absent before filtering can be recovered after application of the fk filter (Reiß *et al.* 2017).

We showed the application of the fk filter to examples of PP and its precursors that arise from underside reflections off mantle discontinuities. This is only one possible application and many other cases are possible. However, it is critical that the signal of interest shows a sufficiently large slowness difference to that of interfering phases. In the cases of Pdiff and PKiKP, which often exhibit long coda windows (e.g. Lessing *et al.* 2014, 2015; Saki *et al.* 2015), the slowness differences are often greater than $2\text{--}3\text{ s deg}^{-1}$ relative to PP wave and its precursors. The fk filter removes a frequency slice of the fk -spectrum, therefore removes the very low frequencies on all phases originating at zero wavenumber and frequency. But this hardly influences the results as we are working with bandpass filtered data which contains minimum low frequency energies. This method would function equally well in the case of SS precursor detection. But in some other cases, where the slowness values of the main phase and interfering phases are too similar, the fk filter either suppress both phases or results in insufficient suppression. In principle, fk filter works for any phase given sufficient spatial resolution from the array data. However, the wavenumber resolution

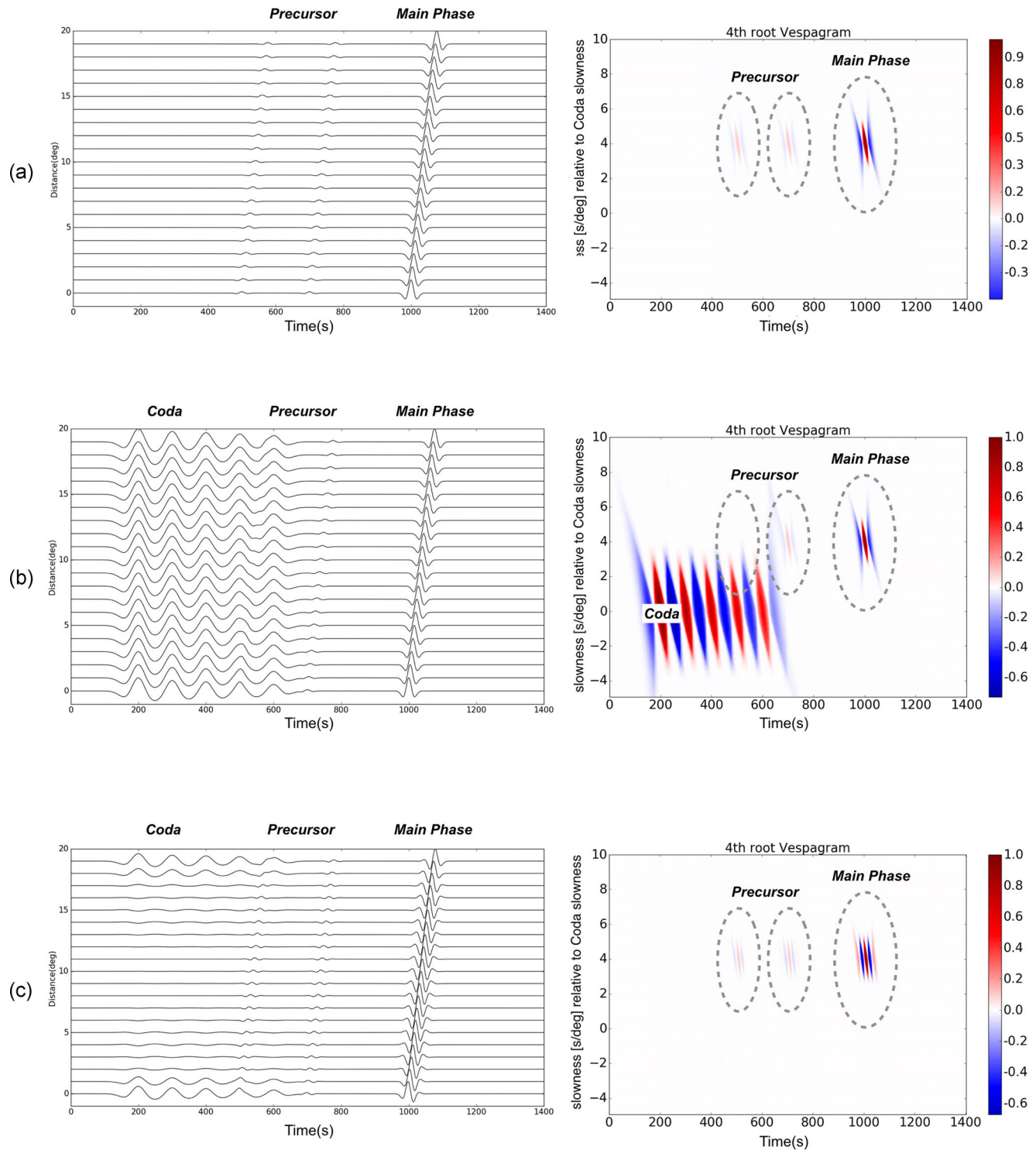


Figure 9. (a) Synthetic data set with a main phase and two precursors with 10 per cent amplitude of the main phase (left) and the corresponding fourth root vespagram (right). (b) Similar to (a), but with an additional long-period signal (coda) masking the early precursor. (c) Seismogram (left) and vespagram (right) of the data using the synthetic seismic section in (b).

must be sufficiently high to distinguish interfering phases from the target phase with similar slowness values.

CONCLUSIONS

We developed and applied an fk filter to enhance data quality and detect smaller phases, such as PP precursors, in vespagrams. For

an application of 2-D FFT the data have to be sampled regularly in space and time. While the time sampling for seismic recordings is usually given, we introduce two methods to re-distribute the recordings of a seismic array onto a regular grid, where the missing traces are filled using a POCs algorithm. We demonstrate the robustness of resampling scheme with synthetic examples and show that the performance of reconstruction is significantly better for data set with

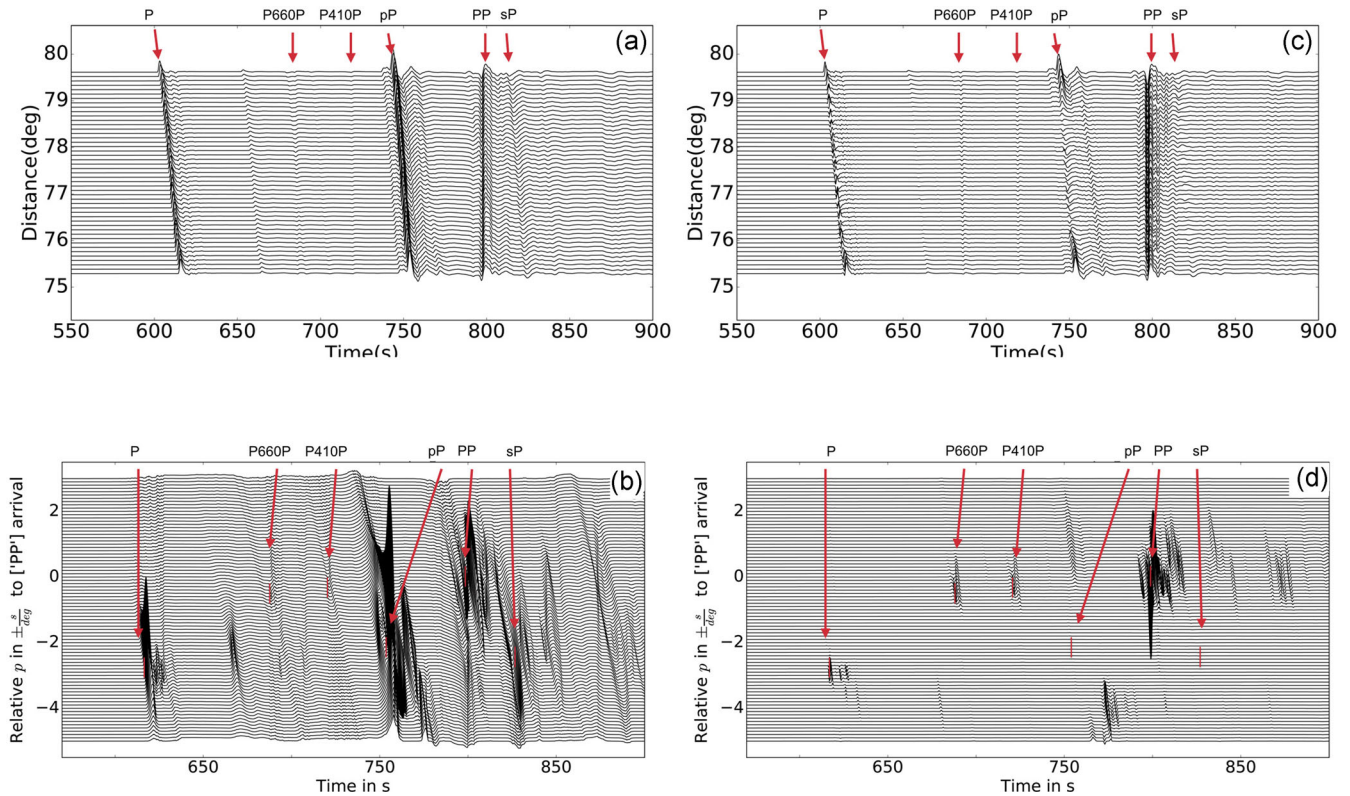


Figure 10. Synthetic data set computed using a source at 700 km depth. (a) Seismograms before filtering. The theoretical arrival times are marked by arrows. (b) The fourth root vespagram of (a), where strong P , pP and sP phases interfere with the waveform of the precursors P410P and P660P. (c) The f_k filtered seismograms. (d) Vespagram of (c). Depth phases of pP and sP are removed from the data, while the onset of both precursors is clearly visible.

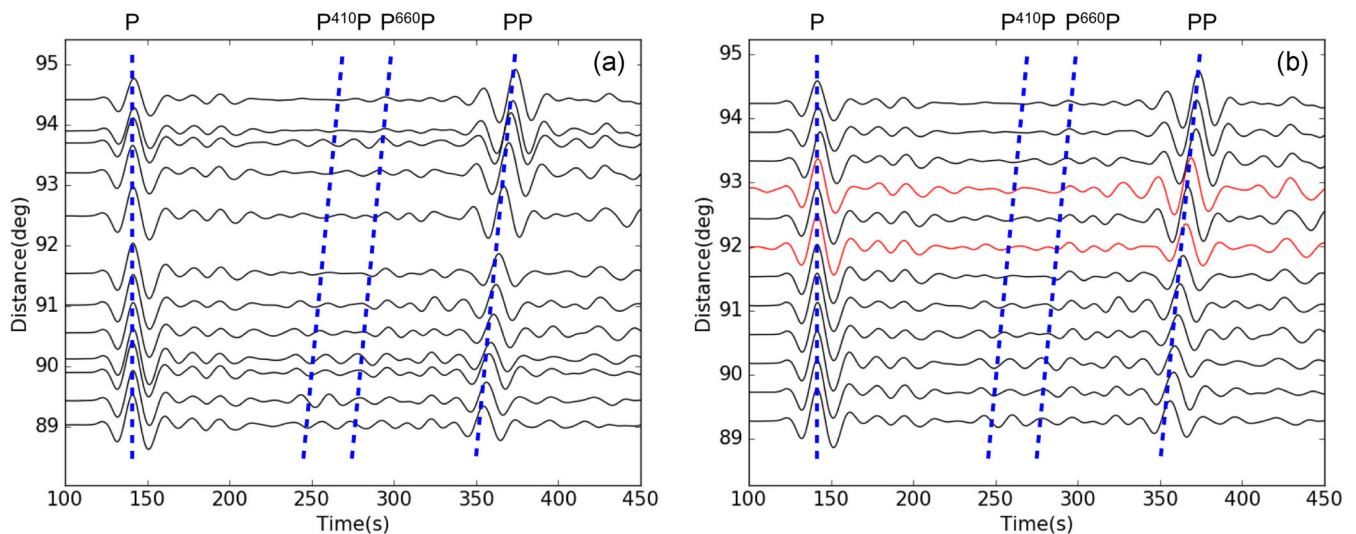


Figure 11. (a) Teleseismic recording of an earthquake that occurred on 1995 May 5 at 3:53:43.59 UTC, latitude 12.268 N, longitude 125.281E at a depth of 18.3 km in the Philippine Sea recorded at an array in Tanzania. The data are bandpass filtered at corner periods of 15 and 75 s. Theoretical arrivals of P and PP are marked by blue dashed lines. (b) Reconstructed data (red traces) into an equally distributed grid using POCS with $\alpha = 0.92$ and 30 iterations.

randomly or sparsely missing traces than those with large spatial gaps.

The f_k filter cuts out an area of the f_k spectrum and allows to remove energy with certain wavenumbers (i.e. slowness). We apply this to synthetic data that resemble a case of PP wave and its precursors with a second main phase with a long coda that interferes with the signals we want to recover. While one of the precursors is apparently missing in the unfiltered vespagram, the application

of the f_k filter allows the identification of the two precursors. Applying the filter to a real data example shows that we are able to improve the identification of PP precursors, thereby allowing an increase in the number of useful events in the data analysis. The f_k filter is not restricted to PP reflections at the underside of mantle discontinuities but applicable also to other data, given that the slowness values of the interfering phases and the phases one wants to recover are sufficiently different.

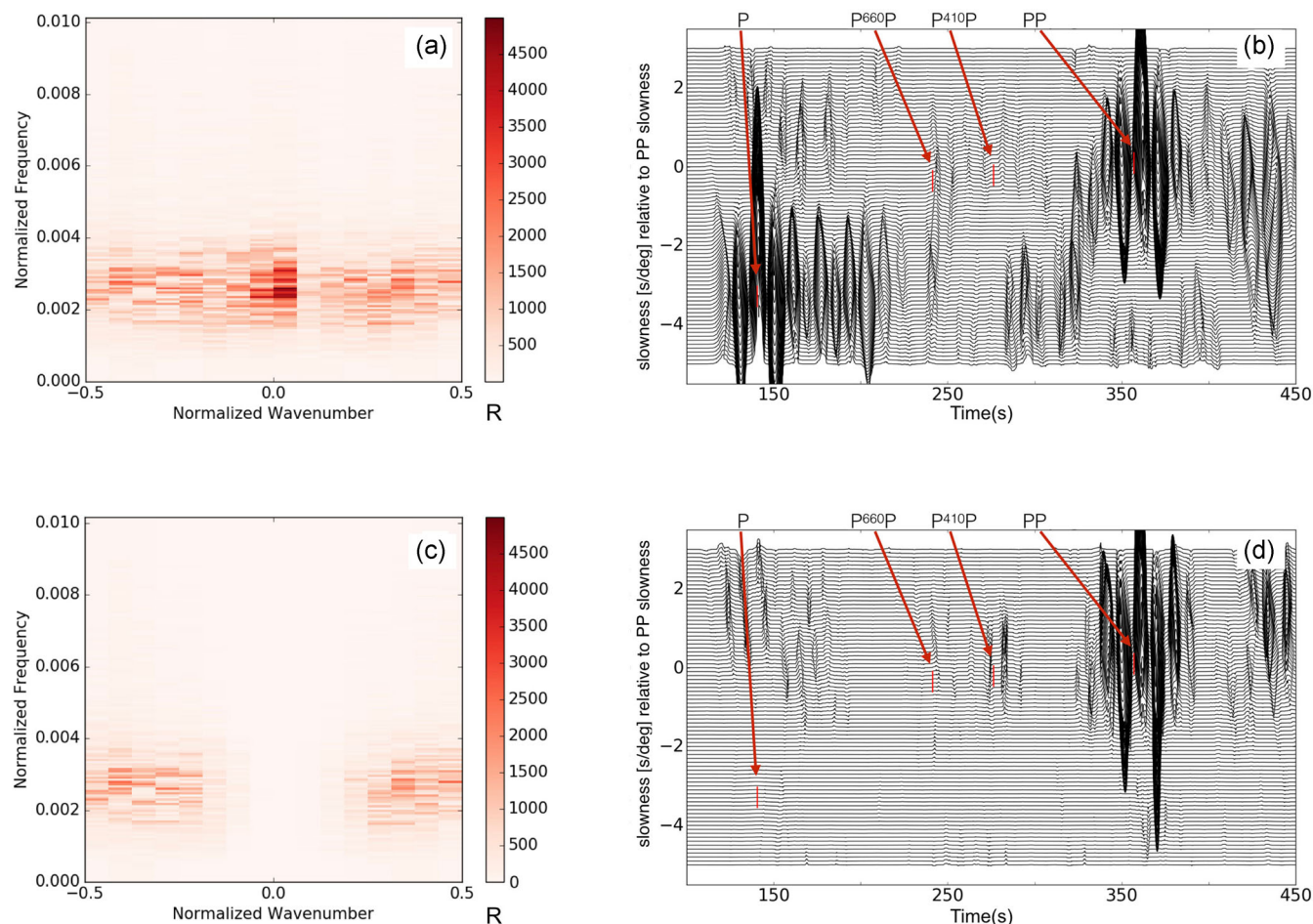


Figure 12. (a) The $f-k$ spectrum of data presented in Fig. 11. The data have been aligned on the P wave before filtering. (b) fourth root vespagram of the event presented in Fig. 11. Theoretical arrivals of P , $P660P$, $P410P$ and PP are marked in red. The energy of the P -coda is visible between 150 and 220 s. (c) The $f-k$ spectrum of the bandstop filtered data with normalized corner wavenumbers of 0.02 and -0.02 and an exponent $a = 2$ (eq. 5) of the data presented in Fig. 5(b). The data were aligned on the P phase before filtering. (d) The vespagram of the $f-k$ filtered data. Phases are marked as in (b).

ACKNOWLEDGEMENTS

We would like to thank Sergi Ventosa, an anonymous reviewer and the Editor Martin Schimmel for constructive and helpful reviews. The facilities of IRIS Data Services, and specifically the IRIS Data Management Center, were used for access to waveforms, related metadata and/or derived products used in this study. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience and EarthScope (SAGE) Proposal of the National Science Foundation under Cooperative Agreement EAR-1261681. Codes, maps and figures were produced with Python (Python Software Foundation (2016)), data were processed using the python library ObsPy (Beyreuther *et al.* 2010). The work was partly financed by DFG (Deutsche Forschungsgemeinschaft) grant TH1530/9-1.

REFERENCES

- Abma, R. & Kabir, N. 2006. 3D interpolation of irregular data with a POCS algorithm, *Geophysics*, **71**(6), E91–E97.
 Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y. & Wassermann, J. 2010. ObsPy: a Python toolbox for seismology, *Seismol. Res. Lett.*, **81**(3), 530–533.

- Candes, E. & Romberg, J. 2007. Sparsity and incoherence in compressive sampling, *Inverse Probl.*, **23**(3), 969–985.
 Chen, Y., Zhang, D., Jin, Z., Chen, X., Zu, S., Huang, W. & Gan, S., 2016. Simultaneous denoising and reconstruction of 5-D seismic data via damped rank-reduction method, *Geophys. J. Int.*, **206**(3), 1695–1717.
 Cobden, L. & Thomas, C. 2013. The origin of D'' reflections: A systematic study of seismic array data sets, *Geophys. J. Int.*, **194**(2), 1091–1118.
 Daubechies, I. 1990. The wavelet transform, time-frequency localization and signal analysis, *IEEE Trans. Inf. Theory*, **36**(5), 961–1005.
 Deuss, A. 2009. Global observations of mantle discontinuities using SS and PP precursors, *Surv. Geophys.*, **30**(4–5), 301–326.
 Dokht, R.M., Gu, Y.J. & Sacchi, M.D. 2017. Singular Spectrum Analysis and its applications in mapping mantle seismic structure, *Geophys. J. Int.*, **208**, 1430–1442.
 Dziewonski, A.M. & Anderson, D.L. 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.*, **25**(4), 297–356.
 Efron, B. & Tibshirani, R. 1991. Statistical data analysis in the computer age, *Science*, **253**, 390–395.
 Estabrook, C.H. & Kind, R. 1996. The nature of the 660-kilometer upper-mantle seismic discontinuity from precursors to the PP phase, *Science*, **274**(5290), 1179–1182.
 Gu, Y.J. & Sacchi, M. 2009. Radon transform methods and their applications in mapping mantle reflectivity structure, *Surv. Geophys.*, **30**(4–5), 327–354.
 Gubbins, D. 2004. *Time Series Analysis and Inverse Theory for Geophysicists*, Cambridge Univ. Press.

- Hewitt, E. & Hewitt, R.E. 1979. The Gibbs-Wilbraham phenomenon: an episode in Fourier analysis, *Arch. Hist. Exact Sci.*, **21**(2), 129–160.
- IRIS DMC 2015. Data Services Products: Synthetics Engine, doi:10.17611/DP/SYNGINE.1.
- Jacobeit, E., Thomas, C. & Vernon, F. 2013. Influence of station topography and Moho depth on the mislocation vectors for the Kyrgyz Broadband Seismic Network (KNET), *Geophys. J. Int.*, **193**(2), 949–959.
- Kabir, M.N. & Verschuur, D. 1995. Restoration of missing offsets by parabolic Radon transform, *Geophys. Prospect.*, **43**(3), 347–368.
- Kazemi, N., Bongajum, E. & Sacchi, M.D. 2016. Surface-Consistent Sparse Multi-channel Blind Deconvolution of Seismic Signals, *IEEE Trans. Geosci. Remote Sens.*, **54**(6), 3200–3207.
- Koroni, M. & Trampert, J. 2015. The effect of topography of upper-mantle discontinuities on SS precursors, *Geophys. J. Int.*, **204**(1), 667–681.
- Kreimer, N., Stanton, A. & Sacchi, M.D. 2013. Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction, *Geophysics*, **78**(6), V273–V284.
- Krüger, F. & Weber, M. 1992. The effect of low-velocity sediments on the mislocation vectors of the GRF array, *Geophys. J. Int.*, **108**(1), 387–393.
- Lessing, S., Thomas, C., Saki, M., Schmerr, N.C. & Vanacore, E. 2014. Evidence of absence of PP precursors, in *EGU General Assembly Conference Abstracts*, Vol. 16, p. 1547.
- Lessing, S., Thomas, C., Saki, M., Schmerr, N. & Vanacore, E. 2015. On the difficulties of detecting PP precursors, *Geophys. J. Int.*, **201**(3), 1666–1681.
- Muirhead, K.J. & Datt, R. 1976. The N -th root process applied to seismic array data, *Geophys. J. Int.*, **47**(1), 197–210.
- Naghizadeh, M. 2012. Seismic data interpolation and denoising in the frequency-wavenumber domain, *Geophysics*, **77**(2), V71–V80.
- Naghizadeh, M. & Sacchi, M.D. 2010. On sampling functions and Fourier reconstruction methods, *Geophysics*, **75**(6), WB137–WB151.
- Nissen-Meyer, T., van Driel, M., Stähler, S.C., Hosseini, K., Hempel, S., Auer, L. & Fournier, A. 2014. AxiSEM: broadband 3-D seismic wavefields in axisymmetric media, *Solid Earth*, **5**(1), 425–445.
- Oropeza, V. & Sacchi, M. 2011. Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis, *Geophysics*, **76**(3), V25–V32.
- Reiß, A., van Driel, J., Thomas, C. & Heyn, B. 2017. A hot mid-mantle anomaly in the area of the Indian Ocean Geoid Low, *Geophys. Res. Lett.*, **44**, 6702–6711.
- Rost, S. & Thomas, C. 2002. Array seismology: methods and applications, *Rev. Geophys.*, **40**(3), 2-1–2-27.
- Rost, S. & Weber, M., 2001. A reflector at 200 km depth beneath the northwest Pacific, *Geophys. J. Int.*, **147**(1), 12–28.
- Saki, M., Thomas, C., Nippres, S.E. & Lessing, S. 2015. Topography of upper mantle seismic discontinuities beneath the North Atlantic: the Azores, Canary and Cape Verde plumes, *Earth planet. Sci. Lett.*, **409**, 193–202.
- Schultz, R. & Gu, Y.J. 2013. Flexible, inversion-based Matlab implementation of the Radon transform, *Comput. Geosci.*, **52**, 437–442.
- Shearer, P.M. & Flanagan, M.P. 1999. Seismic velocity and density jumps across the 410- and 660-kilometer discontinuities, *Science*, **285**(5433), 1545–1548.
- Stark, H. (Ed.). 1987. *Image Recovery: Theory and Application*, Elsevier.
- Thomas, C. & Billen, M.I. 2009. Mantle transition zone structure along a profile in the SW Pacific: thermal and compositional variations, *Geophys. J. Int.*, **176**(1), 113–125.
- Van Driel, M., Hutko, A., Krischer, L., Trabandt, C., Stähler, S. & Nissen-Meyer, T. 2016. Syngine: on-demand synthetic seismograms from the IRIS DMC based on AxiSEM & instaseis, in *EGU General Assembly Conference Abstracts*, Vol. 18, p. 8190.
- Ventosa, S. & Romanowicz, B. 2015. Extraction of weak PoP phases using the slant-stacklet transform—I: Method and examples, *Geophys. J. Int.*, **201**(1), 207–223.
- Ventosa, S., Simon, C. & Schimmel, M. 2012. Window length selection for optimum slowness resolution of the local-slant-stack transform, *Geophysics*, **77**(2), V31–V40.
- Wang, S.-Q., Gao, X. & Yao, Z.-X. 2010. Accelerating POCS interpolation of 3D irregular seismic data with graphics processing units, *Comput. Geosci.*, **36**(10), 1292–1300.
- Xu, S., Zhang, Y., Pham, D. & Lambaré, G. 2005. Antileakage Fourier transform for seismic data regularization, *Geophysics*, **70**(4), V87–V95.
- Yilmaz, Ö. 2001. *Seismic Data Analysis*, Vol. 1, Society of Exploration Geophysicists.
- Zheng, Z., Ventosa, S. & Romanowicz, B. 2015. High resolution upper mantle discontinuity images across the Pacific Ocean from SS precursors using local slant stack filters, *Geophys. J. Int.*, **202**(1), 175–189.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. Amplitude over wavenumber diagrams showing different shapes of filter functions with normalized cut-off wavenumber at $k = -0.2$ and $k = 0.2$. (a) Simple boxcar bandpass filter, (b) tapered bandpass filter, with m indicating the slope at the edges (shown are examples for $m = 1.3, 2, 4$). (c) Example of a Butterworth shaped bandpass filter, with different values for the exponent ($a = 2, 4$ and 8) controlling the steep of the slope. (d) Frequency over wavenumber display of the transfer function of a tapered fk bandpass filter function with normalized corner wavenumbers of 0.04 and -0.04 and an exponent $a = 4$. The amplitude spectrum R is shown in grey scale. A Butterworth wavenumber filter of identical shape is calculated for each frequency slice.

Figure S2. A synthetic data set of a P and PP phases with added Gaussian noise. The waveforms are created with AxiSEM (Nissen Meyer *et al.* 2014) using realistic station-event geometries and an assumed earthquake source at 20 km depth. The phases of interest are marked by the blue dashed lines. (a) The data set with the two centre traces removed, with 70 per cent added Gaussian noise. (b) The reconstructed data (red traces). (c) The difference between the reconstructed and original data sets. (d–f) Similar to (a)–(c) but with data set with 30 per cent of traces randomly removed and 50 per cent added Gaussian noise. (g,h) Same as (a)–(c) but for a large data gap of around 4° and 50 per cent added Gaussian noise.

Figure S3. Fourth root vespagrams of data sets used in Figs 7 and 8, phases are marked with arrows. (a), (d) and (g) show the vespagrams of the seismograms shown in Figs 8(a), (e) and (i), respectively. (b), (e) and (h) show the vespagrams of seismograms after reconstruction and fk filtering shown in Figs 8(c), (g) and (k), respectively. (c), (f) and (i) show the vespagrams of the undisturbed data, which are shown in Figs 7(a), (e) and (i), respectively.

Figure S4. Synthetic data set computed using a source at 500 km depth. (a) Seismograms before filtering. The theoretical arrival times are marked by arrows. (b) The fourth root vespagram of (a), where strong P , pP and sP phases interfere with the waveform of the precursors P410P and P660P. (c) The fk filtered seismograms. (d) Vespagram of (c). Depth phases of pP and sP are removed from the data, while the onset of both precursors is clearly visible.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.