Geophysical Journal International

Geophys. J. Int. (2015) **201,** 1182–1194
GJI Marine geosciences and applied geophysics

doi: 10.1093/gji/ggv072

Simultaneous seismic data interpolation and denoising with a new adaptive method based on dreamlet transform

Benfeng Wang,[1,2] Ru-Shan Wu,[2] Xiaohong Chen[1] and Jingye Li[1]

[1]*State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum, Beijing* 102249, *China. E-mail: chenxh@cup.edu.cn*
[2]*Modeling and Imaging Laboratory, Earth and Planetary Sciences, University of California, Santa Cruz, CA* 95064, *USA*

Accepted 2015 February 11. Received 2015 January 21; in original form 2014 September 23

SUMMARY

Interpolation and random noise removal is a pre-requisite for multichannel techniques because the irregularity and random noise in observed data can affect their performances. Projection Onto Convex Sets (POCS) method can better handle seismic data interpolation if the data's signal-to-noise ratio (SNR) is high, while it has difficulty in noisy situations because it inserts the noisy observed seismic data in each iteration. Weighted POCS method can weaken the noise effects, while the performance is affected by the choice of weight factors and is still unsatisfactory. Thus, a new weighted POCS method is derived through the Iterative Hard Threshold (IHT) view, and in order to eliminate random noise, a new adaptive method is proposed to achieve simultaneous seismic data interpolation and denoising based on dreamlet transform. Performances of the POCS method, the weighted POCS method and the proposed method are compared in simultaneous seismic data interpolation and denoising which demonstrate the validity of the proposed method. The recovered SNRs confirm that the proposed adaptive method is the most effective among the three methods. Numerical examples on synthetic and real data demonstrate the validity of the proposed adaptive method.

Key words: Fourier analysis; Wavelet transform; Inverse theory; Spatial analysis.

1 INTRODUCTION

Observed seismic data is always irregularly sampled in spatial coordinates due to factors such as presence of obstacles, forbidden areas and feathering. Elimination of dead shots and dead traces in the processing stage, is also a major reason for irregularity of observed seismic data. Besides, the observed seismic data always contains random noise because of effects from acquisition equipments or the acquisition environment. The incomplete noisy seismic data can affect the performances of multichannel techniques, such as surface-related multiple elimination (SRME), wave equation based migration and inversion. Therefore, seismic data interpolation and random noise attenuation is an essential step.

Seismic interpolation methods can be divided into four categories (Gao *et al.* 2012; Wang *et al.* 2014b): signal analysis and mathematical transform based methods (Naghizadeh & Sacchi 2010b; Xu *et al.* 2010; Naghizadeh & Innanen 2011; Gao *et al.* 2012; Wu *et al.* 2013; Zhang *et al.* 2013; Xue *et al.* 2014); prediction filters based methods (Spitz 1991; Naghizadeh & Sacchi 2007); wave equation based methods (Ronen 1987) and rank reduction based methods (Gao *et al.* 2013; Kreimer *et al.* 2013; Ma 2013). Missing traces and random noise can increase the rank of matrix which is composed of seismic data at a given frequency, and the interpolated data can be obtained through rank reduction. The least rank, affected by the number of linear events in the window of analysis, should be determined before interpolation and would affect the final performance (Oropeza & Sacchi 2011; Gao *et al.* 2013; Ma 2013). Methods based on wave equation always require distribution of underground parameters and are computationally expensive (Ronen 1987). Methods based on prediction filters, use the predictability of linear events in the frequency–space domain to interpolate aliased high frequency data with filters derived from low frequencies (Spitz 1991; Naghizadeh & Sacchi 2007; Gao *et al.* 2012). Methods based on mathematical transform are easy to handle and have drawn much attention. Fourier based methods (Gao *et al.* 2010; Naghizadeh & Sacchi 2010b; Xu *et al.* 2010; Naghizadeh & Innanen 2011; Gao *et al.* 2012; Zhang *et al.* 2013) can handle linear or quasi-linear events suitably, but they should be handled window by window for curved events. Radon transform based methods (Trad *et al.* 2002; Yu *et al.* 2007; Wang *et al.* 2010; Xue *et al.* 2014) can be divided into three categories: linear radon, parabolic radon and hyperbolic radon based methods, while only linear and parabolic radon transforms are widely used because they are time-invariant and can be implemented in frequency domain efficiently. Curvelet transform based methods (Hennenfent & Herrmann 2008; Herrmann & Hennenfent 2008; Naghizadeh & Sacchi 2010a; Yang *et al.* 2012; Shahidi *et al.*

2013) can handle curved seismic events suitably. In order to use the prior information of curvelet coefficients, Mansour *et al.* (2013) adopted the weighted one-norm minimization in which corrections between locations of significant curvelet coefficients are applied to interpolate the unknown seismic data. While curvelet transform is a mathematical transform which has big redundancy and is time consuming especially for large scale seismic data. Tight frame based method (Liang *et al.* 2014) does not rely on fixed basis and its atom is determined through self-learning. But it is time consuming because a bank of compactly supported filters should be obtained first from observed seismic data through self-learning, which limits its wider applications. Dreamlet transform (Geng *et al.* 2009; Wu *et al.* 2013; Wang *et al.* 2014b) uses a physical wavelet as the basic atom that satisfies wave equation automatically, and can represent seismic data sparsely and efficiently compared with curvelet transform (Wang *et al.* 2014b). Therefore, dreamlet transform is adopted as a sparse transform to decompose seismic data in this paper.

Random noise in the observed seismic data can affect the performances of interpolation methods, therefore, many authors study the random noise attenuation algorithms (Liu & Chen 2013; Beckouche & Ma 2014; Chen & Ma 2014). While most random noise attenuation methods take advantage of the predictability of signals, the irregularity of the observed seismic data can affect this predictability and lead to unsatisfactory denoised results. Since the missing traces in the observed seismic data can affect the final performance of random noise attenuation, simultaneous interpolation and denoising methods are studied in (Oropeza & Sacchi 2011; Naghizadeh 2012; Kumar *et al.* 2013). A mask function was designed based on the dominant dips identified by an angular search in the f–k domain (Naghizadeh 2012), which is suitable for linear or quasi-linear events. It should be handled window by window for complicated events in real case studies. A rank reduction algorithm based on multichannel singular spectrum analysis (MSSA) was proposed for simultaneous reconstruction and random noise attenuation (Oropeza & Sacchi 2011). The spatial data at a given temporal frequency is organized into a block Hankel matrix, which is a matrix of rank k in ideal conditions where k is the plane waves in the window of analysis. The missing traces and additive noise can increase the rank and random noise elimination and seismic data interpolation can be achieved through rank reduction using random singular value decomposition (R-SVD). But it is still just suitable for linear or quasi-linear events. Kumar *et al.* (2013) exploited the low rank structure of seismic data directly in midpoint-offset domain and constructed a functional with nuclear norm minimization for simultaneous interpolation and random noise attenuation. Solving the functional requires a projection onto the nuclear norm ball in each iteration by performing a SVD and then thresholding the singular values. SVD is time consuming, and in order to improve the efficiency, factorization-based approach to nuclear norm minimization was used, which should first parametrize the data as the product of two low rank factors which may affect the final performance. Most interpolation methods use low-rank as a constraint for noise attenuation and interpolation because it has all the information of the original noise-free data. Therefore, Habashy *et al.* (2011) used this idea to compress source-receiver data and did full waveform inversion based on source–receiver compression strategy efficiently. Currently, most full waveform inversion methods still depend on the complete seismic data with high SNR, therefore it is essential to interpolate and denoise observed seismic data, simultaneously.

Projection Onto Convex Sets (POCS) method is an efficient method for seismic data interpolation which belongs to the transform based methods. As it cannot handle noisy data properly, a new adaptive method based on the iterative hard threshold (IHT) method is proposed to overcome that defect in this paper. The POCS method proposed by Bregman (1965), was used in image reconstruction (Stark & Oskoui 1989; Wang *et al.* 2014a) and was introduced into irregular seismic data interpolation by Abma & Kabir (2006). Based on its original idea, many effective strategies are proposed. Gao *et al.* (2010) use Fourier transform based POCS method with an exponential threshold model to obtain interpolated data, and the performances of different threshold models are compared to further improve the convergence rate (Gao *et al.* 2012). Curvelet transform (Candes *et al.* 2006), which is a sparse transform and works better for curved seismic events compared with Fourier-based methods, is also used (Yang *et al.* 2012; Zhang & Chen 2013). Dreamlet transform is also used for seismic data interpolation with the POCS method (Wang *et al.* 2014b) and the performance is better than curvelet-based method. While the POCS method has difficulty in noisy data interpolation, because it inserts the noisy observed data in every iteration, and in order to weaken the noise effects, the weighted strategy is used to reconstruct seismic data (Gao *et al.* 2012). However it still reinserts some random noise into the reconstructed seismic data, affecting the final performance. Besides, the performance is unsatisfactory when the first few iteration solutions are far from the real solution and the value of the weight factor is lower. In order to overcome these defects, a new adaptive method is proposed based on the IHT method, taking advantage of threshold strategy to eliminate random noise (Daubechies *et al.* 2004; Herrmann *et al.* 2007).

In this paper, defects of the POCS and the weighted POCS methods are analysed. The POCS method has an implicit assumption that observed seismic data has high SNR, while observed data always contains random noise which deviates from this assumption. The weighted POCS method still reinserts some random noise, and the performance is unsatisfactory when the first few iteration solutions are far from the real solution and the value of the weight factor is lower, because it lacks data residual constraint. Thus a new weighted POCS method is derived from the IHT point of view, and in order to eliminate the effects of noise, a new adaptive method is proposed. In the procedure, jittered undersampling strategy (Hennenfent & Herrmann 2008) is used to obtain the sampled data and dreamlet transform is used to decompose seismic data. Numerical examples on synthetic and real data demonstrate the validity of the proposed method.

2 THEORY

2.1 Dreamlet theory

Dreamlet transform, which is the tensor product of drumbeat transform and beamlet transform (Geng *et al.* 2009; Wu *et al.* 2013), is a sparse transform and can represent seismic data sparsely. In this transform, the basic atom is a physical wavelet which satisfies the wave equation automatically (Wu *et al.* 2011, 2013). The Gabor frame is used as the local decomposition atom in this paper.

A drumbeat atom is shown in formula (1),

$$g_{\bar{t}\bar{\omega}}(t) = W(t-\bar{t})\,\mathrm{e}^{-\mathrm{i}\bar{\omega}t}, \tag{1}$$

which is a time–frequency atom and $W(t)$ is a Gaussian window function. $\bar{t}$ and $\bar{\omega}$ are local time and frequency, respectively. Then a

time-series signal $f(t)$ can be decomposed by drumbeat transform, shown in eq. (2),

$$f(t) = \sum_{\bar{t}\bar{\omega}} \langle f(t), \hat{g}_{\bar{t}\bar{\omega}}(t)\rangle\, g_{\bar{t}\bar{\omega}}(t) = \sum_{\bar{t}\bar{\omega}} \beta_{\bar{t}\bar{\omega}} g_{\bar{t}\bar{\omega}}(t), \tag{2}$$

where $\hat{g}_{\bar{t}\bar{\omega}}(t)$ is the unitary dual vector of $g_{\bar{t}\bar{\omega}}(t)$ determined by $\langle \hat{g}_{\bar{t}\bar{\omega}}(t), g_{\bar{t}\bar{\omega}}(t)\rangle = 1$ (Chen *et al.* 2006); $\beta_{\bar{t}\bar{\omega}}$ is the drumbeat coefficient, which can be obtained through eq. (3),

$$\beta_{\bar{t}\bar{\omega}} = \langle f(t), \hat{g}_{\bar{t}\bar{\omega}}(t)\rangle = \int f(t)\hat{g}^{*}_{\bar{t}\bar{\omega}}(t)\,\mathrm{d}t, \tag{3}$$

where $*$ stands for complex conjugate.

The expression of a beamlet atom is shown as follows:

$$b_{\bar{x}\bar{\xi}}(x) = B(x-\bar{x})\,\mathrm{e}^{i\bar{\xi}x}, \tag{4}$$

which is a space–wavenumber atom and $B(x)$ is also a Gaussian window function. $\bar{x}$ and $\bar{\xi}$ are space and wavenumber, respectively. Then a spatial signal $h(x)$ can be decomposed by beamlet transform, shown in eq. (5),

$$h(x) = \sum_{\bar{x}\bar{\xi}} \left\langle h(x), \hat{b}_{\bar{x}\bar{\xi}}(x)\right\rangle b_{\bar{x}\bar{\xi}}(x) = \sum_{\bar{x}\bar{\xi}} \gamma_{\bar{x}\bar{\xi}} b_{\bar{x}\bar{\xi}}(x), \tag{5}$$

where $\hat{b}_{\bar{x}\bar{\xi}}(x)$ is the unitary dual vector of $b_{\bar{x}\bar{\xi}}(x)$ determined by $\langle \hat{b}_{\bar{x}\bar{\omega}}(x), b_{\bar{x}\bar{\omega}}(x)\rangle = 1$ (Chen *et al.* 2006); $\gamma_{\bar{x}\bar{\xi}}$ is the beamlet coefficient, which subjects to eq. (6),

$$\gamma_{\bar{x}\bar{\xi}} = \langle h(x), \hat{b}_{\bar{x}\bar{\xi}}(x)\rangle = \int h(x)\hat{b}^{*}_{\bar{x}\bar{\xi}}(x)\,\mathrm{d}x. \tag{6}$$

Therefore, a basic dreamlet atom can be constructed through tensor product of drumbeat atom and beamlet atom,

$$d_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}(x,t) = g_{\bar{t}\bar{\omega}}(t) b_{\bar{x}\bar{\xi}}(x). \tag{7}$$

Substitute eqs (1) and (4) into eq. (7), and a localized wave-packet in time–space plane can be constructed,

$$d_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}(x,t) = W(t-\bar{t})B(x-\bar{x})\,\mathrm{e}^{-\mathrm{i}(\bar{\omega}t-\bar{\xi}x)}, \tag{8}$$

A time–space atom constructed in this way satisfies wave equation causality automatically, and can represent seismic data sparsely and efficiently. Fig. 1 shows some of the dreamlet basis, which are the local basis and can be used to decompose seismic data locally.

Given the dreamlet basis $d_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}(x,t)$, dreamlet coefficients $c_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}$, which are corresponding to seismic data $u(x,t)$, can be obtained uniquely,

$$\begin{aligned} c_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}} &= \langle u(x,t), \hat{d}_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}(x,t)\rangle = \iint u(x,t)\hat{d}^{*}_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}(x,t)\,\mathrm{d}x\,\mathrm{d}t \\ &= \iint u(x,t)\hat{g}^{*}_{\bar{t}\bar{\omega}}(t)\,\hat{b}^{*}_{\bar{x}\bar{\xi}}(x)\,\mathrm{d}x\,\mathrm{d}t, \end{aligned} \tag{9}$$

where $*$ stands for complex conjugate and $\hat{d}_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}(x,t)$ is the unitary dual frame of $d_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}(x,t)$. Then, seismic data $u(x,t)$ can be characterized with the unique dreamlet coefficients $c_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}$ and the basis atoms $d_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}(x,t)$,

$$u(x,t) = \sum_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}} \langle u(x,t), \hat{d}_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}(x,t)\rangle d_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}(x,t) = \sum_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}} c_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}} d_{\bar{t}\bar{\omega}\bar{x}\bar{\xi}}(x,t). \tag{10}$$

2.2 POCS method and its extension

Observed seismic data is always irregular and noisy which can be connected with complete noise free seismic data using eq. (11),

$$\mathbf{d}_{\mathrm{obs}} = \mathbf{R}\mathbf{d}_0 + \boldsymbol{\eta}, \tag{11}$$

where $\mathbf{d}_{\mathrm{obs}}$ is the observed seismic data, $\mathbf{d}_0$ represents the complete noise free data, $\boldsymbol{\eta}$ indicates the random noise and $\mathbf{R}$ denotes the sampling matrix which is a diagonal matrix composed of zero and identity matrix, shown in eq. (12).

$$\mathbf{R} = \begin{bmatrix} \mathbf{I} & & & & & \\ & \mathbf{O} & & & & \\ & & \cdots & & & \\ & & & \mathbf{O} & & \\ & & & & \mathbf{I} & \\ & & & & & \mathbf{I} \\ & & & & & & \mathbf{O} \end{bmatrix}. \tag{12}$$

Each $\mathbf{I}$ in eq. (12) corresponds to sampling a trace and each $\mathbf{O}$ corresponds to missing a trace. It is always ill-posed to solve eq. (11) due to effects of random noise and limited bandwidth of observed seismic data. Because seismic data can be sparsely characterized in dreamlet domain, non-constraint objective functional is constructed as follows,

$$\Phi(\mathbf{x}) = \left\|\mathbf{d}_{\mathrm{obs}} - \mathbf{R}\mathbf{D}^{T}\mathbf{x}\right\|_2^2 + \lambda P(\mathbf{x}), \tag{13}$$

where $\mathbf{x}$ represents dreamlet coefficient vector, $\mathbf{D}^{\mathrm{T}}$ denotes inverse dreamlet transform, λ is the regularization factor and $P(\mathbf{x})$ indicates a sparse constraint, such as L_0 or L_1 norm constraint. Eq. (13) with L_0 norm constraint can be solved using the POCS method with hard threshold (Gao *et al.* 2010, 2012; Yang *et al.* 2012, 2013; Wang *et al.* 2014a,b), and the detailed derivation is shown in Appendix A. The exact formula is shown in eq. (14),

$$\tilde{\mathbf{d}}_{k+1} = \mathbf{d}_{\mathrm{obs}} + (\mathbf{I}-\mathbf{R})\mathbf{D}^{T} T_{\lambda_k}\left(\mathbf{D}\tilde{\mathbf{d}}_k\right), \tag{14}$$

where $\tilde{\mathbf{d}}_k$ is the kth iterative solution and T_{λ_k} represents the hard threshold operator performed element-wise which subjects to,

$$T_{\lambda}(x_i) = \begin{cases} x_i, & |x_i| \geq \tau \\ 0, & |x_i| < \tau \end{cases}, \tag{15}$$

where x_i is the ith element of dreamlet coefficient vector $\mathbf{x}$, $|\bullet|$ represents the absolute value operator and $\tau = \sqrt{\lambda}$ is the threshold which can be determined by the exponential threshold model expressed in eq. (16),

$$\tau^{k} = \tau_{\max}\,\mathrm{e}^{c(k-1)/(N-1)}, \quad c = \ln(\tau_{\min}/\tau_{\max}), \quad k = 1, 2, \ldots, N, \tag{16}$$

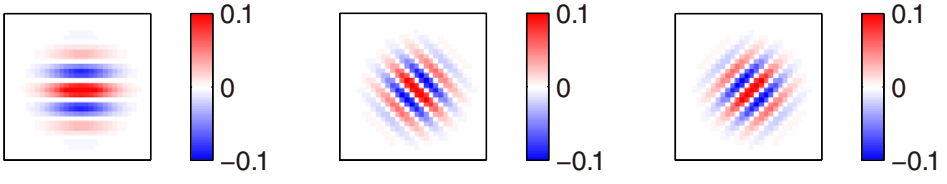

Figure 1. Basic dreamlet atoms.

where $\tau_{\min}$, $\tau_{\max}$ are the minimum and maximum threshold which should be determined according to seismic data.

There is an implicit assumption in the POCS formula that observed seismic data should have high signal-to-noise ratio (SNR) because of the insertion of observed seismic data. However, observed seismic data is always noisy, therefore some authors use the weighted strategy in noisy situations (Oropeza & Sacchi 2011; Gao *et al.* 2012; Stanton & Sacchi 2013; Yang *et al.* 2013). Eq. (14) can be changed into eq. (17) with the weighted strategy,

$$\tilde{\mathbf{d}}_{k+1} = \alpha \mathbf{d}_{\text{obs}} + (\mathbf{I} - \alpha \mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{D}\tilde{\mathbf{d}}_k\right), \tag{17}$$

which could be called the weighted POCS method, and $\alpha \in (0, 1]$ denotes a weight factor determined by the random noise level. Even if the weighted POCS method can weaken the random noise effects, it still inserts the random noise contained in $\alpha \mathbf{d}_{\text{obs}}$, into the updated solution which can affect the final accuracy. When the first few solutions are far from the real solution and the value of the weight factor α is lower, the interpolation performance is unsatisfactory. For example, when α tends to 0 ($\alpha \to 0$), the weighted POCS formula becomes $\tilde{\mathbf{d}}_{k+1} = \mathbf{D}^T T_{\lambda_k}\left(\mathbf{D}\tilde{\mathbf{d}}_k\right)$, therefore the final result is unsatisfactory if the first few solutions are far from the true solution because it lacks data residual constraint. Thus a new weighted POCS method is proposed in the next part, and in order to eliminate the effects of random noise, a new adaptive method is proposed.

2.3 A new adaptive method

From the above analysis, the performances of the POCS and the weighted POCS methods are unsatisfactory because of noisy data ($\mathbf{d}_{\text{obs}}$) insertion or parts of noisy data ($\alpha \mathbf{d}_{\text{obs}}$) insertion. As the α decreases, the inserted components ($\alpha \mathbf{d}_{\text{obs}}$) are much less than the observed data, which leads to the updated solution being far from the true solution in the first few iterations because the weighted POCS method lacks the data residual constraint. Therefore, the performance of the weighted POCS with lower values of α is unsatisfactory if the first few solutions are far from real solution. In order to overcome this defect, we derived a new weighted POCS method based on the IHT method, and in order to eliminate the effects of noise, we exchange the order of projection operator and IHT operator to obtain the novel method which is named as an adaptive method (see Appendix B). The adaptive formula is shown in eq. (18),

$$\begin{aligned} \tilde{\mathbf{d}}_k &= \alpha \mathbf{d}_{\text{obs}} + (\mathbf{I} - \alpha \mathbf{R})\mathbf{d}_{k-1} \\ \mathbf{x}_k &= T_{\lambda_k}\left(\mathbf{D}\left(\tilde{\mathbf{d}}_k + (1-\alpha)\left(\mathbf{d}_{\text{obs}} - \mathbf{R}\mathbf{d}_{k-1}\right)\right)\right) \\ \mathbf{d}_k &= \mathbf{D}^T \mathbf{x}_k \end{aligned} \tag{18}$$

The reconstructed data are controlled by α value as well as the data residual $\mathbf{d}_{\text{obs}} - \mathbf{R}\mathbf{d}_{k-1}$ which guarantees the solution's convergence. Compared with the weighted POCS method, the adaptive method adds a random noise elimination procedure using threshold strategy after each interpolation procedure, and high SNR seismic data can be achieved. With the adaptive method, simultaneous seismic data interpolation and denoising can be obtained.

3 NUMERICAL EXAMPLES

First, a four-layer model is designed to prove the validity of the proposed adaptive method. The weighted POCS method and the adaptive method are tested in a noise free synthetic dataset

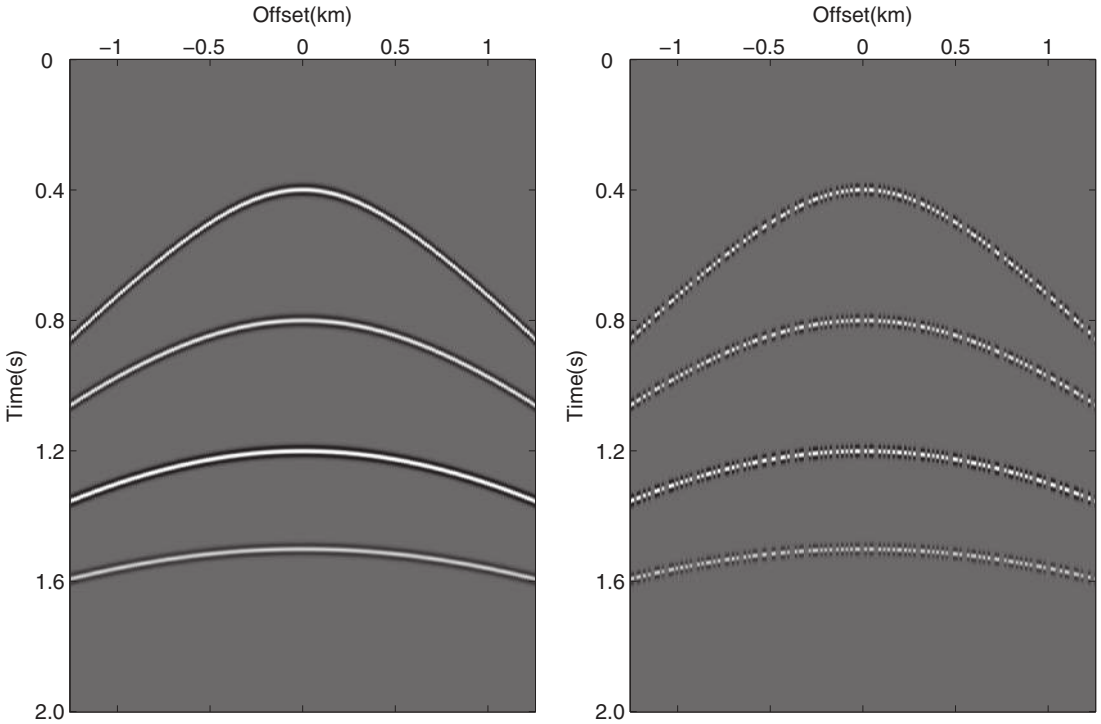

Figure 2. Complete (left-hand panel) and 50 per cent traces missing (right-hand panel) synthetic data.

with different α values. The results demonstrate the validity of the proposed adaptive method and show the defect of the weighted POCS method with lower α value. Tests on noisy synthetic data demonstrate that the adaptive method has superiority to the POCS method and the weighted POCS method. Secondly, numerical examples on real noisy seismic data further demonstrate the validity of the proposed method. In the synthetic and real data examples, 50 per cent traces are missing based on the jittered undersampling strategy (Hennenfent & Herrmann 2008).

3.1 Synthetic example

The synthetic data is shown in Fig. 2 (left-hand panel), which includes 201 traces with 1001 samples per trace. The trace interval and time sampling interval are 12.5 m and 2 ms, respectively. Incomplete seismic data with 50 per cent traces missing is shown in Fig. 2 (right-hand panel). The maximum iteration number is set to 50, and the recovered SNR curves are plotted in Fig. 3. Fig. 3 (left-hand panel) represents recovered SNR curves based on the weighted POCS method and Fig. 3 (right-hand panel) denotes recovered SNR curves based on the new adaptive method with different α values. The definition of SNR is shown below,

$$SNR = 20\ \log_{10} \|\mathbf{d}_0\|_2 / \|\mathbf{d}_{\rm rec} - \mathbf{d}_0\|_2, \tag{19}$$

where $\mathbf{d}_0$ is the original noise free data and $\mathbf{d}_{\rm rec}$ is the reconstructed seismic data.

Fig. 3 demonstrates the validity of the proposed method and points out the defect of the weighted POCS method with lower values of α if the first few solutions are not correct. The thresholds determined by the exponential threshold model decrease with iterations, and the first few thresholds are big which makes the reconstructed seismic data far from the true one in the first few iterations for the weighted POCS method with lower values of α. Since it lacks the data residual constraint, and cannot modify the updated solution into the correct one, therefore the performance is unsatisfactory for lower values of α shown is Fig. 3 (left-hand panel). Fig. 3 (right-hand panel) demonstrates the validity of the proposed adaptive method which is controlled by α value as well as data residual. The weight factor (α) only affects the convergence rate, while the final recovered SNR is almost the same, and the zoomed SNR curves of Fig. 3 (right-hand panel) is shown in Fig. 4 which can better describe the convergence curve.

In order to test the effects of random noise, noisy synthetic data was used to prove that the proposed method is superior to the POCS and the weighted POCS methods. The noisy data with SNR 3.95 dB is shown in Fig. 5 (left-hand panel), and the incomplete one is shown in Fig. 5 (right-hand panel).

Based on the POCS method, the weighted POCS method ($\alpha = 0.6$) and the proposed adaptive method ($\alpha = 0.6$) to interpolate the noisy seismic data, interpolated seismic data can be achieved. The interpolated results and corresponding residuals are shown in Fig. 6. The residual is defined as the difference between the interpolated data and the original noise free data, that is the difference between Fig. 6 (left-hand panel) and Fig. 2 (left-hand panel), and the minor residual means it's more superior of a method.

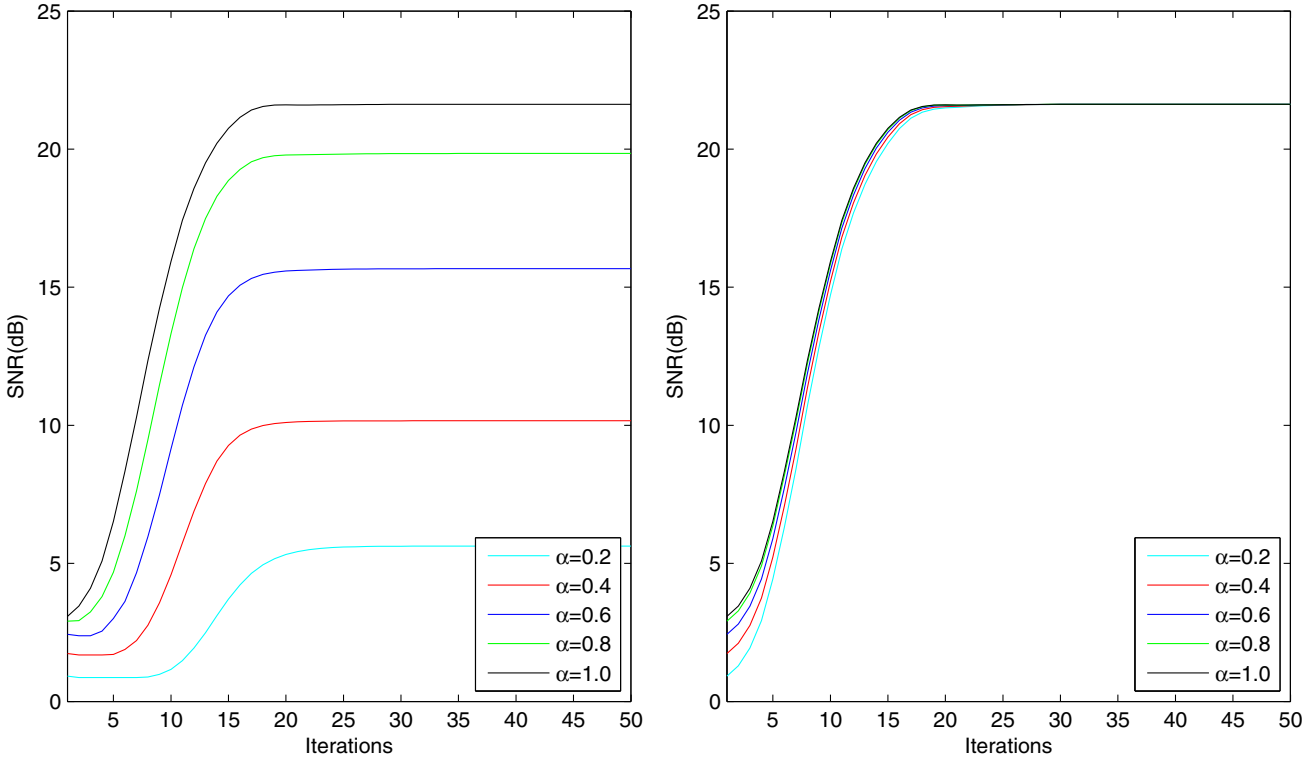

Figure 3. Recovered SNR curves with different weight factor α. Left-hand panel: weighted POCS method; right-hand panel: adaptive method.

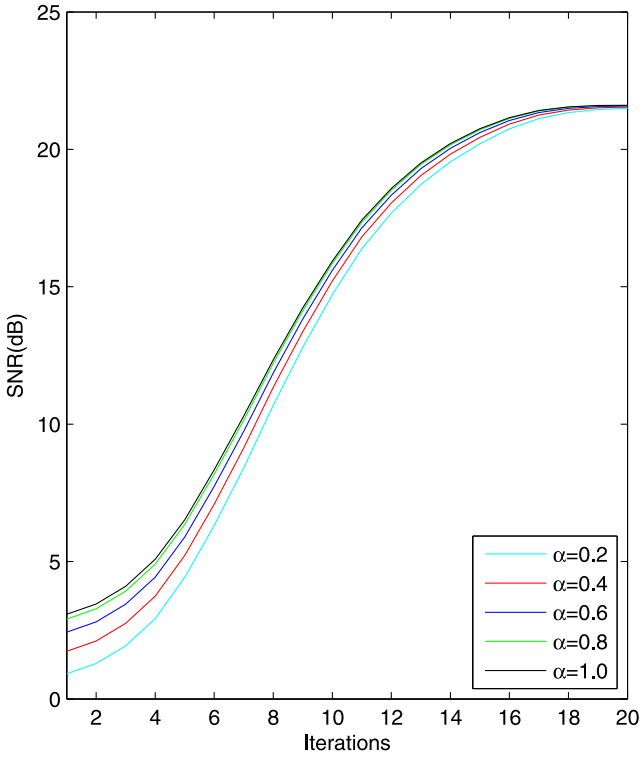

Figure 4. Zoomed figure of Fig. 3 (right-hand panel).

Fig. 6 shows that the weighted POCS can weaken random noise effects to some extent compared with the POCS method, while the adaptive method is the most effective one which can not only interpolate seismic data, but also eliminate the random noise using the threshold operator and its result is consistent with the original noise free data (Fig. 2 left-hand panel). In order to see the details of the interpolated seismic data by the three interpolation methods, we plot single traces of seismic data for better comparisons shown in Fig. 7. Figs 7(a)–(c) represent the 114th trace, 149th trace and 198th trace, respectively. In each subfigure, the 1st trace is from the original noise free data, the 2nd trace is from the interpolated data by the POCS method, the 3rd trace is from the interpolated data by the weighted POCS method and the 4th trace is from the interpolated data by the proposed adaptive method. From Fig. 7, it can be noted that the POCS method cannot handle the noisy data properly because of insertion of noisy data $\mathbf{d}_{\text{obs}}$; the weighted POCS method can weaken the effects of random noise because it just inserts parts of the observed data $\alpha\mathbf{d}_{\text{obs}}$ while the performance is still unsatisfactory; the proposed method can attenuate the random noise using the threshold operator and obtain the satisfactory result. The final recovered SNRs are 6.7, 10.7 and 18.7 dB for the POCS method, the weighted POCS method ($\alpha = 0.6$) and the adaptive method ($\alpha = 0.6$), respectively.

Tests on this simple synthetic data demonstrate the validity of the proposed method, thus an application to a real seismic data is given to further prove its validity.

3.2 Real data application

A marine data is shown in Fig. 8 (left-hand panel), which contains 150 traces and 801 samples per trace. The trace interval is 12.5 m and the time sampling interval is 4 ms. In order to test the superiority of the proposed adaptive method, random noise is added onto the original data, and the noisy data with SNR 6.4 dB is shown in Fig. 8

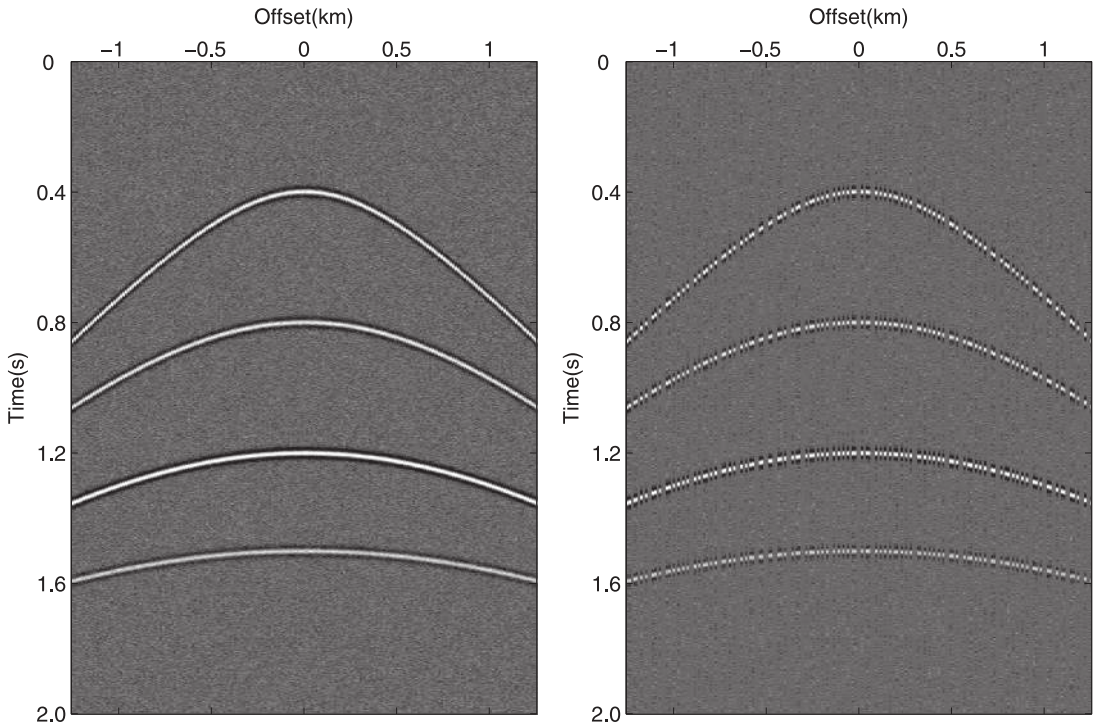

Figure 5. Noisy seismic data (left-hand panel) and incomplete noisy seismic data (right-hand panel).

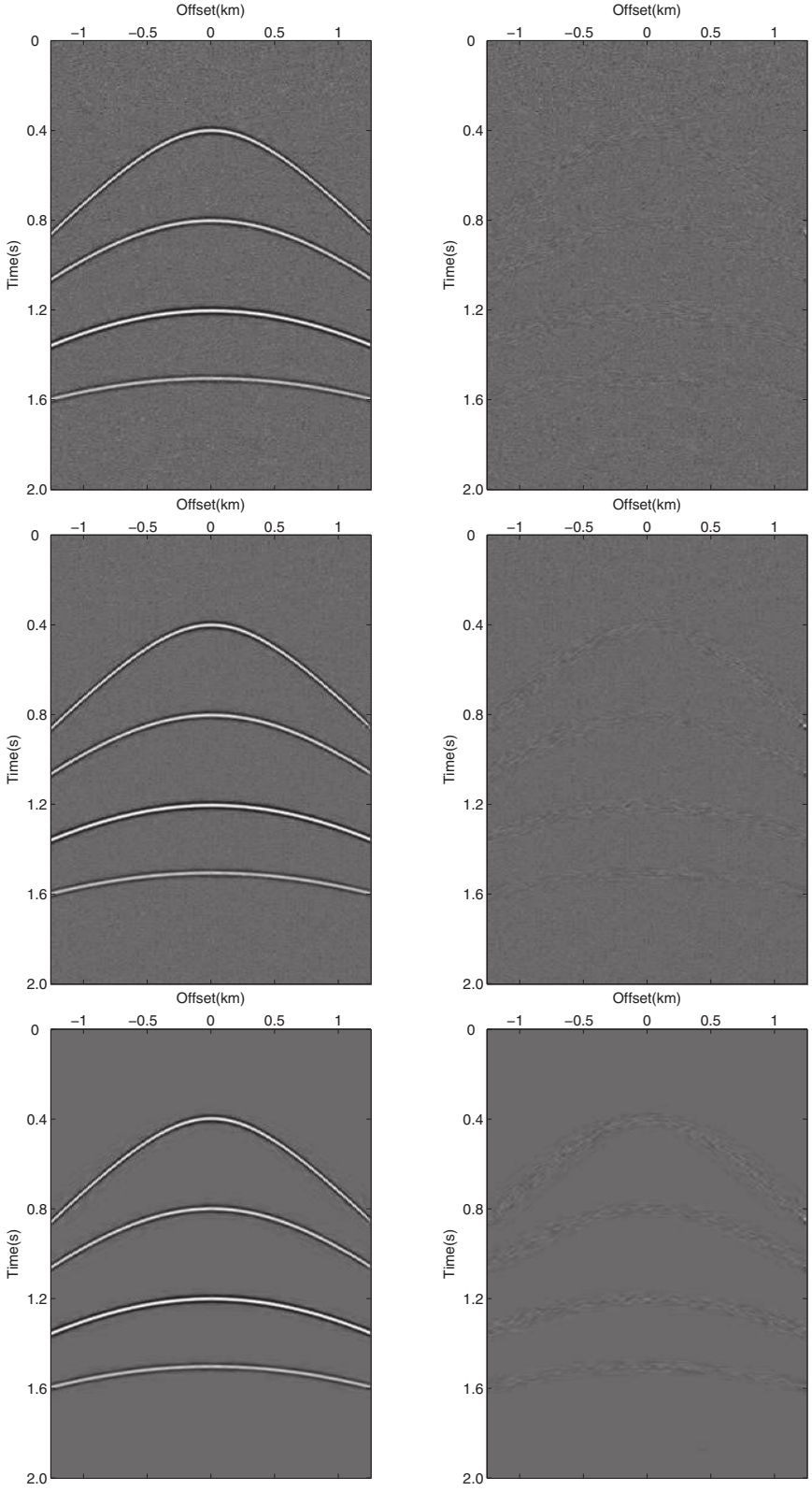

Figure 6. Interpolated results (left-hand column) and corresponding residuals (right-hand column). The top row is for the POCS method, the medium one is for the weighted POCS method ($\alpha = 0.6$) and the bottom one is for the adaptive method ($\alpha = 0.6$).

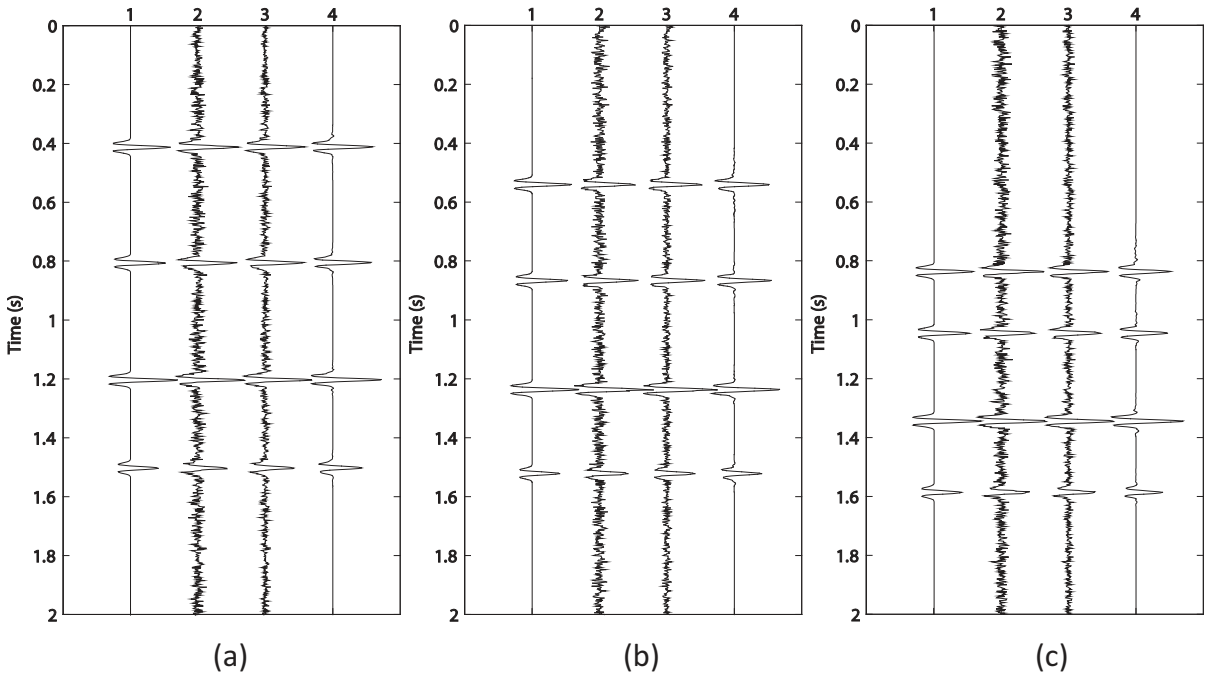

Figure 7. Single trace comparisons of interpolated data and original noise free data. (a) 114th trace; (b) 149th trace and (c) 198th trace from seismic data. In each subfigure, the 1st trace is from original noise free data; the 2nd trace is from interpolated data by the POCS method; the 3rd trace is from interpolated data by the weighted POCS method; the 4th trace is from interpolated data by the proposed method.

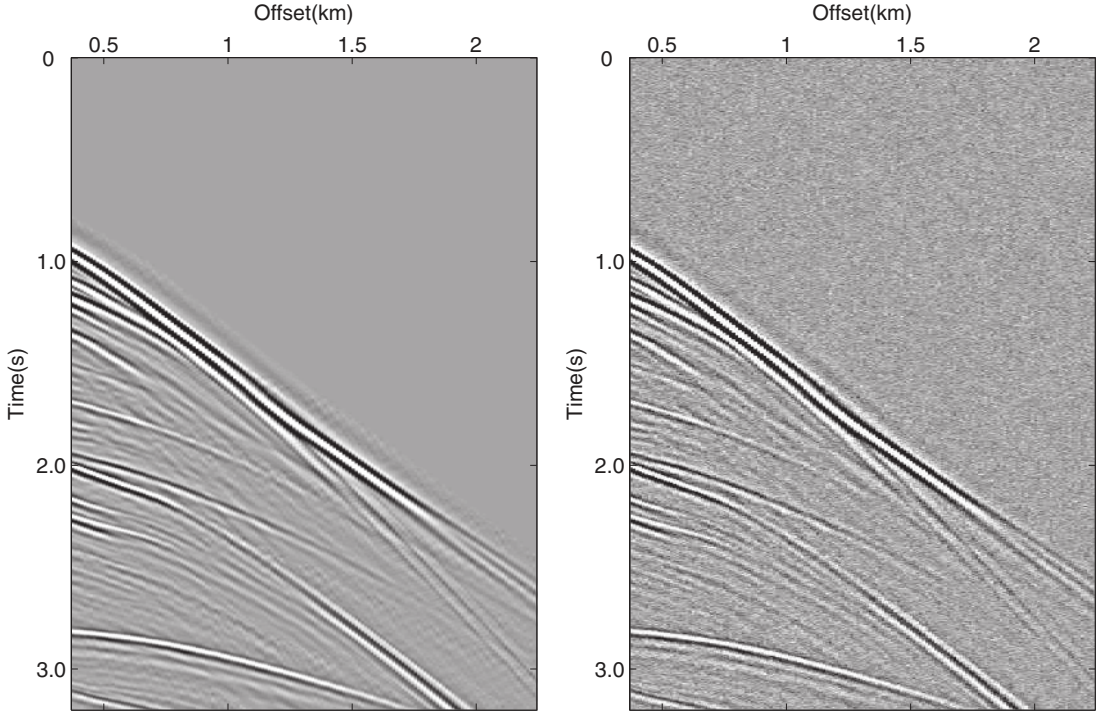

Figure 8. Complete (left-hand panel) and complete noisy (right-hand panel) seismic data.

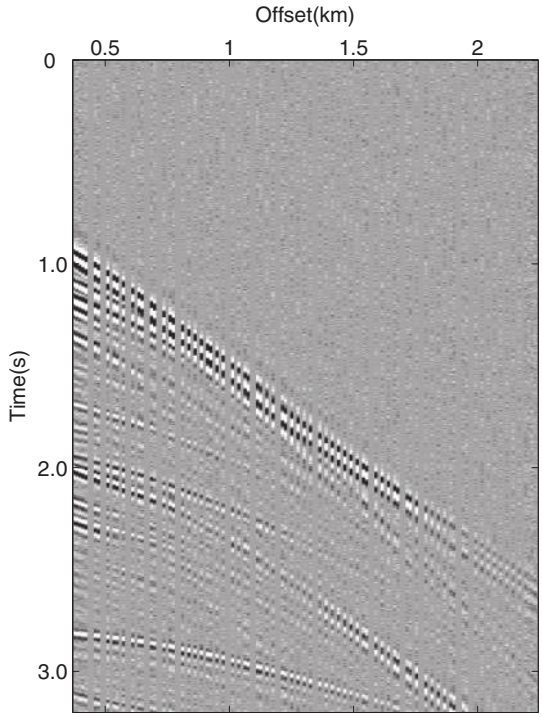

Figure 9. Incomplete noisy seismic data with 50 per cent traces missing.

(right-hand panel) and the incomplete noisy data with 50 per cent traces missing is shown in Fig. 9.

For convenience of comparisons, the POCS method, the weighted POCS method ($\alpha = 0.6$) and the adaptive method ($\alpha = 0.6$) are tested with the same threshold, and the maximum iteration is set to 50. The interpolated results and corresponding residuals are plotted in Fig. 10, and the residual is still defined as the difference between the interpolated data (Fig. 10 left-hand panel) and the original noise free data (Fig. 8 left-hand panel).

From Fig. 10 (left-hand panel), we realize that the interpolated result of the adaptive method is consistent with the original noise free data (Fig. 8 left-hand panel), while the results of the POCS and the weighted POCS methods contain some random noise because of the insertion of noisy data $\mathbf{d}_{\text{obs}}$ or parts of noisy data $\alpha\mathbf{d}_{\text{obs}}$. From the residuals (Fig. 10 right-hand panel), it can be noted that the POCS and the weighted POCS methods insert the random noise which will affect the final SNRs, while the adaptive method eliminates the random noise with threshold strategy, though there is minor signal leaking in the random noise elimination procedure. In order to compare the details of the interpolated results by the three methods, we extract single traces from the interpolated data shown in Fig. 11. Figs 11(a)–(c) represent the 4th trace, 44th trace and 84th trace, respectively. In each sub-figure, the 1st trace is from the original noise free data, the 2nd trace is from the interpolated data by the POCS method, the 3rd trace is from the interpolated data by the weighted POCS method and the 4th trace is from the interpolated data by the proposed adaptive method. From Fig. 11, we notice that the POCS method inserts the observed noisy data $\mathbf{d}_{\text{obs}}$ which makes the final interpolated data with lower SNR; the weighted POCS method can weaken random noise compared with the POCS method because of insertion of parts of the noisy data $\alpha\mathbf{d}_{\text{obs}}$, while the performance is still unsatisfactory; the proposed method can eliminate the random noise using the threshold operator and obtain the satisfactory result which is consistent with the original noise free data (Fig. 8 left-hand panel). The final recovered SNRs are 8.6, 11.3 and 14.7 dB for the POCS method, the weighted POCS method ($\alpha = 0.6$) and the adaptive method ($\alpha = 0.6$), respectively.

From Figs 6, 7, 10 and 11, we can conclude that the proposed adaptive method is more effective in simultaneous interpolation and denoising, compared with the weighted POCS method which is a little superior to the POCS method in weakening random noise. The recovered SNRs in synthetic and real data applications demonstrate the validity of the proposed adaptive method.

4 CONCLUSION

Simultaneous seismic data interpolation and random noise elimination is achieved based on the new adaptive method and dreamlet transform. Defects of the POCS and the weighted POCS methods are analysed for noisy data: the POCS method has an implicit assumption that observed seismic data should have high SNR and has difficulty for noisy data interpolation; the weighted POCS method can weaken the random noise effect, while it still inserts some random noise and its performance is unsatisfactory if the value of weight factor α is lower. Then, based on the IHT method, the adaptive method is proposed to interpolate and denoise seismic data, simultaneously. Numerical examples on synthetic and real data demonstrate that the interpolated results obtained by the adaptive method are better than those obtained by the POCS method and the weighted POCS method in terms of recovered SNRs. The proposed adaptive method can be applied to obtain interpolated data in noisy situations, thus simultaneous interpolation and denoising can be achieved, eventually.

All the interpolated results in this paper are based on jittered undersampling strategy and cannot handle regular sampled data. If regular sampling strategy or random sampling strategy with a big gap is adopted, anti-aliasing strategy should be taken into considerations (Naghizadeh & Sacchi 2010a; Gao *et al.* 2012; Naghizadeh 2012). Because the random noise is eliminated with threshold strategy, there is minor signal leaking and more strategies should be considered to protect the weak signal. Therefore, methods which can achieve simultaneous anti-aliasing interpolation and random noise elimination with weak signal protection, should be developed in the future.

ACKNOWLEDGEMENTS

We thank Yu Geng and Jingjie Cao for helpful and valuable discussions. This research work was conducted during Benfeng Wang's visit to the University of California, Santa Cruz. It is supported by the National Natural Science Foundation of China (Grant No. U1262207), the Key State Science and Technology Project (Grant Nos. 2011ZX05023-005-005, 2011ZX05019-006) and the WTOPI Research Consortium at the University of California, Santa Cruz. The first author would like to thank the China Scholarship Council for their financial support to study abroad.

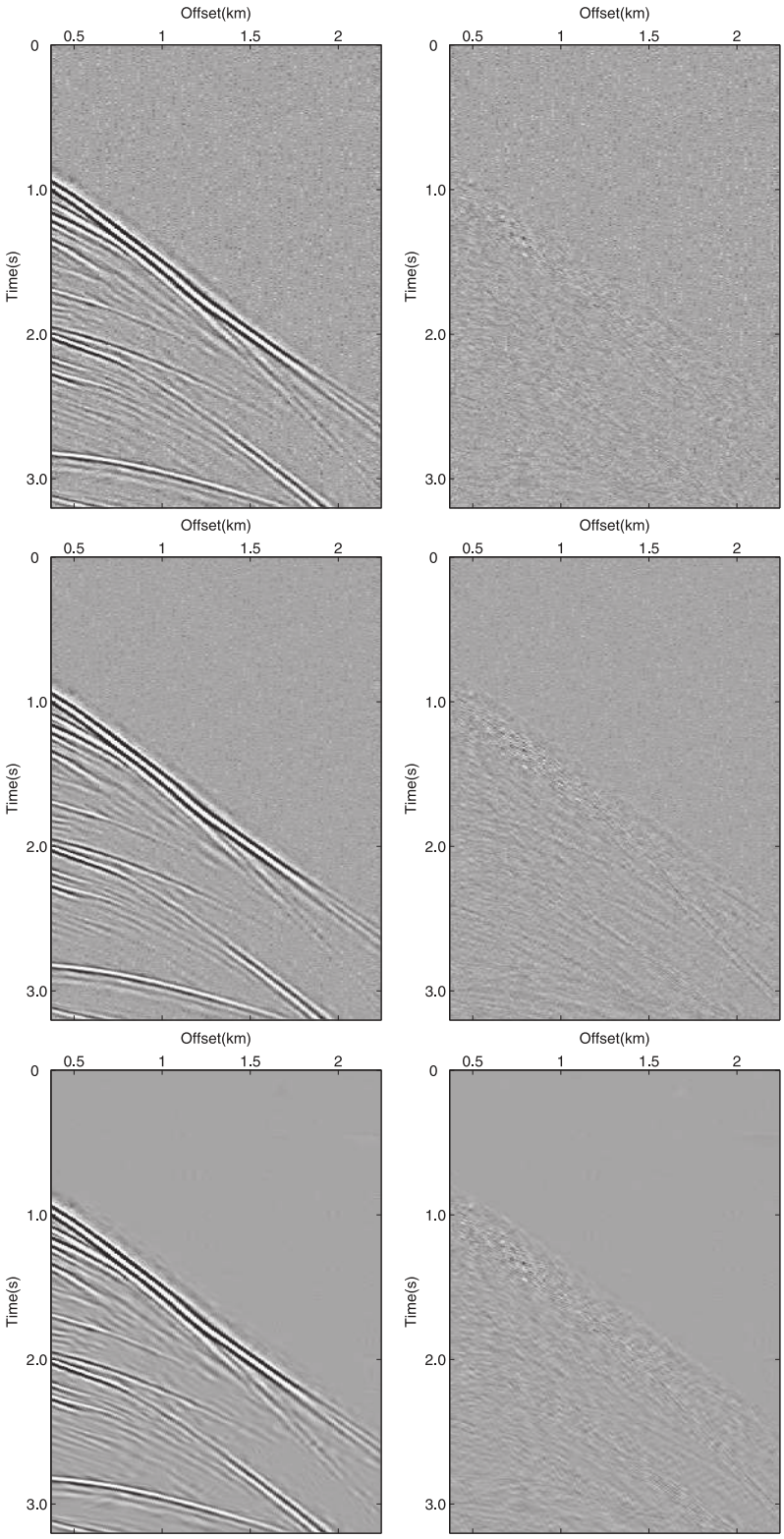

Figure 10. Interpolated results (left-hand column) and corresponding residuals (right-hand column). The top row is for the POCS method, the medium one is for the weighted POCS method ($\alpha = 0.6$) and the bottom one is for the adaptive method ($\alpha = 0.6$).

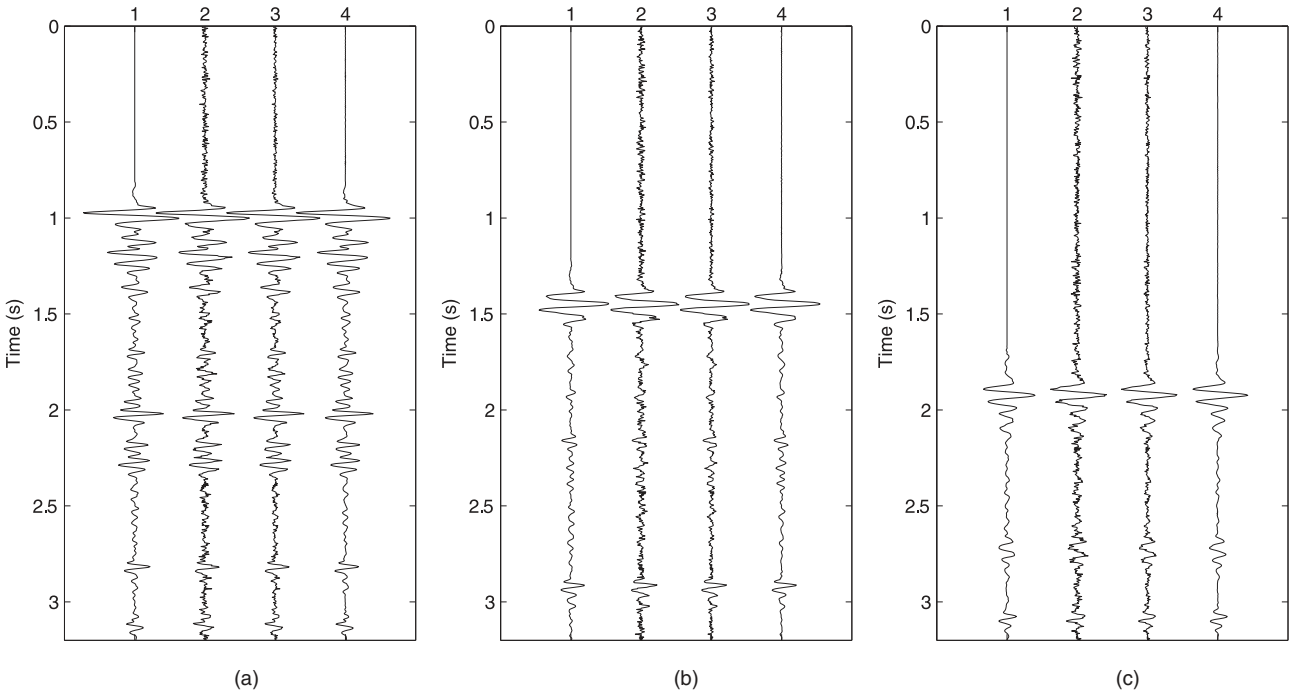

Figure 11. Single trace comparisons of interpolated data and original noise free data. (a) 4th trace; (b) 44th trace and (c) 84th trace from seismic data. In each subfigure, the 1st trace is from original noise free data; the 2nd trace is from interpolated data by the POCS method; the 3rd trace is from interpolated data by the weighted POCS method; the 4th trace is from interpolated data by the proposed method.

REFERENCES

Abma, R. & Kabir, N., 2006. 3D interpolation of irregular data with a POCS algorithm, *Geophysics,* **71**(6), E91–E97.

Beckouche, S. & Ma, J., 2014. Simultaneous dictionary learning and denoising for seismic data, *Geophysics,* **79**(3), A27–A31.

Blumensath, T. & Davies, M.E., 2008. Iterative thresholding for sparse approximations, *J. Fourier Anal. Appl.,* **14**(5–6), 629–654.

Blumensath, T. & Davies, M.E., 2009. Iterative hard thresholding for compressed sensing, *Appl. Comput. Harmonic Anal.,* **27**(3), 265–274.

Bregman, L., 1965. The method of successive projection for finding a common point of convex sets (Theorems for determining common point of convex sets by method of successive projection), *Soviet Math.,* **6,** 688–692.

Candes, E., Demanet, L., Donoho, D. & Ying, L., 2006. Fast discrete curvelet transforms, *Multiscale Modeling Simulat.,* **5**(3), 861–899.

Chen, Y. & Ma, J., 2014. Random noise attenuation by f-x empirical-mode decomposition predictive filtering, *Geophysics,* **79**(3), V81–V91.

Chen, L., Wu, R. & Chen, Y., 2006. Target-oriented beamlet migration based on Gabor-Daubechies frame decomposition, *Geophysics,* **71**(2), S37–S52.

Daubechies, I., Defrise, M. & De Mol, C., 2004. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Commun. Pure appl. Math.,* **57**(11), 1413–1457.

Gao, J., Chen, X., Li, J., Liu, G. & Ma, J., 2010. Irregular seismic data reconstruction based on exponential threshold model of POCS method, *Appl. Geophys.,* **7**(3), 229–238.

Gao, J., Stanton, A., Naghizadeh, M., Sacchi, M.D. & Chen, X., 2012. Convergence improvement and noise attenuation considerations for beyond alias projection onto convex sets reconstruction, *Geophys. Prospect.,* **61**(Suppl. 1), 138–151.

Gao, J., Sacchi, M. & Chen, X., 2013. A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions, *Geophysics,* **78**(1), V21–V30.

Geng, Y., Wu, R. & Gao, J., 2009. Dreamlet transform applied to seismic data compression and its effects on migration, in *Proceedings of the SEG Annual Meeting*, Houston.

Habashy, T.M., Abubakar, A., Pan, G. & Belani, A., 2011. Source-receiver compression scheme for full-waveform seismic inversion, *Geophysics,* **76**(4), R95–R108.

Hennenfent, G. & Herrmann, F.J., 2008. Simply denoise: wavefield reconstruction via jittered undersampling, *Geophysics,* **73**(3), V19–V28.

Herrmann, F.J. & Hennenfent, G., 2008. Non-parametric seismic data recovery with curvelet frames, *Geophys. J. Int.,* **173**(1), 233–248.

Herrmann, F.J., Böniger, U. & Verschuur, D.J.E., 2007. Non-linear primary-multiple separation with directional curvelet frames, *Geophys. J. Int.,* **170**(2), 781–799.

Kreimer, N., Stanton, A. & Sacchi, M., 2013. Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction, *Geophysics,* **78**(6), V273–V284.

Kumar, R., Aravkin, A.Y., Mansour, H., Recht, B. & Herrmann, F.J., 2013. Seismic data interpolation and denoising using svd-free low-rank matrix factorization, in *Proceedings of the 75th EAGE Conference & Exhibition incorporating SPE EUROPEC 2013*.

Liang, J., Ma, J. & Zhang, X., 2014. Seismic data restoration via data-driven tight frame, *Geophysics,* **79**(3), V65–V74.

Liu, G. & Chen, X., 2013. Noncausal f–x–y regularized nonstationary prediction filtering for random noise attenuation on 3D seismic data, *J. appl. Geophys.,* **93,** 60–66.

Loris, I., Douma, H., Nolet, G., Daubechies, I. & Regone, C., 2010. Nonlinear regularization techniques for seismic tomography, *J. Comput. Phys.,* **229**(3), 890–905.

Ma, J., 2013. Three-dimensional irregular seismic data reconstruction via low-rank matrix completion, *Geophysics,* **78**(5), V181–V192.

Mansour, H., Herrmann, F. & Yılmaz, Ö., 2013. Improved wavefield reconstruction from randomized sampling via weighted one-norm minimization, *Geophysics,* **78**(5), V193–V206.

Naghizadeh, M., 2012. Seismic data interpolation and denoising in the frequency-wavenumber domain, *Geophysics,* **77**(2), V71–V80.

Naghizadeh, M. & Sacchi, M.D., 2007. Multistep autoregressive reconstruction of seismic records, *Geophysics,* **72**(6), V111–V118.

Naghizadeh, M. & Sacchi, M.D., 2010a. Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data, *Geophysics,* **75**(6), WB189–WB202.

Naghizadeh, M. & Sacchi, M.D., 2010b. On sampling functions and Fourier reconstruction methods, *Geophysics,* **75**(6), WB137–WB151.
Naghizadeh, M. & Innanen, K.A., 2011. Seismic data interpolation using a fast generalized Fourier transform, *Geophysics,* **76**(1), V1–V10.
Oropeza, V. & Sacchi, M., 2011. Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis, *Geophysics,* **76**(3), V25–V32.
Ronen, J., 1987. Wave-equation trace interpolation, *Geophysics,* **52**(7), 973–984.
Shahidi, R., Tang, G., Ma, J. & Herrmann, F.J., 2013. Application of randomized sampling schemes to curvelet-based sparsity-promoting seismic data recovery, *Geophys. Prospect.,* **61**(5), 973–997.
Spitz, S., 1991. Seismic trace interpolation in the F-X domain, *Geophysics,* **56**(6), 785–794.
Stanton, A. & Sacchi, M.D., 2013. Vector reconstruction of multicomponent seismic data, *Geophysics,* **78**(4), V131–V145.
Stark, H. & Oskoui, P., 1989. High-resolution image recovery from image-plane arrays, using convex projections, *J. Opt. Soc. Am., A,* **6**(11), 1715–1726.
Trad, D.O., Ulrych, T.J. & Sacchi, M.D., 2002. Accurate interpolation with high-resolution time-variant Radon transforms, *Geophysics,* **67**(2), 644–656.
Wang, J., Ng, M. & Perz, M., 2010. Seismic data interpolation by greedy local Radon transform, *Geophysics,* **75**(6), WB225–WB234.
Wang, B., Chen, X., Li, J. & Cao, J., 2014a. Curvelet-based 3D reconstruction of digital cores using POCS method, *Chinese J. Geophys.*, in press.
Wang, B., Wu, R., Geng, Y. & Chen, X., 2014b. Dreamlet-based interpolation using POCS method, *J. appl. Geophys.,* **109,** 256–265.
Wu, R., Geng, Y. & Wu, B., 2011. Physical wavelet defined on an observation plane and the Dreamlet, in *Proceedings of the SEG Annual Meeting*, San Antonio.
Wu, R., Geng, Y. & Ye, L., 2013. Preliminary study on Dreamlet based compressive sensing data recovery, in *Proceedings of the SEG Annual Meeting*, Houston.
Xu, S., Zhang, Y. & Lambaré, G., 2010. Antileakage Fourier transform for seismic data regularization in higher dimensions, *Geophysics,* **75**(6), WB113–WB120.
Xue, Y., Ma, J. & Chen, X., 2014. High-order sparse Radon transform for AVO-preserving data reconstruction, *Geophysics,* **79**(2), V13–V22.
Yang, P., Gao, J. & Chen, W., 2012. Curvelet-based POCS interpolation of nonuniformly sampled seismic records, *J. appl. Geophys.,* **79,** 90–99.
Yang, P., Gao, J. & Chen, W., 2013. On analysis-based two-step interpolation methods for randomly sampled seismic data, *Comput. Geosci.,* **51,** 449–461.
Yu, Z., Ferguson, J., McMechan, G. & Anno, P., 2007. Wavelet-Radon domain dealiasing and interpolation of seismic data, *Geophysics,* **72**(2), V41–V49.
Zhang, H. & Chen, X.-H., 2013. Seismic data reconstruction based on jittered sampling and curvelet transform, *Chinese J. Geophys.,* **56**(5), 1637–1649.
Zhang, H., Chen, X.-H. & Wu, X.-M., 2013. Seismic data reconstruction based on CS and Fourier theory, *Appl. Geophys.,* **10**(2), 170–180.

APPENDIX A: DERIVATION OF THE POCS METHOD

Functional (13) with L_0 constraint can be solved through the iterative hard threshold (IHT) algorithm (Blumensath & Davies 2008, 2009; Loris *et al.* 2010). Then, the iterative solution can be obtained by the IHT method,

$$\mathbf{x}_{k+1} = T_{\lambda_k}\left(\mathbf{x}_k + \left(\mathbf{RD}^T\right)^T\left(\mathbf{d}_{\text{obs}} - \mathbf{RD}^T\mathbf{x}_k\right)\right), \tag{A1}$$

where $\mathbf{x}_k$ is the kth iterative solution in the dreamlet domain and T_{λ_k} is the hard threshold operator performed element-wise which subjects to,

$$T_\lambda(x_i) = \begin{cases} x_i, & |x_i| \geq \tau \\ 0, & |x_i| < \tau \end{cases}, \tag{A2}$$

where x_i is the ith element of dreamlet coefficient vector $\mathbf{x}$, $|\cdot|$ represents the absolute value operator and $\tau = \sqrt{\lambda}$ is the threshold which can be determined by an exponential threshold model. Seismic data interpolation is to obtain the whole seismic data in the data domain, therefore, projecting the updated solution $\mathbf{x}_{k+1}$ onto the convex observation plane $\{\mathbf{d}\,|\mathbf{d}_{\text{obs}} = \mathbf{Rd}\}$ can improve the convergence rate. Then the POCS formula is derived, shown as follows,

$$\begin{aligned}
\tilde{\mathbf{d}}_{k+1} &= \mathbf{d}_{\text{obs}} + (\mathbf{I}-\mathbf{R})\mathbf{D}^T\mathbf{x}_{k+1} \\
&= \mathbf{d}_{\text{obs}} + (\mathbf{I}-\mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{x}_k + \left(\mathbf{RD}^T\right)^T\left(\mathbf{d}_{\text{obs}} - \mathbf{RD}^T\mathbf{x}_k\right)\right) \\
&= \mathbf{d}_{\text{obs}} + (\mathbf{I}-\mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{Dd}_k + \mathbf{D}\left(\mathbf{R}^T\mathbf{d}_{\text{obs}} - \mathbf{R}^T\mathbf{RD}^T\mathbf{x}_k\right)\right) \\
&= \mathbf{d}_{\text{obs}} + (\mathbf{I}-\mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{Dd}_k + \mathbf{D}\left(\mathbf{d}_{\text{obs}} - \mathbf{Rd}_k\right)\right) \\
&= \mathbf{d}_{\text{obs}} + (\mathbf{I}-\mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{D}\left(\mathbf{d}_{\text{obs}} + (\mathbf{I}-\mathbf{R})\,\mathbf{d}_k\right)\right)
\end{aligned} \tag{A3}$$

where $\mathbf{d}_k = \mathbf{D}^T\mathbf{x}_k$ is the solution in data space, $\tilde{\mathbf{d}}_{k+1}$ is the solution after inserting observed data. Sampling matrix $\mathbf{R}$, which is a diagonal matrix shown in eq. (12), subjects to $\mathbf{R} = \mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{RR}$. Denoting $\tilde{\mathbf{d}}_k = \mathbf{d}_{\text{obs}} + (\mathbf{I}-\mathbf{R})\,\mathbf{d}_k$, then the POCS formula can be derived,

$$\tilde{\mathbf{d}}_{k+1} = \mathbf{d}_{\text{obs}} + (\mathbf{I}-\mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{D}\tilde{\mathbf{d}}_k\right). \tag{A4}$$

In fact, the POCS method belongs to two-step methods. In the first step, we get the updated solution $\mathbf{x}$ in the dreamlet domain; then, project the solution onto the convex observation plane $\{\mathbf{d}\,|\mathbf{d}_{\text{obs}} = \mathbf{Rd}\}$ to accelerate the interpolation in data space domain.

In noisy situations, in order to weaken the random noise effects, weighted strategy is adopted and the weighted POCS method can be derived,

$$\tilde{\mathbf{d}}_{k+1} = \alpha\mathbf{d}_{\text{obs}} + (\mathbf{I}-\alpha\mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{D}\tilde{\mathbf{d}}_k\right). \tag{A5}$$

where $\alpha \in (0, 1]$ is a weight factor.

APPENDIX B: DERIVATION OF THE NEW ADAPTIVE METHOD

Appendix A gives the derivation of the POCS method, while it cannot handle noisy data interpolation properly because of observed seismic data insertion. Similar with derivation of the POCS method, we propose a new adaptive method in this Appendix B.

Seismic data interpolation is to obtain the whole seismic data in the data domain, therefore, projecting the updated solution $\mathbf{x}_{k+1}$ obtained in eq. (A1) onto the observation plane $\{\mathbf{d}\,|\mathbf{d}_{\text{obs}} = \mathbf{Rd}\}$ can improve the convergence rate. Since observed data is always noisy, only parts of observed seismic $\alpha\mathbf{d}_{\text{obs}}$ are inserted into the updated solution (A1) using weighted strategy. Then a new weighted POCS formula is derived, shown as follows,

$$\begin{aligned}
\tilde{\mathbf{d}}_{k+1} &= \alpha\mathbf{d}_{\text{obs}} + (\mathbf{I}-\alpha\mathbf{R})\mathbf{D}^T\mathbf{x}_{k+1} \\
&= \alpha\mathbf{d}_{\text{obs}} + (\mathbf{I}-\alpha\mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{x}_k + \left(\mathbf{RD}^T\right)^T\left(\mathbf{d}_{\text{obs}} - \mathbf{RD}^T\mathbf{x}_k\right)\right) \\
&= \alpha\mathbf{d}_{\text{obs}} + (\mathbf{I}-\alpha\mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{Dd}_k + \mathbf{D}\left(\mathbf{R}^T\mathbf{d}_{\text{obs}} - \mathbf{R}^T\mathbf{RD}^T\mathbf{x}_k\right)\right) \\
&= \alpha\mathbf{d}_{\text{obs}} + (\mathbf{I}-\alpha\mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{Dd}_k + \mathbf{D}\left(\mathbf{d}_{\text{obs}} - \mathbf{Rd}_k\right)\right) \\
&= \alpha\mathbf{d}_{\text{obs}} + (\mathbf{I}-\alpha\mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{D}\left(\alpha\mathbf{d}_{\text{obs}} + (\mathbf{I}-\alpha\mathbf{R})\,\mathbf{d}_k\right.\right. \\
&\quad \left.\left. + (1-\alpha)\left(\mathbf{d}_{\text{obs}} - \mathbf{Rd}_k\right)\right)\right),
\end{aligned} \tag{B1}$$

where $\mathbf{d}_k = \mathbf{D}^T\mathbf{x}_k$ is the solution in data space, $\tilde{\mathbf{d}}_{k+1}$ is the solution after inserting parts of observed data $\alpha\mathbf{d}_{\text{obs}}$, and α is a weight factor. Sampling matrix $\mathbf{R}$, which is a diagonal matrix shown in eq. (12), subjects to $\mathbf{R} = \mathbf{R}^T = \mathbf{R}^T\mathbf{R} = \mathbf{R}\mathbf{R}$. Denoting $\tilde{\mathbf{d}}_k = \alpha\mathbf{d}_{\text{obs}} + (\mathbf{I} - \alpha\mathbf{R})\,\mathbf{d}_k$, then the new weighted POCS formula can be derived,

$$\tilde{\mathbf{d}}_{k+1} = \alpha\mathbf{d}_{\text{obs}} + (\mathbf{I} - \alpha\mathbf{R})\mathbf{D}^T T_{\lambda_k}\left(\mathbf{D}\left(\tilde{\mathbf{d}}_k + (1-\alpha)\left(\mathbf{d}_{\text{obs}} - \mathbf{R}\mathbf{d}_k\right)\right)\right). \tag{B2}$$

The new weighted POCS method, which is different from the original one (eq. (17)), can be controlled by α value as well as the data residual term $\mathbf{d}_{\text{obs}} - \mathbf{R}\mathbf{d}_k$. Even if α is lower, it can still converge because of constraint of the data residual. In order to eliminate the effects of random noise, we exchange the order of the IHT operator and projection operator and obtain the new adaptive formula,

$$\begin{aligned}\tilde{\mathbf{d}}_k &= \alpha\mathbf{d}_{\text{obs}} + (\mathbf{I} - \alpha\mathbf{R})\mathbf{d}_{k-1}\\ \mathbf{x}_k &= T_{\lambda_k}\left(\mathbf{D}\left(\tilde{\mathbf{d}}_k + (1-\alpha)\left(\mathbf{d}_{\text{obs}} - \mathbf{R}\mathbf{d}_{k-1}\right)\right)\right).\\ \mathbf{d}_k &= \mathbf{D}^T\mathbf{x}_k\end{aligned} \tag{B3}$$

This adaptive formula can deal with noisy seismic data and can obtain simultaneous seismic data interpolation and denoising. When $\alpha = 0$, the adaptive formula degenerates to IHT method but in data space domain; when $\alpha = 1$, it degenerates to the POCS method plus random noise elimination with threshold strategy. Therefore, the new adaptive method has more flexibility and can be used in wider applications.