Plane-wave orthogonal polynomial transform for amplitude-preserving noise attenuation

Chen, Yangkang; Huang, Weilin; Zhou, Yatong; Liu, Wei; Zhang, Dong

DOI
10.1093/GJI/GGY267

Publication date
2018

Document Version
Final published version

Published in
Geophysical Journal International

Citation (APA)
Chen, Y., Huang, W., Zhou, Y., Liu, W., & Zhang, D. (2018). Plane-wave orthogonal polynomial transform for amplitude-preserving noise attenuation. *Geophysical Journal International*, *214*(3), 2207-2223. https://doi.org/10.1093/GJI/GGY267

Geophysical Journal International

Geophys. J. Int. (2018) **214,** 2207–2223
Advance Access publication 2018 July 04
GJI Seismology

doi: 10.1093/gji/ggy267

RETRACTED

Plane-wave orthogonal polynomial transform for amplitude-preserving noise attenuation

Yangkang Chen,[1] Weilin Huang,[2] Yatong Zhou,[3] Wei Liu[2] and Dong Zhang[4]

[1] *School of Earth Sciences, Zhejiang University, Hangzhou, Zhejiang Province* 310027*, China. E-mail: yangkang.chen@zju.edu.cn*
[2] *State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum-Beijing, Beijing* 102249*, China*
[3] *Tianjin Key Laboratory of Electronic Materials and Devices, School of Electronics and Information Engineering, Hebei Universit*[illegible]5340 *Xiping Road, Beichen District, Tianjin* 300401*, P.R. China. E-mail: zyt_htu@126.com*
[4] *Department of Imaging Physics, Delft University of Technology, Lorentzweg 1, 2628 CJ, Delft, The Netherlands*

Accepted 2018 July 2. Received 2018 January 5; in original form 2018 May 31

SUMMARY

Amplitude-preserving data processing is an imp[illegible]and chall[illegible]opic in many scientific fields. The amplitude-variation details in sei[illegible] are especi[illegible]y important because the amplitude variation is directly related with t[illegible]bsu[illegible]wave impedance and fluid characteristics. We propose a novel seismic noise attenuation ap[illegible]that is based on local plane-wave assumption of seismic events and th[illegible]tude preservi[illegible]capability of the orthogonal polynomial transform (OPT). The OP[illegible]a way for r[illegible]presenting spatially correlative seismic data as a superposition of polynomial[illegible] functio[illegible]y which the random noise is distinguished from the useful energy by the hig[illegible]gonal[illegible]ynomial coefficients. The seismic energy is the most correlative alo[illegible] the struc[illegible]on and thus the OPT is optimally performed in a flattened gather. [illegible]e in detail the flattening operator for creating the flattened dimension, where the[illegible]T ca[illegible]lied subsequently. The flattening operator is created by deriving a plane-wa[illegible]uation relation following the plane-wave equation. We demonstrate[illegible]oth pl[illegible]ave trace continuation and OPT can well preserve the strong amplitud[illegible]existin[illegible]eismic data. In order to obtain a robust slope estimation performanc[illegible]e pr[illegible]of no[illegible]e, a robust slope estimation approach is introduced to substitute the tra[illegible]group of synthetic, pre-stack and post-stack field seismic data are [illegible]to de[illegible]rate the potential of the proposed framework in realistic applications.

[illegible]y words: [illegible]e Processing; Seismic Noise.

1 INTROD[illegible]IO[illegible]

Seismic noise att[illegible]one o[illegible]most significant steps in the whole seis[illegible]data p[illegible]ing [illegible]imaging workflow. It has great influen[illegible] subse[illegible]ocessing tasks, such as amplitude-vari[illegible]-offs[illegible]nversion, [illegible]erse time migration, full waveform in[illegible]a[illegible]interpretation for oil and gas detection (Wa[illegible]2015; [illegible]*et al.*, 2016; Gao *et al.*, 2016; Huang *et al.*, 2016; Q[illegible]l., 2016; Xue *et al.*, 2016b; Asgedom *et al.*, 2017; Zeng *et al*[illegible]; Zhang *et al.*, 2017; Chen, 2018).

In the past several decades, a large number of algorithms have been developed for seismic noise attenuation. Stacking the seismic data along the spatial directions, for example, the offset direction, can enhance the energy of spatially coherent useful waveform signals as well as mitigate the spatially incoherent random noise (Liu *et al.*, 2009a; Xie *et al.*, 2016; Wu & Bai, 2018c). One of the commonly used state-of-the-art algorithms is the prediction-based method, including *t–x* predictive filtering (Abma & Claerbout, 1995), *f–x* deconvolution (Canales, 1984), the polynomial fitting based approach (Liu *et al.*, 2011) and non-stationary predictive filtering (Liu *et al.*, 2012; Liu & Chen, 2013). This type of methods utilize the predictive property of useful signals along spatial direction to create a regression-like model for distinguishing between signal and noise.

Another type of commonly used methods are based on data decomposition. This type of methods assume that noisy seismic data can be decomposed into different components where signal and noise are separated based on their frequency difference or morphological difference (Huang *et al.*, 2017). Empirical mode decomposition (EMD; Huang *et al.* 1998; Chen 2016) and its improved version, for example, ensemble empirical mode decomposition (EEMD; Wu & Huang 2009), complete ensemble empirical mode decomposition (CEEMD; Colominas *et al.* 2012;), have been used intensively for reducing the noise in seismic data (Chen *et al.*, 2016a) . Variational mode decomposition was proposed by Dragomiretskiy & Zosso (2014) for substituting EMD because of its explicit control on the decomposition performance. It has been utilized for noise attenuation in Liu *et al.* (2017) and for time-frequency analysis by Liu

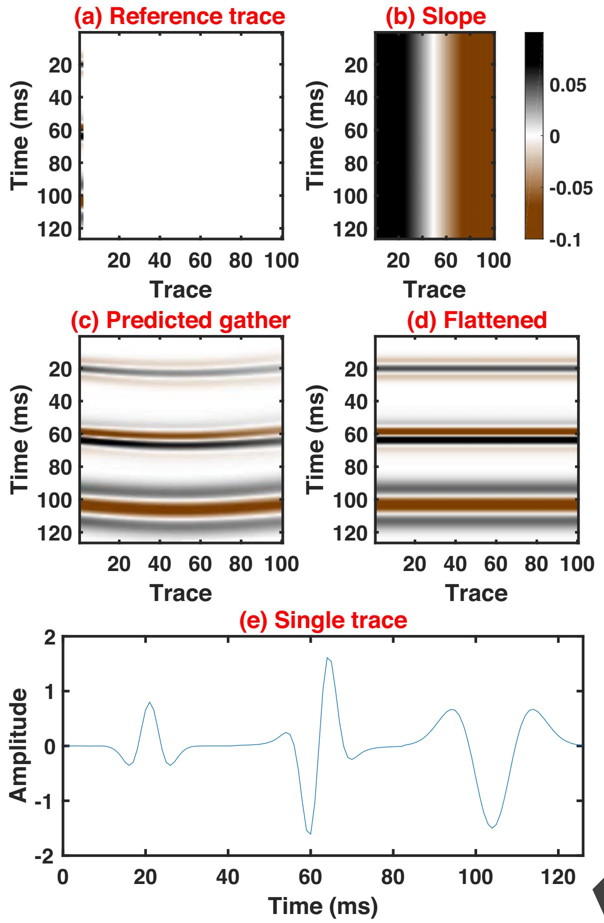
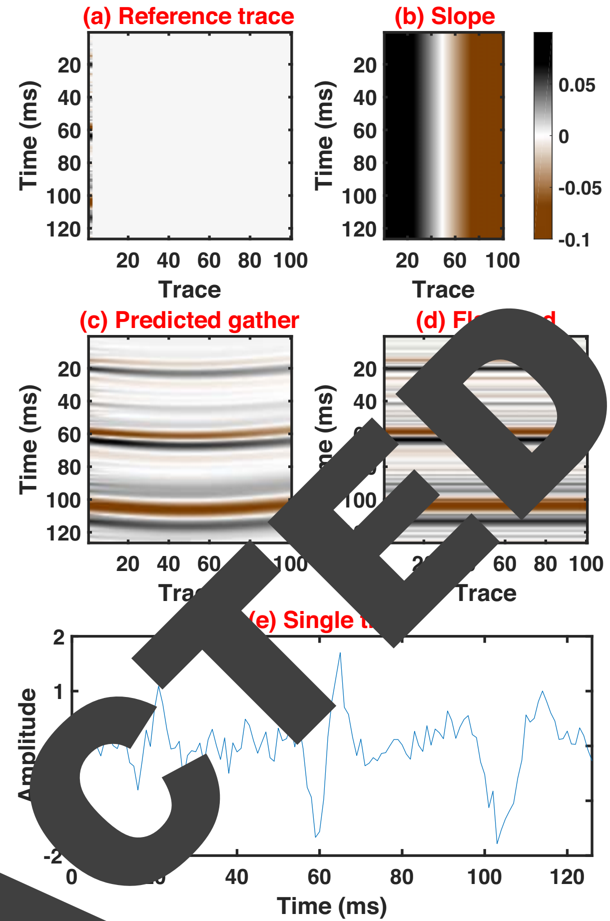


Figure 1. Trace prediction for clean data. (a) Reference trace. (b) Slope field. (c) Predicted gather from the reference trace. (d) Flatt[illegible]er by predicting the reference trace from each trace in Fig. (c[illegible]e in Fig. (a).



[illegible]re 2. Trace prediction for noisy data. (a) Reference trace. (b) Slope [illegible]eld. (c) Predicted gather from the reference trace. (d) Flattened gather by [illegible]dicting the reference trace from each trace in Fig. (c). (e) First trace in [illegible]ig. (a). Note that during trace prediction, the noise is preserved as coherent signal.

et al. (2016a). Regularized non-stationary deco[illegible]on (Ya[illegible] *et al.*, 2014; Wu *et al.* , 2016) is anot[illegible] decompo[illegible]ethod which is also based on a solid mathe[illegible]l model. Sin[illegible]alue decomposition (SVD) can also be [illegible]ct the most spatially coherent components from the m[illegible]dime[illegible] seismic data.

Sparse transform based [illegible]oaches ass[illegible]t multidimensional seismic data can [illegible] compressed in a [illegible]rse transformed domain, where the si[illegible]s re[illegible]nted by high-amplitude coefficients and the noise [illegible]d by [illegible]ll-amplitude coefficients (Gholami, 2013; Bai & [illegible]018). [illegible]ce, by transforming data to the spars[illegible] a so[illegible]ing can be applied to reject those sm[illegible]p[illegible] coeffic[illegible] that correspond to noise, which is fol[illegible] by a[illegible] transform from the thresholded coefficients [illegible] (Chen, 2017). This type of methods are closely co[illegible] with the compressive sensing paradigm (Lorenzi *et al.*, 2016). [illegible]used sparse transforms are Fourier transform, curvelet transfor[illegible] Candès *et al.*, 2006; Herrmann *et al.*, 2007; Herrmann & Hennenfent, 2008); Wang *et al.*, 2011; Zu *et al.*, 2017), seislet transform (Fomel & Liu, 2010; Chen & Fomel, 2015a; Gan *et al.*, 2015a; Gan *et al.*, 2015b), shearlet transform (Kong *et al.*, 2016), Radon transform (Foster & Mosher, 1992) and a variety of sparse wavelet transforms (Mousavi & Langston, 2016b; Anvari *et al.*, 2017), for example, synchrosqueezing (Daubechies *et al.*, 2011; Mousavi *et al.*, 2016; Mousavi & Langston, 2016a, 2017) or empirical wavelet transforms (Liu *et al.*, 2016b), etc. Recently,

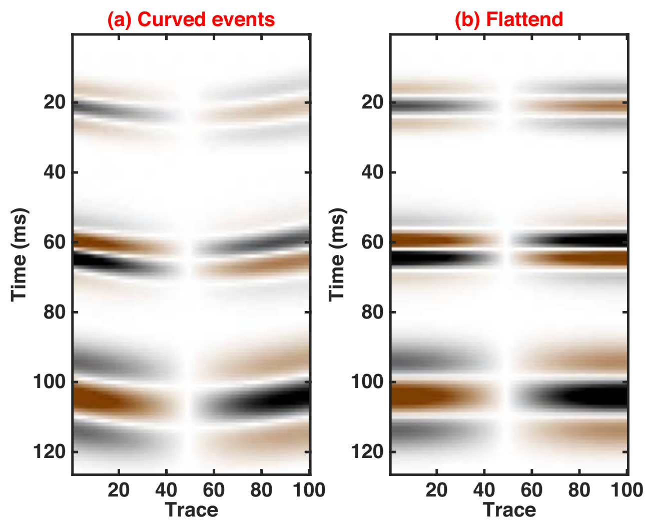


Figure 3. (a) Curved events. (b) Flattened events by predicting the first trace from each trace in (a). Note that during trace prediction, the amplitude is well preserved.

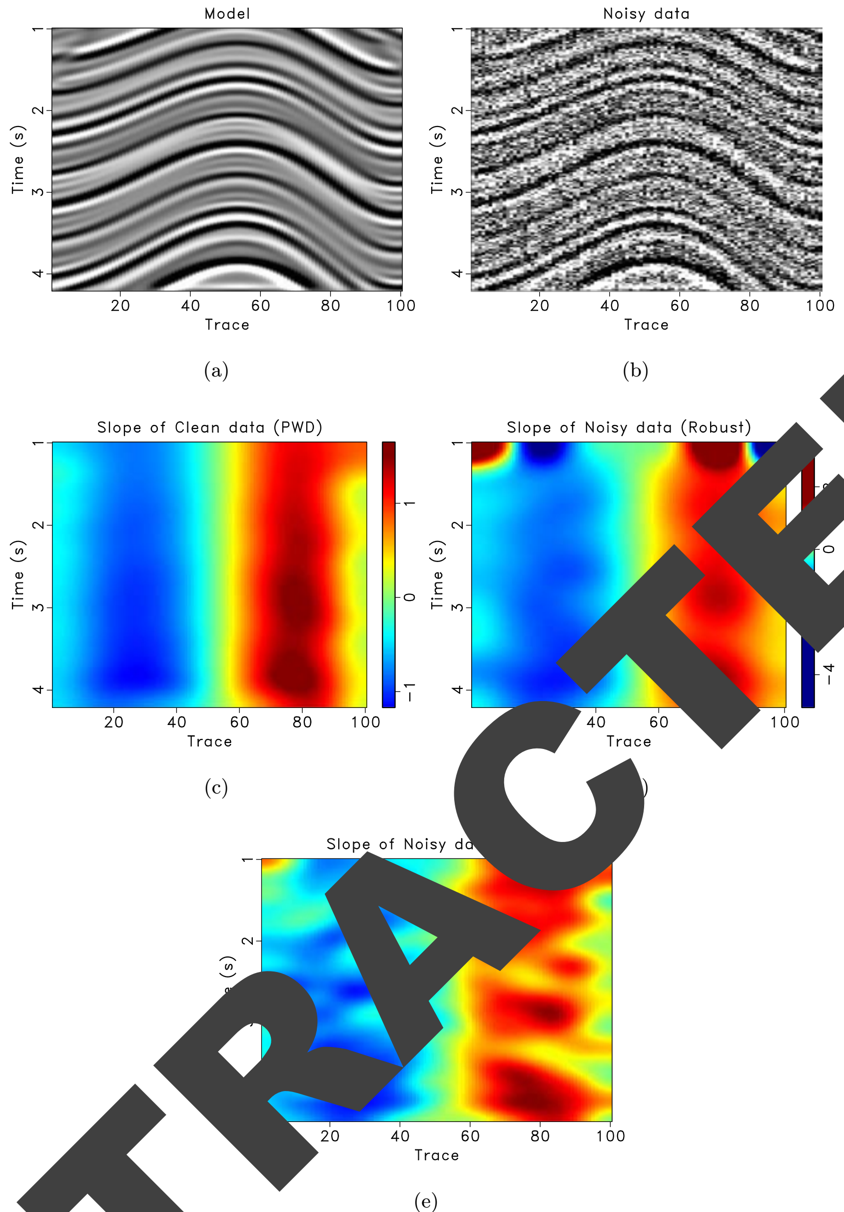


Figure 4. Slope calculati[illegible]t. (a) Clean da[illegible]oisy data. (c) Slope calculated from the clean data using PWD algorithm. (d) Slope calculated from the noisy data using the r[illegible]lope [illegible]lculation algo[illegible]hm. (e) Slope calculated from the noisy data using the PWD algorithm.

the adaptive dic[illegible]ning [illegible]gained a lot of attention in the seismic data pro[illegible] fiel[illegible]en, 2017; Wu & Bai, 2018b). The dic[illegible]earnin[illegible]parse representation differs from the t[illegible]ona[illegible]rse trans[illegible]s in that the basis functions for the s[illegible]ansf[illegible]daptively learned from the data itself, instead of [illegible]tional transforms.

Ran[illegible]tion methods are one of the most effective methods in the seism[illegible]rocessing community, which includes the Cadzow filtering (Tr[illegible]ett, 2008), singular spectrum analysis (SSA) (Vautard *et al.*, 1992; Wu & Bai, 2018a; Zhou *et al.*, 2018), damped SSA (Chen *et al.*, 2016b; Zhang *et al.*, 2016a,b), and multistep SSA (Zhang *et al.* , 2016c). There are two least-squares projection step in the damped SSA method. The first step can be considered as a rank reduction method while the second step can be interpreted as a compensation step for the non-optimal performance of the rank-reduction method, that is, the approximated signal subspace in the traditional rank-reduction framework is a mixture of both signal and noise subspaces. From a different aspect, Xue *et al.* (2016a) proposed a rank-increasing method for iteratively estimating the spike-like noise instead of estimating signals in deblending of simultaneous-source data (Zu *et al.* , 2016a,b; Bai & Wu, 2017; Zhou & Han, 2018; Wu & Bai, 2018d; Bai *et al.*, 2018a,b).

Mean and median filters utilize the statistical difference between signal and noise to reject the Gaussian white noise or impulsive noise (Liu *et al.*, 2009b; Liu, 2013). In addition to these classic noise attenuation methods, some advanced denoising methods have been proposed in recent years. Time-frequency peak filtering (Kahoo & Siahkoohi , 2009; Lin *et al.*, 2013, 2015) based approaches utilize the high-resolution property of time-frequency transform to distinguish between useful signals and random noise. Instead of developing a standalone denoising strategy, Chen & Fomel (2015b) proposed a two-step denoising approach that tries to solve a

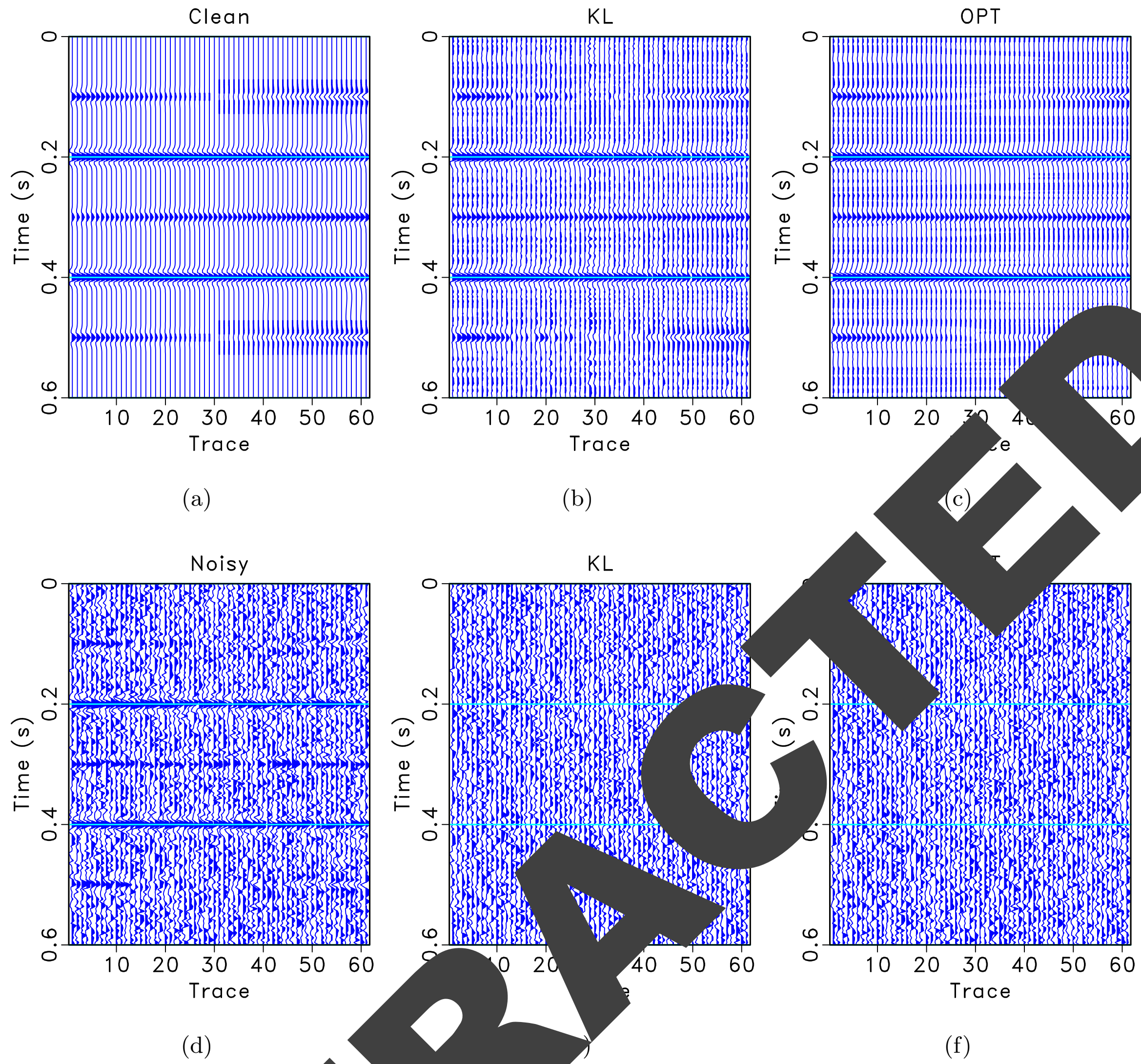


Figure 5. Synthetic example. (a) Clean data. (b) Denoised data using KL filtering. (c) Denoised data using the proposed method. (d) Noisy data. (e) Noise section corresponding to (b). (f) Noise section corresponding to (c).

long-existing problem in almost [illegible] [illegible] approaches: the signal leakage problem. By in[illegible]ing a new c[illegible] called local orthogonalization, Chen & [illegible] (2015b) succe[illegible]ly retrieved the coherent signals from [illegible] remo[illegible] noise section to guarantee no signal leakage in an[illegible] algorithms.

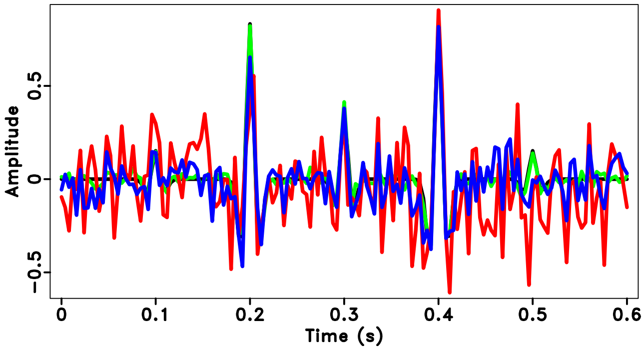
For all the aforemen[illegible] state-o[illegible]-art noise attenuation algorithms, no[illegible]em a[illegible] designed for preserving the strong am[illegible]ation d[illegible] seismic data. As we know, the amplit[illegible] variati[illegible] seismic data greatly affect the subsurface oil and ga[illegible]ction. Hence, the amplitude preserving capab[illegible] one of the backbone features we need to keep in mind when d[illegible]g a new denoising algorithm. In this paper, we are solving the s[illegible]s problem that is often neglected in traditional seismic data processing by proposing the plane-wave orthogonal polynomial transform (OPT) method. Here we want to clarify that the amplitude variation we mention here refers to strong amplitude variation, not simply the edge details or weak signals that are often mentioned in the literature. We first introduce the basic knowledge of the OPT, which is the key component that brings us the amplitude-preserving capability in the proposed framework. We then introduce the theory of plane-wave trace continuation that is used for flattening the seismic events without damaging the amplitude information. We show that both plane-wave trace continuation and OPT can well preserve the amplitude variation details in the seismic data, which accounts for the superb performance of preserving the amplitude details in the real data applications. Considering the strong influence of the slope estimation to the plane-wave flattening, we introduce a robust slope estimation method that can substitute the traditional plane-wave destruction (PWD; Fomel 2002) based methods in the presence of strong noise. A group of synthetic, prestack and post-stack field seismic data are used for demonstrating the performance of the proposed framework.

2 THEORY

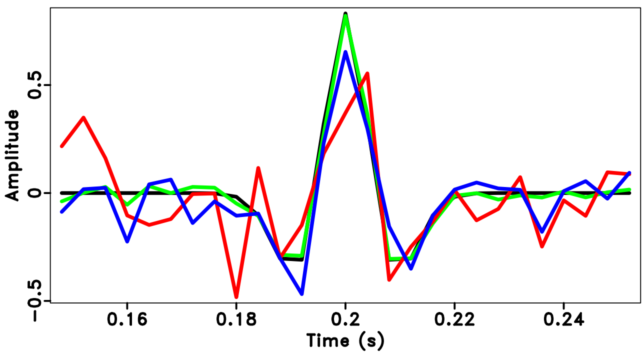
2.1 Orthogonal polynomial transform

In a seismic profile, the amplitude of time t and space x can be expressed as:

$$A(t, x) = \sum_{j=0}^{M-1} C_j(t) P_j(x), \tag{1}$$



(a)



(b)

Figure 6. Comparison of the 20th trace amplitude of each seismic gather in Fig. 5. The black line is from the clean data. The red line is fro[illegible] noisy data. The blue line corresponds to the KL method. The gre[illegible] corresponds to the proposed method. (a) Comparison of the whole [illegible] (b) Zoom-in comparison. Note that the black and green lines are very c[illegible] to each other, thus the reconstruction error using the pro[illegible] approach [illegible] much less than the traditional method for most parts.

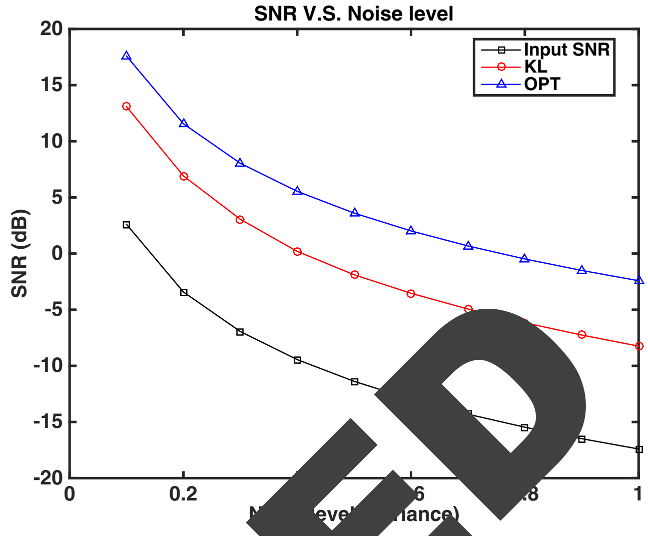
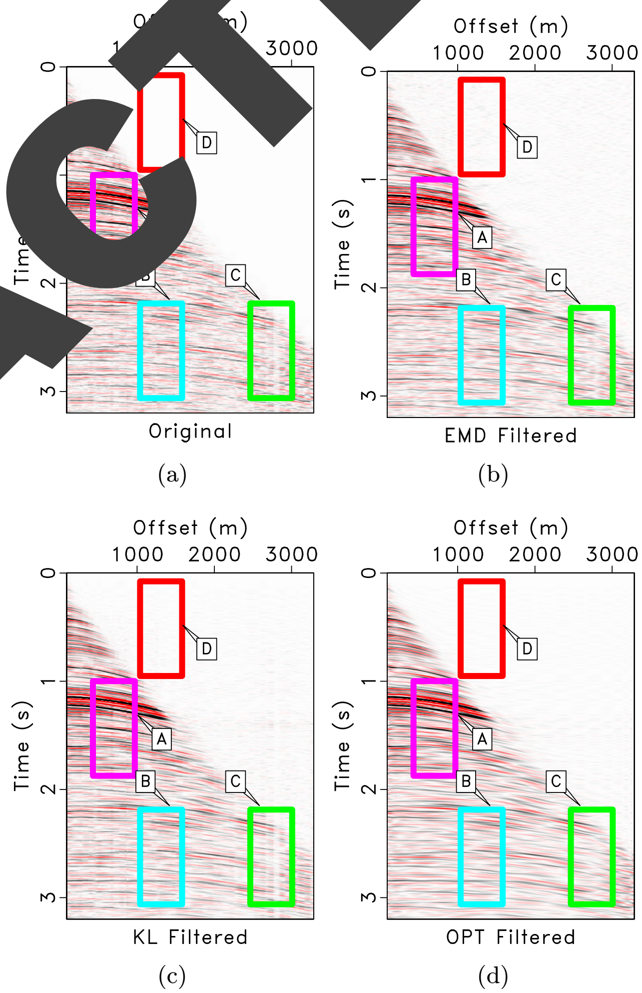


Figure 7. SNR diagra[illegible] synt[illegible] mpl[illegible]



(a) (b)

(c) (d)

Figure 8. Denoising comparison. (a) Raw noisy data. (b) Filtered using EMD method. (c) Filtered using KL method. (d) Filtered using the proposed method.

Table 1. Comparison of SNRs in dB for diffe[illegible] diagram corresponding to this table is show[illegible] in Fig[illegible]

Noise variance	Input data (dB)	KL (dB)	OPT (dB)
0.1	2.60	13.08	17.58
0.2	[illegible]	6.92	11.56
0.3	−6.94	[illegible]05	8.04
0.4	−9.44	[illegible]	5.54
0.5	−11[illegible]3	−1.89	3.60
0.6	[illegible]7	−3.54	2.02
0.7	−14.30	−4.94	0.68
0.8	−15.46	−6.17	−0.48
0.9	[illegible]	−7.26	−1.50
1.0	[illegible]0	−8.25	−2.42

where [illegible] $j = 0, 1, 2, , M-1\}$ is a set of orthogonal polynomials and M is [illegible]mber of basis functions and $\{C_j(t), j = 0, 1, 2, , M - 1\}$ is a s[illegible] coefficients. The $P_j(x)$ is a unit basis function that satisfies the condition:

$$P_j(x)P_i(x) = \delta_{ij}, \quad (2)$$

where δ_{ij} denotes the Kronecker delta. The spectrum defined by $C_j(t)$ denotes the energy distribution of the t–x domain data in the orthogonal polynomials transform domain. Besides, the low-order coefficients represent the effective energy and the high-order coefficients represent the random noise energy. We provide a detailed

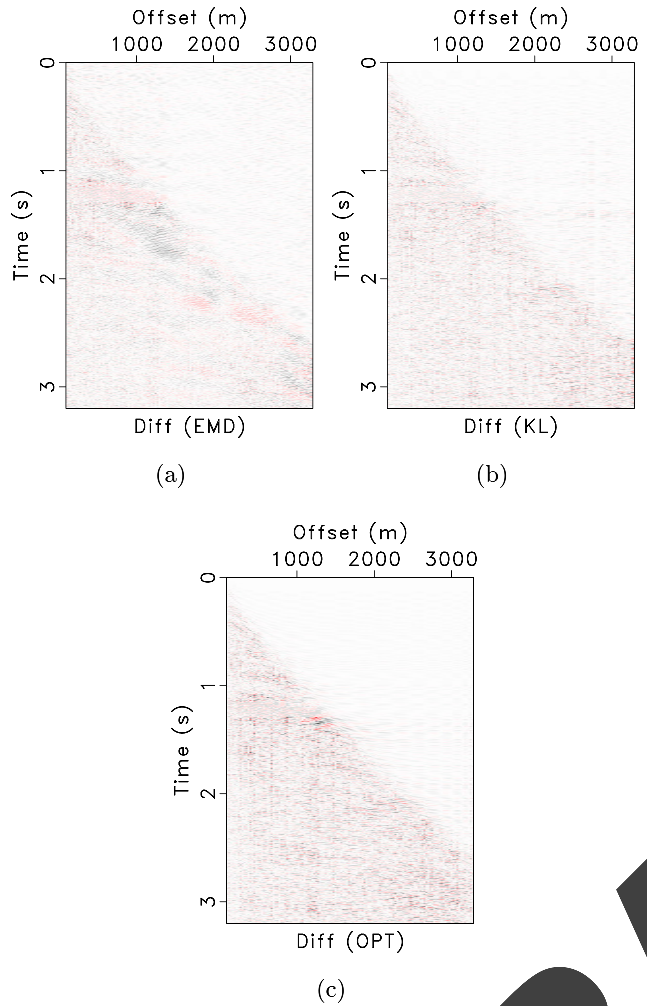
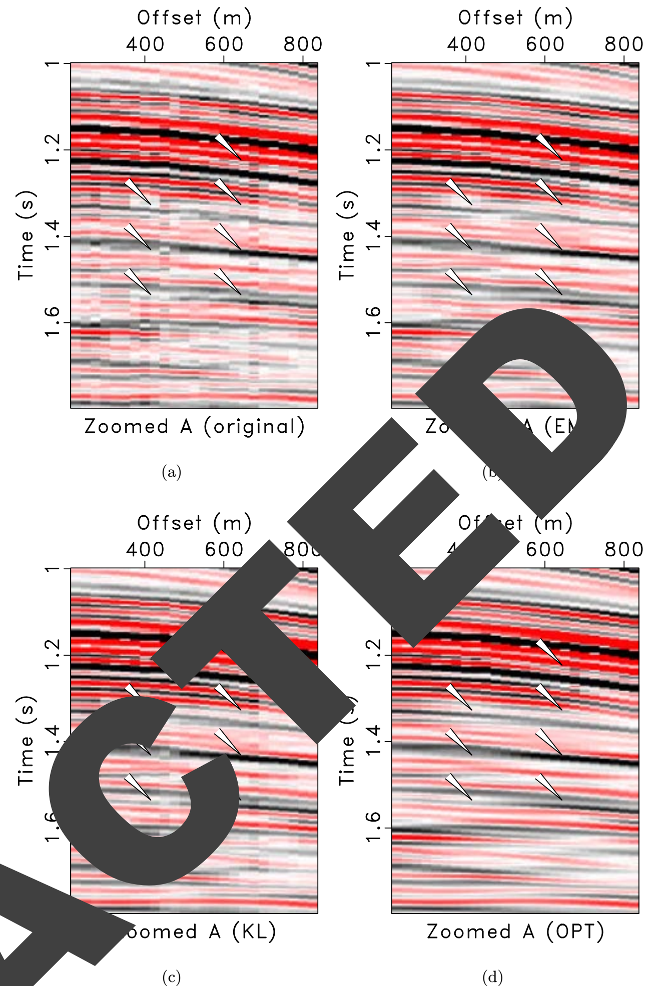


Figure 9. Noise comparison. (b) Removed noise usi[illegible] (b) Removed noise using KL method. (c) Removed no[illegible] method. In this case, the calculated RMSs for [illegible] (c) [illegible] 449.[illegible] and 506.26, respectively. Thus, the proposed [illegible] remo[illegible] per cent more noise than the KL method and 30.0 [illegible] more noise t[illegible] EMD method.



[illegible]**ure 10.** Zoomed frame box A from Fig. 8. (a) Zoomed noisy field data. (b) Zoomed filtered data using EMD method. (c) Zoomed filtered data using KL method. (d) Zoomed filtered data using the proposed method.

introduction about how we [illegible]nstruct the [illegible]nal polynomial basis function in Appendi[illegible]

In a matrix-multipli[illegible]n fo[illegible]q. (1) can be expressed as the following equation:

$$\mathbf{A} = \mathbf{C}\mathbf{P}, \tag{3}$$

where **A** i[illegible]d from [illegible], **C** is constructed from $C_j(t)$ and **P** is c[illegible]cted [illegible]. **A** is known and **P** can be constructed using [illegible]ppendix A. The unknown is **C**. **C** can be obtaine[illegible]verting the eq. (3),

$$\mathbf{C} = \mathbf{A}\mathbf{P}^H(\mathbf{P}\mathbf{P}\text{[illegible]} \tag{4}$$

where $[\cdot]^H$ denotes matrix transpose. In this paper, we choose $M = 20$, which indicates that we select 20 orthogonal polynomial basis function to represent the seismic data. Hence, inverting equation $\mathbf{P}\mathbf{P}^H$ is simply inverting a 20×20 matrix and is computationally efficient.

In the OPT method, we need to define the order of coefficients we want to preserve, the process of which corresponds to applying a mask operator to the orthogonal polynomial coefficients. Mask operator can be chosen to preserve low-order coefficients and reject high-order coefficients. It takes the following form:

$$\mathcal{M}_\tau(C_j(t)) = \begin{cases} C_j(t) & \text{for} \quad j \le \tau \\ 0 & \text{for} \quad j > \tau \end{cases}, \tag{5}$$

where $\mathcal{M}$ denotes the mask operator, $C_j(t)$ denotes the orthogonal polynomial coefficients at time t and order j. τ denotes a threshold for coefficients rejection. In this paper, we chose $\tau = 2$.

The coefficients after applying the mask operator 5 become

$$\hat{\mathbf{C}} = \mathcal{M}_\tau(\mathbf{C}). \tag{6}$$

The useful signals can be reconstructed by

$$\hat{\mathbf{A}} = \hat{\mathbf{C}}\mathbf{P}, \tag{7}$$

where $\hat{\mathbf{A}}$ denotes the denoised data.

2.2 Plane-wave trace continuation

In this section, we will derive a plane-wave flattening operator so that the seismic data can be flattened locally and the OPT can then be applied to the flattened gather.

The key question here is how to map the curved events into flattened events. We do the data mapping by a recursively predicting

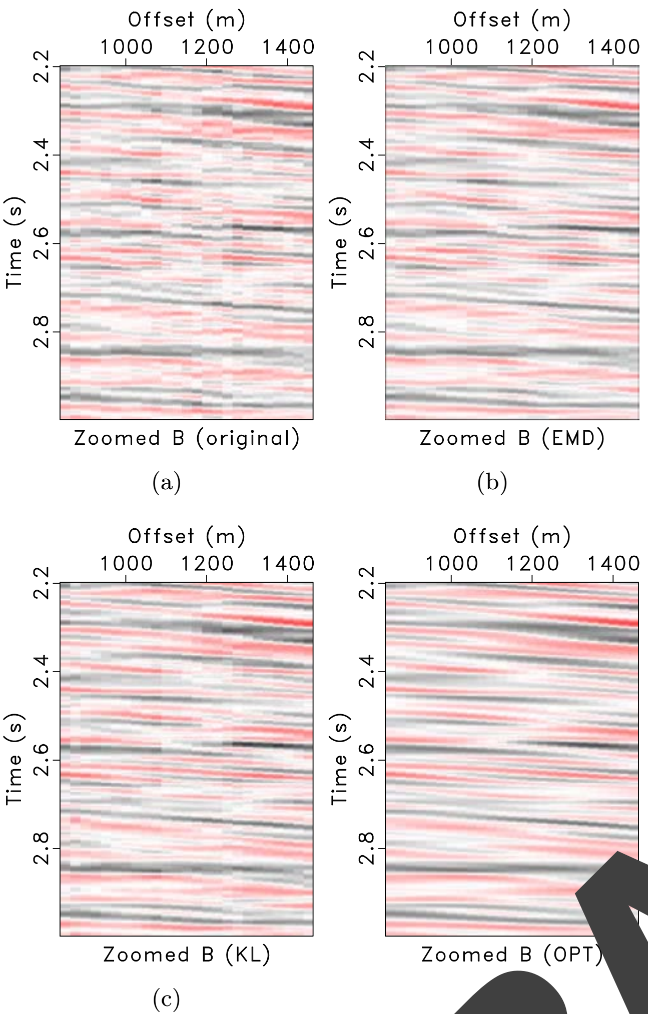


Figure 11. Zoomed frame box B from Fig. 8. (a) [illegible]ed n[illegible] field data. (b) Zoomed filtered data using EMD method. (c) [illegible] KL method. (d) Zoomed filtered data usin[illegible] prop[illegible]hod.

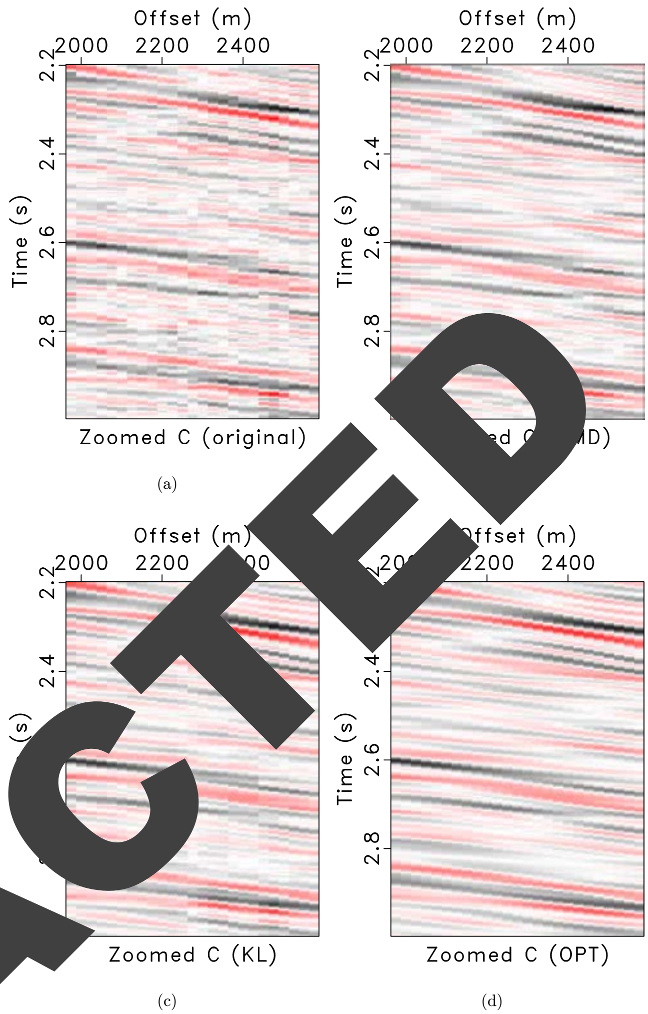


Figure 12. Zoomed frame box C from Fig. 8. (a) Zoomed noisy field data. (b) Zoomed filtered data using EMD method. (c) Zoomed filtered data using KL method. (d) Zoomed filtered data using the proposed method.

strategy. Each trace in a seism[illegible] the[illegible] predicted using neighbour traces. Given a ref[illegible]nce trace, e[illegible]e in the gather can predict the reference [illegible] in some ways, f[illegible]ample, by recursive trace continuation [illegible]lly, a[illegible]ging the predicted reference traces (from all other [illegible] in[illegible] gathe[illegible] constructs a flattened gather. Next, we will intr[illegible]e the[illegible] how we predict traces following th[illegible]e-wa[illegible] which we call plane-wave trace conti[illegible]

F[illegible] simpli[illegible] we always treat the first trace in the gather as the re[illegible] the gather, we need to predict the first trace [illegible]l other traces and arrange them together. In a brief mathem[illegible]ay, predicting the first trace from the jth ($j \neq 1$) trace can be [illegible]essed

$$\mathbf{d}_1 = \mathbf{P}_{2,1}\mathbf{P}_{3,2}\cdots\mathbf{P}_{j,j-1}\mathbf{d}_j, \tag{8}$$

where $\mathbf{P}_{p,q}$ denotes a prediction operator to predict trace $\mathbf{d}_q$ from trace $\mathbf{d}_p$. Specifically, $\mathbf{P}_{p,p-1}$ denotes the prediction between two traces from right to left. In the inverse process, the jth trace can be predicted in a similar recursive formula from left to right:

$$\mathbf{d}_j = \mathbf{P}_{j-1,j}\cdots\mathbf{P}_{2,3}\mathbf{P}_{1,2}\mathbf{d}_1. \tag{9}$$

Predicting from the jth trace to $(j+1)$th trace (or from the $(j+1)$th trace to jth trace) requires solving the plane-wave equation:

$$\frac{\partial u}{\partial x} + \sigma\frac{\partial u}{\partial t} = 0, \tag{10}$$

where $u(t, x)$ is the seismic record and σ is local slope. In the case of the constant local slope, eq. (10) has the following solution:

$$u(t, x) = f(t - \sigma), \tag{11}$$

where f is the waveform function. In the variable-slope case, we can solve eq. (10) by discretizing it. Let u_p^v denote $u(v\Delta t, p\Delta x)$, and then we obtain:

$$\frac{u_{p+1}^{v+1} - u_p^{v+1}}{2\Delta x} + \frac{u_{p+1}^{v} - u_p^{v}}{2\Delta x} + \sigma_p^v\left(\frac{u_{p+1}^{v+1} - u_{p+1}^{v}}{2\Delta t} + \frac{u_p^{v+1} - u_p^{v}}{2\Delta t}\right) = 0. \tag{12}$$

Rearranging the terms in eq. (12), we get

$$\left(\frac{1}{\Delta x} + \frac{\sigma_p^v}{\Delta t}\right)u_{p+1}^{v+1} + \left(-\frac{1}{\Delta x} + \frac{\sigma_p^v}{\Delta t}\right)u_p^{v+1} + \left(\frac{1}{\Delta x} - \frac{\sigma_p^v}{\Delta t}\right)u_{p+1}^{v} + \left(-\frac{1}{\Delta x} - \frac{\sigma_p^v}{\Delta t}\right)u_p^{v} = 0. \tag{13}$$

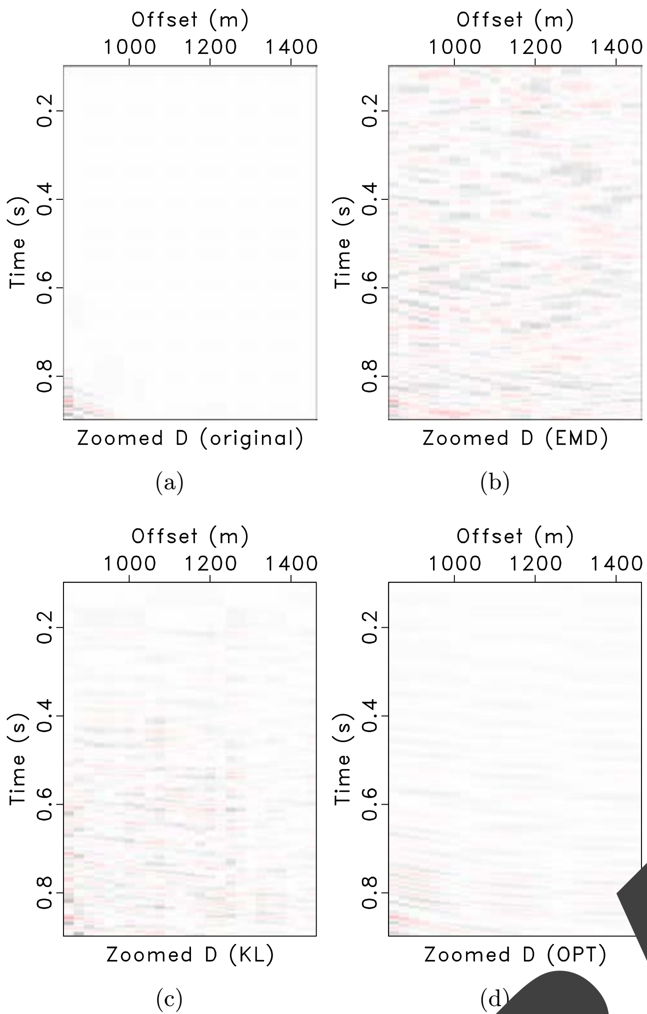
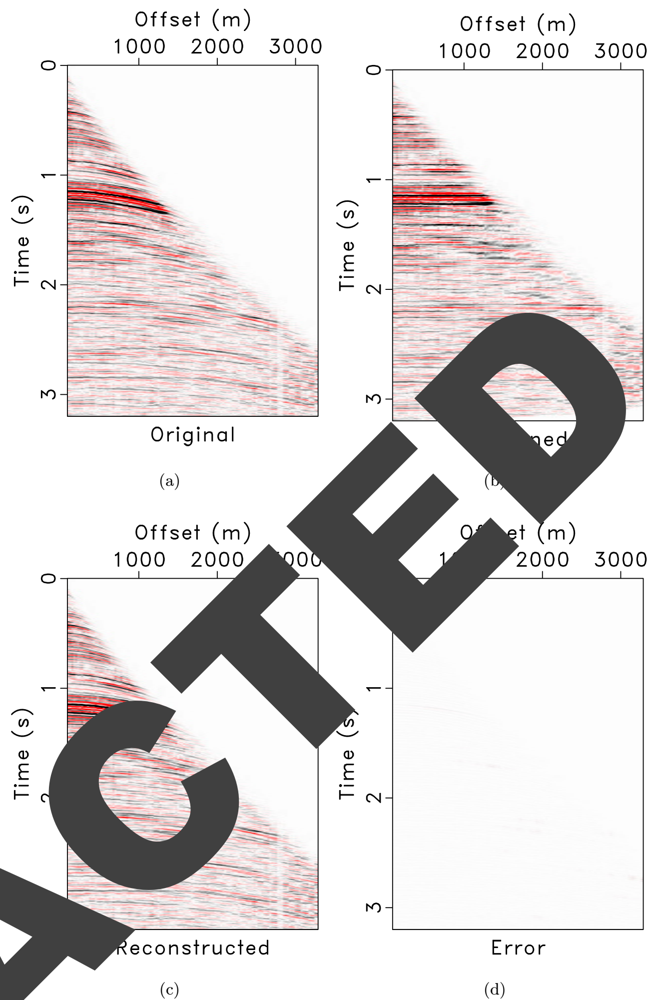


Figure 13. Zoomed frame box D from Fig. 8. (a) Zo[illegible] noisy [illegible] (b) Zoomed filtered data using EMD method. (c) Zoom[illegible] KL method. (d) Zoomed filtered data using the [illegible]osed [illegible]



[illegible]**gure 14.** Pre-stack field data example. (a) Field data. (b) Flattened field data. (c) Reconstructed field data. (d) Reconstruction error.

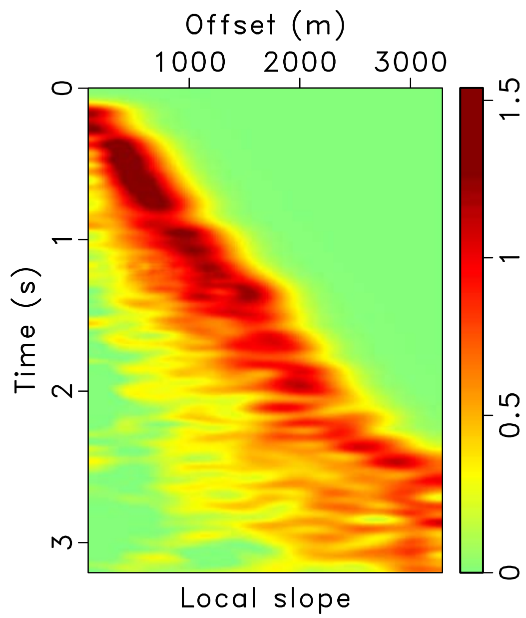


Figure 15. Local slope estimation of the pre-stack field data.

Then we have the following poin[illegible]t recursion [illegible] u_p^v to u_{p+1}^{v+1}:

$$u_{p+1}^{v+1} = \left(\frac{1}{\Delta x} + \frac{\sigma_p^v}{\Delta t}\right)^{-1} \left[\left([illegible] \frac{}{\Delta x} - \frac{\sigma_p^v}{\Delta t}\right) u_p^{v+1} + [illegible] \left(-\frac{1}{\Delta x} + \frac{\sigma_p^v}{\Delta t}\right) u_{p+1}^{v} [illegible] \left([illegible] \frac{\sigma_p^v}{\Delta t}\right) [illegible] \right]. \tag{14}$$

Eq. (14) ca[illegible] for th[illegible]ion from the first trace to all other tra[illegible]lar way, [illegible]other traces can predict the first trace [illegible]nver[illegible]ntinuation process. Since the trace continuatio[illegible] using the local plane-wave assumption of seismic [illegible] call the relation in eq. (14) the plane-wave trace continuation [illegible]

Fig. 1 shows a[illegible]xample of the trace prediction process. We start from the first trace in a gather, and predict each trace in the gather from the first trace, following a given local slope field. Fig. 1(a) shows the initial status of the trace prediction process, where only the first trace is shown. Following the slope field shown in Fig. 1(b), we can predict a complete gather from the first trace using the plane-wave equation. The complete gather is shown in Fig. 1(c). It is clear that the morphology of the predicted gather is consistent with the local slope shown in Fig. 1(b). Fig. 1(d) shows a flattened gather

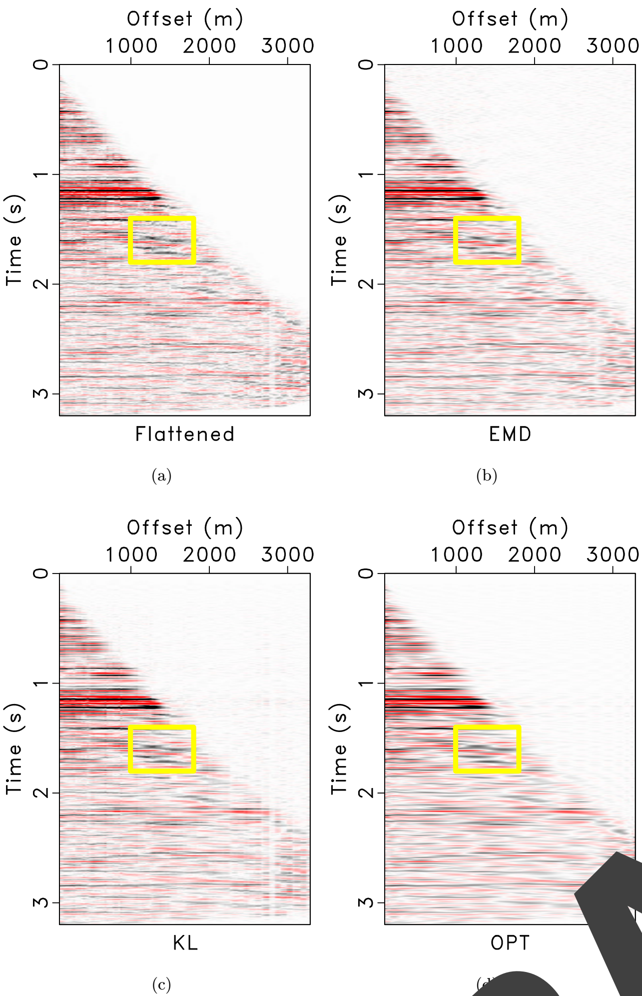


Figure 16. Denoising comparison in the flattened d[illegible]sion. [illegible]aw noisy data. (b) Filtered using EMD method. (b) Filt[illegible]ng K[illegible](c) Filtered using the proposed method.

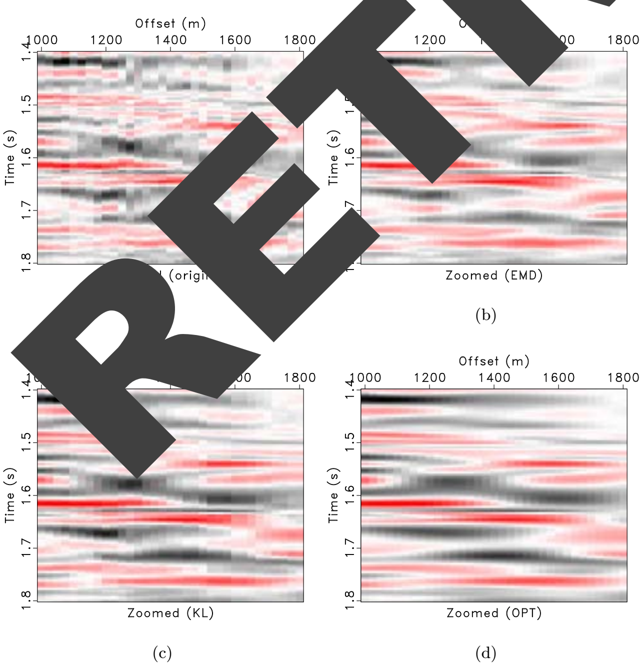


Figure 17. Zoomed sections from Fig. 16. (a) Zoomed noisy field data. (b) Zoomed filtered data using EMD method. (c) Zoomed filtered data using KL method. (d) Zoomed filtered data using the proposed method.

from the curved events shown in Fig. 1(c). We flatten the events by predicting the first trace from each trace shown in Fig. 1(c). For a clear view of the first trace, we plot it in Fig. 1(e).

We then show an example in the presence of random noise. The noisy data is simulated with $\mathbf{d} = \mathbf{s} + \mathbf{n}$, where $\mathbf{s}$ is signal, that is, the solution to a wave equation in some random medium. The signal has a certain mean and variance, and a certain spatial and temporal correlation structure. $\mathbf{n}$ is the noise, for which, presumably the expectation $\langle n \rangle$ is zero. The overall spatial variance of the noise is a certain number, and its covariance with the signal is zero. The noise is distributed in the 2-D plane following a Gaussian rule and has no spatial correlation. Fig. 2(a) shows the nois[illegible] trace. The details of the noisy trace are shown in Fi[illegible]). Fo[illegible] a given slope field shown in 2(b), we predict [illegible]lete gat[illegible]m the first trace, and show the gather in Fig. 2([illegible] be s[illegible]hat the noise is preserved during the pre[illegible]n proce[illegible]2[illegible]hows the flattened gather from Fig. 2(c[illegible] can [illegible]nclud[illegible]his test that trace prediction can preser[illegible] co[illegible]ent in the starting trace.

We then use another ex[illegible]how t[illegible]mplitude preserving feature of the fl[illegible]ing [illegible] Fi[illegible]a) shows a complete gather containing [illegible] curved ev[illegible]mplitude variation along the spatial dir[illegible]g. 3(b) show[illegible]e flattened gather from the curved eve[illegible]n see that the amplitude variation is well preserved du[illegible]ing the t[illegible]diction process.

[illegible] Robust sl[illegible] estimation

A[illegible]r impor[illegible] factor in plane-wave OPT is the local slope calcu[illegible]racy of the slope estimation affects performance of the [illegible]ng operation and the following OPT. In this part, we [illegible]ll introduce a robust slope estimation method that is based on the [illegible] transform (Liu *et al.*, 2015).

R[illegible]arranging eq. (10) we get

$$\sigma = -\frac{\frac{\partial u}{\partial x}}{\frac{\partial u}{\partial t}}. \tag{15}$$

Eq. (15) can be further derived such that

$$\sigma = -\frac{\frac{\partial u}{\partial x}}{\frac{\partial u}{\partial z}} = -\frac{F_x^{-1}[H_{\mathrm{DX}}[F_x u]]}{F_t^{-1}[H_{\mathrm{DT}}[F_t u]]}, \tag{16}$$

where H_{DX} is frequency response function of the partial derivative in the x direction, and H_{DT} is frequency response function of the partial derivative in the t direction. F_x and F_t denote the Fourier transform along the x and t directions, respectively. It can be straightforwardly derived that

$$\sigma = -\frac{\mathcal{H}_x(u)}{\mathcal{H}_t(u)}, \tag{17}$$

where $\mathcal{H}_x(u)$ denotes the Hilbert transform of u along x direction and $\mathcal{H}_t(u)$ denotes the Hilbert transform of u along t direction.

Fig. 4 shows a slope calculation test. We calculate the slope from the noisy data using the traditional PWD method and the robust slope calculation method, respectively. As a comparison, an accurate slope estimation from the clean data using the PWD algorithm is used to evaluate the robustness of different slope estimation approaches in the case of noise. Fig. 4(a) shows the clean data, and Fig. 4(c) shows the slope estimated from the clean data using the PWD algorithm, which is deemed to be the accurate slope. Fig. 4(b) shows the noisy data by adding some Gaussian white noise. Fig. 4(d) shows the slope calculated using the robust slope estimation. It is salient that the slope estimated from the noisy data is fairly close to

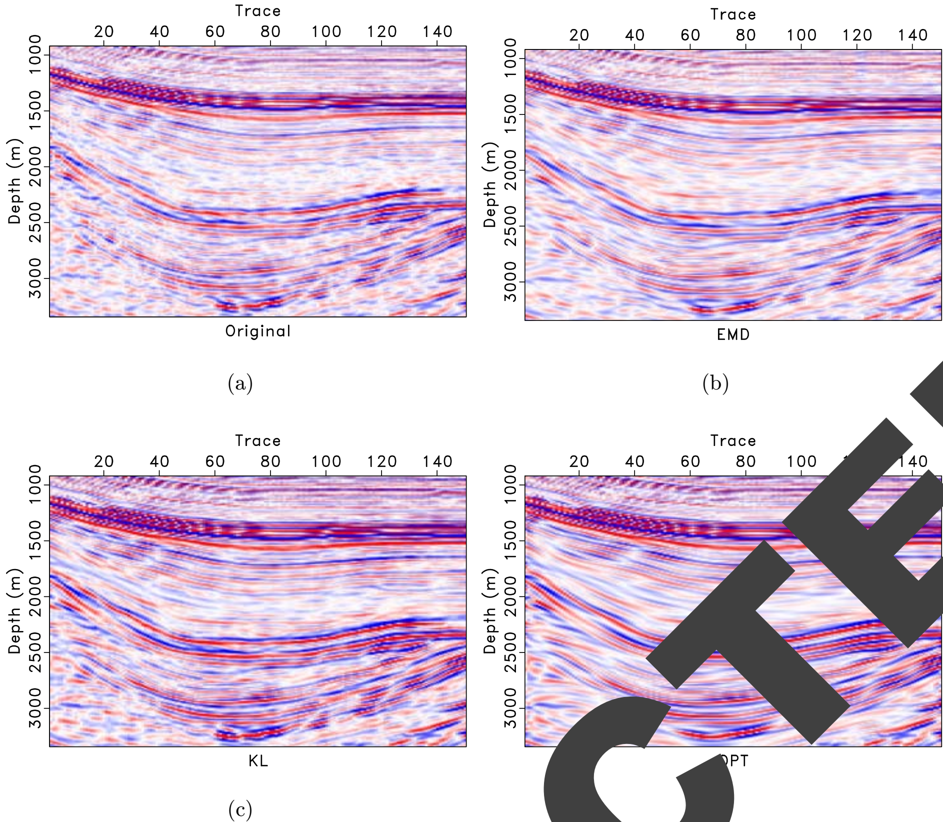


Figure 18. First post-stack field data example. (a) Field data. (b) Filtered d[illegible]sing EMD [illegible] [illegible]iltered data using KL method. (d) Filtered data using the proposed method.

the accurate slope field. However, using the traditional PWD algorithm, it is difficult to obtain an acceptable slope estimati[illegible]rom the noisy data, as can be seen from the result shown in F[illegible] [illegible]rom this test, we conclude that the robust slope estim[illegible] can [illegible]sed to obtain robust slope estimation performance [illegible] the[illegible] of strong random noise.

It is worth mentioning that, by eqs [illegible] and ([illegible] [illegible] do not consider the spatial gradient of amp[illegible] In the case [illegible]ooth spatial amplitude change (e.g. s[illegible] gradient), t[illegible] slope estimation method also works, si[illegible] the [illegible]ion is done locally and the small spatial gradi[illegible]lmost has n[illegible]nce. However, in the case of sharp spa[illegible]mplitude change[illegible]g. large spatial gradient), the method[illegible]not b[illegible]opted. This drawback can be hopefully overcome [illegible] work. [illegible] addition, the problem of spatial gradients of am[illegible] and [illegible]plications for non-plane wave solutio[illegible]entio[illegible]andt (1993).

2.4 Plane-[illegible]rthogonal polynomial transform

We have introdu[illegible] in detail the theory of plane-wave trace continuation, that is, how we predict an arbitrary trace in a seismic gather from a random starting trace. We have shown that by discretizing the plane-wave equation, we can derive the spatial trace continuation relation, which can be used for trace prediction. We have presented that during trace continuation, the amplitude of seismic waveforms can be well preserved. Regarding the slope estimation, which is an important factor in the plane-wave trace continuation operator, we introduce the robust slope estimation approach. We al[illegible] [illegible]hat the robust slope estimation approach can obtain ro[illegible] slope estimation in the presence of strong noise. Considering [illegible]e amplitude-preserving capability of the OPT in a flattened dimen[illegible]on, we can cascade the plane-wave trace continuation operator and the OPT together to obtain a twofolds amplitude-preserving performance during a complete workflow. Thus, we name the cascaded framework as the plane-wave OPT. The complete framework for noise attenuation using the plane-wave OPT is shown in Algorithm 1.

Algorithm 1 Plane-wave OPT

Input: Noisy data $\mathbf{D}$. Order of coefficients to be preserved τ.
Output: Denoised data $\mathcal{D}$.

1: **Forward plane-wave flattening**: $\hat{\mathbf{D}}$=PWF($\mathbf{D}$)
2: **Forward OPT**: $\mathbf{C}$ = OPT($\hat{\mathbf{D}}$)
3: **Mask**: $\hat{\mathbf{C}} = \mathcal{M}(\mathbf{C}, \tau)$
4: **Inverse OPT**: $\hat{\mathcal{D}}$=IOPT($\hat{\mathbf{C}}$)
5: **Inverse plane-wave flattening**: $\mathcal{D}$=IPWF($\hat{\mathcal{D}}$)

The forward OPT corresponds to inverting $\mathbf{P}^H(\mathbf{P}\mathbf{P}^H)^{-1}$. The inverse OPT corresponds to multiplying the orthogonal polynomial coefficients by $\mathbf{P}$. In Algorithm 1, the detailed implementations of the forward plane-wave flattening operator and the inverse plane-wave flattening operator are shown in Algorithms 2 and 3, respectively.

In Algorithms 2 and 3, note that N denotes the number of spatial traces. 1 and -1 in the operator PWTC() denote predicting from a trace to the first trace and predicting the first trace to another trace,

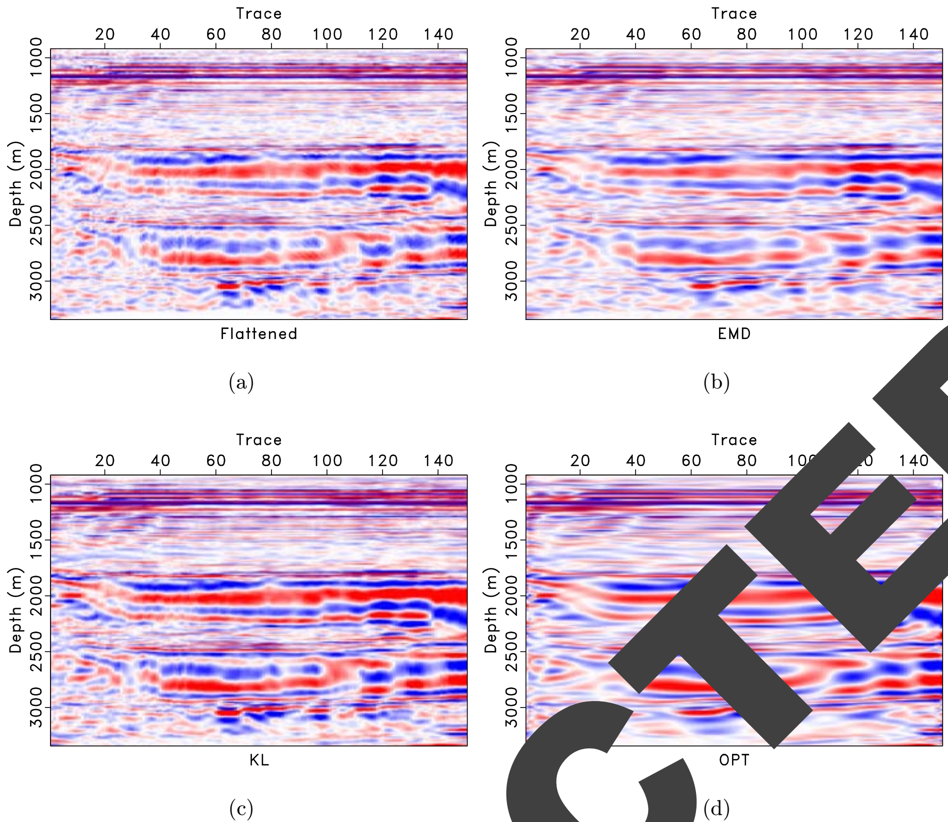


Figure 19. Comparison in the flattened domain. (a) Field data. (b) Filte[illegible]d data using [illegible] (c) Filtered data using KL method. (d) Filtered data using the proposed method.

Algorithm 2 Plane-wave flattening

Input: Matrix containing curved events **D**.
Output: Matrix containing flattened events $\hat{\mathbf{D}}$.
Setting the first trace as the reference trac[illegible]
for $n = 1, \cdots, N$ **do**
  Predict the first trace: $\hat{\mathbf{D}}(n)$=[illegible]TC([illegible]
end for

Algorithm 3 Inverse plane-w[illegible] flatt[illegible]

Input: Matrix includin[illegible]tened events [illegible]
Output: Matrix in[illegible]g cu[illegible]ed events $\hat{\mathbf{D}}$.
Setting the first [illegible] **as t**[illegible]**eference trace.**
for $n = 1, \cdots$ [illegible]
  Pr[illegible] **the** [illegible]**ce**[illegible])=PWTC($\mathbf{D}(n)$, −1)
en[illegible]

resp[illegible] $\mathbf{D}(i)$ [illegible](i) denote the ith column (or trace) in the matrix [illegible] $\hat{\mathbf{D}}$.

3 EXAMPLES

The first example is a synthetic example, as shown in Fig. 5. We apply the Karhunen–Loève (KL) filtering method (Jones & Levy, 1987) and the proposed method to a flattened data set with strong amplitude variation. Fig. 5(a) shows the clean data and Fig. 5(d) shows the noisy data. Figs 5(b) and (c) show the denoised data using the KL filtering method and the proposed method, respectively. [illegible]e) and (f) show the removed random noise using two approaches. We can observe clearly from Figs 5(b) and (c) that the KL filtering causes significant damages to the events, while the proposed method preserves the amplitude-variation details successfully.

In order to compare the amplitude between different seismic profiles in detail, we compare the amplitude for a single trace from each section shown in Fig. 5. The trace is chosen as the 20th trace in each section of Fig. 5. The comparison is presented in Fig. 6. A zoom-in comparison is shown in Fig. 6(b). The black line is from the clean data. The red line is from the noisy data. The blue line corresponds to the KL method. The green line corresponds to the proposed method. It is apparent that the green line is very close to the black line while the blue line deviates from the black line too much in most areas. This trace amplitude comparison further confirms the superior performance of the proposed algorithm.

In order to numerically compare the denoising performance, we use the commonly used signal-to-noise ratio (SNR) defined as follows to quantitatively measure the performance (Chen & Fomel, 2015b):

$$\mathrm{SNR} = 10\log_{10}\frac{\|\mathbf{s}\|_2^2}{\|\mathbf{s}-\hat{\mathbf{s}}\|_2^2}. \quad (18)$$

where $\mathbf{s}$ denotes the noise-free data and $\hat{\mathbf{s}}$ denotes the denoised data. In addition, to quantitatively measure the noise removal in the case of no discernable signal damage, we define the metric as the root-mean-square (RMS),

$$\mathrm{RMS} = \|\mathbf{n}\|_2, \quad (19)$$

where $\mathbf{n}$ denotes the removed noise. Although the amplitude range varies a lot for different data sets, the RMS metric provides us a

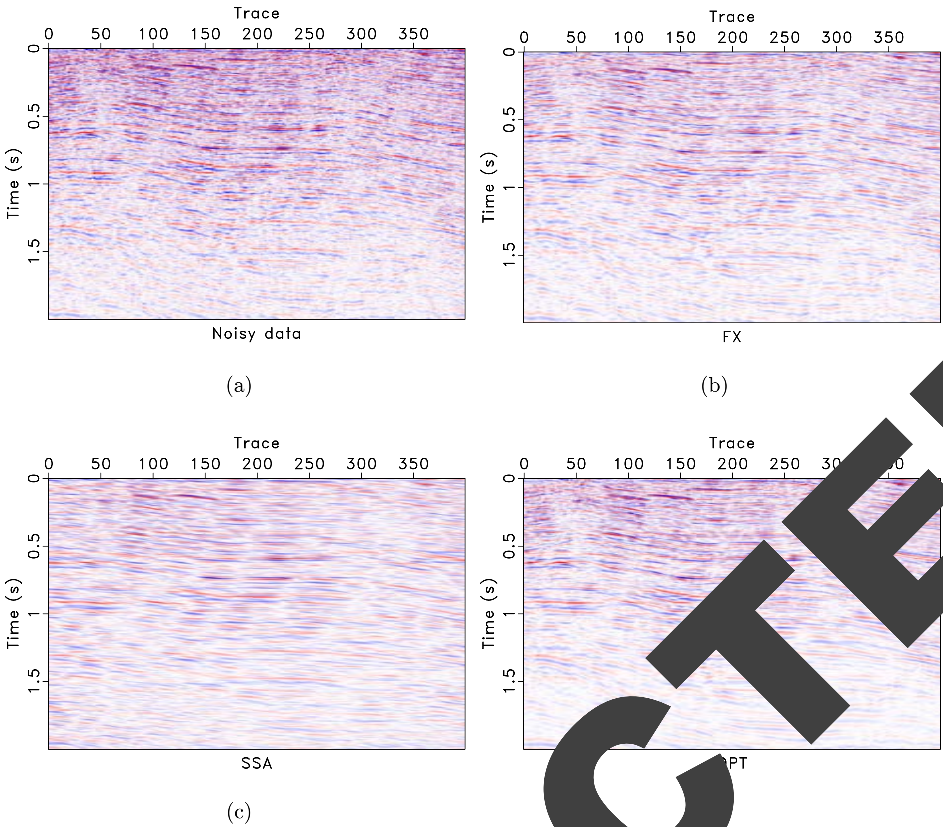


Figure 20. Denoising comparison for the second post-stack field data. (a) T econd post- ata. (b) Filtered data using *f–x* predictive filtering. (c) Filtered data using SSA. (d) Filtered data using the proposed method.

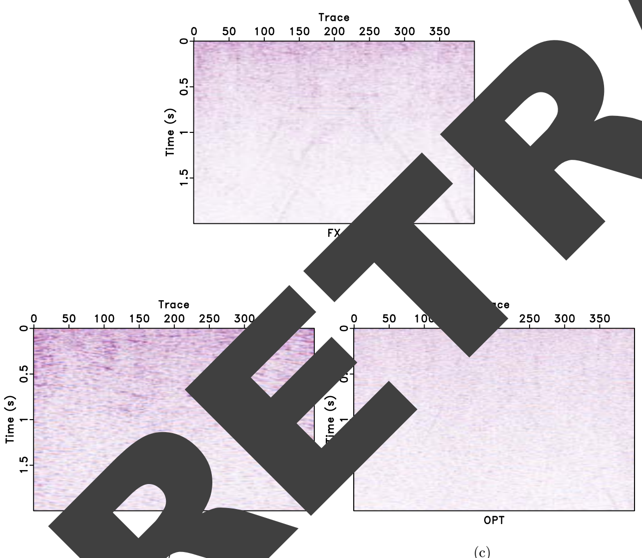


Figure 21. N parison for the second post-stack field data. (a) Removed noise usi predictive filtering. (b) Removed noise using SSA. (c) Removed noise using the proposed method.

quantitative way to evaluate the noise removal performance for one specific data set among different denoising methods, for example, how much better method A performs than method B.

In order to compare the performance of two methods in different noise level, we increase the variance of noise from 0.1 to 1.0, and calculate the SNRs of denoised data of both methods and show th able 1. To see the varied SNRs more vividly, we plot data from Table 1 in Fig. 7. The black line shows the SNRs arying with input noise variances. The red line shows the SNRs orresponding to the KL method. The blue line shows the SNRs of the OPT method. It is obvious that both methods obtain large SNR improvement for all noise levels and the SNRs of the OPT method are always higher than the KL method. We can also observe clearly that the difference between the proposed OPT method and the KL method increases as noise variance becomes larger, which indicates that the proposed method outperforms the KL method more when the seismic data becomes noisier.

For computational cost comparison, the KL method takes 0.62 s for processing the data shown in Fig. 5(d) while the proposed algorithm takes 0.01 s. The data contains 151 samples and 61 traces. The computation is done on a PC station equipped with an Intel Core i7 CPU clocked at 3.1 GHz and 16 GB of RAM. Note that both KL and OPT methods require the events to be flattened in order to obtain the best performance, thus we only compare the cost difference in the filtering stage.

The second example is a pre-stack field data example. Fig. 8(a) shows the original data. Figs 8(b)–(d) show the denoised data using EMD method, KL method and the proposed method, respectively. Fig. 9 shows the removed noise sections using three approaches. Fig. 9(a) shows that some low-frequency energy is damaged while Fig. 9(c) shows that the removed noise is stronger. In this example, the calculated RMSs for Figs 9(a)–(c) are 389.49, 449.22 and 506.26, respectively. Thus, the proposed method removes 12.7 per cent more noise than the KL method and 30.0 per cent more noise than the EMD method.

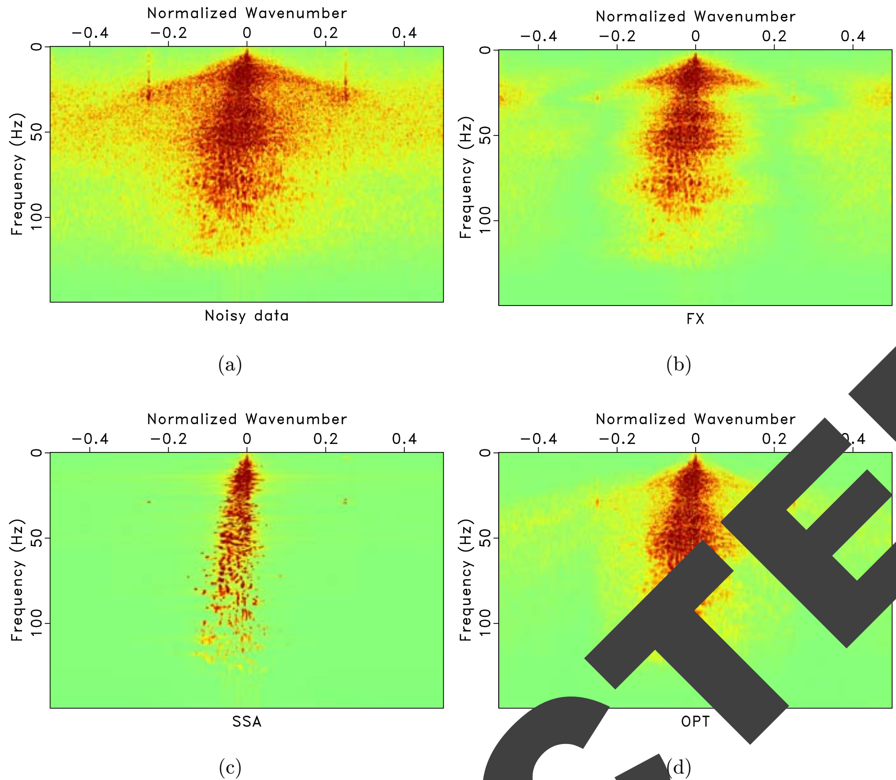


Figure 22. Spectra comparison. (a) Spectrum of the second post-stack seismic data. (b) Spectrum using f–x predictive filtering. (c) Spectrum using SSA. (d) Spectrum using the proposed method.

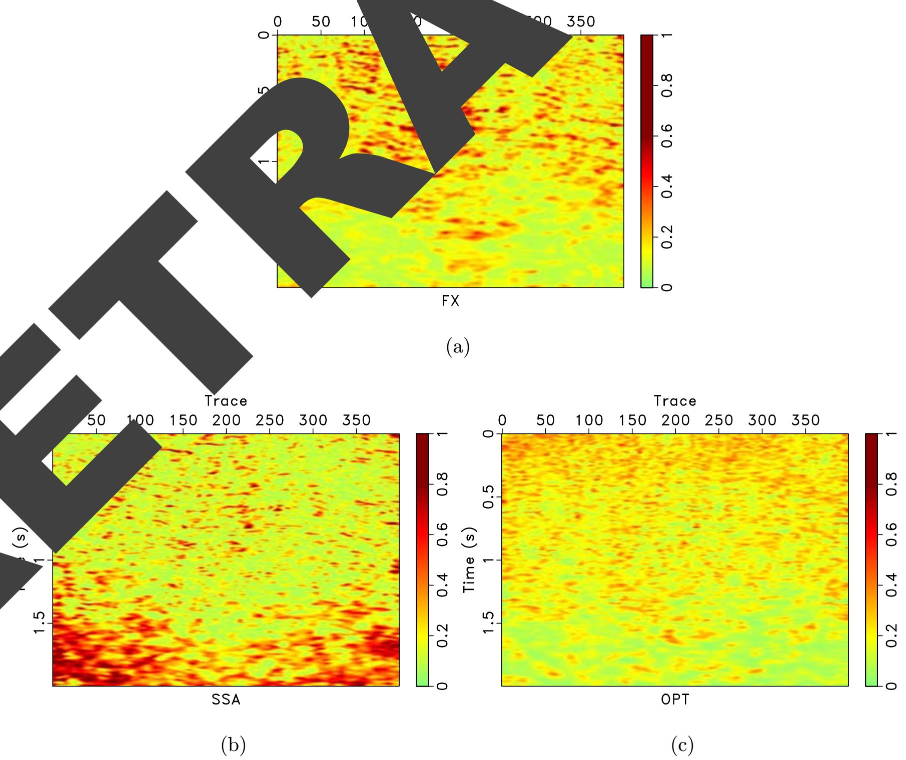


Figure 23. Comparison of local similarity between denoised data and removed noise. (a) Local similarity using f–x predictive filtering. (b) Local similarity using SSA. (c) Local similarity using the proposed method.

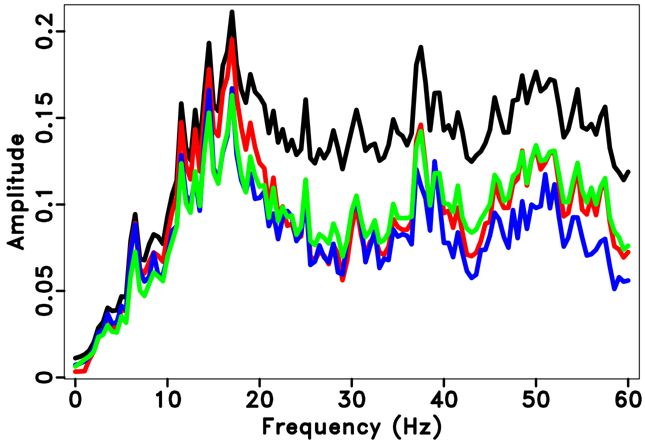


Figure 24. Comparisons of the average spectrum of all the traces. The black line denotes the average spectrum of raw data. The green line corresponds to the proposed approach. The red line corresponds to *f–x* predictive filtering method. The blue line corresponds to the SSA method.

In order to comprehensively compare three different approaches, we zoomed four frame boxes (A, B, C and D) to show the detailed difference. Fig. 10 shows the comparison from frame box A. It is obvious that the KL approach causes some residual noise while EMD and OPT approaches obtain good results, more careful observation can show that the proposed method can obtain a more coherent image. Fig. 11 shows the comparison for frame box B. It is obvious that OPT method obtains the cleanest result. Fig. 12 shows the comparison for frame box C. It is still obvious that OPT method can make the events more coherent and more importantly, preser[illegible] the amplitude-variation-with-offsets details well. Fig. 13 shows th[illegible] comparison for frame box D. Both Figs 13(b) and (c) show obvious amplitude artefacts while the OPT result in Fig. 13(d) s[illegible] nearly zero amplitude in the zoomed section.

For this example, we also demonstrate the pla[illegible]ave f[illegible]ning process in Fig. 14. Fig. 14(b) shows the flatte[illegible] he original data shown in Fig. 14(a) (or Fig.8[illegible]). It [illegible]hat the events have been flattened well. Fig.14[illegible]ws the [illegible]ructed data from the inverse plane-wave fla[illegible] The data is [illegible] the same as the original data. Fig. 14[illegible] difference between the reconstructed data and the o[illegible]nal [illegible] error section is almost zero everywhere, w[illegible] demonstrate[illegible] the plane-wave flattening process does [illegible]trod[illegible]e extra erro[illegible]ig. 15 shows the slope estimated from [illegible]igin[illegible]a. We also show a detailed comparison between diff[illegible] the fla[illegible]ed dimension in Fig. 16. A zoomed comparison [illegible]he f[illegible]ed gathers after filtering is shown in Fi[illegible]ere we[illegible]ude that the proposed method obtains t[illegible]mo[illegible]t result [illegible] best preserving the reflection ampli[illegible]

The [illegible]l post-stack seismic image shown in Fig. 18(a)[illegible] are 140 spatial traces and 194 temporal samples. The se[illegible]age contains highly curved events and the amplitude along [illegible]e events is not continuous, which will make the seismic interpretation difficult. After using three approaches, the EMD method, the KL method and the proposed method, the denoised images are shown in Figs 18(b)–(d), respectively. It is obvious that the proposed OPT based filtering approach can obtain a well smoothed seismic image with the continuity and the amplitude of events enhanced greatly. The EMD based approach, however, cannot effectively smooth the seismic events, and still leaves a lot of discontinuity in the image. The KL filtering approach obtains a much better filtering performance compared with the EMD based approach, however, it is not as successful as the performance of the proposed OPT based approach.

We can find the mechanism that caused the tremendous difference of denoised images from the comparison in the flattened domain, as shown in Fig. 19. It is even more obvious that the OPT based approach obtains a nearly perfect smoothing along the flattened images (equivalent to along the structure in the original domain). The EMD based approach can achieve some smoothing, but remains less continuous than both KL and OPT based approaches. In this example, the calculated RMSs of removed noise are 0.015 for EMD method, 0.016 for KL method and 0.019 for the [illegible]ethod. Thus, the proposed method removes 18.7 per[illegible]hore[illegible]han the KL method and 25.0 per cent more nois[illegible]he EMD[illegible]od.

The next field data example is shown in [illegible](a), w[illegible] is also a post-stack data and contains [illegible]seismic[illegible]ents. Figs 20(b)–(d) show the denoised [illegible]s us[illegible] three [illegible]t methods. For this example, we furt[illegible]mp[illegible]e performance of the proposed method with that of [illegible]edicti[illegible]tering method and the SSA method ([illegible]rd *e*[illegible]992[illegible]g. 21 shows the corresponding noise [illegible]ns. For th[illegible]e, it seems that all three methods ob[illegible] improved re[illegible]s and the performance of different me[illegible]s [illegible]milar. In order to compare the performance in detail and more [illegible]e plot the *F–K* spectra of different denoised[illegible]. The *F–K* sp[illegible]m of the raw data is shown in Fig. 2[illegible] The *F–K* spectra corresponding to different methods are show[illegible] Figs 22(b[illegible]d). Comparing the *F–K* spectra of different meth[illegible]nd *F–K* s[illegible]um of the raw data, it is easy to find that both *f–x* p[illegible]e filt[illegible] method and the proposed method preserve the usef[illegible]ell, but the *f–x* predictive filtering method has [illegible]e residual spectrum energy around the edges (large wavenum[illegible]ents). SSA method causes significant damages to useful s[illegible]

[illegible]n this example, we also calculate the local similarity between the [illegible]enoised data and removed noise for different methods. The local [illegible]milarity is an effective way to detect the lost signals in the removed noise. High local similarity indicates that in the noise section, there are significantly similar components as the useful signals, that is, there is lost energy in the noise. The calculation of local similarity is provided in Appendix B. The local similarity maps for different methods are shown in Fig. 23, where we can clearly observe the high similarity anomalies in the *f–x* predictive filtering and SSA results. Although there are also some similarity anomalies in the result from the proposed method, the similarity value is relatively lower than the other two methods. From this test we conclude that the proposed method causes less damage to useful energy.

We also plot a comparison of the average spectrum of all the traces for different data in Fig. 24. The green line corresponds to the proposed approach. The red line corresponds to *f–x* predictive filtering method. The blue line corresponds to the SSA method. It is quite obvious that the energy preservation of the proposed method in signal frequency band (20–60 Hz) is quite successful. The proposed method mitigates more high-frequency noise than $f-x$ predictive filtering method, which confirms the observation from Fig. 22. We admit that the high-frequency noise of the proposed method is slightly more than the SSA method. However, the proposed method preserves more useful energy than the other two methods in the spectrum. This field data further confirms the superior performance of the presented algorithm.

In this example, to compare the noise removal performance, we need to make sure the removed noise sections do not contain discernable signal energy, as required by the metric defined in eq. (19),

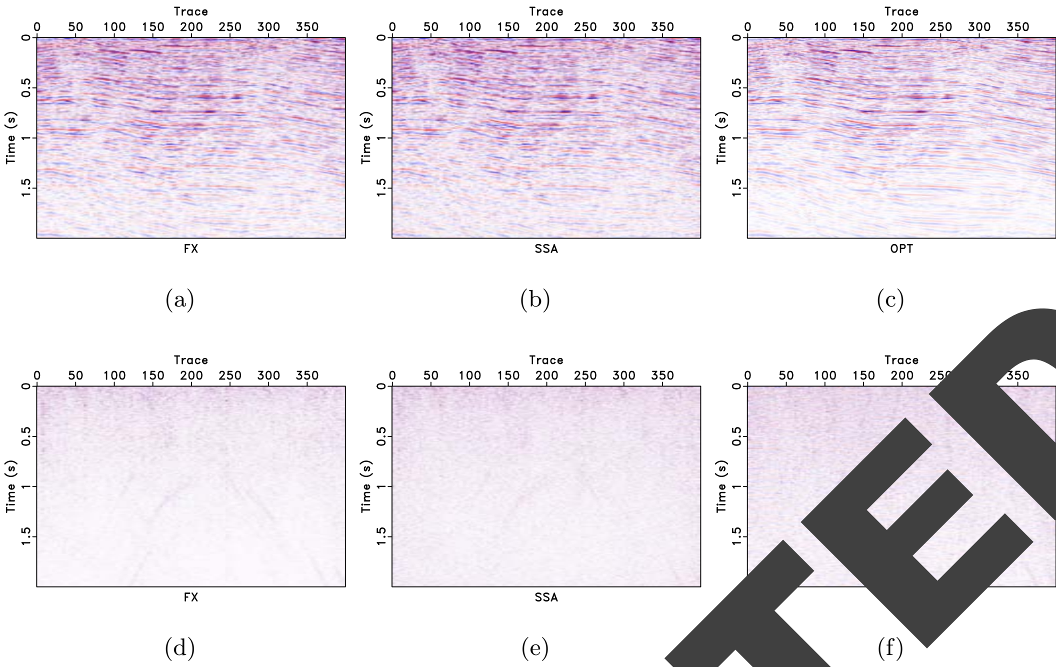


Figure 25. Comparison for the second post-stack field data after adjusting the parameters. (a) Filtered data us[illegible]redictive filtering. (b) Filtered data using SSA. (c) Filtered data using the proposed method. (d) Removed noise using *f*–*x* predi[illegible]ring. (e) Remove[illegible]oise using SSA. (f) Removed noise using the proposed method. In this case, the calculated RMSs for (d)–(f) are 0.059, 0.071 [illegible].086, respectively. Thus, the proposed method removes 21.1 per cent more noise than the SSA method and 38.0 per cent more noise than the $f-x$ predi[illegible]iltering m[illegible]

and have to adjust the parameters for the *f*–*x* and SSA methods. The denoised data and the removed noise sections using the adju[illegible] parameters are shown in Fig. 25. In this case, the calculated [illegible] for Figs 25(d)–(f) are 0.059, 0.071 and 0.086, respectively. Thu[illegible] proposed method removes 21.1 per cent more noise than the S[illegible] method and 38.0 per cent more noise than the *f*–*x* [illegible]

4 CONCLUSIONS

The OPT can be used to effectivel[illegible]arate sp[illegible]correlative signals and spatially incoherent [illegible]ithout losing[illegible]orm amplitude. To create a flattene[illegible] where the OPT can be optimally applied, we der[illegible]plane-[illegible]ce continuation relation for flattening the c[illegible] seismic eve[illegible]e prediction in the flattening process a[illegible]e sub[illegible]equent OPT [illegible]e both demonstrated to be amplitude [illegible]vin[illegible]e robust slope estimation approach can obtain more [illegible]orman[illegible]an the state-of-the-art PWD method in the prese[illegible]stro[illegible]ise. The proposed framework has bee[illegible]to sev[illegible]tic, field pre-stack and post-stack seism[illegible]ata a[illegible]s shown t[illegible]ter preserve the amplitude variations th[illegible]er al[illegible]thods.

ACKN[illegible]EDGEMENTS

This work was supported by the Starting fund at Zhejiang University, National Natural Science Foundation of China (No. 61401307), Hebei Province Foundation of Returned Overseas Scholars (CL201707), Hebei Province Project of Science and Technology R & D (11213565) and Hebei Province Natural Science Foundation (No. E2016202341). The authors would like to thank Yaru Xue, Shaohuan Zu and Wei Chen for helpful comments and suggestions on the topic of noise attenuation.

R[illegible]E[illegible]ES

Abm[illegible]rbout, J., 1995. Lateral prediction for noise attenuation by $t-x$ and $f-x$ techniques, *Geophysics,* **60,** 1887–1896.

[illegible]R., Siahsar, M.A.N., Gholtashi, S., Kahoo, A.R. & Mohammadi, M., [illegible]. Seismic random noise attenuation using synchrosqueezed wavelet transform and low-rank signal matrix approximation, *IEEE Trans. Geosci. Remote Sens.,* **55**(11), 6574–6581.

Asgedom, E.G., Orji, O.C. & Soellner, W., 2017. Rough-sea deghosting of single-sensor seismic data using the knowledge of the sea surface shape, *J. Seismic Explor.,* **26,** 105–123.

Bai, M. & Wu, J., 2017. Efficient deblending using median filtering without correct normal moveout - with comparison on migrated images, *J. Seismic Explor.,* **26,** 455–479.

Bai, M. & Wu, J., 2018. Seismic deconvolution using iteartive transform-domain sparse inversion, *J. Seismic Explor.,* **27**(2), 103–116.

Bai, M., Wu, J., Xie, J. & Zhang, D., 2018a. Least-squares reverse time migration of blended data with low-rank constraint along structural direction, *J. Seismic Explor.,* **27**(1), 29–48.

Bai, M., Wu, J., Zu, S. & Chen, W., 2018b. A structural rank reduction operator for removing artifacts in least-squares reverse time migration, *Comput. Geosci.,* **117,** 9–20.

Canales, L., 1984. Random noise reduction, *54th Annual International Meeting,* SEG, Expanded Abstracts, Atlanta, GA, pp. 525–527.

Candès, E.J., Demanet, L., Donoho, D.L. & Ying, L., 2006. Fast discrete curvelet transforms, *SIAM Multiscale Model. Simul.,* **5,** 861–899.

Chen, Y., 2016. Dip-separated structural filtering using seislet thresholding and adaptive empirical mode decomposition based dip filter, *Geophys. J. Int.,* **206,** 457–469.

Chen, Y., 2017. Fast dictionary learning for noise attenuation of multidimensional seismic data, *Geophys. J. Int.,* **209**(1), 21–31.

Chen, Y., 2018. Automatic microseismic event picking via unsupervised machine learning, *Geophys. J. Int.,* **212**(1), 88–102.

Chen, Y. & Fomel, S., 2015a. EMD-seislet transform, in *85th Annual International Meeting,* SEG, Expanded Abstracts, New Orleans, LA, pp. 4775–4778.

Chen, Y. & Fomel, S., 2015b. Random noise attenuation using local signal-and-noise orthogonalization, *Geophysics,* **80,** WD1–WD9.

Chen, W., Chen, Y., Xie, J., Zu, S. & Zhang, Y., 2016a. Multiples attenuation using trace randomization and empirical mode decomposition, in *86th Annual International Meeting,* SEG, Expanded Abstracts, Dallas, TX, pp. 4498–4502.

Chen, Y., Zhang, D., Jin, Z., Chen, X., Zu, S., Huang, W. & Gan, S., 2016b. Simultaneous denoising and reconstruction of 5-D seismic data via damped rank-reduction method, *Geophys. J. Int.,* **206**(3), 1695–1717.

Colominas, M.A., Schlotthauer, G., Torres, M.E. & Flandrin, P., 2012. Noise-assisted EMD methods in action, *Adv. Adapt. Data Anal.,* **4,** 1250025 .

Daubechies, I., Lu, J. & Wu, H.-T., 2011. Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool, *Appl. Comput. Harmon. Anal.,* **30,** 243–261.

Dragomiretskiy, K. & Zosso, D., 2014. Variational mode decomposition, *IEEE Trans. Signal Process.,* **62,** 531–544.

Fomel, S., 2002. Application of plane-wave destruction filters, *Geophysics,* **67,** 1946–1960.

Fomel, S. & Liu, Y., 2010. Seislet transform and seislet frame, *Geophysics,* **75,** V25–V38.

Foster, D.J. & Mosher, C.C., 1992. Suppression of multiple reflections using the radon transform, *Geophysics,* **57,** 386–395.

Gan, S., Wang, S., Chen, Y. & Chen, X., 2015a. Deblending of distance separated simultaneous-source data using seislet frames in the shot domain, in *85th Annual International Meeting,* SEG Expanded Abstracts, New Orleans, LA, pp. 65–70.

Gan, S., Wang, S., Chen, Y. & Chen, X., 2015b. Seismic data reconstruction via fast projection onto convex sets in the seislet transform domain, in *85th Annual international meeting,* SEG Expanded Abstracts, New Orleans, LA, pp. 3814–3819.

Gan, S., Wang, S., Chen, Y., Qu, S. & Zu, S., 2016. Velocity analysis of simultaneous-source data using high-resolution semblance-coping with the strong noise, *Geophys. J. Int.,* **204,** 768–779.

Gao, Z., Pan, Z. & Gao, J., 2016. Multimutation differential evolution algorithm and its application to seismic inversion, *IEEE Trans. Geosci. Remote Sens.,* **54**(6), 3626–3636.

Gholami, A., 2013. Sparse time-frequency decomposition and some appl[illegible]cations, *IEEE Trans. Geosci. Remote Sens.,* **51,** 3598–3604.

Herrmann, F.J., Böniger, U. & Verschuur, D.J., 2007. Non-lin[illegible] primary-multiple separation with directional curvelet frames, *Geo*[illegible] **170,** 781–799.

Herrmann, F.J. & Hennenfent, G., 2008. Non-param[illegible] seism[illegible]covery with curvelet frames, *Geophys. J. Int.,* **173,** [illegible]

Huang, N.E. *et al.*, 1998. The empirical mode de[illegible]positi[illegible]e Hilbert spectrum for nonlinear and non-station[illegible]e series [illegible] *Phil. Trans. R. Soc. Lond., A.,* **454,** 903–995.

Huang, W., Wang, R., Zu, S. & Chen, [illegible]frequency noise attenuation in seismic and microseismic d[illegible] using [illegible]atical morphological filtering, *Geophys. J. Int.,* **21**[illegible]18–1340.

Huang, Z., Zhang, J., Zha[illegible] & Sun, Y., 2016. [illegible]chrosqueezing S-transform and its app[illegible]n in [illegible]ic spectral decomposition, *IEEE Trans. Geosci. Rem*[illegible] **5**[illegible]817–8[illegible]5.

Jones, I.F. & Levy, S., 198[illegible]o-nois[illegible] enhancement in multichannel seismic d[illegible] the K[illegible] transform, *Geophys. Prospect.,* **35,** 12–32.

Kahoo, A[illegible] Sia[illegible]i, T.R., [illegible]. Random noise suppression from seis[illegible]ta us[illegible]quency peak filtering, in *71st Annual International* [illegible] *Exhibition,* EAGE, Extended Abstracts, doi:10.3[illegible]-4609.201400214.

Kong, D., Peng[illegible] Y. & Fan, H., 2016. Seismic random noise attenuation using directiona[illegible] variation in the shearlet domain, *J. Seismic Explor.,* **25**(4), 321–338.

Lin, H., Li, Y., Yang, B. & Ma, T., 2013. Random denoising and signal nonlinearity approach by time-frequency peak filtering using weighted frequency reassignment, *Geophysics,* **78,** V229–V237.

Lin, H., Li, Y., Ma, H., Yang, B. & Dai, J., 2015. Matching-pursuit-based spatial-trace time-frequency peak filtering for seismic random noise attenuation, *IEEE Geosci. Remote Sens. Lett.,* **12,** 394–398.

Liu, C., Chen, C., Wang, D., Liu, Y., Wang, S. & Zhang, L., 2015. Seismic dip estimation based on the two-dimensional hilbert transform and its application in random noise attenuation, *Appl. Geophys.* **12,** 55–63.

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

Liu, G., Fomel, S., Jin, L. & Chen, X., 2009a. Stacking seismic data using local correlation, *Geophysics,* **74,** V43–V48.

Liu, G., Chen, X., Du, J. & Song, J., 2011. Seismic noise attenuation using nonstationary polynomial fitting, *Appl. Geophys.,* **8,** 18–26.

Liu, G., Chen, X., Du, J. & Wu, K., 2012. Random noise attenuation using *f-x* regularized nonstationary autoregression, *Geophysics*, **77,** V61–V69.

Liu, W., Cao, S. & Chen, Y., 2016a. Applications [illegible] mode decomposition in seismic time-frequency analys[illegible]1(5), V365–V378.

Liu, W., Cao, S. & Chen, Y., 2016b. Seismic ti[illegible]ncy ana[illegible] via empirical wavelet transform, *IEEE Geos*[illegible]*emote*[illegible]*tt.,* **1**[illegible]–32.

Liu, W., Cao, S. & Wang, Z., 2017. A[illegible]ation of va[illegible]de decomposition to seismic random [illegible] reduc[illegible] *J. Ge*[illegible] *Eng.,* **14,** .

Liu, Y., 2013. Noise reductio[illegible] by [illegible]an filt[illegible] *Geophysics,* **78,** V79–V87.

Liu, Y., Liu, C. & Wan[illegible] 2009b. A [illegible]ying median filter for seismic random, s[illegible]oise eliminati[illegible]*eophysics,* **74,** V17–V24.

Lorenzi, L., Mel[illegible]r, G., 2016. Missing-area reconstruction in multispectral [illegible]ages u[illegible]ompressive sensing perspective, *IEEE Trans. G*[illegible] *Remote Sens.,* [illegible]998–4008.

Mousav[illegible] Langston, C.A., 2[illegible]6a. Adaptive noise estimation and suppre[illegible] for improvin[illegible] microseismic event detection, *J. Appl. Geophys.,* **13**[illegible]6–124.

Mou[illegible]M. & La[illegible]n, C.A., 2016b. Hybrid seismic denoising using high[illegible] sta[illegible] and improved wavelet block thresholding, *Bull. seism.* [illegible]6(4), 1380–1393.

[illegible]savi, S.M. & Langston, C.A., 2017. Automatic noise-removal/signal-[illegible] based on general cross-validation thresholding in syn[illegible]zed domain and its application on earthquake data, *Geophysics,* [illegible]2(4), V211–V227.

[illegible]ousavi, S.M., Langston, C.A. & Horton, S.P., 2016. Automatic microseis[illegible]mic denoising and onset detection using the synchrosqueezed continuous wavelet transform, *Geophysics* **81**(4), V341–V355.

Qu, S., Verschuur, D. & Chen, Y., 2016. Full waveform inversion using an automatic directional total variation constraint, in *78th Annual International Conference and Exhibition,* EAGE, Extended Abstracts, doi:10.3997/2214–4609.201701340.

Trickett, S., 2008. F-xy cadzow noise suppression, in *CSPG CSEG CWLS Convention,* Las Vegas, NV, pp. 303–306.

Vautard, R., Yiou, P. & Ghil, M., 1992. Singular-spectrum analysis: a toolkit for short, noisy chaotic signals,*Physica D Nonlinear Phenom.,* **58,** 95–126.

Wang, B., Wu, R., Chen, X. & Li, J., 2015. Simultaneous seismic data interpolation and denoising with a new adaptive method based on dreamlet transform, *Geophys. J. Int.,* **201,** 1180–1192.

Wang, Y., Cao, J. & Yang, C., 2011. Recovery of seismic wavefields based on compressive sensing by an l1-norm constrained trust region method and the piecewise random subsampling, *Geophys. J. Int.,* **187,** 199–213.

Wielandt, E., 1993. Propagation and structural interpretation of non-plane waves, *Geophys. J. Int.,* **113,** 45–53.

Wu, G., Fomel, S. & Chen, Y., 2016. Data-driven time-frequency analysis of seismic data using regularized nonstationary autoregression, in *86th Annual International Meeting,* SEG, Expanded Abstracts, Dallas, TX, pp. 1700–1705.

Wu, J. & Bai, M., 2018a. Adaptive rank-reduction method for seismic data reconstruction, *J. Geophys. Eng.,* **15,** 1688.

Wu, J. & Bai, M., 2018b. Attenuating seismic noise via incoherent dictionary learning, *J. Geophys. Eng.,* **15,** 1327.

Wu, J. & Bai, M., 2018c. Fast principal component analysis for stacking seismic data, *J. Geophys. Eng.,* **15,** 295–306.

Wu, J. & Bai, M., 2018d. Incoherent dictionary learning for reducing crosstalk noise in least-squares reverse time migration, *Comput. Geosci.,* **114,** 11–21.

Wu, Z. & Huang, N.E., 2009. Ensemble empirical mode decomposition: a noise-assisted data analysis method, *Adv. Adapt. Data Anal.,* **1,** 1–41.

Xie, J., Di, B., Wei, J., Xie, Q., Zu, S. & Chen., Y., 2016. Stacking using truncated singular value decomposition and local similarity, in *78th Annual International Conference and Exhibition,* EAGE, Extended Abstracts, doi:10.3997/2214–4609.201601325.

Xue, Y., Chang, F., Zhang, D. & Chen, Y., 2016a. Simultaneous sources separation via an iterative rank-increasing method, *IEEE Geosci. Remote Sens. Lett.,* **13**(12), 1915–1919.

Xue, Z., Alger, N. & Fomel, S., 2016b. Full-waveform inversion using smoothing kernels, in *86th Annual International Meeting,* SEG Expanded Abstracts, Dallas, TX, pp. 1358–1363.

Yang, W., Wang, R., Chen, Y. & Wu, J., 2014. Random noise attenuation using a new spectral decomposition method, in *84th Annual International Meeting,* SEG Expanded Abstracts, DEnver, CO, pp. 4366–4370.

Zeng, Q., Guo, Y., Jiang, R., Ba, J., Ma, H. & Liu, J., 2017. Fluid sensitivity of rock physics parameters in reservoirs: quantitative analysis, *J. Seismic Explor.,* **26,** 125–140.

Zhang, D., Chen, Y. & Gan, S., 2016a. Iterative reconstruction of 3D seismic data via multiple constraints, in *78th Annual International Conference and Exhibition,* EAGE, Extended Abstracts, .

Zhang, D., Chen, Y. & Gan, S., 2016b. Multidimensional seismic data reconstruction with multiple constraints, in *86th Annual International Meeting,* SEG, Expanded Abstracts, Dallas, TX, pp. 4801–4806.

Zhang, D., Chen, Y., Huang, W. & Gan, S., 2016c. Multi-step reconstruction of 3D seismic data via an improved MSSA algorithm, in *CPS/SEG Beijing 2016 International Geophysical Conference & Exposition,* SEG, Expanded Abstracts, Beijing, China, pp. 745–749.

Zhang, P., Dai, Y., Wang, R. & Tan, Y., 2017. A quantitative evaluation method based on EMD for determining the accuracy of time-var[illegible] seismic wavelet extraction, *J. Seismic Explor.,* **26,** 267–292.

Zhou, Y. & Han, W., 2018. Multiples attenuation in the presence of bl[illegible] noise, *J. Seismic Explor.,* **27**(1), 69–88.

Zhou, Y., Li, S., Zhang, D. & Chen, Y., 2018. Seismic [illegible] attenua[illegible] using an online subspace tracking algorithm, *Geo*[illegible] **212**(2[illegible] 1072–1097.

Zu, S., Zhou, H., Chen, Y., Qu, S., Zou, X., C[illegible] & [illegible] 2016a. A periodically varying code for improving de[illegible] sources in marine acquisition, *Geophysi*[illegible]1, V2[illegible]5.

Zu, S., Zhou, H., Chen, Y., Pan, X., G[illegible] Zhang, D[illegible] C., 2016b. Recovering the most from big ga[illegible] least-squares [illegible]n, in *86th Annual International Meeting,* [illegible]d Abstracts, Dallas, TX, pp. 4128–4133.

Zu, S., Zhou, H., Mao, W[illegible]ang, D., Li, C[illegible]Y. & Chen, Y., 2017. Iterative deblending [illegible]ultaneous-source d[illegible]sing a coherency-pass shaping operator, [illegible]*ys. J.* **211**(1), 541–557.

APP[illegible] A: C[illegible]STRUCTION OF P[illegible]NO[illegible]IAL TRANSFORMS

Let [illegible] = 0, [illegible] denote a set of polynomials, which satisfies the ort[illegible]ity condition:

$$\sum_{i=0}^{N} P_k(x_i) P_j(x_i) = \delta_{j,k}. \tag{A1}$$

It is known that as polynomials, $P_j(x_i)$ can be expressed

$$P_j(x_i) = \sum_{k=0}^{j} a_{jk} x_i^k, \tag{A2}$$

a_{jk} denotes polynomial coefficients. It is natural that x_j can be expressed based on superposition of different polynomials:

$$x^j = \sum_{k=0}^{j} \beta_{jk} P_k(x). \tag{A3}$$

Based on eqs (A3) and (A4), jth polynomial can be expressed as lower order polynomials

$$P_j(x_i) = \left\{ x^j - \sum_{k=0}^{j-1} \beta_{jk} P_k(x_i) \right\} / \beta_{jj}, \tag{A4}$$

Get squares of eq. (A3) and co[illegible]e with eq. [illegible]an obtain

$$\beta_{jj} = \sqrt{\sum_{i=0}^{N} x_i^{2j} - \sum^{j-1} \beta_{jk}^2 \ldots} \tag{A5}$$

and

$$\beta_{jk} = \sum^{N} x_i \ldots (x_i). \tag{A6}$$

Fr[illegible] (A4)–(A6), we ca[illegible] construct the set of polynomials. We [illegible]get $\beta_{00} =$ [illegible] based on eq. (A5), and thus $P_0 = 1/\beta_{00}$, then [illegible]ute β_{10}, β[illegible] construct P_1. In the same way, we can construct a[illegible]omia[illegible]

[illegible]PPENDIX B: LOCAL SIMILARITY

[illegible] similarity between vectors **a** and **b** is defined as

$$\mathbf{c} = \sqrt{\mathbf{c}_1 \circ \mathbf{c}_2} \tag{B1}$$

where $\circ$ denotes dot product, $\mathbf{c}_1$ and $\mathbf{c}_2$ come from two least-squares minimization problems:

$$\mathbf{c}_1 = \arg\min_{\mathbf{c}_1} \|\mathbf{a} - \mathbf{B}\mathbf{c}_1\|_2^2 \tag{B2}$$

$$\mathbf{c}_2 = \arg\min_{\mathbf{c}_2} \|\mathbf{b} - \mathbf{A}\mathbf{c}_2\|_2^2, \tag{B3}$$

where **A** is a diagonal operator composed of the elements of **a** and **B** is a diagonal operator composed of the elements of **b**. Note that in eqs (B1)–(B3), **a**, **b** and **c** denote vectorized 2-D matrices. Eqs (B2) and (B3) can be solved using shaping regularization with a local-smoothness constraint:

$$\mathbf{c}_1 = [\lambda_1^2 \mathbf{I} + \mathbf{T}(\mathbf{B}^T\mathbf{B} - \lambda_1^2\mathbf{I})]^{-1}\mathbf{T}\mathbf{B}^T\mathbf{b}, \tag{B4}$$

$$\mathbf{c}_2 = [\lambda_2^2 \mathbf{I} + \mathbf{T}(\mathbf{A}^T\mathbf{A} - \lambda_2^2\mathbf{I})]^{-1}\mathbf{T}\mathbf{A}^T\mathbf{a}, \tag{B5}$$

where **T** is a smoothing operator and λ_1 and λ_2 are two parameters controlling the physical dimensionality and enabling fast convergence when inversion is implemented iteratively. These two parameters can be chosen as $\lambda_1 = \|\mathbf{B}^T\mathbf{B}\|_2$ and $\lambda_2 = \|\mathbf{A}^T\mathbf{A}\|_2$.