submitted to *Geophys. J. Int.*

Variational mode decomposition of seismic data

Siwei Yu[1,2], Jianwei Ma[1], Stanley Osher[2]

[1] *Department of Mathematics, Harbin Institute of Technology, Harbin,* 150001*, China*

[2] *Department of Mathematics, University of California, Los Angeles, CA,* 90095*, USA*

SUMMARY

Seismic decomposition shows people more hidden information in the data than superficially. Seismic decomposition can be use in time-frequency analysis, noise attenuation, interpolation, modeling, inversion and so on. Fourier transform decompose seismic data into harmonic waves with constant frequencies. Field data contain non-stationary signals. Empirical mode decomposition (EMD) method is designed to analysis non-stationary signal, which has been widely used in industry. People tried to apply EMD on seismic data and have achieved promising results on time-frequency analysis, noise attenuation and so on. But EMD results in two drawbacks when dealing with seismic data: (1) Low resolution in frequency spectrum makes it can not separate events with dip angles close to each other. (2) When used on noise attenuation, high wave-number is treated as noise, which also attacks energy of events with high-dip angles. We propose using variational mode decomposition method (VMD) for seismic data decomposition and use it for seismic noise attenuation. VMD is proposed for decomposing a data into an ensemble of band-limited modes. For seismic data consisting of linear events, the frequency slice of frequency-space spectrum is exactly band-limited. The noise attenuation algorithm is summarized as follows: First we apply Fourier transform on the time-axis of the 2D seismic data. Then we apply VMD on each frequency slice and the modes are combined to obtain the denoised frequency slice. At last an inverse Fourier transform is applied on time-axis to obtain the denoising result. The frequency centers of VMD are initialized with matching pursuit method. We also apply 2D VMD on 3D seismic data denoising. Numerical results show the proposed method achieves higher denoising quality than $f-x$ deconvolution and EMD denoising.

Key words: variational mode decomposition – empirical mode decomposition – seismic data – noise attenuation

1 INTRODUCTION

Decomposition of seismic data always shows people more hidden information in the data than superficially. Seismic data can be analyzed by time-frequency analysis for different time-frequency components, or by sparse transforms for different predefined features, or by mode decomposition for different modes derived from the data itself. In this part, we will talk about different decompositions of seismic data and their applications.

All starts from Fourier transform, which gives us the frequency information about seismic data. Other than that, we also concern about direction, scale and other information in seismic data. Sparse decompositions/transforms such as the curvelet transform (Ma & Plonka 2010), the physical wavelet transform (Zhang & Ulrych 2003), the seislet transform (Fomel & Liu 2010) and the dreamlet transform (Geng et al. 2009), help us reveal the additional information by taking advantage of different predefined basis. Adaptive transform (Kaplan et al. 2009; Yu et al. 2015) optimize sparse decomposition of seismic data and explore similarities in the patch space. Sparse decomposition is popularly used in seismic noise attenuation (Neelamani et al. 2008), seismic interpolation (Shahidi et al. 2013), even seismic wave modeling (Gazdag 1981; Sun et al. 2009) and inversion (Pratt 1999).

Fourier transform decomposes the signal into harmonics with constant frequencies. But in seismic data, signal frequency changes along the traces, termed non-stationary and need to be analyzed with time frequency analysis (TFA) methods (Han & van der Baan 2013). This is important because different media show different frequency responses to seismic wave, such that faults and channels can be located (Han & van der Baan 2013; Liu & Marfurt 2007). Ground roll can also be separated by TFA cause ground roll is of low frequency (Liu & Fomel 2013). Short-time Fourier transform (STFT), continuous wavelet transform (CWT) and S transform (ST) are classical TFA tools (Tary et al. 2014).

The STFT, CWT and S transform are all bound by the Heisenberg uncertainty principle with a trade off between time and frequency localizations (Tary et al. 2014). Empirical mode decomposition (EMD) (Huang et al. 1998) is introduced to analysis non-stationary signal and have been widely used in industry (Klügel 2012; Barnhart & Eichinger 2011). TFA based on EMD achieves higher resolution on both time and frequency axis (Tary et al. 2014). In seismic data processing, EMD is not only used for time-frequency analysis, but also used for random and coherent noise attenuation (Bekara & van der Baan 2009). The authors first transform the data into frequency-space domain by fast Fourier transform (FFT) along time axis and then apply EMD on each frequency slice of the $f-x$ spectrum, and then treat the first instinct mode function (IMF) as noise and the sum of the rest IMFs as useful signal. The advantage of EMD over other methods is that (1) it is easy to implement and no parameter is required. (2) it is adaptive to data and can deal with non-stationary signal. But IMF of EMD is not based on bandlimited assumption, so EMD is of low resolution in wave-number axis. Keep in mind here the resolution mean the ability of splitting different wave-number components. In TFA, the high resolution is achieved if only different IMFs are separated by EMD. When using in noise attenuation, EMD attacks all energy at high wave-number, so it fails to preserve the energy of dip-angle events. Regarding to this drawback, we introduce variational mode decomposition (VMD) (Dragomiretskiy & Zosso 2014, 2015) of seismic data, showing VMD achieves higher resolution in wave-number axis and denoising result.

For readers convenient, before going on with VMD, we first introduce existing seismic denoising methods expect for the mentioned sparse decomposition based methods. Nonlocal mean method (Bonar & Sacchi 2012) takes advantage of self similarity of seismic data. Rank reduce methods, such as singular spectrum methods (Sacchi 2009) and Cadzow filtering (Trickett et al. 2008), are carried out by re-arranging frequency slice of seismic data into a low rank Hankel matrix. The $f-x$ deconvolution, also named as prediction error filter (PEF) (Spitz 1991), is a classical method for random noise attenuation with linear regression in $f-x$ domain. As an improvement of EMD based noise attenuation method, Chen & Ma (2014) propose EMDPEF, which can better preserve energy of dip events than EMD method, as we address as the drawback of EMD method.

The VMD was proposed for decomposing a data into an ensemble of band-limited IMFs (we briefly name it as IMFs in both VMD and EMD, but keep in mind the IMFs are band-limited in VMD. EMD decomposes data into IMFs but does not stress the band-limited property.). For seismic data consisting of linear events, the frequency slice of $f-x$ spectrum is exactly band-limited. The band-limited priori guarantees high resolution in splitting events with different dip-angles. Also different from EMD which models the individual modes as signals with explicit IMFs, VMD starts from variational energy minimization of IMFs. The IMFs are extracted concurrently instead of recursively, leading to its high efficiency. VMD has been applied on biomedical image denoising (Lahmiri & Boukadoum 2014) etc. For seismic data, VMD is used for ground roll attenuation (Liu et al. 2015). In this work, we first show how seismic data is decomposed by VMD and then use VMD for random noise attenuation. Instead of dealing real and imaginary parts separately, we work on the complex spectrum of seismic data directly by extending VMD to complex-valued situation. A matching pursuit method is used for initializing the frequency centers of IMFs of VMD. Also, 3D seismic noise attenuation is carried out based on 2D VMD.

In all, the VMD is different from EMD in the following aspects: (1) With the band-limit priori, VMD achieves higher resolution in wave-number axis when splitting events. (2) VMD subtracts the IMFs concurrently instead of recursively, leading to great improvement of efficiency. (3) In noise attenuation, no mode is removed. So events with high-dip angles can be better preserved.

The rest of this paper is arranged as follows. First we introduce the model of seismic data briefly. For the readers who are not familiar with EMD and VMD, we introduce the basic theory and how to apply them on seismic data decomposition and noise attenuation in the second part, illustrated with toy numerical experiments. In the third part we give numerical results of noise attenuation for synthetic data and field data. Finally we conclude and give possible further directions of our work.

2 THEORY

2.1 Seismic data with linear events and its spectrum

The propagation of seismic wave is controlled by wave equations. Seismic portion containing one linear event can be simply modeled by plane wave $d(t, x)$:

$$d(t, x) = w(t - x/c) \tag{1}$$

where x, t stand for coordinates of offset axis and time axis. The w is a waveform, such as Ricker wavelet. The c is the velocity of wave propagation in the medium.

After Fourier transform along t-axis of (1), we obtain the $f-x$ spectrum $D(f, x)$ of original data:

$$D(f, x) = W(f)e^{i2\pi fx/c} \tag{2}$$

where f is frequency, W is the Fourier transform of w. We assume the data is regularly sampled in x-axis and sample interval is $\triangle x$, then we discretize (2) and define:

$$D_f(m) \equiv D(m\triangle x, f), m = 1, 2, ..., M \tag{3}$$

where M is the number of traces. The relationship between adjacent trace is:

$$D_f(m) = b_1(f)D_f(m - 1), m = 2, ..., M \tag{4}$$

where $b_1(f) = \exp(i2\pi f\triangle x/V)$ is a constant for each frequency.

From (4) we can see the frequency slice D_f is composed of one complex harmonic in the $f - x$ domain. It can be shown that the superposition of p linear events in the $t - x$ domain is equivalent to the superposition of p complex harmonics in $f - x$ domain (Bekara & van der Baan 2009).

A synthetic seismic data, its $f - x$ spectrum and $f - k$ spectrum are shown in Figure 1. Here $f - k$ spectrum is obtained by applying 2D FFT on seismic data. The data is composed of four linear events. So the frequency slice of $f - x$ spectrum is composed by four complex harmonics, which can be seen in Figure 1(d), where the frequency slice of $f - k$ spectrum contains four peaks. A noisy version is shown in Figure 2. We can obtain a denoised data by eliminating noise in the $f - x$ spectrum, where EMD and VMD come up.

2.2 Empirical Mode Decomposition

EMD is designed to analyze a non-stationary signal, by decomposing the signal into different 'modes' of oscillations, named IMF. IMF satisfies two conditions: (1) The number of extrema and the number of zero crossings must be equal or differ at most by one, and (2) at any point the mean value of the envelope defined by the local maxima and the envelope defined by the local minima must be zero (Huang et al. 1998). EMD is implemented with a sifting algorithm. For the details of the sifting algorithm, readers may refer to (Huang et al. 1998; Bekara &

van der Baan 2009; Chen & Ma 2014). In this section, we give two examples of EMD on single frequency slice of the data in previous section.

In Figure 3, we show EMD on the real part of the data in Figure 1(d). The data is approximately composed by four harmonics, but the decomposed IMFs are far away from harmonics. In Figure 4, the EMD corresponds to real part of data in Figure 2(d). IMF1 represent the fastest oscillations in the data. So IMF1 is treated as noise, and IMF2 - IMF5 correspond to IMF1-IMF4 in Figure 3. If we summarize IMF2 to IMF5, we can obtain a denoised version of Figure 2(d). If EMD is applied on all frequency slices of the $f-x$ spectrum, we can obtain a denoised $f-x$ spectrum. The algorithm of EMD seismic noise attenuation is summarized as follows (Bekara & van der Baan 2009):

Algorithm 1 EMD based seismic noise attenuation

Input: Noisy data $d(x,t)$.

1: Select a time window and transform the data to the $f-x$ domain.
2: **for** every frequency **do**
3: separate real and imaginary parts in the spatial sequence,
4: compute IMF1, for the real signal and subtract it to obtain the filtered real signal,
5: repeat for the imaginary part, and
6: combine to create the filtered complex signal.
7: **end for**
8: Transform data back to the $t-x$ domain.
9: Repeat for the next time window.

Output: Denoised data.

2.3 Variational Mode Decomposition

The VMD was proposed to decompose a data into an ensemble of band-limited IMFs. The IMFs are extracted concurrently instead of recursively, leading to its high efficiency. VMD is achieved by solving the following optimization problem (Dragomiretskiy & Zosso 2014):

$$\min_{\{u_k\},\{\omega_k\}} \{\sum_k \|\partial_x[(\delta(x)+\frac{j}{\pi x})*u_k(x)]e^{-j\omega_k x}\|_2^2\}, \quad s.t. \sum_k u_k = f \tag{5}$$

where u_k are decomposed modes, w_k are the center of the frequency of the corresponding modes. $\delta(\cdot)$ is a Dirac impulse. f is the original real-valued signal. The term $(\delta(x)+\frac{j}{\pi t})*u_k(x)$ is the Hilbert transform of u_k, which transforms u_k to an analytic signal and makes real valued signal u_k complex valued such that the spectrum has a single peak in the positive frequencies.

The frequency slice of $f-x$ spectrum of seismic data is complex-valued. We can deal with real and imaginary parts separately like EMD. Or we can extend optimization (5) to complex-valued data directly. It is not necessary to apply Hilbert transform on complex-valued data, so we ignore the term $\delta(x) + \frac{j}{\pi t}$ and the problem turns into:

$$\min_{\{u_k\},\{\omega_k\}} \{\sum_k \|\partial_x[u_k(x)e^{-j\omega_k x}]\|_2^2\}, \quad s.t. \sum_k u_k = f. \tag{6}$$

Problem (6) is first written as the form of augmented Lagrangian:

$$L(u_k, \omega_k, \lambda) := \alpha \sum_k \|\partial_x[u_k(x)]e^{-j\omega_k x}\|_2^2 + \|f(x) - \sum_k u_k(x)\|_2^2 + \langle\lambda(x), f(x) - \sum_k u_k(x)\rangle \tag{7}$$

where α is a parameter balancing variational minimization term and constrain term. $\lambda(x)$ is the Lagrange multiplier. Minimization of (7) can be solved with ADMM (Alternating Direction Method of Multipliers) method (See details in appendix). The initial value of w_k can be set uniformly or randomly.

We test VMD on the synthetic model in Figure 1(d) and Figure 2(d). The initial values of w_k are selected uniformly. The results of VMD on clear data is shown in Figure 5. We can see that all the four modes are approximate harmonics. Also the spectrum is perfectly decomposed into four peaks. The result of VMD on noisy data is shown in Figure 6. Only three modes are detected. The fourth mode is far away from real value. Because minimization (5) is non-convex problem, so it is sensitive to initial value and noise. In order to work on noisy situation, we introduce a matching pursuit method to initialize w_k in the next section.

Before going on to the matching pursuit method, we give the VMD-based noise attenuation algorithm here:

Algorithm 2 VMD based seismic noise attenuation

Input: Noisy data $d(x, t)$.

1: Select a time window and transform the data to the $f - x$ domain.
2: **for** every frequency **do**
3: compute complex-valued VMD,
4: combine the IMFs to create the filtered complex signal.
5: **end for**
6: Transform data back to the $t - x$ domain.
7: Repeat for the next time window.

Output: Denoised data.

2.4 Initialize w_k with Matching Pursuit

The matching pursuit algorithm is applied in frequency slice of $f - k$ domain. We summarize the algorithm in Algorithm 3 2.4.

VMD with matching pursuit initialization on signal in Figure 2(d) is shown in Figure 7. We can see all modes are detected successfully. After combining the decomposed modes, we obtain a denoised version of Figure 6(a).

Algorithm 3 Matching pursuit algorithm for initializing w_k

Input: Frequency slice of $f-k$ spectrum

1: Fourier transform along x-axis.
2: Find max amplitude of the amplitude spectrum. Pick the index and apply inverse Fourier transform on selected index, then subtract it from original signal.
3: Use the residual as input and repeat 1,2 until required number of indexes are found.

Output: indexes served as initialized frequency center.

2.5 EMD and VMD for events splitting

In this section, we applied EMD and VMD on all frequency slices and see the decomposed modes. In EMD, the frequency of each IMF decreases according to the order in which it is separated out, by subtracting the IMFs by order, we obtain events from high-dip angles to low-dip angles (Chen & Ma 2014).

First, we test EMD and VMD on a synthetic data consisting of two adjoint linear events with different dip angles, shown in Figure 8(a). The data is first transform into frequency-space domain, and the real part of the $30th$ slice of its $f-x$ spectrum is subtracted and shown in Figure 8(b). Figure 8(c) and Figure 8(e) show VMD on data in Figure 8(b). The two parts with different frequencies are separated. Figure 8(d) shows the decomposed mode by EMD, and Figure 8(f) shows the residual. Only one mode is subtracted. This example shows that VMD can split modes with different frequencies better than EMD. This leads to better decomposition of seismic events than EMD, shown in the following test.

In (Chen & Ma 2014), the author gives an example where the events are separated successfully with EMD. While in Figure 9(a), the synthetic data consists of events which have dip-angles close to each other. Figure 9(b)-Figure 9(e) show the separated events from high-dip angles to low-dip angles. As we can see, adjacent events can not be separated.

In VMD, the IMFs from low-frequency to high-frequency correspond to events from negative-dip-angle to positive-dip-angle. First, the IMFs are sorted from low frequency to high frequency in a sequence. Then we subtract the IMFs in same position of the sequences of all frequency slices to reconstruct one event. In Figure 10, we can see that the four events are successfully separated. The band-limited property of VMD makes it to achieve higher resolution in wave-number axis than EMD.

2.6 2D VMD for 3D seismic data

A 2D plane wave in constant velocity medium can be modeled by $d(t,\vec{x})$:

$$d(\vec{x},t) = w(t - \frac{\langle \vec{x}, \vec{k} \rangle}{c}) \tag{8}$$

where $\vec{x} = (m,h)$ stands for coordinates of midpoint and offset. $\vec{k}$ is the unit direction of wave propagation. $\langle \cdot, \cdot \rangle$ mean inner product or two vectors in Hilbert space.

Using the same derivation as 2D situation, it can be shown that the frequency slice of

D_f is composed of one complex harmonic along the direction of $\vec{k}$ in the $f-m-h$ domain. Here D_f is one frequency slice of 1D Fourier transform of $d(t,\vec{x})$ along time axis. Also the superposition of p plane waves in the $t-m-h$ domain is equivalent to the superposition of p complex harmonics in $f-m-h$ domain.

Dragomiretskiy & Zosso (2015) propose 2D VMD for real-valued data. As in 1D situation, we first extend 2D VMD for complex-valued data. The extended optimization function is as follow:

$$\min_{\{u_k\},\{\vec{\omega}_k\}} \{\sum_k \|\nabla[u_k(\vec{x})e^{-j\langle\vec{\omega}_k,\vec{x}\rangle}]\|_2^2\}, \quad s.t. \ \sum_k u_k = f \tag{9}$$

where ∇ is the gradient operator. $\vec{\omega}_k$ is frequency vector in 2D plane. Same as 1D VMD, (9) is first written as augmented Lagrangian and then solved with ADMM.

A 3D synthetic model consisting of three plane waves is shown in Figure 11(a). Center slices of each direction are shown. There are 150 samples in time axis, 100 samples in midpoint axis and 100 sample in offset axis. Time interval is $4ms$, midpoint and offset interval are both $0.05km$. First we transform the data into $f-m-h$ domain. The 30th frequency slice is shown in Figure 11(b). After applying 2D VMD on the frequency slice, we obtain three modes, shown in Figure 11(c)-Figure 11(e). In Figure 12, we give the $k_m - k_h$ spectrum corresponding to data in Figure 11. We can clearly see that the three modes are separated successfully.

We provide the 2D VMD-based 3D seismic data noise attenuation algorithm here:

Algorithm 4 2D VMD seismic noise attenuation

Input: Noisy data $d(m,h,t)$.

1: Select a time window and transform the data to the $f-m-h$ domain.
2: **for** every frequency **do**
3: compute 2D complex-valued VMD,
4: combine the IMFs to create the filtered complex signal.
5: **end for**
6: Transform data back to the $t-m-h$ domain.
7: Repeat for the next time window.

Output: Denoised data.

3 NUMERICAL RESULTS

In this section, we apply VMD on noise attenuation on both synthetic and field data, compared with $f-x$ deconvolution and EMD based method. A 3D experiment on synthetic data is carried out at the end of this section. For models that does not satisfy the linear event assumption, We deal with the data by splitting windows. The temporal window length is $512ms$ and there is 50% overlapping between adjoint temporal windows for removing edge effects. The space window length is 86 traces and the overlapping is 60%. Parameter α is set as 2000 for the tests.

3.1 Denoising on data consisting of linear events:

The synthetic model in Figure 13 is composed of four linear events, with 501 time samples and 128 traces. The time sample interval is set as 4ms. Figure 13(a)-13(c) show the clean data, the noisy data and the noise separately, with SNR (signal-to-noise ratio) = 2.0. Here, SNR means the ratio of the energy of clean data and the energy of noise. Figure 13(d)-13(f) are the noise attenuation results of $f-x$ deconvolution method, EMD method, and VMD method. Figure 13(g)-13(i) are the corresponding noise portions. From the noise portions we can see that $f-x$ deconvolution method apparently removes the energy of the events. EMD method keeps useful energy better than $f-x$ deconvolution on flat events, but still removes the energy of dip events. VMD methods keep most of the useful energy. The elapsed time of these three methods are 0.17s, 11.94s, 2.00s respectively.

Model in Figure 14(a) consists of three adjoint linear events. The parameters are the same as in previous model. Figure 14(b),14(c) are the noisy version and the noise, with SNR = 2.0. 14(d)-14(f) are the noise attenuation results of $f-x$ deconvolution method, EMD method, and VMD method. Figure 14(g)-14(i) are the corresponding noise portions. Same as previous example, $f-x$ deconvolution method and EMD method remove much of the noise, but also harm the useful energy. The VMD method remove less noise, but keeps most of the useful energy.

3.2 Denoising on data consisting of hyperbolic events:

We test the VMD on model consisting of three hyperbolic events in this example. There are 400 time samples and 190 traces, with time sample interval = 4ms. Figure 15(a) is the noise corrupted data, with SNR = 2.0. 15(b)-15(d) are the noise attenuation results of $f-x$ deconvolution method, EMD method, and VMD method. Figure 15(e)-15(g) are corresponding noise portion. As we can see, $f-x$ deconvolution method harms useful energy of all three events. The EMD method keeps the energy for the two approximately flat events, but not for the dip one. The VMD method preserves most of the useful energy of all three events.

3.3 Denoising on field data:

In this example, we test the VMD on field data. Figure 16(a) shows the data with noise, with 1501 time samples and 333 traces. The time sample interval is 4ms and space sample interval is $0.01km$.

In Figure 17, we show the denoised results of $f-x$ deconvolution method, EMD method, and VMD method and the corresponding noise sections. We can see the details in the zoomed version in Figure 18. $f-x$ deconvolution removes lots of useful energy. The EMD method keeps the energy in horizontal, but there are some dip events still in the noise section shown Figure 18(d). The VMD method keeps most of useful energy. As comparison with state-of-art methods, we give denoising results based on Db6 wavelet (Daubechies et al. 1992) method, curvelet method, empirical curvelet transform (ECT)(Gilles 2013; Gilles et al. 2014) method in Figure 19. In Figure 19(a), DB6 wavelet method results in discontinuous events and dot-like artifact features. In Figure 19(c) and Figure 19(e), curvelet method and ECT method both result in smooth events. The VMD is competitive with the state-of-art methods. By considering spatial redundancy in our next work, the performance of VMD can be further improved.

3.4 Denoising on 3D synthetic data:

In this test, 2D VMD is applied on 3D seismic data for noise attenuation. Figure 20(a) shows noisy version of the data in Figure 11(a), with SNR = 2.0. Figure 20(b) and 20(d) show the denoised results by 2D EMD and 2D VMD. Figure 20(c) and 20(e) are corresponding noise sections. Both methods keep the useful energy well. But in Figure 20(c), we still can notice some weak plane waves. In Figure 20(e), there is only noise left.

3.5 Denoising on 3D field data:

Figure 21(a) shows a 3D post stack seismic data. The data consists of 625 time samples, with time sample interval equals $4ms$. The samples in inline and crossline directions are 701 and 21 separately, and the space sample intervals are both $0.02km$. Figure 21(b) and Figure 21(d) show the denoised results with EMD and VMD methods respectively. Figure 21(c) and Figure 21(e) are the corresponding noise portions. From the left bottom corner of the noise portions, we can see the EMD method obviously eliminates energy dip events. While in the VMD method, the energy of the dip events are mostly kept.

3.6 Discussion on parameter α:

In equation (7), parameter α controls how much noise is removed. In order to keep most of the useful energy, we use small α in the previous tests. VMD removes less noise than $f-x$ deconvolution and EMD method. In this test, we increase α to see how the results is affected. The test is carried on the data in Figure 16(a). Zoomed denoised results and corresponding noise portions are shown in Figure 22.

The α is set as 20000 and 100000 in Figure 22(a) and Figure 22(c) separately. We notice that the result in Figure 22(a) is similar as the result using EMD method in Figure 18(c). Also the result in Figure 22(c) is similar as the result using $f-x$ deconvolution method in Figure 18(a), but Figure 22(c) shows much smoother events than 18(a). Obvious events can be seen in Figure 22(b) and Figure 22(d). EMD method is parameterless and easy to implement, but it is also a drawback when we want to change how much noise to remove. On the contrary, we can adjust the α in VMD for desired denoising results.

4 CONCLUSION

In this paper, we present variational mode decomposition of seismic data and its applications on noise attenuation. The VMD has some good properties than the existing EMD: (1) With the band-limit priori, VMD achieves higher resolution in wave-number axis when splitting events. (2) VMD subtracts the IMFs concurrently instead of recursively, leading to great improvement of efficiency. (3) In noise attenuation, no mode is removed. So events with high-dip angles can be better preserved. We extend the VMD to complex-valued data for processing the $f-x$ spectrum of seismic data. In order to deal with noisy situation, we also introduce matching pursuit method to initialize the center frequencies of VMD. At last, 2D VMD is used for 3D seismic data denoising. Though VMD is efficient in preserving useful signals, it also keeps more noise than EMD and $f-x$, which is a drawback currently. In future work, this drawback should be addressed. Also it is promising to explore more applications of VMD on seismic data such as multiple removal in marine exploration data and earthquake data mode decomposition and prediction. The VMD can also be used for geophysical compressed sensing.

5 ACKNOWLEDGEMENT

The authors thank Dr. Dominique Zosso and Yangkang Chen for valuable discussions. The authors also thank Yangkang Chen for providing the field data. This work is supported by NSFC (grant number: NSFC 91330108, 41374121, 61327013), and the Fundamental Research Funds for the Central Universities (grant number: HIT.PIRS.A201501).

6 APPENDIX: ADMM FOR VMD

We rewrite equation (7) here:

$$L(u_k,\omega_k,\lambda) := \alpha \sum_k \|\partial_x[(\delta(x)) * u_k(x)]e^{-j\omega_k x}\|_2^2 + \|f(x) - \sum_k u_k(x)\|_2^2 + \langle \lambda(x), f(x) - \sum_k u_k(x)\rangle. \tag{10}$$

The ADMM for VMD is as in Algorithm (5):

Minimization (11) is first written as:

$$u_k^{n+1} = \underset{u_k \in X}{\operatorname{argmin}}\, \alpha \|\partial_x[u_k(x)e^{-j\omega_k x}]\|_2^2 + \|f(x) - \sum_i u_i(x) + \frac{\lambda(x)}{2}\|_2^2. \tag{14}$$

It can be solved in Fourier domain and there exists explicit solution:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i\neq k}\hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + \frac{\alpha}{2}(\omega - \omega_k)^2}. \tag{15}$$

Minimization (12) is written as:

$$\omega_k^{n+1} = \underset{\omega_k}{\operatorname{argmin}}\, \|\partial_x[u_k(x)e^{-j\omega_k x}]\|_2^2. \tag{16}$$

Algorithm 5 ADMM for VMD

Input: Initialize $u_k^1, \omega_k^1, \lambda^1, n = 0$

repeat

$n = n + 1$

for $k = 1 : K$ **do**

Update u_k:

$$u_k^{n+1} = arg \min_{u_k} L(\{u_{i<k}^{n+1}\}, \{u_{i\geq k}^n\}, \{\omega_i^n\}, \lambda^n) \tag{11}$$

end for

for $k = 1 : K$ **do**

Update ω_k:

$$\omega_k^{n+1} = arg \min_{\omega_k} L(\{u_i^{n+1}\}, \{\omega_{i<k}^{n+1}\}, \{\omega_{i\geq k}^n\}, \lambda^n) \tag{12}$$

end for

Dual ascent:

$$\lambda^{n+1} = \lambda^n + \tau(f - \sum_k u_k^{n+1}) \tag{13}$$

until convergence: $\sum_k \|u_k^{n+1} - u_k^n\|_2^2 / \|u_k^n\|_2^2 < \epsilon$

Output: Decomposed IMFs.

There also exists explicit solution in Fourier domain:

$$\omega_k^{n+1} = \frac{\int_{-\infty}^{\infty} \omega \|\hat{u}_k(\omega)\|^2 d\omega}{\int_{-\infty}^{\infty} \|\hat{u}_k(\omega)\|^2 d\omega}. \tag{17}$$

REFERENCES

Barnhart, B. & Eichinger, W., 2011. Empirical mode decomposition applied to solar irradiance, global temperature, sunspot number, and co 2 concentration data, *Journal of Atmospheric and Solar-Terrestrial Physics*, **73**(13), 1771–1779.

Bekara, M. & van der Baan, M., 2009. Random and coherent noise attenuation by empirical mode decomposition, *Geophysics*, **74**(5), V89–V98.

Bonar, D. & Sacchi, M., 2012. Denoising seismic data using the nonlocal means algorithm, *Geophysics*, **77**(1), A5–A8.

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

Daubechies, I. et al., 1992. *Ten lectures on wavelets*, vol. 61, SIAM.

Dragomiretskiy, K. & Zosso, D., 2014. Variational mode decomposition, *IEEE Transactions on Signal Processing*, **62**(3), 531–544.

Dragomiretskiy, K. & Zosso, D., 2015. Two-dimensional variational mode decomposition, in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pp. 197–208, Springer.

Fomel, S. & Liu, Y., 2010. Seislet transform and seislet frame, *Geophysics*, **75**(3), V25–V38.

Gazdag, J., 1981. Modeling of the acoustic wave equation with transform methods, *Geophysics*, **46**(6), 854–859.

Geng, Y., Wu, R.-S., Gao, J., et al., 2009. Dreamlet transform applied to seismic data compression and its effects on migration, in *2009 SEG Annual Meeting*, Society of Exploration Geophysicists.

Gilles, J., 2013. Empirical wavelet transform, *IEEE Transactions on Signal Processing*, **61**(16), 3999–4010.

Gilles, J., Tran, G., & Osher, S., 2014. 2d empirical transforms. wavelets, ridgelets, and curvelets revisited, *SIAM Journal on Imaging Sciences*, **7**(1), 157–186.

Han, J. & van der Baan, M., 2013. Empirical mode decomposition for seismic time-frequency analysis, *Geophysics*, **78**(2), O9–O19.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., & Liu, H. H., 1998. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 903–995, The Royal Society.

Kaplan, S. T., Sacchi, M. D., Ulrych, T. J., et al., 2009. Sparse coding for data-driven coherent and incoherent noise attenuation, in *2009 SEG Annual Meeting*, Society of Exploration Geophysicists.

Klügel, N., 2012. Practical empirical mode decomposition for audio synthesis, in *Int. Conference on Digital Audio Effects (DAFx-12)*, no. 2, pp. 15–18.

Lahmiri, S. & Boukadoum, M., 2014. Biomedical image denoising using variational mode decomposition, in *Biomedical Circuits and Systems Conference (BioCAS), 2014 IEEE*, pp. 340–343, IEEE.

Liu, J. & Marfurt, K. J., 2007. Instantaneous spectral attributes to detect channels, *Geophysics*, **72**(2), P23–P31.

Liu, W., Cao, S., & He, Y., 2015. Ground roll attenuation using variational mode decomposition, in *77th EAGE Conference and Exhibition 2015*.

Liu, Y. & Fomel, S., 2013. Seismic data analysis using local time-frequency decomposition, *Geophysical Prospecting*, **61**(3), 516–525.

Ma, J. & Plonka, G., 2010. The curvelet transform, *IEEE Signal Processing Magazine*, **27**(2), 118–133.

Neelamani, R., Baumstein, A. I., Gillard, D. G., Hadidi, M. T., & Soroka, W. L., 2008. Coherent and random noise attenuation using the curvelet transform, *The Leading Edge*, **27**(2), 240–248.

Pratt, R. G., 1999. Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model, *Geophysics*, **64**(3), 888–901.

Sacchi, M. D., 2009. Fx singular spectrum analysis, in *CSPG CSEG CWLS Convention*, pp. 392–395.

Shahidi, R., Tang, G., Ma, J., & Herrmann, F. J., 2013. Application of randomized sampling schemes to curvelet-based sparsity-promoting seismic data recovery, *Geophysical Prospecting*, **61**(5), 973–997.

Spitz, S., 1991. Seismic trace interpolation in the fx domain, *Geophysics*, **56**(6), 785–794.

Sun, B., Ma, J., Chauris, H., & Yang, H., 2009. Solving wave equations in the curvelet domain: A multi-scale and multi-directional approach, *Journal of Seismic Exploration*, **18**(4), 385–399.

Tary, J. B., Herrera, R. H., Han, J., & Baan, M., 2014. Spectral estimation–what is new? what is next?, *Reviews of Geophysics*, **52**(4), 723–749.

Trickett, S. et al., 2008. F-xy cadzow noise suppression, in *78th Annual International Meeting, SEG, Expanded Abstracts*, pp. 2586–2590.

Yu, S., Ma, J., Zhang, X., & Sacchi, M. D., 2015. Interpolation and denoising of high-dimensional seismic data by learning a tight frame, *Geophysics*, **80**(5), V119–V132.

Zhang, R. & Ulrych, T. J., 2003. Physical wavelet frame denoising, *Geophysics*, **68**(1), 225–231.

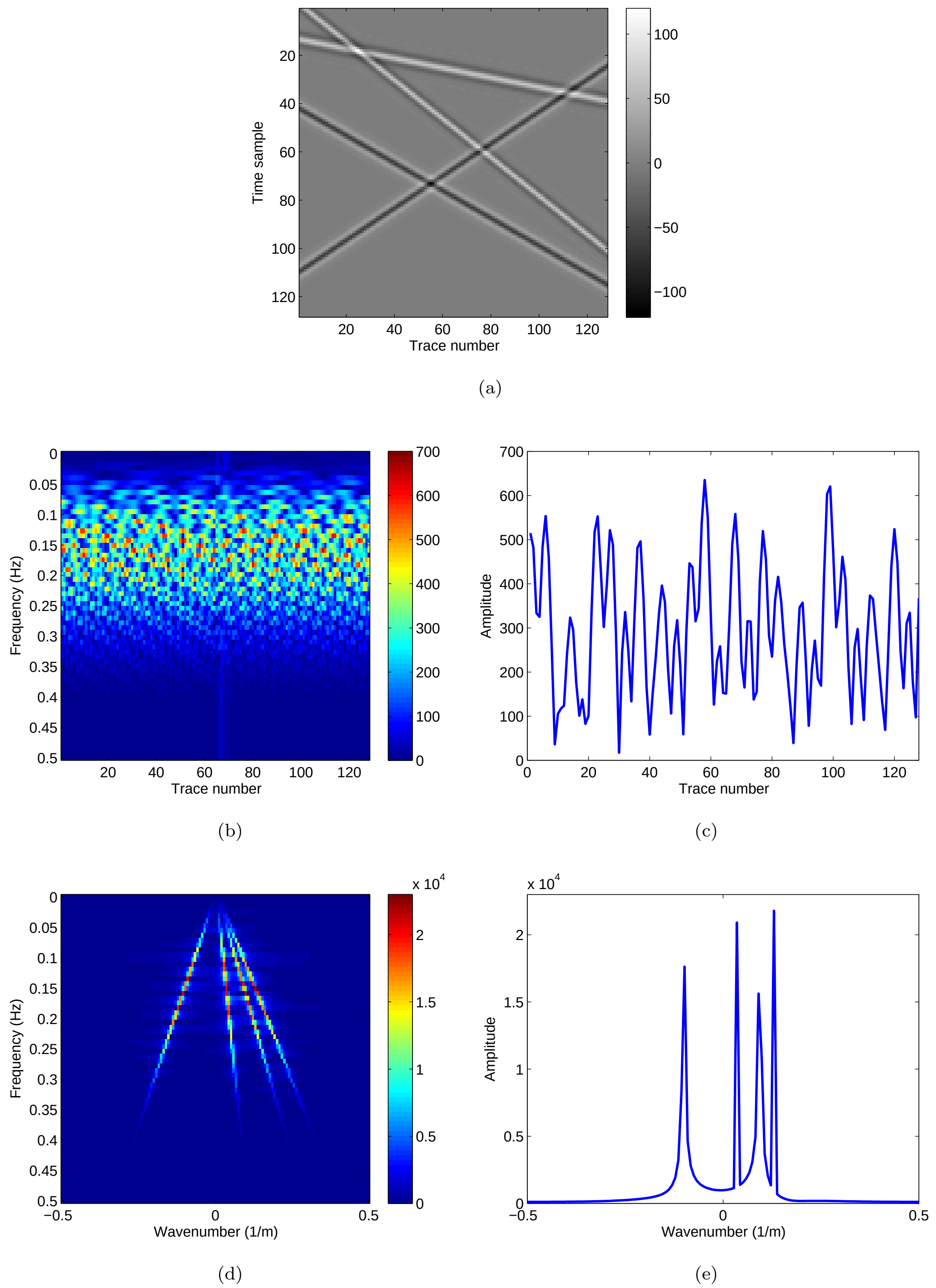


Figure 1. A synthetic seismic model consisting for linear events. (a) Synthetic model. (b) $f-x$ spectrum. (c) The 45th slice of (b). (d) F-K spectrum. (e) The 45th slice of (d).

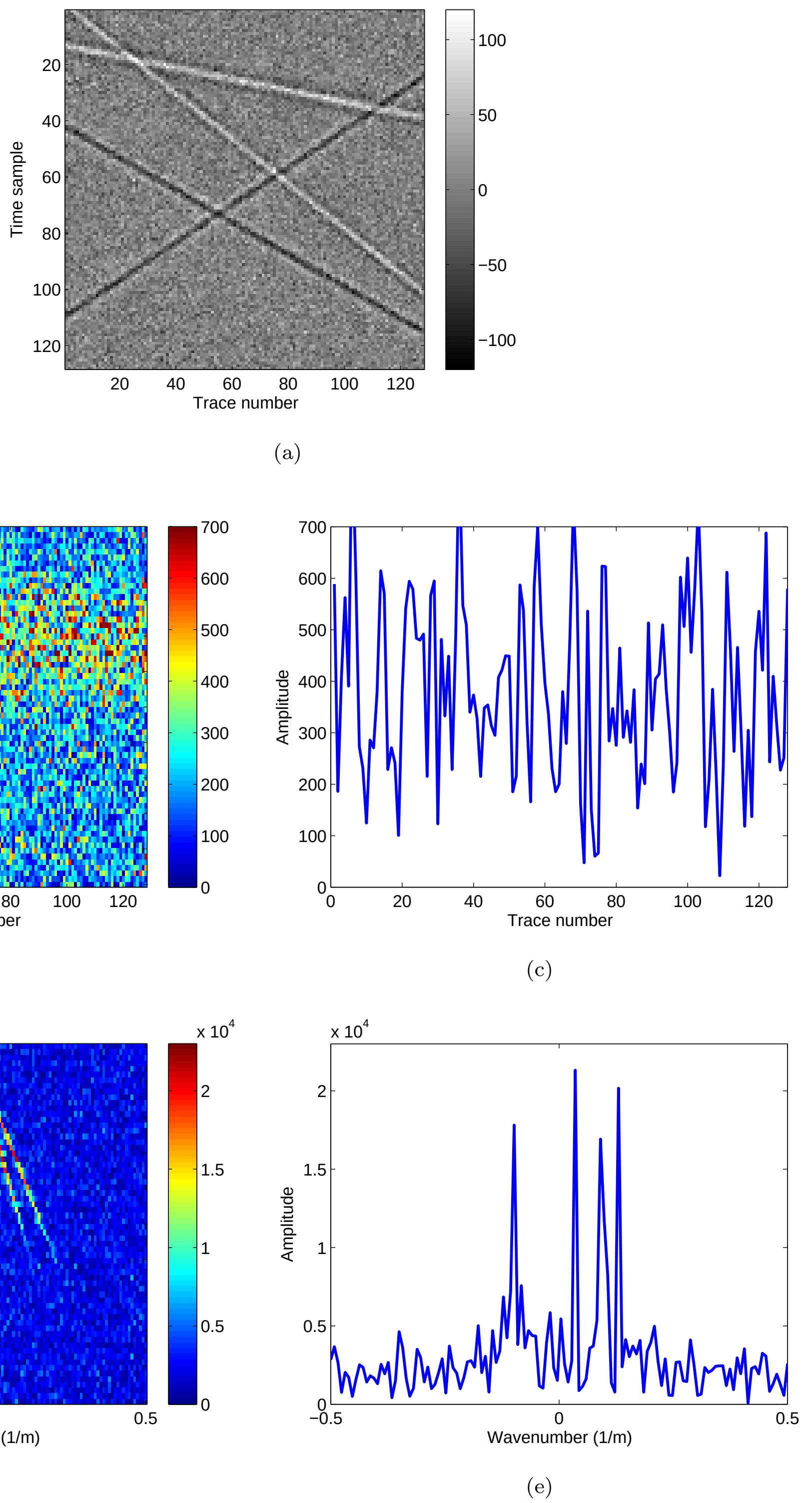


Figure 2. A synthetic seismic model with noise. (a) 2D seismic data contaminated with noise. (b) $f-x$ spectrum. (c) The 45th slice of (b). (d) F-K spectrum. (e) The 45th slice of (d).

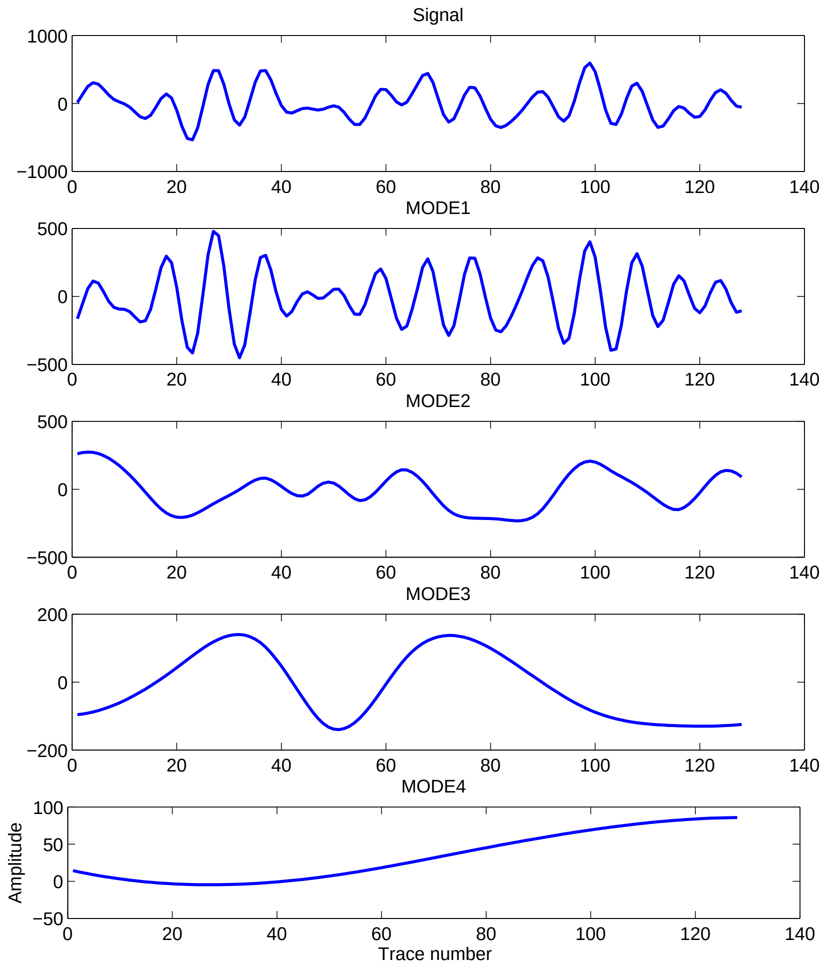


Figure 3. EMD of a clean signal extracted from Figure 1(d).

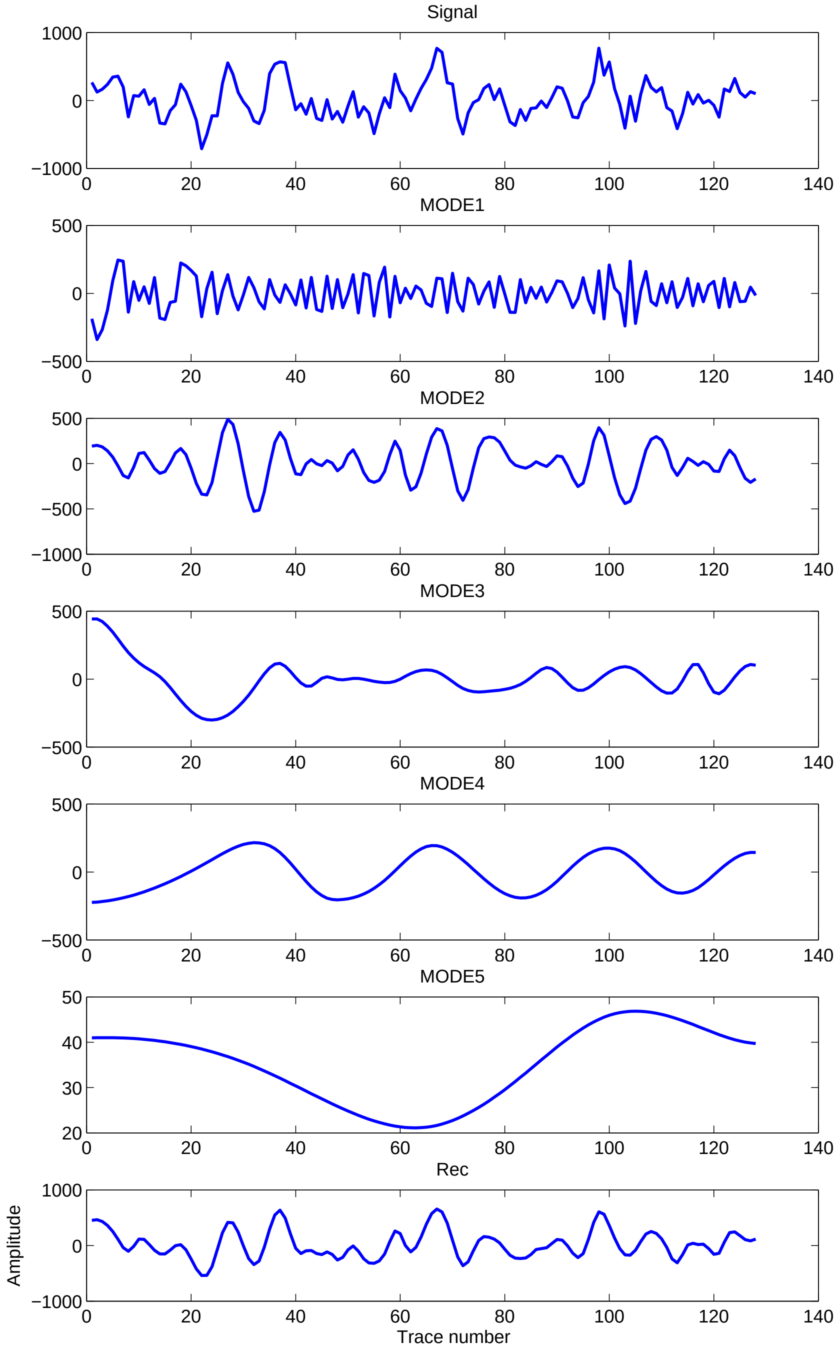


Figure 4. EMD of a noisy signal extracted from Figure 2(d).

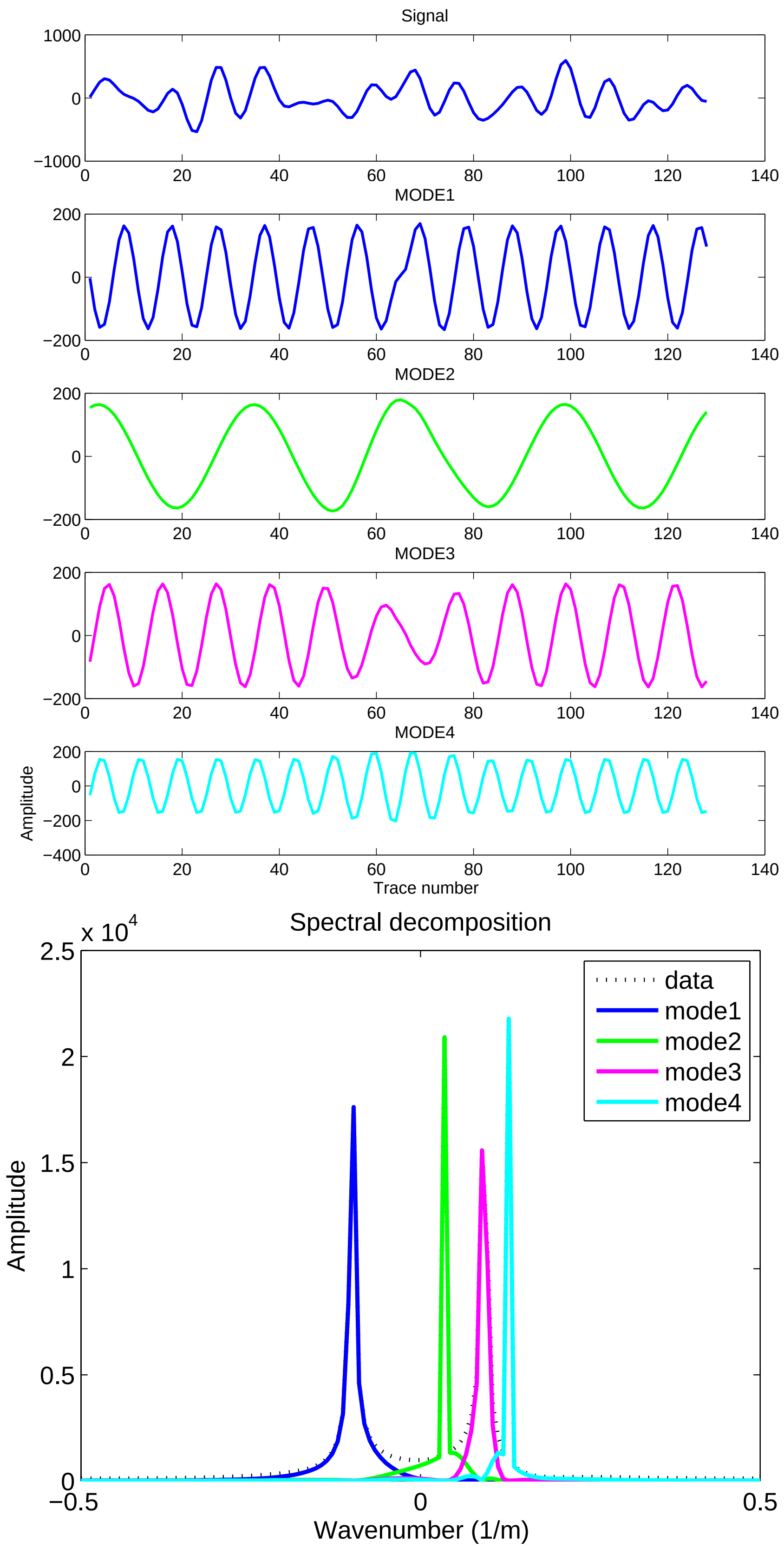


Figure 5. VMD of the clean signal extracted from Figure 1(d). The last figure are amplitude spectrums of the decomposed modes.

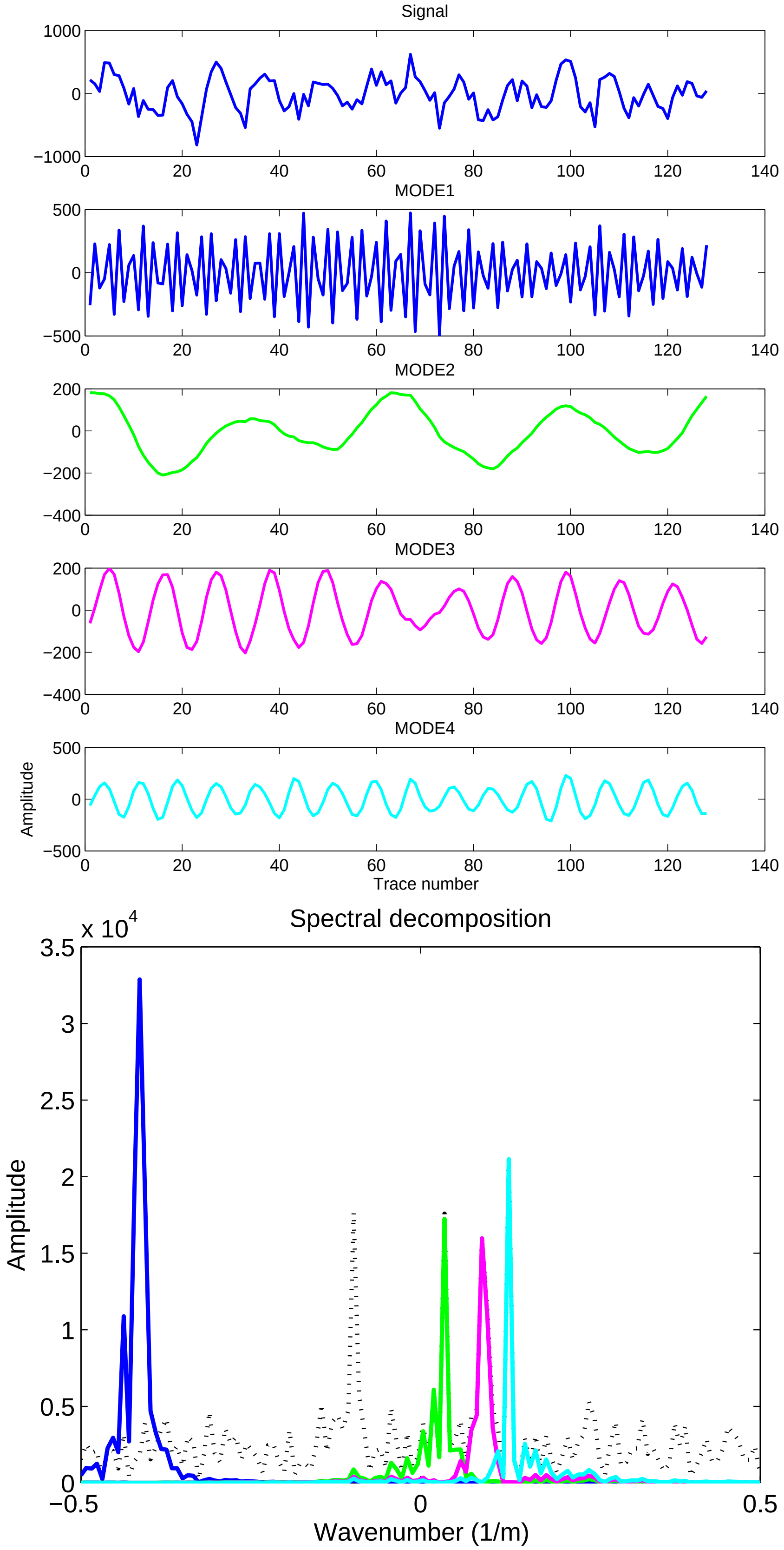


Figure 6. VMD of the noisy signal extracted from Figure 2(d). The last figure are amplitude spectrums of the decomposed modes.

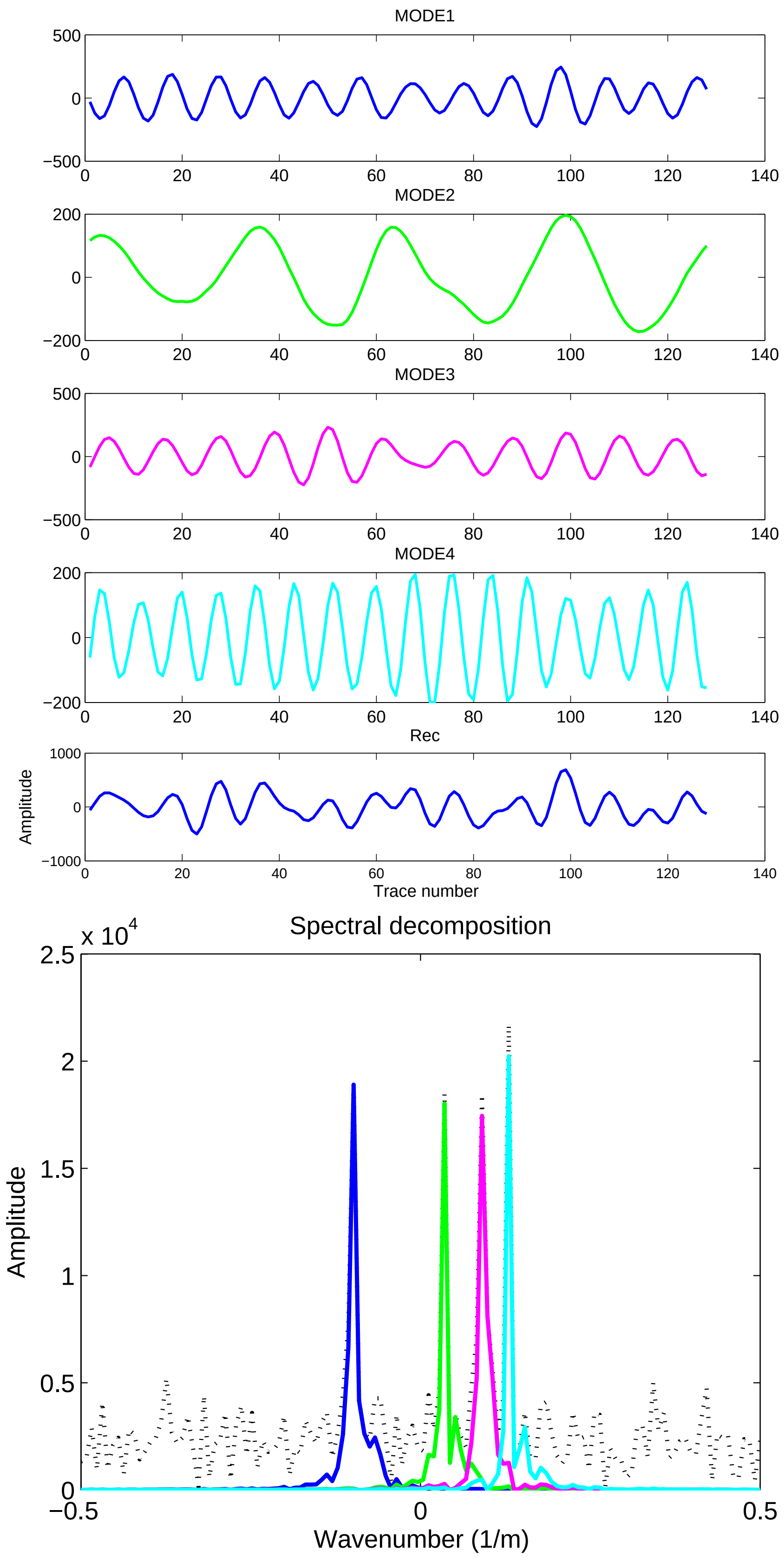


Figure 7. MPF initialized VMD of the noisy signal extracted from Figure 2(d). The last figure are amplitude spectrums of the decomposed modes.

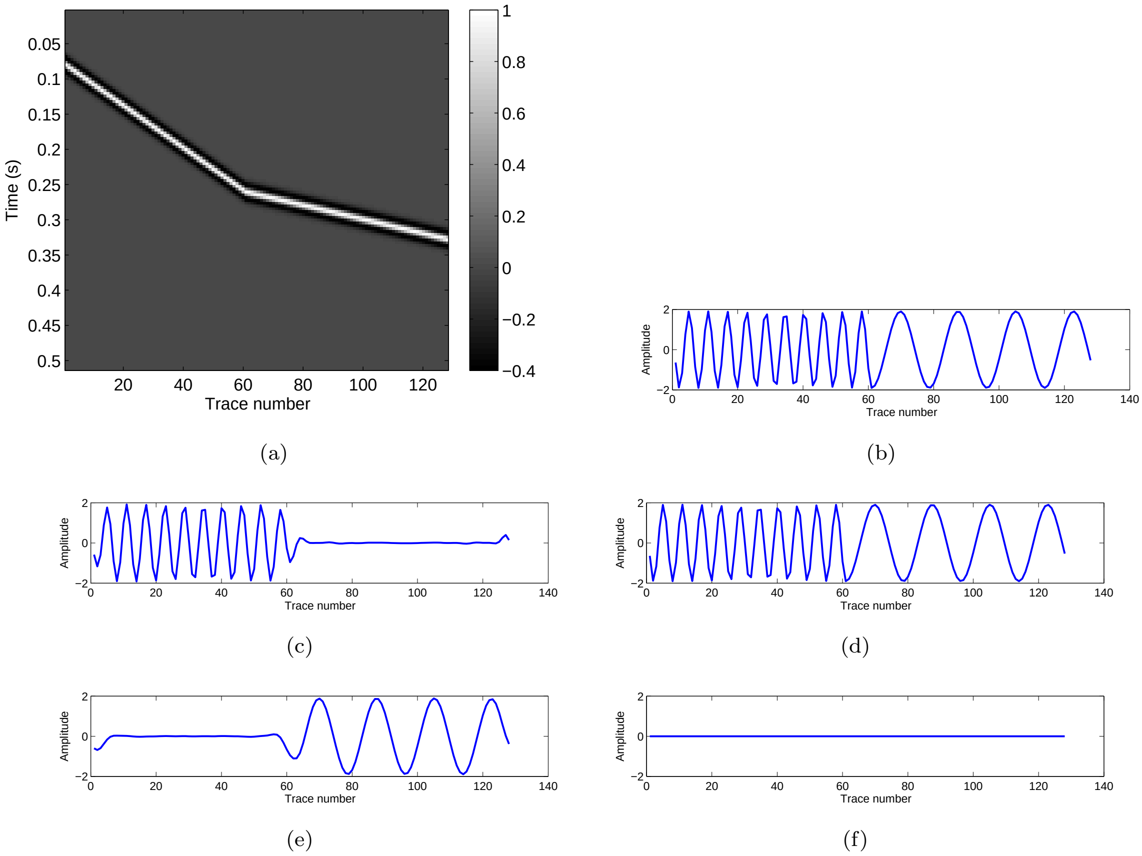


Figure 8. Mode decomposition on a non-stationary signal. (a) A synthetic data consisting of two adjoint linear events with different dip angles. (b) The 30*th* slice of the real part of the frequency-space transform of (a). (c) and (e) Two modes decomposed by VMD. (d) and (f) The decomposed mode by EMD and the residual.

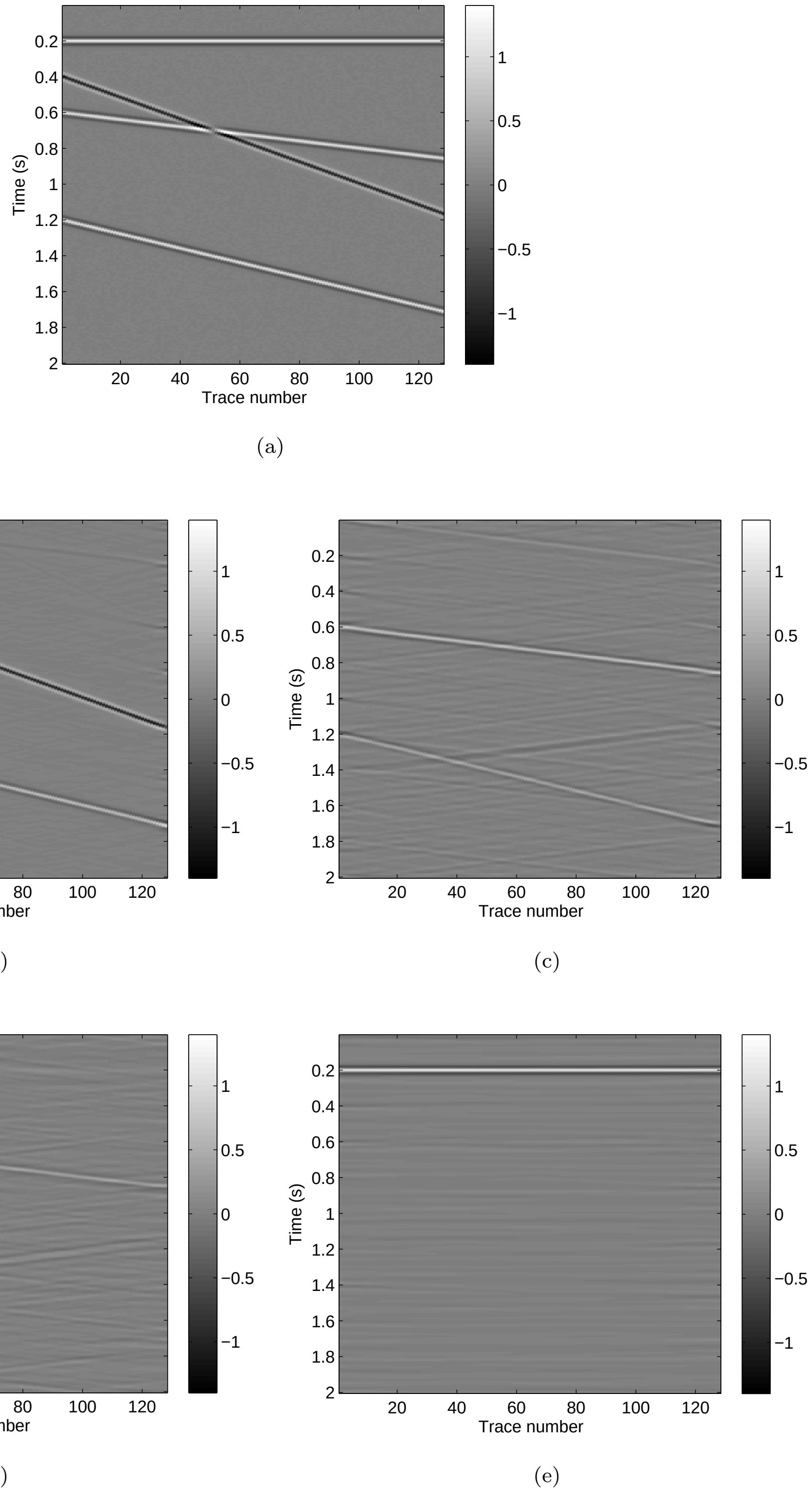
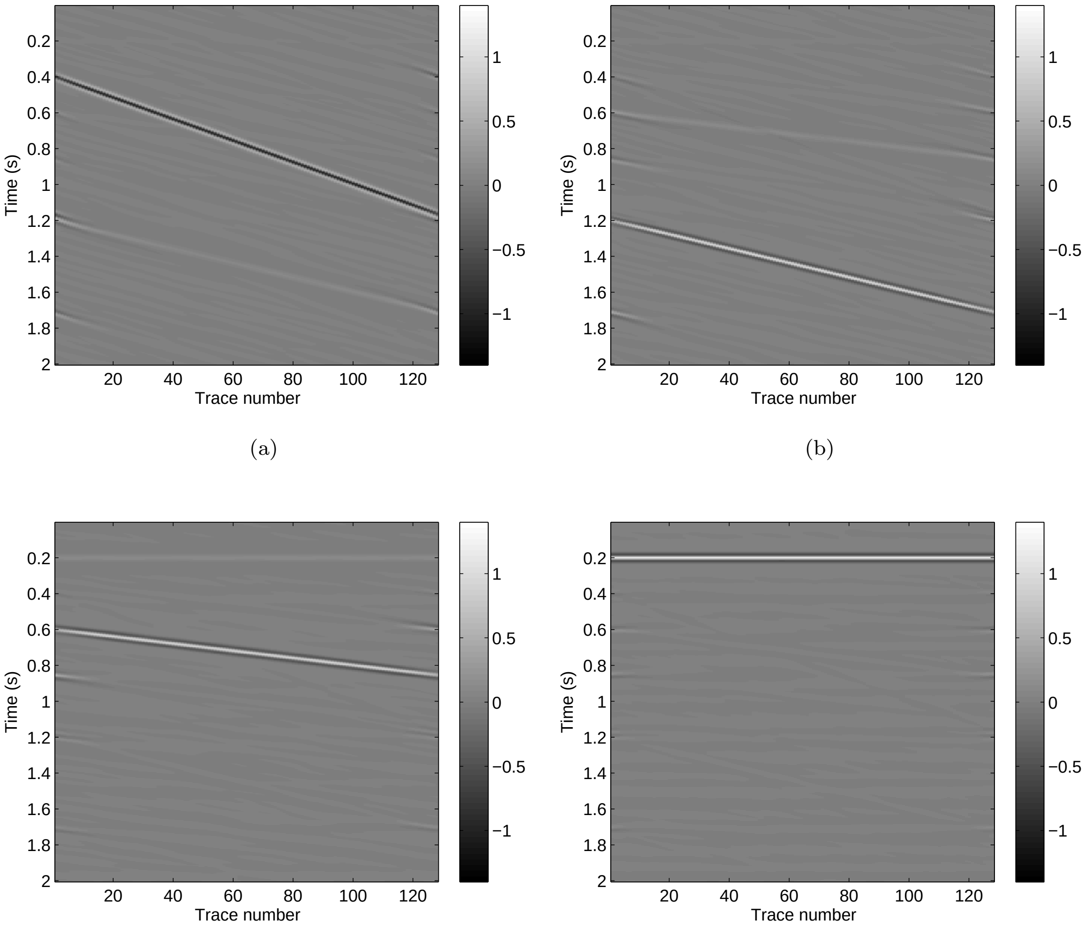


Figure 9. Event separation with EMD. (a) Original data. (b)-(e) Four modes corresponding to IMF1-IMF3 and residual.



(a) (b)

(c) (d)

Figure 10. Event separation with VMD. Four events are separated successfully.

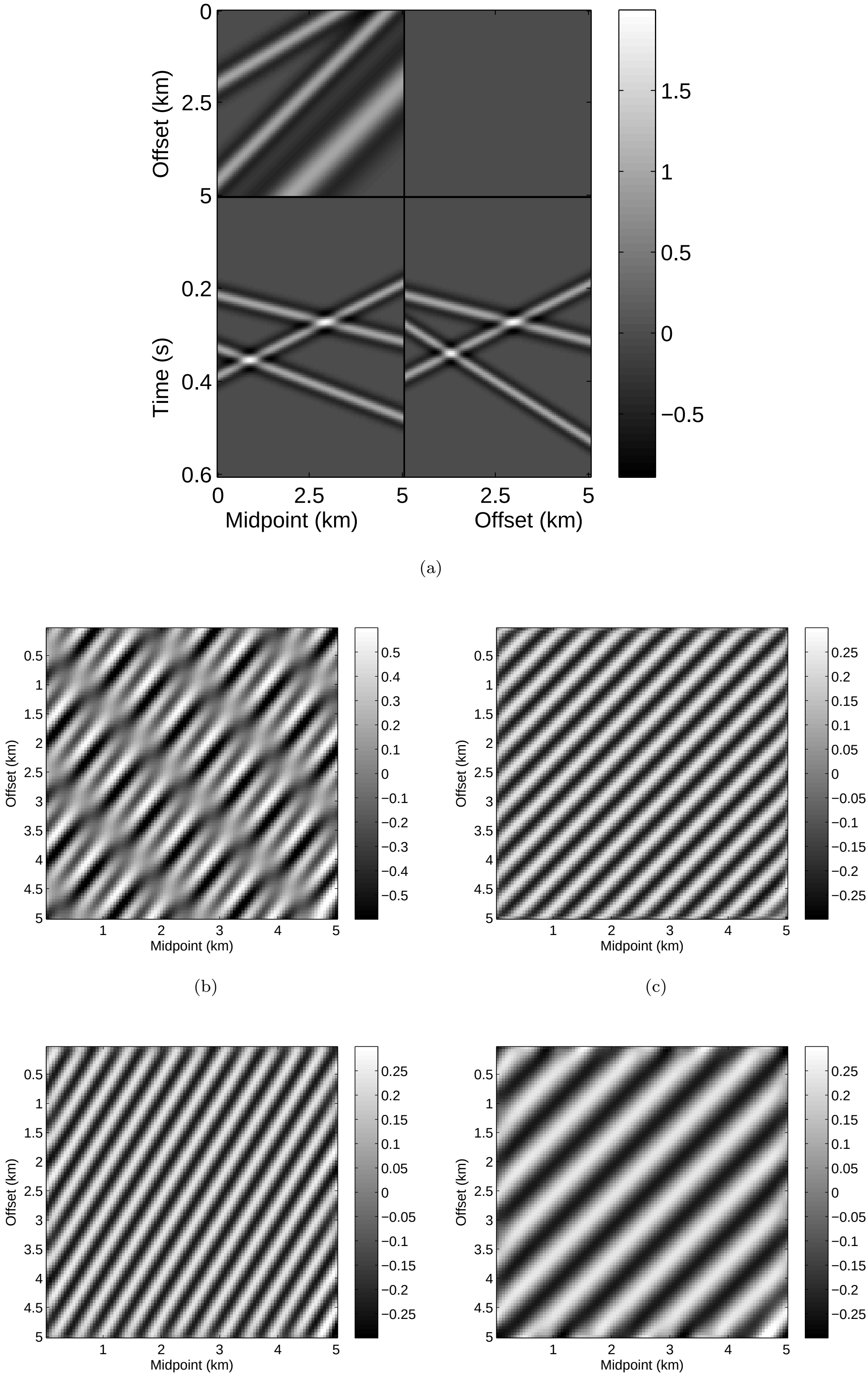


Figure 11. 2D VMD on one slice of the f-midpoint-offset spectrum of a 3D seismic data. (a) 3D synthetic seismic data with three plane waves. (b) The $30th$ slice of the f-midpoint-offset spectrum (a). (c)-(e) Three decomposed modes of (b).

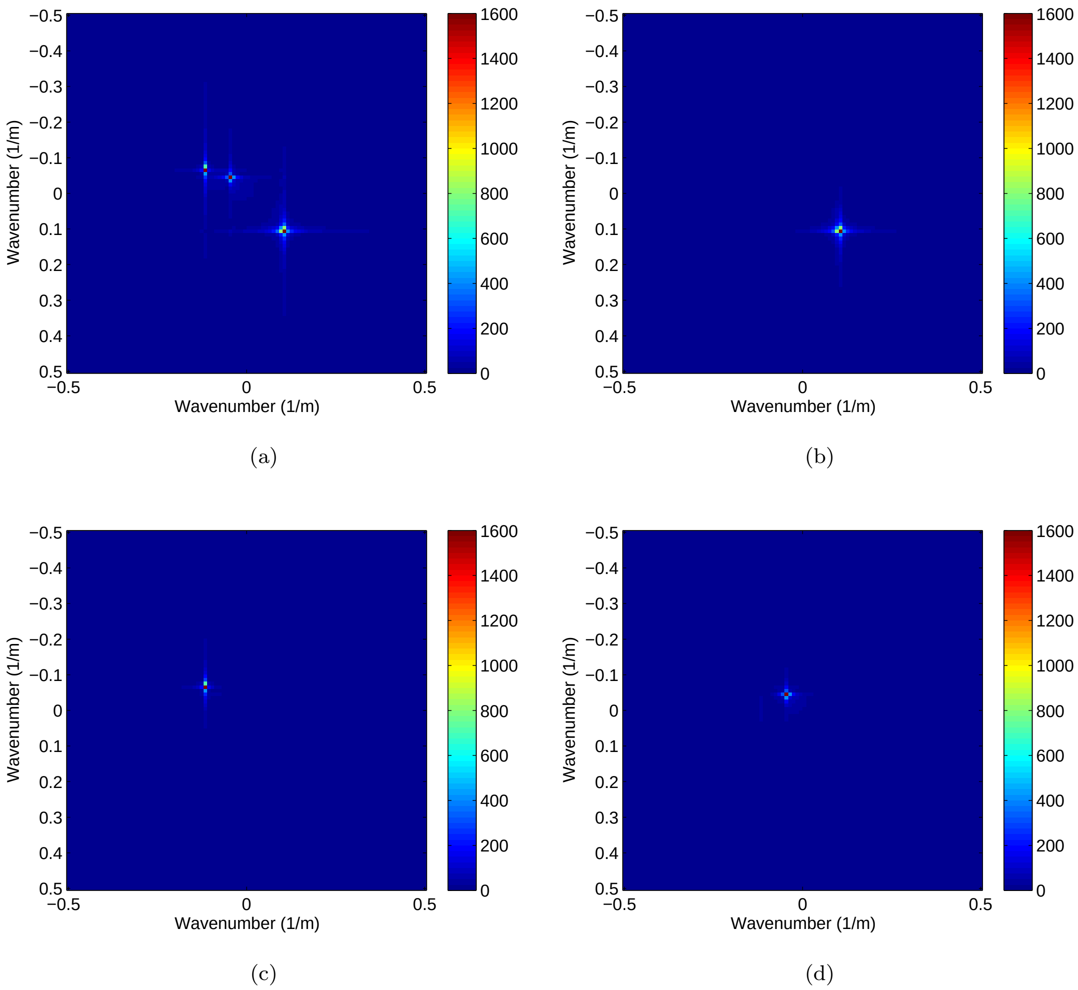


Figure 12. $f - k_m - k_h$ spectrum of the modes in Figure 11(b)-11(e)

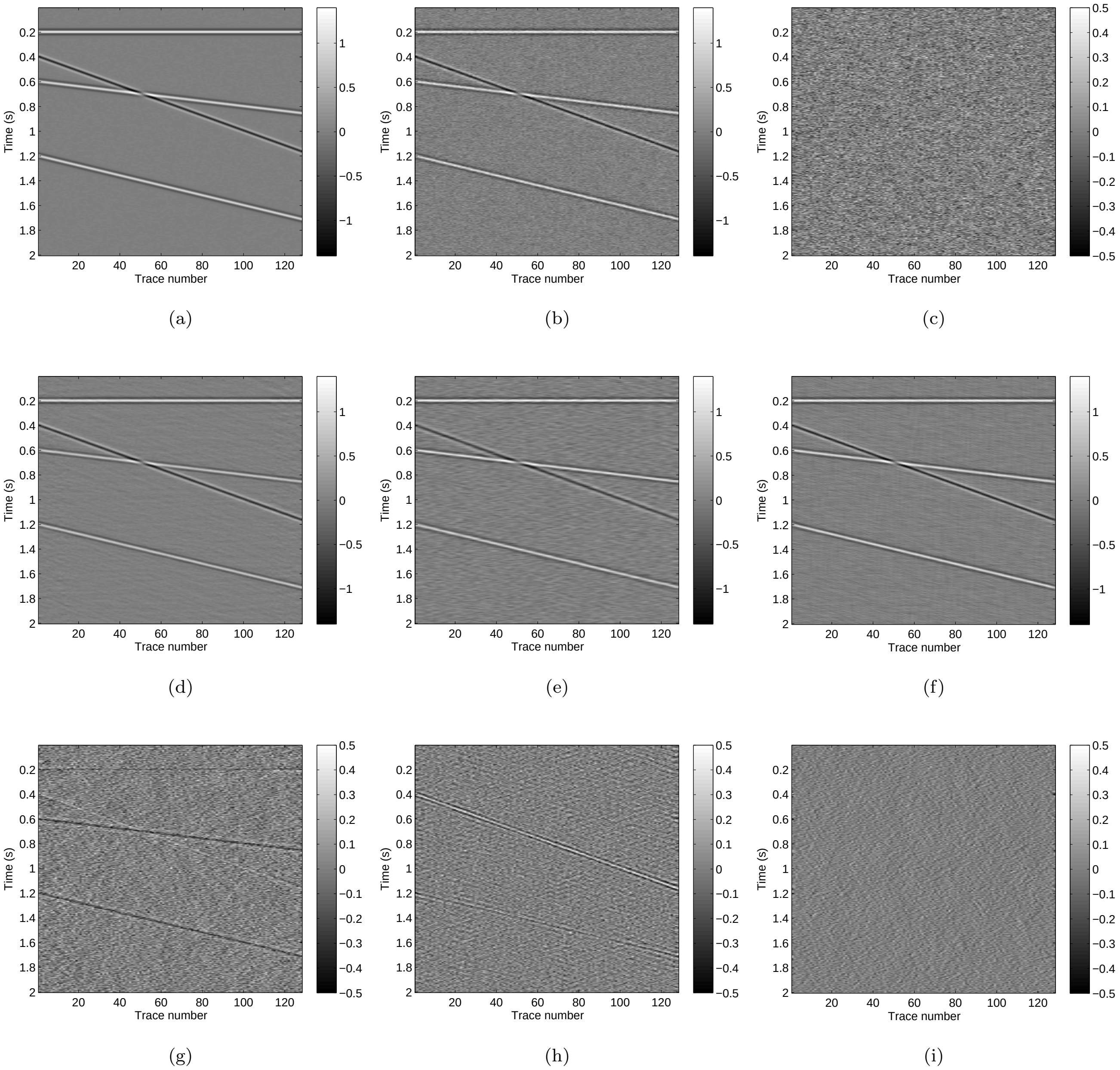


Figure 13. Noise attenuation of a synthetic data containing four linear events. (a)-(c) Synthetic data, noisy data, added noise. (d)-(f) Denoising results of $f-x$ deconvolution, EMD, VMD methods. (g)-(i) Differences between denoised results and noisy data corresponding to (d)-(f).

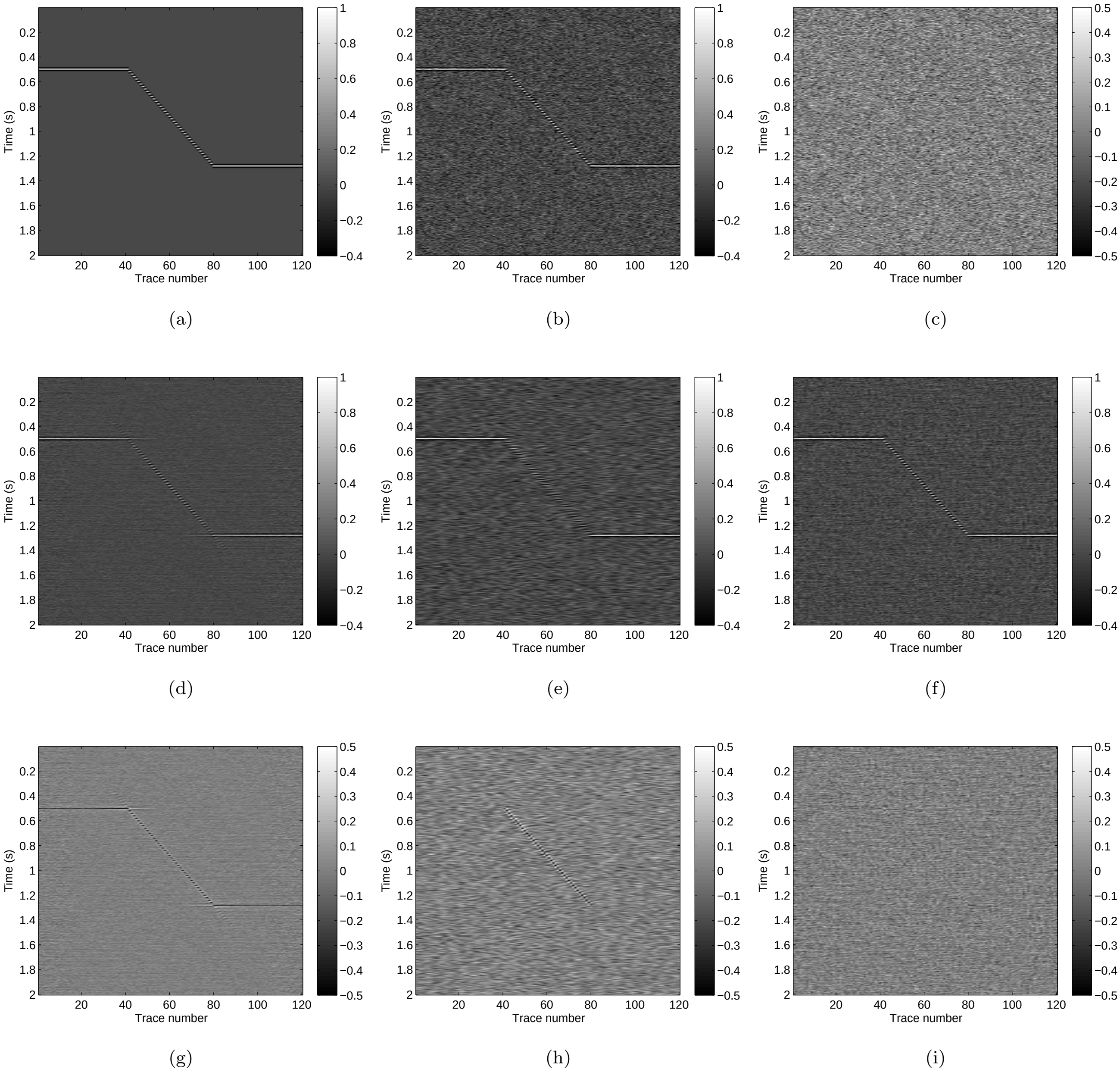
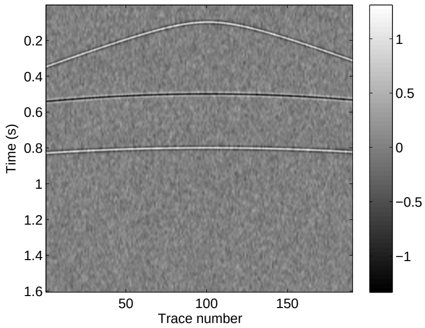
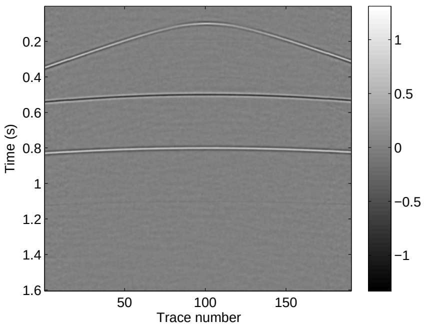


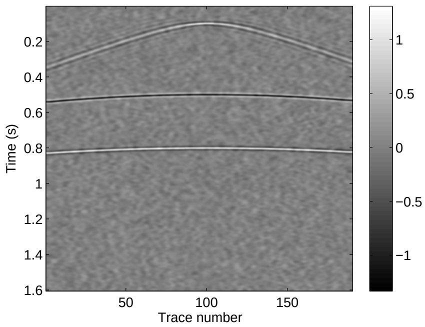
Figure 14. Noise attenuation of a synthetic data containing three adjoint linear events. (a)-(c) Synthetic data, noisy data, added noise. (d)-(f) Denoising results of $f-x$ deconvolution, EMD, VMD methods. (g)-(i) Differences between denoised results and noisy data corresponding to (d)-(f).



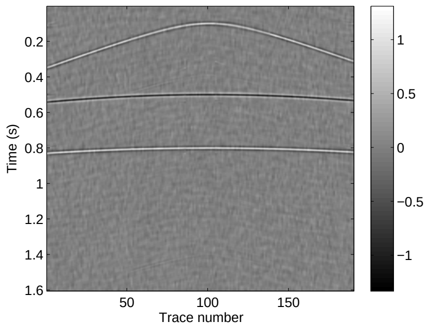
(a)



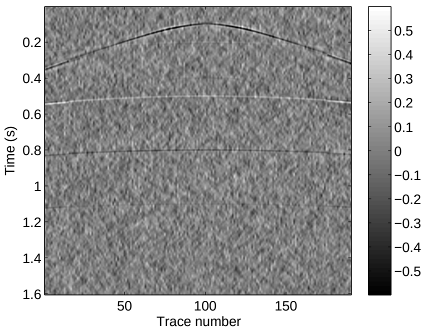
(b)



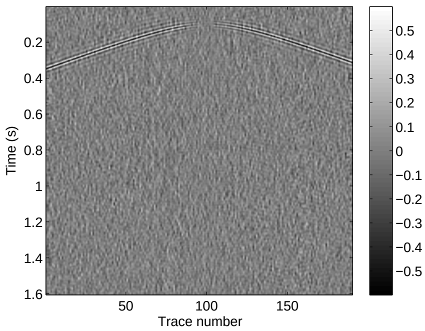
(c)



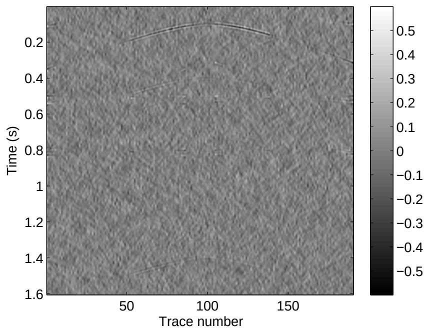
(d)



(e)



(f)



(g)

Figure 15. Noise attenuation of a synthetic noise data containing hyperbolic events. (a) Synthetic noisy model. (b)-(d) Denoising results of $f-x$ deconvolution, EMD, VMD methods. (e)-(g) Differences between denoised results and noisy data corresponding to (b)-(d).

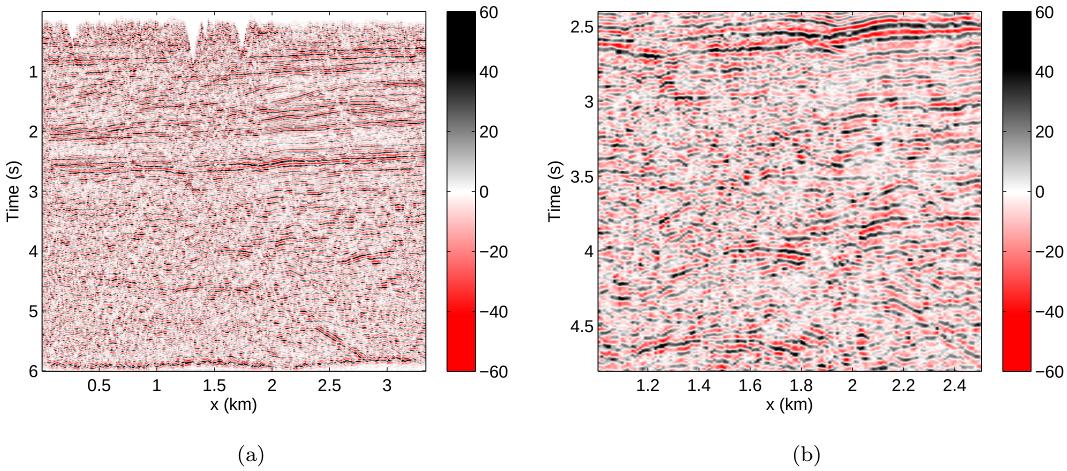


Figure 16. (a) Field data. (b) Zoomed version of (a).

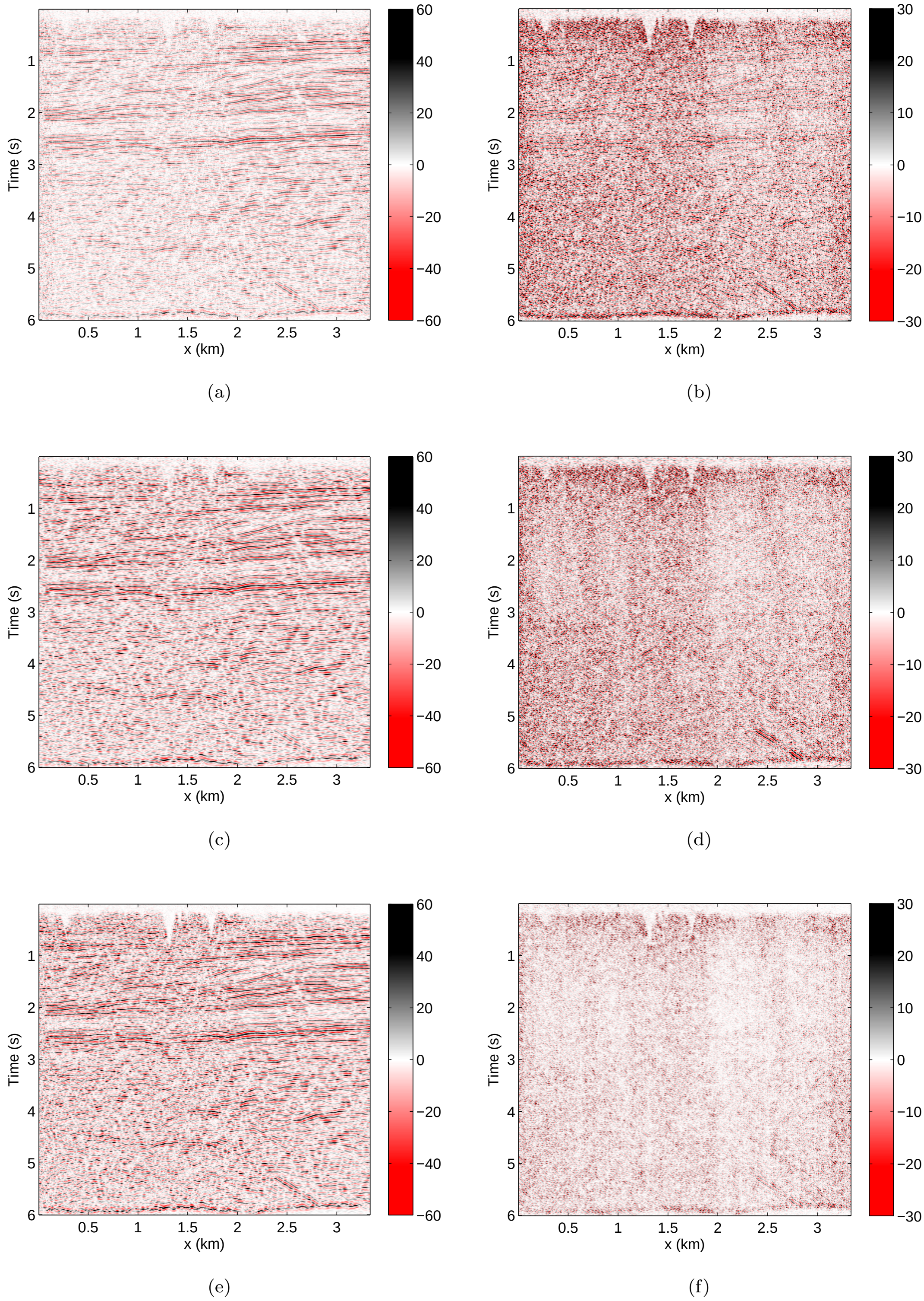


Figure 17. Field data noise attenuation. (a)(c)(e) Denoising results of $f-x$ deconvolution, EMD, VMD methods. (b)(d)(f) Differences between denoised results and noisy data corresponding to (a)(c)(e).

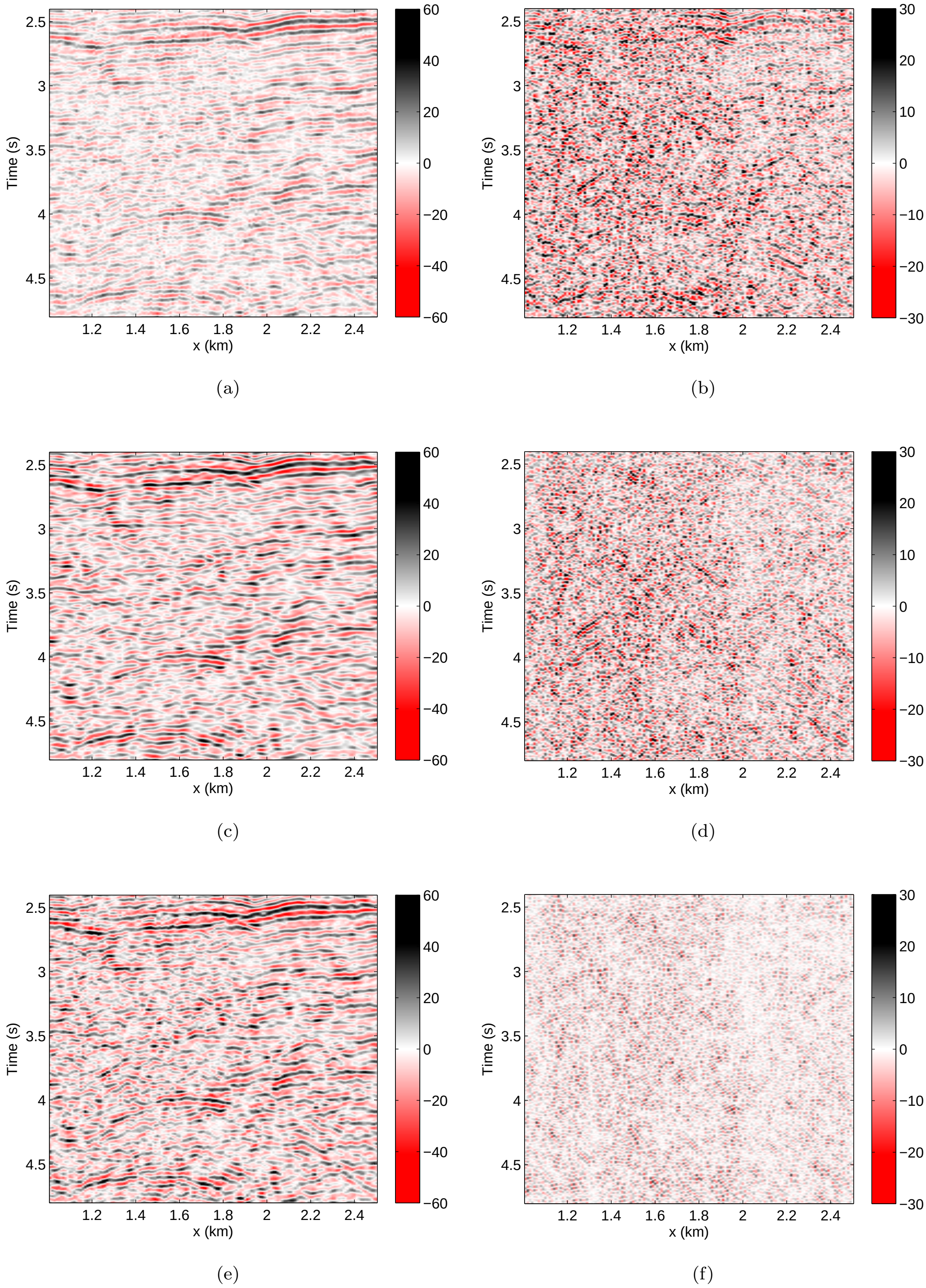


Figure 18. Zoomed version of Figure 17.

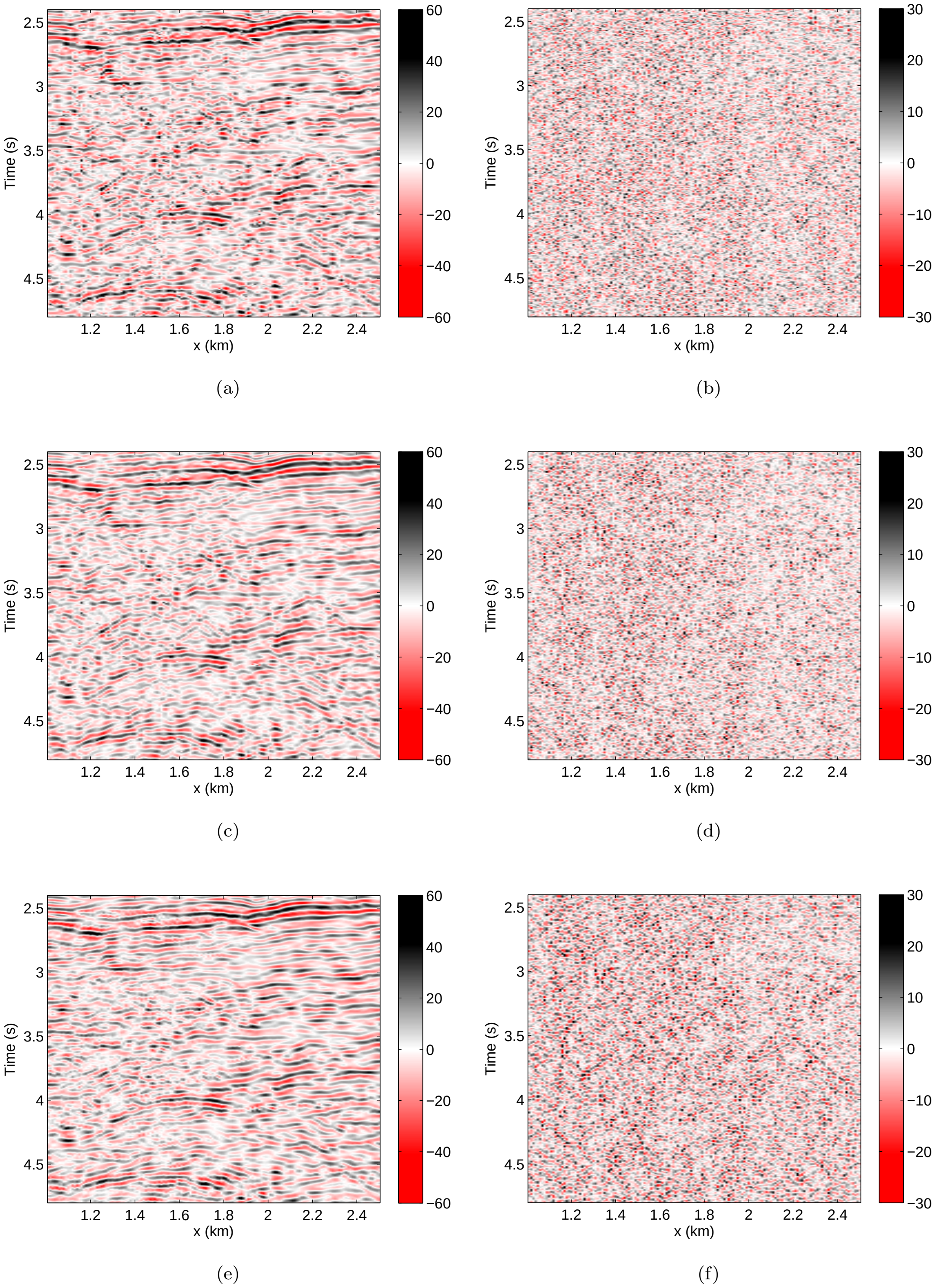
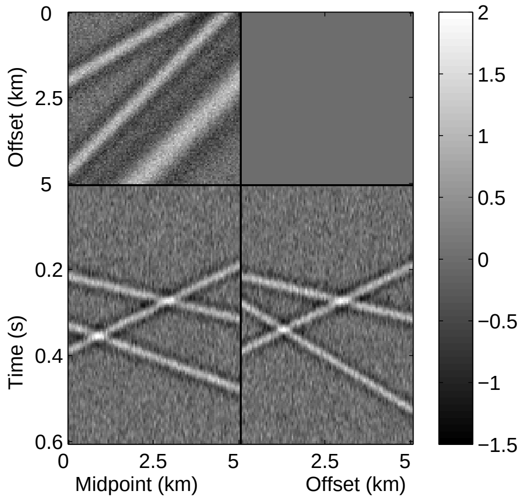
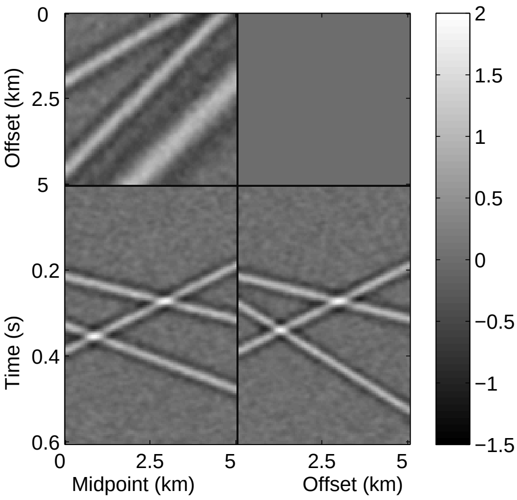


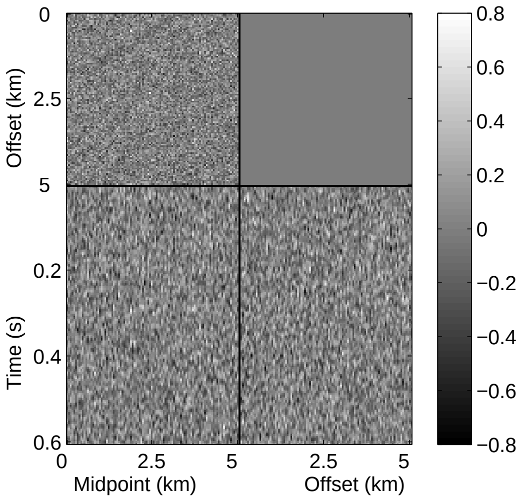
Figure 19. Zoomed version of denoising results of (a) Db6 wavelet method, (c) curvelet method, (e) empirical curvelet transform (ECT). (b) (d) and (f) are their the corresponding noise sections.



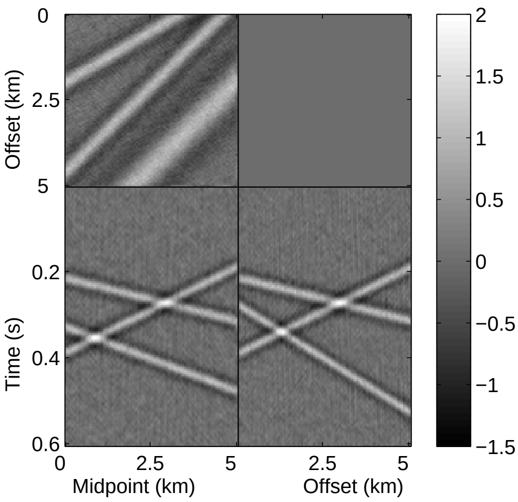
(a)



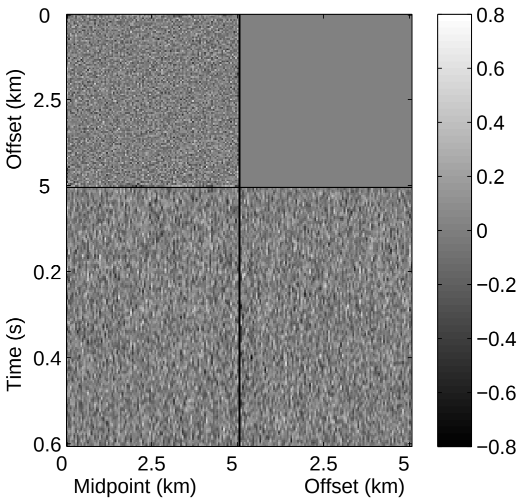
(b)



(c)



(d)



(e)

Figure 20. 3D seismic data noise attenuation. (a) The noisy 3D seismic model. (b)(d) Denoising results of EMD and VMD methods. (c)(e) Differences between denoised data and noisy data corresponding to (b)(d).

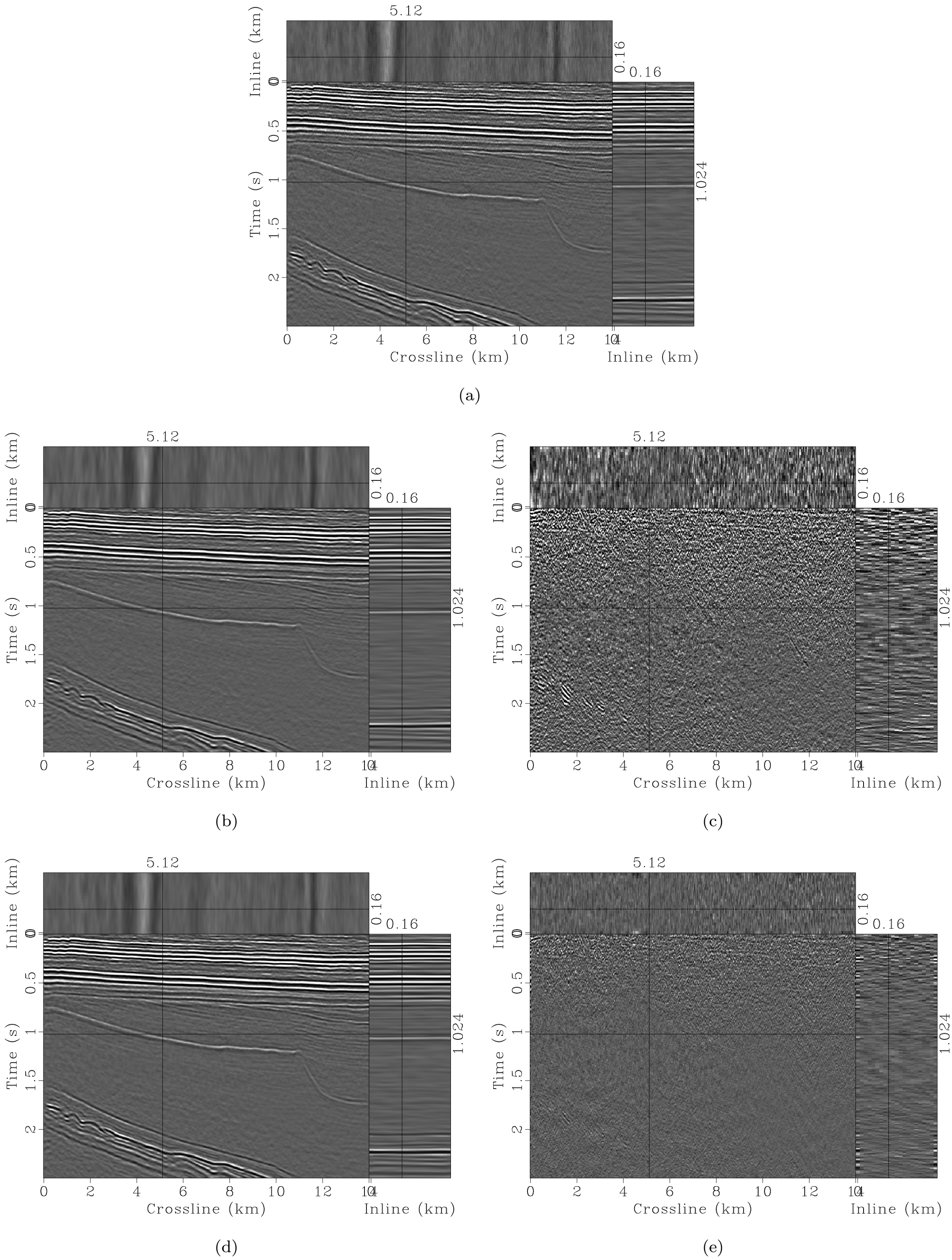


Figure 21. 3D field seismic data noise attenuation. (a) The noisy 3D seismic data. (b)(d) Denoising results of EMD, VMD methods. (c)(e) Differences between denoised data and noisy data corresponding to (b)(d).

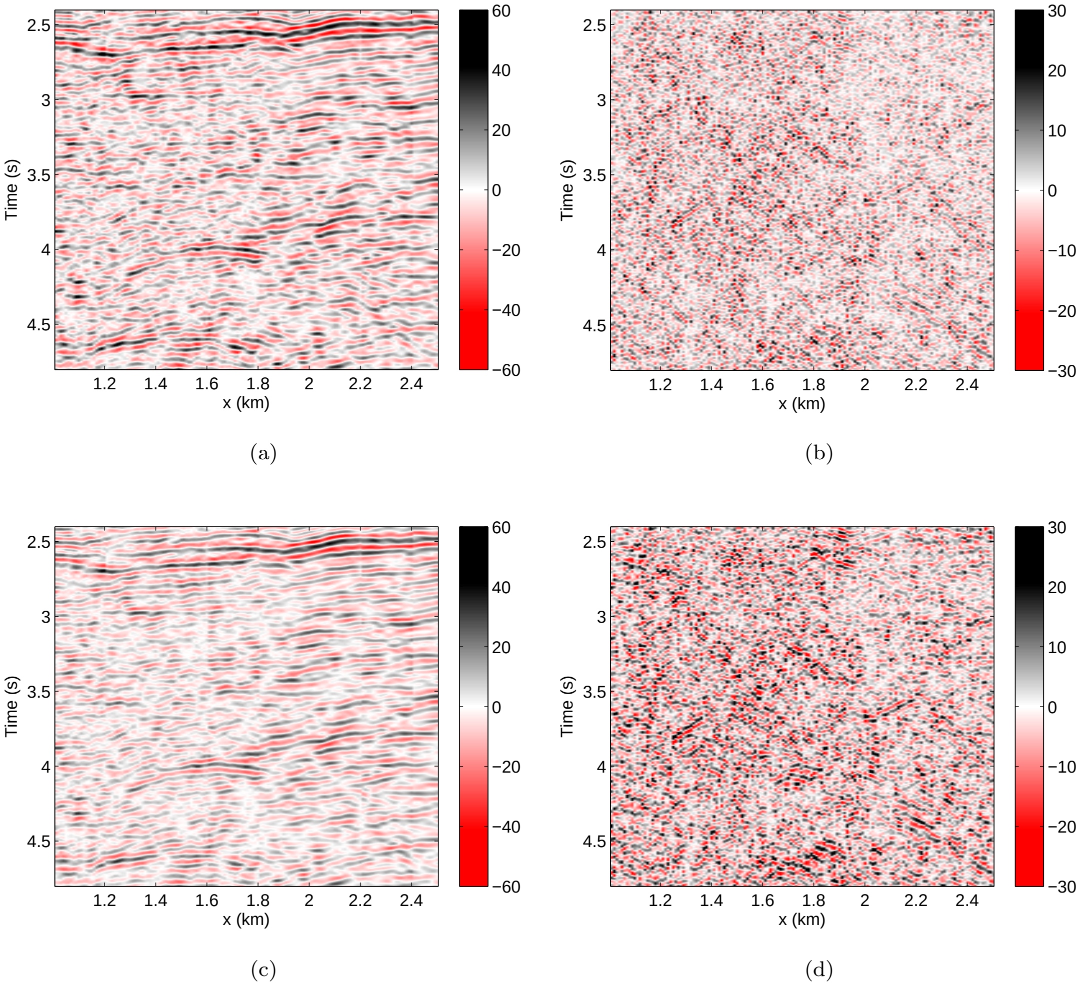


Figure 22. Denoised results of field data with different α. The zoomed denoised results with $\alpha = 20000$ in (a) and $\alpha = 100000$ in (c). The corresponding residuals are shown in (b)(d).