Geometry Mode Decomposition

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*

ABSTRACT

We propose a new decomposition algorithm for seismic data based on a band-limited priori knowledge on the Fourier or Radon spectrum. This decomposition is called geometry mode decomposition (GMD), as it decomposes a 2D signal into components consisting of linear or parabolic features. Rather than using a predefined frame, GMD adaptively obtains the geometry parameters in the data, such as the dominant slope or curvature. GMD is solved by alternatively pursuing the geometry parameters and the corresponding modes in the Fourier or Radon domain. The geometry parameters are obtained from the weighted center of the corresponding mode's energy spectrum. The mode is obtained by applying a Wiener filter, the design of which is based on a certain band-limited property. We applied GMD to seismic events splitting, noise attenuation, interpolation, and demultiple. The results show that our method is a promising adaptive tool for seismic signal processing, in comparisons with the Fourier and curvelet transforms, empirical mode decomposition (EMD) and variational mode decomposition (VMD) methods.

INTRODUCTION

Many image processing tasks take advantage of a priori knowledge of sparsity. The discrete cosine transform and wavelet transform were the transforms most frequently selected as sparse transforms (Do and Vetterli, 2003). The 1D wavelet transform is a multi-scale transform and shows a good performance for representing piece-wise smooth 1D signals. The 2D tensor wavelet transform also can represent 2D images with dot-like features sparsely. However, in fact, 2D images usually contain edge-like features. Candès and Donoho (1999) proposed a new system of representations, named ridgelets, which effectively handle line singularities in 2D, and followed by a large number of studies on multi-scale and multi-directional wavelet transforms, such as the curvelet (Starck et al., 2002; Ma and Plonka, 2010), shearlet (Easley et al., 2008), contourlet (Do and Vetterli, 2005), and surfacelet transform (Lu and Do, 2007).

The sparse transforms mentioned above are all predefined, which does not ensure that the basis are the optimal ones for representing an image in the sparsity manner. There are two techniques for seeking the optimal basis: dictionary learning in the data space and mode decomposition in the frequency space.

Dictionary learning methods that take advantage of the self-similarities inside the data, among which K-SVD is the most well known, have been proposed for image processing (Aharon et al., 2006; Mairal et al., 2009; Beckouche and Ma, 2014). These methods first decompose the 2D data into overlapped patches, which are used as samples to train the dictionary. The dictionary is trained in an optimally sparse representation manner and then used for image processing. Dictionary learning methods achieve improved results as compared to fixed basis methods. A new dictionary learning method, called data driven tight frame (DDTF), was proposed for image denoising (Cai et al., 2014; Bao et al., 2015). This method constructs a tight frame rather than an over-complete dictionary, and as a result, it is very computationally efficient and has been used for seismic data interpolation (Liang et al., 2014; Yu et al., 2015b).

In contrast to dictionary learning methods, mode decomposition methods are aimed to find a tight support of the data in the frequency domain. For instance, if a texture lies between two scales, it may be separated by a conventional wavelet transform. It is better to adaptively decompose the data into some 'modes'. Gilles (2013) and Gilles et al. (2014) proposed the 1D and 2D empirical wavelet transform (EWT) to explicitly build an adaptive wavelet basis to decompose a given signal into adaptive sub-bands. However, the EWT strongly depends on boundary detection methods in the Fourier domain and it may fail when two modes are overlapped. Dragomiretskiy and Zosso (2014, 2015) proposed the 1D and 2D variational mode decomposition (VMD) based on the band-limited properties of the modes. VMD is fully adaptive and robust to noise. 2D VMD is suitable for analyzing oscillation patterns, such as crystal lattices. However, like the 2D wavelet transform, 2D VMD cannot optimally represent geometry features, such as lines or parabolas. Inspired by the framework of VMD, we have designed an adaptive geometry mode decomposition for 2D data.

Mode and geometry mode

Huang et al. (1998) first proposed empirical mode decomposition (EMD) to decompose a signal into principal 'modes'. The mode here is defined as a signal whose number of extrema and the number of zero-crossings must differ at most by one. In most later works, the modes were defined as intrinsic mode functions (IMFs):

$$u_k(t) = A_k(t)\cos(\phi_k(t)), \tag{1}$$

where $\phi_k(t)$ is the phase and $\phi_k'(t) \geq 0$. $A_k(t) \geq 0$ is the envelope. $\phi_k'(t)$ and $A_k(t)$ vary considerably more slowly than $\phi_k(t)$ (Gilles, 2013). The immediate consequence of the new IMF definition is that the mode is an oscillation signal in the time domain and band-limited in the frequency domain. The extension of the definition to 2D is straightforward.

However, in the 2D situation, in addition to the oscillation patterns, we are also interested in geometry information, such as lines, parabolic features, and hyperbolic features, in particular when handling seismic data. We name the signal consisting of certain geometry information the geometry mode. We provide two descriptions of the geometry mode based on the band-limited principle.

The first description is in the Fourier domain. The frequency spectra of the lines are not band-limited in all directions; they are band-limited in one direction, but not in the perpendicular direction, as shown in Figure 1. In this figure, the signal is band-limited only in the direction of the arrow. However, such a band-limited definition in the frequency domain is not suitable for parabolic or hyperbolic features.

The second description is in the Radon domain. For parabolic/hyperbolic features, we can use a parabolic/hyperbolic Radon transform, which results in band-limited support in the Radon spectrum, as shown in Figure 2. We can also define other geometry modes, provided that they can be described with finite parameters, such as polynomial-like curves.

Seismic data processing and role of mode decomposition

Seismic exploration is an efficient method for exploring underground structures, which facilitates the identification of the locations of oil reservoirs. Seismic record processing includes interpolation, noise attenuation, etc. for increasing the resolution of the migrated seismic images (Naghizadeh, 2012). The interpolation of seismic data based on sparse transforms constitutes the most popular methods, such as the Fourier (Spitz, 1991; Naghizadeh, 2012; Liu and Sacchi, 2004), Radon (Wang et al., 2009), and curvelet transform (Naghizadeh and Sacchi, 2010). Noise attenuation can be classified into random noise attenuation (Naghizadeh and Sacchi, 2012; Chen and Ma, 2014) and correlated noise attenuation, such as multiple (Trad et al., 2003; Fomel, 2008) and ground-roll attenuation (Naghizadeh and Sacchi, 2011). Multiples are caused by the multiple reflections between the intersections of the medium. Primaries and multiples are approximately hyperbolic with different curvatures, and therefore, a hyperbolic Radon transform was used for demultiple (Foster and Mosher, 1992; Liu et al., 2002). The normal moveout (NMO) technique results in seismic records that approximately consist of flat events and parabolic events, and therefore, parabolic Radon transform was also used for demultiple on NMO-corrected traces, followed by inverse NMO, to achieve higher efficiency (Kabir and Marfurt, 1999). Usually, for demultiple with a Radon transform, a manual muting method in the Radon spectrum is used to separate primaries and multiples.

Mode decomposition plays an important role in seismic data processing. EMD was used to analyze the time-frequency relationship (Tary et al., 2014), seismic reflection data (Battista et al., 2007), the instantaneous attributes of seismic data (Li et al., 2014), and for random and coherent noise attenuation (Bekara and van der Baan, 2009). In Bekara and van der Baan (2009), the authors first transformed the data into the frequency-space ($f-x$) domain by using fast Fourier transform (FFT) along the time axis and then applied EMD on each frequency slice of the $f-x$ spectrum. Next, the first IMF was treated as noise and the sum of the remaining IMFs was treated as a useful signal. The advantages of EMD over other methods are that (1) it is easily implemented and no parameter is required, and (2) it is adaptive to the data and can handle non-stationary signals. However, when used in noise attenuation, EMD attacks all the energy at high wave numbers, and therefore, fails to preserve the energy of high-dip-angle events. As an improvement, Chen and Ma (2014) proposed EMDPEF, which can preserve the energy of high-dip-angle events better than the EMD method. In our previous work (Yu et al., 2015a), we introduced VMD of seismic data and showed that it can achieve a higher resolution in the wave-number axis and a higher denoising quality than the EMD and $f-x$ deconvolution methods (Spitz, 1991). However, the relationship between different frequency slices was not considered. Under the assumption of linear events, the energy support in different frequency slices should remain on the same line. Considering this a priori information, we can significantly remove more noise while retaining useful energy.

Our work

Three characteristics of 2D VMD inspired us. First, the framework of 2D VMD allows one to explore the directional modes of the data adaptively. Second, the support of 2D IMFs in the Fourier domain is dot-like, which is designed for global oscillation patterns. Usually, practical data are combined with linear features, such as the edge of different regions, or the events in the seismic data. Third, the mode decomposition in the Fourier spectrum can also occur in the Radon spectrum.

Inspired by 2D VMD, we proposed a theoretical framework for adaptively decomposing data into different geometry modes according to a priori information assumed in a specific application. For example, in this paper, in order to represent the modes with directional and line-like geometry features, we present a geometry mode decomposition (GMD) method based on the Fourier transform (GMD-F). We assume the frequency domain is line-supported in GMD-F, unlike in 2D VMD. Therefore, we designed a new objective function and derived a new direction Wiener filter.

Another example of GMD is based on the Radon transform. The common midpoint portions (CMP) and common shot gathers (CSG) are approximately constituted by a combination of hyperbolic events. NMO-corrected traces are approximately constituted by a combination flat events and parabolic events. The NMO-parabolic approximation is computationally more efficient, and therefore, we focus on the parabolic situation in this paper. We designed a second version of GMD based on the parabolic Radon spectrum, denoted by GMD-R. We applied GMD for seismic data event separation, denoising, interpolation, and adaptive demultiple, and achieved results that are better than those of the $f-x$ deconvolution method.

The rest of this paper is arranged as follows. In the second part, we introduce the theory of VMD and GMD. In the third part, we derive the new Wiener filter and provide the algorithm for solving GMD. In the fourth part, we present some applications of GMD to seismic data. Finally, we conclude this paper and give possible directions of future work.

THEORY

Variational mode decomposition

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

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

The optimization problem (2) can be physically described as follows. (1) 1D data f are decomposed as a combination of the band-limited modes u_k. (2) If u_k is real-valued, it should be transformed to an analytic signal for a single-sided spectrum. (3) The center frequency of $(\delta(x) + \frac{j}{\pi t}) * u_k(x)$ is shifted to zero frequency by the term $e^{-j\omega_k x}$. (4) The shifted signal should be smooth along the x-axis, and therefore, we minimize the norm of the derivation along the x direction.

Dragomiretskiy and Zosso (2015) proposed 2D VMD for real-valued data

$$\min_{\{u_k\},\{\vec{\omega}_k\}} \{\sum_k \|\nabla[u_{AS,k}(\vec{x})e^{-j\langle\vec{\omega}_k,\vec{x}\rangle}]\|_2^2\}, \quad s.t. \sum_k u_k = f, \tag{3}$$

where $u_{AS,k}$ is the 2D analytic signal of interest, defined in the frequency domain as

$$\hat{u}_{AS,k}(\vec{\omega}) = (1 + sgn(\vec{\omega} \cdot \vec{\omega}_k))\vec{u}_k(\vec{\omega}). \tag{4}$$

Geometry mode decomposition based on the Fourier spectrum

Both 1D and 2D VMD assume the modes are band-limited. 2D VMD can be applied in situations where oscillation patterns exist, such as in crystal fracture analysis. However, there are situations where the modes are band-limited only in one direction, but not in the perpendicular direction, such as where the 2D seismic data consist of linear events. In our previous work, we addressed the 2D seismic data on each frequency slice of their $f-x$ spectrum, which is 1D band-limited, whereas in this study, we addressed 2D directly. We use the directional derivation in the objective function of GMD-F

$$\min_{\{u_k\},\{\theta_k\}} \{\sum_k \|\langle \nabla u_{AS,k}(\vec{x}), \vec{n}_{\theta_k}\rangle\|_2^2\}, \quad s.t. \sum_k u_k = f, \tag{5}$$

where θ_k is the main direction of the mode u_k in the data space and $\vec{n}_{\theta_k} = (\cos\theta_k, \sin\theta_k)$. The objective function (5) can be described by the following steps. (1) The 2D data f are decomposed as a combination of the modes u_k. (2) The mode u_k should be smooth along the direction θ_k. (3) The norm of the derivation is minimized on the direction θ_k. Step 2 can be split into two steps for consistency with the 1D situation. First, instead of shifting the center frequency of u_k to the original point, we rotate u_k by θ_k clockwise. Second, the rotated mode should be smooth on the x-axis.

Geometry mode decomposition based on the Radon spectrum

GMD-F can handle only linear features. For data with complex structure, a window method can be used for consistency with linear event assumption. As mentioned, the CMP and CSG are approximately constituted by a combination of hyperbolic events. NMO-corrected traces are approximately constitute by a combination of flat events and parabolic events. The NMO-parabolic approximation is computationally more efficient, and therefore, we focus on the parabolic situation in this paper. The definition of the parabolic Radon transform is

$$u(q,\tau) = \int_{-\infty}^{\infty} d(x, t = \tau + qx^2)dx,$$

where $d(x,t)$ is the original seismograph, x is a spatial variable such as the offset, q is the slope of the parabolic feature (for consistency, we use the term 'slope' instead of 'curvature'), and τ is the intercept time. The realization of GMD-R is introduced in the 'Algorithm' section.

GMD-R with one parameter

There exists a situation where we are more concerned with the slope parameter than the intercept parameter, because velocity is more important in seismic data processing. Now, we present an objective function inspired by VMD for GMD-R with one parameter, p, denoted by GMD-R1:

$$\min_{u_k, p_k} \sum_k \|\partial_x \mathrm{NMO}(u_k, p_k)\|_2^2, \text{ s.t. } , \sum_k u_k = f \tag{6}$$

where $\mathrm{NMO}(u, p)$ represents the NMO of seismic data u with a parameter p, which is related to velocity. The NMO of seismic data flattens the events. The objective function (6) can be described as follows. (1) The 2D data f are decomposed as a combination of the modes u_k. (2) NMO is applied to u_k. (3) u_k after NMO is applied should be smooth along the x-axis. (4) The norm of the derivation along the x-axis is minimized after NMO is applied. Note that in (6) we have only the slope parameter p_k, without the intercept parameter τ_k. As mentioned, we focus on the parabolic situation in this paper, and therefore, NMO is approximately achieved with the parabolic transform.

ALGORITHM

Problem (5) is first written in the form of an augmented Lagrangian:

$$L(u_k, \theta_k, \lambda) := \alpha \sum_k \|\langle \nabla u_{AS,k}(\vec{x}), \vec{n}_{\theta_k} \rangle\|_2^2 + \|f(\vec{x}) - \sum_k u_k(\vec{x})\|_2^2 + \langle \lambda(\vec{x}), f(\vec{x}) - \sum_k u_k(\vec{x}) \rangle. \tag{7}$$

The ADMM for GMD-F is summarized in Algorithm (1):

The minimization of (8) can be written as

$$u_k^{n+1} = \operatorname*{argmin}_{u_k \in X} \alpha \|\langle \nabla u_{AS,k}(\vec{x}), \vec{n}_{\theta_k} \rangle\|_2^2 + \|f(\vec{x}) - \sum_i u_i(\vec{x}) + \frac{\lambda(\vec{x})}{2}\|_2^2. \tag{11}$$

Algorithm 1 ADMM for GMD-F

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

repeat

  $n = n + 1$

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

    Update u_k:

$$u_k^{n+1} = arg\min_{u_k} L(\{u_{i<k}^{n+1}\}, \{u_{i\geq k}^{n}\}, \{\theta_i^n\}, \lambda^n) \tag{8}$$

  end for

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

    Update θ_k:

$$\theta_k^{n+1} = arg\min_{\theta_k} L(\{u_i^{n+1}\}, \{\theta_{i<k}^{n+1}\}, \{\theta_{i\geq k}^{n}\}, \lambda^n) \tag{9}$$

  end for

  Dual ascent:

$$\lambda^{n+1} = \lambda^n + \tau(f - \sum_k u_k^{n+1}) \tag{10}$$

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

Output: Decomposed modes.

It can be solved in the Fourier domain

$$\hat{u}_k^{n+1} = \underset{\hat{u}_k}{\operatorname{argmin}}\, \alpha \| j\langle \vec{\omega}, \vec{n}_{\theta_k}\rangle \cdot \vec{n}_{\theta_k}[(1+sgn(\vec{\omega}\cdot\vec{\omega_k}))\hat{u}_k(\vec{\omega})]\|_2^2 + \|\hat{f}(\vec{\omega}) - \sum_i \hat{u}_i(\vec{\omega}) + \frac{\hat{\lambda}(\vec{\omega})}{2}\|_2^2 \tag{12}$$

to obtain an explicit solution:

$$\hat{u}_k^{n+1}(\vec{\omega}) = \frac{\hat{f}(\vec{\omega}) - \sum_{i\neq k}\hat{u}_i(\vec{\omega}) + \frac{\hat{\lambda}(\vec{\omega})}{2}}{1+2\alpha(\vec{\omega}\cdot\vec{n}_{\theta_k})^2} \;\; \forall\vec{\omega}\in\Omega_k : \Omega_k = \{\vec{\omega}|\vec{\omega}\cdot\vec{\omega}_k \geq 0\}. \tag{13}$$

The minimization of (9) can be written as

$$\theta_k^{n+1} = \underset{\theta_k}{\operatorname{argmin}}\, \|\langle \nabla u_{AS,k}(\vec{x}), \vec{n}_{\theta_k}\rangle\|_2^2. \tag{14}$$

It is nontrivial to find the analytic solution for θ_k. As an approximation, we first solve for the weighted center of the corresponding mode's power spectrum, i.e., $\vec{\omega}_k$. Then, we approximate θ_k with the direction of $\vec{\omega}_k$. $\vec{\omega}_k$ is solved analytically by

$$\vec{\omega}_k^{n+1} = \frac{\int_{\Omega_k}\vec{\omega}|\hat{u}_k(\vec{\omega})|^2 d\vec{\omega}}{\int_{\Omega_k}|\hat{u}_k(\vec{\omega})|^2 d\vec{\omega}}. \tag{15}$$

The direction of θ_k should be perpendicular to $\vec{\omega}_k$. Therefore, $\theta_k = \operatorname{atan}(-\frac{\omega_{k,1}}{\omega_{k,2}})$.

Algorithm (1) can be summarized as a clustering problem in the Fourier spectrum, which is similar to k-means. In the first step, we randomly initialize K center frequencies. In the second, a Wiener filter (analogous to the distance criteria in k-means) is used to cluster the modes. In the third, the new center frequencies (directions) are calculated by the weighted center of the power spectrum of the new modes. The final two steps are repeated until some convergence criteria are met. The algorithm of EMD is given in Appendix A for comparison.

The main difference between VMD and GMD-F is the Wiener filters, as we write here

$$H_{\text{VMD}}(\vec{\omega}) = \frac{1}{1+2\alpha(\vec{\omega}-\vec{\omega}_k)^2}; \tag{16}$$

$$H_{\text{GMD-F}}(\vec{\omega}) = \frac{1}{1+2\alpha(\vec{\omega}\cdot\vec{n}_{\theta_k})^2}. \tag{17}$$

Figure 3 shows the Wiener filters used in 2D VMD and GMD-F. It is clear that 2D VMD is point-supported and GMD-F is line-supported. The role of parameter α is also clear. When the value of α is small, the mode support becomes wider and therefore is more tolerant to noise. This is consistent with the optimization function. A small value of α means less regularity, and therefore, the result is more tolerant to noise.

Figure 4 shows GMD-F for synthetic seismic data consisting of three linear events, shown in Figure 4(a). Its frequency spectrum is shown in Figure 4(e). Figure 4(e) also shows the evolution of the center frequencies of the three modes. Here, we use the center frequencies to indicate the main directions. Figure 4(b) - Figure 4(d) show the three decomposed modes. Note the results show an undesired boundary effect due to the truncation in the Fourier domain. Figure 4(f) - Figure 4(h) show the Fourier spectra corresponding to Figure 4(b) - Figure 4(d).

We tested the convergence of GMD-F for the data in Figure 4(a) with a varying initialization of the center frequencies ω_k. In Figure 5, the x-axis represents ω_x and the y-axis the iterations in logarithmic scale. The different lines represent the evolution of ω_x of the different modes. One hundred experiments were performed. As we can see, in a noise-free situation, all the randomly initialized center frequencies in our test converge to the desired points within 1000 iterations. If the data are corrupted by weak noise, our algorithm still converges. However, when the noise energy becomes sufficiently strong, the convergence success ratio decreases. Readers can refer to (Dragomiretskiy and Zosso, 2014) for similar results.

Same as Algorithm (1), GMD-R can be summarized as a clustering problem on the Radon spectrum. In the first step, we randomly initialize K of (τ, p) pairs. In the second, a Wiener filter is used to cluster the modes in the Radon spectrum. In the third, the new (τ, p) pairs are calculated by the weighted center of the new modes' power spectrum. The final two steps are repeated until some convergence criteria are met.

Figure 6 shows GMD-R for synthetic seismic data consisting of three parabolic events, shown in Figure 6(a). Its Radon spectra is shown in Figure 6(e). Figure 6(e) also shows the evolution of the (τ, p) centers of the three modes. Figure 6(b) - Figure 6(d) show the three decomposed modes. Figure 6(f) - Figure 6(h) show the Radon spectra corresponding to Figure 6(b)- Figure 6(d). In order to test the resolution of GMD-R, we applied it to synthetic data with three parabolic events, the slopes of which are similar to each other, as shown in Figure 7(a). Figure 7(b)-Figure 7(d) show that the three events are separated successfully.

The solution of the optimization problem (6) is different from that in the previous GMD-R in that the Wiener filter is a function only of p, not the pair of (p, τ). GMD-R1 is more suitable for data containing many events with only a few different slopes but many different intercepts. GMD-R1 helps reduce the number of the modes needed significantly and increases the efficiency of GMD-R. Figure 8 shows GMD-R1 for the data in Figure 7(a). We show two decomposed modes in Figure 8(a) and Figure 8(b). The mode in Figure 8(a) contains two events with similar slopes.

APPLICATIONS OF GMD

In this section, we first present some applications of GMD to synthetic data. GMD-F is used for denoising seismic data consisting of linear events, as compared with the VMD method and the $f-x$ deconvolution method. GMD-R is used for simultaneous denoising and interpolation of seismic data consisting of parabolic events, as compared with the Spitz method. Then, we use GMD-F to denoise field data and compare this method with the $f-x$ deconvolution method and curvelet method. For field data with a complex structure, we operate on small temporal and spatial windows based on the assumption of linear events. Finally, GMD-R1 is used for demultiple, as compared with the directly muting method.

Figure 9 shows the results of testing GMD-F for seismic noise attenuation. For denoising purposes, the parameter λ is set to zero. Figure 9(a) and Figure 9(b) show the noise corrupted data (SNR = -1.61) and their $f-k$ spectra. The data consist of three linear events, each of which represents an interface of underground. Note that the energy of one event is weaker than that of the other two. Figure 9(c) - Figure 9(e) show the denoised results of the GMD-F, 1D VMD, and $f-x$ deconvolution methods. It is clear that the $f-x$ deconvolution method removes the useful energy of all three events. The 1D VMD method cannot preserve the energy of the weak event. The GMD-F method can preserve the energy more effectively. Figure 9(i) - Figure 9(k) show the corresponding $f-k$ spectra. The Wiener filter in GMD-F remains on a line, and therefore, keeps the energy on a line. However, the noise in high frequency is retained, and therefore, we truncate high frequency energy to zero. 1D VMD handles the frequency slice separately, and therefore, weak energy may be merged by noise.

Figure 10 shows the results of testing GMD-R for seismic anti-aliased interpolation. The data in Figure 7(a) are regularly sub-sampled by 1/4 and contaminated with random noise, shown in Figure 10(a). Its $f-k$ spectra is shown in Figure 10(b). Sub-sampled seismic data is caused by environmental or economic reasons, which restrict the distribution of the receivers. Interpolation of seismic data is essential for improving the resolution of migration and inversion. Figure 10(c) and Figure 10(d) show the interpolated data and the $f-k$ spectra with GMD-R. The achieved results are considerably better than those of the Spitz method (Spitz, 1991), shown in Figure 10(e).

Figure 11(a) shows noisy field data, with 2001 time samples and 776 traces. The time and space sample intervals are 4 ms and 0.01 km, respectively. A zoomed version (0.40 s-1.80 s, 3.51 km-6.00 km) is shown in Figure 11(b). The dip events in the data show the dip structures of underground, which is important for locating the oil reservoirs and should be well preserved while denoising. Figure 12 illustrates the zoomed versions of the denoised results of the GMD-F, curvelet, and $f-x$ deconvolution methods with the corresponding noise sections. The GMD-F method significantly preserves more useful energy than the $f-x$ deconvolution method by taking advantage of the line-supported priori knowledge. The curvelet method tends to over-smooth the denoising result.

Figure 13 shows a demultiple test based on GMD-R1. Multiples are caused by the multiple reflections between the intersections of the medium and may lead to low resolution of migration or inversion. Multiple suppression is an important processing step for marine data seismic data, especially in shallow water environment (Fan et al., 2011). Figure 13(a) shows the NMO-corrected traces. Figure 13(b) shows the parabolic Radon spectrum. The primary and the multiple are located on different slopes and are detected with GMD-R1 adaptively, indicated by the blue and the red line, respectively. Figures 13(c) and 13(d) show the separated multiple and primary with GMD-R1, with $\alpha = 0.005$. Our method can achieve demultiple and random noise attenuation simultaneously, as a Wiener filter is applied in the multiple and primary modes. Figures 14(a) and 14(b) show the demultiple results with GMD-R1, with $\alpha = 10^{-5}$. A small value of α makes the result less smooth, but the primary is still contaminated with notable multiples. Figures 14(c) and 14(d) show the separated multiple and primary when the directly muting method is applied. We manually mute a region of the Radon spectrum to zero to obtain the multiple. The primary is obtained by subtracting the multiple from the original data. As we can see, Figure 14(c) still shows some flat events and Figure 14(d) is contaminated with both notable multiple and random noise.

DISCUSSION

The relationship between GMD-F and 2D VMD

In equation (16) and (17), we show the difference between GMD-F and 2D VMD. In this section, we apply Fourier transform on the terms inside the norm in equation (3) and (5) to explore the relationship between GMD-F and 2D VMD. In equation (5), the norm is:

$$\begin{aligned} & \|\langle \nabla u(\vec{x}), \vec{n}_{\theta_k} \rangle\|_2^2 \\ = \quad & \|j\langle \vec{\omega}, \vec{n}_{\theta_k} \rangle \cdot \vec{n}_{\theta_k} \hat{u}_k(\vec{\omega})\|_2^2 \end{aligned}$$

and in equation (3), the norm is:

$$\begin{aligned} & \|\nabla[u_k(\vec{x})e^{-j\langle\vec{\omega}_k,\vec{x}\rangle}]\|_2^2 \\ = \ & \|j(\vec{\omega}-\vec{\omega}_k)\hat{u}_k(\vec{\omega})\|_2^2 \end{aligned}$$

If we set:

$$\vec{\omega}-\vec{\omega}_k = \langle\vec{\omega},\vec{n}_{\theta_k}\rangle \cdot \vec{n}_{\theta_k}$$

then GMD-F and 2D VMD are the same. Now we get:

$$\vec{\omega}_k = \langle\vec{\omega},\vec{n}_{\theta_k}^{\perp}\rangle \cdot \vec{n}_{\theta_k}^{\perp}$$

where $\vec{n}_{\theta_k}^{\perp}$ is a vector perpendicular to $\vec{n}_{\theta_k}$. So in the Fourier domain, GMD-F is a special case of 2D VMD when setting $\vec{\omega}_k = \langle\vec{\omega},\vec{n}_{\theta_k}^{\perp}\rangle \cdot \vec{n}_{\theta_k}^{\perp}$. Figure 15 shows this relationship in the Fourier domain. $\vec{\omega}$, $\langle\vec{\omega},\vec{n}_{\theta_k}\rangle \cdot \vec{n}_{\theta_k}$ and $\vec{\omega}_k$ are the three sides of a right triangle.

CONCLUSION

We proposed a new decomposition algorithm for seismic data based on a band-limited priori knowledge on the Fourier or Radon spectrum. Different from the existing predefined multi-directional frames, GMD allows adaptive identification of the directional or other dominant geometry information in the data themselves. Our method is also different from EMD and VMD in that EMD is based on an oscillated and symmetric mode definition, VMD is based on a point-supported assumption in the Fourier domain and GMD is based on a line-supported assumption in the Fourier domain or point-supported assumption in the Radon domain. We applied GMD to seismic event splitting, denoising, interpolation, and demultiple. The results show that our method is a promising adaptive tool for seismic signal processing. GMD provides a framework for feature separation, for example, the hyperbolic Radon transform can also be integrated. Future work will focus on the theoretical fundamental of the GMD-R algorithm and applications of GMD involving field seismic data.

APPENDIX A

EMD was designed to analyze a non-stationary signal, by decomposing the signal into different 'modes' of oscillations, named intrinsic mode functions (IMFs). An IMF satisfies two conditions: (1) the number of extrema and the number of zero crossings must be equal or differ by at most 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). IMFs are extracted recursively by using a sifting algorithm:

Algorithm 2 EMD Algorithm

Input: The signal $f(t)$, $k = 1$

1: **repeat**
2: The sifting process:
3: **repeat**
4: Find the local maxima and minima of $f(t)$.
5: Fit these local maxima and minima by cubic spline interpolation in turn to generate the upper and lower envelopes.
6: The mean of the upper and lower envelopes is calculated and subtracted from the initial data and the same interpolation scheme is reiterated on the remainder.
7: **until** The mean envelope is reasonably close to zero everywhere
8: The resultant signal is designated as the kth IMF.
9: Subtract the kth IMF from $f(t)$ and set the difference as new $f(t)$.
10: $k = k + 1$.
11: **until** the last IMF has a small amplitude or becomes monotonic.

Output: Decomposed IMFs.

REFERENCES

Aharon, M., M. Elad, and A. Bruckstein, 2006, K-svd: An algorithm for designing overcomplete dictionaries for sparse representation: Signal Processing, IEEE Transactions on, **54**, 4311–4322.

Bao, C., H. Ji, and Z. Shen, 2015, Convergence analysis for iterative data-driven tight frame construction scheme: Applied and Computational Harmonic Analysis, **38**, 510–523.

Battista, B. M., C. Knapp, T. McGee, and V. Goebel, 2007, Application of the empirical mode decomposition and hilbert-huang transform to seismic reflection data: Geophysics, **72**, H29–H37.

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

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

Cai, J.-F., H. Ji, Z. Shen, and G.-B. Ye, 2014, Data-driven tight frame construction and image denoising: Applied and Computational Harmonic Analysis, **37**, 89–105.

Candès, E. J., and D. L. Donoho, 1999, Ridgelets: A key to higher-dimensional intermittency?: Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, **357**, 2495–2509.

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

Do, M. N., and M. Vetterli, 2003, The finite ridgelet transform for image representation: Image Processing, IEEE Transactions on, **12**, 16–28.

———, 2005, The contourlet transform: an efficient directional multiresolution image representation: Image Processing, IEEE Transactions on, **14**, 2091–2106.

Dragomiretskiy, K., and D. Zosso, 2014, Variational mode decomposition: IEEE Transactions on Signal Processing, **62**, 531–544.

———, 2015, Two-dimensional variational mode decomposition: Energy Minimization Methods in Computer Vision and Pattern Recognition, Springer, 197–208.

Easley, G., D. Labate, and W.-Q. Lim, 2008, Sparse directional image representations using the discrete shearlet transform: Applied and Computational Harmonic Analysis, **25**, 25–46.

Fan, W., H. Keil, V. Spieß, T. Mörz, and C. Yang, 2011, Surface related multiple elimination-application on north sea shallow seismic dataset: Presented at the 73rd EAGE Conference and Exhibition incorporating SPE EUROPEC 2011.

Fomel, S., 2008, Adaptive multiple subtraction using regularized nonstationary regression: Geophysics, **74**, V25–V33.

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

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

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

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

Kabir, M. N., and K. J. Marfurt, 1999, Toward true amplitude multiple removal: The Leading Edge, **18**, 66–73.

Li, X., W. Chen, and Y. Zhou, 2014, A robust method for analyzing the instantaneous attributes of seismic data: The instantaneous frequency estimation based on ensemble empirical mode decomposition: Journal of Applied Geophysics, **111**, 102–109.

Liang, J., J. Ma, and X. Zhang, 2014, Seismic data restoration via data-driven tight frame: Geophysics, **79**, V65–V74.

Liu, B., and M. D. Sacchi, 2004, Minimum weighted norm interpolation of seismic records: Geophysics, **69**, 1560–1568.

Liu, Y., M. D. Sacchi, et al., 2002, De-multiple via a fast least squares hyperbolic radon transform: Presented at the 2002 SEG Annual Meeting, Society of Exploration Geophysicists.

Lu, Y. M., and M. N. Do, 2007, Multidimensional directional filter banks and surfacelets: Image Processing, IEEE Transactions on, **16**, 918–931.

Ma, J., and G. Plonka, 2010, The curvelet transform: IEEE Signal Processing Magazine, **27**, 118–133.

Mairal, J., F. Bach, J. Ponce, and G. Sapiro, 2009, Online dictionary learning for sparse coding: Proceedings of the 26th annual international conference on machine learning, ACM, 689–696.

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

Naghizadeh, M., and M. Sacchi, 2011, Ground-roll elimination by scale and direction guided curvelet transform: Presented at the 73rd EAGE Conference and Exhibition incorporating SPE EUROPEC 2011.

———, 2012, Multicomponent f-x seismic random noise attenuation via vector autoregressive operators: Geophysics, **77**, V91–V99.

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

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

Starck, J.-L., E. J. Candès, and D. L. Donoho, 2002, The curvelet transform for image denoising: Image Processing, IEEE Transactions on, **11**, 670–684.

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

Trad, D., T. Ulrych, and M. Sacchi, 2003, Latest views of the sparse radon transform: Geophysics, **68**, 386–399.

Wang, J., M. Ng, M. Perz, et al., 2009, Fast high-resolution radon transforms by greedy least-squares method: Presented at the 2009 SEG Annual Meeting, Society of Exploration Geophysicists.

Yu, S., J. Ma, and S. Osher, 2015a, Variational mode decomposition of seismic data: ftp://ftp.math.ucla.edu/pub/camreport/cam15-58.pdf, **1**, 1–37.

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

LIST OF FIGURES

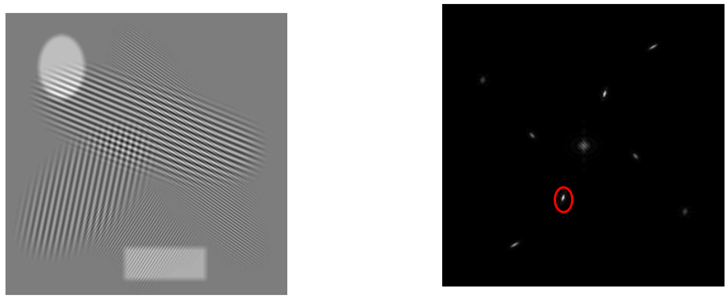
(a).

12 Field data noise attenuation with GMD-F. (a),(c), and (e) are the noise attenuation results of the GMD-F method, the curvelet method, and $f-x$ deconvolution method, respectively. (b),(d), and (f) are the corresponding noise.

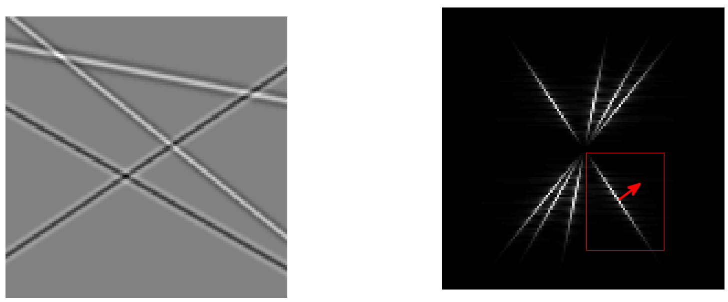
13 Demultiple on NMO-corrected traces. (a) NMO-corrected traces. (b) Parabolic Radon spectrum. The two lines represent the two modes detected. (c) and (d) The separated multiple and primary with GMD-R1. $\alpha = 0.005$.

14 Demultiple on NMO-corrected traces. (a) and (b) The separated multiple and primary with GMD-R1. $\alpha = 10^{-5}$. (c) and (d) The separated multiple and primary by directly muting the Radon spectrum.

15 The relationship between GMD and 2D VMD.

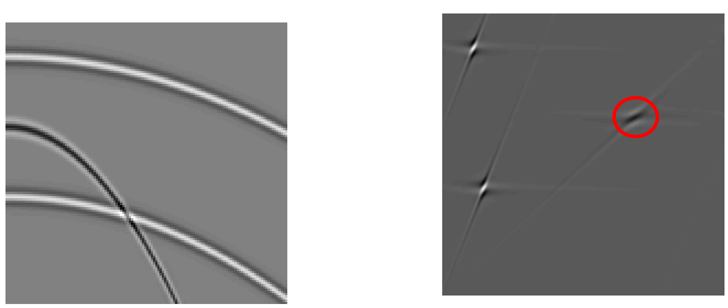


(a) (b)



(c) (d)

Figure 1: Support in the Fourier spectrum. (a) A 'texture' image. (b) Fourier spectrum of (a). The spectrum is band-limited. (c) A 'geometry' image with lines. (d) Fourier spectrum of (c). The spectrum is band-limited in the direction of the arrow.



(a) (b)

Figure 2: Support in the Radon spectrum. (a) A 'geometry' image with parabolic features. (b) Radon spectrum of (a). The spectrum is band-limited.

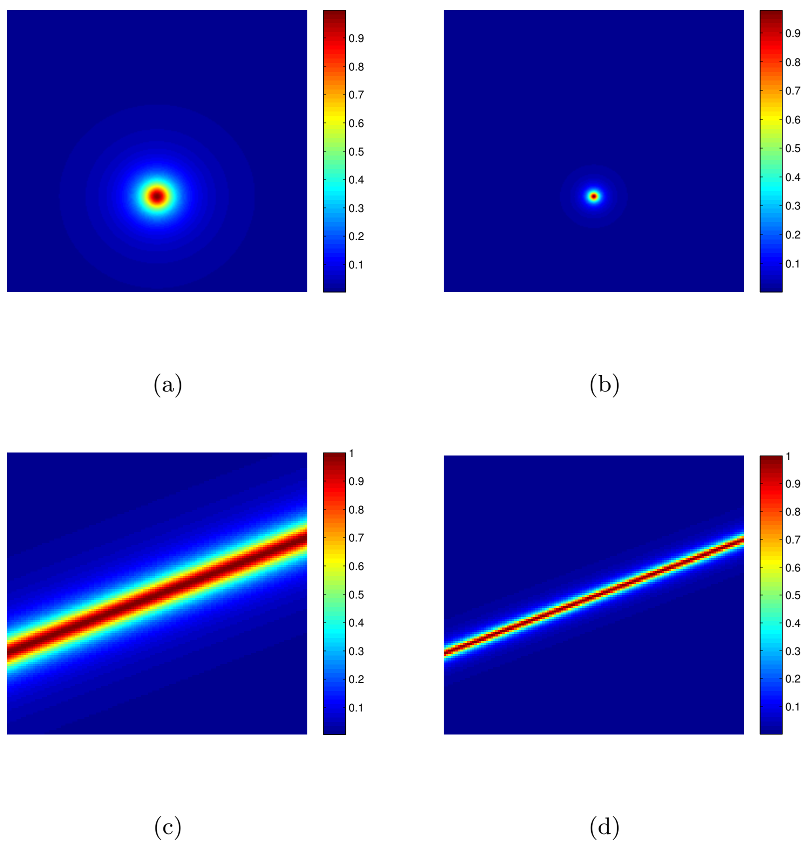
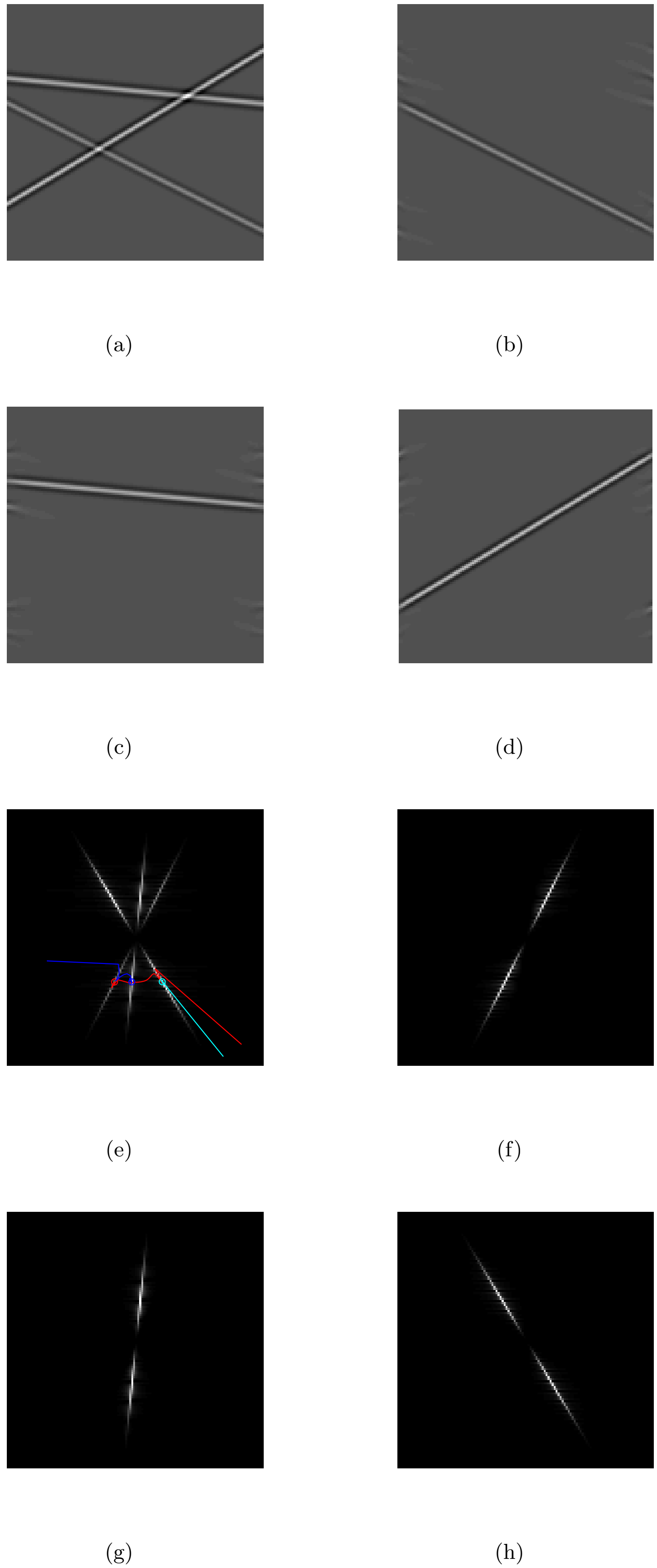


Figure 3: Wiener filter with different a priori information. (a) and (b) Wiener filter with signal a priori $1/(\vec{\omega} - \vec{\omega}_k)^2$, with $\alpha = 500$ and 5000, respectively. (c) and (d) Wiener filter with signal a priori $1/(\vec{\omega} \cdot \vec{n}_{\theta_k})^2$, with $\alpha = 500$ and 5000, respectively.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 4: GMD-F applied to a synthetic seismic model consisting of three linear events. (a) Synthetic model. (b)-(d) Three decomposed modes. (e) Fourier spectrum and the trajectory

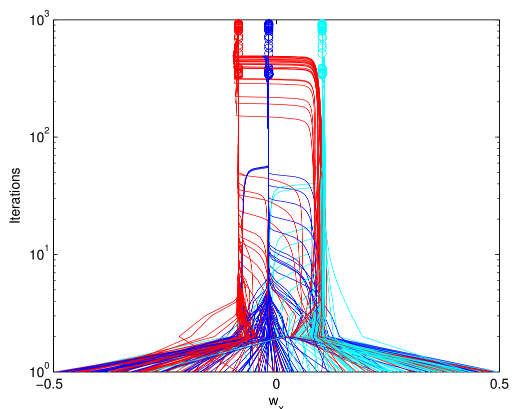
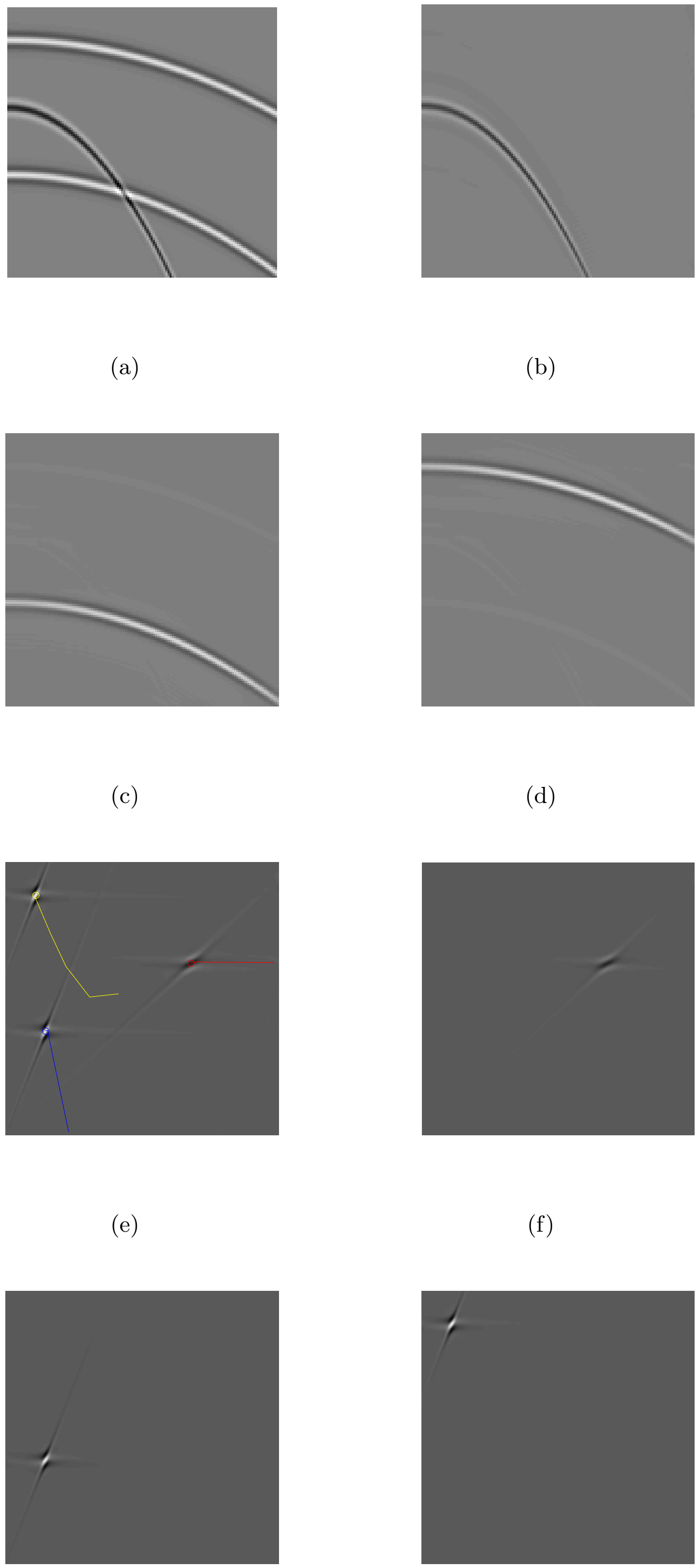


Figure 5: Convergence analysis of ω_x in GMD-F.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 6: GMD-R applied to a synthetic seismic model consisting of three parabolic events. (a) Synthetic model. (b)-(d) Three decomposed modes. (e) Radon spectrum and the

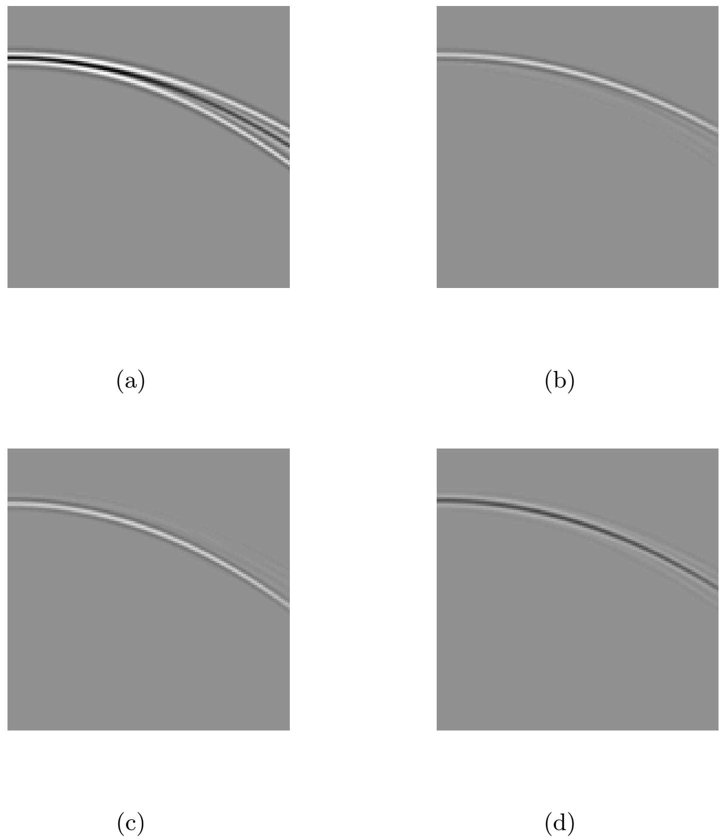


Figure 7: GMD-R applied to a synthetic seismic model consisting of three parabolic events with similar slopes. (a) Synthetic model. (b)-(d) Three decomposed modes.

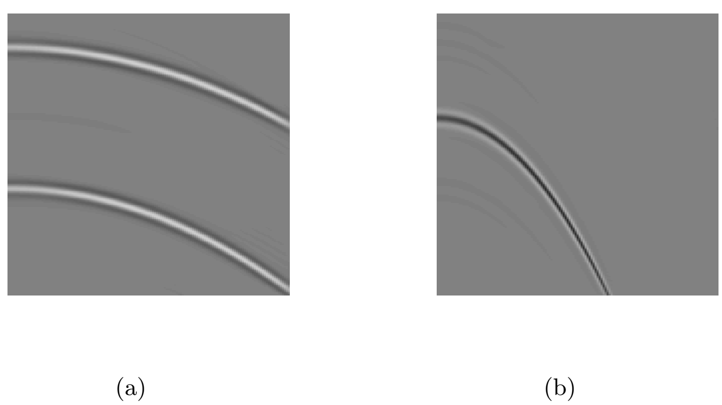


Figure 8: GMD-R1. (a)-(b) The two decomposed modes. The first mode contains two events with similar slopes.

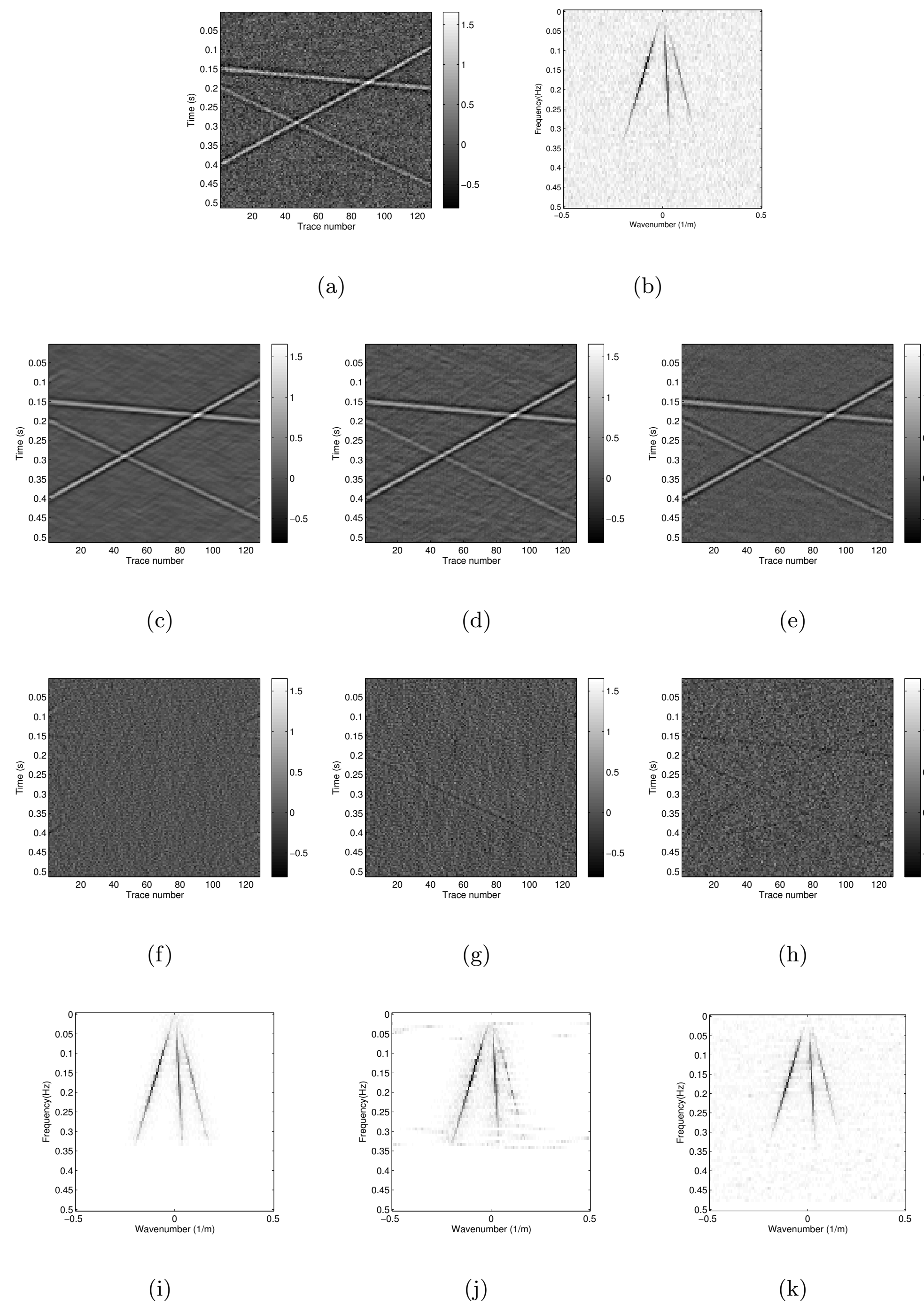


Figure 9: Noise attenuation with GMD-F. (a) Original noisy data. (b) $f-k$ spectrum of (a). (c) - (e) Denoising results of the GMD-F method (SNR=10.77), the 1D VMD method (SNR=6.75), and the $f-x$ deconvolution method (SNR=9.15). (f)-(h) Error between denoising results and noisy data corresponding to (c)-(e). (i)-(k) $f-k$ spectra of (c)-(e).

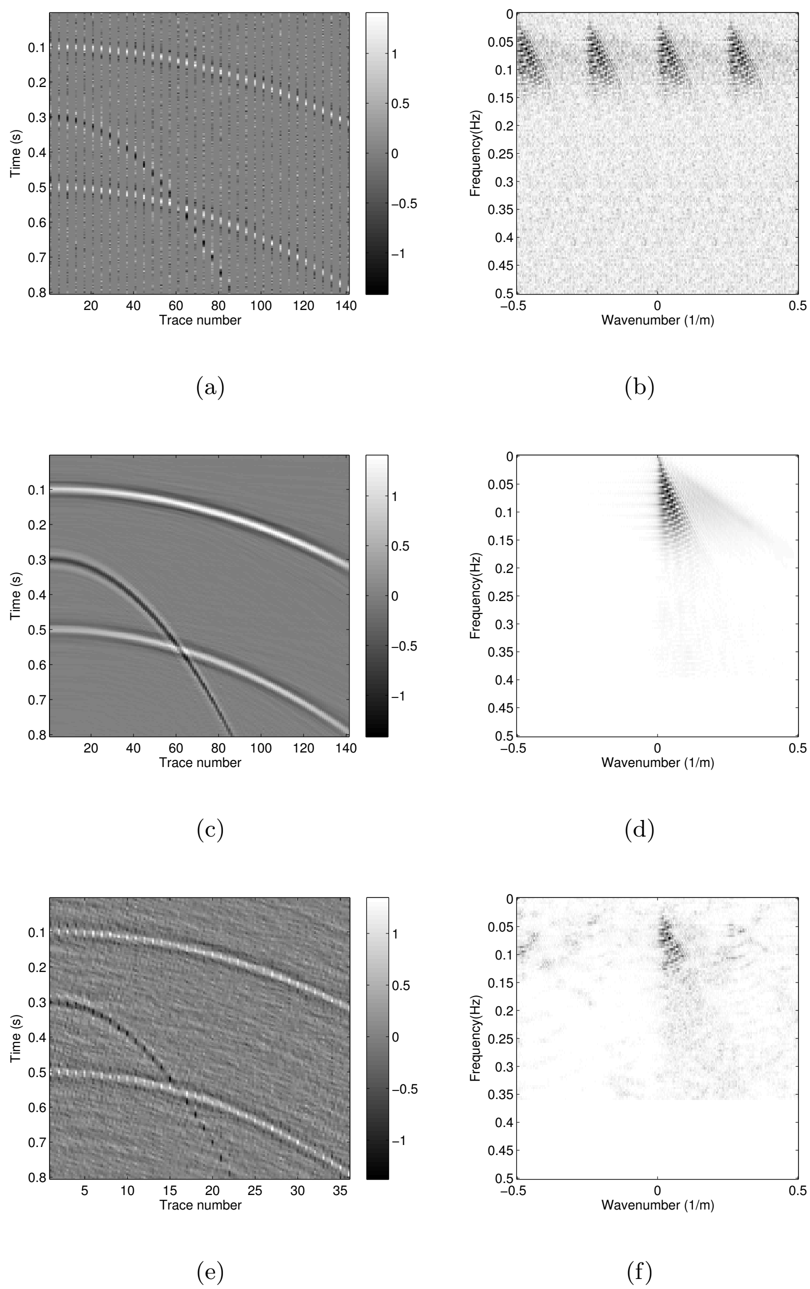
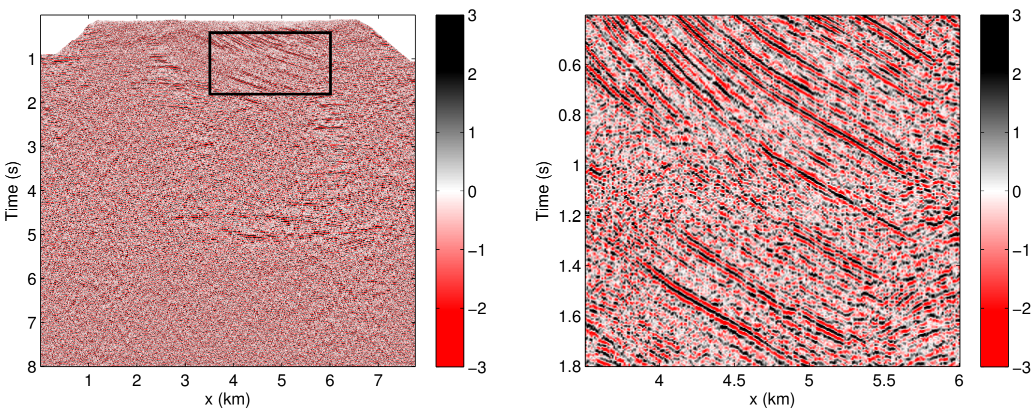


Figure 10: Data interpolation with GMD-R. (a) 25% regularly sub-sampled data. (c) Interpolated data with GMD-R. (e) Interpolated data with Spitz interpolation. (b),(d), and (f) $f-k$ spectra of (a),(c), and (e).



(a) (b)

Figure 11: Field data noise attenuation with GMD-F. (a) Field data. (b) Zoomed version of (a).

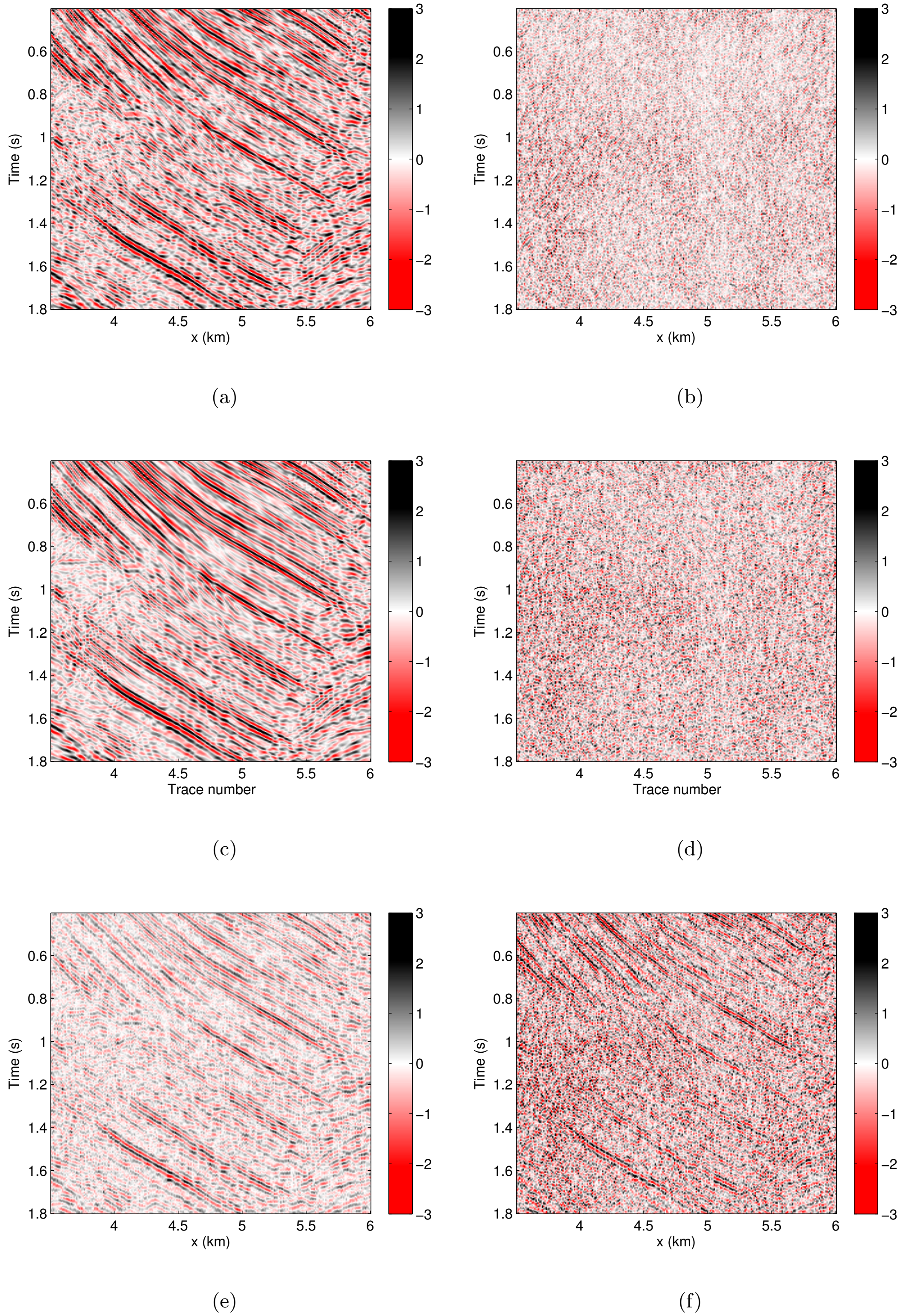
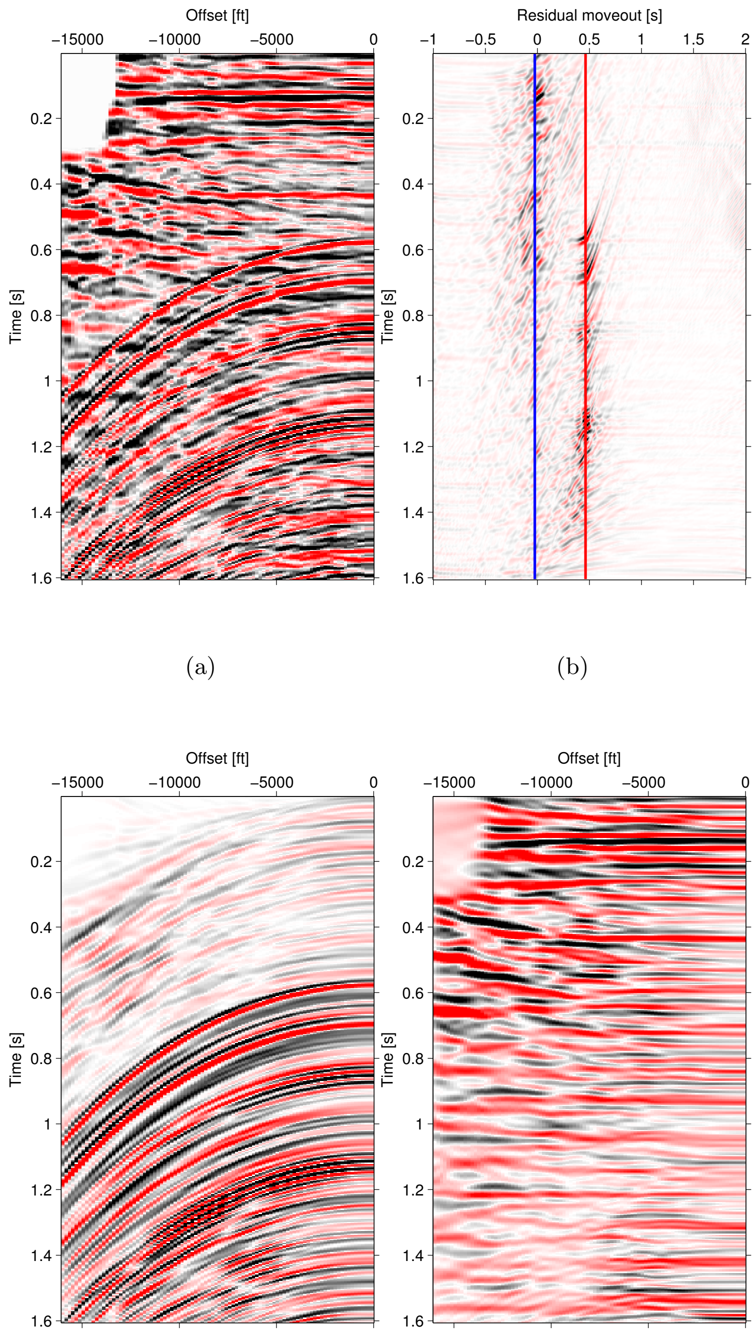


Figure 12: Field data noise attenuation with GMD-F. (a),(c), and (e) are the noise attenuation results of the GMD-F method, the curvelet method, and $f-x$ deconvolution method, respectively. (b),(d), and (f) are the corresponding noise.



(a) (b)

(c) (d)

Figure 13: Demultiple on NMO-corrected traces. (a) NMO-corrected traces. (b) Parabolic Radon spectrum. The two lines represent the two modes detected. (c) and (d) The separated multiple and primary with GMD-R1. $\alpha = 0.005$.

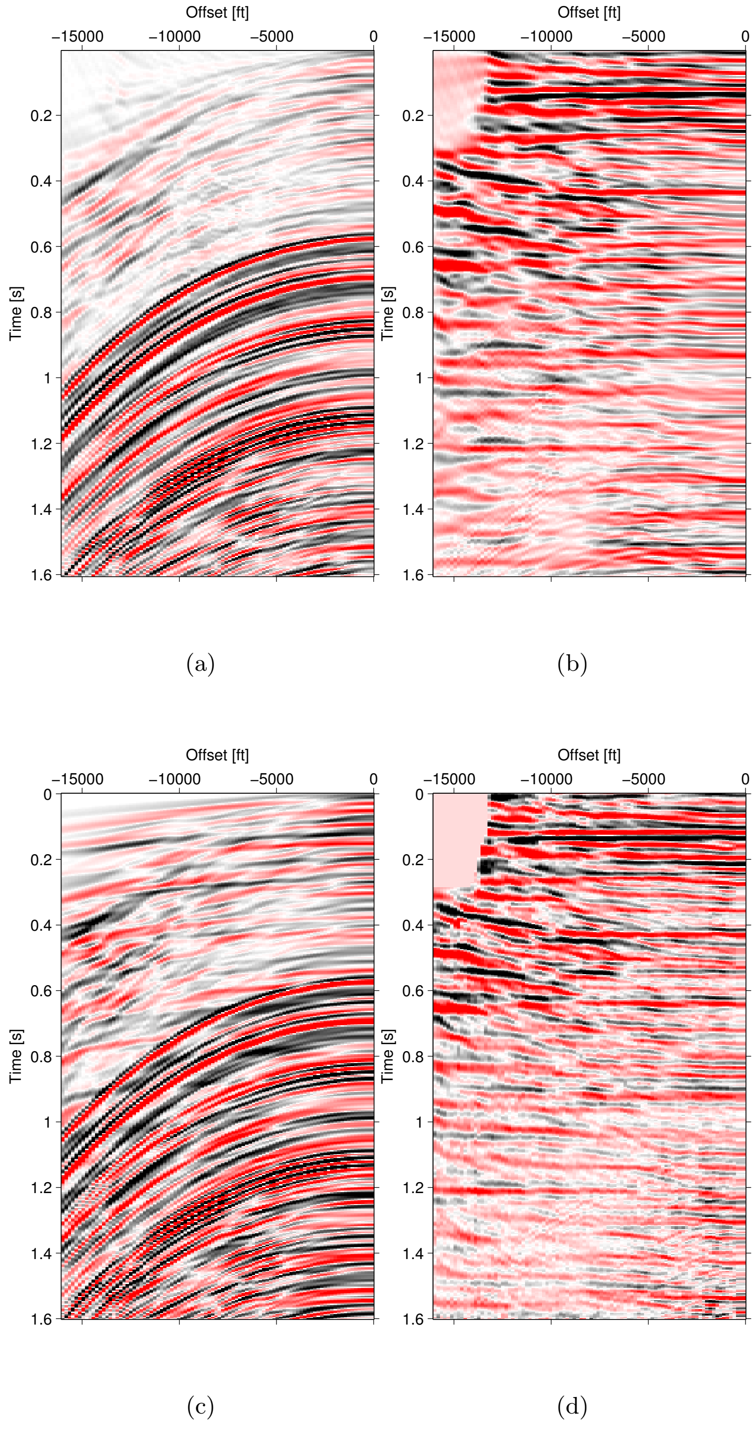


Figure 14: Demultiple on NMO-corrected traces. (a) and (b) The separated multiple and primary with GMD-R1. $\alpha = 10^{-5}$. (c) and (d) The separated multiple and primary by directly muting the Radon spectrum.

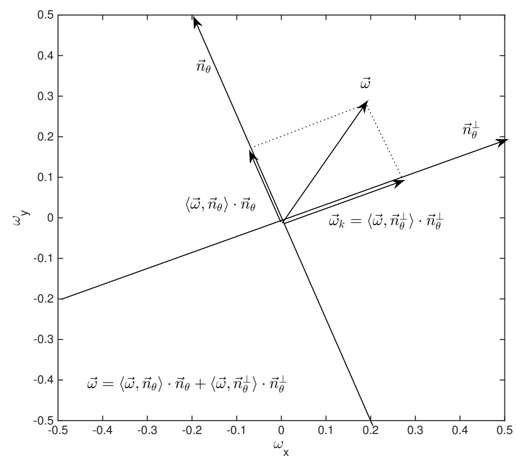


Figure 15: The relationship between GMD and 2D VMD.