Asymmetric chirplet transform for sparse representation of seismic data

Florian Boßmann[1] *and Jianwei Ma*[2]

ABSTRACT

For a fast and accurate extraction of important information in seismic signals, a sparse representation based on physical parameters of the given data is crucial. In this paper we use the Asymmetric Gaussian Chirplet Model (AGCM) and establish a dictionary free variant of the Orthogonal Matching Pursuit (OMP), a Greedy algorithm for sparse approximation. The atoms of AGCM, so-called chirplets, display asymmetric oscillation-attenuation properties, which make the AGCM very suitable for sparse representation of absorption decay seismic signals. Unlike the Fourier transform which assumes that the seismic signals consist of plane waves, the AGCM assumes the seismic signal consists of non-stationary compressed plane waves, i.e., symmetric or asymmetric chirplets. Thus AGCM is a general model for seismic wave simulation, and its model parameters include envelope part and phase part. In this paper, we mainly concentrate on the parameters of envelope part such as envelope amplitude and arrival time. We will show numerical examples using the algorithm for seismic signal approximation and arrive-time detection. The results show a promising performance but may be improved considering also spatial correlations of seismic data.

INTRODUCTION

Seismic signals need to be analyzed in many applications, e.g., the detection of different earth layers. Most of the methods used today are based on the same idea: The measured data

[1]Institute for Numerical and Applied Mathematics, University of Göttingen, Göttingen, 37083, Germany, f.bossmann@math.uni-goettingen.de

[2]Department of Mathematics, Harbin Institute of Technology, Harbin, 150001, China, jma@hit.edu.cn

consists out of a small number of important components plus noise. Thus, decomposition methods are often applied in order to reconstruct the individual components. If we assume the number of components to be small, Greedy methods can be used to obtain a sparse decomposition (Tropp, 2004; Boßmann et al., 2012).

The quality of the reconstructed information is highly dependent on the nature of the components one assumes. Nowadays, many methods use the assumption that seismic data can be sparsely represented by sparse transforms, e.g., the Radon transform (Trad et al., 2003), Fourier transform (Liu and Sacchi, 2004), wavelet transform (Daubechies, 1992), curvelet transform (Herrmann and Hennenfent, 2008; Ma and Plonka, 2010). Some seismic-feature based (e.g., physical wavelet transform (Zhang and Ulrych, 2003) and seislet transform (Fomel and Liu, 2010)) or adaptive learning dictionaries (e.g., data-driven tight frame (Liang et al., 2014)) have been proposed for sparser representation of seismic data. Other sparse transforms, e.g., singular value decomposition (Freire and Ulrych, 1998), empirical mode decomposition (EMD) based Hilbert-Huang transform (Huang and Wu, 2008), Prony transform (Pisarenko, 1973; Mitrofanov et al., 1998; Bath, 1995; Beylkin and Monzon, 2005; Fomel, 2013), have also been used for seismic data processing. The sparsity plays a key role in seismic data denoising (Beckouche and Ma, 2014), data acquisition (Herrmann, 2010), interpolation (Liu and Sacchi, 2004; Herrmann and Hennenfent, 2008; Trad, 2009; Liang et al., 2014), prediction of multiples (Donno et al., 2010; Ventosa et al, 2012), compression (Ma et al., 2010; Duchkov et al., 2010), seismic modeling (Hong and Kennett, 2003; Sun et al., 2009), migration (Chauris and Nguyen, 2008), imaging (de Hoop et al., 2009), simultaneous-source deblending (Chen et al., 2014), and seismic inversion (Li et al., 1996; Yuan and Simons, 2014).

The advantage of e.g. wavelets is, that today wavelets are well understood and there are known decomposition methods that can be applied directly. However, the wavelet itself is more a mathematical structure than based on the physical background of seismic data. Due to the absorption attenuation factor of seismic waves in subsurface, the seismic signal has a basic decay property. The presented method is a model-based method to represent seismic

data that consists of several (overlapping) seismic waves, i.e., symmetric or asymmetric chirplets. The first motivation of our paper is that the AGCM is essentially very suitable to sparsely represent non-stationary decay seismic signals. AGCM uses an adaptive chirplet model function to simulate a seismic wave depending on its physical properties like frequency, bandwidth or other factors. When the seismic signal is decomposed into several waves, these properties can give direct information, e.g., about the underlying material of the earth layer. The second motivation is that the model parameters or coefficients in transform domain such as envelope amplitude and arrival-time can provide not only explicit physical interpretation but also potential use for local attribution (Fomel, 2007), envelope-based inversion (Wu et al., 2014), early arrive-wave detection, and denoising.

To implement the AGCM method, we apply an OMP Greedy algorithm. The OMP seeks a sparse approximation of the given data using a number of basic atoms, the so-called dictionary. We will combine OMP with the AGCM. This model is an extended variant of Gaussian Chirplets or, more basic, Gabor atoms. AGCM introduces a new parameter, the asymmetry factor that allows the model function to have an asymmetric behavior such as absorption decay seismic signals. The AGCM is a general model for seismic wave simulation. However, the flexibility will be lost if we use AGCM to design a dictionary for OMP. Hence, we will introduce a dictionary free version of OMP to overcome this problem. The proposed algorithm will directly calculate the best parameter set in each iteration instead of just choosing from a dictionary. We will concentrate on the slightly simpler case where we only consider the envelope of the seismic data and provide a fast parameter guess that can approximate the optimal parameters. Preliminary tests on numerical examples show the promising performance of the proposed AGCM method on sparse representation of seismic data.

We also refer to previous results in (Lu et al., 2006, 2008) where comparable ideas for the slightly simpler Gaussian chirplet model were presented. The Gaussian chirplet model can only represent waves with a symmetric envelope and thus it is less flexible than AGCM. Similar to the method we present here, the proposed algorithms use Greedy methods

combined with a successive parameter guess for the chirplets. However, those methods make use of the derivative of the model and/or require an evaluation of the chirplets for (many) different parameters. As we will later see, the derivative of AGCM may be unstable and sampling of AGCM for different parameter sets leads to a high computational cost. Thus, we avoid both cases in our method.

Finally, we give a comment of our method in comparison with Prony's method (Mitrofanov et al., 1998). Similar to the Fourier transform, Prony's method extracts valuable information from a sampled signal and builds a series of damped complex exponentials or sinusoids. This allows for the estimation of amplitude, frequency, phase and damping components of a signal. Our method approximates the amplitude using AGCM with 4 unknown parameters (amplitude, bandwidth, asymmetry factor, and time shift). With the Prony method, one can only calculate a sparse approximation where the amplitude and the time shifts are unknown, i.e., the bandwidth and asymmetry factors must be known/fixed. This means the form of the envelope must be fixed first. Thus, our method can reconstruct more information about the given data.

THEORY

Asymmetric Gaussian Chirplet Model

In this paper, we consider a wave model that has been introduced recently in (Demirli and Saniie, 2014) applied to ultrasonic data. It is a generalization of the Gaussian chirplet model used for example in (Fan et al., 2002). The so called asymmetric Gaussian chirplet model (AGCM) is of the form

$$\mathrm{A}_p(t) := \mathrm{E}_p(t)\mathrm{F}_p(t), \tag{1}$$

$$\mathrm{E}_p(t) := \exp^{-\alpha(1-\beta\tanh(C(t-\tau)))(t-\tau)^2}, \tag{2}$$

$$\mathrm{F}_p(t) := \cos(f(t-\tau)+\gamma(t-\tau)^2+\theta), \tag{3}$$

with the envelope part $E_p(t)$ and the frequency part F_p. Here, $\tanh(Ct) = -i\tan(iCt)$ is the hyperbolic tangent function that for large C is approximatively 1 for $t > 0$ and -1 for $t < 0$. The wave form of the AGCM function $A_p(t)$ is defined by a parameter set $p := (\alpha, \beta, \tau, f, \gamma, \theta)$ with the bandwidth factor $\alpha > 0$, the asymmetry factor $\beta \in (-1, 1)$, the time shift $\tau > 0$, the center frequency $f \in [0, 2\pi)$, the chirp rate $\gamma \in [0, 2\pi)$ and the phase $\theta \in [0, 2\pi)$. C is a positive constant chosen a-priori. Note that the envelope $E_p(t)$ decreases exponentially with a bandwidth approximatively equal to $\alpha(1+\beta)$ for $t < \tau$ and $\alpha(1-\beta)$ for $t > \tau$. For $\beta = 0$, the AGCM will be the conventional chirplet model. With the additional degree of freedom, AGCM is able to simulate asymmetric waves that for example have a fast increase of the amplitude in the beginning and then decrease slowly. Thus, this model should be more suitable to approximate seismic waves that are also decay signals because of underground medium absorption. Figure 1 illustrates some elements of the AGCM with different parameter vectors.

The AGCM model decomposes a signal into two parts, envelope $E_p(t)$ and frequency $F_p(t)$. Both parts can be can be further decomposed into its parameter domain (α, β, τ) for $E_p(t)$, and (f, γ, θ) for $F_p(T)$. In this paper, we mainly focus on the parameters of $E_p(t)$ and show how we can obtain the parameters by designed mathematical algorithms. We give simple examples on the physical significance of these parameters and their potential applications. However, we still have full knowledge about $F_p(t)$ since it is given by $D(t)/E_p(t)$ where D is the original data. Thus we can completely reconstruct the data. We do not lose any information on the $F_p(t)$ in our algorithm, but just do not further decompose it into its parameter domain.

The envelope (low-frequency bandwidth) can be used in recently envelop inversion methods that can provide a good initial model for full-waveform inversion (e.g., see R. Wu 2014). Travel time can also be used for joint inversion and other attribute analysis. As you said, the phase part is related to fast varying rate (high-frequency bandwidth), and it is very important. We will focus on the phase parameters in our next work. Further note that, since in this work we concentrate on the envelope part of AGCM, the form and type of the

frequency part F_p is not significant for the here proposed algorithm and may be replaced by other model functions. However, we kept with the chirplet function as it follows the derivation of the model in (Demirli and Saniie, 2014).

Sparse representation of seismic data

Now we assume that seismic data has a sparse representation in AGCM space, that is, the data can be approximated using only few AGCM waves, i.e., given the data $D(t)$ we obtain

$$D(t) \approx \sum_{k=1}^{L} a_k \mathrm{A}_{p_k}(t), \tag{4}$$

where L is a small number, $a_k \in \mathbb{R}$ is the amplitude of each atom $A_{p_k}(t)$ and p_k is a parameter set for each $k \leq L$. Figure 2 compares the amplitude size of seismic data in AGCM space with its Fourier amplitudes using 512^2 elements, i.e., exact representation of the data. Note that the amplitudes in AGCM space decrease exponentially since the x-axis is plotted in logarithmic scale and thus the data can be approximated sparsely in the AGCM space.

Unfortunately the reconstruction of all parameters $\{p_k\}_{k=1}^{L}$ at once can be very unstable due to noisy data. However, the parameters of AGCM can be separated in a simple way using its envelope and phase separately. In this paper, we will concentrate on reconstructing the envelope of the AGCM atoms given by (2). Thus the number of parameters is reduced to 4 for each element including the amplitude a_k. Some crucial information about the underlying seismic structure is already contained in the envelope of the data (Wu et al., 2014). We assume that interference in (4) stays low, i.e., that the AGCM atoms in (4) are well separated and hence R is locally most influenced by only one AGCM atom at once. It follows that the envelope $\mathrm{Env}(D)$ of D can be approximated by

$$\mathrm{Env}(D)(t) \approx \sum_{k=1}^{L} a_k \mathrm{E}_{p_k}(t). \tag{5}$$

However, if the interference in (4) is high the envelope $\mathrm{Env}(D)$ can still be approximated according to (5). Note that in this case the amplitudes a_k and parameter vectors p_k may differ from those in (4) since e.g. overlapping atoms may be merged into one single envelope

E_{p_k} (see e.g. Figure 9 and the discussions in the Numerical Results section). Note that the envelope $\mathrm{Env}(D)$ of D can be calculated using the absolute value of its Hilbert transform as shown in Figure 3. Since the number of used atoms L is small, we can use Greedy algorithms for sparse recovery to reconstruct the parameters. We will adapt a widely used Greedy algorithm for signal decomposition known as OMP (Tropp, 2004). For a given dictionary of atoms $\{f_k\}_{k=1}^N$, where N is the size of the dictionary, OMP tries to find the best sparse approximation of a given data set by choosing the most correlating atom in each iteration and optimizing the amplitudes over all already chosen atoms. Using an AGCM based dictionary the functions f_k may be given by e.g. $f_k = A_{p_k}$ or $f_k = E_{p_k}$ where $\{p_k\}_{k=1}^N$ are a-priori chosen parameter vectors. As we will see later in this work, the choice of these vectors is the most crucial part of the algorithm. For simplicity we assume that the atoms are normalized, then OMP can be expressed by Algorithm 1.

Algorithm 1 Orthogonal Matching Pursuit for normalized atoms

Input: envelope $\mathrm{Env}(D)$, atoms $\{f_k\}_{k=1}^N$

Initialize: approximation $A = 0$, residual $R = \mathrm{Env}(D)$, $J = 1$

While some stopping criterion does not hold

Set $k_J := \mathrm{argmax}|\langle R, f_k\rangle|$

Solve $(a_1, \ldots, a_J) := \mathrm{argmin}\|\mathrm{Env}(D) - \sum_{j=1}^{J} a_j f_{k_j}\|_2$

Set $A = \sum_{j=1}^{J} a_j f_{k_j}$, $R = \mathrm{Env}(D) - A$ and $J = J + 1$

Output: A, R, sparsity J, indices $\{k_j\}_{j=1}^J$, amplitudes $\{a_j\}_{j=1}^J$

As a stopping criterion one can choose an upper bound L for the sparsity $J < L$ or stop when the residual is below some threshold $\|R\| < \varepsilon$. Note that the minimization of the amplitudes is a linear problem and thus can be solved using fast linear optimization methods.

In this form, OMP needs a number of atoms called the dictionary. To apply OMP on seismic data using AGCM waves, the dictionary may consist of several AGCM waves for different parameters $\{p_k\}_{k=1}^N$, i.e., $f_k = A_{p_k}$. Considering that each parameter should be

sampled appropriately, e.g., at M points on its domain where M may become very large, we obtain $O(M^6)$ different parameter sets p_k, i.e., $O(M^6)$ dictionary elements. Even applied to the more simple case $f_k = E_{p_k}$ using only the envelope of both the data and the AGCM atoms the size of the dictionary is $O(M^3)$. Thus, the number of atoms will be too large to implement OMP in a suitable manner.

To overcome this problem we introduce a dictionary free version of OMP. Instead of choosing a predefined dictionary element in the algorithm, we will directly calculate the best parameter set approximating the residual, i.e., we minimize the problem

$$f_{k_J} = E_{p_J} \text{ with } p_J = \operatorname{argmin} \|R - E_p\|_2.$$

Using common minimization methods will require to calculate the gradient of E_p. Unfortunately, this may be very unstable especially for $\beta \neq 0$ because of the tanh function. Therefore, we will use an approximative solution based on the ideas in (Demirli and Saniie, 2014). The authors present a method to successively reconstruct all parameters of the AGCM method. However, they only apply the method to one unknown AGCM atom. In this paper, we will use the ideas to separate several overlapping AGCM atoms. Since this will be very prone to noise when it comes to the reconstruction of the frequency, we will here stick with the envelopes.

The envelope E_p reaches its maximal value at position $t = \tau$. Thus, in the J-th iteration of OMP we may set

$$\tau_J = \operatorname*{argmax}_t \; R(t)$$

$$a_J = R(\tau_J).$$

Given the parameters τ_J and a_J we can calculate α_J and β_J by assuming that

$$R(t) \approx a_J \exp^{-\alpha_J(1-\beta_J \tanh(C(t-\tau_J)))(t-\tau_J)^2}, \tag{6}$$

for t close to τ_J. Plugging in a_J and τ_J and applying the logarithm this becomes a simple quadratic equation that can be evaluated at some test points. Note that these values are

more accurate the less overlapping the seismic waves in the data are. For this reason the test points should not be chosen too far away from $t = \tau_J$. Indeed, if there is no overlap at the test points the parameters will return the exact solution (in the noiseless case). However, in each iteration OMP will perform an optimization of the amplitudes of all selected atoms. This will be more accurate than using $a_J = R(\tau_J)$ since the optimization can also take the correlation between overlapping atoms into account. We obtain the adapted dictionary free OMP (ADOMP) as shown in Algorithm 2.

Note that the sample points $t_1, \ldots, t_K$ in Algorithm 2 should be chosen in a way such that the least square problem becomes stable. We suggest the following restrictions to the sample points:

1. For $k = 1, \ldots, K$ the inequality $R(t_k) > a_J/4$ should hold.
2. $R(t)$ should be monotonically increasing on $[t_1, \tau_J]$ and decreasing on $[\tau_J, t_K]$.
3. For a chosen $\varepsilon > 0$ there should be no $k \leq K$ such that $t_k \in [\tau_J - \varepsilon, \tau_J + \varepsilon]$.

Since for small amplitudes $R(t)$ the relative error due to noise on the data will be big, the first condition ensures that we are only considering sample points with an amplitude above the threshold $a_J/4$. The second condition ensures that the assumption (6) will hold. Because the exponential function in (6) is monotonically increasing for $t < \tau_J$ and decreasing for $t > \tau_J$, other behavior will indicate the influence of other AGCM atoms on the function values. The last restriction is dealing with the problem of noise on the reconstructed values a_J and τ_J, i.e., instead of a_J, τ_J only $a_J + \varepsilon_1, \tau_J + \varepsilon_2$ with small $\varepsilon_1, \varepsilon_2$ is given. Note that $\tanh(C(t - \tau_J))$ is close to 1 if $t > \tau_J$ and close to -1 if $t < \tau_J$. To avoid $t_k > \tau_J$ but $t_k < \tau_J + \varepsilon_2$ or $t_k < \tau_J$ but $t_k > \tau_J - \varepsilon_2$ the sample points should be chosen with some distance ε to τ_J. Moreover, we obtain for all $k = 1, \ldots, K$ that

$$\begin{aligned}
\log(R(t_k)/(a_J + \varepsilon_1)) &= \log(R(t_k)) - \log(a_J + \varepsilon_1) \\
&= \log(R(t_K)) - \log(a_J) - \log(1 + \varepsilon_1/a_J) \\
&= \log(R(t_k)/a_J) - \log(1 + \varepsilon_1/a_J).
\end{aligned}$$

Hence the absolute error is given by $\log(1+\varepsilon_1/a_J)$. For $\varepsilon_1 > 0$ this term will be smaller than ε_1/a_J, i.e., the absolute error is decreasing due to the log term. For $-0.7968a_J < \varepsilon_1 < 0$ (i.e., for a negative relative error up to nearly 80%) the term is bounded by $2\varepsilon_1/a_J$. Thus, the absolute error is stable up to very high noise on a_J. However, since for t_k close to τ_K it will hold that $R(t_k) \approx a_J$ i.e., $\log(R(t_k)/a_J)$ will be small and hence the relative error may be big, we suggest again to chose t_k with some distance ε from τ_K. Considering all three restrictions the least square problem will become stable. Nevertheless, if the number of possible sample points is too low, e.g. we suggest at least 3 on each side of τ_J, the restrictions can be relaxed.

Algorithm 2 Adapted dictionary free OMP for AGCM atoms

Input: envelope $\mathrm{Env}(D)$

Initialize: approximation $A = 0$, residual $R = \mathrm{Env}(D)$, $J = 1$

While some stopping criterion does not hold

Calculate $\tau_J = \mathrm{argmax}\ R$, $a_J = R(\tau_J)$

Choose K sample points $t_1, \ldots, t_K$

For $k = 1, \ldots, K$ consider the K equations

$$\log(R(t_k)/a_J) = -\alpha_J(1 - \beta_J \tanh(C(t_k - \tau_J)))(t_k - \tau_J)^2$$

and solve for α_j, β_j in a least square sense

Set $p_J = (\alpha_J, \beta_J, \tau_J)$

Solve $(a_1, \ldots, a_J) := \mathrm{argmin}\|\mathrm{Env}(D) - \sum_{j=1}^{J} a_j f_{p_j}\|_2$

Set $A = \sum_{j=1}^{J} a_j f_{k_j}$, $R = \mathrm{Env}(D) - A$ and $J = J + 1$

Output: A, R, sparsity J, indices $\{k_j\}_{j=1}^{J}$, amplitudes $\{a_j\}_{j=1}^{J}$

NUMERICAL RESULTS

For our numerical experiments we stop the algorithm when the maximal amplitude of the residual drops below 25%, 10%, 5% or 1% of the original amplitude, i.e., if $\|R\|_\infty <$

$\kappa\|\mathrm{Env}(D)\|_\infty$ where $\kappa = 0.25, 0.1, 0.05, 0.01$. Figure 4 and 5 show the approximation of the envelope using ADOMP with different κ using the original data shown in Figure 3. In Figure 6 the mean squared error (MSE) and signal-to-noise ratio (SNR) of the reconstruction is plotted as a function of the used parameter κ in percent. Here we sampled the function on $\kappa = 0.25, 0.24, \ldots, 0.01$, i.e. the calculated the MSE and SNR when we stop the algorithm after the residual drops below $25\%, 24\%, \ldots, 1\%$. As a comparison we also plotted the MSE and SNR of an approximation using only the largest Fourier coefficients, where the number of coefficients is the same as atoms used in our algorithm. The input data has been modified to have a maximal absolute value of 100 for a better comparison. Since ADOMP returns a sparse approximation of the original data, the method can be used to store given data in an efficient way as Figure 7 illustrates. The number of atoms and the relative benefit is plotted as a function of κ in percent. Here, we assume that storing the complete data pixel by pixel, i.e. 512^2 elements, has a relative storage cost of 1. The relative costs of ADOMP storing all amplitudes and the corresponding parameters are shown in Figure 7 (b). Although we need to store several parameters for a single AGCM atom, the representation is so sparse, i.e. only a few atoms are needed, that the overall storage costs stay low. The number of Fourier coefficients and storage cost needed to obtain the same approximation accuracy is plotted as a comparison.

Besides the sparse representation of seismic data, the proposed method can also be used to directly manipulate the data or extract important information of the source of the data. Here we will demonstrate the utility of AGCM atoms and the ADOMP algorithm for approximation and arrival-time detection.

Using the fact that the data D can be written as $D = \mathrm{Env}(D) \cdot \mathrm{Ph}(D)$ where $\mathrm{Ph}(D)$ is the phase of D, we can implement an approximation algorithm for seismic data by approximating its envelope components (see Algorithm 3). The algorithm is based on the idea, that ADOMP will most likely reconstruct the large major parts of $\mathrm{Env}(D)$, i.e., the parts that obtain important information, and suppress small amplitudes, i.e., noise. Thus the algorithm can also be used for simple denoising. The approximated results using different κ are

shown in Figure 8. Here we just apply the AGCM trace by trace, so it is essentially a one-dimensional problem. The preliminary results show the denoising usability of the AGCM. As an improvement of this method, coherence in spatial dimensions should be considered. This will be part of our future work.

Algorithm 3 Denoising with ADOMP

Input: data D

Calculate Env(D) and Ph(D)

Find an approximation A of Env(D) using ADOMP

Set $D = A \cdot \mathrm{Ph}(D)$

Output: Approximated data D

As another application we want to present the use for seismic arrival-time detection. As stated already in the introduction of this paper, one important difference between mathematical atoms like wavelets and physical based atoms like AGCM is, that the reconstructed parameters carry information about the underlying physical structure. It can give direct conclusion on the type, source or material of detected earth layers. We demonstrate this using the synthetic data shown in Figure 9 (a). The data contains a linear event at its bottom, a parabolic event, and a linear structure at its top that is slowly stretched in time. The stretch was only applied to the decreasing part of the wave (i.e. to the part after the envelope reached its maximum to obtain asymmetric attention waves). Thus, the seismic impulse initialization remains the same but its attenuation in time is slowed. These seismic responses are also observed in real data and hence they are a reasonable assumption for the synthetic data. The seismic impulse/signal is shown in Figure 9 ((b) solid blue line). We manually extracted it from the real data shown in Figure 3 (a). First, we apply the ADOMP algorithm using $\kappa = 0.1$ and a sparsity of at most 3 per trace. As a comparison we apply ADOMP a second time but now force $\beta = 0$, i.e. AGCM is simplified to the Gaussian Chirplet Model (GCM). Both algorithms extract 1337 parameter sets. Figure 9 (b) shows the original envelope (dotted blue line) and the reconstructed envelopes using AGCM (solid red line) and GCM (dashed black line) with the median parameters. Since GCM is

not able to simulate an asymmetric behavior, it overestimates the envelope for $t < 0$ and underestimates for $t > 0$. The AGCM returns a more suitable approximation. This gets even more clear when we consider the approximation error for the complete data envelope. Using GCM the residual norm is 1070.5 while it is only 290.9 for AGCM. Figures 9 (c) and (d) show the extracted arrival times using GCM (c) and AGCM (d). Both algorithms manage to reconstruct the arrival times as long as the signals are not to much overlapping. In this case, the correlation between both signals is too strong and the algorithm tries to approximate the two overlapping signals with only one atom. However, while the parameters may suffer from this, the reconstructed envelope will still be a good approximation. This especially holds for the more flexible AGCM. While GCM adds another atom to approximate the overlapping signals, AGCM recovers a good approximation with only one atom (note that this does not imply that the GCM approximation is better than AGCM since the reconstructed parameters in GCM will not correlate with the real parameters). Thus, the next added atom in AGCM is already used to reconstruct very small structures in the residual (see the artefacts in Figure 9 (d) at time sampling number 140). Anyhow, these artefacts can be filtered out very easy since their corresponding amplitudes are extremely small. In this case, if we use AGCM with increased $\kappa = 0.12$ and the algorithm will only return 1325 elements displayed in Figure 10 (a) (we nevertheless illustrated the algorithm for $\kappa = 0.1$ to compare AGCM and GCM with the same number of atoms used). Figure 10 shows the initialization bandwidth factor $\alpha(1 + \beta)$ of the reconstructed envelope for $t < \tau$ and the attenuation bandwidth factor $\alpha(1 - \beta)$ for $t \geq \tau$ using AGCM and the parameter sets corresponding to the stretched seismic signal. One can clearly see that AGCM is able to detect the constant initialization bandwidth factor and the decreasing attenuation bandwidth factor. The attenuation decreases from 0.0035 to 0.00065, i.e., a decrease of about 81 percent, while $\alpha(1 + \beta)$ changes only by about 7.5 percent from 0.011 to 0.0118. This example demonstrates the benefit of our AGCM algorithm not only in arrival-time detection but also in envelope approximation (e.g., error norms given above), and asymmetric wave detection (e.g., the above linear stretched structure).

Finally, we test the arrive-time detection for real data. We demonstrate this by visualizing the arrival time τ for the atoms used in the reconstruction with different parameters κ. The parameters τ shown in Figure 12 are obtained applying ADOMP on pre-stack marine shot gather data shown in Figure 11, i.e. τ corresponds to the physical arrival time. Figure 13 shows the result when ADOMP is applied to the real post-stack data shown in Figure 3. For post-stack data the τ does not more correlate to the physical arrival time but can now be used as a layer reconstruction.

CONCLUSION

We have presented a new algorithm for sparse representation of seismic data using AGCM. Besides the parameters already known from GCM or Gabor atoms, this new model introduces an additional parameter to simulate asymmetric wave forms. The large number of physical parameters allows a good approximation of acoustic as well as seismic waves but also comes at hand with new challenges. A reconstruction using atoms from a dictionary is no longer suitable. We overcome this problem by adapting the OMP to a dictionary free version. Moreover, the parameter reconstruction was stabilized. Another advantage of our method is, that the reconstructed parameters have a physical interpretation and thus can help e.g., for the reconstruction of earth layers. This does not hold for wavelets or curvelets. The usability of our algorithm was demonstrated showing numerical results for approximation, sparse storage, envelope and arrival-time detection.

In the next work, we will further explore how the parameters (e.g., envelope, arrival-time, phase) benefit for real seismic data processing. How to extend the AGCM to high-dimensional denoising by considering the spatial coherence of seismic data is another important work.

ACKNOWLEDGEMENTS

The authors thank Sergey Fomel and Ru-Shan Wu for their helpful discussions. This work is supported by NSFC (grant number: NSFC 91330108,41374121,61327013), the Fundamental Research Funds for the Central Universities (grant number: HIT.BRETIV.201314), the Program for New Century Excellent Talents in University (grant number: NCET-11-0804) and BMBF joined research project ZeMat (grant number: 05M13MGA). Furthermore, we want to thank the editors and reviewers for their helpful comments to improve this work.

REFERENCES

Bath, M., 1995, Modern Spectral Analysis with Geophysical Applications: Soc. of Expl. Geophys.

Beckouche, S., and J. Ma, 2014, Simultaneously dictionary learning and denoising for seismic data: Geophysics, **79**, no 3, A27-A31.

Beylkin, G., and L. Monzo?n, 2005, On approximation of functions by exponential sums: Applied and Computational Harmonic Analysis, **19**, 17-48.

Boßmann, F., G. Plonka, T. Peter, O. Nemitz, and T. Schmitte, 2012, Sparse deconvolution methods for ultrasonic NDT: Journal of Nondestructive Evaluation, **31**, no. 3, 225-244.

Chauris, H., and T. Nguyen, 2008, Seismic demigration/migration in the curvelet domain: Geophysics, **73**, no. 2, S35-S46.

Chen, Y., S. Fomel, and J. Hu, 2014, Iterative deblending of simultaneous-source seismic data using seislet-domain shaping regularization: Geophysics, **79**, V179-V189.

Daubechies, I., 1992, Ten Lectures on Wavelets: SIAM Press, Philadelphia.

de Hoop, M., H. Smith, G. Uhlmann, and R. van der Hilst, 2009, Seismic imaging with the generalized Radon transform: a curvelet transform perspective: Inverse Problems, **25**, no. 2, 025005.

Demirli, R., and J. Saniie, 2014, Asymmetric Gaussian chirplet model and parameter estimation for generalized echo representation: Journal of the Franklin Institute, **351**, no. 2, 907-921.

Donno, D., H. Chauris, and M. Noble, 2010, Curvelet-based multiple prediction: Geophysics, **75**, no. 6, WB255-WB263.

Duchkov, A., F. Andersson, and M. de Hoop, 2010, Discrete almost-symmetric wave packets and multiscale geometrical representation of (seismic) waves: IEEE Trans. Geosci. Rmote Sens., **48**, no. 9, 3408-3423.

Fan, W., H. Zou,Y. Sun, Z. Li, and R. Shi, 2002, Decomposition of seismic signal via chirplet transform, ICSP02 Proceedings, pp.

Fomel, S., 2007, Local seismic attributes: Geophysics, **72**, no. 3, A29-A33.

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

Fomel, S., 2013, Seismic data decomposition into spectral components using regularized nonstationary autoregression: Geophysics, **78**, O69-O76.

Freire, S., and T. Ulrych, 1998, Application of singular value decomposition to vertical seismic profiling: Geophysics, **53**, 778-785.

Herrmann, F., and G. Hennenfent, 2008, Non-parametric seismic data recovery with curvelet frames: Geophysical Journal International, **173**, 233-248.

Herrmann, F., 2010, Randomized sampling and sparsity: getting more information from fewer samples: Geophysics, **75**, WB173-WB187.

Hong, T., and B. Kennett, 2003, Modelling of seismic waves in heterogeneous media usign a wavelet-based method: application to fault and subdection zones: Geophysics J. Int., **154**, no. 2, 483-498.

Huang, N., and Z. Wu, 2008, A review on Hilbert-Huang transform: method and its applications to geophysical studies: Reviews of Geophysics, **46**, no. 2, DOI: 10.1029/2007RG00022.

Li, X., M. Sacchi, and T. Ulrych, 1996, Wavelet transform inversion with prior scale information: Geophysics, **61**, no. 5, 1379-1385.

Liang, J., J. Ma, and X. Zhang, 2014, Seismic data restoration via data-driven tight frame: Geophysics, **79**, no. 3, V65-V74.

Liu, B., and M. Sacchi, 2004, Minimum weighted norm interpolation of seismic records: Geophysics, **69**, no. 6, 1560-1568.

Lu, Y., R. Demirli, G. Cardoso, and J. Saniie, 2006, A successive parameter estimation algorithm for chirplet signal decomposition, IEEE Trans. Ultrason., Ferroelect., Freq. Control, **53**, no. 11, 2121-2131.

Lu, Y., E. Oruklu, and J. Saniie, 2008, Fast chirplet transform with FPGA-based implementation: IEEE Signal Processing Letters, 15, 577-580.

Ma J., G. Plonka, and H. Chauris, 2010, A new sparse representation of seismic data using adaptive easy-path wavelet transform: IEEE Geoscience and Remote Sensing Letters, **7**,

no. 3, 540-544.

Ma, J., and G. Plonka, 2010, The curvelet transform: IEEE Signal Processing Magazine, **27**, no. 2, 118-133.

Mitrofanov, G., Z. Zhan, and J. Cai, 1998, Using of the Prony transform of Chinese seismic data: 68th Ann. Internat. Mtg, Soc. of Expl. Geophys., pp. 1157-1159.

Pisarenko, V. F., 1973, The retrieval of harmonics from a covariance function: Geophys. J. R. Astr. Soc., 33, 347-366.

Sacchi, M., T. Ulrych, and C. Walker, 1998, Interpolation and extrapolation using a high-resolution discrete Fourier transform: IEEE Trans. Signal Processing, **46**, no. 1, 31-38.

Sun, B., J. Ma, H. Chauris, and H. Yang, 2009, Solving the wave equation in the curvelet domain: a mulit-scale and multi-directional approach: Journal of Seismic Exploration, **18**, 385-399.

Trad, D., 2009, Five-dimensioinal interpolation: recovering from acquisition constraints: Geophysics, **74**, V123-V132.

Trad, D., T. Ulrych, and M. Sacchi, 2003, Latest views of the sparse Radon transform: Geophysicsm, **68**, no. 1, 386-399.

Tropp, J., 2004, Greed is good: Algorithmic results for sparse approximation: IEEE Trans. Inform. Theory, **50**, 2231-2242.

Ventosa, S., S. Le Roy, I. Huard, A. Pica, H. Rabeson, P. Ricarte, and L. Duval, 2012, Adaptive multiple subtration with wavelet-based complex unary Wiener filters: Geophysics, **77**, no. 6, 183-192.

Wu, R., J. Luo, and B. Wu, 2014, Seismic envelope inversion and modulation signal model: Geophysics, **79**, no. 3, WA13-WA24.

Yuan, Y., and F. Simons, 2014, Multiscale adjoint waveform-difference tomography using wavelets: Geophysics, **79**, W79-W95.

Zhang, R., and T. Ulrych, 2003, Physical wavelet frame denoising: Geophysics, **68**, no. 1, 225-231.

Figure Captions:

Figure 1: (a) Symmetric Gabor atom with $p = (0.01, 0, 0, 0.5, 0, 0)$, (b) $\theta = \pi/4$, (c) Gabor atom with $p = (0.0025, 0, 0, 0.5, 0, 0)$, (d) $\beta = 0.8$, (e) $\gamma = 0.015$. All functions sampled on $t \in [-100, 100]$.

Figure 2: (a) Original seismic data with 512^2 pixels, (b) AGCM (red dash line) and Fourier (blue solid line) amplitudes for exact recovery using 512^2 elements in decreasing order using logarithmic scale on the x-axis.

Figure 3: Piece of a seismic image (a) and envelope obtained by Hilbert transform (b).

Figure 4: Reconstructed envelope using ADOMP with decreasing κ (from left-up to right-down).

Figure 5: (a)-(d) Original data (blue), its envelope (black) and its reconstruction (red) with decreasing κ. Larger κ values lead to smoother envelope.

Figure 6: Mean squared error (a) and signal-to-noise ratio (b) of the reconstructed envelope as a function in $\kappa\%$ using ADOMP (red dash line) compared to Fourier approximation using same number of coefficient as atoms used in ADOMP.

Figure 7: (a) Number of elements used in ADOMP (red dash line) for approximation and (b) storage cost relative to a full image storage as a function of $\kappa\%$. The number of Fourier coefficients needed to obtain the same approximation accuracy as well as their storage cost is plotted as a comparison (blue solid line).

Figure 8: Approximated data using the above stated algorithm with decreasing κ.

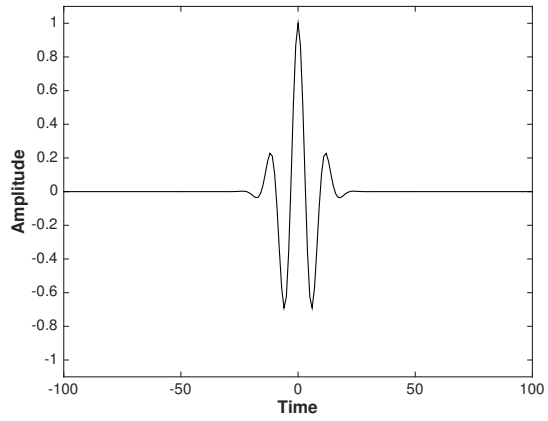
Figure 9: (a) Synthetic seismic data. (b) Seismic impulse (solid blue line) and its envelope (dotted blue line), the envelope reconstruction using GCM (dashed black line) and AGCM (solid red line). (c) and (d) Reconstructed/detection arrival times using GCM and AGCM (red lines plus furcation). (e) Reconstructed ascending bandwidth (solid line) and descending bandwidth (dashed line) of the first seismic response.

Figure 10: (a) Reconstructed arrival times using AGCM with increased $\kappa = 0.12$, (b) reconstructed ascending bandwidth (solid line) and descending bandwidth (dashed line) of the first seismic response in Figure 9 (d).

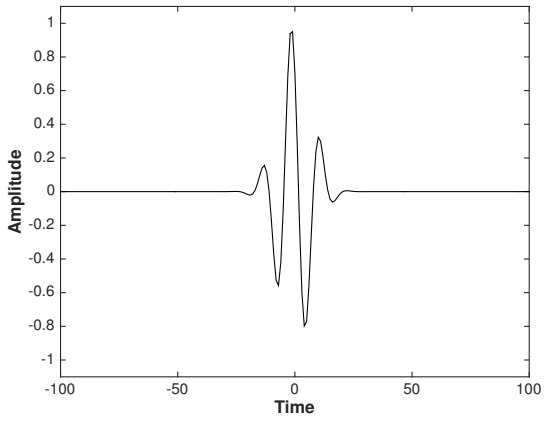
Figure 11: Marine shot gather (a) and its Hilbert transform (b).

Figure 12: Detection of arrival time parameter τ using ADOMP for decreasing κ on real pre-stack data shown in Figure 11.

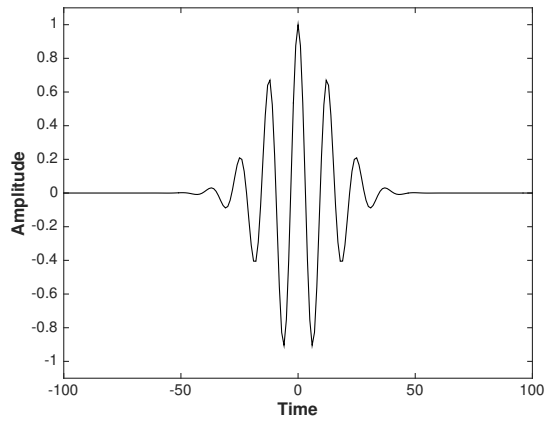
Figure 13: Detection of arrival time parameter τ using ADOMP for decreasing κ on real data shown in Figure 3.



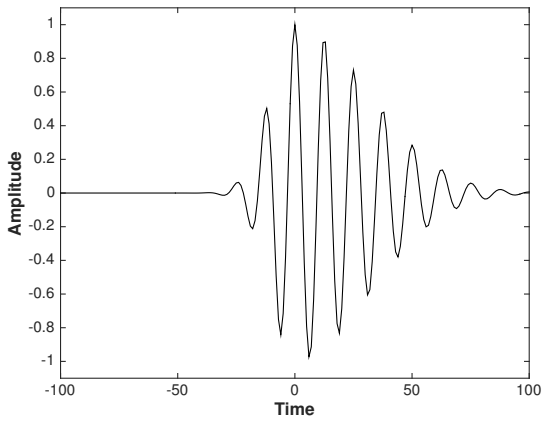
(a)



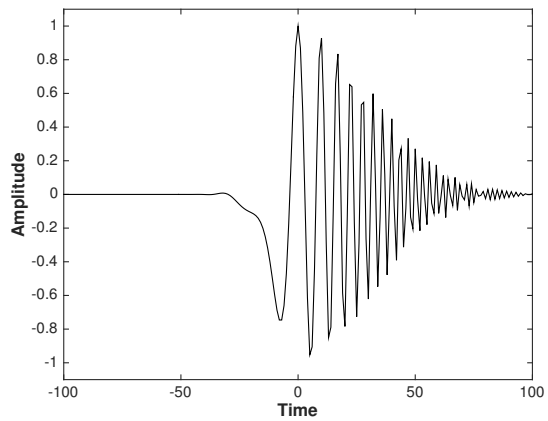
(b)



(c)

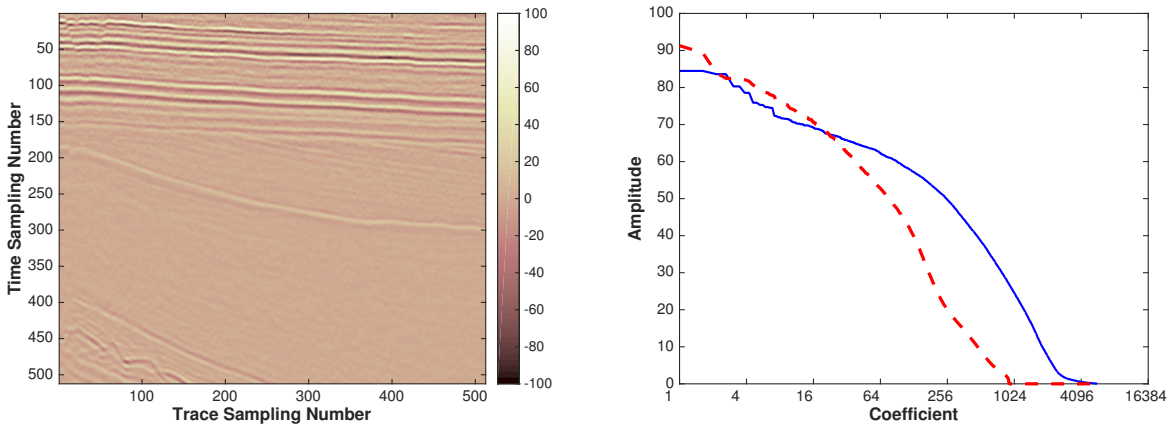


(d)



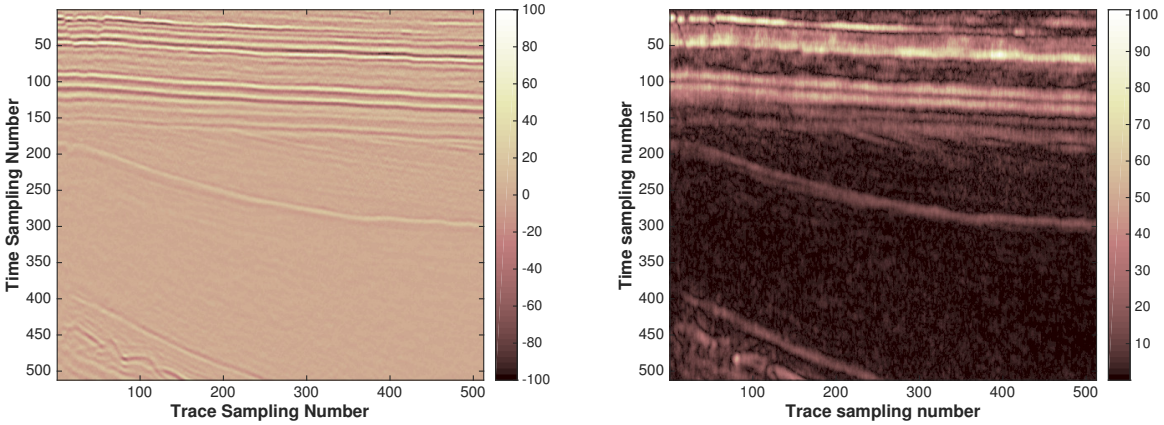
(e)

Figure 1: (a) Symmetric Gabor atom with $p = (0.01, 0, 0, 0.5, 0, 0)$, (b) $\theta = \pi/4$, (c) Gabor atom with $p = (0.0025, 0, 0, 0.5, 0, 0)$, (d) $\beta = 0.8$, (e) $\gamma = 0.015$, sampled on $t \in [-100, 100]$.



(a) (b)

Figure 2: (a) Original seismic data, (b) AGCM (red dash line) and Fourier (blue solid line) amplitudes for exact recovery using 512^2 elements in decreasing order using logarithmic scale on the x-axis.



(a) (b)

Figure 3: Piece of a seismic image (a) and envelope obtained by Hilbert transform (b).

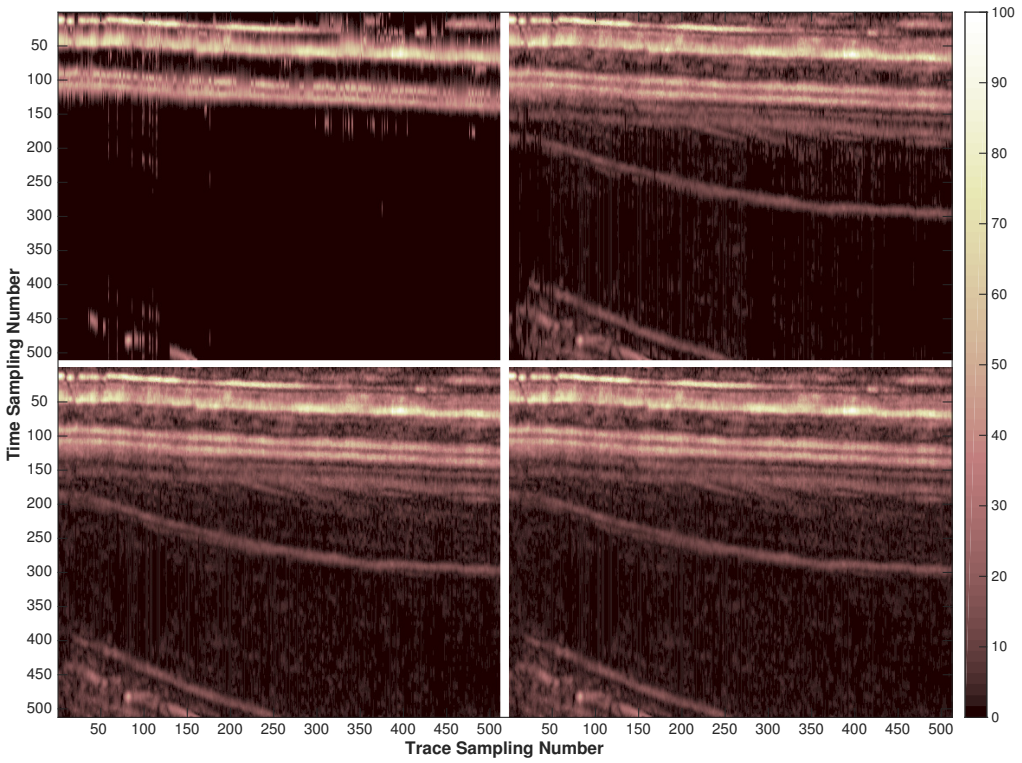


Figure 4: Reconstructed envelope using ADOMP with decreasing κ (from left-up to right-down).

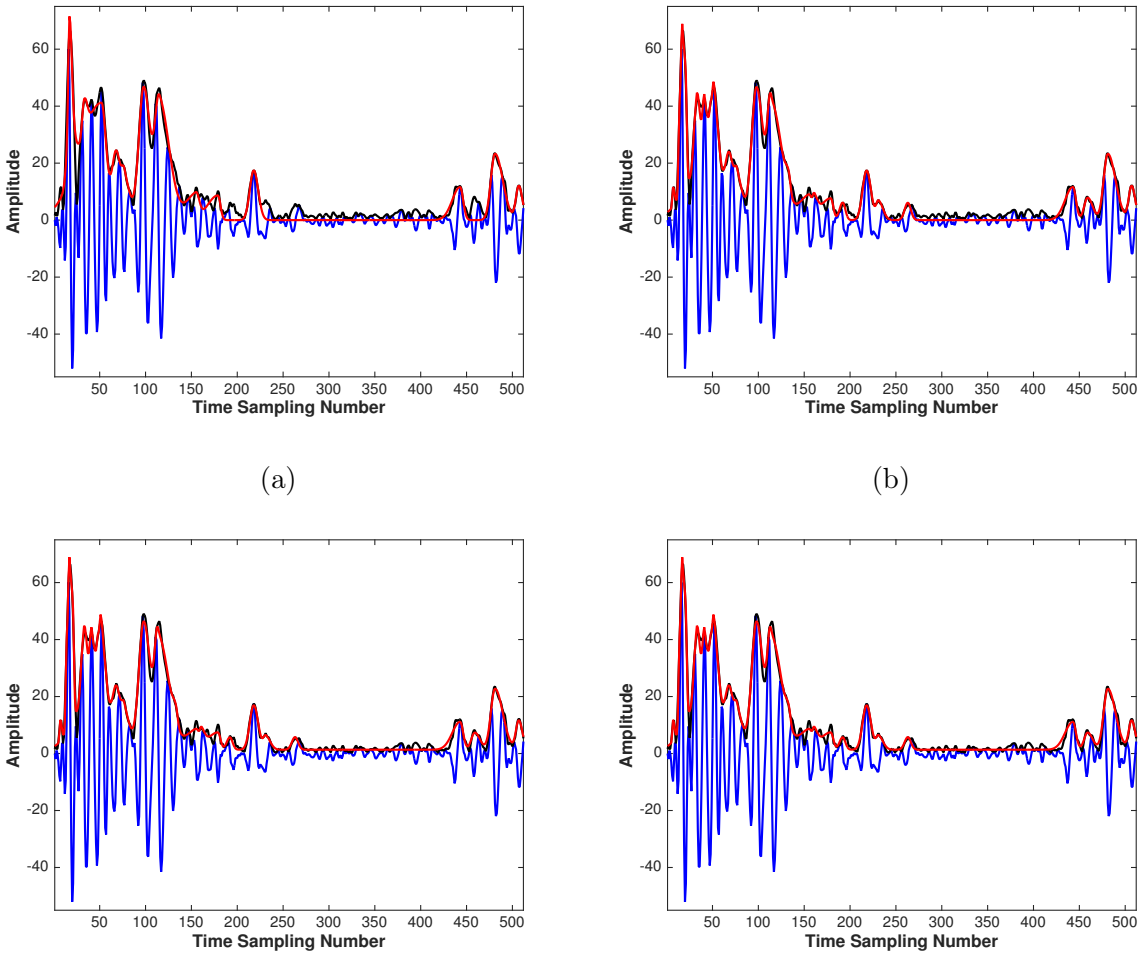
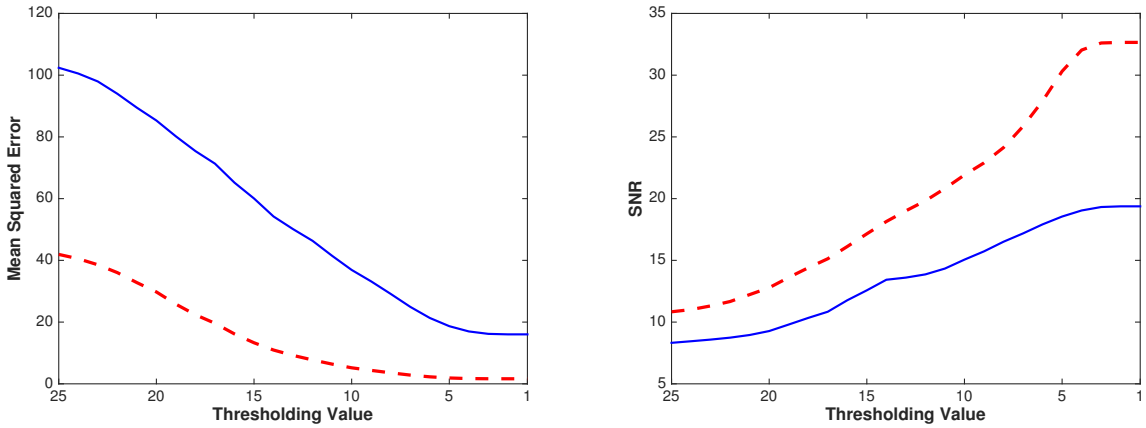
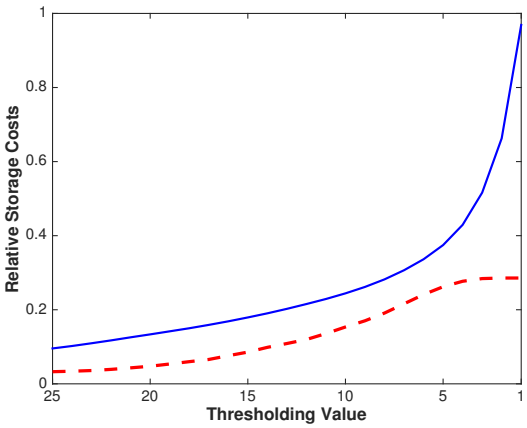


Figure 5: (a)-(d) Original data (blue), its envelope (black) and its reconstruction (red) with decreasing κ. Larger κ values lead to smoother envelope.



(a) (b)

Figure 6: Mean squared error (a) and signal-to-noise ratio (b) of the reconstructed envelope as a function in $\kappa\%$ using ADOMP (red dash line) compared to Fourier approximation using same number of coefficient as atoms used in ADOMP.



(a) (b)

Figure 7: (a) Number of elements used in ADOMP (red dash line) for approximation and (b) storage cost relative to a full image storage as a function of $\kappa\%$. The number of Fourier coefficients needed to obtain the same approximation accuracy as well as their storage cost is plotted as a comparison (blue solid line).

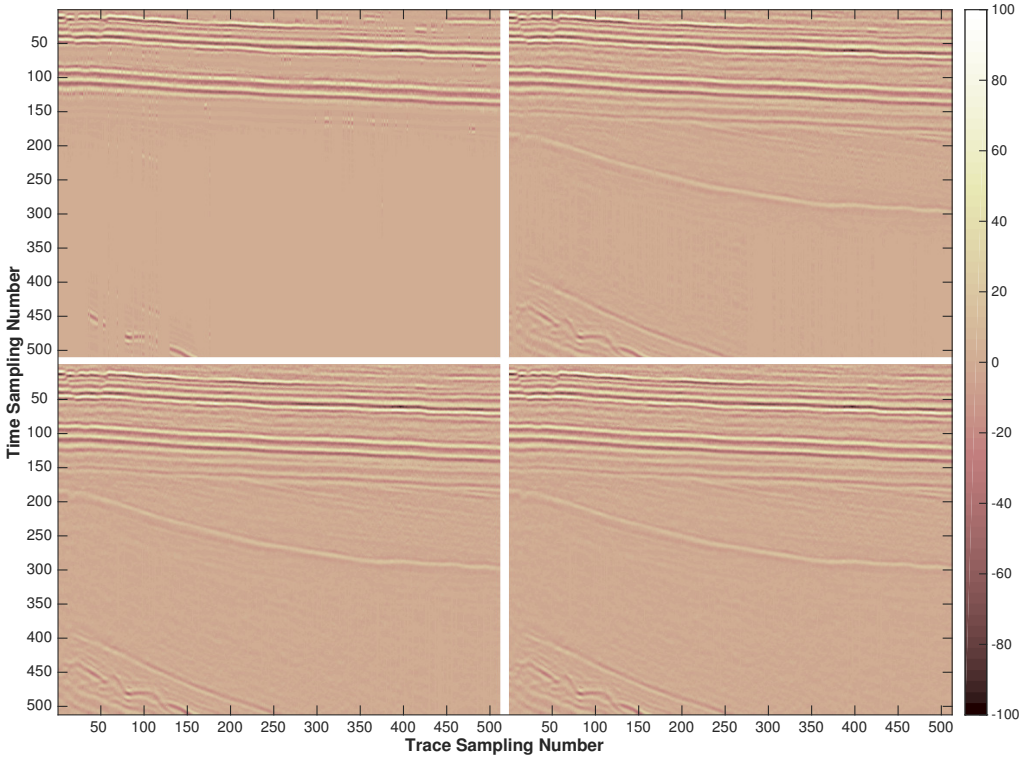


Figure 8: Approximated data using the above stated algorithm with decreasing κ.

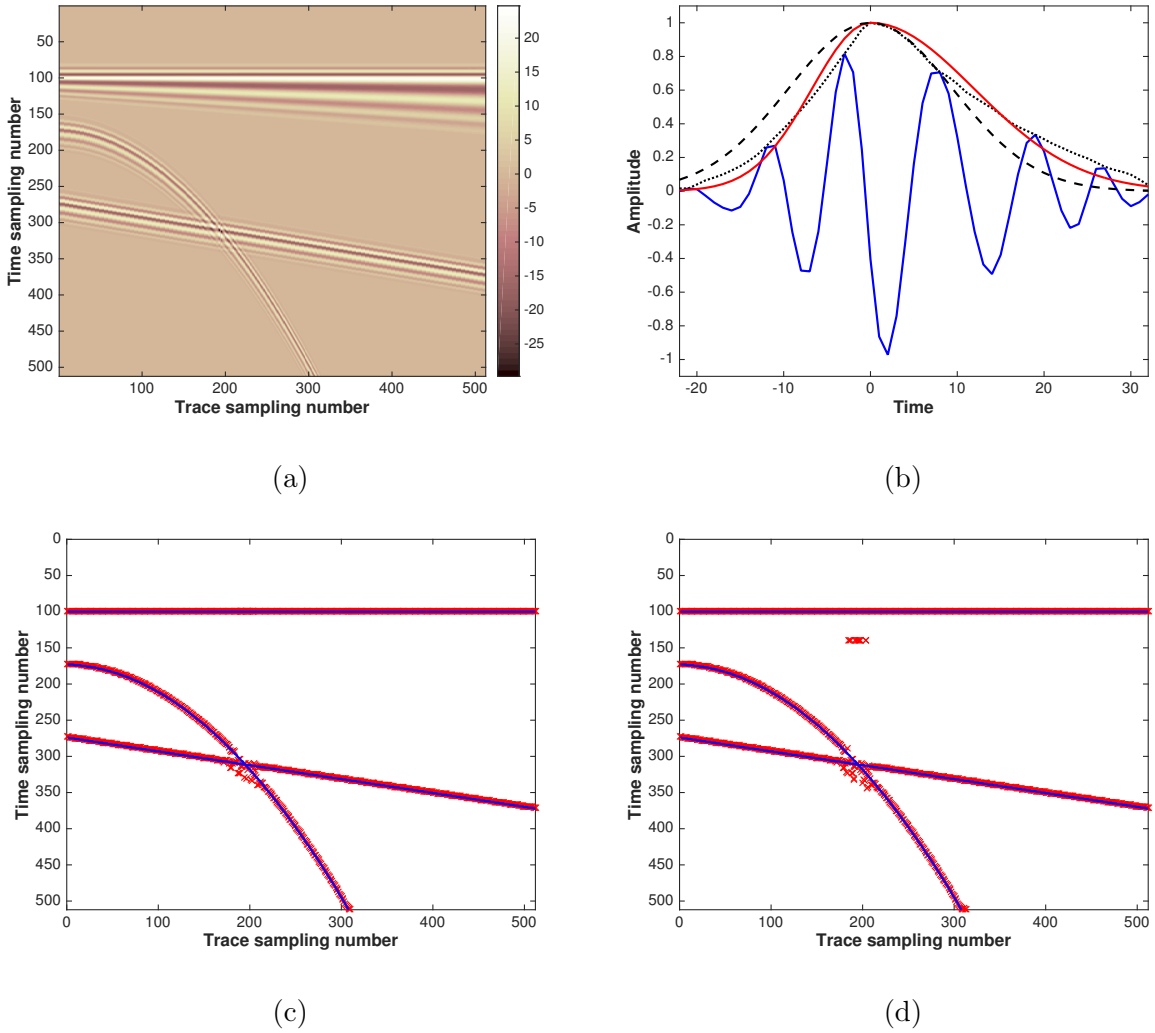
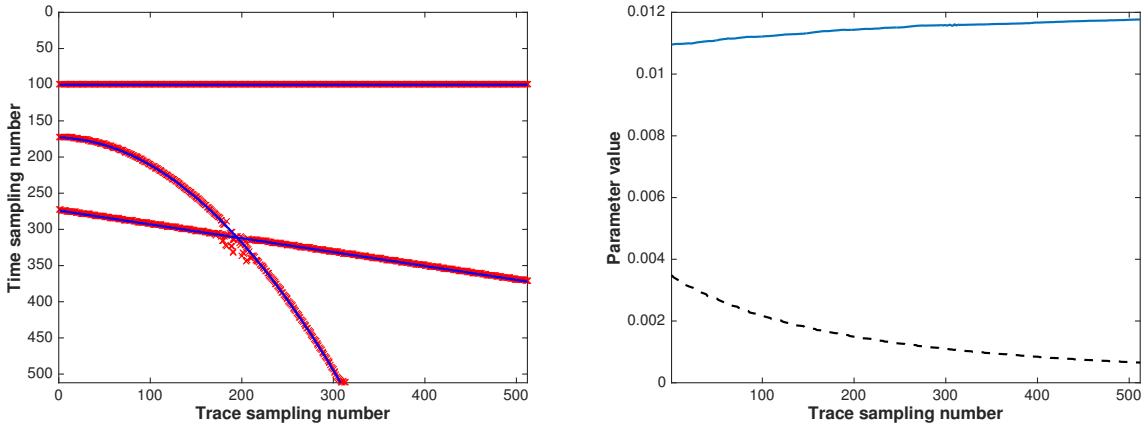
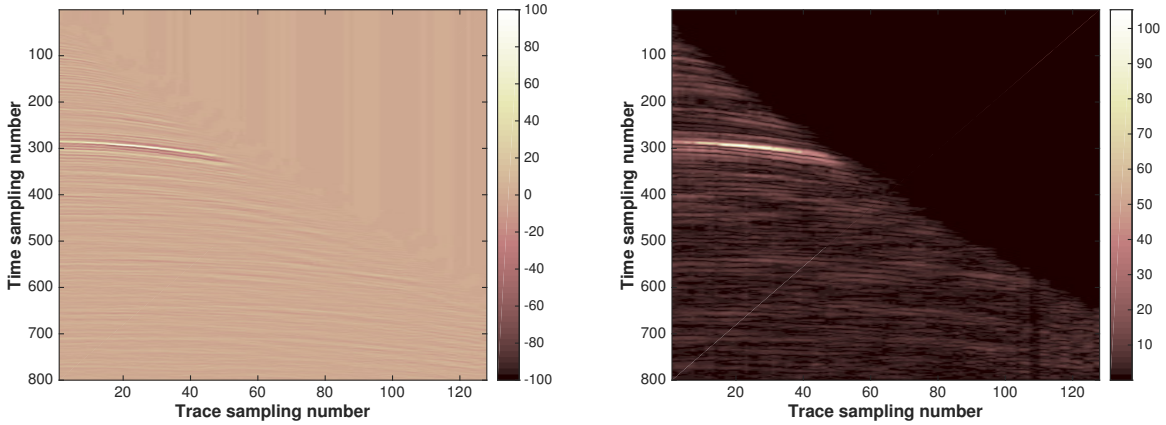


Figure 9: (a) Synthetic seismic data. (b) Seismic impulse (solid blue line) and its envelope (dotted blue line), the envelope reconstruction using GCM (dashed black line) and AGCM (solid red line). (c) and (d) Reconstructed/detection arrival times using GCM and AGCM (red lines plus furcation).



(a) (b)

Figure 10: (a) Reconstructed arrival times using AGCM with increased $\kappa = 0.12$, (b) reconstructed ascending bandwidth (solid line) and descending bandwidth (dashed line) of the first seismic response in Figure 9 (d).



(a) (b)

Figure 11: Marine shot gather (a) and its Hilbert transform (b).

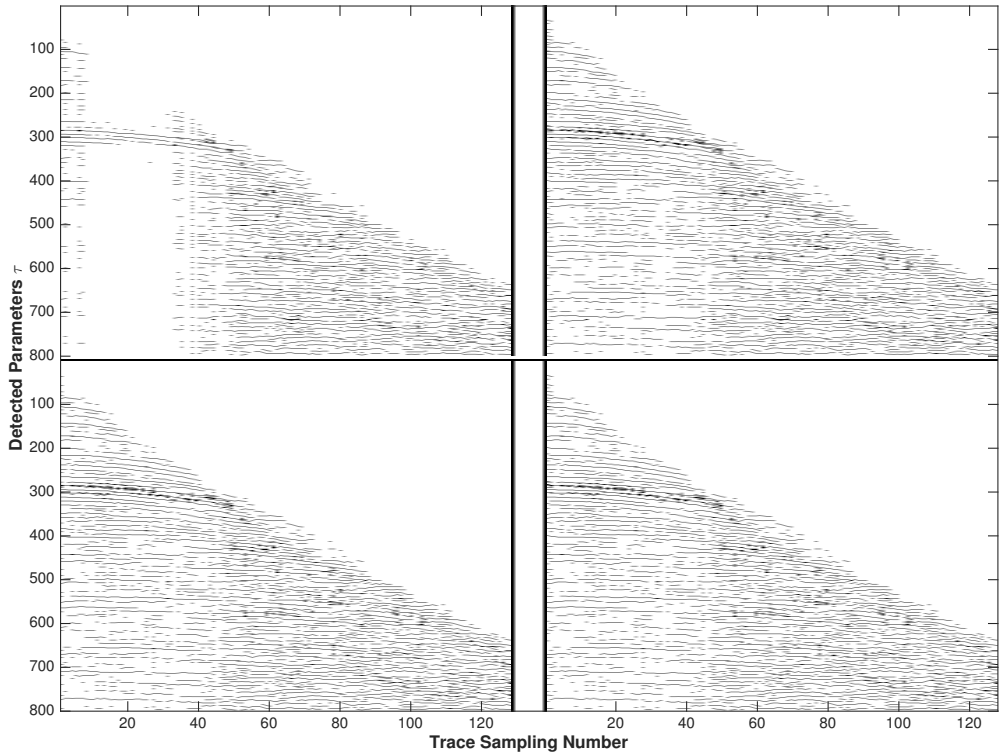


Figure 12: Detection of arrival time parameter τ using ADOMP for decreasing κ on real pre-stack data shown in Figure 11.

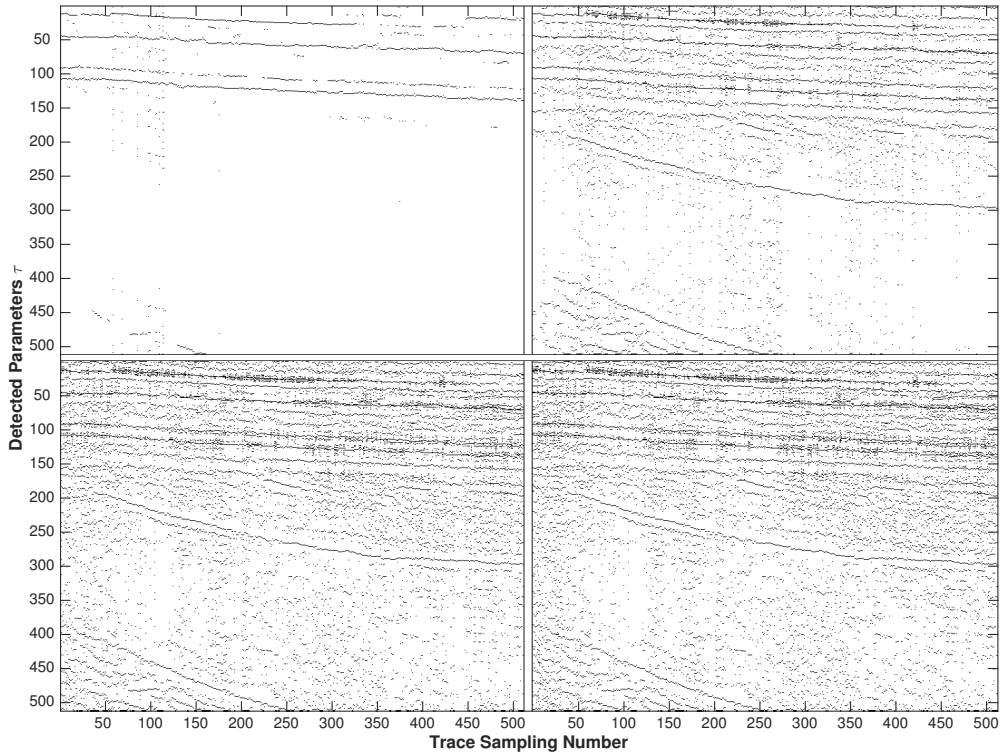


Figure 13: Detection of arrival time parameter τ using ADOMP for decreasing κ on real data shown in Figure 3.