Phase Retrieval Through Regularization for Seismic Problems

Ali Gholami
Institute of Geophysics, University of Tehran
agholami@ut.ac.ir

ABSTRACT

The most advanced data processing/imaging algorithms aim at optimizing a suitably defined cost function consisting of a data misfit term, posed in the data domain, and a regularization term. For strong phase errors in the data, however, such algorithms fail to achieve an acceptable solution because they force to match the incorrect phase information. To remedy this deficiency, we replace the time domain misfit term by a Fourier magnitude fidelity term in order to recover the desired signal from just the amplitude spectrum of the observations. Such problem is known as phase retrieval and is a subject of interest for those who cannot measure the phase information about the system under study. The proposed phase retrieval problem is solved by a fast iterative shrinkage/thresholding algorithm (FISTA) and is used to tackle two common phase problems arising in seismic processing: data recovery in the presence of residual static shifts and deconvolution with missing wavelet phase. Under sparsity constraints, both problems can be solved using just the amplitude spectrum of the observed data. In both cases, using computer simulated data and field data, the regularized phase retrieval algorithm was able to obtain better results than the conventional methods where the misfit term is posed in the time domain.

INTRODUCTION

There are a number of situations in seismic applications where the errors in the acquired wavefield are mainly phase-wise; that means, the desired wavefield is distorted by some mechanism that is known to affect the wavefield Fourier phase only. The phase of a signal carries the structural information and hence phase errors significantly limit the effectiveness of data processing and imaging algorithms.

Such situations can arise for example in land acquisition surveys where a wavefield is measured after propagation through the Earth with laterally strong near surface inhomogeneities. Under the surface-consistent assumption, these inhomogeneities act as an all pass linear phase filter and introduce random time shifts, called residual statics, to the measured traces which demand a challenging processing step to be compensated (see Gholami 2013a, Henley 2012, Ronen and Claerbout 1985 and references therein).

Another situation where we loose the signal phase arises in wavelet deconvolution problem where we attempt to increase the vertical resolution of the data. Statistical deconvolution assumes a random reflectivity for the Earth, although, the correctness of this assumption depends on the geological sequence in the area under study. However, the
random reflectivity assumption allows one to easily estimate the wavelet’s amplitude spectrum/autocorrelation, from the trace amplitude spectrum/autocorrelation (Yilmaz 2001). But in order to perform deconvolution, a good representation of the original wavelet must be employed. That means the phase information also needs to be recovered. The phase recovery is the main and most challenging step in this situation, because, generally, a signal is defined by its amplitude and phase spectra which are independent and can’t be recovered from each other. The minimum phase condition is an exception where the phase spectrum can be uniquely recovered from the amplitude spectrum by the means of Kolmogorov spectral factorization (Robinson and Treitel 1980; Kolmogorov 1939). This is why the minimum phase assumption is essential in statistical seismic deconvolution while in reality this assumption is generally accepted as being invalid (Ulrych and Sacchi 2005, page 67).

In this paper, seismic data recovery or specifically the aforementioned problems are treated by considering just the amplitude spectrum of the data. Given a problem, we look for a solution whose predicted amplitude spectrum matches with the observations amplitude spectrum, up to a constant considered for errors in the data. Obviously, the solution to such problem will be non-unique, because, for a given amplitude spectrum there will be many (up to $2^N$ for one-dimensional (1D) signals of length $N+1$, Osherovich 2011) physically meaningful signals with the same amplitude spectrum but quite different structures. However, by applying general and reasonable constraints on the solution, its structure can be determined from the amplitude spectrum. In other words, suitable regularizations can be used to make the solution uniquely determined (upto a constant shift) from the available magnitude spectrum.

In the last decade, sparsity constraints have proven to be promising tools for seismic data processing and related problems (see Gholami and Sacchi 2012; Gholami and Siahkoohi 2010, Herrmann et al. 2008a,b, and references therein). In a recent paper (Gholami 2013a), the author investigated the problem of residual statics estimation by sparsity maximization. In a second paper (Gholami 2013b), the author has shown that the desirable events contained in a wavefield can be recovered, with a high degree of accuracy, from a partial set of the wavefield frequency samples lying in the range of the source wavelet bandwidth. In (Gholami 2013b), the cost function consists of a frequency misfit term, an $\ell_2$-norm posed on the partial frequency samples in the frequency-space (f-x) domain, and a non-convex sparsity promoting term posed in the curvelet domain (Candés et al. 2006). The work of this paper is an extension of the previous algorithms for seismic data recovery by modifying the misfit term to just consider the amplitude spectrum.

Amplitude only inversion is known as phase retrieval which refers to the process of reconstructing a signal, or equivalently the Fourier phase, from the Fourier amplitude. It has a long history of applications in science and engineering, e.g. optics (Walther 1963), X-ray and crystallography (Millane 1990; Harrison 1993), and astronomy (Dainty and Fienup 1987), where the signal of interest is complex-valued but the receivers can only measure its modulus (magnitude). The phase retrieval is by itself a non-convex problem. The early solution procedures for it are simple projection-based greedy algorithms with constraints on the support of the desired signal (Fienup 1982). Recently, several more sophisticated algorithms have been developed to treat phase retrieval problem, including PhaseLift (Candés et al. 2013) and PhaseCut (Waldspurger et al. 2013). The algorithm in this paper is based on FISTA (Beck and Teboulle 2009), allowing extra constraints while preserving the computational simplicity. Contributions of this paper include the
phase retrieval usage for seismic data recovery (which to the best of the author knowledge
is not considered elsewhere), a fast algorithm developed for the generated problem based
on FISTA, and two important applications of the method: wavefield interpolation in the
presence of residual time shifts and multichannel deconvolution in the presence of residual
time shifts and missing wavelet phase.

In what follows, first, the desired phase retrieval problem is generated in the methodology
section. Then a fast solution algorithm is developed for the corresponding problem. Finally,
the procedure of applying the proposed algorithm to seismic interpolation and deconvolution
are discussed and numerical experiments are presented.

**METHODOLOGY**

Let $y \in \mathbb{R}^M$ denote the acquired wavefield (data) in the time-space (t-x) domain (note that
$y$ is a long trace obtained by stacking $n_x$ traces, each of length $n_t$, one under the other, i.e.
$M = n_x \times n_t$), then the f-x domain representation of $y$ can easily be computed using the
discrete Fourier transform (DFT) formula applied to each trace:

$$
\hat{y}[\omega] = \frac{1}{\sqrt{n_t}} \sum_{t=0}^{n_t-1} y[n_t \lfloor \omega/n_t \rfloor + t] \exp(-i2\pi t \omega/n_t), \quad (1)
$$

with $i^2 = -1$ and $\omega = 0, 1, \cdots$ is the frequency index. In matrix notation, equation (1) can
be written as $\hat{y} = F y$, where the unitary matrix $F$ denotes forward f-x operator. In seismic
applications, $\hat{y}$ is generally complex valued and hence

$$
\hat{y}[\omega] = \text{Re}[\omega] + i \text{Im}[\omega],
$$

$$
|\hat{y}[\omega]| \exp(i\phi[\omega]) \quad (3)
$$

where $\text{Re}[\omega]$ and $\text{Im}[\omega]$ are the real and imaginary parts of $\hat{y}[\omega]$, respectively, $|\hat{y}[\omega]|$ is its
modulus, and $\phi[\omega]$ is its phase. Static shifts affect phase-only information (see Gholami,
2013a), let $\Delta \phi$ be the phase error due to the static shifts, then equation (3) reads as

$$
\hat{y}[\omega] = |\hat{y}[\omega]| \exp(i\phi[\omega] + i\Delta \phi[\omega]) \quad (4)
$$

where $\phi_0 = \phi - \Delta \phi$ denotes the phase of the static-corrected wavefield $y_0$. Accordingly,
the static correction problem in view of phase retrieval reads as

$$
\text{find } y_0 \quad \text{subject to } |Fy_0| = |\hat{y}|,
$$

where the absolute function is applied component-wise. This problem is non-convex with
non-unique solution, thus, additional constraints are required to make the solution uniquely
determined from the magnitude data (see Candès et al., 2013; Waldspurger et al., 2013).
Beside of this, in seismic data processing, usually, we are interested in a surrogate signal
$x \in \mathbb{R}^N$ which is connected to $y_0$ by an ill-posed linear relationship $y_0 = Gx$, where
$G \in \mathbb{R}^{M \times N}$. The word ill-posed means that matrix $G$ filters out some information contained
in the signal $x$ which is not recoverable solely from $y_0$. This makes determination of
the desired signal doubly ill-posed. Due to this, determination of $x$ requires more strong
structural constraints defined in an appropriate transform space so that the desired signal
can be regularized, i.e., matrix $G$ can be considered as $ΨC^{-1}$ with $Ψ$ the filter operator and $C$ a (curvelet) transform. To summarize, $x$ is defined as a solution of the regularized phase retrieval problem

$$\tilde{x} = \arg \min_x \left\{ \frac{1}{2} \| b - |FGx| \|_2^2 + \tau \text{Reg}(x) \right\}, \quad (6)$$

where $b = |\hat{y}|$, Reg is the regularization function and its job is to provide the necessary structural constraints to the problem (Gholami and Hosseini, 2011), and regularization parameter $\tau$ balances between the amount of supplied regularization and magnitude fidelity term. We will discuss about the Reg function in more details in the next sections. As a comparison with the existing data recovery approaches, problem (6) without the modulus function $|·|$ reduces to the conventional data recovery algorithms operating in the time domain. In this case, using a frequency mask operator to restrict the undesired frequencies changes the problem to the algorithm presented in (Gholami, 2013b). In the following a fast algorithm is presented based on FISTA to solve (6).

**Phase Retrieval by FISTA**

The nonregularized case of (6) corresponds to $\tau = 0$. The solution then reduces to

$$\tilde{x} = \arg \min_x \frac{1}{2} \| b - |FGx| \|_2^2, \quad (7)$$

which is the least squares solution to the phase retrieval problem. The right-hand-side of (7) is non-linear in $x$ and hence the usual approach for its minimization is iterative. Starting from an initial estimate $x^0$, a gradient descent algorithm with an appropriate stepsize $\gamma > 0$ constructs a sequence $\{x^j\}_{j \in \mathbb{N}}$ that converges to a minimizer of the problem with the following update scheme

$$x^{j+1} = x^j - \gamma \frac{\partial}{\partial x} \| b - |FGx| \|_2^2 \bigg|_{x = x^j}. \quad (8)$$

The key in each update step is therefore calculation of the gradient of the misfit term. Let $z = FGx$ and $f(z) = \frac{1}{2} \| b - |z| \|_2^2$, then using the chain rule of derivatives we have

$$\frac{\partial}{\partial x} \frac{1}{2} \| b - |FGx| \|_2^2 = \frac{\partial z}{\partial x} \frac{\partial f}{\partial z}. \quad (9)$$

A simple computation shows that

$$\frac{\partial z}{\partial x} = GTF^{-1} \quad (10)$$

where $T$ denotes the transpose and

$$\frac{\partial f}{\partial z[\ell]} = \sum_k (|z[k]| - b[k]) \frac{\partial}{\partial z[\ell]} |z[k]| \quad (11)$$

$$= (|z[\ell]| - b[\ell]) \text{sgn}(z[\ell]) \quad (12)$$

$$= (|z[\ell]| - b[\ell]) \frac{z[\ell]}{|z[\ell]|}. \quad (13)$$
where \( \text{sgn} \) is the signum function. Therefore, we have

\[
\frac{\partial f}{\partial z} = \text{diag} \left( 1 - \frac{b}{|z|} \right) z, \tag{14}
\]

where \( \text{diag}(\mathbf{v}) \) is a diagonal matrix, i.e., \( \{\text{diag}(\mathbf{v})\}[i,i] = v_i, 1 \in \mathbb{R}^M = [1, ..., 1]^T \) and with notation abuse the division operation is done component-wise.

Putting equations (10) and (14) into equation (9) and substituting for \( z \), we arrive at the desired update expression

\[
x^{j+1} = x^j - \gamma G^T F^{-1} \text{diag} \left( 1 - \frac{b}{|FGx^j|} \right) FGx^j. \tag{15}
\]

A simple computation shows that restriction \( \gamma^{-1} > \rho(G^T G) \) (the maximum singular value of \( G^T G \)) is necessary for the convergence of this sequence. However, the solution resulted from (15) has no practical importance without regularization. One of the most simple methods for solving regularized problem (6) is in the class of the iterative shrinkage/thresholding algorithms (ISTA) (Daubechies et al., 2004) where each iteration consists of step (15) followed by thresholding the coefficients of the output as

\[
x^{j+1} = \text{prox}_\tau \left( x^j - \gamma G^T F^{-1} \text{diag} \left( 1 - \frac{b}{|FGx^j|} \right) FGx^j \right). \tag{16}
\]

where

\[
\text{prox}_\tau (\mathbf{v}) = \arg \min_x \left\{ \frac{1}{2} \| \mathbf{v} - \mathbf{x} \|_2^2 + \tau \text{Reg}(\mathbf{x}) \right\}, \tag{17}
\]

is the proximity operator of the \( \text{Reg} \) function and is easily solvable when \( \text{Reg} \) is separable (Combettes and Pesquet, 2011). Clearly, the ISTA is a first-order method (only needs function values and gradient evaluations), so it is suitable for solving large-scale problems (note that each iteration requires one multiplication by \( G, F \) and their adjoints). Despite its simplicity, however, ISTA can suffer from slow convergence rate. Fast ISTA (FISTA) was proposed by Beck and Teboulle (2009) to accelerate the algorithm using a backtracking strategy which ensures that each iterate makes enough progress toward the solution. Instead of the current estimate \( x^j \), FISTA uses a carefully selected point \( \mathbf{u} \) which is a very specific linear combination of the two last estimates \( x^j \) and \( x^{j-1} \). This simple change allows preserving the computational simplicity of ISTA but improving the rate of its convergence (see Algorithm 1). Note, however, that more sophisticated and faster-converging solvers exist today, but would add complexity to the implementation.

**APPLICATIONS**

In this section the proposed phase retrieval algorithm is applied to seismic data recovery. The primary applications of interest are wavefield interpolation in the presence of residual time shifts and source signature deconvolution when only the amplitude information of the wavelet is available and when there are relative static errors between traces. The results of the phase retrieval algorithm are compared with those obtained from similar algorithms where the misfit term is posed in the time domain, and a carefully tuned benchmark solution.

In practice, an appropriate regularization parameter \( \tau \) is needed in order to apply the algorithm. For interpolation problem it is more difficult to determine such \( \tau \) when no
Algorithm 1: FISTA Algorithm for Phase Retrieval Problem (6)

1. Initialize: given $x^0$, set $\alpha^0 = 1$ and $u^0 = x^0$,
2. for $j \leftarrow 1$ to maximum iteration do
3. \[ \gamma \leftarrow \frac{1}{\rho(G^T G)^{\frac{1}{2}}} \]
4. \[ x^{j+1} \leftarrow \text{prox}_{\gamma \tau} \left( u^j - \gamma G^T F^{-1} \text{diag} \left( 1 - \frac{b}{|FGu|^\tau} \right) FGu^j \right) \]
5. \[ \alpha^{j+1} \leftarrow \frac{1 + \sqrt{1 + 4 \alpha^j \alpha^j}}{2} \]
6. \[ u^{j+1} \leftarrow x^{j+1} + \frac{\alpha^j - 1}{\alpha^{j+1}} (x^{j+1} - x^j) \]
7. end

information about the noise level in the data is available. For simplicity and accurate comparison with other methods, here, it is selected as a minimizer of the Euclidean distance between the estimate and the original model which is assumed to be known. But, for deconvolution problem, $\tau$ is determined via the generalized cross validation (GCV) score which has successfully been used for both non-blind and blind deconvolution problems (Gholami and Sacchi, 2012, 2013). For phase retrieval the GCV score reads as

\[
\text{GCV}(\tau) = \frac{M^{-1}||b - |FGx(\tau)||_2^2}{(1 - CM^{-1} \times nnz\{x(\tau)\})^2},
\]

(18)

where $x(\tau)$ is the solution obtained from Algorithm 1 for $\tau$, $nnz$ is a function which returns the number of non-zero components in its argument, and $C \geq 1$ is a stabilizing parameter.

In this paper, we use $C = 1$. The optimum $\tau$ is that minimizes the GCV score. Interestingly, it is observed that the GCV score still works satisfactorily for deconvolution problem with static errors.

Wavefield Interpolation

In seismic exploration, due to limitations during seismic surveys, it is common to incompletely sample the wavefield in spatial direction. Therefore, reconstruction of the original wavefield from the incomplete set of traces is required in order to perform the subsequent processing steps. Furthermore, the recorded wavefield is usually subjected to static shifts which decrease the performance of the reconstruction. Static shifts decrease the spatial coherency of seismic events and hence decrease the sparsity of the data. This characteristic allow statics to be compensated for by a sparsity maximization approach (Gholami, 2013a). Stanton and Sacchi (2012) simultaneously estimated residual statics when interpolating 5D data by using a Projection Onto Convex Sets (POCS) technique, and Stanton et al. (2013) tackled the problem by maximizing sparsity in an alternating fashion. The algorithm in this paper does the same job for interpolation problem but with a quite different approach. Here, the algorithm simply ignores the static errors because it works with the amplitude information.

Curvelet domain regularization is a common and successful approach for seismic interpolation (Gholami, 2013b; Herrmann et al., 2008b,a). Regarding this approach, $G = \Psi C^{-1}$, consist of the inverse curvelet transform operator $C^{-1}$ and a sampling operator $\Psi$, a binary
(0 or 1) matrix whose operation on a wavefield is sub-sampling in the spatial direction, and \( \text{Reg} \) is a sparsity-promoting functional. In this paper, we use \( \text{Reg}(x) = ||x||_1 \) for its simplicity, even though other more appropriate functionals can also be used (see Gholami [2013b] Gholami and Hosseini [2011]). For this choice, the desired seismic wavefield is recovered in the curvelet domain as

\[
\hat{x} = \arg \min_x \left\{ \frac{1}{2} ||b - |F \Psi C^{-1} x||^2_2 + \tau ||x||_1 \right\},
\]

(19)

with the proximity operator

\[
\{\text{prox}_\tau(v)\}[n] = \arg \min_x \left\{ \frac{1}{2} (v[n] - x)^2 + \tau |x| \right\} = \text{sgn}(v[n]) \max(|v[n]| - \tau, 0),
\]

(20)

where \( b \) contains the amplitude spectrum of the acquired traces. By equation (19) we look for a sparse wavefield in the curvelet domain whose predicted amplitude spectrum at the known trace locations matches the amplitude spectrum of the observed traces. Given \( y \), Algorithm 1 initialized with \( x^0 = C \Psi^T y \) provides an approximate \( x^j \) at \( j \)th iteration.

In order to improve the quality of the reconstruction, the algorithm is followed by a phase correction step: at first the time shifts of the observed traces are estimated by cross-correlating each trace of \( y \) with the corresponding trace in \( \Psi C^{-1} x^j \) and picking the maximum index set, then the time shifts are applied to \( y \) to obtain \( P y \). The phase correction step consists of solving the following program:

\[
\hat{x} = \arg \min_x \left\{ \frac{1}{2} ||P y - \Psi C^{-1} x||^2_2 + \tau ||x||_1 \right\},
\]

(21)

where the first term matches phase and amplitude between the statics corrected data and \( x \), after inverse curvelet transform and subsampling, and the second term imposes sparseness in the curvelet domain. Problem (21) is solved by Algorithm 1 with step 4 replaced by

\[
x^{j+1} \leftarrow \text{prox}_{\gamma \tau} \left( u^j - \gamma C \Psi^T (\Psi C^{-1} u^j - P y) \right).
\]

It is observed that only a few iterations are sufficient for the phase correction step when it is started with \( x^j \).

Stanton et al. [2013] also proposed a similar problem as (21) to be optimized for both \( P \) and \( x \). They used the Radon transform for regularization and an alternating minimization approach to estimate both \( P \) and \( x \). Comparing to their approach, here, the static operator \( P \) is estimated more rapidly from the solution of (19).

The functionality of the proposed method is demonstrated by two synthetic and real examples. First, a common midpoint gather (Figure 1a) has been simulated with the time and space sample intervals 4 milliseconds and 12.5 meters, respectively. The traces of the resulting gather have been shifted with Gaussian distributed random errors of standard deviation 12 milliseconds. A small amount of noise has also been added to obtain the gather shown in Figure 1b. The incomplete and inaccurate wavefield (Figure 1c) has been obtained by 2-fold randomly down-sampling (in spatial direction) the gather shown in Figure 1b. Figure 1d shows the gather reconstructed by posing the misfit term in the time domain (problem (21) for \( P \) as identity matrix, and \( \tau = 0.01 \), maximum iteration = 50 in Algorithm 1). Figure 1e is obtained after 10 iterations of alternatively solving problem (21) for \( x \) and
where nine of them were performed using $\tau = 0.2$ and \textit{maximum iteration} = 10 while the last iteration was performed using $\tau = 0.01$ and \textit{maximum iteration} = 50. The phase retrieval algorithm produced the gathers shown in Figure 1f (using $\tau = 0.01$ and \textit{maximum iteration} = 50) and Figure 1g (after 20 iterations of phase correction step). Difference sections between the reconstructed results and the ideal input data are also shown in the last row of Figure 1.

Some points deserve to be mentioned about the recovered results: 1) Amplitude only inversion generated better result (Figure 1f) compared to the conventional amplitude and phase inversion (Figure 1d). 2) The phase correction step improved the quality of the reconstructions (Figure 1e compared with Figure 1d and Figure 1j compared with Figure 1l). 3) The result of the proposed phase retrieval method (Figure 1g) is very similar to the carefully tuned benchmark solution (Figure 1e) with the difference that Figure 1g is obtained more quickly. 4) Some of the curved events in all reconstructions appear as piecewise linear events, instead of smoothly curved events. This can be explained by the fact that a linear event is sparser in the curvelet domain than a curved event and sparse regularization in the curvelet domain will force the solution towards linear events. Note that it is difficult to represent curvatures sparsely, while maintaining sparsity for more general (complicated) signals, under non-adaptive transforms. A possible way to improve this issue is regularization of the solution in a more adaptive domain (see Gholami and Siahkoohi, 2010, and references therein). 5) The residual signal remained in the difference sections corresponding to Figures 1e and 1g are mainly due to the linear approximation of the curved events leading to errors in the statics estimation.

The same procedure discussed above with the same parameters has been applied to a real stacked section corrupted by simulated random statics of standard deviation 16 milliseconds followed by randomly 2-fold down-sampling. Note that, normally we do not try to solve for statics after stack, this section (Figures 2a) has been selected for various structures contained in it. The results of interpolation via different methods are demonstrated in Figure 2. Again the phase correction step improved the quality of the reconstructions. Some curvelet-like artifacts are present in the estimates, specifically in Figure 2f, which arise when forcing a solution (including discontinuous events) to be sparse in the curvelet domain (see Candès and Guo, 2006). Generally, however, it can be seen that for both synthetic and real tests, the proposed algorithm could handle the residual static problem and generated satisfactory reconstructions.

Wavelet Deconvolution

Wavelet/source deconvolution is a long standing problem in seismic applications. It is used primarily to increase temporal resolution and an appropriate wavelet is needed to perform it. In statistical seismic deconvolution the wavelet is estimated from the data, however, it is more easy to estimate its amplitude spectrum (Yilmaz, 2001) and the phase is usually missed or is very inaccurate. This makes deconvolution of mixed-phase wavelets more problematic. Many research has been done and is still in process on extracting a reliable seismic wavelet from the data in order to perform deconvolution (e.g. Ulrych and Sacchi, 2005; Van der Baan and Pham, 2008; Edgar and Van der Baan, 2011; Gholami and Sacchi, 2013). Here, it is shown that the proposed phase retrieval algorithm can appropriately be used for deconvolution of mixed-phase wavelets.
Figure 1: Synthetic data interpolation in the presence of static shifts. Original gather (a) subjected to static shifts plus additive random noises (b), and followed by 2-fold randomly down-sampling in spatial direction (c). (d) The result of conventional interpolation. (e) The result obtained after 10 alternating steps of interpolation followed by static correction. The result of the phase retrieval problem before (f) and after (g) phase correction. The last row shows the difference signals between the original data (a) and the estimates. The RMS amplitude is defined as the square root of the average of the squares of the sample values, multiplied by 500.

Figure 2: Real data interpolation in the presence of static shifts. Explanation is the same as that given for Figure 1.
Under some conditions such as stationarity of the source wavelet, for deconvolution problem $G$ is factored as $F^{-1} \text{diag}(\hat{w}) F$ (Gholami and Sacchi, 2012), where, for a single channel case, $\hat{w}$ is the vector of wavelet Fourier transform and for multichannel deconvolution, $\hat{w}$ is a tall vector obtained by replication of $n_x$ copies of the wavelet Fourier transform. Using just the amplitude information of the wavelet we have $G = F^{-1} \text{diag}(|\hat{w}|) F$ and substituting it into the step 3 of Algorithm 1 produces

$$x^{t+1} \leftarrow \text{prox}_{\gamma \tau} \left( F^{-1} \text{diag}(\hat{d}) Fu \right)$$

where

$$\hat{d} = 1 - \gamma |\hat{w}| \circ \left( |\hat{w}| - \frac{b}{|Fw|} \right),$$

with $\circ$ denoting the Hadamard product operator.

**Single channel deconvolution**

For single channel deconvolution a sparse reflectivity model is desired. Therefore, $\text{Reg}(x) = ||x||_1$ and the corresponding proximity operator is computable via (20). However, other assumptions like whiteness can also be made on the reflectivity by employing an appropriate proximity operator (see Gholami and Sacchi, 2012). The effectiveness of the proposed method in deconvolution problem is illustrated with several examples.

**Example 1:** An appropriate wavelet, resembling the original waveform, is required in order to perform deconvolution. In practice, however, the source is usually unknown or is known only approximately. Suppose that the source function is a maximum-phase ricker wavelet (Figure 3a) while we use its minimum-phase variant (Figure 3b) for deconvolution. The conventional $\ell_1$-norm deconvolution failed to compress the wavelet (Figure 3c), because it requires the reflectivity to match the source structure when passed through the minimum-phase filter (Figure 3e). In other words, it forces the reflectivity to be sparse while predicting the source amplitude and phase spectra (Figure 3g). But the proposed method compressed the wavelet into a spike (Figure 3d) without any concern about its structure (Figure 3h). For both methods the results were obtained with 500 iterations of FISTA and $\tau = 0.1$.

**Example 2:** Now the functionality of the proposed phase retrieval algorithm is tested by deconvolving different sources without phase information. A sparse reflectivity series, shown at the top of figure 4, has been convolved with four different mixed-phase wavelets (shown at the left column). The resulting random noise added traces are depicted at the right column of the same figure. It was assumed that the phase of the wavelets are missed and hence the zero-phase variant of the original wavelets were used to deconvolve the generated traces. The best results (in the mean square error sense, Gholami and Sacchi (2012)) of joint amplitude and phase deconvolution and amplitude only deconvolution, shown in figure 5 confirm that the time domain deconvolution failed to recover an acceptable solution while the proposed method was able to recover the original reflectivity series with a high degree of accuracy.

A further experiment has been designed to show how sensitive the proposed method is to errors in the estimated wavelet spectra and how it deals with close together spikes which is the characteristic of thin layers. A synthetic reflectivity model of 256 samples (sample interval 4 milliseconds) has been simulated including several pairs of opposite sign spikes.
where the distance between each pair varies from 2 to 14 samples. A wavelet with rough spectrum (Figure 6, top left), taken from real data (Figure 10a), has been convolved with the reflectivity model and a small amount of random noise has been added to generate a synthetic trace (Figure 6, top right). Then the proposed method was used to deconvolve the trace using different estimates of the wavelet spectrum. The spectra are obtained by smoothing the original spectrum via convolution with Hamming windows. The results are shown in Figure 6. As seen, the method was able to recover the reflectivity model satisfactorily using just an approximation of the wavelet spectrum. It can also be seen that close together spikes couldn’t be resolved well and that the solution degrades in a non-uniform manner when smoothing the spectrum.

Figure 3: A minimum-phase wavelet (b) is deconvolved from a maximum-phase wavelet (a) via conventional (left column) and proposed (right column) algorithms. Obviously, the conventional sparse deconvolution failed to compress the wavelet into a spike (c) because it requires the reflectivity to match both the amplitude and phase spectra ((e) and (g)). But, the proposed method compressed it into a spike (d) because it just matched the amplitude spectrum ((f) and (h)).

**Multichannel deconvolution**

In practice, we deal with many traces which are discrete samples of a three-dimensional (3D) wavefield. The desired reflectivity model is also a 3D function. It describes hyper-surfaces corresponding to the boundaries of geological layers. Therefore, trace by trace deconvolution results to a reflectivity model which is clearly suboptimal, because it doesn’t take into account the spatial correlation of the information contained in the data (see Gholami and Sacchi 2013). This issue is more problematic when the data are subjected to static shifts. Although it is possible to first resolve the statics and then perform deconvolution, here, we show that both the static and wavelet phase problems can jointly be addressed by the
Figure 4: Phase retrieval for wavelet deconvolution. Top panel shows a sparse reflectivity model which is convolved with four different mixed-phase wavelets (left column) then contaminated with random noises to generate the synthetic traces shown at the right column. The results of deconvolution are shown in Figure 5.

Figure 5: Phase retrieval for wavelet deconvolution. Estimated reflectivity series from deconvolution of the traces shown in Figure 4 by $\ell_1$-norm deconvolution. Left column results are obtained by inverting both the wavelet amplitude and phase spectra while right column results are obtained by amplitude only inversion.
Figure 6: This figure shows the performance of the proposed method with respect to errors in the wavelet spectrum used for deconvolution. A minimum phase wavelet with rough spectrum (top left) has been convolved with a simulated reflectivity, where the distance between its spikes decreases with time, then contaminated with random noises to generate a synthetic trace (top right). The trace has been deconvolved using different estimates of the wavelet spectrum. Left column shows the estimated spectra and right column shows the resulting reflectivity models. Location and amplitude of the true spikes are indicated by circles.
proposed algorithm for multichannel deconvolution.

The algorithm for multichannel deconvolution is the same as that described in equations (22) and (23) for single channel case. To consider the spatial correlation between traces a new regularizer was defined by Gholami and Sacchi (2013):

$$\text{Reg}(x) = (1 - \beta)||x||_1 + \beta \text{TV}_2(x) \quad (24)$$

where $\text{TV}_2(x)$ is the second-order total variation functional defined as

$$\text{TV}_2(x) = \sum_i \sqrt{(\Delta_{11}^i x)^2 + 2(\Delta_{12}^i x)^2 + (\Delta_{22}^i x)^2} \quad (25)$$

and $0 \leq \beta \leq 1$ is the second regularization parameter which balances between the two regularization terms or between the temporal sparsity and lateral continuity. Larger $\beta$ leads to more continuous and less sparse events and vice versa. Here, its value is tuned by trial and error. The linear operators $\Delta_{11}^i$, $\Delta_{12}^i$, and $\Delta_{22}^i$, where the superscripts show the direction of differentiation, correspond to second-order differences at pixel $i$ and are defined such that

$$\Delta_{11}^i x = x[i] - 2x[i + n_t] + x[i + 2n_t], \quad (26)$$
$$\Delta_{22}^i x = x[i] - 2x[i + 1] + x[i + 2], \quad (27)$$
$$\Delta_{12}^i x = x[i] - x[i + 1] - x[i + n_t] + x[i + n_t + 1] \quad (28)$$

(see Lefkimmiatis et al., 2012; Gholami and Sacchi, 2013, for the properties of $\text{TV}_2$). It helps to apply smoothing along reflection events and when combined with the $\ell_1$-norm penalty, as in equation (24), brings the advantage of temporal sparsity and lateral continuity of the estimate. Therefore, we deal with Algorithm 1 and the following proximity operator for multichannel deconvolution in the presence of static shifts and missing wavelet phase.

$$\arg\min_x \left\{ \frac{1}{2}||v - x||_2^2 + \tau(1 - \beta)||x||_1 + \tau \beta \text{TV}_2(x) \right\}. \quad (29)$$

By approximating the absolute function in defining the $\ell_1$-norm as $|x| \approx \sqrt{x^2 + \epsilon}$ for small stabilizing parameter $\epsilon > 0$ (e.g. $10^{-15}$), problem (29) can be solved via a gradient based algorithm. Differentiating it with respect to $x$ and setting the result equal to zero leads to an iteratively reweighed least squares (IRLS) algorithm summarized in Algorithm 2. In the algorithm $D^T \in \mathbb{R}^{N \times 4N} = [I, \Delta_{11}, \sqrt{2}\Delta_{12}, \Delta_{22}]$, where $I$ denotes the identity matrix and matrices $\Delta_{11}$, $\Delta_{22}$, and $\Delta_{12}$ are defined in equations (26)-(28) by their rows, and

$$Q^j \in \mathbb{R}^{4N \times 4N} = \begin{bmatrix} Q_{11}^j & 0 & 0 & 0 \\ 0 & Q_{22}^j & 0 & 0 \\ 0 & 0 & Q_{22}^j & 0 \\ 0 & 0 & 0 & Q_{22}^j \end{bmatrix} \quad (30)$$

where $Q_{11}^j$ and $Q_{22}^j$ are diagonal matrices whose diagonal entries are defined as

$$Q_{11}^j[i, i] = \frac{\tau(1 - \beta)}{\sqrt{(x^j[i])^2 + \epsilon}} \quad (31)$$
and
\[
Q_{22}[i,i] = \frac{\tau\beta}{\sqrt{(\Delta_1^{11}x) + 2(\Delta_1^{12}x)^2 + (\Delta_2^{22}x)^2 + \epsilon}}. \tag{32}
\]

Since the system of equations in step 3 of Algorithm 2 can be very large, it may not be possible to solve it via direct methods. Instead, conjugate gradient (CG) algorithm can be used to solve it iteratively. To increase the convergence rate of CG, the system is preconditioned by a circulant preconditioner \( M \) as the best circulant approximation of the coefficient matrix \( I + D^TQ^D \). In this case, \( M \) is obtained from \( I + D^TQ^D \) by replacing the diagonal elements of \( Q^{11} \) and \( Q^{22} \) with the average value of the diagonal (Lefkimmiatis et al., 2012). Assuming periodic boundary conditions for the reflectivity model, the difference operators are circulant and hence \( M \) is diagonalizeable under Fourier operator and more easily applicable.

**Algorithm 2:** IRLS Algorithm for Solving (29)

1. Initialize: Construct \( D \) as defined in the text and set \( Q^0 = I \).
2. for \( j \leftarrow 1 \) to maximum iteration do
   3. Solve \( (I + D^TQ^D)x^{j+1} = v \)
   4. Compute \( Q^{j+1} \) according to Equations (30)-(32).
5. end

**Example 1:** The reflectivity model, shown in Figure 7a, was used to show the performance of the proposed multichannel deconvolution in the presence of phase problems. It has been convolved with a mixed-phase wavelet shown at the top right corner of the figure. The resulting section (Figure 7b) has been subjected to some random static errors and contaminated by additive white Gaussian random noise, with signal-to-noise-ratio (SNR) 15 decibels, to produce Figure 7c. It was assumed there is no information about the phase of the wavelet and the proposed algorithm with the proximity operator defined in equation (29) and \( \beta = 0.2 \) was used to deconvolve Figure 7c. The estimated reflectivity model is depicted in Figure 7d. This figure clearly shows good performance of the algorithm in multichannel deconvolution with phase errors.

**Example 2:** As a final example the proposed multichannel deconvolution algorithm was used to deconvolve two pre- and post-stack field seismic sections. A portion of an age applied pre-stack section is shown in Figure 8a. It has been deconvolved blindly using the method proposed in (Gholami and Sacchi, 2013) resulted to the reflectivity model in Figure 8c: and a wavelet whose amplitude spectrum is depicted in Figure 9f. To compare the performance of the multichannel deconvolution with that of single channel deconvolution, the extracted wavelet was used to deconvolve each trace separately via \( \ell_1 \)-norm deconvolution (Figure 8b). Despite its significant temporal resolution, static errors make it difficult to track the reflectors. Figure 8d is obtained after 10 alternating iterations of the multichannel deconvolution followed by static correction. The quality of the estimate is improved as expected. The proposed algorithm produced the reflectivity model in Figure 8e using amplitude only information of the wavelet. Figure 9 shows the amplitude spectrum of the middle trace before and after deconvolution. The bandwidth extension after deconvolution can be seen clearly. The same procedure discussed for pre-stack data was used to deconvolve a post-stack section. See Figure 10 for the results.
It can be seen from Figures 8 and 10 that for both pre- and post-stack data, the algorithm greatly increased the quality of the sections. It is observed that the pre-stack data are subjected to static shifts and the algorithm was able to increase the continuity of the reflection events while increasing the temporal resolution. The post-stack data are nearly static free but its temporal resolution has been increased greatly by all deconvolution algorithms.

CONCLUSION

We examined seismic data recovery problem from the Fourier amplitude spectrum of the observed data. A suitable cost function including a frequency modulus misfit term and a sparsity promoting regularization term was defined for this purpose and a fast algorithm, based on FISTA, was developed to solve the generated problem. The proposed algorithm is suitable for use in data recovery situations where there exists significant uncertainty in the phase spectrum compared to amplitude spectrum, such as data recovery in the presence of static time shifts. The primary applications of interest considered were trace interpolation in the presence of residual static shifts and deconvolution in presence of residual static shifts and missing wavelet phase. For both applications, the proposed algorithm was able to produce better results than the same algorithms where the misfit term is posed in the time domain (fitting both the amplitude and phase spectra). Numerical results from computer simulated data and field data confirmed this claim.

For interpolation problem, the proposed method was also compared to a carefully tuned benchmark solution (estimating both the interpolated data and static shifts in an alternating scheme). It was observed that the proposed phase retrieval followed by phase correction is able to arrive at a similar solution more quickly.

The regularized phase retrieval algorithm presented here is general and can be used for estimating solutions with any statistical distribution as long as the corresponding proximity operator is used. But the main assumption behind the presented deconvolution experiments was sparsity of the reflectivity series. In this case, it was observed that the algorithm is able to estimate the original reflectivity, with a high degree of accuracy, from just a rough estimate of the wavelet amplitude spectrum. Furthermore, it was shown that static errors can also be compensated for simultaneously by forcing lateral coherency of the reflection events in multichannel deconvolution by the proposed method.

Finally, it should be noted that the algorithm presented in this paper is flexible and general, it is not limited to seismic applications and can be used in other fields dealing with phase retrieval problem.

ACKNOWLEDGMENTS

The author would like to thank two anonymous reviewers for their constructive comments and suggestions.
Figure 7: (a) a synthetic reflectivity section and (b) the result of its convolution with a mixed-phase wavelet (shown at the top right corner of the reflectivity). (c) section (b) after applying residual static shifts and adding random noise (SNR=15 decibels). (d) the estimated reflectivity via the proposed multichannel deconvolution.

Figure 8: (a) a portion of a pre-stack field seismic data, (b) estimated reflectivity model from conventional trace by trace $\ell_1$-norm deconvolution, (c) estimated reflectivity model by applying multichannel deconvolution for inverting both amplitude and phase spectra. (d) estimated reflectivity model after ten alternating iterations of multichannel deconvolution followed by phase (static) correction. (e) the reflectivity model obtained via the proposed phase retrieval deconvolution.
Figure 9: (a)-(e) are the normalized Fourier magnitude spectrum of the middle trace of the data shown in Figures 8(a)-(e) and (f) shows the wavelet amplitude spectrum used for deconvolution.

Figure 10: (a) a portion of a post-stack field seismic data, (b) estimated reflectivity model from conventional trace by trace $\ell_1$-norm deconvolution, (c) estimated reflectivity model by applying multichannel deconvolution and inverting both amplitude and phase spectra. (d) estimated reflectivity model after ten alternating iterations of multichannel deconvolution followed by static correction. (e) the reflectivity model obtained via the proposed phase retrieval deconvolution. The normalized Fourier magnitude spectrum of the middle trace is shown below each section. The wavelet amplitude spectrum used for deconvolution is depicted at the top left of Figure 6.
REFERENCES


Gholami, A., 2013b, Nonconvex compressed sensing with frequency mask for seismic data reconstruction and denoising: *Geophysical Prospecting*, accepted.


