RECOVERY OF SEISMIC WAVEFIELDS BY AN $l_q$-NORM CONSTRAINED REGULARIZATION METHOD

FENGMIN XU
School of Economics and Finance, Xi’an Jiaotong University
Xi’an 710049, China

*1,2,3* YANFEI WANG
1Key Laboratory of Petroleum Resources Research
Institute of Geology and Geophysics
Chinese Academy of Sciences, Beijing 100029, China

2University of Chinese Academy of Sciences
Beijing 100049, China

3Institutions of Earth Science
Chinese Academy of Sciences
Beijing 100029, China

(Communicated by Hao-Min Zhou)

Abstract. Reconstruction of the seismic wavefield from sub-sampled data is an important problem in seismic image processing, this is partly due to limitations of the observations which usually yield incomplete data. In essence, this is an ill-posed inverse problem. To solve the ill-posed problem, different kinds of regularization technique can be applied. In this paper, we consider a novel regularization model, called the $l_2$-$l_q$ minimization model, to recover the original geophysical data from the sub-sampled data. Based on the lower bound of the local minimizers of the $l_2$-$l_q$ minimization model, a fast convergent iterative algorithm is developed to solve the minimization problem. Numerical results on random signals, synthetic and field seismic data demonstrate that the proposed approach is very robust in solving the ill-posed restoration problem and can greatly improve the quality of wavefield recovery.

1. Introduction. In seismology, due to limitations of the observations, the observed data is usually incomplete, e.g., some traces are lost, or random sampling (insufficient sampling) to save cost of seismic acquisition. In that situation, a key obstacle is how to invert the model using only incomplete, sub-sampled data [14, 18, 25]. Recovery of the original wavefield from incomplete observed data is generally an ill-posed problem. To solve the problem, two main issues are how to establish a proper mathematical model and how to solve the minimization model.

To successfully recover a signal without error, according to Nyquist-Shannon sampling theorem, the signal acquisition systems require that the sampling rate is twice the maximum frequency. This sampling theorem is hard to satisfy in practice.
As an alternative, compressive sensing (CS) has recently received a lot of attention in the signal and image processing community. Instead of relying on the bandwidth of the signal, the CS uses the basic assumption: sparsity. That means using sparse assumption-promoting techniques to produce “compressed modes” that are sparse and localized in space, and then solving $L_q$-norm regularized variational equations in mathematics and physics. Therefore, the (approximately) sparse signals can be recovered efficiently from incomplete information. The sparsity can lead to efficient estimations and compression of the signal via a linear transform, e.g., sine, cosine, wavelet, curvelet and beamlet transforms [41, 15, 21]. The method involves taking a relatively small number of non-traditional samples in the form of projections of the signal onto random basis elements or random vectors [8, 1, 2]. Therefore, if the signal has a sparse representation on some basis, it is possible to reconstruct the signal using few linear measurements. Nevertheless this is a special case of Tikhonov regularization with a priori knowledge, the relationship between the compressive sensing and the Tikhonov regularization has been addressed recently [36].

To address the seismic recovery problem, let us start with a signal $x$ in $N$-dimensional space. Suppose that we have $M$ observation data $d_i = A_i x, \ i = 1, 2, \cdots, M$, where $A_i$ for each $i$ is a row vector, which represents the impulse response of the $i$-th sensor. The product of $A_i$ with $x$ yields the $i$-th component of data $d$. Let the matrix $A$ be composed of $M$ row vectors of $A_i (i = 1, 2, \cdots)$, and we set $A = [A_1^T, A_2^T, \cdots, A_M^T]^T$, where $A_i^T$ stands for the transpose of the row vector $A_i$, then the observation data can be reformulated as $d = Ax$. The aim of the compressive sensing is to use limited observations $d_i (i = 1, 2, \cdots, M)$ with $M \ll N$ to restore the input signal $x$ [36]. There are many ways to choose an orthogonal transform matrix based on some orthogonal bases, e.g., sine curve, wavelet, curvelet and framelet, and so forth [15]. As we are mainly concerned with new methods to solve the linear system $d = Ax$ in this paper, we choose a simple wavelet orthogonal basis to form the transform matrix $\Psi$. The wavelet function we used in this paper is the Haar wavelet [19], which forms an orthonormal system for the space of square-integrable functions on the unit interval.

Suppose $x$ is the original wavefield which can be spanned by a series of orthogonal bases $\Psi_i(t)$. These bases for all $i$ constitute an orthogonal transform matrix $\Psi$ such that

$$x(t) = (\Psi m)(t) = \sum_i m_i \Psi_i(t),$$

where $m_i = (x, \Psi_i)$. Using operator expression, $m = \Psi^* x$, where $\Psi^*$ denotes the adjoint for the operator $\Psi$. The vector $m$ is thought of as the sparse or compressive expression of the signal $x$. Letting $L = A \Psi$, where $A$ is the impulse signal response as mentioned before, the reconstruction problem of the sparse signal $m$ reduces to solving a simple problem $d = Lm$. If we regard $m_i$ as the weight or coefficient of linear combinations for the signal $x$, the reconstruction of the signal $x$ becomes to find the coefficient vector $m$.

Let us begin with a compact problem

$$Lm = d,$$

where $L$ is a measurement matrix of finite rank which maps $m$ from parameter space into observation space. One may readily see that the problem (2) can be regarded as a data recovery problem: $m$ is the model, i.e., the representation of $x$ in the transformed domain and $L$ is the generating matrix that yields the acquired data
d. Problem (2) is usually ill-posed due to the fact that existence, uniqueness and stability of the solution may be violated. Conventional methods for solving such an ill-posed problem are Tikhonov’s regularization [27, 37].

In establishing the regularization model, there are different kinds of ways: smoothing or nonsmoothing. A norm $\| \cdot \|$ usually promotes the assumed structure of the solution. Sparsity is a special structure of the solution. We say that the solution $m$ is sparse, when most of the elements are zero. A natural model to satisfy the sparse solutions of the linear system $Lm = d$ (here we assume that the data $d$ is noiseless) is the equality constrained minimization model with $l_0$ norm. Though the "$l_0$ counting norm" yields the sparsest solution, it is well known that the minimization of $\|m\|_0$ is an NP-Hard problem [22]. The minimization model based on $l_1$ norm approximates the minimization model based on $l_0$ norm quite well, while the sparsity is retained [3, 11, 10]. However, the $l_1$-norm constrained regularization may yield inconsistent selections when applied to variable selection in some situations, e.g., conflict of optimal prediction and consistent variable selection may occur for statistical learning [39]; furthermore, it usually introduces extra bias in parameter estimation in linear regression, where the bias contribution of the estimator can be bounded and cannot be neglected [20], and cannot restore a signal with the least measurements when applied to compressed sensing [4]. Hence, further modifications and improvements are necessary.

As a special case of Tikhonov regularization, we consider sparsity-constrained minimization model in this paper, the sparsity is enforced by an $l_q$-norm ($0 < q < 1$) constraint to the solution. For a general $l_p$-$l_q$ minimization model [33]

$$J_{p,q}^a[m] := \frac{1}{2}\|Lm - d\|_p^p + \alpha\|m\|_q^q \rightarrow \min,$$

for $p, q \geq 0$.

In different cases of choosing values of $p$ and $q$, several mathematical techniques have been developed: e.g., the iterative reweighted algorithm for recovering sparse vectors [17], the split Bregman iteration for L1-regularization [12], the basis pursuit denoising (BPDN) criterion using (orthogonal) matching pursuit method [6, 28] and the least absolute shrinkage and selection operator (LASSO) [26] for $l_1$-norm constrained minimization problems. Efficient optimization algorithms including conjugate gradient methods with preconditioning techniques [16] and gradient projection methods [7, 31, 11, 10] can be applied. For the BPDN problem with $\delta$ the upper bound of the $l_2$ norm of the misfit between simulated data to the observation, a particular method called the interior point (IP) solution method can be employed [34]. However the IP solutions may be physically meaningless for some geophysical problems [32], therefore this method must be carefully used.

Since geophysical inverse problems are usually in large scale, hence, fast solving methods are most welcome. Therefore, in solving the $l_q$-norm constrained minimization problem, we develop a fast convergent iterative algorithm by incorporating a boundedness constraint on the solution. This method provides a fast computation and yields a stable solution. Numerical experiments on signal processing and seismic wavefield restoration problem indicate the robustness and applicability of our algorithms.

2. Modeling and methodology. Inspecting the algorithms effective for $l_0$ and $l_1$ regularization problems, we are particularly interested in finding of a first-order iterative algorithm for $l_q$-norm ($q \in (0, 1)$) constrained regularization. This is promoted not only by the fact that the first-order iterative algorithms are adequate, efficient
for high-dimensional problems (this is crucial for compressed sensing application),
but also by the advantage that it is relatively easy to specify the regularization
parameter in implementation of the algorithm. Therefore, as long as such a fast
iterative algorithm can be developed, the $l_q$-norm constrained regularization could
be applied as powerfully, even more powerfully, as $l_1$-norm constrained regularization,
which then, hopefully, advances an essential step towards the better solution
of sparsity problems.

Our choice of the $l_2 - l_q$ minimization model is based on the following observations:
1. The $l_q$-norm ($q \in (0, 1)$) regularization can assuredly generate more sparse
solutions than the $l_1$-norm regularization;
2. The $l_q$ norm $\| \cdot \|_{l_q}$ for $0 < q < 1$ is neither convex nor Lipschitz continuous;
solving the nonconvex, non-Lipschitz continuous minimization problem should
yield the sparse solution as good as $q = 0$ while retaining fast convergence.

2.1. The fast convergent solving method. To solve the $l_2 - l_q$ ($q \in (0, 1)$)
minimization model using optimization technique, we need to calculate the gradient
and/or the Hessian information of $J_{2,q}^q [m]$ depending on gradient type or Newton
type of methods to be used.

The gradient of the objective function $J_{2,q}^q [m]$ can be expressed as
\begin{equation}
    g(m) = \nabla J_{2,q}^q [m] = L^T L m - L^T d + a q |m|^{q-1} \text{sign}(m),
\end{equation}
where $|m|$ means the componentwise absolute value of $m$, $\text{sign}(\cdot)$ is a symbolic
function defined by $\text{sign}(y) = \begin{cases}
    1, & \text{if } y > 0, \\
    0, & \text{if } y = 0 , \\
    -1, & \text{otherwise},
\end{cases}$
componentwise value of $m$. For non-zero components of the model vector, taking
derivative of the gradient $g(m)$ yield the Hessian matrix of $J_{2,q}^q [m]
\begin{equation}
    H(m) = \nabla g(m) = L^T L + a q (q - 1) \text{diag}(|m|^{q-2}),
\end{equation}
where diag$(\cdot)$ denotes diagonalization of a vector.

With the gradient and Hessian information, fast algorithms can be established.
In the current paper, we only consider the first-order method, i.e., only the gradient
information will be carried out to perform a recovery of the seismic wavefield. The basic iterative formula for the first-order method reads as: denote $m_k$ as the updated
wavefield at the $k$-th iteration, then the next update will be
\begin{equation}
    m_{k+1} = m_k + \tau_k s_k,
\end{equation}
where $s_k = -g_k = -g(m_k)$ is the search direction, $\tau_k$ is the step-length. In our
calculation, the line search technique is based on the Armijo-Goldstein line search,
which requires the step-length $\tau_k$ satisfying [38]
\begin{equation}
    J_{2,q}^q [m_k] - J_{2,q}^q [m_k + \tau_k s_k] \geq -b_1 \tau_k s_k^T g_k,
\end{equation}
\begin{equation}
    J_{2,q}^q [m_k] - J_{2,q}^q [m_k + \tau_k s_k] < -b_2 \tau_k s_k^T g_k,
\end{equation}
where $b_1 < b_2$ are two positive parameters. Typical values of $b_1$ and $b_2$ in our
calculations are that $b_1 = 0.4$ and $b_2 = 0.9$. Under the assumption that the gradient
of the objective function $J_{2,q}^q$ is Lipschitz continuous with Lipschitz constant $C$,
it can be proved that the objective function $\{ J_{2,q}^q [m_k] \}$ is a decreasing sequence
and satisfies $J_{2,q}^q [m_k] - J_{2,q}^q [m_k + \tau_k s_k] \geq \frac{b_1(1-b_2)}{C} \| g_k \|^2 \cos^2 < s_k, -g_k >$, where
<s_k, -g_k> refers to the angle between the search direction $s_k$ and negative gradient vector $-g_k$.

To accelerate convergence of the iterative method, we consider the lower bounds for nonzero entries in the optimal solution of the minimization model (3) with $p = 2$ and $q \in (0, 1)$. The lower bounds are established in [5] for data fitting problem will be used for our seismic data recovery problem. Since $J^2_{2,q}[m] \geq \alpha \|m\|^q_{2,q}$, the objective function $J^2_{2,q}$ is bounded below. From equation (5), we obtain that $(Le_i)^T(Le_i) + \alpha(q-1)|m_i|^q - 2 \geq 0$, for $i = 1, 2, \cdots$, where $e_i$ is the $i$-th column of an identity matrix $I$. This implies that $|m_i| \geq (\frac{\alpha(q-1)}{\|L e_i\|^2})^{1/(2-q)}$. Suppose $\mathcal{M}_q^*$ consists of local minimizers of (3) ($p = 2$, $q \in (0, 1)$), and let $B_i = \left(\frac{\alpha(q-1)}{2\|L e_i\|^2}\right)^{\frac{1}{1-q}}$ for $i = 1, 2, \cdots, N$, where we let $l_i$ be the $i$-th column of the matrix $L$, hence $B_i$ has fixed value for each $i$ ($i = 1, 2, \cdots$), then for any $m^*_i \in \mathcal{M}_q^*$, using the second necessary condition of the minimization of $J^2_{2,q}[m]$, i.e., the positive definiteness of the Hessian at the minima, we have that: if $m^*_i \neq 0$, then $|m^*_i| \geq B_i$. This immediately indicates us that $m^*_i$ must be zero if $m^*_i \in (-B_i, B_i)$ for $i = 1, 2, \cdots, N$. Therefore, using iterative formula (6), its convergence can be accelerated given the current iteration point $m_k$, since $m_{k+1}$ does not need to be updated once components of $m_k + \tau_k s_k$ is bounded by $B_i$.

We outline a realistic fast convergent algorithm as follows:

1. Give initial value $m_0$, input matrix $L$, data $d$, regularization parameter $\alpha > 0$, parameter $1 > q > 0$ and tolerance $\epsilon > 0$; for $i = 1, 2, \cdots$, compute

$$B_i = \left(\frac{\alpha(q-1)}{2\|L e_i\|^2}\right)^{\frac{1}{1-q}}$$

2. At the $k$-th step, compute the gradient of $J^2_{2,q}[m_k]$ and denoted it by $g_k$, let $s_k = -g_k$;

3. Perform a line search using Armijo-Goldstein rule (7) and (8) to get a step-length $\tau_k$;

4. Update the formula $m^{\text{temp}}_{k+1} = m_k + \tau_k s_k$, and compute

$$m_{k+1} = \begin{cases} 0, & \text{if } |m^{\text{temp}}_{k+1}(i)| < B(i), \\ m^{\text{temp}}_{k+1}, & \text{otherwise}; \end{cases}$$

5. If $\|m_{k+1} - m_k\| \leq \epsilon$, then an optimal solution $m^*$ is found, Stop; Otherwise, let $k$ be $k + 1$, go to Step (ii).

2.2. Choosing the regularization parameter $\alpha$. We remark that our method works for any choice of the parameter $q \in (0, 1)$. Choosing the regularization parameter $\alpha$ is also a key issue. To choose a reasonable value of the regularization parameter $\alpha$, we look back at $B_i$ again. Since the denominator of $B_i$ is non-related with $\alpha$, we simply investigate the function $\phi(\alpha, q) = (\alpha(q-1))^{\frac{1}{1-q}}$. The signification of $\phi(\alpha, q)$ is to choose proper values of $q$ and $\alpha$. Note that our aim is to find the lower bounds of $B_i$ to accelerate the convergence, hence to finish this job, $\phi(\alpha, q)$ should be as small as possible. For fixed value of $\alpha$, it is straightforward to calculate the derivative of the function $\phi(\alpha, q)$ with respect to $q$ yielding the stable point of $q = 1/2$. Therefore, in this paper, we simply choose the value of $q$ as 0.5. If $q$ is chosen, then to reach the minimum of $\phi(\alpha, q)$ so as to the lower bound of $B_i$.
Figure 1. $\phi(\alpha, q)$ for $\alpha, q \in (0, 1)$.

for $i = 1, 2, \ldots, N$, $\alpha$ is less than 0.05 will be sufficient. Therefore, in this paper we simply choose the value of $\alpha$ as 0.001. An illustration of this phenomenon is shown in Figure 1.

2.3. Complexity analysis. Regarding to the computational complexity of our method, the main computation is the calculation of the objective gradient. Since our method is a gradient method with the truncation technique, it possesses low computational cost, less storage and it is free to the initial point. Under the exact line search, our method has a linear convergence rate. Therefore our method can be applied for large datasets problem.

3. Numerical results. To verify the feasibility of our algorithm, we consider four numerical examples. We start our simulation from a simple one-dimensional sparse signal reconstruction. Then we consider an example of reconstruction of seismic shot gathers. The synthetic example is similar to Wang et al. (2011). Finally, two practical data applications are performed.

3.1. Random signal reconstruction. In [35], the authors consider a sparse signal $m \in \mathbb{R}^N$, which is measured (sensed) by a random measuring matrix $L \in \mathbb{R}^{M \times N}$ ($M < N$). Then $d = Lm \in \mathbb{R}^M$ is the measurement vector. Every row of the matrix $L$ can be seen as a measuring operator, whose inner product with $m$ is a measurement. $M < N$ means the number of measurements is smaller than the length of the signal, thus the number of measurements is compressed. In their simulation, $M$ is chosen as 140 and $N$ equals 200, then the problem is to recover the original signal $m$ from the measurement $d$. The true signal is marked by legend “o” and the recovered signal is marked by “+”. We apply our algorithm to the same example, the comparing results of the input signal and the restoration, and the difference between the true and restored signals, are shown in Figures 2 (a) and 2 (b), respectively.

Furthermore, we consider a harder signal reconstruction problem than that of [35], i.e., instead of $M$ equaling 140, we consider $M$ as 80. This is a severely ill-posed problem. Since the measurement is random, therefore, the data is randomly recorded. The original sparse signal is shown in Figure 2 (c) with legend “o” lines.
Using our fast convergent iterative algorithm for $l_q$-norm constrained regularization problem, the restoration results (“+” lines) comparing with the original signal is shown in Figure 2 (c). Difference between the true and restored signals is shown in Figure 2 (d). It is evident from the comparison that our algorithm is robust in reconstruction of sparse signals.

This example shows that our method works for randomly generated data using random measurement matrix. Therefore, it would be a reliable and stable method for signal reconstruction problem.

People may wonder whether other choices of the regularization parameter $\alpha$ will yield similar results. From a just glance, we can choose any values of $\alpha$ which is greater than 0. However, either too small or too large values of $\alpha$ will yield unsatisfactory results. Again, we consider the above second signal reconstruction problem. Figure 3 (a) is an illustration of the original signal and its recovery for smaller $\alpha = 1.0 \times 10^{-4}$, and the difference between the true and restored signals is shown in Figure 3 (b); Figure 3 (c) is an illustration of the original signal and its recovery for larger $\alpha = 0.5$, and the difference between the true and restored signals is shown in Figure 3 (d). This confirms our assertion above. For the parameter $q$, we remark that $q = 1/2$ is a representative value for $q \in (0, 1)$. Since we have obtained the optimal value of the parameter $q \in (0, 1)$, there is no need to try other values of $q$.

3.2. Reconstruction of shot gathers. Now we consider a seismogram generated from a 7 layers geologic velocity model [35] where the spatial sampling interval of 15 meters and the time sampling interval of 0.002 second, the shot gathers are shown in Figure 4. The data with missing traces are shown in Figure 5 (a). Using our $l_q$-norm constrained regularization modeling and the fast convergent iterative approach, the recovered wavefield is shown in Figure 5 (b). It is clear from the reconstruction that most of the details of the wavefield are preserved. To show the good performance of our method, we plot the frequency information of the sub-sampled data the restored data in Figures 5 (c) and 5 (d), respectively. It is clear that the aliasing (just like noise) of the sub-sampled data is reduced greatly in the recovered data. The difference of the original data and the recovered data is illustrated in Figure 6 (a). Virtually, all the initial seismic energy is recovered with minor errors. Though there are still the artifacts such as vertical stripes, we consider it might be caused by ill-posed nature of the inversion process and insufficient iterations. As a comparison with other interpolating techniques, we consider the Fourier transform based interpolating technique. This method is well described in literature, e.g., [24, 9, 40, 41]. The reconstruction result, its frequency information and the difference to the original data are shown in Figures 6 (b)-6 (d), respectively.

Finally, to quantitatively describe the ability of our algorithm in restoration of the wavefield from sampled data, we define the signal-to-noise ratio as

$$SNR = 10 \log_{10} \left( \frac{\|d_T\|_2}{\|d_T - d_R\|_2} \right)^2,$$

where $d_T$ refers to the input seismic wavefield (we refer it to the true signal), $d_R$ refers to the restoration. We also define the relative error of the original data to the restored data as

$$Err = \frac{\|d_T - d_R\|_2}{\|d_T\|_2},$$
The norm is taken as $l_2$ norm. From our calculation, for this problem, the SNR is 34.3782, the Err is 0.2692. The SNR and the Err for the Fourier transform based method are 26.1574 and 0.4644, respectively. The high value of SNR and low value of Err indicate stability of our algorithm in seismic data restoration.

To show the speed of our method, we compare it with the standard spectral projected gradient for $l_1$ minimization (SPGL1) method which is well referred in literature [29]. Our calculations show that for our method, it needs 15.8711 seconds to obtain the convergent results; whereas for SPGL1 method, it requires 26.5988 seconds to obtain the convergent results with the SNR equaling 25.1392 and the Err equaling 0.4997.

3.3. Field data. Finally, we examine the efficiency of the new method with field data. The first seismic data about Tarimu oil field is provided in Figure 7 (a),
Figure 3. (a) The input signal and the restoration for a small regularization parameter $\alpha = 1.0 \times 10^{-4}$; (b) difference between the true and restored signals for a small regularization parameter $\alpha = 1.0 \times 10^{-4}$; (c) the input signal and the restoration for a large regularization parameter $\alpha = 0.5$; (d) difference between the true and restored signals for a large regularization parameter $\alpha = 0.5$.

The subsampled gather is shown in Figure 7 (b) with half of the original traces randomly deleted. This sub-sampled gather was used to restore the original gather with suitable solution methods. With our $l_q$-norm constrained minimization model and the fast convergent algorithm to restore the data, the result is shown in Figure 7 (c). Frequency of the recovered data is shown in Figure 7 (d). Comparing the subsampled image with the original image, the restored image can reconstruct most of the details. We apply the Fourier transform based interpolating technique to the field data. The reconstruction result is shown in Figure 8 (a). Frequency of the corresponding recovered data is shown in Figure 8 (b). Comparing the frequency
of the restored seismic data using our proposed method and the Fourier transform based method, it reveals that our method yields a more accurate recovery.

Our next field data is a marine shot gather shown in Figure 9 (a) which consists of 256 traces with spacing 25m and time sampling interval 2ms. There are damaged traces in the gather. The subsampled gather is shown in Figure 9 (b) with half of the original traces randomly deleted. This sub-sampled gather was used to restore the original gather with suitable solution methods. With our $l_0$-norm constrained minimization model and the fast convergent algorithm to restore the data, the results are shown in Figures 9 (c) and 9 (d), respectively. Comparing the subsampled image with the original image, the restored image can reconstruct most of the details. In addition the damaged trace in the original gather was restored as a good trace. Again, we apply the Fourier transform based interpolating technique to the field data. The reconstruction result is shown in Figures 10 (a) and 10 (b), respectively. Again, from comparison of the frequency of the restored seismic data using our proposed method and the Fourier transform based method, it indicates that our method yields a more accurate recovery.

3.4. Conclusion. Sparse optimization has broad applications in seismic data processing, such as time-frequency analysis, de-noising, multiple attenuation, interpolation and migration. This paper focuses on data restoration (interpolation) problem. Major issues affecting the effect of restoration are: sparse transforms, sampling, sparse modeling and solving methods. The sparse transform will influence the restoration results, the sparser of the data in transform domain, the easier to restore the original data; fast and effective solving methods are the key of sparse optimization; and proper sampling methods will save acquisition time and cost greatly. We mainly study sparse modeling and solving methods in this paper, while the sparse transform is based on wavelet transform and the sampling technique is based on the piecewise random sub-sampling developed in [35].

We first reformulate the data restoration problem using an $l_q$-norm constrained minimization model, which is a nonconvex and non-Lipschitz continuous minimization problem. Based on the lower bound of the local minimizers of the minimization
problem $l_2-l_q$ ($0 < q < 1$), we then develop a fast, effective first-order iterative algorithm for realizing the minimization problem. Numerical results on synthetic problems and the field data example indicate potential usage of our method for practical applications. We remark that we only consider the gradient descent method in the present paper. For the seismic wavefield recovery using the $l_q$ norm minimization, the (L-)BFGS method may be applied to this kind of problems, which will be our next stage of investigation.

Acknowledgments. We would like to thank three reviewers very much for their valuable comments and suggestions.

REFERENCES

[1] E. J. Candes, J. Romberg and T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Transactions on Information Theory*, **52** (2006), 489–509.
Figure 6. (a) Difference between the restored data and the original data; (b) recovery results using the Fourier transform based method; (c) frequency of the restored data using the Fourier transform based method; (d) difference between the restored data using the Fourier transform based method and the original data.

[2] E. J. Candes and M. B. Wakin, An introduction to compressive sampling, *IEEE Signal Processing Magazine*, 25 (2008), 21–30.

[3] J. J. Cao, Y. F. Wang, J. T. Zhao and C. C. Yang, A review on restoration of seismic wavefields based on regularization and compressive sensing, *Inverse Problems in Science and Engineering*, 19 (2011), 679–704.

[4] R. Chartrand and V. Staneva, Restricted isometry properties and nonconvex compressive sensing, *Inverse Problems*, 24 (2008), 1–14.

[5] X. J. Chen, F. M. Xu and Y. Y. Ye, Lower bound theory of nonzero entries in solutions of $l_2$-$l_p$ minimization, *SIAM Journal on Scientific Computing*, 32 (2010), 2832–2852.

[6] S. Chen, D. Donoho and M. Saunders, Atomic decomposition by basis pursuit, *SIAM Journal on Scientific Computing*, 20 (1998), 33–61.

[7] Y. H. Dai and R. Fletcher, Projected Barzilai-Borwein methods for large-scale box-constrained quadratic programming, *Numerische Mathematik*, 100 (2005), 21–47.
Figure 7. (a) The field data; (b) seismic data with missing traces; (c) restored seismic data; (d) frequency of the restored seismic data.

[8] D. Donoho, Compressed sensing, *IEEE Transactions on Information Theory*, 52 (2006), 1289–1306.
[9] A. J. W. Duijndam and M. A. Schonewille, Non-uniform fast Fourier transform, *Geophysics*, 64 (1999), 539–551.
[10] V. B. Ewout and P. F. Michael, Probing the pareto frontier for basis pursuit solutions, *SIAM Journal on Scientific Computing*, 31 (2008), 890–912.
[11] M. A. T. Figueiredo, R. D. Nowak and S. J. Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems, *IEEE Journal of Selected Topics in Signal Processing*, 1 (2007), 586–597.
[12] T. Goldstein and S. Osher, The split Bregman method for $l_1$ regularized problems, *SIAM Journal on Imaging Sciences*, 2 (2009), 323–343.
[13] G. Hennenfent and F. J. Herrmann, Simply denoise: Wavefield reconstruction via jittered undersampling, *Geophysics*, 73 (2008), V19–V28.
[14] F. J. Herrmann and G. Hennenfent, Non-parametric seismic data recovery with curvelet frames, *Geophysical Journal International*, 173 (2008), 233–248.
Figure 8. (a) Restored seismic data using the Fourier transform based method; (b) frequency of the restored seismic data using the Fourier transform based method.

[15] F. J. Herrmann, D. L. Wang, G. Hennenfent and P. P. Moghaddam, Curvelet-based seismic data processing: a multiscale and nonlinear approach, Geophysics, 73 (2008), A1–A6.

[16] S. J. Kim, K. Koh, M. Lustig, S. Boyd and D. Gorinevsky, An interior-point method for large-scale $l_1$-regularized least squares, IEEE Journal on Selected Topics in Signal Processing, 8 (2007), 1519–1555.

[17] M. J. Lai, Y. Xu and W. Yin, Improved iteratively reweighted least squares for unconstrained smoothed $l_2$ minimization, SIAM Journal on Numerical Analysis, 51 (2013), 927–957.

[18] B. Liu and M. D. Sacchi, Minimum weighted norm interpolation of seismic records, Geophysics, 69 (2004), 1560–1568.

[19] S. G. Mallat, A Wavelet Tour of Signal Processing. Academic Press, San Diego, 1998.

[20] N. Meinshausen and B. Yu, Lasso-type recovery of sparse representations for high-dimensional data, Annals of Statistics, 37 (2009), 246–270.

[21] M. Naghizadeh and M. D. Sacchi, Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data, Geophysics, 75 (2010), WB189–202.

[22] B. K. Natarajan, Sparse approximate solutions to linear systems, SIAM Journal on Computing, 24 (1995), 227–234.

[23] R. T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ., 1970.

[24] M. D. Sacchi, T. J. Ulrych and C. J. Walker, Interpolation and extrapolation using a high-resolution discrete Fourier transform, IEEE Transactions on Signal Processing, 46 (1998), 31–38.

[25] M. D. Sacchi, D. J. Verschuur and P. M. Zwartjes, Data reconstruction by generalized deconvolution, Expanded Abstracts 74th Annual Meeting SEG, Denver, USA (Denver, Oct. 2004), (2004), 1989–1992.

[26] R. Tibshirani, Regression shrinkage and selection via the lasso, Journal Royal Statistical Society B, 58 (1996), 267–288.

[27] A. N. Tikhonov and V. Y. Arsenin, Solutions of Ill-posed Problems, John Wiley and Sons, New York, 1977.

[28] J. A. Tropp and A. C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, IEEE Transactions on Information Theory, 53 (2007), 4655–4666.

[29] E. van den Berg and M. P. Friedlander, Probing the Pareto frontier for basis pursuit solutions, SIAM Journal on Scientific Computing, 31 (2009), 890–912.

[30] Y. F. Wang, Computational Methods for Inverse Problems and Their Applications, Higher Education Press, Beijing, 2007.

[31] Y. F. Wang and S. Q. Ma, Projected Barzilai-Borwein methods for large scale nonnegative image restorations, Inverse Problems in Science and Engineering, 15 (2007), 559–583.
Figure 9. (a) The field data; (b) seismic data with missing traces; (c) restored seismic data; (d) frequency of the restored seismic data.

[32] Y. F. Wang, S. F. Fan and X. Feng, Retrieval of the aerosol particle size distribution function by incorporating a priori information, *Journal of Aerosol Science*, 38 (2007), 885–901.

[33] Y. F. Wang, J. J. Cao, Y. X. Yuan, C. C. Yang and N. H. Xiu, Regularizing active set method for nonnegatively constrained ill-posed multichannel image restoration problem, *Applied Optics*, 48 (2009), 1389–1401.

[34] Y. F. Wang, Sparse optimization methods for seismic wavefields recovery, *Proceedings of the Institute of Mathematics and Mechanics*, 18 (2012), 42–55.

[35] Y. F. Wang, J. J. Cao and C. C. Yang, Recovery of seismic wavefields based on compressive sensing by an $l_1$-norm constrained trust region method and the piecewise random sub-sampling, *Geophysical Journal International*, 187 (2011), 199–213.

[36] Y. F. Wang, C. C. Yang and J. J. Cao, On Tikhonov regularization and compressive sensing for seismic signal processing, *Mathematical Models and Methods in Applied Sciences*, 22 (2012), 1150008, 24pp.

[37] Y. F. Wang, A. G. Yagola and C. C. Yang (editors), *Computational Methods for Applied Inverse Problems*, (published in “Series: Inverse and Ill-Posed Problems Series 56”), Walter de Gruyter, 2012.
Figure 10. (a) Restored seismic data using the Fourier transform based method; (b) frequency of the restored seismic data using the Fourier transform based method.

[38] Y. X. Yuan, *Numerical Methods for Nonlinear Programming*, Shanghai Science and Technology Publication, Shanghai, 1993.

[39] H. Zou, *The adaptive lasso and its oracle properties*, *Journal of the American Statistical Association*, 101 (2006), 1418–1429.

[40] P. M. Zwartjes and M. D. Sacchi, *Fourier reconstruction of nonuniformly sampled, aliased seismic data*, *Geophysics*, 72 (2007a), V21–V32.

[41] P. M. Zwartjes and A. Gisolf, *Fourier reconstruction with sparse inversion*, *Geophysical Prospecting*, 55 (2007), 199–221.

Received March 2017; 1st revision January 2018; 2nd revision April 2018.

E-mail address: fengminxu@mail.xjtu.edu.cn
E-mail address: yfwang@mail.iggcas.ac.cn