J. Geophys. Eng. **00** (2014) 000000 (8pp) UNCORRECTED PROOF

Seismic data restoration with a fast L_1 norm trust region method

Jingjie Cao[1] **and Yanfei Wang**[2]

[1] Shijiazhuang University of Economics, Shijiazhuang, 050031 Hebei, People's Republic of China
[2] Key Laboratory of Petroleum Resources Research, Institute of Geology and Geophysics, Chinese Academy of Sciences, P O Box 9825, 100029 Beijing, People's Republic of China

E-mail: cao18601861@163.com

Received 8 May 2014, revised 13 June 2014
Accepted for publication 18 June 2014
Published

Abstract
Seismic data restoration is a major strategy to provide reliable wavefield when field data dissatisfy the Shannon sampling theorem. Recovery by sparsity-promoting inversion often get sparse solutions of seismic data in a transformed domains, however, most methods for sparsity-promoting inversion are line-searching methods which are efficient but are inclined to obtain local solutions. Trust region method which can provide globally convergent solutions is a good choice to overcome this shortcoming. A trust region method for sparse inversion has proposed, however, the efficiency should be improved to suitable for large-scale computation. In this paper, a new L_1 norm trust region model is proposed for seismic data restoration and a robust gradient projection method for solving the sub-problem is utilized. Numerical results of synthetic and field data demonstrate that the proposed trust region method can get excellent computation speed and is a viable alternative for large-scale computation.

Keywords: wavefield restoration, sparse optimization, L1 norm trust region, curvelet transform, inverse problem

(Some figures may appear in colour only in the online journal)

1. Introduction

Seismic acquisition often violates the Shannon theorem because of the restrictions of investment, topography, noise, bad traces and so on. The under-sampled data will bring aliasing and artifacts which will influence results of migration (Liu and Sacchi 2004), de-noising (Soubaras 2004), multiple elimination (Naghizadeh 2009) and AVO analysis (Sacchi and Liu 2005). Signal processing based restoration methods, which can provide reliable wavefield for noisy data with fast convergence speed, got great development recently (Schonewille *et al* 2003, Liu *et al* 2004, Zwartjes and Gisolf 2006). During these methods, sparsity of wavefield in some transformed domain (Herrmann and Hennenfent 2008, Wang *et al* 2011) is commonly used. Thus, solutions in a transform domain should be found at first, then transformed into time-space domain.

Seismic data sampling can be expressed as

$$\Phi m = d, \tag{1}$$

where Φ denotes the sampling operator, m is the complete seismic data, and d is the sampled data. Because equation (1) is under-determined, wavefield restoration is an ill-posed problem (Herrmann and Hennenfent 2008). Based on the sparsity assumption, the sparse solutions in a transformed domain can be found by solving:

$$\min \|x\|_1 \quad \text{s.t.}\, Ax = d, \tag{2}$$

where $x = \Psi m$ and $A = \Phi\Psi^H$, Ψ is a transform such that x is sparse. Ψ^H means its Hermitian transpose. Some commonly used sparse transforms are Fourier transform (Sacchi and Ulrych 1996, Liu *et al* 2004, Zwartjes and Gisolf 2006), Radon transform (Verschuur and Kabir 1995, Trad *et al* 2002), local Radon transform (Sacchi *et al* 2004, Wang *et al* 2009) and curvelet transform (Herrmann and Hennenfent 2008). Because curvelets can compress curve-shape events excellently with no assumption of line events and time windowing, curvelet transform is adopted as the sparse transform in this paper. As

JNL:JGE PIPS: 497988 TYPE: PAP TS: NEWGEN DATE:16/7/14 EDITOR: IOP SPELLING: US

a multi-scale, multi-directional and anisotropic frame (Candes *et al* 2006), it was applied in seismic restoration by Herrmann and Hennenfent (2008). Readers can refer to Candes *et al* (2006) for more detailed information of curvelet transform.

Seismic restoration is time consuming, thus fast and robust methods are required to improve the computational efficiency (Trad 2009). Many strategies were proposed to solve problem (2) for seismic restoration: Iterative-reweighed least-squares method is applied to sparse solutions of problem (2) (Sacchi and Ulrych 1996). Abma and Kabir (2006) proposed a POCS method for irregular seismic interpolation. Herrmann and Hennenfent (2008) proposed an iterative soft thresholding algorithm to solve an L_1 norm constrained least square (Daubechies *et al* 2004). van den Berg and Friedlander (2008) developed a spectral projected gradient method (SPGL1) to find the sparse solution; Cao *et al* (2011, 2012) proposed a projected gradient method for a non-convex optimization model of wavefield restoration. These methods which belong to line-searching methods are efficient but easy to get local solutions, trust region method as another optimization strategy can provide globally convergent solutions for non-linear problems (Yuan 1993). For trust region method, a quadratic approximation of the objective function at current iteration point is built as a trust region sub-problem, then a descent direction is obtained by solving the sub-problem (Yuan 1993). Generally, trust region methods are based on L_2 norm trust region, a L_1 norm trust region method for sparse problems was proposed by Wang *et al* (2011) for seismic restoration; however, the computation speed is much lower than line-searching methods. Improving the efficiency of trust region method is the main topic of this paper. In this paper, a novel L_1 norm trust region method is proposed in which the optimization model is simplified and a rapid projected gradient method for solving the sub-problem is proposed to improve the efficiency significantly. Numerical experiments demonstrate that the proposed method can reach excellent computation speed as line-search methods. It should be point out that this exposition aims to explore L_1 norm trust region method as an interesting research avenue for seismic reconstruction.

2. Algorithm

Problem (2) is based on the sparsity of solutions, however, if each descent direction of (2) is sparse, then sparse solutions will be found quickly. In order to obtain sparse descent directions, trust region method should be adopt by using L_1 norm as the trust region constraint. The original problem proposed in Wang *et al* (2011) is the L_1 norm regularized model

$$\min J_1(x) = \frac{1}{2}\|Ax-d\|_2^2 + \lambda\|x\|_1, \tag{3}$$

where λ is the regularization parameter. In order to find sparse descent direction of (3), its L_1 norm trust region sub-problem should be solved:

$$\begin{aligned}&\min \frac{1}{2}s^T H_k s + g_k^T s\\ &\text{s.t. } \|s\|_1 \le \Delta_k,\end{aligned} \tag{4}$$

where H_k and g_k are the Hessian matrix and the gradient of J_1 at the kth iterative point, s is the descent direction at current point, and Δ_k is the trust region radius. Because the non-differentiability, L_1 norm in (3) and (4) should be replaced by its smooth approximation in numerical implementation, such as square root function, then the sub-problem is solved by its unconstrained form

$$\min \frac{1}{2}s^T H_k s + g_k^T s + \tau\left(\sum_i \sqrt{s_i^2+\varepsilon} - \Delta_k\right), \tag{5}$$

where the Lagrangian parameter τ is solved by Newton method (Wang *et al* 2011).

There are two shortcomings in this method. Firstly, the objective function of (3) is in-differentiable, thus its gradient and Hessian matrix must be obtained approximately by using a smooth approximation of L_1 norm, this will affect the accuracy and increase computations. Secondly, similar smooth strategy is also adopted in the sub-problem (5), and the parameter τ is solved by Newton method which is time consuming. In all, this method costs much time and reduce the accuracy of solutions. How to improve the solution accuracy and computational efficiency of L_1 norm trust region method is the main topic of this paper.

The main merit of L_1 trust region method is that the sparsity of the descent directions makes the sparse solutions obtained quickly. The contribution of L_1 norm in problem (3) to get sparse solutions is very small, furthermore, the L_1 norm in (3) makes the original trust region method hard to solved exactly. If it is removed, the least square problem can be solved easily and accurately, then the objective function can be simplified to

$$\min J(x) = \frac{1}{2}\|Ax-d\|_2^2. \tag{6}$$

Here the gradient and Hessian matrix of $J(x)$ can be found exactly; the sub-problem of (6) is the same as problem (4) in form:

$$\begin{aligned}&\min_s \varphi_k(s) = \frac{1}{2}s^T H_k s + g_k^T s\\ &\text{s.t. } \|s\|_1 \le \Delta_k.\end{aligned} \tag{7}$$

However, different from problem (4), H_k and g_k in problem (7) are exact and easy to solve. In order to introduce the idea of trust region strategy, a classical algorithm framework of trust region method for solving non-linear problems is given as follows (Wang 2007):

Algorithm 1.

Step 1: Choose $0<\tau_3<\tau_4<1<\tau_1$, $0\le\tau_0\le\tau_2<1$, give the initial feasible solution x_0, the initial trust region step $\Delta_0>0$, and let $k=0$.

Step 2: If the stopping condition is satisfied, then stop iteration; otherwise, solving problem (7) to get the decent direction s_k.

Step 3: Calculate the ratio r_k. Update the iterative point:

$$x_{k+1} = \begin{cases} x_k & \text{if } r_k \le \tau_0 \\ x_k + s_k & \text{otherwise} \end{cases}$$

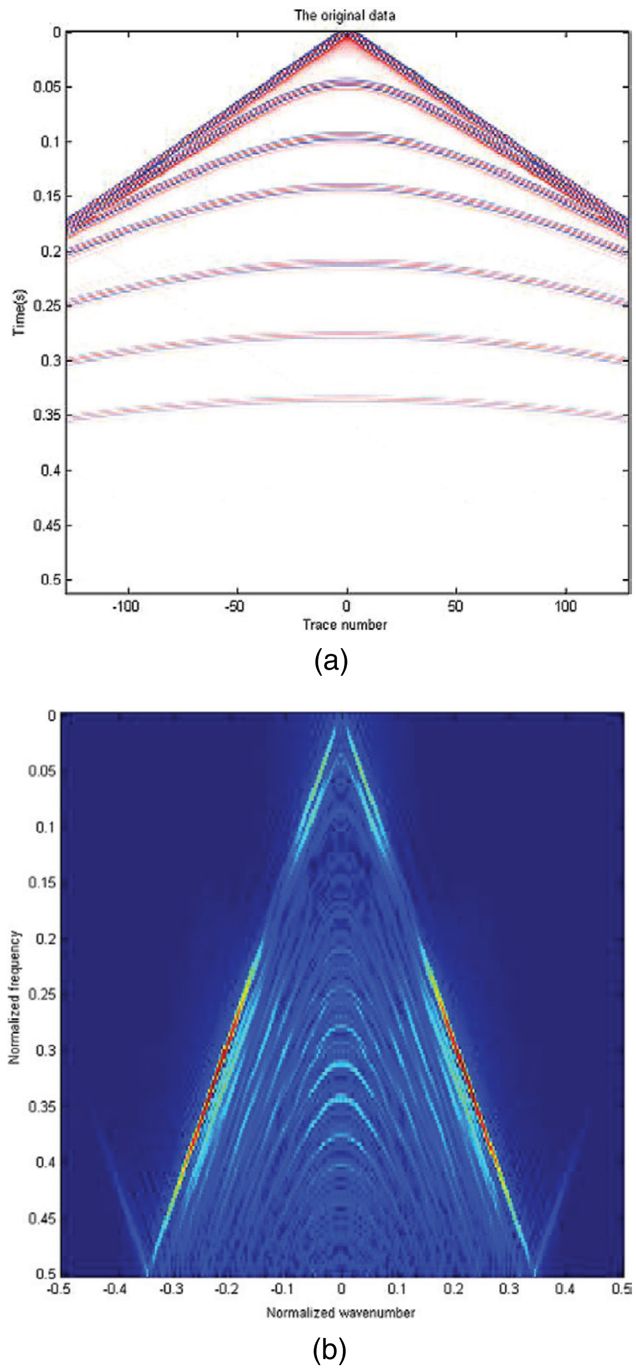

Figure 1. (*a*) Original data; (*b*) F-K spectrum of (*a*).

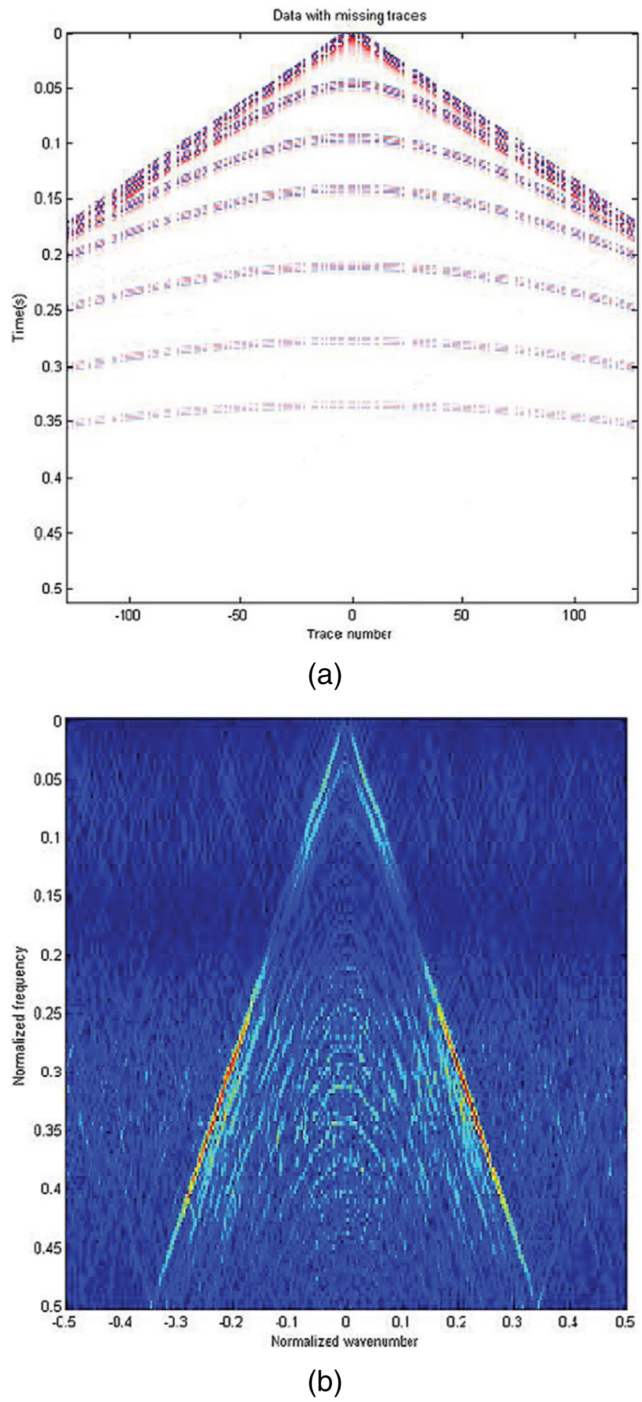

Figure 2. (*a*) Random sub-sampled data; (*b*) F-K spectrum of (*a*). AQ1

Update the trust region radius:

$$\Delta_{k+1} \in \begin{cases} [\tau_3 \|s_k\|, \tau_4 \Delta_k] & \text{if} \quad r_k \leq \tau_2 \\ [\Delta_k, \tau_1 \Delta_k] & \text{otherwise} \end{cases}$$

Step 4: Solve the gradient g_{k+1} at x_{k+1}, and let k=k+1; go to Step 2.

For the stopping condition, this algorithm will not stop until the residual reaches a threshold or the maximum iteration number is reached. r_k controls the updating of x_k and Δ_k which is defined as $r_k = \dfrac{J(x_k) - J(x_k + s_k)}{\varphi_k(0) - \varphi_k(s_k)}$. The initial trust region step Δ_0 is chosen empirically. Refer to Wang (2007) for detailed information about this method.

The efficiency of trust region method hinges on solving problem (7) efficiently. Birgin *et al* (2000) proposed a robust projected gradient method for constrained optimization, here, it is adapted for solving problem (7). By defining $f(s) = \dfrac{1}{2} s^T H_k s + g_k^T s$ and its gradient as $\nabla f(s)$, the projected gradient method is:

Algorithm 2.

Step 1: Give the initial direction s, the trust region step Δ_k, the minimal step length $\alpha_{\min}$, the maximal step length $\alpha_{\max}$,

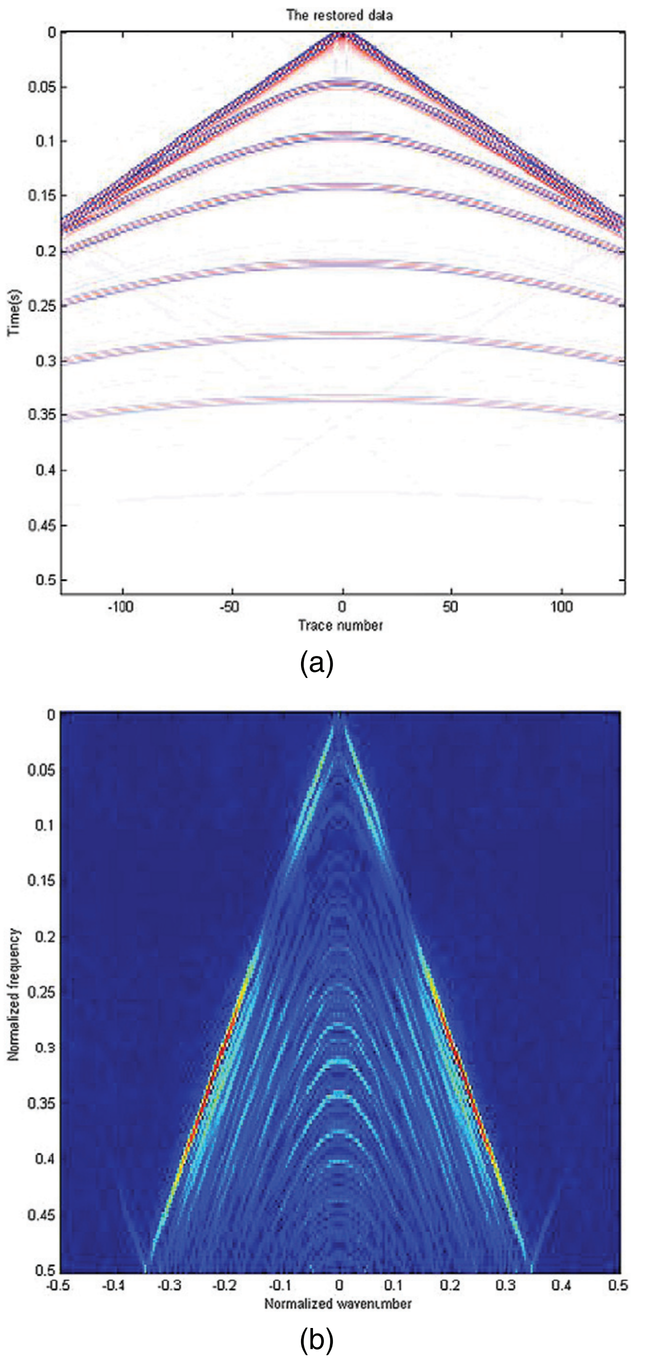

Figure 3. (*a*) Restoration of TRSL1 method; (*b*) F-K spectrum of (*a*).

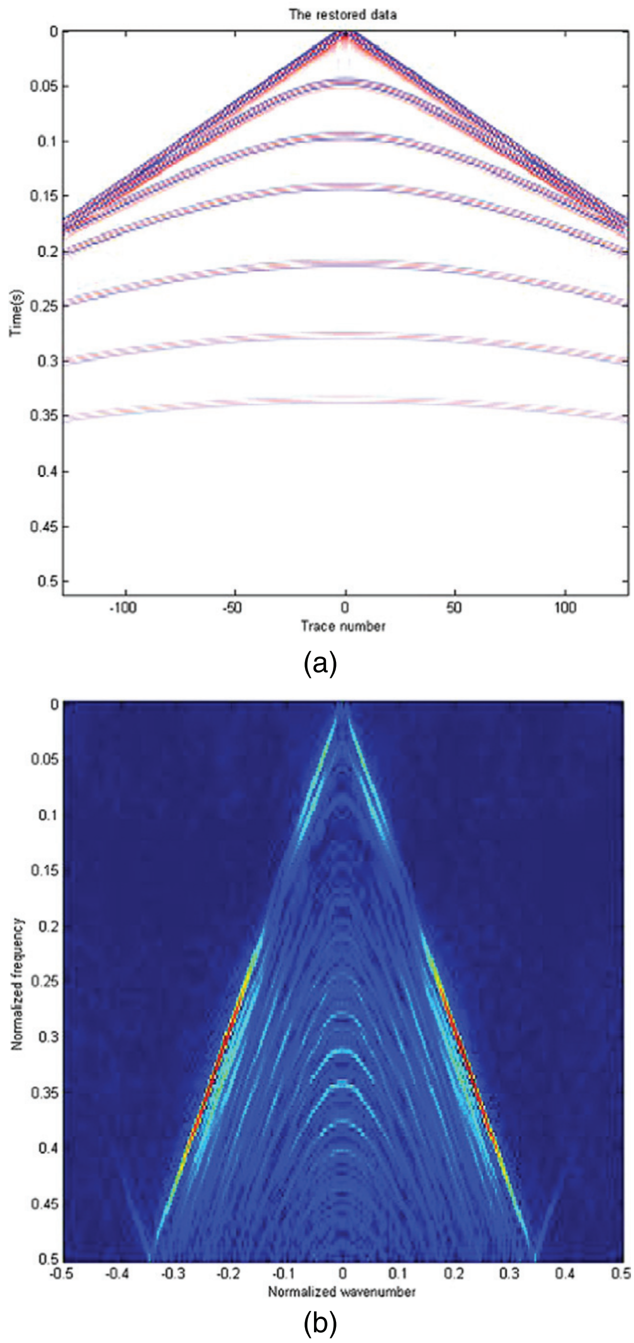

Figure 4. (*a*) Restoration of SPGL1 method; (*b*) F-K spectrum of (*a*).

the initial step length $\alpha_0 \in [\alpha_{\min}, \alpha_{\max}]$, the sufficient decent parameter $\gamma \in (0, 1)$, and back tracing allowed max number of iterations M. Calculate the initial projection $s_0 = P_{S_k}(s)$ (where$P_{S_k}(s)$ is the projection of s onto $S_k = \{s \| \|s\|_1 \leq \Delta_k\}$), the gradient of the objective function $\nabla f(s_0)$, let $l = 0$.

Step 2: If the stopping criterion is satisfied, go to Step 8; otherwise, go to Step 3.

Step 3: Calculate the step length using the back tracing method:

Step 3.1: $\alpha = \alpha_l$;

Step 3.2: Calculate the projection $\bar{s} = P_{\Delta_k}[s_l - \alpha \nabla f(s_l)]$;

Step 3.3: If $f(\bar{s}) \leq \max_{j \in [0, \min[k, M-1]]} f(sl{-}j) + \gamma(\bar{s} - sl)^T \nabla f(sl)$, turns to Step 4; otherwise, $\alpha = \alpha/2$, and turn to Step 3.2.

Step 4: Update the iterative point $s_{l+1} = \bar{s}$, and calculate the new gradient g_{l+1}.

Step 5: Calculate $\Delta s = s_{l+1} - s_l$, $\Delta g = \nabla f(s_{l+1})$-$\nabla f(s_l)$.

Step 6: Update the step length:

If $\Delta s^T \Delta g \leq 0$, $\alpha_{l+1} = \alpha_{\max}$;

otherwise,

$$\alpha_{l+1} = \min\{\alpha_{\max}, \max\{\alpha_{\min}, (\Delta s^T \Delta s)/(\Delta s^T \Delta g)\}\}.$$

Step 7: $l = l + 1$, and go to Step 3.

Step 8: Let $s_{\Delta_k} = s_l$.

In algorithm 2, projections onto $S_k = \{s \| \|s\|_1 \leq \Delta_k\}$ can be calculated according to van den Berg and Friedlander (2008).

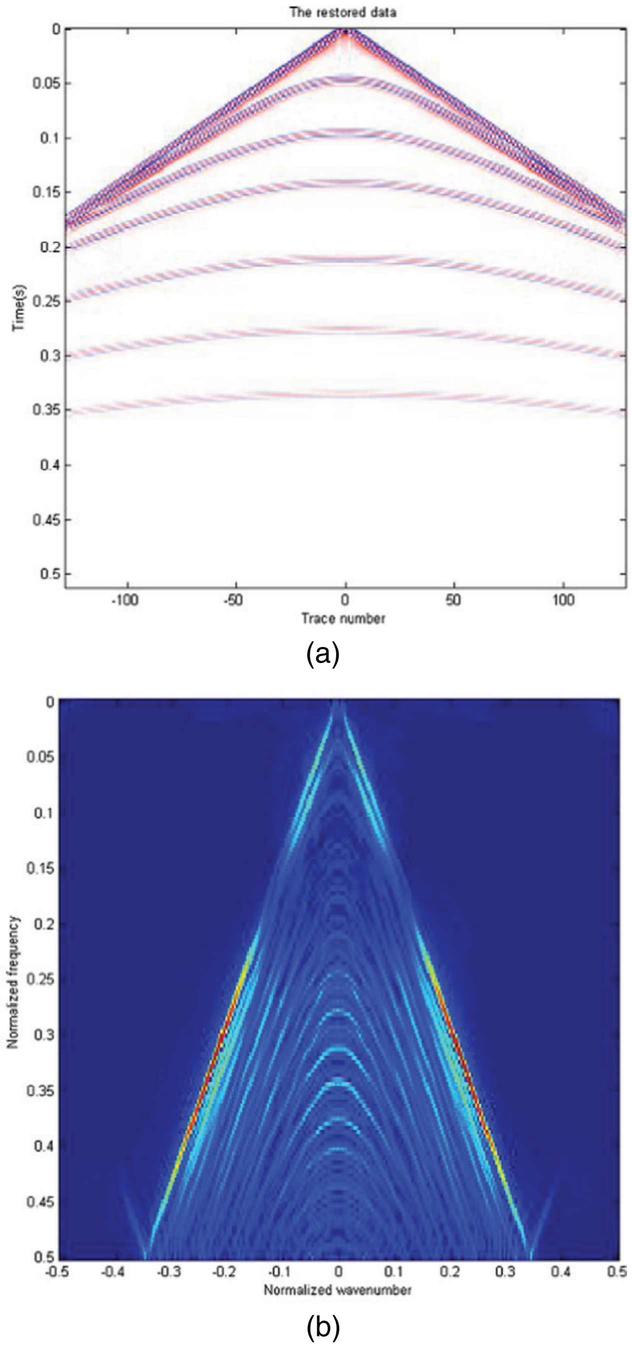

Figure 5. (*a*) Restoration of ISTc method; (*b*) F-K spectrum of (*a*).

For more choice of parameters in this algorithm, refer to van den Berg and Friedlander (2008).

Remark. This method is named as TRSL1. Compared with the original method in Wang *et al* (2011), the objective function of problem (6) is differential and its gradient and Hessian matrix can be calculated exactly which will save computation and improve the accuracy of solutions. In addition, projected gradient method for solving subproblems is very efficient for solving (7). These improvements make the proposed method suitable for large-scale computation.

Table 1. Comparison of synthetic data restoration with TRSL1, SPGL1 and ISTc.

Methods	CPU (s)	Relative error	SNR
TRSL1	337	0.2874	10.8309
SPGL1	268	0.2818	11.0010
ISTc	306	0.2794	11.0765

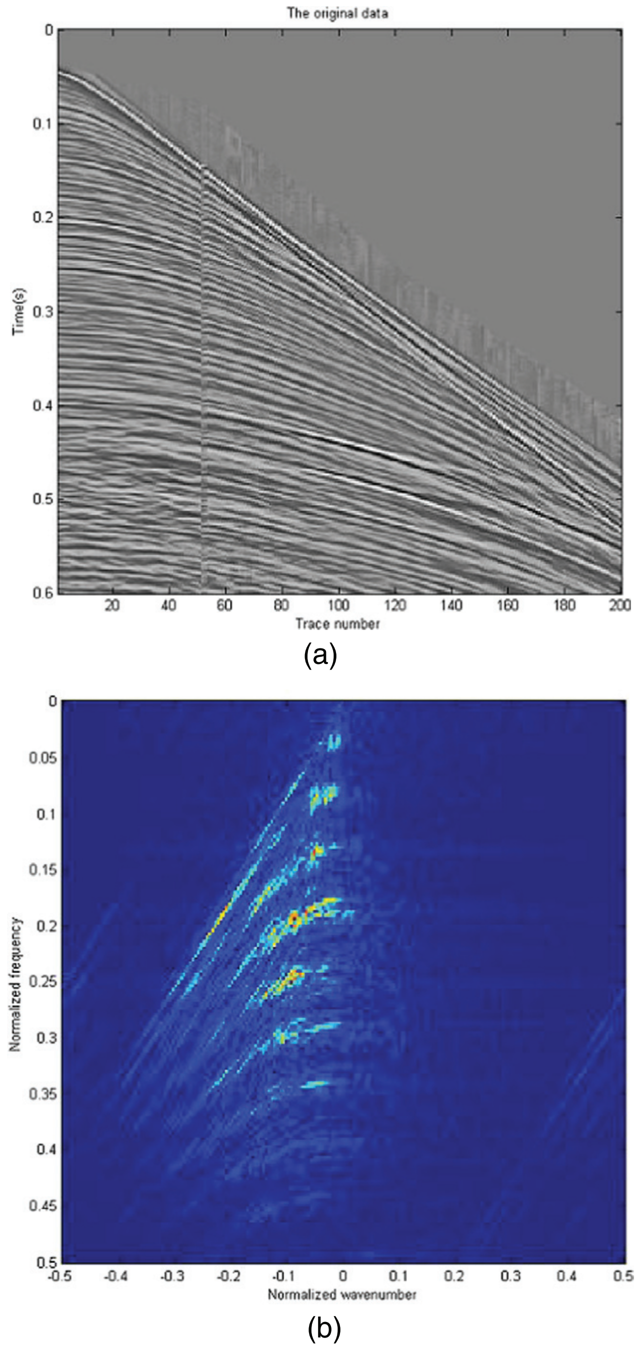

Figure 6. (*a*) The original data; (*b*) F-K spectrum of (*a*).

3. Numerical results

The performance of the proposed method is evaluated on synthetic as well as on real data. The synthetic data experiment aims to demonstrate its potential for large-scale computation; and the field data example can verify its ability for restoring field data.

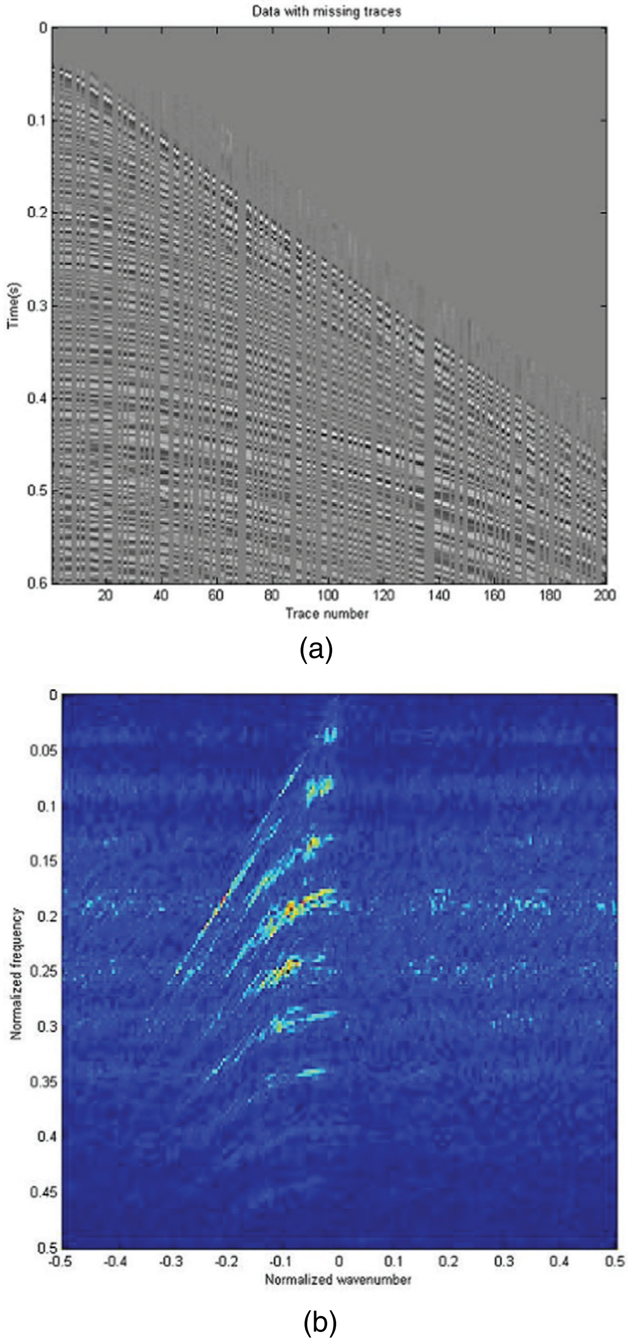

Figure 7. (*a*) The sampled data; (*b*) F-K spectrum of (*a*).

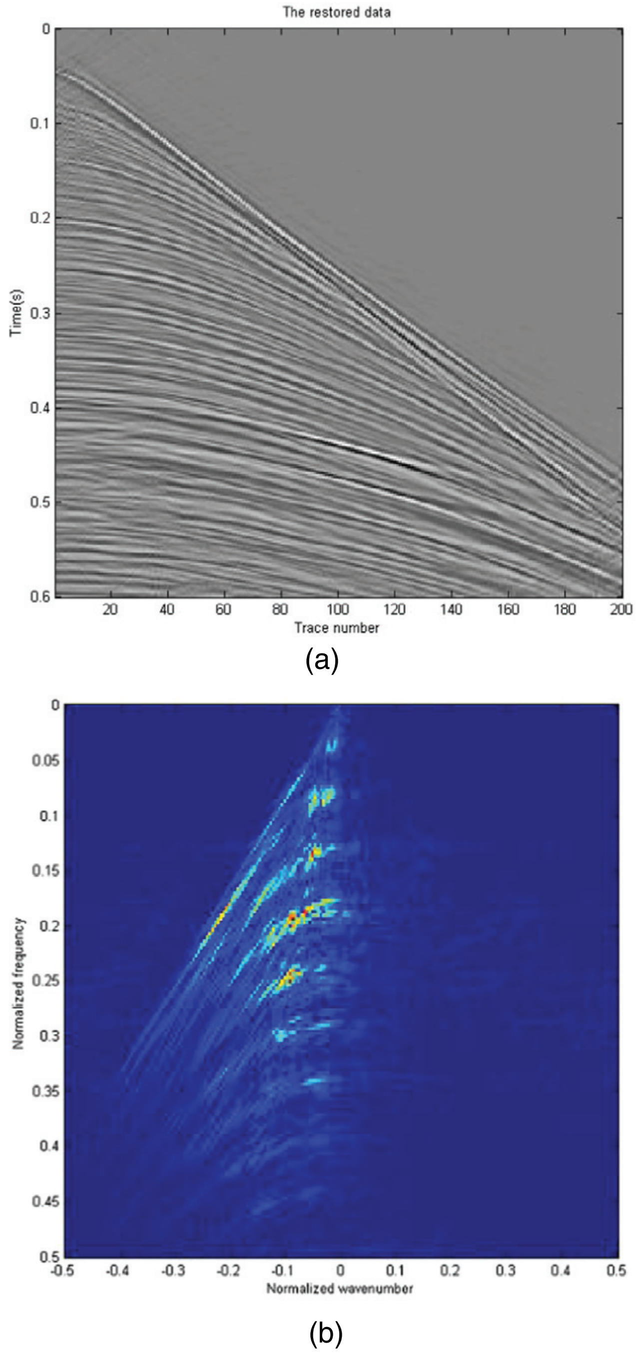

Figure 8. (*a*) Restoration of TRSL1 method; (*b*) F-K spectrum of (*a*).

3.1. Synthetic data example

Considering a seismic data with six layers, modeled with a 15 m receiver interval, 2 ms sampling interval and a source function given by a Ricker wavelet with central-frequency of 25 Hz. The dataset contains 256 traces with 256 time samples in each trace. The incomplete acquisition was simulated by randomly extracting 160 traces from 256 traces. The original data and its F-K spectrum are shown in figure 1; while the sampled data and its F-K spectrum are displayed in figure 2. SPGL1 method is robust and can get high precision results, IST with cooling (ISTc) is present in Herrmann and Hennenfent (2008), both of them are famous in seismic processing. The original L1 norm trust region method is too slow to contrast with these methods, so it can not make comparisons here. Thus, the proposed TRSL1 method is contrast with SPGL1 and ISTc method. In order to get comparable results, the inner loops number and outer loops number of TRSL1 are 5 and 3 respectively, restoration of TRSL1 is shown in figure 3. The max iteration of SPGL1 is 70; restoration with SPGL1 is shown in figure 4. For ISTc method, the inner loops are 5 and the outer loops are 15, restoration of ISTc is shown in figure 5. The CPU time, relative error and SNR of these methods are given in table 1, where SNR is defined as $\mathrm{SNR} = 10\log_{10}\|d_{\mathrm{orig}}\|_2^2 / \|d_{\mathrm{orig}} - d_{\mathrm{rest}}\|_2^2$ (d_{orig} means the original data and d_{rest} is the restored data) and

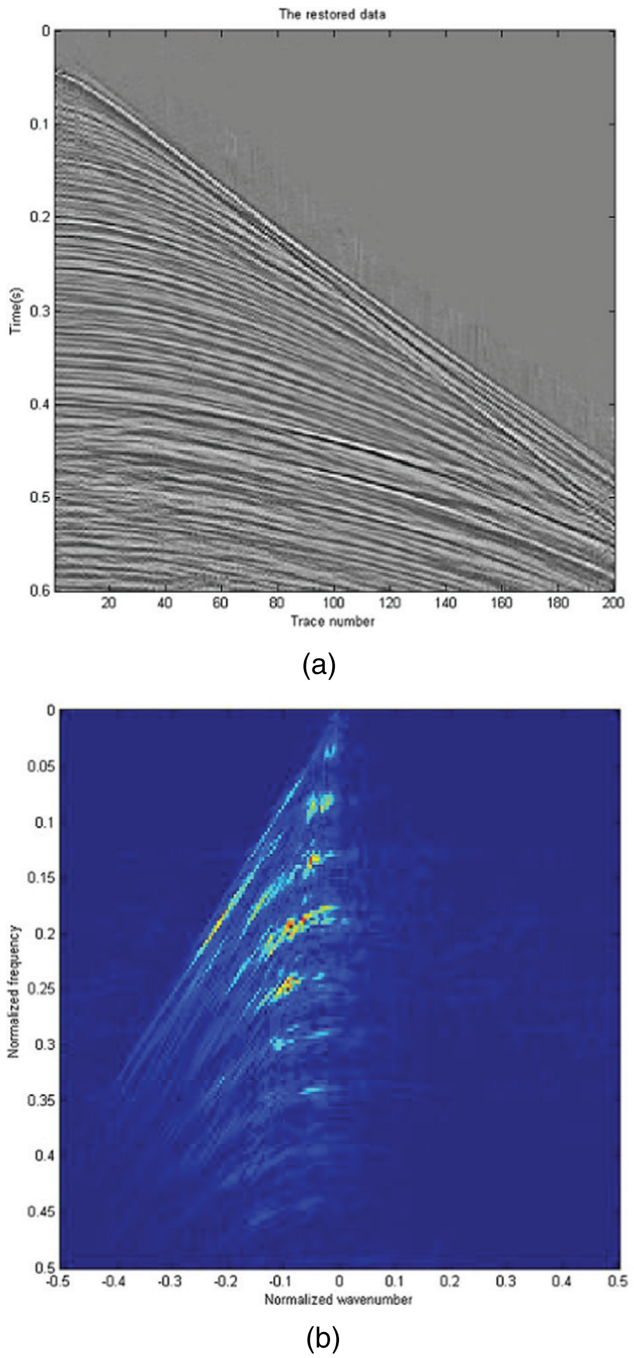

Figure 9. (*a*) Restoration of SPGL1 method; (*b*) F-K spectrum of (*a*).

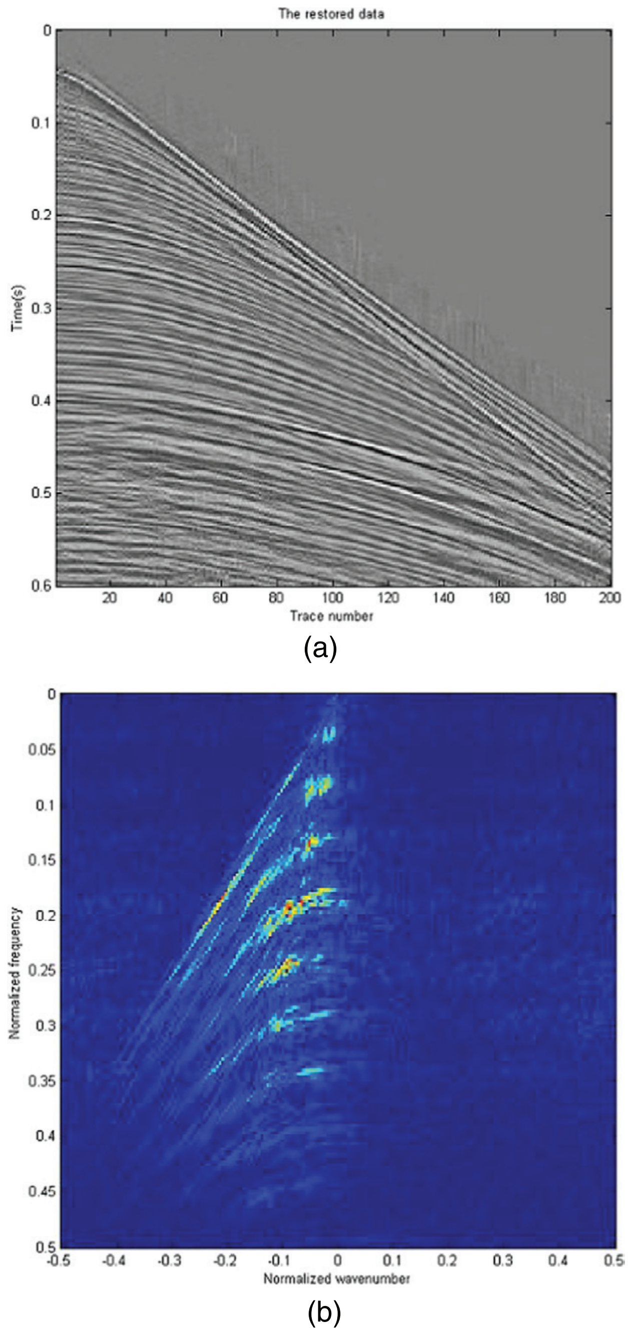

Figure 10. (*a*) Restoration of ISTc method; (*b*) F-K spectrum of (*a*).

the relative error is $\|d_{\mathrm{orig}} - d_{\mathrm{rest}}\|_2 / \|d_{\mathrm{orig}}\|_2$. It can be concluded from table 1 that, when getting the same level reconstruction, TRSL1 can reach the same computation speed as excellent line-searching methods.

3.2. Field data example

A marine data example is given to verify the ability of TRSL1 for field data in this section. Figure 6 displays one shot and its F-K spectrum. It contains 200 traces with the first 0.6 s of data. The sampling rate is 2 ms with a receiver spacing of 15 m. The decimated data with 75 traces randomly removed and its F-K spectrum are shown in figure 7. For TRSL1 method, the inner loops are 5 and the outer loops are 3, restored data with TRSL1 and its F-K spectrum are shown in figure 8. the max iteration of SPGL1 is set as 70; restoration with SPGL1 and their F-K spectrum are given in figure 9. For ISTc method, the inner loops are 5 and the outer loops are 15, restored data with ISTc and its F-K spectrum are shown in figure 10. The CPU time, relative error and SNR of these methods are listed in table 2. Based on these results, TRSL1 method still can reach the similar CPU time of these line-searching methods when obtain the same level results.

The inner loop number and outer loop number of TRSL1 are given according to experience, more information can refer

Table 2. Comparison of field data restoration with TRSL1, SPGL1 and ISTc.

Methods	CPU (s)	Relative error	SNR
TRSL1	293	0.4026	7.9018
SPGL1	308	0.4176	7.5849
ISTc	286	0.4071	7.8050

to Wang (2007). The initial trust region radius is related to the efficiency of this method, it must be carefully chosen. Generally, it should be increased with data scale and amplitude of the data.

4. Conclusions

In this paper, an improved L_1 norm trust region method which is based on sparsity of descent directions is proposed. The greatest difference between the proposed method and line-searching methods is the proposed method using sparse searching directions to get sparse solutions but not based on sparsity constraint of solutions. A simplified trust region model is given and the adopt method for solving sub-problem is suitable for large-scale problems. Numerical experiments of synthetic and field data demonstrate that the proposed method can get comparable computation speed with line-searching methods, so L_1 norm trust region method is feasible for seismic data restoration.

High dimensional restoration is more meaningful than 2D restoration for 3D seismic exploration (Trad 2009). 2D examples are given here, but this method can be easily generated to high dimensional restoration. Besides the optimization methods, less redundancy transforms can also improve computational efficiency, this is currently under consideration.

AQ2
Acknowledgments

Firstly, we are grateful for the valuable suggestions and comments raised by two referees and editor. In addition, the authors would like to thank the authors of Curvelab and Sparco for making their codes available. This work is supported by National Natural Science Foundation of China under grant numbers 41204075 and 11271349, and Natural Science Foundation of Hebei Province under grant number D2014403007.

AQ3
References

Abma R and Kabir N 2006 3D interpolation of irregular data with a POCS algorithm *Geophysics* **71** E91–7

van den Berg, E and Friedlander M P 2008 Probing the pareto frontier for basis pursuit solutions *SIAM J. Sci. Comput.* **31** 890–912

Birgin E G, Martinez J M and Raydan M 2000 Nonmonotone spectral projected gradient methods on convex sets *SIAM J. Optim.* **10** 1196–211

Candes E J, Demanet L, Donoho D and Ying L 2006 Fast discrete curvelet transforms *SIAM Multiscale Model. Simul.* **5** 861–99

Cao J J, Wang Y F and Yang C C 2012 Seismic data restoration based on compressive sensing using regularization and zero norm sparse optimization *Chin. J. Geophys.-Ch.* **55** 596–607

Cao J J, Wang Y F, Zhao J T and Yang C C 2011 A review on restoration of seismic wavefields based on regularization and compressive sensing *Inverse Problems Sci. Eng.* **19** 679–704

Daubechies I, Defriese M and Mol C D 2004 An iterative thresholding algorithm for linear inverse problems with a sparsity constraint *Commun. Pure Appl. Math.* **57** 1413–57

Herrmann F J and Hennenfent G 2008 Non-parametric seismic data recovery with curvelet frames *Geophys. J. Int.* **173** 233–48

Liu B and Sacchi M D 2004 Minimum weighted norm interpolation of seismic records *Geophysics* **69** 1560–8

LiuB, Sacchi M D and Trad D 2004 Simultaneous interpolation of 4 spatial dimensions *SEG Annual Meeting* (*Denver, Oct. 2004*) **23** 2009–12 Expanded Abstract AQ4

Naghizadeh M 2009 Parametric reconstruction of multidimensional seismic records *Phd Thesis* University of Alberta

Naghizadeh M and Sacchi M D 2010 Beyond alias hierarchical scale curvelet interpolation of regularly and irregularly sampled seismic data *Geophysics* **75** WB189–202

Sacchi M D and Liu B 2005 Minimum weighted norm wavefield reconstruction for AVA imaging **53** 787–801

Sacchi M D and Ulrych T J 1996 Estimation of the discrete Fourier transform, a linear inversion approach *Geophysics* **61** 1128–36

Sacchi M D, Verschuur D J and Zwartjes P M 2004 Data reconstruction by generalized deconvolution *SEG Annual Meeting* (*Denver, Oct. 2004*) **23** 1989–92 Expanded Abstracts

Schonewille M A, Romijn R, Duijndam A J W and Ongkiehong L 2003 A general reconstruction scheme for dominant azimuth 3D seismic data *Geophysics* **68** 2092–105

Soubaras R 2004 Spatial interpolation of aliased seismic data *SEG Annual Meeting* (*Denver, Oct 2004*) **23** 1167–70 Expanded Abstracts

Trad D 2009 Five-dimensional interpolation: recovering from acquisition constraints *Geophysics* **74** V123–32

Trad D, Ulrych T J and Sacchi M D 2002 Accurate interpolation with high-resolution time-variant radon transforms *Geophysics* **67** 644–56

Verschuur D J and Kabir M N 1995 Restoration of missing offsets by parabolic radon transform *Geophys. Prospect.* **43** 347–68

Wang J, Ng M and Perz M 2009 Fast high-resolution Radon transforms by greedy least-squares method *SEG Annual Meeting* (*Houston, Oct. 2009*) **28** 3128–32 Expanded Abstracts

Wang Y F 2007 *Computational Methods for Inverse Problems and Their Applications* (Beijing: Higher Education Press)

Wang Y F 2012 Sparse optimization methods for seismic wavefields recovery *Tr. Inst. Mat. Mekh. UrO RAN* **18** 42–55

Wang Y F, Cao J J and Yang C C 2011 Recovery of seismic wavefields based on compressive sensing by an l1-norm constrained trust region method and the piecewise random subsampling *Geophys. J. Int.* **187** 199–213

Wang Y F and Yuan Y X 2005 Convergence and regularity of trust region methods for nonlinear ill-posed inverse problems *Inverse Problems* **21** 821–38

Wang Y H 2010 Multichannel matching pursuit for seismic trace decomposition *Geophysics* **75** V61–6

Yuan Y X 1993 *Numerical Methods for Nonlinear Programming* (Shanghai: Shanghai Science and Technology Publication)

Zwartjes P and Gisolf A 2006 Fourier reconstruction of marine-streamer data in four spatial coordinates *Geophysics* **71** V171–86