Geophysical applications of a novel and robust L1 solver

Yunyue Li, Yang Zhang, and Jon Claerbout

ABSTRACT

L1-norm is better than L2-norm at dealing with noisy data and yielding blocky models, features crucial in many geophysical applications. In this report, we develop a hybrid-norm solver proposed by Claerbout (2009) to perform L1 regressions. The solver is tested on a 1-D field RMS velocity inversion, a 2-D regularized Kirchhoff migration inversion and a 2-D velocity analysis problem. The results of the inversions show that this solver can yield "blocky" models, and has the advantage of straightforward parametrization.

INTRODUCTION

L1 norm optimization is known to be a powerful estimator when the data are noisy or the model is sparse (Claerbout and Muir, 1973; Darche, 1989; Nichols, 0994; Guitton, 2005). However, the most widely used L1 solver–Iterated Reweighted Least-Squares (IRLS)–is cumbersome to use because users must specify numerical parameters with unclear physical meanings. To develop a robust, efficient L1 solver, Claerbout (2009) proposed a hybrid norm function to approximate the L1 norm (absolute value function), and he generalized the conjugate direction (CD) method using Taylor's expansion to guide the plane search in the minimization.

The parametrization of this proposed hybrid norm is straightforward. Users specify thresholds for the data residual and model residual, respectively. These thresholds determine the transition point from L2 to L1. The threshold for the data residual (R_d) can be chosen according to the signal-to-noise ratio in data space; the threshold for the model residual (R_m) can be specified by the desired blockiness in the model space.

In theory, the convergence of this hybrid norm solver is guaranteed, because the objective function is strictly convex. Nevertheless, difficulties may occur as the hybrid-norm approaches the L1 limit.

To test the performance and analyze the stability of this hybrid solver, we apply the solver to a 1-D field RMS velocity inversion, a 2-D regularized Kirchhoff migration inversion, and a 2-D velocity analysis problem. The inversion results show that this hybrid solver is robust and simple to use.

GENERALIZED CONJUGATE DIRECTION METHOD FOR THE HYBRID NORM

The hybrid norm is defined as

$$h(r) = \sqrt{r^2 + R^2} - R. \tag{1}$$

where r is the residual, and R is the corresponding threshold. In the limit, the hybrid norm (1) becomes:

$$h(r) = \begin{cases} |r| - R, & \text{if } R \ll |r| \\ r^2/(2R), & \text{if } R \gg |r|. \end{cases} \tag{2}$$

It is obvious that when R is small, the hybrid norm (1) reduces to L1 norm; when R is big, it becomes the L2 norm. Therefore, threshold R behaves as the turning point where the objective function changes smoothly from L2 to L1.

The Conjugate Direction method is commonly used for solving immense linear regressions in exploration geophysics. The idea of the CD method is to search the plane determined by the gradient and the previous step for the best step direction and length, instead of moving along the gradient direction. The best direction in that plane is the linear combination of the gradient and previous step vector that decreases the measure of the residual the most. Traditionally, the measure is chosen to be L2, for its simplicity; however, we generalize the CD method for any arbitrary convex measure C. Readers can determine which measure to use to satisfy their own objectives.

Now let us examine the generalization of the CD method in detail. At each iteration, we have the residual vector $\bar{r}$, the gradient g and the previous step s. Therefore, the updated residual can be written as:

$$r_i = \bar{r}_i + \alpha g_i + \beta s_i. \tag{3}$$

where α and β are scalars controlling the relative weights of these two directions. To determine these two scaling parameters, we need to minimize the measure of the residual:

$$N(\alpha, \beta) = \sum_i C(\bar{r}_i + \alpha g_i + \beta s_i). \tag{4}$$

The system given by directly setting the partial derivatives of (4) to zero is transcendental, thus difficult to solve. Therefore, we use Taylor's expansion to approximate the original objective function. The polynomial estimation of (4) is given as follows:

$$N(\alpha, \beta) \approx \sum_i \left(C_i + (\alpha g_i + \beta s_i) C_i' + (\alpha g_i + \beta s_i)^2 C_i''/2 \right). \tag{5}$$

Now, taking the derivatives of the parabolic function in (5) with respect to α and β and setting them to zero, we end up with a linear system of α and β:

$$\left\{ \sum_i C_i'' \left[\begin{pmatrix} g_i \\ s_i \end{pmatrix} \begin{pmatrix} g_i & s_i \end{pmatrix} \right] \right\} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = -\sum_i C_i' \left[\begin{matrix} g_i \\ s_i \end{matrix} \right]. \tag{6}$$

where (C_i, C'_i, C''_i) refer to a Taylor expansion of $C(r)$ about r_i.

Then we can obtain α, β by simply solving a set of 2×2 linear equations.

Notice that the α, β we get here is minimizing the approximated function (5), not the original objective function. Therefore, it is necessary to solve for α, β multiple times within each CD iteration. By doing this relatively cheap plane-search loop, we expect to save the number of iterations for the outer loop (Conjugate Direction), which is usually much more computational intensive (requiring application of both the forward and adjoint operator).

DIX INVERSION OF INTERVAL VELOCITY ESTIMATION

The Dix equation (Dix, 1952) inverts interval velocities from RMS velocity, which is picked during velocity scanning in prestack seismic data. The equation can be written as

$$v_{int(k)}^2 \quad = \quad kV_k^2 - (k-1)V_{k-1}^2, \tag{7}$$

or

$$\sum_{i=1}^{k} v_{int(k)}^2 \quad = \quad kV_k^2, \tag{8}$$

where v is interval velocity, V is RMS velocity, and k is the sample number, which can be regarded as travel-time depth. Directly calculating the interval velocity from this formula can easily yield wildly unreasonable results because of the error in the picked RMS velocity. Therefore, it is necessary to solve this problem as a regularized inversion. To linearize the problem, we choose the model space to be the squared interval velocity (v_{int}^2), instead of the interval velocity itself (v_{int}).

Thus we can formulate the Dix inversion problem as follows:

$$\boldsymbol{W_d}(\mathbf{Cu} - \mathbf{d}) \approx \mathbf{0}, \tag{9}$$

$$\epsilon \mathbf{D_z u} \approx \mathbf{0}. \tag{10}$$

In the data-fitting goal (9), $\mathbf{u}$ is the unknown model we are inverting for, $\mathbf{d}$ is the known data computed from the RMS velocity, $\mathbf{C}$ is the causal integration operator and $\mathbf{W}_d$ is a data residual weighting function, which is a measure of to our confidence in the RMS velocity. In the model-styling goal (10), $\mathbf{D}_z$ is the vertical derivative of the velocity model and ϵ is the weight controlling the strength of the regularization.

The input RMS velocity with 1000 samples is shown in Figure 1. It is obvious that the violent variation at the end of the trace is not realistic. Thus, we use the hybrid-norm to ignore the large residuals in the data-fitting, which are considered to be noise. At the same time, to obtain a blocky interval velocity model, the large residual in the derivative of the interval velocity should be "invisible" to the measure. Therefore, the hybrid norm on the model styling appear to be the best choice.

To compare the inversion result, we also use the IRLS and L2 solver on the same data with comparable parameters. The inversion results are shown in Figure 2. The left column is the inverted interval velocity, while the right column is the corresponding reconstructed RMS velocity. The result shows that compared with the IRLS and L2 result, the hybrid solver successfully retrieved the most blocky velocity model, and the corresponding reconstructed RMS velocity contains less noise while keeping the trend of the original data.

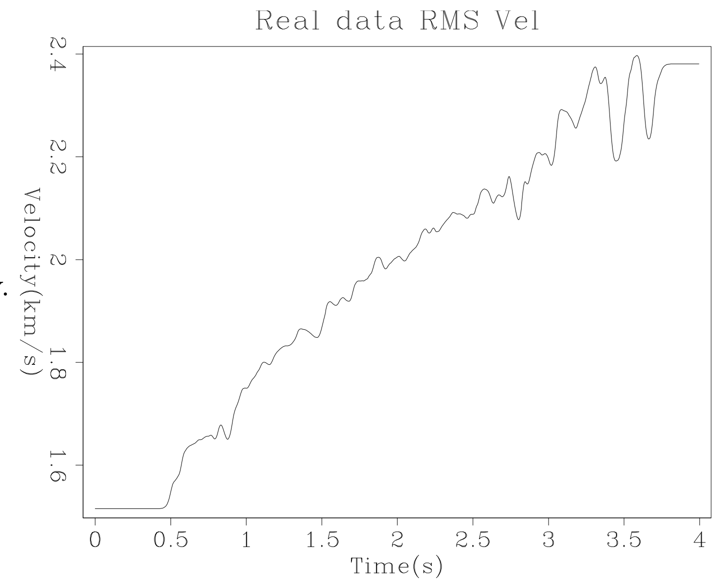

Figure 1: Input 1-D RMS velocity. [**ER**]

SIMPLE KIRCHHOFF INVERSION

Kirchhoff migration was widely used before the era of wave-equation migration for marine data, and is still the principle migration method for land data. It always involves summing over or spreading along certain travel-time surfaces in 3-D, which reduce to curves in 2-D. For the purpose of testing our solver, we define the forward operator to be the Kirchhoff modeling operator, whose adjoint is the traditional Kirchhoff migration operator.

We formulate the inversion problem as follows:

$$\mathbf{Hm} \approx \mathbf{d} \tag{11}$$

$$\epsilon\mathbf{m} \approx \mathbf{0} \tag{12}$$

where **H** is the forward Kirchhoff modeling operator, **m** is the subsurface reflectivity model, and **d** is the seismic response recorded at the surface. The second equation is a damping term, where the hybrid-norm is applied to retrieve the sparse model.

In field acquisition, data usually have denser sampling rates in the in-line direction than the cross-line direction. Therefore, the surface-recorded data are always aliased in the cross line direction. To illustrate the problem in the cross-line direction, figure

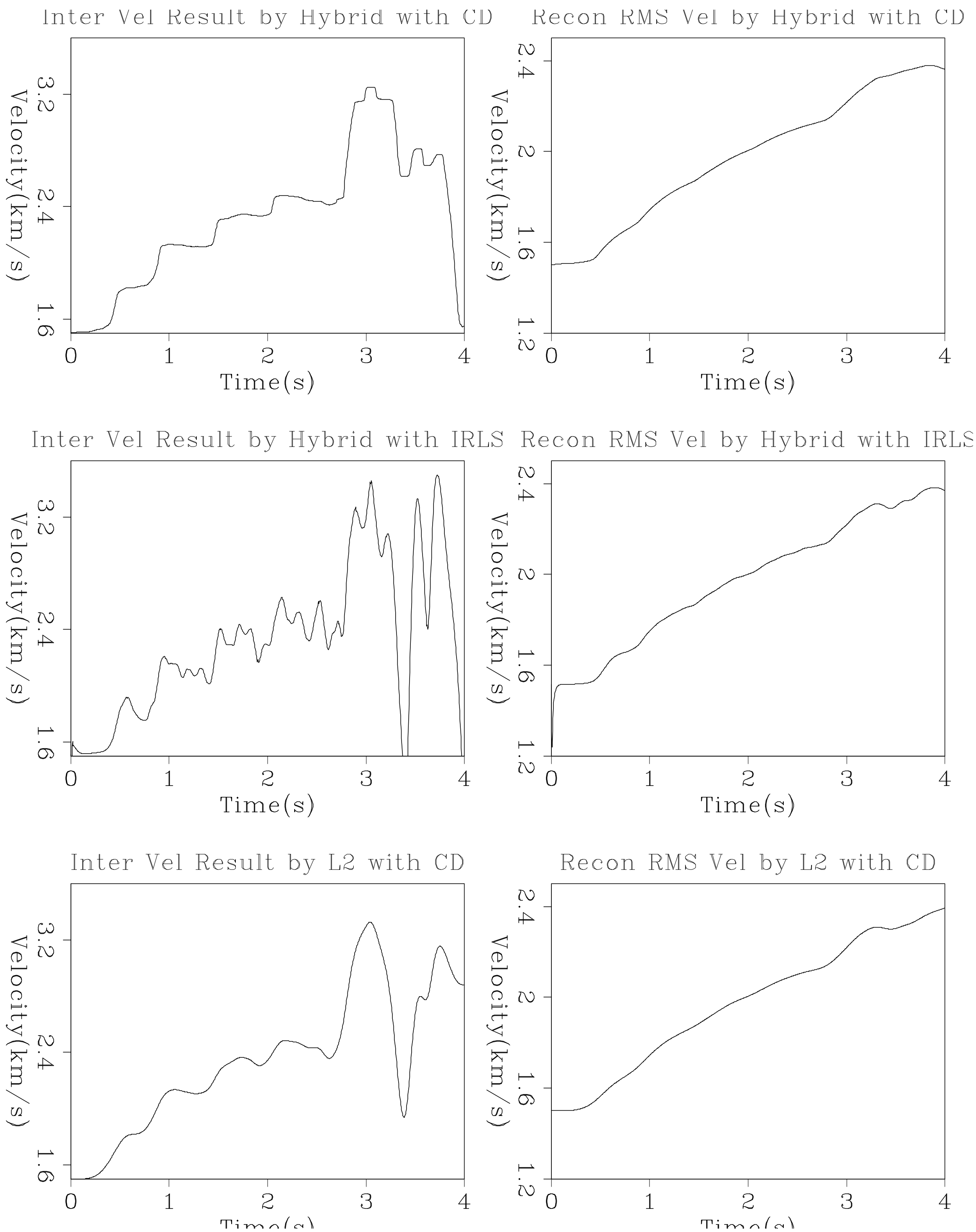

Figure 2: Comparison of the inversion results. Panels in the left column are the estimated interval velocity, while panels on the right are the corresponding reconstructed RMS velocity. Top panels: results of the hybrid with CD; Middle panels: results of the hybrid with IRLS; Bottom panels: results of the L2 norm with CD. Notice that although the reconstructed RMS velocity from the three methods are more or less the same, the interval velocity from hybrid CD is more blocky then the other two. [**ER**]

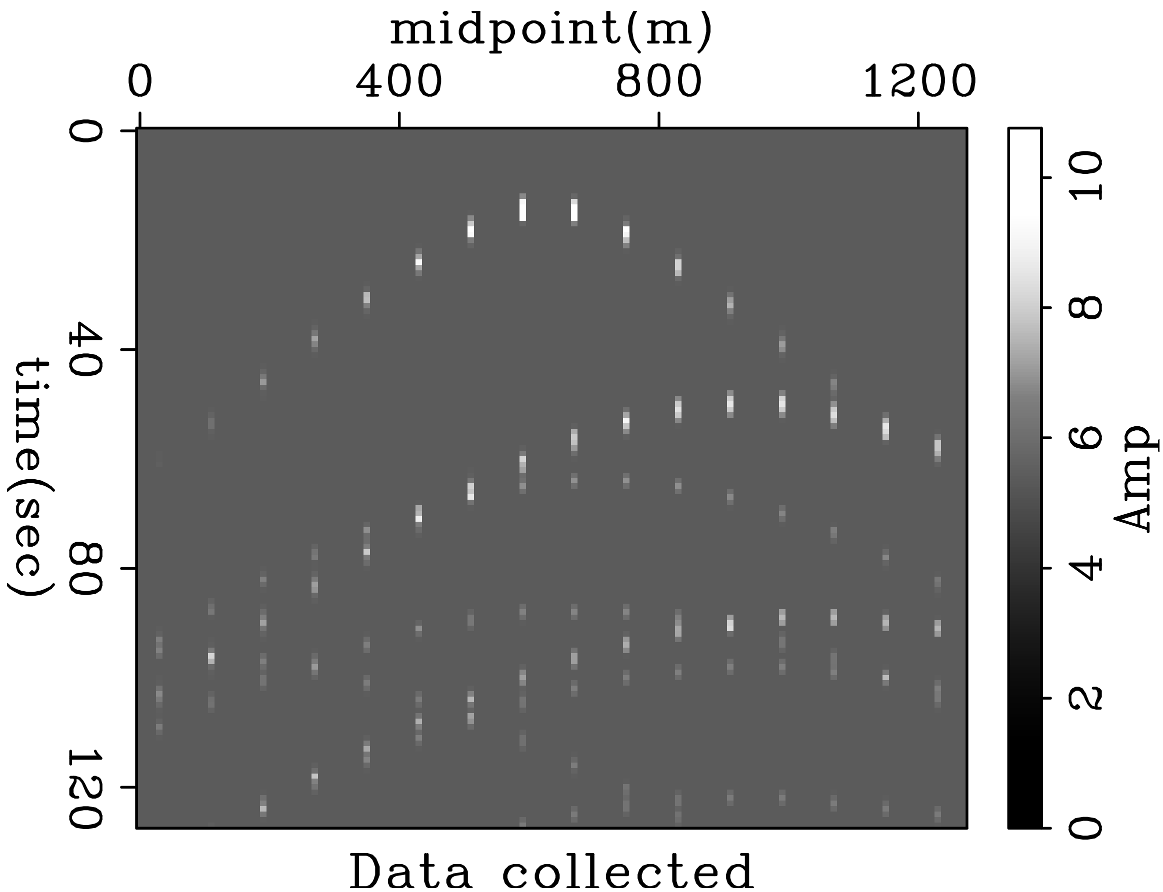

Figure 3: Highly aliased hyperbola. Input data for the Kirchhoff inversion. [**ER**]

3 shows an example of highly aliased hyperbolas.The aliasing makes the inversion problem an underdetermined problem; therefore, the result of the inversion relies heavily on the regularization. With the model space sampling being 128×128, the sampling of data space is only 128×16. Also note that some of the hyperbolas are not symmetric; therefore the tops of the hyperbolas are shifted.

Same as the previous example, we experimented with different solvers: L2, IRLS and hybrid, to compare their results.

Figure 4 shows the inversion results with different schemes. The results show that the hybrid norm is superior for retrieving the spiky result that resembles the original model the best. Although severely aliased, the inversion result recovers the exact position, the correct size and most of the amplitude. Notice that the CD hybrid solver recovers the very low amplitude spike at the left edge. This promising result suggests that by choosing the regularization properly, we can overcome the aliasing problem in the presence of a sparse model.

Figure 5 shows the reconstructed data from the CD hybrid solver. The original data are accurately recovered. Notice the hyperbola with its top at the left edge is

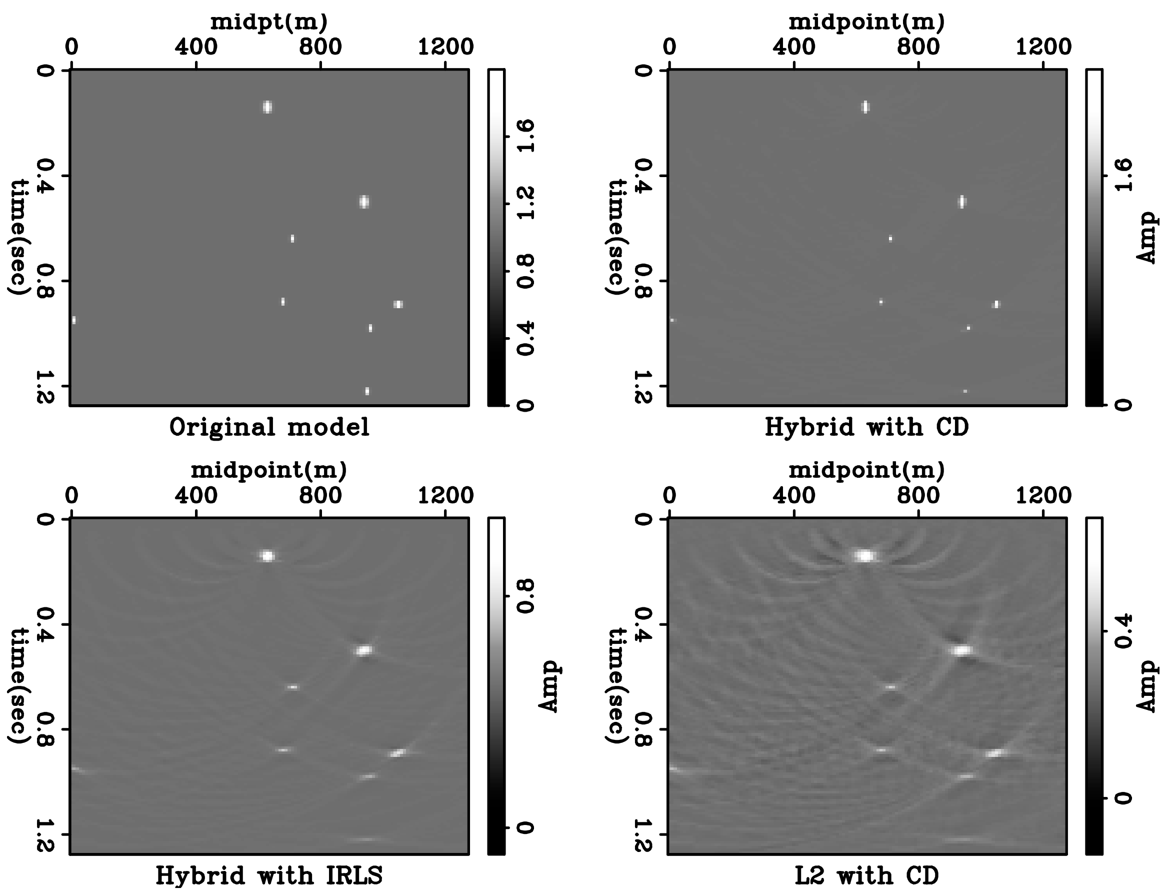

Figure 4: Original model and inversion results by different methods. Top left: original model; Top right: Hybrid result with CD; Bottom left: Hybrid result with IRLS; Bottom right: L2 result with CD. [**ER**]

well resolved.

VELOCITY ANALYSIS AS INVERSION

Velocity analysis is one the most critical and problematic procedures in seismic exploration. In data with noise bursts, velocity analysis is prone to error and even unrealistic results. Therefore, to handle this problem robustly, we formulate velocity analysis as an inversion problem as follows:

$$\mathbf{Hm} \approx \mathbf{d} \tag{13}$$

where **H** is the modeling operator, whose adjoint operator is the slowness scan operator; **m** is the slowness field, and **d** is the data we collect after one shot.

Figure 6 shows a shot gather from Yilmaz's dataset. There are two distinct types of noise in these data: first is the linear noise caused by all kinds of surface waves, which can be attenuated by taking advantage of their physical properties; second is the abnormally high-amplitude burst noise at the near offsets, which is difficult to fit into a statistical model.

Figure 7 shows the inversion results for different methods. Because of the existence of the high amplitude noise at the near-offset, a velocity scan without inversion yields no meaningful result. For the L2 inversion, the noise has contaminated the whole panel, making it impossible to see the velocity trend. The inversion results of both IRLS and HYCD show clear velocity trends, and the near-offset burst noise is successfully removed in the reconstructed data.

CONCLUSIONS

An L1-type optimizer is required to retrieve a blocky model or extract information from noisy data. To avoid the troublesome parameter tuning of IRLS, we develop a hybrid-norm Conjugate Direction solver with straightforward parametrization to achieve the goal in L1 sense. Tests of this solver on three different problems, with both synthetic and field data, show that this hybrid-norm solver is robust and simple to use. These promising results encourage us to apply this solver on large-scale, real-world problems.

ACKNOWLEDGMENTS

We thank Mohammad Maysami, Mandy Wong, and Nader Moussa for the earlier work. We like to thank Robert Clapp, Luis and Ali Almomin for fruitful discussions about the development of this solver and the choice of the test cases.

REFERENCES

Claerbout, J. F., 2009, Blocky models via the l1/l2 hybrid norm: SEP-Report, **139**, 1–10.

Claerbout, J. F. and F. Muir, 1973, Robust modeling with erratic data: Geophysics, **18**, 826–844.

Darche, G., 1989, Iterative l_1 deconvolution: SEP 61, 281–302.

Dix, C. H., 1952, Seismic prospecting for oil.

Guitton, A., 2005, Multidimensional seismic noise attenuation: PhD thesis, Stanford University.

Nichols, D., 0994, Velocity-stack inversion using l^p norms: SEP 82, 1–16.

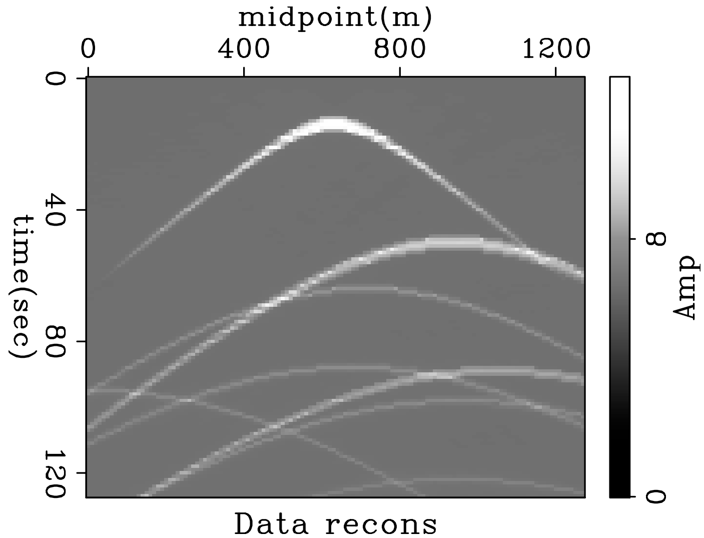

Figure 5: Data reconstructed by the hybrid solver with CD. The CD hybrid solver accurately recovers the original data. Notice the hyperbola with its top at the left edge is well resolved. [**ER**]

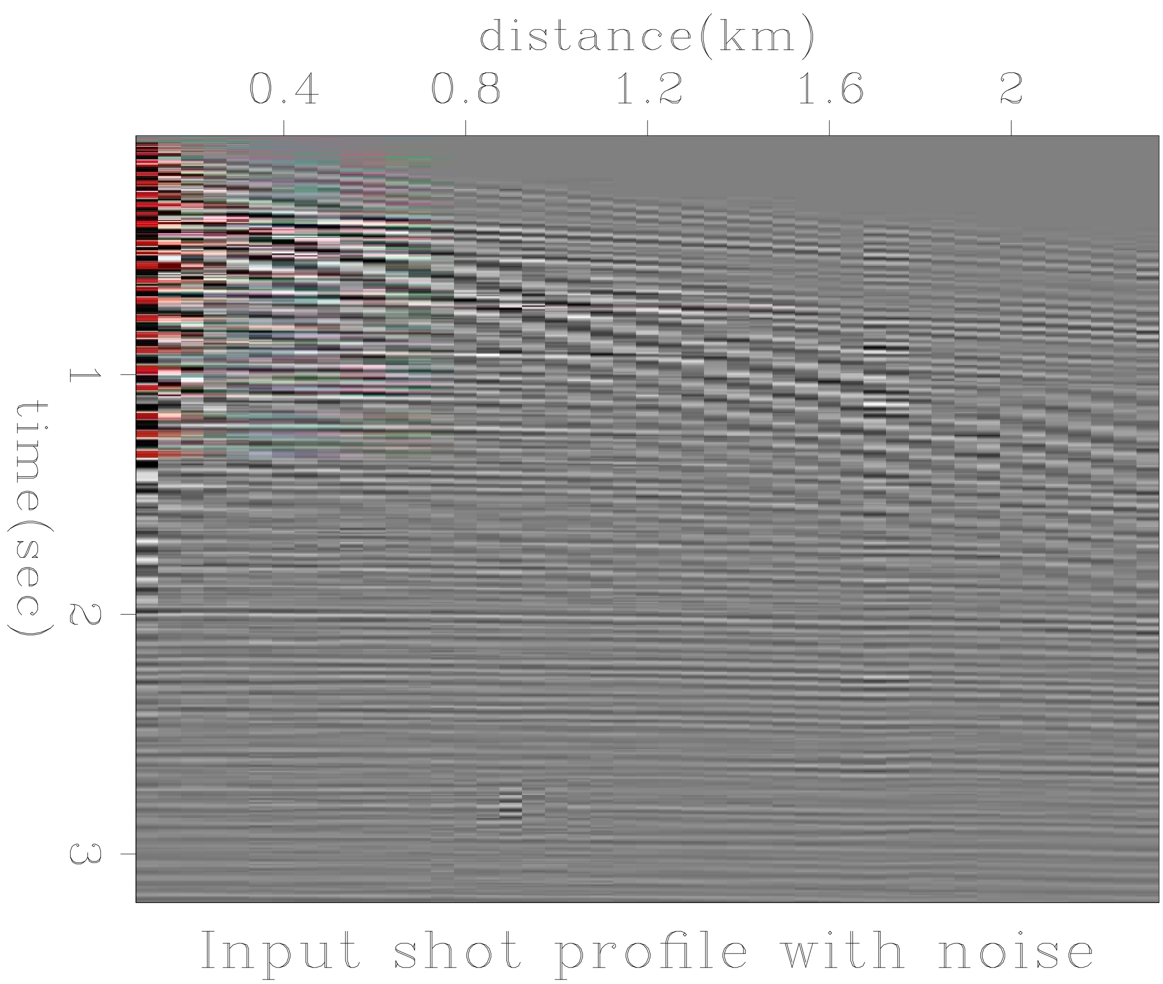

Figure 6: Input shot profile with noise.The *red* points in the figure indicates the data points with very high amplitude. [**ER**]

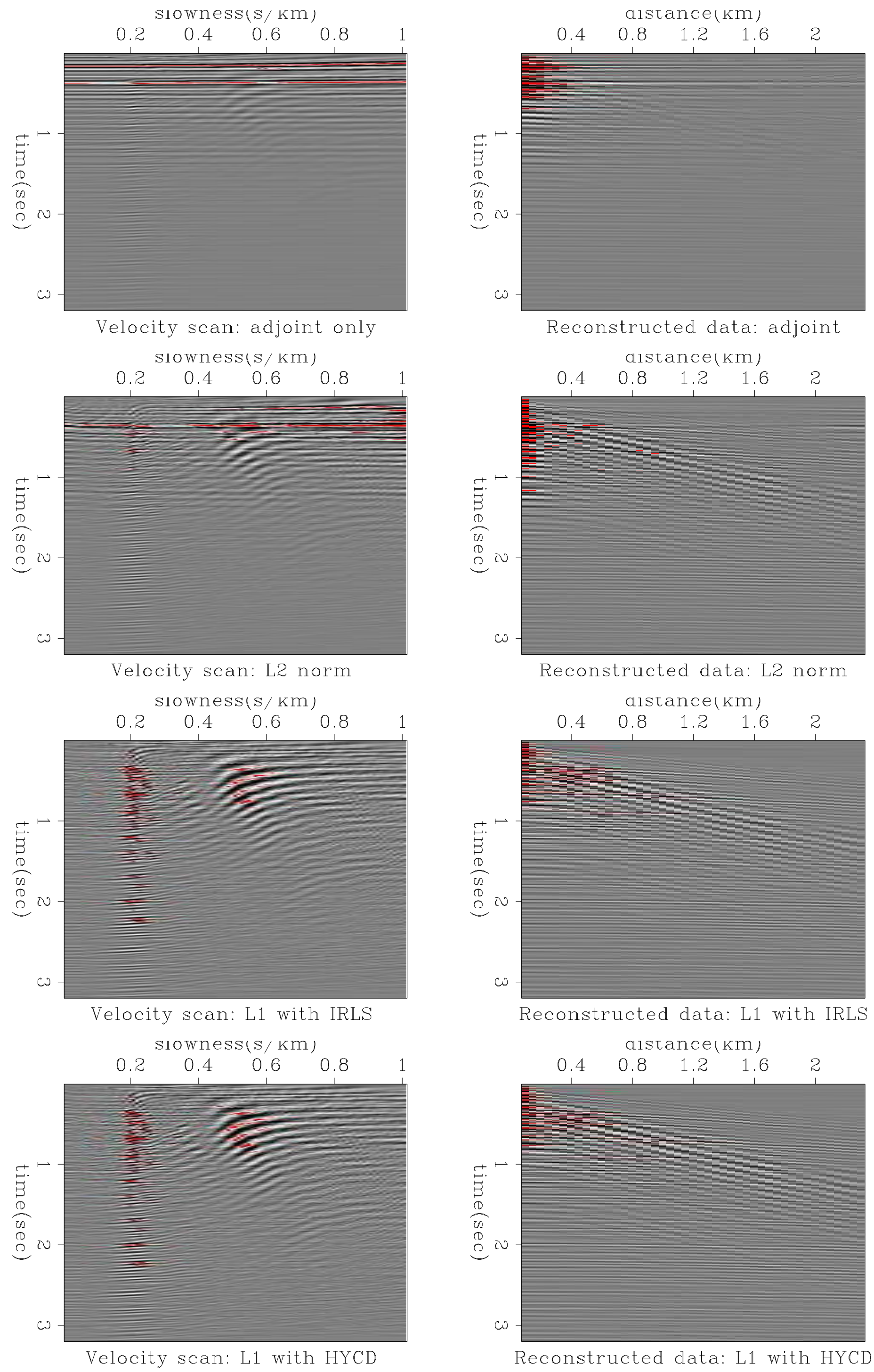

Figure 7: Inversion results of different methods. The *red* points in the figure indicates the data points with very high amplitude. Panels in the left column are the results of velocity scans, while panels on the right are the corresponding reconstructed data. First row: No inversion is applied (adjoint only); Second row: Inversion results by CD with L2; Third row: Inversion results by hybrid with IRLS; Bottom row: Inversion results by CD with L1. [**ER**]