Dix inversion constrained by L1-norm optimization

Yunyue (Elita) Li and Mohammad Maysami

ABSTRACT

To accurately invert for velocity in a model with a blocky interval velocity inversion using Dix inversion, we set up our optimization objective function using $L1$ criterion. In this study, we analyze and test an improved version of the Iterative Reweighted Least Squares (IRLS) solver, a hybrid $L1/L2$ solver and a conjugate direction $L1$ solver. We use a 1-D synthetic velocity data set and a 1-D field RMS velocity data set as test cases. The results of the inversion are promising for applications on realistic geophysical problems.

INTRODUCTION

Dix formula (Dix, 1952) estimates interval velocities from picked stacking velocities. The conventional result of constrained least-squares Dix inversion (Koren and Ravve, 2006; Harlan, 1999; Clapp, 2001) is always a smooth velocity model, because the regularization is imposed in the $L2$ sense. However, to represent a geological environment with sharp velocity contrasts, e.g., carbonate layers, salt bodies and strong faulting, we may need a blocky velocity model rather than a smooth velocity model. Valenciano et al. (2003) proposed to use edge-preserving regularization with Dix inversion in order to get sharp edges in interval velocity. However, one of the solvers they used, IRLS (Iterative Re-weighted Least Squares) is cumbersome to use because users must specify numerical parameters with unclear physical meaning.

$L1$-norm optimization is known to be a robust estimator to yield sparse models. Many works (Claerbout and Muir, 1973; Darche, 1989; Nichols, 1994; Guitton, 2005) has shown that $L1$-norm is not sensitive to outliers, while it penalizes the small residuals down to zero. In theory, when the model space is sparse and the data are noisy, regressions produced by $L1$ optimization always outperform those produced by $L2$ norms.

In this study, we analyze, improve and test different methods on a simple synthetic problem as well as a field-data problem. We aim to develop robust and efficient solvers to perform regressions of an $L1$ nature. We initially improve the traditional IRLS method, explore the conjugate direction $L1$ method, and finally test an $L1/L2$ hybrid method. The inversion results of a 1-D, synthetic, 2-step, interval-velocity model and a 1-D field data example are given at the end of the paper.

DIX INVERSION AS AN $L1$-OPTIMIZATION PROBLEM

The linear relationship between the RMS velocity and the square of the interval velocity is given by Dix Equation:

$$v_k^2 \ = \ kV_k^2 - (k-1)V_{k-1}^2, \tag{1}$$

where v is the interval velocity, V is the stacking velocity or RMS velocity, and k is the sample number. Both velocities run down the traveltime depth axis. If we define $u_k = v_k^2$ and $d_k = kV_k^2$, we can set up the Dix inversion problem in an $L1$ sense as follows:

$$||\boldsymbol{W_d}(\boldsymbol{C}\mathbf{u} - \mathbf{d})||_1 \approx 0, \tag{2}$$

where $\mathbf{u}$ is the unknown model we are inverting for, $\mathbf{d}$ is the known data from velocity scan, $\mathbf{C}$ is the causal integration operator, $\mathbf{W}_d$ is a data residual weighting function, which is proportional to our confidence in the RMS velocity.

Fitting goal (2) itself cannot fully constrain the inversion problem, because the integration operator has a large null space at high frequencies. Therefore, Clapp et al. (1998) supplement this system with a regularization term to take the advantage of the prior geological information, of which smoothness and blockiness are two typical examples. For the case we are interested in, we use blockiness as regularization. In a mathematical form, it can be written as follows:

$$||\epsilon \boldsymbol{D_z}\mathbf{u}||_1 \approx 0, \tag{3}$$

where $\mathbf{D}_z$ is the vertical derivative of the velocity model and ϵ is the weight controlling the strength of the regularization.

DIX INVERSION BY IRLS METHOD

Bube and Langan (1997) and Tang (2006) show that the nonlinear objective functions, such as the regression equation (3), can be solved by the IRLS algorithm. Many authors have demonstrated successful applications of IRLS as a robust estimator to yield sparse models. To take advantage of the well-established $L2$ norm regression, we can transform the problem by introducing a diagonal weighting function $\boldsymbol{W_p}$. Then the fitting goal 2 and the regularization 3 become:

$$||\boldsymbol{W_d}(\boldsymbol{C}\mathbf{u} - d)||_2 \approx 0, \tag{4}$$

$$||\epsilon \boldsymbol{W_p}\boldsymbol{D_z}\mathbf{u}||_2 \approx 0, \tag{5}$$

where $\boldsymbol{W_p}$ is a diagonal weighting function on the model residual. To use the gradient-based method, we have to recompute the weight $\boldsymbol{W_p}$ at each iteration, and the algorithm can be summarized as follows:

1. Set the initial weighting matrix $\boldsymbol{W_p^{(0)}}$ to equal the identity matrix:

$$\boldsymbol{W_p^{(0)}} \ = \ I \tag{6}$$

2. At the $k-th$ iteration, the $i-th$ element of the weighting matrix is recomputed as

$$W_{pi}^{(k-1)} = (|p_i^{k-1}|)^{-1/2} \tag{7}$$

where $\mathbf{p} = \boldsymbol{D_z}\mathbf{u}$ is the model residual. However, applying equation (7) to compute the weighting matrix is dangerous when p_i is zero or too small. One way to avoid this issue is to bound the weights at a certain cutoff σ:

$$\mathbf{W}_{pi}^{(k-1)} = \begin{cases} (|p_i^{k-1}|)^{-1/2} & \text{if } |p_i^{k-1}| \geq \sigma \\ \sigma^{-1/2} & \text{otherwise.} \end{cases} \tag{8}$$

One of the most important disadvantages of IRLS algorithm arises in equation 8: how should we choose the cutoff number σ? We would like to derive this number automatically according to its physical meaning, instead of cumbersome numerical experiments.

Further examining the weighting function, we notice that when applying the truncated weights, we end up treating small residuals in the $L2$ norm, and at the turning point ($p = \sigma$) we have a sharp transition to the $L1$ norm. Thus, σ is the cutoff between the $L1$ region and the $L2$ region, determining the tolerance to the large residuals. Therefore, we can choose σ according to the desired blockiness of the model space. For the synthetic example, which is a 40-point-long interval velocity model with three layers, we expect only two spikes out of those 40 points in the derivative. Therefore we would like σ to be around the 95^{th} percentile of the derivative, allowing 5% of the spikes to be of unlimited size, while the others are small.

DIX INVERSION BY AN $L1/L2$ HYBRID METHOD

Extending the discussion in last section, many authors (Bube and Langan, 1997; Claerbout, 2009) generalize the optimization problem of Dix inversion. Claerbout (2009) points out that any arbitrary norm P can be used as a penalty function. Then the optimization problem can be written as:

$$0 \approx \sum_i P(\sum_j C_{i,j} u_j - d_i), \tag{9}$$

where $P()$ is a convex function of a scalar, $C_{i,j}$, u_j, and d_i are elements in the causal integration operator C, the model u and the known data d.

We have special interests in an $L1/L2$ hybrid norm, because instead of a sharp transition, this norm provides a smooth transition between $L1$ and $L2$. This can be shown in the formulation of the hybrid norm:

$$P = \sigma^2(\sqrt{1 + r^2/\sigma^2} - 1) \tag{10}$$

where r is the data residual and σ is the same threshold as in last section, and thus can be chosen according to the same physical explanation. The hybrid norm approaches the least squares limit as $\sigma \rightarrow \infty$ and approaches the $L1$ limit as $\sigma \rightarrow 0$.

The first and the second derivative of this hybrid norm with respect to the residuals are given in equation 11 and equation 12. We can see the penalty function transits from $L2$ to $L1$ smoothly at $r = \sigma$ since the first derivative is continuous.

$$P' = \frac{r}{\sqrt{1 + r^2/\sigma^2}} \tag{11}$$

$$P'' = \frac{1}{(1 + r^2/\sigma^2)^{3/2}} \tag{12}$$

Claerbout (2009) also proposed a new method based on Taylor's series to search the plane spanned by the gradient and the previous step. He embedded the new iterative bivariate solver in a conjugate direction solver, hoping for significant savings by expending more effort to find a better next step. In this experiment, we use the solver coded by Maysami and Mussa (2009), who adapt Claerbout's theory. For more information, refer to these two papers.

DIX INVERSION BY CONJUGATE DIRECTION $L1$ METHOD

Choosing proper parameters for hybrid $l1/l2$ method and IRLS is still quite empirical, even when we understand their physical meanings. Instead, conjugate direction methods do not need setting parameters. Thus, we develop a similar conjugate direction method in $L1$ sense. The pseudo code of this method is given in Table 1.

The structure of the conjugate-direction $L1$ method is similar to the $L2$ conjugate-direction solver given by Claerbout (2008). The main difference arises in the part of plane search.

Given the gradient $\mathbf{g}^{(d)}$, previous step $\mathbf{s}^{(d)}$, and the current residual $\mathbf{r}^{(d)}$, we construct the $2 \times N$ matrix $\mathbf{B} = [\mathbf{g}^{(d)}\ \mathbf{s}^{(d)}]$ and the column vector $[\alpha\ \beta]'$. We seek to find $[\alpha\ \beta]'$ that minimizes $\mathbf{0} \approx \mathbf{r}^{(d)} + \mathbf{B}[\alpha\ \beta]'$ in the L1-norm sense. This bivariate regression embedded in the plane search is solved in an iterate manner.

At the ultimate solution of the bivariate regression there will be two basis equations that are exactly satisfied. The first one is found by steepest descent. After the first iteration, we do plane searches using the weighted median solver to choose the best equation to be exactly satisfied. "Best equation" is the one that decreases the residual the most, while satisfying the equation chosen by the previous iteration exactly as well.

Table 1: Pseudo Code - Conjugate direction $L1$ solver using Weighted Median

Initialization :			
$\mathbf{m}$	$=$	$\mathbf{m}_{init}$	
$\mathbf{r}^{(d)}$	$=$	$\boldsymbol{F}\mathbf{m} - \mathbf{d}$	
$\mathbf{s}$	$=$	$\mathbf{0}$	
$\mathbf{s}^{(d)}$	$=$	$\mathbf{0}$	
Iteration i :			
$\mathbf{r}^{(t)}$	$=$	$sgn(\mathbf{r}^{(t)})$	
$\mathbf{g}$	$=$	$\boldsymbol{F'}\mathbf{r}^{(t)}$	$*$
$\mathbf{g}^{(d)}$	$=$	$\boldsymbol{F}\mathbf{g}$	$*$
(α, β)	$=$	$\underset{(\alpha,\beta)}{\text{argmin}} \|\|\mathbf{r}^{(t)} + \alpha\mathbf{g}^{(d)} + \beta\mathbf{s}^{(d)}\|\|_1$	
$\mathbf{s}^{(d)}$	$=$	$\alpha\mathbf{g}^{(d)} + \beta\mathbf{s}^{(d)}$	
$\mathbf{r}^{(d)}$	$=$	$\mathbf{r}^{(d)} + \mathbf{s}^{(d)}$	
$\mathbf{s}$	$=$	$\alpha\mathbf{g} + \beta\mathbf{s}$	
$\mathbf{m}$	$=$	$\mathbf{m} + \mathbf{s}$	

To do this, suppose the previous equation is $g_k^{(d)}\alpha + s_k^{(d)}\beta + r_k^{(d)} = 0$. We seek a $(\Delta\alpha, \Delta\beta)$ that still satisfies the equation k. This requirement gives a solution to $(\Delta\alpha, \Delta\beta)$ as $\gamma(s_k^{(d)}, -g_k^{(d)})$, where γ is a scalar. Then the plane search becomes $\mathbf{0} \approx \gamma\mathbf{B}[s_k^{(d)}, -g_k^{(d)}]' + r^{(d)}$, which is a weighted median problem. Thus, using the weighted median solver, we can solve for γ and get a new equation, e.g., equation j. Then we can update (α, β) and $\mathbf{s}^{(d)}$ accordingly, drop the old equation k, and keep the new equation j. We iterate on this process until the inner loop keeps tracking the same equation. The final results of the inner loop are passed out to update the model, residual and the gradient.

The value of expending more effort to find the best step direction will be supported by the real geophysical applications, because the most computationally expensive part of these iterative methods is applying the forward and adjoint operators (steps starred in Table 1). By doing the sophisticate plane search, we hope to decrease the number of outer-loop iterations required for convergence.

However, conjugate direction L1 regression theory is not perfect for a practical problem. The problem of a flat bottom in $L1$ minimization will cause trouble in geophysical practice. Sometimes even where the bottom is not exactly as flat as the median of an even number of points, the slope of the gradient can be so small that we can never reach convergence in a finite number of iterations.

SYNTHETIC AND FIELD DATA EXAMPLE

Figure 1 shows the input synthetic RMS velocities with and without random noise, and the true blocky interval velocity we try to invert for. For the synthetic problem, we experimented on solvers with and without regularization to learn the nature of the solver itself and the nature of its regularization.

Figure 2 shows the inversion results when clean data (free of noise) are fed into different simple solvers without any regularization. The result of $L2$ regression is comparable to the IRLS and hybrid norm. However, the simple conjugate direction $L1$ solver failed to give a satisfactory result (Figure 2(c)). As we have discussed in the previous section, this might be due to the flat bottom caused by the data configuration.

Figure 3 shows the inversion results when clean data are fed into different regularized solvers. As expected, the smoothing effect of $L2$ regularization on the derivative of model produces the round corners at the turning point. In contrast, IRLS and hybrid solvers give perfect exact solutions, which benefit from their $L1$ nature in regularization. We do not fully understand the behavior of conjugate-direction $L1$ solver, but the change in the result can be explained by the change of the data configuration when the regularization term is added.

Figure 4 shows the inversion results when noisy data are fed into different simple solvers without any regularization. When creating the synthetic noisy RMS velocity data, uniform distributed random noise is added. However, $L2$ norm assumes Gaussian noise, and $L1$ minimization is derived under the assumption of exponential distribution. Therefore, the simple $L2$ solver, the IRLS solver or hybrid solver all fail to recognize the noise and attenuate it. Surprisingly, conjugate-direction $L1$ solver successfully eliminates the high-frequency noise and keeps the low-frequency trend of the interval velocity function. It shows great potential for finding the exact solution when the problem is slightly more complicated.

Figure 5 shows the inversion results when noisy data are fed into different regularized solvers. By adding this regularization term to further constrain the problem, we expect better results out of each solver. Comparing the results in Figure 5, IRLS result has the most blocky transition between layers and is almost flat within the layers. However, the big jump at shallower depths is apparently due to its tolerance of the large residuals. The hybrid solver gives result comparable to the IRLS, but it oscillates at deeper depths. The results from $L2$ and conjugate direction $L1$ solver are similar, but the smaller steps in the result of conjugate direction $L1$ (Figure 5(c)) are promising.

Figure 6 shows the 1-D field RMS velocity from the velocity scan. This field data has 1,000 sample points, and the number of blocks in the model space is unknown. In real life, we can never fully constrain a inversion problem without regularization: that is when Dix inversion becomes unstable. Therefore, only regularized solvers are tested.

Figure 7 shows the inversion results when field data in Figure 6 are fed into different regularized solvers. The results from the IRLS and the conjugate direction $L1$ solver have more blocky nature than the other two. The result from IRLS is more flat within layers, which is a nice property in well-log matching and many other geophysical applications. The hybrid and $L2$ solvers give comparable results, although we have chosen a very small σ for hybrid norm to force it towards the $L1$ norm.

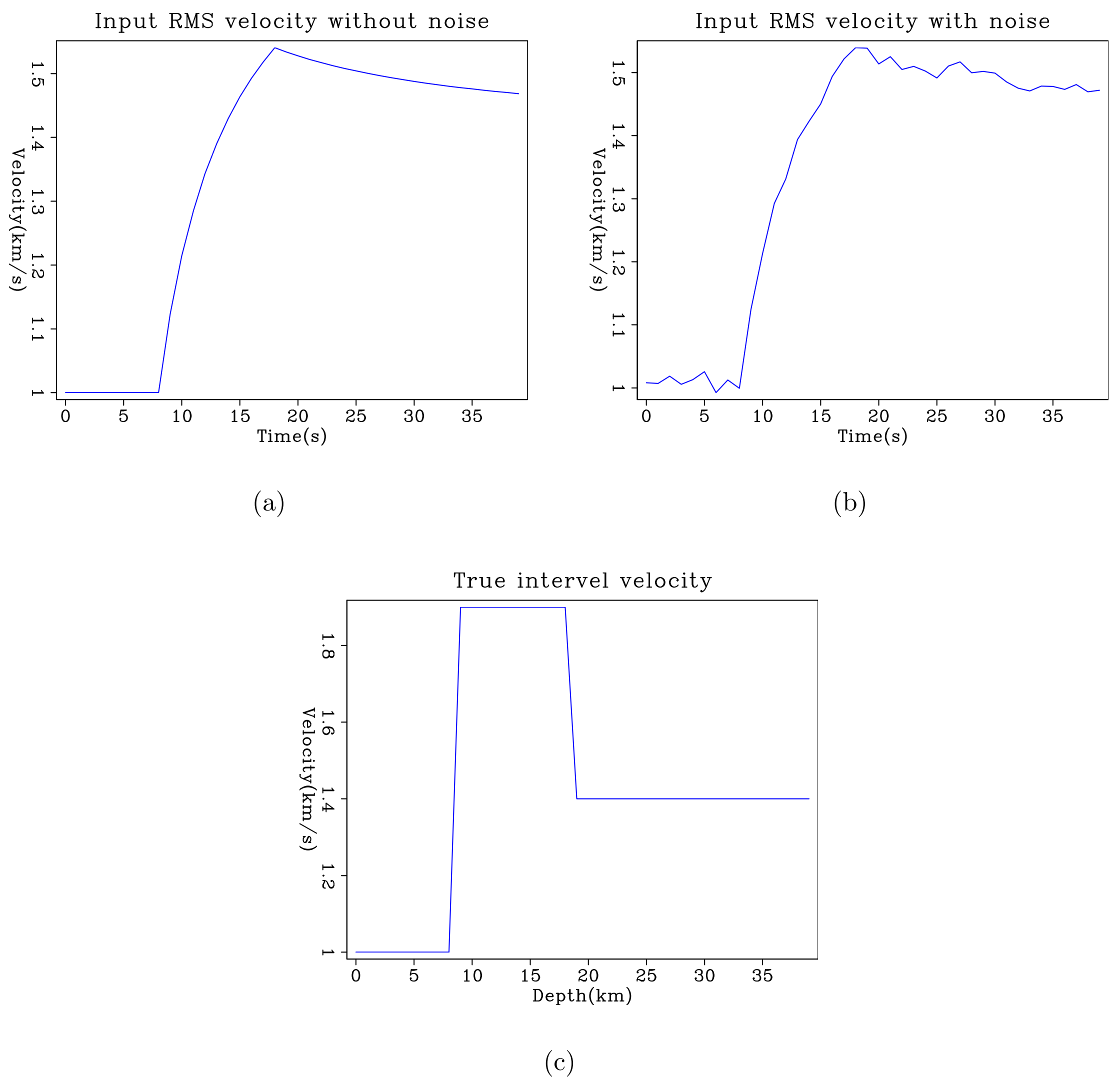


Figure 1: Input synthetic RMS velocity and true interval velocity. The two plots on the top row are the input RMS velocities **(a)** without noise and **(b)** with random noise, respectively. The plot on the bottom is **(c)** the true interval velocity which is true model in the estimation problem. [**ER**]

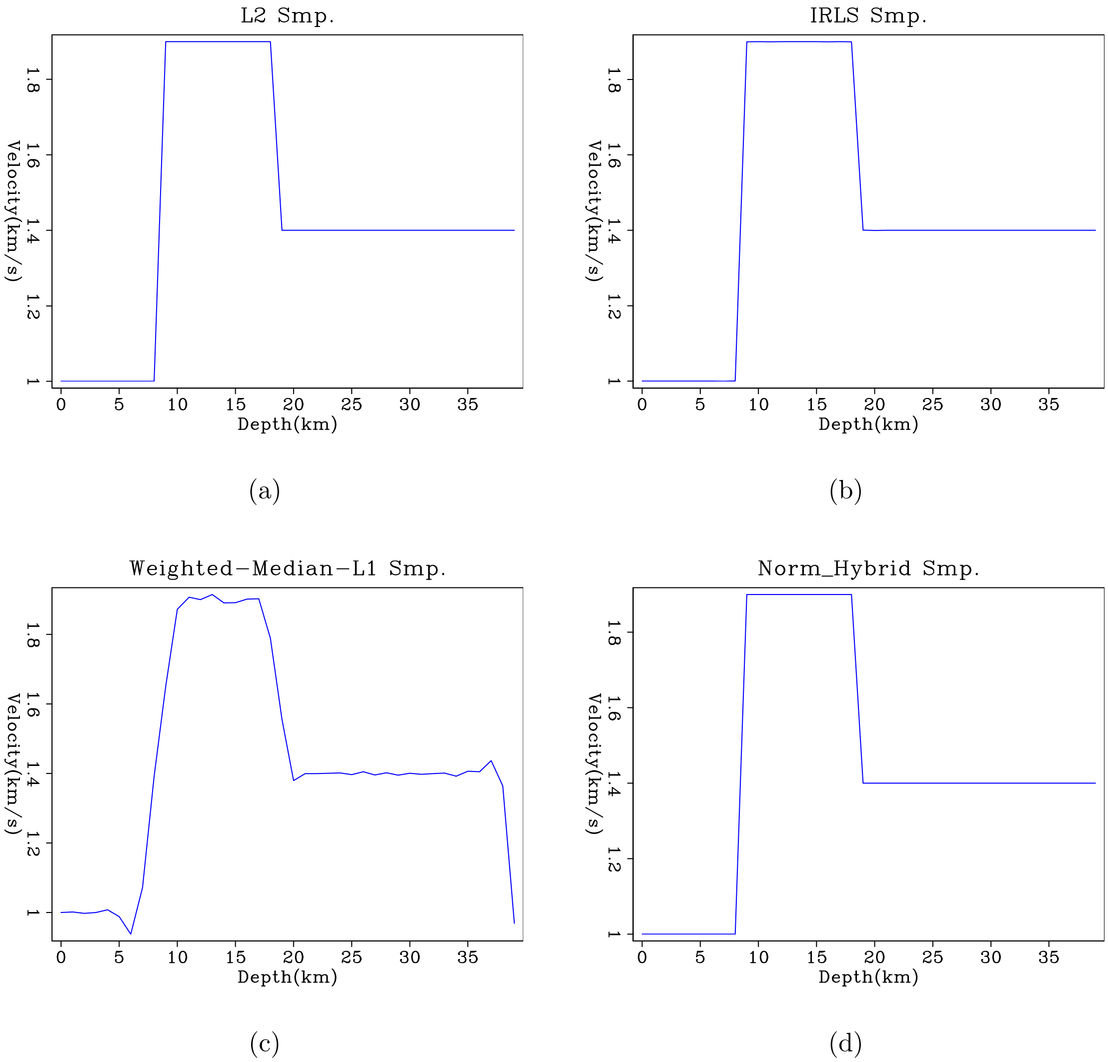


Figure 2: Inversion results of simple **(a)** $L2$ solver; **(b)** IRLS solver; **(c)** conjugate direction $L1$ solver and **(d)** Hybrid solver when clean data are fed in. All the regressions are without regularization. [**ER**]

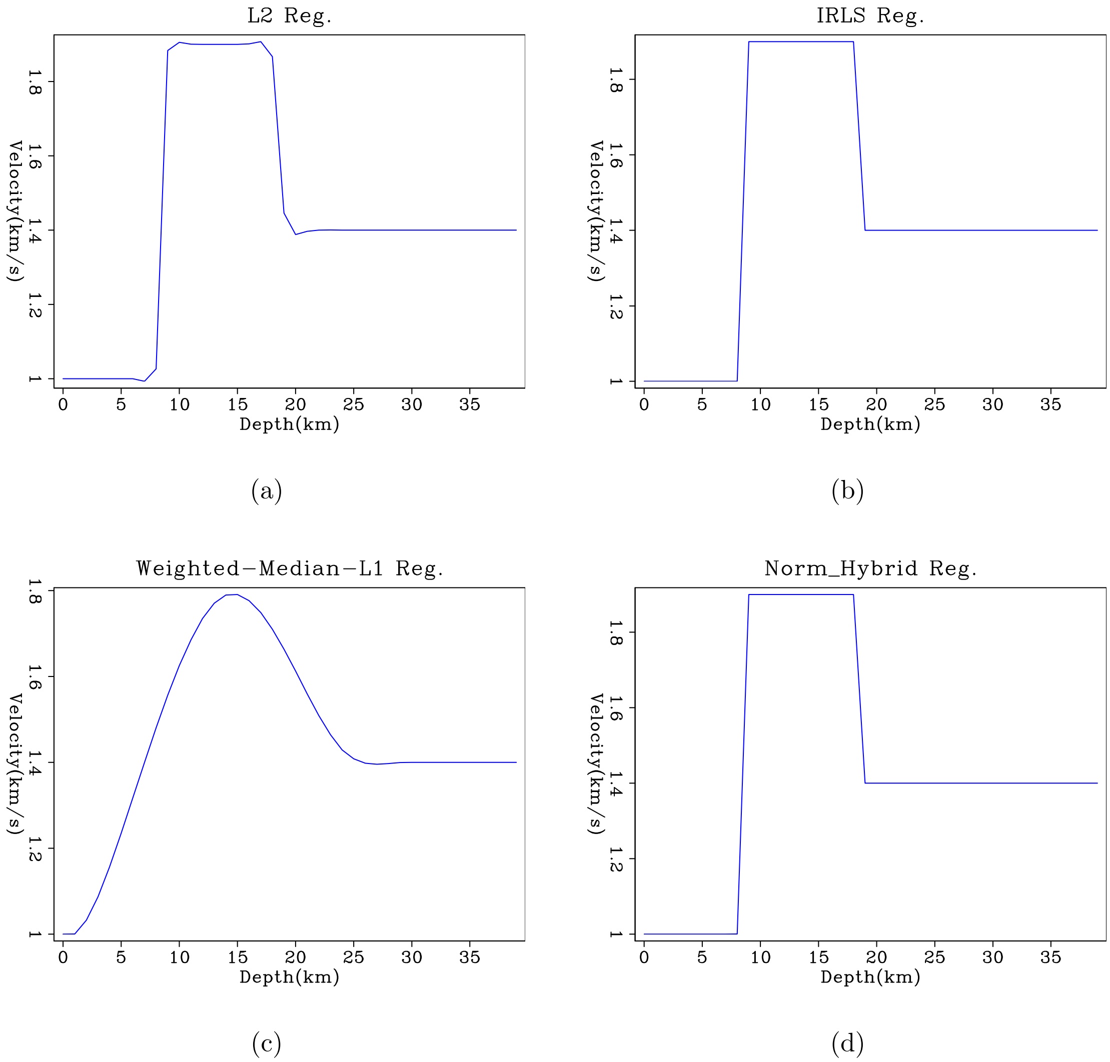


Figure 3: Inversion results of **(a)** $L2$ solver; **(b)** IRLS solver; **(c)** conjugate direction $L1$ solver and **(d)** Hybrid solver when clean data are fed in. All the regressions are regularized.[**ER**]

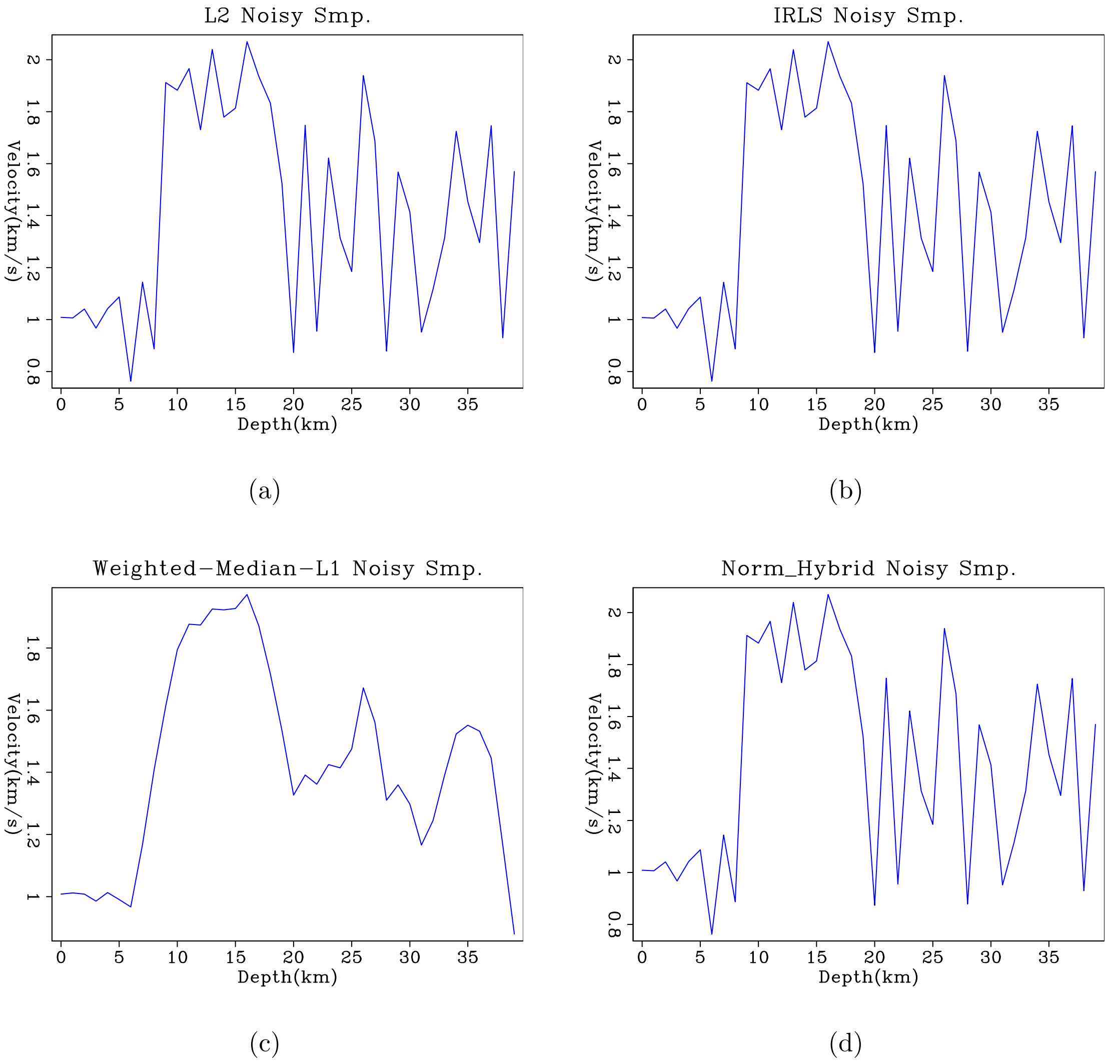
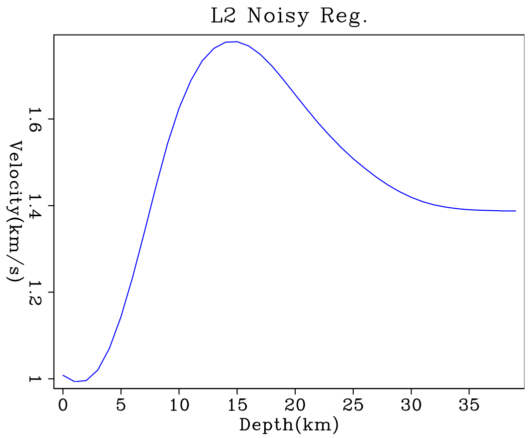
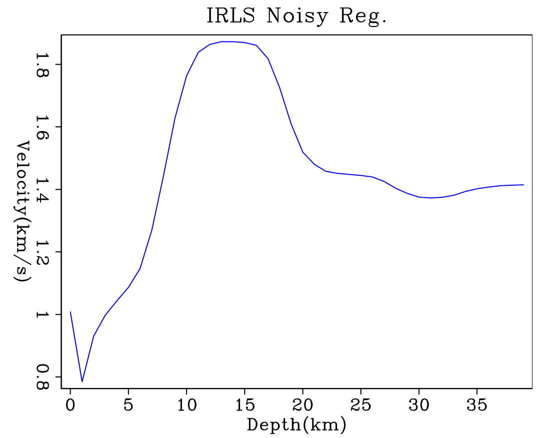


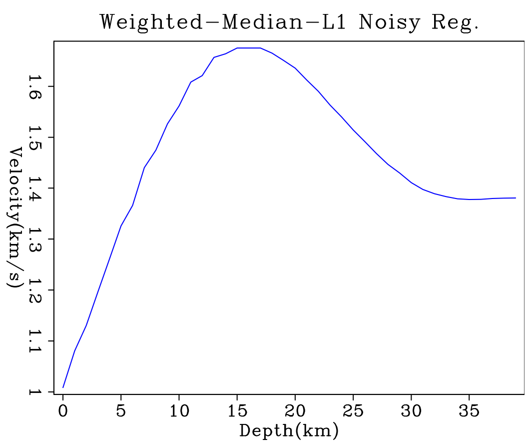
Figure 4: Inversion results of simple **(a)** $L2$ solver; **(b)** IRLS solver; **(c)** conjugate direction $L1$ solver and **(d)** Hybrid solver when noisy data are fed in. All the regressions are without regularization.[**ER**]



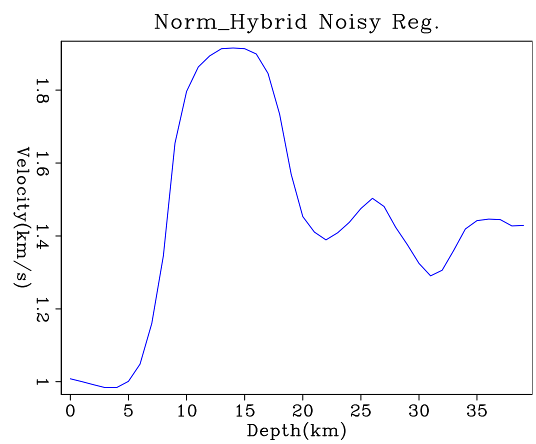
(a)



(b)



(c)



(d)

Figure 5: Inversion results of **(a)** $L2$ solver; **(b)** IRLS solver; **(c)** conjugate direction $L1$ solver and **(d)** Hybrid solver when noisy data are fed in. All the regressions are regularized. [**ER**]

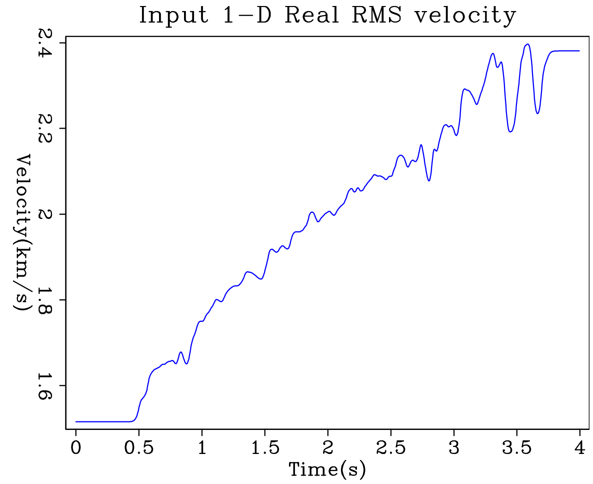


Figure 6: Input 1-D field RMS velocity data from velocity scan. [**ER**]

CONCLUSIONS AND DISCUSSIONS

We explore three methods to constrain Dix inversion in an $L1$ nature: an improved IRLS method, a hybrid $L1/L2$ method and conjugate direction $L1$ method. The IRLS and hybrid methods are implemented in a non-linear least-squares scheme by adding a diagonal weighting function. Conjugate direction $L1$ method is realized by a weighted median solver.

The IRLS method is improved by the physical explanation of the cutoff number σ, allowing this numerical parameter to be determined automatically. The hybrid method has a novel plane search scheme based on Taylor's series at each residual. Conjugate direction $L1$ method has an iterative plane search scheme using a weighted median solver. Both of hybrid and the conjugate direction $L1$ are designed to reduce major computational cost by expending more effort in finding a better next step.

In the numerical experiment, we find that the conjugate direction $L1$ method decreases the iteration number for the outer loop significantly. Hence, the value of spending more to find a better next step is proved. The same concept can be applied to the hybrid method as well. We can expect better inversion results and faster convergence by adding iterations to plane search, which has not been demonstrated before.

In the current study of the conjugate direction $L1$ method, we keep only two equations exactly satisfied in the whole system when searching the plane. In future research, we can add as many equations as needed by Gram Schmidt process. Hopefully, this process can lead us to an even better next step.

ACKNOWLEDGMENTS

I would like to thank Jon Claerbout for the idea of expending the effort in the inner loop to reduce the necessary iteration numbers of the outer loop. I would also like to

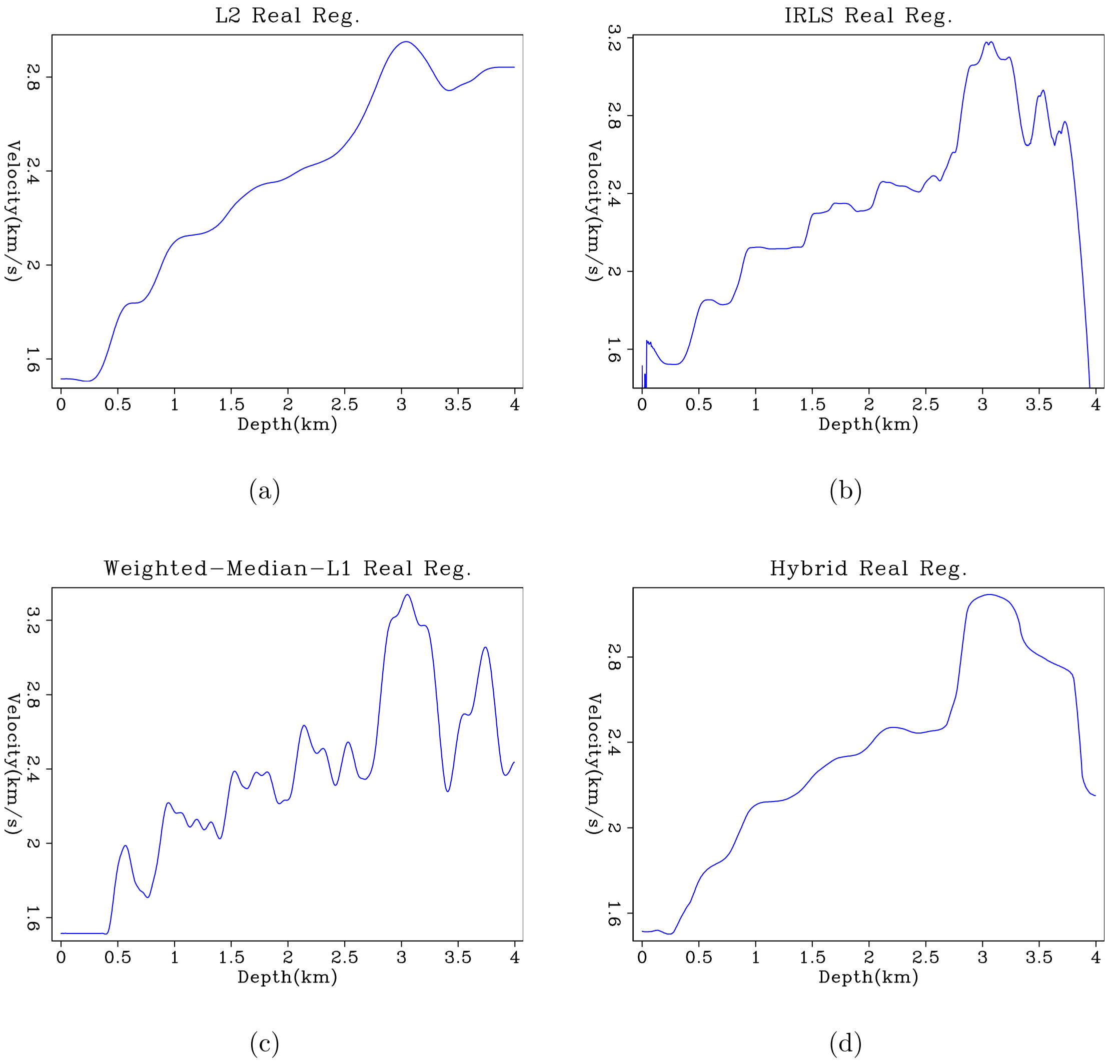


Figure 7: Inversion results of **(a)** $L2$ solver; **(b)** IRLS solver; **(c)** conjugate direction $L1$ solver and **(d)** Hybrid solver when 1-D field RMS velocity data are fed in. All the regressions are regularized. **[ER]**

acknowledge Robert Clapp for the code of conjugate direction $L1$ solver.

REFERENCES

Bube, K. P. and R. T. Langan, 1997, Hybrid l1/l2 minimization with applications to tomography: Geophysics, **62**, 1183–1195.

Claerbout, J. F., 2008, Image estimation by example.

———, 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.

Clapp, R. G., 2001, Geologically constrained migration velocity analysis: PhD thesis, Stanford University.

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

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

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

Harlan, W. S., 1999, Constrained dix inversion.

Koren, Z. and I. Ravve, 2006, Constrained dix inversion: Geophysics, **71**.

Maysami, M. and N. Mussa, 2009, Generalized-norm conjugate direction solver: SEP-Report, **139**, 11–22.

Nichols, D., 1994, Velocity-stack inversion using l^p norms: SEP-Report, **94**, 1–16.

Tang, Y., 2006, Least-squares migration of incomplete data sets with regularization in the subsurface-offset domain: SEP-Report, **125**.

Valenciano, A. A., M. Brown, M. D. Sacchi, and A. Guitton, 2003, Interval velocity estimation using edge-preserving regularization: SEP-Report, **114**, 136–150.