Geophysical application of hyperbolic and hybrid L1/L2 optimization

Yinbin Ma, Musa Maharramov and Biondo Biondi

ABSTRACT

Hyperbolic penalty function method and hybrid L1/L2 methods, often generate better results than conventional least-squares solutions for inverse problem in geophysics. We apply a few hyperbolic and hybrid L1/L2 methods to 2-D Kirchhoff migration inversion and target-oriented linearized waveform inversion. The results demonstrate that we can recover a sparse/blocky model using hyperbolic and hybrid L1/L2 methods with acceptable computational cost.

INTRODUCTION

L1-norm optimization often yields more robost results compared with conventional least-squares optimization (Claerbout and Muir, 1973; Darche, 1989; Nichols, 1994; Guitton, 2005). Different algorithms for hyperbolic penalty function (HPF) optimization or L1/L2 hybrid optimization problems exist. All such algorithms require fine tuning of extra computational parameters. The computational cost and quality of results depends on the quality of parameter tuning. One concern is the physical interpretation of these parameters. Another concern is the cost of the solvers, as the extra computational cost associated with these techniques may not be justified by the improvement in the results. In this paper, we evaluate several hyperbolic and hybrid L1/L2 solvers.

We use several different solvers in our paper: least-squares with conjugate gradient (CGLS), iterative reweighted least squares (IRLS), Split-Bregman (Goldstein and Osher, 2009), Hyperbolic penalty function (HPF) with conjugate directions (Claerbout, 2009; Li et al., 2010; Zhang and Claerbout, 2010). We use IRLS and Split-Bregman methods when the objective function contain both L2-norm and L1-norm, and we call them L1/L2 hybrid methods. To characterize the performance, we first run a few iterations of least-squares methods, obtain the data residual and set it as threshold. Then we run L1 based techniques (HPF, IRLS and Split-Bregman) until the data residuals are below the threshold. The quality of inversion results and computational costs are then analyzed.

We test the solvers on several geophysical examples. First, we apply them to 2-D regularized Kirchhoff migration-inversion. The results show that hyperbolic and hybrid L1/L2 methods can recover the original model within acceptable amount of

computation cost. Next, we apply the solvers to target-oriented linearized waveform inversion. We solve the problem described in Zhang and Claerbout (2010). Along with replicating the hyperbolic penalty function method result, we implement several other hybrid methods. While it is shown that hyperbolic and L1/L2 hybrid methods deliver significant improvement, we can not get pleasing results. We add a new L_1-based regularization term and produce better results.

METHOD

The general form of objective function we try to minimize has the following form:

$$\mathbf{J}(\mathbf{m}) = \frac{1}{2}\left\|\mathbf{Fm}-\mathbf{d}\right\|_2^2 + \frac{\varepsilon}{2}\left\|\mathbf{Am}\right\|_{norm}^{\alpha}. \tag{1}$$

The first term in the objective function is the data fitting term, and the second term is model regularization. Operator $\mathbf{A}$ can be an identity matrix if the model is sparse, and can be a total variation operator if the model is blocky. The second term is necessary in geophysical inverse problems, because we may have insufficient data. Even if we have sufficient data, we still want to keep the regularization term, because real data contains high amplitute burst noise, and overfitting the data may lead to inaccurate results in this case.

We use CGLS, IRLS, the Split-Bregman and Hyperbolic penalty function (HPF) with conjugate direction methods to solve the optimization problem with different norm (or penalty function).

We first use conventional least-squares solver CGLS. The regularization term is in L_2-norm, which means the objective takes the form,

$$\mathbf{J}(\mathbf{m}) = \frac{1}{2}\left\|\mathbf{Fm}-\mathbf{d}\right\|_2^2 + \frac{\varepsilon}{2}\left\|\mathbf{Am}\right\|_2^2. \tag{2}$$

Next, we implement a solver using hyperbolic penalty function (HPF), developed by Li et al. (2010) and Zhang and Claerbout (2010) . The solver solves

$$\mathbf{J}(\mathbf{m}) = \frac{1}{2}\left\|\mathbf{Fm}-\mathbf{d}\right\|_2^2 + \frac{\varepsilon}{2}\left\|\mathbf{Am}\right\|_{HB}, \tag{3}$$

where the second term is replaced by a nonlinear HPF $\|r\|_{HB} = \sqrt{r^2+R^2}-R$, where R is the threshold parameter. The penalty function approaches L1-norm as $r \gg R$, and approaches L2-norm as $r \ll R$. While using the HPF, the problem becomes nonlinear. Claerbout (2009) generalized the idea of the Conjugate Direction method for HPF optimization. At each step, we search the new direction in the plane spanned by the current gradient and the previous step. In this case, we need to solve the new direction by iterative methods. In the case of linear optimization, we can obtain the

new direction directly. The detail of the method has been explained previously (Li et al., 2010; Zhang and Claerbout, 2010).

We also explore the behavior of solvers when the model styling term is L_1 norm as in equation (4). We use two different solvers, IRLS and the Split-Bregman method. Both methods are designed to solve the minimization problem of equation (4),

$$\mathbf{J}(\mathbf{m}) = \frac{1}{2}\left\|\mathbf{Fm}-\mathbf{d}\right\|_2^2 + \varepsilon\left\|\mathbf{Am}\right\|_1. \tag{4}$$

IRLS is an iterative method and at each step we solve a weighted least-squares. For the second term, the residual is now expressed as $\mathbf{r} = \mathbf{W_r Am}$. Where $diag(\mathbf{W_r})_i = |\mathbf{r_i}|^{-0.5}$. It can be found that $\|\mathbf{r}\|_2 = \|\mathbf{Am}\|_1$. Numerically, we minimize the following objective function using CGLS:

$$\mathbf{J}(\mathbf{m}) = \frac{1}{2}\left\|\mathbf{Fm}-\mathbf{d}\right\|_2^2 + \varepsilon\left\|\mathbf{W_r Am}\right\|_2^2. \tag{5}$$

After a few number of CG steps, we recalculate the weight function $\mathbf{W_r}$, and continued minimizing the new objective funcion. It is easy to see that, upon convergence, the $\mathbf{m}$ we obtained from equation (5) approaches the solution to equation (4).

We also implement an algorithm named the Split-Bregman method. The Split-Bregman method tries to minimize equation (4) by minimizing the following objective function:

$$\mathbf{J}(\mathbf{m},\mathbf{z}) = \frac{1}{2}\left\|\mathbf{Fm}-\mathbf{d}\right\|_2^2 + \frac{\rho}{2}\left\|\mathbf{Am}-\mathbf{z}\right\|_2^2 + \varepsilon\left\|\mathbf{z}\right\|_1. \tag{6}$$

With traditional optimization methods, we need to solve the problem with increasing ρ, and as $\rho \to \infty$ the solution converges to the original problem equation (4). Goldstein and Osher (2009) have designed the Split-Bregman method that could solve equation (6) with fixed ρ. They proved upon convergence we obtain the solution to equation (4). The Split-Bregman method is similar to alternating direction method of multipliers (ADMM) (Boyd et al., 2011).

To minimize equation (6) numerically, they introduce an additional parameter $\mathbf{b}$, and solve the following steps iteratively:

$$(\mathbf{m},\mathbf{z}) = \min_{\mathbf{m},\mathbf{z}}\{\frac{1}{2}\left\|\mathbf{Fm}-\mathbf{d}\right\|_2^2 + \lambda\left\|\mathbf{z}\right\|_1 + \frac{\rho}{2}\left\|\mathbf{z}-\mathbf{Am}-\mathbf{b}\right\|_2^2\}, \tag{7}$$

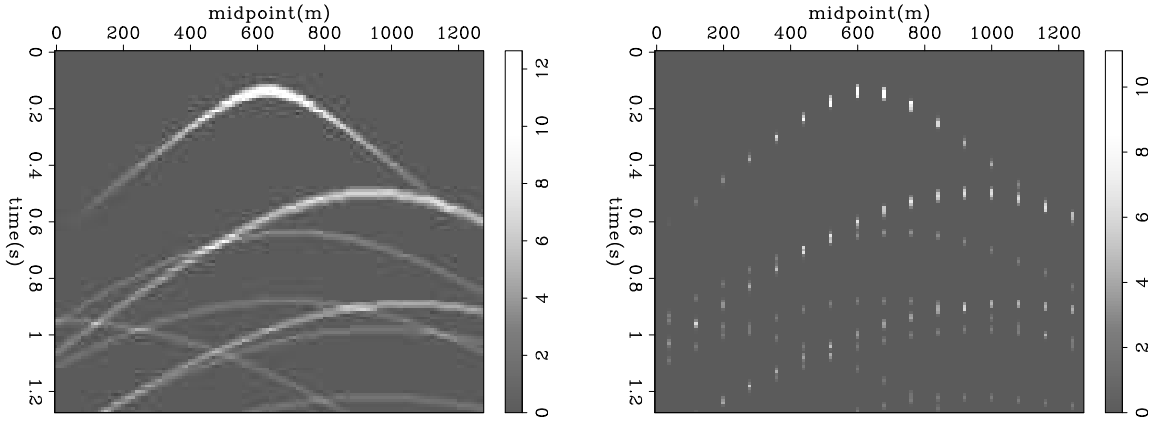
$$\mathbf{b} \leftarrow \mathbf{b} + (\mathbf{Am}-\mathbf{z}). \tag{8}$$

Where equation (7) are solved iteratively via:

$$\textbf{Step } 1 : \mathbf{m} = \min_{\mathbf{m}} \{ \frac{1}{2} \left\| \mathbf{Fm} - \mathbf{d} \right\|_2^2 + \frac{\rho}{2} \left\| \mathbf{z} - \mathbf{Am} - \mathbf{b} \right\|_2^2 \}, \tag{9}$$

$$\textbf{Step } 2 : \mathbf{z} = \min_{\mathbf{m}} \{ \lambda \left\| \mathbf{z} \right\|_1 + \frac{\rho}{2} \left\| \mathbf{z} - \mathbf{Am} - \mathbf{b} \right\|_2^2 \}. \tag{10}$$

When the algorithm converges, from equation (8) we will have $\mathbf{Am} - \mathbf{z} \rightarrow \mathbf{0}$. The solution for the objective function in equation (6) converges to the solution for equation(4).



(a) (b)

Figure 1: Kirchhoff migration data. (a): with sufficient sampling; (b): with insufficient sampling. [**ER**]

KIRCHHOFF MIGRATION INVERSION RESULTS

Kirchhoff migration involves the summation or spreading along travel time surfaces. It is well known that when we have insufficient data, the Kirchhoff migration does not yield good results because of data aliasing. To overcome this problem, we add geophysical constraints to our objective function and reduce the effect of aliasing. For simplicity, we use the Kirchhoff time migration as an example in this section, to test different solvers. The HPF method has been tested on Kirchhoff migration in the previous work (Li et al., 2010). To be consistent, we use the same operator and same model. We formulate our problem as

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

$$\mathbf{Am} \approx \mathbf{0}, \tag{12}$$

to be consistent with the notation in the previous section. $\mathbf{F} = \mathbf{SK}$, where $\mathbf{K}$ is the Kirchhoff modeling operator, and $\mathbf{S}$ is the subsampling operator. $\mathbf{m}$ is reflectivity, and $\mathbf{d}$ is data recorded at the surface. $\mathbf{A}$ is regularization operator, and in the current example, $\mathbf{A} = \mathbf{I}$ as the model is reflectivity and it is sparse.

We choose the model space sampling to be 128×128, and subsampling data space to be 128×16, exactly the same as previous work (Li et al., 2010). As seen in Figure 1, the foward modeling operator maps the spikes in the model space to a series of hyperbolas in the data space if we have enough sampling points (Figure 1(a)), and it is aliased in our setting because we have insufficient data sampling (Figure 1(b)).

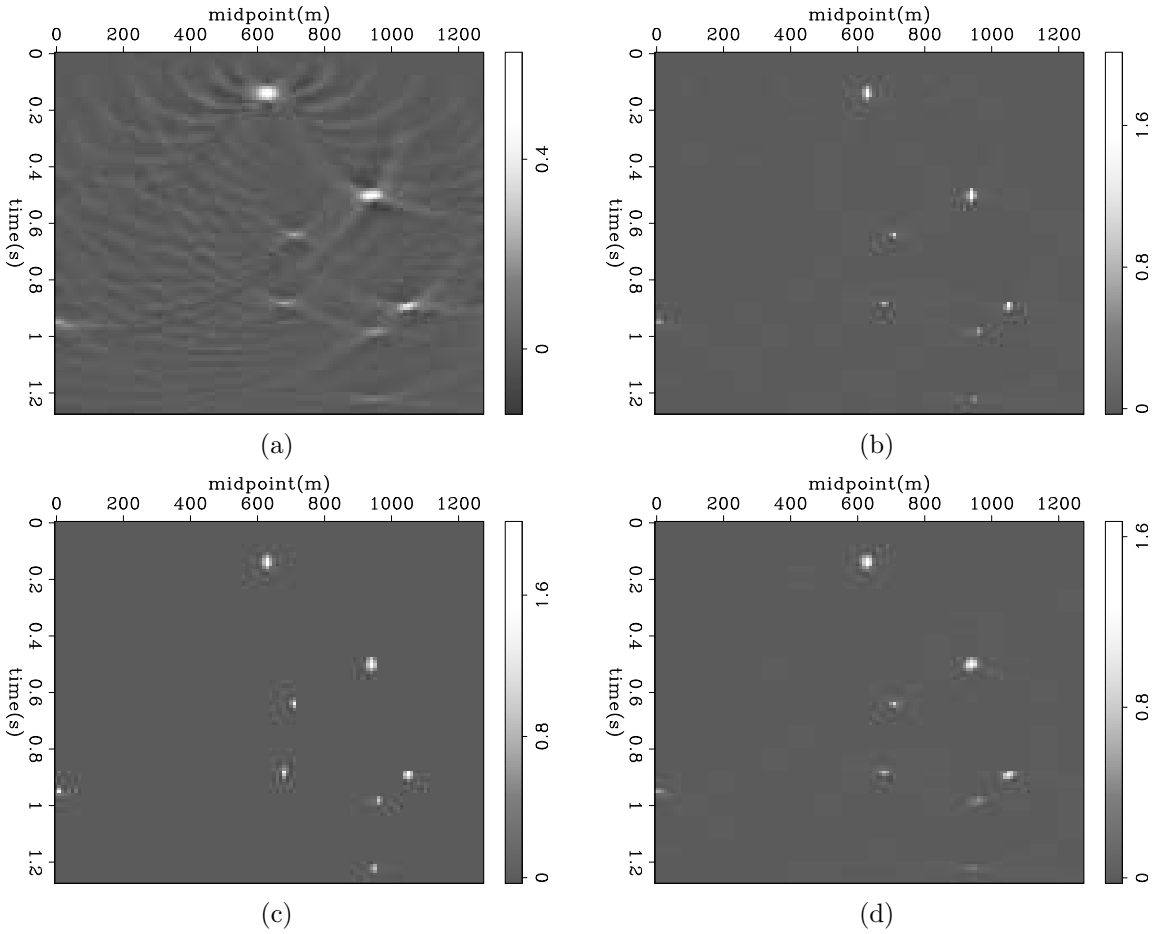


Figure 2: Comparison of the inversion results with different solvers: (a) CGLS, (b) HPF , (c) IRLS and (d) the Split-Bregman method. [**ER**]

We solve the optimization problem with the four different methods mentioned in the previous section. First, we use the CGLS method for equation (2) with parameter $\varepsilon = 0.1$. We run the algorithm for 50 CG steps, and reduced the data residual from 60.93 to 0.33 ($\sim 0.54\%$). Next, we run HPF with equation (3), and successful reduce the data residual to the same level using 408 CG steps. Then, we run IRLS and Split-Bregman method for equation (4). We are able to match the data residual using 330 CG steps for IRLS and 417 CG steps for Split-Bregman. We can see the inversion results from different schemes in Figure 2, and the reconstructed data (Figure 3). The quality of inversion is greatly improved using hyperbolic and hybrid solvers. The additional computational cost is acceptable considering the improvement of inversion results.

We want to emphasis that we do not find optimized parameters for each solver. Thus, we cannot make a conclusion about comparison between different hyperbolic or hybrid solvers based on our numerical results.

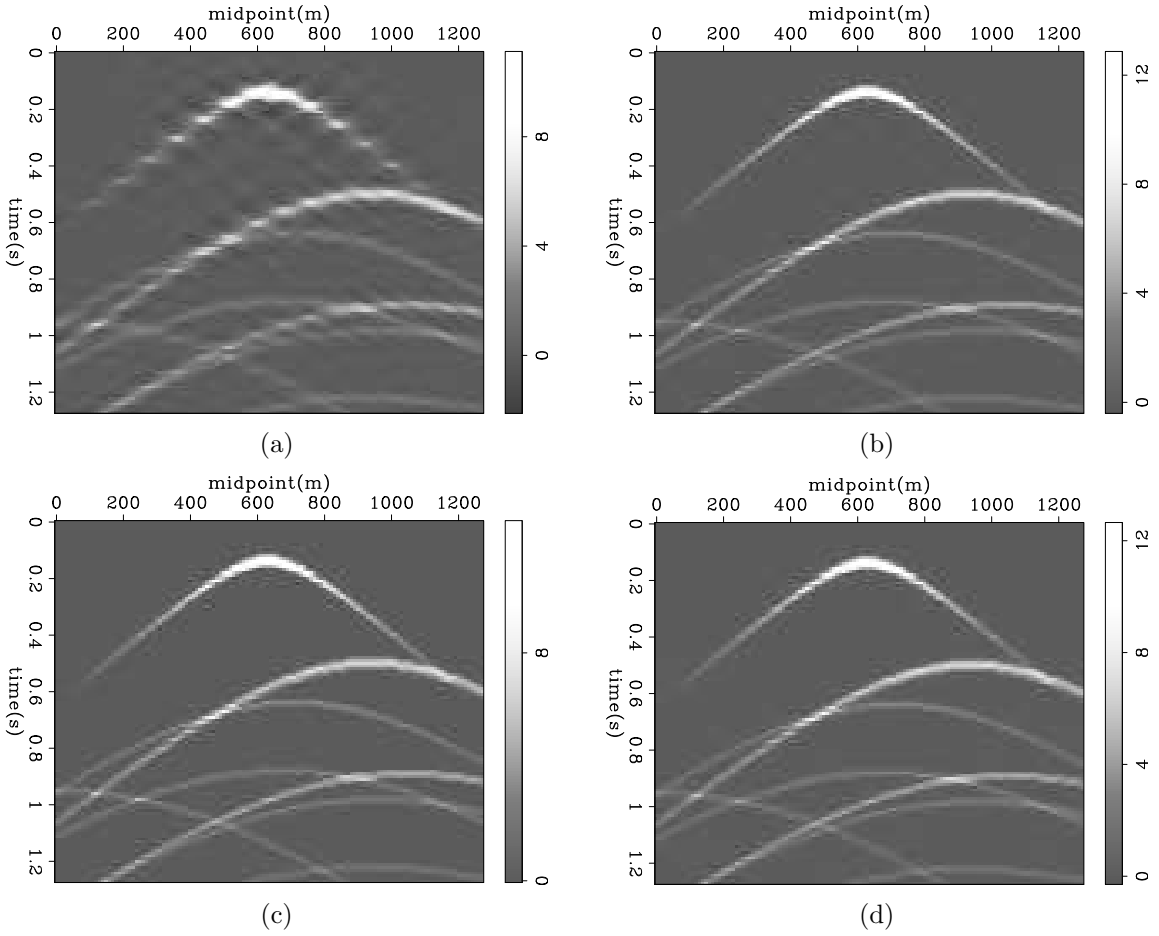


Figure 3: Comparison of modeled data for different schemes: (a) CGLS, (b) HPF, (c) IRLS and (d) the Split-Bregman method. [**ER**]

LINEARIZED WAVEFORM INVERSION RESULTS, PART 1

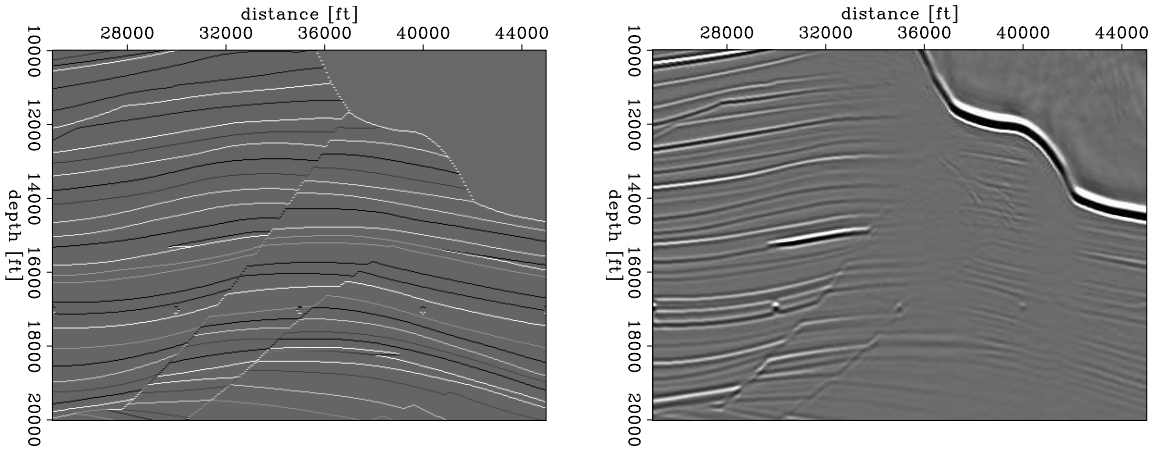
The target-oriented linearized waveform inversion has been explored in several previous works (M. Clapp and Biondi, 2005; Valenciano, 2006; Tang, 2008). The idea is from full waveform inversion(FWI) problem,

$$\mathbf{Lm} \approx \mathbf{d_{data}}. \tag{13}$$

We multiply $\mathbf{L}'$ on both side and obtain the normal equation as follows:

$$\mathbf{Hm} \approx \mathbf{m_{mig}} \tag{14}$$

where $\mathbf{H} = \mathbf{L'L}$ is the Hessian operator, and $\mathbf{m_{mig}} = \mathbf{L'd_{data}}$ is called migrated data.



(a) (b)

Figure 4: (a) Reflectivity model; (b) input migrated image. [**ER**]

The Hessian operator map vectors from the same model space, and we can view it as a convolution operator. It has been shown (Tang, 2008) that the Hessian operator is very sparse and we can afford to save the matrix. The advantage of equation (10) is that we can compute target-oriented inversion. However, under the salt body because of the complexity of wave propogation, the Hessian may be inaccurate or rank-deficient. It is challenging to obtain sparse and clear inversion results in the existence of salt body. As we can see in Figure 4(b), the migrated data has gaps that break the continuity of reflectors and the signal-to-noise ratio under the salt is small.

The previous work (Zhang and Claerbout, 2010) tried to solve the inversion problem with HPF as regularization. The results showd improvement comparing with least square method. In this part, we use $||\mathbf{m}||_{norm}$ as regularization, to be consistent with the previous results.

Similar to previous section, we solve the optimization problem with four different methods. The forward modeling operator is now Hessian $\mathbf{F} = \mathbf{H}$. And we choose the regularization operator to be identity $\mathbf{A} = \mathbf{I}$. The input data is $\mathbf{d} = \mathbf{m_{mig}}$.

First, we use the CGLS method as the benchmark and run 50 CG steps. We reduce the data residual from 3.99 to 0.190($\sim 3.5\%$). Notice that for this problem, focusing on reducing the data residual is not a good choice. The data contains large error due to insufficient illumination under the salt, and the operator is not accurate. We search reasonable parameters for each method and ran less than 300 CG steps(3000 for HPF, 200 for IRLS, and 200 for Split-Bregman). The data residual is 0.199 for the HPF method, 0.311 for IRLS, and 0.254 for the Split-Bregman method.

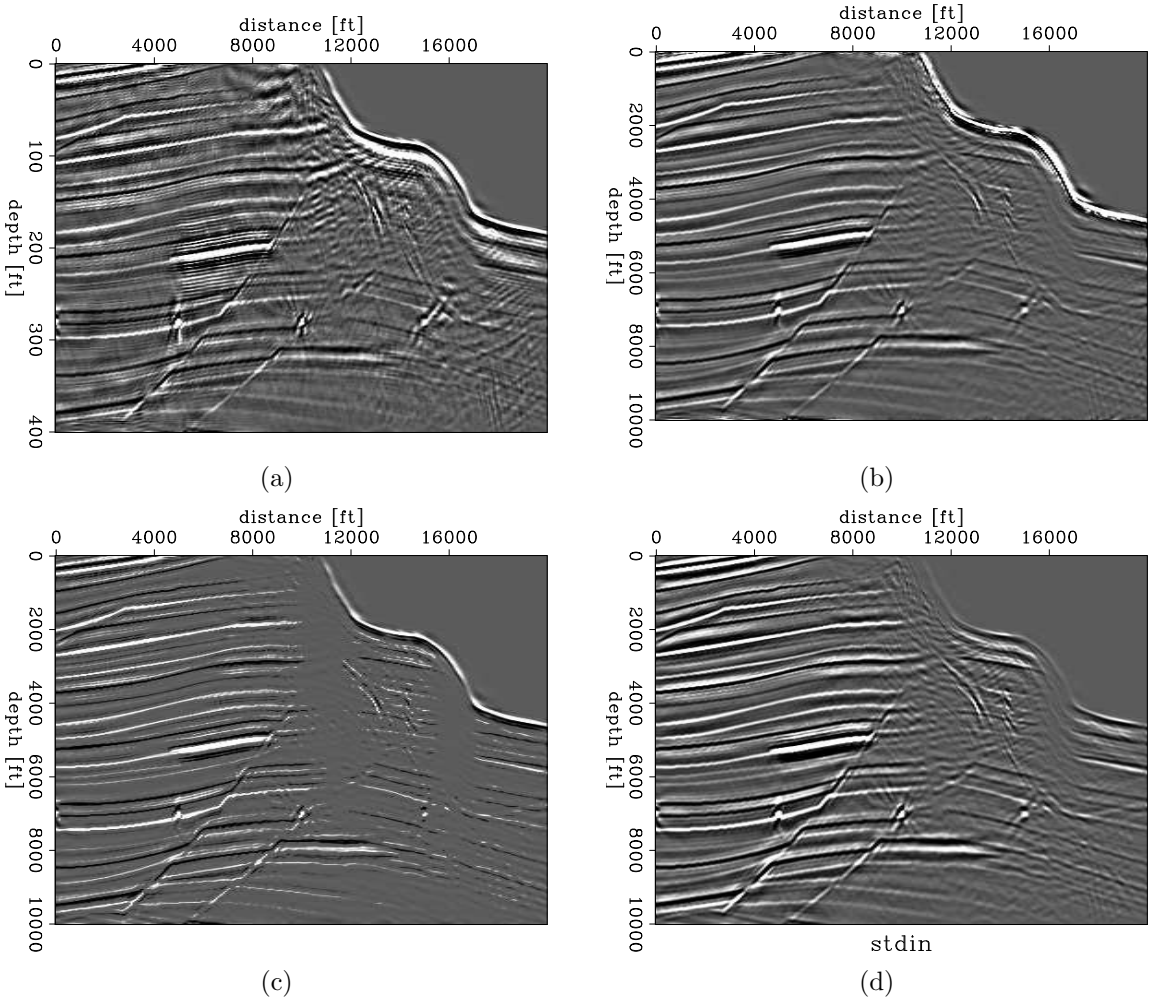


Figure 5: Comparison of linearized waveform inversion inversion results: (a) CGLS, (b) HPF, (c) IRLS and (d) the Split-Bregman method. [**ER**]

All the hyperbolic and hybrid solvers improve the quality of the inversion results. However, none of them could yield satisfying results under the salt body. Figure 6 shows the data residual, suggesting the inversion results are not ideal.

LINEARIZED WAVEFORM INVERSION RESULTS PART 2

We obtain consistent results comparing with previous work. All of the solvers show similar behavior. We believe the poor quality of inversion results are not caused by the inappropriate use of the solvers. We need to check our objective function and make improvement in our model to obtain better results.

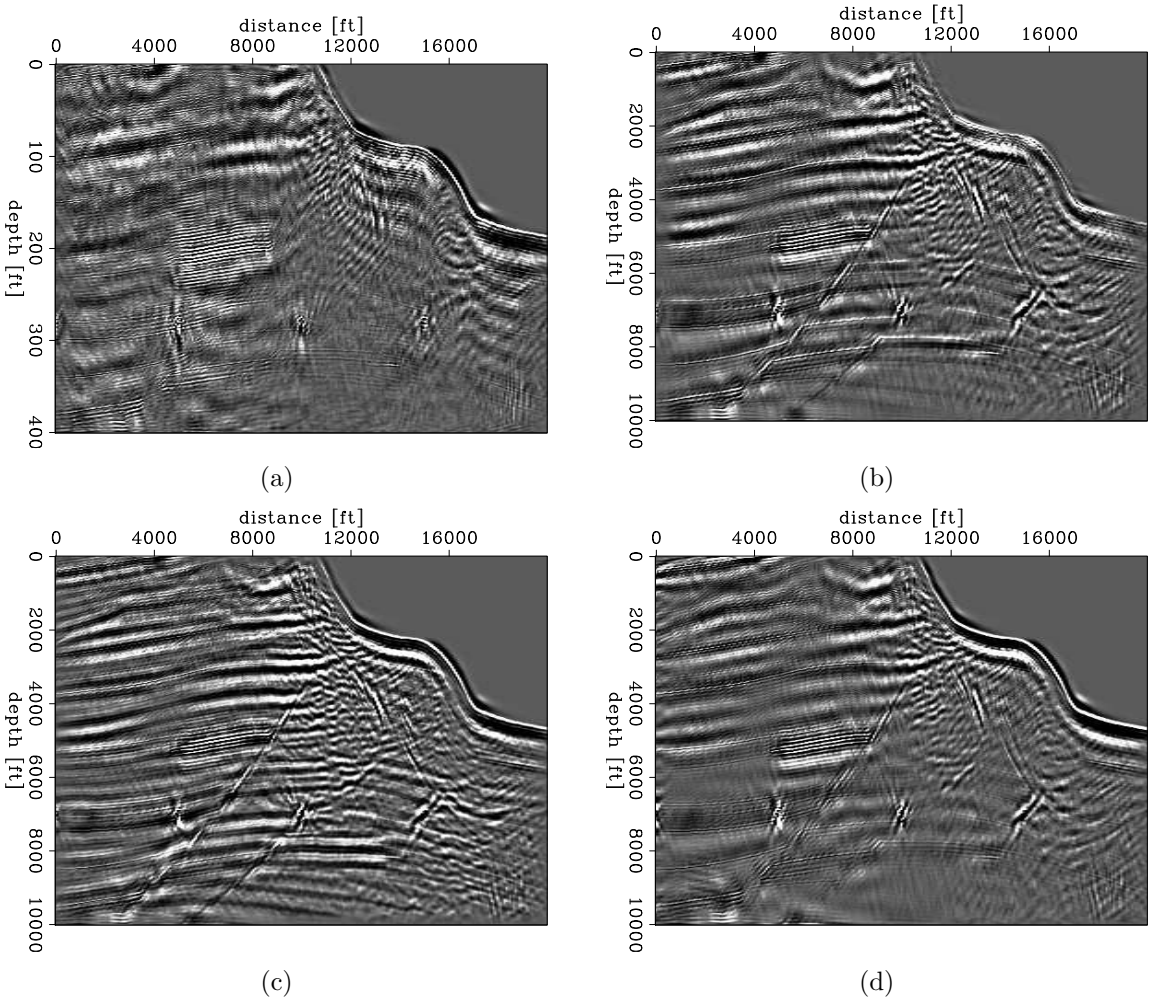


Figure 6: Comparison of linearized waveform inversion data residual: (a) CGLS, (b) HPF, (c) IRLS and (d) the Split-Bregman method. All figures are clipped at the same level. [**ER**]

We use exactly the same regularization for Kirchhoff migration and linearized

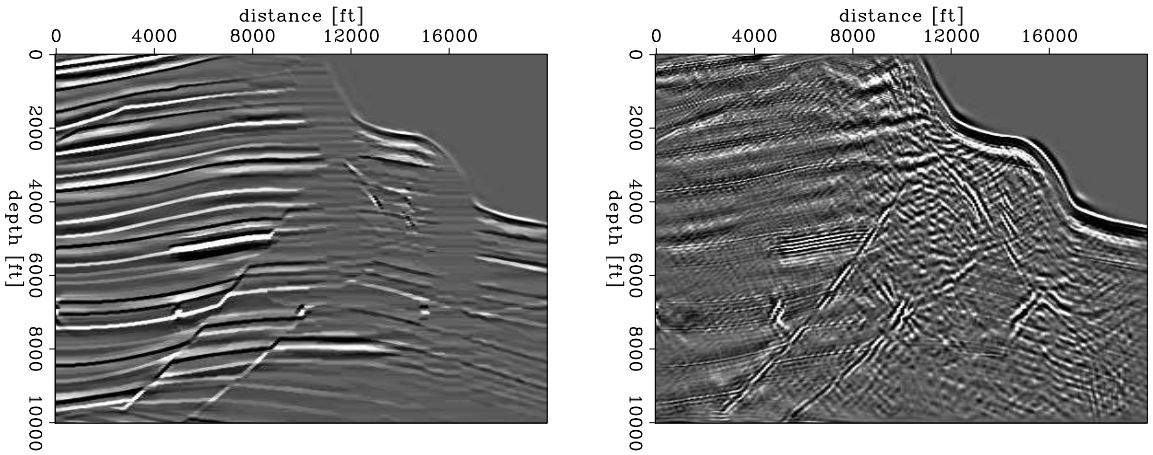
waveform inverion in the previous sections. It is obvious that in the Kirchhoff migration case, all we know is that the reflectors are a series of spikes and they are sparse. We used the correct regularization and obtained high-quality inversion results, as expected.

However, in the linearized waveform inversion problem, we know that the reflectors are spiky, and they are continuous along the **x** direction. Under the salt body, the continuity is more important than sparsity because of the small signal-to-noise ratio. In the previous section, we neglect part of the geophysical property, and should not expect to obtain high-quality results.

To make better use of our geophysical knowledge, we first use derivative operator on the x-axis as regularization,

$$\mathbf{J}(\mathbf{m}) = \frac{1}{2}||\mathbf{Hm} - \mathbf{m_{mig}}||_2^2 + \varepsilon||\nabla_x \mathbf{m}||_1. \tag{15}$$

We can see significant improvement of the quality of inversion results from Figure 7(a). The reflectors under the salt body are cleaner. The gaps where we do not have enough data are filled. However,the reflectors are not smooth along x-direction near 17000 ft. It is not clear if they are improvements or artifacts.



(a) (b)

Figure 7: Linearized waveform inversion result with derivative regularization terms using IRLS: (a) inversion result, (b) data residual. [**ER**]

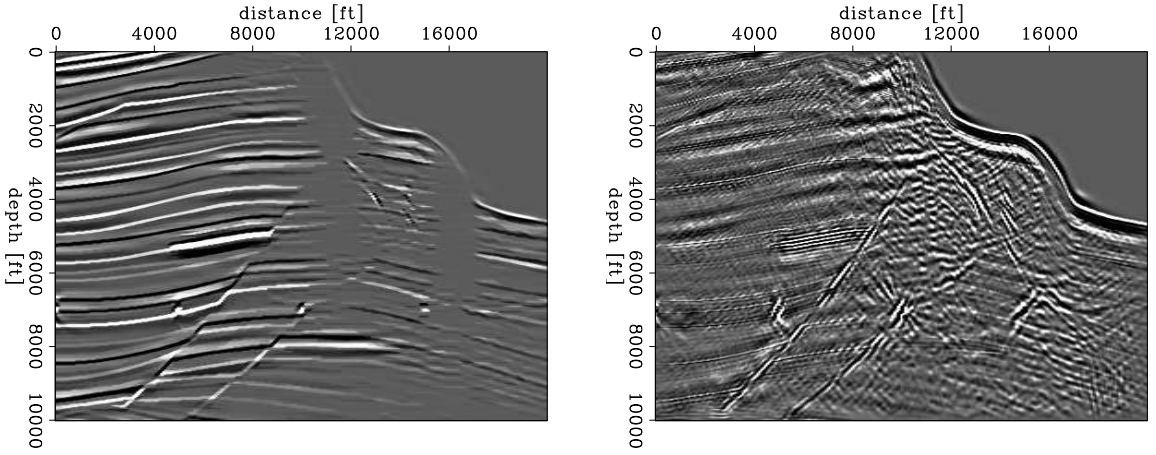
Next, we put both sparsity and continuity information into the objective function as follows:

$$\mathbf{J}(\mathbf{m}) = \frac{1}{2}||\mathbf{Hm} - \mathbf{m_{mig}}||_2^2 + \varepsilon_{\mathbf{1}}||\mathbf{m}||_1 + \varepsilon_{\mathbf{2}}||\nabla_x \mathbf{m}||_1. \tag{16}$$

As the wave propogation is complicated because of the existence of salt body, the data under the salt is diminished and twisted. It makes sense to choose $\varepsilon_{\mathbf{2}} >> \varepsilon_{\mathbf{1}}$ to recover the feature under the salt. Figure 8 shows the inversion result with the new objective function. We can see the reflectors under the salt body are recovered.

CONCLUSIONS

Hyperbolic and hybrid L1/L2 solvers generate more sparse and blocky models, they are also more expensive than traditional least-squares methods. We check a couple of solvers on simple geophysical problems. We demonstrate that great improvement of inversion quality is possible, with acceptable computational cost. However, it is still challenging to find a good L1 (L1-type) solver for generic geophysical problems. In the future, we need to explore the linearized waveform inversion model in detail to get a better understanding. We also need to test a more sophisticated model to get a better understanding of the solvers.



(a) (b)

Figure 8: Linearized waveform inversion result with both regularization terms. Using IRLS. (a) inversion result, (b) data residual. [**ER**]

ACKNOWLEDGEMENT

We thank Elita Li, Mandy Wong and Jon Claerbout for useful suggestions. We also would like to thank Robert Clapp for providing numerious ideas and techinical support on implementing the solvers in the new SEP C++ library.

REFERENCES

Boyd, S., N. Parikh, E. Chu, B. Peleato, and J. Eckstein, 2011, Distributed optimization and statistical learning via the alternating direction method of multipliers: Found. Trends Mach. Learn., **3**, 1–122.

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.

Goldstein, T. and S. Osher, 2009, The split bregman method for l1-regularized problems: SIAM Journal on Imaging Sciences, **2**, 323–343.

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

Li, E., Y. Zhang, and J. F. Claerbout, 2010, Geophysical applications of a novel and robust L1 solver: SEP 140, 155–164.

M. Clapp, R. C. and B. Biondi, 2005, Regularized least-squares inversion for 3-D subsalt imaging: Soc. of Expl. Geophys., 1814–1817.

Nichols, D., 1994, Velocity-stack inversion using L_p norms: SEP 82, 1–16.

Tang, Y., 2008, Wave-equation hessian by phase encoding: SEP-Report, **134**, 1–25.

Valenciano, A., 2006, Target-oriented wave-equation inversion: Geophysics, **71**, A35–38.

Zhang, Y. and J. F. Claerbout, 2010, Least-squares imaging and deconvolution using the hybrid norm conjugate-direction solver: SEP 140, 129–142.