Full waveform inversion and joint migration inversion with an automatic directional total variation constraint[a]

[a]Published in Geophysics, 84, no. 2, R175-R183, (2019)

Shan Qu, Eric Verschuur*, and Yangkang Chen†*

ABSTRACT

As full waveform inversion (FWI) is a non-unique and typically ill-posed inversion problem, it needs proper regularization to avoid cycle-skipping. To reduce the non-linearity of FWI, Joint Migration Inversion (JMI) is proposed as an alternative, explaining the reflection data with decoupled velocity and reflectivity parameters. However, the velocity update may also suffer from being trapped in local minima. To optimally include geologic information, we propose FWI/JMI with directional total variation as an L1-norm regularization on the velocity. We design the directional total variation operator based on the local dip field, instead of ignoring the local structural direction of the subsurface and only using horizontal- and vertical- gradients in the traditional TV. The local dip field is estimated using plane-wave destruction based on a raw reflectivity model, which is usually calculated from the initial velocity model. With two complex synthetic examples, based on the Marmousi model, we demonstrate that the proposed method is much more effective compared to both FWI/JMI without regularization and FWI/JMI with the conventional TV regularization. In the JMI-based example, we also show that L1 directional TV works better than L2 directional Laplacian smoothing. In addition, by comparing these two examples, it can be seen that the impact of regularization is larger for FWI than for JMI, because in JMI the velocity model only explains the propagation effects and, thereby, makes it less sensitive to the details in the velocity model.

INTRODUCTION

Seismic full waveform inversion (FWI) is a powerful method for providing a quantitative description of the subsurface properties by iteratively minimizing an objective function that measures the misfit between observed and predicted data in the least-squares sense (Tarantola, 1984). However, FWI is a non-linear and ill-posed inverse problem and its objective function may suffer from local minima that are not informative about the true parameters (Virieux and Operto, 2009; Chen et al., 2016; Fu and Symes, 2017a,b). Using regularization methods is an effective way to mitigate this ill-posedness and non-uniqueness of FWI.

Joint Migration Inversion (JMI) was proposed as one of the methods to overcome the above-mentioned limitations in FWI (Berkhout, 2014b; Staal, 2015; Verschuur et al., 2016). It is an inverse algorithm to automatically derive both velocity and reflectivity based on the full wavefield modeling (FWMod) process (Berkhout, 2014a) that takes into account transmission effects and surface/internal multiples. In the FWMod procedure, the velocity only affects the kinematics without any scattering effect in the modeling operators and the reflectivity only deals with scattering effects. These characteristics lead to a reduced non-linearity in the inversion process. Even though not as severe as FWI, the velocity update may still suffer from being trapped in local minima. With the help of regularization, JMI can get a more accurate inverted velocity, and thus achieve a better inverted reflectivity (Qu and Verschuur, 2016b, 2017b).

The popular regularization methods include: quadratic L2-norm-based regularization, such as Tikhonov regularization (Hu et al., 2009), and laplacian smoothing (Guitton et al., 2012; Qu and Verschuur, 2016a, 2017a), which tend to produce models with blurred discontinuities; non-quadratic L1-norm-based regularization, such as total variation (TV) (Anagaw and Sacchi, 2011; Qu and Verschuur, 2016b), smooths the model by enhancing the sparsity of the spatial gradient of the velocity, thereby preserving its edges. However, regular TV regularization only tends to reduce the horizontal and vertical gradients of each gridpoint in the model regardless of their structural direction. Thus, TV is not suitable where the local geologic structure has a dominant structural direction. Unlike general digital images, the spatial changes of the seismic model always have some specific geological structures, like tilted layers, faults, or edges of a salt body (Chen et al., 2017; Wu and Bai, 2018; Bai et al., 2018; Chen et al., 2020). Bayram and Kamasak (2012) proposed a directional TV method and applied it to digital image denoising. However, they only consider one single dominant direction for all pixels, which is obviously ineffective for complex-textured geologies. Therefore, we propose a directional TV constraint based on a rough estimate of the subsurface image.

The paper is organized as follows: we first briefly introduce the basics of FWI and JMI. Next, we formulate the conventional TV and the proposed directional TV. Finally, using two complex Marmousi-model-based examples, we show that the proposed method is more effective than the alternative methods, when the model contains tilted layers and steep faults. At the end, using the JMI-based example, we also show that the L1 directional TV works better than the L2 directional Laplacian smoothing regarding the preservation of edges and the steering of the update away from the local minimum. Note that this paper is an extended version of work published in Qu et al. (2017).

THEORY OF FWI

Mathematically, regular FWI can be formulated as a minimization problem with the following objective function:

$$J_{FWI} = ||\mathbf{d} - f(\mathbf{m})||_2^2 + constraint, \tag{1}$$

where $\mathbf{m}$ represents the model, $\mathbf{d}$ is the observed dataset, $f(.)$ is the corresponding modeling operator, and $||.||_2^2$ stands for the L2-norm. In most FWI implementations, $\mathbf{m}$ consists of a gridded velocity distribution that explains both propagation and reflection of the seismic data and forward modeling is done via a finite-difference implementation of the two-way wave equation (Virieux and Operto, 2009). Note that in most FWI implementations, density variations are neglected. Minimizing this misfit function is likely to suffer from ill-posedness and non-uniqueness because of limited input data and non-linearity of the forward modeling operator. Adding regularization to the objective function can be one effective way to mitigate the ill-posedness and non-uniqueness of this inverse problem.

THEORY OF JMI

Joint Migration Inversion (JMI) was proposed as one of the methods to overcome the above-mentioned limitations in FWI. In Figure 1 the flow diagram of the JMI process is shown. The main engine of the JMI method is a forward-modeling process, called Full Wavefield Modeling (FWMod) based on the parameters reflectivity and velocity, which is described in Berkhout (2012), Davydenko and Verschuur (2012), and Berkhout (2014a). With this recursive and iterative two-way modeling process, from the current estimate of the reflectivities and the velocity model, the seismic reflection responses are being generated. In this modeling process, multiples and transmission effects are included. Next, the modeled responses are compared to the measured ones and the resulting difference data, being the residual of the inversion, is back-projected into the parameter space via reverse extrapolation of the residual into the medium and a subsequent transformation of this residual energy into updates of reflectivity and velocity. The parameters are updated, from which new seismic responses are modeled, yielding the next version of the residual data. In this way, the residual is slowly driven to zero (Berkhout, 2012; Staal and Verschuur, 2012, 2013; Berkhout, 2014c). We can treat the whole procedure as minimizing the following objective function:

$$J_{JMI} = \sum_{\omega} ||\mathbf{D}^-(z_0) - \mathbf{P}^-_{mod}(z_0, \mathbf{r}, \mathbf{v})||_2^2 + constraint, \tag{2}$$

where the $||.||^2$ describes the sum of the squares of the values (i.e., the energy), $\mathbf{D}^-(z_0)$ is the collection of all recorded surface seismic shot records in the (x, ω) domain, and $\mathbf{P}^-_{mod}(z_0, \mathbf{r}, \mathbf{v})$ describes the modeled surface shot records as a function of reflectivity $\mathbf{r}$ and velocity $\mathbf{v}$. Note that by using the reflectivity and propagation velocity as

parameters, density variations are implicitly included in $\mathbf{r}$. Even though JMI has a reduced non-linearity, the velocity update still suffers from local minima. With a proper constraint, JMI can lead to a more accurate inverted velocity, and therefore a better inverted reflectivity.

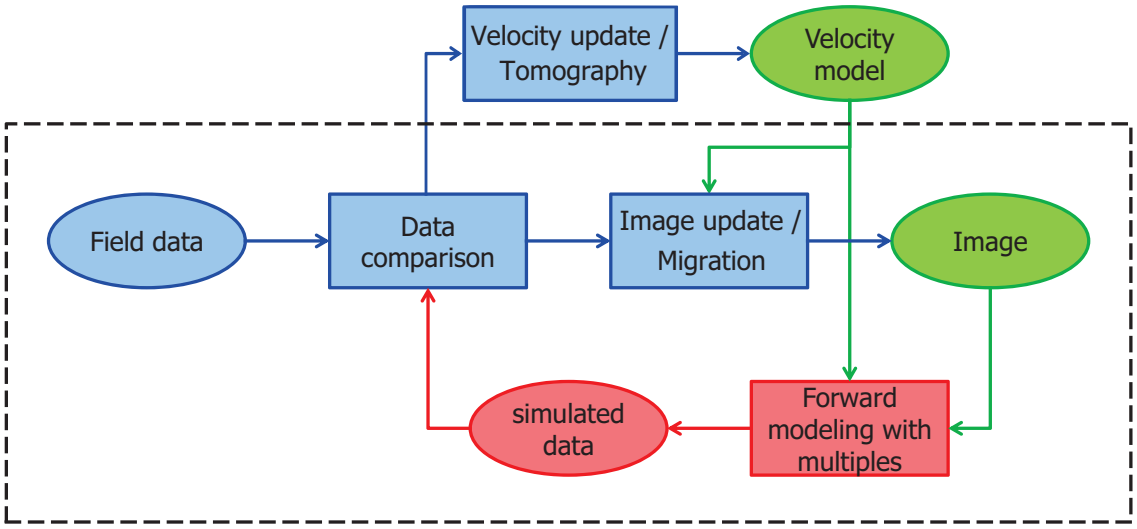

Figure 1: JMI flow chart.

FWI/JMI WITH TV AND DIRECTIONAL TV

In this paper, we consider anisotropic TV as the basic regularization method, since TV can smooth the model and at the same time preserve edges by enhancing the sparsity of the spatial gradient of the velocity difference. In addition, the anisotropic version is easier to minimize compared to the isotropic one. Furthermore, we restrict ourselves to the 2D case, although extension to the full 3D situation is relatively straightforward.

The extended misfit function with a TV constraint can be expressed as

$$J_{tot} = \mu J + \lambda C_{TV}(\mathbf{p}) = \mu J + \lambda ||\nabla_x \mathbf{p}||_1 + \lambda ||\nabla_z \mathbf{p}||_1. \tag{3}$$

Here, J is J_{FWI} or J_{JMI}. $\mathbf{p}$ is the parameter constrained by TV($\mathbf{p} = \mathbf{m}$ for FWI, and $\mathbf{p} = \mathbf{v}$ for JMI). ∇_x and ∇_z represent horizontal- and vertical-gradient operator, respectively. For one gridpoint (i, j) in a cartesian coordinate (x, z), $\nabla_x \mathbf{p}(i, j) = p_{i+1,j} - p_{i,j}$ and $\nabla_z \mathbf{p}(i, j) = p_{i,j+1} - p_{i,j}$ (illustrated in figure 2a with the black dashed arrows). μ is the weight parameter of the fidelity term. λ is the coefficient of the constraint term. The latter two together control the balance between the regularization and the misfit function.

However, this conventional TV regularization only tends to reduce the horizontal- and vertical-gradients of each gridpoint in the model, regardless of the geological direction of the model. Therefore, TV is not suitable where the local structure has a

dominant direction. Unlike general digital images, the spatial changes in the subsurface always follow some specific geological structures, e.g., tilted layers, faults, and edges of a salt body. In this case, we propose FWI/JMI with directional TV and we design the directional TV based on the local dip estimated from a rough reflection image using the plane-wave destruction (PWD) algorithm (Fomel, 2002).

The misfit function with directional TV can be formulated as

$$J_{tot} = \mu J + \lambda C_{DTV}\left(\mathbf{p}\right) = \mu J + \lambda||\nabla_1 \mathbf{p}||_1 + \lambda||\nabla_2 \mathbf{p}||_1, \tag{4}$$

where ∇_1 and ∇_2 are the gradient operators of the dominant direction and the direction perpendicular to the dominant direction, respectively. From the viewpoint of physical meaning, ∇_1 and ∇_2 are the rotated and scaled version of ∇_x and ∇_z, according to the estimated local dip and a weighting parameter. Mathematically, for one point (i,j), $\nabla_1\mathbf{p}\left(i,j\right)$ and $\nabla_2\mathbf{p}\left(i,j\right)$ can be represented as

$$\begin{aligned} \begin{pmatrix} \nabla_1\mathbf{p}\left(i,j\right) \\ \nabla_2\mathbf{p}\left(i,j\right) \end{pmatrix} &= \Lambda\mathbf{R} \begin{pmatrix} \nabla_x\mathbf{p}\left(i,j\right) \\ \nabla_z\mathbf{p}\left(i,j\right) \end{pmatrix} \\ \text{Here } \Lambda = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}, &\mathbf{R} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \end{aligned} \tag{5}$$

where Λ and $\mathbf{R}$ represent scaling matrix and rotation matrix, respectively. α_1 and α_2 represent the weight on the gradient of the dominant direction and its perpendicular direction, respectively, and θ is the dip of the local structure. An illustration of such a directional TV is shown in figure 2a with the red solid arrows.

Please note that if we assume $\alpha_1 = \alpha_2 = 1$ and $\theta = 0^o$, then Λ turns into an identity matrix, which means the same weights are put on both directions, and $\mathbf{R}$ also becomes an identity matrix, indicating that the target directions are horizontal and vertical. Therefore, we can see that the conventional TV is actually a special case of the directional TV, and in turn, the directional TV is a more general version of the conventional TV and more suitable to a model with complex geologic structures. In this paper, we solve both FWI/JMI with the conventional TV and FWI/JMI with the directional TV effectively using the split-Bregman iterative algorithm (Goldstein and Osher, 2009). We only show the framework of solving FWI with the directional TV in Algorithm 1, because, as mentioned before, we treat the conventional TV as a special case of directional the TV, and JMI with the conventional TV/directional TV will follow a similar algorithm.

FWI EXAMPLE

In order to demonstrate the effectiveness of FWI with directional TV, we consider the velocity model shown in Figure 2a, which is scaled from the top half of the Marmousi model. To avoid cycle-skipping, a Ricker wavelet with a dominant frequency of 14 Hz is used as the source wavelet. Using a constant-density acoustic finite-difference

modeling, we generated 23 shots with 334 receivers for each shot. The shot spacing is 180 m and the receiver spacing is 12 m. The horizontal and vertical grid size are 12 m. In addition, some random noise with $SNR = 10$ is added to the modeled data. The middle shot gather is shown in Figure 3. The initial velocity model is shown in Figure 2b, which is a smoothed version of the model in Figure 2a.

First, with the initial model, we apply 6 iterations of Full Wavefield Migration (Berkhout, 2014b; Davydenko and Verschuur, 2017) at a maximum frequency of 25 Hz to the dataset and then denoise the inverted image via a simple soft-thresholding in the curvelet domain (Figure 4a, Donoho (1995)). Note that full wavefield migration can be considered as a JMI process in which the velocity is assumed known and fixed. It will honour all multiples and transmission effects properly. Now, with this inverted reflectivity, we can estimate the dip field using plane-wave destruction algorithm proposed by Fomel (2002), shown in Figure 4b. This estimated dip field is used to build the directional TV operator for each gridpoint.

Next, we compare three methods: regular FWI without any regularization, FWI with conventional TV, and FWI with directional TV. We use the same μ and λ for TV and directional TV. We choose a relaxation strategy to set μ, which is increasing exponentially. λ is chosen as 0.005, which depends on the scale of data. For directional TV, $\alpha_1 : \alpha_2 = 3 : 1$ and $\alpha_1 + \alpha_2 = 2$. For TV, Λ is an identity matrix. After 100 iterations, the inverted results are shown in Figure 8. Note that the regular FWI without any regularization is trapped into local minima very quickly, despite the accurate starting model (Figure 8a). With the help of TV regularization, FWI with TV achieves a better result by smoothing the model via enhancing the sparsity of the spatial gradient of the velocity difference, which allows us to steer away from local minima. However, we can observe that the structures still remain vague in Figure 8b, especially in the deeper area, since traditional TV only uses horizontal- and vertical- and ignores the local structure. Compared to the regular TV, much weaker artifacts can be observed in the result of FWI with the directional TV, shown in Figure 8c, because we consider the structural directions of the spatial gradient and their weights according to the local dip. The convergence diagrams of the misfit function with iteration number corresponding to the three methods are shown in Figure 8d, in which it is visible that FWI with TV works well to mitigate the ill-posedness and non-uniqueness of FWI, and FWI with directional TV behaves clearly better than FWI with the conventional TV. Figure 6 shows a comparison between the different velocities at three different locations. The locations are annotated in Figure 8a-c. We can see the obvious improvement using directional TV.

To further illustrate the contribution of regularization in the inversion, we show the gradients from the residual at the first iteration based on the different methods in Figure 7. Compared to Figure 7a, Figure 7b has sharper structures, especially in the deeper part, by preserving the edges via the TV constraint. The gradient in Figure 7c shows even more blocky structures that correspond to the geologic information. Note that, in Figure 7c, there are imprints introduced by the imperfect raw reflectivity model and dip field (Figure 4). However, these imprints have been compensated and

suppressed during inversion, and the proposed method ends up with a decent result shown in Figure 8c, which shows the proposed method is insensitive to the locally-incorrect dip field. Figure 8 and 9 demonstrate the corresponding depth migration images and common image gathers calculated using full wavefield migration. We can see that the reflectivity based on the velocity from FWI with directional TV has the best focusing resolution and less imprints, and their corresponding common image gathers are flatter than the alternative methods. Please note some obvious improvements pointed out by the red arrows. In the end, we show in Figure 10 the modeled data generated from each of the final inverted velocities and the corresponding differences with the observed data. From this figure, we note that FWI with directional TV approach can restore the observed data much better than the alternatives.

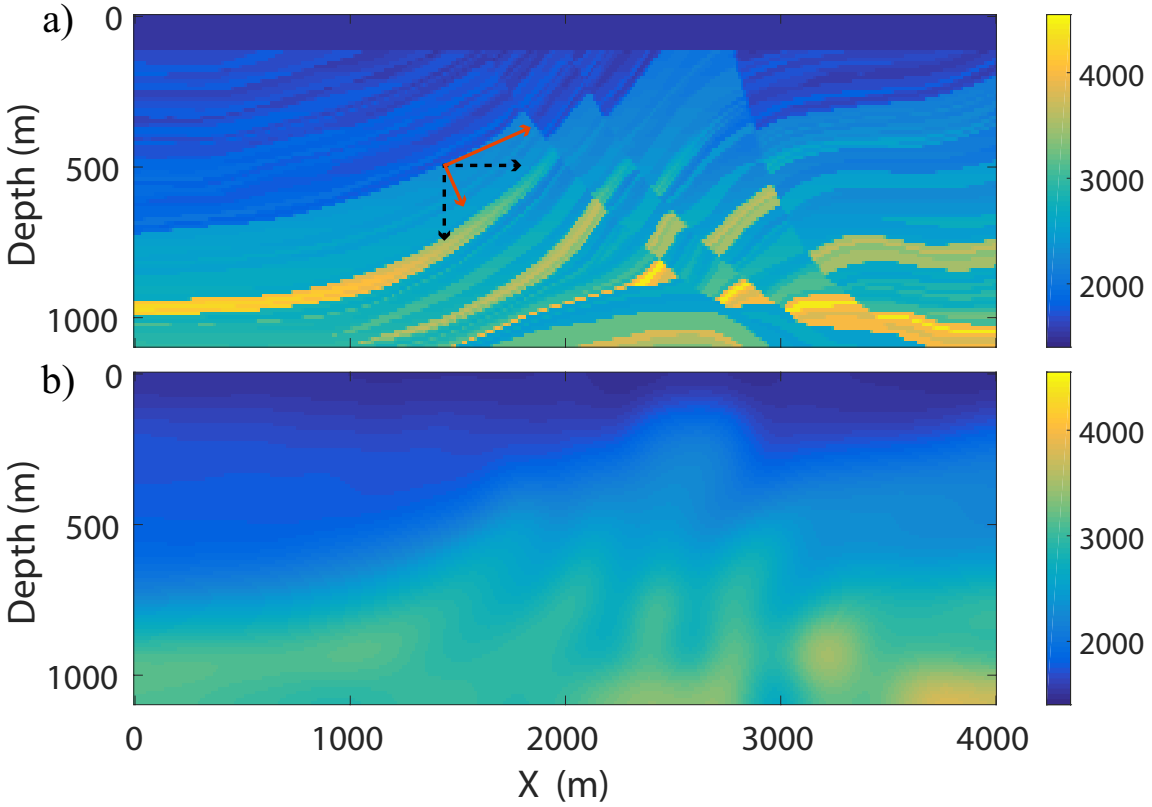

Figure 2: FWI example: (a) Real velocity model, at specific gridpoint (i, j). The black dash arrows illustrate $\nabla_x \mathbf{p}\,(i, j)$ and $\nabla_z \mathbf{p}\,(i, j)$, the red solid arrows illustrate $\nabla_1 \mathbf{p}\,(i, j)$ and $\nabla_2 \mathbf{p}\,(i, j)$, based on the structural dip at (i, j). (b) Initial velocity model.

.

JMI EXAMPLE

In this part, we use the same model as in the previous example to demonstrate the effectiveness of JMI with directional TV. Because the forward modeling in JMI is

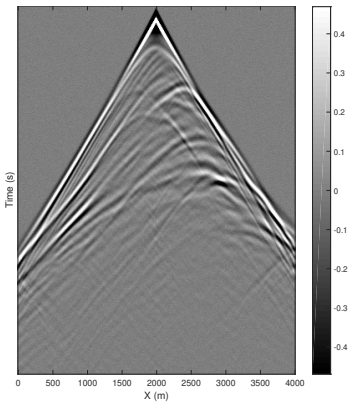

Figure 3: FWI example: Recorded middle shot gather at $X = 2000$ m

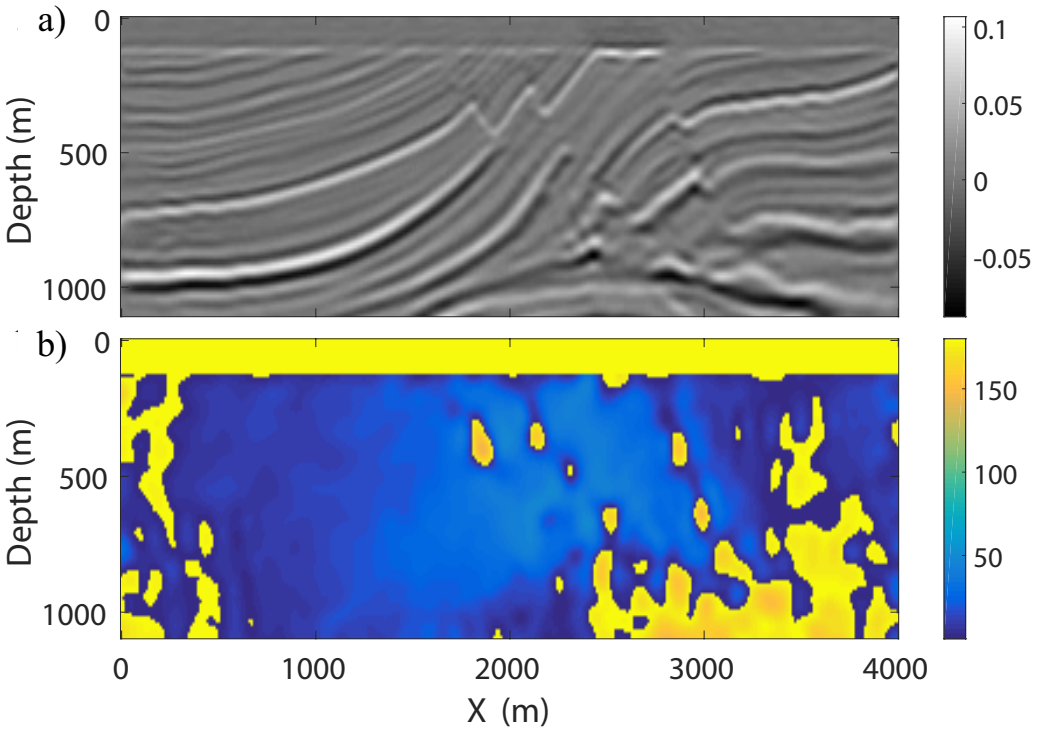

Figure 4: FWI example: (a) The inverted reflectivity model after denoising using thresholding in the curvelet domain. (b) The estimated dip field (in degrees).

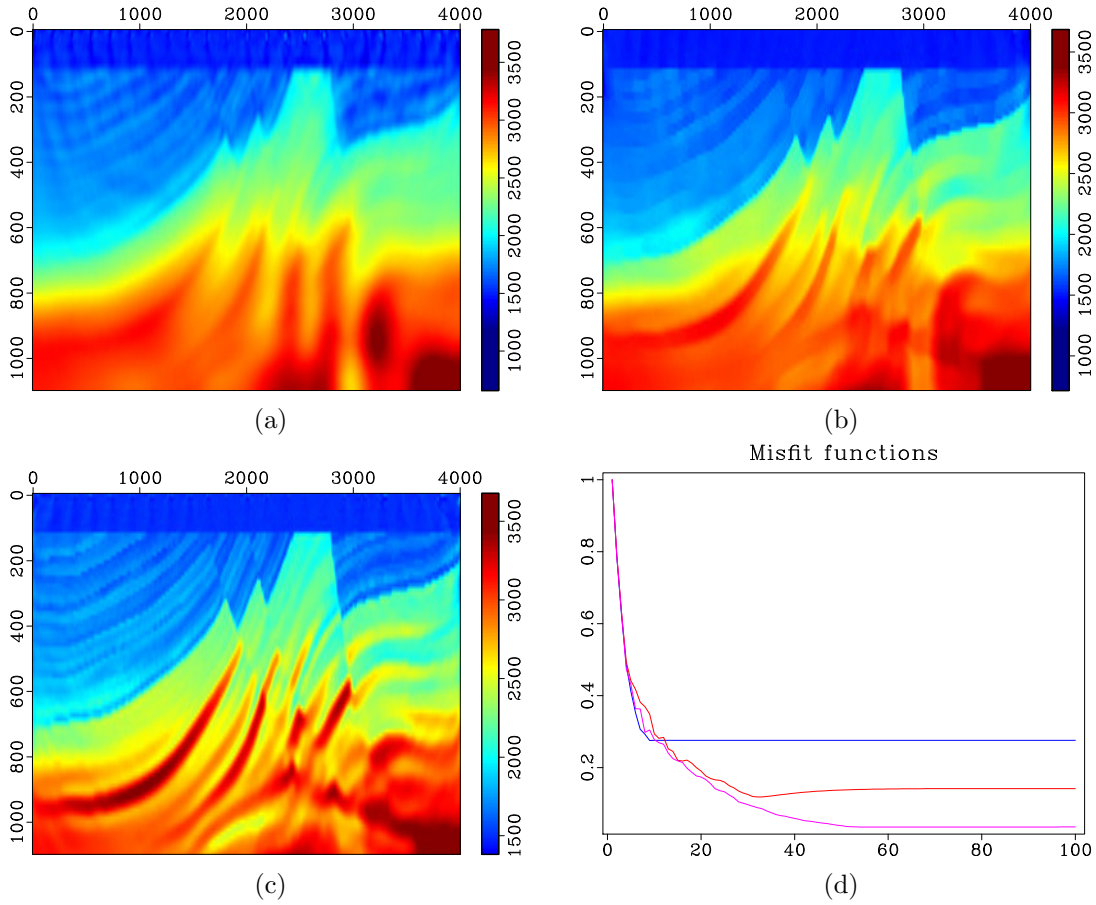

Figure 5: FWI example: The inverted velocity using (a) regular FWI without any regularization, (b) FWI with conventional TV, and (c) FWI with directional TV. (d) The convergence diagrams of the data misfit as a function of iteration. The blue line denotes the inverted velocity using regular FWI without any regularization. The red line denotes the inverted velocity using FWI with conventional TV. The yellow line denotes the inverted velocity using FWI with directional TV.

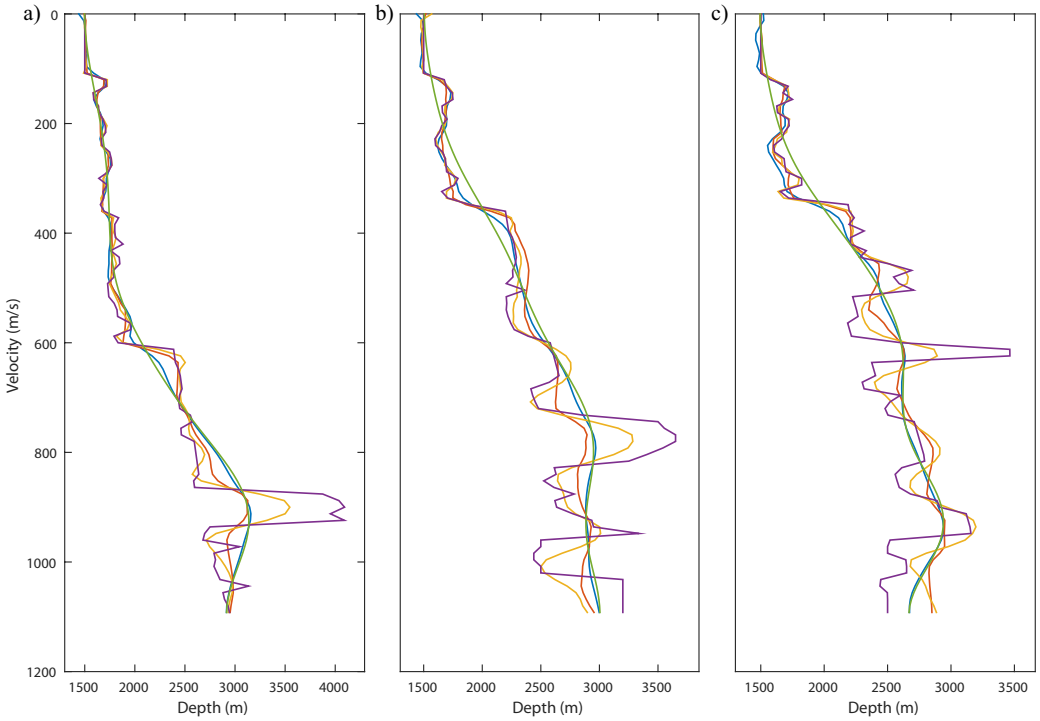

Figure 6: FWI example: Comparison of different velocities at three locations. Velocity comparison at (a) $X = 1000$ m, (b) $X = 2000$ m, (c) $X = 3000$ m. The purple line denotes the true velocity. The green line denotes the initial velocity. The blue line denotes the inverted velocity using regular FWI without any regularization. The red line denotes the inverted velocity using FWI with conventional TV. The yellow line denotes the inverted velocity using FWI with directional TV. The three locations are highlighted by the black dash lines in Figure **??**a-c

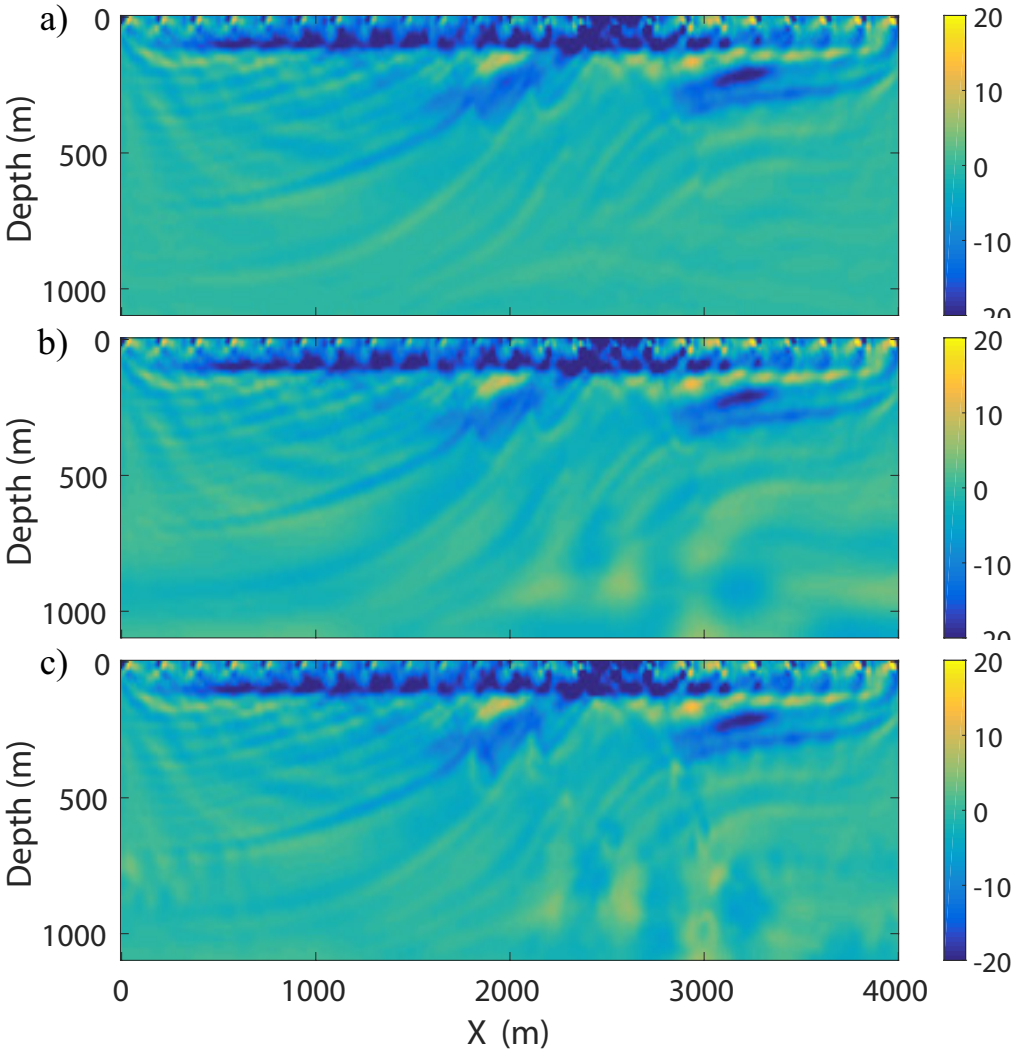

Figure 7: FWI example: The velocity gradient from the residual at the first iteration using (a) regular FWI without any regularization, (b) FWI with conventional TV, and (c) FWI with directional TV.

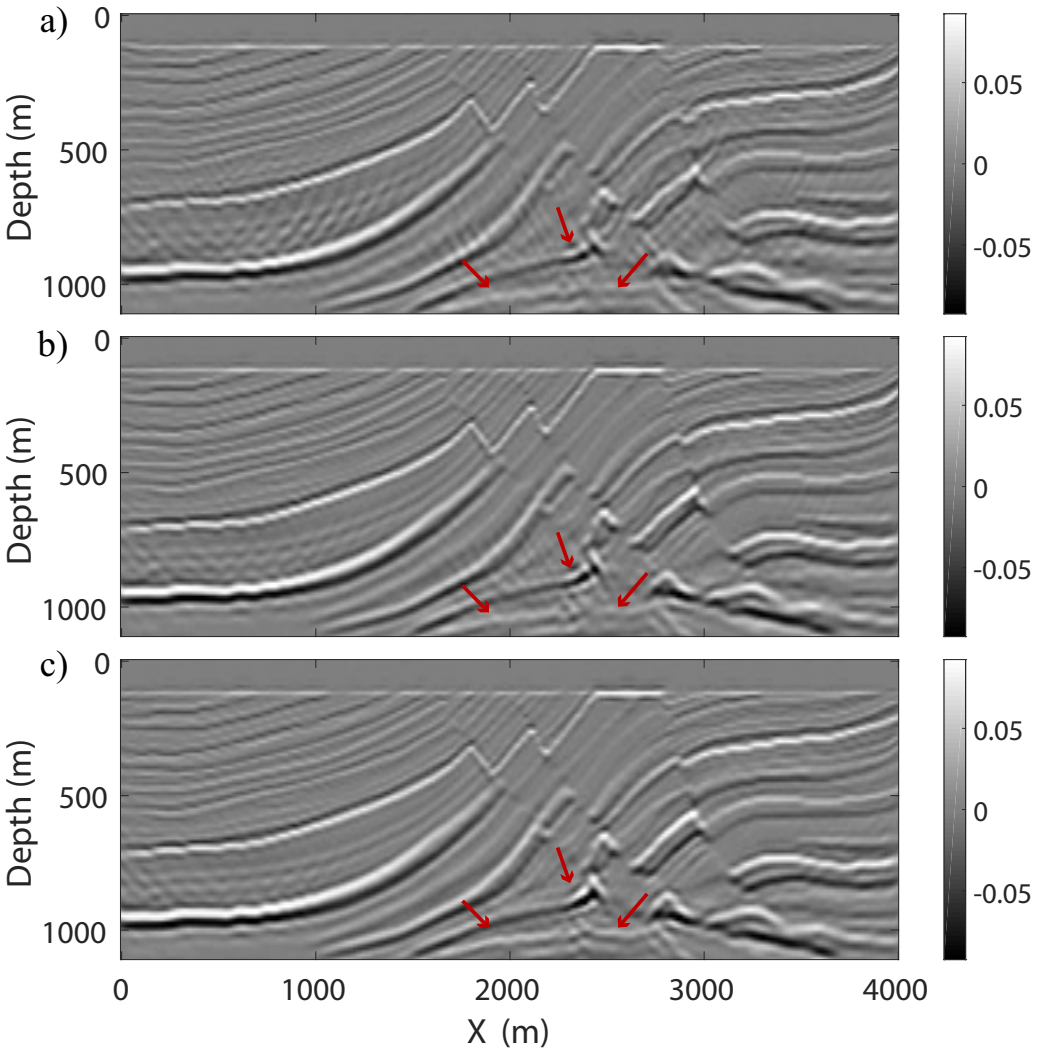

Figure 8: FWI example: The corresponding reflectivity based on the inverted velocity using (a) regular FWI without any regularization, (b) FWI with conventional TV, and (c) FWI with directional TV.

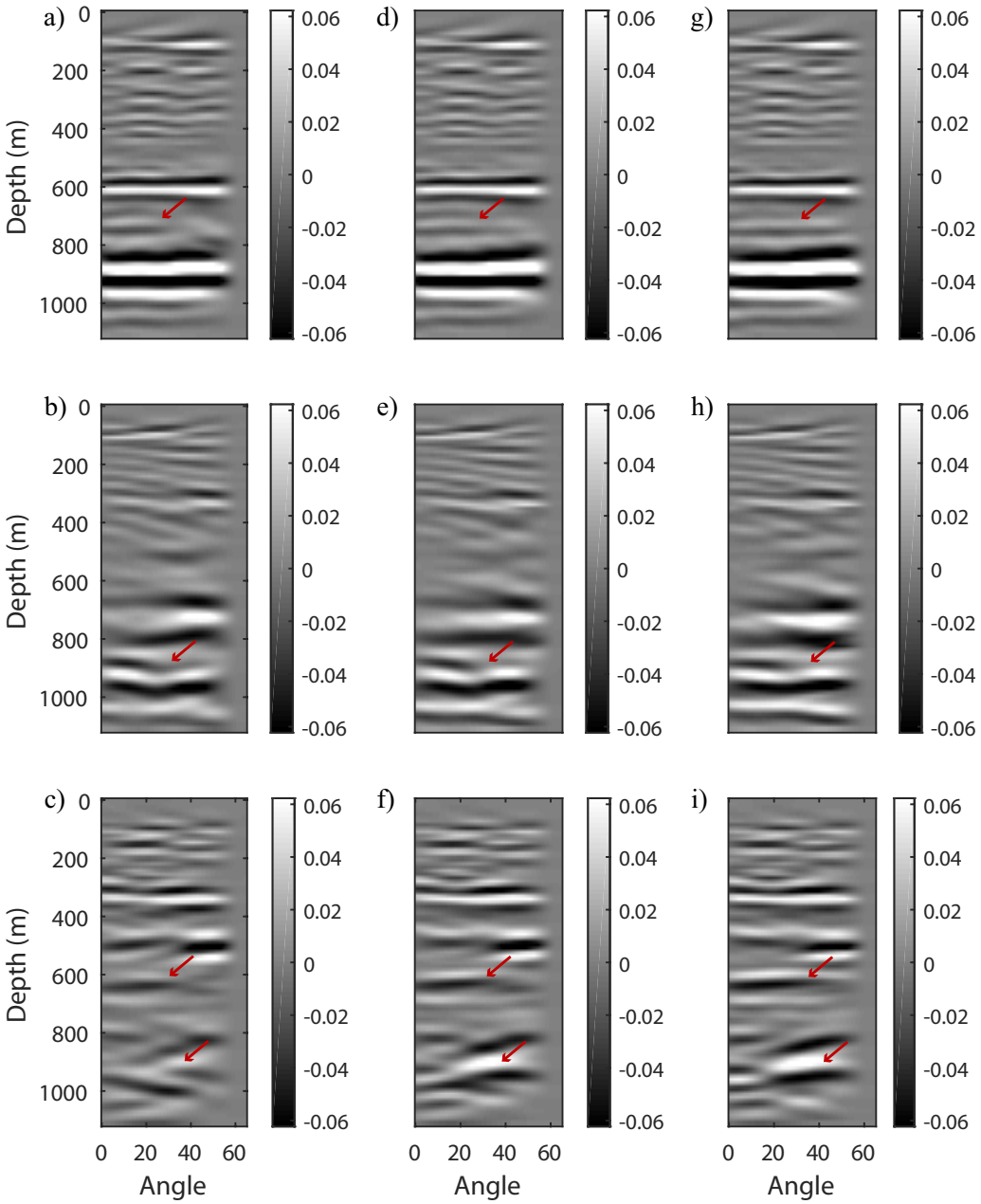

Figure 9: FWI example: The corresponding angle gathers generated at $X = 1000$ m, $X = 2000$ m, and $X = 3000$, based on the inverted velocity using (a, b, c) regular FWI without any regularization, (d, e, f) regular FWI with conventional TV, and (g, h, i) regular FWI with directional TV.

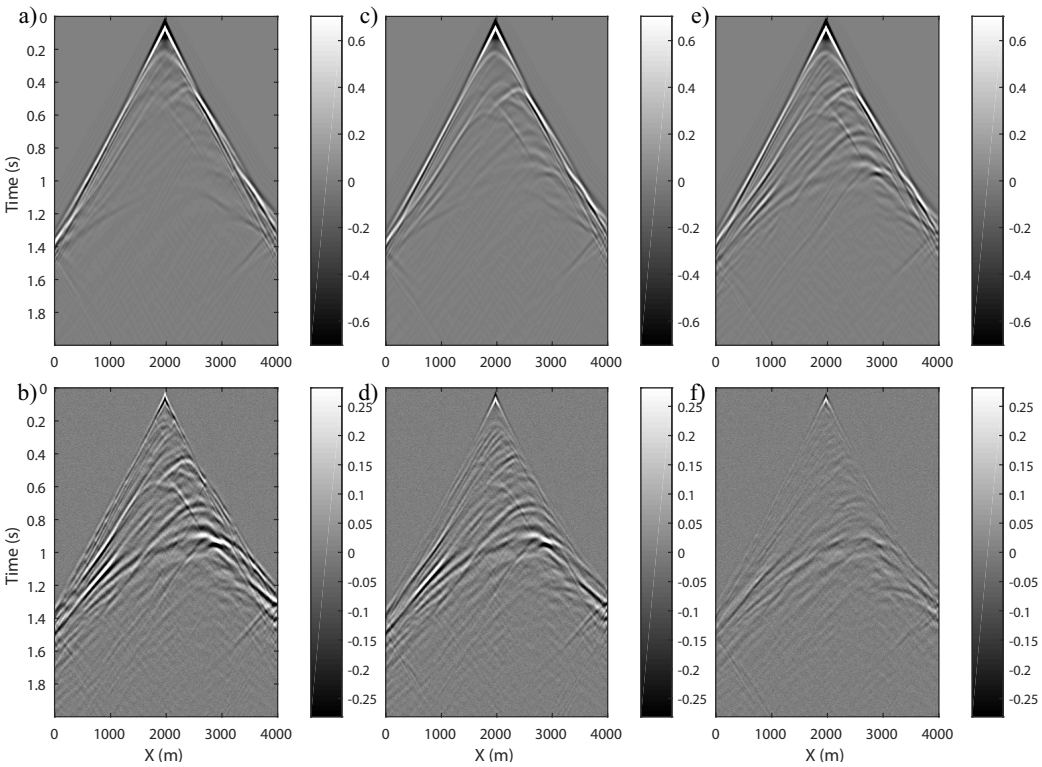

Figure 10: FWI example: The modeled data generated at $X = 2000$ m and the corresponding difference with the observed data based on the inverted velocity using (a, b) regular FWI without any regularization, (c, d) regular FWI with conventional TV, and (e, f) regular FWI with directional TV

computationally less expensive that in FWI, we are able to use a Ricker wavelet with a higher dominant frequency of 20 Hz to acquire higher-resolution reflectivities and velocity. The shot spacing is 200 m and the receiver spacing is 20 m. The horizontal and vertical grid size are 20 m and 10 m, respectively. Surface multiples are excluded in the modeling, but internal multiples and transmission effects are included. The direct wave is removed, as it cannot be explained by JMI. Initially, reflectivities are zero and the initial velocity is a very simple vertical gradient (shown in Figure 11). First, with the initial model, we apply 30 iterations of JMI with 5 Hz $-$ 25 Hz frequency bandwidth of to the dataset and then preprocess the inverted image via a simple soft-thresholding in the curvelet domain. Then, with this preprocessed inverted reflectivity, we estimate the dip field to build the directional TV operator. Meanwhile, the inverted velocity can be used as initial velocity model for the next step.

Next, we compareresults from the regular JMI without any regularization, JMI with conventional TV, and JMI with directional TV. The frequency bandwidth during the second step of JMI is 5 Hz $-$ 40 Hz. We use the same μ and λ for both the conventional TV and directional TV. μ is also increasing with iteration and $\lambda = 1.2$. For directional TV, $\alpha_1 : \alpha_2 = 3 : 1$ and $\alpha_1 + \alpha_2 = 2$. For the conventional TV, Λ is an identity matrix. After 50 iterations for each method, the inverted results are shown in Figure 12 and Figure 13. Because of the inversion process included in JMI, all the images in Figure 13 are quite accurate compared to the true reflectivity structures. Furthermore, all the estimated velocity models in Figure 12 are also surprisingly stable and show some details.

In Figure 12a, the regular JMI without any regularization is slightly trapped in a local minimum, e.g., in the lower right area pointed by the red arrow. With the help of TV regularization, JMI with conventional TV in Figure 12b achieves a better result by smoothing the model via enhancing the sparsity of the spatial gradient of the velocity difference, which allows us to steer away from the local minimum. Instead of using the conventional TV, a much better inverted velocity with clearer edges of structures is obtained in Figure 12c using JMI with directional TV. This is because we consider the structural directions of the spatial gradient and their weights according to the local dip from the associated image. Please note some obvious improvements pointed out by the white arrows. In addition, compared to L1 directional TV, L2 directional Laplacian smoothing results in a smoother velocity model (Figure 12d); however, it intensifies the local minima issue and tends to produce models with blurred discontinuities. That is because the directional Laplacian smoothing may over-smooth the velocity and cannot preserve edges very well; it is also more sensitive to the accuracy of the estimated dip field, compared to L1 directional TV. As a result of the improvement of the inverted velocity, the inverted reflectivity also becomes more accurate (Figure 13): The inverted reflectivities highlighted with white arrows in Figure 13c have better focussing and less distortions than the other alternatives.

Note that the velocity field estimated from JMI has less details compared to that from FWI, as it only needs to describe propagation, not reflection. Similar as in the

FWI example, we show in Figure 14 the modeled data generated from each of the final inverted velocities and the corresponding differences with the observed data. From this figure, we can see that regularizations on velocity do not make much difference in the data residual, because the velocity in JMI only explains propagation effects, and the reflectivities explain the scattering effects, which makes JMI less sensitive to the details in the velocity model compared to FWI.

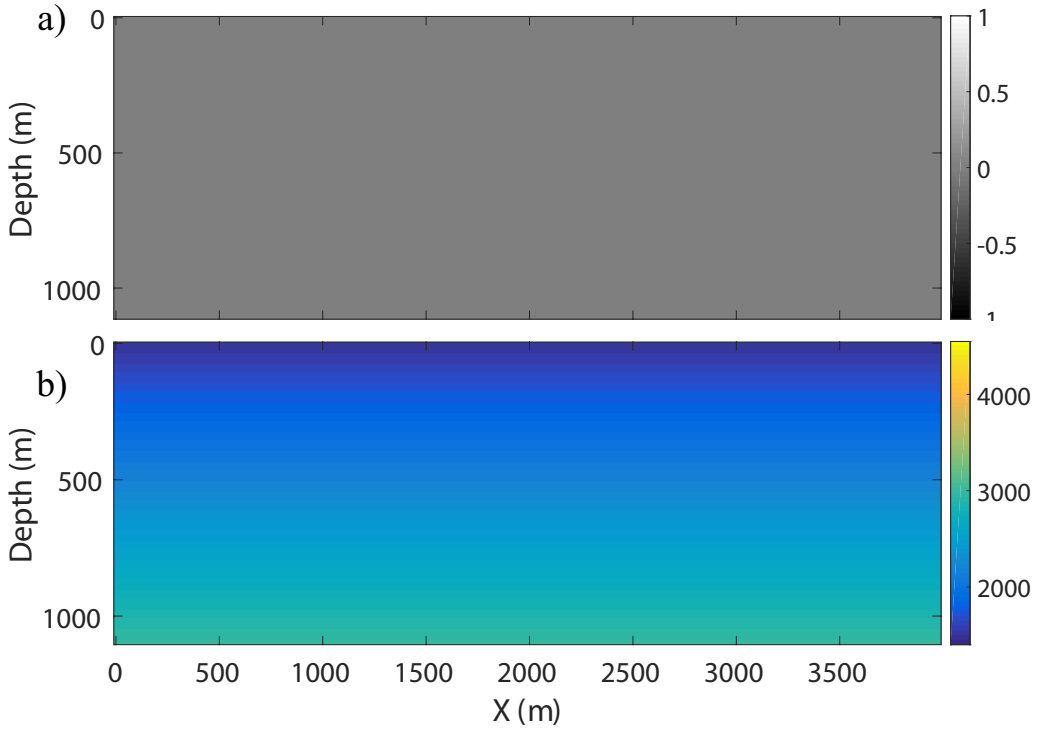

Figure 11: JMI example: (a) Initial reflectivity model. (b) Initial velocity model.

DISCUSSION

FWI/JMI with directional TV has been demonstrated to be a more effective method than the alternatives. We design the directional TV based on the dip field calculated from an initial image. By considering the local structural directions of the spatial gradient and their weights according to the local dip, the proposed method achieves a better result compared to FWI/JMI without regularization or with conventional TV. In the case of complex subsurface structures, the local dip map cannot be estimated properly. However, directional TV regularization is not sensitive to the accuracy of the estimated dip, because even using an arbitrary dominant direction would not be worse than using horizontal- and vertical-gradients like using conventional TV in a complex area.

In terms of the parameter selection, we choose a relaxation strategy for μ, which is increasing exponentially. In this way, we relax the strength of the L1 constraint

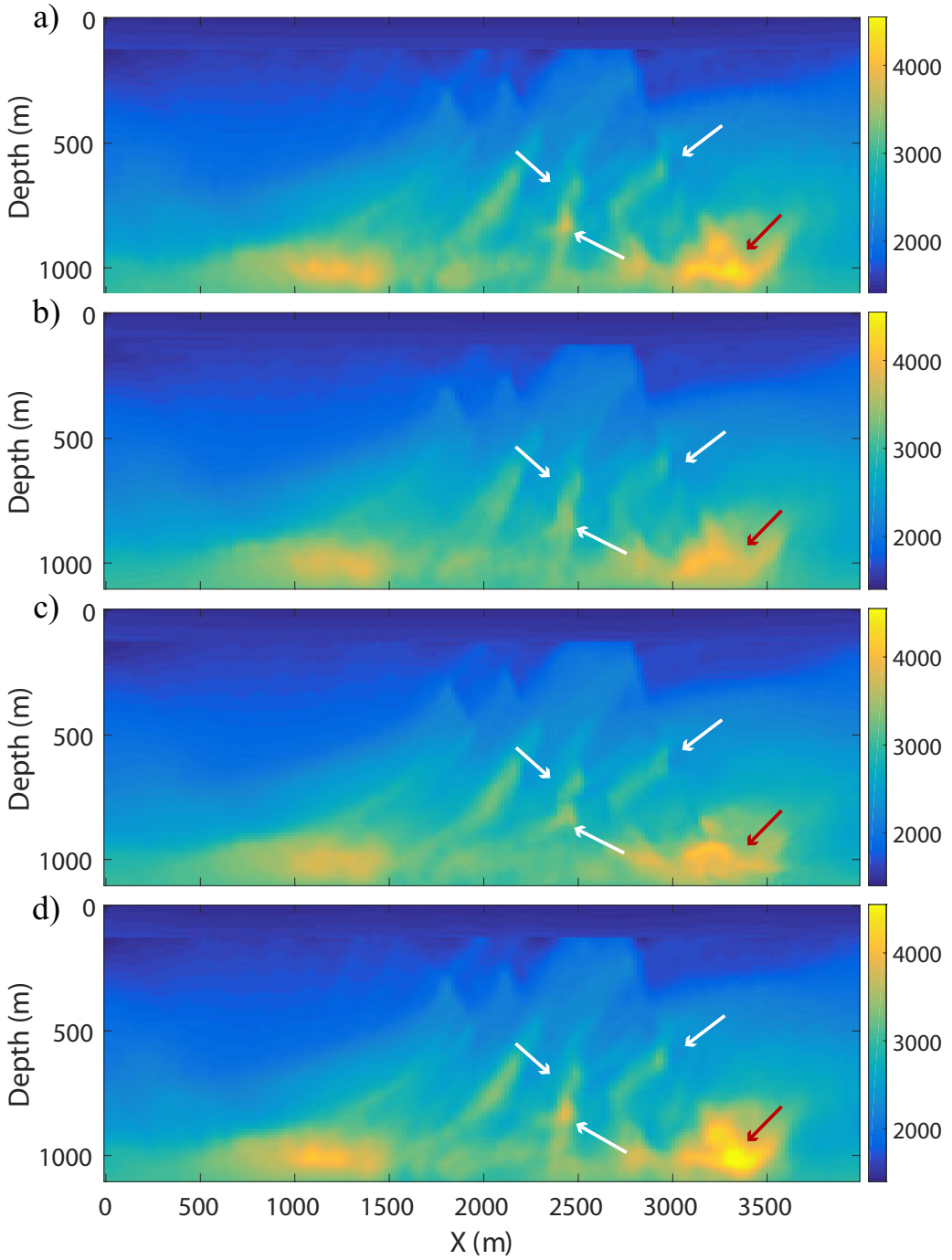

Figure 12: JMI example: The inverted velocity using (a) regular JMI without any regularization, (b) JMI with conventional TV, (c) JMI with directional TV, and (d) JMI with L2 directional laplacian smoothing.

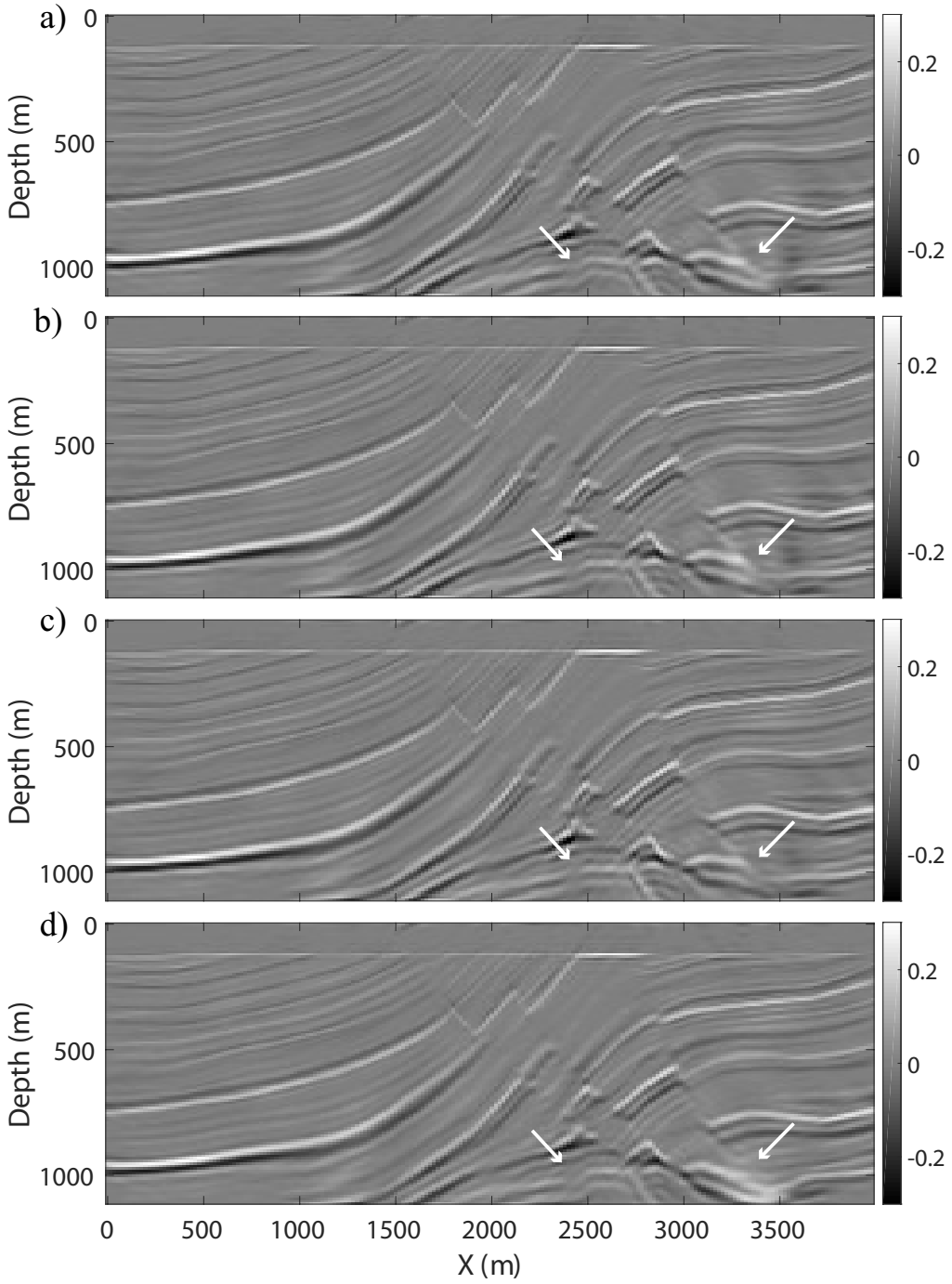

Figure 13: JMI example: The inverted reflectivity using (a) regular JMI without any regularization, (b) JMI with conventional TV, (c) JMI with directional TV, and (d) JMI with L2 directional laplacian smoothing.

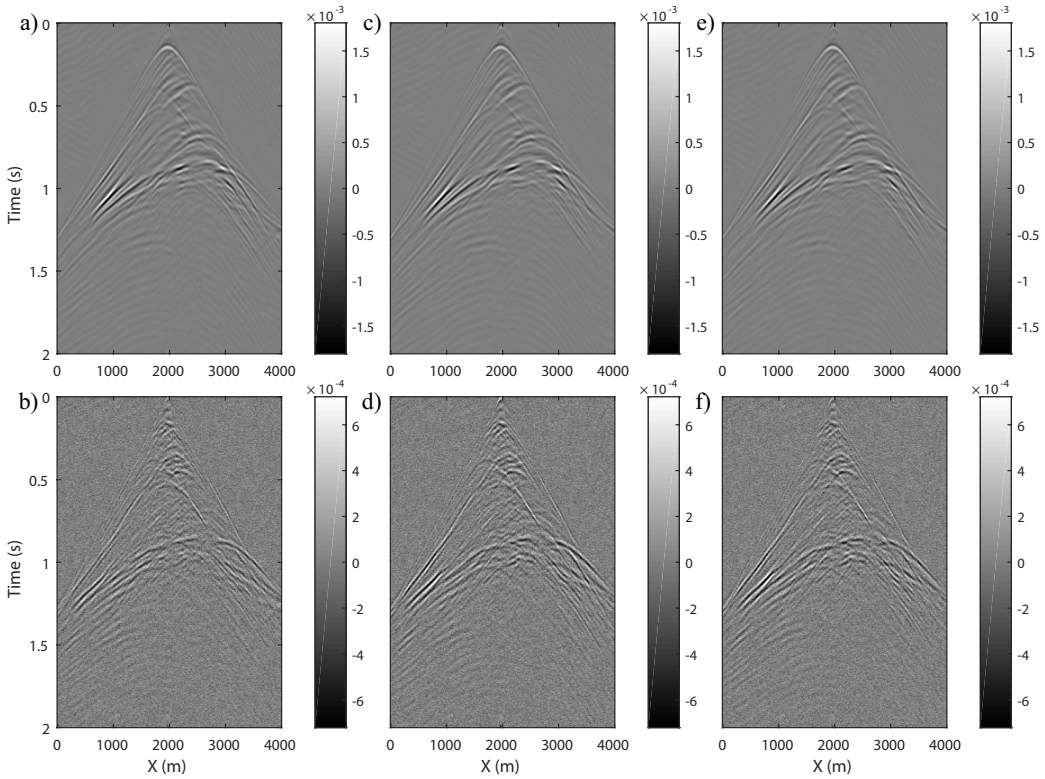

Figure 14: JMI example: the modeled data generated at $X = 2000$ m and the corresponding difference with the observed data based on the inverted velocity and reflectivity using (a, b) regular JMI without any regularization, (c, d) regular JMI with conventional TV , and (e, f) regular JMI with directional TV.

gradually to make the inversion converge. λ is a constant which depends on the scale of the data. We can set a proper λ to make sure around $60\% - 70\%$ of the energy is passed through the shrinkage step in Algorithm 1, in order to improve the stability of the algorithm. Regarding the weights on the dominant direction and its perpendicular direction of gradients, it depends on the accuracy of the estimated dip field and the bias of the subsurface structures. Usually, $\alpha_1 : \alpha_2 = 2 : 1$ is a safe choice. In this paper, we use $\alpha_1 : \alpha_2 = 3 : 1$ for both examples, which puts more weight on the dominant spatial direction of the velocity gradient, because the structures of the Marmousi model are quite tilted and biased.

Regarding the calculation efficiency of JMI, JMI is more cost-effective than FWI. First, it doesn't require a good initial model to start with due to its linearization; Second, it is implemented in the frequency domain and no finite-difference-based method is used, therefore the horizontal and vertical grid size do not have to satisfy a frequency dispersion condition, but are defined by the spatial Nyquist criterion. For instance, in the JMI example, the frequency range is upto 40 Hz and the chosen horizontal and vertical grid size is 20 m and 10 m, respectively.

CONCLUSIONS

Full waveform inversion (FWI) and Joint migration inversion (JMI) with the directional total variation (TV) has been demonstrated to be a more effective method than the alternatives (i.e., FWI/JMI without regularization, with the conventional TV, or directional Laplacian smoothing). We design the directional TV based on the dip field calculated from an initial image. By considering the local structural directions of the spatial gradient and their weights according to the local dip, the proposed method achieves a better result compared to FWI/JMI without regularization or with conventional TV. Finally, it can be concluded that the impact of directional TV is larger for FWI than for JMI, because in JMI the velocity model only explains the propagation effects and, thereby, makes it less sensitive to the details in the velocity model.

ACKNOWLEDGEMENTS

S. Qu and E. Verschuur thank the sponsors of the Delphi consortium for their support. Y. Chen is supported by the Thousand Youth Talents Plan of China, and the Starting Funds from Zhejiang University.

REFERENCES

Anagaw, A. Y., and M. D. Sacchi, 2011, Full waveform inversion with total variation regularization: Presented at the Recovery-CSPG CSEG CWLS Convention.

Bai, M., J. Wu, S. Zu, and W. Chen, 2018, A structural rank reduction operator for removing artifacts in least-squares reverse time migration: Computers and Geosciences, **117**, 9–20.

Bayram, I., and M. E. Kamasak, 2012, Directional total variation: IEEE Signal Processing Letters, **19**, 781–784.

Berkhout, A. J., 2012, Combining full wavefield migration and full waveform inversion, a glance into the future of seismic imaging: Geophysics, **77**, S43–S50.

———, 2014a, Review paper: An outlook on the future of seismic imaging, Part I: forward and reverse modelling: Geoph. Prosp., **62**, 911–930.

———, 2014b, Review paper: An outlook on the future of seismic imaging, Part II: Full-wavefield migration: Geoph. Prosp., **62**, 931–949.

———, 2014c, Review paper: An outlook on the future of seismic imaging, Part III: Joint migration inversion: Geoph. Prosp., **62**, 950–971.

Chen, Y., H. Chen, K. Xiang, and X. Chen, 2016, Geological structure guided well log interpolation for high-fidelity full waveform inversion: Geophysical Journal International, **207**, no. 2, 1313–1331.

———, 2017, Preserving the discontinuities in least-squares reverse time migration of simultaneous-source data: Geophysics, **82**, no. 3, S185–S196.

Chen, Y., W. Huang, Y. Zhou, W. Liu, and D. Zhang, 2020, Plane-wave orthogonal polynomial transform for amplitude-preserving noise attenuation: Geophysical Journal International, **222**, 17891804.

Davydenko, M., and D. J. Verschuur, 2012, Demonstration of full wavefield migration in 2D subsurface models: 74th Ann. Internat. Mtg., EAGE, Expanded Abstracts, Eur. Ass. of Geosc. and Eng., Expanded abstracts, P275.

———, 2017, Full-wavefield migration: using surface and internal multiples in imaging: Geophysical Prospecting, **65**, 7–21.

Donoho, D. L., 1995, De-noising by soft-thresholding: IEEE transactions on information theory, **41**, 613–627.

Fomel, S., 2002, Applications of plane-wave destruction filters: Geophysics, **67**, 1946–1960.

Fu, L., and W. W. Symes, 2017a, An adaptive multiscale algorithm for efficient extended waveform inversion: Geophysics, **82**, R183–R197.

———, 2017b, A discrepancy based penalty method for extended waveform inversion: Geophysics, **82**, 1–69.

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

Guitton, A., G. Ayeni, and E. Díaz, 2012, Constrained full-waveform inversion by model reparameterization 1: Geophysics, **77**, R117–R127.

Hu, W., A. Abubakar, and T. M. Habashy, 2009, Simultaneous multifrequency inversion of full-waveform seismic data: Geophysics, **74**, R1–R14.

Qu, S., and D. J. Verschuur, 2016a, Getting accurate time-lapse information using geology-constrained simultaneous joint migration-inversion, *in* SEG Technical Program Expanded Abstracts 2016: Society of Exploration Geophysicists, 5451–5456.

———, 2016b, Simultaneous time-lapse imaging via joint migration and inversion, *in* 78th EAGE Conference and Exhibition 2016: Eur. Ass. of Geosc. and Eng.,

Expanded abstracts.

———, 2017a, Simultaneous joint migration inversion for accurate time-lapse analysis of sparse monitor surveys, *in* First EAGE Workshop on Practical Reservoir Monitoring: Eur. Ass. of Geosc. and Eng., Expanded abstracts.

———, 2017b, Simultaneous joint migration inversion for semicontinuous time-lapse seismic data, *in* SEG Technical Program Expanded Abstracts 2016: Society of Exploration Geophysicists, 5808–5813.

Qu, S., D. J. Verschuur, and Y. Chen, 2017, Full waveform inversion using an automatic directional total variation constraint, *in* 79th EAGE Conference and Exhibition 2017: Eur. Ass. of Geosc. and Eng., Expanded abstracts.

Staal, X. R., 2015, Combined imaging and velocity estimation by joint migration inversion, *in* Ph.D thesis: Delft University of Technology.

Staal, X. R., and D. J. Verschuur, 2012, Velocity estimation using internal multiples: 82nd Ann. Internat. Mtg., SEG, Expanded Abstracts, Soc. Expl. Geophys., Expanded abstracts, pp. 1–5.

———, 2013, Joint migration inversion, imaging including all multiples with automatic velocity update: 75th Ann. Internat. Mtg., EAGE, Expanded Abstracts, Eur. Ass. of Geosc. and Eng., Expanded abstracts, Tu–02–16, 4pp.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Verschuur, D. J., X. R. Staal, and A. J. Berkhout, 2016, Joint migration inversion: Simultaneous determination of velocity fields and depth images using all orders of scattering: The Leading Edge, **35**, 1037–1046.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC127 – WCC152.

Wu, J., and M. Bai, 2018, Incoherent dictionary learning for reducing crosstalk noise in least-squares reverse time migration: Computers and Geosciences, **114**, 11–21.