Efficient Full Waveform Inversion Subject To A Total Variation Constraint

Yudai INADA, Shingo TAKEMOTO, and Shunsuke ONO
Institute of Science Tokyo

Abstract Full waveform inversion (FWI) aims to reconstruct subsurface properties from observed seismic data. Since FWI is an ill-posed inverse problem, appropriate regularizations or constraints are effective approaches to achieve accurate reconstruction. The total variation (TV) -type regularization or constraint is widely known as a powerful prior that models the piecewise smoothness of subsurface properties. However, the optimization problem of the TV-type regularized or constrained FWI is difficult to solve due to the nonlinearity of the observation process and the non-smoothness of the TV-type regularization or constraint. Conventional approaches to solve the FWI problem rely on inner loops and/or approximations, resulting in high computational cost and/or inappropriate solutions. In this paper, we develop an efficient algorithm with neither inner loops nor approximations to solve the FWI problem based on a primal-dual splitting method. We also demonstrate the effectiveness of the proposed method through experiments using the SEG/EAGE Salt and Overthrust Models.

1 Introduction

Full waveform inversion (FWI) [1, 2] aims to reconstruct subsurface properties from seismic data observed at multiple points. These subsurface properties are used for geological research and resource exploration, including deposits of gas, oil, mineral, and groundwater [2–4]. In addition to geological fields, FWI has also been successfully applied to non-destructive testing in the medical and industrial fields [5, 6].

In FWI, the observation process of seismic data from subsurface properties is nonlinear and complex [2], making an analytic inverse transformation impossible. An effective approach to address this is to formulate FWI as an optimization problem [1, 7–12], such as minimizing the squared error between observed and modeled seismic data. Since FWI is an ill-posed inverse problem, many methods have been proposed that incorporate Tikhonov [13] and Total Variation (TV)-type [14,15] regularizations to capture the piecewise smoothness of subsurface properties [16–20]. However, these regularizations require careful tuning of balance parameters between FWI objective value and these regularization values. Instead of the regularizations, we focus on incorporating TV as a constraint into the FWI problem [21–24]. In contrast to the TV regularizations, the TV constraint has the advantage that its parameter can be determined independently of the objective function value like [25–29]. Specifically, the parameter of TV constraint can be determined based only on prior knowledge of the subsurface properties [30]. This makes the formulation and the reconstructed subsurface properties easier to interpret, which is beneficial for practical applications.

However, the TV-constrained FWI problem is difficult to solve not only because of the nonlinearity of the observation process, but also because of the non-smoothness of the TV constraint. To address this, conventional methods [21–24] adjust the objective variable to satisfy the constraint at each step of an iterative optimization algorithm. This requires an inner loop, which results in high computational cost. In addition, the methods rely on approximations, such as treating nonlinear transformations as linear and satisfying constraints outside the optimization method. If the TV-constrained FWI problem could be solved with neither inner loops nor approximations, more efficient and accurate reconstructions of subsurface properties would be possible.

In this paper, we propose a novel algorithm to solve the TV-constrained FWI problem based on the primal-dual splitting (PDS) method. Our algorithm addresses the challenges posed by both the nonlinearity of the observation process and the non-smoothness of the TV constraint without approximations, resulting in a more accurate reconstruction. Furthermore, it handles the constraint without inner loops to significantly enhance computational efficiency. We demonstrate that our algorithm efficiently handles the constraint while achieving accurate reconstruction.

2 Preliminaries

2.1 Mathematical Tools

Throughout this paper, we denote vectors and matrices by bold lowercase letters (e.g., $\boldsymbol{x}$) and bold uppercase letters (e.g., $\boldsymbol{X}$), respectively.

For $\boldsymbol{x} \in \mathbb{R}^{\mathrm{N}}$, the mixed $l_{1,2}$ norm is defined as follows:

$$\|\boldsymbol{x}\|_{1,2} \coloneqq \sum_{\mathfrak{g} \in \mathfrak{G}} \|\boldsymbol{x}_{\mathfrak{g}}\|_2, \tag{1}$$

where $\mathfrak{G}$ is a set of disjoint index sets, and $\boldsymbol{x}_{\mathfrak{g}}$ is the subvector of $\boldsymbol{x}$ indexed by $\mathfrak{g}$.

For $\boldsymbol{x} \in \mathbb{R}^{\mathrm{N}}$, the total variation (TV) [14] is defined as follows:

$$\mathrm{TV}(\boldsymbol{x}) \coloneqq \|\mathbf{D}\boldsymbol{x}\|_{1,2} = \sum_{i=1}^{\mathrm{N}} \sqrt{d_{h,i}^2 + d_{v,i}^2}, \tag{2}$$

where $d_{h,i}$ and $d_{v,i}$ are the horizontal and vertical differences of the i-th element of $\boldsymbol{x}$, respectively, when the vector $\boldsymbol{x}$ is considered as a matrix.

2.2 Proximal Tools

For $\gamma > 0$, $f \in \mathbb{R}^N \to \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^N$, the proximity operator is defined as follows:

$$\mathrm{prox}_{\gamma f}(\boldsymbol{x}) \coloneqq \underset{\boldsymbol{y} \in \mathbb{R}^N}{\mathrm{argmin}} \left\{ f(\boldsymbol{y}) + \frac{1}{2\gamma}\|\boldsymbol{y} - \boldsymbol{x}\|_2^2 \right\}. \tag{3}$$

For a proper lower-semicontinuous convex function $f \in \mathbb{R}^N \to \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^N$, the convex conjugate function f^* is defined as follows:

$$f^*(\boldsymbol{x}) \coloneqq \sup_{\boldsymbol{y} \in \mathbb{R}^N} \left\{ \boldsymbol{y}^T \boldsymbol{x} - f(\boldsymbol{y}) \right\}. \tag{4}$$

The proximity operator for the convex conjugate function is expressed as follows [31, Theorem 3.1 (ii)]:

$$\mathrm{prox}_{\gamma f^*}(\boldsymbol{x}) = \boldsymbol{x} - \gamma \mathrm{prox}_{\frac{1}{\gamma} f}\left(\frac{1}{\gamma}\boldsymbol{x}\right). \tag{5}$$

For a set $C \subset \mathbb{R}^N$ and $\boldsymbol{x} \in \mathbb{R}^N$, the indicator function ι_C is defined as follows:

$$\iota_C(\boldsymbol{x}) \coloneqq \begin{cases} 0 & \text{if } \boldsymbol{x} \in C, \\ \infty & \text{otherwise.} \end{cases} \tag{6}$$

The proximity operator of ι_C is the projection onto C, given by

$$\mathrm{prox}_{\gamma \iota_C}(\boldsymbol{x}) = P_C(\boldsymbol{x}) \coloneqq \underset{\boldsymbol{y} \in C}{\mathrm{argmin}} \|\boldsymbol{y} - \boldsymbol{x}\|_2. \tag{7}$$

2.3 Primal-Dual Splitting Algorithm

The Primal-Dual Splitting algorithm (PDS) [32–35] is applied to the following problem:

$$\min_{\boldsymbol{x} \in \mathbb{R}^N} \left\{ f(\boldsymbol{x}) + g(\boldsymbol{x}) + h(\boldsymbol{L}\boldsymbol{x}) \right\}, \tag{8}$$

where $\boldsymbol{L} \in \mathbb{R}^{\mathrm{M} \times \mathrm{N}}$, f is a differentiable convex function and g, h are convex functions whose proximity operator can be computed efficiently.

PDS solves Prob. (8) by iteratively updating the following:

$$\left\lfloor \begin{aligned} \boldsymbol{x}^{(k+1)} &= \mathrm{prox}_{\gamma_1 g}\left(\boldsymbol{x}^{(k)} - \gamma_1(\nabla f(\boldsymbol{x}^{(k)}) + \boldsymbol{L}^T \boldsymbol{y}^{(k)})\right), \\ \boldsymbol{y}^{(k+1)} &= \mathrm{prox}_{\gamma_2 h^*}\left(\boldsymbol{y}^{(k)} + \gamma_2 \boldsymbol{L}(2\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)})\right), \end{aligned} \right. \tag{9}$$

where $\gamma_1, \gamma_2 \in \mathbb{R}$ are step sizes.

2.4 Full Waveform Inversion

Typically, FWI is treated as an optimization problem as follows [1]:

$$\underset{\boldsymbol{m} \in \mathbb{R}^N}{\mathrm{argmin}} \;\; E(\boldsymbol{m}) = \frac{1}{2}\|\boldsymbol{u}_{\mathrm{obs}} - \boldsymbol{u}_{\mathrm{cal}}(\boldsymbol{m})\|_2^2, \tag{10}$$

where $\boldsymbol{m} \in \mathbb{R}^N$ is the velocity model representing subsurface properties, $\boldsymbol{u}_{\mathrm{obs}} \in \mathbb{R}^M$ is the observed seismic data, $\boldsymbol{u}_{\mathrm{cal}} \in \mathbb{R}^N \to \mathbb{R}^M$ is the observation process, and $\boldsymbol{u}_{\mathrm{cal}}(\boldsymbol{m})$ is the modeled seismic data with the velocity model. N is the number of grid points, and M is the number of observed signals.

The observation process $\boldsymbol{u}_{\mathrm{cal}}$ is nonlinear and complex, making it difficult to express analytically. However, the gradient ∇E can be computed numerically by simulating the wave equation using the adjoint-state method [36].

3 Proposed Method

We introduce the TV and box constraint into the FWI problem to achieve more accurate reconstruction. As shown in Fig. 1, the velocity model is piecewise smooth, thus introducing the TV constraint to achieve a more accurate reconstruction. Also, by introducing the box constraint, we prevent the velocity model values from being invalid. As mentioned in the introduction, it is easier to determine parameters if the TV is treated as a constraint rather than a regularization.

The optimization problem of the TV and box constrained FWI is formulated as follows:

$$\underset{\boldsymbol{m} \in \mathbb{R}^N}{\mathrm{argmin}} \;\; E(\boldsymbol{m}) \quad \text{s.t.} \quad \begin{cases} \|\mathbf{D}\boldsymbol{m}\|_{1,2} \leq \alpha, \\ \boldsymbol{m} \in [l, u]^N, \end{cases} \tag{11}$$

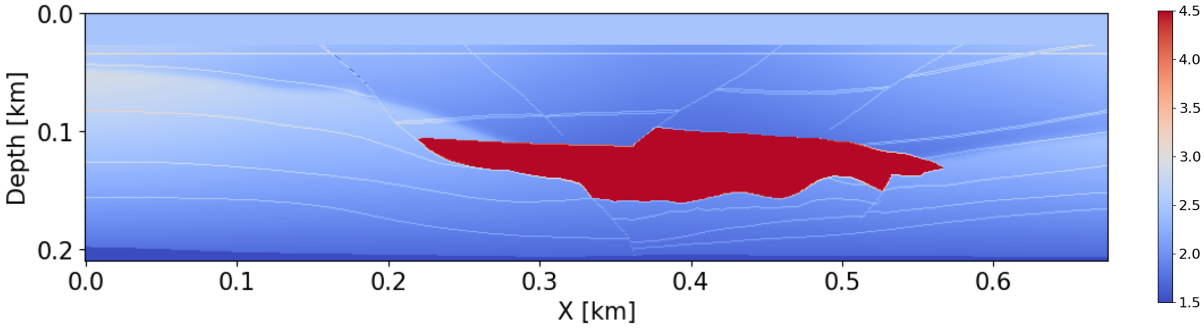


Fig. 1: the velocity model of the Salt [km/s]

Algorithm 1 PDS for (12)

Input: $\boldsymbol{m}^{(0)}, \boldsymbol{y}^{(0)}, \gamma_0 > 0, \gamma_1 > 0$

while A stopping criterion is not satisfied **do**

  $\widetilde{\boldsymbol{m}} \leftarrow \boldsymbol{m}^{(k)} - \gamma_1 \left(\nabla E(\boldsymbol{m}^{(k)}) + \boldsymbol{D}^T \boldsymbol{y}^{(k)}\right)$

  $\boldsymbol{m}^{(k+1)} \leftarrow P_{[l,u]^N} \left(\widetilde{\boldsymbol{m}}\right)$

  $\widetilde{\boldsymbol{y}} \leftarrow \boldsymbol{y}^{(k)} + \gamma_2 \boldsymbol{D} \left(2\boldsymbol{m}^{(k+1)} - \boldsymbol{m}^{(k)}\right)$

  $\boldsymbol{y}^{(k+1)} \leftarrow \widetilde{\boldsymbol{y}} - \gamma_2 P_{\{\|\cdot\|_{1,2} \leq \alpha\}} \left(\frac{1}{\gamma_2} \widetilde{\boldsymbol{y}}\right)$

end while

Output: $\boldsymbol{m}^{(k)}$

where $\alpha \in \mathbb{R}$ is the upper bound of the $l_{1,2}$ norm, and $l, u \in \mathbb{R}$ are the lower and upper bounds of the velocity model values, respectively.

The constraints can be incorporated into the objective function as indicator functions:

$$\underset{\boldsymbol{m} \in \mathbb{R}^N}{\operatorname{argmin}} \; E(\boldsymbol{m}) + \iota_{\|\cdot\|_{1,2} \leq \alpha}(\mathbf{D}\boldsymbol{m}) + \iota_{[l,u]^N}(\boldsymbol{m}). \tag{12}$$

The proximity operator of $\iota_{\|\cdot\|_{1,2} \leq \alpha}$ and $\iota_{[l,u]^N}$ can be computed efficiently. Therefore, these functions of E, $\iota_{[l,u]^N}$ and $\iota_{\|\cdot\|_{1,2} \leq \alpha}$ correspond to f, g and h in (8), respectively, $\mathbf{D}$ is corresponds to $\boldsymbol{L}$, so the problem (12) can be solved using PDS. We show the detailed algorithm in Algorithm 1.

The proximity operators of $\iota_{[l,u]^N}$, $\iota_{\|\cdot\|_{1,2} \leq \alpha}$, that is, the projection onto $[l,u]^N$ and $\|\cdot\|_{1,2} \leq \alpha$ are calculated by

$$P_{[l,u]^N}(\boldsymbol{x}) = \min(\max(\boldsymbol{x}, l), u), \tag{13}$$

$$\left(P_{\{\|\cdot\|_{1,2} \leq \alpha\}}(\boldsymbol{x})\right)_{\mathfrak{g}_i} = \begin{cases} 0 & \text{if } \|\boldsymbol{x}_{\mathfrak{g}_i}\|_2 = 0, \\ \boldsymbol{\beta}_i \frac{\boldsymbol{x}_{\mathfrak{g}_i}}{\|\boldsymbol{x}_{\mathfrak{g}_i}\|_2} & \text{otherwise}, \end{cases} \tag{14}$$

where

$$\boldsymbol{\beta} = P_{\{\|\cdot\|_1 \leq \alpha\}}\left([\|\boldsymbol{x}_{\mathfrak{g}_1}\|_2, \ldots, \|\boldsymbol{x}_{\mathfrak{g}_N}\|_2]^T\right),$$

and in the proposed method, $\mathfrak{g}_i$ is an index set for $\mathbf{D}\boldsymbol{m}$ corresponding to the horizontal and vertical differences of the i-th element of $\boldsymbol{m}$ when the vector $\boldsymbol{m}$ is considered as a matrix.

The proximity operator for the l_1 norm upper bound constraint is expressed as follows [37]:

$$P_{\{\|\cdot\|_1 \leq \alpha\}}(\boldsymbol{x}) = \mathrm{SoftThrethold}(\boldsymbol{x}, \beta), \tag{15}$$

where

$$\boldsymbol{x}_{\mathrm{abs}} = \mathrm{abs}(\boldsymbol{x}),$$
$$\boldsymbol{y} = \mathrm{sort}_{\mathrm{desc}}(\boldsymbol{x}_{\mathrm{abs}}),$$
$$\beta' = \max\left\{ \frac{1}{i}\left(\left(\sum_{j=1}^{i} \boldsymbol{y}_j\right) - \alpha\right) \middle| \, i \in \mathbb{N}, 1 \leq i \leq N \right\},$$
$$\beta = \max\{\beta', 0\}.$$

Our algorithm handles the constraints completely, and it does not require any approximations to incorporate constraints. Furthermore, it can handle constraints without strictly enforcing all of them at each step, so it can be executed efficiently without an inner loop. In fact, the incorporation of the constraints does not significantly increase the overall computational cost, since it is fast enough compared to the ∇E computation, which requires simulation of the wave equation along the time axis.

4 Experiments

4.1 Experimental Setup

To demonstrate the effectiveness of the TV and box constrained FWI, we conducted FWI experiments where we compared with the standard FWI method[1] [1], using the SEG/EAGE Salt and Overthrust Models.

[1]
$$\boldsymbol{m}^{(k+1)} = \boldsymbol{m}^{(k)} - \gamma\left(\nabla E(\boldsymbol{m}^{(k)})\right), \tag{16}$$
where γ is the step size.

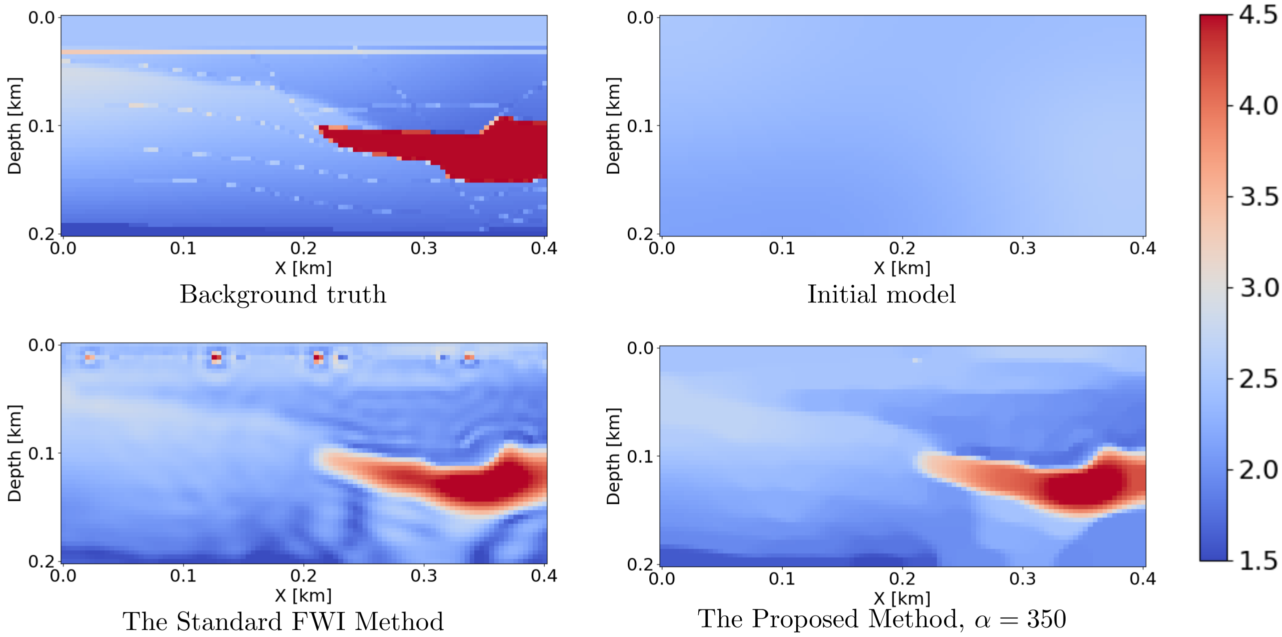


Fig. 2: Velocity models and their corresponding reconstructions.

The velocity model consists of 51 × 101 grid points. The ground truth velocity model is generated by zooming and cropping Fig. 1. The initial velocity model is generated by smoothing the ground truth velocity model with a Gaussian function with a standard deviation of 80. The source waveform is a Ricker wavelet with a peak wavelet frequency of 10 Hz. The number of waveform sources and receivers is 20 and 101, respectively, and they are placed on the surface at equal intervals. The gradient ∇E is computed numerically using the Devito framework [38]. The number of iterations is set to 5000. In the standard FWI, the step size γ is set to 1.0×10^{-4}. In our algorithm, the step size γ_1 and γ_2 are set to 1.0×10^{-4} and 1.0×10^{2}, respectively. The lower and upper bounds of the velocity model l, u are set to 1.5[km/s] and 4.5[km/s], respectively. The experiments are conducted with several α, that is, the upper bound of the $l_{1,2}$ norm.

4.2 Results and Discussion

Fig. 2 shows the ground truth, the initial model, and the reconstructed velocity models using the standard FWI method and the proposed method with the best parameter of $\alpha = 350$. The standard FWI method introduces wave-like artifacts and noise around the source positions, which negatively affects the accuracy of the reconstructed velocity model. In contrast, The enhancements in the proposed method are particularly evident as it successfully eliminates these artifacts and noise, resulting in a more accurate velocity model reconstruction. To quantify this further, we plot the Structural Similarity Index Measure (SSIM) against the number of iterations for both methods in Fig. 3. In this case, the proposed method consistently achieves higher SSIM values than the standard FWI method at every iteration, also indicating enhanced reconstruction accuracy.

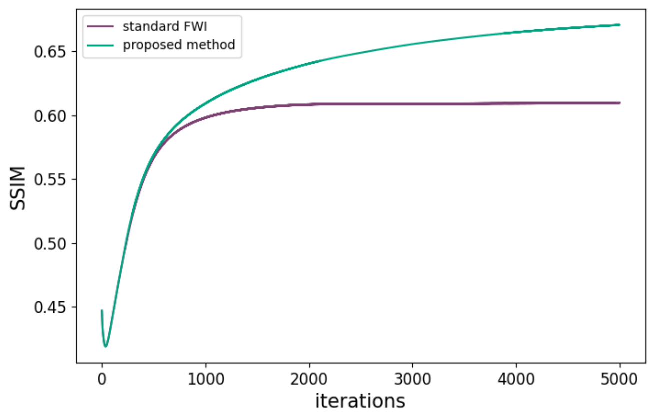
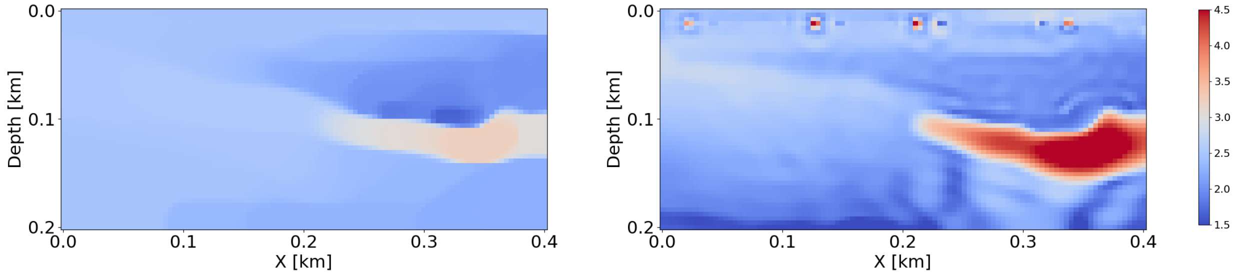


Fig. 3: SSIM against the number of iterations.

To investigate the effect of the parameter α, we show the reconstructed velocity models using the proposed method with multiple α in Fig. 4. When α is as small as 150, the TV constraint is too strong, resulting in an excessively smooth model. Conversely, when α is as large



The Proposed Method, $\alpha = 150$ The Proposed Method, $\alpha = 550$

Fig. 4: Reconstructed velocity models by proposed method with several *alpha*.

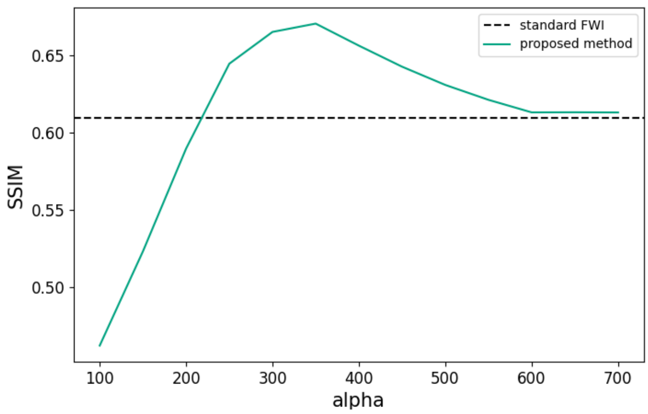


Fig. 5: SSIM against the parameter of alpha.

as 550, the TV constraint is almost meaningless, and the model is similar to that obtained by the standard FWI method. For a more detailed analysis, we plot the SSIM at the last iterations against the parameter α in Fig. 5. As mentioned earlier, the graph shows that when the value of α is small, the last SSIM decreases, and when the value of α is too large, the results become almost the same as the standard FWI method, but not worse, and when the value of *alpha* is appropriate, high SSIM values can be achieved. This demonstrates that the parameter α has a clear and predictable effect on the reconstructed velocity model, allowing for easy adjustment to achieve accurate results.

5 Conclusion

In this paper, we developed an efficient algorithm to solve the TV and box constrained FWI problem based on PDS. Our algorithm handles the constraints completely without approximations, resulting in more accurate reconstructions. Furthermore, by handling the constraints without inner loops, the algorithm significantly enhances computational efficiency. Experimental results demonstrate that our method successfully eliminates wave-like artifacts and noise present in the standard FWI method, resulting in a more accurate velocity model and a higher SSIM value. The predictable effect of the TV constraint parameter allows for easy adjustment to achieve accurate reconstruction, facilitating interpretation of the results and formulation.

Acknowledgement

This work was supported in part by JST PRESTO under Grant JPMJPR21C4 and JST AdCORP under Grant JPMJKB2307, and in part by JSPS KAKENHI under Grant 22H03610, 22H00512, 23H01415, 23K17461, 24K03119, and 24K22291.

References

[1] A. Tarantola, "Inversion of seismic reflection data in the acoustic approximation," *Geophysics*, vol. 49, no. 8, pp. 1259–1266, 1984.

[2] J. Virieux and S. Operto, "An overview of full-waveform inversion in exploration geophysics," *Geophysics*, vol. 74, no. 6, pp. WCC1–WCC26, 2009.

[3] A. Klotzsche, J. van der Kruk, G. A. Meles, J. Doetsch, H. Maurer, and N. Linde, "Full-waveform inversion of cross-hole ground-penetrating radar data to characterize a gravel aquifer close to the thur river, switzerland," *Near surface geophysics*, vol. 8, no. 6, pp. 635–649, 2010.

[4] A. Klotzsche, H. Vereecken, and J. van der Kruk, "Gpr full-waveform inversion of a variably saturated soil-aquifer system," *J. Appl. Geophysics*, vol. 170, p. 103823, 2019.

[5] L. Guasch, O. Calderón Agudo, M.-X. Tang, P. Nachev, and M. Warner, "Full-waveform inversion imaging of the human brain," *NPJ Digit. Med.*, vol. 3, no. 1, p. 28, 2020.

[6] J. Rao, J. Yang, M. Ratassepp, and Z. Fan, "Multi-parameter reconstruction of velocity and density using ultrasonic tomography based on full waveform inversion," *Ultrasonics*, vol. 101, p. 106004, 2020.

[7] C. Shin and D.-J. Min, "Waveform inversion using a logarithmic wavefield," *Geophysics*, vol. 71, no. 3, pp. R31–R42, 2006.

[8] E. Bozdağ, J. Trampert, and J. Tromp, "Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements," *Geophys. J. Int.*, vol. 185, no. 2, pp. 845–870, 2011.

[9] J. Luo and R.-S. Wu, "Seismic envelope inversion: reduction of local minima and noise resistance," *Geophys. Prospecting*, vol. 63, no. 3, pp. 597–614, 2015.

[10] B. Engquist and B. D. Froese, "Application of the wasserstein metric to seismic signals," *arXiv preprint arXiv:1311.4581*, 2013.

[11] L. Metivier, R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, "Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion," *Geophys. J. Int.*, vol. 205, no. 1, pp. 345–377, 2016.

[12] M. Warner and L. Guasch, "Adaptive waveform inversion: Theory," *Geophysics*, vol. 81, no. 6, pp. R429–R445, 2016.

[13] A. Tikhonov, A. Goncharsky, V. Stepanov, and A. Yagola, *Numerical methods for the approximate solution of ill-posed problems on compact sets.* Springer, 1995.

[14] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[15] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 492–526, 2010.

[16] A. Asnaashari, R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, "Regularized seismic full waveform inversion with prior model information," *Geophysics*, vol. 78, no. 2, pp. R25–R36, 2013.

[17] A. Y. Anagaw and M. D. Sacchi, "Full waveform inversion with total variation regularization," in *Recovery-CSPG CSEG CWLS Convention*, 2011, pp. 1–4.

[18] S. Qu, E. Verschuur, and Y. Chen, "Full-waveform inversion and joint migration inversion with an automatic directional total variation constraint," *Geophysics*, vol. 84, no. 2, pp. R175–R183, 2019.

[19] Z. Du, D. Liu, G. Wu, J. Cai, X. Yu, and G. Hu, "A high-order total-variation regularisation method for full-waveform inversion," *J. Geophysics and Eng.*, vol. 18, no. 2, pp. 241–252, 2021.

[20] K. Gao and L. Huang, "Acoustic-and elastic-waveform inversion with total generalized p-variation regularization," *Geophys. J. Int.*, vol. 218, no. 2, pp. 933–957, 2019.

[21] E. Esser, L. Guasch, T. van Leeuwen, A. Y. Aravkin, and F. J. Herrmann, "Total variation regularization strategies in full-waveform inversion," *SIAM J. Imag. Sci.*, vol. 11, no. 1, pp. 376–406, 2018.

[22] E. Esser, L. Guasch, F. J. Herrmann, and M. Warner, "Constrained waveform inversion for automatic salt flooding," *Lead. Edge*, vol. 35, no. 3, pp. 235–239, 2016.

[23] P. Yong, W. Liao, J. Huang, and Z. Li, "Total variation regularization for seismic waveform inversion using an adaptive primal dual hybrid gradient method," *Inverse Problems*, vol. 34, no. 4, p. 045006, 2018.

[24] B. Peters, B. R. Smithyman, and F. J. Herrmann, "Projection methods and applications for seismic nonlinear inverse problems with multiple constraints," *Geophysics*, vol. 84, no. 2, pp. R251–R269, 2019.

[25] M. V. Afonso, J. M. Bioucas-Dias, and M. Figueiredo, "An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, 2011.

[26] G. Chierchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu, "Epigraphical projection and proximal tools for solving constrained convex optimization problems," *Signal Image Video Process.*, vol. 9, no. 8, pp. 1737–1749, 2015.

[27] S. Ono and I. Yamada, "Signal recovery with certain involved convex data-fidelity constraints," *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 6149–6163, 2015.

[28] S. Ono, "Primal-dual plug-and-play image restoration," *IEEE Signal Process. Lett.*, vol. 24, no. 8, pp. 1108–1112, 2017.

[29] S. Ono, "Efficient constrained signal reconstruction by randomized epigraphical projection," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., (ICASSP)*, 2019, pp. 4993–4997.

[30] B. Peters and F. J. Herrmann, "Constraints versus penalties for edge-preserving full-waveform inversion," *Lead. Edge*, vol. 36, no. 1, pp. 94–100, 2017.

[31] P. L. Combettes and N. N. Reyes, "Moreau ' s decomposition in banach spaces," *Math. Program.*, vol. 139, no. 1, pp. 103–114, 2013.

[32] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, pp. 120–145, 2011.

[33] P. L. Combettes and J.-C. Pesquet, "Primal-dual splitting algorithm for solving inclusions with mixtures of composite, lipschitzian, and parallel-sum type monotone operators," *Set-Valued Var. Anal.*, vol. 20, no. 2, pp. 307–330, 2012.

[34] L. Condat, "A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms," *J. Optim. Theory Appl.*, vol. 158, no. 2, pp. 460–479, 2013.

[35] B. C. Vũ, "A splitting algorithm for dual monotone inclusions involving cocoercive operators," *Adv. Comput. Math.*, vol. 38, no. 3, pp. 667–681, 2013.

[36] R.-E. Plessix, "A review of the adjoint-state method for computing the gradient of a functional with geophysical applications," *Geophys. J. Int.*, vol. 167, no. 2, pp. 495–503, 2006.

[37] J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra, "Efficient projections onto the l 1-ball for learning in high dimensions," in *Proceedings of the 25th international conference on Machine learning*, 2008, pp. 272–279.

[38] M. Louboutin, M. Lange, F. Luporini, N. Kukreja, P. A. Witte, F. J. Herrmann, P. Velesko, and G. J. Gorman, "Devito (v3. 1.0): an embedded domain-specific language for finite differences and geophysical exploration," *Geoscientific Model Develop.*, vol. 12, no. 3, pp. 1165–1187, 2019.