# 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. To address this, an effective approach is to formulate FWI as an optimization problem [1], [7]–[12] such as minimizing the squared error between observed and modeled seismic data. Furthermore, 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.

Instead of the regularizations, some methods incorporate TV as a constraint into the FWI problem [21]–[23]. In contrast to the TV regularization, the TV constraint has the advantage that its parameters can be determined based only on prior knowledge of the subsurface properties [24]. This makes the formulation and 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]–[23] 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, approximations are introduced, 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 improve 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:

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

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 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 [25, 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 is defined as follows:

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

Define the proximity operator for the indicator function as $P_C$ as follows.

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

### 2.3 Primal-Dual Splitting Algorithm

The Primal-Dual Splitting (PDS) algorithm [26]–[29] 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.

The PDS algorithm 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]:

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

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, $\boldsymbol{u}_{\mathrm{cal}}$ 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. In general, the velocity model is 2D or 3D grid data, but for simplicity we consider flattened 1D vector.

The observation process $\boldsymbol{u}_{\mathrm{cal}}$ is nonlinear and complex, making it difficult to express analytically. However, the gradient $\nabla E$ can be computed numerically by simulating the wave equation using the adjoint-state method [30].

Therefore, the standard FWI minimizes the objective function and reconstructs the velocity model using the following procedures:

$$\boldsymbol{m}^{(k+1)} = \boldsymbol{m}^{(k)} - \gamma(\nabla E(\boldsymbol{m}^{(k)})), \tag{11}$$

where $\gamma$ is the step size.

## 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 of the Salt is piecewise smooth, thus introducing the TV constraint to achieve a more accurate reconstruction. Also, by introducing the

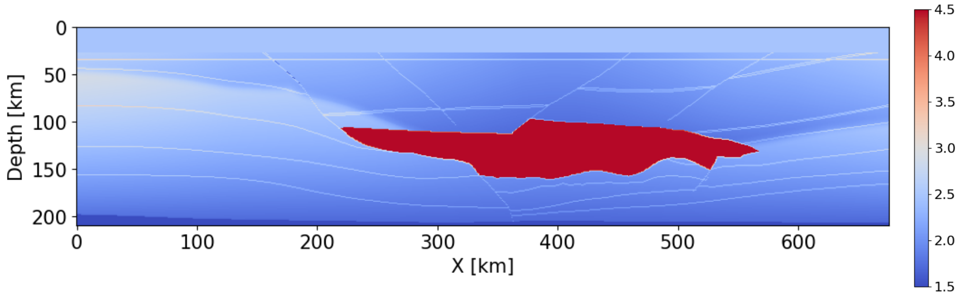

Fig. 1: the velocity model of the Salt [km/s]

box constraint, we can ensure that the velocity model does not take invalid values. 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:

$$\underset{\boldsymbol{m}\in\mathbb{R}^N}{\operatorname{argmin}}\ E(\boldsymbol{m}) \quad \text{s.t. } \|\mathbf{D}\boldsymbol{m}\|_{1,2} \leq \alpha\ ,\ \boldsymbol{m} \in [a,b]^N \quad (12)$$

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

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

$$\underset{\boldsymbol{m}\in\mathbb{R}^N}{\operatorname{argmin}}\ E(\boldsymbol{m}) + \iota_{\|\cdot\|_{1,2}\leq\alpha}(\mathbf{D}\boldsymbol{m}) + \iota_{[a,b]^N}(\boldsymbol{m}) \quad (13)$$

The proximity operator of $\iota_{\|\cdot\|_{1,2}\leq\alpha}$ and $\iota_{[a,b]^N}$ can be computed efficiently. Therefore, these functions of $E$, $\iota_{[a,b]^N}$ and $\iota_{\|\cdot\|_{1,2}\leq\alpha}$ correspond to $f$, $g$ and $h$ in (8), respectively, $\mathbf{D}$ is corresponds to $\boldsymbol{L}$, and the problem (13) can be solved using PDS. The iterative procedures are as follows:

$$\left\lfloor\begin{aligned} \widetilde{\boldsymbol{m}}^{(k+1)} &= \boldsymbol{m}^{(k)} - \gamma_1(\nabla E(\boldsymbol{m}^{(k)}) + \boldsymbol{D}^T\boldsymbol{y}^{(k)}) \\ \boldsymbol{m}^{(k+1)} &= P_{[a,b]^N}(\widetilde{\boldsymbol{m}}^{(k+1)}) \\ \widetilde{\boldsymbol{y}}^{(k+1)} &= \boldsymbol{y}^{(k)} + \gamma_2\boldsymbol{D}(2\boldsymbol{m}^{(k+1)} - \boldsymbol{m}^{(k)}) \\ \boldsymbol{y}^{(k+1)} &= \widetilde{\boldsymbol{y}}^{(k+1)} - \gamma_2 P_{\{\|\cdot\|_{1,2}\leq\alpha\}}(\frac{1}{\gamma_2}\widetilde{\boldsymbol{y}}^{(k+1)}) \end{aligned}\right.$$

The following are the proximity operators of indicator function of the box constraint and the $l_{1,2}$ norm upper bound constraint.

$$P_{[a,b]^N}(\boldsymbol{x}) = \min(\max(\boldsymbol{x}, a), b). \quad (14)$$

$$(P_{\{\|\cdot\|_{1,2}\leq\alpha\}}(\boldsymbol{x}))_{\mathfrak{g}_i} = \begin{cases} 0 & \text{if } \|\boldsymbol{x}_{\mathfrak{g}_i}\|_2 = 0, \\ \boldsymbol{\beta}_i \frac{\boldsymbol{x}_{\mathfrak{g}_i}}{\|\boldsymbol{x}_{\mathfrak{g}_i}\|_2} & \text{otherwise}, \end{cases} \quad (15)$$

where

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

The proximity operator for the $l_1$ norm upper bound constraint is expressed as follows [31]:

$$P_{\{\|\cdot\|_1\leq\alpha\}}(\boldsymbol{x}) = \text{SoftThrethold}(\boldsymbol{x}, \beta), \quad (16)$$

where

$$\begin{aligned} \boldsymbol{x}_{\text{abs}} &= \text{abs}(\boldsymbol{x}), \\ \boldsymbol{y} &= \text{sort}_{\text{desc}}(\boldsymbol{x}_{\text{abs}}), \\ \beta' &= \max\{\frac{1}{i}((\sum_{j=1}^{i}\boldsymbol{y}_j) - \alpha) \mid i = 1, \ldots, N\}, \\ \beta &= \max\{\beta', 0\}. \end{aligned}$$

The computation of $\nabla E$ requires the simulation of the wave equation along the time axis for each grid point. In contrast, the computation of the other parts of the process can be done without time axis simulations. Therefore, the computationally intensive part of the process is primarily the calculation of $\nabla E$, and the introduction of the constraints does not significantly increase the overall computational cost.

## 4 Experiments

### 4.1 Experimental Setup

To demonstrate the effectiveness of the TV and box constrained FWI, we conducted experiments where we compared with the standard FWI with gradient method (11), using the SEG/EAGE Salt and Overthrust Models. The velocity model consists of 51 × 101 grid points. The ground truth velocity model is generated by zooming and cropping Fig.1, and the initial velocity model is generated by smoothing the ground truth velocity model with a Gaussian function with a standard deviation of 80. The number of receivers and source shots are 101 and 20, respectively, and are placed on the surface at equal intervals. The source waveform is a Ricker wavelet with a peak wavelet frequency of 10 Hz. The gradient $\nabla E$ is computed numerically using the Devito framework[32]. The number of iterations is set to 5000.

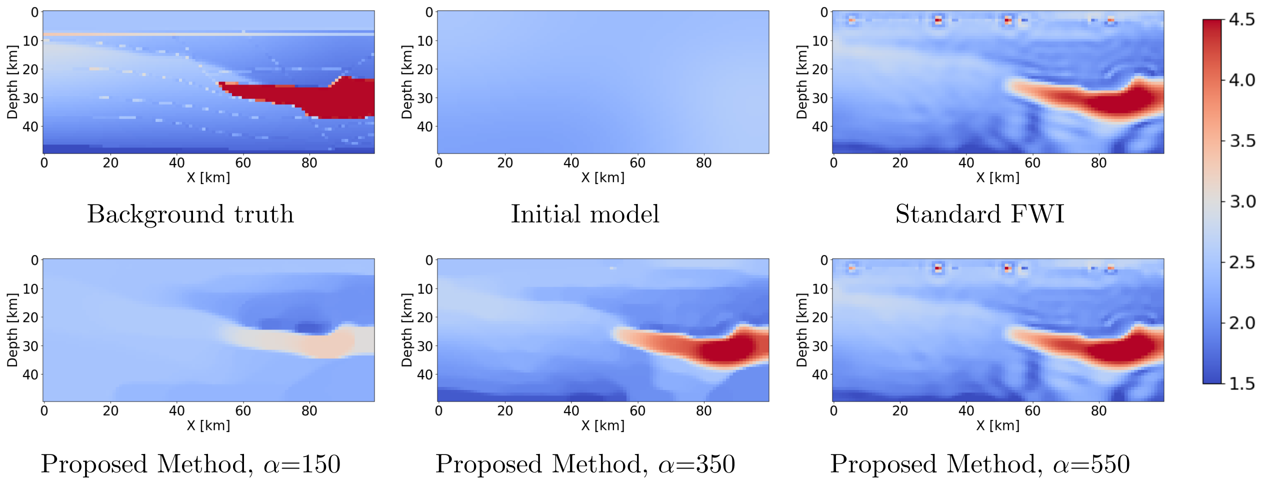


Fig. 2: Velocity models and their corresponding reconstructions.

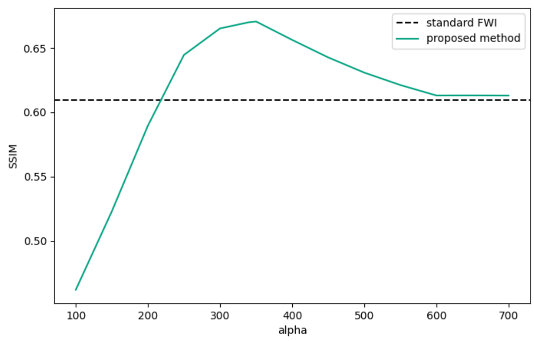


Fig. 3: SSIM against the parameter of alpha.

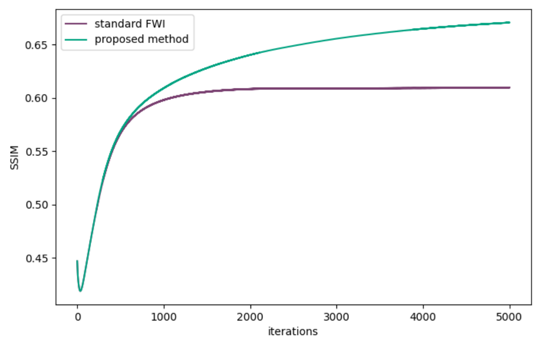


Fig. 4: SSIM against the number of iterations.

In the standard FWI, the step size $\gamma$ is set to $1.0 \times 10^{-4}$. In the TV and box constrained FWI, the step size $\gamma_1$ and $\gamma_2$ are set to $1.0 \times 10^{-4}$ and $1.0 \times 10^{2}$, respectively, and the lower and upper bounds of the velocity model $a$, $b$ are set to 1.5[km/s] and 4.5[km/s], respectively, and experiments are conducted with several $\alpha$ 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 and the TV and box constrained FWI with several $\alpha$. When $\alpha$ is small, such as 150, the TV constraint is too strong, resulting in an excessively smooth model. Conversely, when $\alpha$ is large, such as 550, the TV constraint is almost meaningless, and a model similar to the standard FWI is obtained. When $\alpha$ is appropriate, such as 350, the TV constraint successfully eliminates wave-like artifacts and noise that appear at the source positions, resulting in a more accurate velocity model reconstruction.

In Fig.3, we plot the Structural Similarity Index Measure (SSIM) at the last iterations against the parameter $\alpha$. As mentioned earlier, the graph shows that when the value of $\alpha$ is small, the last SSIM decreases, and when the value of $\alpha$ is too large, the results become almost the same as the standard FWI, but not worse. However, with an appropriately chosen $\alpha$, the graph shows that high SSIM values can be achieved.

In Fig.4, we plot the SSIM against the number of iterations for both methods with $\alpha = 350$. With appropriate parameters, the proposed method consistently achieves higher SSIM values than the standard FWI at every iteration, indicating improved reconstruction accuracy.

## 5 Conclusion

In this paper, we developed an efficient algorithm to solve the TV and box constrained FWI problem with neither inner loops nor approximations based on PDS. We demonstrated that the constrained problem can be fully handled within the PDS algorithm. We also demonstrated that the piecewise smoothness by the TV constraint is well represented even when the PDS algorithm is used, and that efficient and accurate reconstruction is possible. Furthermore, the PDS framework allows for the incorporation of more complex constraints and regularizations, making it a valuable tool for future research.

## 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," *Journal of Applied 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 digital medicine*, 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," *Geophysical Journal International*, 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," *Geophysical 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. Métivier, R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, "Measuring the misfit between seismograms using an optimal transport distance: Application to full waveform inversion," *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, 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.-i. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, A. G. Yagola, 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," *Physica 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. Imaging 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," *Journal of Geophysics and Engineering*, 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," *Geophysical Journal International*, 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. Imaging 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," *The Leading 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 and F. J. Herrmann, "Constraints versus penalties for edge-preserving full-waveform inversion," *The Leading Edge*, vol. 36, no. 1, pp. 94–100, 2017.

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

[26] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of mathematical imaging and vision*, vol. 40, pp. 120–145, 2011.

[27] 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 and variational analysis*, vol. 20, no. 2, pp. 307–330, 2012.

[28] L. Condat, "A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms," *Journal of optimization theory and applications*, vol. 158, no. 2, pp. 460–479, 2013.

[29] B. C. Vũ, "A splitting algorithm for dual monotone inclusions involving cocoercive operators," *Advances in Computational Mathematics*, vol. 38, no. 3, pp. 667–681, 2013.

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

[31] 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.

[32] 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 Development*, vol. 12, no. 3, pp. 1165–1187, 2019.