Efficient Full Waveform Inversion Subject To A Total Variation Constraint

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

アブストラクト 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 useful 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 non-linearity of the observation process and the non-smoothness of the TV-type regularization or constraint. Conventional approaches to solve the problem rely on an inner loop and/or approximations, resulting in high computational cost and/or inappropriate solutions. In this paper, we develop an efficient algorithm with neither an inner loop nor approximations to solve the 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 observed seismic data. These properties are used for geological research and resource exploration, including gas, oil, mineral deposits and groundwater [2]–[4]. FWI has also been applied to non-destructive testing [5], [6].

Since the observed seismic data are generated by subsurface properties, FWI is formulated as an inverse problem. However, it is ill-posed, and the quality of the solution depends significantly on the initial model [2]. To achieve accurate reconstruction, several formulations have been proposed [1], [7]–[12]. Typically, FWI is treated as an optimization problem, where the objective is to minimize the squared error between observed and modeled data.

To enhance stability and accuracy, regularization terms are often added to the objective function, such as Tikhonov regularization [13], Total Variation (TV) [14], and Total Generalized Variation (TGV) [15]. For example, studies have used regularization of Tikhonov [16], TV [17], directional TV [18], high-order TV [19], and TGV [20].

The value of the objective function of FWI depends on the observation method such as input signal type and number of observation equipment, because it contains the squared error between the observed data and the modeled data. Consequently, the regularization parameters must be adapted to the observation method. While, adding constraints to the objective function is advantageous because their parameters can be derived only from prior knowledge of the subsurface properties [21]. Therefore, it has been proposed to add the TV constraint to the objective function [22]–[24].

In conventional methods that apply the TV constraint to FWI [22], [23], parameter updates amount in one step of optimization algorithms are adjusted to satisfy the constraints. This often requires an additional optimization, resulting in an inner loop and increased computational cost. In addition, approximations are introduced to incorporate constraints, such as treating non-linear transformations as linear or imposing constraints outside the optimization method.

In this paper, we develop an efficient algorithm based on a primal-dual splitting method to solve the TV-constrained FWI problem with neither an inner loop nor approximations. ...嬉しさを詳細に書く. We also demonstrate the effectiveness of the proposed method through experiments using the SEG/EAGE Salt and Overthrust Models.

2 PRELIMINARIES

2.1 Mathematical Tools

Throughout this paper, we denote vector and matrix by boldface lowercase letter (e.g., $\boldsymbol{x}$) and boldface uppercase letter (e.g., $\boldsymbol{X}$), respectively. The operator l_X norm of a vector and matrix is denoted by $\|\cdot\|_X$.

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

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

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:

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

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

For 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:

$$f^*(\boldsymbol{x}) := \sup_{\boldsymbol{y} \in \mathbb{R}^N} \left\{ \boldsymbol{y}^T \boldsymbol{x} - f(\boldsymbol{y}) \right\}. \tag{3}$$

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

$$\iota_C(\boldsymbol{x}) := \begin{cases} 0 & \text{if } \boldsymbol{x} \in C, \\ \infty & \text{otherwise.} \end{cases} \tag{4}$$

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:

$$\mathrm{prox}_{\gamma f}(\boldsymbol{x}) := \underset{\boldsymbol{y} \in \mathbb{R}^N}{\mathrm{argmin}} \left\{ f(\boldsymbol{y}) + \frac{1}{2\gamma} \|\boldsymbol{y} - \boldsymbol{x}\|_2^2 \right\}. \tag{5}$$

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

$$\mathrm{prox}_{\gamma \iota_C(\cdot)}(\boldsymbol{x}) = P_C(\boldsymbol{x}) := \underset{\boldsymbol{y} \in C}{\mathrm{argmin}} \|\boldsymbol{y} - \boldsymbol{x}\|_2. \tag{6}$$

The proximity operator for the specific function used in this paper is given below.

$$\mathrm{prox}_{\gamma f^*}(\boldsymbol{x}) = \boldsymbol{x} - \gamma \mathrm{prox}_{f/\gamma}(\boldsymbol{x}/\gamma). \tag{7}$$

$$P_{[a,b]^N}(\boldsymbol{x}) = \min(\max(\boldsymbol{x}, a), b). \tag{8}$$

$$P_{\{\boldsymbol{a} \mid \|\boldsymbol{a}\|_1 \leq \alpha\}}(\boldsymbol{x}) = \mathrm{SoftThrethold}(\boldsymbol{x}, \beta), \tag{9}$$

where

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

$$P_{\{\boldsymbol{a} \mid \|\boldsymbol{a}\|_{1,2} \leq \alpha\}}(\boldsymbol{x}) = [\boldsymbol{p}_1^T, \ldots, \boldsymbol{p}_N^T]^T, \tag{10}$$

where

$$\begin{aligned} \boldsymbol{p}_i &= \begin{cases} 0 & \text{if } \|\boldsymbol{x}_i\|_2 = 0, \\ \boldsymbol{\beta}_i \frac{\boldsymbol{x}_i}{\|\boldsymbol{x}\|_2} & \text{otherwise,} \end{cases} \\ \boldsymbol{\beta} &= P_{\{\boldsymbol{a} \mid \|\boldsymbol{a}\|_1 \leq \alpha\}}([\|\boldsymbol{x}_1^T\|_2, \ldots, \|\boldsymbol{x}_N^T\|_2]). \end{aligned}$$

The proof of equation (7), (9) and (10) can be found in [25, Theorem 3.1 (ii)], [26], [27] accordingly. It is developed a faster algorithm than (9) [28].

2.2 Primal-Dual Splitting Algorithm

The Primal-Dual Splitting (PDS) Algorithm [29]–[32] is applied to the following problem:

$$\min_{\boldsymbol{x} \in \mathbb{R}^N} \left\{ f(\boldsymbol{x}) + g(\boldsymbol{x}) + h(\boldsymbol{L}\boldsymbol{x}) \right\}, \tag{11}$$

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

The PDS algorithm solve by iteratively updating the following:

$$\begin{aligned} \boldsymbol{x}^{(k+1)} &= \mathrm{prox}_{\gamma_1 g}(\boldsymbol{x}^{(k)} - \gamma_1(\nabla f(\boldsymbol{x}^{(k)}) + \boldsymbol{L}^T \boldsymbol{y}^{(k)})), \\ \boldsymbol{y}^{(k+1)} &= \mathrm{prox}_{\gamma_2 h^*}(\boldsymbol{y}^{(k)} + \gamma_2 \boldsymbol{L}(2\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)})), \end{aligned} \tag{12}$$

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

2.3 Full Waveform Inversion

An objective function of FWI is defined as follows[1]:

$$E(\boldsymbol{m}) = \frac{1}{2} \|\boldsymbol{u}_{\mathrm{obs}} - \boldsymbol{u}_{\mathrm{cal}(\boldsymbol{m})}\|_2^2, \tag{13}$$

where $\boldsymbol{m} \in \mathbb{R}^N$ is velocity model representing subsurface properties, $\boldsymbol{u}_{\mathrm{obs}} \in \mathbb{R}^M$ is the observed seismic data, 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, velocity model is 2D or 3D grid data, but for simplicity we consider flattened 1D vector.

The Normal FWI minimize the objective function and reconstruct the velocity model with following procedures:

$$\boldsymbol{m}^{(k+1)} = \boldsymbol{m}^{(k)} - \gamma(\nabla E(\boldsymbol{m}^{(k)})), \tag{14}$$

where γ is the step size.

The gradient ∇E can be computed numerically using the adjoint-state method [33].

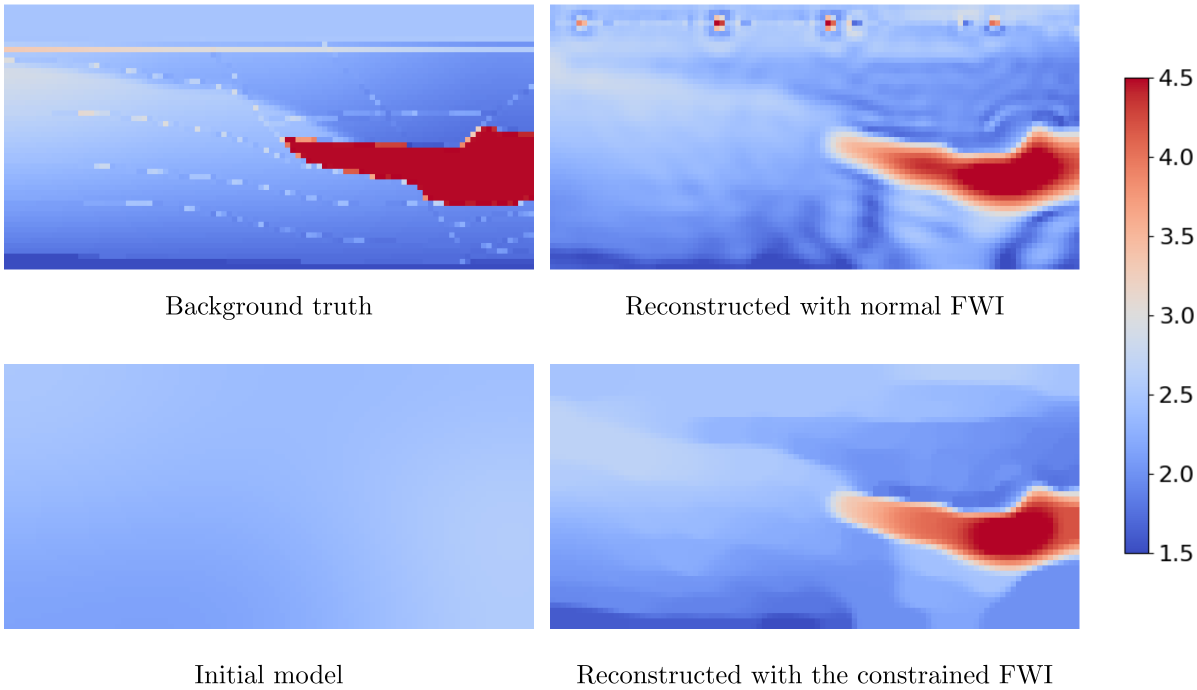

図 1: Velocity models and their corresponding reconstructions.

3 Proposed Method

As shown in fig.1, velocity model can be piecewise smooth. Therefore, we introduce TV constraints to achieve more accurate reconstruction. Also, by introducing a box constraint, we can ensure that the velocity model does not take invalid values.

We minimize the objective function of the TV and box constrained FWI, which is expressed by the following:

$$\operatorname*{argmin}_{\boldsymbol{m}\in\mathbb{R}^N} \ E(\boldsymbol{m}) \quad \text{s.t.} \ \ \|\mathbf{D}\boldsymbol{m}\|_{1,2} \leq \alpha \ , \ \boldsymbol{m} \in [a,b]^N \tag{15}$$

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 bound of the velocity model value, respectively.

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

$$\operatorname*{argmin}_{\boldsymbol{m}\in\mathbb{R}^N} \ E(\boldsymbol{m}) + \iota_{\|\cdot\|_{1,2}\leq\alpha}(\mathbf{D}\boldsymbol{m}) + \iota_{[a,b]^N}(\boldsymbol{m}) \tag{16}$$

As mentioned in section 2.1, $\iota_{\|\cdot\|_{1,2}\leq\alpha}$ and $\iota_{[a,b]^N}$ can be computed efficiently(9)(10). Therefore, these functions of E, $\iota_{[a,b]^N}$ and $\iota_{\|\cdot\|_{1,2}\leq\alpha}$ correspond to f, g and h in (11), respectively, $\boldsymbol{L}$ is corresponds to $\mathbf{D}$, and the problem (16) can be solved using PDS. The iterative procedures are as follows:

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

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 normal FWI with gradient method (14), using the SEG/EAGE Salt and Overthrust Models. The velocity model consists of 101 $\times$ 51 grid points, and the initial velocity model is generated by smoothing the true velocity model with a Gaussian function with a standard deviation of 80. Number of receivers and source shots are 101 and 20, respectively. The source waveform is a Ricker wavelet with a peek wavelet frequency of 10 Hz. The gradient of E is computed numerically using the Devito framework[34]. In normal FWI, the step size γ is set to 1.0×10^{-4}. In TV and box constrained FWI, the step size γ_1 and γ_2 are set to 1.0×10^{-4} and 1.0×10^{2}, respectively, the upper bound of the $l_{1,2}$ norm α is set to 400 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.

4.2 Results and Discussion

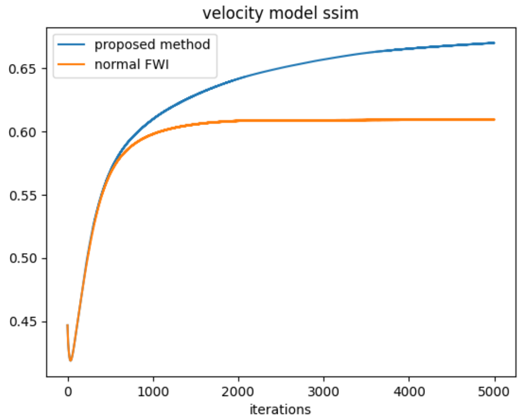

図 2: ssim

5 CONCLUSION

In this paper, we developed an efficient algorithm to solve the TV and box constrained FWI problem with neither an inner loop nor approximations based on PDS. We demonstrated the constrained problem can be fully handled within the PDS algorithm. We also demonstrated 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.

参考文献

[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 Journal on Imaging Sciences*, 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] 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.

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

[23] 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.

[24] 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.

[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] 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.

[27] G. Chierchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu, "Epigraphical projection and proximal tools for solving constrained convex optimization problems," *Signal, Image and Video Processing*, vol. 9, pp. 1737–1749, 2015.

[28] L. Condat, "Fast projection onto the simplex and the l 1 ball," *Mathematical Programming*, vol. 158, no. 1, pp. 575–585, 2016.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.