

Efficient Full Waveform Inversion Subject To A Total Variation Constraint

Yudai INADA
Institute of Science Tokyo
inada.y.af@m.titech.ac.jp

Shingo TAKEMOTO
Institute of Science Tokyo
takemoto.s.af@m.titech.ac.jp

Shunsuke ONO
Institute of Science Tokyo
ono@c.titech.ac.jp

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.

Index Terms—full waveform inversion, total variation, primal-dual splitting method.

I. 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 the 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 linear or quadratic approximation 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, by handling the constraint without inner loops, our algorithm is significantly more efficient compared to existing methods. We demonstrate that our algorithm efficiently handles the constraint while achieving accurate reconstruction.

II. PRELIMINARIES

A. Mathematical Tools

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

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

$$\|\mathbf{x}\|_{1,2} := \sum_{\mathbf{g} \in \mathfrak{G}} \|\mathbf{x}_{\mathbf{g}}\|_2, \quad (1)$$

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

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

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

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

B. Proximal Tools

We denote by $\Gamma_0(\mathbb{R}^N)$ the set of proper lower-semicontinuous convex functions $\mathbb{R}^N \rightarrow (-\infty, \infty]$.

For $\gamma > 0$ and $f \in \Gamma_0(\mathbb{R}^N)$, the proximity operator is defined as follows:

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

For $f \in \Gamma_0(\mathbb{R}^N)$, the convex conjugate function f^* is defined as follows:

$$f^*(\mathbf{x}) := \sup_{\mathbf{y} \in \mathbb{R}^N} \{ \mathbf{y}^T \mathbf{x} - f(\mathbf{y}) \}. \quad (4)$$

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

$$\text{prox}_{\gamma f^*}(\mathbf{x}) = \mathbf{x} - \gamma \text{prox}_{\frac{1}{\gamma} f} \left(\frac{1}{\gamma} \mathbf{x} \right). \quad (5)$$

For a nonempty closed convex set $C \subset \mathbb{R}^N$, the indicator function $\iota_C : \mathbb{R}^N \rightarrow (-\infty, \infty]$ is defined as follows:

$$\iota_C(\mathbf{x}) := \begin{cases} 0 & \text{if } \mathbf{x} \in C, \\ \infty & \text{otherwise.} \end{cases} \quad (6)$$

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

$$\text{prox}_{\gamma \iota_C}(\mathbf{x}) = P_C(\mathbf{x}) := \underset{\mathbf{y} \in C}{\text{argmin}} \|\mathbf{y} - \mathbf{x}\|_2. \quad (7)$$

C. Primal-Dual Splitting Algorithm

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

$$\min_{\mathbf{x} \in \mathbb{R}^N} \{ f(\mathbf{x}) + g(\mathbf{x}) + h(\mathbf{L}\mathbf{x}) \}, \quad (8)$$

where $\mathbf{L} \in \mathbb{R}^{M \times N}$, 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:

$$\begin{cases} \mathbf{x}^{(k+1)} = \text{prox}_{\gamma_1 g} \left(\mathbf{x}^{(k)} - \gamma_1 (\nabla f(\mathbf{x}^{(k)}) + \mathbf{L}^T \mathbf{y}^{(k)}) \right), \\ \mathbf{y}^{(k+1)} = \text{prox}_{\gamma_2 h^*} \left(\mathbf{y}^{(k)} + \gamma_2 \mathbf{L} (2\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) \right), \end{cases} \quad (9)$$

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

D. Full Waveform Inversion

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

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

where $\mathbf{m} \in \mathbb{R}^N$ is the velocity model representing subsurface properties, $\mathbf{u}_{\text{obs}} \in \mathbb{R}^M$ is the observed seismic data, $\mathbf{u}_{\text{cal}} : \mathbb{R}^N \rightarrow \mathbb{R}^M$ is the observation process, and $\mathbf{u}_{\text{cal}}(\mathbf{m})$ is the

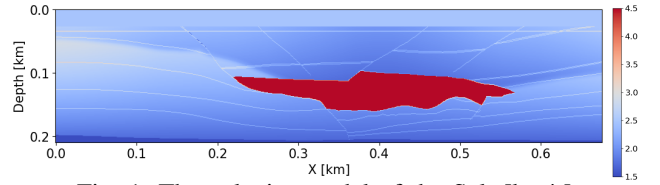


Fig. 1: The velocity model of the Salt [km/s]

modeled seismic data with the velocity model. The number of grid points is N , and the number of observed signals is M .

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

III. 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. The box constraint ensures that the velocity model remains within valid ranges.

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

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

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. By incorporating TV as a constraint, the parameter α can be determined independently of other terms or constraints, which has been highlighted as an advantage in prior works [25–30]. This separation makes it possible to directly control smoothness according to α , providing a clearer interpretation of the reconstructed subsurface properties.

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

$$\underset{\mathbf{m} \in \mathbb{R}^N}{\text{argmin}} E(\mathbf{m}) + \iota_{B_{\text{box}}}(\mathbf{m}) + \iota_{B_{l_{1,2}}}(\mathbf{D}\mathbf{m}), \quad (12)$$

where

$$\begin{aligned} B_{\text{box}} &:= [l, u]^N, \\ B_{l_{1,2}} &:= \{ \|\cdot\|_{1,2} \leq \alpha \}. \end{aligned} \quad (13)$$

The proximity operator of $\iota_{B_{\text{box}}}$ and $\iota_{B_{l_{1,2}}}$ can be computed efficiently. Therefore, these functions of E , $\iota_{B_{\text{box}}}$ and $\iota_{B_{l_{1,2}}}$ correspond to f , g and h in (8), respectively, \mathbf{D} corresponds to \mathbf{L} , so the problem (12) can be solved using PDS. We show the detailed algorithm in Algorithm 1.

The proximity operators of $\iota_{B_{\text{box}}}$ and $B_{l_{1,2}}$, that is, the projection onto B_{box} and $B_{l_{1,2}}$ are calculated by

$$P_{B_{\text{box}}}(\mathbf{x}) = \min(\max(\mathbf{x}, l), u), \quad (14)$$

$$(P_{B_{l_{1,2}}}(\mathbf{x}))_{\mathbf{g}_i} = \begin{cases} 0 & \text{if } \|\mathbf{x}_{\mathbf{g}_i}\|_2 = 0, \\ \beta_i \frac{\mathbf{x}_{\mathbf{g}_i}}{\|\mathbf{x}_{\mathbf{g}_i}\|_2} & \text{otherwise,} \end{cases} \quad (15)$$

Algorithm 1 PDS for (12)

Input: $\mathbf{m}^{(0)}, \mathbf{y}^{(0)}, \gamma_0 > 0, \gamma_1 > 0$
1: **while** A stopping criterion is not satisfied **do**
2: $\tilde{\mathbf{m}} \leftarrow \mathbf{m}^{(k)} - \gamma_1 (\nabla E(\mathbf{m}^{(k)}) + \mathbf{D}^T \mathbf{y}^{(k)})$
3: $\mathbf{m}^{(k+1)} \leftarrow P_{B_{\text{box}}}(\tilde{\mathbf{m}})$
4: $\tilde{\mathbf{y}} \leftarrow \mathbf{y}^{(k)} + \gamma_2 \mathbf{D} (2\mathbf{m}^{(k+1)} - \mathbf{m}^{(k)})$
5: $\mathbf{y}^{(k+1)} \leftarrow \tilde{\mathbf{y}} - \gamma_2 P_{B_{l_{1,2}}} \left(\frac{1}{\gamma_2} \tilde{\mathbf{y}} \right)$
6: **end while**
Output: $\mathbf{m}^{(k)}$

where

$$\beta = P_{\{\|\cdot\|_1 \leq \alpha\}} \left([\|\mathbf{x}_{g_1}\|_2, \dots, \|\mathbf{x}_{g_N}\|_2]^T \right),$$

and g_i is an index set corresponding to the horizontal and vertical differences of the i -th element of \mathbf{m} .

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

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

where

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

Our algorithm can incorporate the constraints without requiring any approximations. Furthermore, since our algorithm can solve the TV and box constrained FWI problem without inner loops, it can be executed efficiently. In fact, the incorporation of the constraints does not significantly increase the overall computational cost, because it is fast enough compared to the ∇E computation, which requires simulation of the wave equation along the time axis.

IV. EXPERIMENTS

A. Experimental Setup

To demonstrate the effectiveness of the TV and box constrained FWI, we conducted FWI experiments where we compared with the standard FWI method¹[1], 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. 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

The standard FWI method uses the following procedures:

$$1 \quad \mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} - \gamma (\nabla E(\mathbf{m}^{(k)})),$$

where γ is the step size.

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 our algorithm, the step sizes 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 using α values ranging from 100 to 700 in steps of 50, representing the upper bound of the $l_{1,2}$ norm. In the standard FWI method, the step size γ is set to 1.0×10^{-4} .

The observed seismic data \mathbf{u}_{obs} is generated using the ground truth velocity model and simulation with Devito. In practice, seismic data often contain noise due to measurement errors and environmental factors. To evaluate the robustness of the proposed method under realistic conditions, we also conducted experiments using noisy observed seismic data, introducing Gaussian noise with mean 0 and variance 1.

B. Results and Discussion

First, we present the experimental results without noise in the observed seismic data. Fig. 2 shows the ground truth, the initial model, and the reconstructed velocity models using the standard FWI method and the proposed methods with $\alpha = 150, 350$, and 550. The best parameter is $\alpha = 350$, where $\alpha = 150$ represents a stronger TV constraint and $\alpha = 550$ represents a weaker one. The standard FWI method generates wave-like artifacts and noise around the source positions. In contrast, our proposed method with the best parameter achieves the accurate velocity model reconstruction without these artifacts and noise. When the TV constraint is too strong, as with $\alpha = 150$, over-smoothing occurs, reducing contrast in the reconstruction. On the other hand, when the TV constraint is too weak, as with $\alpha = 550$, the constraint has little effect, leading to results similar to those of the standard FWI method.

For quantitative evaluation, we plot the Structural Similarity Index Measure (SSIM) against the number of iterations for our proposed method and the standard FWI method in Fig. 3. The proposed method consistently achieves higher SSIM values than the standard FWI method at every iteration, indicating enhanced reconstruction accuracy.

For a more detailed analysis of the TV constraint parameter α , we plot the SSIM of our proposed method against the parameter α and the standard FWI method in Fig. 4. As mentioned earlier, at $\alpha = 350$, the proposed method achieves the highest SSIM. As α is smaller, meaning the TV constraint becomes stronger, the SSIM worsens. On the other hand, when α becomes too large, the SSIM of our proposed method resemble that of standard FWI method. However, thanks to the box constraint, the proposed method still outperforms the standard FWI method. This demonstrates that the parameter α has a clear and predictable effect on the reconstructed velocity model, which can be easily adjusted to achieve accurate results.

Second, we present the experimental results with the noisy observed seismic data. Fig. 2 shows the ground truth, the initial model, and the reconstructed velocity models using

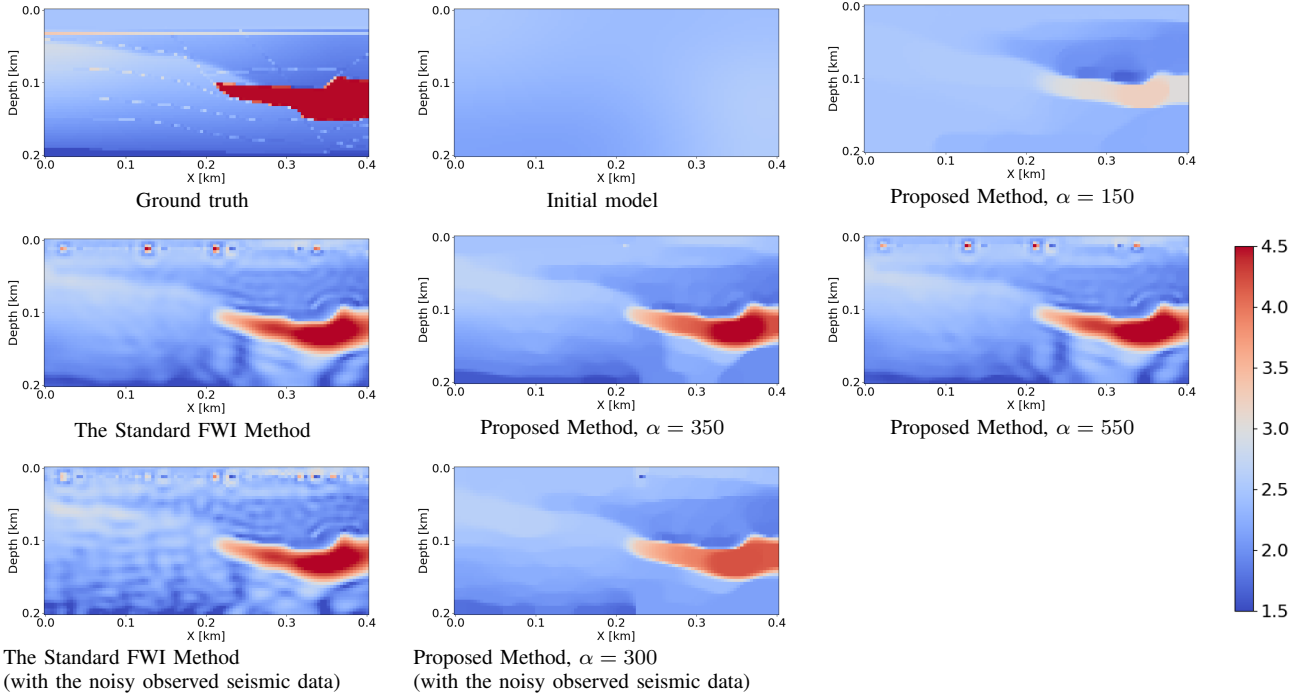


Fig. 2: Velocity models [km/s] and their corresponding reconstructions.

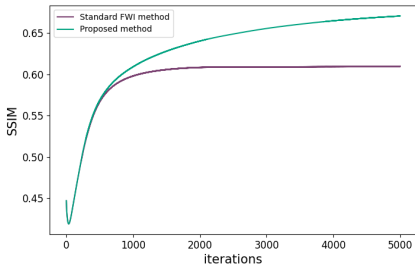


Fig. 3: SSIM against iterations.

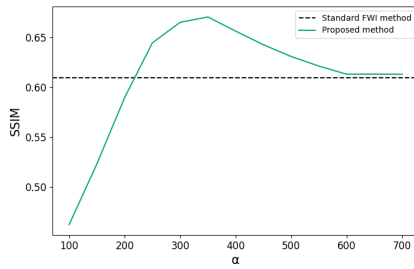


Fig. 4: SSIM against the parameter of α .

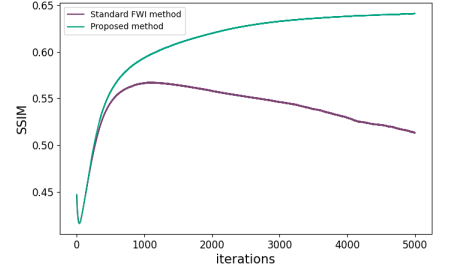


Fig. 5: SSIM against iterations (with the noisy observed seismic data)

the standard FWI method and the proposed methods with $\alpha = 300$, which is the best parameter. In the FWI problem with the noisy observed seismic data, the standard FWI method generates stronger wave-like artifacts and noise around the source positions. In contrast, our proposed method with optimal parameters still achieves accurate velocity model reconstruction without these artifacts and noise, and show robustness to the noise of the observed data.

We also plot the SSIM against the number of iterations for our proposed method and the standard FWI method in Fig. 5. The standard FWI method degrades the reconstruction accuracy after a certain number of iterations. On the other hand, the proposed method maintains a consistently high value without degrading the reconstruction accuracy even after a large number of iterations, indicating enhanced reconstruction accuracy.

The properties of α are the same as when there is no noise in the observed seismic data.

V. CONCLUSION

In this paper, we developed an efficient algorithm to solve the TV and box constrained FWI problem based on PDS. Our algorithm does not require approximations when incorporating the constraints, leading to more accurate reconstructions. Furthermore, the algorithm significantly enhances computational efficiency without inner loops. 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 regardless of the presence of noise in the observed seismic data.

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.