



HAL
open science

Efficient extended-search space full-waveform inversion with unknown source signatures

H Aghamiry, F Mamfoumbi-Ozoumet, A Gholami, Stéphane Operto

► **To cite this version:**

H Aghamiry, F Mamfoumbi-Ozoumet, A Gholami, Stéphane Operto. Efficient extended-search space full-waveform inversion with unknown source signatures. *Geophysical Journal International*, 2021, 227 (1), pp.257-274. 10.1093/gji/ggab202 . hal-03268633

HAL Id: hal-03268633

<https://hal.science/hal-03268633>

Submitted on 30 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Efficient extended-search space full-waveform inversion with unknown source signatures

H.S. Aghamiry¹, F.W. Mamfoumbi-Ozoumet,¹ A. Gholami² and S. Operto¹

¹*Université Côte d'Azur, CNRS, Observatoire de la Côte d'Azur, IRD, Géoazur, 06560 Valbonne, France*

²*Institute of Geophysics, University of Tehran, Tehran 1417466191, Iran. E-mail: agholami@ut.ac.ir*

Accepted 2021 May 17. Received 2021 April 29; in original form 2021 January 20

SUMMARY

Full waveform inversion (FWI) requires an accurate estimation of source signatures. Due to the coupling between the source signatures and the subsurface model, small errors in the former can translate into large errors in the latter. When direct methods are used to solve the forward problem, classical frequency-domain FWI efficiently processes multiple sources for source signature and wavefield estimations once a single lower–upper (LU) decomposition of the wave-equation operator has been performed. However, this efficient FWI formulation is based on the exact solution of the wave equation and hence is highly sensitive to the inaccuracy of the velocity model due to the cycle skipping pathology. Recent extended-space FWI variants tackle this sensitivity issue through a relaxation of the wave equation combined with data assimilation, allowing the wavefields to closely match the data from the first inversion iteration. Then, the subsurface parameters are updated by minimizing the wave-equation violations. When the wavefields and the source signatures are jointly estimated with this approach, the extended wave equation operator becomes source dependent, hence making direct methods and, to a lesser extent, block iterative methods ineffective. In this paper, we propose a simple method to bypass this issue and estimate source signatures efficiently during extended FWI. The proposed method replaces each source with a blended source during each data-assimilated wavefield reconstruction to make the extended wave equation operator source independent. Besides computational efficiency, the additional degrees of freedom introduced by spatially distributing the sources allows for a better signature estimation at the physical location when the velocity model is rough. We implement the source signature estimation with a variable projection method in the recently proposed iteratively refined wavefield reconstruction inversion (IR-WRI) method. Numerical tests on the Marmousi II and 2004 BP salt synthetic models confirm the efficiency and the robustness against velocity model errors of the new method compared to the case where source signatures are known.

Key words: Waveform inversion; Controlled source seismology; Controlled source seismology; Inverse theory; Numerical modelling.

1 INTRODUCTION

Seismic wavefields carry information about subsurface and source, the latter being represented by its location and signature. In controlled source seismic on which this study is focused, the source locations are generally known accurately, while the source signatures are usually unknown and need to be estimated to perform reliable full-waveform inversion (FWI, Tarantola 1984; Pratt *et al.* 1998; Virieux & Operto 2009). Furthermore, it is well acknowledged that the estimation of the source signature is easier in the frequency domain than in the time domain since the time-harmonic wave equation can be solved for each frequency separately (Song *et al.* 1995; Pratt 1999). In frequency-domain seismic modeling, direct methods are the most suitable ones to process a large number of sources efficiently by forward/backward elimination, once a lower–upper (LU) decomposition of the so-called impedance matrix has been performed once (Marfurt 1984). The reader is referred to Operto *et al.* (2015), Amestoy *et al.* (2016), and Operto & Miniussi (2018) for a real 3D frequency-domain FWI case study performed with such forward-modelling engines. When the size of the problem prevents using a direct solver, iterative methods speed-up the processing of multiple right-hand sides with block and recycling methods (Parks *et al.* 2006).

Source signatures may be estimated before FWI or updated jointly with subsurface parameters during FWI iterations. For a fixed velocity model, the source signature estimation can be formulated as a least-squares quadratic data fitting problem (Pratt 1999). The closed-form

expression of the estimated source signature is given by the zero-lag cross-correlation between the calculated and recorded data, scaled by the autocorrelation of the calculated data. In this framework, the source signature estimation can be implemented in the classical reduced-space FWI iterations with two different approaches: In the first, the source signatures and the subsurface parameters are updated in an alternating mode, while the second approach enforces the closed-form expression of the estimated source signature as a function of the subsurface parameters in the objective function (Aravkin & van Leeuwen 2012; Aravkin *et al.* 2012; Li *et al.* 2013) through a variable projection approach (Golub & Pereyra 2003). Plessix & Cao (2011) review these two formulations in the frame of the adjoint-state method and conclude that the variable projection method is more versatile to implement the source signature estimation problem with specific data weighting, while Rickett (2013) showed that the variable projection approach was more resilient to phase errors in the wavelet than the alternated optimization. This source signature estimation does not introduce significant computational overhead in classical FWI since the gradient of the objective function with respect to the subsurface parameters is computed in the same way whether the source signature is available or estimated on the fly during the FWI iterations (Aravkin & van Leeuwen 2012; Rickett 2013).

In its more general form, FWI can be cast as a constrained optimization problem that aims to estimate the wavefields and the subsurface parameters by fitting the recorded data subject that the wave equation is satisfied (Haber *et al.* 2000). Regardless of the source signature estimation issue, it is well acknowledged that FWI is highly nonlinear when the full search space encompassed by the wavefields and the subsurface parameters is projected onto the parameter space after elimination of the wavefield variables. This variable elimination, which is performed by forcing the wavefields to satisfy exactly the wave equation at each FWI iteration, makes FWI prone to cycle skipping as soon as the initial model is not accurate enough to predict recorded traveltimes with an error smaller than half a period (Virieux & Operto 2009). To mitigate the cycle skipping issue, some approaches implement the wave equation as a soft constraint with a penalty method such that the data can be closely matched with inaccurate subsurface models from the early FWI iterations (Abubakar *et al.* 2009; van Leeuwen & Herrmann 2013, 2016). Then, the subsurface model is updated by least-squares minimization of the source residuals generated by the wave equation relaxation. In these extended approaches, the wavefields are the least-squares solution of the overdetermined linear system gathering the wave equation weighted by the penalty parameter and the observation equation relating the simulated wavefield to the data through a sampling operator. In other words, the wavefields are reconstructed with data assimilation. This approach was called wavefield reconstruction inversion (WRI) by van Leeuwen & Herrmann (2013). A variant of WRI, based upon the method of multipliers or augmented Lagrangian method, was proposed by Aghamiry *et al.* (2019b) to increase the convergence rate and decrease the sensitivity of the algorithm to the relaxation (penalty) parameter choice. The augmented Lagrangian method combines a penalty method and a Lagrangian method, where the penalty term is used to implement the initial relaxation of the constraint while the Lagrangian term automatically tunes the sensitivity of the optimization to the constraint in iterations through the gradient ascent update of the Lagrange multipliers with the constraint violations. This method was called iteratively refined (IR)-WRI, where the prefix IR refers to the iterative defect correction action of the Lagrange multipliers.

Recently, Fang *et al.* (2018) tackled the source signature estimation problem in WRI. The formulation of Fang *et al.* (2018) groups together the wavefield and the source signature into a single optimization variable such that WRI can be cast as a separable nonlinear least-squares problem, which can be tackled with the variable projection method. The issue with this approach is that the data assimilation makes the augmented wave-equation (normal) operator source dependent hence making the method expensive since the Cholesky decomposition needs to be performed for each source when a direct method is used. Also, block-processing of multiple right-hand sides is not possible anymore with iterative methods. To bypass this issue, Fang *et al.* (2018) partition the normal-operator matrix into source-dependent and source-independent blocks and apply the block matrix inversion formula (Press *et al.* 2007, page 81) to the partitioned matrix. This block reformulation involves only the inverse of the source-independent normal operator, hence making this approach manageable. The computational overhead generated by the source signature estimation with this approach is however one additional forward/backward elimination per source and frequency, which is not negligible for 3D dense acquisitions (e.g. Amestoy *et al.* 2016), while the number of Cholesky factorization per frequency and iteration remains one.

Another variant of WRI with unknown source signatures was proposed by Huang *et al.* (2018) where WRI is reparametrized in terms of extended sources and subsurface parameters. In this approach, the penalization (or annihilator) term is defined as the distance function from the real source position (Huang *et al.* 2018, Their eq. 9), which means that the source signature estimation is implicitly embedded in the extended source reconstruction. One issue with this approach is related to the presence of the Green functions in the Hessian of the extended source reconstruction subproblem, which makes the normal system very challenging to solve with a good accuracy (Huang *et al.* 2018, their eq. 11). Moreover, they update the subsurface parameters with a variable projection method, which precisely requires an accurate solution of the normal system for the extended sources.

The focus of this paper is to propose an alternative formulation of the source signature estimation problem to that proposed by Fang *et al.* (2018) with the aim to reduce the computational overhead while making the source signature reconstruction potentially more robust when crude background models are used. Our approach relies on a variable projection method, where the closed-form expression of the source signature is projected in the wavefield reconstruction subproblem. To achieve the computational efficiency of the multisource signature and wavefield reconstructions, we reconstruct each individual wavefield with blended sources. This blending makes the normal operator of the wavefield reconstruction subproblems source independent and hence amenable to efficient multisource processing. Although we use source blending, we stress that we estimate one wavefield per physical source thanks to the assimilation of the source-dependent recorded data in the right-hand side (rhs) of the normal system. The source blending implies that, for each reconstructed wavefield, the source signature is a vector of dimension equal to the number of individual sources in the blended source. When the velocity model is accurate, each entry of the

signature vector is zero except the one located at the position of the physical source. Conversely, when the velocity model is inaccurate, the other entries also contribute to decrease the data misfit, although their contribution is much less than the component located at the physical source position. Surprisingly, these additional degrees of freedom provided by the spatially distributed sources help to estimate more accurate source signatures when the velocity model is inaccurate, compared to the case where source blending is not used. Because the extension created by the source blending is artificial, we should correct its effects during the iterations. This is simply performed by keeping only the entry of the signature vectors at the position of the physical sources before proceeding with the parameter updating. Optionally, the wavefields can be also recomputed with these restricted source signatures before parameter updating to remove more efficiently the footprint of the blending at the price of some computational overheads. These different strategies are discussed with several numerical examples.

This paper is organized as follows. In the method section, we first show how the source signature estimation can be combined with the extended wavefield reconstruction subproblem of IR-WRI by variable projection when each source is processed separately. We show that the extended wave equation operator becomes source dependent. The second part of the method section reviews the source blending approach that is used to make the extended wave equation operator source independent and hence amenable to efficient multisource processing with direct methods. Two slightly different algorithms are proposed to implement the method such that the best trade-off between computational efficiency and accuracy is reached. The paper continues with a numerical example section. We first assess the sensitivity of the source signature estimation to several parameters, such as the accuracy of the velocity model, noise and the distance between sources and receivers. Then, we present applications of IR-WRI with the proposed efficient source signature estimation on the Marmousi II model and the 2004 BP salt model and compare the results when the source signatures are known and when they are estimated without the efficient blending strategy.

2 METHOD

2.1 Notation and problem statement

Frequency-domain FWI for multisource acquisition with unknown source signatures can be formulated as the following constrained optimization problem (Aghamiry *et al.* 2019b)

$$\underset{\mathbf{u}_i, s_i, \mathbf{m} \in \mathcal{M}}{\text{minimize}} \quad \mathcal{R}(\mathbf{m}) \quad \text{subject to} \quad \begin{cases} \mathbf{A}(\mathbf{m})\mathbf{u}_i = \phi_i s_i, & i = 1, 2, \dots, n_s, \\ \mathbf{P}\mathbf{u}_i = \mathbf{d}_i, & i = 1, 2, \dots, n_s, \end{cases} \quad (1)$$

where $\mathbf{m} \in \mathbb{R}^{N \times 1}$ is the model parameter vector (squared slowness), N is the number of discretized points of the medium, n_s is the number of sources, $\mathbf{A}(\mathbf{m}) = \Delta + \omega^2 \text{Diag}(\mathbf{m}) \in \mathbb{C}^{N \times N}$ is the Helmholtz operator, ω is the angular frequency, Δ is the Laplacian operator, $\text{Diag}(\bullet)$ denotes a diagonal matrix with the entries of the vector \bullet on its diagonal, $\mathbf{u}_i \in \mathbb{C}^{N \times 1}$ and $\mathbf{d}_i \in \mathbb{C}^{M \times 1}$ denote the wavefield and the recorded data for the i th source, respectively, $\mathbf{P} \in \mathbb{R}^{M \times N}$ is the observation operator and M is the number of receivers. Also, $\phi_i \in \mathbb{R}^{N \times 1}$ is a sparse vector defining the i th source location, and $s_i \in \mathbb{C}$ is the source signature for the i th source at frequency ω . Finally, $\mathcal{R}(\mathbf{m})$ is an appropriate regularization function on the model domain and \mathcal{M} is a convex set defined according to our prior knowledge of \mathbf{m} . For example, if we know the lower and upper bounds on \mathbf{m} then

$$\mathcal{M} = \{\mathbf{m} | \mathbf{m}_{min} \leq \mathbf{m} \leq \mathbf{m}_{max}\}. \quad (2)$$

Here, we limit ourselves to the acoustic approximation of wave-propagation in isotropic media, but the proposed method can be extended for general physics.

IR-WRI solves the constrained problem (1) with the augmented Lagrangian method (or the method of multipliers). The augmented Lagrangian method combines a penalty term to relax the constraints during the early iterations and a Lagrangian term to control how accurately the constraint is satisfied at the convergence point (Nocedal & Wright 2006). In this method, the primal variables and the Lagrange multipliers or dual variables are updated in alternating mode using a primal descent/dual ascent approach. Moreover, to make the computational cost tractable, we update the primal variables \mathbf{u} and \mathbf{m} in an alternating mode in the framework of the alternating-direction method of multipliers (ADMM) (Boyd *et al.* 2010). The reader is referred to Aghamiry *et al.* (2019a, b, 2020a) for more details about the ADMM-based IR-WRI algorithm. In the last two references, $\mathcal{R}(\mathbf{m})$ implements a total-variation (TV) regularization and a hybrid TV+Tikhonov regularization, respectively. Compared to the above references, we extend IR-WRI to update the source signatures jointly with the wavefields during the wavefield reconstruction subproblem through a variable projection.

Beginning with an initial model \mathbf{m}^0 and assume $\hat{\mathbf{d}}_i^0 = \hat{\mathbf{b}}_i^0 = \mathbf{0}$, ADMM solves iteratively the multivariate optimization problem, eq. (1), with alternating directions as (see Boyd *et al.* 2010; Aghamiry *et al.* 2019b, for more details)

$$(\mathbf{u}_i^{k+1}, s_i^{k+1}) = \underset{\mathbf{u}_i, s_i}{\arg \min} \Psi(\mathbf{u}_i, s_i, \mathbf{m}^k, \hat{\mathbf{b}}_i^k, \hat{\mathbf{d}}_i^k), \quad i = 1, 2, \dots, n_s \quad (3a)$$

$$\mathbf{m}^{k+1} = \underset{\mathbf{m} \in \mathcal{M}}{\arg \min} \sum_{i=1}^{n_s} \Psi(\mathbf{u}_i^{k+1}, s_i^{k+1}, \mathbf{m}, \hat{\mathbf{b}}_i^k, \hat{\mathbf{d}}_i^k), \quad (3b)$$

$$\hat{\mathbf{b}}_i^{k+1} = \hat{\mathbf{b}}_i^k + \phi_i s_i^{k+1} - \mathbf{A}(\mathbf{m}^{k+1})\mathbf{u}_i^{k+1}, \quad i = 1, 2, \dots, n_s \quad (3c)$$

$$\hat{\mathbf{d}}_i^{k+1} = \hat{\mathbf{d}}_i^k + \mathbf{d}_i - \mathbf{P}\mathbf{u}_i^{k+1}, \quad i = 1, 2, \dots, n_s \quad (3d)$$

where

$$\Psi(\mathbf{u}_i, s_i, \mathbf{m}, \hat{\mathbf{d}}_i^k, \hat{\mathbf{b}}_i^k) = \mathcal{R}(\mathbf{m}) + \|\mathbf{P}\mathbf{u}_i - \mathbf{d}_i - \hat{\mathbf{d}}_i^k\|_2^2 + \lambda \|\mathbf{A}(\mathbf{m})\mathbf{u}_i - \phi_i s_i - \hat{\mathbf{b}}_i^k\|_2^2 \quad (4)$$

is the scaled form of the augmented Lagrangian (Boyd *et al.* 2010, section 3.1.1), \bullet^k is the value of \bullet at iteration k , the scalar $\lambda > 0$ is the penalty parameter assigned to the wave equation constraint, and $\hat{\mathbf{b}}_i^k \in \mathbb{C}^{N \times 1}$, $\hat{\mathbf{d}}_i^k \in \mathbb{C}^{M \times 1}$ are the scaled Lagrange multipliers, which are updated through a dual ascent scheme by the running sum of the constraint violations (source and data residuals) as shown by eqs (3c)–(3d). Please note that we set the penalty parameter related to the second term ($\|\mathbf{P}\mathbf{u}_i - \mathbf{d}_i - \hat{\mathbf{d}}_i^k\|_2^2$) in eq. (4) equal to one. The penalty parameter λ can be tuned in eq. (4) such that the estimated wavefields approximately fit the observed data from the first iteration at the expense of the accuracy with which the wave equation is satisfied, while the iterative update of the Lagrange multipliers progressively corrects the errors introduced by these penalizations such that both of the observation equation and the wave equation are satisfied at the convergence point with acceptable accuracies.

Here, we focus on the optimization subproblem (3a). The readers are referred to Aghamiry *et al.* (2020a, 2021) for the closed-form expression of the optimization subproblem (3b) with bound constraints and different regularizations and to Aghamiry *et al.* (2020b) for a more robust implementation of this subproblem against velocity model errors with phase retrieval. The governing idea is to use splitting methods to recast the unconstrained subproblem (3b) as a multivariate constrained problem, through the introduction of auxiliary variables. These auxiliary variables are introduced to decouple the ℓ_2 terms from the regularization term such that they can be solved in alternating mode with ADMM. Moreover, a closed form expression of the auxiliary variables is easily obtained by solving the \mathcal{R} -subproblem with proximity operators/denoising algorithms (Combettes & Pesquet 2011; Parikh & Boyd 2013; Venkatakrishnan *et al.* 2013). The main property of such algorithms is that we do not need to solve these subproblems to full convergence. Numerical tests have shown that only a few iterations of the least-squares and \mathcal{R} -subproblem is enough for convergence of the ADMM iteration (Goldstein & Osher 2009). Surprisingly, our numerical tests have shown that only one iteration of this inner iteration is enough for our specific application (Aghamiry *et al.* 2019a, 2020a). Optimization problem (3a) is quadratic in \mathbf{u}_i and s_i and can be written as

$$\underset{\mathbf{u}_i, s_i}{\text{minimize}} \quad \|\mathbf{P}\mathbf{u}_i - \mathbf{d}_i\|_2^2 + \lambda \|\mathbf{A}_k \mathbf{u}_i - \phi_i s_i\|_2^2, \quad (5)$$

where $\mathbf{A}_k \equiv \mathbf{A}(\mathbf{m}^k)$. Fang *et al.* (2018) solve eq. (5) jointly for \mathbf{u}_i and s_i by gathering them in a single vector and solve an $(N+1) \times (N+1)$ linear system (Their eq. 24) instead of the $N \times N$ system of the original WRI (van Leeuwen & Herrmann 2013). Here, we want to solve eq. (5) with the variable projection method (Golub & Pereyra 2003). By taking derivative of eq. (5) with respect to s_i , we get that s_i satisfies $\phi_i^T (\mathbf{A}_k \mathbf{u}_i - s_i \phi_i) = 0$, where \bullet^T denotes the conjugate transpose of \bullet , and

$$s_i = \phi_i^{\text{g}} \mathbf{A}_k \mathbf{u}_i, \quad (6)$$

where $\phi_i^{\text{g}} = (\phi_i^T \phi_i)^{-1} \phi_i^T$ is the generalized inverse of ϕ_i . When the source is on the discretization grid $\phi_i^T \phi_i = 1$ and hence $\phi_i^{\text{g}} = \phi_i^T$. In the rest of this paper, we assume the sources are on the discretization grid just for simplicity of interpretation of the equations.

Substituting the expression of $s_i = \phi_i^T \mathbf{A}_k \mathbf{u}_i$ into eq. (5) leads to a mono-variate optimization problem for the wavefield as

$$\underset{\mathbf{u}_i}{\text{minimize}} \quad \|\mathbf{P}\mathbf{u}_i - \mathbf{d}_i\|_2^2 + \lambda \|\mathbf{Q}_i \mathbf{A}_k \mathbf{u}_i\|_2^2, \quad (7)$$

where $\mathbf{Q}_i = \mathbf{I} - \phi_i \phi_i^T$, and \mathbf{I} is the identity matrix. In eq. (7), the matrix $\phi_i \phi_i^T$ is a diagonal matrix with one nonzero coefficient equal to 1 at the location of the source i , and \mathbf{Q}_i is another diagonal matrix complementary to $\phi_i \phi_i^T$: its diagonal entries equal to 1 except at the source position where the coefficient is zero. The second term in eq. (7) penalizes the predicted source $\mathbf{A}_k \mathbf{u}_i$ at all spatial points except at the physical source location consistently with the elimination (or projection) of s_i , eq. (6), from the optimization variables. Minimization of eq. (7) with respect to \mathbf{u}_i gives (note that $\mathbf{Q}_i^T \mathbf{Q}_i = \mathbf{Q}_i^2 = \mathbf{Q}_i$)

$$\mathbf{u}_i^{k+1} = (\mathbf{P}^T \mathbf{P} + \lambda \mathbf{A}_k^T \mathbf{Q}_i \mathbf{A}_k)^{-1} \mathbf{P}^T \mathbf{d}_i. \quad (8)$$

The explicit relation between the estimated source signature and the data can be obtained as

$$s_i^{k+1} = \phi_i^T \mathbf{A}_k \mathbf{u}_i^{k+1} = \phi_i^T \mathbf{A}_k (\mathbf{P}^T \mathbf{P} + \lambda \mathbf{A}_k^T \mathbf{Q}_i \mathbf{A}_k)^{-1} \mathbf{P}^T \mathbf{d}_i = \phi_i^T (\mathbf{G}_k^T \mathbf{G}_k + \lambda \mathbf{Q}_i)^{-1} \mathbf{G}_k^T \mathbf{d}_i, \quad (9)$$

where $\mathbf{G}_k = \mathbf{P} \mathbf{A}_k^{-1}$ is the rank-deficient forward operator sampling the Green function \mathbf{A}_k^{-1} at receiver positions. Eq. (9) shows that the source signature is estimated by first back propagating the data in time from the receiver positions, i.e. $\mathbf{G}_k^T \mathbf{d}_i$, and then correct the blurring effects induced by the limited bandwidth of the data and the limited spread of the receivers by applying the inverse of the Hessian, that is $(\mathbf{G}_k^T \mathbf{G}_k + \lambda \mathbf{Q}_i)$.

The optimization problem, eq. (7) and its closed-form solution, eq. (8), share some similarities with the extended source reconstruction method described in Huang *et al.* (2018, their eqs 10 and 11) as a source-independent variant of WRI, although there are two differences in their formulation: first, their state variables are the extended sources instead of the extended wavefields, that is $\mathbf{b}_i^e = \mathbf{A}_k \mathbf{u}_i$ where \mathbf{b}_i^e is the i th extended source; Secondly, they used another annihilator function for \mathbf{Q}_i (their eq. 9), which is zero at the source location and linearly increases away from it. The difficulty with the method of Huang *et al.* (2018) is related to the presence of the Green's functions \mathbf{A}_k^{-1} in the Hessian $(\mathbf{G}_k^T \mathbf{G}_k + \lambda \mathbf{Q}_i)$ which makes the system very difficult, if not impossible, to solve exactly.

2.2 Efficient multisource processing with source blending

The proposed method for joint estimation of source signature and data-assimilated wavefield, eq. (7), is robust against velocity model error but it is computationally expensive because the normal operator $(\mathbf{P}^T \mathbf{P} + \lambda \mathbf{A}_k^T \mathbf{Q}_i \mathbf{A}_k)$ is source dependent, hence preventing efficient multisource processing with either direct or iterative solvers. This difficulty also exists in the method of Fang *et al.* (2018) for joint estimation of wavefield and source signature (their eq. 24). Fang *et al.* (2018) propose a block matrix formulation (their eq. 33) to make the computational overhead reasonable. The block partitioning makes the main normal system to be inverted source independent and requires one LU decomposition and two forward-backward substitutions per source at each iteration, while IR-WRI with known sources requires one LU decomposition and one forward-backward substitution per source.

To overcome this computational issue and design a fast and accurate multisource signature and wavefield estimation, we use an alternative approach to Fang *et al.* (2018) by assuming that a virtual blended source generates each wavefield. The true signature of this blended source is the physical source signature at the physical source location and is zero elsewhere. By doing so, each source can be written as $\Phi \mathbf{s}_i$ where $\Phi = [\phi_1 \phi_2 \dots \phi_{n_s}] \in \mathbb{R}^{N \times n_s}$ is a tall matrix including the shifted delta functions (ϕ_i) in its columns and the true $\mathbf{s}_i \in \mathbb{C}^{n_s \times 1}$ is a vector, whose i th entry is the physical source signature s_i while the others are zero. We stress at this stage that this reformulation of the source, $\Phi \mathbf{s}_i$, is equivalent to the original one $\phi_i s_i$ given in eq. (1).

Plugging the new source expression in the objective function (5) and taking the derivative with respect to \mathbf{s}_i gives

$$\mathbf{s}_i = \Phi^T \mathbf{A}_k \mathbf{u}_i. \quad (10)$$

Projecting this expression into eq. (5) and remembering that $\Phi^T \Phi = \mathbf{I}$ give

$$\underset{\mathbf{u}_i}{\text{minimize}} \quad \|\mathbf{P} \mathbf{u}_i - \mathbf{d}_i\|_2^2 + \lambda \|\mathbf{Q} \mathbf{A}_k \mathbf{u}_i\|_2^2, \quad (11)$$

where $\mathbf{Q} = \mathbf{I} - \Phi \Phi^T = \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{Q}_i$.

The closed-form expression of the wavefield \mathbf{u}_i at iteration $k + 1$ is now given by

$$\mathbf{u}_i^{k+1} = (\mathbf{P}^T \mathbf{P} + \lambda \mathbf{A}_k^T \mathbf{Q} \mathbf{A}_k)^{-1} \mathbf{P}^T \mathbf{d}_i. \quad (12)$$

Comparing eqs (12) and (8) shows that the Hessian $(\mathbf{P}^T \mathbf{P} + \lambda \mathbf{A}_k^T \mathbf{Q} \mathbf{A}_k)$ is source independent, hence preserving the benefit of direct solver method to process efficiently multiple sources once one LU factorization has been performed. This solves the computational issue. However, the new parametrization of the source makes the optimization problem blind to the fact that the vector $\mathbf{s}_i \in \mathbb{C}^{n_s \times 1}$ should have only one non zero entry at index i . It gives equal probability to all source positions to reconstruct the wavefield \mathbf{u}_i , eq. (11), hence leading to a blended wavefield. This is highlighted by the fact that $\mathbf{A}_k \mathbf{u}_i$ is potentially dense in the closed-form expression of the reconstructed signature \mathbf{s}_i , eq. (10), when the velocity model is inaccurate. Conversely, when the velocity model is the true one, $\mathbf{A}_{true} \mathbf{u}_i = \phi_i s_i$, is sparse. The blended source assumption gives an extra degree of freedom to the optimization problem to decrease the cost function. Surprisingly, we will show in the *Numerical results* section that the estimated source signature obtained with the source blending approach, eq. (10), is more accurate than the counterpart obtained without blending, eq. (7), when we start IR-WRI from a rough initial velocity model.

The source blending is artificial and its effects (extra non-zero coefficients in \mathbf{s}_i vectors) must be removed during iterations. In the following, we propose two algorithms to achieve this goal. For the sake of compact notations, we recast from now the optimization problem, eq. (1), in matrix form.

Multi sources can be processed efficiently in frequency-domain modeling by gathering them in the rhs of the Helmholtz system in a matrix format. Considering n_s sources, the multi-rhs Helmholtz system is written as $\mathbf{A} \mathbf{U} = \Phi \mathbf{S}$ where $\mathbf{U} = [\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_{n_s}] \in \mathbb{C}^{N \times n_s}$ and $\mathbf{S} \in \mathbb{C}^{n_s \times n_s}$ is a square diagonal matrix with the source signatures on its main diagonal, $\mathbf{S}_{ii} = s_i$. Introducing the data matrix $\mathbf{D} = [\mathbf{d}_1 \mathbf{d}_2 \dots \mathbf{d}_{n_s}] \in \mathbb{C}^{M \times n_s}$ gives the following optimization problem for multisource signature and wavefield estimations

$$\underset{\mathbf{U}, \mathbf{S}}{\text{minimize}} \quad \|\mathbf{P} \mathbf{U} - \mathbf{D}\|_F^2 + \lambda \|\mathbf{A}_k \mathbf{U} - \Phi \mathbf{S}\|_F^2, \quad (13)$$

where $\|\bullet\|_F^2$ denotes the Frobenius norm. Solving eq. (13) with a variable projection method gives

$$\mathbf{S} = \Phi^T \mathbf{A}_k \mathbf{U}, \quad (14)$$

where the diagonal components of \mathbf{S} are dominant, and the off-diagonal coefficients represent the non-physical source components associated with each of the n_s blended sources. Plugging the explicit expression of \mathbf{S} into eq. (13) leads to a mono-variate optimization problem for \mathbf{U} , the closed-form expression of which is given by

$$\mathbf{U}^{k+1} = (\mathbf{P}^T \mathbf{P} + \lambda \mathbf{A}_k^T \mathbf{Q} \mathbf{A}_k)^{-1} \mathbf{P}^T \mathbf{D}. \quad (15)$$

Eq. (15) is the same as eq. (12) but for multidata \mathbf{D} . At this stage, we did not impose any constraint on the structure of \mathbf{S} which is potentially dense when the velocity model is inaccurate. Plugging the expression of \mathbf{U}^{k+1} in eq. (14) gives the explicit expression of \mathbf{S}

$$\mathbf{S}^{k+1} = \Phi^T \mathbf{A}_k \mathbf{U}^{k+1} = \Phi^T \mathbf{A}_k (\mathbf{P}^T \mathbf{P} + \lambda \mathbf{A}_k^T \mathbf{Q} \mathbf{A}_k)^{-1} \mathbf{P}^T \mathbf{D}. \quad (16)$$

Before proceeding with the subsurface parameter updating, we extract the approximate signature of the physical sources as $\mathbf{s} = \text{diag}(\mathbf{S})$, where $\text{diag}(\bullet)$ denotes a vector that contains the diagonal elements of matrix \bullet . The different steps of IR-WRI with source signature estimation are reviewed in Algorithm (1), which begins with an initial model \mathbf{m}^0 and initial dual variables $\hat{\mathbf{B}}^0 = \mathbf{0}$ and $\hat{\mathbf{D}}^0 = \mathbf{0}$.

Algorithm 1 IR-WRI with unknown source signatures

Require: starting point \mathbf{m}^0

- 1: set $\hat{\mathbf{B}}^0 = \mathbf{0}$ and $\hat{\mathbf{D}}^0 = \mathbf{0}$
 - 2: **repeat**
 - 3: $\mathbf{U}^{k+1} = \arg \min_{\mathbf{U}} \|\mathbf{P}\mathbf{U} - \mathbf{D} - \hat{\mathbf{D}}^k\|_{\mathbb{F}}^2 + \lambda \|\mathbf{Q}\mathbf{A}_k \mathbf{U} - \hat{\mathbf{B}}^k\|_{\mathbb{F}}^2$ } Estimation of wavefields
and source signatures.
 - 4: $\mathbf{S}^{k+1} = \text{Diag}(\text{diag}(\Phi^T \mathbf{A}_k \mathbf{U}^{k+1}))$
 - 5: $\mathbf{m}^{k+1} = \arg \min_{\mathbf{m}} \mathcal{R}(\mathbf{m}) + \lambda \|\mathbf{A}(\mathbf{m})\mathbf{U}^{k+1} - \mathbf{S}^{k+1} - \hat{\mathbf{B}}^k\|_{\mathbb{F}}^2$ } Velocity model update.
 - 6: $\hat{\mathbf{B}}^{k+1} = \hat{\mathbf{B}}^k + \mathbf{S}^{k+1} - \mathbf{A}_{k+1} \mathbf{U}^{k+1}$
 - 7: $\hat{\mathbf{D}}^{k+1} = \hat{\mathbf{D}}^k + \mathbf{D} - \mathbf{P}\mathbf{U}^{k+1}$ } Iterative refinement.
 - 8: **until** stopping conditions are satisfied.
-

The \mathbf{U} -subproblem in Algorithm 1 introduce errors in the extended wavefield reconstruction due to the source blending. These errors can be corrected iteratively by the action of the Lagrange multipliers, which are formed by the source residuals computed without source blending (Line 6 of algorithm 1), i.e., with the true spatial point distribution of the sources. This is implemented by using the diagonal components of the source signature matrix, Line 4 of algorithm 1, instead of the whole matrix, eq. (16), to compute the source residuals in Line 6. The rhs correction term $\hat{\mathbf{B}}^k$ in the objective function of the \mathbf{U} -subproblem in Algorithm 1 gathers the running sum of the source residuals of previous iterations. This iterative refinement leads to the error forgetting property discussed by Yin & Osher (2013) in the frame of Bregman iterations, which means that the error correction performed at the current iteration is made independent of the error corrections performed at previous iterations. Here, this iterative solution refinement by rhs updating is necessary to correct three sources of errors: the first results from the fact that each primal subproblem is solved keeping the other primal variable fixed, the second from the fact that we solve a constrained problem with a penalty method keeping the penalty parameter fixed and the third from the non-physical source blending. Another application of iterative refinement in AVO inversion is presented in Gholami *et al.* (2018), where the linearized Zoeppritz equations are used to simplify the primal problem, while the dual problem compensates the linearization-related errors by computing the residuals with the exact Zoeppritz equations.

However, the Algorithm 1 was not able to fully remove the detrimental effects of the source blending and failed to reach the same minimizer as IR-WRI with a known source when we seek to reconstruct a complicated velocity model starting from a rough initial model. This prompts us to propose Algorithm 2 that includes one extra step compared to Algorithm 1 by re-estimating the wavefields (Line 5 of Algorithm 2) with the diagonal restriction of the source signature matrix (Line 4 of Algorithm 2). By doing so, the pollution effects of the non-physical sources are removed from the reconstructed wavefields. The improvement provided by this wavefield refinement is illustrated in the next *Numerical results* section.

Algorithm 2 IR-WRI with unknown sources with wavefield correction

Require: starting point \mathbf{m}^0

- 1: set $\hat{\mathbf{B}}^0 = \mathbf{0}$ and $\hat{\mathbf{D}}^0 = \mathbf{0}$
 - 2: **repeat**
 - 3: $\mathbf{U}^{k+\frac{1}{2}} = \arg \min_{\mathbf{U}} \|\mathbf{P}\mathbf{U} - \mathbf{D} - \hat{\mathbf{D}}^k\|_{\mathbb{F}}^2 + \lambda \|\mathbf{Q}\mathbf{A}_k \mathbf{U} - \hat{\mathbf{B}}^k\|_{\mathbb{F}}^2$ } Source signature estimation.
 - 4: $\mathbf{S}^{k+1} = \text{Diag}(\text{diag}(\Phi^T \mathbf{A}_k \mathbf{U}^{k+\frac{1}{2}}))$
 - 5: $\mathbf{U}^{k+1} = \arg \min_{\mathbf{U}} \|\mathbf{P}\mathbf{U} - \mathbf{D} - \hat{\mathbf{D}}^k\|_{\mathbb{F}}^2 + \lambda \|\mathbf{A}_k \mathbf{U} - \mathbf{S}^{k+1} - \hat{\mathbf{B}}^k\|_{\mathbb{F}}^2$ } Wavefield estimation.
 - 6: $\mathbf{m}^{k+1} = \arg \min_{\mathbf{m}} \mathcal{R}(\mathbf{m}) + \lambda \|\mathbf{A}(\mathbf{m})\mathbf{U}^{k+1} - \mathbf{S}^{k+1} - \hat{\mathbf{B}}^k\|_{\mathbb{F}}^2$ } Velocity model update.
 - 7: $\hat{\mathbf{B}}^{k+1} = \hat{\mathbf{B}}^k + \mathbf{S}^{k+1} - \mathbf{A}_{k+1} \mathbf{U}^{k+1}$
 - 8: $\hat{\mathbf{D}}^{k+1} = \hat{\mathbf{D}}^k + \mathbf{D} - \mathbf{P}\mathbf{U}^{k+1}$ } Iterative refinement.
 - 9: **until** stopping conditions are satisfied.
-

Algorithm 2 requires two LU decompositions at each IR-WRI iteration (one for $\mathbf{U}^{k+\frac{1}{2}}$ and the other one for \mathbf{U}^{k+1}), but still remains much cheaper than the algorithm performing one LU decomposition per source. Numerical tests have shown that the estimated source signature at line 4 of this algorithm is not sensitive to the small changes in the velocity model, and hence it is possible to consider lines 3 and 4 of Algorithm 2 (*source signature estimation* step) only at the first iteration of each frequency batch and still get minimizers very close to the original version of Algorithm 2. Hereafter we refer to this modified algorithm as fast version of Algorithm 2.

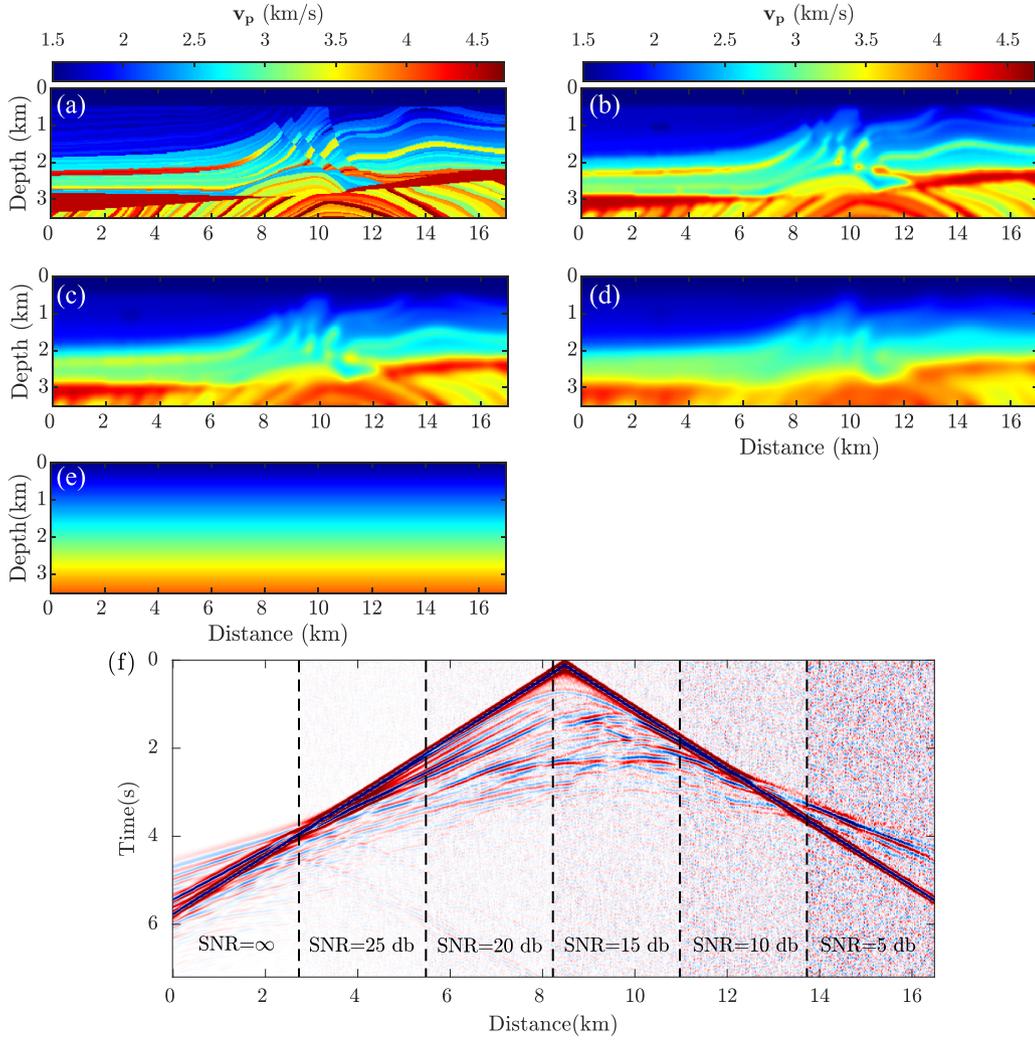


Figure 1. Marmousi II model. (a) True model. (b–e) Smoothed versions of the true model, referred to as models 1 (b), 2 (c), 3 (d) and 4 (e) in the text. (f) Time-domain shot gather for the source located at 8.5 km computed in the true model (a). Each part of this figure is contaminated with different levels of random Gaussian noises to have an insight on the noisy data used in Fig. 5. The signal-to-noise ratio (SNR) of these segments is written in the bottom where SNR is defined in eq. (19).

3 NUMERICAL RESULTS

We first investigate different aspects of the proposed method for efficient source signature estimation in IR-WRI (referred to as *joint* approach in the following) with the Marmousi II (Martin *et al.* 2006) model (Fig. 1a) and compare its performance when each source is processed separately in IR-WRI (referred to as *separate* approach in the following), eq. (3), and when the source signatures are estimated with the method of Pratt (1999) (referred to as *conventional* approach), as

$$\mathbf{S} = (\Phi^T \mathbf{G}^T \mathbf{G} \Phi)^{-1} \Phi^T \mathbf{G}^T \mathbf{D}. \quad (17)$$

Then, we compare the performances of the proposed Algorithms 1 and 2 with those of separate approaches when the source signatures are known (classical IR-WRI) and unknown, eq. (3). We use the Marmousi II model and a scaled version of the left target of the challenging 2004 BP salt model (Billette & Brandsberg-Dahl 2004) for this comparison.

For all the numerical tests, we use a 9-point finite-difference staggered-grid stencil with PML boundary condition (along the model's edges except for the top where the free-surface boundary condition is used) and antilumped mass to solve the Helmholtz equation (Chen *et al.* 2013). Also, for all the numerical tests, λ has been chosen to be a small fraction of the largest eigenvalue of $\mathbf{A}_0^{-T} \mathbf{P}^T \mathbf{P} \mathbf{A}_0^{-1}$ (van Leeuwen & Herrmann 2016).

3.1 Marmousi II test

First we illustrate the performance of the separate, joint and conventional methods for source signature estimation and investigate the robustness of these methods against the accuracy of the initial velocity model, noise in the recorded data, the vertical distance between the

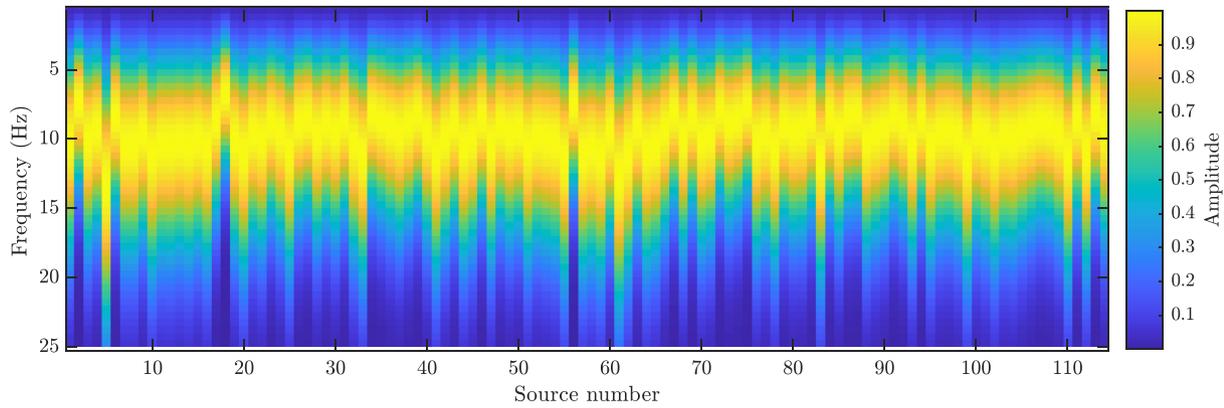


Figure 2. Amplitude spectrum of source signatures used in Marmousi II test.

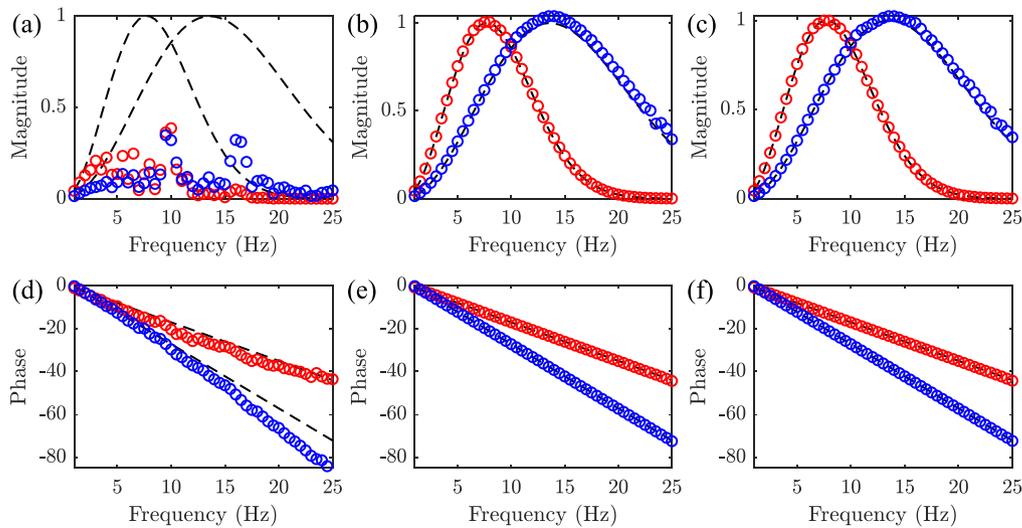


Figure 3. Source signatures (magnitude and unwrapped phase) estimated with different methods for a couple of sources (source numbers 18 and 33 in Fig. 2) when model 4 (Fig. 1e) is used as initial velocity model. (a) Conventional method of Pratt (1999). (b) Separate method. (c) Joint method. The blue and red points show the estimated source signatures for sources 18 and 33, respectively, and the black dashed lines show the true ones. (d–f) Same as (a–c) but for unwrapped phase.

source and the receiver profiles, and the number of sources. We use four initial models for these tests (Figs 1b–e), which are referred to as models 1 to 4. The fixed-spread acquisition contains 114 point sources, the source signatures of which are Ricker wavelets of different central frequency and initial phase (Fig. 2), and a line of receivers spaced 50 m apart at the surface. The central frequency for each source signature is selected randomly between [7–15] Hz, and the peak of each wavelet is centred randomly between [0–0.4] s.

3.1.1 Sensitivity to the background velocity model

First, we put the line of sources at 75 m depth, generate the data with the true velocity model (Fig. 1a) and use the 1-D gradient velocity model (model 4, Fig. 1e) to estimate the source signatures with the conventional method (eq. 17), the separate method (eq. 9), and the joint method (eq. 16). We show the magnitude and phase of the estimated source signatures in Fig. 3 for a couple of sources (source numbers 18 and 33 in Fig. 2). First, the results clearly show the improvement achieved by the relaxed wave-equation methods (the separate and joint methods) (Figs 3b, c, e and f) compared to the conventional method (Figs 3a and d). Secondly, both separate and joint methods estimate accurate source signatures but with a different computational burden (one LU decomposition for the joint method against 114 LU decompositions for the separate method).

To gain more quantitative insights into the accuracy of methods, we plot the relative error (RE) of the source signatures estimated with each method as a function of the source number in Fig. 4. The RE for the estimated source signature is defined as

$$\text{RE} = \frac{\sqrt{\sum_{j=f_1}^{f_n} [\mathbf{x}_j^* - \hat{\mathbf{x}}_j]^2}}{\sqrt{\sum_{j=f_1}^{f_n} \mathbf{x}_j^{*2}}}, \quad (18)$$

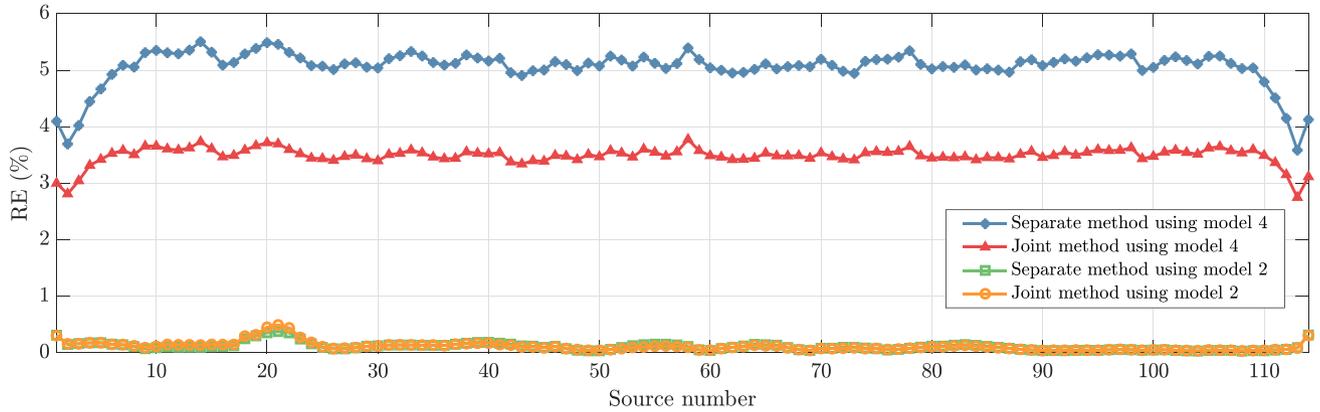


Figure 4. Relative error (RE) of the source signatures estimated with the separate and joint methods when a rough initial model 4 (Fig. 1e) and the kinematically accurate initial model 2 (Fig. 1c) are used.

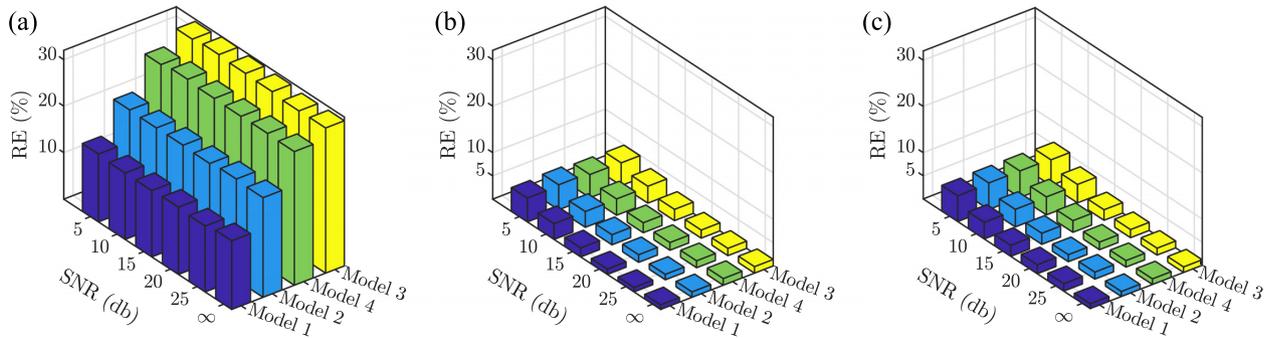


Figure 5. Mean relative error (RE) averaged over all the sources of the source signatures estimated with the (a) conventional, (b) separate and (c) joint methods as a function of SNR of recorded data and initial velocity model.

where \mathbf{x}_j^* and $\hat{\mathbf{x}}_j$ are the true and estimated source signatures at frequency j , respectively, and f_1 and f_n are the minimum and maximum frequencies, respectively. In this figure, we show the RE of the source signatures estimated with the separate and joint methods when the rough velocity model 4 (Fig. 1e) and the kinematically accurate model 2 (Fig. 1c) are used as background velocity model. We do not show the RE of the conventional method in this figure because it is much higher than those obtained with the separate and joint approaches. It is shown that when the initial velocity model is rough, the joint method outperforms the separate method (the blue and red curves in Fig. 4). This is probably due to the extra degrees of freedom available in the joint method compared to the separate counterpart. More precisely, the errors in the estimated source signature induced by the inaccuracy of the initial velocity model are entirely mapped at the physical location of the source in the separate method. In contrast, these errors are distributed across the different components of the blended source in the joint method, hence, providing a better estimation of the source signature at the location of the physical source. On the other hand, as the velocity model becomes more accurate, the separate and joint methods reach the same accuracy for the source signature estimation (the green and orange curves in Fig. 4).

3.1.2 Sensitivity to noise

We continue by assessing the robustness of the methods against the noise in the recorded data and the error in the initial velocity model. For this test, the true data are computed on a finer grid than used for the inversion (half of the grid size of inversion) to avoid the inverse crime. We repeat the same test as before with different initial velocity models and different levels of random Gaussian noises in the data. The time-domain shot gather for the source located at 8.5 km is shown in Fig. 1f when different parts of the gather are contaminated with a different level of noise to gain insight into the magnitude of the noise levels used. The average RE over all the sources for the conventional, separate, and joint methods are shown in Figs 5 a–c, respectively, as functions of the signal-to-noise ratio (SNR) of the data and the initial velocity model. In this paper, the SNR of data is defined as

$$\text{SNR} = 20 \log \left(\frac{\text{Noise-free RMS amplitude}}{\text{Noisy RMS amplitude}} \right). \quad (19)$$

First of all, the conventional method is robust against the noise in the data, but it is sensitive to the errors in the initial velocity model, as illustrated in the previous section. In contrast, the separate and joint methods are robust against inaccuracy of the initial velocity model due to the extended search space allowing for data fitting with an inaccurate model but are sensitive to the noise in recorded data due to the risk

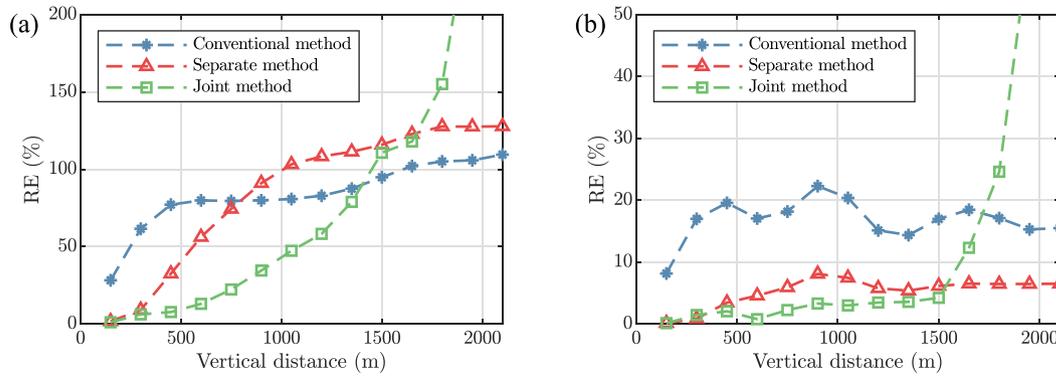


Figure 6. Mean relative error (RE) averaged over all the sources of the estimated source signatures as a function of the vertical distance between the source and receiver lines when (a) the rough velocity model 4 (Fig. 1e) and (b) the kinematically accurate velocity model 2 (Fig. 1c) are used.

of noise overfitting. Although the RE of the separate and joint methods increases with the amount of noise, it remains however far less than the conventional method. It can be seen that even in the worst scenario (with the lowest SNR data and a rough initial model), such methods based on wave-equation relaxation still can estimate an acceptable source signature.

3.1.3 Sensitivity to the distance between the source and the receivers

In IR-WRI, the accuracy of the estimated source signature is directly controlled by the distance between the source and the closest receiver. This results because the extended wavefield, from which the source signature is estimated, eq. (14), matches well the recorded data only near the receivers when the background velocity model is inaccurate. This implies that when the source is close to a receiver, it will be estimated from an accurate estimation of the wavefield at this receiver. Moreover, the impedance matrix in eq. (14), which is built from a potentially inaccurate velocity model, will not generate significant errors when applied to the wavefield to generate the source when the latter and the receiver are close to each other. In this section, we show the robustness of the different methods against this distance. We do the same test as before, but we change the depth of the source line while keeping the receiver line at the surface. The average RE of the estimated source signatures, summed over all the sources, are plotted in Fig. 6 as a function of the depth of the source profile for the conventional, separate, and joint methods and for the rough model 4 (Fig. 6a) and the kinematically accurate model 2 (Fig. 6b). First, it can be seen that for the vertical distances up to 1500 m, the joint method has a better performance for both initial velocity models, but it becomes unstable beyond this vertical distance where the subsurface becomes more complex (green curves in Fig. 6). This suggests that the additional degrees of freedom in the joint method drives the least-squares problem, eq. (13), toward an inaccurate local minimizer when the extended wavefield becomes too inaccurate at the source location. For surface acquisitions, this should however not be an issue in practice. For towed-streamer acquisitions, the sources are close to the nearest receiver, and both of them are in the water. In seabed acquisition, the reciprocal sources are on the seabed and potentially far away from the receivers. However, the medium between the receiver and the source layouts (i.e. the water) is known. On land, areal acquisitions are classically designed with sources and receivers at the surface with short nearest offset.

3.1.4 Sensitivity to the number of sources

The other aspect that we need to investigate is the number of sources and receivers. We repeat the same test as before several times with the rough initial velocity model 4 and a line of 320 receivers with 50 m spacing at the surface. For each of them, we use a line of sources at 75 m depth with a different number of sources ranging from 1 to 360 with a step of 20. The average RE over all the sources for the estimated source signatures as a function of the number of sources for the conventional, separate, and joint methods is shown in Fig. 7. Both the conventional method and the separate method have a stable behavior (blue and red curves), while the joint method becomes unstable as soon as the number of sources exceeds the number of receivers. In this case, the source signature estimation problem becomes underdetermined in the case of the joint method, and eq. (13) converges to the least-norm solution. In practice, this issue can be bypassed by subdividing the sources into patches of suitable dimension, gathering possible closely spaced sources. However, one LU factorization will be needed to be performed for each patch of sources.

3.1.5 Assessment of IR-WRI with source signature estimation

We continue by assessing Algorithms 1 and 2 as well as the separate method for FWI on the Marmousi II model when the inversion is started using the rough initial model 4 (Fig. 1e) and the 3 Hz frequency. The fixed-spread surface acquisition consists of 114 sources spaced 150 m apart with the source signatures depicted in Fig. 2 at the surface, and 340 hydrophone receivers spaced 50 m apart at 75 m depth.

We first show the footprint of the source blending in the matrix S , eq. (16), for the rough initial model 4 (Fig. 8a). The 34th column of the S matrix is plotted in black to give more precise insights into the relative magnitude of the diagonal and off-diagonal elements. Also, the

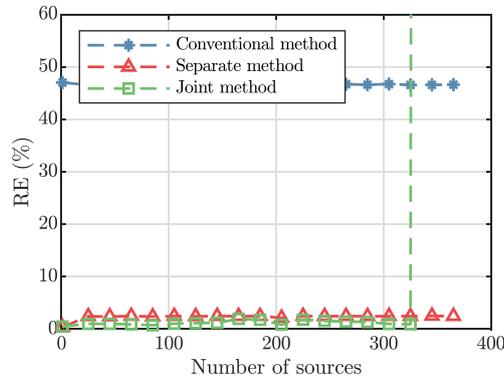


Figure 7. Mean RE averaged over all the sources of the estimated source signatures as a function of the number of the sources when the rough initial velocity model 4 (Fig. 1e) and a line of 320 receivers at the surface are used.

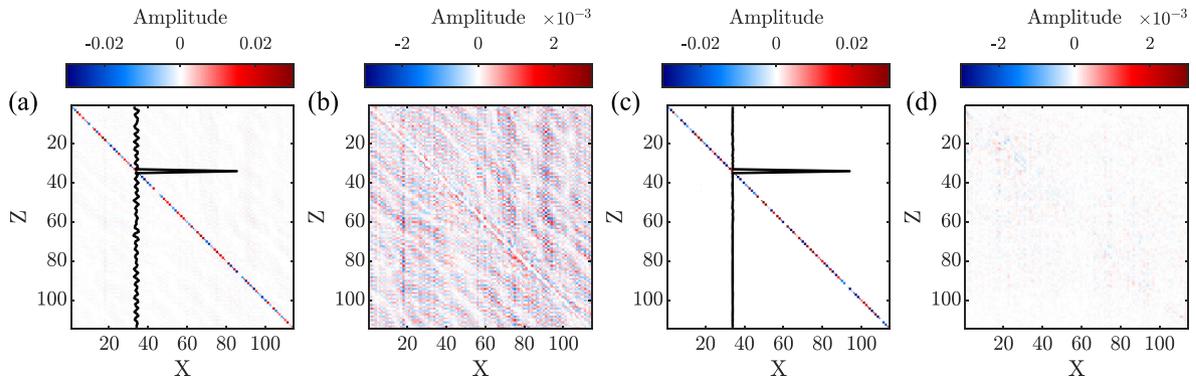


Figure 8. Real components of the source signature matrix, that is $\Phi^T \mathbf{A}_k \mathbf{U}^k$ computed with (a and b) the rough initial model 4 (Fig. 1e) and the 3 Hz frequency, and (c and d) the kinematically accurate velocity model 2 (Fig. 1c) and the 12 Hz frequency. In panels (a, c), the 34th column of $\Phi^T \mathbf{A}_k \mathbf{U}^k$ is plotted in black to gain insight into the relative magnitude of the diagonal and off-diagonal elements. Panels (b, d) show the non-diagonal components of panels (a, c), i.e., $[\Phi^T \mathbf{A}_k \mathbf{U}^k - \text{Diag}(\text{diag}(\Phi^T \mathbf{A}_k \mathbf{U}^k))]$.

off-diagonal elements of this matrix are plotted separately in Fig. 8(b). It is shown that the diagonal elements of this matrix are dominant, the maximum amplitude of the off-diagonal components being <1 percent of the maximum diagonal element. We remind that such off-diagonal elements partially absorb the errors in the estimated physical source signatures when the initial velocity model is rough, but we need to remove these effects during the inversion, which is the goal of Algorithms 1 and 2.

The effects of the source blending are also seen in the reconstructed wavefields. The reconstructed monochromatic wavefield associated with the source located in Fig. 8(a) is shown in Fig. 9(a). To assess the effects of the source blending, we show the difference between this wavefield and the extended wavefield reconstructed with the true source signature in Fig. 9(b). It can be seen that the differences are not significant. This is the worst scenario for the velocity model because of the significant inaccuracy of the velocity model 4. However, the estimated source signature matrix becomes close to a diagonal matrix as soon as the accuracy of the velocity model improves. This statement is verified in Figs 8(c)–(d) and 9(c)–(d), which are similar to Figs 8(a)–(b) and 9(a)–(b), except that the source signature and the wavefield are now estimated from the kinematically accurate velocity model 2 and a frequency of 12 Hz. It can be seen that the source matrix (Fig. 8c), tends to a diagonal matrix, and the differences between the estimated and the true wavefields tend to zero in Fig. 9(d).

We continue by performing the frequency-domain IR-WRI without regularization in the 3–12 Hz frequency band with a frequency interval of 0.5 Hz. Mono-frequency batches are successively inverted following a classical frequency continuation strategy (Sirgue & Pratt 2004). We perform three paths through the frequency batches to improve the inversion results, using one path's final model as the initial model of the next one. The starting and finishing frequencies of the paths are [3, 6], [3, 7], [3, 12] Hz. The stopping criterion of iterations is 10 iterations per batch. Also, the bound constraints using the true values of the minimum and maximum velocities as bounds are applied for velocity-model update.

We compare the results of IR-WRI when performed with known sources (Fig. 10a) and unknown sources (Figs 10(b–e)). In the latter case, the source signatures are estimated with the separate method (Fig. 10b) and the joint method using Algorithm 1 (Fig. 10c), Algorithm 2 (Fig. 10d) and the fast version of Algorithm 2 (Fig. 10e). Also, the model errors (the difference between the estimated and true velocity model) for the different estimated models (Figs 10a–e) are shown in Figs 11(a)–(e). All of these methods perform well but with a different computational cost. Also, IR-WRI with known sources (Fig. 10a) and IR-WRI using normal and fast version of Algorithm 2 (Figs 10d and e) provide similar results with improved accuracy at the reservoir level (compare Figs 11a, d and e).

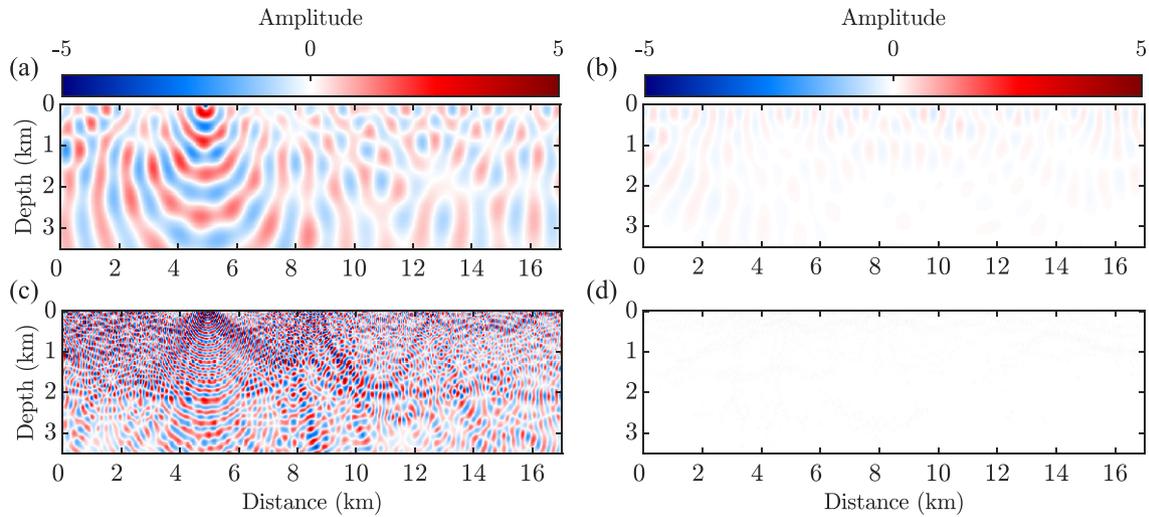


Figure 9. (a) Reconstructed wavefield using eq. (15) using the rough initial model 4 (Fig. 1e) for the 3 Hz frequency. (b) Difference between (a) and the reconstructed wavefield with the true source signature. (c–d) Same as (a–b), but using the kinematically accurate velocity model 2 (Fig. 1c) for the 12 Hz frequency.

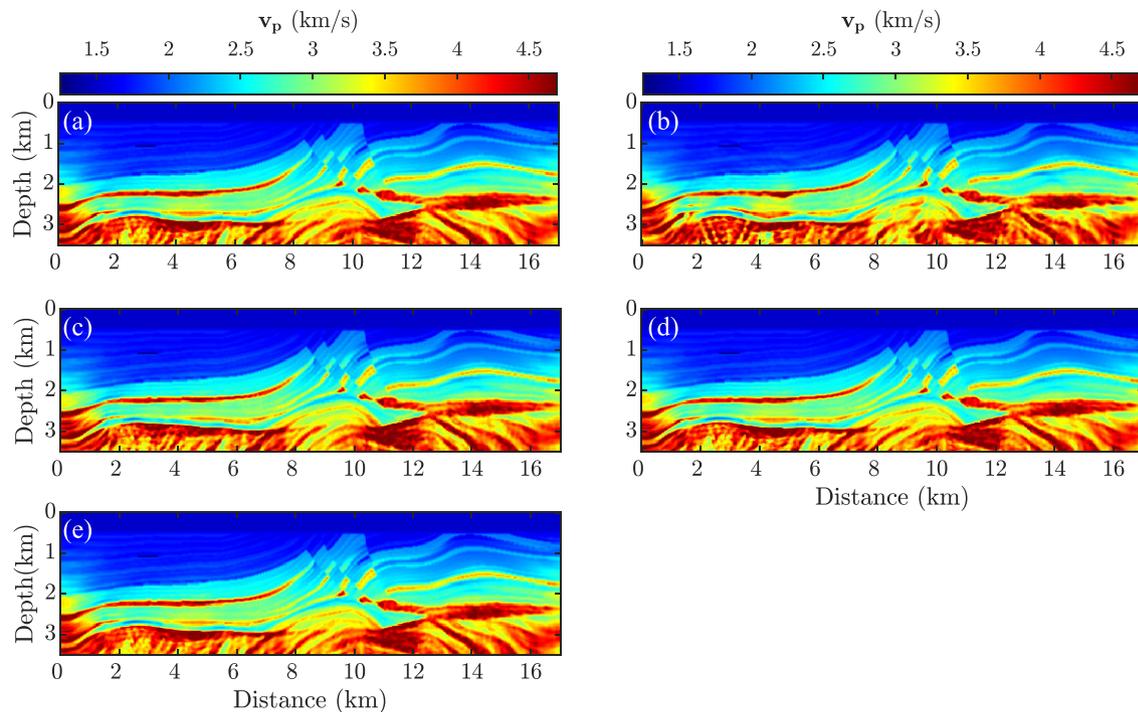


Figure 10. Marmousi II inversion results when the inversion starts from the crude initial velocity model 4 (Fig. 1e). Final IR-WRI velocity models obtained with (a) known sources, (b–d) unknown sources. In (b–d), source estimation is performed with (b) separate method, (c) joint method with Algorithm 1 and (d) joint method with Algorithm 2 and (e) joint method with the fast version of Algorithm 2.

Finally, the wave-equation residuals, data residuals, and RE for IR-WRI with known and unknown sources are shown in Fig. 12. The data residuals (Fig. 12b), for Algorithm 2 (normal and fast versions) are close to the separate method and are less than Algorithm 1 because the wavefield is corrected in Algorithm 2 for the errors of the source blending assumption. The relative error (RE) of the model are close together for all the methods (Fig. 12c); however, the models obtained with the separate and Algorithm 2 (normal and fast versions) are slightly more accurate than the others.

In summary, we see that the separate method and all of the joint-approach algorithms work well and can reconstruct velocity models close to what we get from IR-WRI with known source signatures but with a different computational burden. The test in the next section shows that this conclusion is not valid for a more complicated velocity model.

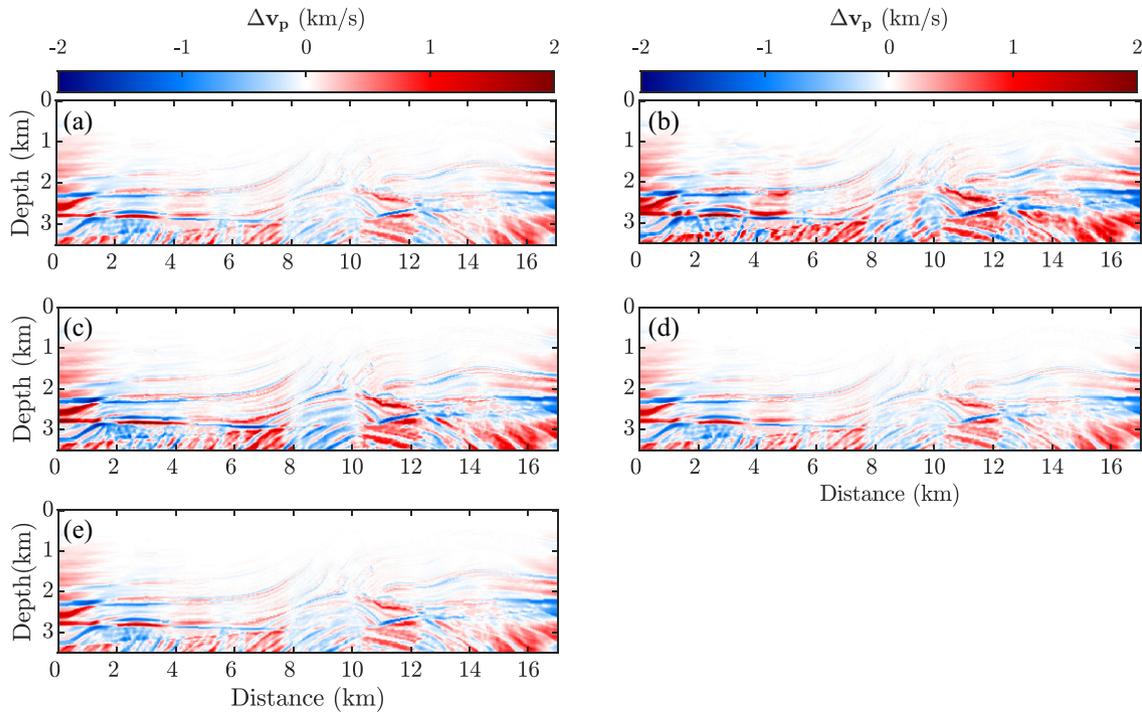


Figure 11. Marmousi II inversion results. (a–e) Model errors (i.e., difference between true and estimated models) for the estimated models shown in Figs 10 (a)–(e).

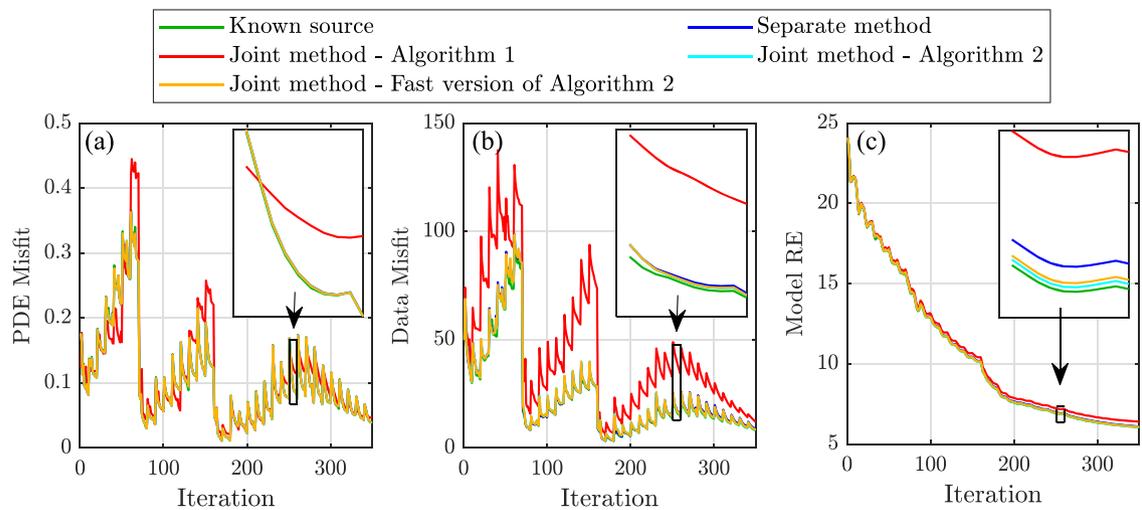


Figure 12. Marmousi II test. Comparison between the convergence history of IR-WRI performed with known sources (green) and unknown sources estimated with separate method (blue), joint method with Algorithm 1 (red), joint method with Algorithm 2 (cyan) and joint method with the fast version of Algorithm 2 (orange) as a function of iteration number. (a) PDE misfit, (b) Data misfit and (c) Model RE. In each panel, a zoomed image from the selected area is shown.

3.2 2004 BP salt model

We continue by assessing the performance of the methods on a rescaled version of left-target of challenging 2004 BP salt model (Billette & Brandsberg-Dahl 2004) (Fig. 16a) when a 1-D gradient initial model (Fig. 16b) and the 3 Hz frequency are used to start the inversion. We use 65 point sources with 250 m spacing and a line of receivers with 50 m spacing at 75 m depth. Like the previous test, the source signatures are random Ricker wavelets with different central frequencies between [8–12] Hz and the initial phases between [0–0.4] s. We apply the inversion in the 3–13 Hz frequency band with a frequency interval of 0.5 Hz. We perform three paths through the frequency batches, using the final model of one path as the initial model of the next one, and each batch contains two frequencies with one frequency overlap. The starting and finishing frequencies of the three paths are [3, 6], [4, 8.5] and [6, 13] Hz, respectively. The stopping criterion of iterations is 45 iterations for the first batch and 10 iterations per batch for the rest of them. Bound-constrained Tikhonov + Total variation (BTT)-regularization (Aghamiry *et al.* 2020a) is applied on IR-WRI for all the cases to decrease the ill-posedness of the problem. We first plot the RE for the b-d), source estimation is performed with the separate and joint methods as a function of the source number in Fig. 13 for the first (3 Hz, using initial

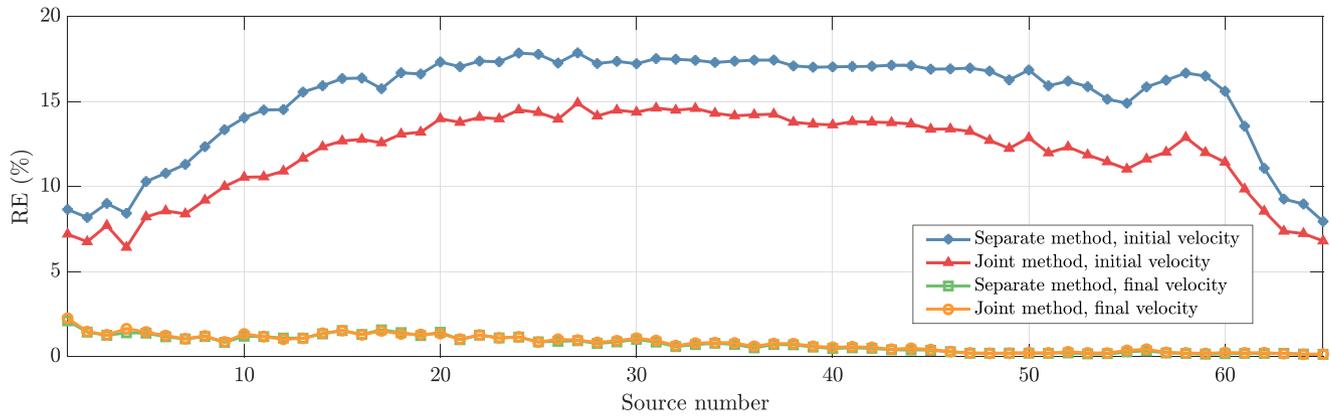


Figure 13. BP Salt test. Relative error (RE) of the source signatures estimated with the separate and joint methods when a rough initial model (Fig. 16b) and the accurate velocity model (Fig. 16g) are used.

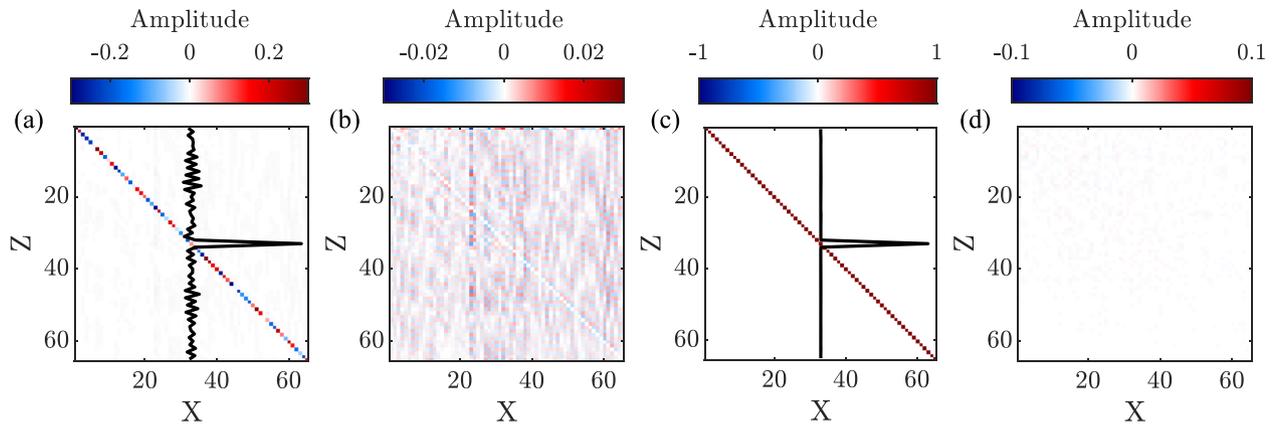


Figure 14. BP Salt test. Real components of source signature matrix, that is $\Phi^T \mathbf{A}_k \mathbf{U}^k$, estimated at the (a–b) first and (c–d) final iterations of IR-WRI. Panels (a, c) The 34th column of $\Phi^T \mathbf{A}_k \mathbf{U}^k$ is plotted in black to give insight into the relative magnitude of the diagonal and off-diagonal elements. Panels (b, d) show the non-diagonal components of panels (a, c), i.e., $[\Phi^T \mathbf{A}_k \mathbf{U}^k - \text{Diag}(\text{diag}(\Phi^T \mathbf{A}_k \mathbf{U}^k))]$.

velocity model Fig. 16b) and the final iteration of the inversion (13 Hz, using updated velocity model Fig. 16g). We see the same effects as those revealed by Marmousi II in Fig. 4 in the sense that the joint approach provides more accurate source signature estimation.

Also, the estimated source signature matrices, eq. (16), at the first and final iterations of IR-WRI, are shown in Fig. 14 and the related reconstructed monochromatic wavefields are shown in Fig. 15. Again, we see similar effects as those revealed by the Marmousi II test, except that the off-diagonal elements are larger relative to the diagonal counterparts, hence revealing the more complex structure of the BP salt model.

We compare the final results of BTT-regularized IR-WRI when performed with the known sources (Fig. 16c) and the unknown sources (Fig. 16(d)–(g)). In the latter case, source estimation is performed with the separate method (Fig. 16d) and the joint method using Algorithm 1 (Fig. 16e), Algorithm 2 (Fig. 16f) and the fast version of Algorithm 2 (Fig. 16g). Also, the model errors for different estimated models are shown in Fig. 17. In contrast to the Marmousi II test, the different methods do not converge to the same minimizer. Let us consider the BTT-regularized IR-WRI with known sources as the benchmark model (Fig. 16c). Only IR-WRI with unknown sources using Algorithm 2 reaches approximately the same results (Fig. 16f). Also, the results of the fast version of this algorithm (Fig. 16g) is close to this benchmark model. The failure of the separate method (Figs 16d and 17b) probably results from the limited quality of the estimated wavelets at the early iterations when the initial velocity model is inaccurate (Fig. 13). Also, the failure of Algorithm 1 (Figs 16e and 17c) may result from the significant amplitudes of the off-diagonal elements of the estimated source signature matrix (Fig. 14), and it seems that the iterative refinement implemented in Algorithm 1 is not enough to correct all of these effects. For such complicated velocity models, we need to recompute the wavefields from the diagonalized source signature with Algorithm 2 (lines 4–5).

4 DISCUSSION

We have proposed an efficient and robust source signature estimation method for frequency-domain FWI with search space expansion. The search space expansion is generated by computing data-assimilated wavefields that jointly satisfy in a least-squares sense the wave equation and the observation equation. In this framework, the reconstructed wavefields satisfy a normal equation, which is ideally solved with direct

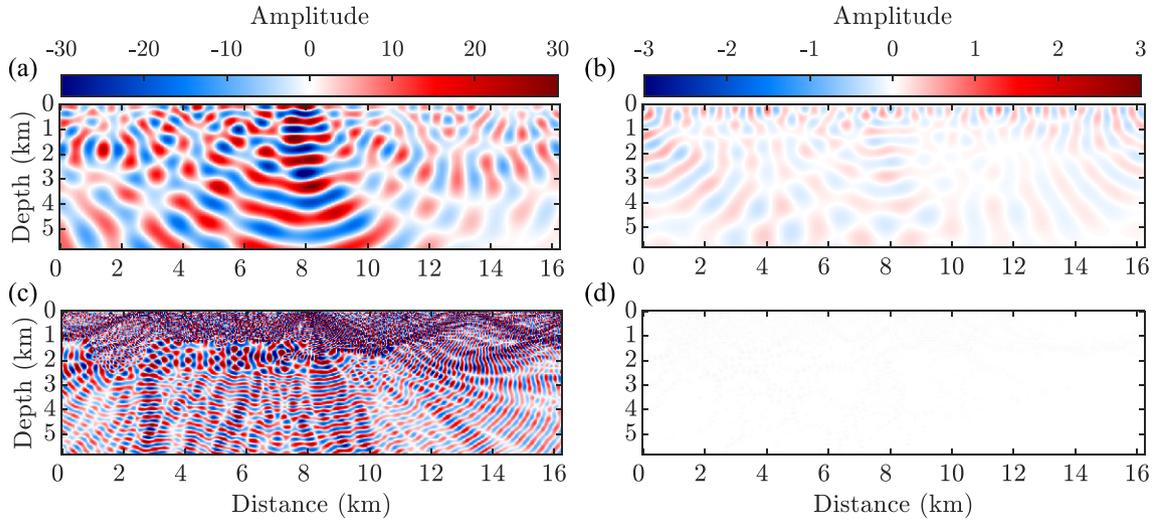


Figure 15. BP Salt test. (a) The reconstructed wavefield using eq. (15) at the first iteration of IR-WRI using joint method with the rough initial model (Fig. 16b) for 3 Hz. (b) The difference between (a) and the reconstructed wavefield with true source. (c–d) Same as (a–b), but for 13 Hz at the final iteration of IR-WRI with updated velocity model Fig. 16(g).

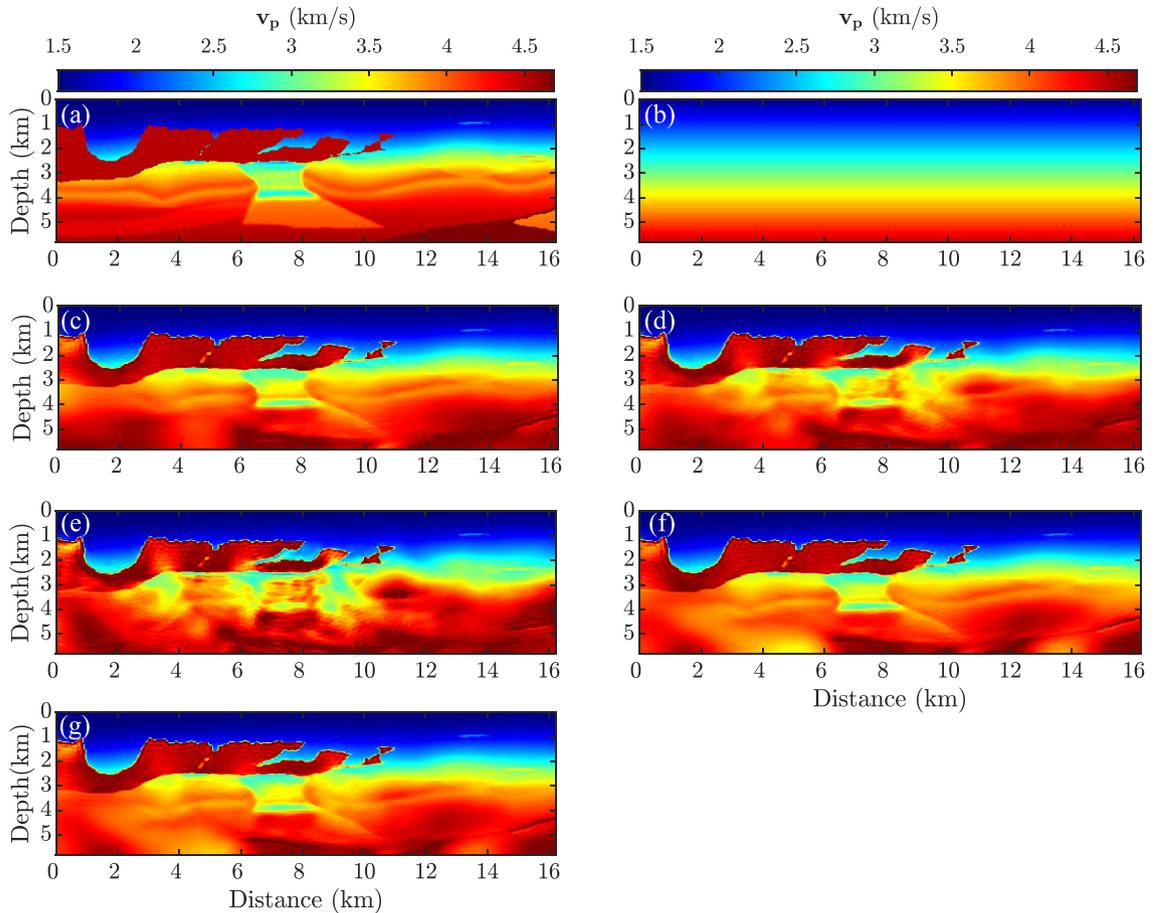


Figure 16. 2004 BP salt test: (a–b) True (a) and initial (b) velocity models. (c–f) Velocity models estimated by BTT-regularized IR-WRI with (c) known sources and (d–g) unknown sources. In (d–g), sources are estimated with (d) separate method, (e–g) joint method with (e) Algorithm 1, (f) Algorithm 2 (f) and (g) the fast version of Algorithm 2.

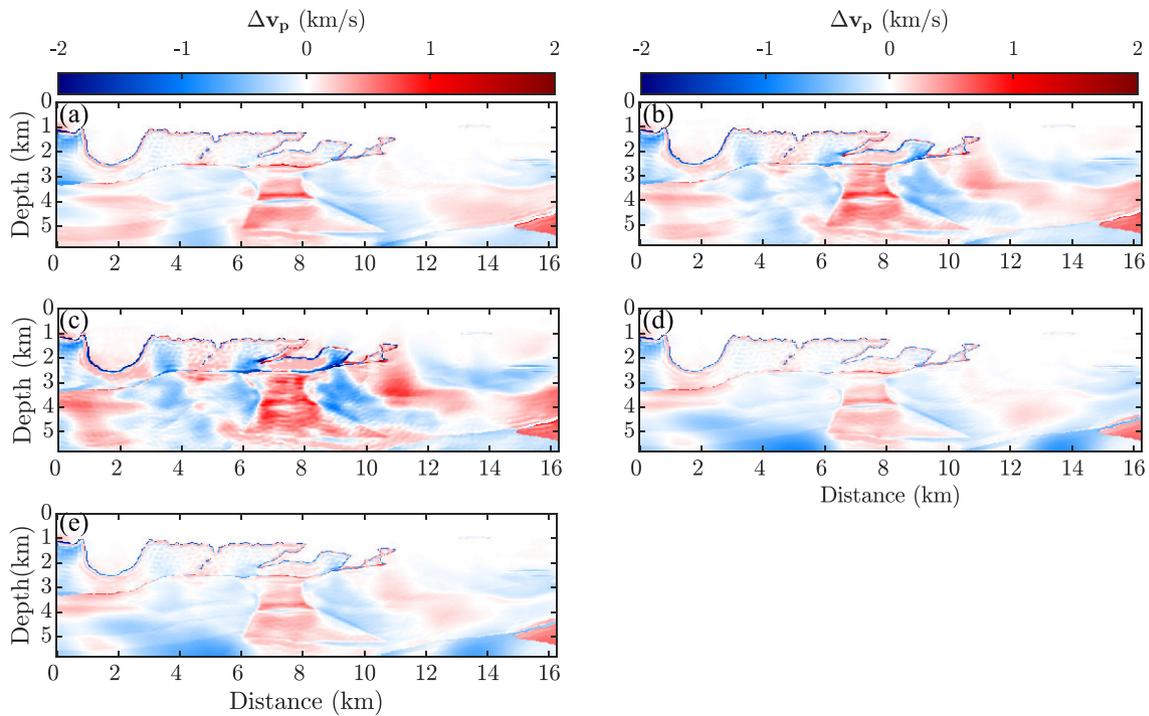


Figure 17. 2004 BP salt test: (a–e) Model errors (difference between the true and estimated models) for the estimated models shown in Figs 16(c)–(g).

methods to process efficiently multiple right-hand sides by forward/backward elimination. We process the source signature as an optimization variable by projecting its closed-form expression in the wavefield reconstruction problem. The issue raised by this variable projection is to make the normal operator source-signature dependent, hence requiring to recompute the Cholesky factorization for each source. To bypass this issue and make the normal operator source independent, we reconstruct each individual wavefield with blended sources so that these wavefields can be efficiently computed in parallel by forward/backward elimination once one Cholesky factorization has been performed. The source blending makes the source signature of each source to be a vector of dimension equal to the number of individual sources involved in the blending. When the velocity model is accurate, each entry of the signature vector is zero except the one located at the position of the physical source. Conversely, when the velocity model is inaccurate, all the entries are non zero, the one located at the physical source position being however dominant. Numerical results show that the additional degrees of freedom provided by the entries of the signature vector located away from the position of the physical source help to absorb the detrimental effects of inaccurate velocity model and, accordingly, estimate a more accurate source signature at the position of the physical source.

This approach however requires to remove the effects of the blending in iterations. To achieve this goal, we keep only the diagonal entries of the source signature matrix, namely the entries at the position of the physical sources, (Algorithm 1, Line 4) to update the model parameters and the Lagrange multipliers (Algorithm 1, Lines 5 and 6) once the wavefields have been reconstructed with the blended sources (Algorithm 1, Line 3). Numerical results suggest that this defect correction is not enough to compensate for the detrimental effects of the source blending and reach the same minimizer as IR-WRI with known sources when the targeted velocity model is complex and the inversion starts from a rough initial model. Therefore, we include an additional step in Algorithm 2 (Line 5) during which we re-compute the wavefields with the diagonal source signature matrix to remove the footprint of the blending in the wavefield reconstruction too.

Although we assume in this study that the source locations match the grid points of the computational domain, the method can be readily used for arbitrary source positions (Hicks 2002). The proposed method can be extended for general physics like visco-acoustic and visco-elastic. Moreover, Operto *et al.* (2015), Amestoy *et al.* (2016), and Operto & Miniussi (2018) show that sparse direct solvers are suitable to tackle efficiently 3D FWI case studies involving several tens of millions of unknowns and a large number of sources for stationary-recording acquisitions (ocean bottom cable and node, land acquisitions). Accordingly, the proposed algorithms can be extended to 3D problems.

The computational cost of IR-WRI (with known or unknown sources) is dominated by the wavefield estimation (\mathbf{u} -subproblem). The theoretical complexity of 2D and 3D factorizations (F) and forward-backward substitution (S) for sparse matrices are reviewed in Nihei & Li (2007). It is reminded that for 3D dense acquisitions, the theoretical time complexity of one LU decomposition is the same as that of the solution step, and hence the latter is far to be negligible (see Amestoy *et al.* (2016) for a real 3D case study). The cost of our algorithms are reviewed in Table 1 and are compared with those of extended FWI with known sources and the approach of Fang *et al.* (2018). The cost of our first algorithm is the same as that of the extended FWI with known sources. The cost of the second algorithm is twice that of the algorithm 1 when the wavefield is recomputed at each iteration of the inversion while the cost of the fast version of Algorithm 2 tends to that of Algorithm 1 when the source signature is estimated at the first iteration only and the number of iteration n_{it} is large ($\gg 1$). The approach of Fang *et al.*

Table 1. Computational burden of different algorithms for IR-WRI with unknown sources.

Method	Known sources	Separate method	Algorithm 1	Algorithm 2	Fast version of Algorithm 2	Fang <i>et al.</i> (2018)
Computational burden	$n_{it}(F + n_s S)$	$n_{it} n_s (F + S)$	$n_{it}(F + n_s S)$	$2n_{it}(F + n_s S)$	$(n_{it} + 1)(F + n_s S)$	$n_{it}(F + 2n_s S)$

(2018) performs the same number of Cholesky factorization as Algorithm 1 but doubles the number of solution steps (this results from the block partitioning of the normal operator). Therefore, the fast version of our Algorithm 2 is significantly faster than that of Fang *et al.* (2018).

Regardless computational-efficiency issue, our approach relies on different optimization strategy relative to Fang *et al.* (2018). Fang *et al.* (2018) lump the wavefield and the source signature into a single variable, which means that they estimate these two variables jointly. Then, they project this lumped variable in the model parameter subproblem. They bypass the issue of the dependency of the normal operator to the source signature by using a block partitioning of the normal operator and the block matrix inversion formula (Press *et al.* 2007, page 81). In our method, we project the closed form expression of the source signature in the wavefield-reconstruction subproblem assuming blended sources to make the normal operator source independent. We review in this paper that this artificial source blending potentially helps to estimate more accurate source signatures when the velocity model is inaccurate. Then, we extract the individual source signature to refine the wavefields and update the model parameters by simply keeping the diagonal entries of the source signature matrix. The relative merits of the method of Fang *et al.* (2018) and of our approach in terms of computational efficiency and reliability clearly deserve further verification against real case studies.

5 CONCLUSIONS

We extended the recently proposed iteratively-refined wavefield reconstruction inversion (IR-WRI) to estimate the unknown source signatures. The source signatures and wavefields are processed with a variable projection method during the extended wavefield reconstruction subproblem. We first show that the source signature estimation generates computational overhead when each source is processed separately because the extended wave equation operator becomes source dependent. This computational overhead becomes prohibitive when the augmented wave equation system is solved with a direct method since the Cholesky factorization needs to be performed for each source. To bypass this issue and make the operator source independent, we proposed a method that blends the sources during each wavefield reconstruction. Accordingly, for each source of the experiment, the proposed method searches for a virtual blended source that best fits each single-source dataset during the wavefield reconstruction. Regardless of the computational efficiency issue, we also show that when the background velocity model is inaccurate, this source blending provides a more accurate source signature estimation at the physical source location than the case where source blending is not used. This probably results from the additional degrees of freedom provided by the extra virtual sources. Once the spatially distributed source signatures have been estimated, we restrict them at the position of the physical source to mimic the true source signatures, and we update the velocity model, and correct the detrimental effects of source blending by the action of the Lagrange multipliers (Algorithm 1). For complicated velocity models, Algorithm 1 is not able to fully correct the effects of source blending. To overcome this issue, we propose Algorithm 2 where it re-estimates the extended wavefields with the localized source signatures. Numerical tests have shown that the estimated source signature using the proposed algorithm is not sensitive to the small changes in the velocity model. Hence, we can perform the source signature estimation only at the first iteration of each frequency batch of IR-WRI to preserve computational efficiency (fast version of Algorithm 2).

ACKNOWLEDGEMENTS

We would like to thank editors F. Simons and L. Alexander, and also thank N. Korta Martiartu and an anonymous reviewer for their careful reading of our manuscript and their insightful comments and suggestions. This study was partially funded by the WIND consortium (<https://www.geoazur.fr/WIND>), sponsored by Chevron, Shell and Total. The authors are grateful to the OPAL infrastructure from Observatoire de la Côte d'Azur (CRIMSON) for providing resources and support. This work was granted access to the HPC resources of IDRIS under the allocation A0050410596 made by GENCI.

DATA AVAILABILITY

The data underlying this article cannot be shared publicly due to the privacy of individuals that participated in the study.

REFERENCES

- Abubakar, A., Hu, W., Habashy, T.M. & van den Berg, P.M., 2009. Application of the finite-difference contrast-source inversion algorithm to seismic full-waveform data, *Geophysics*, **74**(6), WCC47–WCC58.
- Aghamiry, H., Gholami, A. & Operto, S., 2019a. Implementing bound constraints and total-variation regularization in extended full waveform inversion with the alternating direction method of multiplier: application to large contrast media, *Geophys. J. Int.*, **218**(2), 855–872.

- Aghamiry, H., Gholami, A. & Operto, S., 2019b. Improving full-waveform inversion by wavefield reconstruction with alternating direction method of multipliers, *Geophysics*, **84**(1), R139–R162.
- Aghamiry, H., Gholami, A., & Operto, S., 2020a. Compound regularization of full-waveform inversion for imaging piecewise media, *IEEE Trans. Geosci. Remote Sens.*, **58**(2), 1192–1204.
- Aghamiry, H., Gholami, A. & Operto, S., 2020b. Robust wavefield inversion with phase retrieval, *Geophys. J. Int.*, **221**(2), 1327–1340.
- Aghamiry, H., Gholami, A. & Operto, S., 2021. Full waveform inversion by proximal newton methods using adaptive regularization, *Geophys. J. Int.*, **224**(1), 169–180.
- Amestoy, P., Brossier, R., Buttari, A., L'Excellent, J.-Y., Mary, T., Métivier, L., Miniussi, A. & Operto, S., 2016. Fast 3D frequency-domain FWI with a parallel Block Low-Rank multifrontal direct solver: application to OBC data from the North Sea, *Geophysics*, **81**(6), R363–R383.
- Aravkin, A., & van Leeuwen, T., 2012. Estimating nuisance parameters in inverse problems, *Inverse Problems*, **28**(11).
- Aravkin, A.Y., van Leeuwen, T., Calandra, H., & Herrmann, F.J., 2012. Source estimation for frequency-domain FWI with robust penalties, in *Proceedings of the 74th Annual Conference & Exhibition Incorporating SPE EUROPEC, Expanded Abstracts*, Copenhagen, Denmark, 4–7 June 2012.
- Billette, F.J., & Brandsberg-Dahl, S., 2004. The 2004 BP velocity benchmark, in *Proceedings of the 67th Annual EAGE Conference & Exhibition, Extended Abstracts, cp-1-00513*, Madrid, Spain, European Association of Geoscientists & Engineers, doi:10.3997/2214-4609-pdb.1.B035.
- Boyd, S., Parikh, N., Chu, E., Peleato, B., & Eckstein, J., 2010. Distributed optimization and statistical learning via the alternating direction method of multipliers, *Machine Learning*, **3**(1), 1–122.
- Chen, Z., Cheng, D., Feng, W., & Wu, T., 2013. An optimal 9-point finite difference scheme for the Helmholtz equation with PML, *Int. J. Numer. Anal. Model.*, **10**(2), 389–410.
- Combettes, P. L. & Pesquet, J.-C., 2011. Proximal splitting methods in signal processing, in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pp. 185–212, Springer.
- Fang, Z., Wang, R. & Herrmann, F.J., 2018. Source estimation for wavefield-reconstruction inversion, *Geophysics*, **83**(4), R345–R359.
- Gholami, A., Aghamiry, H. & Abbasi, M., 2018. Constrained nonlinear AVO inversion using Zoeppritz equations, *Geophysics*, **83**(3), R245–R255.
- Goldstein, T., & Osher, S., 2009. The split Bregman method for L1-regularized problems, *SIAM J. Imag. Sci.*, **2**(2), 323–343.
- Golub, G., & Pereyra, V., 2003. Separable nonlinear least squares: the variable projection method and its applications, *Inverse Problems*, **19**(2), R1.
- Haber, E., Ascher, U.M. & Oldenburg, D., 2000. On optimization techniques for solving nonlinear inverse problems, *Inverse Problems*, **16**(5), 1263.
- Hicks, G. J., 2002. Arbitrary source and receiver positioning in finite-difference schemes using Kaiser windowed sinc functions, *Geophysics*, **67** (1), 156–166, 10.1190/1.1451454.
- Huang, G., Nammour, R. & Symes, W.W., 2018. Source-independent extended waveform inversion based on space-time source extension: frequency-domain implementation, *Geophysics*, **83**(5), R449–R461.
- Li, M., Rickett, J. & Abubakar, A., 2013. Application of the variable projection scheme for frequency-domain full-waveform inversion, *Geophysics*, **78**(6), R249–R257.
- Marfurt, K., 1984. Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations, *Geophysics*, **49**, 533–549.
- Martin, G.S., Wiley, R. & Marfurt, K.J., 2006. Marmousi2: an elastic upgrade for Marmousi, *Leading Edge*, **25**(2), 156–166.
- Nihei, K.T. & Li, X., 2007. Frequency response modelling of seismic waves using finite difference time domain with phase sensitive detection (TD-PSD), *Geophys. J. Int.*, **169**, 1069–1078.
- Nocedal, J., & Wright, S.J., 2006. *Numerical Optimization*, 2nd edn, Springer.
- Operto, S. & Miniussi, A., 2018. On the role of density and attenuation in 3D multi-parameter visco-acoustic VTI frequency-domain FWI: an OBC case study from the North Sea, *Geophys. J. Int.*, **213**, 2037–2059.
- Operto, S., Miniussi, A., Brossier, R., Combe, L., Métivier, L., Monteiller, V., Ribodetti, A. & Virieux, J., 2015. Efficient 3-D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the visco-acoustic vertical transverse isotropic approximation, *Geophys. J. Int.*, **202**(2), 1362–1391.
- Parikh, N., & Boyd, S., 2013. Proximal algorithms, *Found. Trends Opt.*, **1**(3), 123–231.
- Parks, M. L., De Sturler, E., Mackey, G., Johnson, D. D. & Maiti, S., 2006. Recycling Krylov Subspaces for Sequences of Linear Systems, *SIAM Journal on Scientific Computing*, **28**(5), 1651–1674.
- Plessix, R.E. & Cao, Q., 2011. A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium, *Geophys. J. Int.*, **185**, 539–556.
- Pratt, R.G., 1999. Seismic waveform inversion in the frequency domain, part I: theory and verification in a physical scale model, *Geophysics*, **64**, 888–901.
- Pratt, R.G., Shin, C. & Hicks, G.J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion, *Geophys. J. Int.*, **133**, 341–362.
- Press, W. H., William, H Teukolsky, S. A. Vetterling, W. T. Saul, A. & Flannery, B. P. 2007. *Numerical recipes 3rd edition: The art of scientific computing.*, Cambridge university press.,
- Rickett, J., 2013. The variable projection method for waveform inversion with an unknown source function, *Geophys. Prospect.*, **61**(4), 874–881.
- Sirgue, L. & Pratt, R.G., 2004. Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies, *Geophysics*, **69**(1), 231–248.
- Song, Z., Williamson, P.R. & Pratt, R.G., 1995. Frequency-domain acoustic-wave modeling and inversion of crosshole data. Part 2 : inversion method, synthetic experiments and real-data results, *Geophysics*, **60**(3), 786–809.
- Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**(8), 1259–1266.
- van Leeuwen, T. & Herrmann, F., 2016. A penalty method for PDE-constrained optimization in inverse problems, *Inverse Problems*, **32**(1), 1–26.
- van Leeuwen, T. & Herrmann, F.J., 2013. Mitigating local minima in full-waveform inversion by expanding the search space, *Geophys. J. Int.*, **195**(1), 661–667.
- Venkatakrishnan, S.V., Bouman, C.A. & Wohlberg, B., 2013. Plug-and-play priors for model based reconstruction, in *2013 IEEE Global Conference on Signal and Information Processing*, pp. 945–948, IEEE.
- Virieux, J. & Operto, S., 2009. An overview of full waveform inversion in exploration geophysics, *Geophysics*, **74**(6), WCC1–WCC26.
- Yin, W., & Osher, S., 2013. Error forgetting of Bregman iteration, *J. Scient. Comput.*, **54**, 684–695.