Laplace-Domain Crosstalk-Free Source-Encoded Elastic Full-Waveform Inversion Using Time-Domain Solvers

Zhaolun Liu[*], Jürgen Hoffmann[†], Etienne Bachmann[‡], Congyue Cui[‡],
Frederik J. Simons[‡], and Jeroen Tromp[‡§]

ABSTRACT

Crosstalk-free source-encoded elastic Full-Waveform Inversion (FWI) using time-domain solvers has demonstrated skill and efficiency at conducting seismic inversions involving multiple sources and receivers with limited computational resources. A drawback of common formulations of the procedure is that, by sweeping through the frequency domain randomly at a rate of one or a few sparsely sampled frequencies per shot, it is difficult to simultaneously incorporate time-selective data windows, as necessary for the targeting of arrivals or wave packets during the various stages of the inversion. Here, we solve this problem by using the Laplace transform of the data. Using complex-valued frequencies allows for damping the records with flexible decay rates and temporal offsets that target specific traveltimes. We present the theory of crosstalk-free source-encoded FWI in the Laplace domain, develop the details of its implementation, and illustrate the procedure with numerical examples relevant to exploration-scale scenarios.

INTRODUCTION

Full-waveform inversion (FWI) aims to use all the information in a seismogram to estimate subsurface material properties (Lailly, 1983; Tarantola, 1984, 1986). Over the last few decades, FWI has firmly established itself as an important inversion tool for both acoustic and elastic problems in active-source exploration seismology (Pratt, 1999; Pratt and Shipp, 1999; Plessix, 2006; Operto et al., 2013) as well as for regional and global problems using natural earthquakes as sources (Liu and Tromp, 2008; Liu and Gu, 2012; Tromp, 2020). Adjoint formulations of the inverse problem and spectral-element simulations have fast become an inseparable pairing (Tromp et al., 2008).

Full-Waveform Inversion

Most of the applications of FWI to real data acquired on land (e.g., Brenders and Pratt, 2007; Plessix et al., 2010; Lemaistre et al., 2019; Murphy et al., 2020) or in marine settings (Prieux et al., 2013a; Operto et al., 2015) have used the acoustic approximation to wavefield modeling. While reliable results can be obtained in that case, provided appropriate data preprocessing and inversion preconditioning are applied (Brenders and Pratt, 2007; Prieux et al., 2013a), elastic FWI is desirable for applications when the data include strong elastic effects. Acoustic FWI can lead to erroneous inversion results if applied to elastic data that are sensitive to strong velocity contrasts (Barnes and Charara, 2009; Mora and Wu, 2018; Pan et al., 2018; Pérez Solano and Plessix, 2023; Zhang et al., 2023). Elastic FWI is especially advantageous for the inversion of land seismic data (Pérez Solano and Plessix, 2019), marked by high-amplitude surface waves and elastic effects from near-surface heterogeneities. Elastic FWI also allows for the inversion of particle velocity data of multicomponent ocean-bottom seismometers that record compressional and shear waves directly, not just hydroacoustic pressure conversions. As shown by Cho et al. (2022), for example, the shallow-background shear velocity can be inverted from the horizontal components of motion recorded by ocean-bottom nodes. Multicomponent data help constrain the inversion process and mitigate ill-posedness (Prieux et al., 2013a,b; Pan et al., 2019).

A drawback of elastic FWI is that it remains computationally challenging to solve large-scale optimization systems with three-dimensional (3-D) simulations for hundreds or even thousands of sources. Frequency-domain solvers have proven to be efficient at computing the response of large numbers of

[*] Saudi Aramco, Dhahran, Saudi Arabia; Formerly Department of Geosciences, Princeton University, USA; zhaolun.liu@outlook.com.
[†] DNO, ASA, Norway.
[‡] Department of Geosciences, Princeton University, USA.
[§] Program in Applied and Computational Mathematics, Princeton University, USA.

sources (Pratt, 1999). These methods additionally allow for the selection of just a handful of contributing frequencies (Sirgue and Pratt, 2004), and for the straightforward and cheap incorporation of intrinsic attenuation in the form of complex-valued wave speeds (Marfurt, 1984). Despite these advantages, direct solvers scale poorly (Virieux and Operto, 2009), require vast amounts of storage, and thus are unsuitable for the large 3-D problems that we face today. Iterative frequency-domain solvers may not share these limitations (Operto et al., 2007; Plessix, 2009), but they do not have the benefit of source independence.

In contrast, explicit time-domain solvers scale linearly with the number of sources owing to the absence of all-to-all communications (Komatitsch and Tromp, 2002; Komatitsch et al., 2002) and they do not command a large amount of memory. However, since the compute time for elastic FWI using time-domain solvers is proportional to the number of seismic sources (for anelastic kernels, see Komatitsch et al., 2016), distributing those over the processors of parallel computers claim large time blocks on huge computational platforms, in particular if the source parallelism is combined with a second level of parallelism by domain decomposition of the computational mesh.

Source Encoding

Source encoding remedies the run-away computational burden of wavefield modeling in seismic inversion and migration. It reduces the number of calculations by combining gathers from multiple shots or events into "encoded" "super-gathers" (Morton and Ober, 1998; Etgen, 2005; Zhang et al., 2005; Krebs et al., 2009; Tang and Biondi, 2009; Ben-Hadj-Ali et al., 2011; Schuster et al., 2011). Under the encoded multisource approach the computational cost of FWI or Reverse Time Migration (RTM) no longer grows linearly with the number of sources, allowing for the simultaneous consideration of several thousand sources.

The mathematical effect of source encoding is clear by inspection of the FWI *gradient*, the equivalent RTM *imaging condition* (Luo et al., 2009; Zhu et al., 2009; Luo et al., 2015), or the *sensitivity kernel* (Tromp et al., 2005). Disregarding weighting factors and other subtleties, these objects generally involve the correlation interaction between two scalar wavefields (Romero et al., 2000). Combining all contributions in a survey or experiment by summing the single-frequency wavefields individually over all distinct sources is computationally advantageous but creates cross-term artifacts. Crosstalk negatively affects the quality of the inversion gradient or migration image and may introduce inaccuracy in the model results. A source-encoding strategy amounts to applying an encoding operator prior to expansion over the sources in a manner such that the cross-terms become relatively unimportant.

For every one of N_s seismic sources or shot gathers, and for each of the N_ω frequencies under consideration, let $\tilde{u}_s(\mathbf{x}, \omega_k)$ and $\tilde{u}_s^\dagger(\mathbf{x}, \omega_k)$ be generic forward and adjoint wavefields (the star "$*$" denotes complex conjugation) at position $\mathbf{x}$ and angular frequency ω_k, and $\alpha_s(\omega_k)$ a generic normalized encoding operator. Denoting the real part by $\Re$, the expanded imaging condition (Romero et al., 2000)

$$\widetilde{K}(\mathbf{x}) = \Re \sum_{k=1}^{N_\omega} \left[\sum_{s=1}^{N_s} \alpha_s(\omega_k) \tilde{u}_s(\mathbf{x}, \omega_k) \right] \times \left[\sum_{s'=1}^{N_s} \alpha_{s'}(\omega_k) \tilde{u}_{s'}^\dagger(\mathbf{x}, \omega_k) \right]^* \tag{1}$$

$$= \Re \sum_{k=1}^{N_\omega} \sum_{s=1}^{N_s} \tilde{u}_s(\mathbf{x}, \omega_k) u_s^{\dagger *}(\mathbf{x}, \omega_k) + \Re \sum_{k=1}^{N_\omega} \sum_{s=1}^{N_s} \sum_{s' \neq s}^{N_s} \tilde{u}_s(\mathbf{x}, \omega_k) \left[\alpha_s(\omega_k) \alpha_{s'}^*(\omega_k) \right] \tilde{u}_{s'}^{\dagger *}(\mathbf{x}, \omega_k) \tag{2}$$

suffers from crosstalk unless $\alpha_s(\omega_k)\alpha_{s'}^*(\omega_k) \to \delta_{ss'}\delta_{kk'}$. Encoding operators need to be chosen carefully to produce enough dissimilarity between all sources to mutually decorrelate them in order to reduce the effect of crosstalk (and random data) noise (Aghazade et al., 2022).

A number of encoding strategies have been proposed in the literature. Romero et al. (2000) suggest multiplying the shot gather by $\alpha_s(\omega) = e^{\imath \phi_s(\omega)}$ with a random *phase* $\phi_s(\omega)$. Note that the imaginary unit $\imath$ is to be distinguished from the i used to denote a component index in what follows. Krebs et al. (2009) advocate randomizing *polarity* using $\alpha_s(\omega) = \pm 1$ drawn at random. Tang and Biondi (2009) and Schuster et al. (2011) propose shifting each shot with a random *time* offset. While they all suffer from residual crosstalk to some extent, these strategies furthermore require that the receiver spread be fixed for each shot, hence they are not applicable for, e.g., marine streamer data, with arrays that move with each shot. In contrast, a technique that is applicable to marine streamer acquisition is through plane-wave encoding, whereby shot gathers are linearly time-shifted with respect to source-receiver offset and summed to form a series of plane-wave gathers with different ray parameters (Etgen, 2005; Zhang et al., 2005; Vigh and Starr, 2008; Tang and Biondi, 2009). However, this presupposes a small shot interval to avoid aliasing, which may not be possible for realistic 3-D data sets, and a sufficiently large number of ray parameters.

Crosstalk-Free Source Encoding

Source-encoding strategies designed to be completely crosstalk-*free* have been developed in the last decade (Huang and Schuster, 2012, 2018; Krebs et al., 2013; Schuster and Huang, 2013; Zhang et al., 2018; Tromp and Bachmann, 2019). Both Huang and Schuster (2012) and Krebs et al. (2013) propose an encoding operator that is a narrowband filter, convolved in the time domain with each trace. Different shot gathers are allocated different filters from a nearly orthogonal set. In the context of least-squares RTM, Dai et al. (2013) propose assigning to each source a unique frequency and using an orthogonal encoding operator $\alpha_s(\omega_k) = e^{\imath \omega_k t}$. The encoded wavefields mix frequencies and sources $\{k, s\}$ to form "super-gathers",

$$u(\mathbf{x},t) = \Re \sum_{\{k,s\}=1}^{N_\omega \times N_s} e^{-\imath\omega_k t}\,\tilde{u}_s(\mathbf{x},\omega_k), \qquad (3)$$

which can be computed in the time-domain by simultaneously activating multiple sources driven by specific frequencies (Nihei and Li, 2007). The encoded forward wavefield $u(\mathbf{x}, t)$ and its adjoint $u^\dagger(\mathbf{x}, t)$ are run past the "steady-state time", T. The sensitivity kernel is obtained by zero-lag cross-correlation over $\Delta\tau$, an interval over which the encoding operators are exactly orthogonal, which accomplishes the "deblending" or "decoding",

$$K(\mathbf{x}) = \int_0^{\Delta\tau} u(\mathbf{x}, T + \Delta\tau - t)\, u^\dagger(\mathbf{x}, t + T)\,\mathrm{d}t. \qquad (4)$$

Both T and $\Delta\tau$ require careful consideration, though authors often take $T = \Delta\tau$ for convenience. Huang and Schuster (2018) apply this idea in acoustic FWI. Zhang et al. (2018, 2020) propose a hybrid method that calculates time-domain encoded wavefields but performs the kernel computation in the frequency domain.

Tromp and Bachmann (2019) provide a systematic and didactic overview of crosstalk-free source-encoded FWI methods. They make the distinction between the steady-state time T and the integration interval $\Delta\tau$ required for decoding, which are not typically differentiated. They also formulate different misfits, e.g., for amplitude, phase, and double-difference measurements. Finally, they show that the relative reduction in computational cost of source-encoded to traditional FWI for a large number of sources is approximately given by the original duration of the simulated seismograms (e.g., about 5 s) times their bandwidth (e.g., 20 Hz), a relationship (e.g., a 100× reduction) that holds both for computation time and for the number of input/output (I/O) calls—per iteration (noting that source encoding might require more iterations). Their crosstalk-free source-encoded FWI method has been successfully applied at the global scale of earthquake tomography (Cui et al., 2023), and at the human scale of 3-D ultrasound tomography (Bachmann and Tromp, 2020).

The main drawback of source-encoded formulations arises from the difficulty of *time-windowing* modeled data when inverting one or a few sparsely sampled *frequencies* at one time: time- and frequency-limitation are intrinsically incompatible (Simons, 2010). Yet, time-windowing allows for the balanced selection of specific arrivals during the various stages of the inversion, which is often desirable in order to navigate the misfit surface towards a global optimum. In this paper, we propose the use of complex-valued frequencies through the Laplace transform (Shin et al., 2002; Brenders and Pratt, 2007; Brossier et al., 2009; Prieux et al., 2013a), which damps arrivals at a certain rate γ starting from a given traveltime t_0, to be defined below. We develop the theory of crosstalk-free source-encoded elastic FWI in the Laplace domain and illustrate it with numerical examples.

ELASTIC FWI IN THE LAPLACE DOMAIN

In this paper, we extend the theory of crosstalk-free source-encoded elastic full-waveform inversion. Our treatment largely follows the adjoint formulation laid out by Tromp et al. (2005) and Tromp and Bachmann (2019), but we enable damping of the seismogram beyond a certain time offset (both to target specific seismic phases and to ensure stability), by switching from the Fourier domain to the Laplace domain, involving complex frequencies.

For a generic time-domain function $g(t)$, we write the one-sided Laplace transform at complex argument

$$z = \gamma + \imath\,\omega, \qquad (5)$$

where γ and ω are real, as

$$\widetilde{g}(z) = \mathcal{L}[g(t)](z) = \int_0^\infty g(t)\,e^{-zt}\,\mathrm{d}t. \qquad (6)$$

The objective function

Let us consider a source location $\mathbf{x}_s$ and a source-time function $f(t)$ which drives a body-force point source acting in the $\hat{\mathbf{n}}$ direction with components

$$f_j(\mathbf{x}_s, t) = \hat{n}_j(\mathbf{x}_s) f(t), \qquad (7)$$

which gives rise to displacement components at a receiver located at $\mathbf{x}_r$ through convolution with the elastic Green's function $G_{ij}(\mathbf{x}_r, \mathbf{x}_s; t)$ in the form, whereby Einstein's summation convention is implied,

$$u_i(\mathbf{x}_r, t) = \int_0^t G_{ij}(\mathbf{x}_r, \mathbf{x}_s; t - t') f_j(\mathbf{x}_s, t')\,\mathrm{d}t'. \qquad (8)$$

With Laplace transforms of the synthetic, the source, and the Green's function given by

$$\widetilde{u}_i(\mathbf{x}_r, z) = \mathcal{L}[u_i(\mathbf{x}_r, t)], \qquad (9)$$

$$\widetilde{f}_j(\mathbf{x}_s, z) = \mathcal{L}[f_j(\mathbf{x}_s, t)], \qquad (10)$$

$$\widetilde{G}_{ij}(\mathbf{x}_r, \mathbf{x}_s; z) = \mathcal{L}[G_{ij}(\mathbf{x}_r, \mathbf{x}_s; t)], \qquad (11)$$

the Laplace convolution theorem transforms equation 8 to the equivalent expression, again summed over j,

$$\widetilde{u}_i(\mathbf{x}_r, z) = \widetilde{G}_{ij}(\mathbf{x}_r, \mathbf{x}_s; z)\widetilde{f}_j(\mathbf{x}_s, z). \qquad (12)$$

The scaled Laplace coefficient of $u_i(\mathbf{x}_r, t)$, where the damping is given by γ and the shift is over the arbitrary constant t_0^{sr}, e.g., the first-arrival time of the elastic waves from source $\mathbf{x}_s$ to receiver $\mathbf{x}_r$, is given by

$$\mathcal{L}\left[e^{\gamma t_0^{sr}} u_i(\mathbf{x}_r, t)\right](z) = e^{\gamma t_0^{sr}}\,\widetilde{u}_i(\mathbf{x}_r, z). \qquad (13)$$

Because we will associate every source at $\mathbf{x}_s$ recorded by a receiver at $\mathbf{x}_r$ with a unique complex source frequency z_s, we introduce the shorthand notation

$$\widetilde{f}_j^s \equiv \widetilde{f}_j(\mathbf{x}_s, z_s), \qquad (14)$$

$$\widetilde{G}_{ij}^s(\mathbf{x}_r, \mathbf{x}) \equiv \widetilde{G}_{ij}(\mathbf{x}_r, \mathbf{x}; z_s), \qquad (15)$$

$$\widetilde{u}_i^{sr} \equiv e^{\gamma t_0^{sr}}\,\widetilde{u}_i(\mathbf{x}_r, z_s), \qquad (16)$$

$$\widetilde{d}_i^{sr} \equiv e^{\gamma t_0^{sr}}\,\widetilde{d}_i(\mathbf{x}_r, z_s), \qquad (17)$$

where equation 17 applies to the Laplace transform of the observed data $d_i(\mathbf{x}_r, t)$.

The amplitude damping factor γ and the source-receiver dependent time shift t_0^{sr} (as might be picked from the data or calculated in an initial model) effectively allow for time windowing of the seismograms through the Laplace transform, ensuring Green's function recovery through orthogonality beyond steady-state time, as will be clarified below. To ensure that the discrete Laplace transform of the observed data contains exactly the encoded complex frequencies z_s, we may zero-pad or truncate the time series to be able to evaluate the Laplace transform

$$\widetilde{d}_i^{sr} = e^{\gamma t_0^{sr}} \int_0^{\Delta\tau} d_i(\mathbf{x}_r, t) e^{-z_s t}\, \mathrm{d}t, \tag{18}$$

$$= \int_0^{\Delta\tau} d_i(\mathbf{x}_r, t) e^{-\gamma(t - t_0^{sr})} e^{-\imath\omega_s t}\, \mathrm{d}t, \tag{19}$$

over the integration interval $\Delta\tau$, to be specified below.

The modeling residuals to be minimized are the differences between the scaled Laplace coefficients,

$$\Delta\widetilde{u}_i^{sr} = \widetilde{u}_i^{sr} - \widetilde{d}_i^{sr}. \tag{20}$$

The overall misfit is defined as the sum of squares, with the "$*$" denoting complex conjugation,

$$\chi = \sum_{s=1}^{S}\sum_{r=1}^{R}\sum_{i=1}^{C} w_{sri}\, \Delta\widetilde{u}_i^{sr*} \Delta\widetilde{u}_i^{sr}, \tag{21}$$

accounting for $s = 1, ..., S$ different frequencies or sources, $r = 1, ..., R$ receivers, and $i = 1, ..., C$ sensor components. The weights w_{sri} are nonzero only for those receivers r that record a source (frequency) s on a component i. Note the slight variation in notation compared to equations 2–3.

The objective of FWI is to adjust the Earth model, embodied by the Green's function G, in such a manner that synthetic predictions u made for observations d are ideally matched. When carried out in the least-squares sense, we require ways to minimize equation 21. See Appendix A for alternative misfit formulations.

The misfit variation

To perform the minimization of the misfit criterion χ in equation 21 we begin by considering how it varies under Earth model perturbations. The variation of the misfit is

$$\delta\chi = \Re \sum_{s=1}^{S}\sum_{r=1}^{R}\sum_{i=1}^{C} w_{sri}\, \Delta\widetilde{u}_i^{sr*}\, \delta\widetilde{u}_i^{sr}. \tag{22}$$

Under the Born approximation (Nolet, 2008), in terms of the density and stiffness perturbations $\delta\rho$ and δc_{jklm}, the time-domain perturbed displacement (Tromp et al., 2005)

$$\begin{aligned}\delta u_i(\mathbf{x}_r, t) &= \\ -\int_0^t \int_V &\left[\delta\rho(\mathbf{x})\, G_{ij}(\mathbf{x}_r, \mathbf{x}; t - t')\, \partial_{t'}^2 u_j(\mathbf{x}, t')\right. \\ &\left. + \delta c_{jklm}(\mathbf{x})\, \partial_k G_{ij}(\mathbf{x}_r, \mathbf{x}; t - t')\, \partial_l u_m(\mathbf{x}, t')\right]\, \mathrm{d}^3\mathbf{x}\, \mathrm{d}t',\end{aligned} \tag{23}$$

whereby Einstein's summation convention is now implied for the repeat indices j, k, l, m, and from hereon out. Using the convolution identity and the derivative properties of the Laplace transform, the scaled Laplace coefficients of the perturbed wavefield are

$$\begin{aligned}\delta\widetilde{u}_i^{sr} = -\, e^{\gamma t_0^{sr}} \int_V &\Big[\delta\rho(\mathbf{x})\, \widetilde{G}_{ij}^s(\mathbf{x}_r, \mathbf{x})\, z_s^2\, \widetilde{u}_j(\mathbf{x}, z_s) \\ &+ \delta c_{jklm}(\mathbf{x})\, \partial_k \widetilde{G}_{ij}^s(\mathbf{x}_r, \mathbf{x})\, \partial_l \widetilde{u}_m(\mathbf{x}, z_s)\Big]\, \mathrm{d}^3\mathbf{x},\end{aligned} \tag{24}$$

where initial conditions $u_j(\mathbf{x}, 0) = 0$ and $\partial_t u_j(\mathbf{x}, 0) = 0$ are used to reduce the Laplace transform of the second temporal derivative of the displacement.

Inserting the perturbed displacement $\delta\widetilde{u}_i^{sr}$ of equation 24 into equation 22 yields

$$\begin{aligned}\delta\chi = -\Re\, e^{\gamma t_0^{sr}} \int_V &\Big\{\delta\rho(\mathbf{x}) w_{sri} \Delta\widetilde{u}_i^{sr*} \widetilde{G}_{ij}^s(\mathbf{x}_r, \mathbf{x})\, z_s^2\, \widetilde{u}_j(\mathbf{x}, z_s) \\ &+ \delta c_{jklm}(\mathbf{x}) w_{sri} \Delta\widetilde{u}_i^{sr*} \partial_k \widetilde{G}_{ij}^s(\mathbf{x}_r, \mathbf{x})\, \partial_l \widetilde{u}_m(\mathbf{x}, z_s)\Big\}\, \mathrm{d}^3\mathbf{x},\end{aligned} \tag{25}$$

with the summation convention now additionally applying over the indices s, r and i, i.e., over all sources, receivers, and receiver components.

In the following, we focus on deriving the density kernel for simplicity. Elastic kernels can be derived in a similar way. In equation 25, guided by equations 12 and 7, we first substitute $\widetilde{u}_j(\mathbf{x}, z_s) = \widetilde{G}_{jm}^s(\mathbf{x}, \mathbf{x}_s) \widetilde{f}^s \hat{\mathbf{n}}_m(\mathbf{x}_s)$ and, denoting the Laplace coefficient of the source-time function as $\widetilde{f}^s$, we rewrite $\widetilde{f}^s = -\imath^2 \widetilde{f}^s$ to obtain

$$\begin{aligned}\delta\chi_\rho &= -\int_V \delta\rho(\mathbf{x})\, \Re\left\{e^{\gamma t_0^{sr}} w_{sri} \Delta\widetilde{u}_i^{sr*} \widetilde{G}_{ij}^s(\mathbf{x}_r, \mathbf{x})\right. \\ &\qquad \left. \times\, z_s^2 \widetilde{G}_{jm}^s(\mathbf{x}, \mathbf{x}_s) f^s \hat{n}_m(\mathbf{x}_s)\right\}\, \mathrm{d}^3\mathbf{x}', \end{aligned}\tag{26}$$

$$= -\int_V \delta\rho(\mathbf{x})\, \Re\left\{e^{\gamma t_0^{sr}}\, \bar{\Delta}\widetilde{u}_i^{sr*} \widetilde{G}_{ij}^s(\mathbf{x}_r, \mathbf{x})\, z_s^2 \bar{\widetilde{u}}_j^s(\mathbf{x})\right\}\, \mathrm{d}^3\mathbf{x}', \tag{27}$$

where we have defined a weighted residual as

$$\bar{\Delta}\widetilde{u}_i^{sr} = -\imath\, w_{sri} f^{s*} \Delta\widetilde{u}_i^{sr}, \tag{28}$$

and redefined the Laplace coefficients of a new forward wavefield to be

$$\bar{\widetilde{u}}_j^s(\mathbf{x}) = -\imath\, \widetilde{G}_{jm}^s(\mathbf{x}, \mathbf{x}_s) \hat{n}_m(\mathbf{x}_s). \tag{29}$$

Multiplying equation 27 by $e^{z_s t} e^{-z_{s'} t}$ and integrating over a suitably shifted time interval of interest allows us to exploit their orthogonality to rewrite equation 27 as

$$\begin{aligned}\delta\chi_\rho = -\frac{2}{\Delta\tau} \int_V \delta\rho(\mathbf{x}) \\ \times \int_0^{\Delta\tau} &\underbrace{\Re\left\{e^{\gamma t_0^{sr}}\, \bar{\Delta}\widetilde{u}_i^{sr*} \widetilde{G}_{ij}^s(\mathbf{x}_r, \mathbf{x}) e^{-z_s(t+T)}\right\}}_{\mathrm{I}} \\ &\times \underbrace{\Re\left\{z_{s'}^2 \bar{\widetilde{u}}_j^{s'}(\mathbf{x}) e^{z_{s'}(t+T)}\right\}}_{\mathrm{II}}\, \mathrm{d}t\, \mathrm{d}^3\mathbf{x}.\end{aligned} \tag{30}$$

Other than the damping parameter γ and the trace-dependent temporal shift t_0^{sr}, which provide opportunities for time-windowing, phase selectivity and numerical stability, equation 30 contains the length of the "decoding" integration interval $\Delta\tau$ which commences beyond the "steady-state" time shift T. These two parameters are briefly introduced in equation 4, but here as there, they await further identification and numerical specification.

Interacting forward and adjoint wavefields

In the second factor (II) of equation 30 we recognize the second time derivative of an encoded forward wavefield component, summed over the sources,

$$u_j(\mathbf{x}, t+T) = \Re\left\{\tilde{\tilde{u}}_j^s(\mathbf{x})e^{z_s(t+T)}\right\}, \tag{31}$$

into which we substitute equation 29, the definition of the Laplace transform to reduce $\widetilde{G}^s_{jm}(\mathbf{x}, \mathbf{x}_s)$, and then rearrange (see Appendix B) to obtain

$$u_j(\mathbf{x}, t+T) = \sum_{s=1}^{S}\int_{-\infty}^{T+t} G_{jm}(\mathbf{x}, \mathbf{x}_s, T+t-t') \times \underbrace{\hat{n}_m(\mathbf{x}_s)e^{\gamma t'}\sin\omega_s t'}_{\text{source}}\,\mathrm{d}t', \tag{32}$$

that is, driven by sources $f_m(\mathbf{x}_s, t)$ with source-time functions $f(t) = e^{\gamma t}\sin(\omega_s t)$.

In the first factor (I) of equation 30, we recognize a wavefield, summed over all sources and receivers, which, using the definition of the Laplace transform, the reciprocity of the Green's function, and a rearranging of the integration limits can be rewritten (see Appendix B) as

$$\phi_j(\mathbf{x}, t+T) = \sum_{r=1}^{R}\int_{-\infty}^{T+\Delta\tau-t} G_{ji}(\mathbf{x}, \mathbf{x}_r, T+\Delta\tau-t-t') \times \underbrace{\sum_{s=1}^{S} e^{\gamma t_0^{sr}}\Re\left\{\bar{\Delta}\tilde{u}_i^{sr*}e^{-z_s(2T+\Delta\tau-t')}\right\}}_{\text{adjoint source}}\,\mathrm{d}t', \tag{33}$$

where we now define an adjoint source that can be simplified (see Appendix B) to

$$f_i^\dagger(\mathbf{x}_r, t) = \Re\left\{\sum_{s=1}^{S} e^{\gamma t_0^{sr}}\,\bar{\Delta}\tilde{u}_i^{sr}e^{\imath\omega_s(T+t)}e^{-\gamma(T+t)}\right\}, \tag{34}$$

$$= \Re\left\{\bar{\Delta}u_i^{rs}(T+t)e^{-\gamma(T+t)}\right\}, \tag{35}$$

where we recognized the shifted inverse Fourier transform

$$\bar{\Delta}u_i^{rs}(t) = \sum_{s=1}^{S} e^{\gamma t_0^{sr}}\,\bar{\Delta}\tilde{u}_i^{sr}e^{\imath\omega_s t}. \tag{36}$$

We now introduce the adjoint wavefield

$$u_j^\dagger(\mathbf{x}, t) = \sum_{r=1}^{R}\int_0^t G_{ji}(\mathbf{x}, \mathbf{x}_r, t-t')f_i^\dagger(\mathbf{x}_r, t')\,\mathrm{d}t', \tag{37}$$

from which we obtain the relationship

$$\phi_j(\mathbf{x}, t+T) = u_j^\dagger(\mathbf{x}, T+\Delta\tau-t). \tag{38}$$

Substituting equations 38 and 32 into equation 30 yields

$$\delta\chi_\rho = -\frac{2}{\Delta\tau}\int_V \delta\rho(\mathbf{x}) \times \left(\int_0^{\Delta\tau} u_j^\dagger(\mathbf{x}', T+\Delta\tau-t)\,\partial_t^2 u_j(\mathbf{x}, t+T)\,\mathrm{d}t\right)\mathrm{d}^3\mathbf{x}. \tag{39}$$

As far as the density perturbation is concerned, equation 39 provides an accessible version for the misfit variation in equation 30. To gain further insight we now write it in the form that emphasizes the *sensitivity kernel* in the manner of Tromp et al. (2005).

Fréchet derivative sensitivity kernels

Equation 39 is in the form

$$\delta\chi_\rho = \int_V \delta\log\rho(\mathbf{x})K_\rho(\mathbf{x})\,\mathrm{d}^3\mathbf{x}, \tag{40}$$

which defines the Fréchet derivative with respect to the mass density as

$$K_\rho(\mathbf{x}) = -\frac{2}{\Delta\tau}\int_0^{\Delta\tau}\rho(\mathbf{x})\,u_j^\dagger(\mathbf{x}, T+\Delta\tau-t)\,\partial_t^2 u_j(\mathbf{x}, t+T)\,\mathrm{d}t. \tag{41}$$

For the stiffness perturbations we ultimately have the similar expression

$$\delta\chi_c = \int_V \delta c_{jklm}(\mathbf{x})K_{c_{jklm}}(\mathbf{x})\,\mathrm{d}^3\mathbf{x}, \tag{42}$$

containing, as in Tromp et al. (2005), the Fréchet kernel with respect to the elastic constants

$$K_{c_{jklm}}(\mathbf{x}) = -\frac{2}{\Delta\tau}\int_0^{\Delta\tau}\epsilon_{jk}^\dagger(\mathbf{x}, T+\Delta\tau-t) \times \epsilon_{lm}(\mathbf{x}, t+T)\,\mathrm{d}t, \tag{43}$$

where no summation is implied, and $\epsilon_{lm} = \frac{1}{2}[\partial_l u_m + \partial_m u_l]$ and $\epsilon_{jk}^\dagger = \frac{1}{2}[\partial_j u_k^\dagger + \partial_k u_j^\dagger]$ denote the elements of the strain and the adjoint strain tensors, respectively.

We return to equation 25 for the total misfit variation, $\delta\chi = \delta\chi_\rho + \delta\chi_c$. In an isotropic medium with elastic tensor $c_{jklm} = (\kappa - 2\mu/3)\delta_{jk}\delta_{lm} + \mu(\delta_{jl}\delta_{km} + \delta_{jm}\delta_{kl})$, equations 40 and 43 combine into the expression

$$\delta\chi = \int_V [\delta\log\rho(\mathbf{x}')K_\rho(\mathbf{x}') + \delta\log\kappa(\mathbf{x}')K_\kappa(\mathbf{x}') + \delta\log\mu(\mathbf{x}')K_\mu(\mathbf{x}')]\,\mathrm{d}^3\mathbf{x}', \tag{44}$$

where, in addition to the density kernel 41, we have the bulk and shear modulus kernels

$$K_\kappa(\mathbf{x}) = -\frac{2}{\Delta\tau}\int_0^{\Delta\tau}\kappa(\mathbf{x})\,\partial_i u_i^\dagger(\mathbf{x}, T+\Delta\tau-t) \times \partial_j u_j(\mathbf{x}, t+T)\,\mathrm{d}t, \tag{45}$$

$$K_\mu(\mathbf{x}) = -\frac{2}{\Delta\tau}\int_0^{\Delta\tau} 2\mu(\mathbf{x})D_{ij}^\dagger(\mathbf{x}, T+\Delta\tau-t) \times D_{ij}(\mathbf{x}, t+T)\,\mathrm{d}t, \tag{46}$$

whereby $D_{ij}(\mathbf{x}, t) = \frac{1}{2}(\partial_i U_j + \partial_j U_i) - \frac{1}{3}\partial_k U_k \delta_{ij}$ and $D^\dagger_{ij}(\mathbf{x}, t) = \frac{1}{2}(\partial_i U^\dagger_j + \partial_j U^\dagger_i) - \frac{1}{3}\partial_k U^\dagger_k \delta_{ij}$ are the strain deviators associated with the forward and adjoint wavefields.

The formalism developed so far, in particular equations 41 and 43 for the density and elastic kernels, require numerical evaluation. In particular, they require reconciliation with the philosophy espoused in the Introduction. Following equation 4, single-frequency wavefields calculations can be carried out, and subsequently recombined, to return gradients for the FWI optimization without crosstalk. Naturally, this will entail the discussion of the steady-state time T and the decoding interval $\Delta\tau$.

COMPUTATIONAL CONSIDERATIONS

While frequency-domain numerical modeling methods (Pratt, 1999) can be brought to bear on our problem, Nihei and Li (2007) propose an efficient time-domain method that assigns a single, unique, frequency to each shot and simulates all shots simultaneously without crosstalk. Their approach is key to our source-encoded full-waveform strategy (Zhang et al., 2018; Tromp and Bachmann, 2019), as we further motivate, and specify.

Single-frequency time-domain modeling

Let $G(t)$ denote a Green's function for wave propagation from a certain source to a certain receiver, without specificity to simplify the notation. Its Fourier transform is

$$\tilde{G}(\omega) = \int_0^\infty G(t) e^{-\imath\omega t}\,\mathrm{d}t, \tag{47}$$

where ω denotes the continuous angular frequency. Over a finite time interval $\Delta\tau$, to be specified, the discrete angular frequency will be $\omega_k = 2\pi k/\Delta\tau$, for integer $k = 1, \ldots, K$. The relation between these time and frequency properties will be made explicit below.

The response to a single-frequency source-time function $f_k(t) = \cos(\omega_k t) = \Re\{e^{\imath\omega_k t}\}$, through equation 47, can be equivalently expressed as

$$\begin{aligned} u_k(t) &= \Re\left\{\tilde{G}(\omega_k)\, e^{\imath\omega_k t}\right\}, && (48)\\ &= \int_0^{\Delta\tau} G(t - t')\cos(\omega_k t')\,\mathrm{d}t', && (49)\\ &= [G \star f_k](t). && (50)\end{aligned}$$

The star "$\star$" denotes convolution. Hence we can obtain $u_k(t)$ by solving the wave equation using a time-domain numerical solver driven by a monochromatic source-time function. The single-frequency $\tilde{G}(\omega_k)$ can be recovered via the inner product of equation 48 with respect to the $e^{\pm\imath\omega_k t}/\Delta\tau$ orthonormal basis for the interval $\Delta\tau$, shifted by the steady-state time T, to be defined. By virtue of orthogonality, this integration yields

$$\tilde{G}(\omega_k)\,\delta_{kk'} = \frac{2}{\Delta\tau}\int_T^{T+\Delta\tau} u_k(t)\, e^{-\imath\omega_{k'} t}\,\mathrm{d}t. \tag{51}$$

The numerical evaluation of equation 51 may be performed using the Fast Fourier Transform (FFT), taking care of the factors $(2\mathrm{d}t/\Delta\tau)\exp(-\imath\omega_{k'}T)$. In practice, and to increase numerical stability, we may also change the source-time function from $\cos(\omega_k t)$ to $\sin(\omega_k t)$, for which we multiply equation 51 by $\imath = \sqrt{-1}$.

The steady-state time T is when the value of the Green's function becomes vanishingly small, after which virtually no more seismic waves reach the recording station (Cui et al., 2023). Depending on the Earth model, the source-receiver configuration, and the numbers and types of the seismic sources, steady state may be difficult to achieve, a situation that can be remedied by the damping of the Green's function as discussed further below.

Figure 1 shows a variety of example geometries. In Figure 1a there is precisely one source, at $\mathbf{x}_1$, and only one receiver, at $\mathbf{x}_r$. The single-frequency source-time function $f_1(t) = \cos(\omega_1 t)$, see Figure 2a, the Green's function $G_1(t) = G(\mathbf{x}_r, \mathbf{x}_1; t)$, as shown in Figure 2b, and the seismic response, the convolution of $G_1(t)$ with $f_1(t)$, is $u(t) = u_1(t)$, as shown in Figure 2c. The Fourier terms of the Green's function $\tilde{G}_1(\omega_k)$ at different frequencies k are depicted by blue circles in Figure 2d, and $\tilde{G}_1(\omega_1)$, recovered via equation 51, is shown as a red diamond.

Multiple-source encoding and decoding

Figure 1b shows the situation where one receiver at $\mathbf{x}_r$ is activated simultaneously by two sources at $\mathbf{x}_1$ and $\mathbf{x}_2$. The

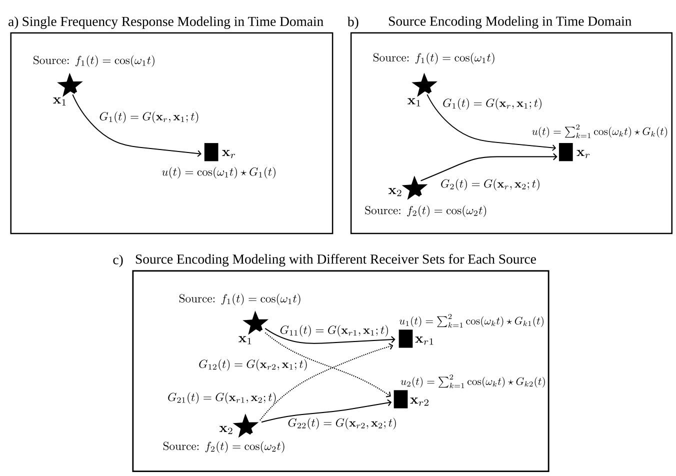

Figure 1: Crosstalk-free time-domain source encoding in a variety of acquisition geometries. (a) Single source $\mathbf{x}_1$, single receiver $\mathbf{x}_r$, single frequency ω_1. (b) Dual sources $\mathbf{x}_1$ and $\mathbf{x}_2$, single receiver $\mathbf{x}_r$, two distinct frequencies, ω_1 and ω_2. (c) Two sources $\mathbf{x}_1$ and $\mathbf{x}_2$, two receivers $\mathbf{x}_{r1}$ and $\mathbf{x}_{r2}$, two unique frequencies ω_1 and ω_2. In (a) and (b), $f_k(t) = \cos(\omega_k t)$, $k = 1, 2$, denote monochromatic source-time functions; $G_k(t) = G(\mathbf{x}_r, \mathbf{x}_k; t)$ denote Green's functions from sources at $\mathbf{x}_k$ to receivers at $\mathbf{x}_r$; $u(t)$ the full seismic response. In (c), $G_{ki}(t) = G(\mathbf{x}_{ri}, \mathbf{x}_k, t)$ denote Green's functions from sources at $\mathbf{x}_k$ to receivers at $\mathbf{x}_{ri}$.

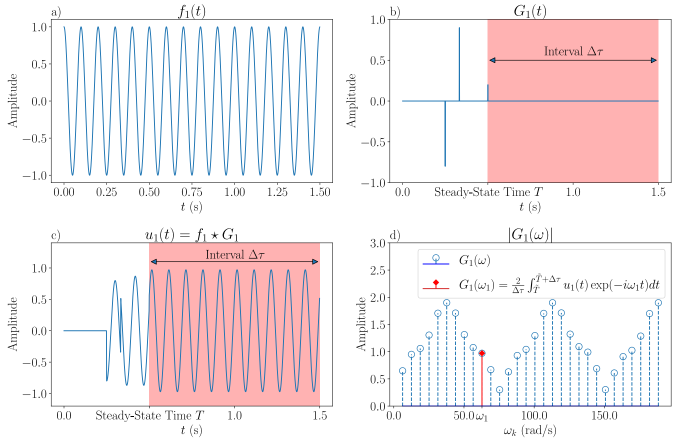

Figure 2: Time-domain source encoding and decoding for the acquisition geometry shown in Figure 1a. (a) Monochromatic source-time function with angular frequency ω_1. (b) Time-domain Green's function $G_1(t)$, which vanishes beyond the steady-state time T. (c) Synthetic observations $u_1(t)$. (d) Frequency-domain Green's function $G_1(\omega)$, highlighting the Fourier coefficient at ω_1 recovered using equation 51. The recording length of the observed data is $\Delta\tau$, and the angular frequency $\omega_1 = 2\pi n_1/\Delta\tau$ for integer $n_1 > 0$.

source-time functions for the two sources $f_1(t) = \cos(\omega_1 t)$ and $f_2(t) = \cos(\omega_2 t)$ are shown in Figures 2a and 3a, respectively. The Green's functions to the receiver $\mathbf{x}_r$ are $G_1(t) = G(\mathbf{x}_r, \mathbf{x}_1; t)$ and $G_2(t) = G(\mathbf{x}_r, \mathbf{x}_2; t)$, respectively, as shown in Figures 2b and 3b. The "supergather" $u(t)$ as shown in Figure 3c is the superposition of two terms of the form of equations 48–50,

$$u(t) = \sum_{k=1}^{2} u_k(t) = \sum_{k=1}^{2} G_k \star f_k. \tag{52}$$

Fourier coefficients $\tilde{G}_1(\omega_1)$ and $\tilde{G}_2(\omega_2)$ can be recovered without crosstalk according to equation 51. The recovery is illustrated numerically in Figure 3d and 3e.

Now suppose that there are K frequencies within the frequency band $[\omega_{\text{min}}, \omega_{\text{max}}]$, with

$$\Delta\omega = (\omega_{\text{max}} - \omega_{\text{min}})/(K-1), \tag{53}$$

which, at last, defines the integration or decoding interval $\Delta\tau$ for the seismic signal as

$$\Delta\tau = 2\pi/\Delta\omega = \frac{2\pi(K-1)}{\omega_{\text{max}} - \omega_{\text{min}}}. \tag{54}$$

Our goal is to perform one "super" forward simulation combining all sources by effectively tagging each individual frequency. If there were as many physical sources as there are discrete frequencies, each source would be randomly assigned a monochromatic source time function with an angular frequency ω_k defined by

$$\omega_k = \omega_{\text{min}} + (k-1)\Delta\omega, \qquad k = 1, \ldots, K. \tag{55}$$

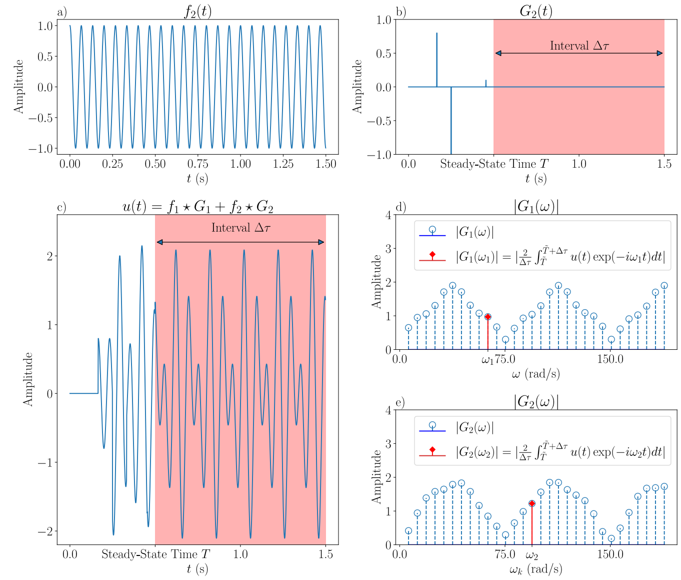

Figure 3: Crosstalk-free time-domain source encoding and decoding for the acquisition geometry shown in Figure 1b. (a) Monochromatic source-time function at frequency ω_2. (b) Time-domain Green's function $G_2(t)$, which vanishes beyond the steady-state time T. (c) Synthetic "supergather", the complete seismic response $u(t)$. (d–e) Frequency-domain Green's functions $\tilde{G}_1(\omega)$ and $\tilde{G}_2(\omega)$, highlighting the Fourier coefficients $\tilde{G}_1(\omega_1)$ in (d) and $\tilde{G}_2(\omega_2)$ in (e), recovered using equation 51. The integration interval is the record length $\Delta\tau$, and the second frequency $\omega_2 = 2\pi n_2/\delta\tau$ with $n_2 > 0$ integer and $n_1 \neq n_2$.

In practical application, the set of frequencies determined by equation 55 is randomly distributed over the available sources at the start of every iteration. As K is the number of frequencies rather than distinct sources, individual sources may be assigned more than one frequency. When the number of sources is very large, as with streamer data or large-scale nodal deployments, the frequency spacing can be reduced by increasing the interval $\Delta\tau$ without loss of orthogonality.

The duration of all observed seismic data may not be equal to $\Delta\tau$, in which case the frequency spacing of the observations may not equal that of the synthetics, $\Delta\omega$. In those cases, we may resort to truncation or zero-padding of the real data before Fourier transformation. If this unduly increases the computation time or storage requirements, we may first apply the discrete Fourier transform to the data directly, without zero-padding or truncating, and subsequently use Fourier interpolation to achieve similar results (Bachmann and Tromp, 2020). In the examples shown we use zero-padding or truncating, as required.

Figure 1c depicts the general case where multiple sources are recorded by multiple receivers, a common acquisition geometry in earthquake seismology or with marine streamer surveys where sources and receivers move for every shot, possibly combined with ocean-bottom node (OBN) deployments. Hence with simultaneous simulation, each receiver will record waves emanating from all sources. When there are two sources and two receivers, the data recorded at $\mathbf{x}_{r1}$ and $\mathbf{x}_{r2}$ are

$$u_1(t) = \sum_{k=1}^{2} G_{k1} \star f_k, \tag{56}$$

$$u_2(t) = \sum_{k=1}^{2} G_{k2} \star f_k, \tag{57}$$

where $G_{k1}(t)$ and $G_{k2}(t)$ are the Green's function from the source at $\mathbf{x}_k$ to the receivers $\mathbf{x}_{r1}$ and $\mathbf{x}_{r2}$, respectively. Again, equation 51 shows how to integrate to $u_1(t)$ and $u_2(t)$ against $e^{-\imath\omega_1 t}$ to obtain $\tilde{G}_{11}(\omega_1)$ and $\tilde{G}_{12}(\omega_1)$, respectively. Similarly, integrating against $e^{-\imath\omega_2 t}$ will yield $\tilde{G}_{22}(\omega_2)$ and $\tilde{G}_{21}(\omega_2)$. Upon doing so we may choose to retain only $\tilde{G}_{11}(\omega_1)$ and $\tilde{G}_{22}(\omega_2)$, choosing to discard $G_{12}(\omega_1)$ and $G_{21}(\omega_2)$.

Time-selective single-frequency modeling

In the methodology outlined above, the steady-state time T should be reached for recovery of the single-frequency Green's function. Reaching steady state may be challenging, to the detriment of source-encoded FWI. Damping by $\gamma > 0$ provides a solution. The Fourier transform of the damped Green's function $G(t)$ is the Laplace transform

$$\widetilde{G}(z = \imath\omega + \gamma) = \mathcal{L}[G(t)](z) \tag{58}$$

$$= \int_0^\infty G(t) e^{-\gamma t} e^{-\imath\omega t}\, \mathrm{d}t \tag{59}$$

$$= \int_0^\infty G(t) e^{-zt}\, \mathrm{d}t. \tag{60}$$

Next consider the response to a cosinusoidal source-time function with exponentially increasing amplitude, $f_k'(t) = e^{\gamma t}\cos(\omega_k t) = \Re\{e^{z_k t}\}$. The corresponding wavefield

$$u_k'(t) = \Re\{\widetilde{G}(z_k)\, e^{z_k t}\}, \tag{61}$$

$$= \int_0^{\Delta\tau} G(t - t')\, e^{\gamma t'} \cos(\omega_k t')\, \mathrm{d}t', \tag{62}$$

$$= [G \star f_k'](t), \tag{63}$$

which again can be calculated using a time-domain solver. In analogy with equation 51, the single complex-frequency Green's function can now be recovered as

$$\widetilde{G}(z_k)\delta_{kk'} = \frac{2}{\Delta\tau} \int_T^{T+\Delta\tau} u_k'(t) e^{-z_{k'} t}\, \mathrm{d}t. \tag{64}$$

As discussed, to provide phase selectivity as well as numerical stability, usually, we will damp $G(t)$ by $e^{-\gamma(t-t_0^{sr})}$, with t_0^{sr} the arrival time for waves traveling from source $\mathbf{x}_s$ to receiver $\mathbf{x}_r$. The Fourier transform of the Green's function damped in time starting from t_0^{sr}, assuming it is zero before that, is its simply scaled Laplace transform,

$$\mathcal{L}\left[e^{\gamma t_0^{sr}} G(t)\right](z) = e^{\gamma t_0^{sr}} \widetilde{G}(z). \tag{65}$$

The Laplace transform of the convolution between functions is the product of their Laplace transforms. The Laplace transform of the nth derivative of a function brings out the nth power of the arguments, in particular,

$$\mathcal{L}\left[\partial_t^2 u(t)\right](z) = z^2 \widetilde{u}(z) - z\,u(0) - \partial_t u(0). \tag{66}$$

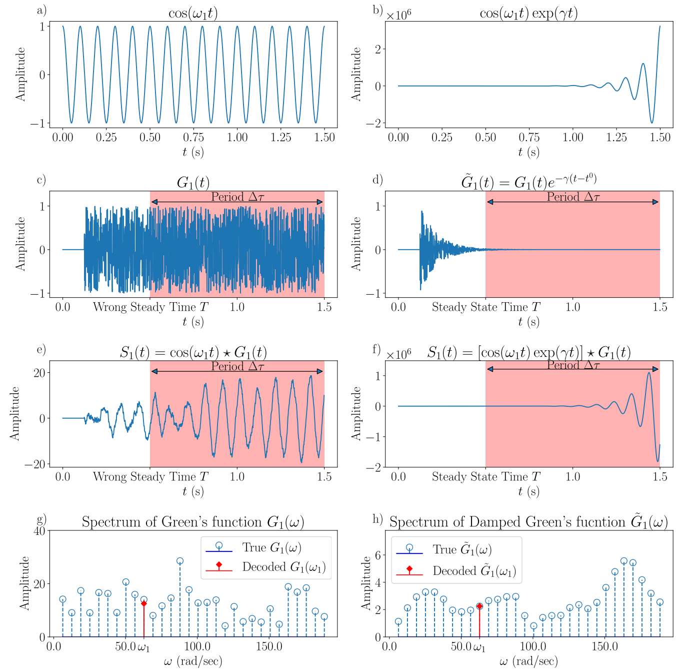

Figure 4: Crosstalk-free damped time-domain source encoding and decoding. (a) Monochromatic source-time function of angular frequency ω_1. (b) Same as (a) but with an exponentially increasing amplitude $e^{\gamma t}$. (c) Green's function $G_1(t)$, and an inappropriate steady-state time T. (d) Green's function $G_1(t)e^{-\gamma(t-t_0^{sr})}$ damped using $\gamma = 10\ s^{-1}$ and with offset time t_0^{sr}. (e)–(f) Synthetics obtained by convolution of the Green's function in (c) with the source-time function in (a) and (b), respectively. (g)–(h) Frequency-domain versions of the Green's function in (c), $|\widetilde{G}_1(\omega)|$, and of the damped Green's function in (d), $|\widetilde{G}_1(\omega)|$, highlighting the recovered Fourier coefficients.

Figure 4 shows a worked example of single complex-frequency forward model. Figure 4a and 4b show source-time functions $f_1(t) = \cos(\omega_1)$ and $f_1'(t) = e^{\gamma t}\cos(\omega_k t)$, Figure 4c and 4d the Green's functions $G_1(t)$ and $G_1(t)e^{-\gamma(t-t_0^{sr})}$, Figure 4e and 4f the wavefields, and Figure 4g and 4h the recovery of the single-frequency Fourier terms. Figure 4g shows the failure of decoding the Fourier coefficient $G_1(\omega_1)$ due to an inappropriate steady-state time T for $G_1(t)$ in Figure 4c. For a damping factor $\gamma = 10\ s^{-1}$, the damped Green's function is shown in Figure 4d and the corresponding Fourier coefficients are shown as blue dashed stems in Figure 4h. Alternatively, we can convolve the source-time function $\cos(\omega_1 t)e^{\gamma t}$ in Figure 4b with $G_1(t)$ in Figure 4c to get the encoded data $u_1'(t)$ in Figure 4f. We can recover the Fourier coefficient at $\omega = \omega_1$ of the damped Green's function using equation 64 (the red stem in Figure 4h).

Damping the Green's function not only mutes late arrivals but also reduces the steady-state time. The time windowing

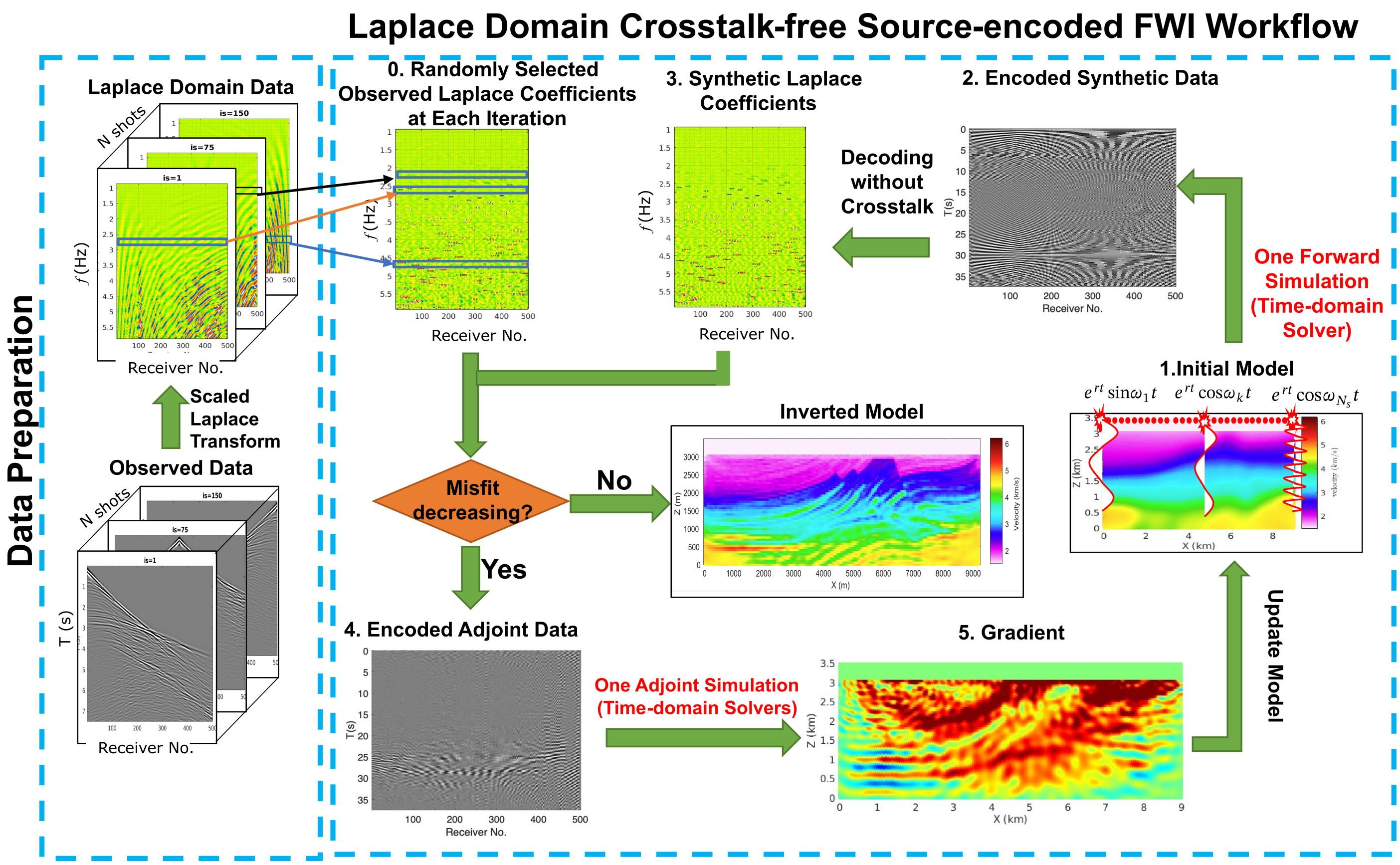

Figure 5: Schematic of the Laplace-domain source-encoded full-waveform inversion method developed in this paper.

operation allows selection of specific arrivals during the stages of the inversion, as an effective means to mitigate the non-linearity of the inversion. For example, we may begin an inversion sequence in the Fourier domain, without any damping ($\gamma = 0$), in order to capture the large-amplitude surface waves and constrain the shallow velocity model, and later switch to the Laplace domain using damping ($\gamma > 0$) to focus on the early arrivals. In any case the damping factor γ should be set appropriately so that the maximum value of the simulated wavefield does not exceed the valid range of the data type in which the calculations are carried out. A rule of thumb is $\gamma(T + \Delta\tau) < 22$ for 4-byte data (`float32`).

Inversion workflow

Our workflow is shown in Figure 5. We begin by transforming the observed data to the Laplace domain, choosing a damping parameter γ and source-receiver-dependent time shift t_0^{sr}, obtaining scaled Laplace coefficients as in equations 17–19. Subsequently, we:

1. select random disjoint sets of scaled data Laplace coefficients for each source, $\widetilde{d}_i^{sr}$;

2. activate all sources with source-time functions in the form of a (co)sine with amplitude exponentially increasing according with the (un-)damping parameter γ, as in equations 61–63;

3. carry out a source-encoded forward simulation until the wavefield reaches steady state after a time T, and obtain the encoded synthetic data that will be decoded owing to the orthogonality between the trigonometric (exponential) phase encoding terms;

4. encode the Laplace-domain data residuals 20 to obtain the adjoint source 34–36;

5. carry out one source-encoded adjoint simulation and calculate the gradients K as in equations 41 and 45–46 through zero-lag cross-correlation between the steady-state encoded forward and adjoint wavefields over interval $\Delta\tau$, proportional to the inverse of the encoded frequency spacing $\Delta\omega$, as in equation 54.

FREQUENCY ASSIGNMENT STRATEGIES

Frequency-domain FWI is traditionally conducted by inverting single frequencies in succession from low to high, for all sources (Pratt, 1999; Sirgue and Pratt, 2004). In contrast, source-encoded FWI assigns to each source only one, or a few, frequencies at a time. A strategy needs to be designed to ensure that an adequate number of frequencies (indeed: all) are used for every source across the iterations. In the approach by Dai et al. (2013), each source is initially allocated a unique frequency within the band of interest. All source frequencies are increased by $\Delta\omega$ in subsequent iterations, wrapping around

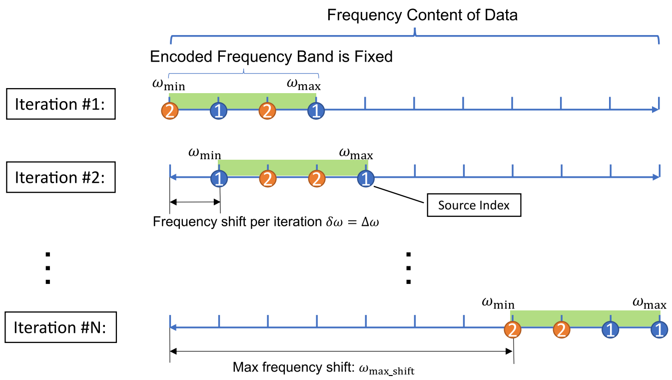

Figure 6: The Moving Frequency Band (MFB) method for source-encoding frequency assignment, a multiscale strategy that involves consecutive inversions of slightly overlapping frequency groups of fixed bandwidth. In this example, the shift $\delta\omega = \Delta\omega$, the frequency spacing. In each iteration, the frequencies assigned to the sources are randomly selected.

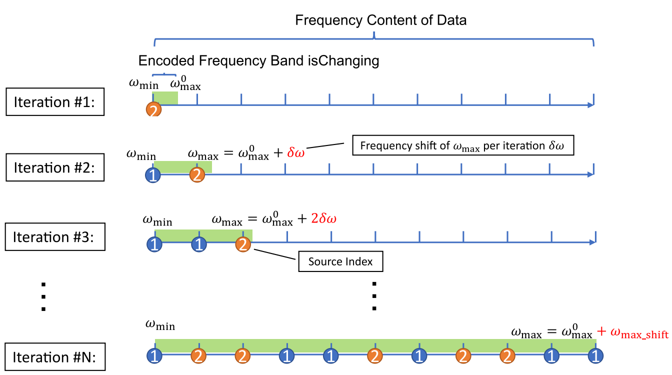

Figure 7: The multiscale strategy for source-encoded frequency assignment, a multiscale strategy that involves successive inversions of overlapping frequency groups whose lowest frequency is fixed. Each subsequent group adds one higher frequency. The frequencies assigned to the sources are randomly selected for each iteration.

upon reaching the bandlimit, back to the minimum frequency. Only $(\omega_{\text{max}} - \omega_{\text{min}})/\Delta\omega$, iterations are needed to sample all frequencies for every source, hence there will be welcome sampling redundancy across the frequency spectrum.

In contrast to this type of systematic frequency sweep, we will adopt a random frequency assignment strategy, as proposed by Huang and Schuster (2018), Zhang et al. (2018) and Tromp and Bachmann (2019). Random frequency assignment is commonly combined with multiscale strategies in order to mitigate the nonlinearity inherent in FWI. We aim to utilize a wide frequency band to accommodate many sources without reducing the spacing $\Delta\omega$, which would increase simulation times. However, due to the nonlinearity of the inverse problem and under the threat of cycle-skipping, it may be necessary to use a narrower band in the early inversion stages. Hence the bandwidth will reflect a compromise between avoiding cycle-skips and simultaneously inverting as many sources and frequencies as possible. Ultimately, all are used.

Tromp and Bachmann (2019) propose the multiscale frequency selection strategy depicted in Figure 6. At the outset they choose a bandwidth and a frequency spacing $\Delta\omega$ that enables encoding all sources using equation 53. This initial interval is shifted up by $\delta\omega$ with each subsequent iteration, until the maximum frequency reaches the overall maximum desired. In our implementation, $\delta\omega$ could be a fraction of $\Delta\omega$, allowing for the bandwidth to stay constant for a set number of iterations. This approach is referred to as the “Moving Frequency Band” (MFB) method. Brossier et al. (2009) adapt the time-domain Bunks et al. (1995) approach to frequency selection as shown in Figure 7. They conduct successive inversions in widening frequency bands. The first iteration involves only the starting frequency interval, to which one frequency is added with every iteration. This approach is known as the “Bunks” method.

To further mitigate the nonlinearity of FWI, our Laplace-domain formulation allows for a subset of specific arrivals, e.g., early or reflected phases, to be used selectively, using the time offset t_0^{sr} and damping factors γ by which to control the time windowing. The damping can be adjusted, e.g., to keep converted waves, free-surface multiples, or surface waves from adding to the nonlinearity of the inversion in the early iterations. For example, FWI of land data presents challenges due to increased nonlinearity caused by free-surface effects, including the propagation of surface waves, in the heterogeneous near-surface. With surface waves, multi-offset strategies may be necessary, as cycle-skipping accumulates nonlinearity with increasing offset (Liu and Huang, 2019; Borisov et al., 2020).

NUMERICAL VALIDATION

Offshore Coupled Acoustic-Elastic FWI of Marmousi Streamer Data

We demonstrate the applicability of our workflow to a synthetic streamer data set generated from the Marmousi model (Versteeg, 1994; Modrak and Tromp, 2016), overlain by a 450 m water layer. Both sources and receivers move with every shot. The shear-wave speed (V_S) model is derived from the compressional (V_P) model using a constant Poisson’s ratio of 0.25, $V_S = V_P/1.732$. The density is assumed to be uniform at 1000 kg/m^3 and is not updated during the inversion. An absorbing boundary condition is applied around the model without adapting it to the free surface at the top of the water layer.

Figure 8a shows the compressional wave speed model, our inversion target. The initial V_P and V_S models for the inversion are formed by smoothing the true models with a 2-D Gaussian with vertical and horizontal standard deviations of 848.5 m, see Figure 8b. We used a lower starting frequency of 1.0 Hz, and employ the exponentiated phase cost function (Fu et al., 2018; Yuan et al., 2019) as defined in equation A-18. A first set of inversion tests is conducted without damping terms ($\gamma = 0$), that is, involving all the arrivals in the inversion.

The model dimensions are 9.2 km$\times$3.5 km. We simulate a 2-D towed-streamer seismic acquisition, consisting of 148 evenly distributed pressure sources excited 10 m below the water surface, with 61 m shot spacing. The streamer length is 8 km, and

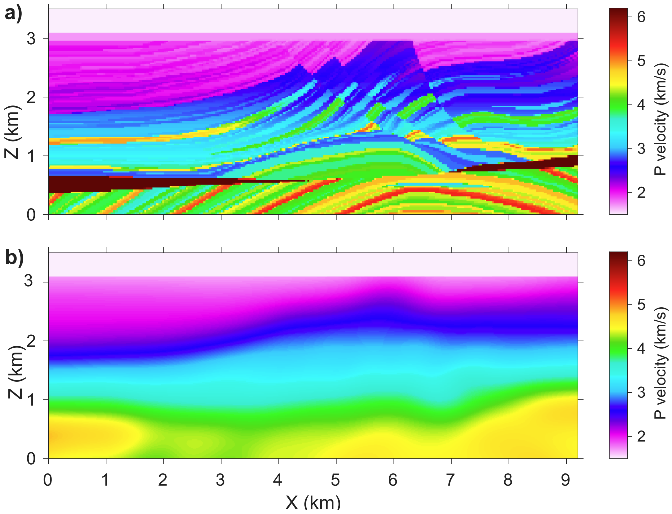

Figure 8: (a) True and (b) initial V_P models (Marmousi).

includes 438 hydrophones with a spacing of 18 m at a depth of 10 m. The distance between the sources and the first recording channel is 30 m. The source-time function is a Ricker wavelet with a center frequency of 5 Hz, considered known. The maximum recording time is 7.5 s, the time step is 0.75 ms. The observed and synthetic data are computed with the 2-D spectral-element method `SPECFEM2D` (Komatitsch and Vilotte, 1998). The corresponding shot gather is depicted in Figure 9a.

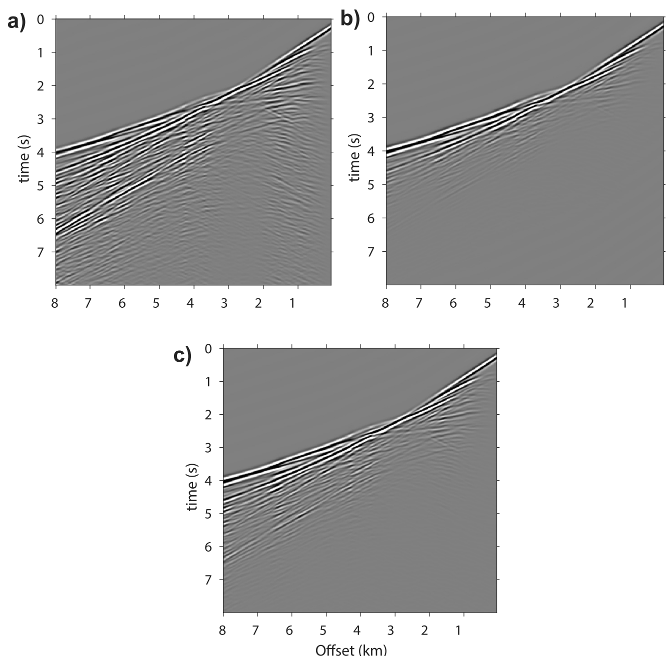

Figure 9: Pressure seismograms in the true model shown in Figure 8 computed with different damping factors: (a) none, $\gamma = 0\ \mathrm{s}^{-1}$, (b) $\gamma = 2\ \mathrm{s}^{-1}$, and (c) $\gamma = 0.8\ \mathrm{s}^{-1}$.

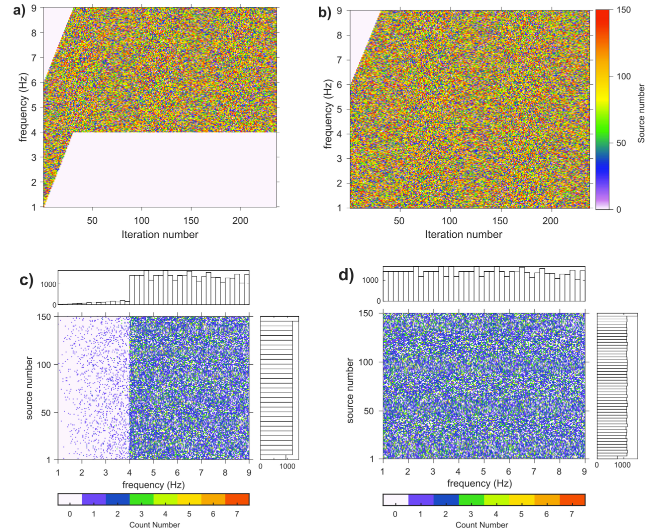

Figure 10: Frequency assignment versus iteration for the undamped inversions using (a) the MFB method in (b) the Bunks approach. Frequency assignment versus source number for (c) MFB and (d) Bunks, histograms and heatmaps.

Moving Frequency Band inversion, undamped

Using the MFB strategy, the first iteration follows Tromp and Bachmann (2019) in utilizing waves between 1–6 Hz. With each of the first 30 iterations, the frequency range shifts by 0.1 Hz. Iterations beyond 30 are performed in the range 4–9 Hz. Each supershot is 37.5 s long ($T + \Delta\tau$), while a non-encoded simulation only requires 7.5 s. For 150 sources and a bandwidth of 5 Hz, this resulted in a speed-up factor of 16× per iteration (see Tromp and Bachmann, 2019, their equation 58). The frequency assignment for each iteration and a frequency heatmap for all sources are shown in Figure 10a and 10c, respectively. Notice the uneven sampling in the low-frequency range. In Figure 11a solid lines represent the normalized model misfits, while the blue line in Figure 11b corresponds to the data phase misfit. The V_P and V_S models after 237 iterations are displayed in Figure 12a and 12c, respectively.

Bunks multiscale inversion, undamped

For the Bunks multiscale strategy we use the same time steps as with the MFB inversion. The first iteration employs waves in the 1–6 Hz range, increased by 0.1 Hz until iteration 30. The encoded source number increases until the 30th iteration. Iterations beyond 30 are performed between 1–9 Hz. A total of 240 frequencies are encoded simultaneously, ensuring that at least 90 sources are assigned two frequencies concurrently. Figure 10b and 10d show the frequency assignment per iteration and the heatmap for each frequency of all sources, respectively. Notice the more balanced sampling of the frequencies. The dashed lines in Figure 11a show the normalized model misfits for V_P and V_S. The red line in Figure 11b shows the phase misfit curve. The Bunks approach uniformly covers the low frequencies, whereas the MFB strategy favors the higher frequencies. The preference for low frequencies pro-

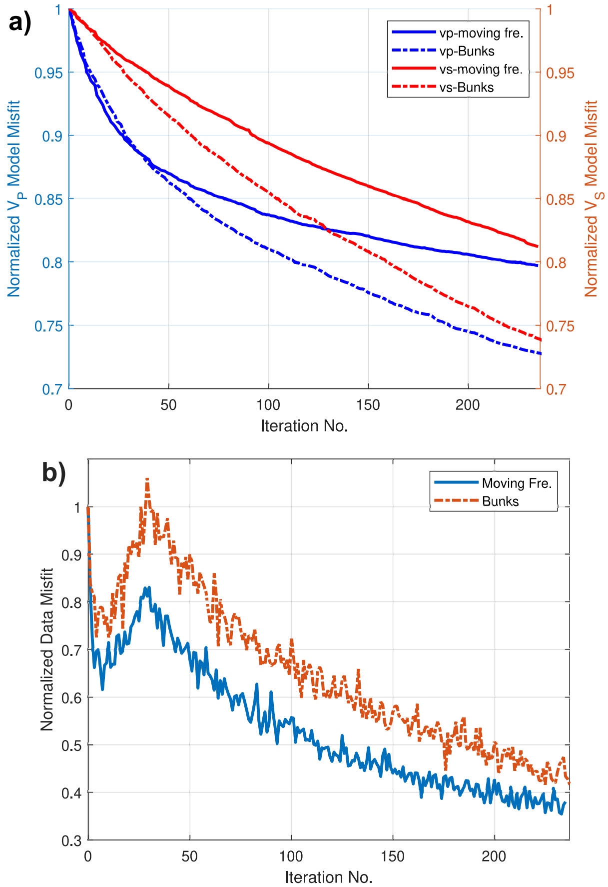

Figure 11: Model and data (phase) misfits. (a) Normalized model misfits for V_P (blue) and V_S (red), for the inversion using the MFB method (solid) and the Bunks (dashed) approach. (b) Normalized phase misfits for the MFB (blue) and Bunks (red) frequency assignment.

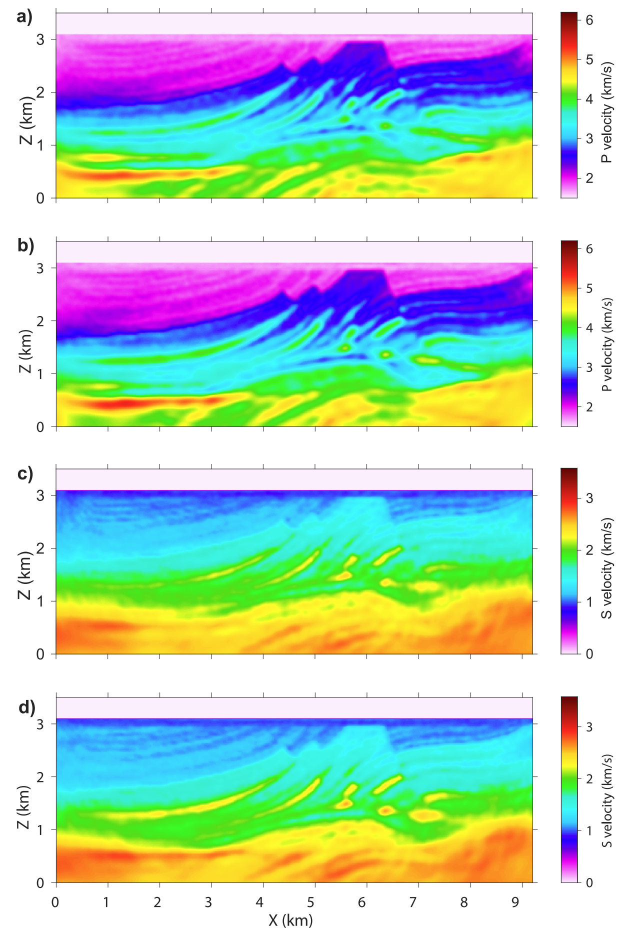

Figure 12: Inversion results after 237 iterations using the MFB in (a) and (c) and the Bunks approach in (b) and (d), for V_P in (a) and (b) and V_S in (b) and (d).

vides a better update to the low-wavenumber components of the models, resulting in improved model updates. Figure 12b and 12d, respectively, show the V_P and V_S models that result, after 237 iterations.

Bunks multiscale inversion, damped

We repeat the experiment under the Bunks frequency assignment using the Laplace-domain approach, in three stages, starting with $\gamma = 2.0$ s^{-1}, then $\gamma = 0.8$ s^{-1}, and finally without damping, $\gamma = 0$, which is equivalent to using the Fourier transform. The damped common-shot gathers are shown in Figure 9b and 9c. The inversion parameters of the three stages are presented in Table 1. Note that damping slows the rate of convergence, offsetting some of the gains in speed made by source-encoding, and thus needs to be carefully considered.

For the damping term $\gamma = 2.0$ s^{-1}, the steady-state time T and the decoding interval $\Delta\tau$ are maintained at the common value, 5.25 s, and only V_P is updated because V_S is not sensitive to early arrivals. For $\gamma = 0.8$ s^{-1}, T and $\Delta\tau$ are set to 7.5 s. As the number of encoded sources is always smaller than the total number of sources available, only a "mini-batch"

Table 1: **Inversion parameters for the Laplace-domain source-encoded FWI of streamer data in the Marmousi model.**

stage	γ (s^{-1})	band (Hz)	$\delta\omega$	$\omega_{\text{max-shift}}$	T (s)	$\Delta\tau$ (s)
1	2.0		0.08		5.25	5.25
2	0.8	1–3	0.08	6	7.50	7.50
3	0.0		0.50		7.50	30.0

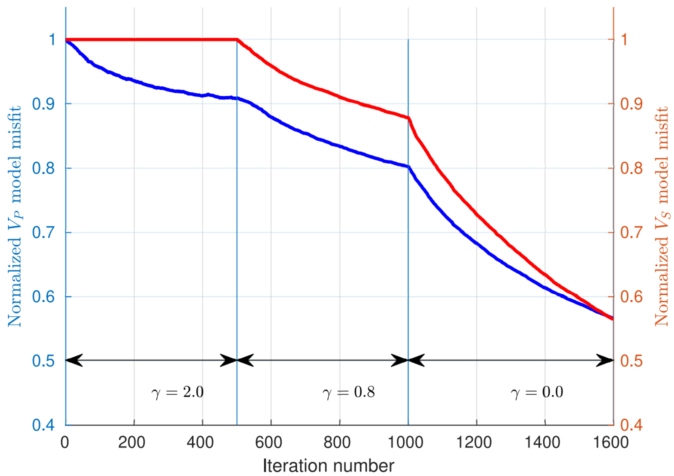

Figure 13: Model misfit evolution for the Laplace-domain FWI of the streamer data for the three-stage inversion tabulated in Table 1.

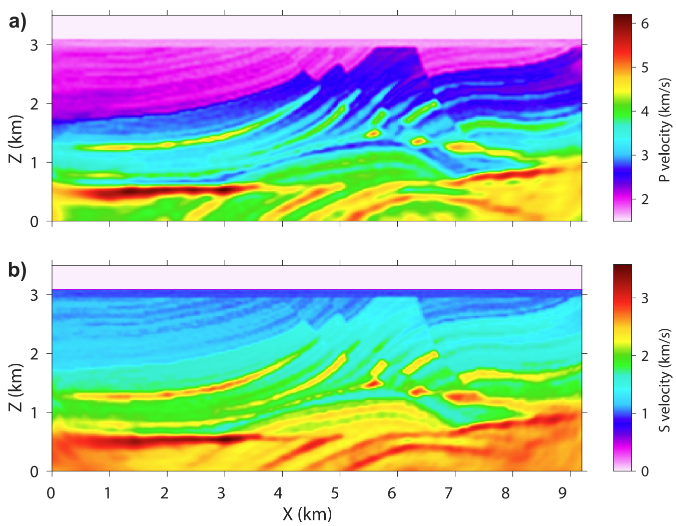

Figure 14: Result after 1600 iterations using the Bunks approach of the three-stage Laplace-domain FWI of Marmousi streamer data: inverted (a) V_P and (b) V_S models.

subset (van Herwaarden et al., 2020) of the whole shot is randomly chosen at each iteration. The model misfits are found in Figure 13. The final FWI V_P and V_S models are shown in Figure 14.

Time-damping applied from the first-arrival time can be viewed as a heuristic way to select aperture angles of P-waves in the data (Brossier et al., 2009). For large γ, only arrivals close in time to the first arrivals are used in the inversion, which corresponds to wide-aperture P-wave events. According to the P-to-P wave scattering pattern from elastic wave-speed perturbations (Tarantola, 1986), those wide-aperture P-wave events mostly come from perturbations in V_P instead of from V_S perturbations. Hence, in Figure 13, we exclusively update the V_P model when $\gamma = 2.0$ s^{-1}. As γ becomes smaller, the phases beyond the first-arrival traveltimes become dominant in the inversion, which corresponds to shorter-aperture P-wave events and converted waves. Because aperture angle is an additional parameter to frequency in controlling the spatial resolution of FWI (Brossier et al., 2009; Sirgue and Pratt, 2004), our Laplace-domain method implements a second level of hierarchy in FWI in addition to that naturally introduced by frequency selection through aperture selection.

We perform our tests on a workstation, Clenardus, equipped with an 8 GB NVIDIA GTX2080Ti GPU and twenty 2.20 GHz Intel Xeon Silver 4210 CPUs with 64 GB of memory. The simulation has a time step of 0.75 ms. Compute times per iteration for stages 1, 2, and 3 are 59.6 s, 73.52 s, and 160.9 s, respectively.

Offshore Couple Acoustic-Elastic FWI of Marmousi Multi-Component OBN Data

In this section, we will examine the Laplace-domain source-encoded acoustic-elastic FWI for the inversion of a multi-component ocean-bottom node (OBN) dataset. The coupled acoustic-elastic spectral-element implementation can accurately simulate the full wavefield at and near the seafloor. It correctly handles fluid-solid boundary conditions (Komatitsch et al., 2000), which has significant advantages for imaging and inversion applications that use amplitude information for model building.

We will investigate how the data damping can affect the inversion process. The true and initial Marmousi models are again those already depicted in Figure 8. The source is positioned 10 m below the sea surface and a total of 150 sources are evenly distributed with a spacing of 60 m. At the seafloor, 46 OBN receivers are placed with a spacing of 200 m and two components per node are recorded. Exploiting reciprocity, as there are twice more physical sources than OBN receivers, it is more efficient to perform two simulations per vertical and horizontal geophone (treated as vertical and horizontal forces) to build the gradient. Therefore, a total of 92 virtual sources will be used. The recording time spans 11.25 s with a time step of 0.75 ms. The simulated common-receiver gather for the vertical and horizontal components is shown in Figure 15a and 15b, respectively. Figures 15c–f show the damped wavefields.

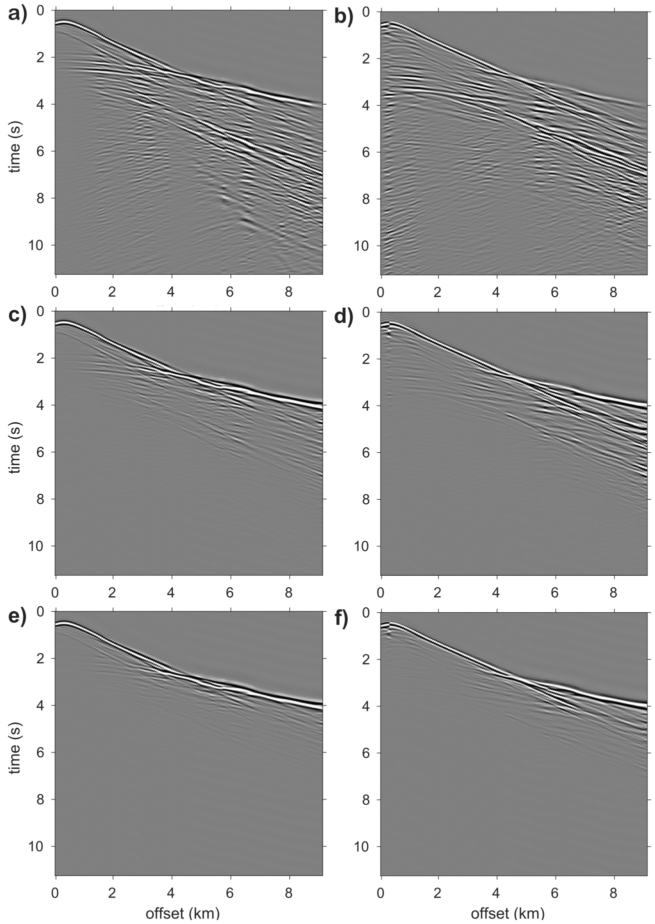

Figure 15: Seismograms for the vertical *(left column)* and horizontal *(right column)* data computed in the Marmousi model with different damping factors: (a–b) zero damping, $\gamma = 0\ \mathrm{s}^{-1}$, (c–d) $\gamma = 1\ \mathrm{s}^{-1}$, (e–f) $\gamma = 2\ \mathrm{s}^{-1}$. Time damping is applied from the first arrival onward to preserve long-offset information.

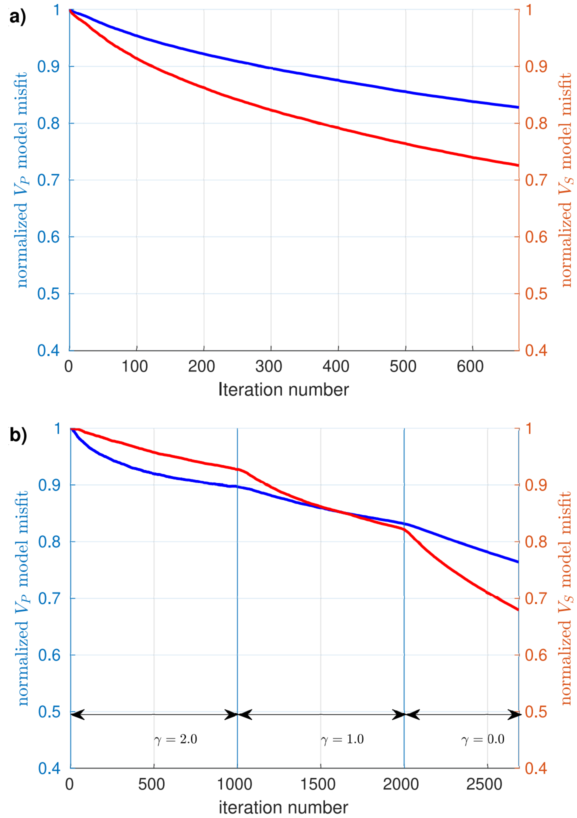

Figure 16: Model misfits for *(a)* the (undamped) Fourier- and the *(b)* (damped) Laplace-domain source-encoded FWI of the OBN Marmousi data.

We first apply the Bunks multiscale strategy without damping, $\gamma = 0$. The first iteration utilizes waves in the 1–3 Hz range with a steady-state time $T = 7.5$ s and an integration interval $\Delta\tau = 30$ s. In the initial iteration, 60 of the virtual sources are randomly chosen for inversion. For each subsequent iteration, the maximum frequency is increased by 0.1 Hz until reaching iteration 60. After that, iterations are performed between 1–9 Hz. A total of 240 frequencies are encoded simultaneously, allowing some sources to be assigned two or three frequencies at the same time. The normalized model misfits are shown in Figure 16a. The V_P and V_S models resulting after 700 iterations are displayed in Figure 17a and 17c, respectively.

We repeated the experiment with two damping terms, $\gamma = 2.0\ \mathrm{s}^{-1}$ and $1.0\ \mathrm{s}^{-1}$. The inversion parameters are unchanged from Table 1, except $\gamma = 1\ \mathrm{s}^{-1}$ in the second stage. The damped vertical and horizontal gathers are shown in Figure 15. The model misfits are shown in Figure 16b. The final V_P and V_S models, obtained using the damped Laplace-domain source-encoded FWI method, are found in Figures 17b and 17d.

In Figure 16a, using (undamped) Fourier-domain source-encoding of OBN data, we notice that the V_S model receives more updates than V_P. In contrast, with the inversion of streamer data in Figure 11a, we observed more updates to V_P. This behavior may indicate that the OBN data contain more S-wave information beneficial to V_S modeling. However, in real-world situations, we may suffer from bad initial V_S models leading the inversion to a local minimum. Thus, a Laplace-domain method is necessary to fit the early arrivals, mainly inverting V_P, and then gradually updating the V_S model as shown in Figure 16.

We perform these tests on a cluster, Della, equipped with an 80 GB NVIDIA A100 GPU and forty-eight 2.8 GHz 1000 GB Intel Ice Lake CPUs with 1000 GB of memory. Compute times for stages 1, 2, and 3 are 59.6 s, 64.46 s, and 105.6 s, respectively, per iteration.

Onshore Elastic FWI of Foothills Data

In this section, we will demonstrate the effectiveness of the Laplace-domain source-encoded elastic FWI method for land data. FWI of land data poses challenges due to increased nonlinearity introduced by free-surface effects such as the propagation of surface waves in the heterogeneous near-surface. Additionally, the presence of short wavelengths in the shear wavefield necessitates an accurate V_S starting model, particularly

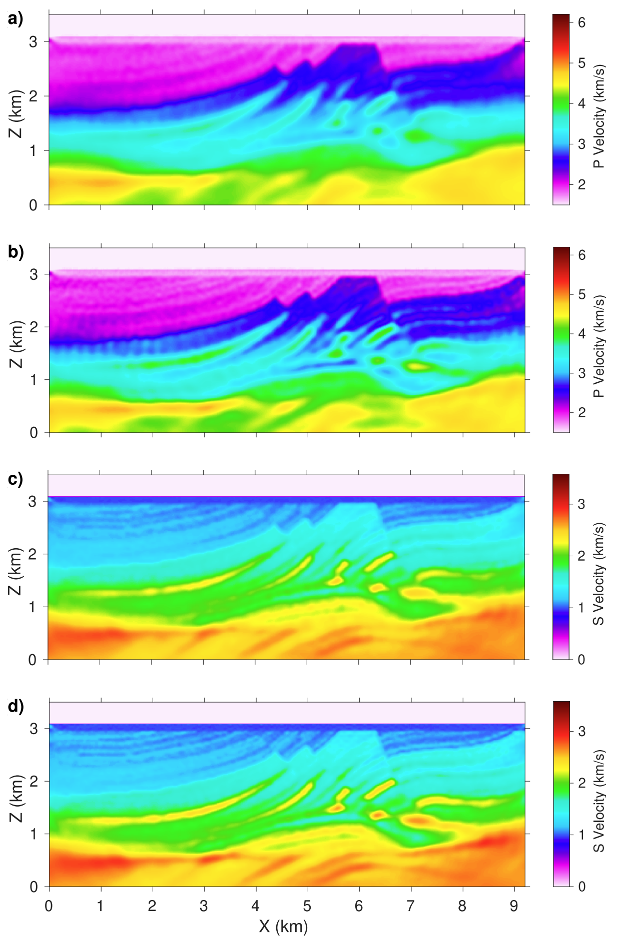

Figure 17: Inverted (a–b) V_P and (c–d) V_S models by (a,c) Fourier (undamped) and (b,d) damped Laplace-domain source-encoding of OBN Marmousi data, 700 iterations.

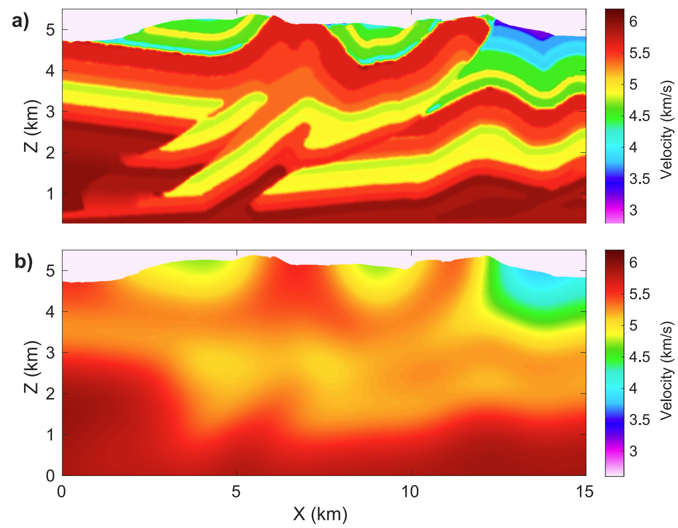

Figure 18: (a) True and (b) initial V_P wave speed Foothills models.

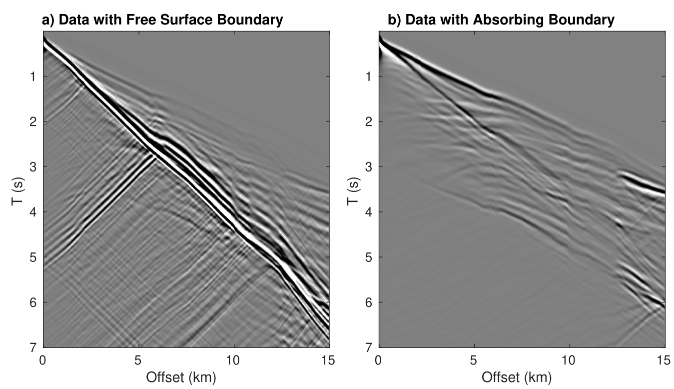

Figure 19: (a) Seismograms computed in the Foothills model for the vertical component particle velocity. The shot is located at a horizontal distance of 0.1 km. A free-surface condition is applied at the top of the model. (b) As in (a) except after imposing an absorbing boundary condition at the top of the model.

when low frequencies are unavailable in the data.

To illustrate the efficacy of Laplace-domain source-encoded elastic FWI, we consider a 2-D onshore Foothills model (Brenders et al., 2008) measuring 5.5 km$\times$15 km. The model incorporates various challenging geological features, including rough topography, alluvial surface deposits, and complex structures resulting from compressive fold-and-thrust tectonics associated with mountain building. The V_P model shown in Figure 18a is modified from Figure 2a of Brenders et al. (2008). A corresponding V_S model is constructed using a constant Poisson ratio, i.e., $V_S = V_P/1.732$. A uniform density of 2600 kg/m^3 is assumed to be known during the inversion. A free-surface is imposed at the top of the model.

We simulate an onshore survey with 300 explosive sources spaced every 50 m, 25 m below the surface. Recordings are made on the surface with 601 vertical geophones spaced every 25 m. Observed and synthetic data are computed with an iden-

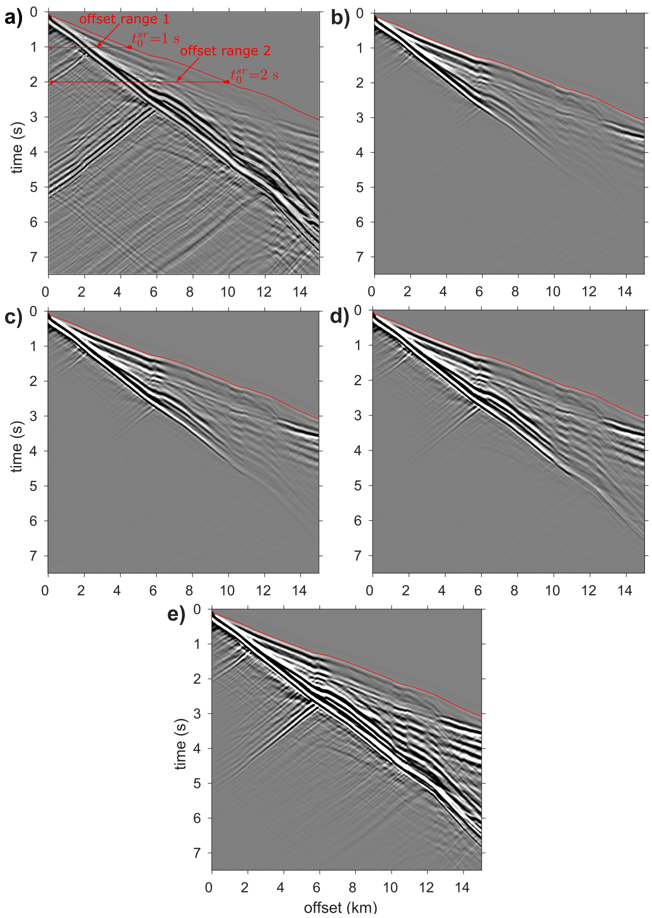

Figure 20: Seismograms for the vertical data computed in the Foothills model with different damping factors: (a) $\gamma = 0\ \mathrm{s}^{-1}$, (b) $\gamma = 3\ \mathrm{s}^{-1}$, (c) $\gamma = 2.35\ \mathrm{s}^{-1}$, (d) $\gamma = 1\ \mathrm{s}^{-1}$, and (e) $\gamma = 0.8\ \mathrm{s}^{-1}$. We mark the first-break times in (a) and show two offset ranges, $t_0^{sr} \leq 1$ s and $t_0^{sr} \leq 2$ s, respectively.

tical algorithm. The source signature is assumed to be known. An elastic shot gather is displayed in Figure 19a. A shot gather computed with an absorbing boundary condition applied to the top of the model is shown in Figure 19b to highlight the additional wave complexities introduced by free-surface effects, such as surface waves and body-wave reflections from the free surface.

For this example we design a combined Fourier/Laplace approach. For the first three stages, we employ (undamped) Fourier-domain source-encoded FWI. However, due to their strong energy, the inversion is dominated by surface waves in the inversion. To address cycle skipping, we limit the maximum offset to traces with a first-break time $t_0^{sr} \leq 1\ s$. Subsequently, we limit the traces to those with $t_0^{sr} \leq 2\ s$, and finally, we include all traces in the inversion (see Figure 20a). The steady-state time $T = 8$ s and decoding interval $\Delta\tau = 60$ s. The starting frequency range is 1–3 Hz, and each subsequent iteration increases the maximum frequency by $\delta\omega = 0.1$ Hz. The maximum frequency shift $\omega_{\text{max-shift}} = 5$ Hz. The inverted V_P and V_S models are presented in Figure 21a and 21c, respectively. The model misfits are shown in Figure 22. Only the shallow part of the V_S model in Figure 21c is well recovered due to the presence of surface waves, while the inversion of the V_P model is not satisfactory.

After these three Fourier-domain stages, we switch to the Laplace domain and we set the damping factors γ to 3.0 s^{-1}, 2.35 s^{-1}, 1.7 s^{-1}, and 0.8 s^{-1}, see Figure 20. For larger values of γ, such as 3.0 s^{-1}, the total simulation time is limited due to the use of a 4-byte data type (`float32`). Consequently, we can only include the near-offset traces with a first-break time smaller than 1 s. For $\gamma = 2.35\ \mathrm{s}^{-1}$, we include traces with a first-break time smaller than 2 s. As γ decreases, a larger simulation time (nominally 99 s) can be set to include all traces in the inversion. The inversion parameters are listed in Table 2.

The inverted V_P and V_S models are displayed in Figures 21b and 21d. Our Laplace-domain source encoding method effectively attenuates surface waves, enabling better updates of the V_P and V_S models compared to traditional Fourier-domain source encoding strategies. Notably, significant reductions in model misfit are observed in stage 4 of the inversion process, as shown in Figure 22.

Another benefit expected from using complex-valued frequencies is the damping during the early FWI iterations of converted *P-S* waves, free-surface multiples, and surface waves, which introduces additional nonlinearities into the inversion. In practice, this second level of hierarchy can be implemented by progressively relaxing the time damping during each frequency-group inversion.

We perform our tests on the aforementioned Della cluster. Compute times, per iteration, for stages 1, 2, and 3 are 262.2 s, 230.5 s, and 496.7 s, respectively. For stages 4, 5, 6, 7, and 8, compute times are 84.7 s, 106.6 s, 98.3 s, 129.56 s, and 202.4 s, respectively.

CONCLUSIONS

To reduce the computational time of elastic full-waveform inversion (FWI) using time-domain solvers, we develop a new

Table 2: **Inversion parameters for the (un)damped source-encoded FWI of land data in the Foothills model.**

	stage	γ (s^{-1})	band (Hz)	$\delta\omega$	$\omega_{\text{max-shift}}$	T (s)	$\Delta\tau$ (s)	max t_0^{sr} (s)
	1							1.0
Fourier	2	0.0	1–3	0.1	5	8.0	60	2.0
	3						99	
	4	3.0	1–3	0.0063	5	3.0	3.6	1.0
	5	2.35	1–3	0.05	5	4.0	4.4	2.0
Laplace	6	1.7	1–3	0.1	7	6.0	7.0	99
	7	0.8	1–3	0.1	8	7.5	18.5	99
Fourier	8	0.0	1–6	0.1	5	8.0	60	99

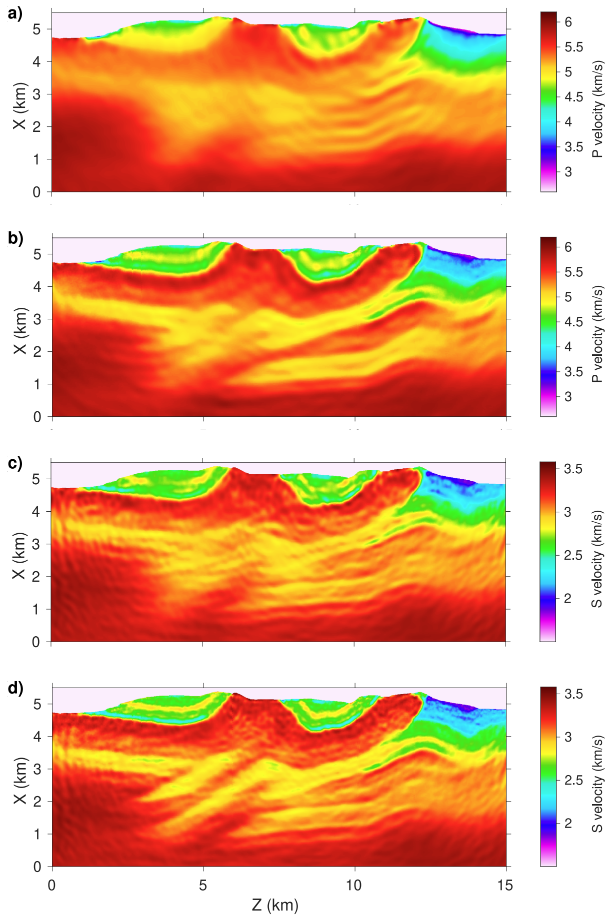

Figure 21: Inverted (a-b) V_P and (c-d) V_S models by (a,c) Fourier (undamped) and (b,d) damped Laplace-domain source-encoded FWI of Foothills land data.

version of source encoding that assigns to each source at random for each iteration a unique complex frequency, and that includes a damping factor to attenuate late arrivals. The misfit criterion is the sum of squared errors in the scaled Laplace coefficients between observed and synthetic data. The source-time function takes the form of a weighted cosine or sine with exponentially increasing amplitude.

We simultaneously activate all the sources and carry out one source-encoded forward simulation followed by one source-encoded adjoint simulation, using a time-domain solver. The encoded forward and adjoint wavefields are run past their steady state. The gradient is calculated through zero-lag cross-correlation between the steady-state encoded forward and adjoint wavefields over a decoding time interval proportional to the inverse of the encoded frequency spacing. Owing to the orthogonality between the trigonometric terms of the encoding operator, no crosstalk is introduced during gradient calculation, and there are no requirements or limitations on acquisition geometry. By tuning the damping factor, we can time-window the data even when only one or a few sparse frequencies are being sampled. Time-windowing allows for the selection of specific arrivals during the various stages of the inversion.

The new version of source encoding can reduce the computational cost of elastic FWI with time-domain solvers. Using our method, a practicing geoscientist can implement elastic FWI on a small workstation. All shot gathers are "encoded" into one "supergather", instead of needing to distribute the seismic sources over the processors of multiple computers in parallel systems that require high-performance computational platforms.

ACKNOWLEDGEMENTS

We thank the three anonymous reviewers, the Associate Editor, Scott Keating, and the Assistant Editor, Yang Liu, for their constructive comments and productive advice. We thank the Princeton Institute for Computational Science and Engineering (PICSciE) for computational facilities and support.

APPENDIX A

ALTERNATIVE MISFITS

We consider three formulations in addition to the misfit criterion in equation 21.

Phase measurement

We define

$$s^s(\mathbf{x}_r, t_0^{sr}) = A^s(\mathbf{x}_r)\exp(\imath\theta^s(\mathbf{x}_r)), \tag{A-1}$$

$$d^s(\mathbf{x}_r, t_0^{sr}) = A^s_{\text{obs}}(\mathbf{x}_r)\exp(\imath\theta^s_{\text{obs}}(\mathbf{x}_r)), \tag{A-2}$$

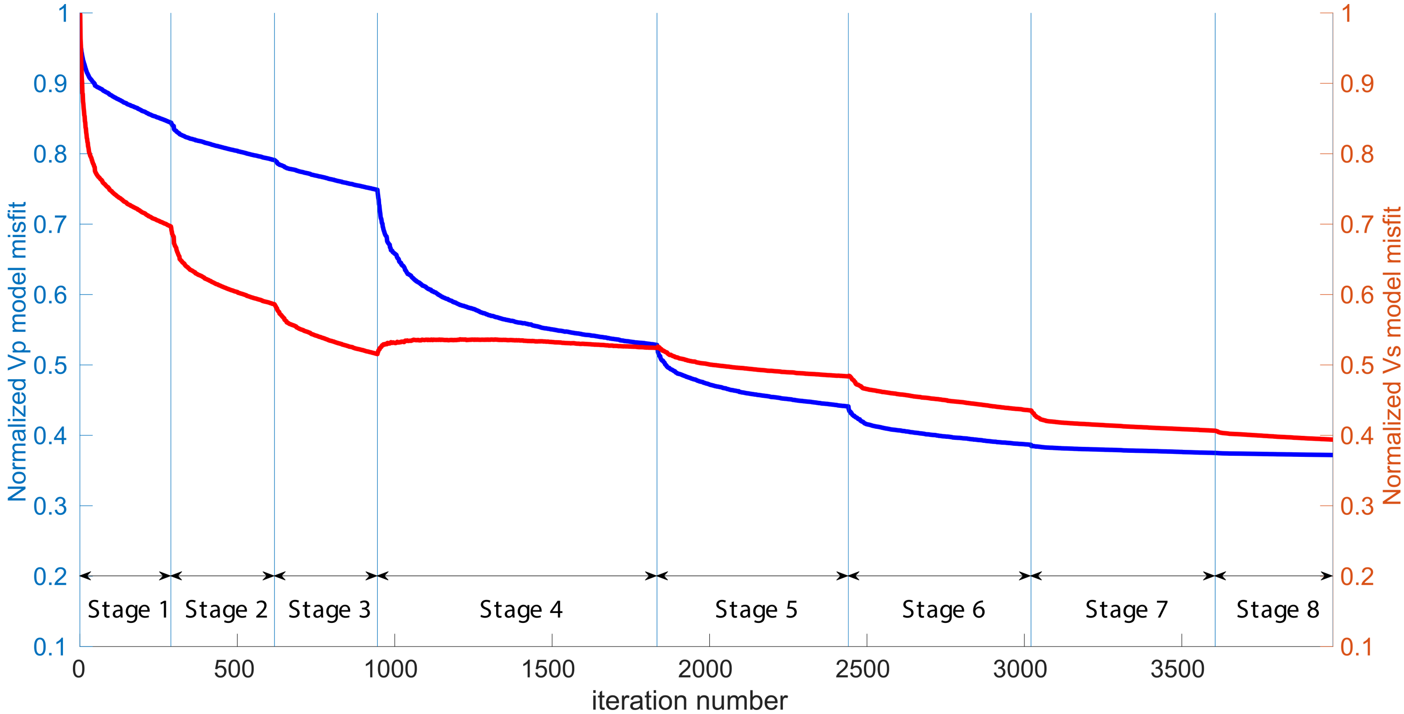

Figure 22: Model misfits for the Fourier-domain (stages 1–3, 8) and Laplace-domain (stages 4–7) source-encoded FWI of land data from the Foothills model.

where

$$\theta^s(\mathbf{x}_r) = \arctan\left[\frac{\mathcal{I}\{s^s(\mathbf{x}_r, t_0^{sr})\}}{\mathcal{R}\{s^s(\mathbf{x}_r, t_0^{sr})\}}\right], \tag{A-3}$$

$$\theta^s_{\text{obs}}(\mathbf{x}_r) = \arctan\left[\frac{\mathcal{I}\{d^s(\mathbf{x}_r, t_0^{sr})\}}{\mathcal{R}\{d^s(\mathbf{x}_r, t_0^{sr})\}}\right], \tag{A-4}$$

and

$$\begin{aligned} A^s(\mathbf{x}_r) &= |s^s(\mathbf{x}_r, t_0^{sr})| \qquad \text{(A-5)} \\ &= \sqrt{[\mathcal{R}\{s^s(\mathbf{x}_r, t_0^{sr})\}]^2 + [\mathcal{I}\{s^s(\mathbf{x}_r, t_0^{sr})\}]^2}, \\ A^s_{\text{obs}}(\mathbf{x}_r) &= |d^s(\mathbf{x}_r, t_0^{sr})| \qquad \text{(A-6)} \\ &= \sqrt{[\mathcal{R}\{d^s(\mathbf{x}_r, t_0^{sr})\}]^2 + [\mathcal{I}\{d^s(\mathbf{x}_r, t_0^{sr})\}]^2}. \end{aligned}$$

The phase misfit function is defined as

$$\chi_\theta = \frac{1}{2}\sum_{s=1}^{S}\sum_{r=1}^{R_s}(\Delta\theta_r^s)^2, \tag{A-7}$$

where $\Delta\theta_r^s$ is the differential phase measurement,

$$\begin{aligned} \Delta\theta_r^s &= \arctan\left\{\frac{\mathcal{I}\left[\frac{s^s(\mathbf{x}_r, t_0^{sr})}{d^s(\mathbf{x}_r, t_0^{sr})}\right]}{\mathcal{R}\left[\frac{s^s(\mathbf{x}_r, t_0^{sr})}{d^s(\mathbf{x}_r, t_0^{sr})}\right]}\right\} \qquad \text{(A-8)} \\ &= \theta^s(\mathbf{x}_r) - \theta^s_{obs}(\mathbf{x}_r), \qquad \text{(A-9)} \end{aligned}$$

where the last equality holds modulo 2π. This phase misfit function has variation,

$$\begin{aligned} \delta\chi_\theta &= \mathcal{R}\sum_{s=1}^{S}\sum_{r=1}^{R_s}\Delta\theta_r^s\delta\theta^s(\mathbf{x}_r), \qquad \text{(A-10)} \\ &= \mathcal{R}\sum_{s=1}^{S}\sum_{r=1}^{R_s}\Delta\theta_r^s[A^s(\mathbf{x}_r)]^{-2} \qquad \text{(A-11)} \\ &\quad \times [-\imath s^{s*}(\mathbf{x}_r, t_0^{sr})]\delta s^s(\mathbf{x}_r, t_0^{sr}), \end{aligned}$$

where we can define the weighted data residual as

$$\begin{aligned} \Delta s^{s*}(\mathbf{x}_r, t_0^{sr}) &= \Delta\theta_r^s[A^s(\mathbf{x}_r)]^{-2} \qquad \text{(A-12)} \\ &\quad \times [-\imath s^{s*}(\mathbf{x}_r, t_0^{sr})](if^s) \\ &= \Delta\theta_r^s[A^s(\mathbf{x}_r)]^{-2}s^{s*}(\mathbf{x}_r, t_0^{sr})f^s \qquad \text{(A-13)} \end{aligned}$$

so that

$$\Delta s^s(\mathbf{x}_r, t_0^{sr}) = \Delta\theta_r^s[A^s(\mathbf{x}_r)]^{-2}s^s(\mathbf{x}_r, t_0^{sr})f^{s*}, \tag{A-14}$$

so we can get the adjoint

$$\begin{aligned} \tilde{f}^\dagger(\mathbf{x}_r, t) &= f_i^\dagger(\mathbf{x}_r, T + \Delta\tau - t), \qquad \text{(A-15)} \\ &= \mathcal{R}\left\{\sum_{s=1}^{S}\delta\tilde{s}^s(\mathbf{x}_r, t_0^{sr})e^{\imath\omega_s(T+t)}\right\}e^{-\gamma(T+t)}, \qquad \text{(A-16)} \end{aligned}$$

where we can define

$$\Delta\tilde{s}_i^s(\mathbf{x}_r, t_0^{sr}) = e^{\gamma t_0^{sr}}\Delta s_i^s(\mathbf{x}_r, t_0^{sr}). \tag{A-17}$$

Exponentiated Phase

We also consider the exponentiated phase misfit function (Yuan et al., 2019),

$$\chi_\theta = \frac{1}{2}\sum_{s=1}^{S}\sum_{r=1}^{R_s} |\exp[\imath\theta^s(\mathbf{x}_r)] - \exp[\imath\theta^s_{obs}(\mathbf{x}_r)]|^2 \quad \text{(A-18)}$$

$$= 2\sum_{s=1}^{S}\sum_{r=1}^{R_s} \sin^2\left(\frac{1}{2}\Delta\theta^s_r\right), \quad \text{(A-19)}$$

and its perturbation

$$\delta\chi_\theta = \sum_{s=1}^{S}\sum_{r=1}^{R_s} \sin(\Delta\theta^s_r)\delta\theta^s(\mathbf{x}_r), \quad \text{(A-20)}$$

$$= \mathcal{R}\sum_{s=1}^{S}\sum_{r=1}^{R_s} \sin(\Delta\theta^s_r)[A^s(\mathbf{x}_r)]^{-2} \quad \text{(A-21)}$$
$$\times [-\imath s^{s*}(\mathbf{x}_r, t_0^{sr})]\delta s^s(\mathbf{x}_r, t_0^{sr}).$$

Now define

$$\Delta s^{s*}(\mathbf{x}_r, t_0^{sr}) = \sin(\Delta\theta^s_r)[A^s(\mathbf{x}_r)]^{-2}[-\imath s^{s*}(\mathbf{x}_r, t_0^{sr})](if^s),$$
$$= \sin(\Delta\theta^s_r)[A^s(\mathbf{x}_r)]^{-2}s^{s*}(\mathbf{x}_r, t_0^{sr})f^s, \quad \text{(A-22)}$$

so that

$$\Delta s^{s}(\mathbf{x}_r, t_0^{sr}) = \sin(\Delta\theta^s_r)[A^s(\mathbf{x}_r)]^{-2}s^{s}(\mathbf{x}_r, t_0^{sr})f^{s*}. \quad \text{(A-23)}$$

We can then get the adjoint,

$$\tilde{f}^\dagger(\mathbf{x}_r, t) = f_i^\dagger(\mathbf{x}_r, T + \Delta\tau - t), \quad \text{(A-24)}$$
$$= \mathcal{R}\left\{\sum_{s=1}^{S}\Delta\tilde{s}^s(\mathbf{x}_r, t_0^{sr})e^{\imath\omega_s(T+t)}\right\}e^{-\gamma(T+t)}, \quad \text{(A-25)}$$

where

$$\delta\tilde{s}^s_i(\mathbf{x}_r, t_0^{sr}) = e^{\gamma t_0^{sr}}\Delta s^s_i(\mathbf{x}_r, t_0^{sr}). \quad \text{(A-26)}$$

Double-Difference Phase

We next consider the misfit function (Yuan et al., 2016)

$$\chi_\theta^{\text{DD}} = \frac{1}{2}\sum_{s=1}^{S}\sum_{r=1}^{R_s}\sum_{r'>r}^{R_s} w_{rr'}[\Delta\Delta\theta^s_{rr'}]^2, \quad \text{(A-27)}$$

where $w_{rr'}$ is a suitably chosen weighting function, and $\Delta\delta\theta^s_{rr'}$ is the "double-difference" phase measurement,

$$\Delta\delta\theta^s_{rr'} = \Delta\theta^s_r - \Delta\theta^s_{r'}, \quad \text{(A-28)}$$

$$= \arctan\left\{\frac{\mathcal{I}\left[\frac{s^s(\mathbf{x}_r,t_0^{sr})d^s(\mathbf{x}_{r'},t_0^{sr'})}{d^s(\mathbf{x}_r,t_0^{sr})s^s(\mathbf{x}_{r'},t_0^{sr'})}\right]}{\mathcal{R}\left[\frac{s^s(\mathbf{x}_r,t_0^{sr})d^s(\mathbf{x}_{r'},t_0^{sr'})}{d^s(\mathbf{x}_r,t_0^{sr})s^s(\mathbf{x}_{r'},t_0^{sr'})}\right]}\right\}, \quad \text{(A-29)}$$

where the last equality holds modulo 2π. This phase misfit function has the variation

$$\delta\chi_\theta^{\text{DD}} = \mathcal{R}\sum_{s=1}^{S}\sum_{r=1}^{R_s}\sum_{r'>r}^{R_s} w_{rr'}\Delta\Delta\theta^s_{rr'}[\delta\theta^s(\mathbf{x}_r) - \delta\theta^s(\mathbf{x}_{r'})], \quad \text{(A-30)}$$

$$= \mathcal{R}\sum_{s=1}^{S}\sum_{r=1}^{R_s}\sum_{r'>r}^{R_s} w_{rr'}\Delta\Delta\theta^s_{rr'}[A^s(\mathbf{x}_r)]^{-2} \quad \text{(A-31)}$$
$$\times [-\imath s^{s*}(\mathbf{x}_r, t_0^{sr})]\delta s^s(\mathbf{x}_r, t_0^{sr})$$
$$- \mathcal{R}\sum_{s=1}^{S}\sum_{r=1}^{R_s}\sum_{r'>r}^{R_s} w_{rr'}\Delta\Delta\theta^s_{rr'}[A^s(\mathbf{x}_{r'})]^{-2} \quad \text{(A-32)}$$
$$\times [-\imath s^{s*}(\mathbf{x}_{r'}, t_0^{sr'})]\delta s^s(\mathbf{x}_{r'}, t_0^{sr'}),$$
$$= \mathcal{R}\sum_{s=1}^{S}\sum_{r=1}^{R_s}\sum_{r'\neq r}^{R_s} w_{rr'}\Delta\Delta\theta^s_{rr'}[A^s(\mathbf{x}_r)]^{-2} \quad \text{(A-33)}$$
$$\times [-\imath s^{s*}(\mathbf{x}_r, t_0^{sr})]\delta s^s(\mathbf{x}_r, t_0^{sr}), \quad \text{(A-34)}$$

where we can define the weighted data residual as

$$\Delta s^{s*}(\mathbf{x}_r, t_0^{sr}) = \sum_{r'\neq r}^{R_s} w_{rr'}\Delta\Delta\theta^s_{rr'}[A^s(\mathbf{x}_r)]^{-2} \quad \text{(A-35)}$$
$$\times [-\imath s^{s*}(\mathbf{x}_r, t_0^{sr})](if^s),$$
$$= \sum_{r'\neq r}^{R_s} w_{rr'}\Delta\Delta\theta^s_{rr'}[A^s(\mathbf{x}_r)]^{-2}s^{s*}(\mathbf{x}_r, t_0^{sr})f^s, \quad \text{(A-36)}$$

so that

$$\Delta s^{s}(\mathbf{x}_r, t_0^{sr}) = \sum_{r'\neq r}^{R_s} w_{rr'}\Delta\Delta\theta^s_{rr'}[A^s(\mathbf{x}_r)]^{-2}s^{s}(\mathbf{x}_r, t_0^{sr})f^{s*}, \quad \text{(A-37)}$$

so we can get the adjoint,

$$\tilde{f}^\dagger(\mathbf{x}_r, t) = f_i^\dagger(\mathbf{x}_r, T + \Delta\tau - t) \quad \text{(A-38)}$$
$$= \mathcal{R}\left\{\sum_{s=1}^{S}\delta\tilde{s}^s(\mathbf{x}_r, t_0^{sr})e^{\imath\omega_s(T+t)}\right\}e^{-\gamma(T+t)}, \quad \text{(A-39)}$$

where we can define

$$\delta\tilde{s}^s_i(\mathbf{x}_r, t_0^{sr}) = e^{\gamma t_0^{sr}}\Delta s^s_i(\mathbf{x}_r, t_0^{sr}). \quad \text{(A-40)}$$

APPENDIX B

SIMULTANEOUS-SOURCE FORWARD AND ADJOINT WAVEFIELD

Foward Wavefield and Source

To reduce the second factor (II) of equation 30, we derive, via equation 31

$$u_j(\mathbf{x}, t+T) = \Re\left\{\bar{\widetilde{u}}_j^s(\mathbf{x})e^{z_s(t+T)}\right\}, \tag{B-1}$$

$$= \Re\left\{(-\imath)\hat{n}_m(\mathbf{x}_s)e^{z_s(t+T)}\widetilde{G}_{jm}^s(\mathbf{x}, \mathbf{x}_s)\right\}, \tag{B-2}$$

$$= \Re\Bigg\{(-\imath)\hat{n}_m(\mathbf{x}_s)e^{z_s(t+T)} \tag{B-3}$$

$$\times \int_0^{\infty} G_{jm}(\mathbf{x}, \mathbf{x}_s, t')e^{-z_s t'}\,\mathrm{d}t'\Bigg\},$$

$$= \Re\Bigg\{(-\imath)\hat{n}_m(\mathbf{x}_s) \tag{B-4}$$

$$\times \int_0^{\infty} G_{jm}(\mathbf{x}, \mathbf{x}_s, t')e^{z_s(T+t-t')}\,\mathrm{d}t'\Bigg\},$$

$$= \Re\Bigg\{(-\imath)\hat{n}_m(\mathbf{x}_s) \tag{B-5}$$

$$\times \int_{-\infty}^{T+t} G_{jm}(\mathbf{x}, \mathbf{x}_s, T+t-t')e^{z_s t'}\,\mathrm{d}t'\Bigg\},$$

$$= \int_{-\infty}^{T+t} G_{jm}(\mathbf{x}, \mathbf{x}_s, T+t-t') \tag{B-6}$$

$$\times \Re\left\{(-\imath)\hat{n}_m(\mathbf{x}_s)e^{z_s t'}\right\}\,\mathrm{d}t',$$

$$= \int_{-\infty}^{T+t} G_{jm}(\mathbf{x}, \mathbf{x}_s, T+t-t') \tag{B-7}$$

$$\times \underbrace{\Re\left\{(-\imath)\hat{n}_m(\mathbf{x}_s)e^{z_s t'}\right\}}_{\text{source}}\,\mathrm{d}t',$$

where the source term is defined as

$$f_m(\mathbf{x}_s, t) = \Re\left\{(-\imath)\hat{n}_m(\mathbf{x}_s)e^{z_s t}\right\}, \tag{B-8}$$

$$= \hat{n}_m(\mathbf{x}_s)e^{\gamma t}\Re\left\{(-\imath)e^{\imath\omega_s t}\right\}, \tag{B-9}$$

$$= \hat{n}_m(\mathbf{x}_s)e^{\gamma t}\sin\omega_s t. \tag{B-10}$$

Adjoint Wavefield and Adjoint Source

To reduce the first factor (I) of equation 30, we define a wavefield, summed over all sources and receivers, using the Laplace transform definition and the reciprocity of the Green's function, and for $t \in [0, \delta\tau]$,

$$\phi_j(\mathbf{x}, t+T) = \Re\left\{e^{\gamma t_0}\bar{\Delta}\widetilde{u}_i^{sr*}e^{-z_s(t+T)}\widetilde{G}_{ij}^s(\mathbf{x}_r, \mathbf{x})\right\}, \tag{B-11}$$

$$= \Re\Bigg\{e^{\gamma t_0}\bar{\Delta}\widetilde{u}_i^{sr*}e^{-z_s(T+t)} \tag{B-12}$$

$$\times \int_0^{\infty} G_{ji}(\mathbf{x}, \mathbf{x}_r, t')e^{-z_s t'}\,\mathrm{d}t'\Bigg\},$$

$$= \Re\Bigg\{e^{\gamma t_0}\bar{\Delta}\widetilde{u}_i^{sr*}e^{-z_s T} \tag{B-13}$$

$$\times \int_0^{\infty} G_{ji}(\mathbf{x}, \mathbf{x}_r, t')e^{-z_s(t+t')}\,\mathrm{d}t'\Bigg\},$$

$$= \Re\Bigg\{e^{\gamma t_0}\bar{\Delta}\widetilde{u}_i^{sr*}e^{-z_s T} \tag{B-14}$$

$$\times \int_t^{\infty} G_{ji}(\mathbf{x}, \mathbf{x}_r, t'-t)e^{-z_s t'}\,\mathrm{d}t'\Bigg\},$$

$$= \int_t^{\infty} G_{ji}(\mathbf{x}, \mathbf{x}_r, t'-t) \tag{B-15}$$

$$\times \Re\left\{\sum_{s=1}^{S} e^{\gamma t_0}\bar{\Delta}\widetilde{u}_i^{sr*}e^{-z_s(T+t')}\right\}\,\mathrm{d}t'.$$

Letting $t' = T + \Delta\tau - t'$, we have

$$\phi_j(\mathbf{x}, t+T) = \sum_{r=1}^{R}\int_{-\infty}^{T+\Delta\tau-t} G_{ji}(\mathbf{x}, \mathbf{x}_r, T+\Delta\tau-t-t')$$

$$\times \underbrace{\sum_{s=1}^{S} e^{\gamma t_0}\Re\left\{\bar{\Delta}\widetilde{u}_i^{sr*}e^{-z_s(2T+\Delta\tau-t')}\right\}}_{\text{adjoint source}}\,\mathrm{d}t', \tag{B-16}$$

where the adjoint source term is defined as

$$f_i^{\dagger}(\mathbf{x}_r, T+\Delta\tau-t') = \sum_{s=1}^{S} e^{\gamma t_0}\Re\Bigg\{\bar{\Delta}\widetilde{u}_i^{sr*}e^{-z_s(2T+\Delta\tau-t')}\Bigg\}, \tag{B-17}$$

$$= \sum_{s=1}^{S} e^{\gamma t_0}\Re\Bigg\{\bar{\Delta}\widetilde{u}_i^{sr}e^{-z_s^*(2T+\Delta\tau-t')}\Bigg\}, \tag{B-18}$$

$$= \sum_{s=1}^{S} e^{\gamma t_0} \tag{B-19}$$

$$\times \Re\Bigg\{\bar{\Delta}\widetilde{u}_i^{sr}e^{(\imath\omega_s-\gamma)(2T+\Delta\tau-t')}\Bigg\},$$

$$= \Re\Bigg\{\sum_{s=1}^{S} e^{\gamma t_0}\bar{\Delta}\widetilde{u}_i^{sr}e^{\imath\omega_s(2T+\Delta\tau-t')}$$

$$\times e^{-\gamma(2T+\Delta\tau-t')}\Bigg\}. \tag{B-20}$$

REFERENCES

Aghazade, K., H. S. Aghamiry, A. Gholami, and S. Operto, 2022, Randomized source sketching for full waveform inversion: IEEE Transactions on Geoscience and Remote Sensing, **60**, 5909412, doi: 10.1109/TGRS.2021.3131039.

Bachmann, E., and J. Tromp, 2020, Source encoding for viscoacoustic ultrasound computed tomography: Journal of the Acoustical Society of America, **147**, 3221–3235, doi: 10.1121/10.0001191.

Barnes, C., and M. Charara, 2009, The domain of applicability of acoustic full-waveform inversion for marine seismic data: Geophysics, **74**, no. 6, WCC91–WCC103, doi: 10.1190/1.3250269.

Ben-Hadj-Ali, H., S. Operto, and J. Virieux, 2011, An efficient frequency-domain full waveform inversion method using simultaneous encoded sources: Geophysics, **76**, no. 4, R109–R124,10. doi: 1190/1.3581357.

Borisov, D., F. Gao, P. Williamson, and J. Tromp, 2020, Application of 2D full-waveform inversion on exploration land data: Geophysics, **85**, no. 2, R75–R86, doi: 10.1190/geo2019–0082.1.

Brenders, A. J., S. Charles, and R. G. Pratt, 2008, Velocity estimation by waveform tomography in the Canadian Foothill – A synthetic benchmark study: 71st Conference and Exhibition, EAGE, Extended Abstracts, cp–40–00138, doi: 10.3997/2214–4609.20147678.

Brenders, A. J., and R. G. Pratt, 2007, Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model: Geophysical Journal International, **168**, 133–151, doi: 10.1111/j.1365–246X.2006.03156.x.

Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion: Geophysics, **74**, no. 6, WCC105–WCC118, doi: 10.1190/1.3215771.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473, doi: 10.1190/1.1443880.

Cho, Y., C. Pérez Solano, J. Kimbro, Y. Yang, R.-E. Plessix, and K. Matson, 2022, Influence of shear velocity on elastic full-waveform inversion: Gulf of Mexico case study using multicomponent ocean-bottom node data: Geophysics, **87**, no. 5, R391–R400, doi: 10.1190/geo2022–0014.1.

Cui, C., E. Bachmann, D. B. Peter, Z. Liu, and J. Tromp, 2023, Source-encoded waveform inversion in the Northern Hemisphere: Geophysical Journal International, **235**, 2305–2322, doi: 10.1093/gji/ggad363.

Dai, W., Y. Huang, and G. T. Schuster, 2013, Least-squares reverse time migration of marine data with frequency-selection encoding: Geophysics, **78**, no. 4, S233–S242, doi: 10.1190/10.1190/GEO2013–0003.1.

Etgen, J. T., 2005, How many angles do we really need for delayed-shot migration?: 75th Annual International Meeting, SEG, Expanded Abstracts, 1985–1988, doi: 10.1190/1.2148097.

Fu, L., B. Guo, and G. T. Schuster, 2018, Multiscale phase inversion of seismic data: Geophysics, **83**, no. 2, R159–R171, doi: 10.1190/geo2017–0353.1.

Huang, Y., and G. T. Schuster, 2012, Multisource least-squares migration of marine streamer and land data with frequency-division encoding: Geophysical Prospecting, **60**, 663–680, doi: 10.1111/j.1365–2478.2012.01086.x.

———, 2018, Full-waveform inversion with multisource frequency selection of marine streamer data: Geophysical Prospecting, **66**, 1243–1257, doi: 10.1111/1365–2478.12588.

Komatitsch, D., C. Barnes, and J. Tromp, 2000, Wave propagation near a fluid-solid interface: A spectral-element approach: Geophysics, **65**, 623–631, doi: 10.1190/1.1444758.

Komatitsch, D., J. Ritsema, and J. Tromp, 2002, The spectral-element method, Beowulf computing, and global seismology: Science, **298**, 1737–1742, doi: 10.1126/science.1076024.

Komatitsch, D., and J. Tromp, 2002, Spectral-element simulations of global seismic wave propagation — I. Validation: Geophysical Journal International, **149**, 390–412, doi: 10.1046/j.1365–246X.2002.01653.x.

Komatitsch, D., and J. P. Vilotte, 1998, The spectral element method: An efficient tool to simulate the seismic response of 2D and 3D geological structures: Bulletin of the Seismological Society of America, **88**, 368–392, doi: 10.1785/BSSA0880020368.

Komatitsch, D., Z. Xie, E. Bozdağ, E. S. de Andrade, D. Peter, Q. Liu, and J. Tromp, 2016, Anelastic sensitivity kernels with parsimonious storage for adjoint tomography and full waveform inversion: Geophysical Journal International, **206**, 1467–1478, doi: 10.1093/gji/ggw224.

Krebs, J. R., J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse, 2009, Fast full-wavefield seismic inversion using encoded sources: Geophysics, **74**, no. 6, WCC177–WCC188, doi: 10.1190/1.3230502.

Krebs, J. R., Y. H. Cha, S. Lee, P. Dimitrov, A. A. Mullur, N. J. Downey, and P. S.Routh, 2013, Orthogonal source and receiver encoding: U. S. Patent, **10012745**.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering: Theory and Application, 206–220.

Lemaistre, L., J. Brunellière, F. Studer, and C. Rivera, 2019, FWI on land seismic datasets with topography variations: Do we still need to pick first arrivals?: 88th Annual International Meeting, SEG, Expanded Abstracts, 1078–1082, doi: 10.1190/segam2018–2995924.1.

Liu, Q., and Y. J. Gu, 2012, Seismic imaging: From classical to adjoint tomography: Tectonophysics, **566–567**, 31–66, doi: 10.1016/j.tecto.2012.07.006.

Liu, Q., and J. Tromp, 2008, Finite-frequency sensitivity kernels for global seismic wave equation based upon adjoint methods: Geophysical Journal International, **174**, 265–286, doi: 10.1111/j.1365–246X.2008.03798.x.

Liu, Z., and L. Huang, 2019, Multiscale and layer-stripping wave-equation dispersion inversion of Rayleigh waves: Geophysical Journal International, **218**, 1807–1821, doi: 10.1093/gji/ggz215.

Luo, Y., R. Modrak, and J. Tromp, 2015, Strategies in adjoint tomography, *in* Handbook of Geomathematics, 2 ed.: Springer, 1943–2001, doi: 10.1007/2F978–3–642–27793–1_96–2.

Luo, Y., H. Zhu, T. Nissen-Meyer, C. Morency, and J. Tromp, 2009, Seismic modeling and imaging based upon spectral-element and adjoint methods: The Leading Edge, **28**, 568–

574, doi: 10.1190/1.3124932.

Marfurt, K. J., 1984, Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations: Geophysics, **49**, 533–549, doi: 10.1190/1.1441689.

Modrak, R., and J. Tromp, 2016, Seismic waveform inversion best practices: regional, global and exploration test cases: Geophysical Journal International, **206**, 1864–1889, doi: 10.1093/gji/ggw202.

Mora, P., and Z. Wu, 2018, Elastic versus acoustic inversion for marine surveys: Geophysical Journal International, **214**, 596–622, doi: 10.1093/gji/ggy166.

Morton, S. A., and C. C. Ober, 1998, Fastershot-record depth migrations using phase encoding: 68th Annual meeting, SEG, Expanded Abstracts, Soc. Explor. Geophys., 1131–1134, doi: 10.1190/1.1820088.

Murphy, G., V. Brown, and D. Vigh, 2020, Land FWI in the Delaware Basin, west Texas: A case study: The Leading Edge, **39**, 324–331, doi: 10.1190/tle39050324.1.

Nihei, K. T., and X. Li, 2007, Frequency response modelling of seismic waves using finite difference time domain with phase sensitive detection (TD–PSD): Geophysical Journal International, **169**, no. 3, 1069–1078, doi:10.1111/j.1365–246X.2006.03262.x.

Nolet, G., 2008, A breviary of seismic tomography: Cambridge Univ. Press.

Operto, S., Y. Gholami, R. Brossier, L. Métivier, V. Prieux, A. Ribodetti, and J. Virieux, 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice: The Leading Edge, **32**, 1040–1054, doi: 10.1190/tle32091040.1.

Operto, S., A. Miniussi, R. Brossier, L. Combe, L. Métivier, V. Monteiller, A. Ribodetti, and J. Virieux, 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: Geophysical Journal International, **202**, 1362–1391, doi: 10.1093/gji/ggv226.

Operto, S., J. Virieux, P. Amestoy, J.-Y. L'Excellent, L. Giraud, and H. B. H. Ali, 2007, 3D finite-difference frequency-domain modeling of visco-acoustic wave propagation using a massively parallel direct solver: A feasibility study: Geophysics, **72**, no. 5, SM195–SM211, doi: 10.1190/1.2759835.

Pan, W., K. A. Innanen, and Y. Geng, 2018, Elastic full-waveform inversion and parameterization analysis applied to walk-away vertical seismic profile data for unconventional (heavy oil) reservoir characterization: Geophysical Journal International, **213**, 1934—1968, doi: 10.1093/gji/ggy087.

Pan, W., K. A. Innanen, Y. Geng, and J. Lu, 2019, Interparameter trade-off quantification for isotropic-elastic full-waveform inversion with various model parameterizations: Geophysics, **84**, no. 2, R185–R206, doi: 10.1190/geo2017–0832.1.

Pérez Solano, C., and R.-E. Plessix, 2019, Velocity-model building with enhanced shallow resolution using elastic waveform inversion—an example from onshore Oman: Geophysics, **84**, no. 6, R977–R988, 10.1190/geo2018–0736.1.

———, 2023, Can elastic waveform inversion benefit from inverting multicomponent data?: The Leading Edge, **42**, 184–189, doi: 10.1190/tle42030184.1.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503, doi: 10.1111/j.1365–246X.2006.02978.x.

———, 2009, Three-dimensional frequency-domain full-waveform inversion with an iterative solver: Geophysics, **74**, no. 6, WCC149–WCC157, doi: 10.1190/1.3211198.

Plessix, R.-E., G. Baeten, J. W. de Maag, M. Klaassen, R. Zhang, and Z. Tao, 2010, Application of acoustic full waveform inversion to a low-frequency large-offset land data set: 80th Annual Meeting, SEG, Expanded Abstracts, 930–934, doi: 10.1190/1.3513930.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901, doi: 10.1190/1.1444597.

Pratt, R. G., and R. M. Shipp, 1999, Seismic waveform inversion in the frequency domain, Part 2: Fault delineation in sediments using crosshole data: Geophysics, **64**, 902–914, doi: 10.1190/1.1444598.

Prieux, V., R. Brossier, S. Operto, and J. Virieux, 2013a, Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 1: imaging compressional wave speed, density and attenuation: Geophysical Journal International, **194**, 1640–1664, doi: 10.1093/gji/ggt177.

———, 2013b, Multiparameter full waveform inversion of multicomponent ocean-bottom-cable data from the Valhall field. Part 2: imaging compressive-wave and shear-wave velocities: Geophysical Journal International, **194**, 1665–1681, doi: 10.1093/gji/ggt178.

Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, 426–436, doi: 10.1190/1.1444737.

Schuster, G. T., and Y. Huang, 2013, Multisource full waveform inversion of marine streamer data with frequency selection: 75th Conference and Exhibition, EAGE, Extended Abstracts, cp–348–00697, doi: 10.3997/2214–4609.20130833.

Schuster, G. T., X. Wang, Y. Huang, W. Dai, and C. Boonyasiriwat, 2011, Theory of multisource crosstalk reduction by phase-encoded statics: Geophysical Journal International, **184**, 1289–1303, doi: 10.1111/j.1365–246X.2010.04906.x.

Shin, C., D. Min, K. J. Marfurt, H. Y. Lim, D. Yang, Y. Cha, S. Ko, K. Yoon, T. Ha, and S. Hong, 2002, Traveltime and amplitude calculations using the damped wave solution: Geophysics, **67**, 1637–1647, doi: 10.1190/1.1512811.

Simons, F. J., 2010, Slepian functions and their use in signal estimation and spectral analysis, *in* Handbook of Geomathematics: Springer, 30, 891–923, doi: 10.1007/978–3–642–01546–5_30.

Sirgue, L., and R. G. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.

Tang, Y., and B. Biondi, 2009, Least-squares migration/inversion of blended data: 79th Annual Meeting, SEG, Expanded Abstracts, 2859–2863, doi: 10.1190/1.3255444.

Tarantola, A., 1984, Linearized inversion of seismic reflection data: Geophysical Prospecting, **34**, 998–1015, doi: 10.1111/j.1365–2478.1984.tb00751.x.

———, 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903, doi: 10.1190/1.1442046.

Tromp, J., 2020, Seismic wavefield imaging of Earth's interior across scales: Nature Reviews Earth and Environment, **1**, 40–53, doi: 10.1038/s43017–019–0003–8.

Tromp, J., and E. Bachmann, 2019, Source encoding for adjoint tomography: Geophysical Journal International, **218**, 2019–2044, doi: 10.1093/gji/ggz271.

Tromp, J., D. Komatitsch, and Q. Liu, 2008, Spectral-element and adjoint methods in seismology: Communications in Computational Physics, **3**, 1–32.

Tromp, J., C. Tape, and Q. Liu, 2005, Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: Geophysical Journal International, **160**, 195–216, doi: 10.1111/j.1365–246X.2004.02453.x.

van Herwaarden, D. P., C. Boehm, M. Afanasiev, S. Thrastarson, L. Krischer, J. Trampert, and A. Fichtner, 2020, Accelerated full-waveform inversion using dynamic mini-batches: Geophysical Journal International, **221**, 1427–1438, doi :10.1093/gji/ggaa079.

Versteeg, R., 1994, The Marmousi experience: Velocity model determination on a synthetic complex data set: The Leading Edge, **13**, 927–936, doi: 10.1190/1.1437051.

Vigh, D., and E. W. Starr, 2008, 3D prestack plane-wave, full-waveform inversion: Geophysics, **73**, no. 5, VE135–VE144, doi: 10.1190/1.2952623.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC1–WCC26, doi: 10.1190/1.3238367.

Yuan, Y. O., E. Bozdağ, C. Ciardelli, F. Gao, and F. J. Simons, 2019, The exponentiated phase measurement, and objective-function hybridization for adjoint waveform tomography: Geophysical Journal International, **221**, 1145–1164, doi: 10.1093/gji/ggaa063.

Yuan, Y. O., F. J. Simons, and J. Tromp, 2016, Double-difference adjoint seismic tomography: Geophysical Journal International, **206**, 1599–1618, doi: 10.1093/gji/ggw233.

Zhang, Q., W. Mao, and J. Fang, 2020, Elastic full waveform inversion with source-independent crosstalk-free source-encoding algorithm: IEEE Transactions on Geoscience and Remote Sensing, **58**, 2915, doi: 10.1109/TGRS.2019.2957829.

Zhang, Q., W. Mao, H. Zhou, H. Zhang, and Y. Chen, 2018, Hybrid-domain simultaneous-source full waveform inversion without crosstalk noise: Geophysical Journal International, **215**, 1659–1681, doi: 10.1093/gji/ggy366.

Zhang, Y., J. Sun, C. Notfors, S. H. Gray, L. Chernis, and J. Young, 2005, Delayed-shot 3D depth migration: Geophysics, **70**, no. 5, E21–E28, doi: 10.1190/1.2057980.

Zhang, Z., Z. Wu, Z. Wei, J. Mei, R. Huang, and P. Wang, 2023, Enhancing salt model resolution and subsalt imaging with elastic FWI: The Leading Edge, **42**, 207–215, doi: 10.1190/tle42030207.1.

Zhu, H., Y. Luo, T. Nissen-Meyer, C. Morency, and J. Tromp, 2009, Elastic imaging and time-lapse migration based on adjoint methods: Geophysics, **74**, no. 6, WCA167–WCA177, doi: 10.1190/1.3261747.