

LATTE: Open-source, high-performance acoustic and elastic travelttime computation, tomography, and source location

Kai Gao^{*,1} and Ting Chen¹

¹Earth and Environmental Sciences Division, Los Alamos National Laboratory,
Los Alamos, NM 87545, USA

¹ Abstract

Travelttime-based tomography and source location are classical but important approaches to revealing subsurface structures and understanding spatiotemporal distribution of seismicity ranging from local to global scales. We develop an open-source, high-performance implementation of eikonal equation solving and adjoint-state theory to perform travelttime computation, velocity tomography, and source location in 2D/3D acoustic and elastic media. Specially, we develop novel regularization schemes based on total generalized p -variation, structural similarity, and multitask machine learning models to improve the fidelity and interpretability of inverted model and source parameters. Additionally, our implementation encloses several notable features: its exploits both absolute-difference or double-difference misfit of travelttime to achieve high-fidelity velocity tomography, source location, and source origin time estimation; it enables flexible travelttime computation, tomography, and location in both 2D/3D acoustic and elastic media by allowing arbitrary source and receiver distribution; and we develop a perturbation-based optimal step size computation method to reduce the computational cost. Leveraging both shared-memory and distributed memory parallel programming models, our implementation provides a highly efficient framework for travelttime-based computation, tomography, and source location. We demonstrate the efficacy and efficiency of our method and implementation through several synthetic data examples.

*Corresponding author; Email: kaigao@lanl.gov

18 1 Introduction

19 Traveltime-based tomography and source location are classical but also important approaches
20 to revealing subsurface structures and understanding seismicity ranging from local scale to global
21 scales. In an era of full-waveform-based imaging and inversion, traveltome-based methods still hold
22 their advantage especially in terms of computational efficiency in different applications.

23 Traveltome computation is the foundation for traveltome-based subsurface characterization.
24 Early works of traveltome computation were mostly based on ray tracing by solving an one-point
25 initial value problem or two-point boundary value problem (Pereyra et al., 1980; Grechka and
26 McMechan, 1996; Sadeghi et al., 1999; Meléndez et al., 2015). While ray-based methods are
27 efficient for sparse source-receiver geometry, the computational complexity is directly proportional
28 to the number of source-receiver pairs as well as the complexity of velocity model in an inversion.
29 Wavefront construction methods (Vinje et al., 1993; Lambaré et al., 1996; Gibson et al., 2005;
30 Chambers and Kendall, 2008) compute traveltome and amplitude by approximating wavefronts
31 and interpolating new rays on-the-fly, but is also more computationally demanding compared
32 with conventional ray-based approaches. Vidale (1988) developed the first eikonal-equation-based
33 approach to computing first-arrival traveltome. Later developments based on the eikonal equation
34 include expanding wavefront methods (Podvin and Lecomte, 1991; Qin et al., 1992), fast marching
35 methods (Sethian and Popovici, 1999; Rawlinson and Sambridge, 2004; Zhang et al., 2006), fast
36 sweeping methods (Tsai et al., 2003; Zhao, 2004; Kao et al., 2004, 2005; Fomel et al., 2009; Luo and
37 Qian, 2011; Waheed et al., 2015b), etc. Recent developments of eikonal solving also include solvers
38 on triangular or unstructured mesh (e.g., Qian et al., 2007; Le Bouteiller et al., 2019), high-order and
39 non-oscillatory solvers (e.g., Kim and Cook, 1999; Kim, 2002; Zhang et al., 2006; Luo and Qian,
40 2011; Luo et al., 2012), solvers for anisotropic media (Qian and Symes, 2002; Wang et al., 2006;
41 Waheed et al., 2015a; Waheed and Alkhalifah, 2017), and so on. To solve the source singularity
42 issue, one of the major problems intrinsic to eikonal solving, Fomel et al. (2009) and a number of
43 subsequent works (e.g., Luo and Qian, 2011, 2012; Luo et al., 2012) decompose traveltome field
44 through addition or multiplication and achieve accurate traveltome computation for near-source
45 region and correspondingly higher accuracy in the far-field region.

46 Traveltome tomography is a classical inversion method to estimate subsurface medium properties
47 using traveltome information of seismic signals (e.g., Wu and Toksöz, 1987; Schuster and Quintus-
48 Bosz, 1993; Zelt and Barton, 1998; Zhang and Toksöz, 1998). In contrast to more recent full-
49 waveform inversion (FWI) (e.g. Tarantola, 1984; Mora, 1987; Virieux and Operto, 2009) that
50 uses both traveltome and the amplitude information for estimating medium parameters, traveltome
51 tomography uses only traveltome information, therefore leads to more convex objective function
52 but usually lower-resolution results. In addition, traveltome-based tomography usually enjoys a

53 sheer advantage of low computational cost compared with full-waveform methods, which require
54 solving full wave equations. Early traveltome tomography methods rely on seismic rays, and the
55 model update concentrates merely on ray paths. Studies show that it is also important to consider
56 the band-limited ray path effect and update the model parameters that are near the ray paths, or
57 even use the “full” traveltome information inherited in full seismic wavefields (e.g., Michelena and
58 Harris, 1991; Woodward, 1992; Yomogida, 1992; Schuster and Quintus-Bosz, 1993; Vasco et al.,
59 1995; Snieder and Lomax, 1996; Marquering et al., 1999; Spetzler and Snieder, 2004; Pyun et al.,
60 2005; Xu et al., 2006; Liu et al., 2009; Luo et al., 2016; Zelt and Chen, 2016). These approaches
61 take the finite-frequency effect into consideration, and lead to more accurate tomography results
62 compared with classical ray tomography, but usually require higher computational costs (Luo and
63 Schuster, 1991; Michelena and Harris, 1991; Woodward, 1992; Schuster and Quintus-Bosz, 1993;
64 Vasco et al., 1995; Liu et al., 2009; Luo et al., 2016).

65 Ray-based traveltome tomography is usually formulated in a large linear system and is solved
66 using iterative strategies, the computational cost of which is directly proportional to the number
67 of source-receiver pairs and the number of model parameters. For active-source applications, the
68 number of source-receiver pairs can be prohibitively large, therefore can result in high computational
69 cost. In addition, as ray computation can fail in complex heterogeneous media (e.g., Rawlinson
70 et al., 2010), ray-based traveltome tomography may suffer from applicability and accuracy issue in
71 complex geological models. In observation of these issues, traveltome tomography was formulated
72 in a nonlinear inversion framework using the adjoint-state method (Sei and Symes, 1994; Leung
73 and Qian, 2006; Taillandier et al., 2009; Huang and Bellefleur, 2012), which was previously applied
74 to computing FWI gradients (Plessix, 2006; Fichtner et al., 2006; Liu and Tromp, 2006). The main
75 advantage of the adjoint-state traveltome tomography is that the problem dimension is independent
76 of the number of source-receiver pairs and is only determined by the number of sources and the
77 dimension of the discretized model. For each common-shot gather, the adjoint-state method only
78 requires solving an eikonal equation and an adjoint-state equation in a similar fashion with FWI
79 gradient computation. Because the traveltome fields in the adjoint-state traveltome methods are
80 computed for the whole space, the kernel in the adjoint-state traveltome naturally approximate the
81 band-limit effect in the sophisticated fat-ray or wave-path approaches (Taillandier et al., 2009;
82 Bretaudeau et al., 2014). In addition, one can use efficient eikonal solvers to obtain the traveltome
83 field and the adjoint-state field, e.g., the fast-sweeping method (Zhao, 2004) with a computational
84 complexity of $\mathcal{O}(N)$, where N is the number of grid points in the discretized model. Traveltome
85 tomography can also use reflection traveltome information (e.g., Zhang et al., 1998; Korenaga et al.,
86 2000; Huang and Bellefleur, 2012; Meléndez et al., 2015; Zhang et al., 2023) for deriving deep
87 subsurface structures when there is no sufficiently wide-aperture traveltome available.

88 Depending on the complexity of the target model, traveltome tomography may need proper

89 preconditioning and regularization schemes for improving the convergence. Regularized geophysical
90 inversion has a long history (Zhdanov, 2002), of which the Tikhonov regularization is frequently
91 used (Tikhonov et al., 1995; Asnaashari et al., 2013). Rudin et al. (1992) developed the methodology
92 of the total variation (TV) to reconstruct sharp edges of images in the context of image analysis
93 and processing. Anagaw (2011) applied this method to geophysical inverse problems to promote
94 sharp interfaces of models. In the context of FWI, Guitton (2012) developed a blocky regularization
95 scheme to promote interface reconstruction. Lin and Huang (2014) developed a modified TV (MTV)
96 regularization scheme to obtain clean and accurate TV regularization results with the split-Bregman
97 technique (Goldstein and Osher, 2009). Lin et al. (2015) applied this regularization to double-
98 difference traveltome tomography (Zhang and Thurber, 2003). Esser et al. (2016) developed an
99 asymmetric TV regularized FWI, in which they penalizes the model discontinuities only in the
100 vertical direction in an asymmetric way, resulting in high-quality reconstruction of deep regions and
101 large medium parameter contrasts. Gao and Huang (2019) developed a total generalized p -variation
102 regularization scheme that preserves both sharp interfaces and piecewise smooth medium property
103 variations. In the context of ground-penetrating radar imaging, Gao et al. (2022) developed a
104 machine-learning (ML) based regularizer to improve the resolution, structure coherence, and fault
105 delineation.

106 Source location has been one of the most classical problems in seismology. Accurately locating
107 seismic event to their correct spatiotemporal locations is the key to understand the evolution of
108 earthquakes and their correlation with faults. Geiger's method has been the classical principle of
109 source location, which computes the source location through the partial derivatives with respect to
110 location in a framework of linear traveltome equation based on Taylor series expansion. Initially
111 formulated within the framework of ray method (Thurber, 1983), the principle was later applied in
112 the context of wave equation and eikonal equation (Tong et al., 2016; Tong, 2021a) for determining
113 source location. In a departure from Geiger's location method, Waldhauser and Ellsworth (2000)
114 leveraged the similarity between the ray paths of two close events, and attributed the time difference
115 of two events observed at a same station to the spatial separation of the two events through the
116 so-called double difference (DD). The method exploits both absolute traveltome and differential time,
117 and provides an effective way to remove the receiver-side structure uncertainties and obtain high-
118 resolution source location. The principle was later applied to joint velocity tomography and source
119 location (Zhang and Thurber, 2003, 2006) in the framework of seismic rays. The methodology of
120 DD, along with the open-source implementation (hypoDD and tomoDD), gained a wide variety of
121 applications ranging from earthquake seismology (e.g., Guo and Thurber, 2021; Zeng et al., 2016)
122 to CO₂ reservoir microseismicity monitoring (Dando et al., 2021), to list a few. Yuan et al. (2016)
123 applied DD to FWI. Tong et al. (2024) extended the methodology of DD or differential traveltome to
124 adjoint-state traveltome tomography and hypocenter location, allowing for complex velocity models.

125 In this work, we develop an open-source, high-performance implementation of traveltime
126 computation, tomography, and source location. Specially, we develop novel regularization schemes
127 for updating model and source parameters, aiming to improve the fidelity and interpretability of
128 inversion results. The motivation for us to develop this work is two-fold.

129 Firstly, current traveltime-based tomography and source location method do not present a
130 systematic approach to regularizing model and source parameters. The damping strategy used by
131 tomoDD ([Zhang and Thurber, 2003, 2006](#)) is for stabilizing the inversion by resolving the imbalance
132 between ray path density and grid spacing, but does not primarily regularize model parameters. The
133 MTV regularization developed by [Lin et al. \(2015\)](#) is a more modern approach to model parameter
134 regularization, yet the work is developed in the framework of ray-based tomoDD. As to the source
135 parameter, we are unaware of any systematic method that regularizes source locations. For both
136 global to regional-scale earthquakes and local-scale microseismicity applications, the fundamental
137 observation is that seismicity is strongly correlated with faults or fractures, which is consistent with
138 well-established seismic moment source theory ([Aki and Richards, 2002](#)). In response to these
139 two issues, in this work, we develop a novel regularization scheme to model parameter update (for
140 both first-arrival travelttime tomography and joint tomography-location). The new regularization
141 scheme consists of a total generalized p -variation (TGpV) regularizer and a P-S wave velocity
142 structure similarity regularizer. The first regularizer results in piecewise smooth updated velocity
143 models by penalizing both the first- and second-order total variations ([Rudin et al., 1992; Goldstein
144 and Osher, 2009; Knoll et al., 2011; Gao and Huang, 2019](#)), while second regularizer results in
145 structurally consistent v_p and v_s models by imposing v_p/v_s ratio limits and applying median and
146 Gaussian smoothing. Our test results show that such a joint regularizer results in elastic parameter
147 models of higher fidelity. Meanwhile, we develop a novel ML-based source parameter regularizer
148 for improving the spatial consistency between faults/fractures and inverted seismic locations. In
149 specific, we develop a supervised multitask ML model to infer faults from a source image, a
150 supervised multitask ML model to refine the inferred faults, and then use the inferred and refined
151 faults as a “guidance” to guide the update of source locations over iterations. Both regularization
152 schemes lead to better interpretability of inversion results.

153 Secondly, as of today, there have been numerous open-sources codes for travelttime computation
154 (e.g., [de Kool et al., 2006; White et al., 2020; Chen et al., 2023](#)). On the contrary, the available open-
155 source implementation of travelttime-based tomography and source location is limited ([Rawlinson
156 et al., 2006; Fang et al., 2019](#)) in addition to hypoDD ([Waldhauser and Ellsworth, 2000](#)) and
157 tomoDD ([Zhang and Thurber, 2003](#)), with different levels of convenience to use. When considering
158 eikonal-equation and adjoint-state-equation-based travelttime tomography and source location, we
159 know few available open-source, fully-functional codes, except for instance, RAJZEL ([Koehn and
160 De Nil, 2022](#)), ATT_Training ([Tong, 2021b](#)), both apply only to 2D acoustic media. Our work

161 therefore aims specifically to build an open-source, systematic, high-performance implementation
 162 enclosing traveltime computation, first-arrival traveltime tomography, source location, as well as
 163 joint tomography-location, based on the eikonal equation (Fomel et al., 2009) and the adjoint-state
 164 equation (Leung and Qian, 2006; Taillandier et al., 2009) for both 2D/3D acoustic and elastic
 165 media. In addition, we intend to include the novel model and source parameter regularization
 166 schemes developed in this paper to this package, thus enhancing its capability in inverting complex
 167 models and fault-related source locations. An additional intention is that we aim to provide a hybrid
 168 shared- and distributed memory parallel implementation of both factorized eikonal equation and
 169 adjoint-state equation, thus, to achieve high computational efficiency for large inversion problems
 170 on modern high-performance computing platforms. We also design user-friendly parameter input
 171 and result output in our implementation to enable convenient setup for inversions involving complex
 172 model, observation geometry, data, and parameter tuning. To distinguish our work with existing
 173 works, we name our open-source implementation as LATTE – Los Alamos Travel-Time package
 174 based on Eikonal equation.

175 The rest of the paper is organized as follows. In the Methodology section, we detail the methods
 176 and algorithms we use for traveltime computation, first-arrival traveltime tomography (FATT), and
 177 source location, and joint tomography-location (TLOC). Specially, we describe the algorithms
 178 associated with optimal step size computation in FATT and TLOC based on a small-perturbation
 179 approach and detail a TGpV- and $v_p - v_s$ similarity-based model parameter regularization scheme
 180 and an ML-based fault-constrained source parameter regularization scheme to improve the fidelity
 181 and interpretability of inversion results. In the Numerical Results section, we use five examples to
 182 validate the efficacy and accuracy of our methods and implementation. We summarize our work in
 183 Conclusions.

184 **2 Methodology**

185 **2.1 Travelttime computation**

186 The eikonal equation in isotropic media is an infinitely high-frequency approximation to wave
 187 equation:

$$v^2 |\nabla t|^2 = 1, \quad t(\mathbf{x}_s) = 0, \quad (1)$$

188 where $t = t(\mathbf{x})$ is the travelttime field associated with a velocity model $v = v(\mathbf{x})$ and a source
 189 location \mathbf{x}_s . When there is a non-zero source initiation time t_0 , we add t_0 to $T(\mathbf{x})$ as $T(\mathbf{x}) + t_0$ to
 190 obtain its true arrival time.

191 A known issue associated with solving equation (1) is that the travelttime field curvature at

192 the source location is infinite, resulting in notable inaccuracy near the source location, which can
 193 propagate outwards. In our LATTE implementation, we adopt the factorized eikonal equation
 194 (Fomel et al., 2009) to mitigate the issue:

$$v^2 \sum_{i=1}^3 \left(t_0 \frac{\partial \tau}{\partial x_i} + \frac{\partial t_0}{\partial x_i} \tau \right)^2 = 1, \quad (2)$$

$$t(\mathbf{x}) = t_0(\mathbf{x})\tau(\mathbf{x}) + \eta_0, \quad t_0(\mathbf{x}_s) = 0, \quad \tau(\mathbf{x}_s) = 1, \quad (3)$$

195 where x_i with $i = 1, 2, 3$ represent spatial coordinates, $t_0 = t_0(\mathbf{x})$ is a background traveltimes field
 196 computed by analytical expression for avoid source singularity, $\tau = \tau(\mathbf{x})$ is a multiplicative field,
 197 which is also the field to be solved with through equation (2). In addition, η_0 is a scalar value
 198 representing the origin time of the source and is added to the computed traveltimes field after t is
 199 obtained; \mathbf{x}_s represent the source location.

200 We generalize equation (2) to the scenario of simultaneous multiple point sources, where each
 201 source may have a nonzero origin time. Achieving an ensemble source is trivial with non-factorized
 202 eikonal equation, but the methodology is not straightforward for factorized eikonal equation. If
 203 a source is an ensemble source consisting of multiple single-point sources $\{\mathbf{s}_i\}$, each point with
 204 different origin time $\{\eta_0(\mathbf{s}_i)\}$, then the background traveltimes field can be computed as

$$t_0(\mathbf{x}) = \min_{i=1, \dots, M_s} \{t_0(\mathbf{x}; \mathbf{s}_i) + \eta_0(\mathbf{s}_i)\}, \quad (4)$$

205 while the initial condition for τ remains the same: $\tau(\mathbf{x}; \mathbf{s}_i) = 1$ ($i = 1, 2, \dots, M_s$), where M_s is
 206 the number of single point sources in the ensemble source. Constructing t_0 for an ensemble source
 207 requires computing all $t_0(\mathbf{x}; \mathbf{s}_i)$ associated with the single-point sources. Therefore, we solve the
 208 forward modeling problem using

$$v^2 \sum_{i=1}^3 \left(t_0 \frac{\partial \tau}{\partial x_i} + \frac{\partial t_0}{\partial x_i} \tau \right)^2 = 1, \quad (5)$$

$$t(\mathbf{x}) = t_0(\mathbf{x})\tau(\mathbf{x}), \quad t_0(\mathbf{x}) = \min_{i=1, \dots, M_s} \{t_0(\mathbf{x}; \mathbf{s}_i) + \eta_0(\mathbf{s}_i)\}, \quad \tau(\mathbf{x}; \mathbf{s}_i) = 1. \quad (6)$$

209 To simplify notations, we denote equation (5) as a functional:

$$\mathcal{E}(v, t; \mathbf{s}, \eta_0) = \mathcal{E}(v, t; \{\mathbf{s}_i\}, \{\eta_0(\mathbf{s}_i)\}) = 0, \quad (7)$$

210 where $v = v(\mathbf{x})$ is the P- or S-wave velocity of a medium, and $\{\cdot\}$ represents an ensemble.

211 In elastic media, we solve two decoupled eikonal equations:

$$\mathcal{E}(v_p, t_p; \mathbf{s}, \eta_0) = 0, \quad \mathcal{E}(v_s, t_s; \mathbf{s}, \eta_0) = 0, \quad (8)$$

212 where the P- and S-traveltime fields share the same source location and origin time.

213 For the purpose of future extensibility, our LATTE also implements seismic reflection
214 traveltimes computation, where we assume that a reflector can be represented as an ensemble of
215 points, say, $I = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l\}$, i.e., a collection of l spatial coordinates. Then we solve the
216 following eikonal equations to obtain the PP reflection (i.e., P incidence to P reflection) traveltimes
217 associated with a reflector I^h ($h = 1, 2, \dots, H$):

$$\mathcal{E}(v_p, t_p; \mathbf{s}, \eta_0) = 0, \quad \mathcal{E}(v_p, t_{pp}^h; I^h, t_p(I^h)) = 0, \quad (9)$$

218 where the discrete traveltime $t_p(I)$ is the spatially varying origin time associated with the reflector.
219 The equation can be extended to arbitrary number of reflectors, but the incident traveltimes t_p
220 only needs to be computed once. Similarly, we identify the following eikonal equations to obtain
221 PS reflection (i.e., P incidence to S reflection) traveltimes in elastic media:

$$\mathcal{E}(v_p, t_p; \mathbf{s}, \eta_0) = 0, \quad \mathcal{E}(v_s, t_{ps}^h; I^h, t_p(I^h)) = 0, \quad (10)$$

222 where in the second equation, the velocity is v_s rather than v_p . This can be straightforwardly
223 extended to the case of S-wave incidence and SS/SP reflections.

224 In LATTE, we accelerate the fast sweeping eikonal solving (and the adjoint-state equation
225 solving detailed below) by adopting and slightly modifying the parallel strategy developed by
226 [Detrixhe et al. \(2013\)](#). For completeness, we detail our parallel fast-sweeping algorithms in
227 Appendix A.

228 2.2 First-arrival traveltimes tomography

229 Adjoint-state first-arrival traveltimes tomography (FATT) based on eikonal equation was estab-
230 lished by [Leung and Qian \(2006\)](#) and [Taillardier et al. \(2009\)](#). For completeness, we first review
231 the fundamental definitions of FATT. The original FATT is developed based on minimizing L_2 -
232 norm absolute-difference (AD) misfit function between the observed and the synthetic first-arrival
233 traveltimes ([Taillardier et al., 2009](#)), i.e.,

$$\mathcal{J}(v_p) = \frac{1}{2} \sum_{i=1}^{N_s} \sum_{j=1}^{N_r^i} [t(v_p, \mathbf{s}_i, \mathbf{r}_j) - T(\mathbf{s}_i, \mathbf{r}_j)]^2, \quad (11)$$

234 where N_s is the number of sources, and N_r^i is the number of receivers associated with the i -th
 235 source. AD-FATT is not immune to nonzero source origin time, i.e., if there is a nonzero origin
 236 time η_0 associated with T , it must be estimated before AD-FATT.

237 One can also formulate FATT using double-difference or differential time misfit function ([Zhang and Thurber, 2003; Tong et al., 2024](#))

$$\mathcal{J}(v_p) = \frac{1}{2} \sum_{i=1}^{N_s} \sum_{j=1}^{N_r^i} \sum_{k=1}^{N_r^i} [(t_p(v_p, \mathbf{s}_i, \mathbf{r}_j) - t_p(v_p, \mathbf{s}_i, \mathbf{r}_k)) - (T_p(\mathbf{s}_i, \mathbf{r}_j) - T_p(\mathbf{s}_i, \mathbf{r}_k))]^2. \quad (12)$$

239 DD-FATT is immune to nonzero source origin time. Both misfit functions can be used for inverting
 240 v_p and v_s . To simplify notations, we use a generalized symbol $\mathcal{D}(t, T; \mathbf{s}, \mathbf{r})$ to represent the misfit
 241 associated with a source \mathbf{s}_i :

$$\mathcal{D}(t, T; \mathbf{s}_i, \mathbf{r}) = \sum_{j=1}^{N_r^i} [t(v, \mathbf{s}_i, \mathbf{r}_j) - T(\mathbf{s}_i, \mathbf{r}_j)]^2, \quad (13)$$

$$\mathcal{D}(t, T; \mathbf{s}_i, \mathbf{r}) = \sum_{j=1}^{N_r^i} \sum_{k=1}^{N_r^i} [(t(v, \mathbf{s}_i, \mathbf{r}_j) - t(v, \mathbf{s}_i, \mathbf{r}_k)) - (T(\mathbf{s}_i, \mathbf{r}_j) - T(\mathbf{s}_i, \mathbf{r}_k))]^2, \quad (14)$$

242 where v is the velocity associated with t , and a symbol $\Delta(t, T; \mathbf{s}, \mathbf{r})$ to represent the traveltime
 243 residual associated with a trace \mathbf{r}_j :

$$\Delta(t, T; \mathbf{s}_i, \mathbf{r}_j) = t(v, \mathbf{s}_i, \mathbf{r}_j) - T(\mathbf{s}_i, \mathbf{r}_j), \quad (15)$$

$$\Delta(t, T; \mathbf{s}_i, \mathbf{r}_j) = \sum_{k=1}^{N_r^i} [(t(v, \mathbf{s}_i, \mathbf{r}_j) - t(v, \mathbf{s}_i, \mathbf{r}_k)) - (T(\mathbf{s}_i, \mathbf{r}_j) - T(\mathbf{s}_i, \mathbf{r}_k))]. \quad (16)$$

244 Based on our arguments in Appendix B, for a source \mathbf{s}_i , the adjoint-state equation for AD- or
 245 DD-FATT reads

$$\nabla \cdot [\lambda_p(\mathbf{x}) \nabla t_p(v_p, \mathbf{x})] = \sum_{j=1}^{N_r^i} \Delta(t_p, T_p, v_p, \mathbf{s}_i, \mathbf{r}_j). \quad (17)$$

246 The gradient of $\mathcal{J}(v_p)$ with respect to v_p associated with all sources is

$$\frac{\partial \mathcal{J}}{\partial v_p}(\mathbf{x}) = - \sum_{i=1}^{N_s} \frac{\lambda_{p,i}(\mathbf{x})}{v_p^3(\mathbf{x})}, \quad (18)$$

247 where $\lambda_{p,i}$ represents the adjoint-state variable associated with \mathbf{s}_i . The above equations can be
 248 straightforwardly extended to elastic case where both P and S-arrival traveltimes are used.

249 We adopt a ray-density preconditioning scheme to the common-source gradients (Li et al.,
 250 2017; Zhang et al., 2023) to improve traveltime field illumination and thus accelerate convergence.
 251 Specifically, for each source, in addition to computing the adjoint field λ , we also compute an
 252 adjoint energy field λ_e using

$$\nabla \cdot [\lambda_e(\mathbf{x}) \nabla t_p(v_p, \mathbf{x})] = \sum_{j=1}^{N_r} C(\mathbf{r}_j), \quad (19)$$

253 where $C(\mathbf{r}_j) = 1$, and compute the gradient as

$$\frac{\partial \mathcal{J}}{\partial v_p}(\mathbf{x}) = - \sum_{i=1}^{N_s} \frac{\lambda_{p,i}(\mathbf{x})}{\lambda_e + \epsilon \max(|\lambda_e|)}, \quad (20)$$

254 where we choose $\epsilon = 10^{-3} \sim 10^{-4}$ to avoid division of zero.

255 We implement three types of nonlinear inversion schemes in LATTE to perform FATT: steepest
 256 decent (SD), nonlinear conjugate gradient (NCG), and limited-memory Broyden-Fletcher-Goldfarb-
 257 Shanno (*l*-BFGS) schemes. The three schemes differ in the way of computing the search direction
 258 based on the gradient. Detailed algorithms can be found in Nocedal and Wright (2006).

259 All the three schemes require determination of an optimal step size at each iteration for properly
 260 updating the velocity model. In linear search or quadratic search method, an optimal step size is
 261 computed based on finding the root of quadratic equation. In such a case, usually the minimal
 262 number of forward modeling to determine the optimal step size is $M \geq 3$. Inspired by an early
 263 work for FWI (Gauthier et al., 1986), here we develop a perturbation-based method to compute the
 264 optimal step size. Specifically, at each FATT iteration l , after obtaining the search direction $\Delta m^{(l)}$,
 265 we compute the optimal step size as

$$\alpha_m^{(l)} = \frac{\sum_{i=1}^{N_s} |\beta_i| \cdot |\gamma_i|}{\sum_{i=1}^{N_s} \|\beta_i\|^2}. \quad (21)$$

266 The time residual vectors are

$$\beta_{i,j}(m) = \Delta \left(t(m^{(l-1)} + \epsilon \Delta m^{(l)}), t(m^{(l-1)}); \mathbf{s}_i, \mathbf{r}_j \right), \quad (22)$$

$$\gamma_{i,j}(m) = \Delta \left(T, t(m^{(l-1)}); \mathbf{s}_i, \mathbf{r}_j \right), \quad (23)$$

267 where the small perturbation trial coefficient $\epsilon = 0.05$, or equals to a small value that ensures
 268 $m_{\min} \leq m^{(l-1)} + \epsilon \Delta m^{(l)} \leq m_{\max}$, where m_{\min} and m_{\max} are box bounding limits set as hyper-
 269 parameters for the inversion, if necessary. In the above equations, Δm represents the search

270 direction computed using steepest decent, NCG, or l-BFGS. Therefore, at each iteration, in addition
 271 to solving the eikonal equation once and the adjoint-state equation once, we only need to solve
 272 the eikonal equation for one additional time based on the perturbed model $m^{(l-1)} + \epsilon \Delta m^{(l)}$ for
 273 determining $\alpha_m^{(l)}$. After that, the updated model in the l -th FATT iteration is

$$m^{(l)} = m^{(l-1)} + \alpha_m^{(l)} \Delta m^{(l)}. \quad (24)$$

274 Our strategy effectively reduces the total computational complexity of each tomography iteration
 275 from $\mathcal{O}((M+2)N_s)$ where $M \geq 3$ to $\mathcal{O}(3N_s)$. It is worth noting that in some cases, the number of
 276 required forward modeling can be $M > 1$, especially when the valley of loss is narrow and a large
 277 step may cause “overshoot,” i.e., the resulting step size causes the misfit in the l -th iteration (this
 278 iteration) to be higher than that at the $(l-1)$ -th iteration (previous iteration). In this case, given
 279 an initial step size $\alpha_m^{(l)}$ computed from the perturbation-bases strategy, we reduce the step size by
 280 half every time, and check if the resulting misfit is smaller than that of the last iteration. When it is
 281 smaller than the misfit of the last iteration, we then choose the reduced $\alpha_m^{(l)}$ as the optimal step size
 282 for this iteration. At early iterations, this check is almost not necessary as a search direction can
 283 always reduce the misfit. However, at later iterations where the inversion reaches a local minimum,
 284 such a trial-and-error may become necessary, making the number of additional forward modeling
 285 $M > 1$.

286 For elastic FATT, we compute the optimal step size in the l -th iteration as

$$\alpha^{(l)} = \frac{\sum_{i=1}^{N_s} (|\beta_i(v_p)| \cdot |\gamma_i(v_p)| + |\beta_i(v_s)| \cdot |\gamma_i(v_s)|)}{\sum_{i=1}^{N_s} (\|\beta_i(v_p)\|^2 + \|\beta_i(v_s)\|^2)}, \quad (25)$$

287 and update v_p and v_s using this step size:

$$v_p^{(l)} = v_p^{(l-1)} + \alpha^{(l)} \Delta v_p^{(l)}, \quad (26)$$

$$v_s^{(l)} = v_s^{(l-1)} + \alpha^{(l)} \Delta v_s^{(l)}. \quad (27)$$

288 Note that the optimal step size computed with equation (25) is more like an “averaged” or
 289 “balanced” optimal step size rather than a direct extension from equation (21) for multi-component
 290 traveltimes data, which can avoid unstable results for unbalanced data misfit during inversion.

291 2.3 Joint first-arrival traveltimes tomography and source location

292 Our LATTE contains a functionality, TLOC, to perform joint tomography, hypocenter location,
 293 and origin time inversion. The same functionality can also perform DD-based simultaneous

294 tomography and hypocenter location, mitigating the need of estimating an origin time for each
 295 source.

296 The misfit function for joint tomography-location in elastic media reads

$$\mathcal{J}(v_p, v_s, \mathbf{s}, \eta_0) = \frac{1}{2} \sum_{i=1}^{N_s} [\mathcal{D}(t_p + \eta_0(\mathbf{s}_i), T_p; \mathbf{s}_i, \mathbf{r}) + \mathcal{D}(t_s + \eta_0(\mathbf{s}_i), T_s; \mathbf{s}_i, \mathbf{r})]. \quad (28)$$

297 which generalizes the cases of joint FATT and hypocenter based on AD or DD traveltimes data
 298 misfit \mathcal{D} . Obtaining the gradients of \mathcal{J} with respect to source location requires the determination
 299 of $\nabla t(\mathbf{s}_i)$. Because the traveltimes field at the source location is a singularity, a direct computation
 300 of ∇t at \mathbf{s}_i is not mathematically meaningful. Therefore in LATTE, we exchange the location of
 301 source and receivers (\mathbf{s}, \mathbf{r}) to $(\hat{\mathbf{r}}, \hat{\mathbf{s}})$ during inversion and invert for the location of the virtual receiver
 302 $\hat{\mathbf{r}}$ instead:

$$\mathcal{J}(v_p, v_s, \hat{\mathbf{r}}, \zeta_0) = \frac{1}{2} \sum_{i=1}^{N_{\hat{s}}} [\mathcal{D}(t_p + \zeta_0(\hat{\mathbf{r}}), T_p; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}) + \mathcal{D}(t_s + \zeta_0(\hat{\mathbf{r}}), T_s; \hat{\mathbf{s}}_i, \hat{\mathbf{r}})], \quad (29)$$

303 where $\zeta_0(\hat{\mathbf{r}})$ is the virtual receiver's base time, which varies from virtual receiver to virtual receiver,
 304 but each virtual receiver's ζ_0 is consistent over different virtual sources. Exchanging sources and
 305 receivers requires an additional step prior to tomography; especially, one has to find all the unique
 306 virtual sources and virtual receivers and assign correspondingly the original traveltimes to these
 307 sources and receivers. In LATTE, we implement this step by leveraging message passing interface
 308 (MPI) based distributed-memory parallelism to reduce computational time.

309 The AD misfit function applies to joint velocity tomography and source location in equation (28)
 310 and equation (29). However, we remark that in this case, because the source and receivers are
 311 exchanged, for a common-virtual-source gather, the virtual receiver base time $\zeta_0(\hat{\mathbf{r}})$ (or real origin
 312 time $\eta_0(\mathbf{s}_i)$) differ from trace to trace. Therefore, in this case, $\zeta_0(\hat{\mathbf{r}})$ must be inverted, and the DD
 313 misfit function can no longer eliminates the common origin time for common-virtual-receiver gather
 314 as for common-real-source gather. The observation means that for joint tomography and source
 315 location, if one uses AD misfit function and if in practice the actual $\eta_0(\mathbf{s})$ are unknown, one must
 316 estimate η_0 along with estimating source location.

317 To apply DD to source location meanwhile avoiding inverting for the unknown η_0 , we define a
 318 similar but different misfit function. Taking the acoustic case as an example, the DD misfit function
 319 for joint tomography and location should read

$$\mathcal{J}_{DD}(v_p, \hat{\mathbf{r}}) = \frac{1}{2} \sum_{i=1}^{N_{\hat{s}}} \sum_{j=1}^{N_{\hat{r}}} \sum_{k=1}^{N_{\hat{s}}} [(t_p(v_p, \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) - t_p(v_p, \hat{\mathbf{s}}_k, \hat{\mathbf{r}}_j)) - (T_p(\hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) - T_p(\hat{\mathbf{s}}_k, \hat{\mathbf{r}}_j))]^2. \quad (30)$$

320 Note that the innermost summation for each virtual receiver is defined to integrate all the virtual
 321 sources. This contrasts with DD-FATT where the summation for each virtual receiver is defined
 322 to integrate all the virtual receivers. In fact, it is no longer possible to invert for η_0 with source-
 323 receiver-exchanged DD-TLOC because η_0 is not a part of the misfit function.

324 In this case, the residual for a virtual receiver \mathbf{r}_j is

$$\Gamma(t, T, \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) = \sum_{k=1}^{N_{\hat{s}}} [(t_p(v_p, \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) - t_p(v_p, \hat{\mathbf{s}}_k, \hat{\mathbf{r}}_j)) - (T_p(\hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) - T_p(\hat{\mathbf{s}}_k, \hat{\mathbf{r}}_j))]. \quad (31)$$

325 In the following, we distinguish the two cases with \mathcal{J}_{AD} and \mathcal{J}_{DD} , respectively.

326 Leveraging the derivations in [Tong \(2021a\)](#) developed for acoustic media, we derive the gradients
 327 of \mathcal{J}_{AD} with respect to the source parameters for elastic media as

$$\frac{\partial \mathcal{J}_{AD}}{\partial s_x}(\mathbf{s}_j) = \frac{\partial \mathcal{J}}{\partial \hat{r}_x}(\hat{\mathbf{r}}_j) = \sum_{i=1}^{N_{\hat{s}}} \left[\frac{\partial t_p}{\partial x} \Delta(t_p + \zeta_{0,j}, T_p; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) + \frac{\partial t_s}{\partial x} \Delta(t_s + \zeta_{0,j}, T_s; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) \right] \delta(\mathbf{x} - \hat{\mathbf{r}}_j), \quad (32)$$

$$\frac{\partial \mathcal{J}_{AD}}{\partial s_y}(\mathbf{s}_j) = \frac{\partial \mathcal{J}}{\partial \hat{r}_y}(\hat{\mathbf{r}}_j) = \sum_{i=1}^{N_{\hat{s}}} \left[\frac{\partial t_p}{\partial y} \Delta(t_p + \zeta_{0,j}, T_p; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) + \frac{\partial t_s}{\partial y} \Delta(t_s + \zeta_{0,j}, T_s; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) \right] \delta(\mathbf{x} - \hat{\mathbf{r}}_j), \quad (33)$$

$$\frac{\partial \mathcal{J}_{AD}}{\partial s_z}(\mathbf{s}_j) = \frac{\partial \mathcal{J}}{\partial \hat{r}_z}(\hat{\mathbf{r}}_j) = \sum_{i=1}^{N_{\hat{s}}} \left[\frac{\partial t_p}{\partial z} \Delta(t_p + \zeta_{0,j}, T_p; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) + \frac{\partial t_s}{\partial z} \Delta(t_s + \zeta_{0,j}, T_s; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) \right] \delta(\mathbf{x} - \hat{\mathbf{r}}_j), \quad (34)$$

$$\frac{\partial \mathcal{J}_{AD}}{\partial \eta_0}(\mathbf{s}_j) = \frac{\partial \mathcal{J}}{\partial \zeta_0}(\hat{\mathbf{r}}_j) = \sum_{i=1}^{N_{\hat{s}}} [\Delta(t_p + \zeta_{0,j}, T_p; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) + \Delta(t_s + \zeta_{0,j}, T_s; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j)] \delta(\mathbf{x} - \hat{\mathbf{r}}_j), \quad (35)$$

328 The above equations indicate that in either acoustic or elastic media, the gradients of \mathcal{J}_{AD} with
 329 respect to each of the source parameters (s_x, s_y, s_z, η_0) for each virtual receiver (or true source) is a
 330 scalar value summing from the contributions of all virtual sources (or true receivers).

331 In the case of DD-TLOC, the gradients of misfit function with respect to source location based
 332 on both P- and S-arrival traveltimes are given by

$$\frac{\partial \mathcal{J}_{DD}}{\partial s_x}(\mathbf{s}_j) = \frac{\partial \mathcal{J}_{DD}}{\partial \hat{r}_x}(\hat{\mathbf{r}}_j) = \sum_{i=1}^{N_{\hat{s}}} \left[\frac{\partial t_p}{\partial x} \Gamma(t_p, T_p; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) + \frac{\partial t_s}{\partial x} \Gamma(t_s, T_s; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) \right] \delta(\mathbf{x} - \hat{\mathbf{r}}_j), \quad (36)$$

$$\frac{\partial \mathcal{J}_{DD}}{\partial s_y}(\mathbf{s}_j) = \frac{\partial \mathcal{J}_{DD}}{\partial \hat{r}_y}(\hat{\mathbf{r}}_j) = \sum_{i=1}^{N_{\hat{s}}} \left[\frac{\partial t_p}{\partial y} \Gamma(t_p, T_p; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) + \frac{\partial t_s}{\partial y} \Gamma(t_s, T_s; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) \right] \delta(\mathbf{x} - \hat{\mathbf{r}}_j), \quad (37)$$

$$\frac{\partial \mathcal{J}_{\text{DD}}}{\partial s_z}(\mathbf{s}_j) = \frac{\partial \mathcal{J}_{\text{DD}}}{\partial \hat{\mathbf{r}}_z}(\hat{\mathbf{r}}_j) = \sum_{i=1}^{N_{\hat{\mathbf{s}}}} \left[\frac{\partial t_p}{\partial z} \Gamma(t_p, T_p; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) + \frac{\partial t_s}{\partial z} \Gamma(t_s, T_s; \hat{\mathbf{s}}_i, \hat{\mathbf{r}}_j) \right] \delta(\mathbf{x} - \hat{\mathbf{r}}_j), \quad (38)$$

with the DD misfit function Δ defined in equation (30). Again, the gradients of \mathcal{J}_{DD} with respect to each of the source location (s_x, s_y, s_z) for each virtual receiver (or true source) is a scalar value summing from the contributions of all virtual sources (or true receivers).

In LATTE, we invert for source parameters (including spatial location, and source origin time, if necessary) in the same manner as for model parameters. Therefore, the inversion scheme developed for model parameters seamlessly apply to source parameters inversion as well, resulting in a more consistent inversion scheme. This is contrast to the hybrid local-global inversion scheme by Tong (2021a). More importantly, such a consistent inversion scheme enables a more flexible way to regularize an inversion using the model and source regularization schemes that will be detailed in the next section.

It should be noted that in the joint tomography-location, we again compute the optimal step size using the small-perturbation strategy. However, in this case, we misfit vectors are

$$\beta_{i,j}(m, \mathbf{s}) = \Delta \left(t(m^{(l-1)} + \epsilon \Delta m^{(l)}, \hat{\mathbf{r}}_j^{(l-1)} + \Delta \hat{\mathbf{r}}_j^{(l)}), t(m^{(l-1)}, \hat{\mathbf{r}}_j^{(l-1)}); \hat{\mathbf{s}}_i \right), \quad (39)$$

$$\gamma_{i,j}(m, \mathbf{s}) = \Delta \left(T, t(m^{(l-1)}, \hat{\mathbf{r}}_j^{(l-1)}); \hat{\mathbf{s}}_i \right), \quad (40)$$

where $\Delta \hat{\mathbf{r}}_j^{(l)}$ represents the search direction of the virtual receiver (or real source) locations of the l -th iteration. Correspondingly, the update of the model and source parameters are

$$v_p^{(l)} = v_p^{(l-1)} + \alpha^{(l)} \Delta v_p^{(l)}, \quad (41)$$

$$v_s^{(l)} = v_s^{(l-1)} + \alpha^{(l)} \Delta v_s^{(l)}, \quad (42)$$

$$\hat{\mathbf{r}}_j^{(l)} = \hat{\mathbf{r}}_j^{(l-1)} + \alpha^{(l)} \Delta \hat{\mathbf{r}}_j^{(l)}. \quad (43)$$

Similarly, if one needs to solve for the origin time η_0 , then β and γ should be computed by properly considering the perturbation of η_0 under search direction $\Delta \eta_0$ like $\hat{\mathbf{r}}_j$: $t(m^{(l-1)} + \epsilon \Delta m^{(l)}, \hat{\mathbf{r}}_j^{(l-1)} + \Delta \hat{\mathbf{r}}_j^{(l)}, \eta_{0,j}^{(l-1)} + \Delta \eta_{0,j}^{(l)})$.

2.4 Model and source parameter regularization

For both functionalities that involve model parameter update (i.e., FATT and TLOC), we develop a novel model parameter regularization to improve the geological fidelity of the inversion results. The model parameter regularizer consists of a total generalized p -variation regularizer (Gao and Huang, 2019) and a P/S wave velocity structure similarity regularizer. In addition, for TLOC, we

355 introduce a novel source parameter regularizer based on a end-to-end, supervised ML model to
 356 improve the geological fidelity of inverted source locations.

357 Specifically, we define the regularized joint tomography-location as a hybrid optimization
 358 problem:

$$\begin{aligned} \mathcal{J}(v_p, v_s, \mathbf{s}, \eta_0) = & \sum_{i=1}^{N_s} \mathcal{D}(t_p, T_p; \mathbf{s}, \eta_0) + \sum_{i=1}^{N_s} \mathcal{D}(t_s, T_s; \mathbf{s}, \eta_0) \\ & + \omega_{v_p} \mathcal{T}(v_p) + \omega_{v_s} \mathcal{T}(v_s) + \omega_{v_p/v_s} \|1 - \mathcal{S}(v_p, v_s)\|^2 \\ & + \omega_s \|\mathcal{F}(\mathbf{s})\|^2, \end{aligned} \quad (44)$$

359 where for convenience, we drop the coefficient $\frac{1}{2}$ associated with every misfit term.

360 The operator \mathcal{T} is an ℓ_p -norm minimization problem defined as (Knoll et al., 2011; Gao and
 361 Huang, 2019):

$$\mathcal{T}(v_p) = \min_m \{\alpha_1 \|\nabla v_p - m\|_p^p + \alpha_2 \|\varepsilon(m)\|_p^p\} \quad (45)$$

362 with the norm $0 \leq p \leq 1$, and α_1 and α_2 are weighting factors for the first- and second-order total
 363 variations, respectively. In 3D, the gradient matrix $\varepsilon(m)$ for a vector field $m = (m_x, m_y, m_z)$ reads

$$\varepsilon(m) = \begin{bmatrix} \nabla_x m_x & \frac{1}{2}(\nabla_x m_y + \nabla_y m_x) & \frac{1}{2}(\nabla_x m_z + \nabla_z m_x) \\ \frac{1}{2}(\nabla_x m_y + \nabla_y m_x) & \nabla_y m_y & \frac{1}{2}(\nabla_y m_z + \nabla_z m_y) \\ \frac{1}{2}(\nabla_x m_z + \nabla_z m_x) & \frac{1}{2}(\nabla_y m_z + \nabla_z m_y) & \nabla_z m_z \end{bmatrix}. \quad (46)$$

364 The operator \mathcal{S} measures the similarity between v_p and v_s , by which we intend to improve the
 365 structural similarity between updated v_p and v_s . Although there are sophisticated structure similarity
 366 operators, in LATTE, we impose the similarity simply through constraining and smoothing the ratio
 367 between v_p and v_s , i.e., v_p/v_s .

368 We also impose a source regularization term through \mathcal{F} , which is a misfit function that minimizes
 369 the spatial spreading of source locations. In other words, we want the inverted seismicity locations
 370 to be “focused” as much as possible. However, we do not want all the inverted seismicity converges
 371 to a single spatial location by the regularizer \mathcal{F} ; otherwise, the solution will be of low seismological
 372 fidelity. In LATTE, we focus on fault/fracture-related seismicity, and therefore intend to develop a
 373 regularizer that improve the consistency between the inverted source locations with one or multiple
 374 fault/fracture surfaces.

375 It is essentially difficult to compute the gradient $\partial \mathcal{T}/\partial v_p$ or $\partial \mathcal{T}/\partial v_s$ as \mathcal{T} itself is defined
 376 through an optimization problem rather than an analytical equation. The same challenge occurs
 377 to computing $\partial \mathcal{F}/\partial \mathbf{s}$, because \mathcal{F} does not have an analytical expression. To solve the regularized
 378 inversion problem, we convert regularized optimization in equation (44) to an alternating-direction

379 optimization:

$$\begin{aligned} v_p^{(l+1)}, v_s^{(l+1)}, \mathbf{s}^{(l+1)} &= \arg \min_{v_p, v_s, \mathbf{s}} \sum_{i=1}^{N_s} \mathcal{D}(t_p + \eta_0, T_p) + \sum_{i=1}^{N_s} \mathcal{D}(t_s + \eta_0, T_s) \\ &\quad + \omega_{v_p} \|v_p - m_p^{(l)}\|^2 + \omega_{v_s} \|v_s - m_s^{(l)}\|^2 + \omega_{v_p/v_s} \left\| \frac{v_p}{v_s} - r^{(l)} \right\|^2 \\ &\quad + \omega_s \|\mathbf{s} - \sigma^{(l)}\|^2, \end{aligned} \quad (47)$$

$$m_p^{(l+1)} = \arg \min_{m_p} \gamma_{v_p} \mathcal{T}(m_p) + \omega_{v_p} \|v_p^{(l+1)} - m_p\|_2^2, \quad (48)$$

$$m_s^{(l+1)} = \arg \min_{m_s} \gamma_{v_s} \mathcal{T}(m_s) + \omega_{v_s} \|v_s^{(l+1)} - m_s\|_2^2, \quad (49)$$

$$r^{(l+1)} = \arg \min_r \omega_{v_p/v_s} \left\| \mathcal{S} \left(\frac{v_p^{(l+1)}}{v_s^{(l+1)}} \right) - r \right\|_2^2, \quad (50)$$

$$\sigma^{(l+1)} = \arg \min_{\sigma} \omega_s \|\mathcal{F}(\mathbf{s}^{(l+1)}) - \sigma\|^2. \quad (51)$$

380 The first optimization problem is simply first-arrival traveltimes tomography (or joint tomography-
 381 location) by adding $v_p - m_p^{(l)}$, $v_s - m_s^{(l)}$, or $\mathbf{s} - \sigma^{(l)}$ to the gradients of model parameters or source
 382 parameters in each iteration, respectively. This will gradually guide the model and source parameters
 383 converge to $m_p^{(l)}$, $m_s^{(l)}$, and $\sigma^{(l)}$, respectively, which are solved via the following optimization
 384 problems.

385 The second and third optimization problems are TGpV image denoising problems. We solve \mathcal{T}
 386 optimization (TGpV 2D and 3D image denoising) using the algorithm in [Gao and Huang \(2019\)](#),
 387 with an open-source implementation we developed in [Gao and Chen \(2024\)](#). The input to the
 388 optimization is the updated v_p or v_s model, while the output is “denoised” or “regularized” v_p or v_s
 389 model.

390 The fourth optimization is not a strict minimization problem, but to impose a constraint \mathcal{S} on v_p
 391 and v_s so that they are structurally similar. There are many choices for this constraint. In LATTE, \mathcal{S}
 392 composes of three operations: box limiting, median filtering, and Gaussian smoothing. The box
 393 limiting \mathcal{B}_a^b constraints the ceiling and floor values for the ratio v_p/v_s , which mimics the fact that
 394 in practice this value is generally not arbitrary but lies within a range $[a, b]$. While for different
 395 geologies or materials this range can be different, in practice the approximate values of a and b are
 396 not completely unknown. The other operations, including a median filtering \mathcal{M} and a Gaussian
 397 smoothing \mathcal{G}_σ , reduce abrupt spatial variations of v_p/v_s , making v_p and v_s closer in structures.
 398 The standard deviation σ is a hyper-parameter that can be chosen differently for different models
 399 depending on a user’s preference. Alternatively, one can also use more sophisticated smoothing,
 400 such as structure-oriented nonlinear anisotropic diffusion ([Wu and Guo, 2018](#)), to smooth the ratio.

401 In our code, we find the composite operation $\mathcal{S} = \mathcal{G}_\sigma \circ \mathcal{M} \circ \mathcal{B}_a^b$ suffices the purpose of similarizing
 402 v_p and v_s .

403 Similarly, the fifth optimization, i.e., the source parameter regularization problem \mathcal{F} , is not a
 404 strict optimization problem. The purpose of this optimization is to improve the spatial correlation
 405 among the inverted seismicity locations. For geophysical applications, seismicity does not occur
 406 randomly, and in general the locations of seismicity are strongly correlated with faults or fractures.
 407 Therefore, we want the inverted source locations fall on one or multiple faults/fracture surfaces
 408 as much as possible, effectively making \mathcal{F} a fault geometry constraint for seismicity. For other
 409 types of applications where sources do not essentially correlate with faults/fractures, and we can
 410 conveniently ignore this fault geometry constraint by setting $\omega_s = 0$.

411 The fault constraint is not trivial to solve. For instance, one can use automatically clustering
 412 algorithms to cluster inverted seismicity locations, and then use some surface fitting algorithm to
 413 move the clustered events to a plane or surface. However, it is very challenging to develop an
 414 adaptive clustering algorithm that works generally well for different scenarios, especially when the
 415 locations distribute irregularly with drastically different densities in space. Instead, we solve this
 416 fault constraint problem using a supervised machine learning model. At each iteration, we convert
 417 the source locations $\mathbf{s}^{(l+1)}$ to a 2D or 3D grid-based image using a maximum-limiting summation
 418 of Gaussian functions:

$$I(\mathbf{x}) = \max_{i=1, N_f} \exp\left(-\frac{\|\mathbf{x} - \mathbf{s}_i\|^2}{2\sigma^2}\right), \quad (52)$$

419 where we omit the superscript $(l + 1)$ for simplicity, $\|\cdot\|$ represents L_2 norm, and σ represents
 420 the standard deviation of the Gaussian function. In our code, we set $1/2\sigma^2 = 0.3$, resulting in an
 421 annihilating amplitude approximately eight or nine grid points away from the source location.

422 We develop an iterative, multitask ML model shown in Figure 32 of Appendix C to infer and
 423 refine faults and fault attributes (including probability, dip, strike) from a source image. Specifically,
 424 this ML models contains two neural networks (NN): a multitask inference NN and a multitask
 425 refinement NN. The multitask inference ML model is an end-to-end model where the output has the
 426 same dimensions as the input source image. The multitask refinement ML model is also an end-
 427 to-end model, but the input to this refinement NN contains the source image and the inferred fault
 428 attributes, which might be “noisy” or “broken” due to imbalanced source locations. By applying
 429 the refinement NN several times based on the results obtained from a previous iteration, we obtain
 430 cleaner, continuous, and thus more interpretable faults compared with the ones generated from
 431 the multitask inference NN. Upon obtaining the fault attributes, we generate a fault-constrained
 432 source image by moving every source \mathbf{s}_i to the nearest fault point in terms of Euclidean distance
 433 and use the fault-constrained source locations to guide the iterative source location. Therefore,
 434 over iterations, the source locations become more topologically meaningful in terms of correlation

435 with fault/fracture surfaces, resulting in better seismological or geological interpretability. The
 436 two objects (source image and fault image) mutually improve each other through the ML model,
 437 facilitating the update of source locations for TLOC.

438 In a qualitative manner, our ML-based source parameter regularization for the l -th iteration
 439 could be represented as

$$\begin{aligned} \text{updated source location } \mathbf{s}^{(l)} &\rightarrow \text{multitask inference} \\ &\rightarrow \underbrace{\text{multitask refinement}}_{\text{repeat } N \text{ times}} \rightarrow \text{regularized location } \sigma^{(l)}. \end{aligned} \quad (53)$$

440 The regularization procedure will generate $\sigma^{(l)}$ that essentially falls on a fault/fracture surface, and
 441 act as a guidance for updating $\mathbf{s}^{(l+1)}$ in the next iteration. Eventually, $\mathbf{s}^{(N)} \approx \sigma^{(N)}$ at the end of a
 442 TLOC inversion.

443 3 Numerical results

444 3.1 Traveltime computation

445 We first use a synthetic velocity model to validate the traveltime computation functionalities of
 446 LATTE.

447 LATTE enables both first-arrival traveltime computation and reflection traveltime computation.
 448 Figures 1a and b display the v_p and v_s models with low and high velocity anomalies, respectively.
 449 Figure 1c displays two reflectors (represented by values 1 and 2, respectively).

450 Figures 2 displays the first P-arrival traveltime field for a source placed at the horizontal position
 451 of 0.3 km. The gradient in the velocity model causes notable diving wave features in the traveltime
 452 field. Figures 2b and c display the PP-arrival traveltime fields associated with the first and second
 453 reflectors, respectively. Note that below each reflector, the computed traveltime field does not
 454 represent reflection but transmission traveltime. Therefore, the traveltime field below each reflector
 455 is the same with the first-arrival traveltime in Figure 2a. Figures 2d and e display the PS-arrival
 456 traveltime fields associated with the first and second reflectors, respectively. In this case, the
 457 traveltime field below each reflector represents PS transmission arrival traveltime, and they are not
 458 the same with the PP transmission arrival traveltime displayed in Figure 2a.

459 Figure 3a displays the first-arrival traveltime field. Figures 3b and c display the SP reflection
 460 traveltime fields, where the traveltime field below each reflector represents SP transmission arrival
 461 traveltime field. Similarly, Figures 3d and e display the SS reflection traveltime fields, where the
 462 traveltime field below each reflector represents SS transmission arrival traveltime, and is consistent

463 with the values displayed in Figure 3a.

464 Figures 4a and b display the traveltime recorded at the surface of the validation model. We
465 observe that in this case because $v_p/v_s > 1$ in the entire model, PP1 and PP2 (representing the
466 PP reflection traveltimes from reflectors 1 and 2, respectively) arrivals are always earlier than PS
467 reflection traveltimes (PS1 and PS2); at large offsets, the PP reflections may arrive earlier than the
468 first-arrival traveltime, which is essentially a mixture of direct wave at near offsets and diving wave
469 arrivals at large offsets.. By contrast, for SS source, the SP reflections arrive notably earlier than
470 both the “first-arrival” traveltime and the SS reflection traveltime field. The results indicate that for
471 elastic media characterization, picking first-arrival S traveltimes can be very challenging depending
472 on the geometry. At large offsets, SP component may arrive much earlier than other components,
473 and the true “first-arrival” S-wave may be obscured in noisy waveforms.

474 3.2 First-arrival traveltime tomography

475 We use a near-surface faulted model to demonstrate the efficacy of AD-FATT and DD-FATT
476 implementation in our LATTE.

477 Figure 5a displays a faulted velocity model with a horizontal span of 4 km and a maximum depth
478 of 500 m. We set a number of structural complexities of the layers including anticlines, inclines,
479 and faults. We also set a low velocity value (600 m/s) for the three faults. The model consists of
480 51 grid points in the depth direction, with a grid spacing of 10 m, and with 401 grid points in the
481 horizontal direction, with a grid spacing of 10 m. Figure 5b displays a smooth velocity model (1D
482 linear gradient) model as the initial model for both AD-FATT and DD-FATT. All the important
483 features of the ground-truth velocity model are invisible on this initial velocity model. We place a
484 total of 40 sources on the top surface of the model, starting from 50 m and with a uniform horizontal
485 spacing of 100 m. We place a total of 401 receivers on the top surface, with a uniform horizontal
486 spacing of 10 m. For both inversions, we apply energy preconditioning to the gradient and adopt
487 the NCG inversion scheme to obtain the search direction.

488 Figures 6a and b display the inverted v_p models by AD-FATT and DD-FATT, respectively, both
489 after 100 iterations. Meanwhile, Figure 7 displays the normalized data misfit convergence curves of
490 AD-FATT and DD-FATT using blue and red curves, respectively. Both inversions correctly recover
491 the low-wavenumber features of the ground-truth model. However, both inversion results indicate
492 that it could be very challenging to accurately delineate high-resolution features based solely on
493 first-arrival traveltimes, even though both AD-FATT and DD-FATT converge to a low data misfit.
494 For instance, both inversions miss the deep part of two low-velocity faults between 2 km to 3.5 km
495 at the horizontal position, although both correctly recover the shallow part of the faults.

496 Figure 8a displays a comparison between the ground-truth and synthetic traveltimes in the initial

497 1D velocity model for the second source at 20 m. Figures 8b and c show the ground-truth and
498 synthetic traveltimes for the same source in the inverted velocity models obtained using AD-FATT
499 and DD-FATT, respectively. We observe that both inversions generate an accurate first-arrival
500 traveltimes after 100 iterations. We also display the distribution of traveltime misfit in the
501 inverted model for AD-FATT and DD-FATT in Figures 8d and e, respectively. Through the statistics,
502 we find that DD-FATT results in slightly more consistent traveltimes for a total of 16,040 traveltime
503 measurements, even though it uses differential traveltimes rather than absolute traveltimes for residual
504 and misfit computation.

505 The results demonstrate the efficacy of AD-FATT and DD-FATT functionalities implemented in
506 LATTE.

507 3.3 Source location and joint tomography-location

508 Next, we validate our method and implementation of traveltime-based source location, as well
509 as joint tomography-location, in LATTE. Same as in the last example, we validate both AD and DD
510 misfit functions, and we use AD-TLOC and DD-TLOC to denote these two cases. As we described
511 in the methodology, we exchange sources and receivers for source location, indicating that the
512 receivers are not placed on the surface of the model. The tomography results displayed below will
513 therefore demonstrate the validity of our arguments on placing receivers at arbitrary positions of a
514 model as detailed in Appendix B.

515 We use an elastic checkerboard model and both P- and S-arrival traveltimes in this test. Figure 9a
516 displays the ground-truth v_p model, and the S-wave velocity model is set to $v_s = v_p/\sqrt{3}$ for
517 simplicity. The dimension of the model is 3 km in the depth direction and 4 km in the horizontal
518 direction. The grid spacing is 10 m in both directions. We set a total of 300 randomly distributed
519 sources within the model, and a total of 50 receivers on the top surface, starting from 50 m and with
520 a uniform horizontal spacing of 100 m. In addition, we set random origin time η_0 ranging from 0 to
521 100 s for the 300 sources. We generate traveltimes using the parallel fast-sweeping elastic eikonal
522 solver implemented in LATTE.

523 In the first test for this checkerboard model, we use AD-TLOC to simultaneously invert for
524 the source location and origin time by assuming known velocity models. Figure 10a displays the
525 ground-truth location of the 300 sources in space. For validating AD-TLOC, we set the initial guess
526 of location of all sources at the center of the model as denoted by the red dot. We also set the initial
527 guess of origin time η_0 to be the mean of all ground-truth η_0 which is approximately 50 s.

528 Figures 10b-c display the inverted source locations at the 5th, 10th, and 100th iterations,
529 respectively, where we use gray lines to connect the ground-truth and inverted source locations.
530 We observe that the inverted source locations gradually converge to their ground-truth positions.

531 Sources in the deep and boundary regions of the model appear to have slightly larger errors because
532 of the insufficient traveltimes field coverage and stacking in these regions.

533 Figure 11a displays the comparison among the ground-truth origin time, the initial guess, and
534 the inverted origin time of the 300 sources. The inversion result indicates that AD-TLOC correctly
535 estimates the origin time given a trivial initial guess (a same value for all sources). Figure 11b
536 displays the comparison among the ground-truth T_p subtracting the ground-truth origin time (i.e.,
537 $T_p - \eta_0$, blue curve), the synthetic t_p in the initial smooth velocity model (i.e., $t_p^{(l=0)}$, green curve),
538 and the synthetic t_p in the inverted velocity in the inverted model ((i.e., $t_p^{(l=100)}$ red curve), all
539 corresponding to the second virtual source. Similarly, Figure 11c displays the comparison for
540 S-arrival traveltimes, where we observe a similar level of accuracy. Common-virtual-source gathers
541 at other locations show similar level of error with the ones displayed in Figures 11b and c. The
542 consistency between the traveltimes computed in the inverted model and the ground-truth traveltimes
543 validates the efficacy and accuracy of AD-TLOC in LATTE.

544 In the second test for this checkerboard model, we perform DD-TLOC by again assuming known
545 velocity models. We set a trivial initial guess for the location of all sources – the center of the model.
546 In the test, we only invert for the source locations. As we described in the text, using DD-only
547 misfit functions, we cannot invert for the origin time as η_0 is eliminated by the DD misfit function
548 for each real-source gather (or virtual-receiver gather).

549 Figure 12a displays the ground-truth and initial guess of the source locations, while Figures 12b-
550 d show the inverted source locations in the 5th, 10th, and 100th iterations, respectively. Comparing
551 with the AD-TLOC inversion results shown in Figures 10b-d, we find that by eliminating the
552 common origin time, DD-TLOC results in a more accurate estimation of source location for almost
553 all the sources. There are several sources in the deep part of the model that are not well located, but
554 these sources are also furthest away from the receivers. The differential traveltime misfits associated
555 with these sources are also the smallest, resulting in a suboptimal update of these deep sources.

556 Figure 13a displays the comparison among the ground-truth $t_p - \eta_0$, $t_p^{(l=0)}$, and $t_p^{(l=100)}$. Even
557 though we use DD misfit function rather than AD misfit function, we observe a good consistency
558 between the observed and the synthetic traveltimes. In fact, the traveltime misfits associated with
559 this common-virtual-source gather is better than that in the AD-TLOC displayed in Figure 11b. The
560 consistency of S-arrival traveltimes displayed in Figure 13b is at a similar level with the P-arrival
561 traveltimes, and again is higher than that generated by AD-TLOC displayed in Figure 11c. The
562 results validate the efficacy and accuracy of DD-TLOC in LATTE, and demonstrate the advantage
563 of DD-TLOC over AD-TLOC in leveraging differential time to improve source location accuracy.

564 In the third test for this checkerboard model, we perform simultaneous velocity tomography
565 and source location using DD-TLOC. We assume homogeneous initial velocity models $v_p^{(l=0)} =$
566 2000 m/s and $v_s^{(l=0)} = 2000/\sqrt{3}$ m/s, and set the initial guess of source location to be $(x_0, z_0) =$

567 $(\mu(s_x), 2980)$ m where we use $\mu(s_x)$ to denote the average value of the horizontal positions of
568 all sources. Because seismic velocity and source location are strongly coupled in terms of
569 traveltimes, and in this test we only have surface receivers, we anticipate a poorer source location
570 result compared with those of the first two tests.

571 Figures 14a and b display the inverted v_p and v_s models by DD-TLOC, respectively. Compared
572 with the ground-truth model in Figure 9, we find that the central part of the model is relatively
573 better recovered than the regions in the deep and boundary regions. This is probably because
574 the background model is a homogeneous model for both v_p and v_s , therefore there is not diving
575 wave/traveltimes field to leverage for updating the deep region.

576 Figure 15a compares the ground-truth and the initial source locations, while Figures 15b-c
577 compared the ground-truth and inverted source locations in the 5th, 10th, and 100th iterations,
578 respectively. Comparing the inverted source locations with that in Figure 10d where the velocity
579 model is known, we observe that inaccurate velocity models introduce a notable challenge to source
580 location. In this case, the sources in the deep and boundary regions show notably higher level of
581 error than those in the known-velocity case. The inaccuracy is consistent with the low accuracy of
582 inverted velocity models displayed in Figure 15.

583 Figure 16 compares the ground-truth traveltimes $t_p - \eta_0$ and $t_s - \eta_0$ with synthetic traveltimes
584 associated with the second virtual source. We observe visually higher standard deviations for both
585 P- and S-arrival traveltimes compared with those in the first two tests.

586 The above results indicate the limitation of DD-TLOC for joint tomography-location in an elastic
587 model with poor initial guesses of velocity models and source locations. Comparing the results
588 with those in the location-only tests, we find that velocity uncertainty can deteriorate the accuracy
589 of source location. Because velocity and source location are strongly coupled, the influence is
590 essentially mutual and cannot be straightforwardly decoupled.

591 In the last part of the Methodology section, we introduced model parameter regularization
592 consisting of TGpV and $v_p - v_s$ similarity regularizers to FATT and TLOC. For this model, because
593 the source locations are purely random, the ML-based source parameter regularization does not
594 apply – there is no fault that the sources can align to. In the fourth test, therefore, we validate the
595 efficacy of the model parameter regularization by setting $\omega_s = 0$.

596 Figure 17 display the inverted v_p and v_s models in the 100th iterations using the regularized
597 DD-TLOC joint tomography-location functionality. Compared with those without model parameter
598 regularization displayed in Figure 14, we find that model parameter regularization notably reduces
599 random-noise-like artifacts in the inversion results. The pattern of checkerboard in this case becomes
600 clearer, more closely resembling the ground-truth model in Figure 9. Similar with the case without
601 model parameter regularization, the most well-recovered region is the central part of the model,
602 with less accurate recovery of velocity perturbation in the deep and boundary regions. We must

remind that the inaccuracy is not intrinsic to FATT or TLOC in LATTE. Any tomography methods may encounter similar issue as the inaccuracy is essentially determined by the poor illumination of these regions with a surface-only receiver distribution.

Figure 18 display the initial and inverted source locations using regularized DD-TLOC. Although in this case the sources in the deep and boundary regions still cannot be well located, visually the errors are smaller compared with those in Figure 15. Because we do not regularize source parameters in this test, the improvement of location accuracy is essentially introduced by the better-resolved velocity models.

Lastly, the traveltime comparison in Figure 19 further demonstrates the improvement in traveltime consistency introduced by regularized DD-TLOC functionality compared with plain DD-TLOC.

3.4 Fault-constrained source location

In the fourth example, we demonstrate the efficacy of fault-constrained source location functionality in LATTE.

Figure 20a displays a 2D v_p model of 3 km in the depth direction and 5 km in the horizontal direction. The model consists of a smoothly varying upper part and faulted structures in the lower part. Figure 20b displays a Gaussian-smoothed velocity model for locating the sources.

We set a total of 50 receivers on the surface, starting from 50 m and with a uniform interval of 100 m, and also set 30 receivers at the horizontal position of 2 km, starting from 100 m in depth, and with a uniform interval of 60 m. The vertically distributed receivers mimic the scenario of receivers placed in a well. We set a total of 1,200 sources along the faults in the lower half, mimicking the scenario of fracturing-induced seismicity. Same with the previous tests, we assume a trivial initial source location in the center of the lower half model at $(x_0, z_0) = (2500, 2500)$ m.

To improve reality, we assign random values ranging from 0 to 10 s as the origin time for these 1,200 sources and use DD-TLOC to invert for the source locations. In this test, we do not update the velocity. However, we add smoothed random noise to the computed traveltime in the ground-truth model to mimic imperfect traveltime picking in practice, as displayed in Figure 21. The maximum value of the added noise in all common-source gathers is 20 ms. Translating to spatial distance under this velocity model, the noise generates up to approximately 50-m random errors in space for each source.

As we point out in the Methodology, one may want to avoid an “early kick-in” when the source locations are still far away from the truth locations. In practice, one may need to use experience to decide when to regularize source parameters. We start to regularize source location update using the ML-based regularizer starting from the 16th iteration. Figures 22a and b display the initial

source location and the updated source locations in the 5th iteration. Figures 22c and d display the inverted source locations by plain DD-TLOC and fault-constrained DD-TLOC, respectively. It is visually evident that, without fault constraint for source location, the noise in the data can affect the accuracy of location, even though both inversions can converge in terms of data misfit. In the deep and boundary regions, located sources can smear into each other, making it difficult to correlate the located sources with individual faults. By contrast, the fault-constrained DD-TLOC results in a more interpretable source location result, where the located sources are mostly aligned with faults and close to their ground-truth locations. In fact, in the fault-constrained location result, we only observe few notable mislocated sources in the lower right corner, and two in the lower left corner. Given better illumination of these regions, it is likely these errors can be further reduced.

Figures 23a and b display the source image and ML-inferred fault dip image in the 5th iteration. At early iterations, the sources are not yet well located, and therefore the inferred faults do not resemble the ground-truth faults denoted by the red points. By contrast, Figures 22c and d display the source image and corresponding ML-inferred fault dip image at the 50th iteration. We observed a good consistency between the source image and the ground-truth source locations, as well as a good consistency between the inferred faults and ground-truth source locations. The results demonstrate that our ML-based source parameter regularization can gradually guide or constrain updated source locations towards faults inferred from the source locations themselves, eventually leading to higher fidelity and interpretability of located source locations.

3.5 Regularized joint tomography and source location

In the final example, we demonstrate the efficacy and accuracy of DD-TLOC joint tomography-location with both model and source parameter regularization for a 3D elastic model.

Figure 24a displays a 3D v_p model. The background variation of this model is a smooth model displayed in Figure 24b with several intersecting faults. The range of background velocity is [1000, 3000] m/s. We add a 3D checkerboard velocity perturbation with a range of [-300, 300] m/s to the background velocity model and obtain the velocity model in Figure 24a. The model is 1 km in depth, 2 km along the Y direction, and 3 km along the X direction. The v_s model follows the same background spatial variation pattern, yet with a different value range from [500, 2000] m/s; the checkerboard perturbation added to the background v_s model has a range of $[-300, 300]/\sqrt{3}$ m/s. Therefore, the resulting v_p and v_s velocity models have nonuniform ratios in space. The background velocity models also serve as the initial velocity model in the following tests.

We display four faults in the upper right corner of Figure 24a. We set a total of 1,200 sources randomly distributed on the faults. We set a total of $R_x \times R_y = 15 \times 10$ receivers on the surface of the model. For this DD-TLOC test, we exchange the sources and receivers for simultaneous velocity

671 update and source location inversion, therefore there are effectively 150 common-virtual-source
 672 gathers after reciprocity traveltime data rearrangement.

673 Similar with the previous example, to mimic practical noise caused by inaccurate phase picking,
 674 we add smoothed random noise to the traveltime data simulated in the ground-truth models as
 675 displayed in Figure 25. The difference is that for this test, we set a higher maximum amplitude for
 676 the noise (50 ms), which intuitively may result in higher uncertainties to the inversions.

677 Figure 26 display the inverted v_p and v_s models without model or source parameter regulariza-
 678 tion. We observe that the general heterogeneity pattern of the ground-truth velocity models are
 679 revealed in both v_p and v_s . However, there are numerous random velocity perturbations in the
 680 inverted models. These artifacts are possibly due to the uneven coverage of the sources and receivers,
 681 as well as the random noise in data. By contrast, in Figure 27, we display the inverted models and
 682 source locations with simultaneous model and source parameter regularization as described in the
 683 Methodology section. The inverted velocity models are notably cleaner than the ones without model
 684 parameter regularization, where we observe almost no random artifacts.

685 We further compare horizontal slices at two depths among the ground-truth and the inverted
 686 velocity perturbations without and with model regularization. Figure 28 display the comparison
 687 of a horizontal slices at a depth of 100 m. The regularized DD-TLOC generates a horizontal
 688 velocity perturbation with notably piecewise smooth velocity variations than those without model
 689 regularization. Figure 29 display a similar comparison for the depth of 340 m, which shows similar
 690 improvement by our TGpV model regularization.

691 We display the map view of the ground-truth source locations in Figure 30a and the initial guess
 692 for both inversions in Figure 30. We define the horizontal and vertical errors of an inverted source
 693 location as

$$E_h = \sqrt{(s_x - s_{x,0})^2 + (s_y - s_{y,0})^2}, \quad (54)$$

$$E_z = |s_z - s_{z,0}|, \quad (55)$$

694 where (s_x, s_y, s_z) represents the inverted source location and $(s_{x,0}, s_{y,0}, s_{z,0})$ represents the ground-
 695 truth location. Figures 30c and d display the horizontal and vertical errors of the inverted source
 696 locations without source parameter regularization. The results show that with only surface receivers,
 697 the horizontal locations are more accurately estimated than the vertical locations. If well receivers
 698 are available, in principle the vertical errors should reduce. Figures 30e and f display the horizontal
 699 and vertical errors of the source locations inverted with DD-TLOC with our ML-based source
 700 parameter regularization, which show an improved consistency with the faults compared with those
 701 without source regularization, especially at the ends and intersection regions of the faults.

702 In Figures 31a and b, we compared the ground-truth source locations (blue balls) and the

703 inverted source locations (red balls) without and with ML-based source parameter regularization. It
704 is evident the ML-based source parameter regularization result in an improved location accuracy.
705 Meanwhile, Figures 31c and d display the inferred and refined faults generated by our multitask ML
706 model in the 50th iteration based on the inverted source locations. The consistency between the
707 ground-truth source locations and the estimated faults demonstrates the efficacy of our multitask
708 ML model in serving as an adaptive guide for source location.

709 The results for this example demonstrate the efficacy of our model and source parameter
710 regularization. We remark that the fault-constrained source parameter regularization does not apply
711 to scenarios where the sources do not correlate to faults or fault-like structures. In those scenarios,
712 it may be possible to develop other types of regularization to improve source location. Investigating
713 the feasibility of such regularization schemes is beyond the scope of this paper.

714 4 Conclusions

715 We have developed an open-source, systematic, high-performance implementation of travel-
716 time computation, travelttime-based tomography, and travelttime-based source location based on
717 the eikonal equation and adjoint-state tomography theory for 2D/3D acoustic and elastic media.
718 Specially, to improve the fidelity and interpretability of inverted model parameters and source
719 parameters, we have developed a novel model parameter regularization scheme based on total gen-
720 eralized p -variation and P- and S-wave velocity structure similarity, as well as a source parameter
721 regularization scheme based on multitask machine learning models. We have demonstrated the
722 efficacy and accuracy of our methods and implementation using several synthetic data examples.
723 The results indicate that our implementation can serve as an adaptive computational framework for
724 travelttime computation, velocity tomography, and source location in 2D/3D acoustic and elastic
725 media.

726 Acknowledgments

727 The work was supported by Los Alamos National Laboratory (LANL) Laboratory Directory
728 Research and Development (LDRD) project 20240322ER. LANL is managed and operated by Triad
729 National Security, LLC for the U.S. Department of Energy (DOE) National Nuclear Security Admin-
730 istration (NNSA) under contract No. 89233218CNA000001. The research used high-performance
731 computing resources provided by LANL’s Institutional Computing (IC) program.

732 **Data Availability**

733 The codes (LATTE) implementing the methods developed in this work, as well as the parameter
734 files and scripts for reproducing the results, are open-source available at [github.com/lanl/
735 latte_traveltime](https://github.com/lanl/latte_traveltime). Faulted models in the examples are generated using our open-source
736 package RGM (Random Geological Model generation package) available at [github.com/lanl/
737 rgm](https://github.com/lanl/rgm). Datasets for training the multitask machine learning models are also generated using RGM.
738 Figures in this work are generated using our open-source plotting package pymplot (open-source
739 available at github.com/lanl/pymplot), visualization toolkit (VTK), and matplotlib.

740 **References**

- 741 Aki, K., and P. G. Richards, 2002, Quantitative Seismology: University Science Books, Sausalito,
742 California.
- 743 Anagaw, A., 2011, Full waveform inversion with total variation regularization: Proceedings of the
744 2011 CSPG CSEG CWLS Joint Annual Convention, 1–4.
- 745 Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, 2013, Regularized
746 seismic full waveform inversion with prior model information: Geophysics, **78**, no. 2, R25–R36,
747 doi: [10.1190/geo2012-0104.1](https://doi.org/10.1190/geo2012-0104.1).
- 748 Bretaudeau, F., R. Brossier, J. Virieux, and L. Métivier, 2014, First-arrival delayed tomography
749 using 1st and 2nd order adjoint-state method: SEG Technical Program Expanded Abstracts 2014,
750 4757–4762, doi: [10.1190/segam2014-0987.1](https://doi.org/10.1190/segam2014-0987.1).
- 751 Chambers, K., and J.-M. Kendall, 2008, A practical implementation of wave front construction for
752 3-D isotropic media: Geophysical Journal International, **173**, no. 3, 1030–1038, doi: [10.1111/j.1365-246X.2008.03790.x](https://doi.org/10.1111/j.1365-246X.2008.03790.x).
- 753 Chen, Y., Y. Chen, S. Fomel, A. Savvaidis, O. M. Saad, and Y. A. S. I. Oboué, 2023, Pyekfmm: A
754 Python package for 3D fast-marching-based travel-time calculation and its applications in seis-
755 mology: Seismological Research Letters, **94**, no. 4, 2050–2059, doi: [10.1785/0220230042](https://doi.org/10.1785/0220230042).
- 756 Dando, B. D. E., B. P. Goertz-Allmann, D. Kühn, N. Langet, A. M. Dichiarante, and V. Oye, 2021,
757 Relocating microseismicity from downhole monitoring of the Decatur CCS site using a modified
758 double-difference algorithm: Geophysical Journal International, **227**, no. 2, 1094–1122, doi:
759 [10.1093/gji/ggab255](https://doi.org/10.1093/gji/ggab255).
- 760 de Kool, M., N. Rawlinson, and M. Sambridge, 2006, A practical grid-based method for tracking
761 multiple refraction and reflection phases in three-dimensional heterogeneous media: Geophysical
762 Journal International, **167**, no. 1, 253–270, doi: [10.1111/j.1365-246X.2006.03078.x](https://doi.org/10.1111/j.1365-246X.2006.03078.x).

- 764 Detrixhe, M., F. Gibou, and C. Min, 2013, A parallel fast sweeping method for the eikonal equation:
 765 *Journal of Computational Physics*, **237**, 46–55, doi: [10.1016/j.jcp.2012.11.042](https://doi.org/10.1016/j.jcp.2012.11.042).
- 766 Esser, E., L. Guasch, F. J. Herrmann, and M. Warner, 2016, Constrained waveform inversion for
 767 automatic salt flooding: *The Leading Edge*, **35**, no. 3, 235–239, doi: [10.1190/tle35030235.1](https://doi.org/10.1190/tle35030235.1).
- 768 1.
- 769 Fang, H., R. D. van der Hilst, M. V. de Hoop, K. Kothari, S. Gupta, and I. Dokmanić, 2019, Parsi-
 770 monious Seismic Tomography with Poisson Voronoi Projections: Methodology and Validation:
 771 *Seismological Research Letters*, **91**, no. 1, 343–355, doi: [10.1785/0220190141](https://doi.org/10.1785/0220190141).
- 772 Fichtner, A., H.-P. Bunge, H. Igel, A. Fichtner, H.-P. Bunge, and H. Igel, 2006, The adjoint method
 773 in seismology: I. Theory: *Physics of the Earth and Planetary Interiors*, **157**, no. 1, 86 – 104, doi:
 774 [10.1016/j.pepi.2006.03.016](https://doi.org/10.1016/j.pepi.2006.03.016).
- 775 Fomel, S., S. Luo, and H. Zhao, 2009, Fast sweeping method for the factored eikonal equation:
 776 *Journal of Computational Physics*, **228**, no. 17, 6440–6455, doi: [10.1016/j.jcp.2009.05.029](https://doi.org/10.1016/j.jcp.2009.05.029).
- 777 029.
- 778 Gao, K., 2024, Iterative multitask learning and inference from seismic images: *Geophysical Journal
 779 International*, **236**, no. 1, 565–592, doi: [10.1093/gji/ggad424](https://doi.org/10.1093/gji/ggad424).
- 780 Gao, K., and T. Chen, 2024, FLIT: A generic Fortran library based on interfaces and templates:
 781 GitHub Open-source Repository, doi: github.com/lanl/flit.
- 782 Gao, K., C. Donahue, B. G. Henderson, and R. T. Modrak, 2022, Deep-learning-guided high-
 783 resolution subsurface reflectivity imaging with application to ground-penetrating radar data:
 784 *Geophysical Journal International*, **233**, no. 1, 448–471, doi: [10.1093/gji/ggac468](https://doi.org/10.1093/gji/ggac468).
- 785 Gao, K., and L. Huang, 2019, Acoustic- and elastic-waveform inversion with total generalized
 786 p-variation regularization: *Geophysical Journal International*, **218**, no. 2, 933–957, doi: [10.1093/gji/ggz203](https://doi.org/10.1093/gji/ggz203).
- 787
- 788 Gauthier, O., J. Virieux, and A. Tarantola, 1986, Two-dimensional nonlinear inversion of seismic
 789 waveforms: Numerical results: *Geophysics*, **51**, no. 7, 1387–1403, doi: [10.1190/1.1442188](https://doi.org/10.1190/1.1442188).
- 790 Gibson, R. L., V. Durussel, and K.-J. Lee, 2005, Modeling and velocity analysis with a wavefront-
 791 construction algorithm for anisotropic media: *Geophysics*, **70**, no. 4, T63–T74, doi: [10.1190/1.1988188](https://doi.org/10.1190/1.1988188).
- 792
- 793 Goldstein, T., and S. Osher, 2009, The split Bregman method for L1-regularized problems: *SIAM
 794 Journal on Imaging Sciences*, **2**, no. 2, 323–343, doi: [10.1137/080725891](https://doi.org/10.1137/080725891).
- 795 Grechka, V. Y., and G. A. McMechan, 1996, 3-D two-point ray tracing for heterogeneous, weakly
 796 transversely isotropic media: *Geophysics*, **61**, no. 6, 1883–1894, doi: [10.1190/1.1444103](https://doi.org/10.1190/1.1444103).
- 797 Guitton, A., 2012, Blocky regularization schemes for full-waveform inversion: *Geophysical
 798 Prospecting*, **60**, no. 5, 870–884, doi: [10.1111/j.1365-2478.2012.01025.x](https://doi.org/10.1111/j.1365-2478.2012.01025.x).
- 799 Guo, H., and C. Thurber, 2021, Double-difference seismic attenuation tomography method and its

- application to The Geysers geothermal field, California: *Geophysical Journal International*, **225**, no. 2, 926–949, doi: [10.1093/gji/ggab017](https://doi.org/10.1093/gji/ggab017).
- Huang, J.-W., and G. Bellefleur, 2012, Joint transmission and reflection traveltime tomography using the fast sweeping method and the adjoint-state technique: *Geophysical Journal International*, **188**, no. 2, 570–582, doi: [10.1111/j.1365-246X.2011.05273.x](https://doi.org/10.1111/j.1365-246X.2011.05273.x).
- Kao, C. Y., S. Osher, and J. Qian, 2004, Lax-Friedrichs sweeping scheme for static Hamilton-Jacobi equations: *Journal of Computational Physics*, **196**, no. 1, 367 – 391, doi: <https://doi.org/10.1016/j.jcp.2003.11.007>.
- Kao, C.-Y., S. Osher, and Y.-H. Tsai, 2005, Fast Sweeping Methods for Static Hamilton–Jacobi Equations: *SIAM Journal on Numerical Analysis*, **42**, no. 6, 2612–2632, doi: [10.1137/S0036142902419600](https://doi.org/10.1137/S0036142902419600).
- Kim, S., 2002, 3-D eikonal solvers: First-arrival traveltimes: *Geophysics*, **67**, no. 4, 1225–1231, doi: [10.1190/1.1500384](https://doi.org/10.1190/1.1500384).
- Kim, S., and R. Cook, 1999, 3-d travelttime computation using second-order eno scheme: *Geophysics*, **64**, no. 6, 1867–1876, doi: [10.1190/1.1444693](https://doi.org/10.1190/1.1444693).
- Knoll, F., K. Bredies, T. Pock, and R. Stollberger, 2011, Second order total generalized variation (TGV) for MRI: *Magnetic Resonance in Medicine*, **65**, no. 2, 480–491, doi: [10.1002/mrm.22595](https://doi.org/10.1002/mrm.22595).
- Koehn, D., and D. De Nil, 2022, RAJZEL – 2D parallel first-arrival travelttime modelling and inversion code using an Eikonal solver and the adjoint state method: GitHub Repository, doi: github.com/daniel-koehn/RAJZEL.
- Korenaga, J., W. S. Holbrook, G. M. Kent, P. B. Kelemen, R. S. Detrick, H. Larsen, J. R. Hopper, and T. Dahl-Jensen, 2000, Crustal structure of the southeast Greenland margin from joint refraction and reflection seismic tomography: *Journal of Geophysical Research: Solid Earth*, **105**, no. B9, 21591–21614, doi: [10.1029/2000JB900188](https://doi.org/10.1029/2000JB900188).
- Lambaré, G., P. S. Lucio, and A. Hanyga, 1996, Two-dimensional multivalued travelttime and amplitude maps by uniform sampling of a ray field: *Geophysical Journal International*, **125**, no. 2, 584–598, doi: [10.1111/j.1365-246X.1996.tb00021.x](https://doi.org/10.1111/j.1365-246X.1996.tb00021.x).
- Le Bouteiller, P., M. Benjemaa, L. Métivier, and J. Virieux, 2019, A discontinuous Galerkin fast-sweeping eikonal solver for fast and accurate travelttime computation in 3D tilted anisotropic media: *Geophysics*, **84**, no. 2, C107–C118, doi: [10.1190/geo2018-0555.1](https://doi.org/10.1190/geo2018-0555.1).
- Leung, S., and J. Qian, 2006, An adjoint state method for three-dimensional transmission travelttime tomography using first-arrivals: *Communications in Mathematical Sciences*, **4**, no. 1, 249–266.
- Li, Y., L. Dong, and Y. Liu, 2017, First-arrival travelttime tomography based on a new preconditioned adjoint-state method: *Chinese Journal of Geophysics*, **228**, no. 10, 3934–3941, doi: [10.6038/cjg20171021](https://doi.org/10.6038/cjg20171021).

- 836 Lin, Y., and L. Huang, 2014, Acoustic- and elastic-waveform inversion using a modified total-
837 variation regularization scheme: *Geophysical Journal International*, **200**, no. 1, 489–502, doi:
838 [10.1093/gji/ggu393](https://doi.org/10.1093/gji/ggu393).
- 839 Lin, Y., E. M. Syracuse, M. Maceira, H. Zhang, and C. Larmat, 2015, Double-difference traveltime
840 tomography with edge-preserving regularization and a priori interfaces: *Geophysical Journal
841 International*, **201**, no. 2, 574–594, doi: [10.1093/gji/ggv047](https://doi.org/10.1093/gji/ggv047).
- 842 Liu, Q., and J. Tromp, 2006, Finite-frequency kernels based on adjoint methods: *Bulletin of the
843 Seismological Society of America*, **96**, no. 6, 2383–2397, doi: [10.1785/0120060041](https://doi.org/10.1785/0120060041).
- 844 Liu, Y., L. Dong, Y. Wang, J. Zhu, and Z. Ma, 2009, Sensitivity kernels for seismic Fresnel volume
845 tomography: *Geophysics*, **74**, no. 5, U35–U46, doi: [10.1190/1.3169600](https://doi.org/10.1190/1.3169600).
- 846 Luo, S., and J. Qian, 2011, Factored singularities and high-order Lax-Friedrichs sweeping schemes
847 for point-source traveltimes and amplitudes: *Journal of Computational Physics*, **230**, no. 12,
848 4742–4755, doi: [10.1016/j.jcp.2011.02.043](https://doi.org/10.1016/j.jcp.2011.02.043).
- 849 ———, 2012, Fast sweeping methods for factored anisotropic eikonal equations: Multiplicative
850 and additive factors: *Journal of Scientific Computing*, **52**, no. 2, 360–382, doi: [10.1007/
851 s10915-011-9550-y](https://doi.org/10.1007/s10915-011-9550-y).
- 852 Luo, S., J. Qian, and H. Zhao, 2012, Higher-order schemes for 3D first-arrival traveltimes and
853 amplitudes: *Geophysics*, **77**, no. 2, T47–T56, doi: [10.1190/geo2010-0363.1](https://doi.org/10.1190/geo2010-0363.1).
- 854 Luo, Y., Y. Ma, Y. Wu, H. Liu, and L. Cao, 2016, Full-travelttime inversion: *Geophysics*, **81**, no. 5,
855 R261–R274, doi: [10.1190/geo2015-0353.1](https://doi.org/10.1190/geo2015-0353.1).
- 856 Luo, Y., and G. T. Schuster, 1991, Wave-equation travelttime inversion: *Geophysics*, **56**, no. 5,
857 645–653, doi: [10.1190/1.1443081](https://doi.org/10.1190/1.1443081).
- 858 Marquering, H., F. Dahlen, and G. Nolet, 1999, Three-dimensional sensitivity kernels for finite-
859 frequency traveltimes: the banana–doughnut paradox: *Geophysical Journal International*, **137**,
860 no. 3, 805–815, doi: [10.1046/j.1365-246x.1999.00837.x](https://doi.org/10.1046/j.1365-246x.1999.00837.x).
- 861 Meléndez, A., J. Korenaga, V. Sallarès, A. Miniussi, and C. Ranero, 2015, TOMO3D: 3-D joint
862 refraction and reflection travelttime tomography parallel code for active-source seismic data—
863 synthetic test: *Geophysical Journal International*, **203**, no. 1, 158–174, doi: [10.1093/gji/ggv292](https://doi.org/10.1093/gji/ggv292).
- 865 Michelen, R. J., and J. M. Harris, 1991, Tomographic travelttime inversion using natural pixels:
866 *Geophysics*, **56**, no. 5, 635–644, doi: [10.1190/1.1443080](https://doi.org/10.1190/1.1443080).
- 867 Mora, P., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: *Geophysics*,
868 **52**, no. 9, 1211–1228, doi: [10.1190/1.1442384](https://doi.org/10.1190/1.1442384).
- 869 Nocedal, J., and S. J. Wright, 2006, Numerical Optimization (2nd ed.): Springer, New York.
- 870 Pereyra, V., W. H. K. Lee, and H. B. Keller, 1980, Solving two-point seismic-ray tracing problems
871 in a heterogeneous medium: Part 1. A general adaptive finite difference method: *Bulletin of the*

- 872 Seismological Society of America, **70**, no. 1, 79–99.
- 873 Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional
874 with geophysical applications: *Geophysical Journal International*, **167**, no. 2, 495–503, doi:
875 [10.1111/j.1365-246X.2006.02978.x](https://doi.org/10.1111/j.1365-246X.2006.02978.x).
- 876 Podvin, P., and I. Lecomte, 1991, Finite difference computation of traveltimes in very contrasted
877 velocity models: a massively parallel approach and its associated tools: *Geophysical Journal International*
878 **105**, no. 1, 271–284, doi: [10.1111/j.1365-246X.1991.tb03461.x](https://doi.org/10.1111/j.1365-246X.1991.tb03461.x).
- 879 Pyun, S., C. Shin, D.-J. Min, and T. Ha, 2005, Refraction travelttime tomography using damped
880 monochromatic wavefield: *Geophysics*, **70**, no. 2, U1–U7, doi: [10.1190/1.1884829](https://doi.org/10.1190/1.1884829).
- 881 Qian, J., and W. W. Symes, 2002, Finite-difference quasi-P traveltimes for anisotropic media:
882 *Geophysics*, **67**, no. 1, 147–155, doi: [10.1190/1.1451438](https://doi.org/10.1190/1.1451438).
- 883 Qian, J., Y.-T. Zhang, and H.-K. Zhao, 2007, Fast sweeping methods for eikonal equations on
884 triangular meshes: *SIAM J. Numer. Anal.*, **45**, no. 1, 83–107, doi: [10.1137/050627083](https://doi.org/10.1137/050627083).
- 885 Qin, F., Y. Luo, K. B. Olsen, W. Cai, and G. T. Schuster, 1992, Finite-difference solution of the
886 eikonal equation along expanding wavefronts: *Geophysics*, **57**, no. 3, 478–487, doi: [10.1190/1.1443263](https://doi.org/10.1190/1.1443263).
- 888 Rawlinson, N., M. d. Kool, and M. Sambridge, 2006, Seismic wavefront tracking in 3D hetero-
889 geneous media: Applications with multiple data classes: *Exploration Geophysics*, **37**, no. 4,
890 322–330, doi: [10.1071/EG06322](https://doi.org/10.1071/EG06322).
- 891 Rawlinson, N., S. Pozgay, and S. Fishwick, 2010, Seismic tomography: A window into deep Earth:
892 *Physics of the Earth and Planetary Interiors*, **178**, no. 3, 101–135, doi: [10.1016/j.pepi.2009.10.002](https://doi.org/10.1016/j.pepi.2009.10.002).
- 894 Rawlinson, N., and M. Sambridge, 2004, Wave front evolution in strongly heterogeneous layered
895 media using the fast marching method: *Geophysical Journal International*, **156**, no. 3, 631–647,
896 doi: [10.1111/j.1365-246X.2004.02153.x](https://doi.org/10.1111/j.1365-246X.2004.02153.x).
- 897 Rudin, L. I., S. Osher, and E. Fatemi, 1992, Nonlinear total variation based noise removal algorithms:
898 *Physica D: Nonlinear Phenomena*, **60**, no. 1, 259 – 268, doi: [10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F).
- 900 Sadeghi, H., S. Suzuki, and H. Takenaka, 1999, A two-point, three-dimensional seismic ray tracing
901 using genetic algorithms: *Physics of the Earth and Planetary Interiors*, **113**, no. 1, 355 – 365, doi:
902 [10.1016/S0031-9201\(99\)00011-4](https://doi.org/10.1016/S0031-9201(99)00011-4).
- 903 Schuster, G. T., and A. Quintus-Bosz, 1993, Wavepath eikonal travelttime inversion: Theory:
904 *Geophysics*, **58**, no. 9, 1314–1323, doi: [10.1190/1.1443514](https://doi.org/10.1190/1.1443514).
- 905 Sei, A., and W. W. Symes, 1994, Gradient calculation of the travelttime cost function without ray trac-
906 ing: *SEG Technical Program Expanded Abstracts*, 1351–1354, doi: [10.1190/1.1822780](https://doi.org/10.1190/1.1822780).
- 907 Sethian, J. A., and A. M. Popovici, 1999, 3-D travelttime computation using the fast marching

- method: Geophysics, **64**, no. 2, 516–523, doi: [10.1190/1.1444558](https://doi.org/10.1190/1.1444558).
- Snieder, R., and A. Lomax, 1996, Wavefield smoothing and the effect of rough velocity perturbations on arrival times and amplitudes: Geophysical Journal International, **125**, no. 3, 796–812, doi: [10.1111/j.1365-246X.1996.tb06024.x](https://doi.org/10.1111/j.1365-246X.1996.tb06024.x).
- Spetzler, J., and R. Snieder, 2004, The Fresnel volume and transmitted waves: Geophysics, **69**, no. 3, 653–663, doi: [10.1190/1.1759451](https://doi.org/10.1190/1.1759451).
- Taillardier, C., M. Noble, H. Chauris, and H. Calandra, 2009, First-arrival traveltome tomography based on the adjoint-state method: Geophysics, **74**, no. 6, WCB1–WCB10, doi: [10.1190/1.3250266](https://doi.org/10.1190/1.3250266).
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, no. 8, 1259–1266, doi: [10.1190/1.1441754](https://doi.org/10.1190/1.1441754).
- Thurber, C. H., 1983, Earthquake locations and three-dimensional crustal structure in the Coyote Lake Area, central California: Journal of Geophysical Research: Solid Earth, **88**, no. B10, 8226–8236, doi: [10.1029/JB088iB10p08226](https://doi.org/10.1029/JB088iB10p08226).
- Tikhonov, A., A. Goncharsky, V. Stepanov, and A. Yagola, 1995, Numerical methods for the solution of ill-posed problems: Springer.
- Tong, P., 2021a, Adjoint-state traveltome tomography: Eikonal equation-based methods and application to the Anza Area in Southern California: Journal of Geophysical Research: Solid Earth, **126**, no. 5, e2021JB021818, doi: [10.1029/2021JB021818](https://doi.org/10.1029/2021JB021818).
- , 2021b, Adjoint-state traveltome tomography software package (Training version): GitHub Repository, doi: doi.org/10.21979/N9/YPGQRF.
- Tong, P., T. Li, J. Chen, and M. Nagaso, 2024, Adjoint-state differential arrival time tomography: Geophysical Journal International, **236**, no. 1, 139–160, doi: [10.1093/gji/ggad416](https://doi.org/10.1093/gji/ggad416).
- Tong, P., D. Yang, Q. Liu, X. Yang, and J. Harris, 2016, Acoustic wave-equation-based earthquake location: Geophysical Journal International, **205**, no. 1, 464–478, doi: [10.1093/gji/ggw026](https://doi.org/10.1093/gji/ggw026).
- Tsai, Y., L. Cheng, S. Osher, and H. Zhao, 2003, Fast sweeping algorithms for a class of Hamilton-Jacobi equations: SIAM Journal on Numerical Analysis, **41**, no. 2, 673–694, doi: [10.1137/S0036142901396533](https://doi.org/10.1137/S0036142901396533).
- Vasco, D. W., J. John E. Peterson, and E. L. Majer, 1995, Beyond ray tomography: Wavepaths and Fresnel volumes: Geophysics, **60**, no. 6, 1790–1804, doi: [10.1190/1.1443912](https://doi.org/10.1190/1.1443912).
- Vidale, J., 1988, Finite-difference calculation of travel times: Bulletin of the Seismological Society of America, **78**, no. 6, 2062–2076.
- Vinje, V., E. Iversen, and H. Gjøystdal, 1993, Traveltome and amplitude estimation using wavefront construction: Geophysics, **58**, no. 8, 1157–1166, doi: [10.1190/1.1443499](https://doi.org/10.1190/1.1443499).
- 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](https://doi.org/10.1190/1.3238367).

- 944 Waheed, U. B., and T. Alkhalifah, 2017, A fast sweeping algorithm for accurate solution of the tilted
945 transversely isotropic eikonal equation using factorization: *Geophysics*, **82**, no. 6, WB1–WB8,
946 doi: [10.1190/geo2016-0712.1](https://doi.org/10.1190/geo2016-0712.1).
- 947 Waheed, U. B., T. Alkhalifah, and H. Wang, 2015a, Efficient traveltime solutions of the acoustic
948 TI eikonal equation: *Journal of Computational Physics*, **282**, 62 – 76, doi: [10.1016/j.jcp.2014.11.006](https://doi.org/10.1016/j.jcp.2014.11.006).
- 950 Waheed, U. B., C. E. Yarman, and G. Flagg, 2015b, An iterative, fast-sweeping-based eikonal
951 solver for 3D tilted anisotropic media: *Geophysics*, **80**, no. 3, C49–C58, doi: [10.1190/geo2014-0375.1](https://doi.org/10.1190/geo2014-0375.1).
- 953 Waldhauser, F., and W. L. Ellsworth, 2000, A double-difference earthquake location algorithm:
954 Method and application to the Northern Hayward Fault, California: *Bulletin of the Seismological
955 Society of America*, **90**, no. 6, 1353–1368, doi: [10.1785/0120000006](https://doi.org/10.1785/0120000006).
- 956 Wang, Y., T. Nemeth, and R. T. Langan, 2006, An expanding-wavefront method for solving
957 the eikonal equations in general anisotropic media: *Geophysics*, **71**, no. 5, T129–T135, doi:
958 [10.1190/1.2235563](https://doi.org/10.1190/1.2235563).
- 959 White, M. C. A., H. Fang, N. Nakata, and Y. Ben-Zion, 2020, PyKonal: A Python package for
960 solving the eikonal equation in spherical and Cartesian coordinates using the fast marching
961 method: *Seismological Research Letters*, **91**, no. 4, 2378–2389, doi: [10.1785/0220190318](https://doi.org/10.1785/0220190318).
- 962 Woodward, M. J., 1992, Wave-equation tomography: *Geophysics*, **57**, no. 1, 15–26, doi: [10.1190/1.1443179](https://doi.org/10.1190/1.1443179).
- 963 Wu, R., and M. N. Toksöz, 1987, Diffraction tomography and multisource holography applied to
964 seismic imaging: *Geophysics*, **52**, no. 1, 11–25, doi: [10.1190/1.1442237](https://doi.org/10.1190/1.1442237).
- 966 Wu, X., and Z. Guo, 2018, Detecting faults and channels while enhancing seismic structural and
967 stratigraphic features: *Interpretation*, **7**, no. 1, T155–T166, doi: [10.1190/INT-2017-0174.1](https://doi.org/10.1190/INT-2017-0174.1).
- 969 Xu, S., Y. Zhang, and T. Huang, 2006, Enhanced tomography resolution by a fat ray technique:
970 SEG Technical Program Expanded Abstracts 2006, 3354–3358, doi: [10.1190/1.2370229](https://doi.org/10.1190/1.2370229).
- 971 Yomogida, K., 1992, Fresnel zone inversion for lateral heterogeneities in the earth: *Pure and Applied
972 Geophysics*, **138**, no. 3, 391–406, doi: [10.1007/BF00876879](https://doi.org/10.1007/BF00876879).
- 973 Yuan, Y. O., F. J. Simons, and J. Tromp, 2016, Double-difference adjoint seismic tomography:
974 *Geophysical Journal International*, **206**, no. 3, 1599–1618, doi: [10.1093/gji/ggw233](https://doi.org/10.1093/gji/ggw233).
- 975 Zelt, C. A., and P. J. Barton, 1998, Three-dimensional seismic refraction tomography: A comparison
976 of two methods applied to data from the Faeroe Basin: *Journal of Geophysical Research: Solid
977 Earth*, **103**, no. B4, 7187–7210, doi: [10.1029/97JB03536](https://doi.org/10.1029/97JB03536).
- 978 Zelt, C. A., and J. Chen, 2016, Frequency-dependent traveltime tomography for near-surface seismic
979 refraction data: *Geophysical Journal International*, **207**, no. 1, 72–88, doi: [10.1093/gji/ggv103](https://doi.org/10.1093/gji/ggv103).

- 980 ggw269.
- 981 Zeng, X., C. H. Thurber, D. R. Shelly, R. M. Harrington, E. S. Cochran, N. L. Bennington,
 982 D. Peterson, B. Guo, and K. McClement, 2016, 3-D P- and S-wave velocity structure and
 983 low-frequency earthquake locations in the Parkfield, California region: Geophysical Journal
 984 International, **206**, no. 3, 1574–1585, doi: [10.1093/gji/ggw217](https://doi.org/10.1093/gji/ggw217).
- 985 Zhang, H., and C. Thurber, 2006, Development and applications of double-difference seis-
 986 mic tomography: pure and applied geophysics, **163**, no. 2, 373–403, doi: [10.1007/s00024-005-0021-y](https://doi.org/10.1007/s00024-005-0021-y).
- 988 Zhang, H., and C. H. Thurber, 2003, Double-difference tomography: The method and its application
 989 to the Hayward Fault, California: Bulletin of the Seismological Society of America, **93**, no. 5,
 990 1875–1889, doi: [10.1785/0120020190](https://doi.org/10.1785/0120020190).
- 991 Zhang, J., L. Dong, J. Wang, and C. Huang, 2023, Preconditioned transmission + reflection
 992 joint traveltimes tomography with adjoint-state method for subsurface velocity model building:
 993 Geophysical Prospecting, **71**, no. 2, 171–190, doi: [10.1111/1365-2478.13287](https://doi.org/10.1111/1365-2478.13287).
- 994 Zhang, J., U. S. ten Brink, and M. N. Toksöz, 1998, Nonlinear refraction and reflection travel time
 995 tomography: Journal of Geophysical Research: Solid Earth, **103**, no. B12, 29743–29757, doi:
 996 [10.1029/98JB01981](https://doi.org/10.1029/98JB01981).
- 997 Zhang, J., and M. N. Toksöz, 1998, Nonlinear refraction traveltimes tomography: Geophysics, **63**,
 998 no. 5, 1726–1737, doi: [10.1190/1.1444468](https://doi.org/10.1190/1.1444468).
- 999 Zhang, Y.-T., H.-K. Zhao, and J. Qian, 2006, High order fast sweeping methods for static
 1000 Hamilton-Jacobi equations: Journal of Scientific Computing, **29**, no. 1, 25–56, doi: [10.1007/s10915-005-9014-3](https://doi.org/10.1007/s10915-005-9014-3).
- 1002 Zhao, H., 2004, A fast sweeping method for eikonal equations: Mathematics of Computation, **74**,
 1003 no. 250, 603–627, doi: [10.1090/S0025-5718-04-01678-3](https://doi.org/10.1090/S0025-5718-04-01678-3).
- 1004 Zhdanov, M., 2002, Geophysical inverse theory and regularization problems, 1 ed.: Elsevier Science.

1005 **Appendix A: Parallel fast sweeping algorithms for the eikonal
 1006 and adjoint-state equations**

1007 We slightly modify the parallel fast sweeping algorithm presented in Detrixhe et al. (2013). For
 1008 the purpose of completeness, we detail our algorithms as below.

1009 Fast sweeping relies on sweeping all possible orders of dimension directions. For 2D, the
 1010 possible ordering of directions are

$$I = 1 : N_x, \quad K = 1 : N_z, \quad (56a)$$

$$I = 1 : N_x, \quad K = N_z : 1, \quad (56b)$$

$$I = N_x : 1, \quad K = 1 : N_z, \quad (56c)$$

$$I = N_x : 1, \quad K = N_z : 1. \quad (56d)$$

1011 We denote each of the orderings as $S(x_0, x_1, z_0, z_1)$, where we use subscripts 0 and 1 to represent
 1012 the starting and end elements, respectively. For example, for the second ordering, $(x_0, x_1, z_0, z_1) =$
 1013 $(1, N_x, N_z, 1)$.

1014 Then we implement the parallel fast sweeping in 2D for a specific ordering $S(x_0, x_1, z_0, z_1)$
 1015 with Algorithm 1.

Algorithm 1: Algorithm for 2D parallel fast sweeping adopted and modified from [Detrixhe et al. \(2013\)](#) for LATTE.

Input: Velocity model, traveltimes field τ or adjoint-state field λ

Parameters : Ordering $S(x_0, x_1, z_0, z_1)$, dimensions of the model N_x and N_z , and grid spacings d_x and d_z

for $1 \leq l \leq N_x + N_z - 1$ **do**

Compute the starting and ending element indices for z dimension as:

$$j_a = \begin{cases} j_0, & l \leq N_x \\ j_0 + (l - N_x) \times c_j, & \text{otherwise.} \end{cases} \quad (57)$$

$$j_b = \begin{cases} j_0 + (l - 1) \times c_j, & l \leq N_z \\ j_1, & \text{otherwise,} \end{cases} \quad (58)$$

where

$$c_j = \begin{cases} 1, & \text{if } j_0 \leq j_1, \\ -1, & \text{otherwise.} \end{cases} \quad (59)$$

for parallel $j_a \leq j \leq j_b$ **do**

(1) Compute x index i from

$$|i - i_0| + |j - j_0| = l - 1. \quad (60)$$

(2) Update the multiplicative traveltimes field $\tau_{i,j}$ or the adjoint-state field $\lambda_{i,j}$.

end

end

Output: τ or λ

1016 For 3D, the possible ordering of directions are

$$I = 1 : N_x, \quad J = 1 : N_y, \quad K = 1 : N_z, \quad (61a)$$

$$I = 1 : N_x, \quad J = 1 : N_y, \quad K = N_z : 1, \quad (61b)$$

$$I = 1 : N_x, \quad J = N_y : 1, \quad K = 1 : N_z, \quad (61c)$$

$$I = 1 : N_x, \quad J = N_y : 1, \quad K = N_z : 1, \quad (61d)$$

$$I = N_x : 1, \quad J = 1 : N_y, \quad K = 1 : N_z, \quad (61e)$$

$$I = N_x : 1, \quad J = 1 : N_y, \quad K = N_z : 1, \quad (61f)$$

$$I = N_x : 1, \quad J = N_y : 1, \quad K = 1 : N_z, \quad (61g)$$

$$I = N_x : 1, \quad J = N_y : 1, \quad K = N_z : 1. \quad (61h)$$

1017 We denote each of the orderings as $S(x_0, x_1, y_0, y_1, z_0, z_1)$, where we use subscripts 0 and 1
 1018 to represent the starting and end elements, respectively. For example, for the second ordering,
 1019 $(x_0, x_1, y_0, y_1, z_0, z_1) = (1, N_x, 1, N_y, N_z, 1)$.

1020 The algorithm for 3D parallel fast sweeping is not straightforwardly available from [Detrixhe et al. \(2013\)](#). Therefore, here we provide a complete algorithm for achieving parallel fast sweeping
 1021 with an arbitrary number of threads for 3D eikonal and adjoint-state equations. We implement 3D
 1022 parallel fast sweeping for a specific ordering $S(x_0, x_1, y_0, y_1, z_0, z_1)$ with [Algorithm 2](#).

1024 Then we go to the next ordering and repeat the procedure until all orderings are computed. We
 1025 repeat the entire procedure (fast sweeping of 4 orderings in 2D and 8 orderings in 3D) until the
 1026 threshold of field difference is reached. Here we ignore the outer loop algorithm as the details have
 1027 been described by a number of existing works (e.g., [Zhao, 2004](#); [Taillandier et al., 2009](#); [Detrixhe et al., 2013](#)). The algorithm has the same computational complexity with serial fast sweeping yet
 1028 can be accelerated with OpenMP shared-memory parallelism.

1030 Appendix B: The adjoint-state equation for arbitrary receiver 1031 location

1032 In the original works of adjoint-state FATT by [Leung and Qian \(2006\)](#) and later by [Taillandier et al. \(2009\)](#), the authors developed the formulation for adjoint-state equation. However, in both
 1033 works, solving the adjoint-state equation requires the determination of the adjoint-state variable λ
 1034 on the boundaries through
 1035

$$\lambda(\mathbf{x}_r) \nabla t(\mathbf{x}_r) \cdot \mathbf{n}(\mathbf{x}_r) = \Delta T(\mathbf{x}_r), \quad (69)$$

Algorithm 2: Algorithm for 3D parallel fast sweeping adopted and modified from Detrixhe et al. (2013) for LATTE.

Input: Velocity model, traveltime field τ or adjoint-state field λ

Parameters : Ordering $S(x_0, x_1, y_0, y_1, z_0, z_1)$, dimensions of the model N_x , N_y , and N_z , and grid spacings d_x , d_y , and d_z

for $1 \leq l \leq N_x + N_y + N_z - 2$ **do**

Compute the starting and ending element indices for z and y dimensions as:

$$k_a = \begin{cases} k_0, & l \leq N_x + N_y \\ k_0 + (l - (N_x + N_y - 1)) \times c_k, & \text{otherwise.} \end{cases} \quad (62)$$

$$k_b = \begin{cases} k_0 + (l - 1) \times c_k, & l \leq N_z \\ k_1, & \text{otherwise.} \end{cases} \quad (63)$$

$$j_a = \begin{cases} j_0, & l \leq N_x + N_y \\ j_0 + (l - (N_x + N_y - 1)) \times c_j, & \text{otherwise.} \end{cases} \quad (64)$$

$$j_b = \begin{cases} j_0 + (l - 1) \times c_j, & l \leq N_y \\ j_1, & \text{otherwise,} \end{cases} \quad (65)$$

where

$$c_k = \begin{cases} 1, & \text{if } k_0 \leq k_1, \\ -1, & \text{otherwise,} \end{cases}, \quad (66)$$

$$c_j = \begin{cases} 1, & \text{if } j_0 \leq j_1, \\ -1, & \text{otherwise.} \end{cases} \quad (67)$$

for parallel $k_a \leq k \leq k_b$, $j_a \leq j \leq j_b$ **do**

(1) Compute x index i from

$$|i - i_0| + |j - j_0| + |k - k_0| = l - 1. \quad (68)$$

(2) If $i < 1$ or $i > N_x$, then skip Step (3).

(3) Update the multiplicative traveltime field $\tau_{i,j,k}$ or the adjoint-state field $\lambda_{i,j,k}$.

end

end

Output: τ or λ

1036 where \mathbf{n} is the normal to the surface (or boundary) of the model, $\partial\Omega$. This condition introduces
 1037 nontrivial restriction on the applicability of adjoint-state FATT to arbitrary source-receiver geometry
 1038 in a rigorous sense. For instance, rigorously, adjoint-state FATT does not apply to the scenario

1039 where the receivers are placed in a well or below the ground surface.

1040 We argue that such a restriction is not necessary. The emergence of this condition is in fact
 1041 caused by the assumption that the receivers are placed on the surface $\partial\Omega$. Rather than defining the
 1042 misfit function using the surface integral of the traveltime misfit on $\partial\Omega$, we assume the following
 1043 constrained L_2 -norm optimization problem:

$$\mathcal{J}(m) = \min_m \frac{1}{2} \int_{\Omega} [t(m, \mathbf{x}) - T(\mathbf{x})]^2 \delta(\mathbf{x} - \mathbf{x}_r) d\mathbf{x}, \quad \text{s.t. } m^2 |\nabla t|^2 = 1, \quad (70)$$

1044 where the receivers can be in arbitrary location in Ω .

1045 Defining the augmented Lagrangian,

$$\mathcal{L}(m, t, \lambda) = \frac{1}{2} \int_{\Omega} (t - T)^2 \delta(\mathbf{x} - \mathbf{x}_r) d\mathbf{x} + \frac{1}{2} \int_{\Omega} \lambda \left(|\nabla t|^2 - \frac{1}{m^2} \right) d\mathbf{x}, \quad (71)$$

1046 the first-order optimality conditions ([Nocedal and Wright, 2006](#)) of which read

$$\frac{\partial \mathcal{L}}{\partial m} = \frac{\partial \mathcal{J}}{\partial m} + \int_{\Omega} \frac{\lambda}{m^3} d\mathbf{x} = 0, \quad (72)$$

$$\frac{\partial \mathcal{L}}{\partial t} = 0, \quad (73)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = |\nabla t|^2 - \frac{1}{m^2} = 0. \quad (74)$$

1047 The second equation, $\partial \mathcal{L} / \partial t = 0$, gives

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial t} &= \int_{\Omega} (t - T) \delta(\mathbf{x} - \mathbf{x}_r) d\mathbf{x} + \int_{\Omega} \left(\nabla t \cdot \frac{\partial \nabla t}{\partial t} \right) \lambda d\mathbf{x} \\ &= \int_{\Omega} (t - T) \delta(\mathbf{x} - \mathbf{x}_r) d\mathbf{x} + \int_{\Omega} \left[\nabla t \cdot \frac{\partial}{\partial t} \left(\frac{\partial t}{\partial \mathbf{x}} \right) \right] \lambda d\mathbf{x} \\ &= \int_{\Omega} (t - T) \delta(\mathbf{x} - \mathbf{x}_r) d\mathbf{x} + \int_{\Omega} \nabla t \cdot \nabla \lambda d\mathbf{x}. \end{aligned} \quad (75)$$

1048 If we assume $\lambda = 0$ on $\partial\Omega$, then we can add an arbitrary boundary integral term of λ to
 1049 equation (75). By adding $\int_{\partial\Omega} \lambda (\mathbf{n} \cdot \nabla t) ds$ where \mathbf{n} is the normal vector of $\partial\Omega$ and using integration
 1050 by parts, we have

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial t} &= \int_{\Omega} (t - T) \delta(\mathbf{x} - \mathbf{x}_r) d\mathbf{x} + \int_{\Omega} \nabla t \cdot \nabla \lambda d\mathbf{x} - \int_{\partial\Omega} \lambda \mathbf{n} \cdot \nabla t ds, \\ &= \int_{\Omega} (t - T) \delta(\mathbf{x} - \mathbf{x}_r) d\mathbf{x} - \int_{\Omega} \nabla \cdot (\lambda \nabla t) d\mathbf{x} \\ &= 0, \end{aligned} \quad (76)$$

1051 which indicates that, for an arbitrary t and traveltime difference $(t - T)\delta(\mathbf{x} - \mathbf{x}_r)$, the following
 1052 adjoint-state equation must be satisfied:

$$\nabla \cdot (\lambda \nabla t) = (t - T)\delta(\mathbf{x} - \mathbf{x}_r). \quad (77)$$

1053 This is developed in the augmented Lagrangian functional framework as in [Leung and Qian](#)
 1054 ([2006](#)) and [Taillandier et al. \(2009\)](#), yet is consistent with the results obtained based on a perturbation
 1055 approach ([Tong, 2021a](#)).

1056 In practice, \mathbf{x}_r can contain multiple nonzero values (multiple receivers). Therefore, writing in a
 1057 clearer way with more informative notations, we need to solve the adjoint-state equation in the form
 1058 of

$$\nabla \cdot [\lambda(\mathbf{x}) \nabla t(m, \mathbf{x})] = \sum_{i=1}^{N_r} [t(m, \mathbf{r}_i) - T(\mathbf{r}_i)], \quad (78)$$

1059 where we use \mathbf{r}_i to indicate the spatial location of the i -th receiver. The traveltime field $t(m, \mathbf{x})$ is the
 1060 traveltime corresponding to a model m in some inversion iteration, and is precomputed beforehand.

1061 Equation (78) can be solved using exactly the same method as that described in Appendix A of
 1062 the work by [Taillandier et al. \(2009\)](#). However, the major difference is that an arbitrary number of
 1063 \mathbf{x}_r can be at any position of Ω . The initial condition for equation (78) is the traveltime difference,
 1064 $t(m, \mathbf{r}_i) - T(\mathbf{r}_i)$, at the position of each receiver. For double-difference misfit, it is not difficult to
 1065 obtain that the adjoint-state equation is

$$\nabla \cdot [\lambda(\mathbf{x}) \nabla t(m, \mathbf{x})] = \sum_{i=1}^{N_r} \left(\sum_{j=1}^{N_r} [(t(m, \mathbf{r}_i) - t(m, \mathbf{r}_j)) - (T(\mathbf{r}_i) - T(\mathbf{r}_j))] \right). \quad (79)$$

1066 It is straightforward to derive the adjoint-state equations for the elastic case where $m = (v_p, v_s)$.
 1067 For brevity, we omit the details here.

1068 **Appendix C: Multitask machine learning models for inferring 1069 and refining fault attributes from a source image**

1070 We develop a multitask supervised ML method to infer and refine fault and fault attributes from
 1071 a source image. We display the architectures of the multitask inference and refinement NNs in
 1072 Figure 32. The input to the multitask inference NN is a source image computed using equation (52),
 1073 while the output from this NN includes the fault probability, fault dip, and fault strike (in 2D, fault
 1074 strike does not apply). In some cases, the fault surfaces estimated by this inference NN can be
 1075 “noisy” and contain “cheese holes” ([Gao, 2024](#)) because of insufficient source density. Using these

1076 fault surfaces as a guidance for source relocation may not be optimal. Therefore, the inference
1077 results are then transferred to the multitask refinement NN for refinement, and we use the refined
1078 fault attributes in LATTE as a fault/fracture constraint for source location.

1079 In both multitask inference and refinement NNs, we use a residual U-Net (ResUNet) as encoders
1080 and decoders. We leverage the ResUNet architecture developed in [Gao \(2024\)](#) to achieve a large
1081 inception field. The open-source codes associated with the multitask inference and refinement NNs
1082 based on a source image, the training strategy, as well as the algorithms and codes for generating
1083 training data and labels, are available in the repository of LATTE.

1084 **List of Figures**

1085 1	(a) A P-wave velocity model, (b) an S-wave velocity model, and (c) two characteris-	44
1086 tic reflectors for validating LATTE's traveltime computation functions.		
1087 2	Travelttime fields computed using LATTE eikonal solver for the model shown in	
1088 Figure 1. (a) First-arrival travelttime field t_p , (b-c) PP-reflection travelttime fields		
1089 t_{pp}^1 and t_{pp}^2 associated with the first and second reflectors, respectively, and (d-e)		
1090 PS-reflection travelttime fields t_{ps}^1 and t_{ps}^2 associated with the first and the second		
1091 reflectors, respectively.	45	
1092 3	Travelttime fields computed using LATTE eikonal solver for the model shown in	
1093 Figure 1. (a) First-arrival travelttime field t_s , (b-c) SS-reflection travelttime fields		
1094 t_{ss}^1 and t_{ss}^2 associated with the first and second reflectors, respectively, and (d-e)		
1095 SP-reflection travelttime fields t_{sp}^1 and t_{sp}^2 associated with the first and the second		
1096 reflectors, respectively.	46	
1097 4	Travelttime computed using LATTE eikonal solver for (a) P incident wave, and (b)	
1098 S incident wave, respectively.	47	
1099 5	(a) A v_p model used for validating LATTE's FATT functionality, and (b) smooth	
1100 1D v_p as the initial model for FATT.	48	
1101 6	Inverted v_p models obtained with (a) AD-FATT (absolute travelttime misfit) and	
1102 (b) DD-FATT (double-difference travelttime misfit), respectively. The results are		
1103 plotted on the same color scale.	49	
1104 7	A comparison between the normalized data misfit convergence curves associated	
1105 with AD-FATT (blue) and DD-FATT (red). The curves are plotted on the logarithmic		
1106 scale.	50	
1107 8	Comparisons between the observed travelttime and the synthetic travelttime of the	
1108 third source simulated in (a) the initial v_p model, and (b) the AD-FATT-updated v_p		
1109 model, and (c) DD-FATT-updated v_p model, respectively. Bottom panels display		
1110 the probability distributions and the fitted Gaussian of travelttime misfit for all		
1111 the $N_s \times N_r = 40 \times 401$ traces obtained using (d) AD-FATT and (e) DD-FATT,		
1112 respectively, where a smaller absolute value of μ , a smaller σ , and a narrower error		
1113 range represent more accurate travelttime fit.	51	
1114 9	A checkerboard model for validating the source location and joint tomography-	
1115 location functionalities of LATTE. We set $V_s = V_p / \sqrt{3}$ for simplicity.	52	
1116 10	(a) Initial source locations and (b-d) inverted source locations in the 5th, 10th, and	
1117 100th iterations obtained using AD-TLOC, respectively.	53	
1118 11	Comparisons among the ground-truth, initial guess, and synthetic values in the	
1119 final inversion model regarding (a) the origin time, (b) P-arrival travelttime, and		
1120 (c) S-arrival travelttime. The gray dots in the panels represent absolute differences		
1121 between the synthetic and ground-truth values.	54	
1122 12	(a) Initial source locations and (b-d) inverted source locations in the 5th, 10th, and	
1123 100th iterations using DD-TLOC, respectively.	55	

1124	13	(a) Comparisons among the ground-truth traveltime $t_p - \eta_0$ (blue curve), the synthetic traveltime in the initial model $t_p^{(l=0)}$ (green curve), and the inverted model $t_p^{(l=100)}$ (red curve). Panel (b) displays the S-arrival traveltme result. The gray dots in the panels represent absolute differences between the synthetic and ground-truth values.	56
1125	14	Inverted (a) v_p and (b) v_s models by DD-TLOC.	57
1126	15	(a) Initial source locations and (b-d) inverted source locations in the 5th, 10th, and 100th iterations using DD-TLOC, respectively.	58
1127	16	(a) Comparisons among the ground-truth traveltme $t_p - \eta_0$ (blue curve), the synthetic traveltme in the initial model $t_p^{(l=0)}$ (green curve), and the inverted model $t_p^{(l=100)}$ (red curve). Panel (b) displays the S-arrival traveltme result. The gray dots in the panels represent absolute differences between the synthetic and ground-truth values.	59
1128	17	Inverted (a) v_p and (b) v_s models using regularized DD-TLOC.	60
1129	18	(a) Initial source locations and (b-d) inverted source locations in the 5th, 10th, and 100th iterations with regularized DD-TLOC, respectively.	61
1130	19	(a) Comparisons among the ground-truth traveltme $t_p - \eta_0$ (blue curve), the synthetic traveltme in the initial model $t_p^{(l=0)}$ (green curve), and the inverted model $t_p^{(l=100)}$ (red curve). Panel (b) displays the S-arrival traveltme result. The gray dots in the panels represent absolute differences between the synthetic and ground-truth values.	62
1131	20	(a) A v_p model overlain by sources and receivers, and (b) a v_p model by smoothing the model in Panel (a) with a Gaussian filter for validating TLOC.	63
1132	21	Two examples of clean and noisy data. For clarity, the origin time is subtracted from the data. In both panels, the noise data are generated by adding smoothed random noise displayed as a gray curve (consisting of 80 points) on the top. All the 1,200 common-source gathers are added with similar noise as in this figure to mimic time picking error in practice.	64
1133	22	(a) Initial source locations, (b) inverted source locations using TLOC at the 5th iteration (for both cases of with and without ML-based source regularization), (c) inverted source locations using TLOC at the 50th iteration without ML-based source regularization, and (d) with ML-based source regularization.	65
1134	23	(a-b) Source image and ML-inferred and refined fault dip image in the 5th iteration, and (c-d) in the 50th iteration.	66
1135	24	(a) A 3D heterogeneous v_p model with four intersecting faults designed for validating LATTE's DD-TLOC joint tomography-location. (b) The background smooth velocity model. The 3D plots at the top-right corner of both panels show the ground-truth fault surfaces and source locations.	67
1136	25	Two examples of clean and noisy data. For clarity, the origin time is subtracted from the data. In both panels, the noise data are generated by adding smoothed random noise displayed as a gray curve (consisting of 150 points) on the top. All the 1,200 common-source gathers are added with similar noise as in this figure to mimic time picking error in practice.	68

1167	26	Inverted (a) v_p and (b) v_s models using DD-TLOC without model or source parameter regularization.	69
1168	27	Inverted (a) v_p and (b) v_s models using DD-TLOC with model and source parameter regularization.	70
1169	28	(a-b) Ground-truth Δv_p and Δv_s at a depth of 100 m, (c-d) inverted Δv_p and Δv_s by DD-TLOC without model parameter regularization, and (e-f) inverted Δv_p and Δv_s with model parameter regularization.	71
1170	29	(a-b) Ground-truth Δv_p and Δv_s at a depth of 340 m, (c-d) inverted Δv_p and Δv_s by DD-TLOC without model parameter regularization, and (e-f) inverted Δv_p and Δv_s with model parameter regularization.	72
1171	30	Map views of (a) ground-truth source locations colored by their depth, (b) initial source locations, (c-d) inverted source locations colored by horizontal/depth errors at the 50th iteration without ML-based source parameter regularization, and (e-f) inverted source locations colored by horizontal/depth errors with ML-based source parameter regularization.	73
1172	31	(a) A comparison between ground-truth source locations (blue balls) and DD-TLOC-inverted source locations without source parameter regularization (red balls). (b) A similar comparison with that in Panel (a) but the red balls represent the source locations inverted by DD-TLOC with source parameter regularization. (c-d) 3D views of the ground-truth source locations (blue balls) and the faults inferred and refined using our multitask NNs in the 50th iteration.	74
1173	32	Architecture of our multitask fault inference and refinement NNs for inferring/refining fault attributes from a source image.	75
1174			
1175			
1176			
1177			
1178			
1179			
1180			
1181			
1182			
1183			
1184			
1185			
1186			
1187			
1188			
1189			

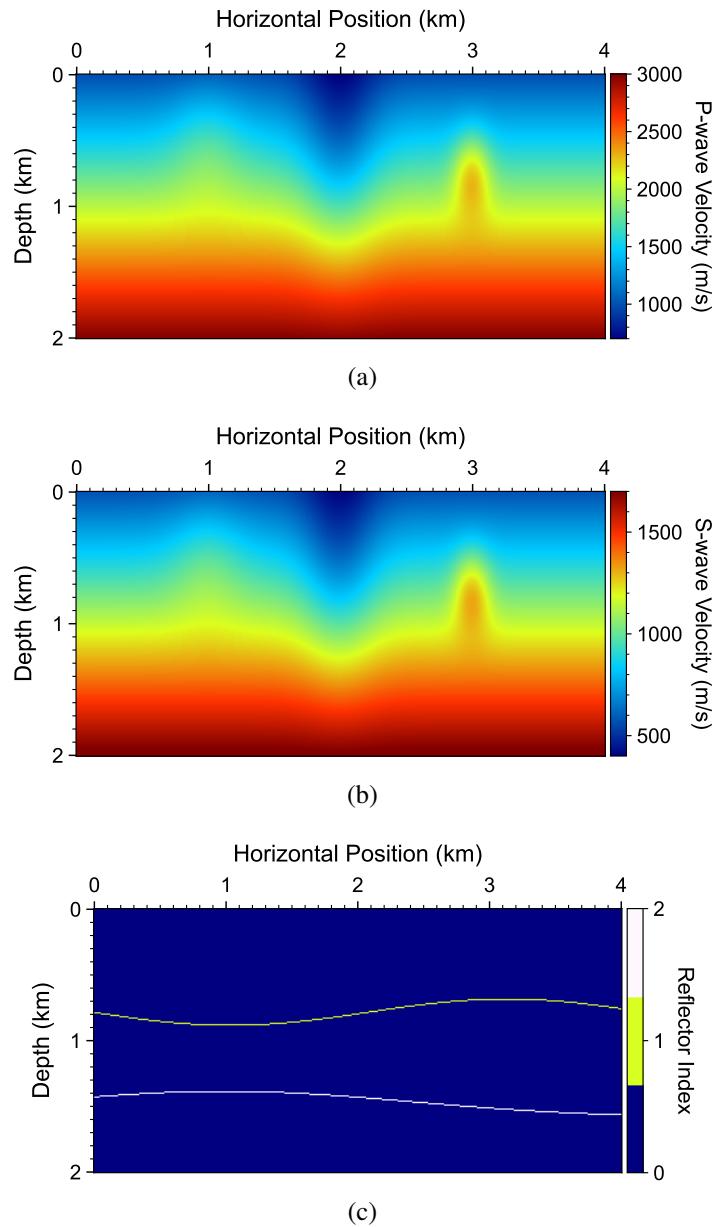


Figure 1: (a) A P-wave velocity model, (b) an S-wave velocity model, and (c) two characteristic reflectors for validating LATTE's travelttime computation functions.

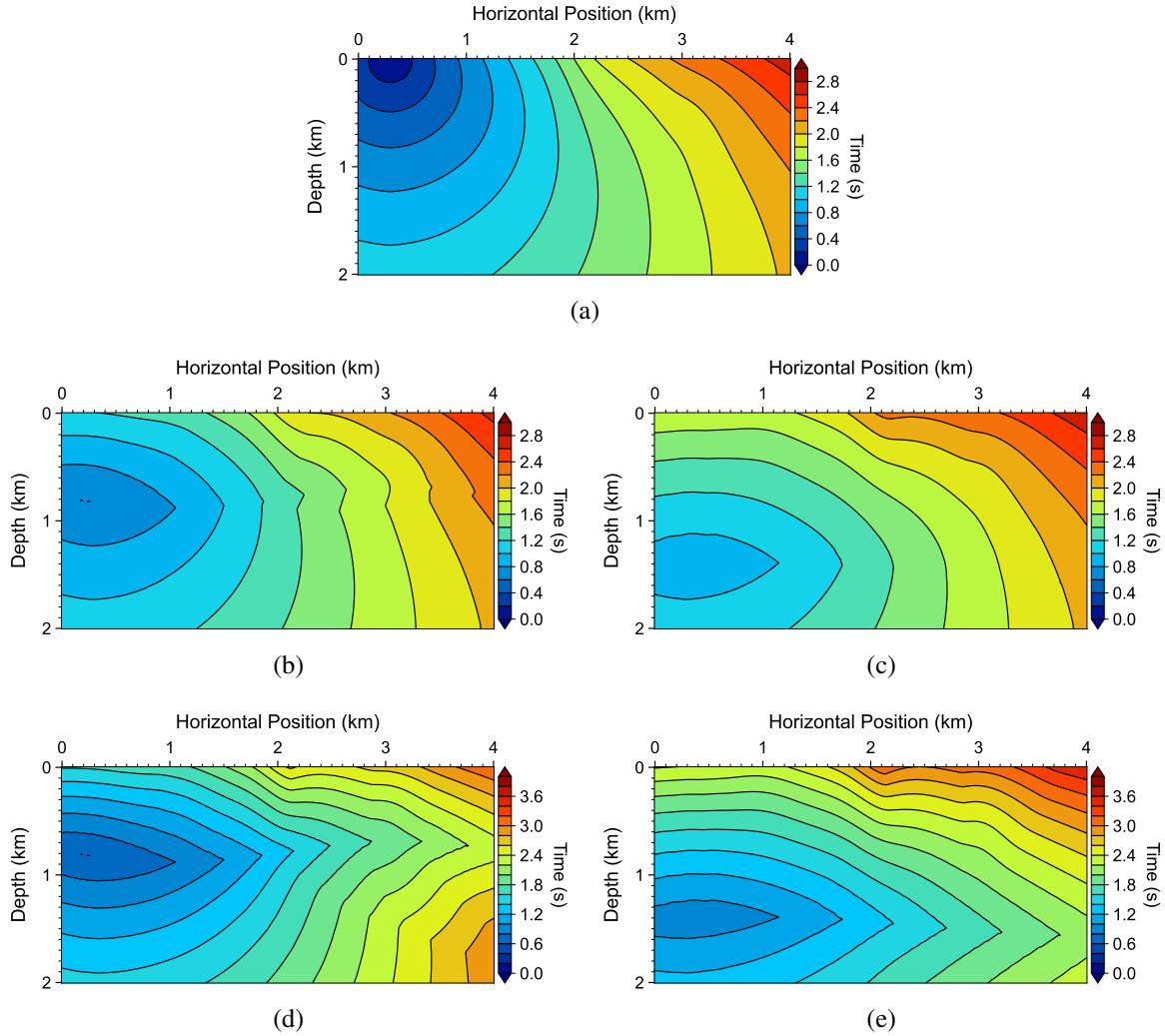


Figure 2: Traveltimes computed using LATTE eikonal solver for the model shown in Figure 1. (a) First-arrival traveltimes t_p , (b-c) PP-reflection traveltimes t_{pp}^1 and t_{pp}^2 associated with the first and second reflectors, respectively, and (d-e) PS-reflection traveltimes t_{ps}^1 and t_{ps}^2 associated with the first and the second reflectors, respectively.

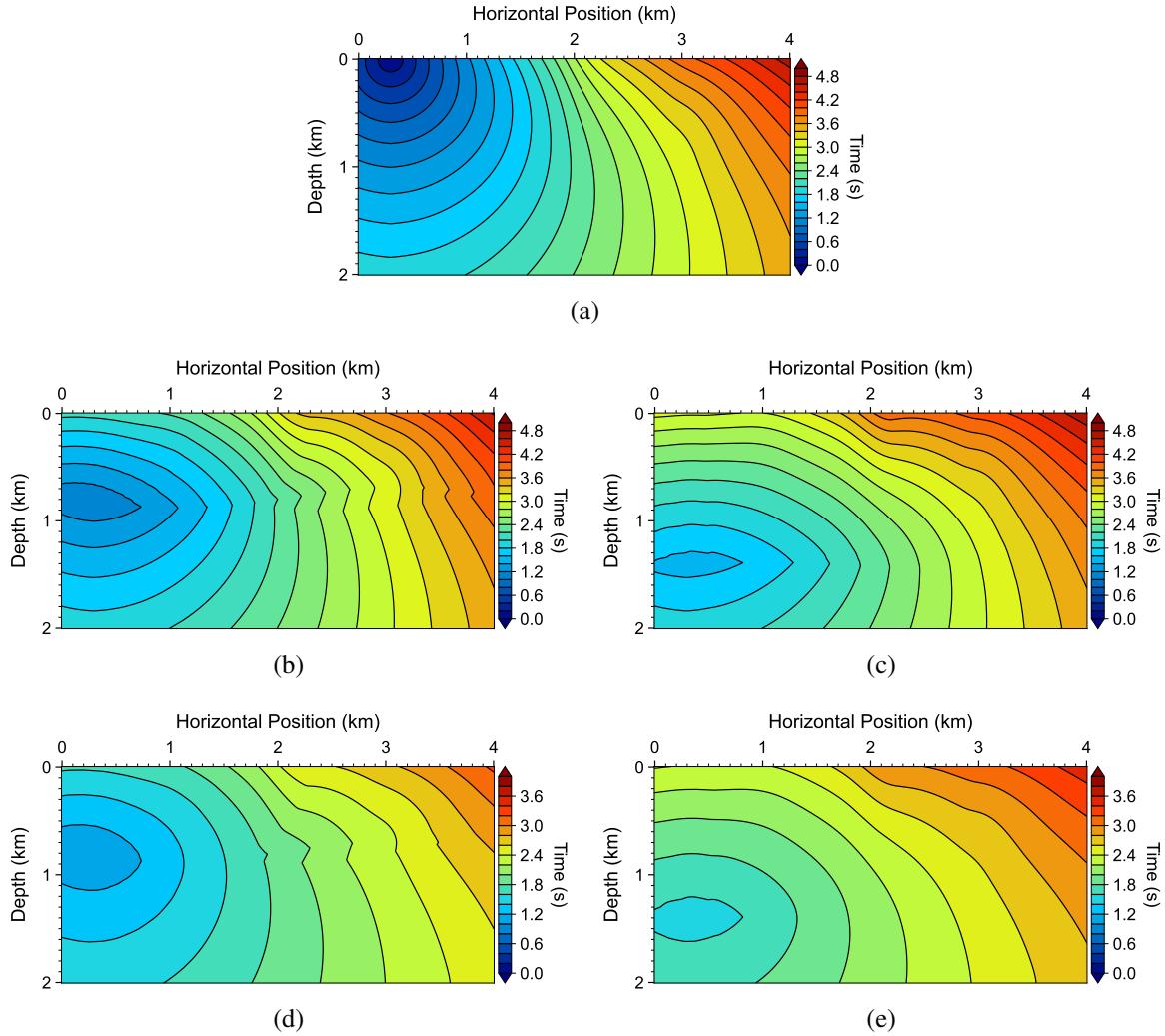


Figure 3: Traveltimes computed using LATTE eikonal solver for the model shown in Figure 1. (a) First-arrival traveltime field t_s , (b-c) SS-reflection traveltimes t_{ss}^1 and t_{ss}^2 associated with the first and second reflectors, respectively, and (d-e) SP-reflection traveltimes t_{sp}^1 and t_{sp}^2 associated with the first and the second reflectors, respectively.

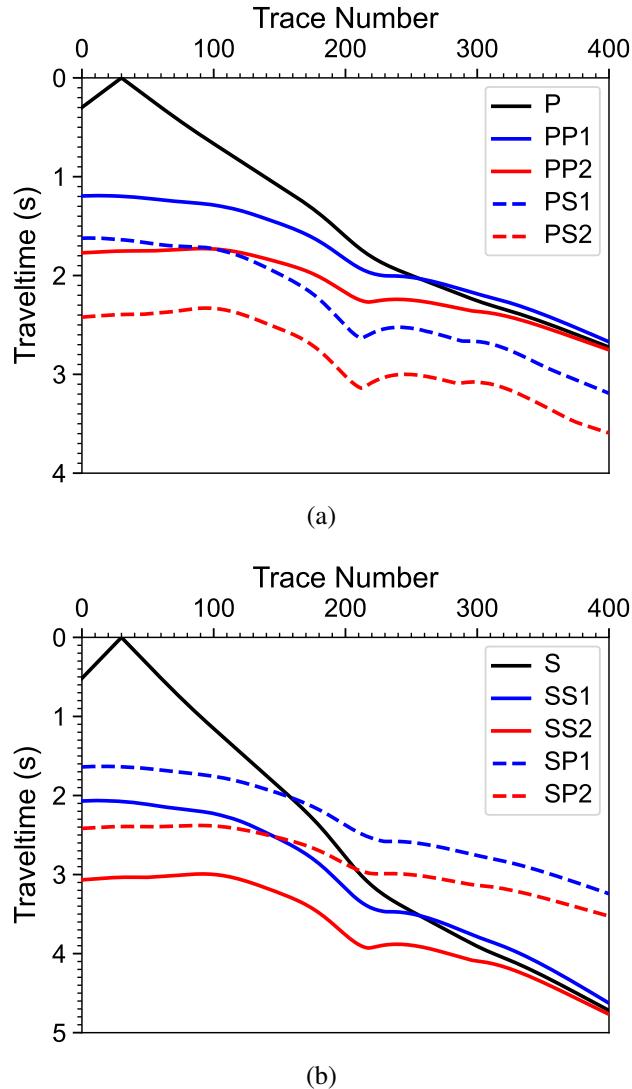


Figure 4: Traveltime computed using LATTE eikonal solver for (a) P incident wave, and (b) S incident wave, respectively.

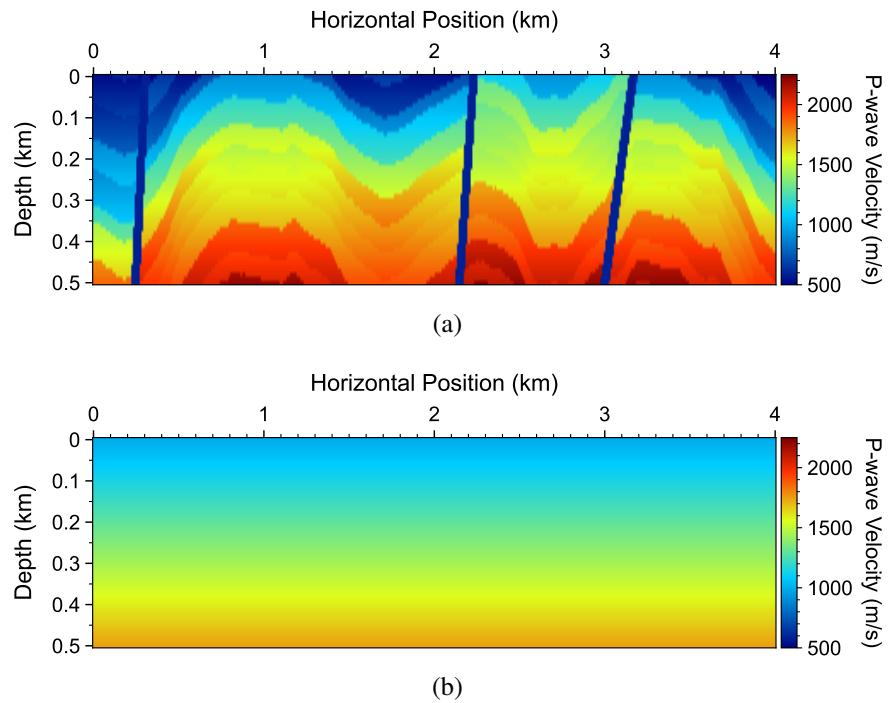


Figure 5: (a) A v_p model used for validating LATTE's FATT functionality, and (b) smooth 1D v_p as the initial model for FATT.

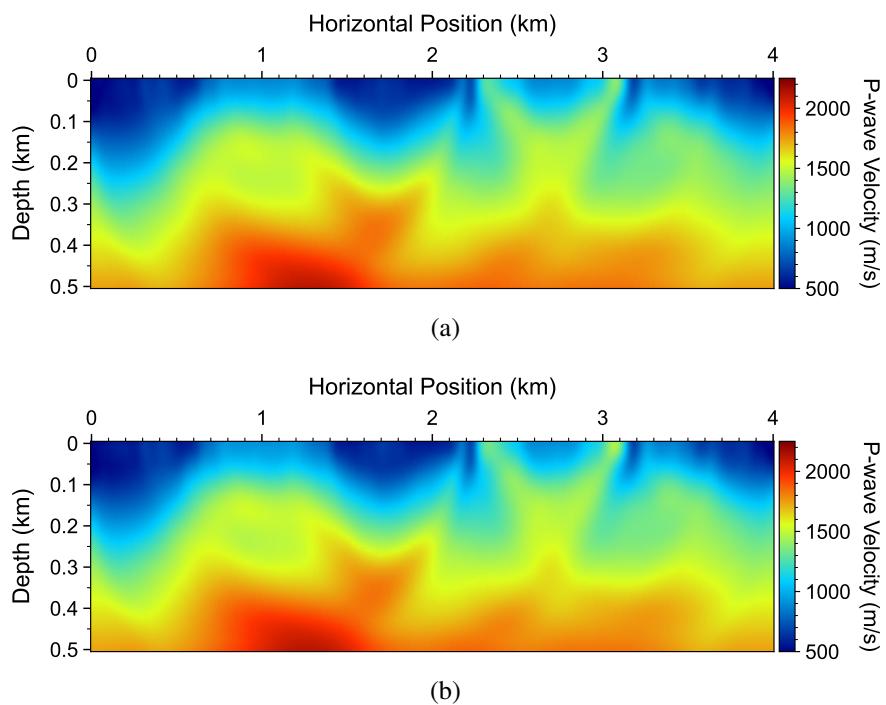


Figure 6: Inverted v_p models obtained with (a) AD-FATT (absolute traveltime misfit) and (b) DD-FATT (double-difference traveltime misfit), respectively. The results are plotted on the same color scale.

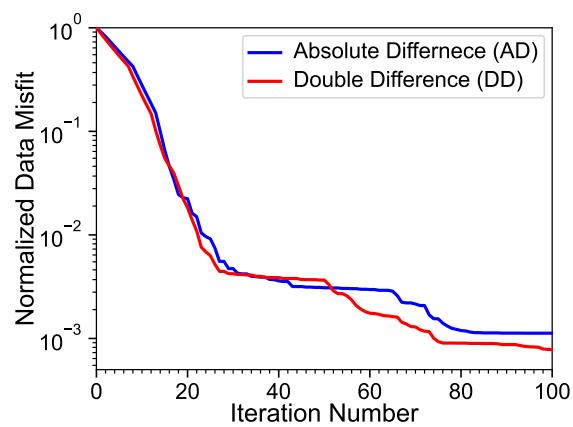


Figure 7: A comparison between the normalized data misfit convergence curves associated with AD-FATT (blue) and DD-FATT (red). The curves are plotted on the logarithmic scale.

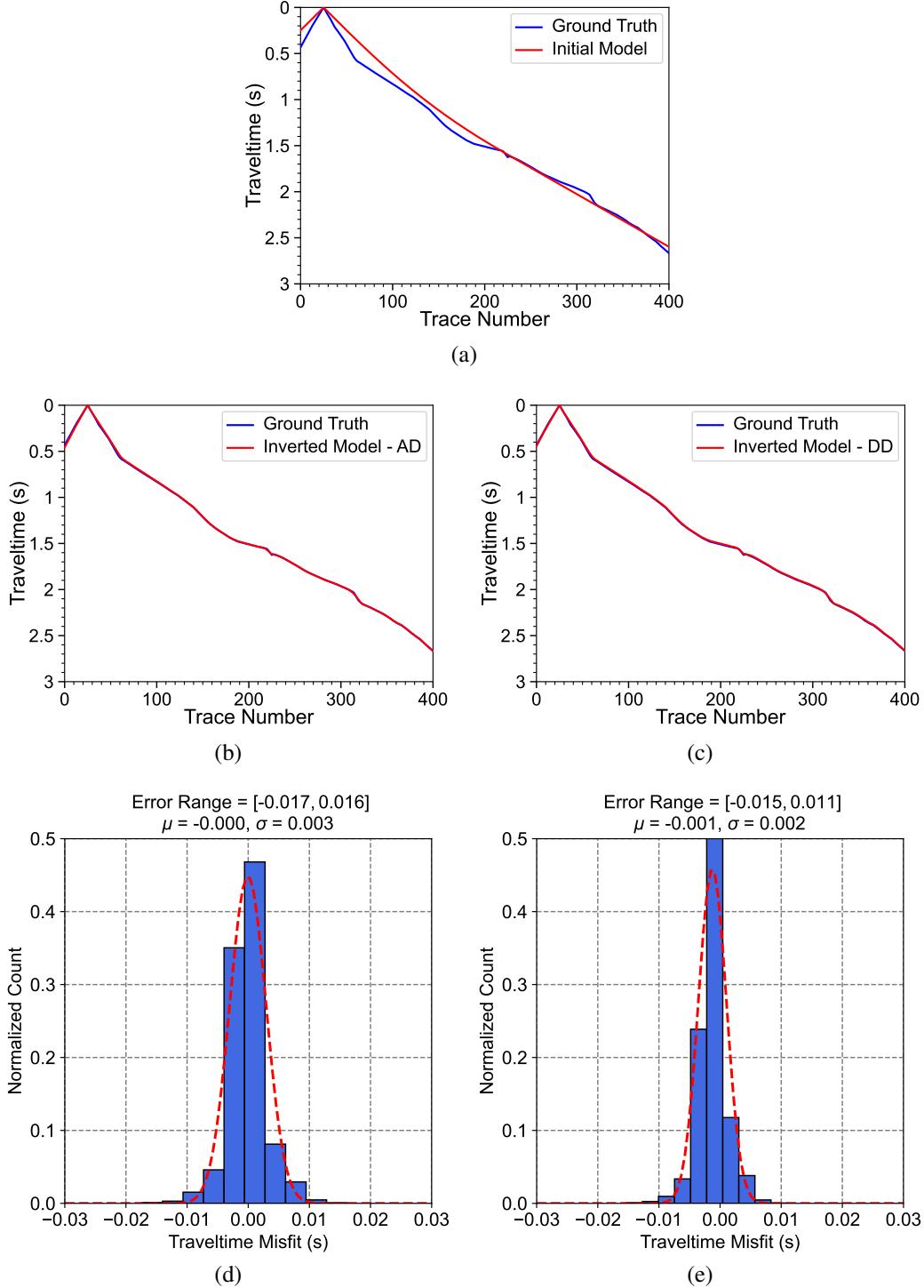


Figure 8: Comparisons between the observed traveltimes and the synthetic traveltimes of the third source simulated in (a) the initial v_p model, and (b) the AD-FATT-updated v_p model, and (c) DD-FATT-updated v_p model, respectively. Bottom panels display the probability distributions and the fitted Gaussian of traveltime misfit for all the $N_s \times N_r = 40 \times 401$ traces obtained using (d) AD-FATT and (e) DD-FATT, respectively, where a smaller absolute value of μ , a smaller σ , and a narrower error range represent more accurate traveltime fit.

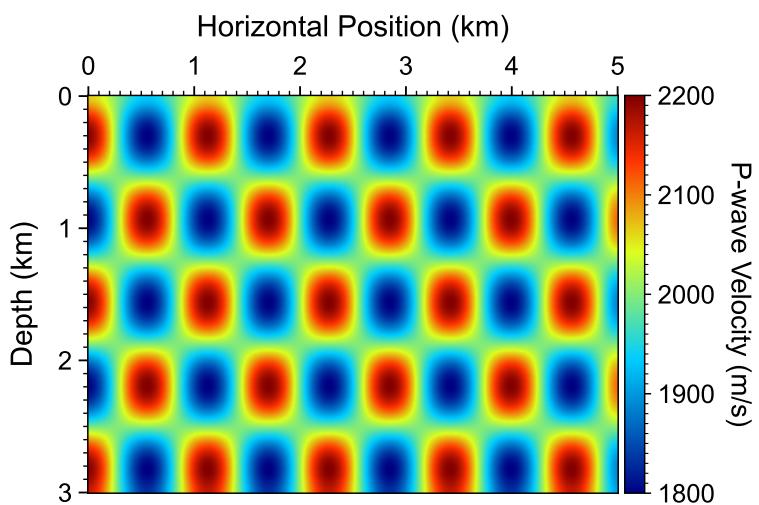


Figure 9: A checkerboard model for validating the source location and joint tomography-location functionalities of LATTE. We set $V_s = V_p / \sqrt{3}$ for simplicity.

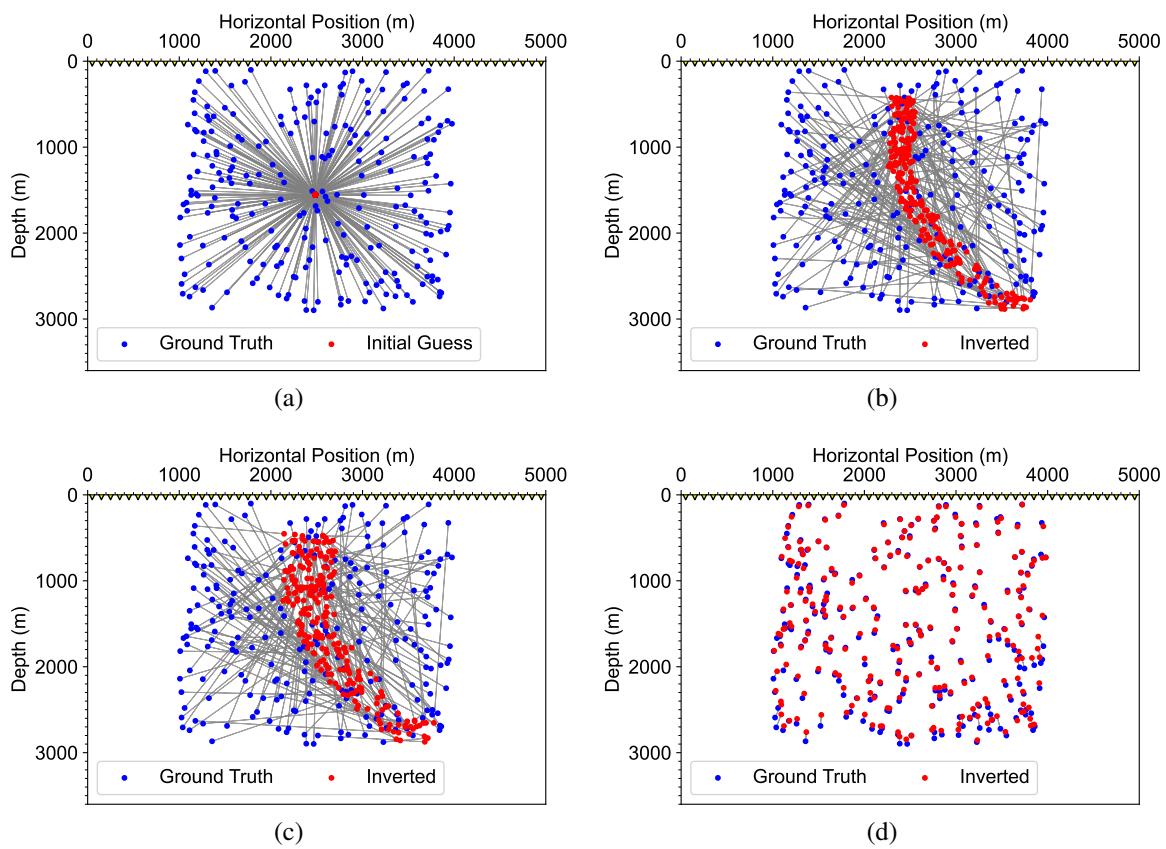
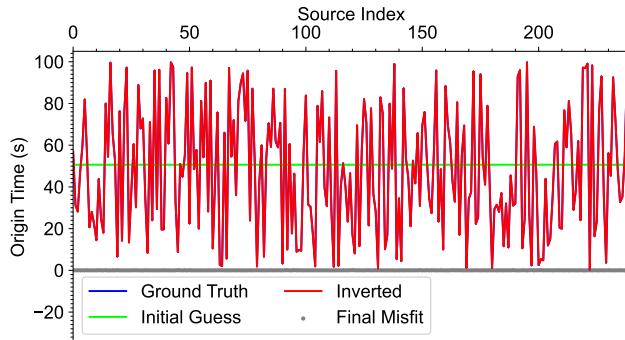
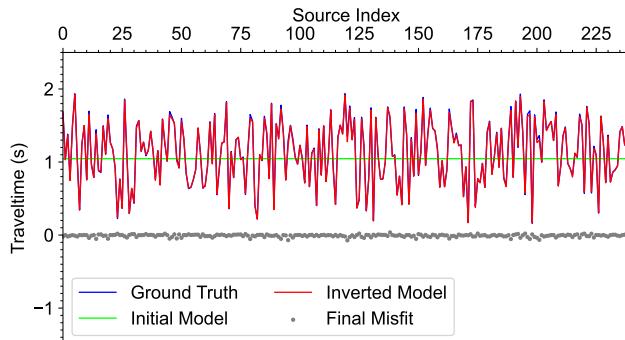


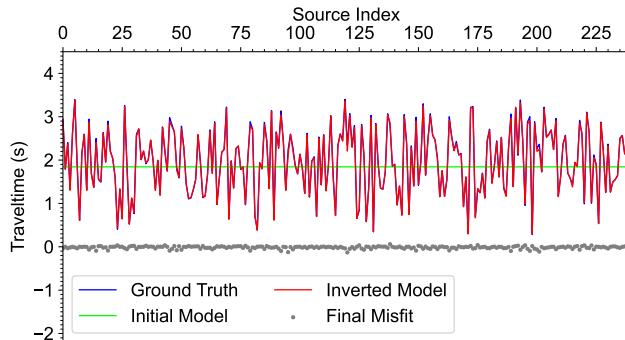
Figure 10: (a) Initial source locations and (b-d) inverted source locations in the 5th, 10th, and 100th iterations obtained using AD-TLOC, respectively.



(a)



(b)



(c)

Figure 11: Comparisons among the ground-truth, initial guess, and synthetic values in the final inversion model regarding (a) the origin time, (b) P-arrival traveltimes, and (c) S-arrival traveltimes. The gray dots in the panels represent absolute differences between the synthetic and ground-truth values.

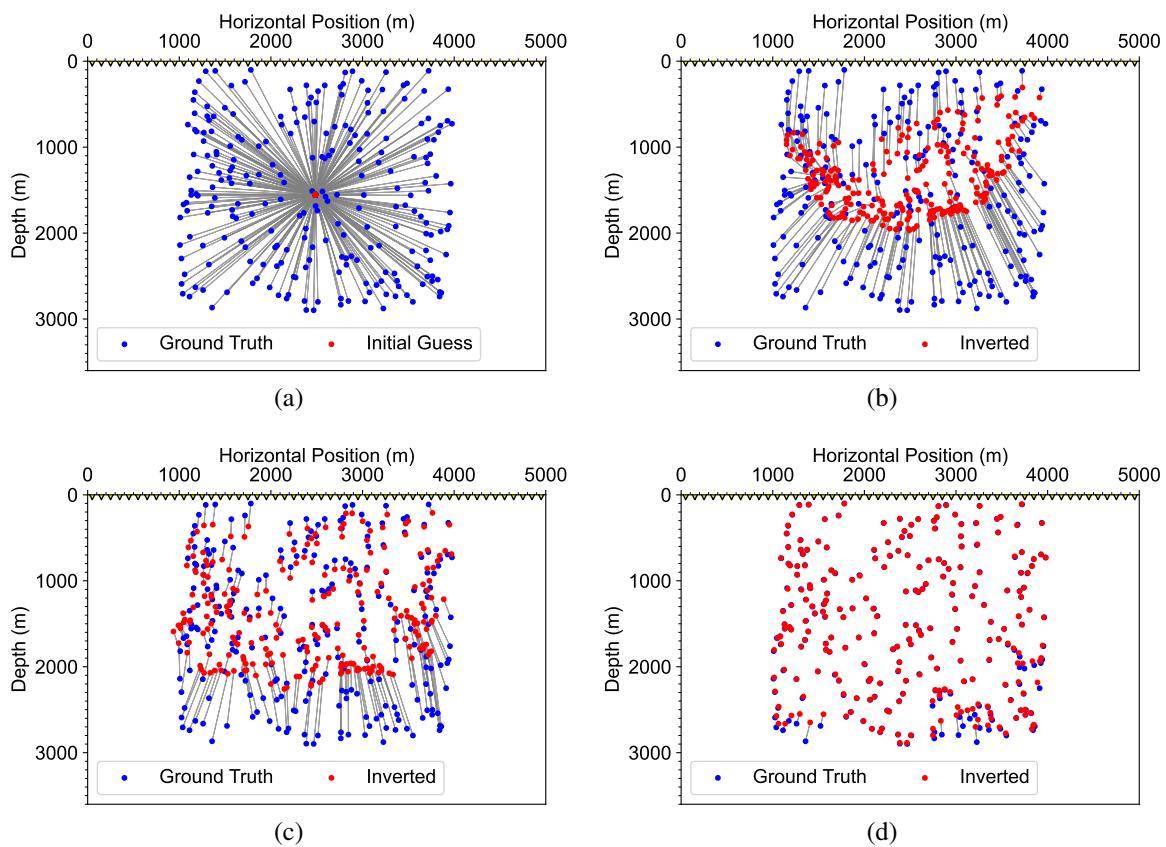
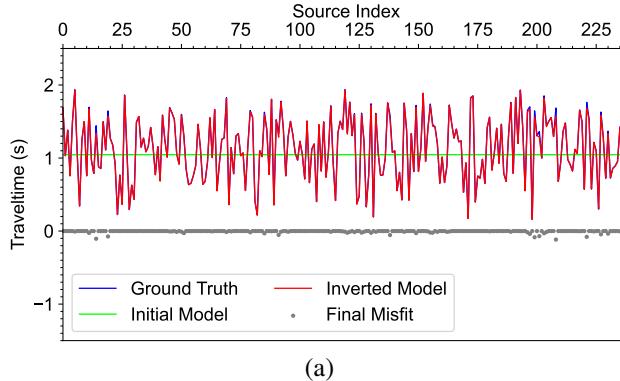
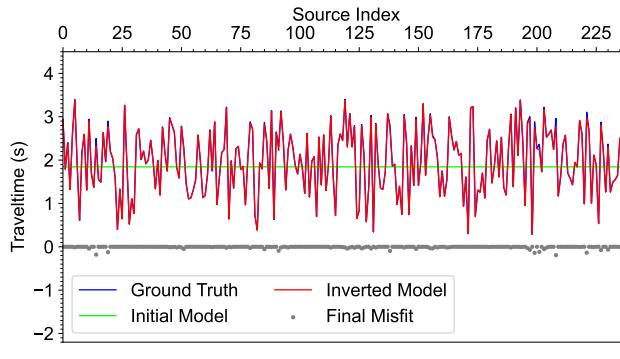


Figure 12: (a) Initial source locations and (b-d) inverted source locations in the 5th, 10th, and 100th iterations using DD-TLOC, respectively.

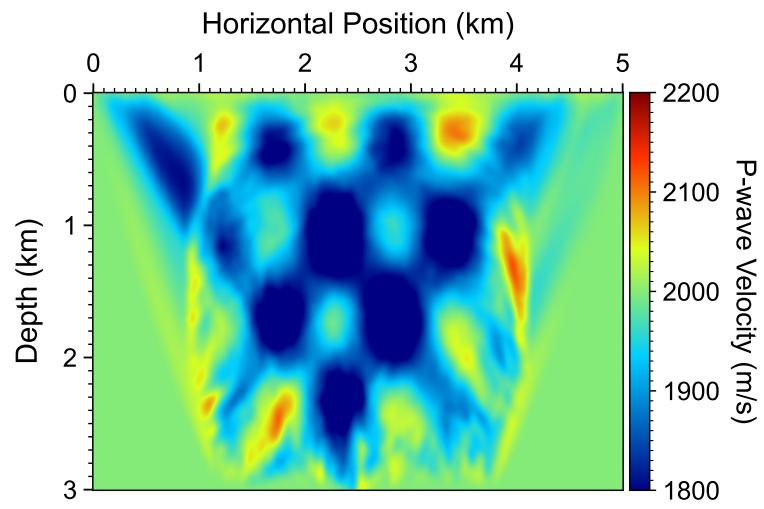


(a)

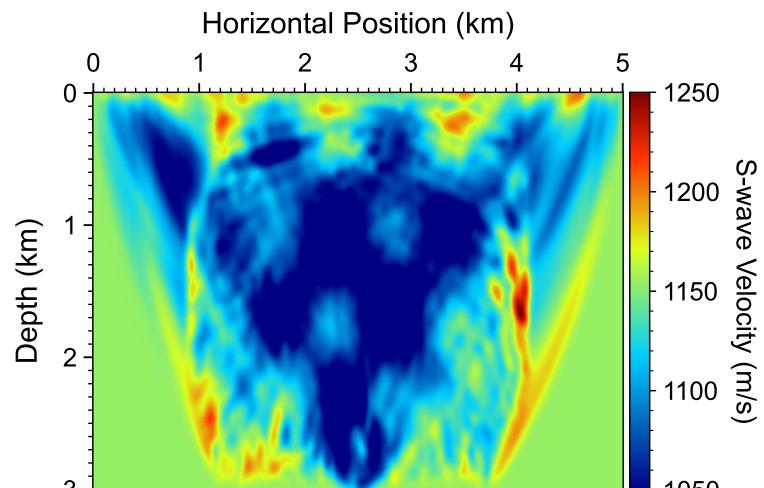


(b)

Figure 13: (a) Comparisons among the ground-truth traveltime $t_p - \eta_0$ (blue curve), the synthetic traveltine in the initial model $t_p^{(l=0)}$ (green curve), and the inverted model $t_p^{(l=100)}$ (red curve). Panel (b) displays the S-arrival traveltine result. The gray dots in the panels represent absolute differences between the synthetic and ground-truth values.



(a)



(b)

Figure 14: Inverted (a) v_p and (b) v_s models by DD-TLOC.

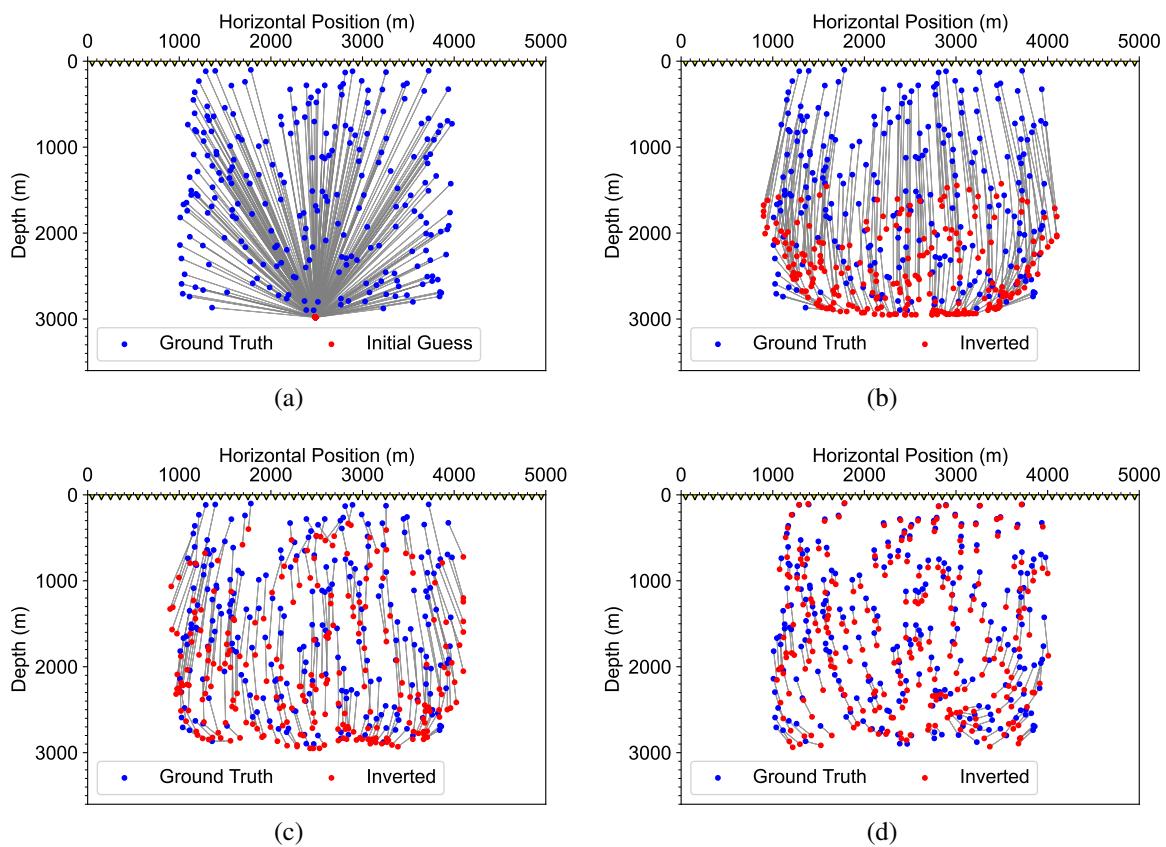
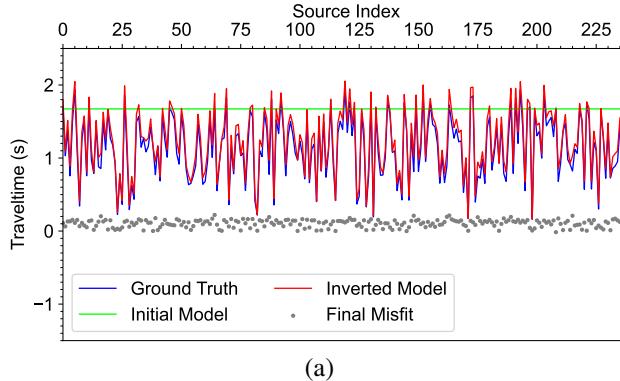
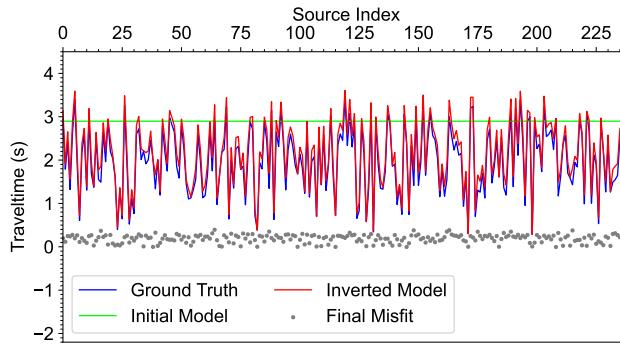


Figure 15: (a) Initial source locations and (b-d) inverted source locations in the 5th, 10th, and 100th iterations using DD-TLOC, respectively.



(a)



(b)

Figure 16: (a) Comparisons among the ground-truth traveltimes $t_p - \eta_0$ (blue curve), the synthetic traveltimes in the initial model $t_p^{(l=0)}$ (green curve), and the inverted model $t_p^{(l=100)}$ (red curve). Panel (b) displays the S-arrival traveltimes result. The gray dots in the panels represent absolute differences between the synthetic and ground-truth values.

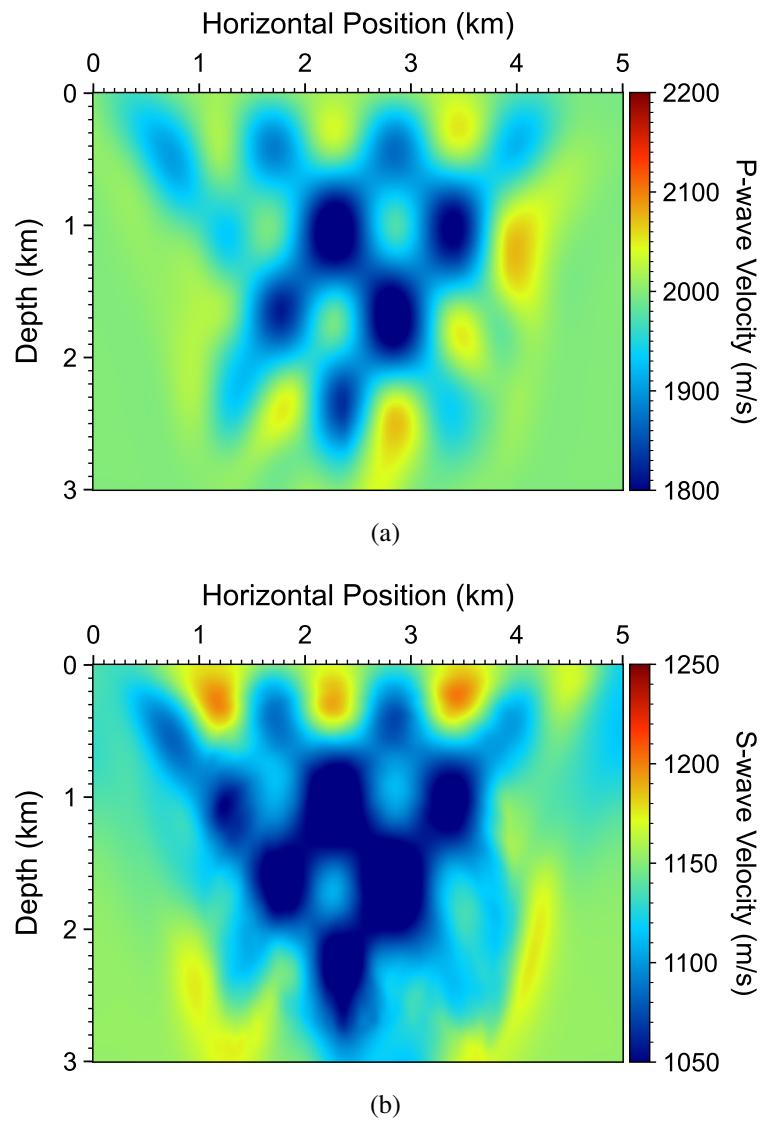


Figure 17: Inverted (a) v_p and (b) v_s models using regularized DD-TLOC.

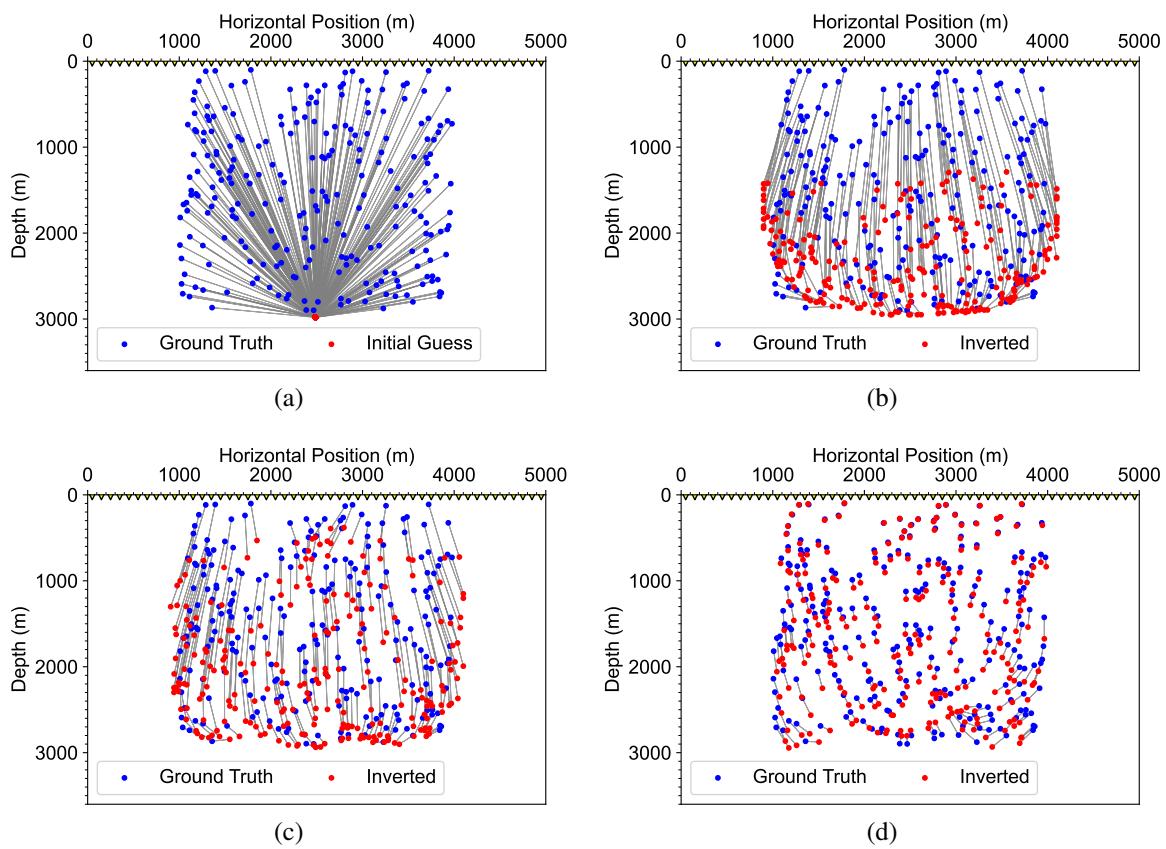
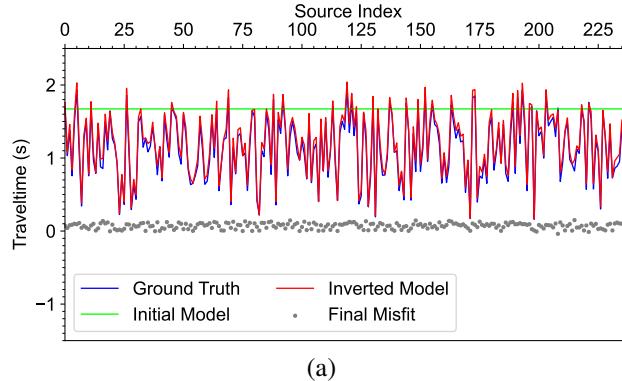
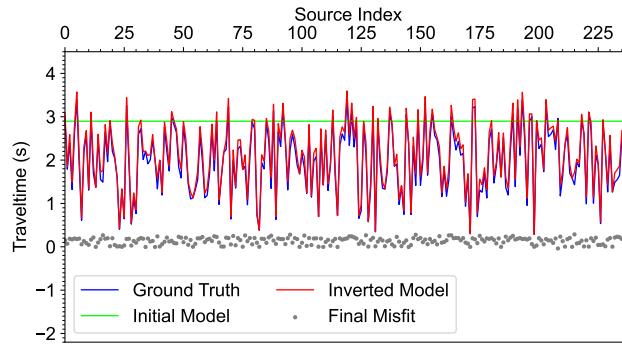


Figure 18: (a) Initial source locations and (b-d) inverted source locations in the 5th, 10th, and 100th iterations with regularized DD-TLOC, respectively.



(a)



(b)

Figure 19: (a) Comparisons among the ground-truth traveltimes $t_p - \eta_0$ (blue curve), the synthetic traveltimes in the initial model $t_p^{(l=0)}$ (green curve), and the inverted model $t_p^{(l=100)}$ (red curve). Panel (b) displays the S-arrival traveltimes result. The gray dots in the panels represent absolute differences between the synthetic and ground-truth values.

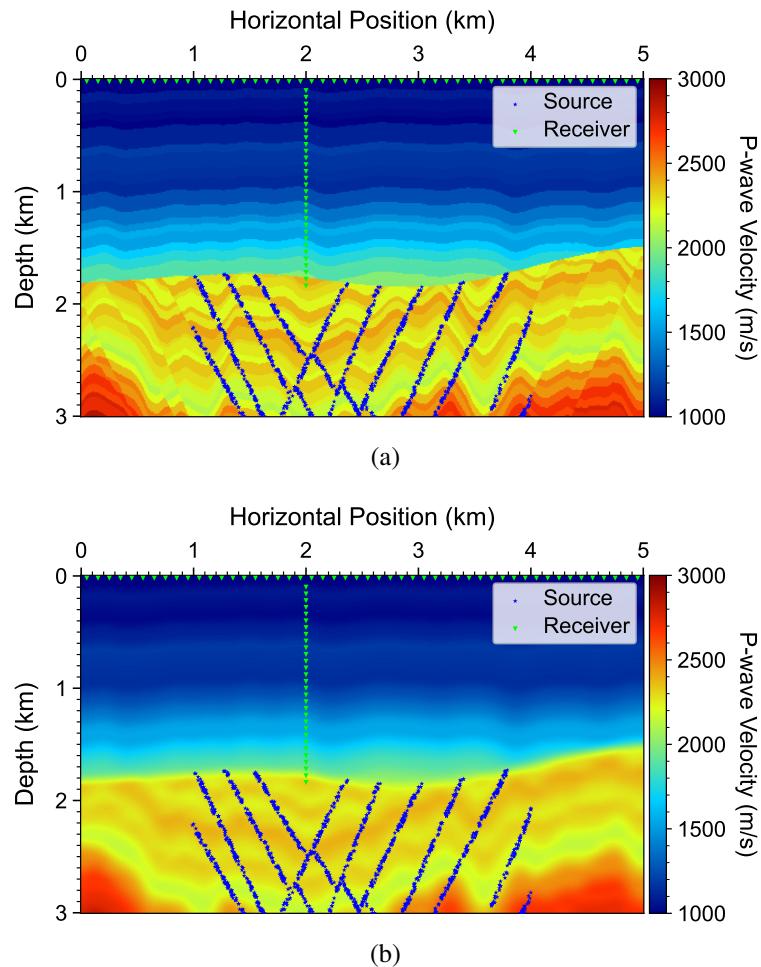
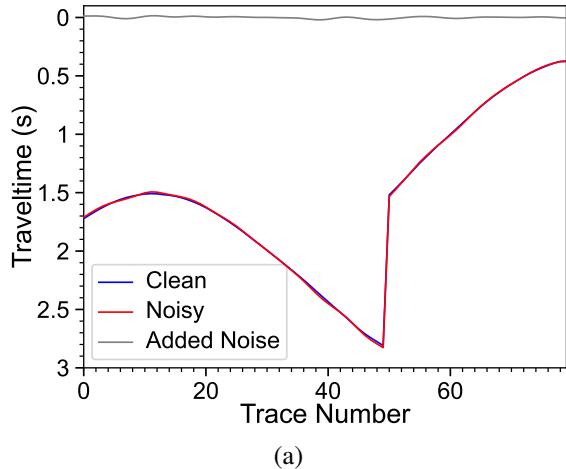
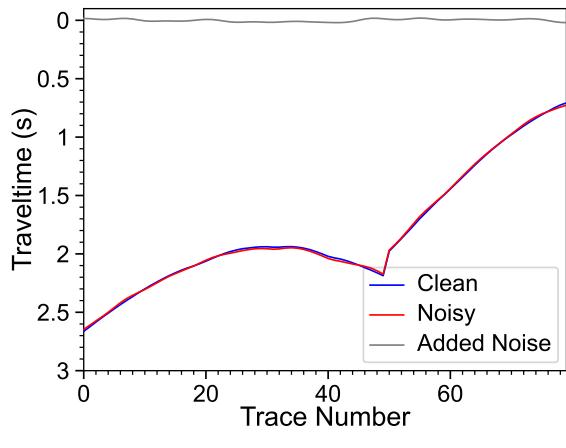


Figure 20: (a) A v_p model overlain by sources and receivers, and (b) a v_p model by smoothing the model in Panel (a) with a Gaussian filter for validating TLOC.



(a)



(b)

Figure 21: Two examples of clean and noisy data. For clarity, the origin time is subtracted from the data. In both panels, the noise data are generated by adding smoothed random noise displayed as a gray curve (consisting of 80 points) on the top. All the 1,200 common-source gathers are added with similar noise as in this figure to mimic time picking error in practice.

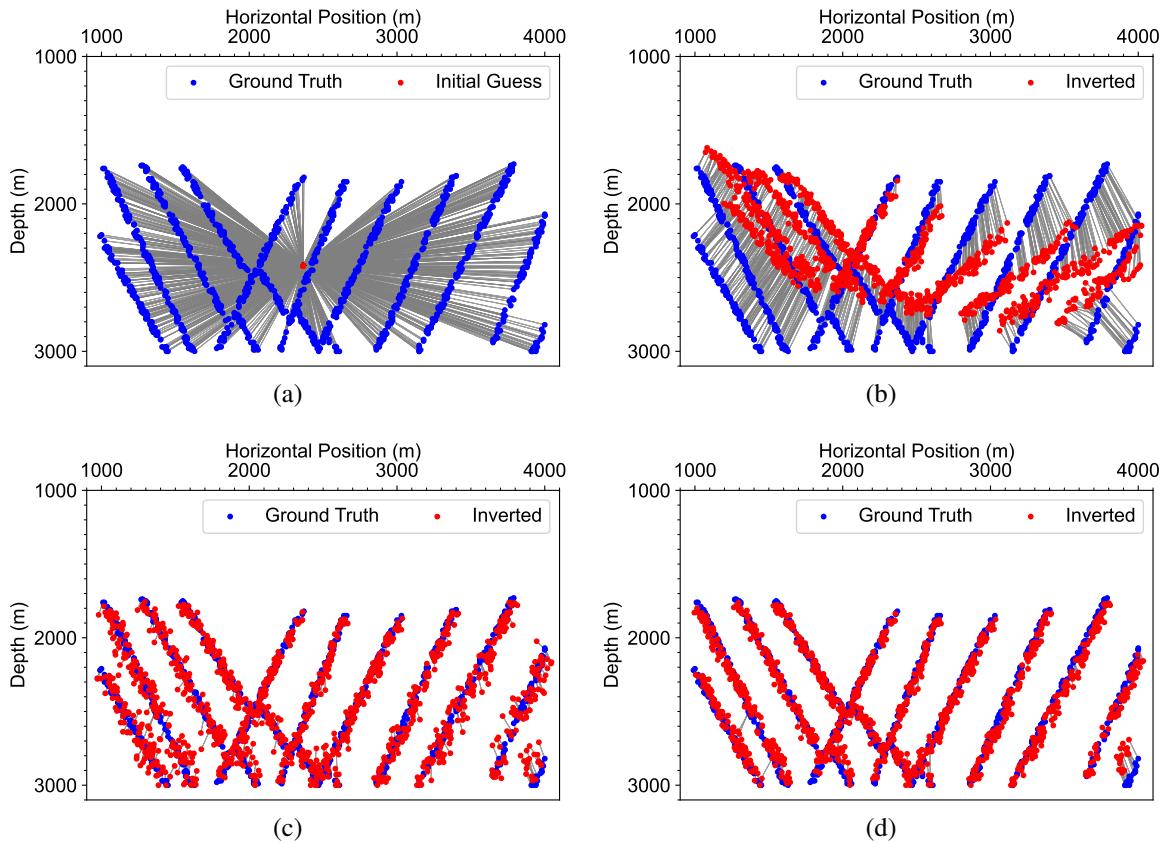


Figure 22: (a) Initial source locations, (b) inverted source locations using TLOC at the 5th iteration (for both cases of with and without ML-based source regularization), (c) inverted source locations using TLOC at the 50th iteration without ML-based source regularization, and (d) with ML-based source regularization.

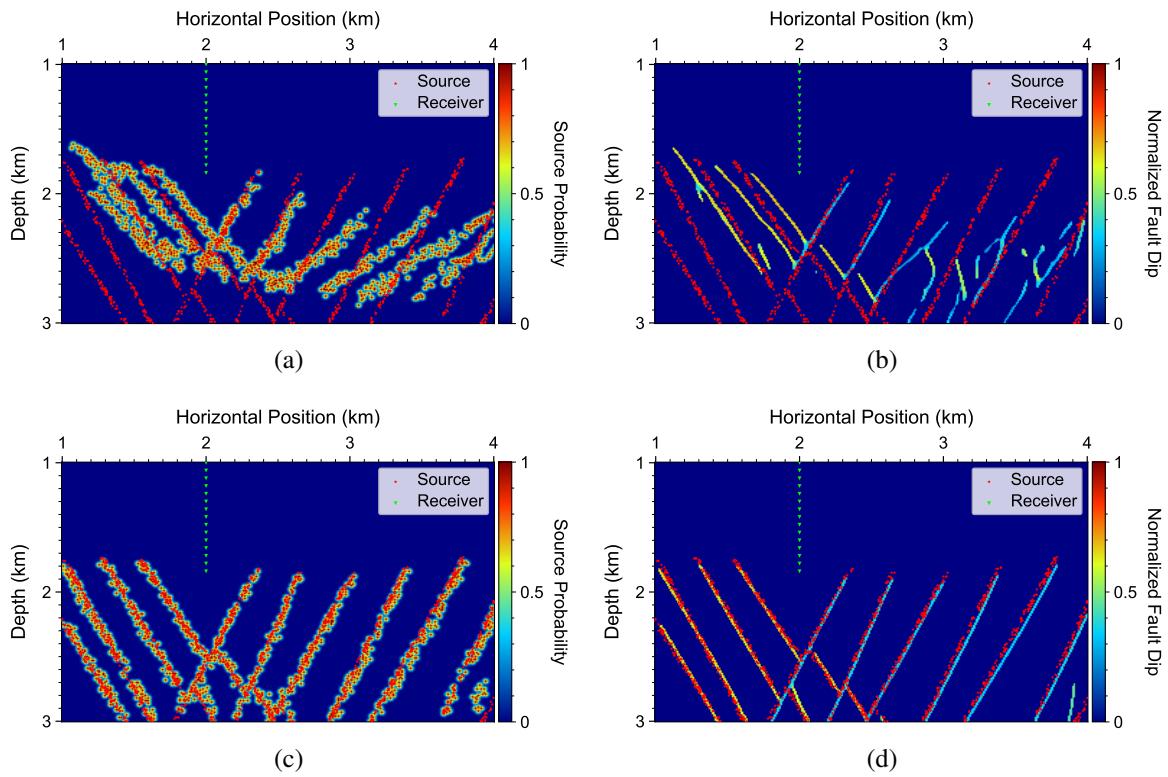


Figure 23: (a-b) Source image and ML-inferred and refined fault dip image in the 5th iteration, and (c-d) in the 50th iteration.

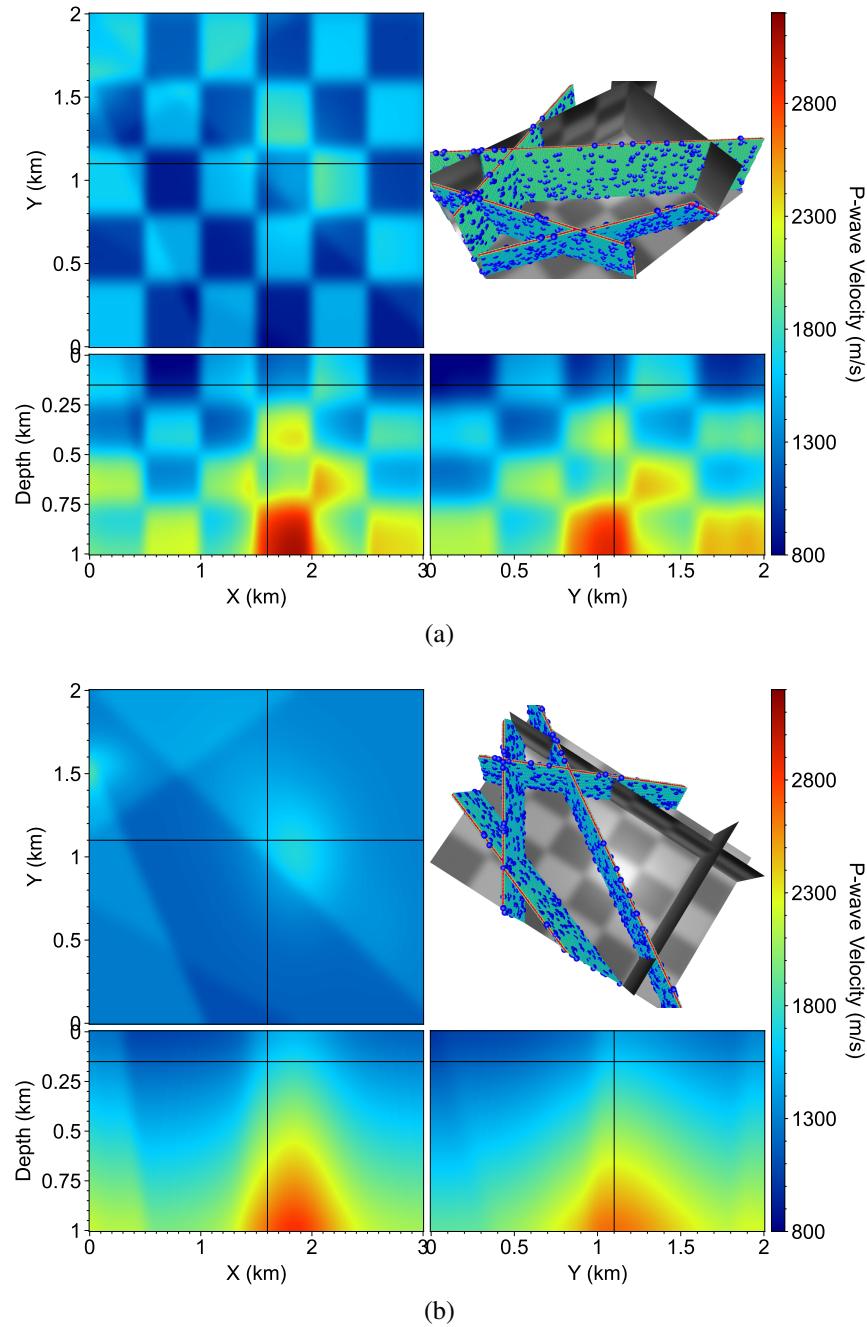
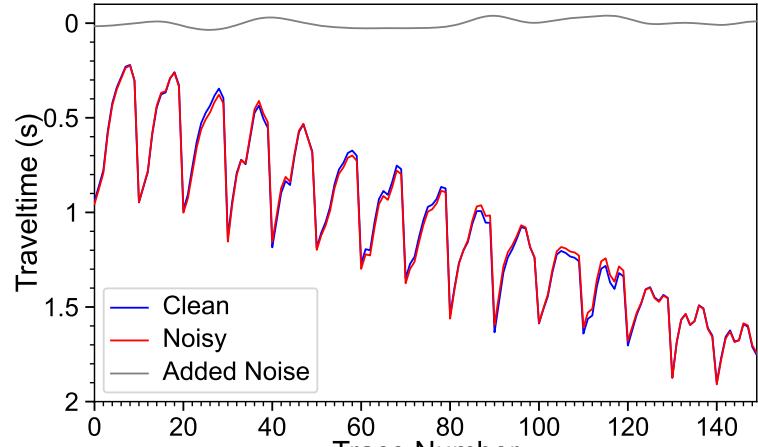
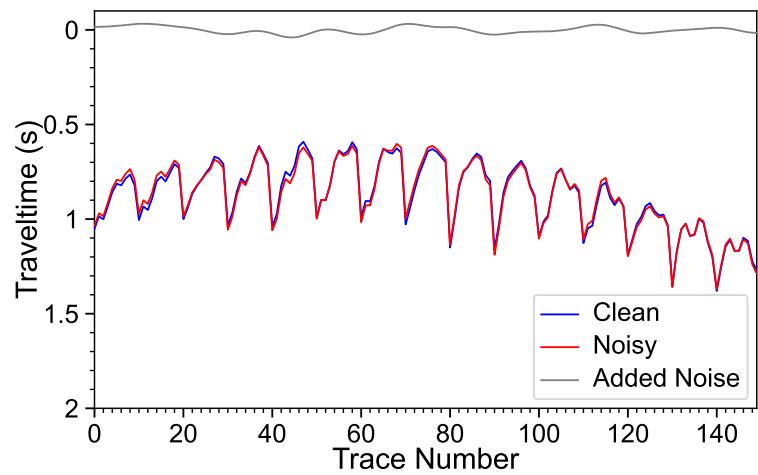


Figure 24: (a) A 3D heterogeneous v_p model with four intersecting faults designed for validating LATTE's DD-TLOC joint tomography-location. (b) The background smooth velocity model. The 3D plots at the top-right corner of both panels show the ground-truth fault surfaces and source locations.



(a)



(b)

Figure 25: Two examples of clean and noisy data. For clarity, the origin time is subtracted from the data. In both panels, the noise data are generated by adding smoothed random noise displayed as a gray curve (consisting of 150 points) on the top. All the 1,200 common-source gathers are added with similar noise as in this figure to mimic time picking error in practice.

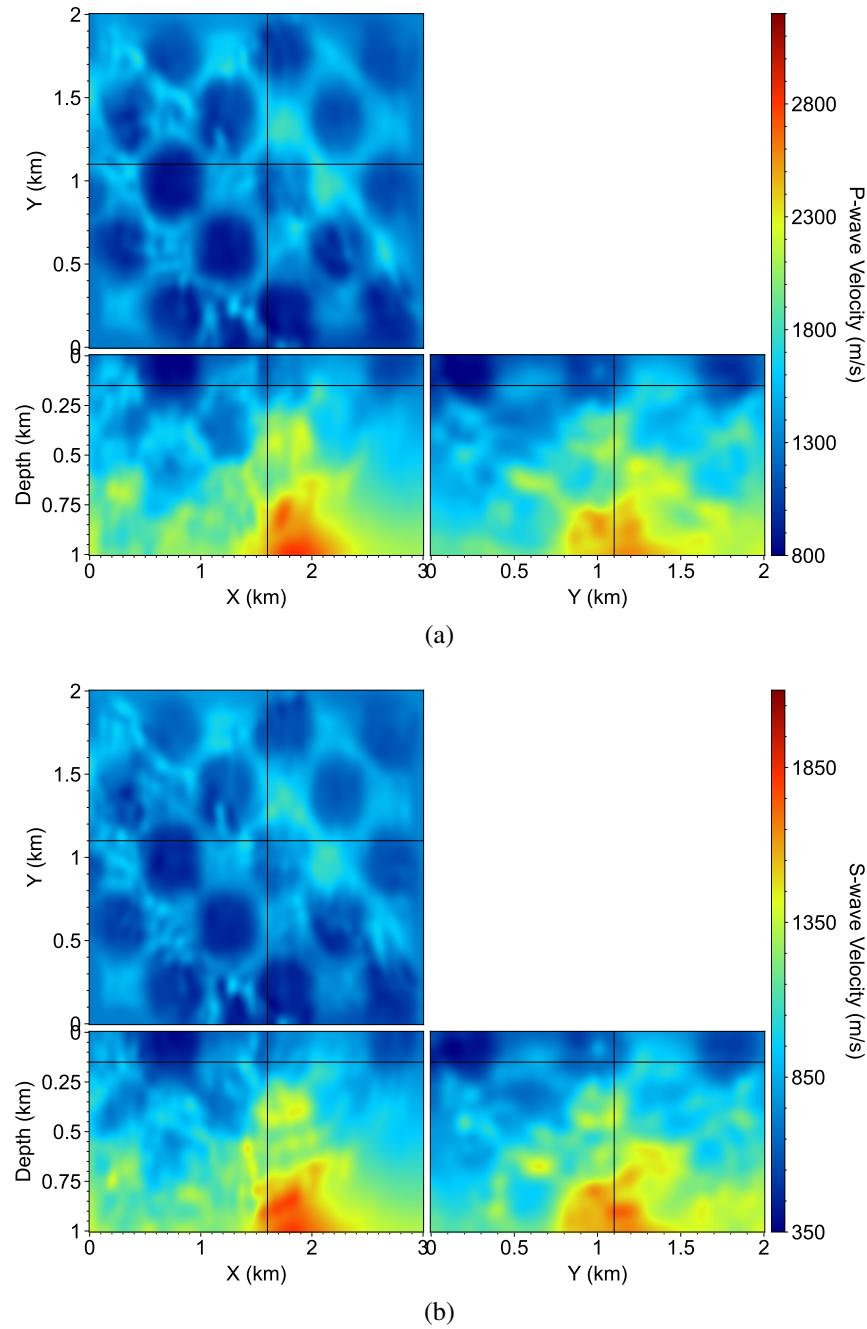


Figure 26: Inverted (a) v_p and (b) v_s models using DD-TLOC without model or source parameter regularization.

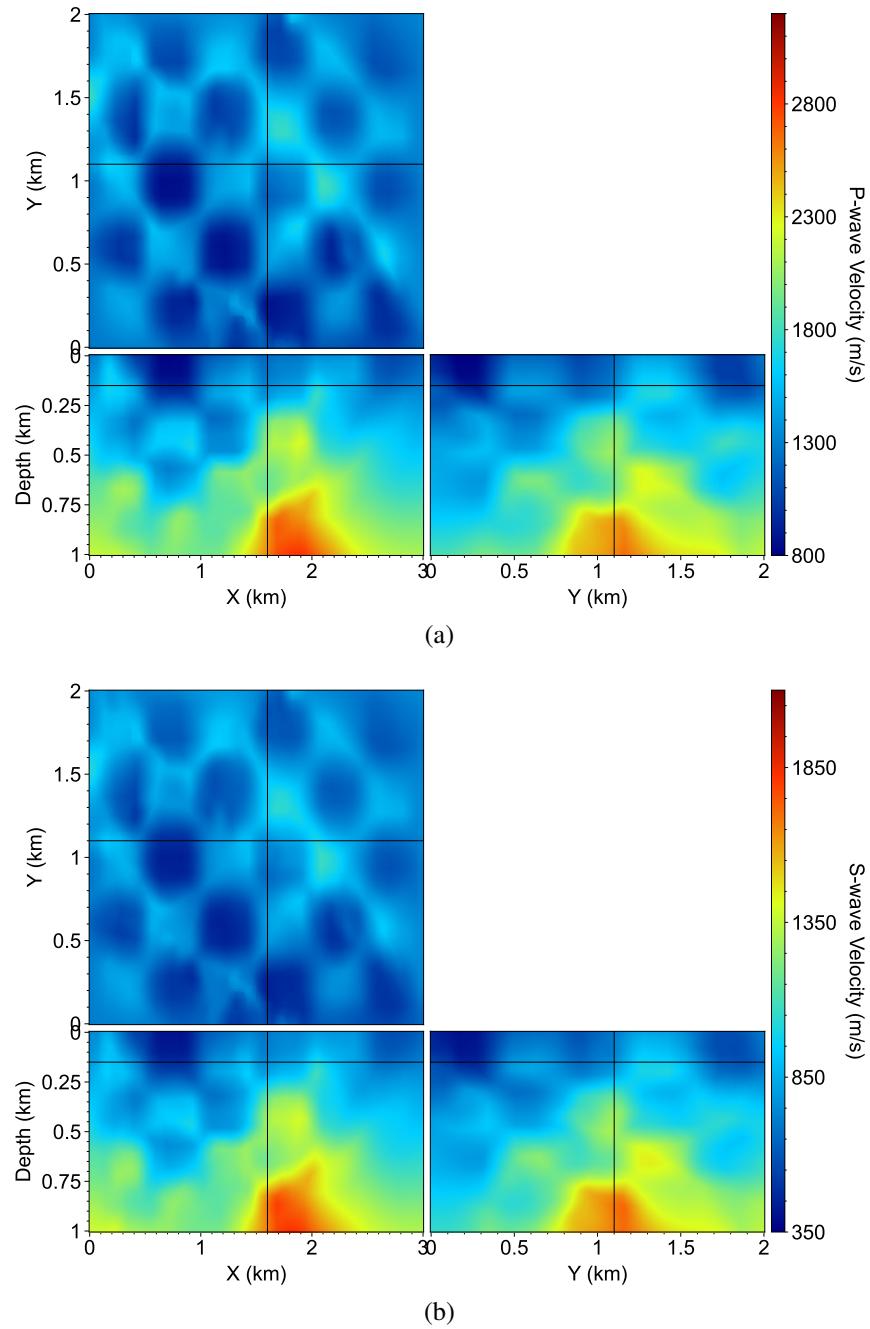


Figure 27: Inverted (a) v_p and (b) v_s models using DD-TLOC with model and source parameter regularization.

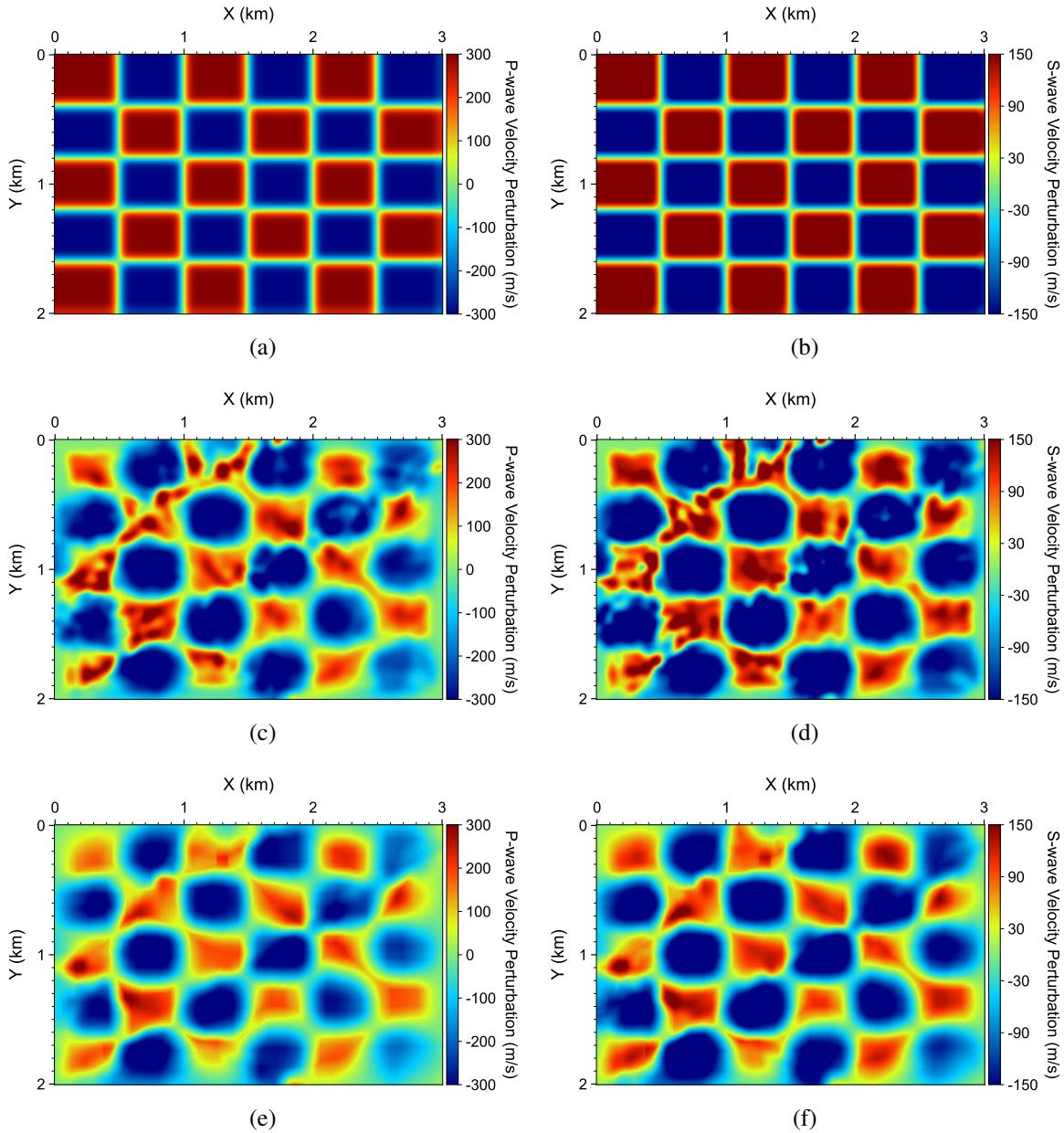


Figure 28: (a-b) Ground-truth Δv_p and Δv_s at a depth of 100 m, (c-d) inverted Δv_p and Δv_s by DD-TLOC without model parameter regularization, and (e-f) inverted Δv_p and Δv_s with model parameter regularization.

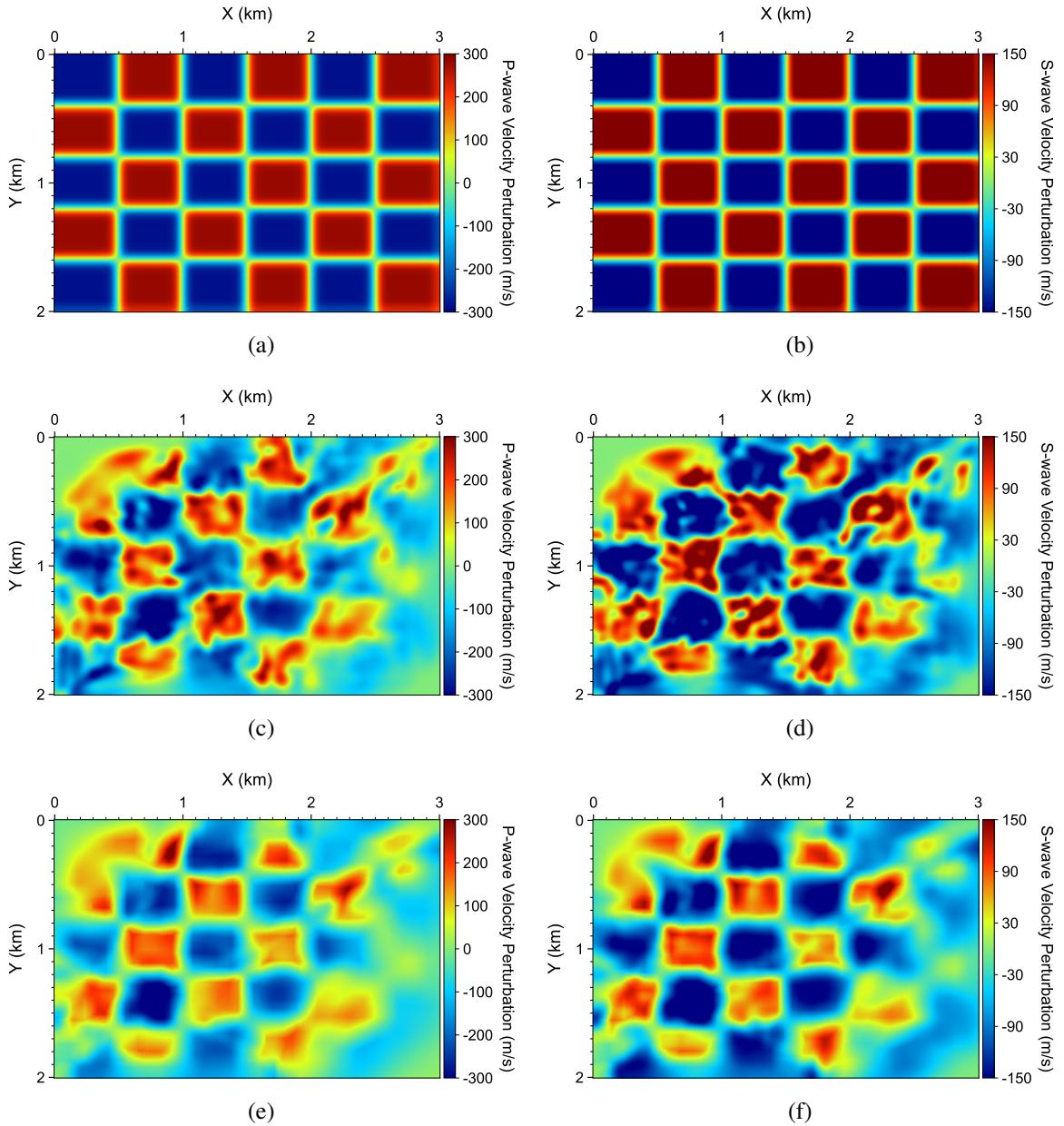


Figure 29: (a-b) Ground-truth Δv_p and Δv_s at a depth of 340 m, (c-d) inverted Δv_p and Δv_s by DD-TLOC without model parameter regularization, and (e-f) inverted Δv_p and Δv_s with model parameter regularization.

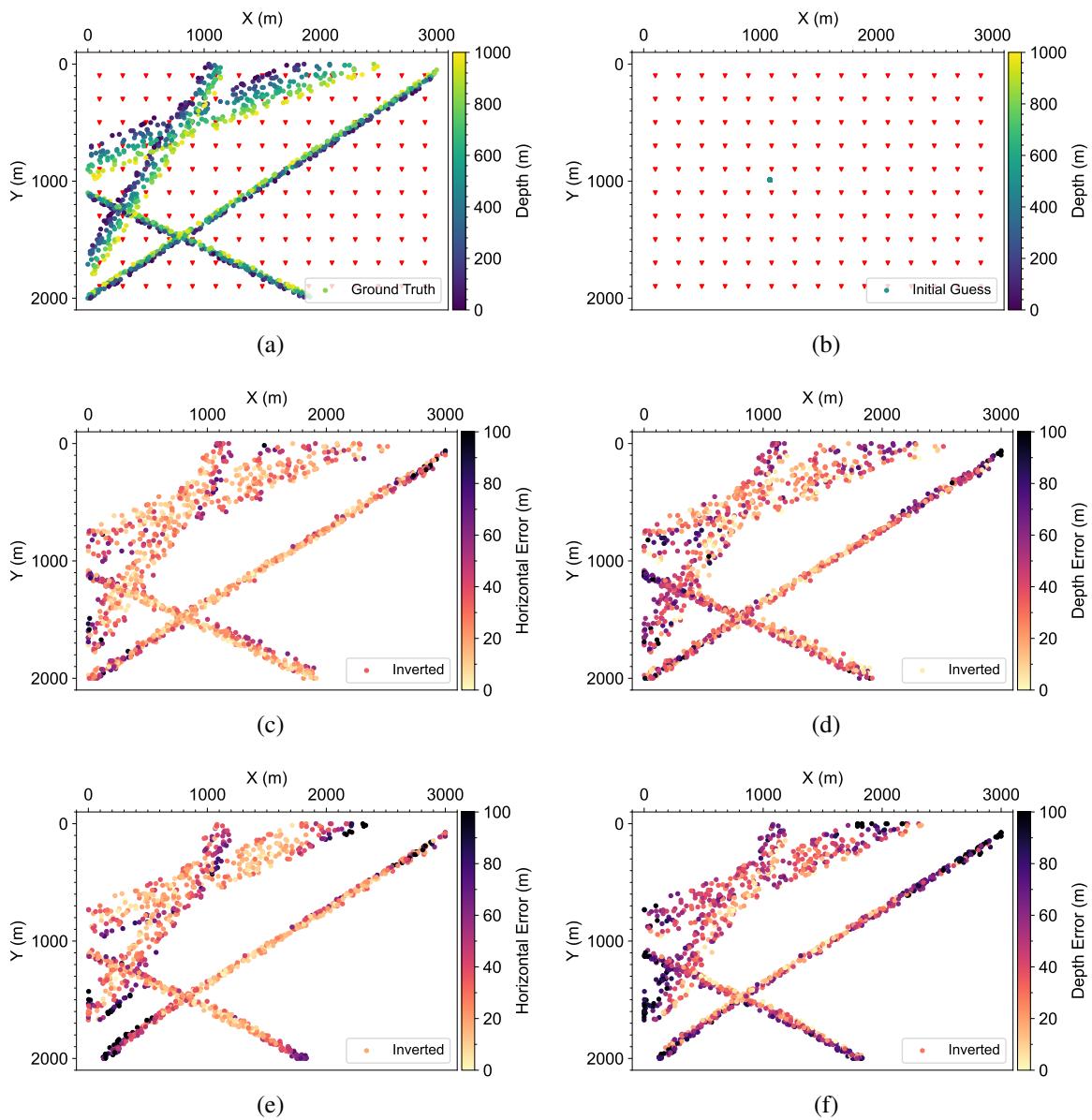


Figure 30: Map views of (a) ground-truth source locations colored by their depth, (b) initial source locations, (c-d) inverted source locations colored by horizontal/depth errors at the 50th iteration without ML-based source parameter regularization, and (e-f) inverted source locations colored by horizontal/depth errors with ML-based source parameter regularization.

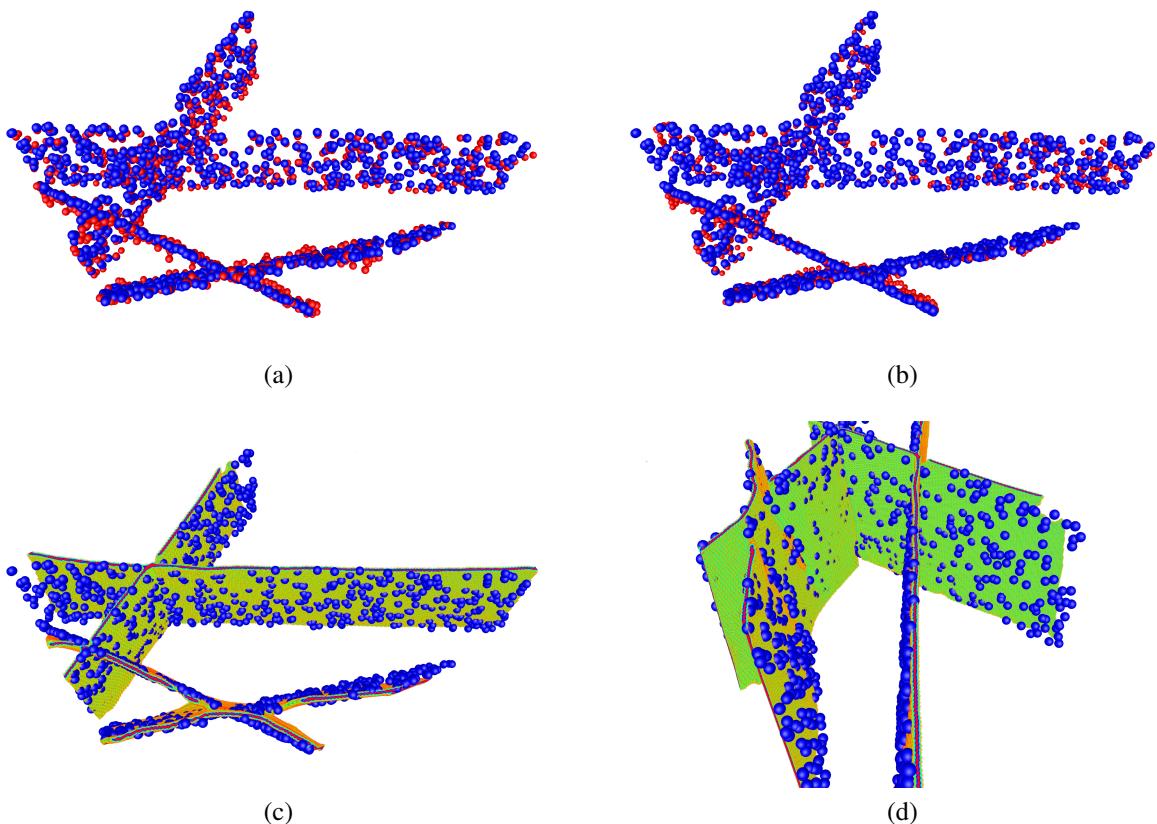


Figure 31: (a) A comparison between ground-truth source locations (blue balls) and DD-TLOC-inverted source locations without source parameter regularization (red balls). (b) A similar comparison with that in Panel (a) but the red balls represent the source locations inverted by DD-TLOC with source parameter regularization. (c-d) 3D views of the ground-truth source locations (blue balls) and the faults inferred and refined using our multitask NNs in the 50th iteration.

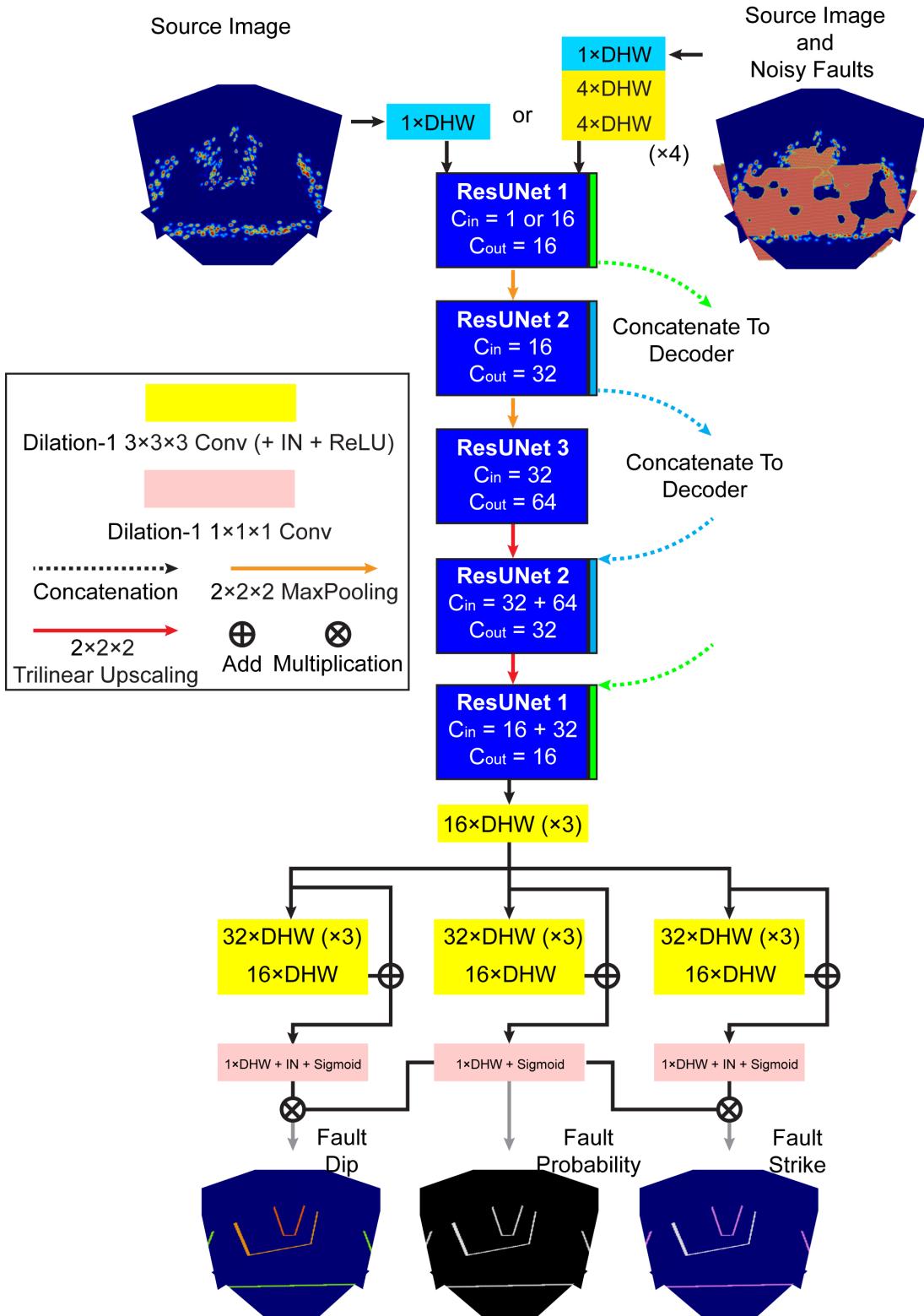


Figure 32: Architecture of our multitask fault inference and refinement NNs for inferring/refining fault attributes from a source image.