

The auxiliary function method for waveform based earthquake location



Jing Chen^a, Hao Jing^b, Ping Tong^{c,d}, Hao Wu^{a,*}, Dinghui Yang^a

^a Department of Mathematical Sciences, Tsinghua University, Beijing, 100084, China

^b Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China

^c Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore

^d Asian School of the Environment, Nanyang Technological University, Singapore

ARTICLE INFO

Article history:

Received 3 February 2019

Received in revised form 29 October 2019

Accepted 31 March 2020

Available online 7 April 2020

Keywords:

Computational seismology

Inverse theory

Waveform inversion

Earthquake location

ABSTRACT

This paper introduces the auxiliary function method, a novel, fast and simple approach for waveform based earthquake location. From any initial hypocenter and origin time, we can construct the auxiliary function, whose zero set contains the real earthquake hypocenter and the origin time. This translates the earthquake location problem into the problem of searching the common zeros of the auxiliary functions.

The computational cost of constructing the auxiliary functions is close to the cost of one single iteration of the traditional iterative method. And the cost of searching the common zeros of the auxiliary functions is almost negligible. Thus, the overall cost of this new method is significantly less than that of the iterative methods. Moreover, there is only one common zero point of the auxiliary functions in most practical situations. This means that the new method only requires one round of calculation to obtain an accurate earthquake hypocenter and origin time from arbitrary initial values. According to our numerical tests, even for large data noise, the method can still achieve good location results.

© 2020 Elsevier Inc. All rights reserved.

1. Introduction

In this work, we present a novel approach to solving the nonlinear optimization problem with PDE constraints [26] to determine the real earthquake hypocenter ξ_T and the origin time τ_T

$$(\xi_T, \tau_T) = \underset{\xi, \tau}{\operatorname{argmin}} \sum_{r \in \mathcal{R}} \chi_r(\xi, \tau), \quad (1.1)$$

in which $\chi_r(\xi, \tau)$ is the misfit function

$$\chi_r(\xi, \tau) = \frac{\int_0^{t_f} |d_r(t) - s(\eta_r, t)|^2 dt}{2 \int_0^{t_f} |d_r(t)|^2 dt}, \quad (1.2)$$

* Corresponding author.

E-mail addresses: jing-che16@mails.tsinghua.edu.cn (J. Chen), jinghao0320@gmail.com (H. Jing), tongping@ntu.edu.sg (P. Tong), hwu@tsinghua.edu.cn (H. Wu), ydh@mail.tsinghua.edu.cn (D. Yang).

for the r -th receiver. The time-series $d_r(t)$ is the real earthquake signal which occurred at (ξ_T, τ_T) and was recorded at receiver r . The simulated time interval is $[0, t_f]$. Here η_r denotes the location of the r -th receiver. The set \mathcal{R} contains all of the receivers that we use for inversion. And the synthetic earthquake signal $s(\mathbf{x}, t)$ is corresponding to the initial hypocenter ξ and the initial origin time τ . For model simplicity, they can be regarded as the solutions

$$d_r(t) = u(\eta_r, t; \xi_T, \tau_T), \quad s(\mathbf{x}, t) = u(\mathbf{x}, t; \xi, \tau), \quad (1.3)$$

of the following acoustic wave equation

$$\frac{\partial^2 u(\mathbf{x}, t; \xi, \tau)}{\partial t^2} = \nabla \cdot (c^2(\mathbf{x}) \nabla u(\mathbf{x}, t; \xi, \tau)) + f(t - \tau) \delta(\mathbf{x} - \xi), \quad \mathbf{x}, \xi \in \Omega, \quad (1.4)$$

with initial-boundary conditions

$$u(\mathbf{x}, 0; \xi, \tau) = \partial_t u(\mathbf{x}, 0; \xi, \tau) = 0, \quad \mathbf{x} \in \Omega, \quad (1.5)$$

$$\mathbf{n} \cdot (c^2(\mathbf{x}) \nabla u(\mathbf{x}, t; \xi, \tau)) = 0, \quad \mathbf{x} \in \partial\Omega. \quad (1.6)$$

Here $c(\mathbf{x})$ is the wave speed. The simulated domain $\Omega \subset \mathbb{R}^d$, d is the dimension of the problem and \mathbf{n} is the unit outer normal vector to the boundary $\partial\Omega$. The point source hypothesis $\delta(\mathbf{x} - \xi)$ is considered in (1.4) since we investigate the situation that the temporal and spatial scales of the seismic wave propagated are large enough compared to the scales of seismic rupture [1,15]. The source time function has the form of Ricker wavelet

$$f(t) = A \left(1 - 2\pi^2 f_0^2 t^2 \right) e^{-\pi^2 f_0^2 t^2}, \quad (1.7)$$

in which f_0 is the dominant frequency, and A is the normalization factor. Since we focus on the earthquake location problem on a large computational domain Ω , we can simply consider the reflection boundary condition (1.6). The other type of boundary conditions, e.g., the perfectly matched layer absorbing boundary condition [11], can be similarly considered here without any difficulty.

The equations (1.1)-(1.7) provide a mathematical model to the waveform based earthquake location problem, which is a fundamental problem [20] with various applications in seismology [12,18,24]. Traditionally, the earthquake location problem is solved within the framework of the ray theory, see for example [5–7,16,19]. However, the earthquake location results are not satisfactory since the ray theory is low accuracy when the seismic wavelength is not small enough compared to the scale of wave propagation region [4,9,17,27]. Thus, it is necessary to develop the waveform based earthquake location method. This direction is becoming more and more popular in recent years [2,8,10,14,22,26,29], together with the fast increase in computational power.

In general, the optimization problem (1.1)-(1.4) can be solved by iterative methods [23,26]. This approach requires that the initial value should be close enough to the global optimization solution. In particular, due to the highly singular nature of the delta function $\delta(\mathbf{x} - \xi)$ in the wave equation (1.4), the convergence domain could be very small. Although this issue has been studied in [26], the convergence domain is still restricted. The other important problem is the computational cost of the iterative method. To update the value of hypocenter and origin time, the sensitivity kernel needs to be computed, which costs several times wave equation computations. Considering the number of iterations and initial values, the total computational cost could be very large. Thus, it is difficult to meet the demand for practical applications, e.g., the real-time earthquake location and many earthquakes relocation.

In this study, we would present an accurate and efficient algorithm to solve the difficulty mentioned above. The proposed algorithm is based on a set of newly introduced auxiliary functions $\Xi_r(\xi, v)$, $r \in \mathcal{R}$. We will prove that

$$\Xi_r(\xi_T, \tau_T) = 0, \quad \forall r \in \mathcal{R}.$$

This shows that the real earthquake location ξ_T and origin time τ_T is one of the common zeros of the auxiliary functions. Thus, the earthquake location problem is translated into the problem of searching the common zeros of the auxiliary functions. The computational cost of searching the common zeros is almost negligible since the value of the auxiliary functions has been already obtained during the construction of these auxiliary functions. According to our numerical experiments, there is only one common zero point of the auxiliary functions, which reduces the uncertainty of the practical earthquake location problems. Moreover, the cost of constructing the auxiliary functions is close to the cost of one single iteration of the traditional iterative method. From the above discussion, we can believe the total computational cost of the auxiliary function method is much less than that of the traditional iterative methods [14,23,26]. It could be very competitive for practical real-time earthquake location problems and many earthquakes relocation problems.

The paper is organized as follows. In Section 2, we prove the main theorem and propose the algorithms. The numerical experiments are presented to illustrate the features and highlights of the auxiliary function method in Section 3. In Section 4, we make some conclusive remarks.

2. The auxiliary function method

2.1. The main theorem

We begin this section with the following theorem. In fact, all the discussions in this paper depend on this theorem.

Theorem 1. For any given initial hypocenter ξ and origin time τ , define the auxiliary functions

$$\Xi_r(\xi, \nu) = 2\chi_r(\xi, \tau) - \int_0^{t_f} f(t - \nu) w_r(\xi, t) - f(t - \tau) w_r(\xi, t) dt, \quad \forall r \in \mathcal{R}, \quad (2.1)$$

in which $\chi_r(\xi, \tau)$ has been given in (1.2), and $w_r(\mathbf{x}, t)$ satisfies the adjoint equation with terminal-boundary conditions

$$\begin{cases} \frac{\partial^2 w_r(\mathbf{x}, t)}{\partial t^2} = \nabla \cdot (c^2(\mathbf{x}) \nabla w_r(\mathbf{x}, t)) + \frac{d_r(t) - s(\eta_r, t)}{\int_0^{t_f} |d_r(t)|^2 dt} \delta(\mathbf{x} - \eta_r), & \mathbf{x} \in \Omega, \\ w_r(\mathbf{x}, t_f) = \frac{\partial w_r(\mathbf{x}, t_f)}{\partial t} = 0, & \mathbf{x} \in \Omega, \\ \mathbf{n} \cdot (c^2(\mathbf{x}) \nabla w_r(\mathbf{x}, t)) = 0, & \mathbf{x} \in \partial\Omega. \end{cases} \quad (2.2)$$

Then, for real earthquake hypocenter ξ_T and origin time τ_T , we have

$$\Xi_r(\xi_T, \tau_T) = 0, \quad \forall r \in \mathcal{R}. \quad (2.3)$$

Proof. Let us first define the difference function

$$\delta s(\mathbf{x}, t) = u(\mathbf{x}, t; \xi_T, \tau_T) - u(\mathbf{x}, t; \xi, \tau).$$

According to (1.3), it follows

$$\delta s(\eta_r, t) = d_r(t) - s(\eta_r, t), \quad \forall r \in \mathcal{R}.$$

And the difference function $\delta s(\mathbf{x}, t)$ satisfies the wave equation with initial-boundary conditions

$$\begin{cases} \frac{\partial^2 \delta s(\mathbf{x}, t)}{\partial t^2} = \nabla \cdot (c^2(\mathbf{x}) \nabla \delta s(\mathbf{x}, t)) + f(t - \tau_T) \delta(\mathbf{x} - \xi_T) - f(t - \tau) \delta(\mathbf{x} - \xi), & \mathbf{x} \in \Omega, \\ \delta s(\mathbf{x}, 0) = \frac{\partial \delta s(\mathbf{x}, 0)}{\partial t} = 0, & \mathbf{x} \in \Omega, \\ \mathbf{n} \cdot (c^2(\mathbf{x}) \nabla \delta s(\mathbf{x}, t)) = 0, & \mathbf{x} \in \partial\Omega. \end{cases} \quad (2.4)$$

Multiply the wave function $w_r(\mathbf{x}, t)$ given in (2.2), integrate it on $\Omega \times [0, t_f]$ and use the integration by parts, we obtain

$$\int_0^{t_f} \int_{\Omega} \frac{\partial^2 w_r}{\partial t^2} \delta s d\mathbf{x} dt = \int_0^{t_f} \int_{\Omega} \delta s \nabla \cdot (c^2 \nabla w_r) d\mathbf{x} dt + \int_0^{t_f} f(t - \tau_T) w_r(\xi_T, t) - f(t - \tau) w_r(\xi, t) dt. \quad (2.5)$$

On the other hand, the misfit function $\chi_r(\xi, \tau)$ in (1.2) can be rewritten as

$$\chi_r(\xi, \tau) = \frac{\int_0^{t_f} (d_r(t) - s(\eta_r, t)) \delta s(\eta_r, t) dt}{2 \int_0^{t_f} |d_r(t)|^2 dt} = \frac{\int_0^{t_f} \int_{\Omega} (d_r(t) - s(\eta_r, t)) \delta s(\mathbf{x}, t) \delta(\mathbf{x} - \eta_r) d\mathbf{x} dt}{2 \int_0^{t_f} |d_r(t)|^2 dt}. \quad (2.6)$$

Multiplying both sides of the above equation by 2, and adding equation (2.5), we get

$$2\chi_r(\xi, \tau) = \int_0^{t_f} f(t - \tau_T) w_r(\xi_T, t) - f(t - \tau) w_r(\xi, t) dt.$$

This completes the proof. \square

According to the theorem, the problem of determining the real earthquake hypocenter ξ_T and origin time τ_T is translated into the problem of searching the common zeros of the auxiliary functions

$$\Xi_r(\xi, \nu) = 0, \quad \forall r \in \mathcal{R}. \quad (2.7)$$

Assume the solution of the above equations is unique, we only need one round of operation, which is constructing the auxiliary functions $\Xi_r(\xi, \nu)$ and searching the common zero of the equation (2.7), to obtain the real earthquake hypocenter

ξ_T and origin time τ_T . In the following subsections, we will show that the cost of constructing the auxiliary functions is comparable to the cost of an iteration of the traditional method and the cost of searching the common zero is minor. Thus, the theorem leads to a highly efficient non-iterative numerical method.

Moreover, this theorem holds for any initial earthquake hypocenter ξ and origin time τ , which overcomes the limitation of the traditional iterative methods that the initial earthquake hypocenter ξ and origin time τ should be close enough to the real earthquake hypocenter ξ_T and origin time τ_T . In this sense, the obtained numerical method for the earthquake location problem has global convergence for the optimization problem (1.1).

Remark 1. In practice, we prefer to solve the equations (2.7) in the least square sense, i.e.

$$(\xi_T, \tau_T) = \underset{\xi, \nu}{\operatorname{argmin}} \Gamma(\xi, \nu), \quad (2.8)$$

in which

$$\Gamma(\xi, \nu) = \sum_{r \in \mathcal{R}} \Xi_r^2(\xi, \nu). \quad (2.9)$$

Remark 2. The theorem does not guarantee the uniqueness of the solution. This may lead to incorrect inversion result. Fortunately, more constraints $r \in \mathcal{R}$ may improve the uniqueness of the solution. According to the numerical experiments, we are not suffering from the problem of uniqueness.

Remark 3. For the elastic wave equation, we can obtain similar results as Theorem 1. The theorem and proof can be found in the appendix.

2.2. Algorithm

According to the previous theorem and discussions, we present the detailed implementation as follows:

Algorithm 1 (*The auxiliary function method*).

- Initialization. Given a searching domain Ω_s and a searching time interval I_s , we wish that $\xi_T \in \Omega_s$ and $\tau_T \in I_s$. Select a mesh size $h > 0$ and a time step $\sigma > 0$.
- Discretization. Select a uniform or quasi-uniform grid $\Omega_h \subset \Omega_s$ of mesh size h . Select a uniform or quasi-uniform time division $I_\sigma \subset I_s$ of time step σ .
- Forward and Backward evolution. Given the initial hypocenter ξ and origin time τ , compute the wave equation (1.4)-(1.6) to get

$$s(\eta_r, t) = u(\eta_r, t; \xi, \tau),$$

and the misfit function $\chi_r(\xi, \tau)$, $\forall r \in \mathcal{R}$ in (1.2). Next, compute the adjoint wave equations (2.2) to get $w_r(x, t)$ for $r \in \mathcal{R}$.

- Construction. Evaluate all the values of the auxiliary functions $\Xi_r(\xi, \nu)$ on the mesh size $\xi \in \Omega_h$, time division $\nu \in I_\sigma$ and $r \in \mathcal{R}$ by using equation (2.1). Thus, $\Gamma(\xi, \nu)$ defined in (2.9) on the mesh grid and time division $(\xi, \nu) \in \Omega_h \times I_\sigma$ is directly obtained.
- Output. Finally, we can easily obtain the approximated solution of the optimization problem (2.8)

$$(\xi_*, \tau_*) = \underset{\xi \in \Omega_h, \nu \in I_\sigma}{\operatorname{argmin}} \Gamma(\xi, \nu),$$

by direct search. Output (ξ_*, τ_*) and stop. \square

The auxiliary function method relies on three components which have not been specified yet, namely, the numerical solve of the forward and backward wave equations (1.4), (2.2), the numerical integration for computing the auxiliary function (2.1), and the selection of the mesh grid Ω_h and time division I_σ . We will specify these issues in the later part of this paper.

When the algorithm terminates, we have an approximation of the earthquake hypocenter and the origin time, that is the numerical solution (ξ_*, τ_*) . However, the above approximation does not hold in an extreme situation. This is the case where the real earthquake hypocenter and the origin time is not in the searching domain and searching time interval, i.e.,

$$\xi_T \notin \Omega_s \text{ or } \tau_T \notin I_s.$$

However, this situation can also be easily avoided. We need to simply enlarge the searching domain Ω_s and searching time interval I_s . Moreover, the resulting increase in computational cost is minor.

Remark 4. We further discuss the possibility of the extreme situation

$$\xi_T \notin \Omega_s \quad \text{or} \quad \tau_T \notin I_s.$$

It should be emphasized that the searching domain Ω_s and the searching time interval I_s are subsets of the simulated domain Ω and the simulated time interval $[0, t_f]$:

$$\Omega_s \subseteq \Omega, \quad I_s \subseteq [0, t_f].$$

Obviously, the earthquake hypocenter and the origin time must be in the simulated domain Ω and the simulated time interval $[0, t_f]$,

$$\xi_T \in \Omega, \quad \tau_T \in [0, t_f],$$

otherwise, any method will not converge correctly. The computational cost can be saved by selecting a smaller searching domain Ω_s and searching time interval I_s . However, this choice is based on our prior estimates of the earthquake hypocenter and the origin time. We can just consider the situation

$$\Omega_s = \Omega, \quad I_s = [0, t_f],$$

to avoid suffering from the extreme situation. The extra computational cost of this consideration is minor.

We now discuss the computational cost of Algorithm 1, which consists of three parts:

1. The simulation of the wave equation (1.4)-(1.6) and the adjoint wave equations (2.2). It needs to solve the wave equation for $\#\mathcal{R} + 1$ times. Here $\#\mathcal{R}$ denotes the number of elements in the set \mathcal{R} .
2. The computation of the misfit functions $\chi_r(\xi, \tau)$ in (1.2). It needs to compute the 1-d integral for $2\#\mathcal{R}$ times.
3. The computation of the auxiliary functions $\Xi_r(\xi, \nu)$ in (2.1). It needs to compute the 1-d integral for $\#\mathcal{R} \cdot (\#\Omega_h \cdot \#I_\sigma + 1)$ times. \square

Adding these three parts, the total computational cost is $\#\mathcal{R} + 1$ times wave equation computations plus $\#\mathcal{R} \cdot (\#\Omega_h \cdot \#I_\sigma + 3)$ times 1-d integral calculations. Below, we will illustrate that the cost of $\#\mathcal{R} \cdot (\#\Omega_h \cdot \#I_\sigma + 3)$ times 1-d integral calculations is comparable to the cost of one-time wave equation computation. Thus, the overall cost of Algorithm 1 is comparable to $\#\mathcal{R} + 2$ times wave equation computations.

To obtain the misfit functions $\chi_r(\xi, \tau)$ and the auxiliary functions $\Xi_r(\xi, \nu)$, we need to calculate one dimensional integrals for $\#\mathcal{R} \cdot (\#\Omega_h \cdot \#I_\sigma + 3)$ times. However, the support of integrand functions is small. This implies a very small cost of 1-d integral computations. Moreover, the mesh size and the time step of the searching domain and the searching time interval can be coarser than that of the wave field simulations. Thus, we believe that the total cost of this part is comparable to the cost of one-time wave equation computations.

For the iterative method, the overall computational cost is about $(\#\mathcal{R} + 1) \cdot N_{\text{iter}} \cdot N_{\text{init}}$ times wave equation computations. Here N_{iter} denotes the average number of iterations for each initial data and N_{init} denotes the average number of initial data for each problem. In contrast, an accurate approximation of the earthquake hypocenter and origin time can be obtained by the auxiliary function method with only one iteration and one initial data. Thus, the auxiliary function method is much more efficient. In the later section, we will illustrate this through several numerical experiments.

2.3. The auxiliary function preprocessing method

In order to achieve a high location accuracy in Algorithm 1, the mesh size h and time step σ should be small enough. This requires a large computational cost of the Construction step. In extreme cases, the total cost of the auxiliary function method may exceed the iterative method. To avoid this situation, we can first get a reasonable approximation of the earthquake hypocenter and origin time by using the auxiliary function method. After that, we will take this approximation as the initial value of the iterative method to get a more accurate approximation. The sketch of the algorithm is as follows:

Algorithm 2 (*The auxiliary function preprocessing method*).

- Initialization. Given a searching domain Ω_s and a searching time interval I_s .
- Preprocessing. Execute Algorithm 1 on a coarse mesh grid $\Omega_h \subset \Omega_s$ and time division $I_\sigma \subset I_s$ to obtain (ξ_*, τ_*) .
- Iteration. Execute the iterative method, see e.g. in [2,26], for the inverse problem (1.1)-(1.7) with initial hypocenter and origin time (ξ_*, τ_*) to get (ξ_A, τ_A) . \square

In the above algorithm, the auxiliary function method (Algorithm 1) is used as a preprocessing method to obtain an accurate initial hypocenter and origin time for the iterative. Thus, we refer this algorithm as the auxiliary function preprocessing

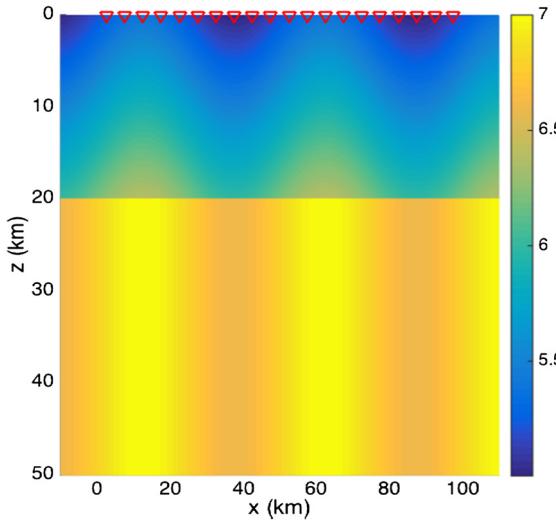


Fig. 1. Illustration of two-layer velocity model. The red triangles indicate the receivers. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

method. When the above algorithm terminates, we have a more accurate approximation of the earthquake hypocenter and origin time, that is (ξ_A, τ_A) .

The total computational cost of this algorithm is about

$$(\#\mathcal{R} + 2) + (\#\mathcal{R} + 1) \cdot N'_{iter} = (\#\mathcal{R} + 1) \cdot (N'_{iter} + 1) + 1,$$

times wave equation computations. Here N'_{iter} denotes the average number of iterations of the iterative method. This number could be small, since the initial value (ξ_*, τ_*) is very close to the optimization solution (ξ_T, τ_T) . Overall, the auxiliary function preprocessing method is still more efficient than the iterative methods.

3. Numerical experiments

In this section, two examples are presented to demonstrate the efficiency and effectiveness of our method. In all the numerical examples, the finite difference schemes [3,13,28] are applied to solve the acoustic wave equation (1.4) with the initial condition (1.5). On the surface of the earth, we consider the reflection boundary condition (1.6). And the perfectly matched layer boundary condition [11] is used for the other boundaries within the earth. The point source $\delta(x - \xi)$ is numerically discretized as follows [25]:

$$\delta_h(x) = \begin{cases} \frac{1}{h} \left(1 - \frac{5}{4} |\frac{x}{h}|^2 - \frac{35}{12} |\frac{x}{h}|^3 + \frac{21}{4} |\frac{x}{h}|^4 - \frac{25}{12} |\frac{x}{h}|^5 \right), & |x| \leq h, \\ \frac{1}{h} \left(-4 + \frac{75}{4} |\frac{x}{h}| - \frac{245}{8} |\frac{x}{h}|^2 + \frac{545}{24} |\frac{x}{h}|^3 - \frac{63}{8} |\frac{x}{h}|^4 + \frac{25}{24} |\frac{x}{h}|^5 \right), & h < |x| \leq 2h, \\ \frac{1}{h} \left(18 - \frac{153}{4} |\frac{x}{h}| + \frac{255}{8} |\frac{x}{h}|^2 - \frac{313}{24} |\frac{x}{h}|^3 + \frac{21}{8} |\frac{x}{h}|^4 - \frac{5}{24} |\frac{x}{h}|^5 \right), & 2h < |x| \leq 3h, \\ 0, & |x| > 3h. \end{cases}$$

Here h is a numerical parameter which is related to the mesh size.

3.1. The two-layer velocity model

Consider the two-layer model in the bounded domain $\Omega = [-10 \text{ km}, 110 \text{ km}] \times [0 \text{ km}, 50 \text{ km}]$, the wave speed is

$$c(x, z) = \begin{cases} 5.2 + 0.05z + 0.2 \sin \frac{\pi x}{25}, & 0 \text{ km} \leq z \leq 20 \text{ km}, \\ 6.8 + 0.2 \sin \frac{\pi x}{25}, & z > 20 \text{ km}. \end{cases}$$

The unit is ‘km/s’. The computational time interval $I = [0, 25 \text{ s}]$. The dominant frequency of the earthquakes is $f_0 = 2 \text{ Hz}$. There are 20 equidistant receivers on the surface

$$\boldsymbol{\eta}_r = (x_r, z_r) = (5r - 2.5 \text{ km}, 0), \quad r = 1, 2, \dots, 20,$$

see Fig. 1 for illustration.

First, we test the auxiliary function preprocessing method (Algorithm 2) using 500 experiments. The searching domain is $\Omega_s = [0, 100 \text{ km}] \times [0, 40 \text{ km}]$, and the searching time interval is $I_s = [0, 25 \text{ s}]$. The mesh sizes for the searching grid are

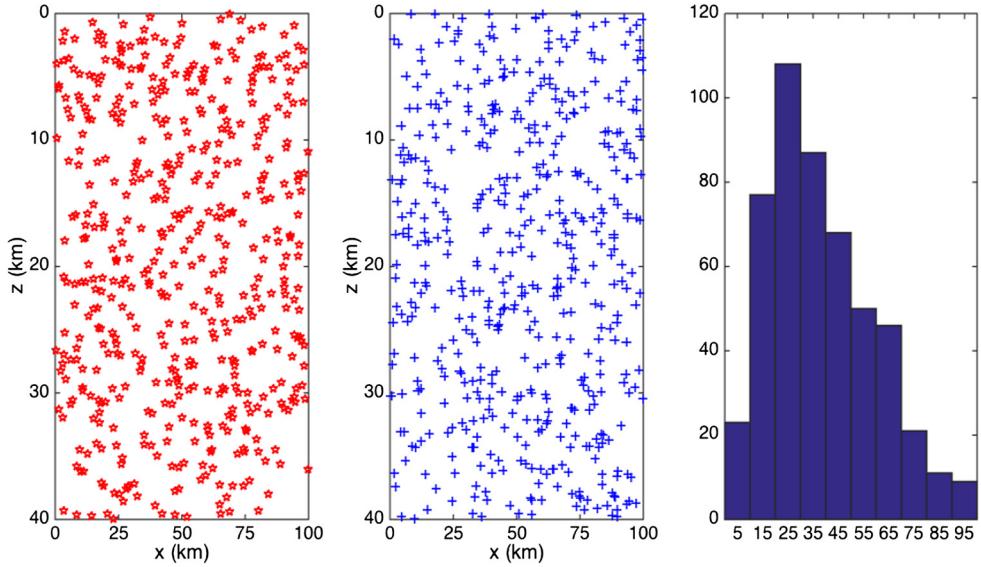


Fig. 2. The two-layer velocity model. Left: the spatial distribution of the real earthquake hypocenter ξ_T^i ; Middle: the spatial distribution of the initial earthquake hypocenter ξ^i ; Right: the distance distribution histogram between the real and the initial earthquake hypocenter d^i .

Table 1

The two-layer velocity model. Convergent results for the auxiliary function preprocessing method (AFPM) and the iterative method (IM).

	Correct convergence	Diverge	Error convergence	Total
AFPM	500	0	0	500
IM	117	355	28	500

$h_x = 0.5$ km and $h_z = 0.4$ km. The time step for the searching time interval is $\sigma = 0.1$ s. As a comparison, we also compute these experiments by the iterative method proposed in [26].

The experiments are designed as follows: the real and initial earthquake hypocenter ξ_T^i , ξ^i are both uniformly distributed over $[0, 100\text{ km}] \times [0, 40\text{ km}]$, the real and initial original time τ_T^i , τ^i are both uniformly distributed over $[5\text{ s}, 20\text{ s}]$. Their spatial distribution and the histogram of the distance between the real and the initial hypocenter

$$d^i = \|\xi_T^i - \xi^i\|_2,$$

are presented in Fig. 2.

In the auxiliary function preprocessing method, we randomly select five receivers for inversion, e.g., $r = 3, 5, 9, 14, 18$. In Table 1, we can see the convergent results of the two methods. From which, we can conclude that the auxiliary function preprocessing method converges globally here. For the iterative method proposed in [26], only 23.4% experiments converge. We have to remark that the range of convergence of the iterative method has been enlarged by several tens of time. In contrast, the auxiliary function preprocessing method has an absolute advantage in terms of convergence.

In Fig. 3, we output the histogram of the iterations and the computational time for the two methods. The mean and standard deviation of iterations and computational time for the two methods are also presented in Table 2. In particular, we present the mean and standard deviation of the time consuming for the preprocessing step and the single iteration step of the auxiliary function preprocessing method in Table 3. It is obvious that the time consuming of the preprocessing step and the single iteration step are almost the same. Thus, we consider the preprocessing step to be an iteration step. Taking account of all the above issues, the total computational cost of the iterative method is about

$$\frac{500}{117} \times \frac{1523}{689} \approx 9.45 \text{ times}$$

of the auxiliary function preprocessing method. Thus, we can conclude that the auxiliary function preprocessing method is more efficient than the iterative method. This agrees with the theoretical discussions in the previous section.

Next, two examples are specifically presented. The parameters are selected as follows:

$$(i) \xi_T = (90.36 \text{ km}, 35.67 \text{ km}), \tau_T = 10 \text{ s}, \xi = (18.23 \text{ km}, 13.13 \text{ km}), \tau = 15.5 \text{ s};$$

$$(ii) \xi_T = (87.252 \text{ km}, 8.842 \text{ km}), \tau_T = 10 \text{ s}, \xi = (12.75 \text{ km}, 32.87 \text{ km}), \tau = 17.4 \text{ s}.$$

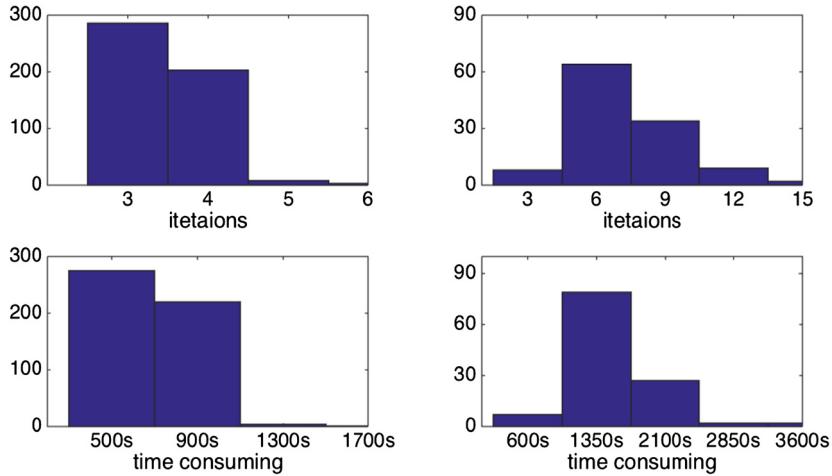


Fig. 3. The two-layer velocity model. The histogram of the iterations and computational time for the two methods. Up: the histogram of iterations; Down: the histogram of computational time; Left: the auxiliary preprocessing method; Right: the iterative method.

Table 2

The two-layer velocity model. The Mean(M) and Standard Deviation(SD) of iterations and computational time for the auxiliary function preprocessing method (AFPM) and the iterative method (IM).

	Iterations		Computational time	
	M	SD	M	SD
AFPM	3.46	0.61	689 s	132.7 s
IM	8.33	2.35	1523 s	501.0 s

Table 3

The two-layer velocity model. The mean and standard deviation of the time consuming for the preprocessing step and the single iteration step of the auxiliary function preprocessing method.

	Mean	Standard deviation
The preprocessing step	224.9 s	13.2 s
The single iteration step	222.7 s	11.8 s

In Figs. 4–5, we output the cross-sections of the functions $\Gamma(\xi, v)$ in (2.9), from which we can see the global minimum is unique in both cases. The convergent history of the auxiliary preprocessing method is also illustrated in Fig. 6. Here, we randomly select five receivers for inversion, e.g., $r = 3, 5, 9, 14, 18$. We can see that the global minimum of the function $\Gamma(\xi, v)$ is very close to the optimization solution of (1.1). Thus, when the accuracy requirement is not high, the solution of the auxiliary function method is directly applicable. On the other hand, when the accuracy requirement is high, the solution of the auxiliary function method can provide excellent initial values for the iterative methods.

At last, we test the influence of the noise. The same parameters (i) and (ii) are selected here. The real earthquake signal can be regarded as

$$d_r(t) = u(\eta_r, t; \xi_T, \tau_T) + N_r(t),$$

with $N_r(t)$ is subject to the normal distribution with mean $\mu = 0$ and the standard deviation

$$\sigma = R \times \max_t |u(\eta_r, t; \xi_T, \tau_T)|.$$

Here R denotes the ratio, which will be selected as 10%, 15%, 20%, and 25% respectively. The real earthquake signal with noise $d_r(t)$ and the noise-free signal $u(\eta_r, t; \xi_T, \tau_T)$ are illustrated in Fig. 7. In order to reduce the impact of noise, we can select a time window that contains the main part of $u(\eta_r, t; \xi_T, \tau_T)$. In Table 4, the mean and standard deviation of the errors between the location results (ξ_*, τ_*) computed via the auxiliary function method (Algorithm 1) and the exact solution (ξ_T, τ_T) are presented. The auxiliary function preprocessing method (Algorithm 2) is not considered here since the iterative method may fail even for small ratio R . For each parameter group and ratio R , we test the algorithm with ten different noises. We note that all the standard deviations are zero. This implies that all the tests converge to the same solution. They are

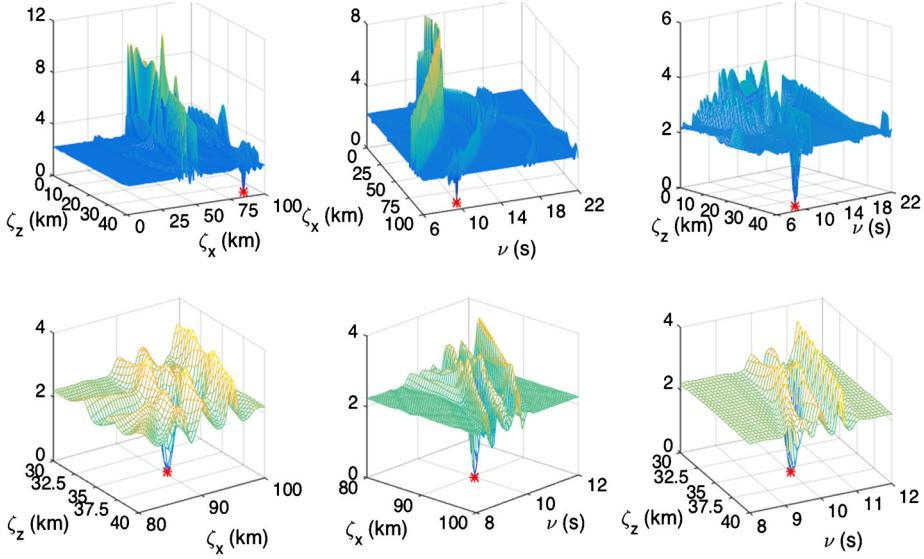


Fig. 4. The two-layer velocity model, case (i). The $\xi_x - \xi_z$ (Left, $\nu = 10$ s), $\nu - \xi_x$ (Middle, $\xi_z = 35.67$ km) and $\nu - \xi_z$ (Right, $\xi_x = 90.36$ km) cross-sections of the function $\Gamma(\xi, \nu)$. Up for large-scale image and Down for zoom-in image.

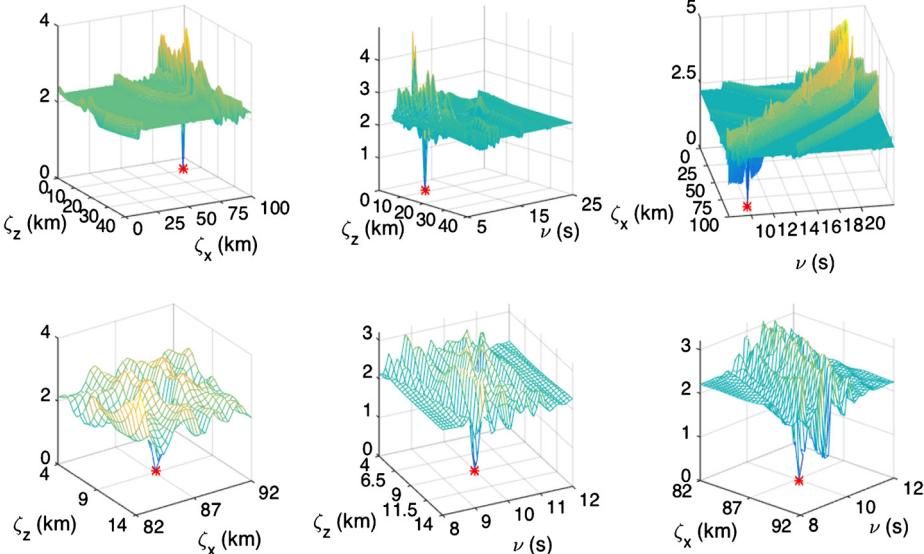


Fig. 5. The two-layer velocity model, case (ii). The $\xi_x - \xi_z$ (Left, $\nu = 10$ s), $\nu - \xi_x$ (Middle, $\xi_z = 8.842$ km) and $\nu - \xi_z$ (Right, $\xi_x = 87.252$ km) cross section plan of function $\Gamma(\xi, \nu)$. Up for large-scale image and Down for zoom-in image.

(i) $\xi_* = (90.40 \text{ km}, 35.60 \text{ km})$, $\tau_* = 10.01 \text{ s}$;

(ii) $\xi_* = (87.20 \text{ km}, 8.80 \text{ km})$, $\tau_* = 9.99 \text{ s}$.

This is because we are using the same mesh grid Ω_h and time division I_σ in the program. It should be noted that the numerical results may become better or worse as the mesh grid and time division change. However, the location errors are always in the same order of magnitude, which is caused by the noise. We can also observe that the algorithm failed when $R = 25\%$ in the parameter group (i). But the algorithm works for all the ratios in the parameter group (ii). Thus, we tend to believe that the algorithm can achieve success when R below 20% in this example. This is a huge advantage of the auxiliary function method over the traditional iterative methods. Accordingly to our numerical tests, the traditional iterative methods fail when R reaches 10% and above. Moreover, the computational cost of the auxiliary function method is much less than that of the conventional iterative methods.

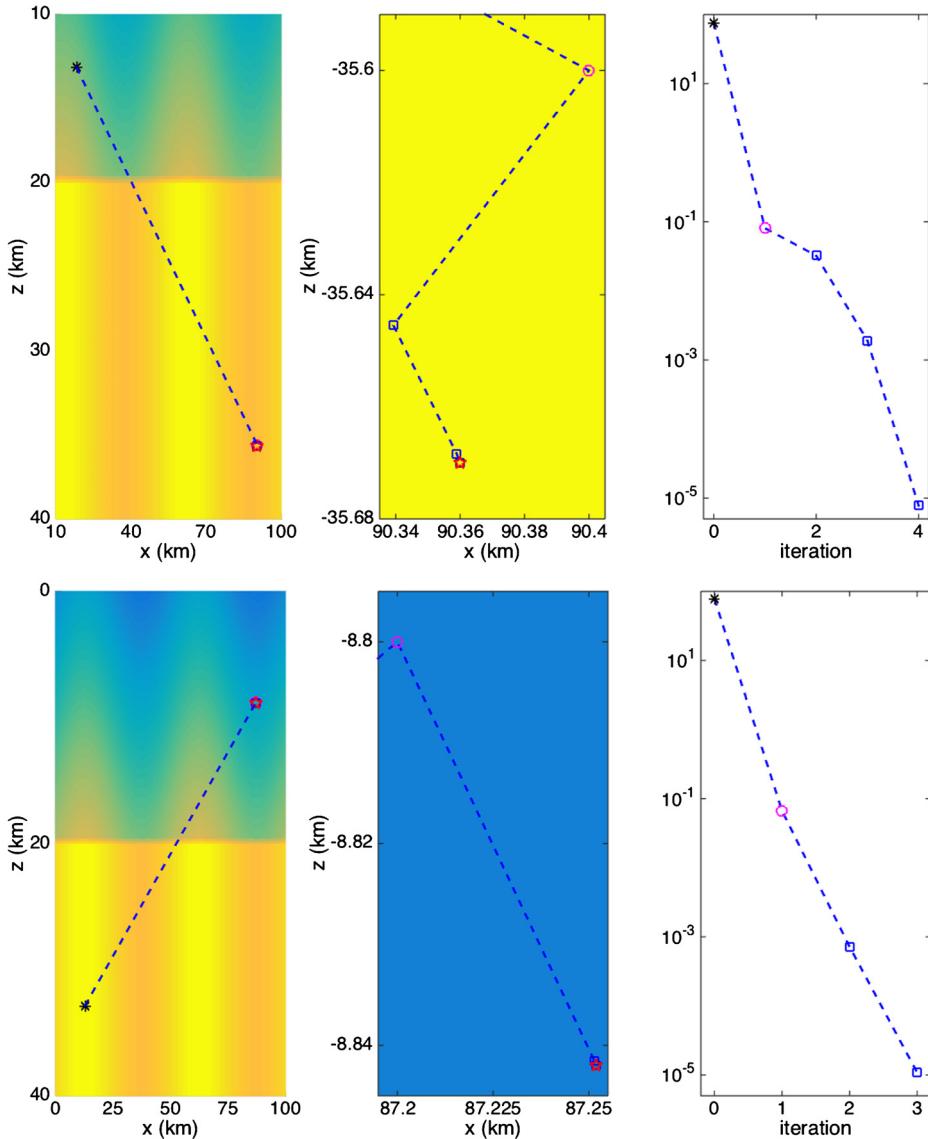


Fig. 6. Convergence history of the two-layer velocity model. Up for case (i), and Down for case (ii). Left (large-scale image) and Middle (zoom-in image): the convergent trajectories; Right: the absolute errors with respect to iteration steps between the real and computed hypocenter of the earthquake. The black star is the initial hypocenter, the magenta circle denotes the hypocenter obtained by auxiliary function method (Algorithm 1), the blue square indicate the hypocenters in the iterative process of the auxiliary function preprocessing method (Algorithm 2), and the red pentagon is the real hypocenter.

3.2. The practical velocity model

Let's consider a more practical model, the computational domain is $[0 \text{ km}, 200 \text{ km}] \times [0 \text{ km}, 200 \text{ km}]$, and the wave speed is

$$c(x, z) = \begin{cases} 5.5, & 0 < z \leq 33 + 5 \sin \frac{\pi x}{40}, \\ 7.8, & 33 + 5 \sin \frac{\pi x}{40} < z \leq 45 + 0.4x, \\ 7.488, & 45 + 0.4x < z \leq 60 + 0.4x, \\ 8.268, & 60 + 0.4x < z \leq 85 + 0.4x, \\ 7.8, & \text{others,} \end{cases}$$

with unit 'km/s'. The model consists of the crust, the mantle, and the undulating Moho discontinuity. In the mantle, there is a subduction zone with a thin low-velocity layer atop a fast velocity layer [23,26], see Fig. 8 for illustration. This model is a typical seismogenic zone [21]. Taking account into the complexity of the velocity structure, it is much more difficult to locate the earthquake hypocenter. The computational time interval $I = [0, 55 \text{ s}]$ and the dominant frequency is $f_0 = 2 \text{ Hz}$.

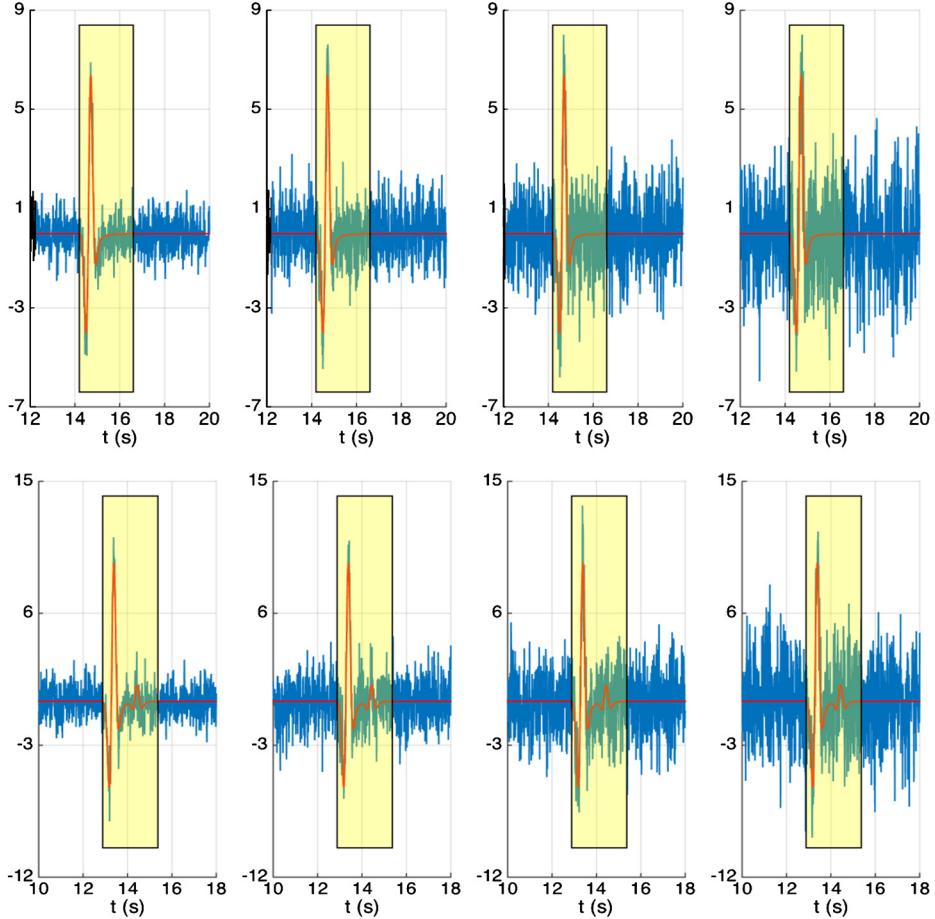


Fig. 7. Illustration of signals with noise in the two-layer velocity model. The signal with noise $d_r(t)$ (blue line) and the noise-free signal $u(\eta_r, t; \xi_T, \tau_T)$ for receivers $r = 9$. The horizontal axis is the time t . Up: parameters (i); Down: parameters (ii); From left to right, the ratio $R = 10\%, 15\%, 20\%, 25\%$ respectively.

Table 4

The two-layer velocity model. The Mean(M) and Standard Deviation(SD) of the errors between the location results (ξ_*, τ_*) and the exact solution (ξ_T, τ_T) .

R	Errors of AFM, case (i)		Errors of AFM, case (ii)	
	M	SD	M	SD
10%	0.0812	0	0.0676	0
15%	0.0812	0	0.0676	0
20%	0.0812	0	0.0676	0
25%	fail		0.0676	0

Table 5

The practical velocity model: the horizontal positions of receivers, with unit 'km'.

r	1	2	3	4	5	6	7	8	9	10	11	12
x_r	21	33	39	58	68	74	86	98	126	132	158	197

Consider 12 randomly distributed receivers $\eta_r = (x_r, z_r)$ on the surface $z_r = 0$ km, and their horizontal positions are given in Table 5.

We first investigate the noise-free situation. Consider the following four cases: (i) An earthquake occurs in the subduction zone, but the initial hypocenter is chosen in the mantle

$$\xi_T = (168.352 \text{ km}, 142.849 \text{ km}), \quad \tau_T = 10 \text{ s},$$

$$\xi = (53.494 \text{ km}, 47.113 \text{ km}), \quad \tau = 13.79 \text{ s};$$

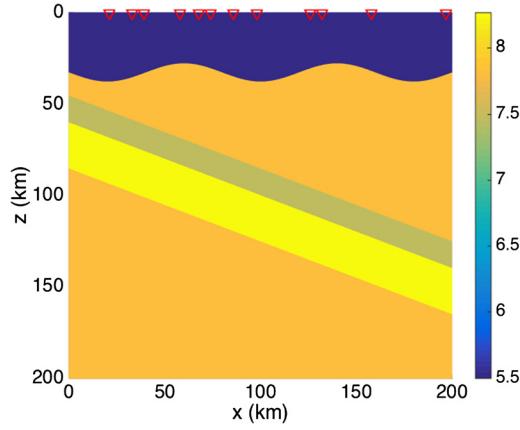


Fig. 8. The practical velocity model. The red triangles indicate the receivers.

(ii) The contrary case of (i)

$$\begin{aligned}\xi_T &= (53.494 \text{ km}, 47.113 \text{ km}), \quad \tau_T = 10 \text{ s}, \\ \xi &= (168.352 \text{ km}, 142.849 \text{ km}), \quad \tau = 13.79 \text{ s};\end{aligned}$$

(iii) An earthquake occurs in the crust and close to the Moho discontinuity, but the initial hypocenter is chosen in the subduction zone

$$\begin{aligned}\xi_T &= (163.326 \text{ km}, 32.877 \text{ km}), \quad \tau_T = 10 \text{ s}, \\ \xi &= (26.497 \text{ km}, 69.235 \text{ km}), \quad \tau = 15.32 \text{ s};\end{aligned}$$

(iv) The contrary case of (iii)

$$\begin{aligned}\xi_T &= (26.497 \text{ km}, 69.235 \text{ km}), \quad \tau_T = 10 \text{ s}, \\ \xi &= (163.326 \text{ km}, 32.877 \text{ km}), \quad \tau = 15.32 \text{ s};\end{aligned}$$

The searching domain is $\Omega_s = [0, 200 \text{ km}] \times [0, 200 \text{ km}]$, and the searching time interval is $I_s = [\tau - 10 \text{ s}, \tau + 10 \text{ s}]$. The mesh sizes for the searching grid are $h_x = 0.2 \text{ km}$ and $h_z = 0.2 \text{ km}$. The time step for the searching time interval is $\sigma = 0.05 \text{ s}$.

The cross-sections of the function $\Gamma(\xi, v)$ in (2.9) are output in Figs. 9–12. In these figures, it is easy to observe the uniqueness of the global minimum. We also present the convergent history of the auxiliary preprocessing method in Fig. 13. We can see that the global minimum of the function $\Gamma(\xi, v)$ is very close to the optimization solution of (1.1). From which, we can draw the same conclusion as in Subsection 3.1.

Finally, the noise is taken into consideration. We test the cases (i)–(iv). The noise is added in the same way as in Subsection 3.1. The real earthquake signal with noise $d_r(t)$ and the noise-free signal $u(\eta_r, t; \xi_T, \tau_T)$ are illustrated in Fig. 14. We also select a time window that contains the main part of $u(\eta_r, t; \xi_T, \tau_T)$ to reduce the impact of noise. In Fig. 14 (iii), two waveforms are observed, where the first is the head wave, and the latter is the direct wave. In Fig. 14 (iv), only one waveform is observed. Since the source location is very close to the discontinuity between the low-velocity layer and the fast velocity layer in the subduction zone, the direct wave, and the reflected wave arrive almost at the same time. The preceding two cases are very typical.

In Tables 6–9, the location results (ξ_*, τ_*) computed via the auxiliary function method (Algorithm 1) and their errors with respect to different ratio R are presented. From which, we can see that the method can obtain satisfactory location results for $R = 10\%$ and 15% . For $R = 20\%$, the auxiliary function method fails to obtain correct results in case (i) and (iii). Nevertheless, the auxiliary function method is still much better than the iterative methods, which is only valid for $R \leq 5\%$ according to our numerical tests. Taking into account that the computational cost of the auxiliary function method is almost the same as the single iteration step of the iterative method. The computation efficiency of our method is also obvious.

Remark 5. In the above numerical examples, we can see the critical point of convergence of the auxiliary function method is the global minimum of the function $\Gamma(\xi, v)$ in (2.9) coincides with the location of earthquake hypocenter and origin time. However, the global minimum of the function $\Gamma(\xi, v)$ may deviate under data noise. When the noise intensity is large, the global minimum of the function $\Gamma(\xi, v)$ deviates far from the earthquake hypocenter and origin time, which causes the method to fail. This explains the reason why the auxiliary function method fails in case (i) and (iii). However, we have to remark that this method is still much better than the iterative method.

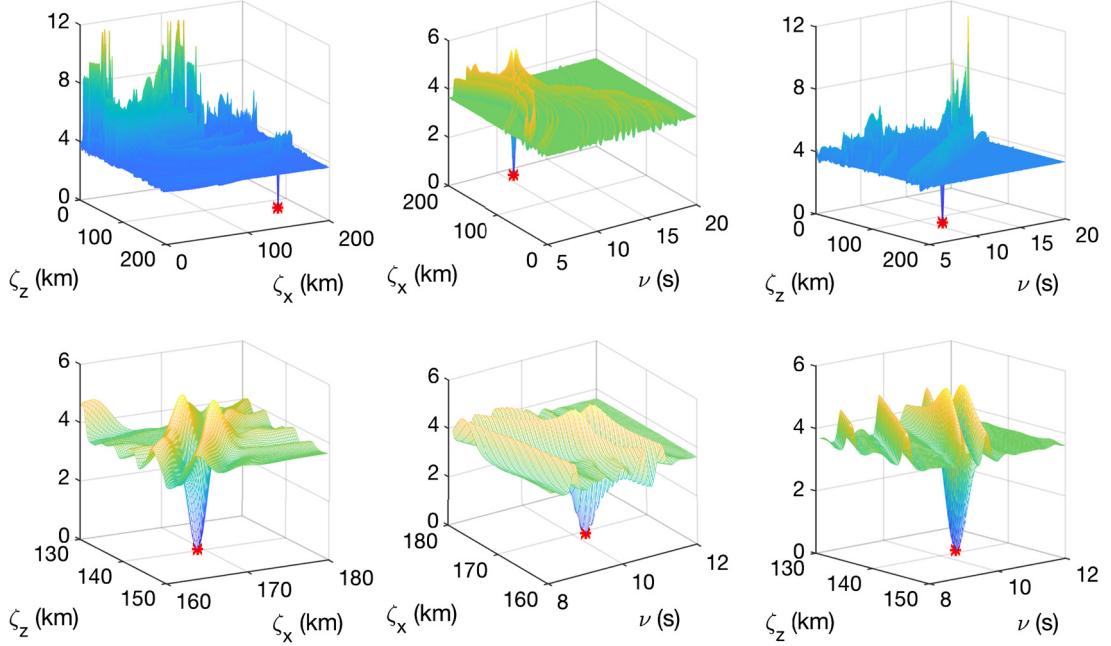


Fig. 9. The practical velocity model, case (i). The $\xi_x - \xi_z$ (Left, $\nu = 10$ s), $\nu - \xi_x$ (Middle, $\xi_z = 142.849$ km) and $\nu - \xi_z$ (Right, $\xi_x = 168.352$ km) cross-sections of the function $\Gamma(\xi, \nu)$. Up for large-scale image and Down for zoom-in image.

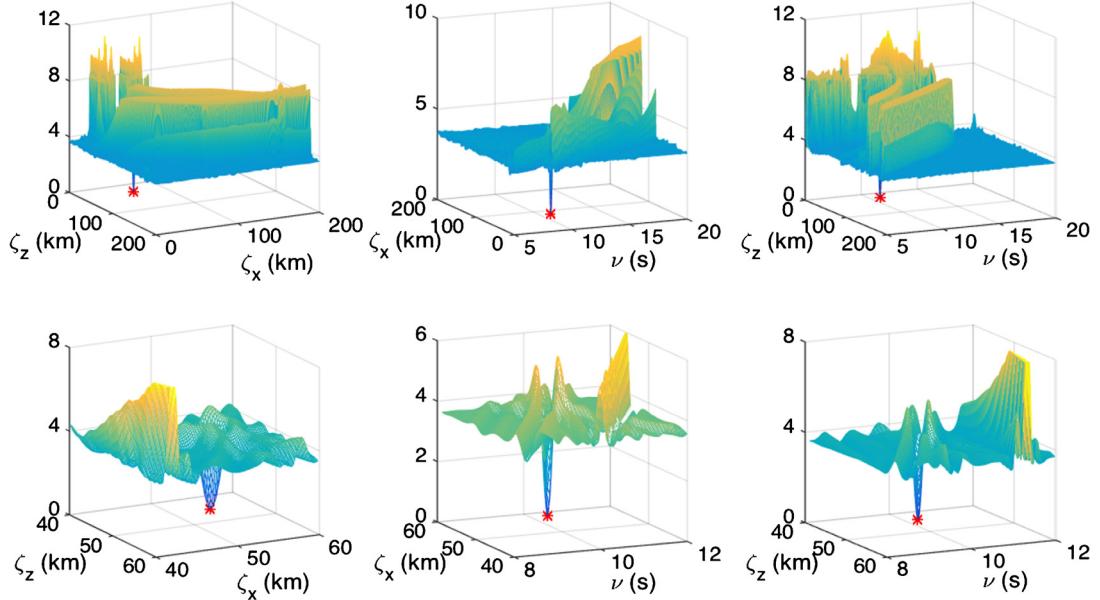


Fig. 10. The practical velocity model, case (ii). The $\xi_x - \xi_z$ (Left, $\nu = 10$ s), $\nu - \xi_x$ (Middle, $\xi_z = 47.113$ km) and $\nu - \xi_z$ (Right, $\xi_x = 53.494$ km) cross-sections of the function $\Gamma(\xi, \nu)$. Up for large-scale image and Down for zoom-in image.

Table 6

The practical velocity model, case (i). The location results (ξ_*, τ_*) and the errors.

R	ξ_* (km)	τ_* (s)	$\ \xi_T - \xi_*\ _2$ (km)	$ \xi_{*Z} - \xi_{Tz} / \xi_{Tz}$
10%	(168.2, 142.6)	10.04	0.292	0.17%
15%	(168.2, 142.2)	10.09	0.667	0.45%
20%			fail	

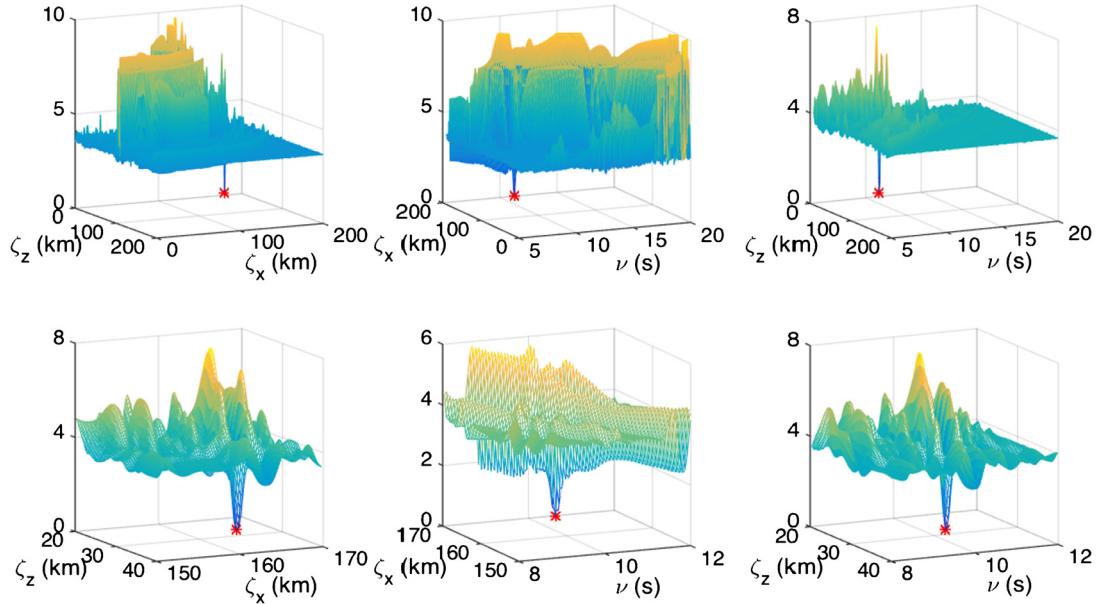


Fig. 11. The practical velocity model, case (iii). The $\zeta_x - \zeta_z$ (Left, $\nu = 10$ s), $\nu - \zeta_x$ (Middle, $\zeta_z = 32.877$ km) and $\nu - \zeta_z$ (Right, $\zeta_x = 163.326$ km) cross-sections of the function $\Gamma(\zeta, \nu)$. Up for large-scale image and Down for zoom-in image.

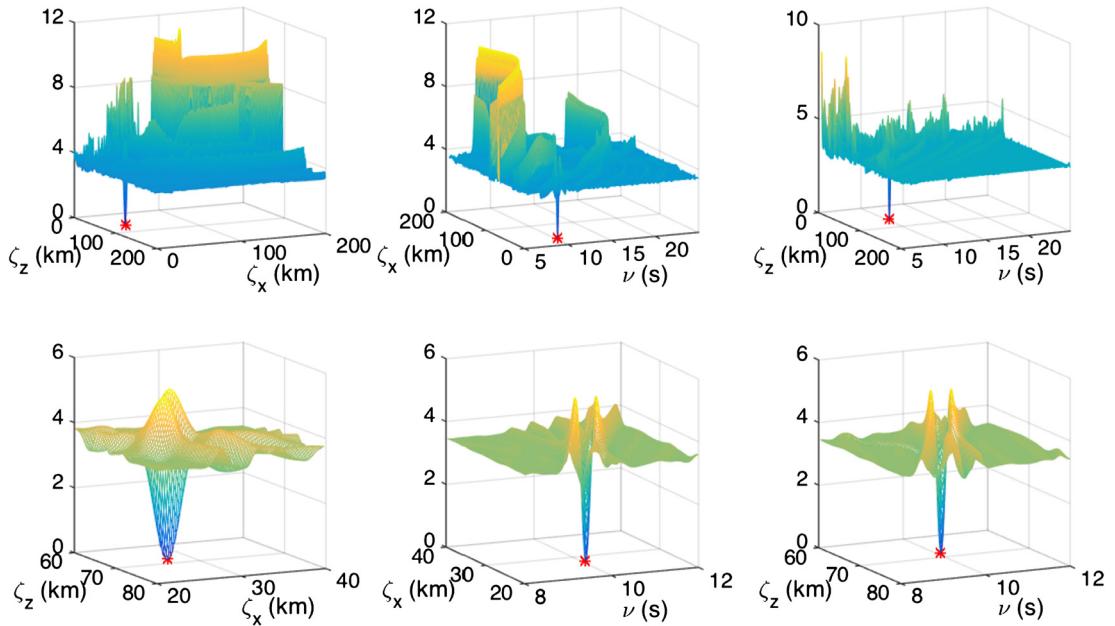


Fig. 12. The practical velocity model, case (iv). The $\zeta_x - \zeta_z$ (Left, $\nu = 10$ s), $\nu - \zeta_x$ (Middle, $\zeta_z = 69.235$ km) and $\nu - \zeta_z$ (Right, $\zeta_x = 26.497$ km) cross-sections of the function $\Gamma(\zeta, \nu)$. Up for large-scale image and Down for zoom-in image.

Table 7

The practical velocity model, case (ii). The location results (ξ_*, τ_*) and the errors.

R	ξ_* (km)	τ_* (s)	$\ \xi_T - \xi_*\ _2$ (km)	$ \xi_{*Z} - \xi_{Tz} / \xi_{Tz}$
10%	(53.4, 47.2)	9.99	0.128	0.18%
15%	(53.4, 47.2)	9.99	0.128	0.18%
20%	(53.4, 47.2)	9.99	0.128	0.18%

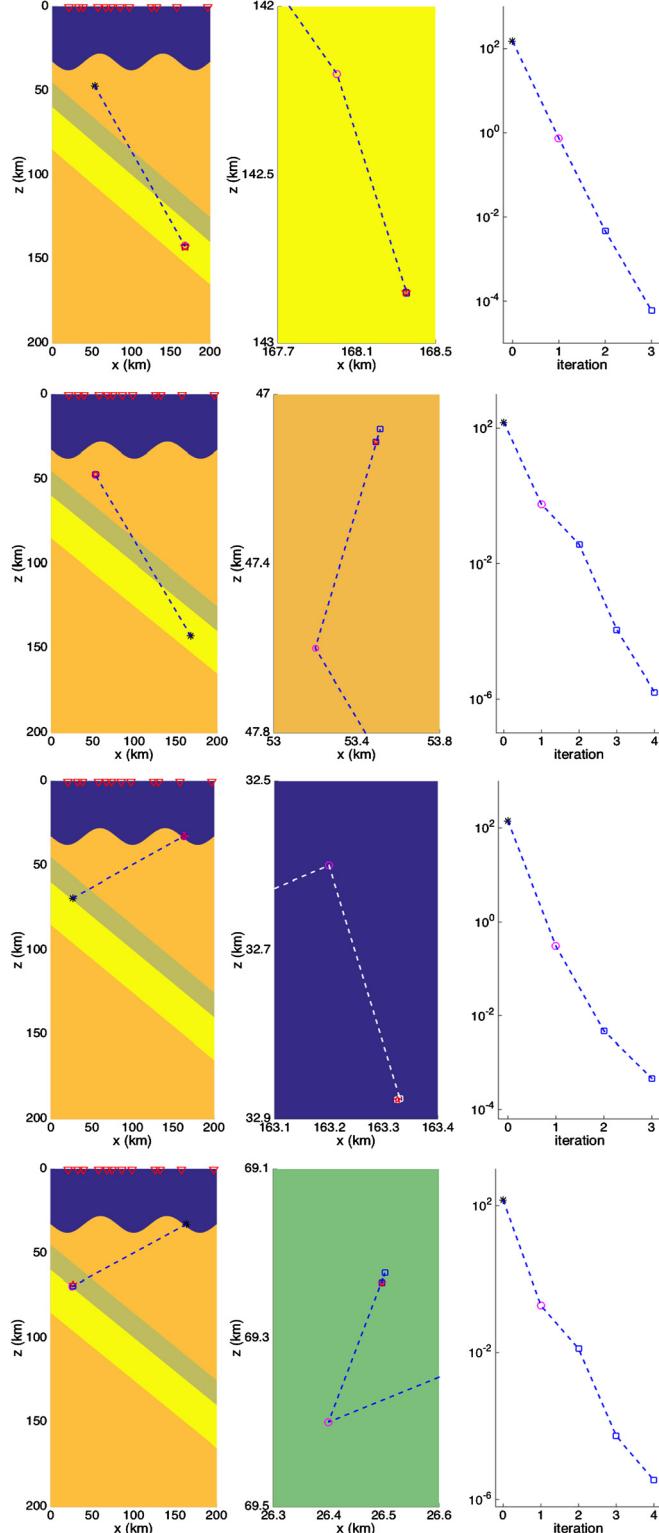


Fig. 13. Convergence history of the practical velocity model. From Up to Down corresponding the cases (i)-(iv). Left (large-scale image) and Middle (zoom-in image); the convergent trajectories; Right: the absolute errors with respect to iteration steps between the real and computed hypocenter of the earthquake. The black star is the initial hypocenter, the magenta circle denotes the hypocenter obtained by auxiliary function method (Algorithm 1), the blue square indicate the hypocenters in the iterative process of the auxiliary function preprocessing method (Algorithm 2), and the red pentagram is the real hypocenter.

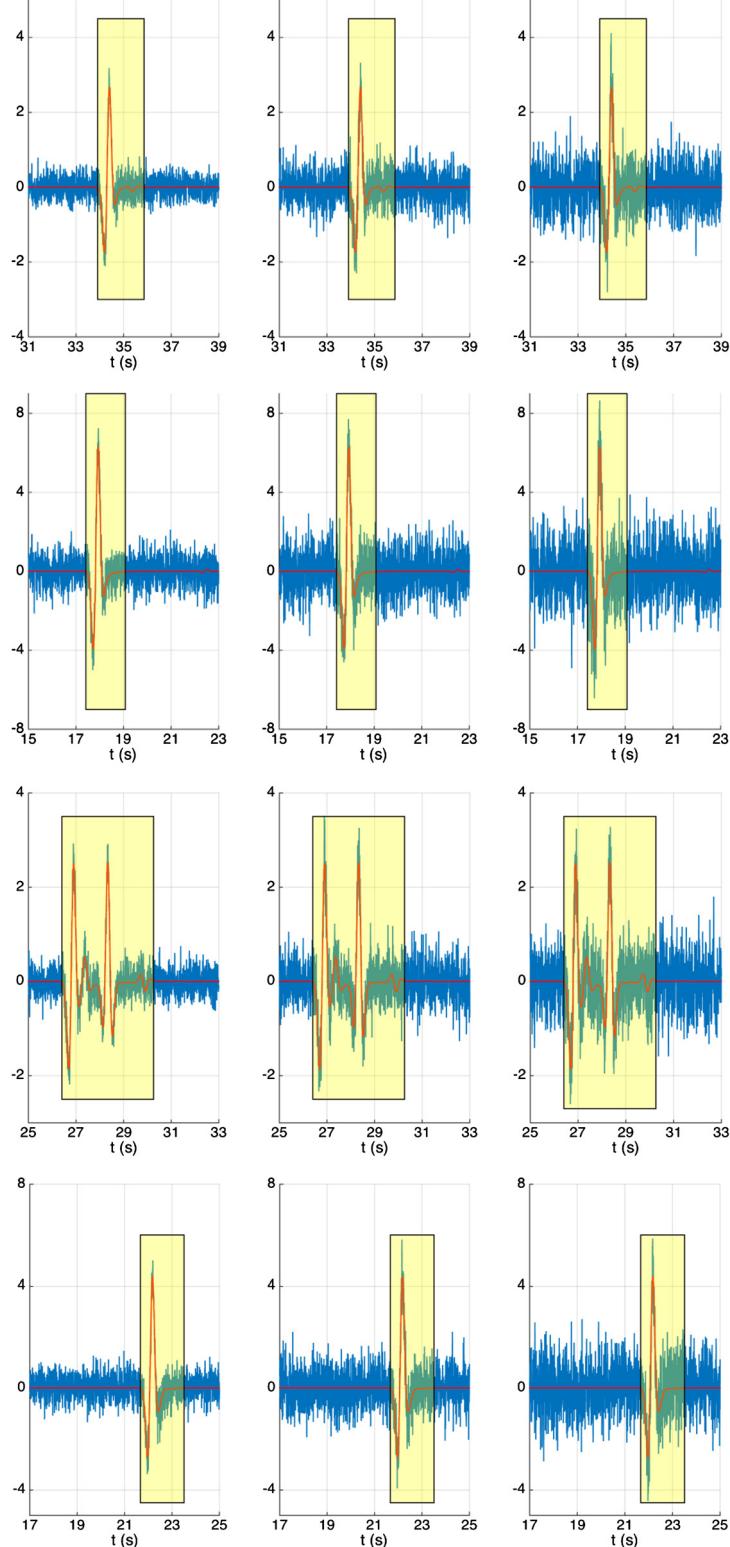


Fig. 14. Illustration of signal with noise in the practical velocity model. The signal with noise $d_r(t)$ (blue line) and the noise-free signal $u(\eta_r, t; \xi_T, \tau_T)$ for receivers $r = 5$. The horizontal axis is the time t . From Up to Down corresponding the cases (i)-(iv). From left to right, the ratio $R = 10\%, 15\%, 20\%$ respectively.

Table 8

The practical velocity model, case (iii). The location results (ξ_* , τ_*) and the errors.

R	ξ_* (km)	τ_* (s)	$\ \xi_T - \xi_*\ _2$ (km)	$ \xi_{*Z} - \xi_{Tz} / \xi_{Tz}$
10%	(163.2, 32.6)	10.02	0.304	0.84%
15%	(163.2, 32.6)	10.02	0.304	0.84%
20%			fail	

Table 9

The practical velocity model, case (iv). The location results (ξ_* , τ_*) and the errors.

R	ξ_* (km)	τ_* (s)	$\ \xi_T - \xi_*\ _2$ (km)	$ \xi_{*Z} - \xi_{Tz} / \xi_{Tz}$
10%	(26.4, 69.4)	9.97	0.191	0.24%
15%	(26.4, 69.4)	9.97	0.191	0.24%
20%	(26.4, 69.4)	9.97	0.191	0.24%

4. Conclusion and discussion

The first conclusion to be drawn from the numerical evidence presented earlier is that the auxiliary function pre-processing method (Algorithm 2) can determine the earthquake hypocenter and the origin time very efficient and accurate when the seismic signals are noise-free. Secondly, the auxiliary function method (Algorithm 1) can locate the earthquake hypocenter and the origin time with reasonable accuracy in the situation of noise. The above advantages are based on the fact that the real hypocenter and origin time is the root of the new constructed auxiliary functions. Moreover, the total computational cost of constructing these functions is comparable to the single iteration step of the iterative method.

It should be noted that there are still many issues need to be further investigated: (a) Currently, the uniqueness of the solution cannot be proved, but we have observed the uniqueness of the solution numerically. It is exciting to present an intuitive study. (b) We are currently working on the 2-D problem and the accurate velocity model. For 3-D problem and the inaccurate velocity model, the situations may be more complicated. These all require much more effort.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (Grant Nos. 11871297, U1839206, 91730306), the National Key Research and Development Program of China on Monitoring, Early Warning and Prevention of Major Natural Disaster (Grant No. 2017YFC1500301), the Startup Grant of Nanyang Technological University (Grant No. M4081881.110), China Postdoctoral Science Foundation (Grant No. 2018M641323) and Tsinghua University Initiative Scientific Research Program. The authors are grateful to Prof. Shi Jin for his helpful suggestions and discussions that significantly improve the presentation. Hao Wu would like to thank Prof. Jingwei Hu and Ms. Meixia Wang for their valuable comments.

Appendix A

In this appendix, we give the detailed derivation of the auxiliary function method for the elastic wave equation

$$\rho(\mathbf{x}) \frac{\partial^2 \mathbf{u}(\mathbf{x}, t; \xi, \tau)}{\partial t^2} = \nabla \cdot (\mathbf{C} : \nabla \mathbf{u}(\mathbf{x}, t; \xi, \tau)) + \mathbf{f}(t - \tau) \delta(\mathbf{x} - \xi), \quad \mathbf{x}, \xi \in \Omega, \quad (\text{A.1})$$

with initial-boundary conditions

$$\mathbf{n} \cdot (\mathbf{C} : \nabla \mathbf{u}(\mathbf{x}, t; \xi, \tau)) = \mathbf{0}, \quad \mathbf{x} \in \partial\Omega, \quad (\text{A.2})$$

$$\mathbf{u}(\mathbf{x}, 0; \xi, \tau) = \partial_t \mathbf{u}(\mathbf{x}, 0; \xi, \tau) = \mathbf{0}, \quad \mathbf{x} \in \Omega, \quad (\text{A.3})$$

where $\rho(\mathbf{x})$ is the density, $\mathbf{u}(\mathbf{x}, t; \cdot, \cdot)$ is the displacement vector at time t and location \mathbf{x} , \mathbf{f} is the force-source vector, and \mathbf{C} is the fourth-order elastic tensor satisfies the following symmetry conditions:

$$C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klji}.$$

Similar to the acoustic wave situation, the real earthquake signal $\mathbf{d}_r(t)$ occurred at (η_T, τ_T) and recorded at receiver r , and the synthetic earthquake signal $\mathbf{s}(\mathbf{x}, t)$ corresponding to the initial hypocenter ξ and the initial origin time are regarded as the solutions of the elastic wave equation

$$\mathbf{d}_r(t) = \mathbf{u}(\boldsymbol{\eta}_r, t; \boldsymbol{\xi}_T, \tau_T), \quad \mathbf{s}(\boldsymbol{\eta}_r, t) = \mathbf{u}(\boldsymbol{\eta}_r, t; \boldsymbol{\xi}, \tau) \quad (\text{A.4})$$

The earthquake location problem under the framework of elastic wave equation can also be formulated as (1.1):

$$(\boldsymbol{\xi}_T, \tau_T) = \underset{\boldsymbol{\xi}, \tau}{\operatorname{argmin}} \sum_{r \in \mathcal{R}} \chi_r(\boldsymbol{\xi}, \tau),$$

with a slight modification of the misfit function $\chi_r(\boldsymbol{\xi}, \tau)$ at the r -th receiver

$$\chi_r(\boldsymbol{\xi}, \tau) = \frac{\int_0^{t_f} \|\mathbf{d}_r(t) - \mathbf{s}(\boldsymbol{\eta}_r, t)\|_2^2 dt}{2 \int_0^{t_f} \|\mathbf{d}_r(t)\|_2^2 dt}. \quad (\text{A.5})$$

Based on the aforementioned discussions, the traditional iterative methods for the elastic waveform based earthquake location problem (1.1) and (A.1)-(A.5) are numerically expensive. Thus, we also propose the following elastic wave version of the auxiliary function method.

Theorem 2. For any given initial hypocenter $\boldsymbol{\xi}$ and origin time τ , define the auxiliary functions

$$\Xi_r(\boldsymbol{\xi}, \nu) = 2\chi_r(\boldsymbol{\xi}, \tau) - \int_0^{t_f} \mathbf{w}_r(\boldsymbol{\xi}, t) \cdot \mathbf{f}(t - \nu) - \mathbf{w}_r(\boldsymbol{\xi}, t) \cdot \mathbf{f}(t - \tau) dt, \quad \forall r \in \mathcal{R},$$

in which $\chi_r(\boldsymbol{\xi}, \tau)$ has been given in (A.5), and $\mathbf{w}_r(\mathbf{x}, t)$ satisfies the adjoint elastic equation with terminal-boundary conditions

$$\begin{cases} \rho(\mathbf{x}) \frac{\partial^2 \mathbf{w}_r(\mathbf{x}, t)}{\partial t^2} = \nabla \cdot (\mathbf{C} : \nabla \mathbf{w}_r(\mathbf{x}, t)) + \frac{\mathbf{d}_r(t) - \mathbf{s}(\boldsymbol{\eta}_r, t)}{\int_0^{t_f} \|\mathbf{d}_r(t)\|_2^2 dt} \delta(\mathbf{x} - \boldsymbol{\eta}_r), & \mathbf{x} \in \Omega, \\ \mathbf{n} \cdot (\mathbf{C} : \nabla \mathbf{w}_r(\mathbf{x}, t)) = 0, & \mathbf{x} \in \partial\Omega, \\ \mathbf{w}_r(\mathbf{x}, t_f) = \partial_t \mathbf{w}_r(\mathbf{x}, t_f) = 0, & \mathbf{x} \in \Omega. \end{cases} \quad (\text{A.6})$$

Then, for real earthquake hypocenter $\boldsymbol{\xi}_T$ and origin time τ_T , we have

$$\Xi_r(\boldsymbol{\xi}_T, \tau_T) = 0, \quad \forall r \in \mathcal{R}.$$

Proof. Let us first define the difference function

$$\delta \mathbf{s}(\mathbf{x}, t) = \mathbf{u}(\mathbf{x}, t; \boldsymbol{\xi}_T, \tau_T) - \mathbf{u}(\mathbf{x}, t; \boldsymbol{\xi}, \tau).$$

According to (A.4), it follows

$$\delta \mathbf{s}(\boldsymbol{\eta}_r, t) = \mathbf{d}_r(t) - \mathbf{s}(\boldsymbol{\eta}_r, t)$$

And the difference function $\delta \mathbf{s}(\mathbf{x}, t)$ satisfies the elastic wave equation with initial-boundary conditions

$$\begin{cases} \rho(\mathbf{x}) \frac{\partial^2 \delta \mathbf{s}(\mathbf{x}, t)}{\partial t^2} = \nabla \cdot (\mathbf{C} : \nabla \delta \mathbf{s}(\mathbf{x}, t)) + \mathbf{f}(t - \tau_T) \delta(\mathbf{x} - \boldsymbol{\xi}_T) - \mathbf{f}(t - \tau) \delta(\mathbf{x} - \boldsymbol{\xi}), & \mathbf{x} \in \Omega, \\ \mathbf{n} \cdot (\mathbf{C} : \nabla \delta \mathbf{s}(\mathbf{x}, t)) = 0, & \mathbf{x} \in \partial\Omega, \\ \delta \mathbf{s}(\mathbf{x}, 0) = \partial_t \delta \mathbf{s}(\mathbf{x}, 0) = 0, & \mathbf{x} \in \Omega. \end{cases}$$

Multiply the adjoint wave function $\mathbf{w}_r(\mathbf{x}, t)$ given in (A.6) in both sides and integrate them on $\Omega \times [0, t_f]$, we obtain

$$\begin{aligned} & \int_{\Omega} \int_0^{t_f} \rho(\mathbf{x}) \mathbf{w}_r(\mathbf{x}, t) \cdot \frac{\partial^2 \delta \mathbf{s}(\mathbf{x}, t)}{\partial t^2} dt d\mathbf{x} = \int_0^{t_f} \int_{\Omega} \mathbf{w}_r(\mathbf{x}, t) \cdot (\nabla \cdot (\mathbf{C} : \nabla \delta \mathbf{s}(\mathbf{x}, t))) d\mathbf{x} dt \\ & + \int_0^{t_f} \int_{\Omega} \mathbf{w}_r(\mathbf{x}, t) \cdot \mathbf{f}(t - \tau_T) \delta(\mathbf{x} - \boldsymbol{\xi}_T) d\mathbf{x} dt - \int_0^{t_f} \int_{\Omega} \mathbf{w}_r(\mathbf{x}, t) \cdot \mathbf{f}(t - \tau) \delta(\mathbf{x} - \boldsymbol{\xi}) d\mathbf{x} dt \end{aligned}$$

By using the integration by parts and the symmetry properties of \mathbf{C} , we can obtain

$$\begin{aligned} & \int_{\Omega} \int_0^{t_f} \rho(\mathbf{x}) \delta \mathbf{s}(\mathbf{x}, t) \cdot \frac{\partial^2 \mathbf{w}_r(\mathbf{x}, t)}{\partial t^2} dt d\mathbf{x} = \int_{\Omega} \int_0^{t_f} \delta \mathbf{s}(\mathbf{x}, t) \cdot (\nabla \cdot (\mathbf{C} : \nabla) \mathbf{w}_r(\mathbf{x}, t)) dt d\mathbf{x} \\ & + \int_0^{t_f} \int_{\Omega} \mathbf{w}_r(\boldsymbol{\xi}_T, t) \cdot \mathbf{f}(t - \tau_T) dt - \int_0^{t_f} \int_{\Omega} \mathbf{w}_r(\boldsymbol{\xi}, t) \cdot \mathbf{f}(t - \tau) dt \end{aligned} \quad (\text{A.7})$$

On the other hand, the misfit function $\chi_r(\xi, \tau)$ can be rewritten as

$$\chi_r(\xi, \tau) = \frac{\int_0^{t_f} (\mathbf{d}_r(t) - \mathbf{s}(\eta_r, t)) \cdot \delta \mathbf{s}(\eta_r, t) dt}{2 \int_0^{t_f} \|\mathbf{d}_r(t)\|_2^2 dt} = \frac{\int_0^{t_f} \int_{\Omega} (\mathbf{d}_r(t) - \mathbf{s}(\eta_r, t)) \cdot \delta \mathbf{s}(\mathbf{x}, t) \delta(\mathbf{x} - \eta_r) d\mathbf{x} dt}{2 \int_0^{t_f} \|\mathbf{d}_r(t)\|_2^2 dt}.$$

Multiplying both sides of the above equation by 2, and adding equation (A.7), we get

$$2\chi_r(\xi, \tau) = \int_0^{t_f} \mathbf{w}_r(\xi_T, t) \cdot \mathbf{f}(t - \tau_T) - \mathbf{w}_r(\xi, t) \cdot \mathbf{f}(t - \tau) dt.$$

This completes the proof. \square

References

- [1] K. Aki, P.G. Richards, Quantitative Seismology: Theory and Methods Volume II, W.H. Freeman & Co (Sd), 1980.
- [2] J. Chen, Y.F. Chen, H. Wu, D.H. Yang, The quadratic Wasserstein metric for Earthquake Location, *J. Comput. Phys.* 373 (2018) 188–209.
- [3] M.A. Dablain, The application of high-order differencing to the scalar wave equation, *Geophysics* 51 (1) (1986) 54–66.
- [4] B. Engquist, O. Runborg, Computational high frequency wave propagation, *Acta Numer.* 12 (2003) 181–266.
- [5] M.C. Ge, Analysis of source location algorithms Part I: overview and non-iterative methods, *J. Acoust. Emiss.* 21 (2003) 14–28.
- [6] M.C. Ge, Analysis of source location algorithms Part II: iterative methods, *J. Acoust. Emiss.* 21 (2003) 29–51.
- [7] L. Geiger, Probability method for the determination of earthquake epicenters from the arrival time only, *Bull. St. Louis Univ.* 8 (1912) 60–71.
- [8] X.Y. Huang, D.H. Yang, P. Tong, J. Badal, Q.Y. Liu, Wave equation-based reflection tomography of the 1992 Landers earthquake area, *Geophys. Res. Lett.* 43 (2016) 1884–1892.
- [9] S. Jin, H. Wu, X. Yang, Gaussian beam methods for the Schrödinger equation in the semi-classical regime: Lagrangian and Eulerian formulations, *Commun. Math. Sci.* 6 (4) (2008) 995–1020.
- [10] Y.H. Kim, Q.Y. Liu, J. Tromp, Adjoint centroid-moment tensor inversions, *Geophys. J. Int.* 186 (2011) 264–278.
- [11] D. Komatitsch, J. Tromp, A perfectly matched layer absorbing boundary condition for the second-order seismic wave equation, *Geophys. J. Int.* 154 (2003) 146–153.
- [12] W.H.K. Lee, S.W. Stewart, Principles and Applications of Microearthquake Networks, Academic Press, 1981.
- [13] J.S. Li, D.H. Yang, H. Wu, X. Ma, A low-dispersive method using the high-order stereo-modelling operator for solving 2-D wave equations, *Geophys. J. Int.* 210 (2017) 1938–1964.
- [14] Q.Y. Liu, J. Polet, D. Komatitsch, J. Tromp, Spectral-element moment tensor inversion for earthquakes in southern California, *Bull. Seismol. Soc. Am.* 94 (5) (2004) 1748–1761.
- [15] R. Madariaga, Seismic source theory, in: S. Gerald (Ed.), Treatise on Geophysics, second edition, Elsevier B.V., 2015, pp. 51–71.
- [16] A.F. Prugger, D.J. Gendzwill, Microearthquake location: a nonlinear approach that makes use of a simplex stepping procedure, *Bull. Seismol. Soc. Am.* 78 (1988) 799–815.
- [17] N. Rawlinson, S. Pozgay, S. Fishwick, Seismic tomography: a window into deep Earth, *Phys. Earth Planet. Inter.* 178 (2010) 101–135.
- [18] C. Satriano, A. Lomax, A. Zollo, Real-time evolutionary earthquake location for seismic early warning, *Bull. Seismol. Soc. Am.* 98 (3) (2008) 1482–1494.
- [19] C.H. Thurber, Nonlinear earthquake location: theory and examples, *Bull. Seismol. Soc. Am.* 75 (3) (1985) 779–790.
- [20] C.H. Thurber, Earthquake, location techniques, in: Gupta (Ed.), Encyclopedia of Earth Sciences Series, H.K., Springer, 2014, pp. 201–207.
- [21] P. Tong, D.P. Zhao, D.H. Yang, Tomography of the 1995 Kobe earthquake area: comparison of finite-frequency and ray approaches, *Geophys. J. Int.* 187 (2011) 278–302.
- [22] P. Tong, D. Zhao, D.H. Yang, X. Yang, J. Chen, Q. Liu, Wave-equation-based travel-time seismic tomography – Part 1: method, *Solid Earth* 5 (2014) 1151–1168.
- [23] P. Tong, D.H. Yang, Q.Y. Liu, X. Yang, J. Harris, Acoustic wave-equation-based earthquake location, *Geophys. J. Int.* 205 (1) (2016) 464–478.
- [24] F. Waldhauser, W.L. Ellsworth, A double-difference earthquake location algorithm: method and application to the northern Hayward Fault, California, *Bull. Seismol. Soc. Am.* 90 (6) (2000) 1353–1368.
- [25] X. Wen, High order numerical quadratures to one dimensional delta function integrals, *SIAM J. Sci. Comput.* 30 (4) (2008) 1825–1846.
- [26] H. Wu, J. Chen, X.Y. Huang, D.H. Yang, A new earthquake location method based on the waveform inversion, *Commun. Comput. Phys.* 23 (1) (2018) 118–141.
- [27] H. Wu, X. Yang, Eulerian Gaussian beam method for high frequency wave propagation in the reduced momentum space, *Wave Motion* 50 (6) (2013) 1036–1049.
- [28] D.H. Yang, M. Lu, R.S. Wu, J.M. Peng, An optimal nearly analytic discrete method for 2D acoustic and elastic wave equations, *Bull. Seismol. Soc. Am.* 94 (5) (2004) 1982–1991.
- [29] D.T. Zhou, J. Chen, H. Wu, D.H. Yang, L.Y. Qiu, The Wasserstein-Fisher-Rao metric for waveform based earthquake location, arXiv:1812.00304, 2018.