Overcoming challenges in time reversal for passive seismic source localization

Dissertation

with the aim of achieving a doctoral degree

at the Faculty of Mathematics, Informatics and Natural Sciences

Department of Earth Sciences

at Universität Hamburg

Submitted by Peng Yang

Hamburg, 2024

Accepted as Dissertation at the Department of Earth Sciences

Day of oral defense:	30.04.2024
Reviewers:	Prof. Dr. Dirk Gajewski
Chair of the Subject Doctoral Committee:	Prof. Dr. Hermann Held
Dean of Faculty of MIN:	Prof. Dr.-Ing. Norbert Ritter

Abstract

Passive seismic source localization plays a crucial role in understanding geodynamics, monitoring geological activities, forecasting geological hazards, and managing geological fracturing processes. This doctoral thesis presents three consecutive approaches to enhance the robustness and accuracy of time-reversal source localization. The proposed methods address challenges such as source imaging artifacts, low-resolution source images, sparse and small-aperture seismic data acquisitions, and unknown seismic velocity models.

In the first approach, the focus is on refining time-reversal imaging. Conventional time-reversal source imaging methods like autocorrelation imaging or grouped crosscorrelation imaging often suffer from source imaging artifacts due to the use of low-quality seismic data or less constrained velocity models. These artifacts typically degrade the quality of the image and can lead to subsequent misinterpretation, resulting in false source location estimation. To overcome this, the Gaussian-weighted crosscorrelation imaging condition is proposed. Each time step in this method includes dividing the back-propagated wavefield, weighting seismic amplitudes using Gaussian functions, and using a zero-lag crosscorrelation. This process effectively minimizes source imaging artifacts, resulting in high-resolution, low-noise source images. Numerical examples of complex models and field examples illustrate the method's performance, demonstrating its effectiveness in identifying sources, even within clusters and under conditions of noisy and sparse-sampled data.

The second approach addresses another challenge posed by sparse and small-aperture seismic data acquisition in time-reversal imaging techniques. Such acquisitions often lead to false wave focusing due to insufficient wave illumination. To address the issues, the maximum-amplitude path method is introduced for source localization using maximum-amplitude paths. The paths are constructed from back-projected wavefields using receiver patches selected from the acquisition. They include the maximum amplitudes of the back-projected wavefronts for each considered time step. The proposed method exploits the continuity of the maximum amplitudes of the back-projected wavefields. The point of closest proximity (or crossing point) of the paths is the source location and the corresponding time is the source time. The maximum-amplitude path method successfully overcomes the acquisition problems and provides accurate source location and source excitation time even in challenging scenarios.

In the final approach, a data-driven hybrid workflow is presented to address the challenge of the lack of velocity model in time-reversal localization methods including the proposed first and second time-reversal localization approaches. The proposed workflow can simultaneously invert the source location, excitation times, and velocity model using wavefront attributes of passive seismic data. By combining wavefront tomography and time-reversal methods, the workflow eliminates the need for detailed prior information, making it particularly applicable in practical scenarios. The proposed workflow comprises the following steps. First, a set of user-defined vertical gradient velocity models is designed. Time reversal is then used to estimate the source excitation times for each model. After that, these source times and gradient models are used in wavefront tomography. The second step uses an optimization procedure to refine the velocity model and the source excitation time. Each iteration of the optimization involves a sequential application of time reversal and wave-

front tomography. In the final step, the source location is refined using the optimal velocity model and the Gaussian-weighted crosscorrelation imaging condition. The proposed workflow overcomes the limitations of time reversal in the absence of a velocity model, providing good velocity models and fairly accurate source locations and excitation times.

Zusammenfassung

Die genaue Lokalisierung passiver seismischer Quellen spielt eine entscheidende Rolle für das Verständnis der Geodynamik, die Überwachung geologischer Aktivitäten, die Vorhersage geologischer Gefahren und die Kontrolle von Frakturprozessen. In dieser Dissertation werden drei aufeinander aufbauende Ansätze zur Verbesserung der Robustheit und Genauigkeit der zeitumgekehrten Quellenlokalisierung vorgestellt. Die vorgeschlagenen Methoden befassen sich mit Herausforderungen wie Artefakten bei der Quellenabbildung, niedrig aufgelösten Quellenbildern, unzureichender Akquisition und unbekannten seismischen Geschwindigkeitsmodellen.

Beim ersten Ansatz liegt der Schwerpunkt auf der Verfeinerung der zeitumgekehrten Bildgebung. Herkömmliche zeitumgekehrte Quellenabbildungsmethoden wie die Autokorrelationsabbildung oder die gruppierte Kreuzkorrelationsabbildung leiden häufig unter Artefakten bei der Quellenabbildung, die auf die Verwendung von seismischen Daten geringer Qualität oder von zu ungenauen Geschwindigkeitsmodellen zurückzuführen sind. Diese Artefakte verschlechtern in der Regel die Qualität des Bildes und können zu einer Fehlinterpretation führen, die eine falsche Schätzung der Quellenposition zur Folge hat. Um dieses Problem zu lösen, wird die Gauß-gewichtete Kreuzkorrelationsabbildung vorgeschlagen. Bei dieser Methode wird in jedem Zeitschritt das sich rückwärts ausbreitende Wellenfeld geteilt, die seismischen Amplituden mit Gaußfunktionen gewichtet und eine Kreuzkorrelation mit Nullverzögerung verwendet. Dieses Verfahren minimiert effektiv Artefakte bei der Quellenabbildung und führt zu hochauflösenden, rauscharmen Quellenbildern. Numerische Beispiele komplexer Modelle und Feldbeispiele veranschaulichen die Leistung der Methode und zeigen ihre Wirksamkeit bei der Identifizierung von Quellen, sogar innerhalb von Clustern und unter Bedingungen verrauschter und spärlich abgetasteter Daten.

Der zweite Ansatz begegnet einer weiteren Herausforderung, die sich aus der Erfassung von spärlichen und kleinen Seismikdaten bei den Zeitumkehrbildgebungstechniken ergibt. Solche Erfassungen führen häufig zu einer falschen Wellenfokussierung aufgrund unzureichender Abdeckung. Um diese Probleme zu lösen, wird die Maximum-Amplituden-Pfad-Methode zur Quellenlokalisierung unter Verwendung von Maximum-Amplituden-Pfaden eingeführt. Die Pfade werden aus rückprojizierten Wellenfeldern unter Verwendung von aus der Erfassung ausgewählten Empfängerfeldern konstruiert. Sie enthalten die maximalen Amplituden der rückprojizierten Wellenfelder für jeden betrachteten Zeitschritt. Die vorgeschlagene Methode nutzt die Kontinuität der maximalen Amplituden. Der Punkt, an dem sich die Pfade am nächsten kommen (oder kreuzen), ist der Ort der Quelle, und die entsprechende Zeit ist die Quellzeit. Diese Methode überwindet erfolgreich Akquisitionsprobleme und liefert selbst in schwierigen Szenarien eine genaue Quellenortung und Quellenanregungszeit.

Im letzten Ansatz wird ein datengesteuerter Hybrid-Workflow vorgestellt, um die Herausforderung des fehlenden Geschwindigkeitsmodells in Zeitumkehrungslokalisationsmethoden, einschließlich der vorgeschlagenen ersten und zweiten Zeitumkehrungslokalisationsansätze, anzugehen. Dieser Ablauf kann gleichzeitig die Quellenposition, die Anregungszeiten und das Geschwindigkeitsmodell unter Verwendung von Wellenfrontattributen der passiven seismischen Daten umkehren. Durch die Kombination von Wellenfronttomographie und Zeitum-

kehrverfahren macht der Arbeitsablauf detaillierte Vorabinformationen überflüssig und ist daher besonders in praktischen Szenarien anwendbar. Zunächst wird eine Reihe von benutzerdefinierten vertikalen Gradientengeschwindigkeitsmodellen entworfen. Anschließend wird die Zeitumkehr verwendet, um die Quellenanregungszeiten für jedes Modell zu schätzen. Danach werden diese Quellzeiten und Gradientenmodelle in der Wellenfronttomographie verwendet. Im zweiten Schritt wird ein Optimierungsverfahren eingesetzt, um das Geschwindigkeitsmodell und die Quellenanregungszeit zu verfeinern. Jede Iteration der Optimierung beinhaltet eine sequentielle Anwendung von Zeitumkehr und Wellenfronttomographie. Im letzten Schritt wird die Quellenposition unter Verwendung des finalen Geschwindigkeitsmodells und der Gauß-gewichteten Kreuzkorrelationsabbildungsbedingung verfeinert. Der vorgeschlagene Arbeitsablauf überwindet die Beschränkungen der Zeitumkehr in Abwesenheit eines initialen Geschwindigkeitsmodells und liefert gute Geschwindigkeitsmodelle und genaue Quellenstandorte und Zeiten.

Contents

List of Figures

1 Introduction

A passive seismic source can be defined as the origin of seismic waves generated by natural processes or human activities in the Earth's interior. Passive seismic source localization is the process of determining the epicenter, depth, and onset time of a source, which is important in both seismology and geoengineering (Shearer, 2019; Li et al., 2020). The sources can be triggered by natural causes like tectonic events or by human activities such as fluid injection/extraction (Eisner et al., 2009; Suckale, 2009). In seismology, earthquakes often occur when the Earth's crust is fractured or cracked (Shearer, 2019). Accurate localization of earthquake sources serves as an important step in understanding the propagation of fractures and facilitates subsequent seismological analysis, including magnitude assessment (Grigoli et al., 2014), the delineation of seismicity models (Sawires et al., 2016), monitoring of seismic source areas (Chen et al., 2022), and inversion of source mechanisms (Li et al., 2020; Kwiatek et al., 2016; Anikiev et al., 2014). In geoengineering, activities such as hydrocarbon production (Grechka et al., 2017; Li and Einstein, 2019), geothermal development (Dyer et al., 2008), coal mining (Cheng et al., 2017; Dong et al., 2016), and CO2 injection (Oye et al., 2013; Zhou et al., 2010; Mazzoldi et al., 2012) cause induced seismicities due to the rock fracturing, underground explosion, or fault activation. High-precision estimation of the excitation time and location of these fractures helps to assess the potential impact on groundwater and nearby subsurface structures (Anderson, 2017; Schultz et al., 2021), which helps to provide timely responses, manage induced seismicity, and optimize industrial processes (Maxwell et al., 2010; Warpinski, 2009; Eisner et al., 2009) to reduce potential environmental hazards and improve oil/gas production.

Early established localization methods are based on traveltime inversion, which uses the arrival times of seismic waves to determine the precise location and onset time of the event. These methods, exemplified by the work of Geiger (1912); Waldhauser and Ellsworth (2000); Font et al. (2004), aim to minimize differences between observed and model-predicted arrival times, resulting in accurate source positions and excitation times. Waldhauser and Ellsworth (2000) proposed a source localization method by iteratively fitting the synthetic differential traveltime to the observed one, known as the "hypoDD" method. The principle of hypoDD is to use the differential traveltime between stations to determine the location of an earthquake. Zhang and Thurber (2003) implemented the double-difference approach to estimate event locations and the underlying velocity model. This type of approach searches for the minimum misfit between modeled and observed traveltimes using either inversion techniques or grid-based searches. While the approach is known for its efficiency and reasonable reliability, it requires accurate picking of seismic phases and arrival times, which can be challenging when dealing with data characterized by a low signal-to-noise ratio (S/N) or data acquired from sparsely distributed receiver networks.

With advances in computing power and storage capacity, time reversal has developed as a common and powerful tool for passive seismic source localization that does not require picking seismic arrival times and phases. Time reversal relies on the seismic wave equation, and it focuses the energy of seismic waves back to the hypocenter if a reasonably accurate velocity model is available (Fink et al., 2002; Gajewski and Tessmer, 2005; Artman et al.,

2010; Chen et al., 2021). This type of method is robust to high noise levels and strong heterogeneity, and can accurately localize sources by performing numerical backpropagation of the recorded waveforms Capdeville and Maron (2006); Yang and Gajewski (2022). Fink (1996) introduced time reversal in acoustics, which laid the foundation for backpropagation simulation in passive seismics. Gajewski and Tessmer (2005) presented the reverse modeling for seismic event characterization. The recorded wavefield of the seismic acquisition was reversed in time and then considered as the boundary value for the reverse modeling. Afterward, time-reversal imaging was introduced for source localization. The imaging technique is independent of the source excitation time and has been successfully applied to various seismic events ranging from glacial earthquakes to microseismic activity to improve the reliability of time-reversal source localization (Artman et al., 2010; Larmat et al., 2008). Larmat et al. (2008) approximated the location and focal mechanism of relatively large glacial earthquakes from time-reversal images. Artman et al. (2010) introduced a migration strategy for source localization, which adopts the autocorrelation imaging condition. The time-reversal imaging method is robust and reliable for long-period event localization (O'Brien et al., 2011; Folesky et al., 2015; Price et al., 2015).

1.1 Gaussian-weighted imaging condition

However, the established time-reversal imaging method like auto-correlation imaging (Steiner et al., 2008; Artman et al., 2010), suffers from strong imaging artifacts and low spatial resolution (Yang and Gajewski, 2021b, 2022). To improve the quality of time-reversal images, Sun et al. (2016); Chen et al. (2021) proposed the grouped crosscorrelation imaging method. Although the spatial resolution of the image is improved, artifacts are still present. The artifacts typically degrade the image quality and can potentially lead to subsequent misinterpretation, resulting in incorrect source location. In addition, the loss of data illumination in the grouped crosscorrelation imaging is inevitable because of receiver grouping. Selecting receivers for each group is a challenging task.

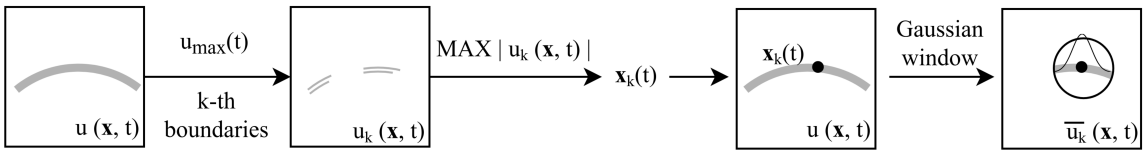


Figure 1.1: Schematic illustration of the Gaussian-weighted wavefield calculation. For each time step, the full back-propagated wavefield $u(\boldsymbol{x}, t)$ is first divided into parts, $u_k(\boldsymbol{x}, t)$, according to the amplitude boundaries calculating by equation 2.3 given in **Chapter 2**. Then the spatial location $\boldsymbol{x}_k(t)$ for the maximum absolute amplitude of the k-th part of the wavefield $u_k(\boldsymbol{x}, t)$ is calculated by equation 2.5 given in **Chapter 2**. The Gaussian-weighted wavefield $\overline{u}_k(\boldsymbol{x}, t)$ is finally computed by applying a Gaussian window to the full wavefield $u(\boldsymbol{x}, t)$. The Gaussian window centered at the position $\boldsymbol{x}_k(t)$ corresponds to the spatial extent of the dominant wavelength of the event.

To address the issues, the Gaussian-weighted crosscorrelation imaging condition (Yang et al., 2023) is presented in **Chapter 2**. The calculation of each Gaussian-weighted wavefield at each time step is sketched in Figure 1.1. Zero-lag cross correlating the weighted wavefields at each spatial point produces an image for this time step. The Gaussian-weighted crosscor-

relation image is obtained by summing the images for all time steps. The proposed method eliminates the need to group receivers and can provide clear and high-resolution source images by time-reversal imaging. The details of the method are described in **Chapter 2**.

1.2 Maximum-amplitude path method

Aside from the issue of poor imaging quality mentioned above, time-reversal imaging methods, including those mentioned in **Chapter 2**, usually require a large-aperture seismic acquisition to ensure sufficient data illumination. The established time-reversal imaging methods (Artman et al., 2010; Nakata and Beroza, 2016; Yang et al., 2023) struggle to accurately localize sources using the passive seismic data acquired from sparse and small-aperture seismic acquisitions, however, such challenging acquisitions are often met in field cases.

To solve the problem of acquisition limitation, the maximum-amplitude path method (Yang and Gajewski, 2022) is introduced in **Chapter 3**. The construction of the maximum-amplitude path is shown in Figure 1.2. The path is obtained by tracking the maximum amplitudes of the wavefront formed by the back-projected wavefields. I construct three maximum-amplitude paths from the time-reversed data of three receiver patches selected from the recorded traces. The paths cross or display the closest proximity to each other, which defines the source location. The time at this location corresponds to the source excitation time. The proposed method allows the accurate estimation of source location and excitation time in the sparse and small-aperture seismic acquisitions. I will describe the details of the method in **Chapter 3**.

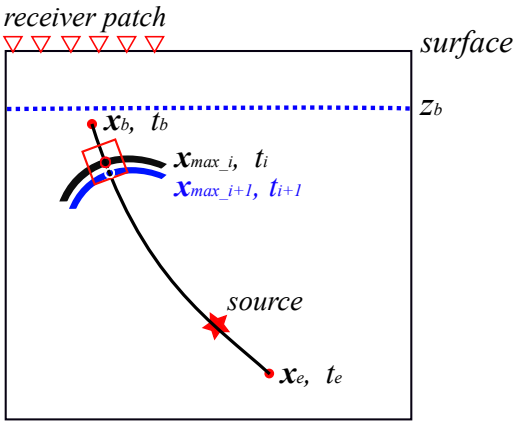


Figure 1.2: Schematic sketch of the maximum-amplitude path construction. The threshold Z_b is a user-defined parameter for determining the maximum-amplitude path start $(\mathbf{x}_b, t_b)$. Two snapshots at subsequent time steps t_i and t_{i+1} are displayed. The black curve describes the maximum-amplitude path. Assuming that the maximum amplitude for the time step t_i is located at $\mathbf{x}_{max\_i}$, the maximum for the next time step $\mathbf{x}_{max\_i+1}$, denoted by the white circle, is determined by scanning the amplitude of the wavefront for this time step, that is, t_{i+1}, in the neighborhood of the previous maximum. This neighborhood, indicated by the red rectangular box, is centered at the previous maximum and oriented in propagation direction with side lengths corresponding to the prevailing wavelength of the signal.

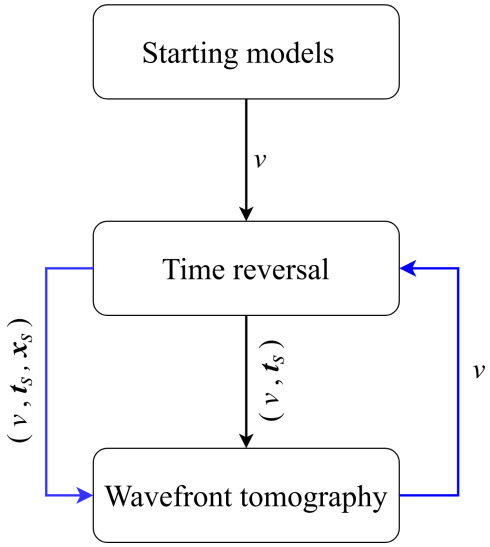


Figure 1.3: The flowchart of the hybrid workflow, sketching main procedures and their respective outcomes. The black and blue arrows mark the first and second steps of the hybrid workflow, respectively. The first step starts with a set of user-defined vertical gradient models, each using time reversal to determine source excitation times $\boldsymbol{t}_s$. Using these source times, along with the corresponding user-defined gradient models, wavefront tomography is performed to produce the initial model for the second step. In the second step, an optimization process unfolds: time reversal and wavefront tomography are sequentially performed, with time reversal estimating source excitation times $\boldsymbol{t}_s$ and locations $\boldsymbol{x}_s$, and wavefront tomography providing the velocity model v.

1.3 Source localization and joint velocity model invesion

The accuracy of source location and onset time obtained by time-reversal methods, including those proposed in **Chapters 2 and 3** (Gajewski and Tessmer, 2005; Artman et al., 2010; Yang and Gajewski, 2022; Yang et al., 2023), is highly dependent on the quality of the provided underlying velocity model, however, the velocity model is often unknown in field work. Image-domain wavefield tomography methods, using the PS energy imaging function Oren and Shragge (2021) or the source-focusing function Song et al. (2019), aim to minimize the image residues derived from these functions to simultaneously invert for the source location and velocity model. However, these tomography methods rely on high-quality source images, which require dense and well-distributed seismic data for effective source imaging. Another approach to joint inversion is full waveform inversion (e.g., Wang and Alkhalifah, 2018; Sun et al., 2016), which yields fine-scale velocity variations, but is sensitive to the initial velocity model, susceptible to cycle skipping, and requires high-quality data. Additionally, these methods typically do not include inversion of the source excitation time. To overcome challenges arising from poor data quality and unknown source time, Diekmann et al. (2019) introduced a robust alternative using wavefront attributes (Hubral, 1983; Schwarz et al., 2016) for the joint inversion of source location, excitation time, and velocity model. This data-driven approach requires only prior information of the near-surface velocity and is capable of handling data with high noise levels and/or sparsely

sampled (Diekmann et al., 2019; Yang and Gajewski, 2021a). However, the approach faces difficulties in accurately estimating the source time, particularly in scenarios with strong lateral heterogeneity.

To address the issues of the lack of velocity model information and the limitations of the conventional approach presented by Diekmann et al. (2019), I combine the potential to determine the velocity model by wavefront attributes with the independence on the source times of time-reversal methods. Figure 1.3 display the main procedures and their respective outcomes of the hybrid workflow (for details, see **Chapter 4**). The proposed workflow (Yang et al., 2024) performs time reversal and wavefront tomography sequentially. Time reversal (Gajewski and Tessmer, 2005; Yang et al., 2023) determines the source parameters, while wavefront tomography (Duveneck, 2004; Bauer et al., 2019a) provides the velocity model. It is worth mentioning that I applied the Gaussian-weighted crosscorrelation imaging conditions from **Chapter 2**.

Structure of the thesis

Following the introduction given in this chapter, **Chapter 2** contains the paper *Gaussian-weighted crosscorrelation imaging condition for microseismic source localization*, which was published in 2023 in the journal *Geophysics* (Yang et al., 2023). This work introduces the Gaussian-weighted crosscorrelation imaging condition to improve passive seismic source imaging quality. The proposed imaging condition generates high-resolution images displaying little to no imaging artifacts, which is particularly suitable for noisy data acquired from sparse acquisitions and provides good results for source clusters. The work successfully addresses the problem of imaging artifacts in time-reversal imaging.

Chapter 3 consists of the paper *Seismic source localization with time-reversal and maximum amplitude path for sparse and small-aperture acquisitions*, which was published in 2022 in the journal *Geophysics* (Yang and Gajewski, 2022). Since the time-reversal imaging methods discussed in **Chapter 2**, often face the challenge associated with sparse and small-aperture seismic acquisition, this work proposes the maximum-amplitude path method to overcome this challenge by constructing maximum-amplitude paths. The paths consist of the maximum amplitudes of the back-projected wavefronts for each time step considered. The source location and excitation time are determined by calculating the closest proximity of the paths.

In **Chapter 4**, the paper *Integrated wavefront tomography and time reversal for source location, excitation time, and velocity model building* is presented, which has already been submitted to the journal *Geophysics* in 2024 (Yang et al., 2024) and is under review (since February 2024). This work introduces a data-driven hybrid workflow by combining wavefront tomography and time reversal methods. The time-reversal methods used in the workflow were proposed in previous chapters, which are independent of source time. The proposed workflow can simultaneously invert for source location, excitation time, and velocity model. It overcomes the common limitation of the previous time-reversal localization methods, that is, the previous methods require a reasonably accurate velocity model, which is often unknown in practice, for accurate source location estimation.

In **Chapter 5**, I summarize the results and conclusions of the thesis. This chapter serves to consolidate the findings and implications discussed throughout the research. Finally, in

Chapter 6, I offer an outlook on potential future work, suggesting directions for continued research and development.

Contributions of co-authors

Chapters 2 and 3 of this thesis are papers that have been published in scientific journals, and Chapter 4 is based on a manuscript that will be published in the near future. In the following, I briefly describe the contributions of the co-authors of each paper.

The results presented in Yang and Gajewski (2022) were generated by my efforts. I wrote the computer programs and the original manuscripts for the journal paper. I authored the computer programs and original manuscripts for the journal paper, with supervision from my supervisor Prof. Dr. Dirk Gajewski, who provided continuous guidance and participated in discussions. After substantial results were achieved, he made efforts to improve the structure and clarity of my written materials and refined the manuscripts. I made the final refinements to the manuscript. Although not listed as co-authors, I would like to explicitly acknowledge the work of Dr. Alex Bauer and Dr. Claudia Vanelle, who produced comments and suggestions that helped to improve the manuscript.

The idea and the implementation of Gaussian-weighted crosscorrelation imaging presented in Yang et al. (2023) were my own work. Based on the previous research on time-reversal localization methods, I developed the 2D and 3D codes for time-reversal imaging. Prof. Dr. Dirk Gajewski offered insights by discussing ideas and refining the manuscript. Yujiang Xie contributed to discussions on mathematical methods and proposed methods, refining the manuscript. The final adjustments to this manuscript were also made by me.

Recently, I submitted the third paper (Yang et al., 2024) presented in Chapter 4. This paper addressed a common issue of the previous two published papers and constituted the final chapter of my dissertation. I programmed this work and wrote the original manuscripts for this journal paper. Prof. Dr. Dirk Gajewski supervised the entire work and refined the manuscript. Yujiang Xie contributed to the continuous discussion of the proposed methods and refined the written materials. I made the final refinements to the manuscripts and composed this thesis by combining these three journal papers.

2 Gaussian-weighted crosscorrelation imaging condition for microseismic source localization

Abstract

Time-reversal imaging is a powerful method for localization of microseismic events. Conventional time-reversal imaging methods such as autocorrelation imaging or grouped crosscorrelation imaging may suffer from imaging artifacts, e.g., caused by less refined velocity models, noise-contaminated data, and data acquired from sparse receiver networks. These artifacts typically reduce imaging quality and may cause subsequent misinterpretation leading to false positives. To address these issues, we develop a new imaging condition that comprises three steps for each time step. First, we divide the back-propagated wavefield into parts according to their maximum absolute amplitudes. Second, the amplitudes of the back-propagated wavefield are weighted by a Gaussian function with the spatial extent of the prevailing wavelength of the event, centered at the absolute maximum of that part of the wavefield. Finally, we zero-lag crosscorrelate these weighted wavefields at each space point to obtain an image for this time step. The final image is gained by summing the images for each time step. This process collects all energy concentrations along the back projection process, and the energy on the wavefront overlaps and collapses at the hypocenter leading to high-resolution images displaying little to no imaging artifacts. Numerical examples using the Marmousi-II and the 3D SEG overthrust models and a 3D field data example indicate the performance of our method. High-resolution low-noise source images allow unique identification of sources even for source clusters, noisy data, and sparse acquisitions. The source localization errors are smaller than the dominant wavelength of the signal, where a smooth model with a mean velocity error of approximately 5% was considered in the synthetic examples, and a homogeneous model was used in the field data example.

2.1 Introduction

Microseismic source localization is an important tool, e.g., for hydraulic fracturing or reservoir monitoring. Observing fluid-induced seismicity can help to characterize reservoirs for oil and gas exploration. Accurate fracture localization helps to control fracturing processes (Maxwell et al., 2010; Warpinski, 2009). Conventional source localization methods are traveltime inversions, which use arrival times for microseismic monitoring. By minimizing the differences between the observed and modeled arrival times, accurate source locations and origin times of the events are obtained (Geiger, 1912; Waldhauser and Ellsworth, 2000; Eisner et al., 2009). To identify seismic events from continuously recorded data, arrival times are obtained by crosscorrelating template waveforms with time-windowed data out of the continuous recording (Shelly et al., 2007). Zhang and Thurber (2003) use the double-difference technique to approximate event locations and the velocity model. This method searches for minimum misfit between the modeled and observed traveltimes based on either

an inversion or a grid search. Although efficient and reasonably reliable, this method requires accurately picking seismic phases and arrival times, which may be compromised for data with a low signal-to-noise ratio (S/N) or data acquired from sparse receiver networks.

Advances in computational power and imaging techniques enable the use of waveforms rather than the arrival times for source locations. The common and powerful types of methods are time-reversal imaging and full-waveform inversion methods using wave equations to backpropagate seismic waves to the hypocenter. Time-reversal imaging techniques require a reasonably accurate velocity model (Fink et al., 2002; Yang et al., 2019), and full-waveform inversion methods simultaneously invert for the source image and velocity distribution (Kaderli et al., 2015; Wang and Alkhalifah, 2018; Song and Alkhalifah, 2019). In this paper, assuming the velocity model is known, only time-reversal imaging methods are considered. Because the back-propagation process starts when the event is injected into the model, it is source-time independent and does not require the picking of arrival times. These types of methods are more robust and reliable for long-period event localization (O'Brien et al., 2011) and are generally suitable for passive seismic imaging (Folesky et al., 2015; Price et al., 2015). Artman et al. (2010) introduce a migration strategy for source locations, which adopts the auto-correlation imaging condition. However, the image displays strong artifacts and suffers from low spatial resolution (Yang and Gajewski, 2021b; Artman et al., 2010; Yang and Gajewski, 2022). Sun et al. (2016) and Chen et al. (2021) propose the grouped cross-correlation imaging method, which first groups several receivers and then extrapolates the wavefield for each group. Although the spatial resolution of the image is improved, artifacts are still present. The lack of wave illumination in the grouped crosscorrelation imaging is inevitable because of receiver grouping. Moreover, selecting receivers for each group is a challenging task.

We propose a method to mitigate the previously mentioned issues via a Gaussian-weighted crosscorrelation imaging condition. It does not require grouping receivers and can provide clear and high-resolution source images by time-reversal imaging. Numerical examples using the Marmousi-II (Martin et al., 2002) and the 3D SEG overthrust models (Aminzadeh et al., 1997), and a field 3D data example illustrate the proposed method for laterally heterogeneous media. Effects caused by velocity errors, or low S/N data, and data acquired from sparse receiver networks are investigated. The proposed method is compared with two published time-reversal imaging methods, namely, the autocorrelation imaging method (Steiner et al., 2008; Artman et al., 2010) and the grouped crosscorrelation imaging method (Sun et al., 2016; Chen et al., 2021). Because multiple sources in close spatial proximity may occur during hydraulic fracturing (Maxwell et al., 2010), in synthetic tests, we consider this situation as well.

2.2 Time-reversal source imaging

The velocity-stress acoustic wave equation with constant density is used for forward modeling and time-reversal imaging (Aki and Richards, 2002; Cerjan et al., 1985). The recorded data are reversed in time and served as boundary values for time-reversal imaging. We assume that a velocity model consistent with the data is available for the imaging. In the following sections, we first consider two published imaging conditions and then introduce the new Gaussian-weighted imaging condition.

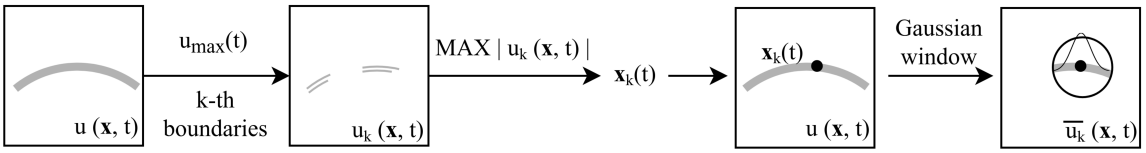


Figure 2.1: Schematic illustration of the Gaussian-weighted wavefield calculation. For each time step, the full back-propagated wavefield $u(\boldsymbol{x}, t)$ is first divided into parts, $u_k(\boldsymbol{x}, t)$, according to the amplitude boundaries calculating by equation 2.3. Then the spatial location $\boldsymbol{x}_k(t)$ for the maximum absolute amplitude of the k-th part of the wavefield $u_k(\boldsymbol{x}, t)$ is calculated by equation 2.5. The Gaussian-weighted wavefield $\overline{u}_k(\boldsymbol{x}, t)$ is finally computed by applying a Gaussian window to the full wavefield $u(\boldsymbol{x}, t)$. The Gaussian window centered at the position $\boldsymbol{x}_k(t)$ corresponds to the spatial extent of the dominant wavelength of the event.

2.2.1 Imaging condition

Imaging conditions represent an important step in time-reversal imaging. We review briefly two well established variants: the autocorrelation imaging condition and the grouped cross-correlation imaging condition, which are subsequently used for comparison in the example sections. Then we introduce the Gaussian-weighted imaging condition.

Auto-correlation imaging condition

The autocorrelation imaging condition (Artman et al., 2010) is a widely used time-reversal imaging condition for microcosmic monitoring. The back-propagated wavefield $u(\boldsymbol{x}, t)$ is obtained by simultaneously back-propagating the time-reversed recorded wavefields at all receivers. Zero-lag autocorrelation of the wavefield for each time step is performed. The summation of the autocorrelated wavefields for all times provides the autocorrelation image I_1, which may be expressed as

$$I_1(\boldsymbol{x}) = \sum_t \left(u(\boldsymbol{x}, t)\right)^2 \quad , \tag{2.1}$$

where $\boldsymbol{x}$ represents the spatial coordinates and t is the wave-propagation time. The maximum value of this image is considered to be the event location. The autocorrelation imaging condition is sensitive to data with strong noise and/or data acquired from sparse receiver networks, which may result in strong imaging artifacts.

Grouped cross-correlation imaging condition

The grouped cross-correlation imaging condition (Nakata and Beroza, 2016; Sun et al., 2016; Chen et al., 2021) improves the image quality when compared with the autocorrelation imaging condition. This imaging condition requires first grouping the receivers. The adjacent grouping strategy (Bai et al., 2022) is commonly used to acquire the receiver groups when the grouped crosscorrelation imaging condition is applied. The wavefields recorded for each group are back propagated. Assuming that the receivers are divided into N groups, the expression for the grouped zero-lag crosscorrelation imaging I_2 is

$$I_2(\boldsymbol{x}) = \sum_t \prod_g u_g(\boldsymbol{x}, t) \quad , \quad 1 \leq g \leq N \quad , \tag{2.2}$$

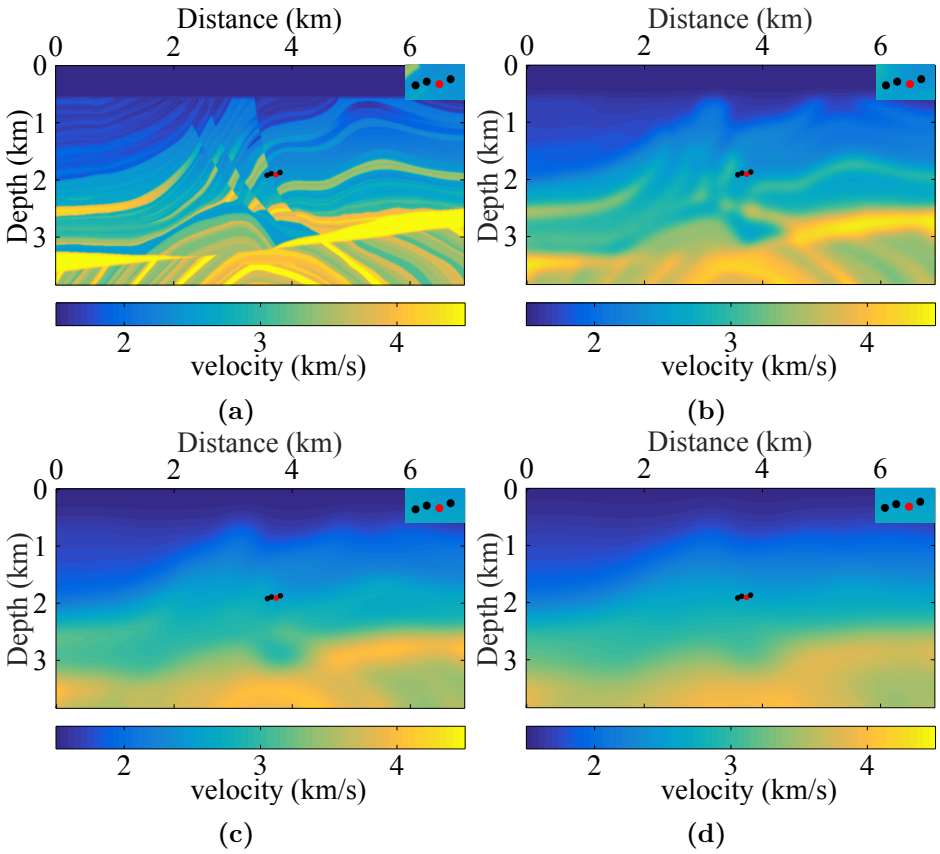


Figure 2.2: Velocities for a subset of the Marmousi-II model. (a) True velocity model. Smoothed models with the 2D Gaussian window of 300×300 m^2 (b), 600×600 m^2 (c), and 900×900 m^2 (d). Dots indicate the source locations. The right top panel is a close-up of four clustered sources.

where g denotes the group number and $u_g(\boldsymbol{x}, t)$ corresponds to the back-propagated wavefield obtained by the g-th receiver group. The wavefield of each receiver group goes through the hypocenter. Multiplying these wavefields for each space and time sample and summing over all times provides high-resolution images where imaging artifacts are reduced. The event location is estimated by scanning the maximum value of the image. Because the receiver groups used for the back propagation comprise only a small portion of the total data, illumination is decreased and some artifacts may still remain, particularly for subsurface structures with strong heterogeneity or sparse receiver acquisitions. Moreover, choosing the receiver groups is a nontrivial task.

Gaussian-weighted imaging condition

To circumvent the limitations of the preceding time-reversal imaging methods, we suggest the Gaussian-weighted crosscorrelation imaging condition. The motivation of the procedure is based on complicated amplitude patterns on the back-propagated wavefront. This pattern may be caused by focusing and defocusing the effects of the back-propagated wavefront in heterogeneous media and/or complicated radiation patterns of the source. For each time

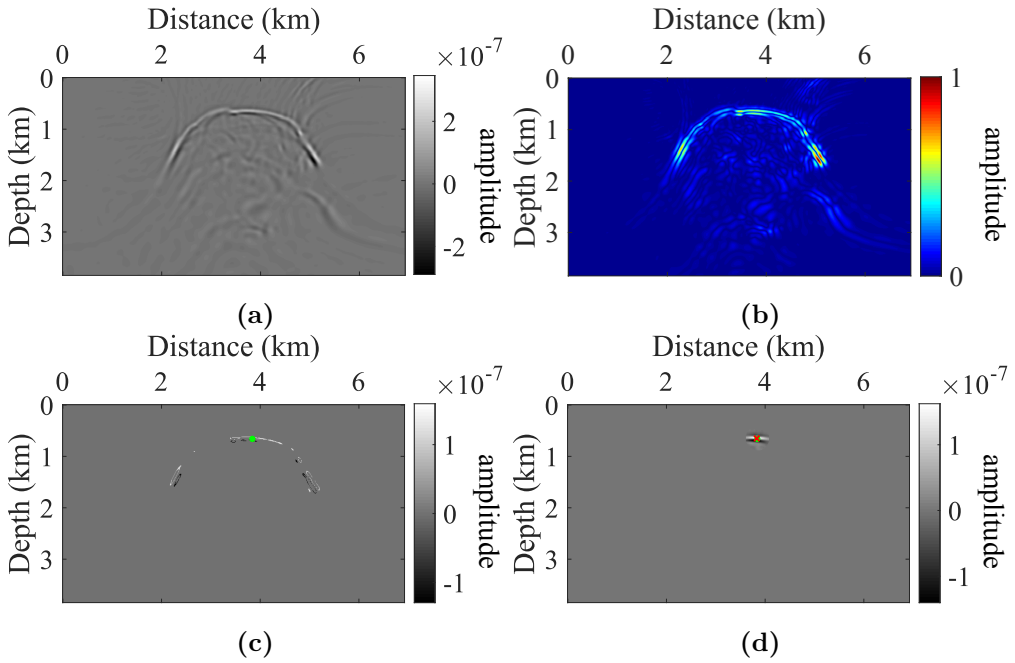


Figure 2.3: Images of the full back-propagated wavefield (a), the normalized absolute full back-propagated wavefield (b), the first selected wavefield (c), and the first Gaussian-weighted wavefield (d) at the time of 0.6 s. The green dot and red × represent the position of the maximum absolute amplitude of the first selected wavefield and Gaussian-weighted wavefield.

step, we decompose the total back-propagated wavefield in Gaussian-shaped contributions, where the spatial extent of these contributions is linked to the prevailing wavelength of the event. The individual wavefield contributions are classified according to their maximum absolute amplitudes. The normalized wavefield contributions are multiplied (zero-lag crosscorrelation) to provide the image for this time step. The final image is obtained by summing the images for all time steps. Due to focusing at the hypocenter, the wavefield contributions increasingly overlap, resulting in large values of image amplitudes, whereas the false maxima caused by image artifacts or arbitrary focusing are considerably smaller.

The formal realization of the procedure is as follows. For each time step, we scan the maximum absolute amplitude of the back-propagated wavefield $u(\boldsymbol{x}, t)$. Then, we select N wavefields from the total wavefield according to predefined lower and upper amplitude bounds. The normalized absolute amplitude of the k-th selected wavefield $u_k(\boldsymbol{x}, t)$ from the full wavefield satisfies the following expression for each time step:

$$e^{-R^2} + (k-1)\left(\frac{1-e^{-R^2}}{N}\right) \leq \frac{|u_k(\boldsymbol{x},t)|}{u_{max}(t)} \leq e^{-R^2} + k\frac{1-e^{-R^2}}{N} \quad , \quad 1 \leq k \leq N \quad , \tag{2.3}$$

where $u_{max}(t)$ denotes the maximum absolute amplitude of the full back-propagated wavefield $u(\boldsymbol{x}, t)$:

$$u_{max}(t) = \max\left\{|u(\boldsymbol{x},t)|\right\} \quad , \tag{2.4}$$

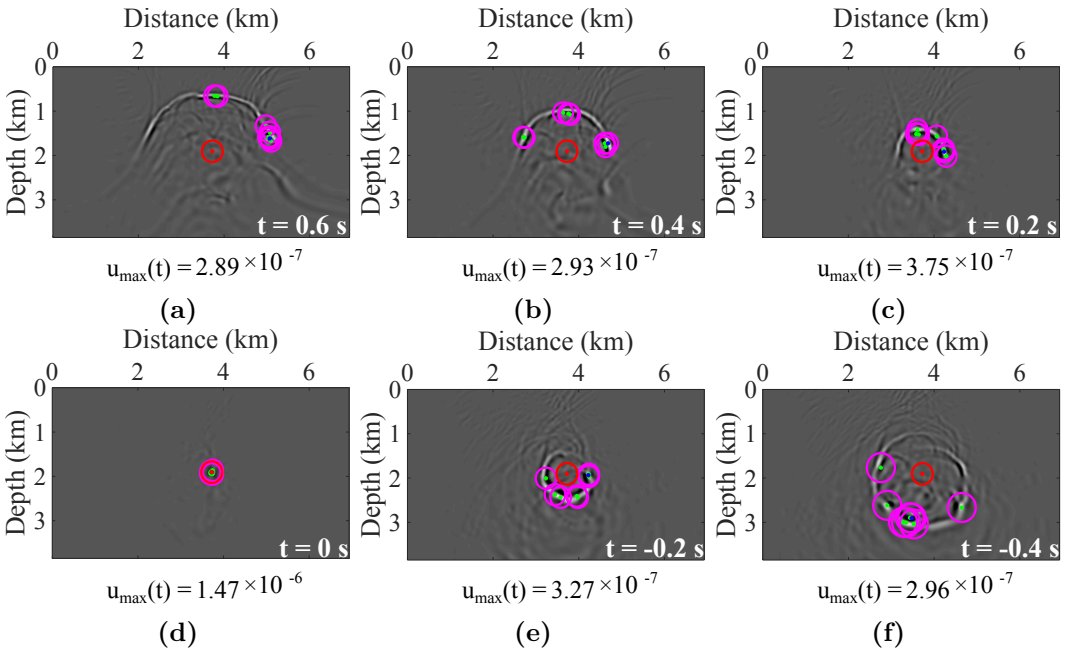


Figure 2.4: Images of the back-projected wavefields at the time of 0.6 s (a), 0.4 s (b), 0.2 s (c), 0 s (d), -0.2 s (e), and -0.4 s (f) for the Marmousi-II model. Blue and green dots denote picks, corresponding to the maximum absolute amplitude of the whole back-propagated wavefield and the selected back-propagated wavefields at the considered time step, respectively. Red dots represent the exact source location. Red and purple circles, with radii of the dominant signal wavelength, are centered at the picks and the exact source location, respectively.

where e^{-R^2} is a Gaussian shape factor and R and N are tuning parameters. A common strategy for tuning the parameters is a control variate. We first set the parameter N to a small value, and then we fix the N and tune the parameter R in the range $(0, 1]$. The same process is applied to other N and R sets until the preferred result is obtained, which is based on the user's experience to judge the quality of the image, the same as the other imaging methods. In principle, the larger the parameter N, the higher the image resolution and the fewer the imaging artifacts. As a good general rule, R is set to one. We can decrease R or N, or decrease both when there is insufficient focusing of the image, and increase N when the image has strong artifacts.

After the wavefield selection process, the spatial position of the maximum absolute amplitude of each selected wavefield for each time step is picked:

$$\max\left\{\left|u_k(\boldsymbol{x},t)\right|\right\} \Rightarrow \boldsymbol{x}_k(t) \quad , \quad 1 \leq k \leq N \quad , \tag{2.5}$$

where $\boldsymbol{x}_k(t)$ represents the position of the maximum absolute amplitude of the k-th selected wavefield for the corresponding time step. For each picked location of the maxima $\boldsymbol{x}_k(t)$, the surrounding wavefield selected from the full back-propagated wavefield is weighted by a Gaussian function centered at that point. The spatial extent of each Gaussian window is a circle in two dimensions and a sphere in three dimensions with a radius equal to

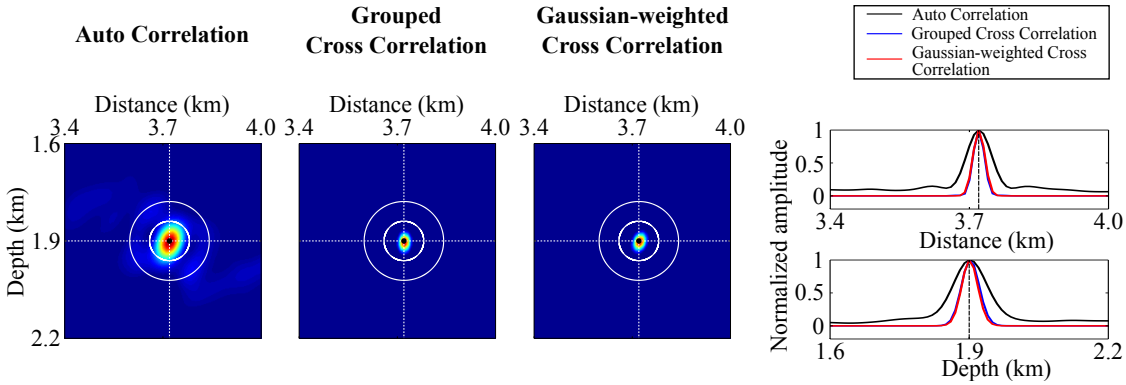


Figure 2.5: Comparison of three time-reversal imaging methods for the source at $x = 3.72$ km, $z = 1.9$ km in the Marmousi-II model. The black dots denote the spatial locations of the maximum amplitude of the images. White circles are centered at the true source location with radii of a quarter and a half of the dominant wavelength. The right column shows horizontal (top) and vertical (bottom) amplitude sections through the source location.

the dominant wavelength. The dominant wavelength is calculated at the position of the maximum absolute amplitude of the total wavefield. The spatial extension is considered the resolution limit of the considered event. The kth-weighted wavefield $\overline{u}_k$ is written as

$$\overline{u}_k(\boldsymbol{x}, t) = u(\boldsymbol{x}, t) G_k(\boldsymbol{x}, t) \quad , \quad 1 \leq k \leq N \quad , \tag{2.6}$$

where the weighting function G_k for the k-th point is given by

$$G_k(\boldsymbol{x}, t) = \begin{cases} e^{-\frac{\left(\boldsymbol{x}-\boldsymbol{x}_k(t)\right)^2}{\alpha\left(\lambda(t)\right)^2}} & , \quad |\boldsymbol{x} - \boldsymbol{x}_k(t)| \leq \lambda(t) \\ 0 & , \quad \text{otherwise} \end{cases} \tag{2.7}$$

For each time step, the dominant wavelength $\lambda(t)$ describes the spatial extent of the back-propagated wavefield, which is preserved. The parameter α is another tuning parameter to control the shape of the window function. Note that when we adjust the R and N using the preceding scheme, the α is fixed, likewise, we fix the R and N when tuning the α. To assure each selected and weighted wavefield has the same maximum, we choose α to be one. In addition, the calculation of the kth-weighted wavefield $\overline{u}_k$ at each time step is sketched in Figure 2.1.

For each time step, we compute the zero-lag crosscorrelation (corresponding to a simple multiplication) of the weighted wavefields $\overline{u}_k$. The final image is obtained by summing the images of all time steps up to the predefined maximum time T, i.e.,

$$I_3(\boldsymbol{x}) = \sum_t \left| \prod_k \overline{u}_k(\boldsymbol{x}, t) \right| \quad , \quad 1 \leq k \leq N \quad , \tag{2.8}$$

This process provides a new imaging condition, which we call the Gaussian-weighted cross-correlation imaging condition. The new imaging condition exploits the concentration of the wave energy during the back-propagation process, which significantly enhances the energy

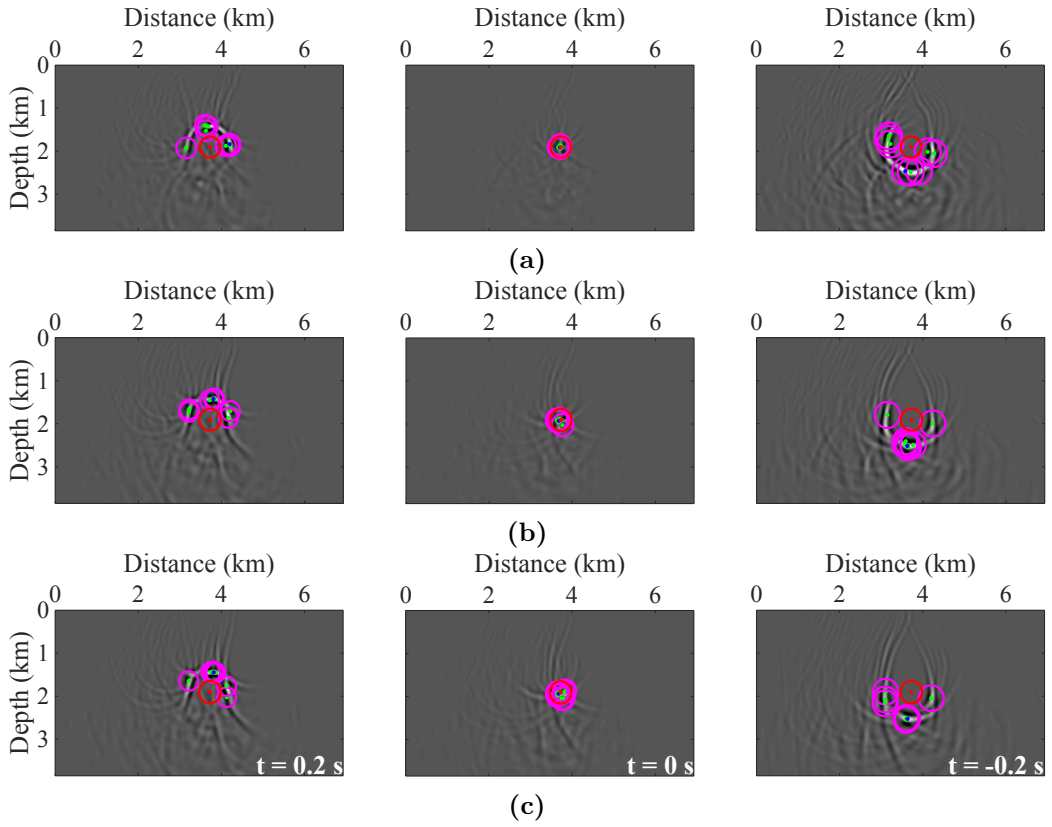


Figure 2.6: Images of the back-projected wavefields at the time of 0.2 s, 0 s, and -0.2 s for the model smoothed with the 2D Gaussian window of 300×300 m^2 (a), 600×600 m^2 (b), and 900×900 m^2 (c). Blue and green dots denote picks, corresponding to the maximum absolute amplitude of the whole back-propagated wavefield and the selected back-propagated wavefields at the considered time step, respectively. The red dots represent the exact source location. Red and purple circles, with radii of the dominant event wavelength, are centered at the picks and the exact source location, respectively.

in the neighborhood of the source position within the radius of the dominant wavelength. Moreover, it greatly decreases the amplitudes outside the hypocentral area. False maxima caused by arbitrary focusing of scattered energy in heterogeneous media are unlikely to be observed. Because absolute amplitudes are considered, phase changes caused by complex source radiation are no an issue. In the following section, we investigate the performance of the new imaging condition under controlled conditions.

2.3 Numerical examples

We perform numerical experiments using the Marmousi-II and 3D SEG overthrust models. For all examples, we use explosive sources and the standard Ricker wavelet. A dominant frequency of 10 Hz is applied for the source-time function. A time shift is applied to simulate the minimum phase signal. Synthetic data are generated using the velocity-stress acoustic

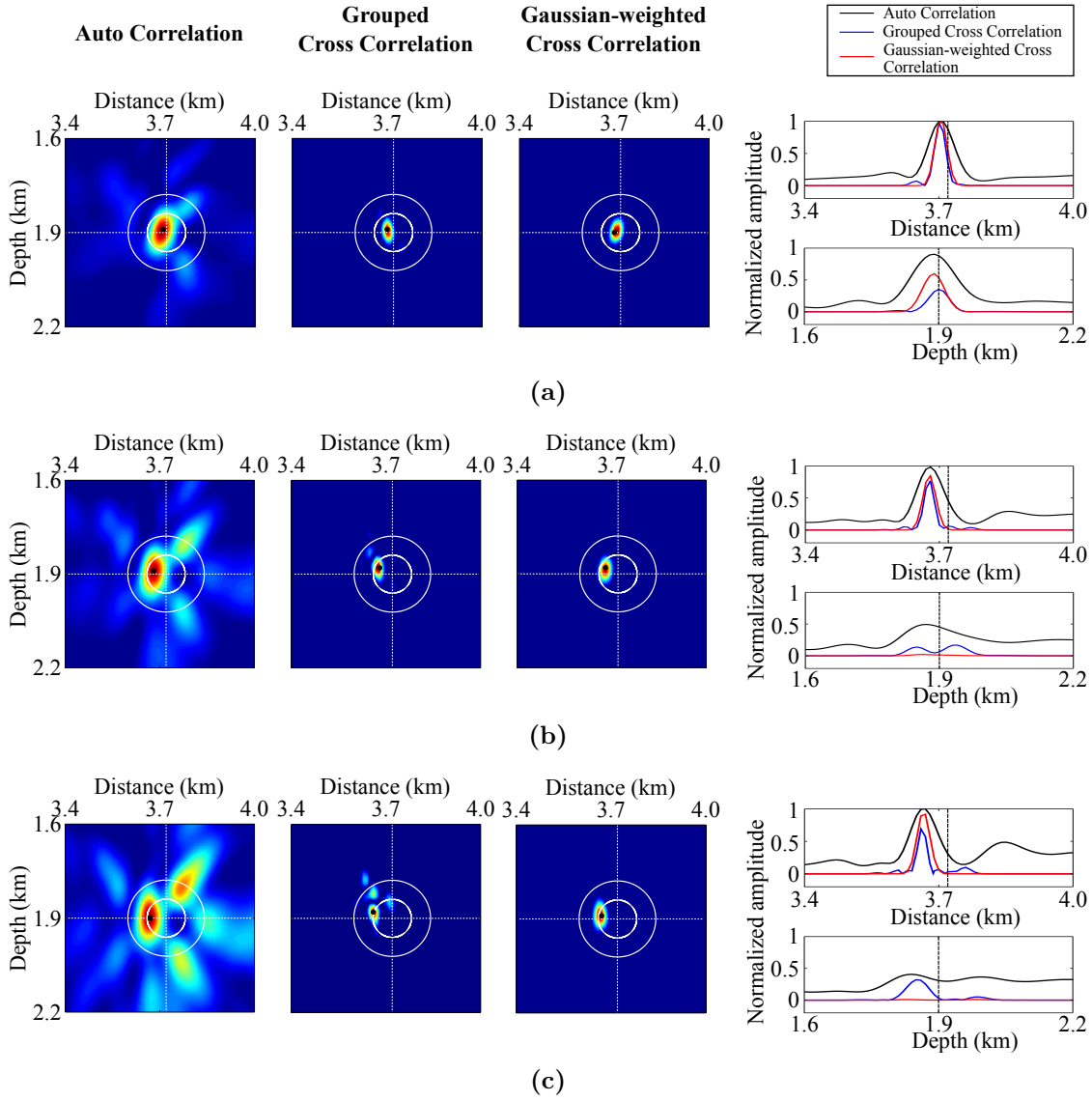


Figure 2.7: Comparison of three time-reversal imaging methods in the presence of velocity errors. Images at the true source location for the Marmousi-II model smoothed with the 2D Gaussian window of 300×300 m^2 (a), 600×600 m^2 (b), and 900×900 m^2 (c), are obtained by auto-correlation imaging, grouped cross-correlation imaging, and Gaussian-weighted cross-correlation imaging. The black dots denote the estimated source locations. The white circles, centered at the actual source location, have radii of a quarter and a half of the dominant wavelength. The right column shows horizontal (top) and vertical (bottom) amplitude sections through the source location.

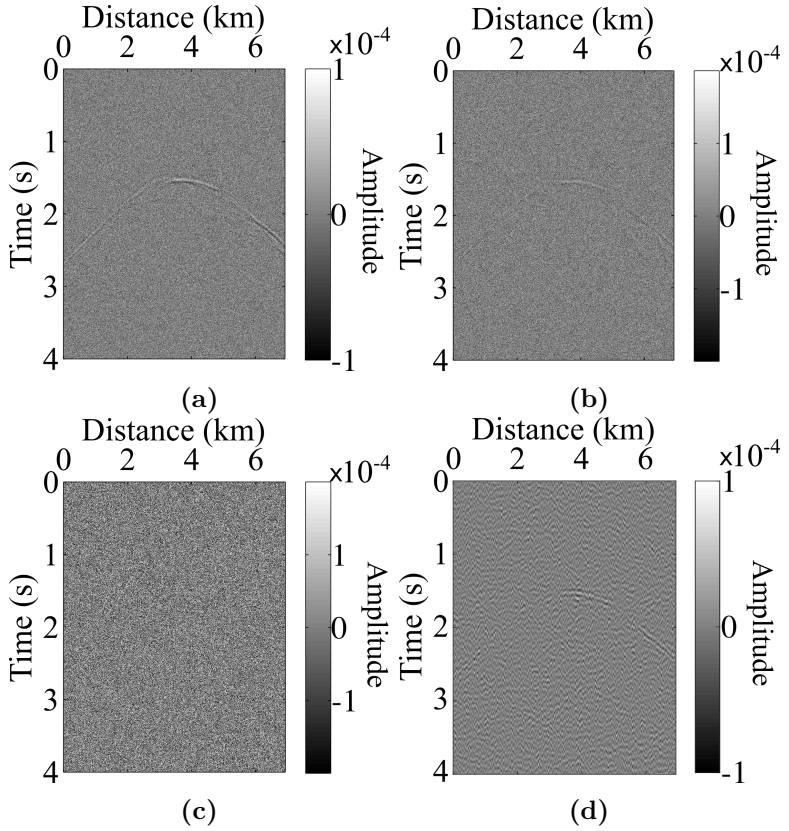


Figure 2.8: Data for the source at $x = 3.72$ km, $z = 1.9$ km in the Marmousi-II model. Noisy data with S/N of 0.4 (a), 0.2 (b), and 0.1 (c). Band-pass filtered data (6 - 17 Hz) with S/N of 0.1 (d).

wave equation (Aki and Richards, 2002; Gajewski and Tessmer, 2005) with constant density. The time-reversed data are used as input to back propagate the wavefield using the same wave equation. Unless noted, random noise with an S/N of 30 was added to the data.

Forward modeling was performed using the true velocity model to simulate the observed data, whereas, for the time-reversal imaging, a smoothed version of the true model is used. An optimized tuning parameter set of R, N and α can be found through some interactive tests. In all examples, R, N, and α are 1, 8, and 1, respectively, and the imaging results are set to the same clip $[0.1, 1]$, unless stated otherwise. For the grouped crosscorrelation imaging, we use eight receiver groups. Results of the two conventional time-reversal imaging methods, the autocorrelation imaging (Artman et al., 2010) and the grouped crosscorrelation imaging method (Chen et al., 2021), are compared with the proposed method. To produce the image for the grouped crosscorrelation imaging method, in all examples, we use the absolute value of the image computed in equation 2.2.

Surface observations are considered in the examples. For input data with poor S/N, the back-propagated wavefield displays insufficient focusing of the back-propagated energy. A noise mask is present close to the injection points of the time-reversed data. The imaging process starts at a depth when sufficient focusing is observed. A reasonable threshold for

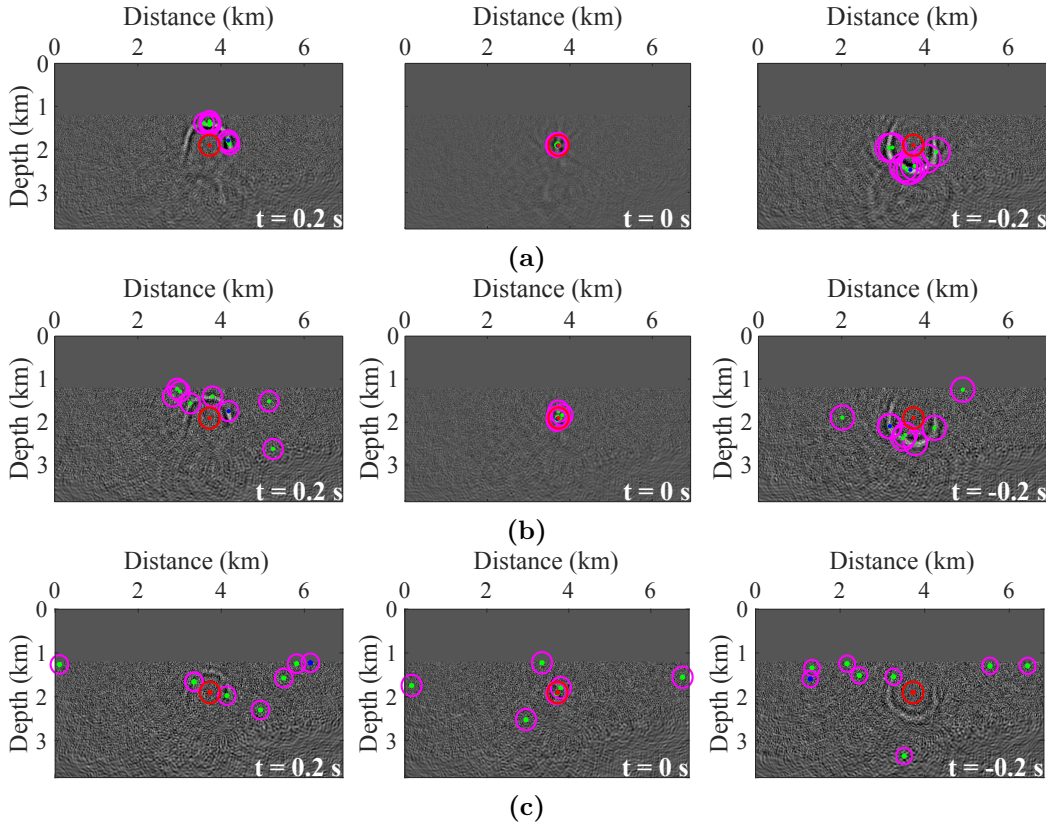


Figure 2.9: Images of the back-projected wavefields at the time of 0.2 s, 0 s, and -0.2 s for noisy data with S/N of 0.4 (a), 0.2 (b), and 0.1 (c). Blue and green dots denote the picks, corresponding to the maximum absolute amplitude of the whole back-propagated wavefield and the selected back-propagated wavefields at the considered time step, respectively. Red dots represent the exact source location. Red and purple circles with radii of the dominant event wavelength are centered at the picks and the exact source location, respectively.

this depth may be determined by visual inspection of images. For the numerical study, controlled conditions are applied, i.e., source locations, source excitation times, and the velocity model are known. The present work does not include an alternative approach (Oren and Shragge, 2021) to mitigate imaging artifacts and minimize amplitude transitions in depth. However, this approach will be integrated into future research.

Data with strong noise and sparse acquisition are considered for the 2D examples. The depth threshold to avoid the noise mask is set to 1.2 km. In the 3D examples, the depth threshold is set to 1.4 km. All times mentioned in the following examples are relative to the actual source time, which is 0 s (zero time). The negative numbers indicate the time prior to the real excitation time, and the positive numbers indicate the time after the real excitation time.

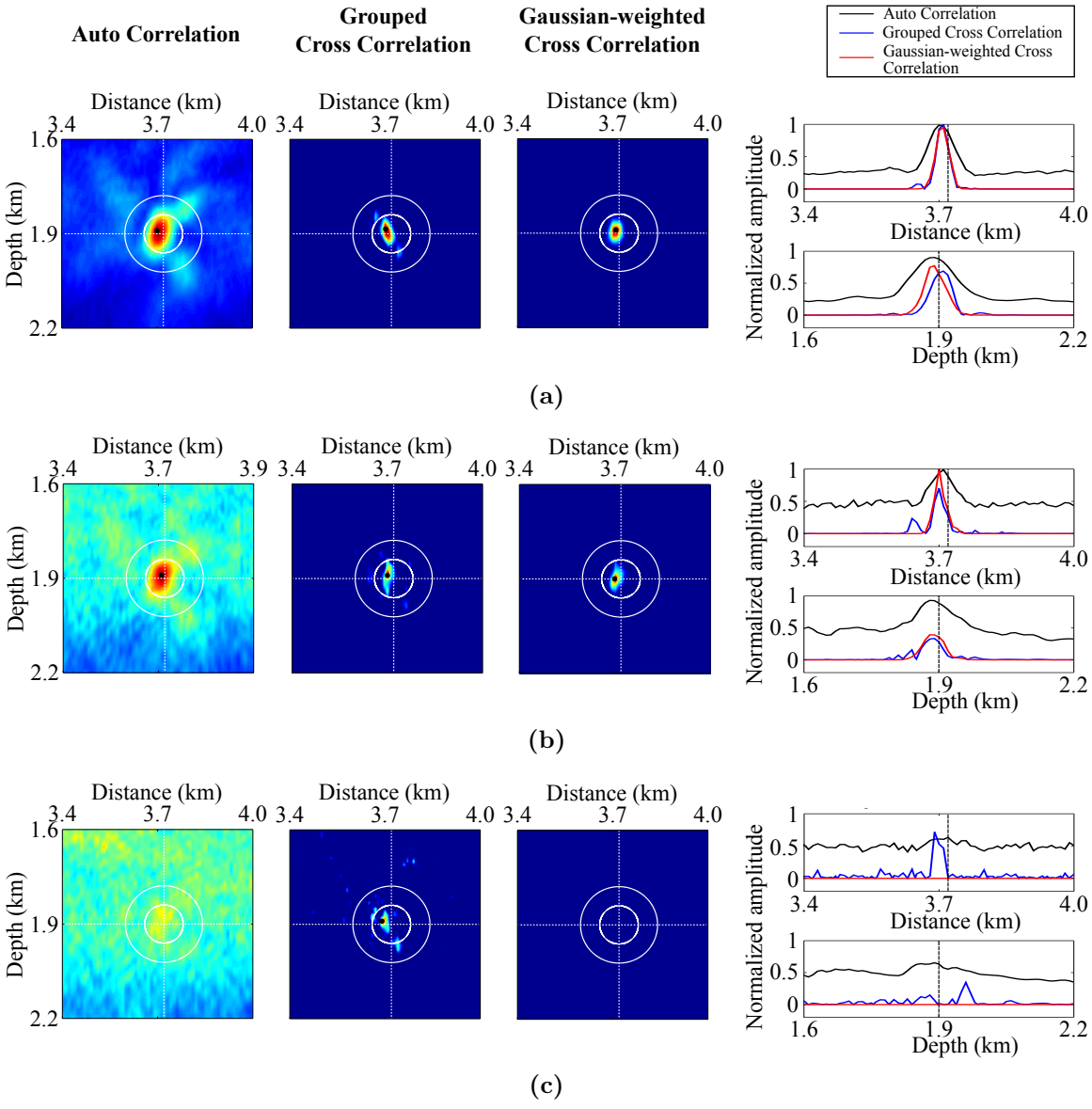


Figure 2.10: Comparison of three time-reversal imaging methods for noisy data with S/N of 0.4 (a), 0.2 (b), and 0.1 (c). Black dots indicate the source locations by the time-reversal imaging methods. White circles, centered at the correct source location, have radii of a quarter and a half of the dominant wavelength. The right column shows horizontal (top) and vertical (bottom) amplitude sections through the source location.

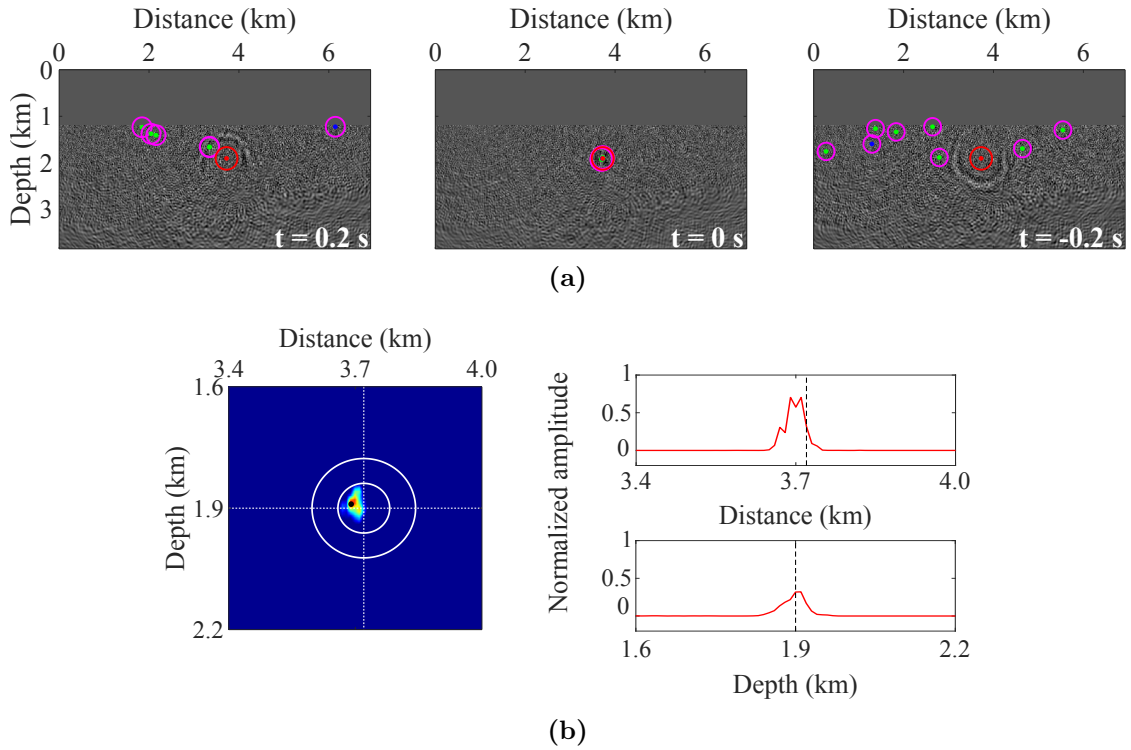


Figure 2.11: (a) Images of back-projected wavefields at the time of 0.2 s, 0 s, and -0.2 s for noisy data with S/N of 0.1. Gaussian-weighted wavefields are obtained for R of 0.5. Blue and green dots denote the picks, corresponding to the maximum absolute amplitude of the whole back-propagated wavefield and the selected back-propagated wavefields at the considered time step, respectively. Red dots represent the exact source location. Red and purple circles, with radii of the dominant signal wavelength, are centered at the picks and the exact source location, respectively. (b) Left: the result of the Gaussian-weighted cross-correlation imaging method; Right: horizontal (top) and vertical (bottom) amplitude sections through the source location.

2.3.1 Marmousi-II model

We consider a subset of the Marmousi-II model, defined by a 693×385 grid with a spacing of 10 m in the x- and z- directions. Figure 2.2 a shows the actual model, and Figures 2.2 b– 2.2 d denote the model smoothed by 300 m $\times$ 300 m, 600 m $\times$ 600 m, and 900 m $\times$ 900 m 2D Gaussian windows, respectively. The true model is used to obtain the synthetic data, and the smoothed models are used to simulate the back-propagated wavefield. The time sampling for all synthetic data is 1 ms. Sources denoted by the black and red dots are explosive sources placed near a fault in the presence of strong velocity variations.

A single source located at $x = 3.72$ km and $z = 1.9$ km, denoted by the red dot in Figure 2.2, is used to evaluate the influence in terms of velocity errors, random noise, and sparse receiver distribution for the time-reversal imaging. All of these factors are crucial for field data applications.

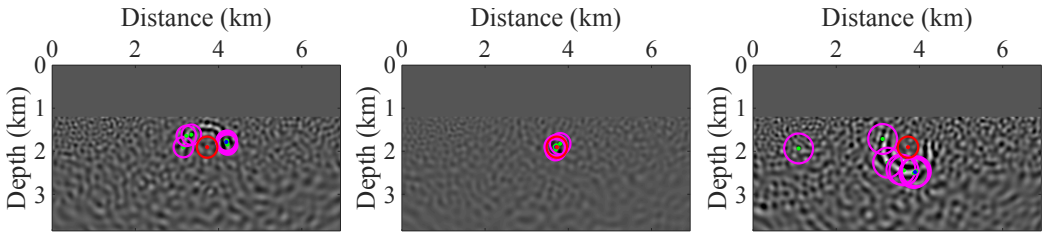


Figure 2.12: Images of the back-projected wavefields at the time of 0.2 s, 0 s, and -0.2 s for the band-pass filtered data (6 - 17 Hz) with S/N of 0.1. Blue and green dots denote the picks, corresponding to the maximum absolute amplitude of the whole back-propagated wavefield and the selected back-propagated wavefields at the considered time step, respectively. Red dots represent the exact source location. Red and purple circles, with radii of the dominant signal wavelength, are centered at the picks and the exact source location, respectively.

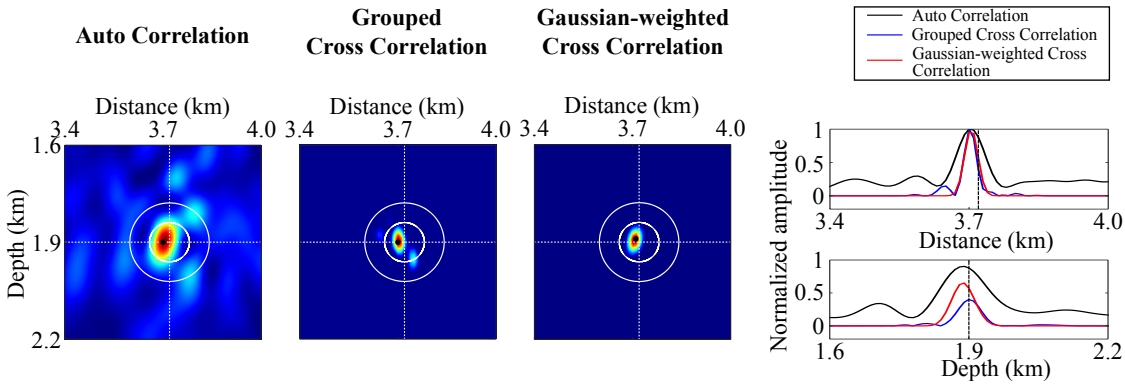


Figure 2.13: Comparison of three time-reversal imaging methods for the band-pass filtered data (6 - 17 Hz) with S/N of 0.1. The black dots indicate the estimated source location by the time-reversal imaging methods. White circles, centered at the correct source location, have radii of a quarter and a half of the dominant wavelength. The right column shows horizontal (top) and vertical (bottom) amplitude sections through the source location.

Influence of velocity models

In the first example, forward modeling and time-reversal imaging are performed in the true model as shown in Figure 2.2 a. Receivers are evenly distributed with a distance of 10 m at the surface. Eight receiver groups are chosen for the grouped crosscorrelation imaging. The first group comprises the receivers between 0 and 0.5 km and the second receiver group contains the receivers from 0.5 to 1.0 km. The receivers between 1.0 to 6.0 km are evenly divided into five receiver groups and the last group comprises the remaining receivers. This uniform, dense, and larger aperture acquisition provides good illumination for time-reversal imaging.

Figure 2.3 a and 2.3 b shows the full back-propagated wavefield at the time of 0.6 s and its corresponding normalized absolute value, respectively. Figure 2.3 c and 2.3 d shows the first selected wavefield and Gaussian-weighted wavefield for this time step. By setting

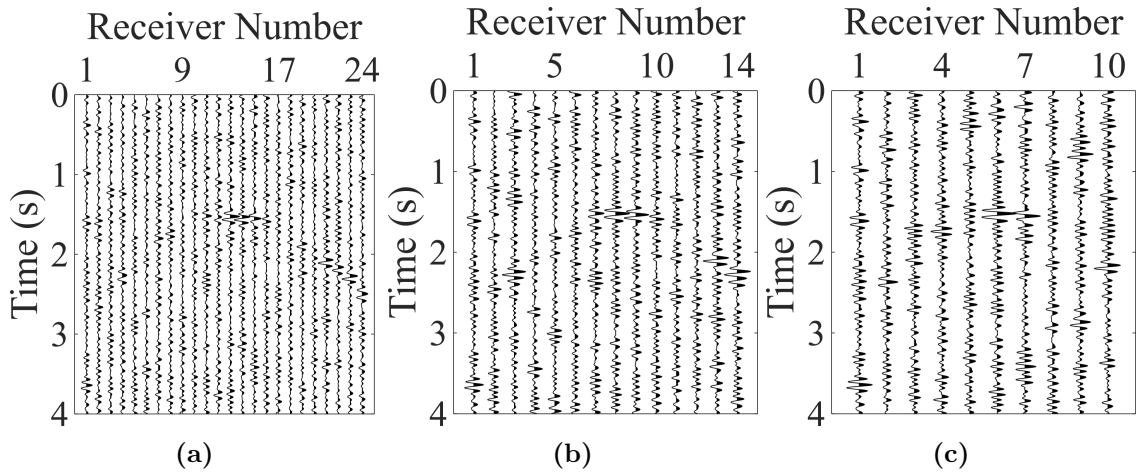


Figure 2.14: Band-pass filtered sparse data (6 - 17 Hz) for a receiver spacing of 300 m (a), 500 m (b), and 700 m (c), respectively, for the source located at $x = 3.72$ km, $z = 1.9$ km in the Marmousi-II model.

k to one in equation 2.3, the first wavefield is selected from the full wavefield according to the lower and upper bounds as described in this equation. The first Gaussian-weighted wavefield centered at the green dot, i.e., the position of the maximum amplitude of the first selected wavefield, is obtained by applying a Gaussian window to the full wavefield using equation 2.6. Other selected wavefields and Gaussian-weighted wavefields are acquired in the same fashion. In the following, only the picked position of the maximum absolute amplitude of each selected wavefield and the spatial extension of the window of each picked position are displayed in the images of the full back-propagated wavefield at a few time steps.

Figure 2.4 shows the full back-propagated wavefields at the time of 0.6, 0.4, 0.2, 0, -0.2, and -0.4 s. In Figure 2.4, at the actual excitation time, the Gaussian-weighted wavefields focus at the source location, whereas at other times, they scatter on the back-projected wavefronts. Figure 2.5 shows the time-reversal imaging results. The grouped and Gaussian-weighted crosscorrelation imaging methods deliver images of high quality, i.e., good resolution, whereas a few weak artifacts are observed in the autocorrelation image, which is less focused. The resolution of the Gaussian-weighted crosscorrelation imaging is similar to that of the grouped crosscorrelation imaging. All estimated locations (the black dots) coincide with the correct source location. This example confirms the feasibility of the Gaussian-weighted crosscorrelation imaging method, which requires only a single run of the back-propagated wavefield to provide a clear and high-resolution image for accurate source localization, whereas, for the grouped crosscorrelation method, the number of runs to back propagate the wavefield is equivalent to the number of groups.

Velocity uncertainty is a fact in field data applications. We explore time-reversal imaging for several smoothed versions of the true models. These smoothed models are shown in Figures 2.2 b– 2.2 d, which have a mean deviation of 5 %, 7 %, and 8 %, respectively. Data acquisition and receiver groups used here are the same as for the first example. Back-propagated wavefields are computed for the smoothed versions of the true model. Figure 2.6

Receiver Group Number	Receiver Number
1	1 - 3
2	3 - 5
3	5 - 8
4	8 - 11
5	11 - 14
6	14 - 17
7	17 - 21
8	21 - 24

(a)

Receiver Group Number	Receiver Number
1	1 - 2
2	2 - 3
3	4 - 5
4	6 - 7
5	7 - 8
6	9 - 10
7	11 - 13
8	13 - 14

(b)

Receiver Group Number	Receiver Number
1	1 - 2
2	2 - 3
3	3 - 4
4	4 - 5
5	5 - 6
6	6 - 7
7	7 - 8
8	8 - 10

(c)

Figure 2.15: Receiver groups for the band-pass filtered (6 - 17 Hz) data with a receiver spacing of 300 m (a), 500 m (b), and 700 m (c) for the source at $x = 3.72\,\text{km}$, $z = 1.9\,\text{km}$ in the Marmousi-II model. The receiver groups are used for the grouped cross-correlation imaging.

shows the wavefields at the time of 0.2, 0, and -0.2 s. The time-reversal imaging results using the respective smoothed models (Figures 2.2 b – 2.2 d) are shown in Figures 2.7 a – 2.7 c. In comparison to Figure 2.4, the back-propagated wavefields display artifacts and do not focus as well as for the true model. The results obtained by the autocorrelation imaging method are of poor quality. These imaging artifacts caused by velocity errors display strong amplitudes and may lead to false positives. The Gaussian-weighted crosscorrelation imaging provides the best image quality with high resolution and almost no artifacts. Because of velocity errors, the source locations show deviations from the true location. This applies to all methods, however, the Gaussian-weighted crosscorrelation is the most robust approach with respect to imaging artifacts, which in turn decrease the risk of false positives.

Poor data quality

To examine the effect of uncorrelated noise on the Gaussian-weighted crosscorrelation imaging technique, random noise with the S/N of 0.4, 0.2, and 0.1, was added to the data. The respective records are shown in Figures 2.8 a – 2.8 c. In Figure 2.8 c, the data are poor, and no events are observed. In the following 2D examples, the back-propagated wavefields are computed for the smoothed model as shown in Figure 2.2 b, i.e., 300 m $\times$ 300 m Gaussian window. The seismic acquisition and receiver groups used are the same as in the previous example. Time-reversal imaging is applied using the three noisy data sets. Wavefields at time of 0.2, 0, and -0.2 s are shown in Figure 2.9 a – 2.9 c.

The results show that, with increasing noise levels, the wavefields selected are not located on the actual wavefront. For the S/N of 0.4 and 0.2, very good foci at the source time are observed. For the S/N of 0.1, most selected wavefields still focus near the source but no distinct maximum is observed (Figure 2.10). This issue can be addressed by either decreasing the R value to 0.5 or band-pass filtering of the data. Figure 2.11 shows the results when R is set to 0.5. At the exact source time, the Gaussian-weighted crosscorrelation image provides good focusing with high resolution and negligible artifacts. Figure 2.12 and Figure 2.13 show the results using band-pass filtered data (6 - 17 Hz). Band-pass filtering

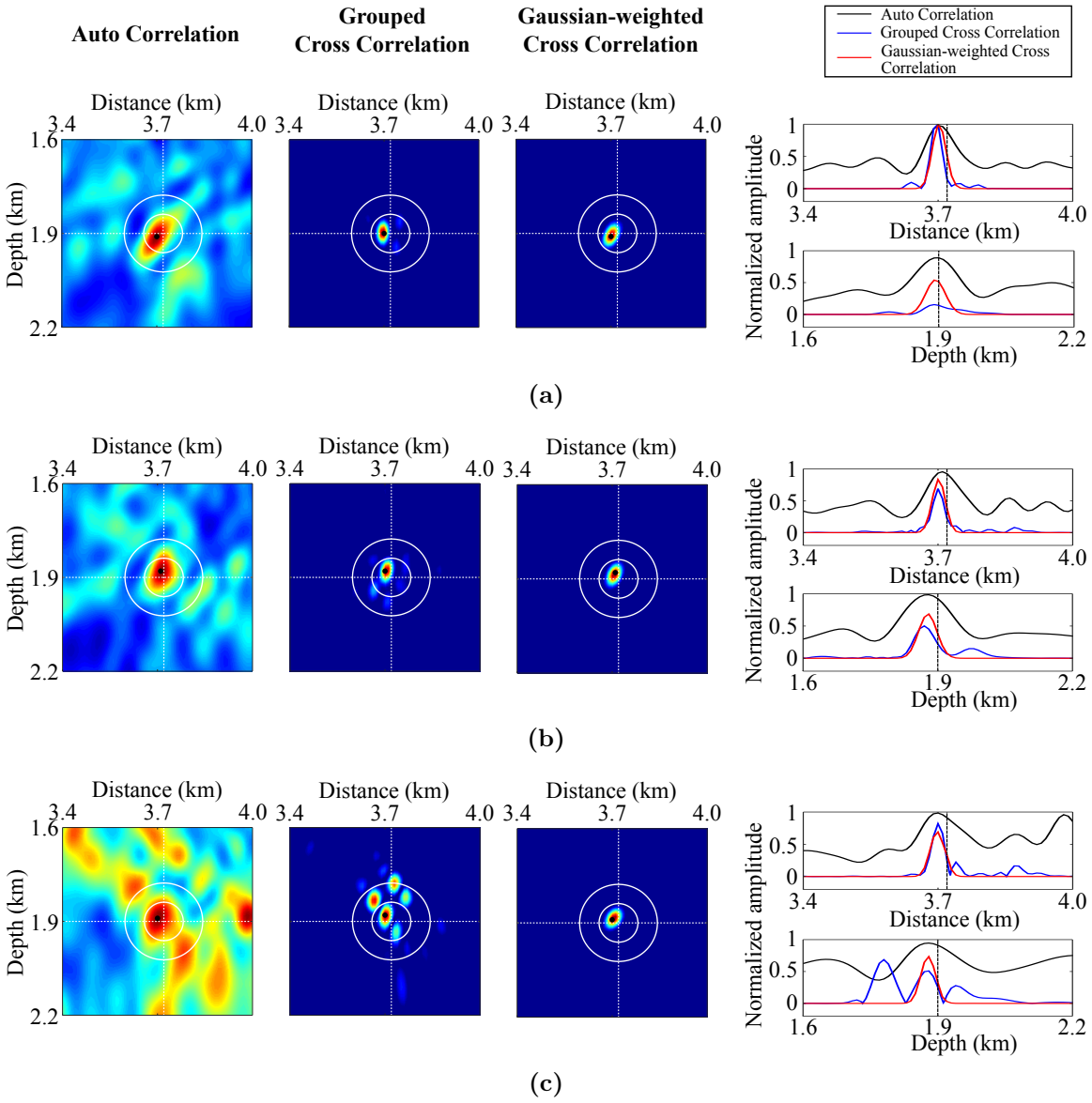


Figure 2.16: Comparison of three time-reversal imaging methods for the band-pass filtered data (6 - 17 Hz) with a receiver spacing of 300 m (a), 500 m (b), and 700 m (c), respectively. Black dots indicate the estimated source location by the time-reversal imaging methods. White circles, centered at the correct source location, have radii of a quarter and a half of the dominant wavelength. The right column shows horizontal (top) and vertical (bottom) amplitude sections through the source location.

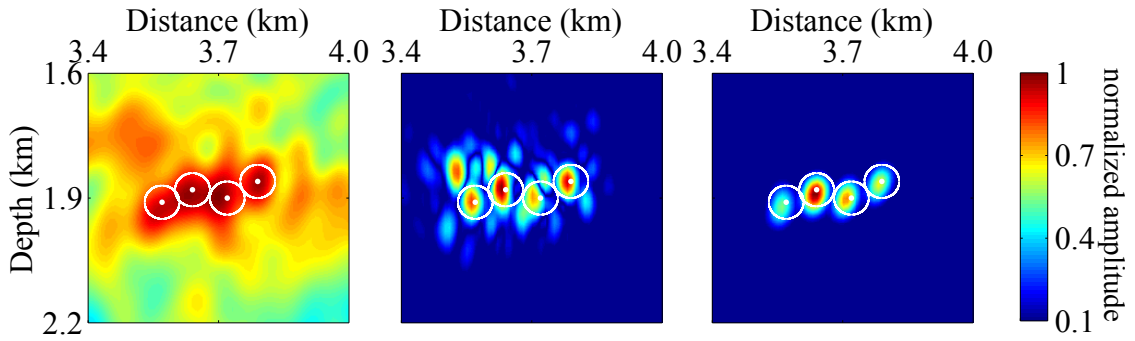


Figure 2.17: Comparison of three time-reversal imaging methods for a source cluster in the Marmousi-II model. Left: auto-correlation imaging; Middle: grouped cross-correlation imaging; Right: Gaussian-weighted cross-correlation imaging. White circles, centered at the true source location, have radii of 40 m.

improves the S/N to approximately 0.77. The filtered data are shown in Figure 2.8 d. The event is now clearly visible. Using the filtered data, all three imaging methods provide a reasonable source location, but the Gaussian-weighted cross-correlation imaging provides the best focusing with weak to no artifacts.

Sparse receiver networks

For passive seismic acquisitions, the number of receivers is considerably smaller than for the previous examples. In the following, we discuss the influence of sparse receiver networks on imaging methods. For this example, receivers are distributed at the surface with a spacing of 300, 500, and 700 m. Random noise with an S/N of 0.4 was added to the sparse data and the band-pass filter (6 - 17 Hz) was applied to increase the S/N to approximately 3. In Figure 2.14, the resulting seismograms with the receiver number are shown. For the grouped crosscorrelation imaging with sparse data, the choice of the receiver groups is shown in Figure 2.15. Figure 2.16 shows the results obtained for the three filtered sparse data sets. For the autocorrelation and grouped crosscorrelation imaging methods, imaging artifacts increase substantially with increasing receiver spacing. Although the autocorrelation and grouped crosscorrelation imaging methods deliver reasonably accurate source locations, strong artifacts still exist, which may cause ambiguities for event identification leading to false positives. The Gaussian-weighted crosscorrelation imaging method yields clear images when compared to the two conventional imaging methods. Moreover, almost no artifacts are observed in the images. The Gaussian-weighted crosscorrelation imaging method is thereby effective for data with noise from a sparse receiver network.

Multiple sources

As shown in previous examples, the Gaussian-weighted crosscorrelation imaging method can produce clear and high-resolution source images. In the following, we will investigate its performance for multiple sources which are spatially clustered. As shown in Figure 2.2, four explosive sources with the same source characteristics are located around the fault. They are excited at the time of 0.4, 1.25, 2.3, and 3.3 s. The velocity at the source cluster is approximately 2.6 km/s. The minimum distance between the adjacent sources is

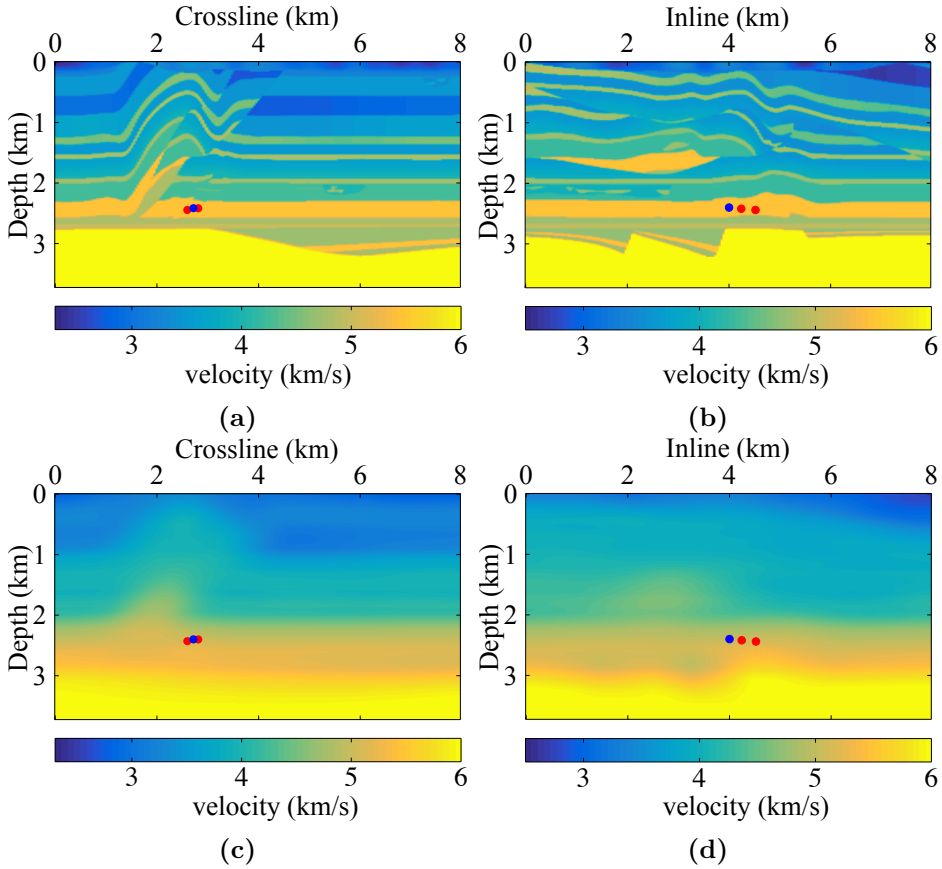
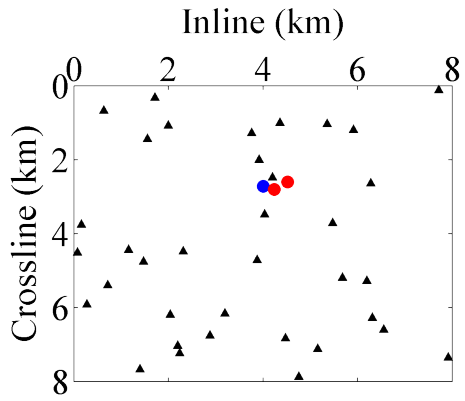


Figure 2.18: Velocity sections of the 3D SEG overthrust model for the source located at $x = 2.72$ km, $y = 4.0$ km and $z = 2.4$ km. (a) and (b) show the crossline and inline sections of the true model, respectively. (c) and (d) show the crossline and inline sections of the smoothed model. Blue dot denotes the source located at $x = 2.72$ km, $y = 4.0$ km and $z = 2.4$ km. Red dots denote the projected source located at $x = 2.8$ km, y= 4.24 km and $z = 2.42$ km, and $x = 2.6$ km, $y = 4.52$ km and $z = 2.44$ km, respectively.



(a)

Receiver Group Number	Receiver Number
1	1 - 5
2	5 - 9
3	9 - 13
4	13 - 17
5	17 - 21
6	21 - 25
7	25 - 29
8	29 - 36

(b)

Figure 2.19: (a) Seismic network for the 3D SEG overthrust model. Dots show the projection of the sources to the surface. Black triangles denote receivers randomly placed at the surface. (b) Receiver groups for the grouped cross-correlation imaging.

approximately 70 m, i.e., less than half of the dominant wavelength. Receivers are evenly distributed at the surface with a spacing of 300 m. Random noise with the S/N of 0.4 was added to the sparse data. A band-pass filter (6 - 17 Hz) was applied. The S/N after filtering is approximately 3. The targeted region of the time-reversal imaging is beneath 1.2 km. The square root of image amplitudes is considered in this example to enhance weaker events. Figure 2.17 shows the time-reversed images. The results of the autocorrelation imaging method suffer from strong artifacts with a low resolution of the image, where no real foci are observed, except for a cluster of high energy.

The same set of receiver groups is used for the first sparse data example of the grouped crosscorrelation imaging (Figure 2.15 a). Although the resolution is improved, the grouped crosscorrelation image is ambiguous due to strong artifacts hampering interpretation. As shown in Figure 2.17, the Gaussian-weighted crosscorrelation image shows a high-resolution image and displays four distinct foci close to the correct source locations. The maximum errors for the location are smaller than 40 m, which is far below the dominant wavelength. The proposed Gaussian-weighted crosscorrelation imaging method demonstrates clearly discerning sources with concentrating distribution in comparison with the two conventional imaging methods. In the next section, we consider a 3D example.

2.3.2 3D SEG overthrust model

The 2D examples shown in the previous section are suitable to investigate the general performance of the Gaussian-weighted crosscorrelation method under the condition of noisy data and sparse acquisitions. Because the localization of passive seismic events is an inherently 3D problem, we consider this situation. We use the 3D SEG overthrust model, where velocities are defined on a $400 \times 400 \times 187$ grid. The grid spacing is 40 m in the x- (crossline-) and y- (inline-) directions and 20 m in z- (depth-) direction. In Figure 2.18 a, three sources cluster in this model. Figure 2.18 a and 2.18 b shows the crossline and inline sections of the model at the position $x = 2.72$ km, $y = 4.0$ km, and $z = 2.4$ km, i.e., the location of the source denoted as a blue dot. The red dots denote the projection of the other

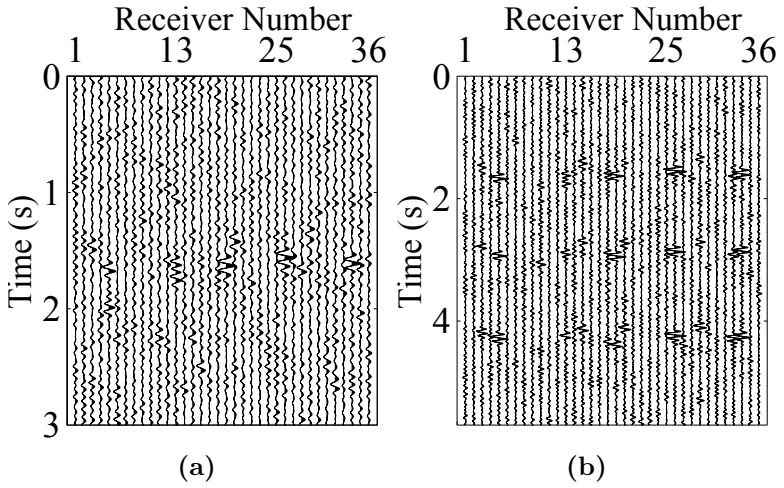


Figure 2.20: Band-pass filtered sparse data (6 - 17 Hz) for the single source at x = 2.72 km, y = 4.0 km and z = 2.4 km (a), and for the source cluster in the 3D SEG overthrust model (b).

two sources on the sections. We perform the smoothing by a 1200 × 1200 × 600 m^3 3D Gaussian window, and the smoothed version of the sections is shown in Figure 2.18 c and 2.18 d. Data are computed again for the true model and the back-propagated wavefields are computed in the smoothed model. The time sampling is 1 ms. In Figure 2.19 a, 36 receivers are randomly distributed at the surface, which reflects a sparse acquisition. The surface projection of the sources are denoted by the dots. Receivers are sorted into eight groups (Figure 2.19 b) for the grouped crosscorrelation imaging.

We first consider a single source located at $x = 2.72$ km, $y = 4.0$ km, and $z = 2.4$ km, which is denoted by the blue dot. Random noise with an S/N of 0.4 was again added to the data and the band-pass filter (6 - 17 Hz) was applied, leading to an S/N of approximately 3. The resulting waveforms sorted by receiver number are shown in Figure 2.20 a. The locations retrieved from the autocorrelation and grouped crosscorrelation images (Figure 2.21 a and 2.21 b) display limitations for these time-reversal imaging methods. Figure 2.21 c shows the results obtained by the Gaussian-weighted crosscorrelation imaging method. The source location is obtained with good accuracy. The location error is smaller than one quarter of the dominant wavelength.

To investigate the quality of the time-reversal images, we sum the time-reversal images over the crossline and depth directions and use the inline, crossline, and depth slices of the sum at the actual source location (see Figure 2.22 and Figure 2.23). Compared with the conventional time-reversal imaging methods, the Gaussian-weighted crosscorrelation imaging method yields the best image with almost no artifacts. It provides good source locations by decreasing the risk of false positives.

In the following example, we consider multiple sources. In Figure 2.18 a and 2.18 b, three explosive sources are excited at 0.6, 1.9, and 3.3 s with the same amplitude and frequency range. The velocity at the locations is approximately 5.1 km/s. The minimum distance between the adjacent sources is approximately 250 m, i.e., smaller than half of the prevailing wavelength, reflecting the clustered sources. Random noise with the S/N of

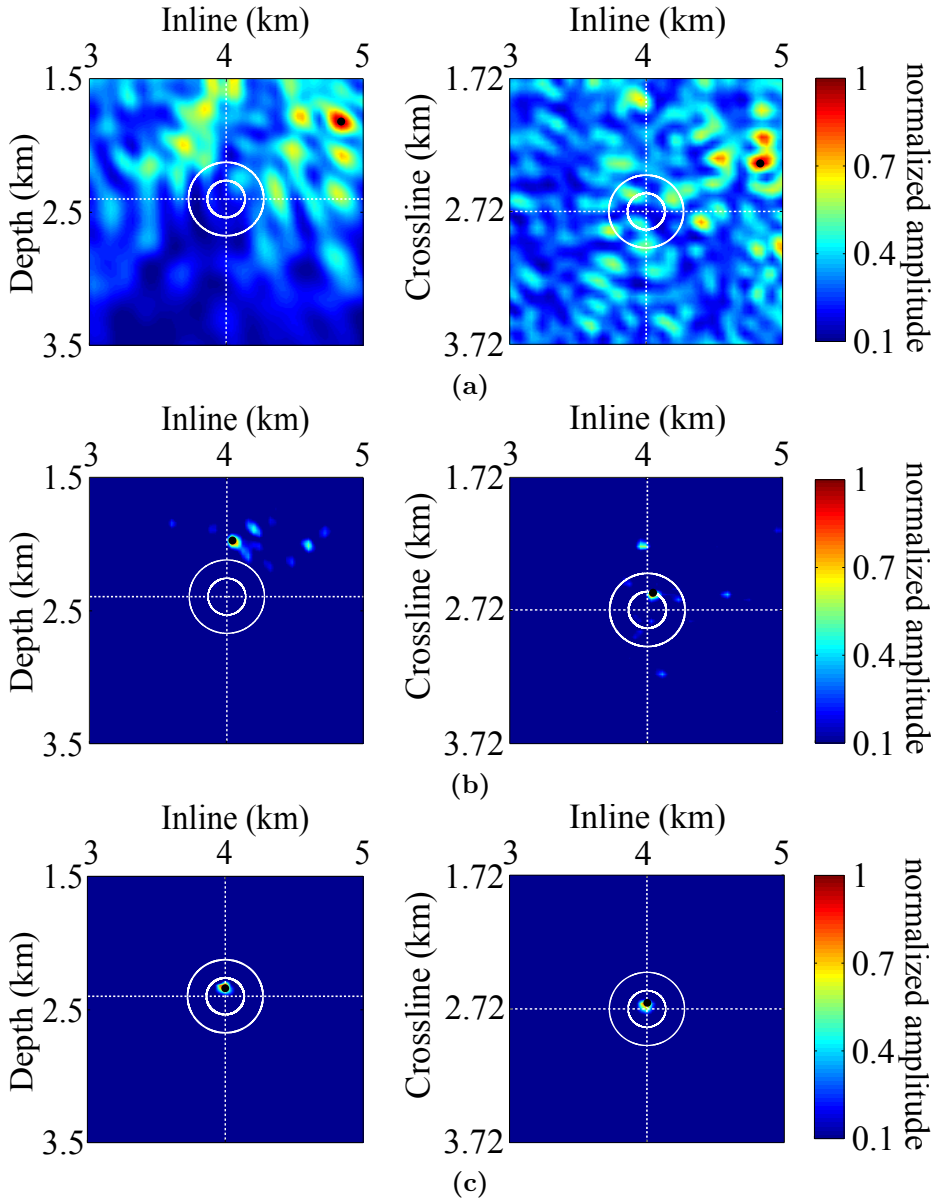


Figure 2.21: Comparison of three time-reversal imaging methods for the source at $x = 2.72$ km, $y = 4.0$ km and $z = 2.4$ km in the 3D SEG overthrust model. The images at the true source location are obtained by auto-correlation imaging (a), grouped cross-correlation imaging (b), and Gaussian-weighted cross-correlation imaging (c). Black dots indicate the estimated source locations. White circles, centered at the true source location, have radii of a quarter and a half of the dominant wavelength.

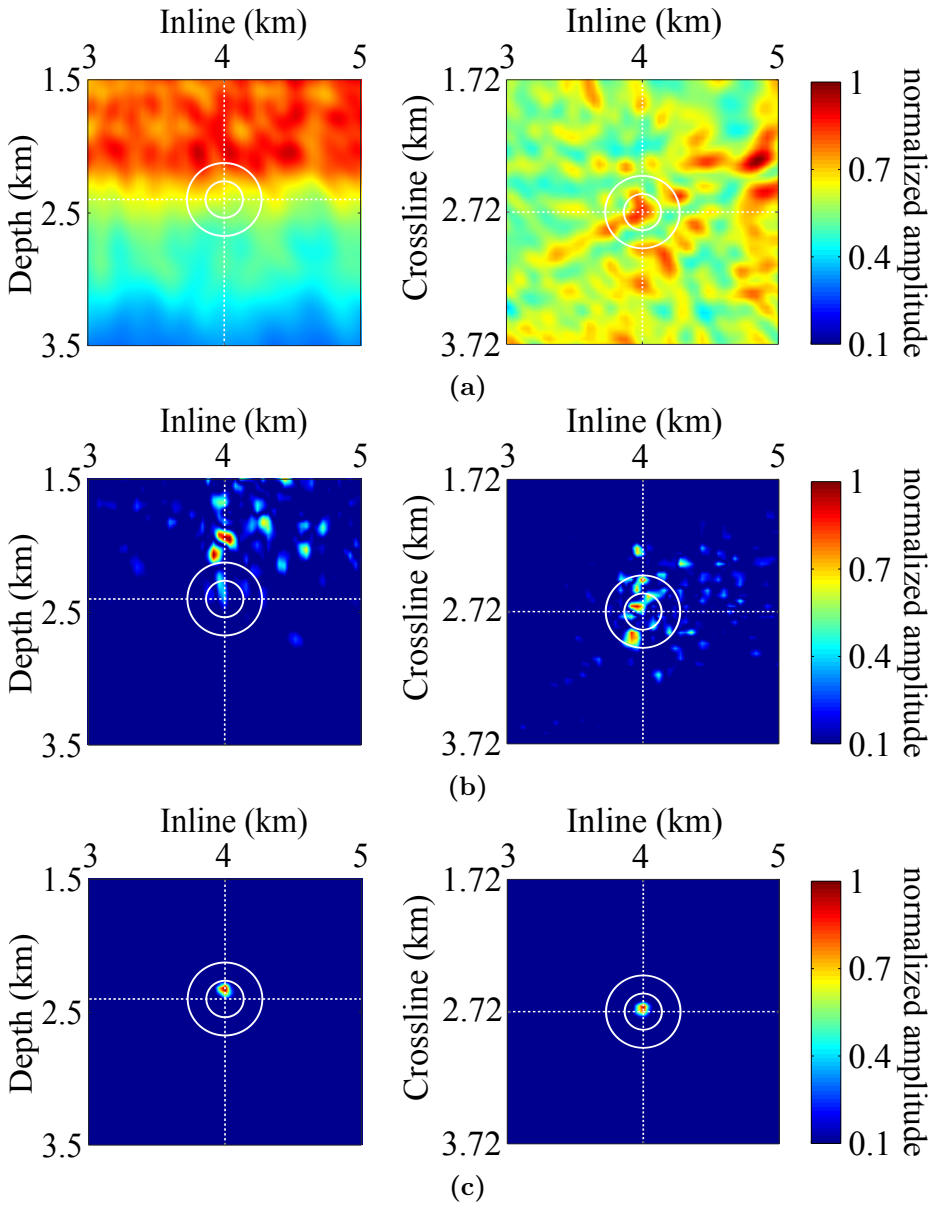


Figure 2.22: Results of auto-correlation imaging (a), grouped cross-correlation imaging (b), and Gaussian-weighted cross-correlation imaging (c). Left: the sum of the images over the cross-line direction. Right: the sum of the images over the depth direction. White circles are centered at the true source location with radii of a quarter and a half of the dominant wavelength.

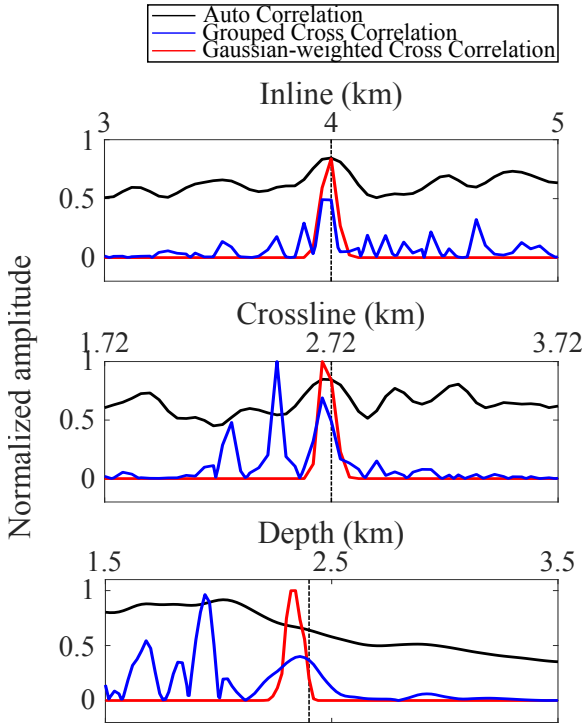


Figure 2.23: Horizontal and vertical amplitude sections of the images shown in Figure 2.22 at fixed depth (top) and fixed distance (bottom) of the source at $x = 2.72$ km, $y = 4.0$ km and $z = 2.4$ km in the 3D SEG overthrust model.

0.5 was added to the data and then a band-pass filter (6 - 17 Hz) was applied. The S/N after the filtering was approximately 3.8. The filtered data sorted by receiver number are shown in Figure 2.20 b. Figure 2.24 shows the inline slices of the images at the crossline distance of 3.76, 4.0, 4.24, 4.52, and 4.76 km, and the slices of the images at the depth of 2.28, 2.4, 2.42, 2.44, and 2.56 km are shown in Figure 2.25. Figures 2.24 a, 2.24 b, 2.25 a, 2.25 b, 2.26 a, and 2.26 b show the results of the autocorrelation imaging and the grouped crosscorrelation imaging methods. Many false positives are observed, which may lead to image misinterpretation. The grouped crosscorrelation imaging generates many false foci near the exact source locations because of noise and imaging artifacts. The Gaussian-weighted crosscorrelation imaging method provides the image with almost no artifacts and reduces the risk of false positives considerably (Figures 2.24 and 2.26 c). Three distinct foci are imaged located close to the exact source locations. The maximum location errors are about 90 m, which is smaller than half the dominant wavelength. In addition, for the 3D complex examples, the Gaussian-weighted crosscorrelation imaging method provides the best performance which is preserved even for noisy data and sparse acquisitions.

2.4 Field data example

We use these time-reversal imaging methods to locate a weak event of mining-induced seismicity monitored by the HAMNET network (Wehling-Benatelli et al., 2013). To estimate

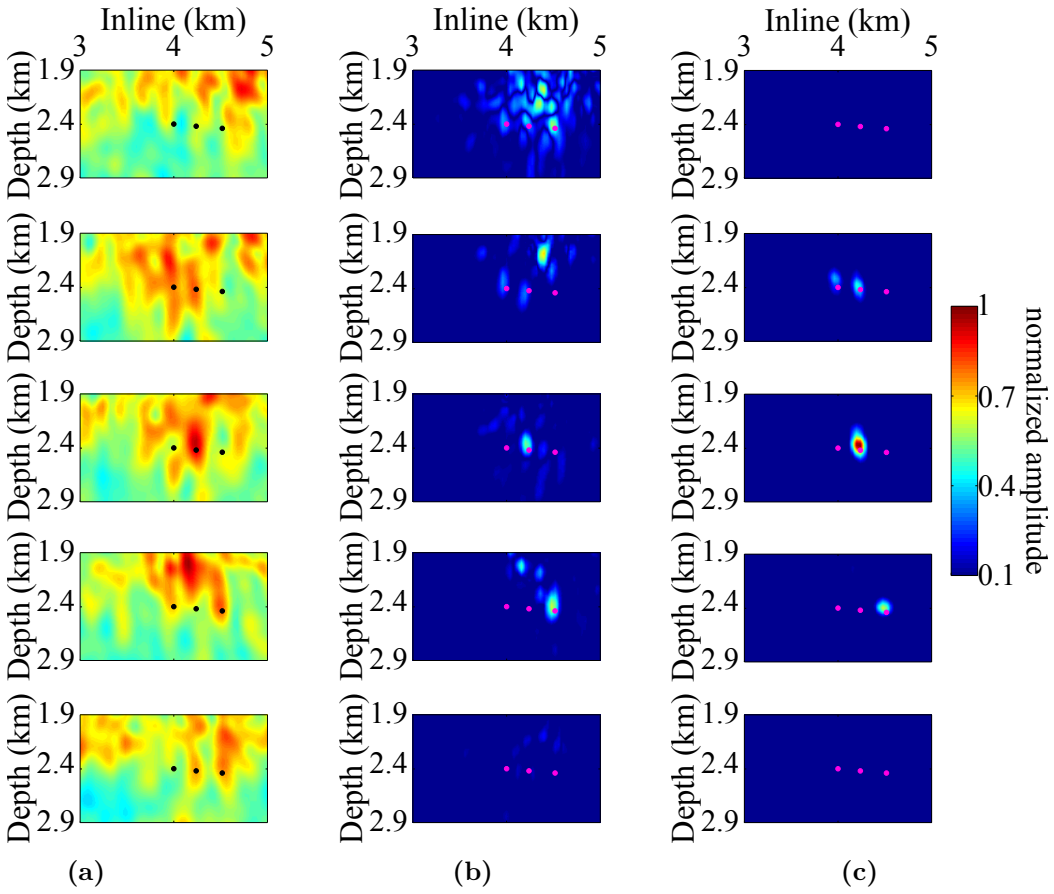


Figure 2.24: Inline slice comparison of three time-reversal imaging methods for a source cluster in the 3D SEG overthrust model. (a) Auto-correlation imaging. (b) Grouped cross-correlation imaging. (c) Gaussian-weighted cross-correlation imaging. The five inline slices of each method are shown at 3.76 km, 4.0 km, 4.24 km, 4.52 km, and 4.76 km of the crossline direction, respectively. The black and purple dots indicate the exact source locations.

the location error, we take the location inverted from the picking-based traveltime inversion (Grigoli et al., 2013; Li et al., 2018) as a reference. Figure 2.27 a shows 13 receivers selected from the HAMNET network, which are randomly distributed at the surface. The surface projection of the event is denoted by the red dot. The weak event here has a magnitude of -0.8. Figure 2.28 a shows the vertical component of the raw data sorted by the receiver number recorded for the weak event, which contains strong P-wave energy and noise. A frequency analysis (e.g., Priestley, 1996) was performed to estimate the dominant frequency, and a band-pass filter (6 - 17 Hz) was then applied to the data. The resulting seismograms are shown in Figure 2.28 b, where the S/N was significantly increased and the P phase information is kept for the frequency band of 6 - 17 Hz, which is sufficient for the localization. Receivers are sorted into five groups (Figure 2.27 b) for the grouped crosscorrelation imaging. The same α and R values as for the previous 3D examples are used here, whereas the N value is set to five.

We use the same homogeneous P-wave velocity model, 3880 m/s, as in the traveltime

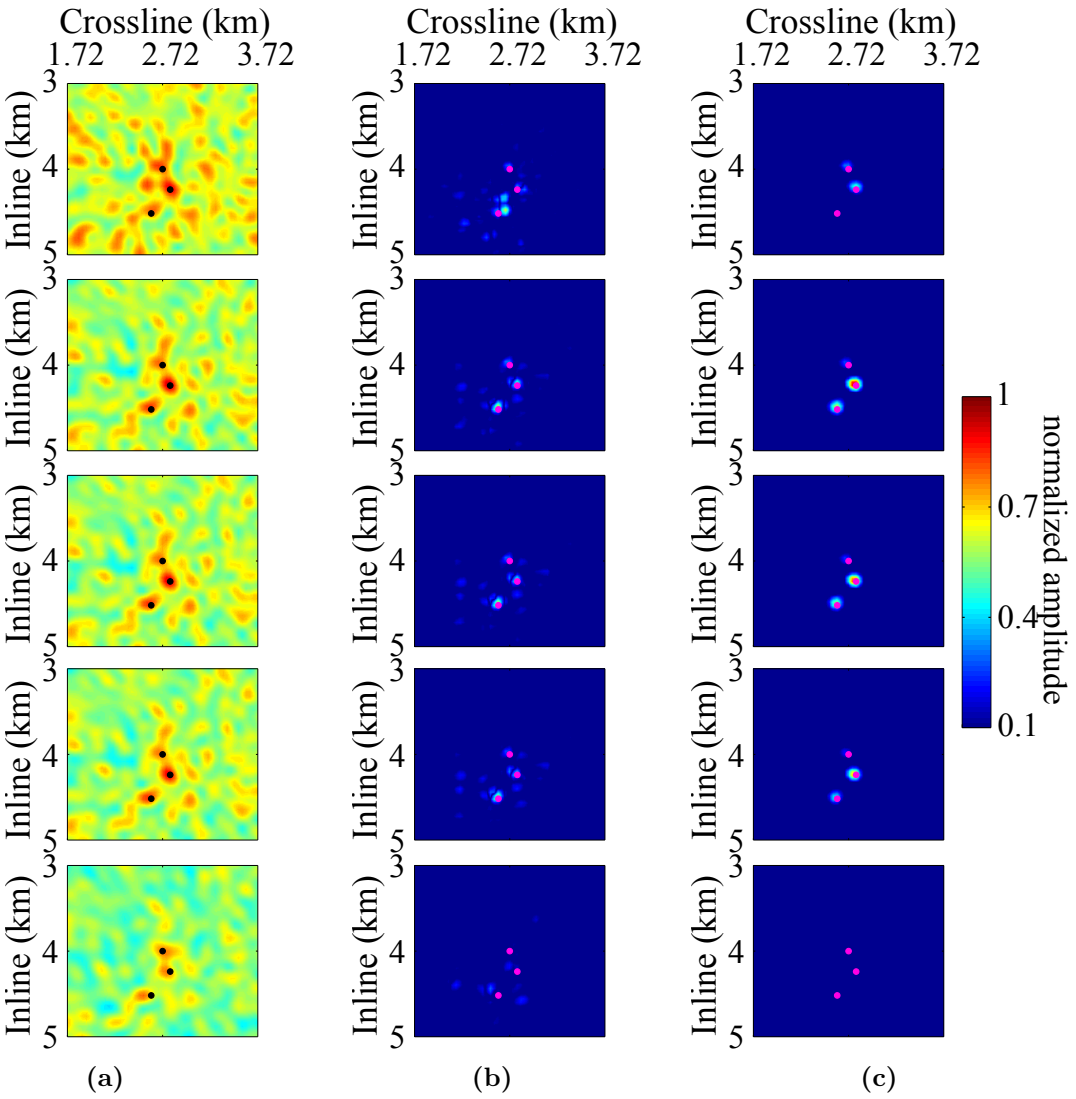


Figure 2.25: Depth slice comparison of three time-reversal imaging methods for a source cluster in the 3D SEG overthrust model. (a) Auto-correlation imaging. (b) Grouped cross-correlation imaging. (c) Gaussian-weighted cross-correlation imaging. The five horizontal slices of each method are shown at the depth of 2.28 km, 2.4 km, 2.42 km, 2.44 km, and 2.56 km, respectively. The black and purple dots indicate the exact source locations.

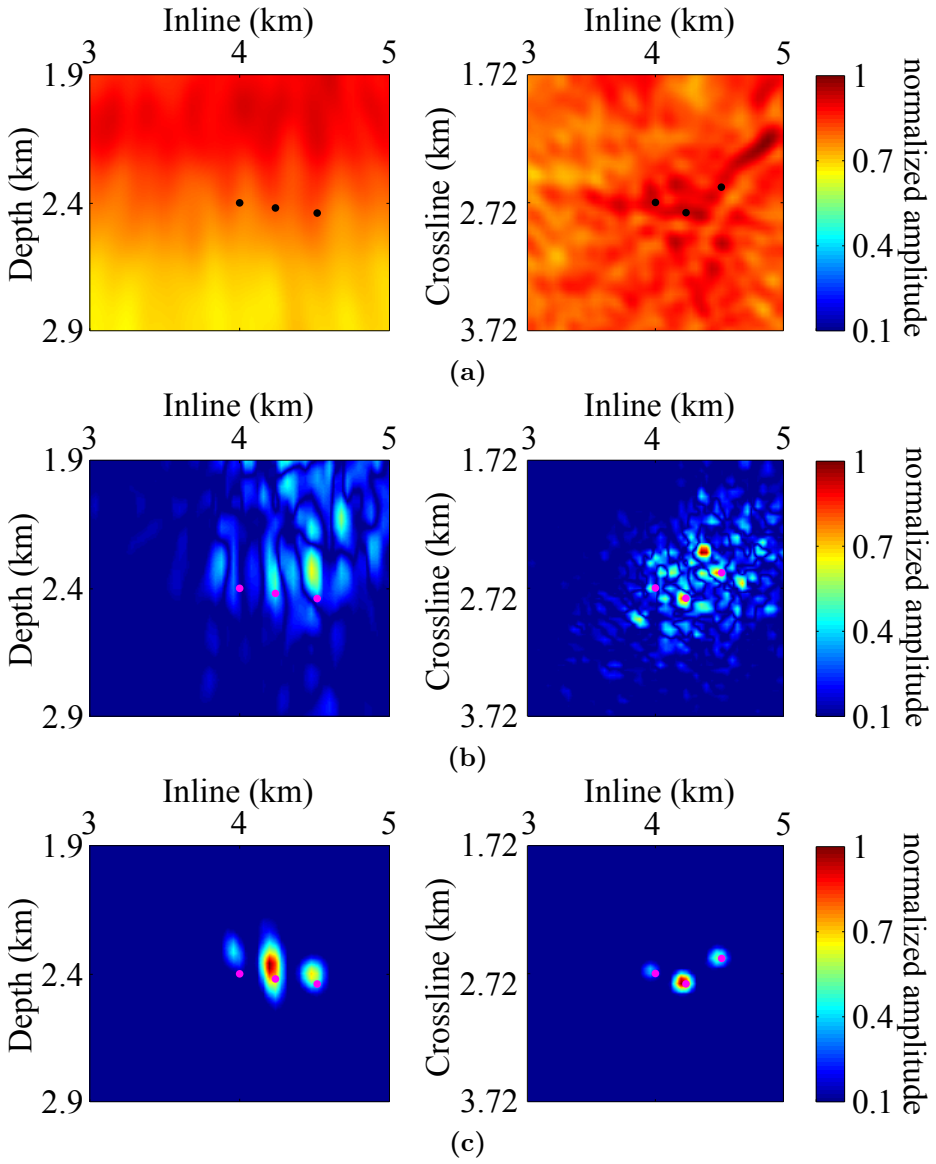
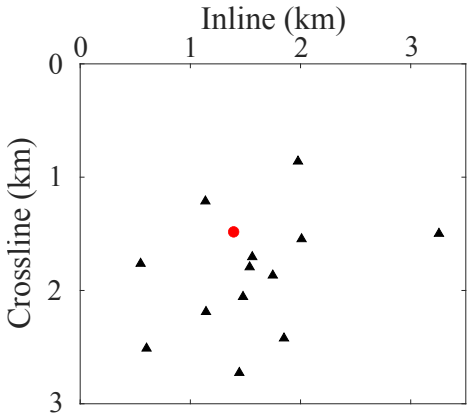


Figure 2.26: Comparison of three time-reversal imaging methods for a source cluster in the 3D SEG overthrust model. Images at the true source location are obtained by auto-correlation imaging (a), grouped cross-correlation imaging (b), and Gaussian-weighted cross-correlation imaging (c). Left: the square root of the sum of the images over the cross-line direction. Right: the square root of the sum of the images over the depth direction. The black and purple dots indicate the exact source locations.



(a)

Receiver Group Number	Receiver Number
1	1 - 2
2	3 - 5
3	6 - 8
4	9 - 11
5	12 - 13

(b)

Figure 2.27: (a) Seismic network for mining-induced seismicity monitoring. The red dot shows a projection of the reference source location to the surface which is provided by the travel-time inversion. Black triangles denote the receiver located at the surface. (b) A set of receiver groups for the grouped cross-correlation imaging.

inversion and consider only P waves to perform the time-reversal source imaging. The same acoustic wave equation as used for the synthetic examples is adopted for the back-propagation simulation. A reasonable depth threshold of 0.8 km is chosen to avoid the noise mask. Figure 2.29 shows the results of the time-reversal imaging methods using the band-pass filtered data. We can find that the Gaussian-weighted crosscorrelation imaging method can still yield the best location result that has the clearest focus. A few artifacts are observed in the image (Figure 2.29 c) that might be due to the less accurate velocity model. The location errors of the autocorrelation imaging and the grouped crosscorrelation imaging methods (Figure 2.29 a and 2.29 b) are approximately 1912 and 558 m, respectively, beyond the dominant wavelength. The error of the proposed imaging method is approximately 240 m, i.e., approximately 60% of the dominant wavelength. The results are consistent with the previous numerical examples. In this example, the computational time of the autocorrelation imaging and the Gaussian-weighted crosscorrelation imaging is comparable, which is about one-third of the cost used by the grouped crosscorrelation imaging. The results of this field example illustrate the Gaussian-weighted crosscorrelation imaging method can serve as an advantageous alternative to the autocorrelation imaging and grouped crosscorrelation imaging methods.

2.5 Discussion

In complex heterogeneous media, amplitudes on a propagating wavefront change smoothly with time but may strongly vary along the wavefront, i.e., for a constant time. Due to scattering in the back projection process, random focusing may be observed combined with imaging artifacts because of insufficient receiver sampling and/or noisy recording conditions. By selecting wavefields and sorting them according to their amplitudes, the Gaussian-weighted crosscorrelation imaging method "collects" these energy concentrations. However, only contributions on the wavefront collapse with increasing propagation time at

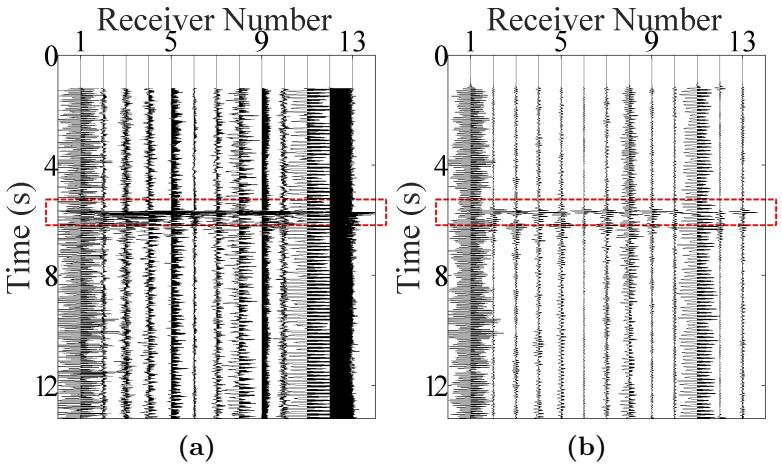


Figure 2.28: Data for the weak event from mining-induced seismicity (a). Band-pass filtered data (6 - 17 Hz) (b). The waveforms within the red dashed windows contain strong P-wave energy.

the hypocenter leading to high crosscorrelation results of these wavefields, whereas random foci are considerably smaller in amplitude since they are not on the wavefront and do not overlap with the wavefields at the source location.

This behavior results in a passive seismic source image of high resolution, which is almost free from imaging artifacts. Interpreting these images significantly reduces the risk of false positives. The computational effort is comparable with other time-reversal imaging techniques, which back project data in a single run. The selection of wavefields, the determination of maxima, and the crosscorrelation represent a small overhead when compared with the computational expense of the back projection process.

It is obvious that velocity errors would lead to localization errors. This applies to all methods that rely on the velocity model. In the numerical examples, all localization results are carried out for smooth models. The mean deviations from the true velocity model are in the order of 5% to 8% which appear fairly high for an acquisition in exploration. In the field data example, the results are obtained by using an homogeneous model. Velocity errors as considered in the examples provide localization results, which are in the resolution limit of the data. One may suspect that velocity errors may increase random focusing during the back projection process. Because of its superiority in suppressing imaging artifacts, the Gaussian weighted cross-correlation method can outperform the considered conventional techniques with respect to resolution and imaging noise.

The excellent performance in suppressing imaging artifacts also contributed positively to the case of clustered sources, a case observed in hydraulic stimulation or for earthquake swarms. The Gaussian-weighted crosscorrelation method spatially resolved a source cluster with high accuracy, where the spacing between the sources is below the prevailing wavelength of the data.

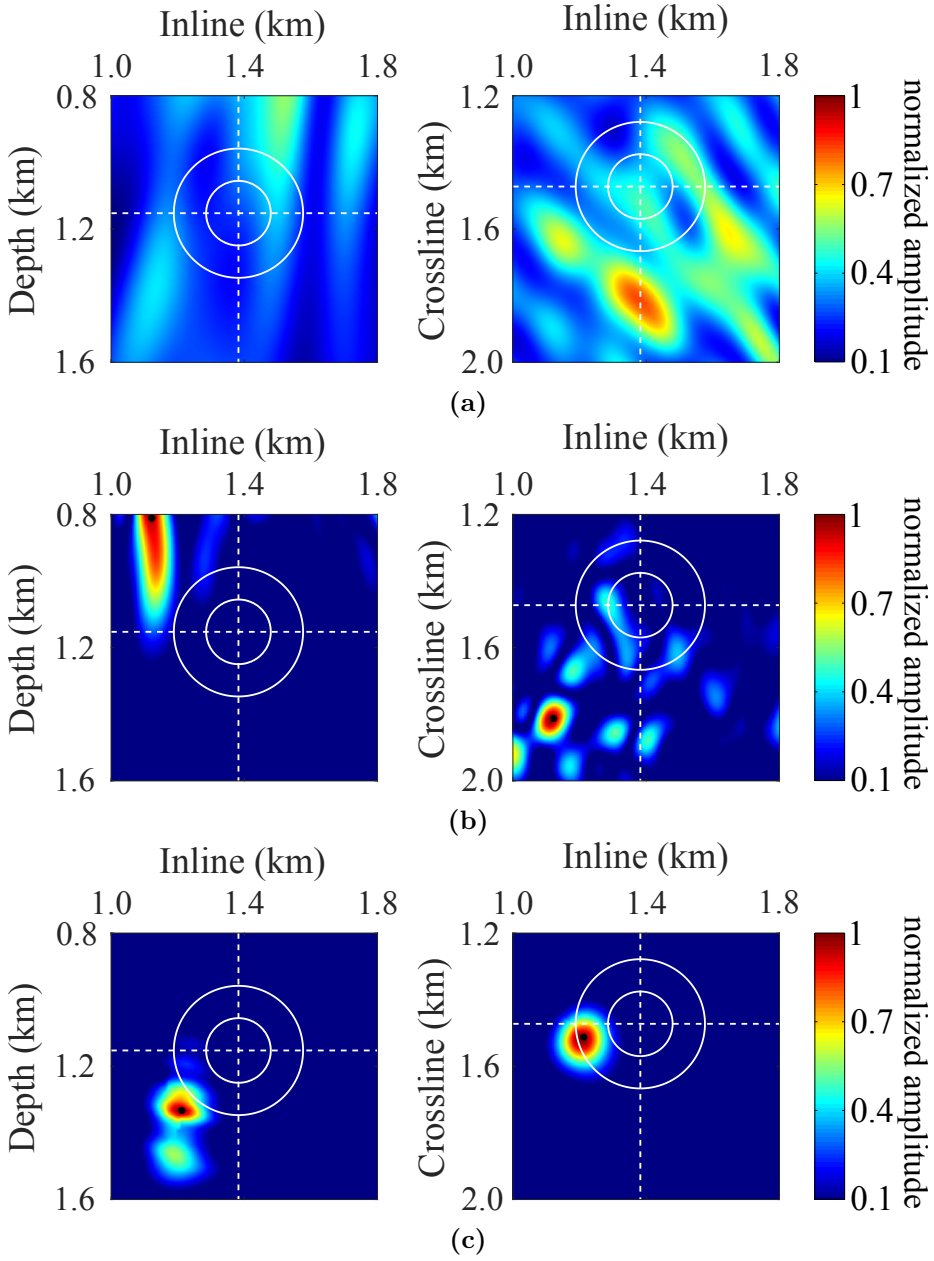


Figure 2.29: Comparison of three time-reversal imaging methods for a weak event from mining-induced seismicity. Images at the reference source locations provided by the travel-time inversion are obtained by auto-correlation imaging (a), grouped cross-correlation imaging (b), and Gaussian-weighted cross-correlation imaging (c). Black dots indicate the estimated source locations. The white circles, centered on the true source position, have radii of a quarter and a half of the dominant wavelength.

2.6 Conclusion

We propose a new imaging condition for passive seismic source imaging, which we call the Gaussian-weighted crosscorrelation imaging condition. The Gaussian-weighted crosscorrelation imaging method for microseismic source localization creates higher-resolution source images than the conventional autocorrelation imaging method. Next to good localization results, the method generates very few imaging artifacts in comparison with the autocorrelation and grouped crosscorrelation imaging methods, which highly reduces the risk of false positives in passive seismic source imaging interpretation. The new imaging condition is particularly suitable for noisy data acquired from sparse acquisitions and provides good results for source clusters. The method requires very little computational overhead compared with the efforts needed for the back projection process. Similar to other time-reversal source imaging methods, the proposed technique can be applied to multiple scales, ranging from acoustic emissions to microseismic events and global earthquakes.

2.7 Acknowledgments

We thank the Applied Seismics Group at the University of Hamburg for continuous discussions. We are grateful to the editor and reviewers for their comments and suggestions to improve the paper. Partial financial support from the University of Hamburg and the sponsors of the Wave Inversion Technology (WIT) Consortium is highly appreciated. P. Yang is also supported through a stipend from the China Scholarship Council. The SOFI package (https://www.gpi.kit.edu/english/498.php) is used for forward modeling and we developed codes for the time-reversal imaging. Other software packages for computing and data processing used in this work are developed through the support of the WIT consortium (https://www.wit.uni-hamburg.de/aboutwit/consortium.html).

3 Seismic source localization with time-reversal and maximum amplitude path for sparse and small-aperture acquisitions

Abstract

Time-reversal imaging is a powerful technique for localizing passive seismic sources. It is robust independent of the source time and applicable to heterogeneous media. However, time-reversal imaging usually requires a dense, uniform, and large-aperture acquisition that provides sufficient illumination to accurately localize passive seismic sources. This requirement often is not met in practice. In many field data cases, apertures are an insufficient and spatial sampling of the recorded wavefield is irregular and sparse because of complex topographical conditions, technical issues, budget problems, etc. For a sparse and small-aperture acquisition, it is difficult to identify events using common time-reverse imaging procedures because of strong imaging artifacts leading to false positives. To mitigate these problems, we have proposed a new approach that evaluates maximum-amplitude paths constructed from back-projected wavefields using receiver patches selected from the acquisition. These paths are built by scanning for the maximum absolute amplitude of the back-propagated wavefronts within a spatial dimension of the dominant wavelength of the considered event. The paths comprise maximum amplitudes of the back-projected wavefronts for each considered time step. The methods exploit the continuity of maximum amplitudes of the back-projected wavefields. The point of closest proximity (or crossing point) of the paths denotes the source location and the corresponding time is the source time. Numerical examples for the Marmousi-II and SEG 3-D overthrust models indicate that the proposed approach can estimate the source excitation time and source localization simultaneously for sparse and small-aperture acquisitions of noisy data with high accuracy. Source excitation time errors are below the prevailing period, and maximum localization errors are below the dominant wavelength, that is, within the resolution of band-limited signals. Using a wide aperture acquisition, these results are further improved and lead to errors in the order of the spatial and temporal sampling.

3.1 Introduction

The accurate location of seismic events is a key component in passive seismic monitoring. It can provide valuable information on the seismicity caused by tectonic activity (e.g., Waldhauser and Ellsworth, 2000) or induced by anthropogenic actions, such as hydraulic fracturing (e.g., Maxwell et al., 2010; Ellsworth, 2013; Kamei et al., 2015) or CO2 storage monitoring (e.g., Goertz Allmann et al., 2014; Jones et al., 2014). There are many approaches available for the location of seismic events, which can be mainly categorized into two types: ray-based methods and waveform-based methods(e.g., Bardainne et al., 2009; Li

et al., 2020). Most ray-based methods need to identify seismic phases and arrival times. For waveform-based methods, such as stacking/migration approaches, no picking is required. Waveform data are stacked to back-project recorded data to the source. In seismology, subsurface velocities often are not known and techniques for simultaneous location of sources and velocity model building are needed (Diekmann et al., 2019; Schwarz et al., 2016; Yang and Gajewski, 2021a). Most localization methods require the velocity structure between the passive seismic source and the receiver. A common approach is to estimate the locations and origin times of events by minimizing predicted arrival-time differences between pairs of events or stations (Waldhauser and Ellsworth, 2000; Eisner et al., 2009). Though efficient and reasonably reliable, this type of approach is limited by low signal-to-noise ratios (S/Ns), strong heterogeneity, and small-aperture seismic acquisition, resulting in large errors for the estimation of source excitation time and source location (Yang et al., 2019).

Time-reversal relies on the seismic wave equation and focuses the energy of seismic waves back to the hypocenter by reverse modeling if a reasonably accurate velocity model is available (Fink et al., 2002; Gajewski and Tessmer, 2005; Artman et al., 2010; Chen et al., 2021). Because the back-projection process starts as soon as the event is injected into the model, it is source time independent and does not require arrival-time picking or polarization estimation of seismic phases. It is more robust and reliable for long-period event localization (O'Brien et al., 2011) and suitable for passive seismic imaging (Folesky et al., 2015; Price et al., 2015). Larmat et al. (2008) approximate the location and focal mechanism of relatively large glacial earthquakes from time-reversal images. To improve the temporal and spatial resolution of images, Douma and Snieder (2015) perform focusing of microseismic images by deconvolution. Because time-reverse imaging methods focus on the wavefronts of recorded data, they are also suitable for localizing weak or noisy seismic events. However, in a sparse and small-aperture acquisition, reverse imaging approaches face challenges. If only a few traces or traces with a spacing considerably larger than the prevailing wavelength of the considered event are injected, no unique focusing of the wavefront is observed. To improve the spatial resolution, Nakata and Beroza (2016) propose an alternative imaging technique called geometric-mean reverse time migration (RTM)(GmRTM), which reduces imaging artifacts and works well in a sparse network. However, this approach also still relies on a large-aperture acquisition to obtain good estimates for the seismic events. We can summarize that time-reversal imaging techniques usually require uniform and/or large-aperture acquisitions.

Here, we propose a methodology to mitigate these issues by introducing the maximum-amplitude path technique, which provides source location and source time simultaneously for sparse and small aperture acquisitions. After the "Introduction", we describe how to obtain the maximum-amplitude path from back-propagated wavefields. Synthetic examples using the Marmousi-II and SEG 3D overthrust models are discussed for proof of concept and to quantify the method's location and source time estimation potential in laterally heterogeneous media. Results are compared with conventional time-reverse imaging techniques, namely, the auto-correlation imaging method (Steiner et al., 2008; Artman et al., 2010) and the GmRTM approach (Nakata and Beroza, 2016). Moreover, issues related to poor S/N and a wide-aperture acquisition also are quantified and discussed by numerical examples.

3.2 Outline of methodology

In this section, we provide a brief overview of the methodology and explain the details of its implementation in the subsequent sections. The maximum-amplitude path technique to localize passive seismic sources is a waveform-based method. The path is built by scanning for the maximum absolute amplitude of time-reversed wavefields. The search area is chosen according to the dominant wavelength of the considered event and is aligned in the direction of propagation of the back-propagated wavefront. This procedure is performed for each time step and exploits the continuity of amplitudes, which do not change discontinuously in the dimension of the search area.

The maximum amplitude paths are not classical rays obtained by ray tracing. These paths describe how the maximum amplitude of the wavefront back-propagates with time, that is, the driving engine of the method is time-reverse imaging. We start the determination of maximum amplitude when the injected time-reversed traces are forming a detectable wavefront by constructive interference, that is, focusing occurs. We construct three maximum-amplitude paths from the time-reversed data of three receiver patches selected from the recorded traces. The back-projected wavefront and therefore the maximum-amplitude paths focus on the hypocenter. Here, the paths cross or display the closest proximity to each other, which defines the source location. The time at this location corresponds to the source excitation time.

Thus, the proposed method comprises two main steps:

(1) maximum-amplitude path construction

(2) seismic event and source time estimation.

The exploitation of the continuity of the maximum-amplitude path and the restriction of the search area for the maximum to an area with the dimension of the prevailing wavelength of the considered events distinguishes this approach from other reverse modeling techniques using the maximum-amplitude imaging condition. Details of the previously described outline are provided in the following sections.

3.3 Maximum-amplitude path construction

We first perform some preprocessing on the recorded data, comprising the selection of traces for the considered receiver patches, and frequency analysis (e.g., Priestley, 1996) followed by band-pass filtering. Receiver patches and dominant frequency provided by the frequency analysis are used for constructing maximum-amplitude paths.

There are many options to select receiver patches out of the seismic acquisition. We can select patches from the first to the last receiver, or we may consider overlapping receiver patches. Alternatively, we may randomly select three receiver patches. A good choice is overlapping patches providing different illumination. Local coherence of the wavefield (Schwarz et al., 2016; Diekmann et al., 2019) can aid the selection process. The recorded wavefield in the patches should include the seismic event that we wish to localize. At least two maximum-amplitude paths are constructed from the data of the selected receiver patches. Here, we construct three paths from three patches, in which we use the following selection strategy for the receiver patches: the first and second receiver patch are overlapping

by a few traces with each other while the third receiver patch comprises all receivers of the acquisition. The acquisition may be sparse and of limited aperture, that is, a situation often observed in field data studies. After the selection of receiver patches, the seismograms are reversed in time and serve as boundary values for the reverse modeling. The reverse modeling requires a velocity model that is consistent with the data. We assume here that such a model is available. Some details on the requirements of the velocity model are discussed in the following numerical examples.

To localize passive seismic sources using the maximum-amplitude path method, we apply reverse modeling, that is, back projecting the recorded seismic energy to its origin. There are various types of wave equations for the propagation algorithm. In the present work, the first-order velocity-stress acoustic wave equation for constant density (Aki and Richards, 2002; Cerjan et al., 1985) is adopted. The spatial and time derivatives are computed by first-order finite differences. To avoid backscattering, we separate up- and downgoing wavefields using the Hilbert transform (Zheng et al., 2018), which provides the downgoing back-propagated pressure fields. Unless otherwise noted, back-propagated wavefields refer to the downgoing fields.

The determination of the maximum-amplitude path starts after focusing caused by constructive interference of the back-projected wavefield is observed and a wavefront is formed. Before sufficient interference is obtained, maximum absolute amplitude locations of the back-projected wavefields are scattered close to the surface where the data are injected at the receivers. This is displayed in Figure 3.1 in which the maximum-amplitude path is built without any constraints, that is, only by picking the locations of the absolute maximum amplitudes for each time step. The black, purple and white curve denotes the path for the receiver patches comprising receivers between 0.0 and 0.8, 0.4 and 1.2, and 0.0 and 1.2 km at the surface, respectively. The red circle denotes the source at $x = 3.0$ km and $z = 3.0$ km. We observe that the maximum-amplitude path starts to be spatially continuous at a certain depth. This continuity reflects the focusing of the injected wavefields and allows the construction of the path. The determination of the maximum amplitude path thus should not start above a threshold "depth" Z_b, displayed in Figure 3.2. As a good general rule of thumb for this threshold, we use a depth corresponding to two times the prevailing wavelength of the considered event. When the depth of the maximum amplitude exceeds this threshold, its corresponding time is denoted by t_b. At this time, the determination of the maximum-amplitude paths start. The prevailing wavelength is obtained from the dominant frequency of the event as determined by spectral analysis in the preprocessing and an average of the velocities at the receiver patches. During the determination of the maximum amplitude paths, we also evaluate the distances between them. These distances decrease with increasing time until a minimum is reached. This minimum corresponds to the source location and the corresponding time is the hypo-central time. After this time is reached, the paths start to diverge and the determination of the paths is stopped. This time is denoted by t_e. After the general description of the method, we provide additional technical information on the implementation of the maximum amplitude path method in the following paragraphs.

The path for each receiver patch is constructed by scanning maximum absolute amplitudes of the back-propagated wavefields for each time step between t_b and t_e, which is schematically visualized in a 2D sketch shown in Figure 3.2. Two snapshots of adjacent time steps are displayed. The first maximum is determined from the amplitudes of the

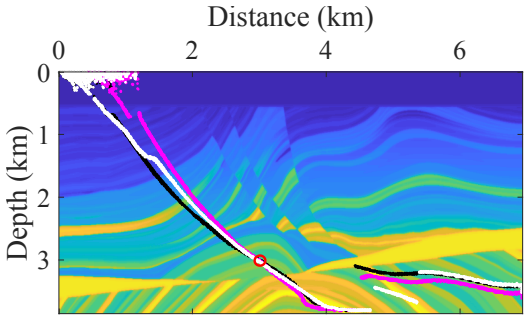


Figure 3.1: Maximum-amplitude paths for the source at $x = 3.0$ km, $z = 3.0$ km without constraints of the time window. The black, purple and white curves are the maximum-amplitude paths for the first, second and third receiver patch containing receivers between 0.0 and 0.8, 0.4 and 1.2, and 0.0 and 1.2 km, respectively. Receiver sampling is 20 m. Maximum amplitudes are scattered close to the surface since the injected wavefields are not yet interfering and focusing.

back-projected wavefield at the time t_b. The following maxima are determined by scanning a rectangular box centered at the previous maximum where the axes of the box have a length corresponding to the prevailing wavelength of the considered event and the box is aligned in the direction of propagation, which is computed numerically (e.g., Jia and Wu, 2009).

The wave propagation angle is used for a rotation matrix R to align the rectangular box in 2D (or cube in 3D) into the direction of propagation, see also Figure 3.2. The position for the maximum absolute amplitude within the rectangular box is picked using the following expression,

$$u_j(\mathbf{x}_m, t_{i+1}) = \max\left\{|u_j(\mathbf{x}, t_{i+1})\, H(\mathbf{x}_m - R\cdot\mathbf{x}, t_i)|\right\} \quad , \quad 1 \le j \le 3 \quad , \tag{3.1}$$

with

$$H(\mathbf{x}_m - R\cdot\mathbf{x}, t_i) = \begin{cases} 1, & |\mathbf{x}_m - R\cdot\mathbf{x}| \le \lambda_{ji} \\ 0, & \text{otherwise} \end{cases} \quad , \tag{3.2}$$

where u_j is the back-propagated wavefield generated in the j-th receiver patch; the vector $\mathbf{x}_m$ denotes the position of the maximum absolute amplitude of the wavefield, whereas t_i and t_{i+1} represent the time at the i-th and $(i+1)$-th time step, respectively; λ_{ji} is the prevailing wavelength at the maximum $\mathbf{x}_m$ obtained at the i-th time step; and H is the step or Heaviside function. The maximum-amplitude path comprises all picked maxima $\mathbf{x}_m$ between starting and ending time for the considered patch.

3.4 Seismic event localization and source time estimation

The back-propagated wavefront shrinks to its spatial minimum at the source location. Therefore, the maximum-amplitude paths should focus on the source location, and the corresponding time is the source excitation time. The constructed maximum-amplitude paths do not necessarily cross at the source location because these paths do not correspond

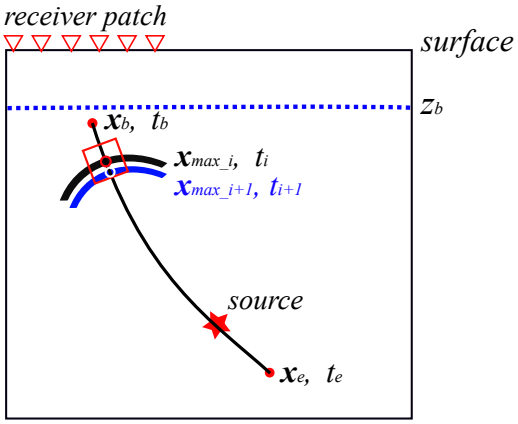


Figure 3.2: Schematic sketch of the maximum-amplitude path construction. The threshold Z_b is a user-defined parameter for determining the maximum-amplitude path start $(\mathbf{x}_b, t_b)$. Two snapshots at subsequent time steps t_i and t_{i+1} are displayed. The black curve describes the maximum-amplitude path. Assuming that the maximum amplitude for the time step t_i is located at $\mathbf{x}_{max\_i}$, the maximum for the next time step $\mathbf{x}_{max\_i+1}$, denoted by the white circle, is determined by scanning the amplitude of the wavefront for this time step, that is, t_{i+1}, in the neighborhood of the previous maximum. This neighborhood, indicated by the red rectangular box, is centered at the previous maximum and oriented in propagation direction with side lengths corresponding to the prevailing wavelength of the signal.

to classical rays but to the paths following maximum amplitude of the propagating wavefront. Therefore, we consider the closest proximity of the maximum-amplitude paths as the source location. We compute the distance between the paths for each time step. For the case of three paths, we obtain three distances, which are summed. The position with the minimum of these sums represents the source location and the time at this location is the source time. The expression for calculating the sum, denoted by D, of the distances reads

$$D(t) = \sum_j \sum_k \sqrt{(\mathbf{x}_{jm}(t) - \mathbf{x}_{km}(t)) \cdot (\mathbf{x}_{jm}(t) - \mathbf{x}_{km}(t))^{\mathrm{T}}} \quad , \qquad j \neq k \quad , \tag{3.3}$$

where the subscripts j and k (with $j \neq k$) denote the receiver patch and $\mathbf{x}_{jm}$ and $\mathbf{x}_{km}$ correspond to the locations with the maximum absolute amplitude for the considered time steps. After the maximum-amplitude paths start, that is, at starting time t_b, the sum of the distances D would first decrease, minimize at the source location, and then increase, which can be visualized in Figure 3.5 c and 3.5 d. Therefore, the paths should stop when the sum of the distances increases. Accounting for model complexity and strong noise, instead of the sum of the distances D, we consider its average in a user-defined time window. It reads

$$\bar{D}_l = \frac{\sum_t D(t)}{N} \quad , \quad t_b + l \cdot w \leq t < t_b + (l+1) \cdot w \quad , \tag{3.4}$$

where $\bar{D}_l$ denotes the average of D in the l-th time window, $N = w/dt$ denotes the number of time steps in the time window, dt indicates the time step used for reverse modeling, and w represents the user-defined time window. A good choice of its size is two prevailing periods of the considered event. When $\bar{D}_l$ increases, the determination of the paths stop

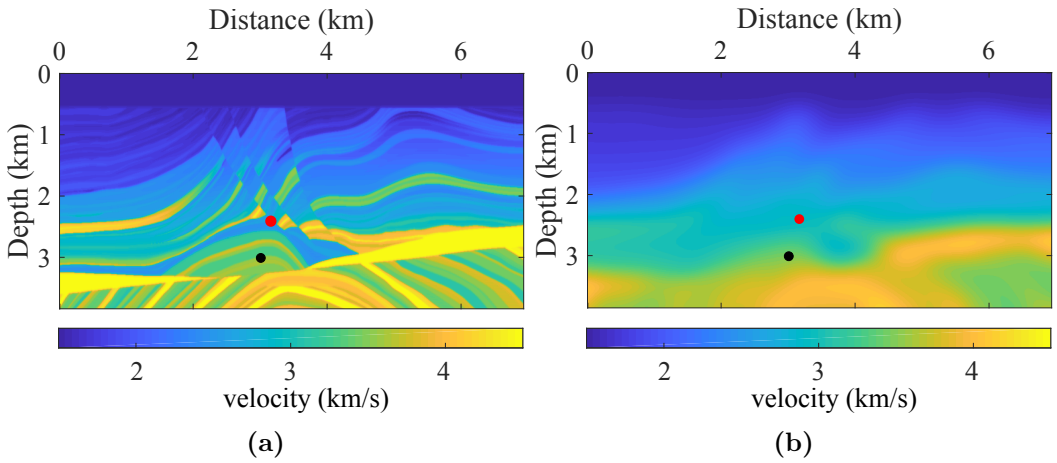


Figure 3.3: Velocities for a part of the Marmousi-II model. The black and red dots denote the two considered sources located at $x = 3.0\,\text{km}$, $z = 3.0\,\text{km}$ and $x = 3.15\,\text{km}$, $z = 2.4\,\text{km}$, respectively. (a) The true velocity model used to generate data and (b) the smoothed version used for reverse modeling.

and defines t_e. The minimum of the sum of the distances D over the time range L, that is, from t_b to t_e, provides the source location and source time:

$$\min_{L} [D(t)] \quad \Rightarrow \quad (\mathbf{x}_s,\ t_s) \quad , \tag{3.5}$$

where t_s indicates the source excitation time and $\mathbf{x}_s$ is the corresponding source location.

3.5 Numerical studies

For proof of concept for the localization method, Marmousi-II and SEG 3D overthrust models are used. The synthetic data for the tests have been computed using the constant-density two-way first-order velocity-stress acoustic wave equation (Aki and Richards, 2002; Gajewski and Tessmer, 2005) and the respective time-reversed data serve as input for reverse modeling using the same acoustic wave equation with a scheme for up- and downgoing wavefield separation using the Hilbert transform (Zheng et al., 2018). Forward modeling of the data is performed with the true velocity model whereas reverse modeling is performed in a smoothed version of the true model. For the synthetic examples, the exact locations and excitation times are known. Unless otherwise noted, all times mentioned in the following tests are relative to the exact source time, which is 0 s (zero time). Negative numbers correspond to times prior to the real excitation time and positive numbers indicate times after the real excitation time. We consider first a subset of the Marmousi-II model.

3.5.1 Marmousi-II model

In Figure 3.3 a, a part of the Marmousi-II model is displayed. Velocities are defined on a 693×385 grid with a spacing of 10 m in the x- and z- directions. Explosive sources denoted by a black and red dot are located at $x = 3.0\,\text{km}$, $z = 3.0\,\text{km}$ and $x = 3.15\,\text{km}$, $z = 2.4\,\text{km}$, respectively. The latter source is placed on a fault in the presence of strong

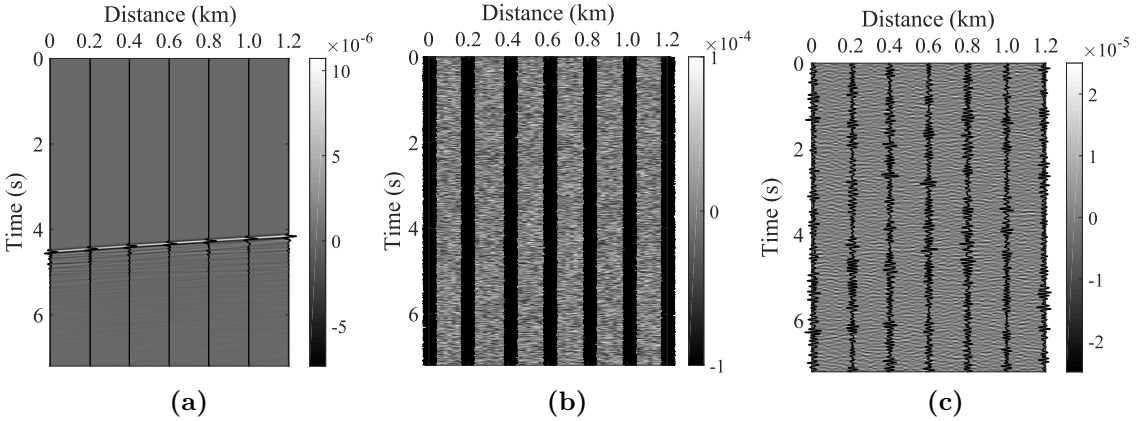


Figure 3.4: Data for the source at $x = 3.0\,\text{km}$, $z = 3.0\,\text{km}$ in the Marmousi-II model. The background image is the wavefield recorded at the surface, and the black lines illustrate waveforms at seven receivers used below for a sparse data acquisition. (a) Noise-free data, (b) noisy data with S/N of 0.12, and (c) band-pass filtered (6 - 17 Hz) data.

velocity variations. The source time function is a Ricker wavelet with a dominant frequency of 10 Hz. A time shift was applied to simulate a minimum phase signal. The time sampling is 1 ms.

A data example is shown in Figure 3.4. The time reversed data serve as input for reverse modeling. Reverse modeling is performed in the smoothed version of the Marmousi-II model (Figure 3.3 b). Smoothing is performed with a $600 \times 600\,\text{m}^2$ 2D Gaussian window. This model is used for reverse modeling to construct maximum-amplitude paths, as well as for GmRTM and auto-correlation imaging. We consider a smoothed version of the velocity model with a smoothing window considerably larger than the prevailing wavelength of the event because, in practice, it is unlikely to achieve a velocity resolution comparable or close to the true velocity model.

We will use the synthetic data for numerous tests beginning with densely sampled noise-free data. In further examples, we investigate noisy and sparse data with small apertures. Finally, we present a wide-aperture acquisition. Results of two time-reverse imaging methods, namely auto-correlation imaging (Artman et al., 2010) and the GmRTM method (Nakata and Beroza, 2016), are used for comparison with the results of the maximum-amplitude path approach. Both time-reverse imaging methods localize sources by scanning for the spatial position with the maximum absolute value of the image.

Densely sampled noise-free data

In a first example, we consider the source at $x = 3.0\,\text{km}$ and $z = 3.0\,\text{km}$. Receivers are evenly distributed at the surface between 0.0 to 1.2 km with a spacing of 20 m. The data example is shown as an image plot in Figure 3.4 a. This example reflects a densely sampled acquisition with limited aperture.

The threshold for determining the starting time is set to 650 m below the surface, and the ending time of the paths is automatically estimated by equation 3.3 and equation 3.4. This provides the duration L of the time window for the construction of maximum-amplitude paths.

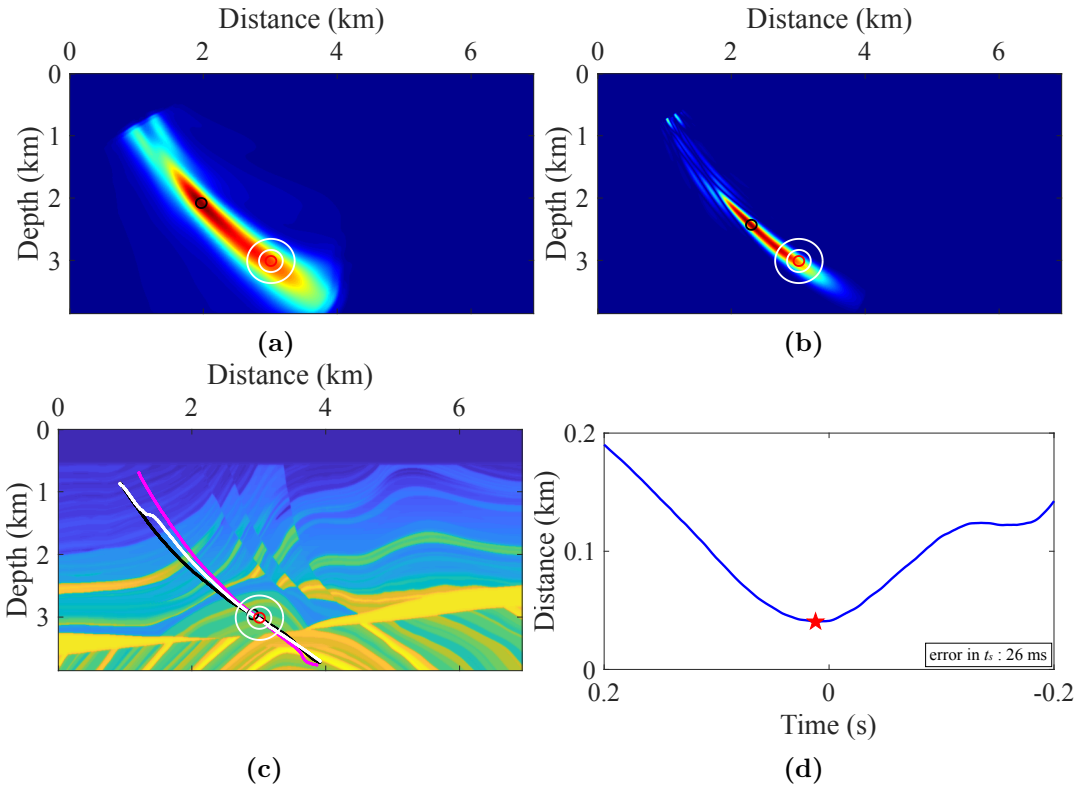


Figure 3.5: Comparison of methods for the source at $x = 3.0$ km and $z = 3.0$ km. Receivers are regularly distributed at the surface between 0.0 and 1.2 km with a spacing of 20 m. The red and white circles are centred at the true source location and with radii of half and one dominant signal wavelength. The black circles represent the localization results by (a) the auto-correlation imaging method, (b) the GmRTM method, and (c) the maximum-amplitude path method using three receiver patches comprising receivers between 0.0 and 0.8, 0.4 and 1.2, and 0.0 and 1.2 km, respectively. (d) The distance between maximum-amplitude paths against time with a red star as a minimum, determines source excitation time.

Figure 3.5 a and 3.5 b shows the time-reverse images, obtained by auto-correlation imaging and the GmRTM method, respectively. Unless otherwise noted, time-reverse images for these methods are obtained by stacking all snapshots over duration L in which only the area beyond the threshold is considered. In Figure 3.5 a and 3.5 b, the source region is smeared because of the limited aperture of the surface observations. The spatial resolution of the images is low, resulting in localization uncertainties and incorrect source times. The estimated source locations in the images are denoted by black circles, whereas the true locations are indicated by red circles. The white circles around the true source location display a radius of half and one dominant wavelength of the considered event. In Figure 3.5 a and 3.5 b, the distances between the true and the estimated source locations are approximately four times the dominant wavelengths leading to substantial deviations in source time. We conclude that the two common time-reverse imaging methods fail to properly localize the source for the dense acquisition with small aperture used in this example.

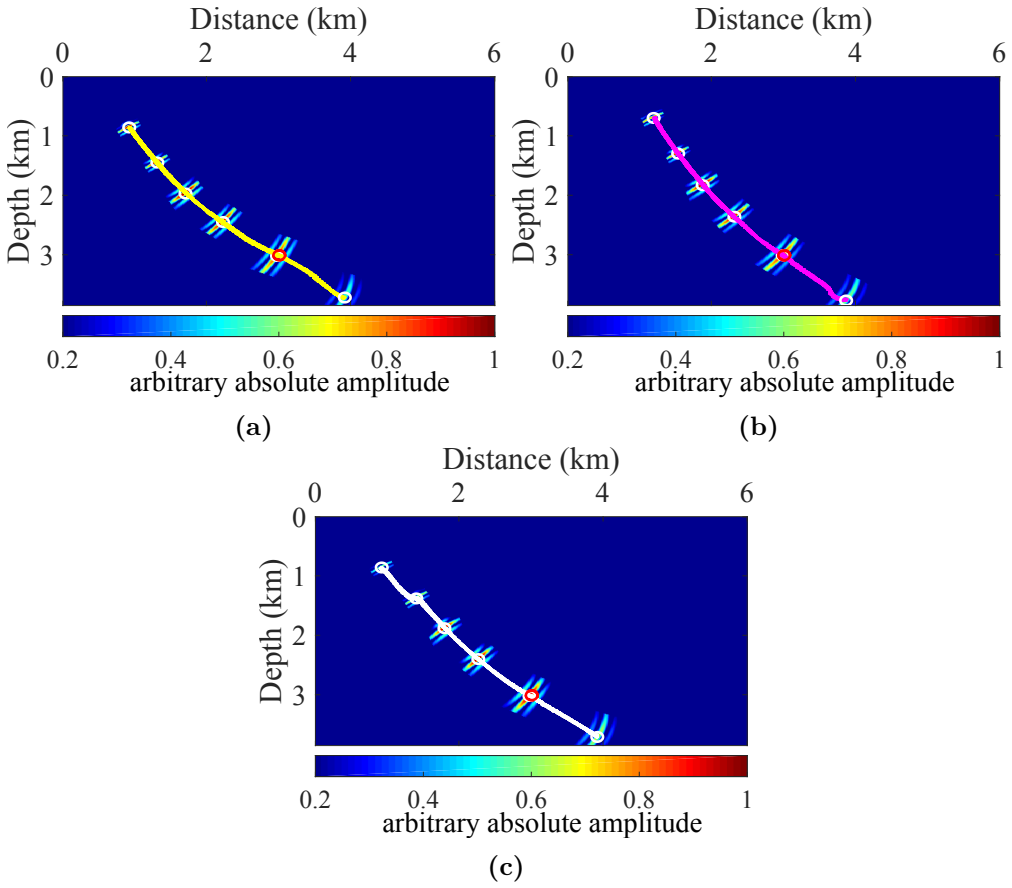


Figure 3.6: Images of the analysis window along the maximum-amplitude path for (a) the first receiver patch, (b) the second receiver patch, and (c) the third receiver patch at the time of 1.25, 0.85, 0.55, 0.3, 0 and –0.3 s. Receivers are regularly distributed at the surface between 0.0 and 1.2 km with a spacing of 20 m. The yellow, purple, and white curve represent the maximum-amplitude path using three receiver patches comprising receivers between 0.0 and 0.8, 0.4 and 1.2, and 0.0 and 1.2 km, respectively. The red circles are centered at the actual source location with $x = 3.0$ km and $z = 3.0$ km, and the white circles are centered at the point on the path at the respective time.

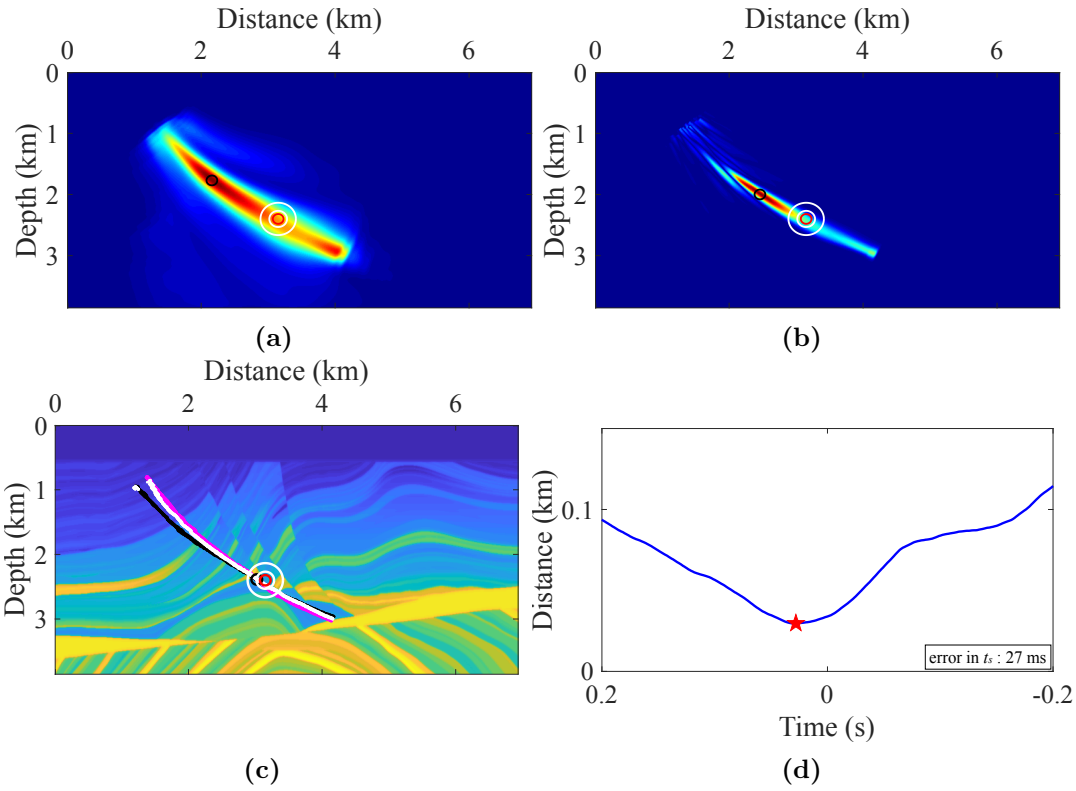


Figure 3.7: Comparison for the source at the fault ($x = 3.15$ km and $z = 2.4$ km). Receivers are regularly distributed at the surface between 0 and 1.2 km with a spacing of 20 m. The red and white circles are centred at the true source location. The white circles have radii of half and one dominant signal wavelength. The black circles represent the localization results by (a) the auto-correlation imaging method, (b) the GmRTM method, and (c) the maximum-amplitude path method using three receiver patches comprising receivers between 0.0 and 0.8, 0.4 and 1.2, and 0.0 and 1.2 km, respectively. (d) The distance between maximum-amplitude paths against time, with the red star denoting a minimum, determines source excitation time.

For the maximum-amplitude path method we choose three receiver patches. The patches comprise receivers between 0.0 to 0.8, 0.4 to 1.2, and 0.0 to 1.2 km, respectively. Together with the maximum-amplitude path, images with the size of the analysis window evaluated by equation 3.2 at time steps 1.25, 0.85, 0.55, 0.3, 0, and –0.3 s are displayed in Figure 3.6. The paths start at 1.25 s and end at –0.3 s. At the actual source time, the paths are close to the actual source location. Figure 3.5 c shows that the maximum-amplitude path (the black, purple, and white lines) is focusing close to the actual source location. In Figure 3.5 d, the blue curve corresponds to the sum of the distance between paths as defined previously. The minimum sum is indicated by a red star at 26 ms. It corresponds to the determined source excitation time which is within the prevailing period of the source signal. The source location error is about 100 m, that is, about 32 % of the dominant wavelength. For this small-aperture dense acquisition, we conclude that the maximum-amplitude path approach yields a good estimate of source excitation time and location. Errors are within

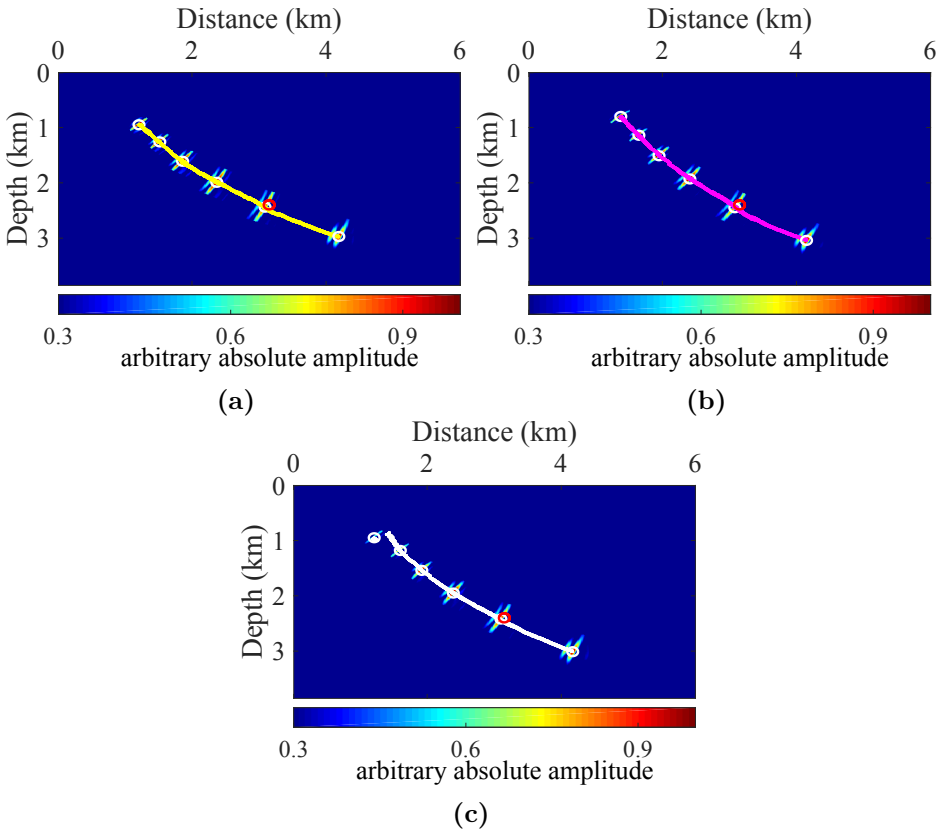


Figure 3.8: (a-c) Maximum-amplitude paths for three receiver patches. The patches include receivers between 0.0 and 0.8, 0.4 and 1.2, and 0.0 and 1.2 km, respectively. Receiver spacing is 20 m. Images of the analysis window at times of 1.05, 0.8, 0.55, 0.3, 0, and –0.42 s also are displayed. The red circles are centered at the actual source location with $x = 3.15$ km and $z = 2.4$ km, and the white circles denote the point on the path at the respective time. The yellow, purple, and white curve represent the maximum-amplitude path using three receiver patches comprising receivers between 0.0 and 0.8, 0.4 and 1.2, and 0.0 and 1.2 km, respectively.

the resolution of the seismic event.

In a second example we consider a passive seismic event on a fault with strong variations in velocity in its vicinity. This source is located at $x = 3.15$ km and $z = 2.4$ km (the red dot in Figure 3.3). The acquisition as well as the selection of receiver patches are the same as for the previous numerical test. The common time-reverse imaging methods displayed in Figure 3.7 a and 3.7 b again show strong deviations from the true location. Figure 3.8 displays the images within the analysis window evaluated by equation 3.2 along the maximum-amplitude paths at times 1.05, 0.8, 0.55, 0.3, 0, and –0.42 s. Unlike classic rays, the maximum-amplitude path displays a discontinuous change at the beginning of the white path in Figure 3.8 c . Choosing a different threshold to estimate the starting time of the path may have avoided this. However, with increasing time, the continuity of the path is established. The result of the maximum-amplitude path method is displayed in Figure 3.7 c and 3.7 d. Although the source is located in a challenging geologic region,

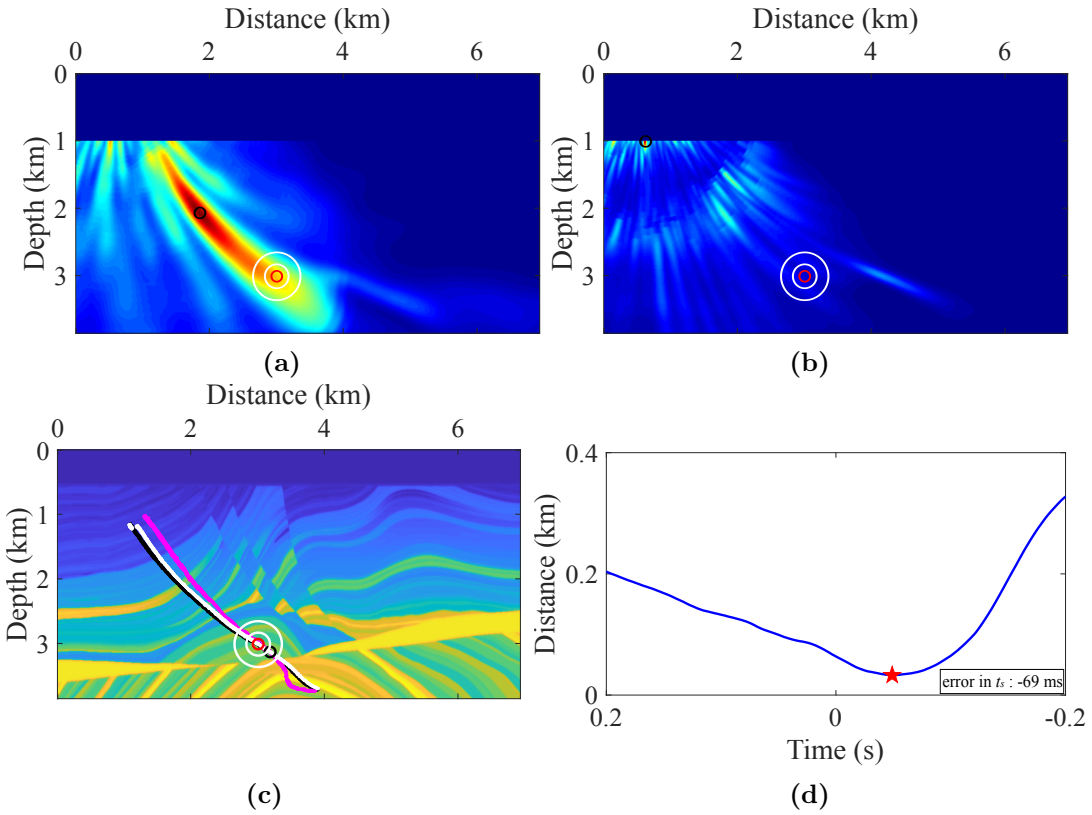


Figure 3.9: Comparison of methods for noisy data. The source is located at $x = 3.0$ km and $z = 3.0$ km. Receivers are regularly distributed at the surface between 0.0 and 1.2 km with a spacing of 20 m. The red and white circles are centred at the true source location. The white circles have radii of half and one dominant signal wavelength. The black circles represent the localization results by (a) the auto-correlation imaging method, (b) the GmRTM method, and (c) the maximum-amplitude path method using three receiver patches comprising receivers between 0.0 and 0.8, 0.4 and 1.2, and 0.0 and 1.2 km, respectively. (d) The distance between maximum-amplitude paths against time, with the red star denoting a minimum, determines source excitation time.

the distance between the actual source location and the location where the paths focus is again smaller than the prevailing wavelength. Furthermore, the source excitation time error is smaller than the prevailing period of the signal. This promising result emphasizes the potential of the maximum-amplitude path method for limited aperture networks even for strongly heterogeneous media.

Noisy data

To examine the effect of uncorrelated noise on the localization accuracy of the maximum-amplitude path method, random noise with an S/N of 0.12 is applied to the data. The resulting data for a source at $x = 3.0$ km and $z = 3.0$ km and receivers with a 20 m spacing between 0.0 and 1.2 km at the surface are displayed in Figure 3.4 b. The seismic event might be too weak to obtain focusing in the back-propagation process. A conventional

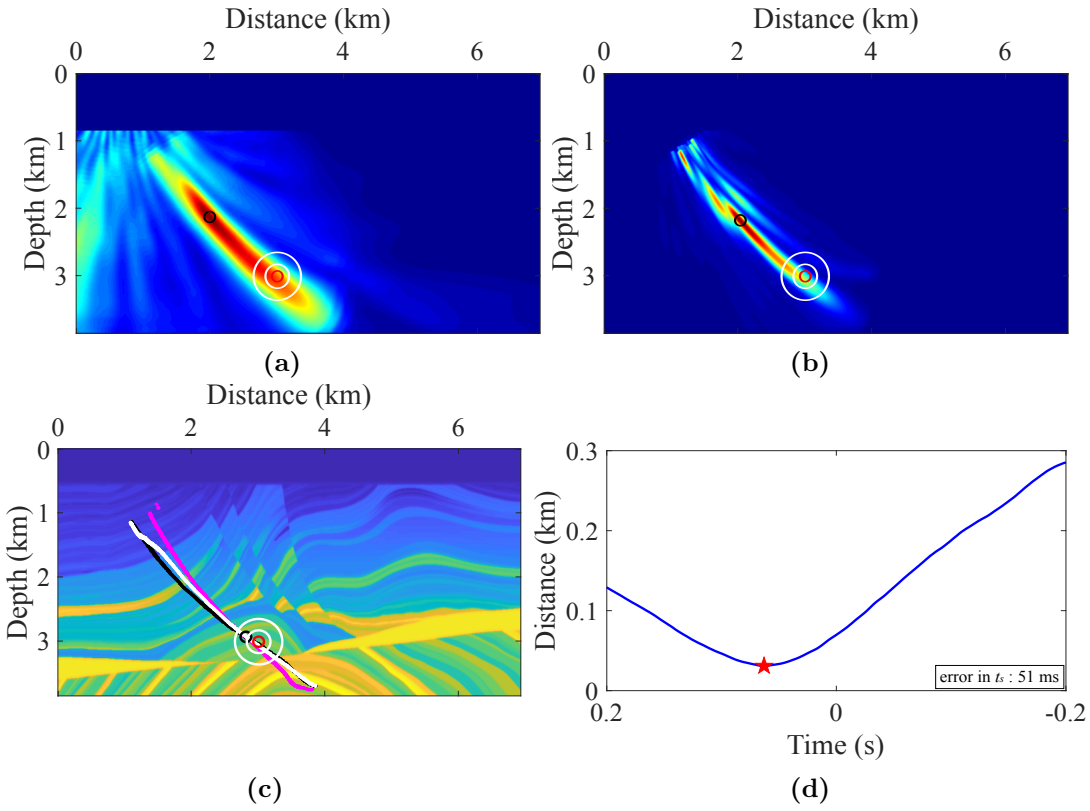


Figure 3.10: Comparison of methods for sparse and noisy data comprising seven traces. The source is located at $x = 3.0$ km and $z = 3.0$ km. Receivers are placed at the surface between 0.0 and 1.2 km with a spacing of 200 m. The red and white circles are centred at the true source location. The white circles have radii of half and one dominant signal wavelength. The black circles represent the localization results by (a) auto-correlation, (b) GmRTM, and (c) maximum-amplitude path using three receiver patches comprising receivers between 0.0 and 0.8, 0.4 and 1.2, and 0.0 and 1.2 km, respectively. (d) The distance between maximum-amplitude paths against time, with the red star denoting a minimum, determines source excitation time.

preprocessing step is band-pass filtering. We apply a $6-17$ Hz band-pass to the data, which are displayed in Figure 3.4 c. The S/N ratio after filtering is 0.9. Although the event is now weakly visible in the densely sampled data, the waveforms of the sparse data used in the examples next do not allow a correlation of the event. In this example, the threshold for estimating the starting time is chosen to be 1.0 km. The acquisition as well as the receiver patches are the same as for the first numerical test in the previous section. In Figure 3.9 a and 3.9 b, only the area below 1.0 km is considered. Auto-correlation and GmRTM result in incorrect estimates far off the true source location.

The results of the maximum-amplitude paths approach are shown in Figure 3.9 c and 3.9 d. The paths focus approximately 200 m off the true source location, which is within a dominant wavelength of the event (about 62 % of the dominant wavelength). The corresponding estimated source excitation time error is –0.069 s, which is smaller than the

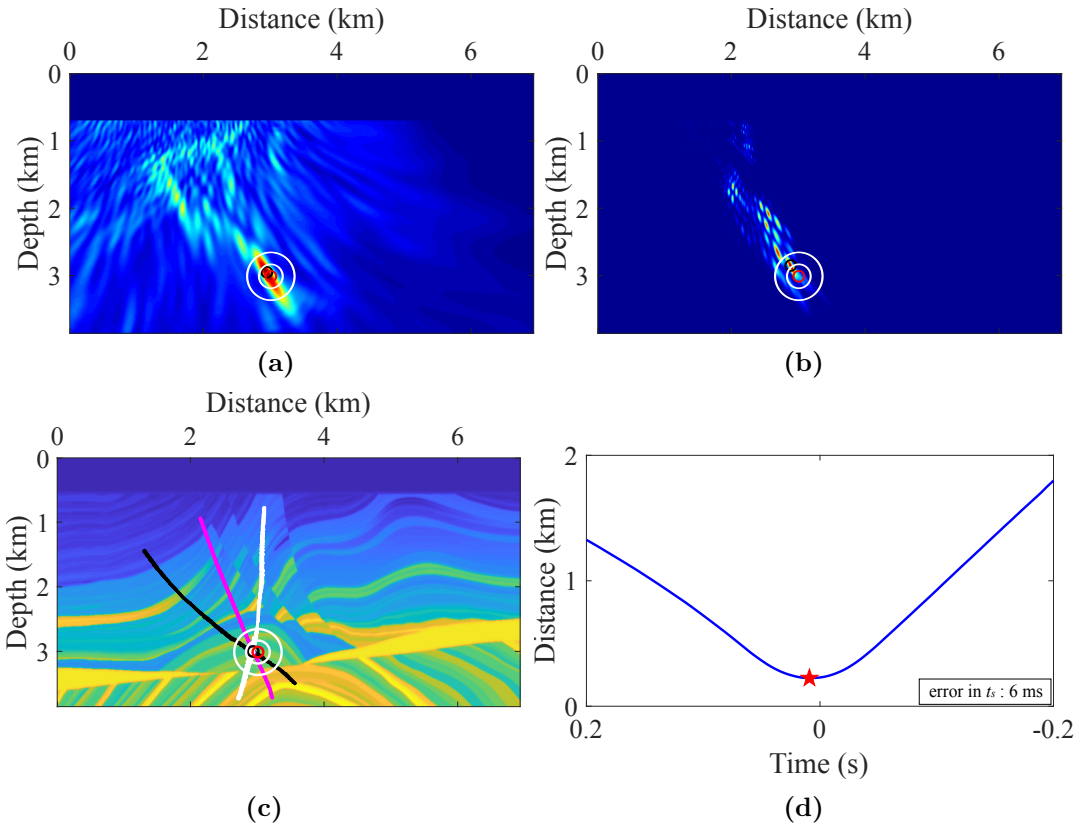


Figure 3.11: Comparison of methods for wide-aperture data. The source is located at $x = 3.0$ km and $z = 3.0$ km. Receivers are located between 0.0 and 0.8, 1.4 and 2.2, and 2.8 and 3.6 km with a spacing of 200 m. Localization results for (a) auto-correlation (b) GmRTM, and (c) maximum-amplitude path. The red and white circles are centred at the true source location. The white circles have radii of half and one the dominant signal wavelength. The black circles indicate the estimated source locations. (d) The distance between maximum-amplitude paths against time, with the red star denoting a minimum, determines source excitation time.

prevailing period of the signal. This result illustrates that the maximum-amplitude path approach yields good source estimates for small-aperture acquisitions even if the event is weak.

Sparse and noisy data

For more realistic acquisitions, the number of receivers is usually considerably smaller than in the previous examples. Therefore, we carry out another test for a source at $x = 3.0$ km and $z = 3.0$ km where the number of receivers is only approximately one-tenth of the first example. This means that only seven receivers are regularly distributed at the surface with a spacing of 200 m. The receivers are located at 0.0, 0.2, 0.4, 0.6, 0.8, 1.0, and 1.2 km, respectively. Again, random noise with S/N 0.6 is added to the sparse data. Band-pass filtering (6 – 17 Hz) establishes an S/N of 4.35 in this case. The seven waveforms of the data

are displayed in Figure 3.4. Figure 3.10 a and 3.10 b shows the images below the threshold depth at 0.85 km and the estimated locations for auto-correlation and GmRTM, which are far off the true source positions. Results of the maximum-amplitude path approach are shown in Figure 3.10 c and 3.10 d. The first receiver patch contains five receivers from 0.0 to 0.8 km, whereas the second receiver patch contains five receivers from 0.4 to 1.2 km. The third receiver patch includes all seven receivers of the acquisition. The threshold of 0.85 km leads to a starting time of 1.05 s to ensure sufficient focusing and continuity of maximum amplitudes. The estimates of the source excitation time and location are similar to the previous examples. Therefore, we conclude that the maximum-amplitude path localization scheme provides promising results even for very sparse small-aperture acquisitions.

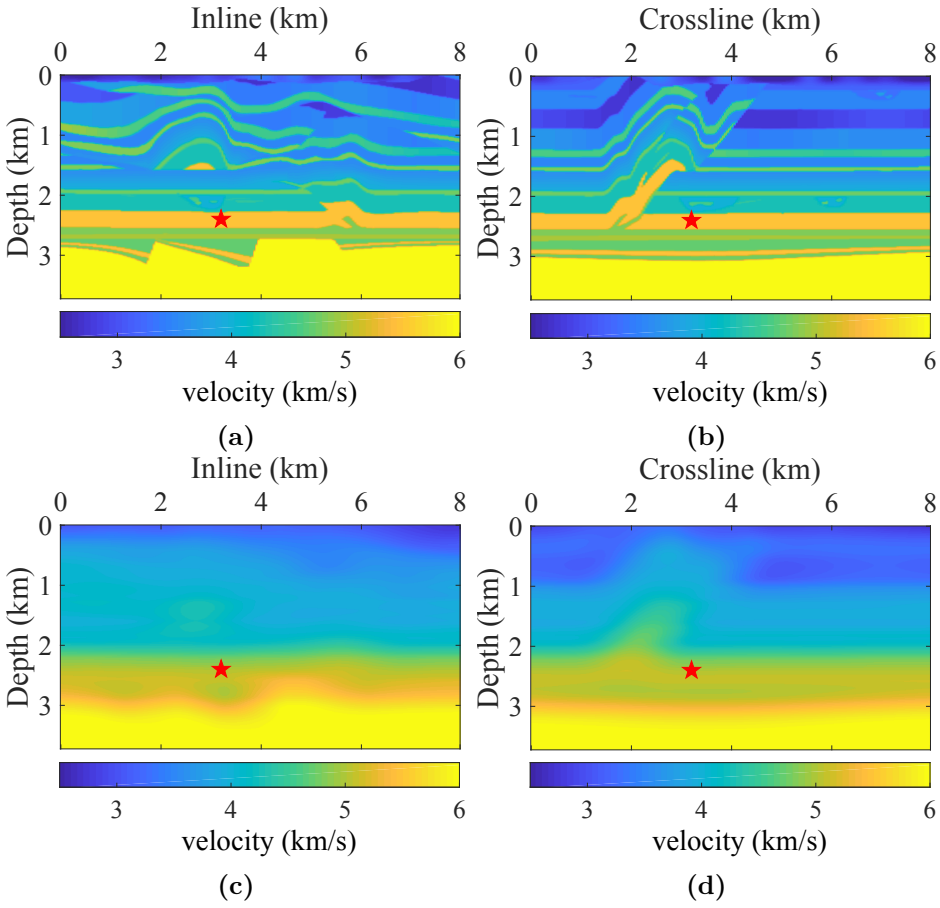


Figure 3.12: Velocity sections for the SEG 3-D overthrust model at the source location. (a) Inline section and (b) crossline section. (c) Inline section and (d) crossline section of the smoothed model. The red star denotes the source located at $x = 3.2$ km, $y = 3.2$ km, and $z = 2.4$ km.

Wide-aperture acquisition

In the previous examples, we have considered small apertures and sparse acquisitions to be close to most field data situations. The maximum amplitude path method, however, is not

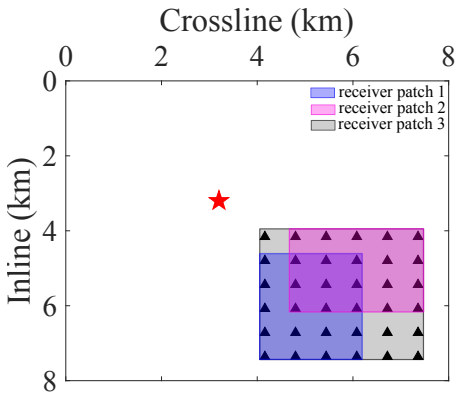


Figure 3.13: Seismic network for the source at $x = 3.2$ km, $y = 3.2$ km, and $z = 2.4$ km. The red star shows the surface projection of the source. The black triangles denote receivers placed between 4.16 and 7.36 km in both x- (crossline-) and y- (inline-) direction. Receiver spacing is 640 m. The first, second, and third receiver patches comprise receivers of the blue, purple, and gray rectangular boxes, respectively.

restricted to these cases but works equally well for wide-aperture acquisitions. We consider three separated receiver lines distributed over the surface where the first line covers the area from 0.0 to 0.8 km lateral distance, the second line covers the area from 1.4 to 2.2 km, and the third line covers the area from 2.8 to 3.6 km. The receiver distribution has a spacing of 200 m for each line for the source at $x = 3.0$ km and $z = 3.0$ km. The three recording lines form three receiver patches for the maximum-amplitude path approach. Each receiver patch has a sparse and small-aperture setting. This acquisition geometry provides a wide aperture layout and delivers a much more favorable situation for source localization because of the improved illumination. Random noise and a band-pass filtering are the same as for the previous example.

This acquisition requires less effort and costs if compared to a full-aperture acquisition with a single 6.4 km receiver line. For this case, all methods provide good locations. Based on the previous small-aperture results, it is no surprise that the maximum-amplitude path method provides very good locations and times for this type of acquisition. The source location and source time are very close to the true values (Figure 3.11 c and 3.11 d). The source time deviates by only 6 ms, and the source location deviates by approximately 30 m, which is far below the dominant wavelength of the signal. It has to be emphasized, however, that very good results already have been obtained for cost-effective sparse small-aperture acquisitions. This is the particular strength of the maximum-amplitude path method if compared with other wave-equation imaging or time-reverse imaging techniques for source localization.

3.5.2 SEG 3-D overthrust model

As a 3-D example, we consider the SEG 3-D overthrust model. Velocities are defined on a $400 \times 400 \times 187$ grid. The grid spacing is 40 m in the x- (crossline) and y- (inline) directions, and 20 m in z- (depth) direction. A source, denoted by a red star in Figure 3.12, is located at $x = 3.2$ km, $y = 3.2$ km, and $z = 2.4$ km. The source is an explosive source with a Ricker

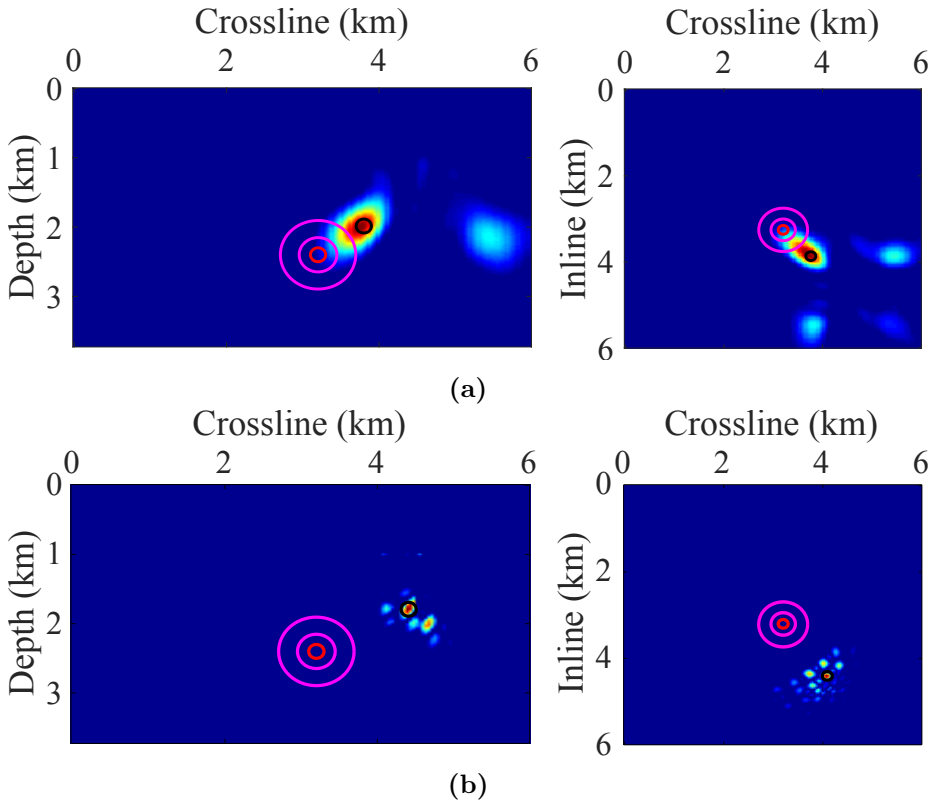


Figure 3.14: Results for (a) auto-correlation imaging and (b) GmRTM. The red and purple circles are centred at the true source location. The purple circles have radii of half and one dominant signal wavelength. The black circles indicate estimated source locations.

wavelet as time signal and a dominant frequency of 10 Hz. A time shift is applied to simulate a minimum phase signal. The time sampling is 1 ms.

Figure 3.12 a and 3.12 b shows the inline and crossline sections of the actual model at the source location, respectively. The corresponding smoothed version of the sections is shown in Figure 3.12 c and 3.12 d. Smoothing is performed with a $600 \times 600 \times 300\,\mathrm{m}^3$ 3D Gaussian window. Data are again computed for the true model and back projections are performed in the smooth model. In Figure 3.13, 36 receivers are evenly distributed at the surface between 4.16 and 7.36 km in the x- and y- directions with a spacing of 640 m. The surface projection of the source is denoted by the red star. This example reflects a sparse acquisition with limited aperture and therefore limited illumination in a complex 3-D model. The source location is laterally displaced with respect to the area covered by the seismic network.

Figure 3.14 displays the results of auto-correlation imaging and GmRTM. The localization results deviate from the actual source location by more than the dominant wavelength, that is, beyond the resolution limit of the source signal. For estimating the seismic event by the maximum-amplitude path method (Figure 3.13), the first receiver patch comprises receivers in the blue rectangular box, the second receiver patch contains receivers in the purple rectangular box, and the third receiver patch includes the receivers in the gray rectangular

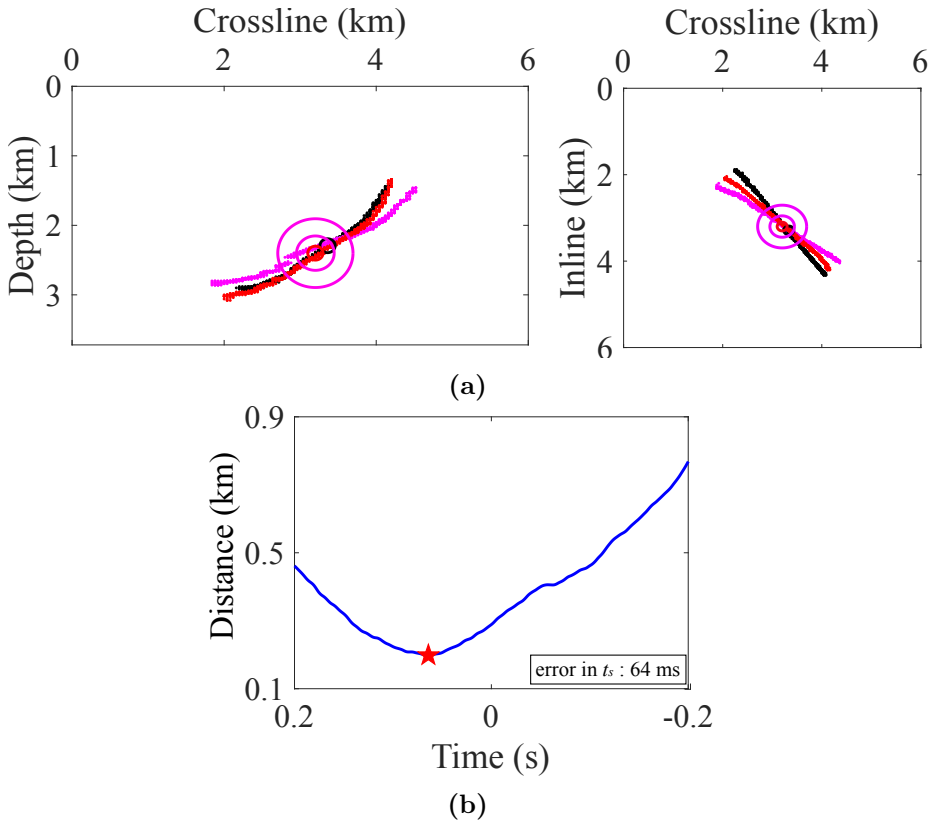


Figure 3.15: (a) Localization results for the maximum-amplitude path method. The black, purple, and red lines represent the maximum-amplitude path for the first, second, and third receiver patch, respectively. The red and purple circles are centred at the true source location. The purple circles have radii of half and one dominant signal wavelength. The black circles indicate estimated source locations. (b) The distance between maximum-amplitude paths against time, with the red star denoting a minimum, determines source excitation time.

box. Figure 3.15 shows the paths that start at 0.45 s and end at –0.3 s. They focus on a small region close to the actual source location. The location error is approximately 40% of the dominant signal wavelength, whereas the source-time errors are 64 ms. The 3D example shows that the maximum-amplitude path approach provides a good location and source excitation time for the sparse small-aperture acquisition with limited illumination because the source is located outside the seismic network.

3.6 Discussion

In a sparse and small-aperture acquisition, classic time-reverse imaging methods, such as auto-correlation imaging or GmRTM (among others), usually have issues because of insufficient focusing of the back-projected wavefield. The images display artifacts and no unique focus is obtained to identify the source location, which may lead to false positives.

To circumvent these limitations of time-reverse imaging methods, we propose a novel approach for simultaneously estimating source excitation time and location using maximum-amplitude paths. Because of the wavefield back-propagation step in constructing the path, the maximum-amplitude path approach belongs to the class of time-reverse imaging techniques for source localization.

The maximum-amplitude path is constructed by scanning for the maximum absolute amplitude of the back-propagated wavefields. The scan for maximum absolute amplitude is performed in a limited area aligned in propagation direction with a spatial extent of the dominant wavelength of the considered event for each time step. The maximum-amplitude path then comprises the maximum absolute amplitudes of the wavefronts for each considered time step. The methods exploit the continuity of the maximum amplitude in the back-projection process as soon as focusing of the injected wavefields is achieved. Moreover, the limitation of the analysis to a close proximity of the maximum avoids picks of false positives. These aspects basically distinguish the proposed approach from other techniques using back-projection methods. We choose maximum absolute amplitude as an imaging condition; however, other imaging conditions can be considered as well at the expense of computational efficiency.

In the presented examples, we have constructed three maximum-amplitude paths from three receiver patches where a patch denotes a collection of different traces out of the considered acquisition. The location where the three considered paths are closest to each other or cross indicates the source location. The corresponding time of this point provides the excitation time of the source.

Any point on the wavefront could be used to construct a path because, for a data consistent velocity model, the wavefront collapses at the source. Because the maximum is easiest to identify, we follow the propagation of the maximum amplitude. Rather than being described as rays in the classical sense, that is, rays obtained by ray tracing, the resulting maximum-amplitude paths connect the locations of the maximum amplitude of the wavefront for each considered time step. The wavefront collapses to a focal "point" at the source location, where "point" has to be physically interpreted in the sense of a band-limited seismic event. Because of this focusing, the maximum-amplitude paths either cross or are close to each other at the source location. The maximum-amplitude path method will work best for such point-type sources. It has not yet been tested on extended sources as observed for large earthquakes.

For quality control and improvement of results, more than three paths can be considered at the costs of additional computing time. In the construction of the maximum-amplitude path, it may be possible to skip some time steps by only considering snap shots which differ in time by, for example, half the prevailing period of the signal. This, however, is not yet adopted in our current work. Furthermore, it may lead to a decreased accuracy in location and source time determination. To avoid the latter, a refinement approach for the time steps between snapshots down to the sampling rate may be considered, if the distance between the paths decreases to, for example, half the prevailing wavelength.

Numerical examples based on the Marmousi-II and SEG 3-D overthrust models have shown that the maximum-amplitude path approach provides accurate locations and source times. In the considered examples, the determined source locations and source times deviate from the true locations and times by less than the prevailing wavelength and less than the prevailing period of the considered event, that is, they stay in the resolution limit of the

event. This observation holds even for very sparse acquisitions with only seven receivers in the receiver patch for 2D examples and 36 receivers in the receiver patch for the 3D examples. A denser receiver distribution in the patch leads to even better results. Because of the improved illumination, the locations and source times for a wide-aperture acquisition show only little errors. Because of the focusing during the back-propagation process, poor data quality has not heavily influenced the results. For very sparse acquisitions and noisy data, careful preprocessing should be carried out to optimize data quality.

Data enhancement techniques (e.g., Diekmann et al., 2019; Bauer et al., 2019b; Xie and Gajewski, 2017; Saad and Chen, 2020) also can be applied to improve passive seismic data quality. The data enhancement methods often require local coherence of the recorded events, which is not met for sparse data. It is one of the particular strengths of the maximum-amplitude path method that it is applicable to sparse and noisy data with limited-aperture acquisition.

As for any time-reverse imaging method, the maximum-amplitude path approach also requires a velocity model. We have used the true model for computing the synthetic data and a smoothed version of the true velocity model for the reverse modeling in the numerical examples. The resulting resolution of the smoothed velocity model turns out to be sufficient for localization and source time determinations within the resolution limit of the considered event. The requirements on the quality of the velocity model for passive seismic source imaging are less strict if compared to, for example, reverse time migration of reflection data. Particularly for the wide-aperture acquisition, locations and source times have been obtained with high accuracy. Because source times play a crucial role in passive seismic wavefront tomography (Diekmann et al., 2019; Yang and Gajewski, 2021a), the maximum-amplitude path approach may be helpful to better estimate source times and to improve passive seismic velocity model building by cascading both techniques.

The major driver in the computational effort of the maximum-amplitude path technique is the reverse time imaging step because the construction of the paths consumes only a small fraction of the total central processing unit (CPU) time. In this work, the reverse time imaging step has been applied for each receiver patch separately. Using phase encoding (Romero et al., 2000), it appears feasible to perform the reverse time imaging process for all receiver patches simultaneously. This would lead to computational costs similar to conventional time-reverse imaging methods for source location.

3.7 Conclusion

We have presented a methodology for seismic source localization using time-reverse modeling. By tracking the maximum amplitudes of the wavefront formed by the back-projected wavefields, we obtain the maximum-amplitude path. This path is constructed from three or more receiver patches comprising a selection of traces out of the recorded data. Because the back-projected wavefields focus on the source, the closest proximity or crossing point of these paths specifies the source location and the corresponding time is the source time. The methodology is capable of providing localization results within the resolution limit of the considered event. This conclusion holds even for small-aperture sparse acquisitions and noisy data, which provide major challenges to other current time-reversal imaging techniques.

3.8 Acknowledgments

We are grateful to the Applied Seismic Team at the University of Hamburg for continuous discussions and suggestions. We are particularly grateful to A. Bauer and C. Vanelle for proofreading and comments. Three anonymous reviewers provided constructive comments and suggestions that improved the manuscript. The partial financial support of the University of Hamburg and the sponsors of the Wave Inversion Technology (WIT) Consortium is highly appreciated. P.Yang also is supported through a stipend of the China Scholarship Council.

4 Integrated wavefront tomography and time reversal for source location, excitation time, and velocity model building

Abstract

The locations and excitation times of passive seismic sources are crucial for understanding Earth's physical processes. Most established localization methods require a reasonable velocity model, which is usually unknown in practice. In this work, we propose a data-driven hybrid workflow for simultaneously inverting the velocity model, source location, and source onset times using wavefront attributes. This workflow combines wavefront tomography and time-reversal methods. The latter are independent of source time and wavefront tomography is used for velocity inversion. The hybrid workflow has distinct advantages over the conventional workflow for joint inversion of velocities and source parameters using wavefront attributes. The proposed workflow comprises the following steps. First, a set of user-defined vertical gradient velocity models is designed. For these models time reversal is carried out to estimate the source excitation times for each model. These source times and gradient models are considered next for wavefront tomography. The second step uses an optimization procedure to find the optimal velocity model and the optimal source excitation time. In each iteration of the optimization, time reversal and wavefront tomography are sequentially performed, with time reversal producing the source time and location, and wavefront tomography providing the velocities. In a final step the source location is refined using the optimal velocity model and the Gaussian-weighted cross-correlation imaging condition. Numerical examples using the overthrust model illustrate the performance of the hybrid workflow. For a source in a model region with strong lateral heterogeneity and data acquired in a noisy environment or for a sparse source-receiver geometry, and sources distributed in a cluster, the hybrid workflow provides good velocity models, and fairly accurate source locations and excitation times. The errors of source locations and excitation times are smaller than the dominant wavelength and prevailing period, i.e., within the resolution limit of band-limited passive seismic signals.

4.1 Introduction

Precisely determining the excitation time and location of the passive seismic sources is important in both seismology and geoengineering (Shearer, 2019; Li et al., 2020). These sources can be triggered by natural causes like tectonic events or human activities such as fluid injection/extraction (Suckale, 2009; Mazzoldi et al., 2012; Maxwell and Urbancic, 2001; Warpinski, 2009; Eisner et al., 2009). In seismology, earthquakes originate from crustal fractures (Shearer, 2019) and accurately determining the onset time and location of earthquake sources offers fundamental information for seismological analyses like magnitude

calculations and seismic source mechanism inversion (Li et al., 2020). In geoengineering applications, like hydraulic fracturing and CO2 sequestration, induced seismicity is triggered during fluid injection. High-precision estimation of the excitation time and location of these events aids in assessing the potential impacts on groundwater and nearby subsurface structures (Anderson, 2017; Schultz et al., 2021), which helps in making timely responses, managing induced seismicity and optimizing industrial processes to reduce potential environmental hazards.

With advances in computing power and storage capacity, time reversal has developed as a common and powerful tool for passive seismic source localization that does not require picking seismic arrival times and phases. This type of method is robust to high noise levels and strong heterogeneity, and can accurately localize sources by performing numerical backpropagation of the recorded waveforms through a given velocity model (Fink et al., 2002; Gajewski and Tessmer, 2005; Larmat et al., 2008; Yang and Gajewski, 2022). Gajewski and Tessmer (2005) introduced the reverse modeling approach to characterize passive seismic events, which can estimate the excitation time and location of the events. Yang et al. (2023) proposed the Gaussian-weighted cross-correlation imaging condition for source localization, which provides higher-resolution and clearer source focusing compared to the auto-correlation imaging method presented by Artman et al. (2010). Since the observed waveforms are injected at the receiver locations, time reversal techniques are independent of source time. The accuracy of the source location and excitation time obtained from the time-reversal methods highly depends on the quality of the provided underlying velocity model which is often unknown in field monitoring projects.

To address the issue regarding the velocity model, waveform-based approaches for joint inversion of source location and velocity distributions have been developed. Image-domain wavefield tomography methods using the PS energy imaging function (Oren and Shragge, 2021) or the source-focusing function (Song and Alkhalifah, 2019) simultaneously invert for the source location and velocity model by minimizing the image residues calculated from the functions. Yet, these tomography methods require the use of high-quality source images in the inversion process, and therefore they need dense and well-distributed seismic data for source imaging. Another waveform-based approach to the joint inversion task is full waveform inversion (e.g., Wang and Alkhalifah, 2018; Sun et al., 2016). In this approach modeled waveforms are fitted to observed data. Full waveform inversion techniques yield fine-scale velocity variations, however, they are sensitive to the initial velocity models, prone to cycle skip, and require high-quality data. Moreover, these methods typically do not include inversion of the source excitation time.

To overcome the limitations of poor data quality and unknown source time, a powerful alternative presented by Diekmann et al. (2019) uses wavefront attributes (Hubral, 1983; Schwarz et al., 2016) to jointly invert source location, excitation time, and velocity model, which can handle noisy and/or sparsely sampled data (Diekmann et al., 2019; Yang and Gajewski, 2021a). It is a data-driven workflow that requires only prior information of the near-surface velocity. However, this workflow struggles to produce precise source excitation times, particularly in scenarios with strong lateral heterogeneity.

To address these issues of the conventional workflow presented by Diekmann et al. (2019) we combine the potential to determine the velocity model by wavefront attributes with the independence on the source times of time reversal methods. This hybrid workflow performs time reversal and wavefront tomography sequentially: time reversal (Gajewski

and Tessmer, 2005; Yang et al., 2023) determines the source parameters, while wavefront tomography (Duveneck, 2004; Bauer et al., 2019a) provides the velocity model.

In the following sections we first review the determination of wavefront attributes and its use in wavefront tomography which is followed by a brief description of time reversal imaging. The description of the hybrid approach is followed by numerical examples which lead to discussion and conclusions of the method.

4.2 Wavefront attributes

Wavefront attributes and first estimates of the source excitation time need to be determined before performing the joint inversion workflow. Estimating the source excitation time requires wavefront attributes estimated by the parabolic traveltime operator (Schwarz et al., 2016; Yang and Gajewski, 2021a). In the 2D case, it is expressed as:

$$t(x) \approx t(x_0) + \frac{\sin(\alpha)}{v_0} \cdot \Delta x + \frac{1}{2} \cdot \frac{\cos(\alpha)^2}{v_0 \cdot R_{xx}} \cdot \Delta x^2 \quad , \tag{4.1}$$

where the arrival time t of the passive seismic source signal recorded at the receiver location x is approximated by a Taylor series expansion (Castle, 1994) in the vicinity of a given receiver location x_0. The displacement $\Delta x = x - x_0$, represents the distance from the receiver position x_0, and v_0 indicates the locally constant near-surface velocity around x_0, which is assumed to be known in this work. The wavefront attribute α represents the emergence angle of the locally observed wavefront, while R_{xx} represents the radii of wavefront curvature (Hubral, 1983). The parabolic traveltime operator has also been formulated for the 3D cases (Schwarz et al., 2016; Diekmann et al., 2019). Because the expression proposed by Diekmann et al. (2019) for source excitation time estimation is derived from the parabolic traveltime operator (equation 4.1), we use the above-mentioned parabolic traveltime operator instead of the hyperbolic traveltime operator (Ursin, 1982; Jäger et al., 2001) for wavefront attribute estimation in the numerical examples below.

Wavefront attribute determination can be considered as an optimization problem, aiming to maximize the coherent energy for each target sample, in which the shape of the event such as the wavefront curve passed through the sample of the zero-off section is approximated by a parabolic traveltime operator. Global optimization algorithms, such as particle swarm optimization (Kennedy and Eberhart, 1995) and differential evolution (Storn and Price, 1997), can be used to estimate the wavefront-related quantities accurately (Walda and Gajewski, 2017; Xie and Gajewski, 2018). In the present work, we use particle swarm optimization, to simultaneously search for the optimal set of wavefront attributes. The set of wavefront attributes (α, R_{xx}) is extracted from the data in an automatic, data-driven fashion by searching the largest local semblance coefficient (Taner and Koehler, 1969) for a subset of traces located in a predefined aperture around the central point (t_0, x_0) (Yang and Gajewski, 2021a). The process of estimating wavefront attributes involves a stacking procedure, contributing to the identification of weak events in controlled-source seismology (Mayne, 1962; Taner and Koehler, 1969).

Afterward, the wavefront attributes obtained by equation 4.1, are used to enhance the S/N and data resolution by creating a stacked section by summing all the samples along the traveltime trajectories within a certain range around x_0. It can suppress the noise

and enable data interpolation and regularization (Baykulov and Gajewski, 2009; Xie and Gajewski, 2017). In addition, using the obtained wavefront attributes, the source excitation time t_s can be approximated by a hyperbolic expression (Diekmann et al., 2019), which is derived from the parabolic expression (equation 4.1):

$$t(x) \approx \left[(t(x_0) - t_s + \frac{\sin(\alpha)}{v_0} \cdot \Delta x)^2 + (t(x_0) - t_s) \cdot \frac{\cos^2(\alpha)}{v_0 \cdot R_{xx}} \cdot \Delta x^2\right]^{\frac{1}{2}} + t_s \quad , \tag{4.2}$$

The 3D expression of the hyperbolic traveltime operator is also formulated by Diekmann et al. (2019). The source excitation time estimated in this way is used as an initial guess of the source excitation time in the conventional workflow presented by Diekmann et al. (2019).

4.3 Wavefront tomography and time-reversal methods

The proposed hybrid workflow used for joint inversion of the source location, excitation time, and velocity model, consists of both wavefront tomography and the time-reversal methods, while the conventional workflow used for the joint inversion task primarily relies on wavefront tomography.

4.3.1 Wavefront tomography

Wavefront tomography is a robust, data-driven approach for inverting smooth velocity models down to the scale of the prevailing wavelength of the considered data. The picking is performed in the stacked section which has a high S/N ratio and is easier than picking in pre-stack data (Duveneck, 2004). The input data vector of wavefront tomography contains four components in 2D and eight components in 3D (Diekmann et al., 2019). In the 2D cases, it comprises the actual traveltime τ, the lateral coordinates x, the horizontal slowness p_x, and the curvature-dependent parameter M_{xx}:

$$\boldsymbol{d}_i = (x, \tau, p_x, M_{xx})_i \quad , \quad i = 1, ..., n_{data} \quad , \tag{4.3}$$

where $\tau = t - t_s$ is the time of the considered sample minus the corresponding source excitation time, and n_{data} represents the total number of picked data points. The horizontal slowness p_x and the wavefront curvature-related quantity M_{xx}, are calculated using the wavefront attributes (α, R_{xx}) (Schwarz et al., 2016):

$$p_x = \frac{\sin(\alpha)}{v_0} \quad , \quad M_{xx} = \frac{\cos^2(\alpha)}{v_0 \cdot R_{xx}} \quad , \tag{4.4}$$

The input data $\boldsymbol{d}$ contains all data points $\boldsymbol{d}_i$, which are automatically extracted within a selected window as displayed in Figure 4.4. The length of the selected window is chosen to be smaller than the width of the event in the coherence section. The coherence section and the stacked seismic section are obtained during the search of wavefront attributes. Since our study only investigates individual events, tagging of events is not required as used by Bauer et al. (2019b).

The model vector $\boldsymbol{m}$ of wavefront tomography, comprises the velocity model and the starting positions and directions of the upward propagating rays, whose expressions are

given by Diekmann et al. (2019) in two and three dimensions. The $\boldsymbol{m}$ in 2D is represented by

$$\begin{array}{c} v_{jk} \quad , \quad j = 1, ..., n_x \quad , \quad k = 1, ..., n_z \\ (x_0, z_0, e_x)_i \quad , \quad i = 1, ..., n_{data} \quad , \end{array} \tag{4.5}$$

where v_{jk} denotes the velocity model defined in terms of B-spline coefficients (De Boor, 1978), n_x and n_z mark the number of knots used in x- and z-direction, respectively. The location $(x_0, z_0)_i$ represents the starting locations for the upwards propagating rays corresponding to the i-th data subset $\boldsymbol{d}_i$ while the respective ray direction is determined by the horizontal components $(e_x)_i$. Under the initial velocity model, we perform downward kinematic ray tracing using the picked data to complete the initial set of model vector $\boldsymbol{m}$, including the initial ray starting locations and directions. Then using the starting locations and directions of the rays, we conduct upward dynamic ray tracing to derive the modeled data $\boldsymbol{d}_i^{mo}$ for each picked data point. This data vector contains modeled values for the surface location, traveltime, and wavefront attributes of each data point in the registration surface:

$$\boldsymbol{d}_i^{mo}(\boldsymbol{m}) = (x(\boldsymbol{m}),\ \tau(\boldsymbol{m}),\ p_x(\boldsymbol{m}),\ M_{xx}(\boldsymbol{m}))_i^{mo} \quad , \tag{4.6}$$

The inversion is formulated in a damped weighted least-squares sense (Tarantola, 2005). The misfits between the modeled and the picked observed quantities are calculated using the following objective function:

$$\boldsymbol{S}(\boldsymbol{m}) = \frac{1}{2}\left\| (\boldsymbol{d} - \boldsymbol{d}^{mo}(\boldsymbol{m})) \cdot \boldsymbol{W} \right\|_2^2 + \Lambda(v_{jk}) \quad , \tag{4.7}$$

Here, the modeled dataset $\boldsymbol{d}^{mo}$ comprises all $\boldsymbol{d}_i^{mo}$, $\boldsymbol{W}$ denotes the vector scaling the physical quantities, and $\Lambda(v_{jk})$ represents the regularization term that favors the smooth velocity model. In each iteration of the inversion, dynamic ray tracing is performed with an updated model vector. The optimal velocity model is determined when the data misfits derived by equation 4.7 are minimized.

This velocity inversion method using wavefront attributes, known as wavefront tomography, originally introduced by Duveneck (2004), has been further developed with an enforced focusing strategy presented by Bauer et al. (2019a). This enhanced wavefront tomography method, termed wavefront tomography with enforced focusing, uses the mean event locations of all equally picked data points to enforce focusing and initializing the upward ray tracing from these positions. It reduces the dependence of wavefront tomography on second-order wavefront attributes, thereby improving the stability of the inversion in complex situations with lateral heterogeneity. We introduce this advanced inversion approach into the second step of the proposed workflow. However, in passive seismic applications, this advanced inversion approach may fail when the excitation time and location of the considered event are less accurately determined. Therefore, the first step of the proposed workflow is used to provide a good initial velocity model for the enforced focusing step.

4.3.2 Time-reversal localization

The Gaussian-weighted cross-correlation imaging condition, proposed by Yang et al. (2023), is particularly effective for handling noisy data and data acquired from sparse source-receiver geometry. Each time step includes dividing the total back-propagated wavefield

into Gaussian-shaped contributions that are classified by their maximum absolute amplitudes. The spatial extent of these contributions is linked to the dominant wavelength of the source signal. Multiplying these wavefield contributions, which are weighted by Gaussian functions, yields the image for the considered time step. The final image is attained by summing the images of all time steps, and the location corresponding to the maximum of the final image is the source location. The final image I is calculated by:

$$I(\boldsymbol{x}) = \sum_{t=1}^{T} \left| \prod_{k} \overline{u}_k(\boldsymbol{x}, t) \right| \quad , \quad 1 \leq k \leq N_w \quad , \tag{4.8}$$

where $\overline{u}_k(\boldsymbol{x}, t)$ represents the k-th Gaussian-weighted wavefield at the considered time step, and T denotes the predefined maximum time for the backpropagation. The value of the parameter N_w, which indicates the count of the number of Gaussian-weighted wavefields, is consistently set to 2 in all numerical tests unless otherwise stated. The other tuning parameters used in this imaging condition follow the conventional set (Yang et al., 2023).

The more accurate the velocity model, the more energy is focussed during the backpropagation. We use this imaging condition as the basis for developing a new objective function in the optimization of the second step of the hybrid workflow, discussed in the following section. The final source location in both, hybrid and conventional workflows, are determined using the Gaussian-weighted cross-correlation imaging condition. However, the strategy for source excitation time estimation based on the Gaussian-weighted cross-correlation imaging condition is not applied in the present work.

We use reverse modeling to estimate the excitation time of the source in the hybrid workflow because reverse modeling is a well-established and easily implemented technique for estimating source excitation time. Reverse modeling introduced by Gajewski and Tessmer (2005), is a time-reversal localization method. Assuming a reasonably accurate velocity model, the method uses the observed data that is reversed in time to simulate a backpropagation and focuses the energy of the back-propagated wavefield toward the hypocenter. The time of the maximum absolute amplitude of the back-propagated wavefield over the entire backpropagation time is the source excitation time:

$$\max \left\{ \left| u(\boldsymbol{x}, t) \right| \right\} \Rightarrow t_s \quad , \quad 1 \leq t \leq T \quad , \tag{4.9}$$

Here, $u(\boldsymbol{x}, t)$ represents the back-propagated wavefield at the considered time step, and t_s denotes the estimated source excitation time.

4.4 Hybrid workflow

The conventional workflow (Figure 4.1 a) proposed by Diekmann et al. (2019) suffers from poor source time estimation. In addition, the workflow faces challenges in accurately localizing sources in scenarios where models exhibit strong lateral heterogeneity, data obtained from noisy environments, or data obtained from sparse networks, which are often met in practice. Thus, we propose a hybrid workflow to improve the reliability of localization results in these scenarios.

This hybrid workflow provides a technique to evaluate the quality of initial models, potentially improving the accuracy of inversion results in wavefront tomography. Moreover,

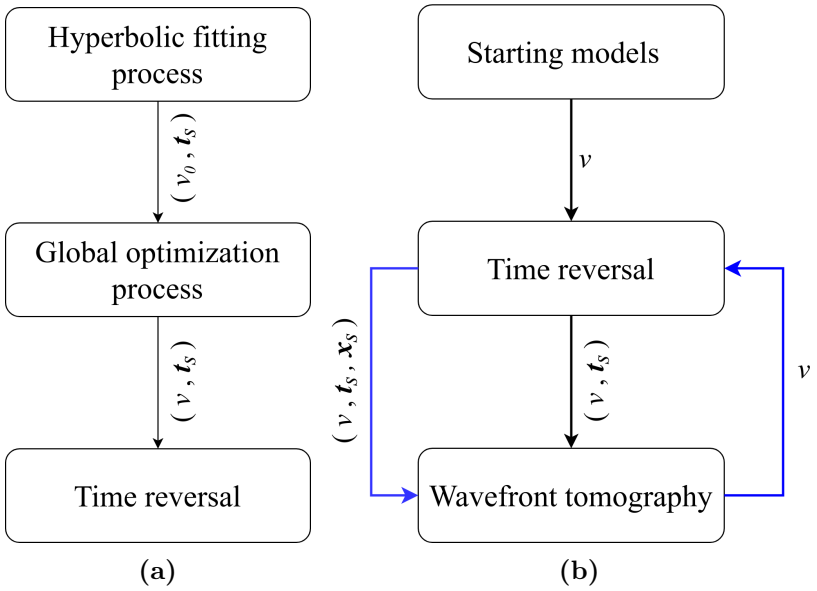


Figure 4.1: (a) The flowchart of the conventional workflow (for details, see text) (b) The flowchart of the hybrid workflow, sketching main procedures and their respective outcomes. The black and blue arrows mark the first and second steps of the hybrid workflow, respectively (for details, see text). The first step starts with a set of user-defined vertical gradient models, each using time reversal to determine source excitation times $\boldsymbol{t}_s$. Using these source times, along with the corresponding user-defined gradient models, wavefront tomography is performed to produce the initial model for the second step. In the second step, an optimization process unfolds: time reversal and wavefront tomography are sequentially performed, with time reversal estimating source excitation times $\boldsymbol{t}_s$ and locations $\boldsymbol{x}_s$, and wavefront tomography providing the velocity model v.

by employing time-reversal methods (i.e., reverse modeling (Gajewski and Tessmer, 2005) and Gaussian-weighted imaging method (Yang et al., 2023)), the source localization within the workflow becomes independent of the source excitation time, thereby improving overall accuracy. The hybrid workflow, displayed in Figure 4.1 b, consists of two main steps:
(1) The first step involves selecting a suitable velocity model for source excitation time estimation and velocity inversion. We begin by creating a set of vertical gradient velocity models using the following equation:

$$\begin{aligned} v_{l_{jk}} = v_0 + w_l \cdot (Z_k - Z_1) \quad , \quad 0 \leq w_l \leq \frac{v_m - v_0}{N_v \cdot (Z_{n_z} - Z_1)} \quad , \\ l = 1, ..., N_v \quad , \quad j = 1, ..., n_x \quad , \quad k = 1, ..., n_z \end{aligned} \tag{4.10}$$

The velocity models are defined using B-splines, with n_x and n_z holding the same definitions as in equation 4.5. The parameter Z_k denotes the depth of the k-th nodes in the z-direction, and v_0 stands for the surface velocity that is assumed to be known. The v_m represents the user-defined maximum velocity value, which is determined using a few interactive tests, and it can be initialized from a value greater than v_0, such as 8.0 km/s, and gradually increases if the time-reversal image lacks unique focusing or if the objective function value remains large. The N_v denotes the total number of models. Each model is created using a vertical

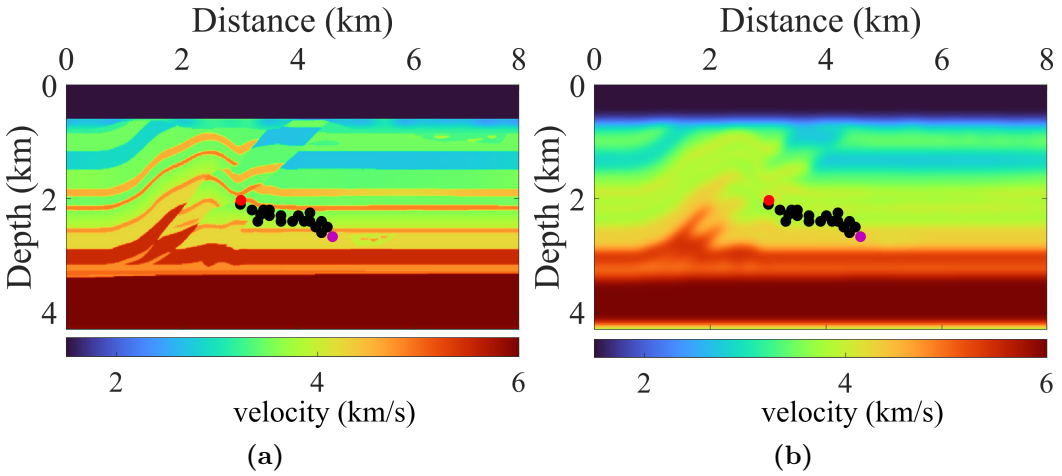


Figure 4.2: The overthrust velocity model. (a) True velocity model. (b) Smoothed version of the true velocity model after applying a Gaussian smoothing window of 300×300 m^2. The dots indicate the sources, where the red and purple dots indicate the fault-side and layer-in sources located at the point (x = 3.01 km, z = 2.03 km) and point (x = 4.59 km, z = 2.67 km), respectively.

gradient w_l. Reverse modeling is then applied to each gradient model to estimate the source excitation time. By employing each gradient model as an initial model and with the fixed corresponding estimated source excitation time, we conduct wavefront tomography (Duveneck, 2004) to obtain the respective inversion velocity model. The inversion model displaying the minimum objective function value is selected as the initial model for the second step.

(2) The second step optimizes the velocity model and source excitation time. This optimization aims to maximize the velocity-dependent objective function J_G, the expression of which reads:

$$J_G(v) = \sum_{\boldsymbol{x}} \frac{I(\boldsymbol{x}, v)}{I_m} \quad , \tag{4.11}$$

where $I(\boldsymbol{x}, v)$ indicates the Gaussian-weighted cross-correlation image obtained through equation 4.8 using the velocity model v, and I_m denotes the maximum value of this image. Each iteration of the optimization process involves the following steps: reverse modeling and Gaussian-weighted time-reversal imaging are first conducted to determine the source excitation time and source location while calculating the objective function J_G (equation 4.11). Subsequently, by fixing the source excitation time and using the source location as the initial ray starting location, we perform wavefront tomography with enforced focusing (Bauer et al., 2019a).

4.5 Numerical examples

We test the proposed workflow using the overthrust model (Aminzadeh et al., 1997). The overthrust models, as displayed in Figure 4.2, are defined on an 800 × 432 grid with a spacing of 10 m in both the x- and z- directions. The actual model in Figure 4.2 a is used to generate synthetic data using the density-constant two-way first-order velocity-stress

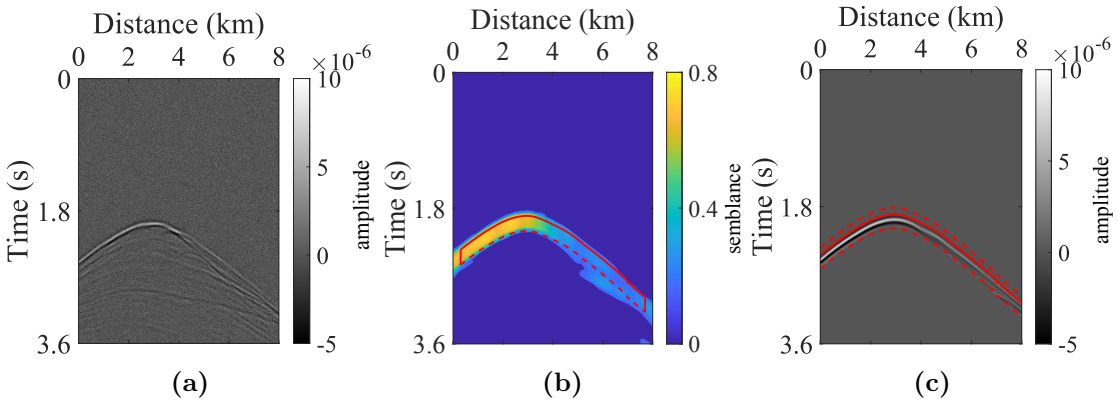


Figure 4.3: Seismic data and semblance for the fault-side source excited at $x = 3.01$ km, $z = 2.03$ km. (a) Seismic section with S/N of 2 generated by the fault-side source. (b) Maximum semblance obtained by the parabolic traveltime operator searched for the seismic data of (a). (c) Stacked section with S/N enhanced based on the traveltime operator determined by the maximum-semblance-determined wavefront attributes, α and R_{xx}. The red solid curve indicates the first pick sample on each trace. The red dashed curve indicates the selected window around the red solid curve. Samples within the window are selected.

acoustic wave equation (e.g., Aki and Richards, 2002; Gajewski and Tessmer, 2005), while the smoothed version of the overthrust model in Figure 4.2 b is employed to simulate back propagation using the same wave equation. Smoothing is carried out using a 2D Gaussian window of dimensions $300 \times 300\,\mathrm{m}^2$. We use explosive sources with a Ricker source-time function with a dominant frequency of 10 Hz. A time shift is applied to the waveform signal to simulate a minimum phase source wavelet. The prevailing signal period is 0.1 s in all examples.

Reverse modeling using the same acoustic wave equation with a scheme for up-going and down-going wavefield separation using the Hilbert transform (Zheng et al., 2018; Yang and Gajewski, 2022) is conducted to obtain the down-going wavefield. The down-going wavefield serves as the back-propagated wavefield in all examples. In addition, we generate seismic data of improved quality by stacking with the parabolic operator (equation 4.1) for reverse modeling in all tests.

For all numerical examples, we chose a maximum velocity v_m of 12000 m/s and a total number of models $N_v = 20$ to generate a set of gradient models for the first step of the proposed workflow to obtain first estimates of source time and velocities. All initial constant models for the conventional workflow are built with the provided near-surface velocity of 1.5 km/s. Moreover, all the time errors mentioned in the following examples are calculated relative to the actual source time, where negative numbers indicate that the estimated source time is prior to the real excitation time and positive numbers denote the time after the real excitation time.

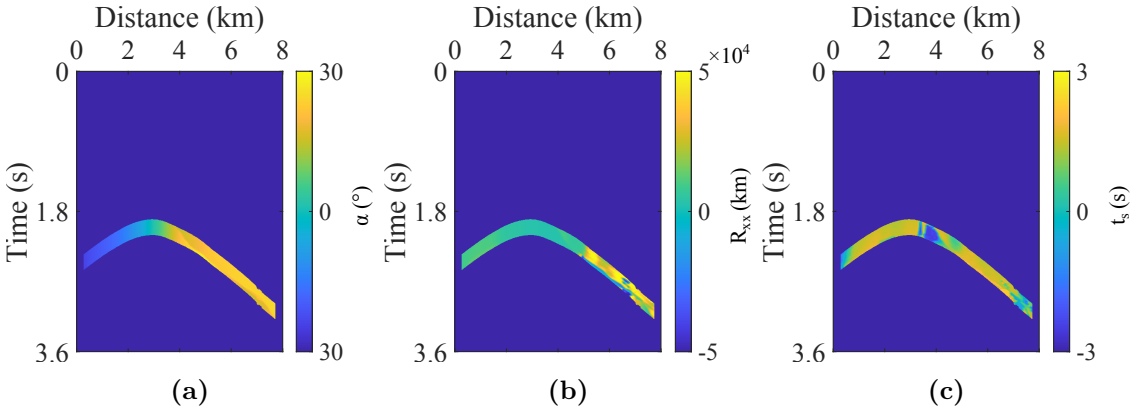


Figure 4.4: Wavefront attributes determined by maximum semblance using the parabolic traveltime operator. (a) Emergence angle α. (b) Radii of curvature R_{xx}. (c) Sample-wise source excitation times t_s. The samples beyond the selected window are shown in blue.

4.5.1 Fault-side source

To analyze the hybrid workflow for joint source localization and velocity inversion in the model region of strong lateral heterogeneity, we consider a fault-side source located at $x = 3.01$ km, $z = 2.03$ km near a geological fault. This source, marked by a red dot in Figure 4.2, is triggered at the time of 1.144 s. In this example, a densely uniformly sampled network with a receiver distance of 10 m is used to record the seismic event. Random noise with an S/N of 2 was added to the data. Figure 4.3 a displays the resulting data.

Wavefront attribute estimation

In the fault-side source example, we use the parabolic traveltime operator (equation 4.1) to determine the wavefront attributes from the recorded seismic data using a coherence analysis with an optimization aperture of 800 m. Afterward, the estimated wavefront attributes are used to enhance the S/N of the recorded events where the stacking aperture is set to 1000 m. The aperture for coherent analysis and stacking in the following examples is set to be the same as in this example, unless otherwise noted.

Figure 4.3 b shows the coherence section of the densely sampled data generated by the fault-side source signal, with specific samples marked by a continuous red solid curve. This curve aids in setting selection windows for wavefront attributes and stack sections. The wavefront attribute α and R_{xx} in the selected windows are depicted in Figure 4.4 a and 4.4 b, respectively. The stacked data is presented as an image plot in Figure 4.3 c. Using these estimated wavefront attributes, a hyperbolic best fit is sought for each sample to estimate source time t_s, as seen in Figure 4.4 c. Source time is a global property of the event and should be constant over the whole aperture. We consider the median value t_s of all estimated source excitation times. The conventional workflow uses this estimate as the initial guess of the source excitation time, which deviates from the true source excitation time by about 0.354 s, i.e., roughly three and a half prevailing signal periods.

Using the enhanced data obtained by stacking, the Gaussian-weighted cross-correlation imaging condition is used to determine the source location using the true and smoothed

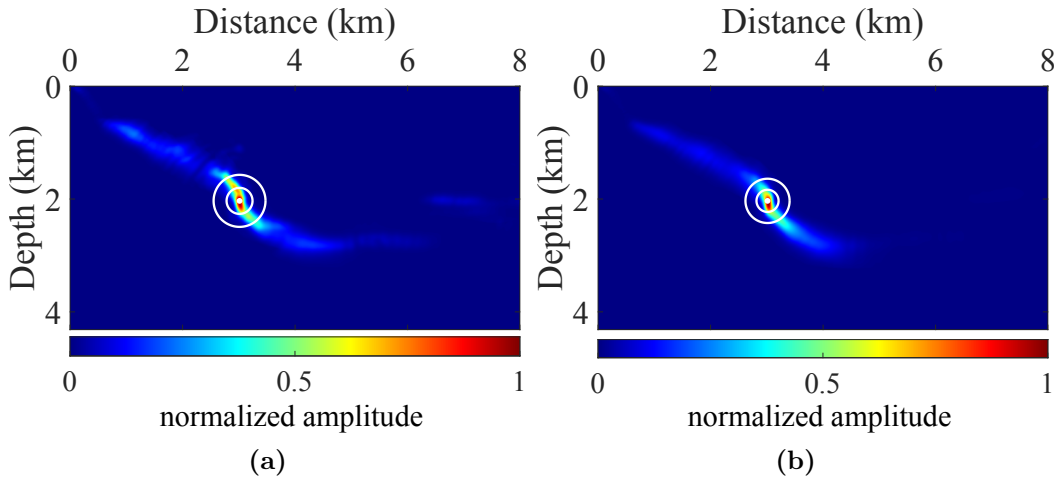


Figure 4.5: Source location results for the fault-side source. (a) Images of the source located at x = 3.01 km and z = 2.03 km near the fault, obtained by the Gaussian-weighted cross-correlation imaging condition using the true velocity model. (b) Same as (a) but using a smoothed model. The white dots indicate the true source locations. The two white circles, centered at the actual source location, have radii of half and one dominant signal wavelength, respectively.

overthrust models for reverse modeling. The respective results, as displayed in Figure 4.5 a and 4.5 b, demonstrate estimates with the location errors of only 10 m, and 30 m for the true and smoothed models, respectively, which are notably smaller than the dominant signal wavelength (392 m), along with negligible errors of only about -0.009 s in the source time estimate. Because the smoothed velocity model allows an accurate determination of source location and excitation time, in the following tests we evaluate the inversion velocity errors by calculating the relative velocity error w.r.t. the smoothed model, considering the illuminated part of the model. The location errors are evaluated using the dominant wavelength of the source signal calculated in the smoothed model.

Conventional workflow

We first perform the conventional workflow to localize the fault-side source using an initial constant model and an initial gradient model, respectively. The initial constant model is established with the near-surface velocity of 1.5 km/s, while the initial gradient model is generated using equation 4.10 with a gradient w_l of 1.5. Figure 4.6 a and 4.6 b illustrate the ray starting locations obtained by the initial constant model and the initial gradient model, respectively. During the velocity inversion process, we focused only on the illumination part of the model where rays cover, as displayed in Figure 4.6 and Figure 4.7. The illuminated portion of the true and smoothed model are illustrated in Figure 4.6 c and 4.6 d, respectively. Figure 4.7 a and 4.7 b display the outcomes of the conventional workflow using the initial constant model and the initial gradient model. The estimated source locations deviate from the true source position by 471 m, and 36 m, respectively, with the source time errors of -0.976 s and -0.201 s. This example demonstrates the superiority of the initial gradient model over the initial constant model in the conventional workflow, which significantly

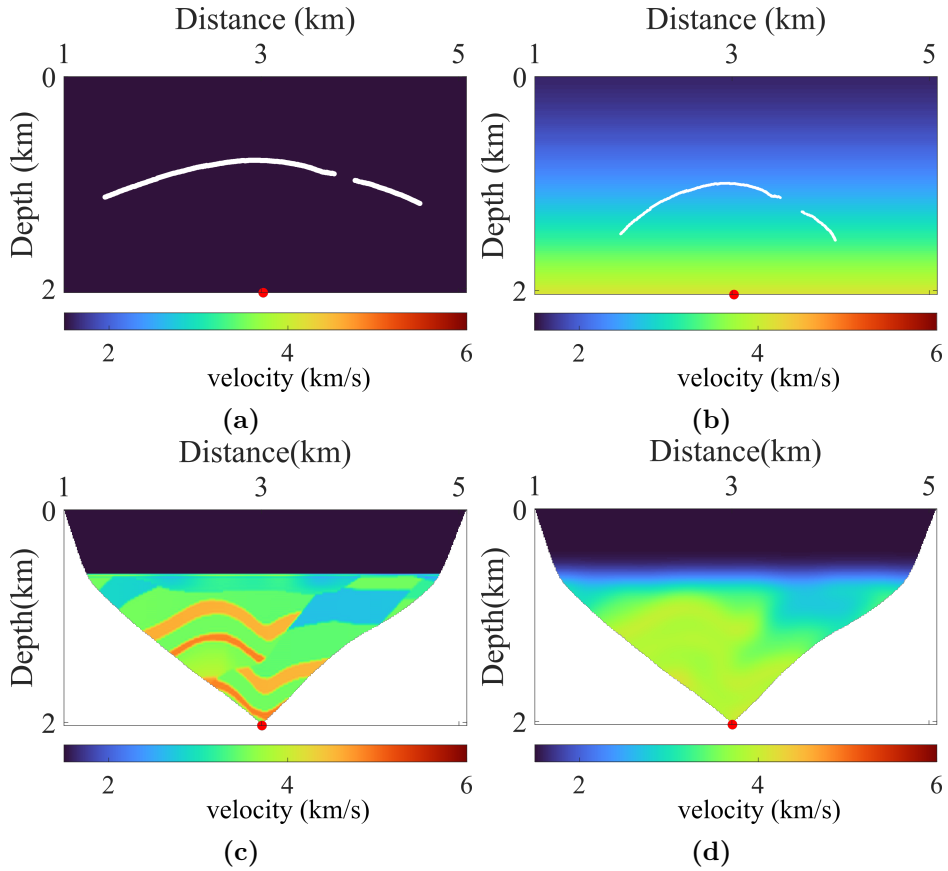


Figure 4.6: Initial constant model (a) and initial vertical-gradient model (b) of the conventional workflow for the fault-side source. The illuminated parts of the true model (c) and smoothed model (d) are shown. The white dots in (a) and (b), indicate the ray starting locations. The red dots denote the true source positions.

improves localization accuracy. When using the initial constant model, the joint inversion of the conventional workflow fails, resulting in localization errors far beyond the resolution limit of the source signal. However, finding a reasonable initial gradient model is a non-trivial task. Furthermore, even with a well-chosen initial gradient model, the conventional workflow provides a less accurate estimate of the source excitation time.

Hybrid workflow

We then apply the hybrid workflow to the fault-side source, exhibiting the results at each step of the workflow to evaluate its effectiveness. The first step involves selecting a suitable initial gradient model, generated using equation 4.10, for wavefront tomography. Figure 4.8 demonstrates the cost function values obtained from wavefront tomography for each defined initial gradient model, with the minimum value indicating the final velocity model of the first step that is used as the initial model for the second step.

Figure 4.9 a and 4.9 b illustrate the final results of the first and second steps of the hybrid

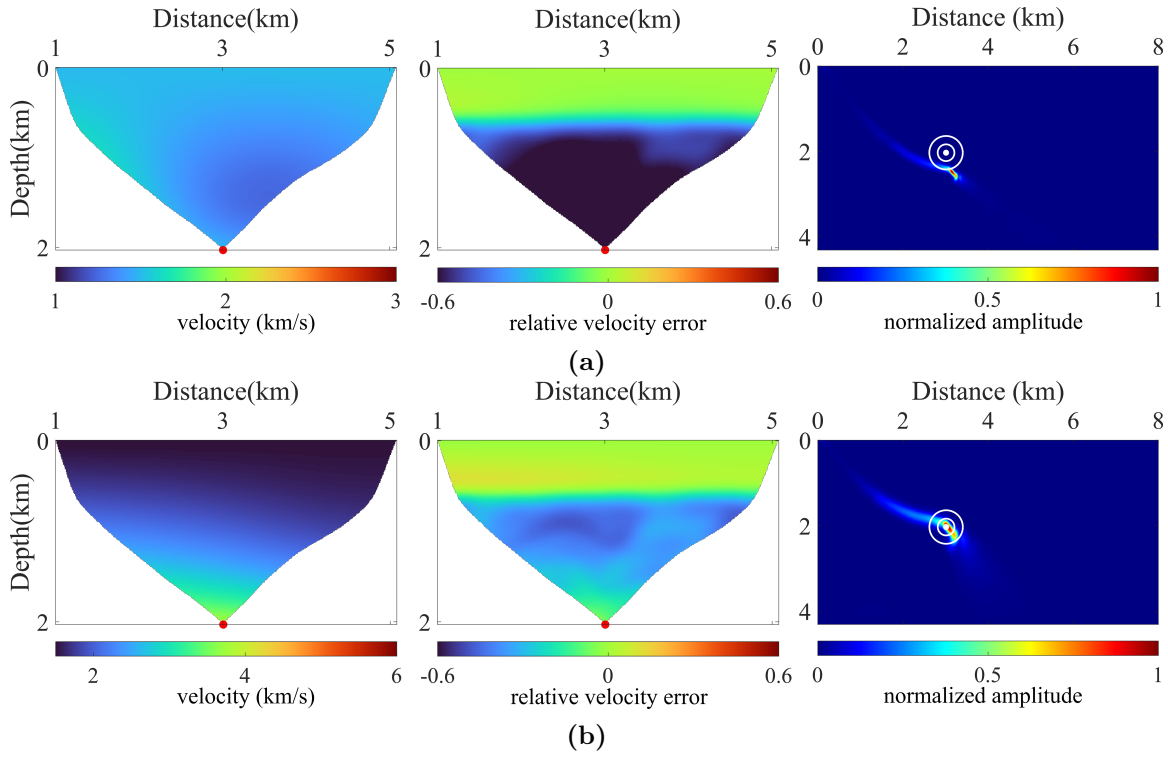


Figure 4.7: Comparison of results obtained for the conventional workflow using the initial constant model (a), and the initial gradient model (b), for the fault-side source. The left panel exhibits the final inverted velocity models, the middle panel displays the relative velocity error of the inverted velocity models to the smoothed velocity models, and the right panel represents the Gaussian-weighted cross-correlation source location images. The red dots in the left- and middle-panel images, and the white dots in the right-panel images, indicate the true source locations. The white circles, centered at the actual source location, have radii of half and one dominant signal wavelength. Only the illuminated part of the model is shown.

workflow, respectively. The results of the first step are quite similar to those obtained by the conventional workflow using the initial gradient model above. The source location error remains at about 36 m, smaller than half the dominant signal wavelength, and the source time error is at -0.207 s, about two prevailing signal periods. Subsequent refinement is evident in the second step of the hybrid workflow, which particularly improves the quality of the arc-like structures of the inversion model while reducing the source excitation time error to -0.089 s, smaller than the dominant signal period. Besides, the source location error is also smaller than half of the dominant signal wavelength, that is, the errors of the source location and excitation time provided by our hybrid workflow are within the resolution limit of the source signal. It reflects that the hybrid workflow is more advantageous in dealing with cases where the source is excited in the model region of strong lateral heterogeneity as demonstrated in this section in comparison to the conventional workflow.

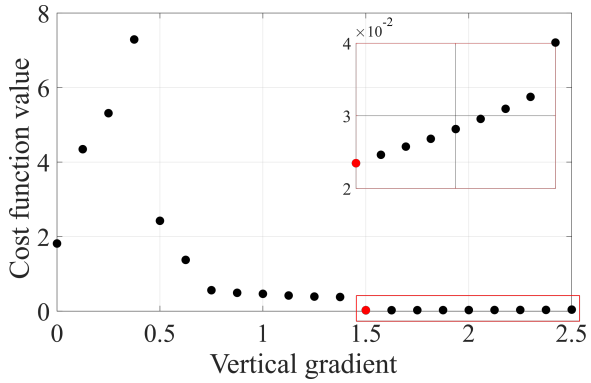


Figure 4.8: The cost function value shown for the first step of the hybrid workflow. The top-right red panel shows a close-up view of the bottom-right red panel. An increase in cost occurs after the red dot.

4.5.2 Layer-in source

To evaluate the effect in terms of random noise and sparse receiver acquisition for the joint inversion of the conventional and hybrid workflows, we consider a layer-in source located at $x = 4.59$ km, $z = 2.67$ km, denoted by the purple dot in Figure 4.2. This layer-in source is excited at the time of 1.099 s in the model region of weak lateral heterogeneity.

Dense data

In the first example of the layer-in source, we use the same dense seismic receiver network as in the previous examples of the fault-side source to record the seismic event, obtaining dense data. Random noise with an S/N of 2 was added to the data. The respective record is shown in the top panel of Figure 4.10 a. The source excitation time estimation of the conventional workflow, wavefront attribute determination, and data enhancement are performed the same as in the previous examples. The enhanced data is displayed in the bottom panel of Figure 4.10 a. The error of the source excitation time obtained from the best hyperbolic fit used as the initial estimate of the source time in the conventional workflow is about 0.422 s. The illuminated part of the model displayed in Figure 4.11 a for this dense data example is not as complex as the fault-side source example. Thus, both workflows yield fairly good source locations and resolve reasonable velocity models. Despite that, the hybrid workflow outperforms the conventional workflow, as displayed in Figure 4.12. The hybrid workflow improves the quality of velocities of the structures around the faults. Besides, the error of the source location from the hybrid workflow is smaller than half the dominant wavelength and the source time error is approximately -0.055 s, i.e., the errors are within the resolution limit of the source signal. However, the final source excitation time from the conventional workflow is still of poor quality with an error of about -0.201 s, though the corresponding source location deviates from the true source position by only 268 m, about 60 % of the dominant signal wavelength. This example illustrates that the hybrid workflow can still yield better inversion results than the conventional workflow for source localization within weak lateral heterogeneity.

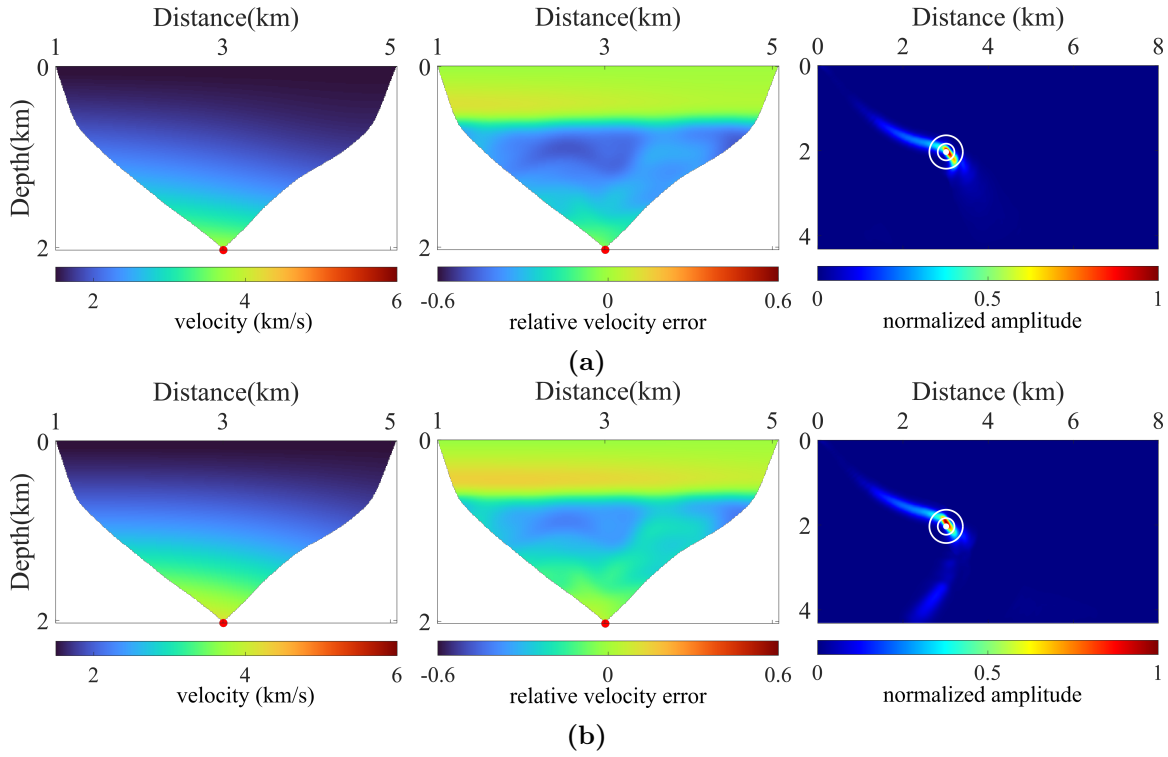


Figure 4.9: Comparison of results obtained from the first step (a) and the second step (b) of the hybrid workflow for the fault-side source. The left panel exhibits the final inverted velocity models, the middle panel displays the relative velocity error of the inverted velocity models to the smoothed velocity models, and the right panel represents the Gaussian-weighted cross-correlation source location images. The red dots in the left- and middle-panel images, and the white dots in the right-panel images, indicate the true source locations. The white circles, centered at the actual source location, have radii of half and one dominant signal wavelength. Only the illuminated part of the model is shown.

Noisy data

As a next step, we investigate the influence of strong random noise on the performance of the joint inversion of the workflows. The same seismic source setting and seismic receiver acquisition as for the dense data example is used to obtain the synthetic data. Then we added random noise to the synthetic data to simulate noisy data with an S/N of 0.2 (Figure 4.10 b), where the event is difficult to identify. To improve the S/N of the data, we stacked the data using the parabolic traveltime operator (equation 4.1). The stacked section is shown in the bottom panel of Figure 4.10 b, where the event is clearly observed. This enhanced data is used for reverse modeling and time-reversal imaging in both conventional and hybrid workflows. In this example of noisy data, the conventional workflow initiates with the source excitation time estimated from the best hyperbolic fit, displaying an error of 0.313 s. As displayed in Figure 4.11, compared to the dense data example, the part of the model where the ray coverage becomes smaller. As seen from Figure 4.13, the hybrid

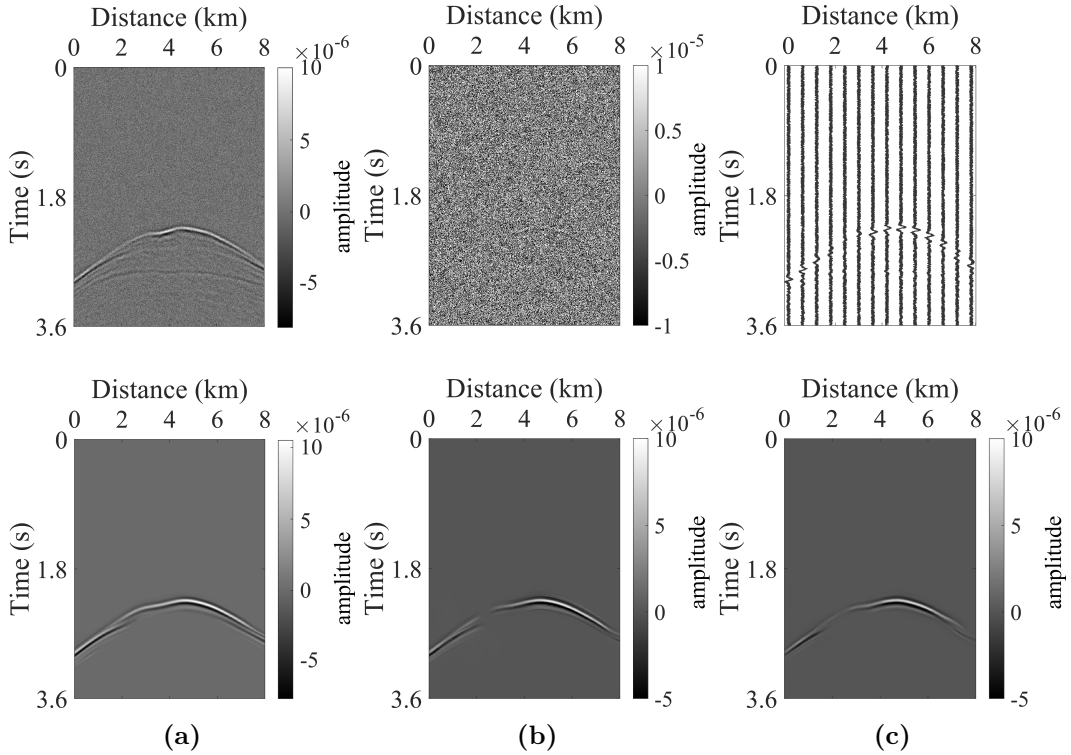


Figure 4.10: Data for the layer-in source excited at $x = 4.59$ km, $z = 2.67$ km. (a) Dense data from a dense receiver acquisition with an S/N of 2 (top) and its enhanced stacked data (bottom). (b) Noisy data from a dense receiver acquisition with an S/N of 0.2 (top) and its enhanced stacked data (bottom). (c) Sparse data from a sparse acquisition with an S/N of 2 (top) and its enhanced stacked data (bottom).

workflow maintains its superior performance over the conventional approach, consistently delivering fairly accurate results for the source location, excitation time, and velocity inversion. The hybrid workflow estimates the source excitation time with an error of -0.062 s (smaller than one prevailing signal period) and the source position with a deviation of 93 m (smaller than half the dominant signal wavelength), while the conventional workflow shows errors of -0.212 s (greater than a prevailing period) and a deviation of 335 m (about 75 % of the dominant wavelength) for the source excitation time and position, respectively. This noisy data example shows that the hybrid workflow performs well on poor-quality data with an S/N of 0.2 and the localization error is within the resolution limit of the source signal.

Sparse data

In the last example of this layer-in source, we consider a sparse network where receivers are regularly placed on the surface with a spacing of 600 m to record the event. Random noise was added to the data to make the S/N equal to 2 (Figure 4.10 c). We use the wavefront attributes to interpolate the recorded wavefield by stacking samples along the traveltime trajectories defined by equation 4.1. A stacking aperture of 1200 m is used in the interpolation process. Despite some minor amplitude deviations caused by noise and

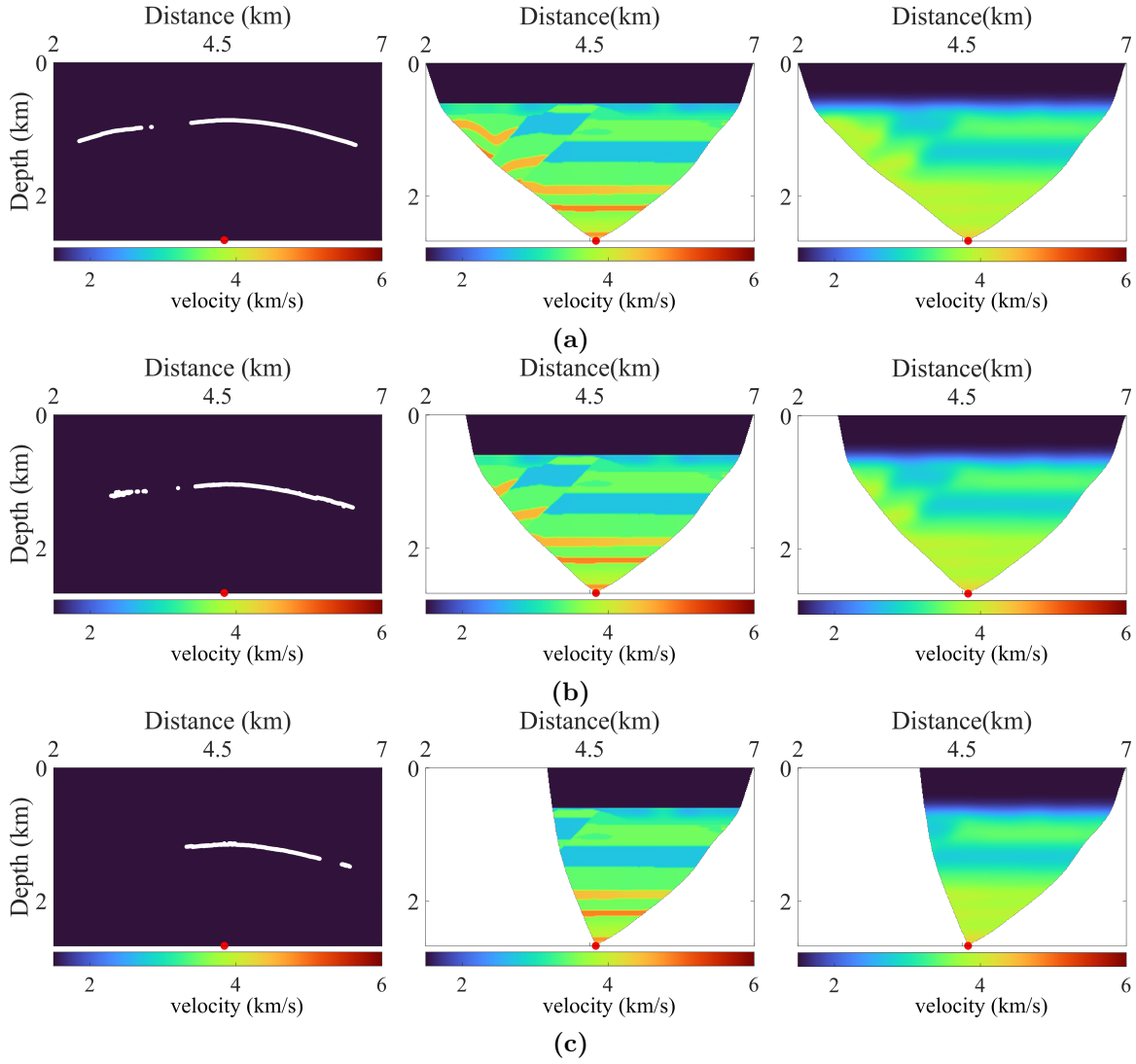


Figure 4.11: Initial setting of the conventional workflow for the layer-in source in the example of densely sampled data (a), noisy data (b), and sparse and noisy data (c). The images in the left panel show the initial constant models, while the middle- and right-panel images display the illuminated part of the true model and smoothed model in the inversion process, respectively. The white dots in the left-panel images indicate the ray starting locations. The red dots denote the true source positions. Only the illuminated part of the model is shown.

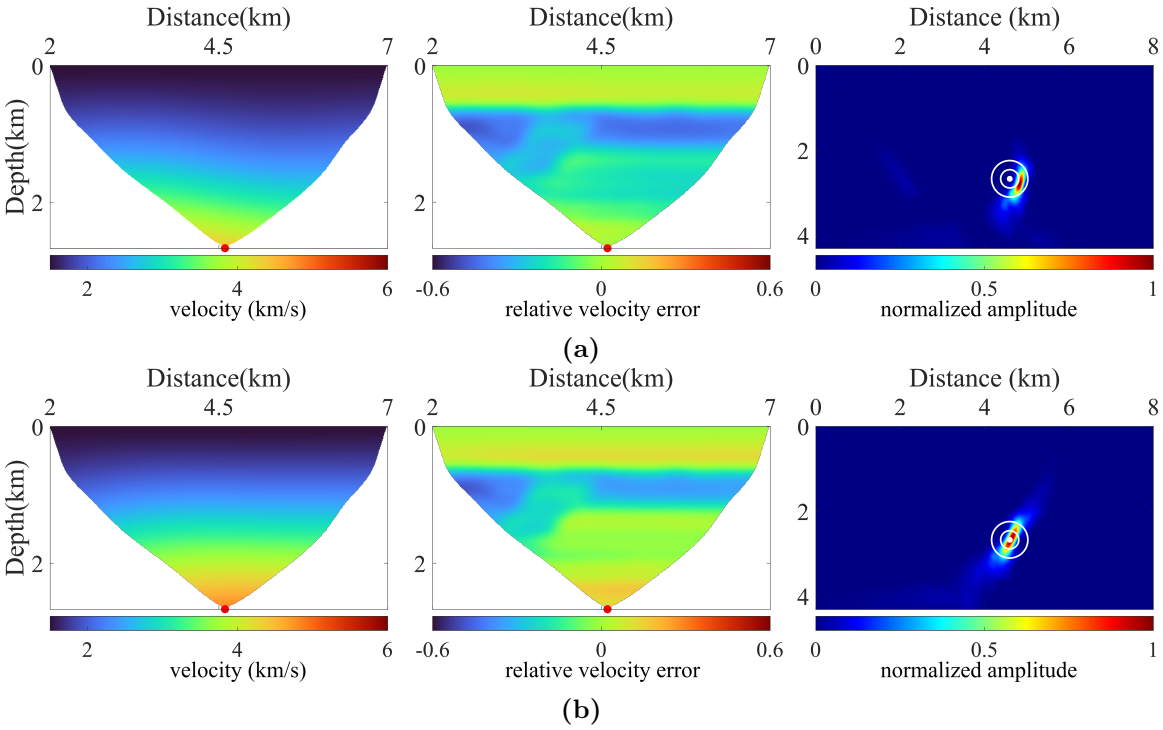


Figure 4.12: Comparison of results obtained from the conventional (a) and hybrid workflow (b) using densely sampled data for the layer-in source. The left panel exhibits the final inverted velocity models, the middle panel displays the relative velocity error of the inverted velocity models to the smoothed velocity models, and the right panel represents the Gaussian-weighted cross-correlation source location images. The red dots in the left- and middle-panel images, and the white dots in the right-panel images, indicate the true source locations. The white circles, centered at the actual source location, have radii of half and one dominant signal wavelength. Only the illuminated part of the model is shown.

spatial aliasing, most of the missing data is well recovered (bottom panel of Figure 4.10 c). The initial value of the source excitation time in the conventional workflow, which is estimated using equation 4.2, deviates from the actual source excitation time by 0.243 s. In this sparse data example, only a limited portion of the left-side model could be illuminated during the inversion. Figure 4.14 a and 4.14 b show the final inversion results obtained by the conventional and hybrid workflows, respectively. The hybrid workflow performs better than the conventional workflow. The velocity errors of the illuminated model region are significantly reduced (Figure 4.14 b), leading to a more accurate determination of the source location and excitation time. The hybrid workflow achieves a source location error of 220 m, about 51 % of the dominant signal wavelength, and a source excitation time error of 0.079 s, smaller than the prevailing signal period. In contrast, the conventional workflow results in final source location and excitation time errors of 583 m and 0.246 s, respectively, exceeding the resolution limit of the source signal. This sparse data example shows that the hybrid workflow is more robust for source localization with noisy and very sparse data.

From these good inversion results obtained by the hybrid workflow for the fault-side source

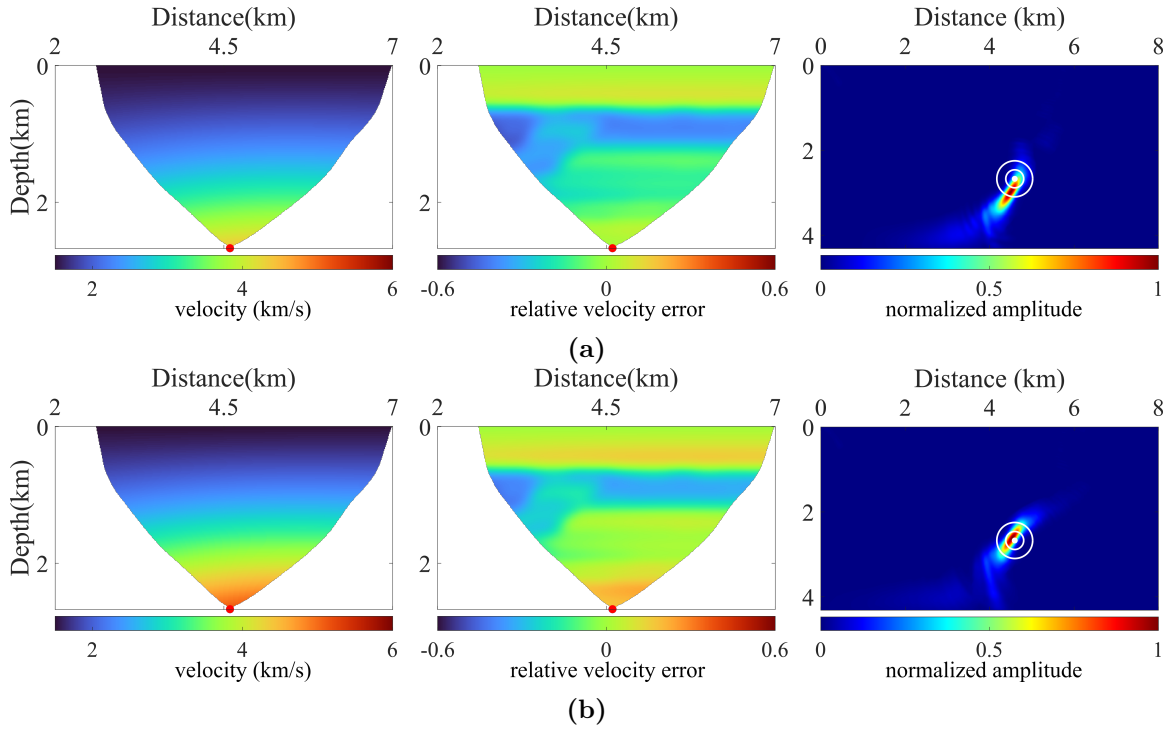


Figure 4.13: Comparison of results obtained from the conventional (a) and hybrid workflow (b) using noisy data for the layer-in source. The left panel exhibits the final inverted velocity models, the middle panel displays the relative velocity error of the inverted velocity models to the smoothed velocity models, and the right panel represents the Gaussian-weighted cross-correlation source location images. The red dots in the left- and middle-panel images, and the white dots in the right-panel images, indicate the true source locations. The white circles, centered at the actual source location, have radii of half and one dominant signal wavelength. Only the illuminated part of the model is shown.

and layer-in source examples, we conclude that the hybrid workflow provides improved results if compared to the conventional workflow for event localization in strongly laterally heterogeneous regions and is better able to handle strong random noise data and very sparse data. Even with a single source, it provides a reasonable velocity model and produces fairly accurate localization results. Source localization errors are all within the signal resolution.

4.5.3 Multiple sources

The proposed hybrid workflow works well for a single source and performs equally well for multiple sources distributed in a cluster. Because the sources are often clustered in practice, we consider a source cluster of 20 seismic sources, represented by red, black, and purple dots in Figure 4.2. Compared to the examples of the fault-side and layer-in sources, the illuminated part of the velocity model (Figure 4.15) expands in this example. We use the same dense seismic network as in the fault-side source example to record the seismic events. Random noise is added to the data from each source to ensure that the S/N is

equal to 2 for all sources. Figure 4.16 a and 4.16 b show the results of the conventional and hybrid workflows, respectively. Both workflows provide high-quality velocity models, with the hybrid workflow providing much better lateral resolution results. It improves the accuracy of inverted velocities, especially for arc-like structures in the vicinity of faults. The mean relative velocity error for the hybrid workflow is approximately -2 %, which is much better than the mean relative velocity error for conventional workflow of about -5 %. Consequently, the hybrid workflow provides more accurate source excitation times compared to the conventional workflow, with errors smaller than a prevailing period, as illustrated in Figure 4.17 a. While both workflows provide good source location estimates (Figure 4.17 b), the hybrid workflow demonstrates good accuracy, with most of the estimated source position errors being less than half the dominant wavelength. These accurate estimates of source excitation times, locations, and velocities emphasize the superiority of the hybrid workflow over the conventional workflow. Moreover, the use of multiple sources allows the hybrid workflow to generate a long-wavelength velocity model of higher accuracy.

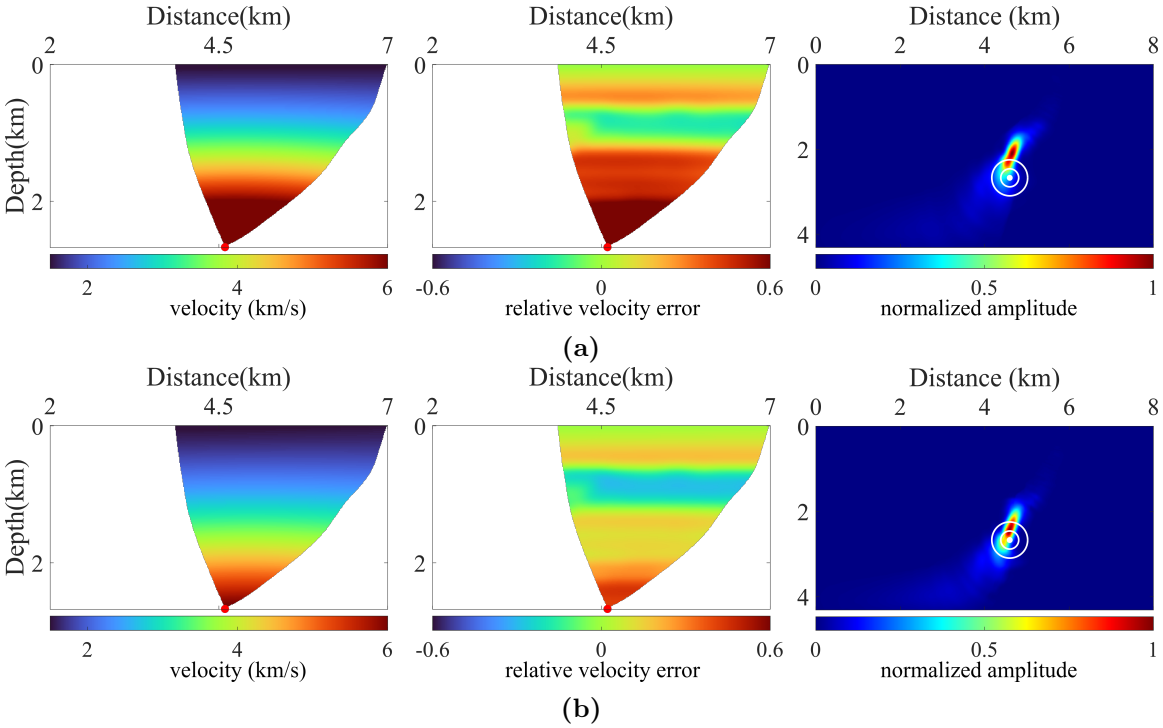


Figure 4.14: Comparison of results obtained from the conventional (a) and hybrid workflow (b) using sparsely sampled and noisy data for the layer-in source. The left panel exhibits the final inverted velocity models, the middle panel displays the relative velocity error of the inverted velocity models to the smoothed velocity models, and the right panel represents the Gaussian-weighted cross-correlation source location images. The red dots in the left- and middle-panel images, and the white dots in the right-panel images, indicate the true source locations. The white circles, centered at the actual source location, have radii of half and one dominant signal wavelength. Only the illuminated part of the model is shown.

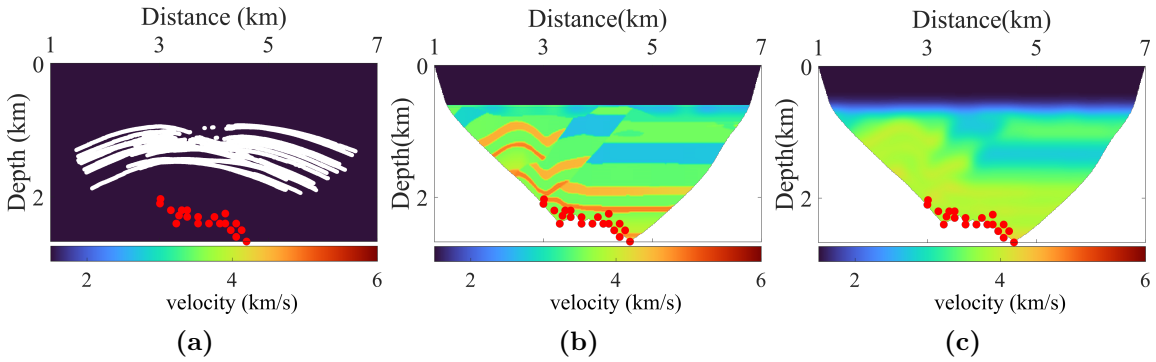


Figure 4.15: (a) Initial model of the conventional workflow for the source cluster of 20 sources. The part of the true model (b) and smoothed model (c) that could be illuminated by rays in the inversion process. The white dots indicate the ray starting locations. The red dots denote the true source positions.

4.6 Discussion

The conventional workflow for joint source localization and velocity model building from passive seismic data described by Diekmann et al. (2019) suffers from inaccurate estimation of source excitation time. When an event is excited within the model region of strong lateral heterogeneity or when an event is recorded in a sparse seismic network, the conventional workflow may lead to significant errors in the inversion results of source time, location, and velocities. This paper presents a robust hybrid workflow for the joint inversion of source location, excitation time, and velocity model, which integrates wavefront tomography with time reversal methods. The proposed workflow involves a search process and an optimization process. The search is designed to establish the optimal initial model for the optimization process, while the optimization process refines the quality of the velocity model and the source excitation time. Both processes involve a sequential use of wavefront tomography and time-reversal methods. Wavefront tomography provides the velocities that are required for the time-reversal methods, and the time-reversal methods determine the source excitation time and source position for wavefront tomography. The improvement of the results of the proposed workflow is caused by the independence of time reversal methods on source time and the potential of wavefront tomography to determine velocities.

Numerical examples display the performance of the proposed workflow in comparison to the conventional workflow. In scenarios where the source is excited in proximity to a complex geological fault, the proposed workflow provides fairly accurate source location and excitation time estimates with errors smaller than the dominant wavelength and the prevailing signal period, respectively. These errors are within the resolution limits of the used passive seismic data. In contrast, the conventional workflow produces source localization results with errors significantly exceeding the signal resolution limit. Furthermore, the proposed workflow outperforms the conventional workflow when applied to noisy data and data from sparse acquisitions. Despite the limited model illumination, the presence of strong random noise and sparse sampling (Figure 4.11 b and 4.11 c), the quality of source localization remains quite good (Figure 4.13 and 4.14). The source location and excitation

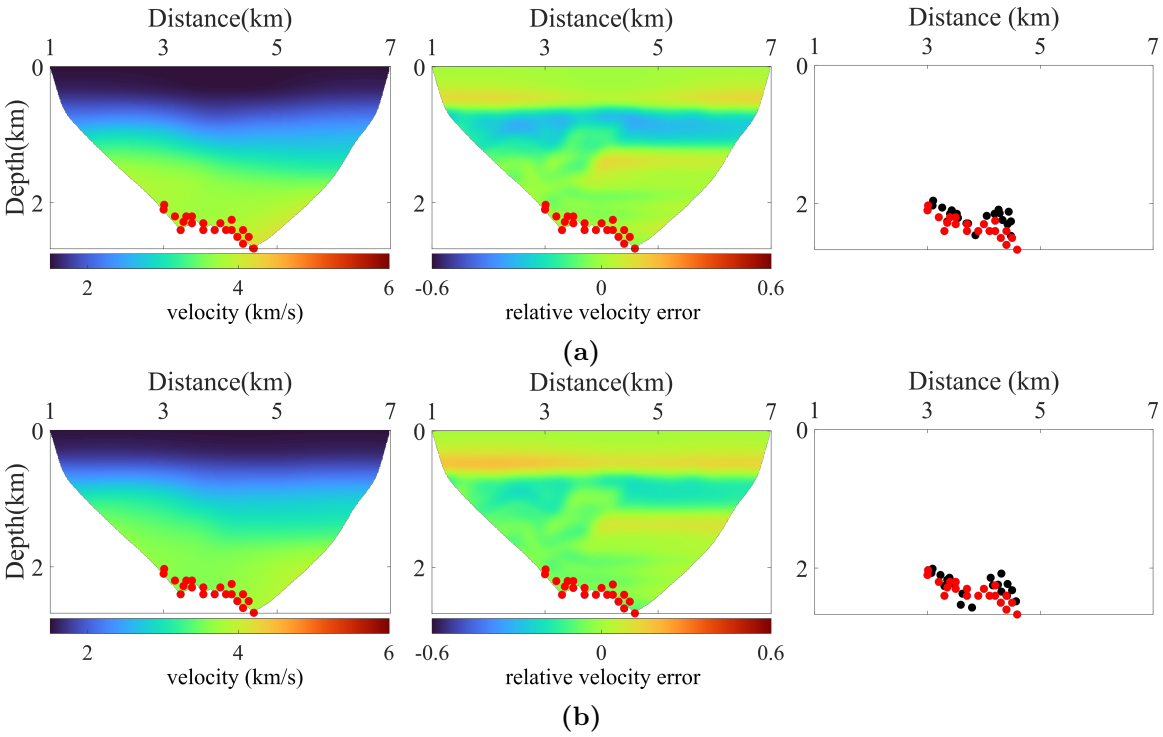


Figure 4.16: Comparison of results obtained from the conventional (a) and hybrid workflow (b) for the example of the source cluster. The left panel exhibits the final inverted velocity models, the middle panel displays the relative velocity error of the inverted velocity models to the smoothed velocity models, and the right panel represents the estimated source locations. The red and black dots in the images indicate the true and estimated source locations, respectively. Only the illuminated part of the model is shown.

time errors of the proposed workflow are within the resolution limit of the signal. The proposed workflow consistently provides highly accurate source locations and times, with errors well within the resolution limits of the signal in all examples. It should be emphasized that these results are obtained despite limited ray coverage.

Compared to the conventional workflow proposed by Diekmann et al. (2019), the proposed workflow circumvents both, the hyperbolic fitting for estimating the source excitation time and the global optimization for finding the optimal source excitation time refinement. These processes become computationally expensive in scenarios with many sources. Although the first step of the proposed workflow requires several runs of reverse modeling, the number of runs depends on the number of predefined gradient velocity models, which accounts for most of the workflow's computational cost. Thus, we can reasonably reduce the number of models to reduce the computation cost of interactive tests. Besides, in comparison to the inversion methods using either recorded waveforms or back-propagated wavefields Song et al. (2019); Wang and Alkhalifah (2018), the proposed workflow does not require detailed prior model information, and it can deal with poor-quality seismic data due to the built-in data enhancement capability using the wavefront attributes.

The time-reversal localization method using the maximum amplitude path (Yang and

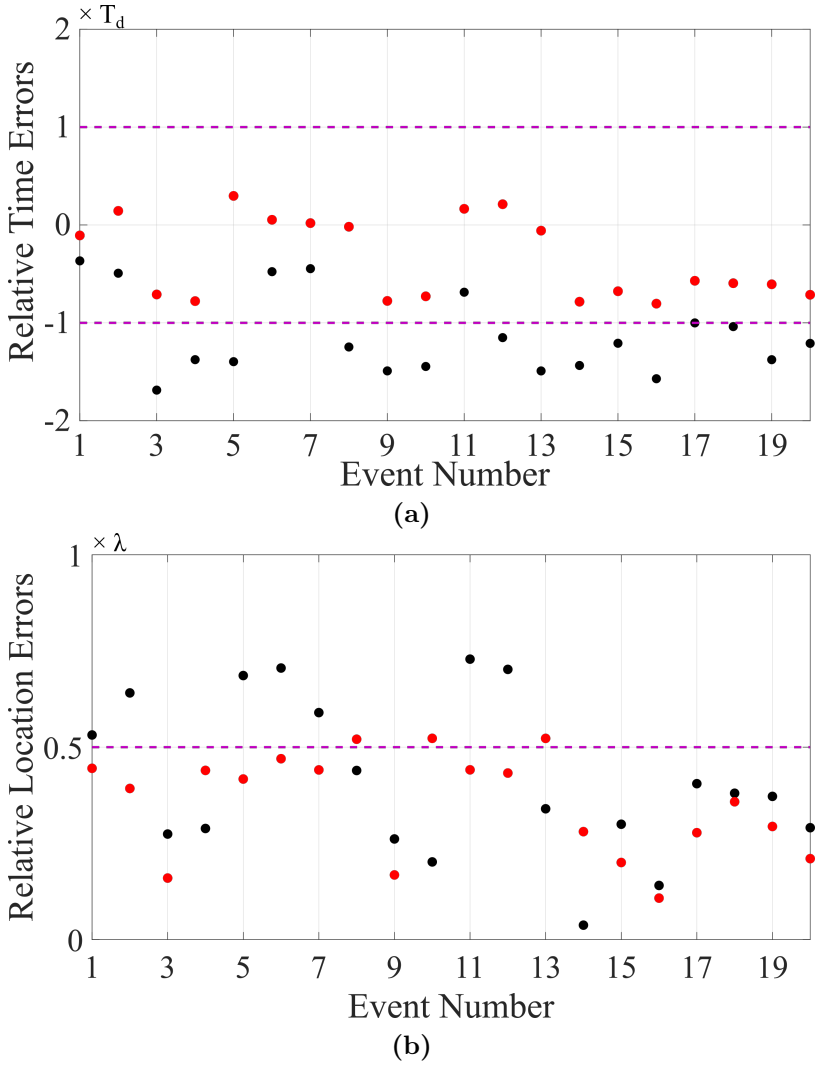


Figure 4.17: Relative errors of the source excitation time (a) and location (b) for the example of the source cluster. The black and red dots indicate the errors in the localization results obtained by the conventional and hybrid workflows, respectively. The λ and T_d denote the dominant wavelength and prevailing period of the signals.

Gajewski, 2022) offers robust results for sparsely sampled data with noise and proves to be effective even in scenarios with small aperture seismic networks. If a good velocity model is provided, e.g., through the workflow proposed here, the localization method using the maximum amplitude paths holds the substantial potential to enhance the accuracy of source time and location. Moreover, the smooth velocity model obtained by the proposed workflow can serve as an initial model in subsequent applications, such as full waveform inversion for both, active-source and passive-source seismic data.

4.7 Conclusion

We present a data-driven hybrid workflow for jointly determining source location, excitation time, and velocity model from passive seismic data. The proposed workflow integrates wavefront tomography and time-reversal methods. Wavefront tomography provides a good velocity model and time-reversal methods yield source excitation time and source location. It is automatic and independent of a priori information. The investigated examples show the improved performance of the proposed workflow in the presence of strong lateral heterogeneity, for noisy data, data acquired for sparse acquisition, and a cluster of events if compared to the conventional approach. The proposed workflow clearly has superior performance in source time estimation. It successfully resolves a velocity model down to the scale of the prevailing wavelength of the data and simultaneously yields fairly accurate source location and source excitation time. The errors of the source times and locations are within the resolution limit of the source signals. The proposed workflow offers a promising and efficient alternative to the currently established techniques for joint inversion of the source location, source excitation time, and velocity model.

Acknowledgments

We thank the Applied Seismics Group, Hamburg, for continuous discussions. The funding by the sponsors of the Wave Inversion Technology (WIT) Consortium and the partial financial support of the University of Hamburg is highly appreciated. P. Yang also thank China CSC for the scholarship. The software packages for forward modeling, data processing, time-reversal imaging, and velocity inversion used in this work were developed through the support of the WIT consortium (https://www.wit.uni-hamburg.de/aboutwit/consortium.html).

5 Conclusions

Time reversal emerged as a powerful tool in passive seismic source localization during the late 20th century and its prominence began to grow in the late 1980s and early 1990s with the development of advanced seismic imaging techniques and computational capabilities. Time-reversal imaging is a common type of time-reversal localization method. It does not require picking seismic arrival times and seismic phases and is robust to high noise levels and strong heterogeneity. However, this type of method often faces challenges, such as strong imaging artifacts due to noisy sparse data, insufficient data illumination due to sparse and small-aperture data acquisitions, and the lack of a sufficiently constrained velocity model.

To clearly identify seismic sources even for source clusters and to reduce imaging noise for data obtained from sparse data acquisitions, I proposed the Gaussian-weighted crosscorrelation imaging condition in **Chapter 2**. This imaging condition yields high-resolution, low-noise source images and it successfully solves the challenges of strong imaging artifacts and the low-resolution issues in the classical time-reversal localization methods. Nonetheless, the Gaussian-weighted crosscorrelation imaging method has the same problem as the common time-reversal imaging methods such as the autocorrelation imaging method or the geometric-mean reverse time migration (GmRTM), that is, the time-reversal imaging requires a large-aperture acquisition that provides sufficient illumination to accurately localize passive seismic sources. This requirement for seismic data acquisition may not meet in practice. The maximum-amplitude path method proposed in **Chapter 3** successfully addresses this type of seismic data acquisition problem. It can produce accurate source location and excitation time using the passive seismic data obtained from sparse and small-aperture acquisitions. Although the methods proposed in **Chapters 2 and 3** overcome the challenges related to poor image quality and limited-aperture acquisitions, they require a reasonably constrained velocity model for accurate passive seismic source localization. However, the velocity model is often unknown in practice. This remains an issue for most established time-reversal localization techniques. In **Chapter 4**, I combined wavefront tomography and time-reversal methods (mentioned in **Chapter 2**) and proposed a hybrid data-driven workflow that allows simultaneous inversion of source location, excitation time, and velocity model. It overcomes the limitation of the lack of velocity model information in time-reversal localization.

5.1 Gaussian-weighted crosscorrelation imaging condition

The Gaussian-weighted crosscorrelation imaging condition (**Chapter 2**) contributes to the improvement of the image quality of passive seismic sources, and enhances the adaptability of time-reversal localization to challenges such as source clusters and noisy data obtained from sparse acquisitions. When dealing with source clusters and noisy and sparsely sampled passive seismic data, the proposed imaging condition can provide high-quality source images for clearly identifying sources while the common time-reversal imaging methods, i.e, the autocorrelation and grouped crosscorrelation imaging methods, cannot. Numerical examples using the Marmousi-ii and 3D SEG overthrust models, and a field example using the HAMNET network in this chapter, demonstrate the advantages of the Gaussian-weighted

crosscorrelation imaging condition. The proposed imaging condition outperforms the autocorrelation and grouped crosscorrelation imaging methods in source imaging, even using passive seismic data of good quality. Compared to the autocorrelation imaging method, the proposed method can produce higher-resolution source images. When compared to the autocorrelation and grouped crosscorrelation imaging methods, the proposed method stands out for its minimal generation of imaging artifacts, reducing the risk of incorrect source location estimation. The errors of source locations obtained by the proposed imaging condition in the examples are all smaller than the dominant wavelength, i.e., within the resolution limit of band-limited passive seismic signals.

5.2 Maximum-amplitude path method

Aside from the problem associated with poor source imaging quality, time-reversal imaging is usually limited by sparse and small-aperture seismic data acquisitions. The maximum-amplitude path method (**Chapter 3**) contributes to improving the accuracy of the time-reversal localization in such acquisitions. It allows accurate estimation of the source location and excitation time using the maximum-amplitude paths. The maximum-amplitude paths are constructed from three or more receiver patches containing a selection of traces from the recorded data. Because the back-projected wavefields focus on the source, the closest proximity or crossing point of these paths specifies the source location, and the corresponding time is the source time. Examples using the Marmousi-ii and 3D SEG overthrust models in the chapter show that the proposed method performs better than the autocorrelation imaging method and the geometric-mean reverse time migration (GmRTM). Despite insufficient data illumination due to such challenging acquisitions, the maximum-amplitude path method still provides accurate source locations and excitation times. The errors of source locations and excitation times estimated in the examples are smaller than the dominant wavelength and prevailing period, i.e., within the resolution limit of band-limited source signals.

5.3 Joint inversion of source location, excitation time and velocity model

In **Chapter 4**, I presented a data-driven hybrid workflow to solve the common issue of unknown velocity models in time-reversal localization. The time-reversal localization methods (mentioned in **Chapters 2, 3**), require a reasonably accurate velocity model to perform numerical backpropagation of the recorded waveforms for accurate source location estimation, however, the velocity model is often unavailable in field cases. The proposed hybrid workflow integrates wavefront tomography and time-reversal methods for simultaneously inverting source location, excitation time, and velocity model from passive seismic data. The wavefront tomography and time-reversal methods are performed sequentially in the workflow. The time-reversal methods, including the Gaussian-weighted crosscorrelation imaging condition proposed in **Chapter 2** and reverse modeling, are used for estimating source location and excitation time, while wavefront tomography is used for velocity inversion. The proposed workflow, independent of a priori information, successfully resolves

a velocity model down to the scale of the prevailing wavelength of the data and simultaneously yields fairly accurate source location and source excitation time. The examples using the overthrust models in the chapter demonstrate the improved performance of the proposed workflow over the conventional approach proposed by Diekmann et al. (2019) (for details, see **Chapter 4**). The proposed workflow clearly exhibits superior performance in source time estimation. It effectively handles situations involving source clusters, data obtained with random noise, and data acquired from sparse seismic acquisitions. The errors of source locations and excitation times estimated in the examples are within the resolution limit of band-limited passive seismic source signals. The proposed workflow offers promising and efficient alternatives to the currently established source localization techniques and improves the robustness of time-reversal methods in passive seismic source localization when no velocity models are provided.

The proposed methods (**Chapters 2, 3, 4**) in this thesis consecutively address the challenges associated with passive seismic source image quality, passive seismic data acquisition limitation, and the lack of seismic velocity model, in time-reversal localization. These advances have direct implications for practical applications, especially in situations where data quality is compromised, which can ultimately enhance our ability to understand Earth's processes and monitor microseismic events with high accuracy. Similar to other source localization methods, the three proposed methods can all be applied over different scales, from acoustic emissions to microseismic events to global earthquakes.

6 Outlook

The proposed methods in this thesis offer robust solutions for determining source locations and excitation times even in complex scenarios. These methodologies hold potential for application in the development of real-time microseismic monitoring systems, facilitating the management of fracturing processes in geoengineering endeavors. In addition, accurate source locations and excitation times improves the ability to analyze the earthquake source mechanism. This detailed knowledge will help to elucidate the characteristics and patterns of microseismicity in the Earth's interior, thus contributing to our understanding of microseismic phenomena.

Seismicity monitoring system

Accurate localization of microseismic sources plays an important role in the development and optimization of real-time microseismic monitoring systems. These systems, equipped with the capability to quickly determine the precise location and excitation time of microseisms, are instrumental in facilitating rapid response protocols during emergencies, such as mine collapses. Improving the localization accuracy of such systems is significant because it enables rapid estimation of fault focus mechanisms (Wamriew et al., 2020). This understanding of fault behavior is crucial for mitigating potential environmental hazards and increasing oil/gas production in geoengineering projects.

The time-reversal localization methods proposed in this thesis (Yang and Gajewski, 2022; Yang et al., 2023) demonstrate their efficacy and robustness in source localization. Future research endeavors should focus on further refining the localization accuracy of real-time monitoring systems by integrating these proposed methods into deep learning frameworks (Zhang et al., 2021; Saad et al., 2021). This integration holds the potential to significantly enhance the capabilities of microseismic monitoring systems, thereby advancing our ability to predict and respond to microseismic events effectively.

Source mechanism inversion

Seismic mechanism inversion improves our understanding of physical processes within the Earth's interior. This knowledge not only contributes to hazard assessment and risk mitigation, but also plays a crucial role in predicting passive seismic events. These insights have far-reaching implications for various fields including petroleum engineering and disaster management (Abercrombie, 2013; Stein, 2003; Hanks and Kanamori, 1979). In petroleum engineering projects involving hydraulic fracturing, accurate determination of location and excitation timing of microseisms can reveal details about fault strike and behavior. This information can also help to infer fault mechanics and aid in the understanding of fault geometry, sliding dynamics, and the modeling of rupture propagation and fault sliding dynamics (Gomberg et al., 2010). The use of source location and onset time as constraints can enhance the inversion algorithms, leading to more accurate mechanism estimation (Wald et al., 1999).

Future research should explore the integration of source parameters estimated by the proposed methods (Yang and Gajewski, 2022; Yang et al., 2023) into source-mechanism inversion techniques, such as the hybrid moment tensor (HMT) technique (Kwiatek et al., 2016) and the scanning algorithm (SSA) (Kao et al., 2006). Such integration has the potential to further refine our understanding of earthquake processes and improve the accuracy of source mechanism determination.

Velocity analysis

The development of a cascade workflow that combines the proposed hybrid workflow (Yang et al., 2024) with the full waveform inversion (FWI) method (Kaderli et al., 2015; Wang and Alkhalifah, 2018; Sun et al., 2016) is a promising research. This study has the potential to improve the resolution of velocity models, while also refining the the accuracy of source localization. FWI techniques (Wang and Alkhalifah, 2018; Sun et al., 2016) are able to simultaneously invert source location, excitation time, and velocity model by fitting modeled data to observed data. While this type of method can produce fine-scale velocity variations, however, it is sensitive to the quality of the initial velocity models, prone to cycle skip, and requires high-quality data. Moreover, it does not include inversion of the source excitation time. The hybrid workflow (Yang et al., 2024) overcomes these limitations of the FWI techniques, which does not require detailed a priori information. Additionally, it can improve the quality of the recorded data. The improved data can be used as the observed data for the FWI, while the inversion results obtained from the hybrid workflow can be used as the initial model setup for the FWI. This integration allows for a more complete understanding of the velocity distribution beneath the Earth's surface and a more accurate characterization of events occurring in the Earth's interior.

Appendix A

Illustrating 2D and 3D rotation in Cartesian coordinates

In **Chapter 3**, the maximum amplitude path is built by scanning for the maximum absolute amplitudes of the back-projected wavefronts within the spatial extent of the dominant wavelength of the considered event. Apart from the first maximum, which is determined from the back-projected wavefield amplitudes at a specified starting time (see **Chapter 3**), the following maxima are identified by scanning for the maximum absolute amplitude of the back-projected wavefield within a rectangular box or cube. The box or cube is centered at the previous maximum, with its axes having a length corresponding to the dominant wavelength of the event.

The orientation of the box in 2D or the cube in 3D is consistent with the wave propagation angle, which is determined by rotating the coordinates. The illustration for calculating the rotation in 2D and 3D Cartesian coordinates is given by the following matrices:
(1) 2D Rotation Matrix: The 2D rotation matrix R for a counterclockwise rotation by an angle θ, i.e., the direction of propagation, is given by

$$R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}, \tag{A.1}$$

where θ denotes the wave propagation angle,

$$\theta = \arctan\left(\frac{u_x^{'}}{u_z^{'}}\right), \tag{A.2}$$

In this formula, $u_x^{'}$ and $u_z^{'}$ represent the first derivative of the seismic wavefield $u(x, z, t)$ with respect to x- and z- axis at the considered time step, respectively. To perform a rotation on a point (x, z) by an angle θ, we can apply the following formula:

$$\begin{bmatrix} x' \\ z' \end{bmatrix} = R(\theta) \begin{bmatrix} x \\ z \end{bmatrix}, \tag{A.3}$$

(2) 3D Rotation Matrices: In 3D, the rotations are around the x, y, and z axes. The corresponding rotation matrices R_x, R_y, and R_z, are defined as follows:

$$\begin{aligned} R_x(\alpha) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix}, \\ R_y(\beta) &= \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix}, \\ R_z(\gamma) &= \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \end{aligned} \tag{A.4}$$

where α, β, and γ, Euler angles, describe the orientation of a coordinate system in three-dimensional space. The α, the roll angle, representing a rotation around the x-axis, corresponds to a tilting motion along the direction of the x-axis in the context of a seismic wave. The β, the pitch angle, which represents a rotation around the y-axis, corresponds to an upward/downward tilting motion along the direction of the y-axis. The γ, the yaw angle, denotes a rotation around the z-axis. To perform a rotation at a point (x, y, z) by these angles, the following formula is applied

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R_x(\alpha) \cdot R_y(\beta) \cdot R_z(\gamma) \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tag{A.5}$$

These Euler angles (α, β, γ), representing the rotation of the seismic wavefield in 3D, are required to align the coordinate system with the principal axes of the seismic wavefield. They are related to the gradients of the wavefield, which can be obtained by performing an eigenanalysis on the gradient tensor or covariance matrix of the seismic wavefield. The eigenanalysis comprises the following steps. Firstly, we numerically calculate the gradient tensor by differentiating the seismic wavefield $u(x, y, z, t)$ in 3D, resulting in a 3×3 matrix

$$G = \begin{bmatrix} u''_{xx} & u''_{xy} & u''_{xz} \\ u''_{yx} & u''_{yy} & u''_{yz} \\ u''_{zx} & u''_{zy} & u''_{zz} \end{bmatrix}, \tag{A.6}$$

where u''_{ij} represents the gradient component in the i-th direction with respect to the j-th coordinate at the considered time step. Then, we compute the covariance matrix C from the gradient tensor G using the given formula:

$$C = G^T \cdot G, \tag{A.7}$$

, and we perform an eigenanalysis on the covariance matrix C to obtain its eigenvectors and eigenvalues. The eigenvectors denote the principal axes of the seismic wavefield, and the eigenvalues indicate the magnitude of the gradients along these axes.

$$C \cdot \mathbf{v} = \lambda \cdot \mathbf{v}, \tag{A.8}$$

Where $\mathbf{v}$ represents the eigenvector, and λ indicates the corresponding eigenvalue. Finally, the eigenvectors obtained from the eigenanalysis can be used to calculate the Euler angles (α, β, γ) using the following expressions:

$$\alpha = \arctan\left(\frac{v_{y,1}}{v_{x,1}}\right), \quad \beta = \arcsin(v_{z,1}), \quad \gamma = \arctan\left(\frac{v_{z,2}}{v_{z,3}}\right), \tag{A.9}$$

Here, $(v_{x,1}, v_{y,1}, v_{z,1})$ represents components of the first eigenvector, and $(v_{z,2}, v_{z,3})$ indicate the second and third components of the third eigenvector. Note that, the specific notation may vary depending on the convention used.

Appendix B

Calculating modeled wavefront attributes

In **Chapter 4**, the wavefront attributes, determined from a parabolic fitting process in a data-driven fashion, serve as observed data for the velocity inversion, i.e., wavefront tomography (Duveneck, 2004; Bauer et al., 2019a), while the modeled data used for calculating the data misfits in the inversion is obtained by kinematic and dynamic ray tracing. In the following, I will show how to compute the wavefront attributes by kinematic and dynamic ray tracing in the general Cartesian coordinates x_1, x_2, x_3, assuming a certain smooth, laterally inhomogeneous background velocity model is given. Reducing the general Cartesian coordinates from three to two dimensions makes it a two-dimensional case.

Kinematic ray tracing is important for forward modeling to compute travel time and the slowness vector along a reference ray. While kinematic ray tracing is in general expressed using curvilinear coordinates ξ_1, ξ_2, and ξ_3, it can be complicated to perform in such coordinates because of the need to analytically express the scale factors h_1, h_2, and h_3 in orthogonal coordinates (Červený, 2001). Spherical polar coordinates r, θ, and ϕ are commonly used, but can be problematic in some regions. Therefore, ray tracing in ray-centered coordinates is used, requiring parameter transformations between these and general Cartesian coordinates. In this work, I focus on kinematic and dynamic ray tracing in inhomogeneous isotropic media, similar relations for anisotropic media can be found elsewhere (Červený, 2001).

Here, I provide a solution for performing kinematic and dynamic ray tracing in the general Cartesian coordinate system and explain how the slowness vector and ray-propagation matrix computed by kinematic and dynamic ray tracing are related to the data-driven wavefront attributes. In general, the Cartesian coordinate system for isotropic media simplifies the eikonal equation to

$$\sum_{i=1}^{3}(p_i^{(x)})^2 - v^{-2}(x_1, x_2, x_3) = 0, \quad p_i^{(x)} = \frac{\partial \tau}{\partial x_i}, \quad i = 1, 2, 3, \tag{B.1}$$

where τ represents the traveltime along a given ray and $v(x_1, x_2, x_3)$ denotes the background medium velocity. By assuming no turning point with respect to the x_3- direction and eliminating one space variables, the reduced Hamiltonian $\mathcal{H}$ is defined as

$$p_3^{(x)} = \left[v^{-2} - (p_1^{(x)})^2 - (p_2^{(x)})^2\right]^{1/2} = -\mathcal{H}\left(x_1, x_2, x_3, p_1^{(x)}, p_2^{(x)}\right), \tag{B.2}$$

with associated reduced ray-tracing systems reads

$$\begin{aligned}
\frac{dx_1}{dx_3} &= \frac{\partial \mathcal{H}}{\partial p_1^{(x)}} = \frac{p_1^{(x)}}{p_3^{(x)}}, \\
\frac{dx_2}{dx_3} &= \frac{\partial \mathcal{H}}{\partial p_2^{(x)}} = \frac{p_2^{(x)}}{p_3^{(x)}}, \\
\frac{dp_1^{(x)}}{dx_3} &= -\frac{\partial \mathcal{H}}{\partial x_1} = -\frac{1}{v^3 p_3^{(x)}} \frac{\partial v}{\partial x_1}, \\
\frac{dp_2^{(x)}}{dx_3} &= -\frac{\partial \mathcal{H}}{\partial x_2} = -\frac{1}{v^3 p_3^{(x)}} \frac{\partial v}{\partial x_2},
\end{aligned} \tag{B.3}$$

The traveltime variation along the given ray is integrated as

$$\frac{d\tau}{dx_3} = \frac{1}{v^2 p_3^{(x)}}, \tag{B.4}$$

The reduced Hamiltonian offers several advantages, such as reducing the differential equations in kinematic ray tracing from seven to five and requiring derivatives of velocity only in the x_1- and x_2- directions, which improves computational efficiency in the three-dimensional case. The reduced Hamiltonian is particularly beneficial since rotating normals rarely occur in practice (Duveneck, 2004).

Dynamic ray tracing, also known as paraxial ray tracing, involves the computation of the 4×4 ray propagator matrix $\underline{\mathbf{\Pi}}^{(x)}$ or the wavefront curvature matrix $\underline{\mathbf{M}}^{(x)}$. The previously mentioned reduced Hamiltonian is applied in dynamic ray tracing, and the dynamic ray tracing system in the Cartesian coordinate system is written as

$$\frac{d}{dx_3} \Delta \boldsymbol{w} = \underline{\mathbf{S}} \Delta \boldsymbol{w}, \tag{B.5}$$

where $\underline{\mathbf{S}}$ is a 4×4 matrix

$$\underline{\mathbf{S}} = \begin{pmatrix}
\dfrac{\partial^2 \mathcal{H}}{\partial p_1^{(x)} \partial x_1} & \dfrac{\partial^2 \mathcal{H}}{\partial p_1^{(x)} \partial x_2} & \dfrac{\partial^2 \mathcal{H}}{\partial p_1^{(x)} \partial p_1^{(x)}} & \dfrac{\partial^2 \mathcal{H}}{\partial p_1^{(x)} \partial p_2^{(x)}} \\
\dfrac{\partial^2 \mathcal{H}}{\partial p_2^{(x)} \partial x_1} & \dfrac{\partial^2 \mathcal{H}}{\partial p_2^{(x)} \partial x_2} & \dfrac{\partial^2 \mathcal{H}}{\partial p_2^{(x)} \partial p_1^{(x)}} & \dfrac{\partial^2 \mathcal{H}}{\partial p_2^{(x)} \partial p_2^{(x)}} \\
-\dfrac{\partial^2 \mathcal{H}}{\partial x_1 \partial x_1} & -\dfrac{\partial^2 \mathcal{H}}{\partial x_1 \partial x_2} & -\dfrac{\partial^2 \mathcal{H}}{\partial x_1 \partial p_1^{(x)}} & -\dfrac{\partial^2 \mathcal{H}}{\partial x_1 \partial p_2^{(x)}} \\
-\dfrac{\partial^2 \mathcal{H}}{\partial x_2 \partial x_1} & -\dfrac{\partial^2 \mathcal{H}}{\partial x_2 \partial x_2} & -\dfrac{\partial^2 \mathcal{H}}{\partial x_2 \partial p_1^{(x)}} & -\dfrac{\partial^2 \mathcal{H}}{\partial x_2 \partial p_2^{(x)}}
\end{pmatrix}, \tag{B.6}$$

with elements calculated using equation ((B.2)), which is given by

$$
\begin{aligned}
&\frac{\partial^2\mathcal{H}}{\partial p_1^{(x)}\partial x_1}=\frac{p_1^{(x)}}{v^3(p_3^{(x)})^3}\frac{\partial v}{\partial x_1},\quad \frac{\partial^2\mathcal{H}}{\partial p_1^{(x)}\partial x_2}=\frac{p_1^{(x)}}{v^3(p_3^{(x)})^3}\frac{\partial v}{\partial x_2},\\
&\frac{\partial^2\mathcal{H}}{\partial p_1^{(x)}\partial p_1^{(x)}}=\frac{(p_1^{(x)})^2}{(p_1^{(x)})^3}+\frac{1}{p_3^{(x)}},\quad \frac{\partial^2\mathcal{H}}{\partial p_1^{(x)}\partial p_2^{(x)}}=\frac{p_1^{(x)}p_2^{(x)}}{(p_3^{(x)})^3},\\
&\frac{\partial^2\mathcal{H}}{\partial p_2^{(x)}\partial x}=\frac{p_2^{(x)}}{v^3(p_3^{(x)})^3}\frac{\partial v}{\partial x_1},\quad \frac{\partial^2\mathcal{H}}{\partial p_2^{(x)}\partial x_2}=\frac{p_2^{(x)}}{v^3(p_3^{(x)})^3}\frac{\partial v}{\partial x_2},\\
&\frac{\partial^2\mathcal{H}}{\partial p_2^{(x)}\partial p_1^{(x)}}=\frac{p_1^{(x)}p_2^{(x)}}{(p_3^{(x)})^3},\quad \frac{\partial^2\mathcal{H}}{\partial p_2^{(x)}\partial p_2^{(x)}}=\frac{(p_2^{(x)})^2}{(p_1^{(x)})^3}+\frac{1}{p_3^{(x)}},\\
&-\frac{\partial^2\mathcal{H}}{\partial x_1\partial x_1}=-\frac{1}{v^4p_3^{(x)}}\big[v\frac{\partial^2 v}{\partial x_1^2}+(\frac{1}{v^2(p_3^{(x)})^2})(\frac{\partial v}{\partial x_1})^2\big],\\
&-\frac{\partial^2\mathcal{H}}{\partial x_1\partial x_2}=-\frac{1}{v^4p_3^{(x)}}\big[v\frac{\partial^2 v}{\partial x_1\partial x_2}+(\frac{1}{v^2(p_3^{(x)})^2})(\frac{\partial v}{\partial x_1})(\frac{\partial v}{\partial x_2})\big],\\
&-\frac{\partial^2\mathcal{H}}{\partial x_1\partial p_1^{(x)}}=-\frac{p_1^{(x)}}{v^3(p_3^{(x)})^3}\frac{\partial v}{\partial x_1},\quad -\frac{\partial^2\mathcal{H}}{\partial x_1\partial p_2^{(x)}}=-\frac{p_2^{(x)}}{v^3(p_3^{(x)})^3}\frac{\partial v}{\partial x_1},\\
&-\frac{\partial^2\mathcal{H}}{\partial x_2\partial x_1}=-\frac{1}{v^4p_3^{(x)}}\big[v\frac{\partial^2 v}{\partial x_1\partial x_2}+(\frac{1}{v^2(p_3^{(x)})^2})(\frac{\partial v}{\partial x_1})(\frac{\partial v}{\partial x_2}\big],\\
&-\frac{\partial^2\mathcal{H}}{\partial x_2\partial x_2}=-\frac{1}{v^4p_3^{(x)}}\big[v\frac{\partial^2 v}{\partial x_1^2}+(\frac{1}{v^2(p_3^{(x)})^2})(\frac{\partial v}{\partial x_1})^2\big],\\
&-\frac{\partial^2\mathcal{H}}{\partial x_2\partial p_1^{(x)}}=-\frac{p_1^{(x)}}{v^3(p_3^{(x)})^3}\frac{\partial v}{\partial x_2},\quad -\frac{\partial^2\mathcal{H}}{\partial x_2\partial p_2^{(x)}}=-\frac{p_2^{(x)}}{v^3(p_3^{(x)})^3}\frac{\partial v}{\partial x_2},
\end{aligned}
\tag{B.7}
$$

The matrix $\underline{\mathbf{\Pi}}^{(x)}$ satisfies

$$
\frac{d}{dx_3}\underline{\mathbf{\Pi}}^{(x)}=\underline{\mathbf{S}}\ \underline{\mathbf{\Pi}}^{(x)},
\tag{B.8}
$$

and is defined as

$$
\underline{\mathbf{\Pi}}^{(x)}=\begin{pmatrix}\mathbf{Q}_1^{(x)} & \mathbf{Q}_2^{(x)}\\ \mathbf{P}_1^{(x)} & \mathbf{P}_2^{(x)}\end{pmatrix},
\tag{B.9}
$$

The matrices $\mathbf{Q}_1^{(x)}$, $\mathbf{Q}_2^{(x)}$, $\mathbf{P}_2^{(x)}$, and $\mathbf{P}_2^{(x)}$ can be integrated along the reference ray with the increment of dx_3 starting from a ray source to a receiver. In the general scenario, the solutions of $(\mathbf{Q}^{(x)},\mathbf{P}^{(x)})^T$ in the dynamic ray-tracing system can be employed to determine the wavefront curvatures $\underline{\mathbf{M}}^{(x)}$

$$
\underline{\mathbf{M}}^{(x)}=\mathbf{P}^{(x)}\mathbf{Q}^{(x)-1},
\tag{B.10}
$$

with $M_{11}=\partial^2\tau/\partial x_1^2$, $M_{12}=M_{21}=\partial^2\tau/\partial x_1\partial x_2$, and $M_{22}=\partial^2\tau/\partial x_2^2$. The expression of paraxial traveltimes of rays near the central ray belonging to a specified wave reads

$$
\tau(x_1+\Delta x_1,x_2+\Delta x_2,x_3)=\tau(x_1,x_2,x_3)+\sum_{i=1}^{2}p_i^{(x)}\Delta x_i+\sum_{i,j=1}^{2}M_{ij}^{(x)}\Delta x_i\Delta x_j,
\tag{B.11}
$$

Bibliography

Abercrombie, R. (2013). Earthquake source scaling relationships from -1 to 5 ML using seismograms recorded at 2 km depth. *Journal of Geophysical Research: Solid Earth*, 118:1539–1554.

Aki, K. and Richards, P. (2002). *Quantitative seismology*. Sausalito, Calif: University Science Books.

Aminzadeh, F., Brac, J., and Kunz, T. (1997). *SEG/EAGE 3-D Salt and Overthrust Models. SEG/EAGE 3-D Modeling Series, No. 1: Distribution CD of Salt and Overthrust models.* SEG book series.

Anderson, S. (2017). Risk, Liability, and Economic Issues with Long-Term CO2 Storage - A Review. *Natural Resources Research*, 26:89–112.

Anikiev, D., Valenta, J., Staněk, F., and Eisner, L. (2014). Joint location and source mechanism inversion of microseismic events: benchmarking on seismicity induced by hydraulic fracturing. *Geophysical Journal International*, 198:249–258.

Artman, B., Podladtchikov, I., and Witten, B. (2010). Source location using time-reverse imaging. *Geophysical Prospecting*, 58:861–873.

Bai, T., Lyu, B., Williamson, P., and Nakata, N. (2022). Receiver grouping strategies for hybrid geometric-mean reverse time migration. *Geophysics*, 87:KS45–KS55.

Bardainne, T., Gaucher, E., Cerda, F., and Drapeau, D. (2009). Comparison of picking-based and waveform-based location methods of microseismic events: Application to a fracturing job. *SEG Technical Program Expanded Abstracts*, pages 1547–1551.

Bauer, A., Schwarz, B., Diekmann, L., and Gajewski, D. (2019a). Wavefront tomography with enforced diffraction focusing. *in Proceedings of the 81st EAGE Conference and Exhibition*, pages 1–3.

Bauer, A., Schwarz, B., Werner, T., and Gajewski, D. (2019b). Unsupervised event identification and tagging for diffraction focusing. *Geophysical Journal International*, 217:2165–2176.

Baykulov, M. and Gajewski, D. (2009). Prestack seismic data enhancement with partial common-reflection-surface (CRS) stack. *Geophysics*, 74(3):V49–V58.

Capdeville, Y. and Maron, C. (2006). Time-reversal imaging for the passive seismic experiment of AlpArray. *Geophysical Journal International*, 165:443–451.

Castle, R. (1994). A theory of normal moveout. *Geophysics*, 59(6):983—999.

Cerjan, C., Kosloff, D., Kosloff, R., and Reshef, M. (1985). A nonreflecting boundary condition for discrete acoustic and elastic wave equations. *Geophysics*, 50:705–708.

Chen, Y., Saad, O., Liu, X., and Fomel, S. (2021). A Compact Program for 3D Passive Seismic Source-Location Imaging. *Seismological Research Letters*, 92:3187—3201.

Chen, Y., Saad, O., Savvaidis, A., Chen, Y., and Fomel, S. (2022). 3D microseismic monitoring using machine learning. *Journal of Geophysical Research: Solid Earth*, 127:e2021JB023842.

Cheng, G., Ma, T., Tang, C., Liu, H., and Wang, S. (2017). A zoning model for coal mining-induced strata movement based on microseismic monitoring. *International Journal of Rock Mechanics and Mining Sciences*, 94:123–138.

De Boor, C. (1978). *A Practical Guide to Splines*. Springer.

Diekmann, L., Schwarz, B., Bauer, A., and Gajewski, D. (2019). Source localization and joint velocity model building using wavefront attributes. *Geophysical Journal International*, 219:995–1007.

Dong, L., Wesseloo, J., Potvin, Y., and Li, X. (2016). Discrimination of mine seismic events and blasts using the Fisher classifier, naive Bayesian classifier and logistic regression. *Rock Mechanics and Rock Engineering*, 49:183–211.

Douma, J. and Snieder, R. (2015). Focusing of elastic waves for microseismic imaging. *Geophysical Journal International*, 200:390–401.

Duveneck, E. (2004). Velocity model estimation with data-derived wavefront attributes. *Geophysics*, 69(1):265–274.

Dyer, B., Schanz, U., Ladner, F., Häring, M., and Spillman, T. (2008). Microseismic imaging of a geothermal reservoir stimulation. *The Leading Edge*, 27:856–869.

Eisner, L., Duncan, P., Heigl, W., and Keller, W. (2009). Uncertainties in passive seismic monitoring. *The Leading Edge*, 28:648–655.

Ellsworth, W. (2013). Injection-induced earthquakes. *Science*, 341:1225942.

Fink, M. (1996). *Time reversal in acoustics*, volume 37.

Fink, M., Kuperman, W., Montagner, J., and Tourin, A. (2002). *Imaging of complex media with acoustic and seismic waves*. Springer.

Folesky, J., Kummerow, J., and Shapiro, S. (2015). Microseismic rupture propagation imaging. *Geophysics*, 80:WC107–WC115.

Font, Y., Kao, H., Lallemand, S., Liu, C., and Chiao, L. (2004). Hypocentre determination offshore of eastern Taiwan using the maximum intersection method. *Geophysical Journal International*, 158:655–675.

Gajewski, D. and Tessmer, E. (2005). Reverse modeling for seismic event characterization. *Geophysical Journal International*, 163:276–284.

Geiger, L. (1912). Probability method for the determination of earthquake epicenters from the arrival time only. *Bulletin St. Louis University*, 8:56–71.

Goertz Allmann, B., Kühn, D., Oye, V., Bohloli, B., and Aker, E. (2014). Combining microseismic and geomechanical observations to interpret storage integrity at the In Salah CCS site. *Geophysical Journal International*, 198:447–461.

Gomberg, J., Bodin, P., and Larson, K. (2010). Earthquake nucleation by transient deformations caused by the M= 7.9 Denali, Alaska, earthquake. *Nature Geoscience*, 3:310–313.

Grechka, V., Li, Z., Howell, B., Garcia, H., and Wooltorton, T. (2017). High-resolution microseismic imaging. *The Leading Edge*, 36:822–828.

Grigoli, F., Cesca, S., Amoroso, O., Emolo, A., Zollo, A., and Dahm, T. (2014). Automated seismic event location by waveform coherence analysis. *Geophysical Journal International*, 196:1742–1753.

Grigoli, F., Cesca, S., Vassallo, M., and Dahm, T. (2013). Automated seismic event location by travel-time stacking: An application to mining induced seismicity. *Seismological Research Letters*, 84:666–677.

Hanks, T. and Kanamori, H. (1979). A moment magnitude scale. *Journal of Geophysical Research: Solid Earth*, 84:2348–2350.

Hubral, P. (1983). Computing true amplitude reflections in a laterally inhomogeneous Earth. *Geophysics*, 48(8):1051–1062.

Jia, X. and Wu, R. (2009). Superwide-angle one-way wave propagator and its application in imaging steep salt flanks. *Geophysics*, 74:S75–S83.

Jones, G., Kendall, J., Bastow, I., and Raymer, D. (2014). Locating microseismic events using borehole data. *Geophysical Prospecting*, 62:34–49.

Jäger, R., Mann, J., Höcht, G., and Hubral, P. (2001). Common-reflection- surface stack: Image and attributes. *Geophysics*, 66(1):97–109.

Kaderli, J., McChesney, W., and Minkoff, S. (2015). Microseismic event estimation in noisy data via full waveform inversion. *SEG Technical Program Expanded Abstracts*, pages 1159–1164.

Kamei, R., Nakata, N., and Lumley, D. (2015). Introduction to microseismic source mechanisms. *The Leading Edge*, 34:876–880.

Kao, H., Shan, S., Dragert, H., Rogers, G., Cassidy, J., Wang, K., James, T., and Ramachandran, K. (2006). Spatial-temporal patterns of seismic tremors in northern Cascadia. *Journal of Geophysical Research*, 111:B03309.

Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks*, pages 1942–1948.

Kwiatek, G., Martinez Garzon, P., and Bohnhoff, M. (2016). HybridMT: A MATLAB/Shell Environment Package for Seismic Moment Tensor Inversion and Refinement. *Seismological Research Letters*, 87:964–976.

Larmat, C., Tromp, J., Liu, Q., and Montagner, J. (2008). Time reversal location of glacial earthquakes. *Journal of Geophysical Research*, 113:B09314.

Li, B. and Einstein, H. (2019). Direct and microseismic observations of hydraulic fracturing in barre granite and opalinus clayshale. *Journal of Geophysical Research: Solid Earth*, 124:11900–11916.

Li, L., Becker, D., Chen, H., Wang, X., and Gajewski, D. (2018). A systematic analysis of correlation-based seismic location methods. *Geophysical Journal International*, 212:659–678.

Li, L., Tan, J., Schwarz, B., Staněk, F., Poiata, N., Shi, P., Diekmann, L., Eisner, L., and Gajewski, D. (2020). Recent advances and challenges of waveform-based seismic location methods at multiple scales. *Reviews of Geophysics*, 58:e2019RG000667.

Martin, G., Marfurt, K., and Larsen, S. (2002). Marmousi-2 : An updated model for the investigation of AVO in structurally complex areas. *SEG Technical Program Expanded Abstract*, pages 1979–1982.

Maxwell, S., Rutledge, J., Jones, R., and Fehler, M. (2010). Petroleum reservoir characterization using downhole microseismic monitoring. *Geophysics*, 75:75A129–75A137.

Maxwell, S. and Urbancic, T. (2001). The role of passive microseismic monitoring in the instrumented oil field. *The Leading Edge*, 20(6):636–63.

Mayne, W. (1962). Common reflection point horizontal data stacking techniques. *Geophysics*, 27(6):927–938.

Mazzoldi, A., Rinaldi, A., Borgia, A., and Rutqvist, J. (2012). Induced seismicity within geological carbon sequestration projects: Maximum earthquake magnitude and leakage potential from undetected faults. *International Journal of Greenhouse Gas Control*, 10:434–442.

Nakata, N. and Beroza, G. (2016). Reverse time migration for microseismic sources using the geometric mean as an imaging condition. *Geophysics*, 81:KS51–KS60.

Oren, C. and Shragge, J. (2021). PS energy imaging condition for microseismic data — Part 1: Theory and applications in 3D isotropic media. *Geophysics*, 86:KS37–KS48.

Oye, V., Aker, E., Daley, T., Kühn, D., Bohloli, B., and Korneev, V. (2013). Microseismic monitoring and interpretation of injection data from the in Salah CO2 storage site (Krechba), Algeria. *Energy Procedia*, 37:4191–4198.

O'Brien, G., Lokmer, I., De Barros, L., Bean, C., Saccorotti, G., Metaxian, J., and Patane, D. (2011). Time reverse location of seismic long-period events recorded on Mt Etna. *Geophysical Journal International*, 184:452–462.

Price, D., Angus, D., Chambers, K., and Jones, G. (2015). Surface microseismic imaging in the presence of high-velocity lithologic layers. *Geophysics*, 80:WC117–WC131.

Priestley, M. (1996). Wavelets and time-dependent spectral analysis. *Journal of Time Series Analysis*, 17:85–103.

Romero, L., Ghiglia, D., Ober, C., and Morton, S. (2000). Phase encoding of shot records in prestack migration. *Geophysics*, 65:426–436.

Saad, O. and Chen, Y. (2020). Automatic waveform-based source-location imaging using deep learning extracted microseismic signals. *Geophysics*, 85:KS171–KS183.

Saad, O. M., Huang, G., Chen, Y., Savvaidis, A., Fomel, S., Pham, N., and Chen, Y. (2021). SCALODEEP: A highly generalized deep learning framework for real-time earthquake detection. *Journal of Geophysical Research: Solid Earth*, 126:e2020JB021473.

Sawires, R., Peláez, J., and Ibrahim, H. e. a. (2016). Delineation and characterization of a new seismic source model for seismic hazard studies in Egypt. *Natural Hazards*, 80:1823–1864.

Schultz, R., Beroza, G., and Ellsworth, W. (2021). A risk-based approach for managing hydraulic fracturing–induced seismicity. *Science*, 372:504–507.

Schwarz, B., Bauer, A., and Gajewski, D. (2016). Passive seismic source localization via common-reflection-surface attributes. *Studia Geophysica et Geodaetica*, 60:531–546.

Shearer, P. (2019). *Introduction to Seismology*.

Shelly, D., Beroza, G., and Ide, S. (2007). Complex evolution of transient slip derived from precise tremor locations in western Shikoku, Japan. *Geochemistry Geophysics Geosystems*, 8:Q10014.

Song, C. and Alkhalifah, T. (2019). Microseismic event estimation based on an efficient wavefield inversion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12:4664–4671.

Song, C., Alkhalifah, T., and Wu, Z. (2019). Microseismic event estimation and velocity analysis based on a source-focusing function. *Geophysics*, 84:KS85–KS94.

Stein, R. (2003). The role of stress transfer in earthquake occurrence. *Nature*, 402:605–609.

Steiner, B., Saenger, E., and Schmallholz, S. (2008). Time reverse modeling of low-frequency microtremors: Application to hydrocarbon reservoir localization. *Geophysical Research Letters*, 35:L03307.

Storn, R. and Price, K. (1997). Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11:341–359.

Suckale, J. (2009). Induced seismicity in hydrocarbon fields. *Advances in Geophysics*, 51:55–106.

Sun, J., Xue, Z., Fomel, S., and Nakata, N. (2016). Full-waveform inversion of passive seismic data for sources and velocities. *SEG Technical Program Expanded Abstracts*, pages 1405–1410.

Taner, M. and Koehler, F. (1969). Velocity spectra-digital computer derivation applications of velocity functions. *Geophysics*, 34(6):859–881.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*.

Ursin, B. (1982). Quadratic wavefront and traveltime approximations in inhomogeneous layered media with curved interfaces. *Geophysics*, 47:1012–1021.

Wald, D., Quitoriano, V., Heaton, T., and Kanamori, H. (1999). Relationships between peak ground acceleration, peak ground velocity, and modified Mercalli intensity in California. *Earthquake Spectra*, 15:557–564.

Walda, J. and Gajewski, D. (2017). Determination of wavefront attributes by differential evolution in the presence of conflicting dips. *Geophysics*, 82(4):V229–V239.

Waldhauser, F. and Ellsworth, W. (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California. *Bulletin of the Seismological Society of America*, 90:1353–1368.

Wamriew, D., Charara, M., and Evgenii, M. (2020). Deep Neural Network for Real-Time Location and Moment Tensor Inversion of Borehole Microseismic Events Induced by Hydraulic Fracturing. *SPE Russian Petroleum Technology Conference*.

Wang, H. and Alkhalifah, T. (2018). Microseismic imaging using a source function independent full waveform inversion method. *Geophysical Journal International*, 214:46–57.

Warpinski, N. (2009). Microseismic monitoring: Inside and out. *Journal of Petroleum Technology*, 61(11):80–85.

Wehling-Benatelli, S., Becker, D., Bischoff, M., Friederich, W., and Meier, T. (2013). Indications for different types of brittle failure due to active coal mining using waveform similarities of induced seismic events. *Solid Earth Discussions*, 5:655–698.

Xie, Y. and Gajewski, D. (2017). 5-D interpolation with wave-front attributes. *Geophysical Journal International*, 211:897–919.

Xie, Y. and Gajewski, D. (2018). 3D wavefront attribute determination and conflicting dip processing. *Geophysics*, 83:V325–V343.

Yang, P. and Gajewski, D. (2020). An optimized workflow for source localization and joint velocity inversion using wavefront attributes. *American Geophysical Union, Fall Meeting*, pages S063–0008.

Yang, P. and Gajewski, D. (2021a). An optimized workflow for source localization and joint velocity inversion using wavefront attributes. *in Proceedings of the 82nd EAGE Conference and Exhibition*, pages 1–5.

Yang, P. and Gajewski, D. (2021b). High-resolution time reverse imaging for microseismic event location. *in Proceedings of the 82nd EAGE Conference and Exhibition*, pages 1–5.

Yang, P. and Gajewski, D. (2022). Seismic source localization with time-reversal and maximum amplitude path for sparse and small-aperture acquisitions. *Geophysics*, 87:KS113–KS123.

Yang, P., Gajewski, D., and Xie, Y. (2023). Gaussian-weighted crosscorrelation imaging condition for microseismic source localization. *Geophysics*, 88:L65–L78.

Yang, P., Gajewski, D., and Xie, Y. (2024). Integrated wavefront tomography and time reversal for source location, excitation time, and velocity model building. *Geophysics*, under review.

Yang, P., Sergius, D., Gajewski, D., and Znak, P. (2019). Improving focusing and estimation of excitation time for passive seismic events: Sparse and limited-aperture data examples. *SEG Technical Program Expanded Abstracts*, pages 3091–3095.

Zhang, H. and Thurber, C. (2003). Double-difference tomography: The method and its application to the Hayward Fault, California. *Bulletin of the Seismological Society of America*, 93:1875–1889.

Zhang, X., Zhang, M., and Tian, X. (2021). Real-time earthquake early warning with deep learning: Application to the 2016 M 6.0 Central Apennines, Italy Earthquake. *Geophysical Research Letters*, 48:e2020GL089394.

Zheng, Y., Wang, Y., and Chang, X. (2018). 3D forward modeling of upgoing and downgoing wavefields using Hilbert transform. *Geophysics*, 83:F1–F8.

Zhou, R., Huang, L., and Rutledge, J. (2010). Microseismic event location for monitoring CO2 injection using double-difference tomography. *The Leading Edge*, 29:183–211.

Červený, V. (2001). *Seismic ray theory*. Cambridge Univ. Press.

Publications

Journal papers

- Yang, P., Gajewski, D., and Xie, Y. (2024). Integrated wavefront tomography and time reversal for source location, excitation time, and velocity model building. *Geophysics*, under review

- Yang, P., Gajewski, D., and Xie, Y. (2023). Gaussian-weighted crosscorrelation imaging condition for microseismic source localization. *Geophysics*, 88:L65–L78

- Yang, P. and Gajewski, D. (2022). Seismic source localization with time-reversal and maximum amplitude path for sparse and small-aperture acquisitions. *Geophysics*, 87:KS113–KS123

Expanded abstracts

- Yang, P. and Gajewski, D. (2021b). High-resolution time reverse imaging for microseismic event location. *in Proceedings of the 82nd EAGE Conference and Exhibition*, pages 1–5

- Yang, P. and Gajewski, D. (2021a). An optimized workflow for source localization and joint velocity inversion using wavefront attributes. *in Proceedings of the 82nd EAGE Conference and Exhibition*, pages 1–5

- Yang, P. and Gajewski, D. (2020). An optimized workflow for source localization and joint velocity inversion using wavefront attributes. *American Geophysical Union, Fall Meeting*, pages S063–0008

- Yang, P., Sergius, D., Gajewski, D., and Znak, P. (2019). Improving focusing and estimation of excitation time for passive seismic events: Sparse and limited-aperture data examples. *SEG Technical Program Expanded Abstracts*, pages 3091–3095

Acknowledgments

- I'm truly thankful to **Prof. Dr. Dirk Gajewski** for the invaluable opportunity to be a part of his team at the University of Hamburg. His insightful vision and passion for passive seismic applications, reverse modeling, wavefront attributes, and wave focusing have been a constant source of inspiration, shaping the core of my work. Prof. Gajewski's guidance has not only fostered my understanding of scientific writing and oral reporting, but also left a lasting impact on my research.

- I thank **PD Dr. Claudia Vanelle** for improving my presentation style, motivation, figures, and font size, as well as insightful guidance on focusing on the physical meaning of my work, significantly impacted the quality of my research.

- I acknowledge **Yujiang Xie** for consistently providing support, attentively considering my ideas, and dedicating considerable time to engaging in detailed discussions about intricate technical details.

- I thank **Dr. Pavel Znak** for his assistance, engaging in motivating and philosophical conversations, and encouraging my ideas to enhance the stability and reproducibility of time reversal and wavefront tomography.

- I thank **Lei Li** for generously sharing the HAMNET data and engaging in constructive discussions on microseismic localization algorithms.

- I thank **Stefan Knispel** for his assistance in enhancing my writing materials for his kindness in life.

- I thank **Dr. Jan Walda**, **Dr. Sergius Dell**, **Yan Yang**, and **Dr. Ekkehart Tessmer** for offering programming support during my early days in Hamburg.

- I'm grateful to **Dr. Alexander Bauer** and **Dr. Leon Diekmann** for sharing his experience in determining wavefront attributes and offering valuable insights into wavefront tomography in passive seismics.

- Next, I would like to thank Society of Exploration Geophysicists (SEG) for providing the Marmousi-ii model, the salt velocity model, and the data. I also appreciate China Scholarship Council (CSC) and the University of Hamburg, and the sponsors of the Wave Inversion Technology (WIT) Consortium for partial funding of my Ph.D. studies.

The most significant support throughout this journey has come from my dearest family members, particularly my sister, **Fen Yang**, and my Father, **Chenggang Yang**. Their unwavering love, understanding, and patience have been invaluable to me. I extend my heartfelt gratitude to them for their continuous support and encouragement.

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation mit dem Titel Overcoming challenges in time reversal for passive seismic source localization selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel – insbesondere keine im Quellenverzeichnis nicht benannten Internetquellen – benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass ich die Dissertation oder Teile davon vorher weder im In- noch im Ausland in einem anderen Prüfungsverfahren eingereicht habe und die eingereichte schriftliche Fassung der auf dem elektronischen Speichermedium entspricht.

Hamburg, den 26. Januar 2024

Peng Yang