WAVEFIELD RECONSTRUCTION USING SEISMIC INTERFEROMETRY AND COMPRESSIVE SENSING

by

Patipan Saengduean

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Geophysics).

Golden, Colorado

Date _____

Signed: _____

Patipan Saengduean

Signed: _____

Dr. Roel Snieder
Thesis Advisor

Signed: _____

Dr. Michael B. Wakin
Thesis Advisor

Golden, Colorado

Date _____

Signed: _____

Dr. Paul Sava
Department Head
Department of Geophysics

ABSTRACT

Seismic data acquisition and processing are essential steps for seismic exploration, the determination of deep earth structure, and subsurface monitoring. Common problems in seismic data acquisition are missing or unusable receivers, and a low signal-to-noise ratio of recorded signals. These problems apply to seismic interferometry, a passive-source technique to estimate inter-receiver wavefields. In this thesis, I propose novel methods that use seismic interferometry and compressive sensing to mitigate these problems.

I use inter-source interferometry to estimate body waves that propagate between two earthquakes from waves that are recorded at a surface receiver array. Inter-source interferometry, which turns an earthquake into a virtual receiver, allows receivers to be virtually placed at earthquake locations (e.g. inside volcanoes and subduction zones) that are normally inaccessible for receiver installation. To accurately estimate inter-source body wavefields, I derive a criterion for how densely receivers must be spaced to retrieve the waves that propagate between earthquakes. Using waves recorded near the San Andreas Fault, I estimate the inter-earthquake wavefields.

The problem of missing or unusable receivers is important for receiver arrays that do not regularly sample the wavefield at the surface. I propose a compressive-sensing-based multi-source wavefield reconstruction to alleviate the problem of missing or unusable receivers. Using the Fourier and Curvelet domains for sparse transforms, I show that a multi-source method, which reconstructs correlated wavefields at the locations of unavailable seismometers from correlograms of all available virtual sources, improves wavefield recovery compared to single-source reconstruction, which uses a correlogram from a single virtual source. I show successful applications of multi-source reconstruction for wavefield-recovery improvement over single-source reconstruction using synthetic data on linear and areal arrays and present a data example using distributed acoustic sensing data. The methodology is, in principle, applicable to active source seismic surveys.

Also, I develop weighted compressive sensing to mitigate the imprint of noise on the reconstructed wavefield, and to impose a priori information about the nature of the recovered wavefields. Wavefields obtained from seismic interferometry may contain spurious arrivals. Using the Fourier basis as a sparse transform, I show that weighted compressive sensing can suppress these spurious arrivals.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank the Development and Promotion of Science and Technology Talents Project (a scholarship from Thai Royal Government) that supported me financially for my PhD study at the Colorado School of Mines and the PhD projects included in this thesis.

I would like to acknowledge my thesis advisor (Dr. Roel Snieder) and co-advisor (Dr. Michael Wakin) who supported me academically throughout my PhD experiences at Mines and helped me with life problems I encountered during my PhD period. I also thank my colleagues from Center for Wave Phenomena and the department of Geophysics at Mines who shared their useful academic discussions and experiences with me as well as spend leisure time together with me for extracurricular activities.

I thank my family members, particularly my parents, who supported me both financially and spiritually during my PhD period at Mines. I thank Jihyun Yang, Harpreet Sethi, and Manuel Alejandro who always helped and listened to my life problems and worries during my PhD period.

Finally, I would like to thank Dr. Ilya Tsvankin, Dr. Ebru Bozdag, Dr. Stephen Pankavich, and Dr. Gabriel Walton for being part of my thesis committee and for useful discussions and suggestions to improve and complete my research and my thesis.

To my mom Pratumma and dad Rattana,

thank you for believing in me that I can complete my long-period studies abroad.

CHAPTER 1

GENERAL OVERVIEW

Seismic data acquisition is an essential part of seismic signal processing workflow and seismic interpretation. One acquires seismic signals by deploying seismic receiver arrays and these arrays can be sparse, dense, regular, or irregular. For practical reasons, one usually installs receivers on the Earth's surface because the installation of receivers in the subsurface is difficult and expensive. In addition, receiver arrays are often sparse and irregular because of problems such as inoperative receivers and restricted locations for receiver installation. Parts of acquired seismic signals can also be low in quality (low signal-to-noise ratio (SNR)) due to noise contamination of the signals or receiver malfunction during the array deployment. These issues of receiver installation and low-quality seismic signals appear in the forms of irregular random missing traces and low-SNR seismic traces.

One can mitigate the difficulty and expense of receiver installation in the subsurface using seismic redatuming techniques and seismic signals recorded from surface receivers. This involves redatuming techniques based on methods that include the Kirchhoff integral (Wapenaar, 1993), wave-equation based redatuming (William A. Schneider et al., 1995), amplitude correction (Tegtmeier et al., 2004), interferometric redatuming (Wapenaar et al., 2010; Xiao and Schuster, 2006), and recently, Marchenko imaging which is an inverse-scattering method (Diekmann and Vasconcelos, 2021; Kiraz et al., 2021; Rose, 2001; Wapenaar et al., 2014). One can apply these redatuming methods to move the positions of sources and receivers from the surface to the subsurface.

Irregularity of seismic arrays is another issue of seismic data acquisition that may affect the subsequent seismic signal processing steps. For example, seismic imaging including the Kirchhoff migration requires regular seismic profiles to reduce post-migration noise and enhance subsurface imaging (Al-Gain et al., 2020; Cao et al., 2018; Poole and Herrmann, 2007). Irregular seismic shots also affect removal of ground roll and noise reduction (Mann and Emanuel, 2006). Regularizing seismic profiles can also enhance the alignment of reflection events and the SNR of seismic imaging (Chopra and Marfurt, 2013). In addition, parts of acquired seismic signals from seismic arrays can be unreliable or of low quality for subsequent signal processing steps because the signals are noisy or corrupted by unexpected factors such as an incorrect orientation of receiver installation. Thus, regularization or reconstruction of seismic wavefields is required to alleviate the limitations of irregular and unreliable seismic records.

Seismic data reconstruction helps handling these issues by filling the data gaps and replacing the low-SNR records with interpolated traces before subsequent processing and interpretations. Different techniques of

seismic data reconstruction have been performed to interpolate sampled data. There are 5 main categories of these interpolation methods: wave-equation-based methods, linear prediction filter methods, rank-reduction methods, deep-learning approaches, and mathematical-transform methods (Liu et al., 2019; Wang et al., 2016). Wave-equation-based methods use migration and de-migration operators to interpolate sampled data, but these methods carry a heavy computational cost (Ronen, 1987). Linear prediction filter methods interpolate seismic records, assuming the data are the superposition of linear or quasi-linear events in the f-k domain. To interpolate sampled data, the prediction filter methods need a local window to ensure the linearity of complex and curved events (Naghizadeh and Sacchi, 2007; Spitz, 1991; Wang and Lu, 2017). Rank-reduction methods assume that the data are of low rank to interpolate sampled seismic data (Trickett et al., 2010). The rank-reduction techniques can predict NMO corrected data and interpolate curved events with the decrease in the flexibility of the techniques (Ma, 2013). The missing traces and noise in sampled seismic data also increase the rank of the data (Trickett et al., 2010). Interpolation methods using deep learning (e.g. generative adversarial networks or GANs) require large training data sets and usually uses synthetic data for the training because the acquisition of enough field seismic data is expensive and difficult. Training of GANs with synthetic data leads to low-quality seismic data reconstruction due to the difference in the features of test and training data. The bias between the two data sets increases with the increases in the difference of the two data sets (Alwon, 2018; Wang et al., 2019). GANs can interpolate strongly under-sampled seismic data but the technique requires some sampled traces in the training (Siahkoohi et al., 2018). The interpolation of complex-structure seismic data using conditional GANs may also give a low resolution and non-geological artefacts (Oliveira et al., 2018).

Interpolation of seismic signals using mathematical transform-based approaches is more flexible and widely used to reconstruct signals from irregular and fewer signal samples. The methods transform the data into a domain where the data is sparse and use sparse-recovery algorithms for the data reconstruction. The method can handle irregularly sampled data. The interpolation has been applied using different transforms, including the Fourier transform (Abma and Kabir, 2006; Gao et al., 2013), Radon transform (Kabir and Verschuur, 1995; Wang et al., 2010), wavelet transform (Pawelec et al., 2019; Yu et al., 2007), seislet transform (Gan et al., 2015; Liu et al., 2016a), dreamlet transform (Wang et al., 2014), and curvelet transform (Herrmann et al., 2008; Wang et al., 2015; Yang et al., 2012). A sparse-recovery algorithm that combines different transforms also improves the efficiency and performance of the data reconstruction (Kim et al., 2015). In my studies, I use a reconstruction method called compressive sensing to retrieve seismic profiles from few and irregular seismic traces, addressing the limitations of the irregularity of seismic receiver arrays due to missing or inoperative receivers.

1.1 Seismic Signal Recovery using Compressive Sensing

The Nyquist-Shannon sampling theorem requires the sampling rate of a signal to be twice the maximum frequency of the signal, the Nyquist frequency; signals are uniformly sampled at the Nyquist frequency. Candès et al. (2006c) and Donoho (2006) introduced a tranform-based method for signal recovery called compressive sensing (CS) as a new sampling/sensing paradigm that violates the traditional data acquisition methods which always follow the Shannon sampling theorem as CS can recover signals using fewer samples than those used by the common sampling methods. CS has been applied to several fields including compressive imaging that include radio interferometry (Vijay Kartik et al., 2017), radar imaging (Gurbuz et al., 2009; Patel et al., 2010), image processing (Li and Yang, 2011), medical applications (Liu et al., 2015; Otazo et al., 2015), and communication systems (Liu et al., 2017; Sharma et al., 2016). CS exploits the sparsity or compressibility of signals to recover the signals and requires irregularly sampled measurements for accurate signal reconstruction (Candès and Wakin, 2008).

The main applications of CS in seismology are data acquisition, seismic data reconstruction and denoising, sparsity-promoting based migration, and seismic imaging and inversion. The acquisition of seismic data using dense arrays is expensive, and in practice, some receivers in arrays can be inoperative or restricted for installation, resulting in irregular arrays. CS makes it possible to reconstruct dense and regular seismic data using records from fewer sensors in an irregular array, increasing the efficiency of seismic data acquisition (Mosher et al., 2014). Mansour et al. (2012) propose a randomized marine-data acquisition scheme that utilizes the sparse-promoting recovery and reduces overall time to conduct marine surveys. Pawelec et al. (2019) and Jiang et al. (2019b) also demonstrate the CS reconstruction of land seismic profiles from fewer randomized samples that are comparable to the dense seismic profiles. Using CS, Herrmann et al. (2008) and Hennenfent et al. (2010) demonstrate the interpolation and denoising of exploration seismic profiles. In addition, Boehm et al. (2016) illustrate the memory-requirement reduction in 3D adjoint full-waveform inversion, and Li et al. (2012) obtain full-waveform inversion comparable to that from conventional methods, while reducing the computational cost. The wavefield reconstruction using CS helps improve the resolution and robustness of surface-wave tomography (Zhan et al., 2018).

1.2 Seismic Interferometry and its Connection to Compressive Sensing

Common problems of seismic data acquisition include (1) difficult or restricted locations for seismometer deployment, (2) missing or inoperative seismometers, and (3) low quality (low-SNR) seismic signals recorded by seismometers. These issues apply to subsequent signal processing techniques including seismic interferometry (SI), which is a method to estimate the wavefield (the Green's function) that accounts for wave propagation between seismometers. Each seismometer can act as a virtual source. SI has been developed

for several applications to investigate and image the subsurface using ambient noise (Draganov et al., 2007; Shapiro et al., 2005), traffic and production noise (Nakata et al., 2011; Vasconcelos and Snieder, 2008b), and earthquake or active sources (Matzel et al., 2016, 2017). The principles of SI have been explained by several review papers (Curtis et al., 2006; Larose et al., 2006; Snieder and Larose, 2013a; Wapenaar et al., 2010).

Even though one can apply SI to recover the wavefield between seismometers using deconvolution (Pianese et al., 2018; van Dalen et al., 2015; Vasconcelos and Snieder, 2008b), cross-coherence (Nakata et al., 2011; Prieto et al., 2009), or convolution (Curtis and Halliday, 2010; Entwistle et al., 2015), applications of SI have been usually based on using cross-correlation (Asano et al., 2017; Miyazawa et al., 2008; Mordret et al., 2010; Shapiro et al., 2005). Passive seismic exploration is an application of cross-correlation SI that utilizes SI to retrieve inter-receiver wavefields. Receivers used for passive surveys are usually deployed on the Earth's surface because the installation of subsurface receivers is difficult and expensive. Rather than using seismic redatuming techniques, Curtis et al. (2009) have developed a version of SI called inter-source interferometry that addresses the limitation of installing subsurface receivers by estimating inter-source wavefields, where the locations of sources may be in the subsurface.

Using inter-source interferometry, one can estimate the wavefields that account for the wave propogation between two seismic sources (e.g. earthquakes), one of which turns into a virtual receiver. Recovering inter-source wavefields has been used for many applications. For example, Matzel et al. (2016) and Morency and Matzel (2017) monitor the micro-seismic events via the observation of the change in the source-mechanism inversion of the Green's function retrieved from these micro events. Other applications include imaging redatumed reflection responses using drill-bit noise as virtual receivers (Liu et al., 2016b), crustal tomography (Shirzad et al., 2019), and localization of olivine wedge (Shen and Zhan, 2020). Previous applications of inter-source interferometry involve the cross-correlation of surface wavefields, measuring the dispersion of the waves. However, estimating body waves using interferometry requires stricter conditions than the requirements of surface-wave retrieval (Forghani and Snieder, 2010).

In Chapter 2, rather than recovering surface waves as a common application of SI, we apply inter-source interferometry for body-wave retrieval and investigate the requirements associated with recovery of inter-source wavefields when sources are located in the subsurface. Using a numerical model, we show that one can recover direct body wavefields using inter-source interferometry if one satisfies two conditions: (1) seismic waves are recorded at receivers located inside the stationary phase zone and (2) there are 4 receivers per wavelength inside the stationary phase location of the interferometric integral. The interferometric stationary phase zone is explained in Chapter 2. Satisfying these two conditions, we recover direct body waves that account for wave propagation between two earthquakes in the area of San Andreas Fault and show that we can use the retrieved direct body waves to estimate the width of the damage zone of the fault, which is

approximately 4 km.

Apart from the difficulty of installing subsurface receivers, seismic noise interferometry, which is an application of cross-correlation SI, typically requires much data storage to store long records of seismic noise that can be as long as 3 months (Nakata et al., 2015) and even 1 year (Lin et al., 2008) for passive seismic surveys. In addition, seismic receiver arrays for these surveys are often irregular because some receivers may be inoprative or missing and may record low-quality seismic signals. One can apply wavefield reconstruction to mitigate the problem of irregular array and to save data storage by discarding some parts of seismic records. For example, Jayne et al. (2022) use CS to reduce the storage required for cross-correlating long noise records by reducing the number of samples for cross-correlation in the Fourier domain. Wang et al. (2009), Hanafy and Schuster (2014), and Xu et al. (2018) use a technique called interferometric interpolation, that utilizes a matched filter and virtual source gathers provided by SI, to extrapolate and interpolate missing marine seismic profiles. Since Zhan et al. (2018) demonstrate the reconstruction of surface wavefields generated from a source on a dense receiver array, we can apply CS to reconstruct interferometric surface wavefields retrieved from SI on seismic receiver arrays when some receivers are missing or inoperative.

In Chapter 3, we focus on using CS to reconstruct interferometric surface wavefields (cross-correlation profiles) when gaps of the correlation wavefields appear on the profiles because of inoperative or absent receivers. We propose and develop an alternative wavefield-recovery technique called multi-source wavefield reconstruction, where we apply CS to recover the missing correlation profiles recorded from a linear receiver array using the correlation traces provided from all available virtual sources. Using SNR as a diagnostic for wavefield-recovery comparison, we show that recovering interferometric wavefields using correlation profiles from all available virtual sources is more accurate than using correlation profiles from an available virtual source to reconstruct the wavefields.

The improvement of wavefield reconstruction using information from multiple sources is beneficial to passive seismic studies. One can apply our multi-source method to existing linear sensor arrays, where some sensors are missing or inoperative. A recent application of noise interferometry in seismic exploration is to image and monitor subsurface using signals collected from optical fiber networks using distributed acoustic sensing (DAS) technology. Commonly, seismic signals collected using DAS cables are highly sensitive to the orientation of the cables and the direction of incident wavefields (Bakku et al., 2014; Martin et al., 2021; Shragge et al., 2021). Thus, parts of the collected signals can be inadequate in signal quality and may be discarded from subsequent signal processing steps. In Chapter 4, we apply our multi-source wavefield reconstruction to DAS signals collected from an optical fiber located in the area of Perth, Australia. We show that the multi-source method, which uses correlation profiles from all available virtual sources, is applicable to field DAS data. The method also shows an improvement in recovering interferometric wavefields, compared

to the wavefiled reconstruction using a correlation profile from an individual virtual source.

In Chapters 3 and 4, we show the application of our multi-source wavefield reconstruction of seismic signals collected from linear receiver arrays. In the appendix, we apply our multi-source method to reconstruct interferometric wavefields collected from an areal receiver array. We show that the multi-source method is applicable to reconstruct wavefields collected from areal arrays. Using a numerical model, we demonstrate that similar to linear receiver array examples, the multi-source method gives more accurate reconstruction of interferometric wavefields, compared to wavefield reconstruction using correlated wavefields from an available virtual source.

We show in Chapters 3 and 4 and the appendix that one can use interferometry and wavefield reconstruction techniques to mitigate the problems of difficult receiver installation and missing/inoperative sensors. Another problem of seismic data acquisition is low-quality (low SNR) signals due to noise assimilation and signal sensitivity. For example, signals collected from DAS fibers can be sensitive to weather conditions, where colder air and ground temperatures can vary the velocity of surface waves monitored by near-surface DAS sensors (Yang et al., 2021). One needs to filter these unreliable signals for subsequent processing steps. In SI applications, some interferometric traces may be unreliable because these traces include interferometric noise, which is the wave arrivals prior to the surface waves retrieved by SI. One can reduce this interferometric noise by satisfying the sampling criteria of source distribution provided by Fan and Snieder (2009), or for noise interferometry, one can stack cross-correlated windows of very long seismic noise records.

In Chapter 5, we investigate how we can use these unreliable signals for wavefield reconstruction. We develop a weighted compressive sensing wavefield reconstruction technique that helps reduce interferometric noise. Using a cross-correlation profile that includes both traces with and without interferometric noise, we show that one can apply the weighted method to reconstruct a correlation profile without interferometric noise. The quality of wavefield reconstruction using our weighted algorithm depends on the weight and the degree of unreliability of signals (more details in Chapter 5). Finally, we give a brief conclusion of the previous chapters in Chapter 6. We also provide potential research directions and applications of inter-source interferometry, multi-source wavefield reconstruction, and weighted compressive sensing. These applications include deep-earth investigations, noise reduction, DAS data compression, and interpolation for seismic exploration.

CHAPTER 2

INTER-SOURCE INTERFEROMETRY OF SEISMIC BODY WAVES: REQUIRED CONDITIONS AND EXAMPLES

Reproduced with the permission for The Journal of Pure and Applied Geophysics.[1]

Patipan Saengduean[2], Morgan P. Moschetti[3], and Roel Snieder[4]

2.1 Abstract

Seismic interferometry is widely applied to retrieve wavefields propagating between receivers. Another version of seismic interferometry, called inter-source interferometry, uses the principles of seismic reciprocity and expands interferometric applications to retrieve waves that propagate between two seismic sources. Previous studies of inter-source interferometry usually involve surface-wave and coda-wave estimations. We use inter-source interferometry to estimate the P-waves propagating between two sources rather than the estimation of surface waves and coda waves. We show that the recovered arrival times are dependent on the accuracy of the earthquake catalog of the two sources. Using inter-source interferometry, one can recover the waveform of the direct body waves and potentially reconstruct the waveform of coda waves, depending on the source-receiver geometry. The retrieval of these waveforms is accurate only when the wavefield is sampled with approximately 4 receivers per wavelength in the stationary phase zone. We show that using only receivers inside the stationary phase region for inter-source interferometry introduces the phase error of approximately 0.3 radians. In our study, we show an example of the P-wavefield reconstruction between two earthquakes using the seismic records from an array along San Andreas Fault. The retrieved P waves give a qualitative estimation of the thickness of the low-velocity zone of San Andreas Fault of approximately 4 km.

2.2 Introduction

Seismic interferometry is a technique to estimate the Green's function or the wavefield that accounts for wave propagation between receivers. The inter-receiver technique has been developed for many applications using ambient noise (Draganov et al., 2007; Shapiro et al., 2005), traffic and oilfield production noise (Miyazawa et al., 2008; Nakata et al., 2011), and earthquakes or active sources. The principles of seismic interferometry are explained in several review papers (Curtis et al., 2006; Larose et al., 2006; Snieder and Larose, 2013b; Wapenaar et al., 2010).

[1]Reproduced with the permission of Pure Appl. Geophys., 178, 3441-3460. See Appendix for the permission.
[2]Graduate student, primary researcher, and author at Department of Geophysics, Colorado School of Mines
[3]Research Geophysicist at U.S. Geological Survey, Golden, CO
[4]Supervisor at Department of Geophysics, Colorado School of Mines

In general, one can retrieve the Green's function between receivers using cross-correlation, deconvolution, and cross-coherence of the signals recorded at receivers (Nakata et al., 2011; Snieder et al., 2009). Although applications of seismic interferometry are usually based on cross-correlation (Asano et al., 2017; Miyazawa et al., 2008; Mordret et al., 2010; Shapiro et al., 2005), some applications also use deconvolution (Nakata et al., 2013; Pianese et al., 2018; van Dalen et al., 2015; Vasconcelos and Snieder, 2008a,b), cross-coherence (Nakata et al., 2011; Prieto et al., 2009), and convolution (Curtis and Halliday, 2010; Entwistle et al., 2015). These data processing methods for Green's function retrieval have diverse advantages and disadvantages (Snieder et al., 2009). One has to select the process that is best suited to particular data and research requirements. In theory, cross-correlation of ambient seismic noise interferometry suppresses strong amplitudes of earthquakes. In this study, we implement cross-correlation for Green's function retrieval because, unlike in the case of noise interferometry, we use real earthquakes, and the cross-correlation correctly handles the strong amplitudes of the waves that are excited by these earthquakes.

Recovering the Green's function using inter-receiver interferometry, in particular with the retrieval of surface waves from ambient noise cross-correlation, has become an accepted technique (Shapiro et al., 2005). For practical reasons, the receivers used are usually installed on the Earth's surface. A version of seismic interferometry, called inter-source interferometry (Curtis et al., 2009), addresses this limitation because it yields the waves that propagate between two sources, which may be located in the subsurface.

Inter-source interferometry, or the virtual receiver method (VRM), is a technique to estimate the Green's function that accounts for wave propagation between two earthquakes, one of which acts as a virtual receiver. The Green's function retrieved from the inter-source technique is the wavefield measured through the dynamic strain induced by the wavefield at one of the earthquakes excited by the other earthquake (Curtis et al., 2009). Even though the theory of inter-source interferometry for Green's function retrieval requires receivers on a closed surface surrounding the earthquakes, the dominant contribution of the retrieved Green's function comes from the receivers located inside the stationary phase zones of the pair of sources (Curtis et al., 2009; Snieder, 2004).

Inter-source interferometry as introduced by Curtis et al. (2009) has been applied in several studies. One application involves the monitoring of micro-seismic events by observing the change in the source mechanism inversion of the Green's function retrieved from these events (Matzel et al., 2016; Morency and Matzel, 2017). Other applications of inter-source interferometry include the improvement of crustal structure tomography using inter-event interferometry (Shirzad et al., 2019), the localization and identification of the geometry of the metastable olivine wedge (Shen and Zhan, 2020), and the determination of the shear-wave velocity in different source clusters (Eulenfeld, 2020). In addition, Schuster (2009) uses interferometry to convert VSP data into surface seismic data and thus extracts the waves that propagate between active sources

at the surface. Curtis and Halliday (2010) use inter-source interferometry as a part of a new form of interferometry called source-receiver interferometry. Poliannikov et al. (2012) image a subduction slab using virtual responses between earthquakes located inside the slab. Liu et al. (2016b) turn non-transient sources such as drill-bit noise into virtual receivers, imaging the redatumed reflection responses close to targeted areas.

Previous studies of inter-source interferometry usually estimate the Green's function using the cross-correlation of surface waves, measuring the dispersion of the waves. However, the estimation of body waves using seismic interferometry involves stricter requirements on the source-receiver geometry than the extraction of surface waves (Forghani and Snieder, 2010; Nakata et al., 2011). Since inter-source interferometry can yield the waves that propagate between earthquakes, this technique has the potential to extract the body waves that propagate between these earthquakes.

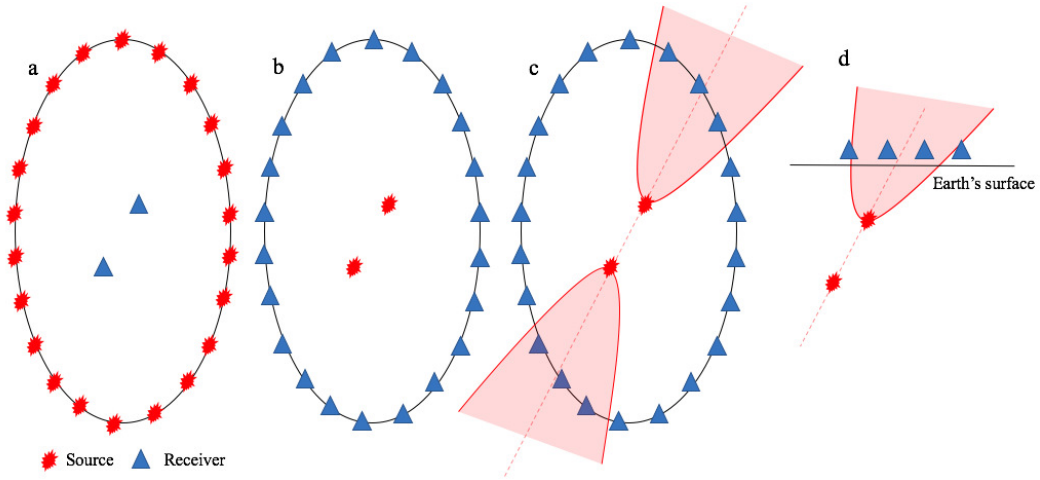


Figure 2.1 Cartoon of a) inter-receiver interferometry with noise or active sources on a closed surface, b) inter-source interferometry with receivers on a closed surface, c) inter-source interferometry with receivers in stationary phase zones (pink areas), and d) inter-source interferometry with surface receivers in one stationary phase zone.

In order to construct the Green's function using seismic interferometry, one has to satisfy the theoretical requirements on the source or reciever distribution. Fan and Snieder (2009) have derived sampling criteria for inter-receiver interferometry. Due to seismic reciprocity, which allows for the exchange of the source and receiver positions, the same criteria holds for the inter-source technique. However, noise sources in the inter-receiver interferometry tend to have better spatial distribution than that of receivers in inter-source interferometry (Figure 2.1a and Figure 2.1d). At the global scale, dominant noise sources (e.g., oceanic waves) are omnipresent, while smaller-scale noise sources (e.g., traffic and cultural noise) are concentrated at particular locations such as roads and construction sites. In inter-source interferometry, the reconstruction

quality of the retrieved waves depends on the spatial distribution of surface receivers. Thus, the applicability of inter-source interferometry depends on the geometry of the used receiver array (Figure 2.1d).

Instead of measuring the arrival time and the dispersion of surface waves, we seek to retrieve body waves that can be characterized by the following information: 1) the arrival time of the direct wave, 2) the waveform of the direct wave, and 3) the waveform of coda waves. In our study, we test to what extent we can recover this information; we also identify the required sampling criteria (i.e., sampling density and station locations) for the retrieval of this information using inter-source interferometry. We present the Green's function retrieval from the cross-correlation of 1) 2D synthetic acoustic waves excited from two sources in a homogeneous medium with an embedded scatterer and 2) earthquake seismograms recorded at an array deployed along San Andreas Fault. The retrieval uses inter-source interferometry and receivers located in the stationary phase zones.

We first introduce the basic equations of cross-correlation based interferometry and the data processing steps. We further demonstrate the Green's function retrieval using different stationary phase zones and sampling density for the synthetic model. Next, we show a field example of Green's function reconstruction of P waves and the qualitative estimation of the low-velocity zone thickness of the San Andreas Fault (SAF) using the reconstructed P waves. Finally, we explain the reason why we can recover parts of the waveform, yet cannot recover the arrival time and the full waveform of body waves using inter-source interferometry; we also discuss the application of the inter-source technique with field examples.

2.3 Theory of Cross-Correlation Based Interferometry

Using the Fourier convention $f(t) = \int F(\omega)e^{i\omega t}d\omega$ and assuming a far-field integration boundary oriented perpendicularly to wave propagation, Wapenaar and Fokkema (2006) have derived seismic interferometry for acoustic waves:

$$G(\mathbf{x}_A, \mathbf{x}_B, \omega) - G^*(\mathbf{x}_A, \mathbf{x}_B, \omega) \approx -\frac{2i\omega}{\rho c} \oint G(\mathbf{x}_B, \mathbf{x}, \omega) G^*(\mathbf{x}_A, \mathbf{x}, \omega) d^2\mathbf{x}, \tag{2.1}$$

where $G(\mathbf{x}_A, \mathbf{x}_B, \omega)$ is the frequency-domain representation of the Green's function that accounts for the wave propagation from $\mathbf{x}_B$ to $\mathbf{x}_A$, ρ is the mass density, c is the wave velocity, ω is the angular frequency, and the asterisk denotes the complex conjugation. The factor $i\omega$ corresponds to a differentiation in the time domain.

In the frequency domain, the wavefield $u(\mathbf{x}_A, \mathbf{x}, \omega)$ excited from a point source at $\mathbf{x}$ and recorded at $\mathbf{x}_A$ is the Green's function $G(\mathbf{x}_A, \mathbf{x}, \omega)$ convolved with the source-time function $W(\mathbf{x}, \omega)$. The cross-correlation of the wavefields measured at $\mathbf{x}_A$ and $\mathbf{x}_B$ is, in the frequency domain, given by

$$C_{BA} = u(\mathbf{x}_B, \mathbf{x}, \omega) u^*(\mathbf{x}_A, \mathbf{x}, \omega) = |W(\mathbf{x}, \omega)|^2 G(\mathbf{x}_B, \mathbf{x}, \omega) G^*(\mathbf{x}_A, \mathbf{x}, \omega). \tag{2.2}$$

Integrating the cross-correlation in equation (2) over a closed surface that includes uncorrelated sources on the closed surface surrounding the receivers (Snieder et al., 2007) gives

$$\oint C_{BA} d^2\mathbf{x} = \langle |W(\mathbf{x},\omega)|^2 \rangle \oint G(\mathbf{x}_B,\mathbf{x},\omega) G^*(\mathbf{x}_A,\mathbf{x},\omega) d^2\mathbf{x}, \tag{2.3}$$

where $\langle |W(\mathbf{x},\omega)|^2 \rangle$ is the average of the spectrum of the squared source-time functions. In practice, we carry out the integration by stacking all sources on the closed surface (Figure 2.1a). Inserting the integration of cross-correlation of equation (3) into equation (1) yields:

$$\big[G(\mathbf{x}_A,\mathbf{x}_B,\omega) - G^*(\mathbf{x}_A,\mathbf{x}_B,\omega)\big] \langle |W(\mathbf{x},\omega)|^2 \rangle \approx -\frac{2i\omega}{\rho c} \oint C_{BA} d^2\mathbf{x}, \tag{2.4}$$

where $G(\mathbf{x}_A,\mathbf{x}_B,\omega) - G^*(\mathbf{x}_A,\mathbf{x}_B,\omega)$ denotes the difference of the causal and acausal parts of the Green's function that accounts for wave propagation between receivers at $\mathbf{x}_A$ and $\mathbf{x}_B$. Equations (1)-(4) represent the theory of cross-correlation based inter-receiver interferometry. With source-receiver reciprocity, the equations also hold for inter-source interferometry as Curtis et al. (2009) show that one can estimate the same Green's function that accounts for the wave propagation between two sources at the same $\mathbf{x}_A$ and $\mathbf{x}_B$ locations, using the records at receivers on a closed surface (Figure 2.1b) or surface receivers located in the stationary phase zone (Figure 2.1c and Figure 2.1d). In our study, we define the width of the stationary phase zone as the region where the phase change of the interferometric integral compared to the point where the phase is at an extremum is a specified fraction of the dominant period. In our study, we use a quarter of the dominant period (the appendix (Section 2.9)). The width of the phase zone depends on the source-receiver geometry and is explained in the following sections.

As shown by Curtis et al. (2009), inter-source interferometry for moment tensor sources $\mathbf{M}^A$ and $\mathbf{M}^B$ at locations $\mathbf{x}^A$ and $\mathbf{x}^B$, respectively, gives an interferometric measurement

$$v = M^B_{ip} M^A_{mq} \partial^B_p \partial^A_q G_{im}(\mathbf{x}^B,\mathbf{x}^A) \, . \tag{2.5}$$

As shown by Aki and Richards (2002), the displacement generated by a moment tensor source $\mathbf{M}^A$ at $\mathbf{x}^A$ is given by

$$u^{mom}_i(\mathbf{x}) = M^A_{mq} \partial^A_q G_{im}(\mathbf{x}^B,\mathbf{x}^A) \, , \tag{2.6}$$

where G denotes the elastic wave Green's tensor. According to expression (2.5), inter-source interferometry thus extracts

$$v = M^B_{ip} \partial^B_p u^{mom}_i(\mathbf{x}^B) \, . \tag{2.7}$$

Since the moment tensor M^B_{ip} is symmetric, this can also be written as

$$v = M^B_{ip} \varepsilon^{mom}_{pi}(\mathbf{x}^B) \, , \tag{2.8}$$

where $\varepsilon_{pi}^{mom} = (1/2)\left(\partial_p u_i^{mom} + \partial_i u_p^{mom}\right)$ is the strain field. This means that inter-source interferometry gives the projection of the strain field at $\mathbf{x}^B$ onto the moment tensor $\mathbf{M}^B$ that is associated with elastic waves excited by a moment tensor source $\mathbf{M}^A$ at location $\mathbf{x}^A$.

2.4 Synthetic Model and Green's Function Retrieval

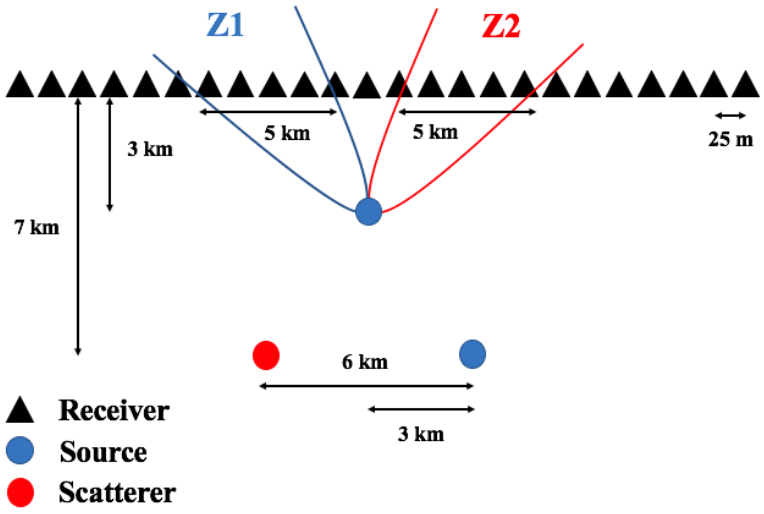


Figure 2.2 Cartoon of 2D synthetic model used to compute acoustic wavefields. The model includes two sources and a scatterer. The stationary phase zones for the direct and scattered waves propagated between the two sources are Z1 and Z2, respectively. The receiver separation is 25 m and is not drawn to scale.

In a 2D synthetic example, we compute acoustic wavefields in a homogeneous model with a single embedded scatterer. As shown in Figure 2.2, two seismic sources are located at depths of 3 km and 7 km with a horizontal separation of 3 km. The scatterer is located at a depth of 7 km with a horizontal separation of 3 and 6 km from the two sources, respectively. In our model, 801 receivers are located on the surface with a uniform separation distance of 25 m between receivers. The wave velocity of this medium is 4 km/s and the frequency range is 1-10 Hz; thus, the shortest wavelength of this model is 400 m. Figure 2.2 is a sketch of our synthetic model with two 5 km-wide stationary phase zones; Z1 and Z2 represent the zones for the direct wave between two earthquakes and the scattered wave from the scatterer, respectively.

We use the cross-correlation based method to construct the Green's function between two seismic sources. We first compute the cross-correlation of full wavefields excited from the two sources that are recorded at all receivers located in the two stationary phase zones. Next, we apply a Gaussian taper to the cross-correlated traces at the receivers before computing the integration of cross-correlations; the tapering minimizes the truncation artifacts resulting from the truncation of the integration interval (Burdick and Orcutt, 1979). The tapering employs a Gaussian window, $w(x) = e^{-(x-x_0)^2/2\sigma^2}$, where x is the receiver location away from

the center of the stationary phase zone (x_0) and σ is 40% of the width of the receivers array located within the stationary phase region. For a meaningful comparison, we use equation (4) to match the integrated correlogram with the direct forward-modelled wavefield between the two sources convolved with the source-time function. To recover the Green's function, one needs to take the time derivative of the integrated correlogram and deconvolve with the squared source-time function. In order to avoid a deconvolution, we follow equation (4) and compare the scaled time derivative of the correlation (the right hand side of equation (4)) with the Green's function's convolved with the power spectrum of the source-time functions (the left hand side of equation (4)). Because our comparison uses normalized waveform, we do not account for the constants 2, mass density, and wave velocity in equation (4).

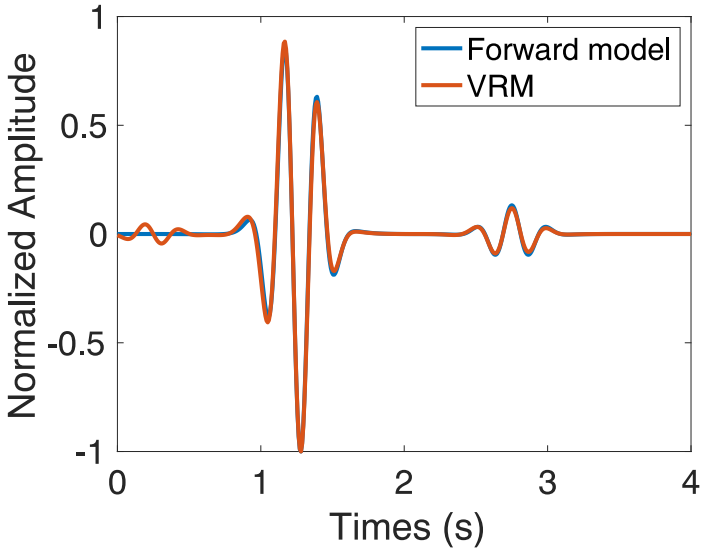


Figure 2.3 Comparison between the forward-modeled wavefield convolved with the source wavelet (blue) and the Green's function retrieved from the inter-source technique (VRM) with a time derivative correction (red), using all 801 receivers. The amplitude of the two traces are normalized.

The ability to accurately reconstruct the Green's function depends on 1) the location of the receiver array and 2) the sampling density of the array. We test the ability to reconstruct the Green's function using inter-source interferometry, given these two requirements. We first show the comparison between the forward-modeled and the interferometric wavefields when we use all 801 receivers on the surface for inter-source interferometry. Using all the receivers, the waveform extraction is accurate, except for an early arrival at 0.25 s (Figure 2.3). This early arrival at 0.25 s with an amplitude of approximately 2.5% of the direct arrival is the contribution from the cross-correlation of scattered waves and has the same arrival time for all receivers (Figure 2.4). Snieder et al. (2008) show that a full aperture of receiver array can eliminate the early arrival at 0.25 s. Since we use a limited aperture with receivers only at the surface, we cannot completely

eliminate this early arrival.

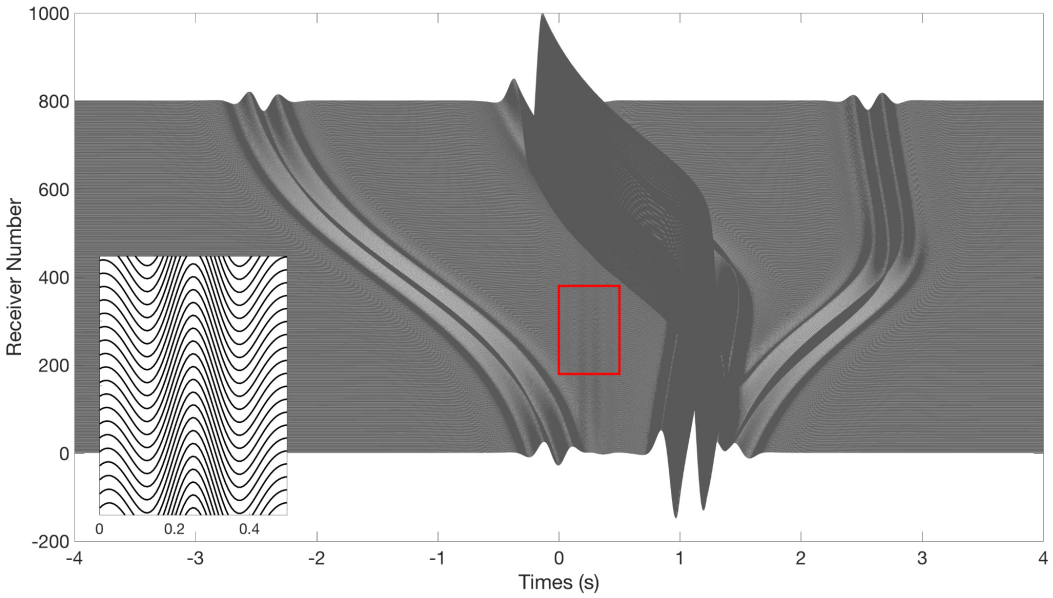


Figure 2.4 Cross-correlation of wavefields recorded at each individual receiver. The bottom-left plot magnifies the area in the red rectangle, showing weak waves that arrive consistently at 0.25 s at all receivers.

As the wavefield reconstructed by inter-source interferometry depends on the receiver array location, we test Green's function retrieval using cross-correlation of the wavefields recorded at receivers located in different stationary phase zones. Figure 2.5 and Figure 2.6 compare the forward-modeled wavefield and the interferometric Green's function retrieved using only receivers in the stationary phase locations, Z1 and Z2, respectively. Figure 2.5 illustrates that we can only retrieve the direct wave when we use receivers in the stationary phase zone (Z1) of the direct wave. By contrast, Figure 2.6 shows the same comparison, using stations located in the stationary phase zone (Z2) of the scattered wave, indicating that in this case, we can only retrieve the scattered wave. Notice from Figure 2.5 and Figure 2.6 that the limited aperture of the used receiver array located inside the stationary phase zone leads to a small phase distortion between the forward-modelled and interferometric wavefields.

We show in the appendix (Section 2.9) that the small phase distortion in Figure 2.5 is due to an integration over receivers that is limited to the stationary phase zone. We also show in the appendix that the stationary phase integral converges slowly in the sense that one needs to integrate over an interval that is much larger than the width of the stationary phase zone as defined earlier, and that an integration over the stationary phase zone–defined as the region where the maximum phase delay is a quarter of a period–leads to a phase error of about 0.273 radians. This phase error does not depend on the distance between the sources.

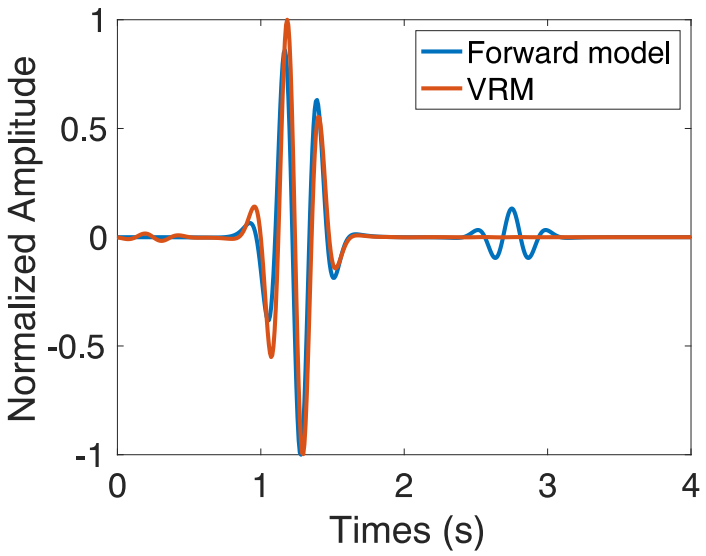


Figure 2.5 Comparison between the forward-modeled wavefield convolved with the source wavelet (blue) and the Green's function retrieved from the inter-source technique (VRM) with a time derivative correction (red), using receivers located in stationary phase region for inter-source path (Z1). The amplitudes of the traces are normalized.

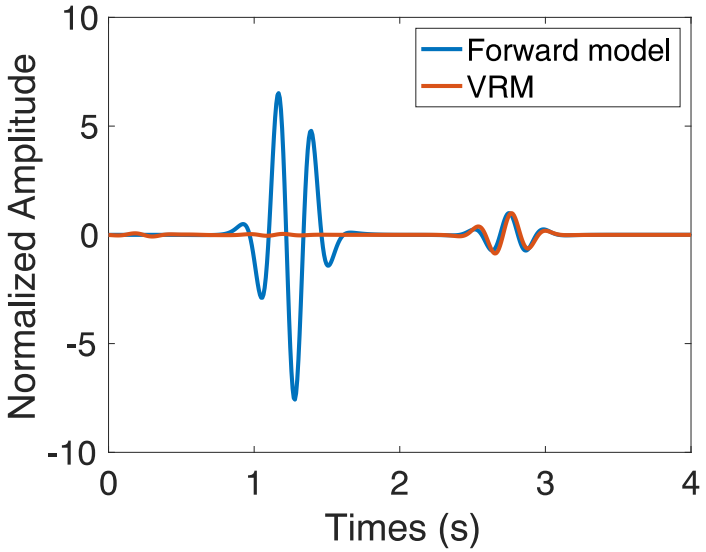


Figure 2.6 Comparison between the forward-modeled wavefield convolved with the source wavelet (blue) and the Green's function retrieved from the inter-source technique (VRM) with a time derivative correction (red), using receivers located in stationary phase zone for source-scatterer path (Z2). The amplitudes of the two traces are normalized using the amplitude of the scattered wave.

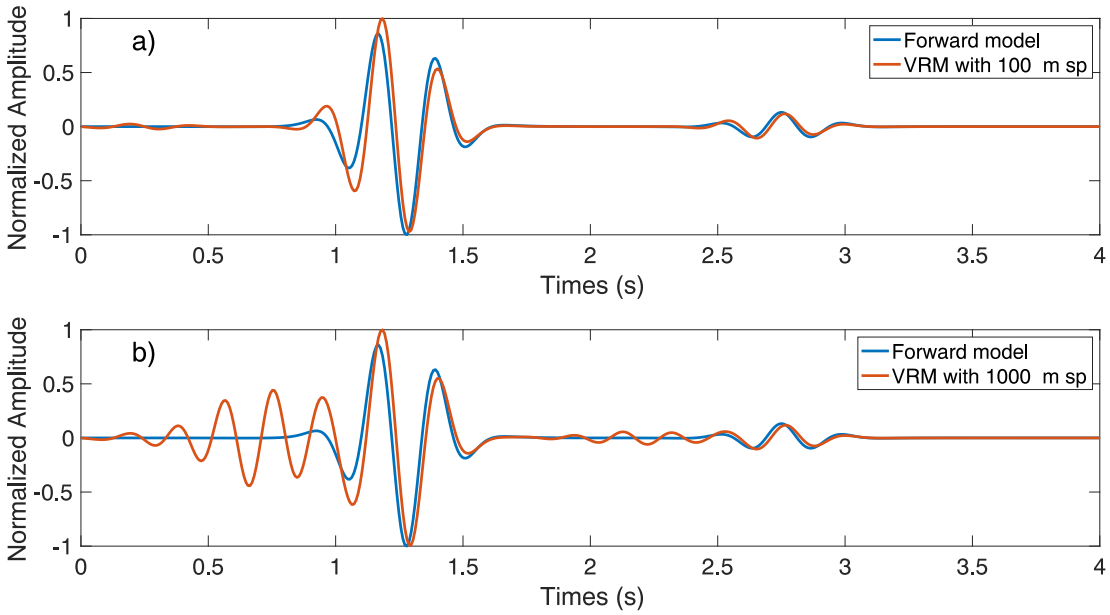


Figure 2.7 Comparison between the forward-modeled wavefield convolved with the source wavelet (blue) and the Green's function retrieved from the inter-source technique (VRM) with a time derivative correction (red), using receivers located in Z1 and Z2. The amplitude of the two traces are normalized. Figure 2.7a and Figure 2.7b have uniform receiver spacings of 100 and 1000 m, respectively, in the stationary phase zones.

In addition to the different receiver array locations, we also test the connection between the sampling density of the station array and the retrieved Green's function using inter-source interferometry by lowering the number of stations and thereby increasing the station spacing inside the two stationary phase zones (Z1 and Z2). The comparison of the forward-modeled and the retrieved traces in Figure 2.7 involves stations located in the two stationary phase zones (Z1 and Z2). Using 41 stations with a uniform station spacing of 100 m in each zone, we can recover both the direct and the scattered waves (Figure 2.7a). By contrast, Figure 2.7b shows the retrieved direct and scattered waves with oscillations prior to the arrival of these waves, using 5 stations in each zone with a uniform station spacing of 1 km. The direct and scattered waves in Figure 2.7 are well-reconstructed using the different station spacing because the width of the stationary phase locations is well-covered. In the case of Figure 2.7b, the retrieved signal is under-sampled away from the stationary phase regions as the station spacing is much larger than the smallest wavelength of this synthetic model (400 m). Thus, the oscillations prior to the direct and scattered waves are not completely eliminated, compared to the retrieved signal in Figure 2.7a. Similar to Figure 2.5 and Figure 2.6, Figure 2.7 shows a phase distortion between the forward-modelled and interferometric wavefields due to the limited aperture of the used receiver array located inside the stationary phase zones. We explain the detail of the phase shift in the appendix (Section 2.9). One can mitigate this phase distortion using all receivers on the

surface (Figure 2.3), but this is usually not possible with realistic surface array.

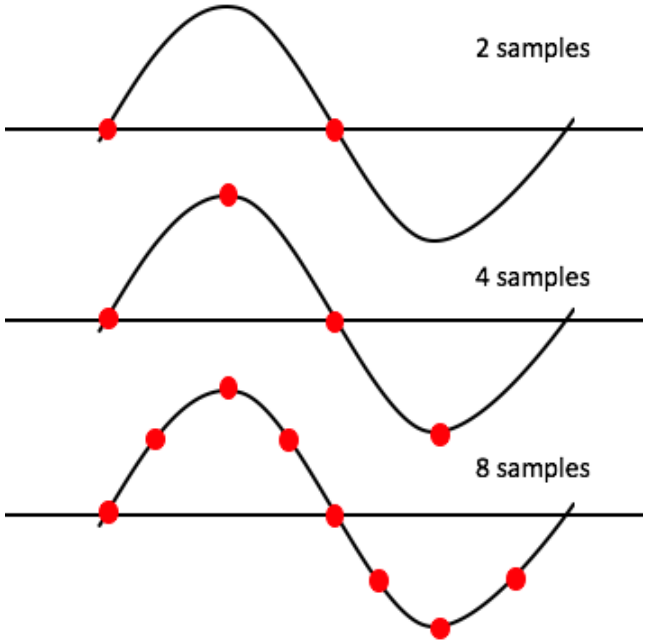


Figure 2.8 Different numbers of spatial samples for an integrated oscillatory function.

In this section, we showed that one can reconstruct the waveform of direct body waves, which accounts for the wave propagation between two sources, using inter-source interferomety and seismic records of receivers located in the stationary phase zones. The reconstruction is accurate only when the wavefield is sampled with sufficient spatial sampling density. For the direct summation of the cross-correlated wavefields over field array receivers, one requires 4 sampling points in a wavelength because two sampling points are not adequate to numerically integrate an oscillatory function when using a summation over the sampling points; as shown in the top function in Figure 2.8, these sampling points of an integrated oscillatory function might be located at zero crossings of the integrand and would thus erroneously suggest that the function vanishes. However, using techniques such as interpolations of the integrand prior to the summation of the integrand (Entwistle et al., 2015) or Filon-Trapizoidal rule for numerically integrating oscillatory integral functions (Tuck, 1967), one can reduce the number of required sampling points and only needs more than 2 sampling points in a wavelength. With only the direct integration of the cross-correlations over receivers, one needs at least 4 points per wavelength to numerically integrate such an oscillatory function, although even in that case the numerical quadrature might not be very accurate.

2.5 Green's Function Retrieval using Seismograms Recorded near the San Andreas Fault (SAF)

Section 2.4 shows that for the direct summation of the cross-correlated wavefields over array receivers, one requires at least 4 receivers located within the stationary phase locations to reconstruct the waveform of the direct body waves. Here, we exemplify the recovery of direct body waves using the inter-source method and seismic records from two earthquakes recorded at a seismic array inside the stationary phase locations of the two earthquakes.

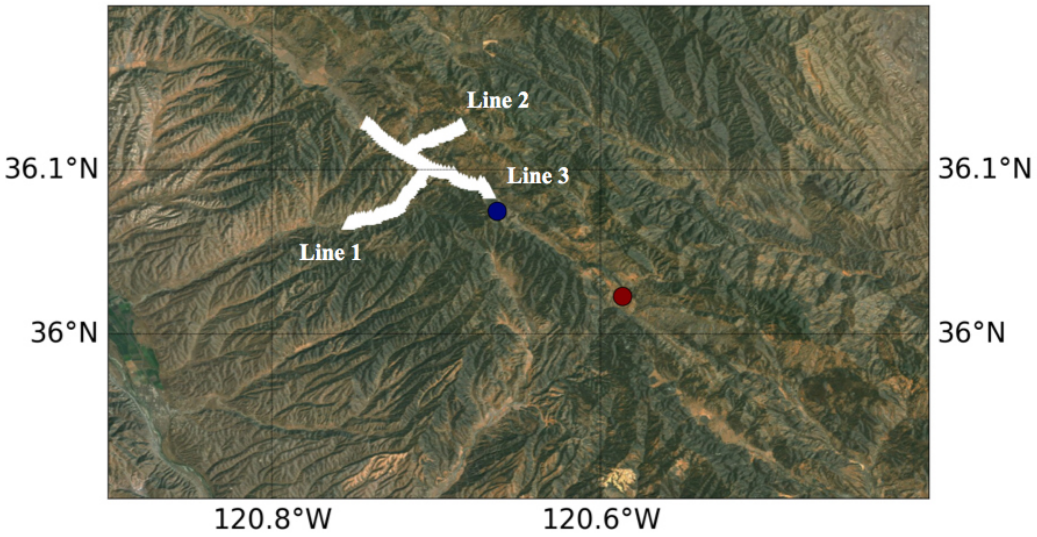


Figure 2.9 Map view of seismic line arrays (white triangles) along and across San Andreas Fault and earthquakes (circles). Blue and red circles represent the shallow and deeper earthquakes, respectively.

We select two earthquakes from the U.S. Geological Survey comprehensive earthquake catalog (Guy et al., 2015). The first earthquake ($M_d = 1.48$) occurred on July 11, 2014 at 6:56am with depth of approximately 1 ± 1.4 km. The second earthquake ($M_d = 2.47$) occurred on July 11, 2014 at 12:14pm with depth of approximately 4 ± 0.2 km. We use the seismograms of the two earthquakes recorded at an array installed in Peachtree Valley, San Andreas Fault (SAF). The array consists of 116 ZLand nodes, along 3 line arrays (Figure 2.9). For the interferometry, we use the seismic records of the two local earthquakes from the line array along the SAF (Line 3); seismograms of the earthquakes start at the origin times of the earthquakes from the catalog (Figure 2.10). The seismograms are bandpassed between 1-3 Hz to accommodate the station spacing of the array and the sampling criteria.

In this study, we perform inter-source interferometry using the waves arriving in the P-wave time windows. We first manually pick the arrival times of P and S waves from the shallow earthquake of each station. Then, we visually pick the P-wave arrival time of the deeper earthquake and determine the local P-wave velocity

from the picked P-wave arrivals; assuming straight ray paths between the sources and the surface receivers and using the moveout along the array (local P-wave velocity approximately 4.8 km/s). We assume that the wave-velocity ratio V_p/V_S is constant. Thus, we can estimate the arrival times of S waves for the deeper earthquake and separate the P- and S-wave time windows of the two earthquakes. The windows we use in our interferometric study are from the origin times of the earthquakes up to the S-wave arrivals. We then apply a Hanning taper of 2 periods on both ends of the P-wave signals. Here, all of the seismic stations of the SAF data set are visually located in the stationary phase zone of the two selected earthquakes (Figure 2.11). The station spacing of approximately 200 m is less than a quarter of the wavelength for the highest-frequency waves in the selected 1-3 Hz bandpass ($\lambda/4$ is approximately 400 m). The station spacing of the SAF data set thus satisfies the sampling requirement. Since the data at stations 5 and 40 are missing, we cross-correlate the waves in the tapered P-wave windows for stations 6-39. (Figure 2.10). Next, we stack the tapered P-wave windows and then taper the sum of the windows, using the Gaussian window that accounts for approximately 40% of the width of the receivers array used. The two tapers minimize the truncation effects (Burdick and Orcutt, 1979). The final process is to take the time derivative of the stacked tapered cross-correlated P-waves. Assuming step functions for the source-time function and using the P-wave seismic records, this provides an estimate of the P-waves that propagate between the two selected earthquakes (Figure 2.12a).

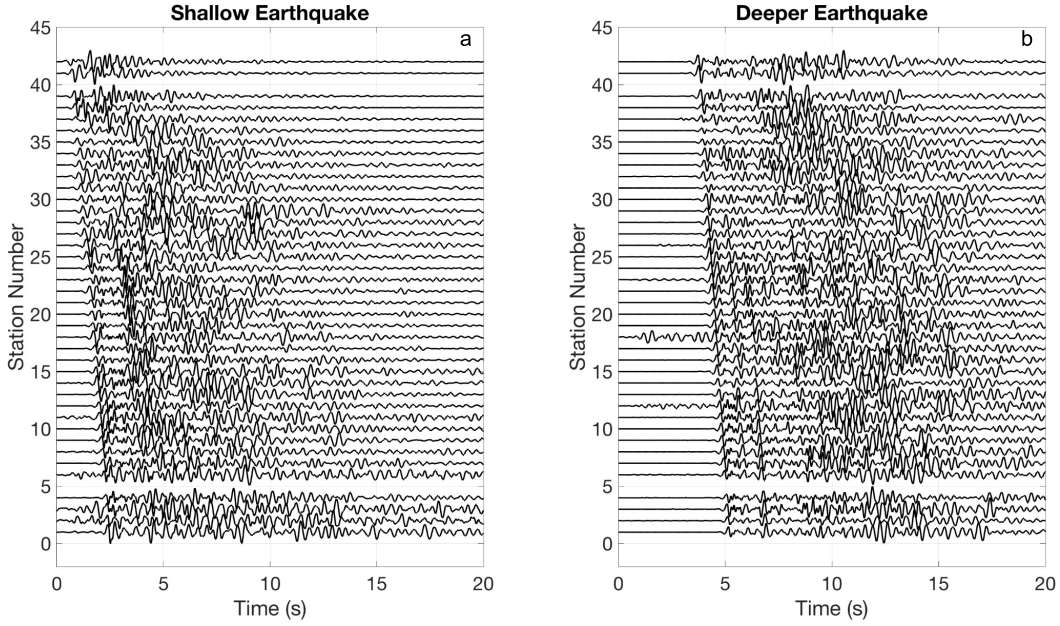


Figure 2.10 Seismic records of a) shallow earthquake and b) deeper earthquake from the line array along SAF (Line 3 in Figure 2.9). The records are bandpass-filtered at 1-3 Hz and start at the origin times of the two earthqaukes from the USGS catalog. Station 1 is the farthest receiver from the two earthquakes.

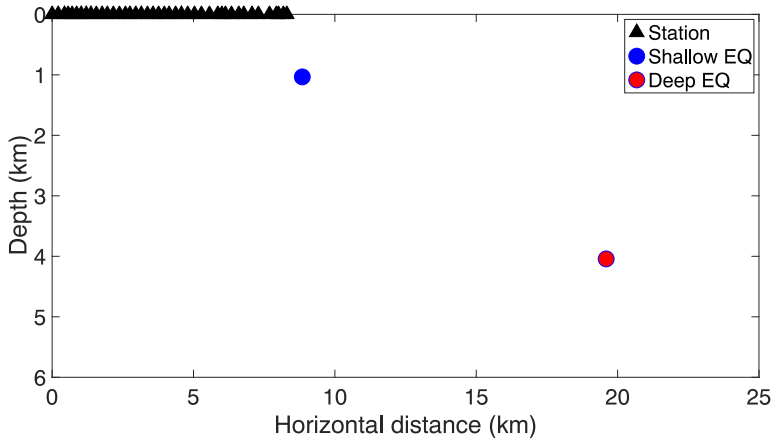


Figure 2.11 Cross-section of the selected earthquakes and seismic array along the SAF.

Using receivers only along a line, the interferometric P-waves (Figure 2.12a) provide a qualitative estimation of the P-wave propagation between the two sources. The interferometric signal in Figure 2.12a consists of 1) a direct arrival at approximately 2.5-3 s and 2) a later arrival after approximately 3 s. The timing of the direct arrival is consistent with the expected arrival time determined from the earthquake catalog and the local P-wave velocity of the area.

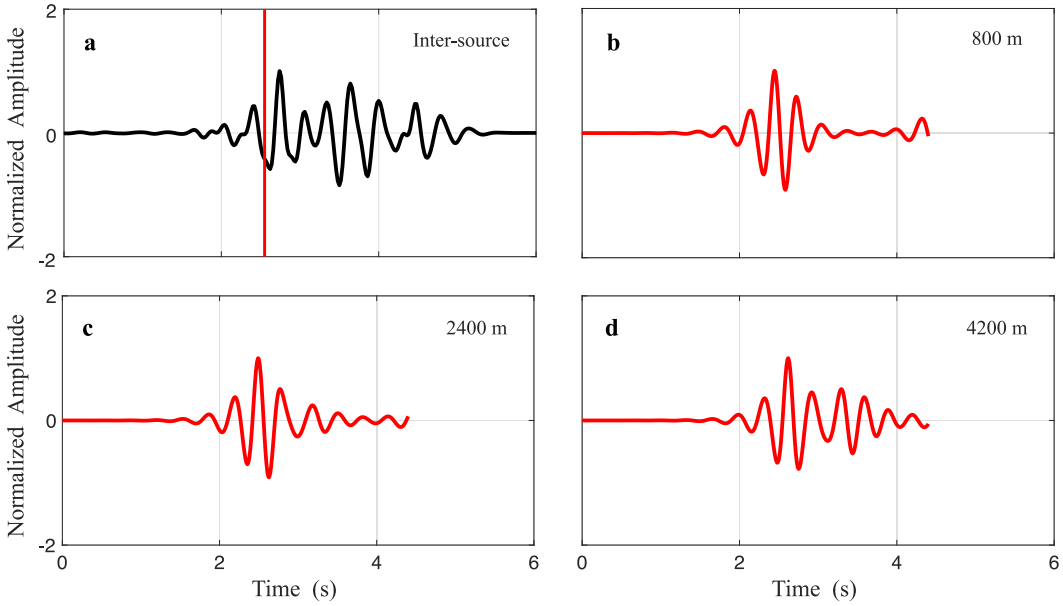


Figure 2.12 a) Inter-source interferometry of 1-3 Hz bandpass-filtered P-waves. Numerical simulations of wave propagation in a 3-layer model between a source and a receiver with a thickness of the low-velocity layer of b) 800 m, c) 2.4 km, and d) 4.2 km.

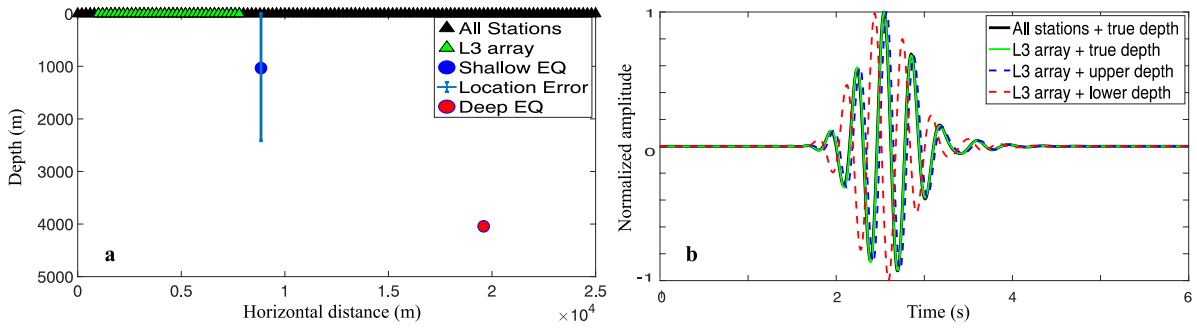


Figure 2.13 a) Synthetic model to compute acoustic wavefields. The model includes surface receivers with receiver spacing of 200 m and two sources located at the same depth of the selected earthquakes from the SAF. The errorbar indicates the depth uncertainty of the shallow earthquake from the earthquake catalog. The green receivers indicate the location of Line 3 (L3) array along the SAF in Figure 2.7. b) The inter-source interferometric waveforms reconstructed using all surface receivers and the true depth of the shallow earthquake (solid black), L3 receiver and the true depth of the shallow earthquake (solid green), L3 receiver and the shallow earthquake located at the surface (dashed blue), L3 receiver and the shallow earthquake located at the lower end of the errorbar in a (dashed red).

We recognize the large uncertainty in the depth of the selected earthquakes from the catalog, and this high uncertainty may contribute to the error of the wave retrieval using inter-source interferometry because when the depth of the earthquakes is significantly different from the depths in the catalog, the receiver array may not adequately sample the stationary phase zone of the interferometric integral. To test the effect of the depth sensitivity on the interferometric P-wave retrieval, we use a synthetic homogeneous model with two sources at the same locations of the two selected SAF earthquakes and surface receivers, including the Line 3 (L3) array of 200 m station spacing (Figure 2.13a). Figure 2.13b shows that although the L3 array does not cover the whole stationary phase zone of the two earthquakes, we can reconstruct the same waveforms using the records from the L3 array and the full line surface array. By contrast, varying the depth of the shallow earthquake, we can retrieve the similar waveforms but with different phases of the direct arrivals (red and blue dashed lines on Figure 2.13b). The depth uncertainty only affects the phase shift of the waveforms but does not reshape the waveform.

Since we verified that the depth uncertainty of the selected earthquakes does not affect the shape of the retrieved waveform, we speculate that the recovered later arrival after approximately 3 s is a guided wave traveling inside the low-velocity fault zone of SAF. To investigate whether the fault-zone wave guide qualitatively produces the observed waveform characteristics, we simulate the wave propagation between a source and a receiver in a 3-layer model, both of which are located in the middle slow-velocity layer. This is a crude model (Figure 2.14) for the host rock and the slow-velocity layers of SAF (Jeppson and Tobin, 2015; Li et al., 2004). The elastic model is the plan view of SAF and accounts for P-S wave conversions. We use a source-time function given by a Ricker wavelet with the dominant frequency at 3.3 Hz. Assuming

the two earthquakes occurred inside the fault zone of SAF, we model the source and receiver inside the low-velocity zone with horizontal and vertical separation distance of 12 km and 350 m, respectively (Figure 2.14). The horizontal distance of 12 km is approximately the same separation distance between the two earthquakes (Figure 2.11). The vertical distance of 350 m is arbitrarily chosen to be less than the smallest fault-zone thickness of SAF, which is less than 1 km (Holdsworth et al., 2011; Jeppson and Tobin, 2015; Korneev et al., 2004; Li et al., 1990, 2004; Zoback et al., 2011). This arbitrarily chosen vertical distance also minimizes reflection interference effects from the two boundaries of the slow-velocity layers in our 3-layer model (Figure 2.14). We simulated the wavefield with a 2D spectral element method (Tromp et al., 2008) for low-velocity layers with a thickness from 700 m to 7 km. The 2D simulation cannot be expected to quantitatively explain the waveforms; it is a crude model to explain the qualitative features of the extracted waveforms.

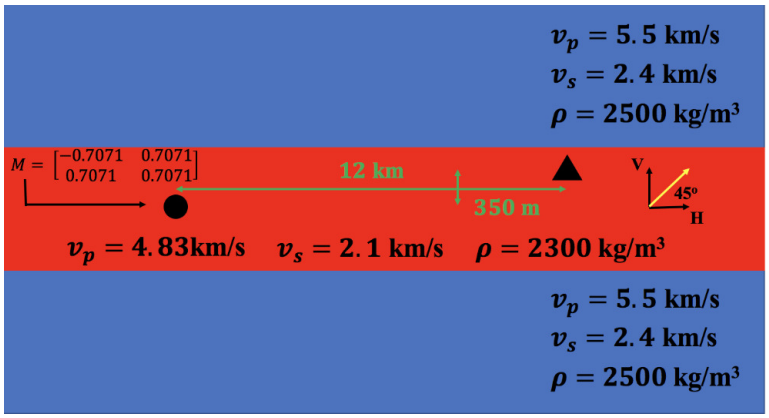


Figure 2.14 Numerical model of 3 layers used to simulate 2D wave propagation between a source and a receiver in a low-velocity layer (red) The yellow arrow indicates the direction in which the wave motion is recorded.

Examples of the wavefield with the middle layer thickness of 800 m, 2.4 km, and 4.2 km, respectively, are shown in Figure 2.12; the wavefields are bandpass-filtered between 2.5 Hz and 3.8 Hz to account for the dominant frequency of the interferometric signal at approximately 3.2 Hz. The modeled seismograms (Figure 2.12b, c, and d) illustrate that a low-velocity zone with a thickness of 4 km best explains the qualitative features of the extended wavetrains obtained from inter-source interferometry (Figure 2.12a). Figure 2.12 suggests that the later arrival is likely to consist of guided waves travelling in the damage zone of SAF with a thickness of approximately 4 km. The crude model used cannot be expected to lead to a fit of the synthetic waveforms with the waves extracted with the VRM, but the qualitative similarity between the wavefield estimated by the interferometry and the numerical simulation illustrates that inter-source interferometry can qualitatively estimate the wavefield, which accounts for the wave propagation between

two earthquake sources.

2.6 Discussion

Body waves contain three pieces of information: 1) the arrival time of the direct wave, 2) the waveform of the direct wave, and 3) the waveform of the coda wave. Using inter-source interferometry, one can reconstruct the direct wave and the coda of a body wave that propagates between two seismic sources, but the retrieved arrival time of the direct body wave may not be accurate.

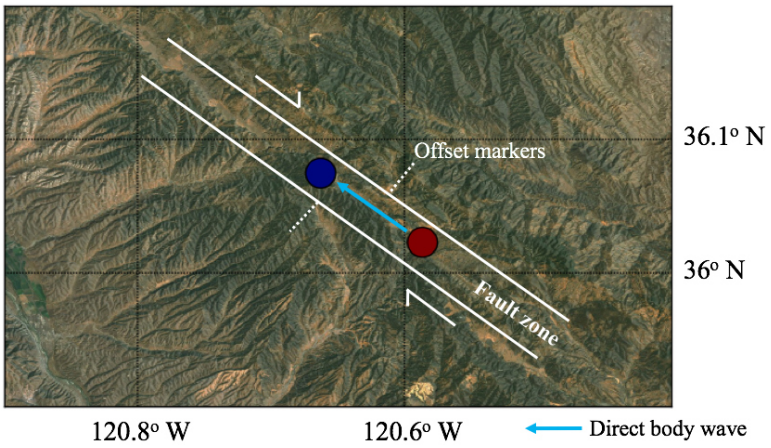


Figure 2.15 Cartoon of the map view of the fault-zone-guided waves that travels in a fault zone of the SAF. The shown fault zone width is not to scale. The cartoon is the plan view of the numerical model in Fig 2.14.

The retrieved arrival time depends on the difference between the origin time of the two sources, which comes from the earthquake catalog. Using inter-source interferometry, one cross-correlates earthquake seismograms with start time at the origins as recorded in the event catalog. However, the origin time information taken from the catalog cannot be assumed to be precise, and errors in the origin time lead to errors in the arrival time of the extracted direct wave. In addition, the depth uncertainty of the selected two earthquake and the limited aperture of used seismic array also contribute to the phase shift of the retrieved waveform (Section 2.3, Section 2.4, Figure 2.13, and the appendix) and thus maps the phase error to the errors in the arrival time retrieval.

We can reconstruct the waveform of the direct body wave by using receivers located within the stationary phase zone of the two sources. Retrieving this waveform offers a variety of benefits. For example, by selecting two earthquakes occurring in a fault or subduction zone, and using the inter-source technique, we can reconstruct the waveform of the direct body wave that travels along the zones and is influenced by the structure of these zones (Figure 2.15); the direct wave travels in the low-velocity zone of the fault zone or subduction zone. This direct wave helps to define properties such as the width (thickness) and the wave

velocity of the zones, for example by using fault-zone guided wave as a diagnostic of structure of fault zones (Li et al., 1990).

Using VRM, one needs an adequate sampling density of receivers to reconstruct waveforms of direct body waves. In practice, if the waves propagate horizontally along surface receivers, one needs adequate sampling points of 4 samples/wavelength to directly integrate the interferometric wavefield (Fan and Snieder, 2009). However, additional processes prior to the integration of oscillatory functions such as interpolation and Filon-Trapezoidal rule (Entwistle et al., 2015; Tuck, 1967) can reduce the sampling requirements and one only needs more than 2 sampling points in a wavelength to adequately sample the interferometric wavefields. This sampling density can be larger if the wave arrives at an angle to the receivers because only the wavelength along the surface contributes to the wave reconstruction; the required receiver spacing changes to $\lambda/sin(i)$, where λ and i denote the wavelength and the incident angle of the wavefront to the receiver array, respectively.

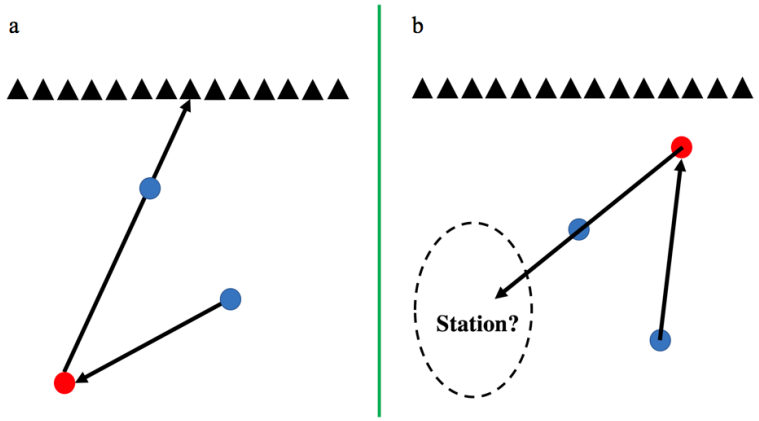


Figure 2.16 Cartoon of receiver locations needed for the retrieval of coda waves. Figure 2.16a and Figure 2.16b have a scatterer (red) located below and above two seismic events (blue).

As equation (1) illustrates, to recover the waveform of coda waves using inter-source interferometry, one needs receivers on a closed surface surrounding the sources. Because the dominant contribution of Green's function retrieval comes from the receivers located inside stationary phase zones, one can reconstruct scattered waves using surface receivers only if scatterers are located below seismic events (Figure 2.16a). However, one needs underground receivers rather than surface receivers to recover the waveform of coda waves if scatterers are located above the events (Figure 2.16b). Although Figure 2.16 shows a sketch with a scatterer, one can apply similar reasoning when retrieving coda waves from a reflector (e.g., core-mantle boundary and Moho discontinuity).

Our synthetic examples show that to retrieve the waveform of direct body waves, one needs to use surface receivers located in the stationary phase zone of the selected sources with enough sampling density.

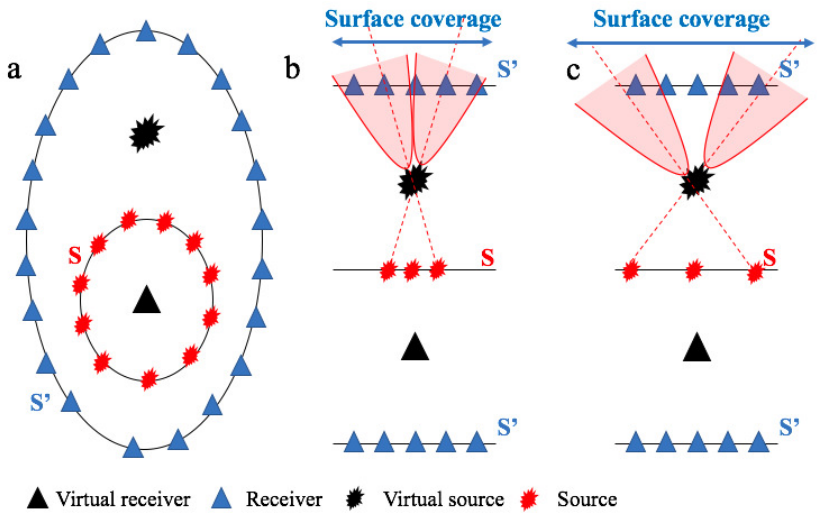


Figure 2.17 Different geometries of source-receiver interferometry, modified from Figure 5 in Curtis and Halliday (2010). a) source-receiver geometry with two closed surfaces: S surrounding a virtual receiver and S' surrounding a virtual source. b) source-receiver geometry with limited surface receiver coverage controlled by the stationary phase zones of sources on the surface S and the virtual source. c) similar source-receiver geometry to b) but with more sparsely distributed sources on the surface S, leading to a larger surface receiver coverage. Pink hyperbolae denote the stationary phase zones.

These limitations apply to any advances of inter-source interferometry. Curtis and Halliday (2010) apply inter-source interferometry as an operation for source-receiver interferometry, which estimates the Green's function accounted for the wave propagation between a virtual source and a vitual receiver. The operation requires the stationary phase zone to be parallel to the virtual source and a set of sources located on the closed surfaces surrounding the virtual receiver (Figure 2.17). Thus, the geometry of the sources and the virtual source control the size and the location of the stationary phase zone and the surface receiver coverage. Depending on the distribution of the sources, we may need a large aperture for the stationary phase zone and the surface coverage, or we may need the entire surface array if the set of sources are very sparsely distributed away from the virtual source (Figure 2.17c). The requirement for surface receiver coverage also applies to other interferometric advances. For example, in the virtual source method application, the surface coverage controls the geometry of the target area and source-receiver positions as well as the resolution of the reconstructed wavefields (Bakulin and Calvert, 2006).

In section 2.5, we illustrate a field example of inter-source interferometry of P waves recorded from two earthquakes in the SAF area, in which the recorded seismograms of the example satisfy the stationary phase location and sampling criteria. Because we lack seismic wavefields across the SAF and only use seismograms from the array along the fault (Line 3, Figure 2.9) for the inter-source method, we can only qualitatively reconstruct the P wavefield accounting for the wave propagation between the two earthquakes.

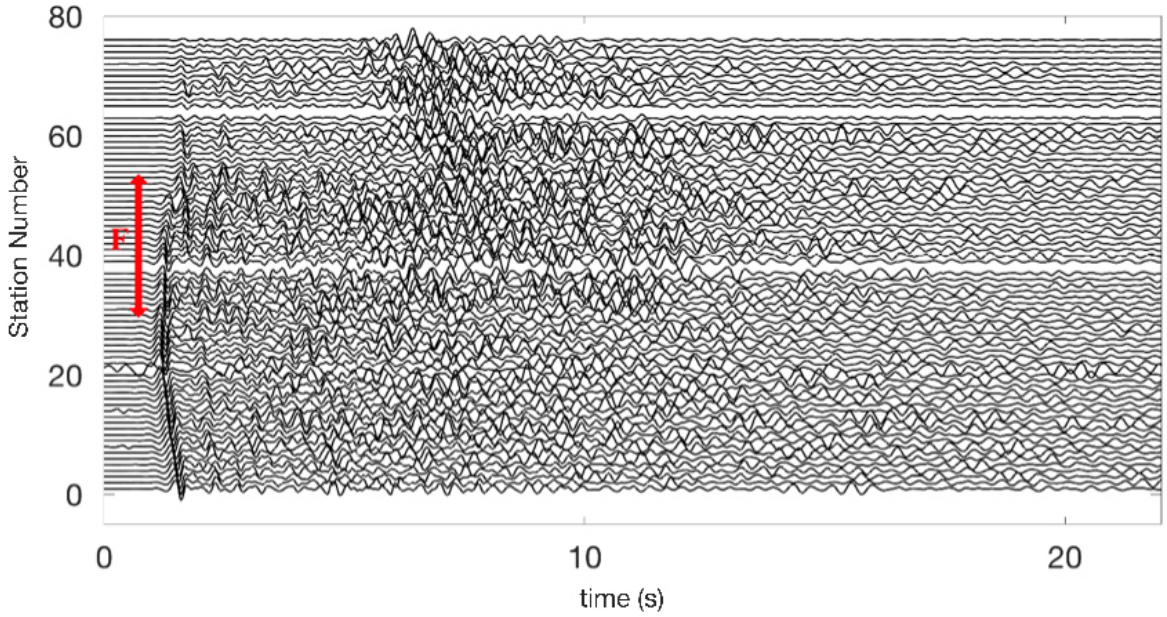


Figure 2.18 Seismograms with 1-3 Hz bandpass filter of the selected deeper earthquake in Figure 2.11, recorded at receivers across the SAF (Line 1 and 2, Figure 2.9). The red arrow shows the approximately wide region that straddles the SAF where the waveforms show an increased complexity.

The retrieved wavefield and the wave simulation using a 2D fault model suggest that the damage zone of the SAF is approximately 4 km thick. This damage zone is the region straddling the central fault zone where the surrounding rock is softened by deformation. This concept is explained in detail by Choi et al. (2016) who call this the *cross-fault damage zone*. The thickness inferred from the interferometric waveform is different than the thickness of the SAF damage zone that is estimated to be less than 1 km, in previous studies of the nearby areas of SAF (Holdsworth et al., 2011; Jeppson and Tobin, 2015; Korneev et al., 2004; Li et al., 1990, 2004; Zoback et al., 2011). The difference in the thickness estimates may be from the diversity of our investigated seismic frequency of 1-3 Hz compared to the other studies. Korneev et al. (2004) investigate the fault zone guide wave of the SAF areas at approximately 3-8 Hz while the investigated trap waves of Li et al. (1990) contain frequency contents approximately below 100 Hz. Although Li et al. (2004) study trapped waves with a frequency content below 5 Hz (similar frequency content to our study), the investigated guided waves are part of the S coda which is different to our investigated P coda. Thus, the difference in the frequency content and the wave types considered may explain the difference of the estimated thickness of the SAF damage zone. Other explanations for our different estimates for the thickness may be found in the following studies of nearby SAF areas and other strike slip faults: 1) the shallow structure of the SAF may contain local flower structures and splay faults that expand approximately up to 1-3 km (Delong et al., 2010; Unsworth et al., 1997) and 2) off-fault deformations may occur and extend up to 1-2 km away from

the primary rupture zone (Gold et al., 2015; Vallage et al., 2015; Zinke et al., 2014). In addition, the 1-3 Hz bandpass-filtered wavefield of the selected deeper earthquake recorded at receivers across the SAF (Line 1 and 2, Figure 2.9) illustrates a high energy and complicated waveforms in a band with a width of approximately 4 km across the SAF (Figure 2.18), supporting the thickness determined by the interferometric wavefield.

In our SAF example, we used the inline receivers only and only have receivers with a limited aperture. As a result, we can only recover the qualitative waveform but do not retrieve the complete waveform of the P wavefield. Halliday and Curtis (2008) show that using an in-line array for interferometry of surface waves induces phase shifts compared to the true Green's function. Seismic interferometry implicitly involves integration of oscillatory integrals where the dominant contribution comes from the stationary phase point. According to expression (24.40) of Snieder and van Wijk (2015a), each stationary phase integral gives a phase contribution $\exp(\pm i\pi/4)$, where the sign is determined by the curvature of the phase function. Leaving out the stationary phase integration in the transverse direction gives a phase error of $\pm\pi/4$. Since we only seek to compare the character of the waveforms qualitatively, this phase error is not significant.

Our work shows that inter-event interferometry of two earthquakes cannot recover the complete waveform of coda waves but can retrieve the waveform of the direct body waves, and the accuracy of the retrieved arrival time of the direct waves depends on the accuracy of the earthquake catalog. The lack of transverse receiver array also contributes to the incomplete recovery of the direct body waves. However, in active-source explorations, particularly on land, one has a well-defined catalog of the locations and timing of active sources. Using inter-source interferometry and well-cataloged active sources, one can better recover the arrival time of the direct body waves with smaller timing errors, compared to the uses of global or local earthquakes.

2.7 Conclusion

Although one can reconstruct a body wave propagating between two seismic sources (e.g., earthquakes) using inter-source interferometry, the retrieved arrival time of the direct body wave depends on the difference between the origin times of the earthquakes that are extracted from the earthquake catalog, and parts of the catalog could be inaccurate. Moreover, we show that the depth uncertainty influences the phases of the retrieved direct arrival and maps the error to the retrieved arrival time. Thus, the retrieved arrival time may be useful only if the origin times and depths of the used events in the catalog are accurate. However, when one knows the source-time function of at least one of the sources (the virtual sensor), the waveforms of the direct and coda waves, which are part of the body wave propagated between the two sources, can be recovered using the inter-source method with receivers on a closed surface.

The retrieval of the waveforms of the direct and coda waves can also be approximated using surface receivers located in the associated stationary phase zones, but one can retrieve coda waves using surface

receivers only if the waves are scattered and reflected below the shallowest source. Theoretically, one needs receivers in the stationary phase zone, which is a patch on the surface, to recover the direct or coda waves but in practice one may use receivers along a line array.

To reconstruct accurate body waves propagated between two sources using inter-source interferometry, one needs to satisfy two conditions: 1) using receivers inside the stationary phase location and 2) the sampling criteria of approximately 4 samples/wavelength for the direct summation of interferometric wavefields over receivers. In this paper, we illustrate an example of direct P-wave reconstruction using the inter-source method and real-earthquake wavefields from the SAF satisfying the two requirements. Using finite-element modelling and a simplified seismic velocity model for the SAF region with a slow-velocity fault zone embedded in a homogeneous medium, we test the effect of different fault zone thicknesses on the waveform character. We find that a fault zone thickness of approximately 4 km is needed to qualitatively match the wavetrains of the P waves recovered using VRM. The width of the fault zone determined from the interferometric waveform and our modelled solution is consistent with the distance over high-energy, highly scattered, and complicated waveforms manifested on recordings from surface receivers across the fault.

2.8 Acknowledgements

We thanks colleagues from Center for Wave Phenomena (CWP) for useful discussions, Robert Clayton (Caltech) for the seismic records used in this study, and Ridvan Orsvuran for helping with the 2D spectral element method modelling. The authors thank the editors from the journal of Pure and Applied Geophysics for handling this manuscript and thank Andrew Curtis for constructive and critical comments that helped to improve this work. P Saengduean is financially supported by DPST, the Development and Promotion of Science and Technology Talents Project (Royal Government of Thailand scholarship). The authors declare that they have no competing interests.

2.9 Appendix: Phase Shift and Fresnel Integral

In section 2.4, we show that one can reconstruct direct and scattered body waves using receivers in the stationary phase zones. Using an array with a limited aperture in the stationary phase region, the phase of the recovered waveforms differs slightly from the forward-modelled waveforms (Figure 2.5, Figure 2.6, and Figure 2.7).

When we use the source-receiver geometry in Figure 2.2, we can determine the time differences between the direct arrivals of the two sources at each receiver. We define the size of the stationary phase zone (SPZ) using a faction of the dominant period $T_{dominant}$ of our synthetic signals, where $T_{dominant} = 0.25$ s. In Section 2.4, our SPZ covers the receivers within $T_{dominant}/4$ from the maximum time difference (Figure 2.19).

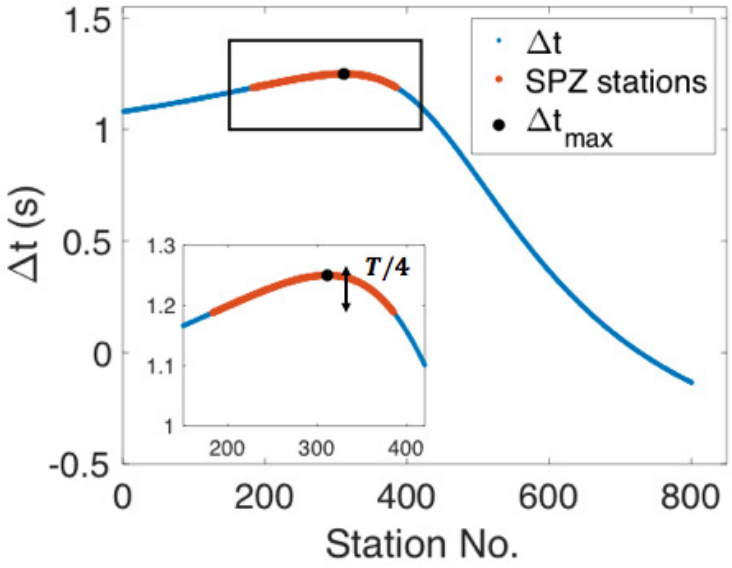


Figure 2.19 Time differences between the direct arrivals of the two sources at each surface receiver in Figure 2.2. The graph on the bottom left corner is the enlarged version of the area inside the black box.

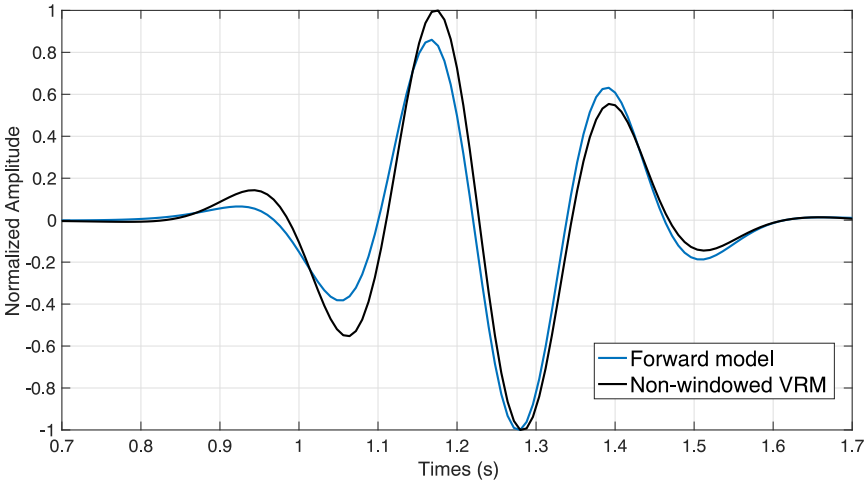


Figure 2.20 The forward-modeled and the interferometric waveforms of the direct body waves.

Figure 2.20 shows the direct wave that propagates between the two sources in Figure 2.2, and the waves extracted from inter-source interferometry take only receivers into account with a delay time $T_{dominant}/4$ relative to the delay time at the stationary phase zone. One should note the slight phase change between the two waveforms. We show in this appendix that this phase error is associated with limiting the used receivers to the stationary phase region.

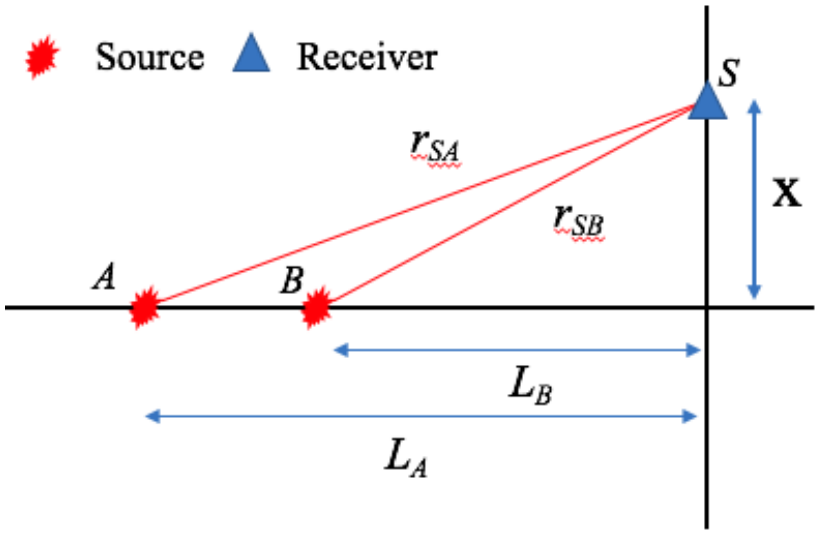


Figure 2.21 Source and receiver positions in interferometry. x is the half-width of the stationary phase zone.

Consider the source-receiver geometry in Figure 2.21. Seismic interferometry for waves in a 2D homogeneous medium gives

$$G(r_A, r_B, \omega) - G^*(r_A, r_B, \omega) \approx \frac{i}{4\pi\rho} \oint_{S'} \sqrt{\frac{1}{r_{SA} r_{SB}}} e^{ik(r_{SA} - r_{SB})} dS', \tag{2.9}$$

where $G(r_A, r_B, \omega) - G^*(r_A, r_B, \omega)$ is the difference between the causal and acausal parts of the Green's function accounting for the wave propagation between the points A and B, ω is the angular frequency, ρ is the mass density, k is the wave number, and S is the source position on a surface S' (Fan and Snieder, 2009). We can approximate the term $r_{SA} - r_{SB}$ in equation (2.9) using a second-order Taylor series:

$$r_{SA} - r_{SB} \approx (L_A - L_B) + \frac{1}{2}\left(\frac{1}{2L_A} - \frac{1}{2L_B}\right) x^2. \tag{2.10}$$

Using this second order Tayor expansion for the phase and replacing the geometrical spreading in equation (2.9) by $1/\sqrt{L_A L_B}$, the integral in expression (2.9) is in the stationary phase approximation (Snieder and van Wijk, 2015a) given by

$$G(r_A, r_B, \omega) - G^*(r_A, r_B, \omega) \approx \frac{i}{4\pi\rho} \frac{e^{ik(L_A - L_B)}}{\sqrt{L_A L_B}} \int_{-\infty}^{+\infty} \exp\left(\frac{ik}{2}\left(\frac{1}{L_A} - \frac{1}{L_B}\right) x^2\right) dx\,, \tag{2.11}$$

where we replaced the integration over a closed surface by an integration over the transverse x-coordinate. The integral in the right-hand side is of the same form as the Fresnel integral (Sandoval-Hernandez et al., 2018), which is defined as

$$F(X) = \int_0^X \exp\left(i\frac{\pi}{2}x'^2\right) dx' \; . \tag{2.12}$$

This integral is complex and is shown in the complex plane in Figure 2.22 where the the origin corresponds to the integral for $X = 0$, and the asymptotic point $0.5 + 0.5i$ is reached for $X \to \infty$. This graphical representation is called the *Cornu spiral*. The many windings of this spiral around the asymptotic point reflect that the Fresnel integral converges slowly.

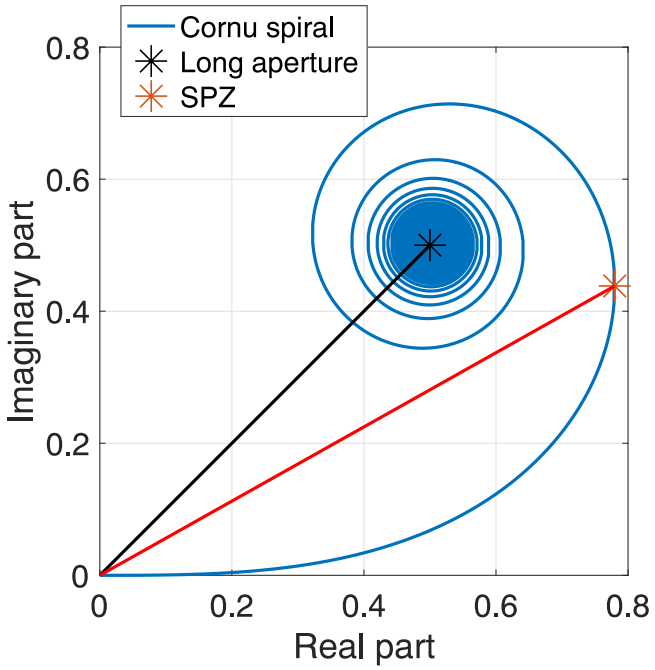


Figure 2.22 Fresnel integral from $X = 0$ to $X \to \infty$ in the complex domain (blue). Black and red asterisks represent the integral over a large part of surface receivers and surface receivers located in the stationary phase zone, respectively.

The real part of the integrand of the Fresnel integral (2.12) is shown in Figure Figure 2.23. The oscillatory nature of this integrand explains the shape of the Cornu spiral and the slow convergence of the Fresnel integral: the alternating positive and negative values of the integrand lead to the slow spiraling of the Fresnel integral in the complex plane, and one needs to integrate over a large interval for the integral to be close to its asymptotic value for $X \to \infty$.

In the waveforms of Figure 2.20 the summation over receivers was limited to receivers that have a delay time of less than a 1/4 of the dominant period. That corresponds to a phase delay of less than $\pi/2$. In

the Fresnel integral (2.12), this integration interval corresponds to the upper limit $X = 1$. As shown by the black dot in Figure 2.23, this upper limit is at the point where the real part of the Fresnel integral has its first zero crossing. (This is actually the rationale for restricting the integration to delay times less than 1/4 of the dominant period.) Since the real part of the integrand of the Fresnel integral changes sign, the real part of the Fresnel integral has a maximum at this point, and hence this upper limit corresponds to the rightmost point of the Cornu spiral that is indicated by the red star in Figure 2.22.

The important point to note is that the phase of the Fresnel integral for $X = 1$ differs from the phase of the integral computed for its asymptotic value for $X \to \infty$; the phase angle for the red star in Figure 2.22 differs by about 0.273 rad from the phase angle for the black star. This phase difference leads to the phase difference in the waveforms shown in Figure 2.20. The phase difference in the waveforms corresponds to about 0.201 rad. This is slightly less than the phase difference of 0.273 rad predicted above. This discrepancy is caused by the fact that the argument in this appendix ignores the decay of the interferometric integral that is due to variations of the geometrical spreading in the integration. Furthermore, the analysis in this appendix is applicable in the frequency domain, whereas the time-domain waveforms in Figure 2.20 contain contributions from a range of frequencies. However, the main points from this appendix are that (1) the interferometric integral converges slowly and (2) truncating this integral can leads to phase errors of about 0.3 rad. Tapering of the interferometric integral will lead to a faster convergence of the integral. Since the phase error does not depend on the distance between the sources at locations A and B, the impact of the phase error on estimated velocities decreases with increasing distance between the sources.

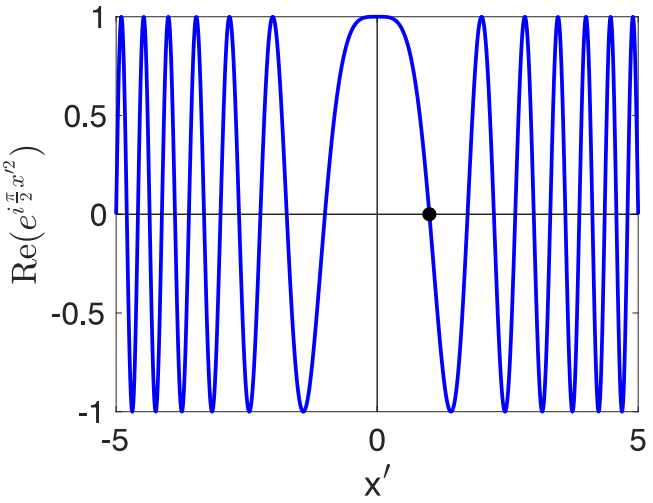


Figure 2.23 Real part of the integrand of the Fresnel integral (2.12) for $-5 < x' < 5$. Black dot denotes the first position where the real part changes from positive to negative values.

CHAPTER 3

MULTI-SOURCE WAVEFIELD RECONSTRUCTION COMBINING INTERFEROMETRY AND COMPRESSIVE SENSING: APPLICATION TO A LINEAR RECEIVER ARRAY

Modified from a paper submitted to Geophysical Journal International.

Patipan Saengduean[5], Roel Snieder[6], and Michael B. Wakin[7]

3.1 Summary

Seismic interferometry is a technique that allows one to estimate the wavefields accounting for the wave propagation between seismometers, any of which can act as a virtual source. Interferometry, particularly passive noise interferometry, has been applied to several geophysical disciplines such as passive monitoring and distributed acoustic sensing. In practice, one requires long recordings of seismic noise for noise interferometry. Additionally, some receivers in seismic arrays may be absent or inoperative due to issues of receiver installation and malfunction. Reducing the storage for seismic noise records and alleviating the limitations of receiver operation and installation require wavefield reconstruction and regularization techniques. Compressive sensing is one such method that can reconstruct seismic wavefields and help mitigate the limitations by exploiting the sparsity of seismic waves. Using numerical examples, we show that one can apply compressive sensing to recover interferometric wavefields resulting from interferometry of a linear seismic array. Traditionally, one can interpolate interferometric wavefields using correlograms provided by one virtual source. This method is called *single-source wavefield reconstruction*. We propose an alternative technique called *multi-source wavefield reconstruction*, which applies compressive sensing to reconstruct multiple interferometric wavefields using correlograms provided from all available virtual sources. To exploit the sparsity of interferometric wavefields, we apply the Fourier and Curvelet transforms to the two reconstruction schemes. Using the signal-to-noise ratio (SNR) to compare the wavefield reconstructions, the Fourier multi-source method improves the recovery of interferometric wavefields by approximately 50 dB compared to the Fourier and Curvelet single-source wavefield reconstructions.

3.2 Introduction

Seismic interferometry (SI) is widely applied to estimate the wavefields propagating between seismometers. The technique allows each seismometer to become a virtual source (VS), where wavefields from the VS are recorded by the other seismometers. Applications and principles of SI have been explained in several

[5]Graduate student, primary researcher, and author at Department of Geophysics, Colorado School of Mines
[6]Supervisor at Department of Geophysics, Colorado School of Mines
[7]Supervisor at Department of Electrical Engineering, Colorado School of Mines

review papers (Curtis et al., 2006; Larose et al., 2006; Snieder and Larose, 2013a; Wapenaar et al., 2010). For SI, one can estimate the interferometric wavefields using cross-correlation (Asano et al., 2017; Miyazawa et al., 2008; Mordret et al., 2010; Shapiro et al., 2005), deconvolution (Nakata et al., 2011; Pianese et al., 2018; van Dalen et al., 2015; Vasconcelos and Snieder, 2008a,b), cross-coherence (Nakata et al., 2011; Prieto et al., 2009), or convolution (Curtis and Halliday, 2010; Entwistle et al., 2015).

Seismic noise interferometry is an application of cross-correlation SI that can be used in passive seismic surveys. Noise interferometry typically requires long seismic records, which requires much data storage. For example, Lin et al. (2008) and Nakata et al. (2015) cross-correlate noise recordings over 1-year and 3-month records, respectively. Jayne et al. (2022) handle the long-recording issue by reducing the number of the Fourier coefficients required for the cross-correlation in the Fourier domain (Snieder and Wakin, 2022). In addition, seismic arrays (e.g. the US array from IRIS) are usually sparse and irregular. In practice, dense and regular seismic profiles are not typically acquired due to limitations of installation and operation of seismic sensors, for example because of restricted locations for receiver installation and low-quality seismic records (low SNR) in some receivers. These irregular seismic records may affect the performance of subsequent seismic data processing steps. For example, seismic imaging such as the Kirchhoff migration requires regular seismic shot lines to enhance imaging of the subsurface and to mitigate post-migration noise (Al-Gain et al., 2020; Cao et al., 2018; Poole and Herrmann, 2007). Irregular sampling of seismic data can also affect the attenuation of noise for guided wave and ground roll removal (Mann and Emanuel, 2006). Regularizing seismic shots can improve the SNR of seismic signals and the alignment reflection events (Chopra and Marfurt, 2013). Thus, wavefield reconstruction or regularization techniques are required to alleviate the issue of irregular sampling of seismic data.

Wavefield reconstruction techniques include wave-equation-based methods (Kim et al., 2015; Peters et al., 2014), linear prediction filter methods (Islam et al., 2015; Vaidyanathan, 2007), rank-reduction methods (Innocent Oboué et al., 2021; Wu and Bai, 2018), mathematical-transform methods (Wang et al., 2016), and deep learning (Liu et al., 2019). These techniques help handle the limitations of receiver malfunction and installation by recovering the missing signals within the gaps of the seismic receiver arrays. Among these techniques, compressive sensing (CS) is a transform-based method that has gained popularity and been applied in many signal processing disciplines (e.g. medical and geophysical) (Candès and Wakin, 2008). The CS method is a sampling paradigm that reduces the traditional sampling requirements of the Shannon sampling theorem in which the sampling rate of a signal has to be twice the maximum frequency of the signal (Donoho, 2006). CS has been shown to be effective for interpolating seismic profiles (Hennenfent et al., 2010; Herrmann et al., 2008) and surface waves (Zhan et al., 2018) using signals recorded on an irregular seismic array.

Because SI can estimate the interferometric wavefields accounting for wave propagation between receivers, SI can provide sampling and coverage of wavefields in the area under receiver arrays. Thus, SI may provide new perspectives or improvements to seismic signal reconstruction. A technique called interferometric interpolation utilizes a matched filter and virtual source gathers provided by SI (Hanafy and Schuster, 2014; Wang et al., 2009; Xu et al., 2018). This technique can extrapolate near-offset seismic profiles and improve the interpolation quality of missing seismic traces compared to conventional interpolation methods without SI. Since Zhan et al. (2018) demonstrate the reconstruction of surface waves on a dense surface array generated from a source using CS, we can use CS reconstruction to recover surface waves retrieved using SI on a dense receiver array when some receivers are missing.

In this work, we use a linear receiver array to record noise wavefields when some traces are missing. Using SI, we retrieve cross-correlation profiles from available traces and then aim to reconstruct the missing cross-correlation wavefields using CS. We propose a joint reconstruction technique called multi-source wavefield reconstruction where we use CS to reconstruct the missing interferometric wavefields (i.e. cross-correlations). Our work focuses on the estimation of fundamental-mode surface waves using SI which is illustrated by synthetic surface waves. In Section 3.3, we describe the basic theory of cross-correlation interferometry and CS. We then explain the processes to simulate interferometric surface waves and the Fourier and Curvelet transforms used in our wavefield reconstruction. In Section 3.4, we show the different wavefield reconstruction techniques using different sparse transforms and compare these reconstructions to our proposed multi-source wavefield reconstruction.

3.3 Theory and Synthetic Example

In this section, we describe the theory of cross-correlation interferometry, compressive sensing, and the synthetic model used in our wavefield reconstruction examples.

3.3.1 Cross-Correlation Interferometry

Using a frequency-domain formulation, one can express the wavefield $u(\mathbf{x_r}, \mathbf{x}, \omega)$ excited from a point source at $\mathbf{x}$ and recorded at $\mathbf{x_r}$ by

$$u(\mathbf{x_r}, \mathbf{x}, \omega) = W(\mathbf{x}, \omega) G(\mathbf{x_r}, \mathbf{x}, \omega), \tag{3.1}$$

where ω is the angular frequency, $W(\mathbf{x}, \omega)$ is the source-time function, and $G(\mathbf{x_r}, \mathbf{x}, \omega)$ is the frequency-domain representation of the Green's function that accounts for the wave propagation from $\mathbf{x}$ to $\mathbf{x_r}$. The cross-correlation of two wavefields recorded at receivers $\mathbf{x_a}$ and $\mathbf{x_b}$ is given by

$$C_{ba} = u(\mathbf{x_b}, \mathbf{x}, \omega) u^*(\mathbf{x_a}, \mathbf{x}, \omega) = |W(\mathbf{x}, \omega)|^2 G(\mathbf{x_b}, \mathbf{x}, \omega) G^*(\mathbf{x_a}, \mathbf{x}, \omega), \tag{3.2}$$

where the asterisk denotes complex conjugation. Integrating the cross-correlation equation (eq 3.2) over a closed surface that includes uncorrelated sources on the surface surrounding the receivers (Snieder et al., 2007) gives

$$\oint C_{ba} d^2\mathbf{x} = \langle |W(\mathbf{x},\omega)|^2 \rangle \oint G(\mathbf{x_b},\mathbf{x},\omega) G^*(\mathbf{x_a},\mathbf{x},\omega) d^2\mathbf{x}, \tag{3.3}$$

where $\langle |W(\mathbf{x},\omega)|^2 \rangle$ is the average of the amplitude spectrum of the source. Because the integrated cross-correlation in equation 3.3 is proportional to $G(\mathbf{x_a},\mathbf{x_b},\omega)$ (Wapenaar and Fokkema, 2006), one can estimate the approximate Green's function that accounts for the wave propagation between $\mathbf{x_a}$ and $\mathbf{x_b}$ given by

$$\big[G(\mathbf{x_a},\mathbf{x_b},\omega) - G^*(\mathbf{x_a},\mathbf{x_b},\omega)\big] \langle |W(\mathbf{x},\omega)|^2 \rangle \approx -\frac{2i\omega}{\rho c} \oint C_{ba} d^2\mathbf{x}, \tag{3.4}$$

where $i\omega$ corresponds to the time derivative and $G(\mathbf{x_a},\mathbf{x_b},\omega)$ and $G^*(\mathbf{x_a},\mathbf{x_b},\omega)$ are the causal and the acausal parts of the Green's function that accounts for the wave propagation between $\mathbf{x_a}$ and $\mathbf{x_b}$, respectively. In practice, one can estimate the integral in equation 3.4 by stacking the correlograms of all sources on the closed surface. For noise interferometry, one estimates the integral by stacking cross-correlated noise windows taken from long noise records. Our work compares seismic correlation profiles recovered using different reconstruction algorithms. Thus, we only estimate the integrated correlograms, ignoring the constants, the time derivative, and the source-time function in equation 3.4.

3.4 Signal Recovery using CS and Sparse Transforms

The Nyquist sampling theorem formulates the required sampling rate for a given frequency or wavenumber content of a signal. One can use compressive sensing (CS) to recover a signal from fewer random measurements than the sampling theorem requires by exploiting the sparsity or compressibility of signals in a basis or frame (Donoho, 2006; Wakin, 2017).

Using CS, one can represent an N-dimensional signal f via its sparse coefficients α in a basis or frame Ψ: $f = \Psi\alpha$. For example, one can transform seismic data f using the Fourier basis Ψ with Fourier coefficients α. The signal f is sparse when $K \ll N$, where K is the number of non-zero coefficients in α, and f is compressible when the sorted coefficients decay rapidly to approximately zero (Wakin, 2017). A K-sparse signal can be recovered without information loss if a sensing or sampling function Φ, which prescribes how the signals are sampled, satisfies the restricted isometry property (RIP). This property ensures the stability and energy conservation of sparse signals; examples of sampling matices that satisfy the RIP are random Gaussian and sub-Gaussian matrices (Candès and Wakin, 2008; Donoho, 2006; Wakin, 2017). In our example, the sampling function Φ specifies at which pairs of receivers on a linear array with randomly missing sensors the correlogram is available.

To satisfy RIP, one needs at least M compressive measurements d, where $d = \Phi f$ and $M = O(Klog(N/K))$, to reconstruct a K-sparse signal (Candès and Wakin, 2008; Wakin, 2017). Several different algorithms to seek the sparse solution have been reviewed by Rani et al. (2018). One common method is to solve the following l_1 optimization problem to recover the signal from noisy data d with noise n, where $d = \Phi f + n$ and σ is the constant bound of noise and $||n||_2 \leq \sigma$,

$$min||\alpha||_1 \text{ subject to } ||\Phi\Psi\alpha - d||_2 \leq \sigma. \tag{3.5}$$

Once the coefficient vector α is recovered, one can reconstruct the signal by using $f = \Psi\alpha$.

For our CS wavefield reconstruction examples, we use the 2D and 3D Fourier transforms associated with space-time and space-space-time dimensions of seismic data, respectively. An example of recovering space-space-time wavefields using the 3D Fourier transform is the reconstruction of the Green's function $G(\mathbf{r_i}, \mathbf{r_j}, t)$ for locations $\mathbf{r_i}$ and $\mathbf{r_j}$ at time t. The discrete 2D and 3D Fourier transforms are equivalent to the 1D Fourier transform along each dimension of the signals (Nussbaumer, 1982). To avoid wrap-around of the Fourier transform, we pad zeroes at the front and the end of our data matrix for each dimension. We also apply the 2D Curvelet transform in our wavefield reconstruction using CS. The Curvelet transform represents signals using three parameters: scales, location, and direction. We use the discrete Curvelet transform with wedge wrapping, previously written and developed by Candès et al. (2006a), to reconstruct interferometric wavefields.

3.4.1 Numerical Model Used for Wavefield Reconstruction

CS wavefield reconstruction can be applied to both active-source and noise interferometry. We use active-source interferometry as an example for our CS wavefield reconstruction. Our synthetic model includes a linear seismic array consisting of 100 receivers and 2000 impulsive sources. The sensors are separated by 10 m from each other and the sources are uniformly distributed on a rectangle surrounding the array (Figure 5.2). In our study, we simulate surface waves using the 2D Green's function of the Helmholz equation (equation 18.46 in Snieder and van Wijk (2015b)) given by

$$u(x,t) = \int W(\omega) H_0^{(1)}(k(\omega)x) e^{-i\omega t} d\omega, \tag{3.6}$$

where u is the wavefield at time t, x is the distance between the receiver and the source, $W(\omega)$ is the source-time function, $H_0^{(1)}$ is the first Hankel function of degree zero, k is the wavenumber, and ω is the angular frequency. One can compute the wavenumber in equation 3.6 by $k(\omega) = \omega/c(\omega)$, where c is the phase velocity.

In our simulation, the sources are delta functions bandlimited between 5-20 Hz and we use the phase velocity $c(\omega)$ of surface waves for a laterally homogeneous layered medium, determined by Xia et al. (1999),

to estimate the wavenumber in equation 3.6. Using interferometry, we select a master receiver and cross-correlate the wavefields recorded at each station to the master sensor. We then stack the cross-correlation for all sources. Each sensor thus can act as a virtual source providing surface waves recorded by the other receivers.

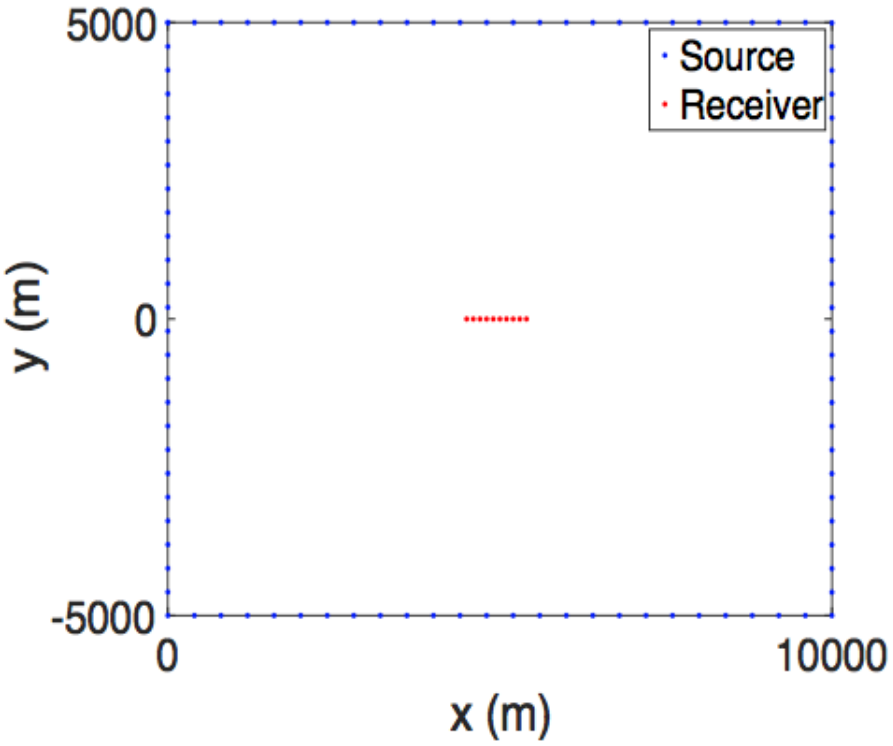


Figure 3.1 Geometry of sources and receivers used for the numerical examples. The sources (blue dots) are uniformly distributed on a rectangle surrounding a linear receiver array (red dots). For clarity the figure does not show all the sources and receivers. In the example, 100 receivers and 2000 sources are uniformly spaced.

3.5 Wavefield Reconstruction

Our work involves two steps of signal processing where cross-correlation interferometry of numerical surface waves is performed prior to CS reconstruction of interferometric wavefields. One can reconstruct interferometric wavefields using signals provided from a virtual source by cross-correlating the noise recorded on one master receiver with the noise recorded on all other receivers. We call this traditional method *single-source reconstruction*. By contrast, our proposed recovery method of interferometric wavefields is called *multi-source reconstruction*, which uses the cross-correlations of all available virtual sources together to fill the gaps of cross-correlation profiles.

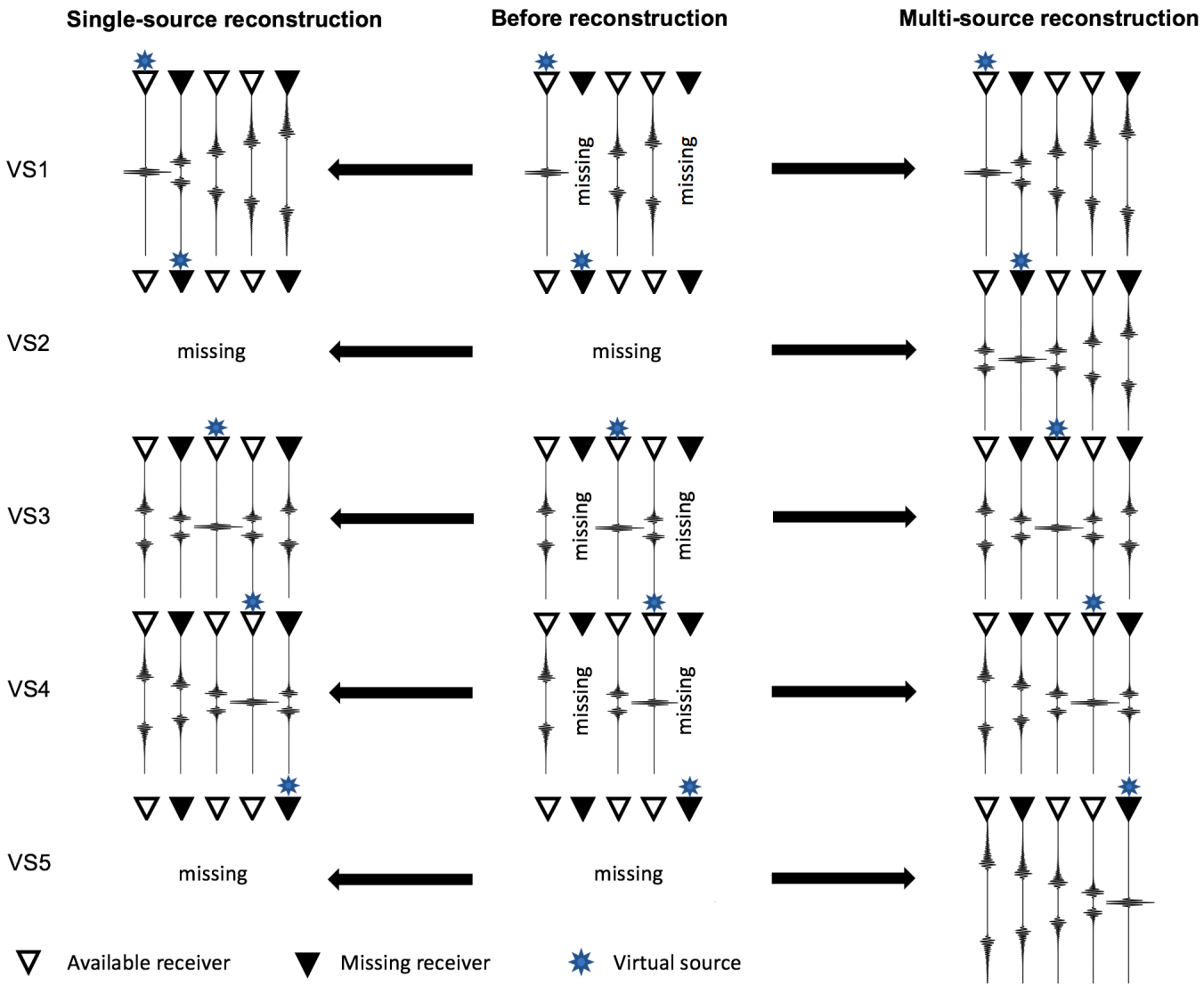


Figure 3.2 Difference between the two schemes of wavefield recovery of correlation profiles of a linear array. The middle column shows the correlation gathers before wavefield reconstruction. Single-source reconstruction (left column) of interferometric wavefields uses correlograms with gaps (i.e. missing the 2^{nd} and 5^{th} receivers) provided only from a virtual source to recover interferometric wavefields inside the gaps. The wavefield recovery for each VS is performed separately. Note that single-source reconstruction cannot recover the correlation profiles for VS2 and VS5 because the master receiver for this virtual source gather is missing. By contrast, multi-source reconstruction (right column) of interferometric wavefields uses correlograms with gaps provided from all available virtual sources to recover interferometric wavefields inside the gaps of all virtual sources including the two missing correlation profiles. The wavefield recovery of all virtual sources is performed simultaneously.

Using a linear array consisting of 3 available and 2 missing receivers as an example, Figure 3.2 illustrates the processes and differences of the single- and multi- source wavefield reconstructions. Each receiver in the array can act as a master receiver (VS) producing a cross-correlation profile that accounts for the cross-correlation of wavefields recorded at the VS and the other receivers. Figure 3.2 shows 3 correlation profiles associated with the 3 available receivers (VS1, VS3, and VS4) before wavefield recovery. Note that for these available profiles, the cross-correlated wavefields at the missing 2^{nd} and 5^{th} receivers are absent. Single-source wavefield reconstruction is a source-by-source recovery operation and the method only uses the cross-correlated wavefields of the available 1^{st}, 3^{nd}, and 4^{th} receivers to reconstruct the correlated wavefields of the missing 2^{nd} and 5^{th} receivers. Thus, using the single-source method, one can reconstruct 3 correlation profiles of the 3 available virtual sources but cannot reconstruct the correlation profiles of the 2 missing virtual sources (Figure 3.2). By contrast, multi-source wavefield reconstruction uses the correlation profiles provided by the 3 available VSs together to recover 5 correlation profiles for all virtual sources (Figure 3.2).

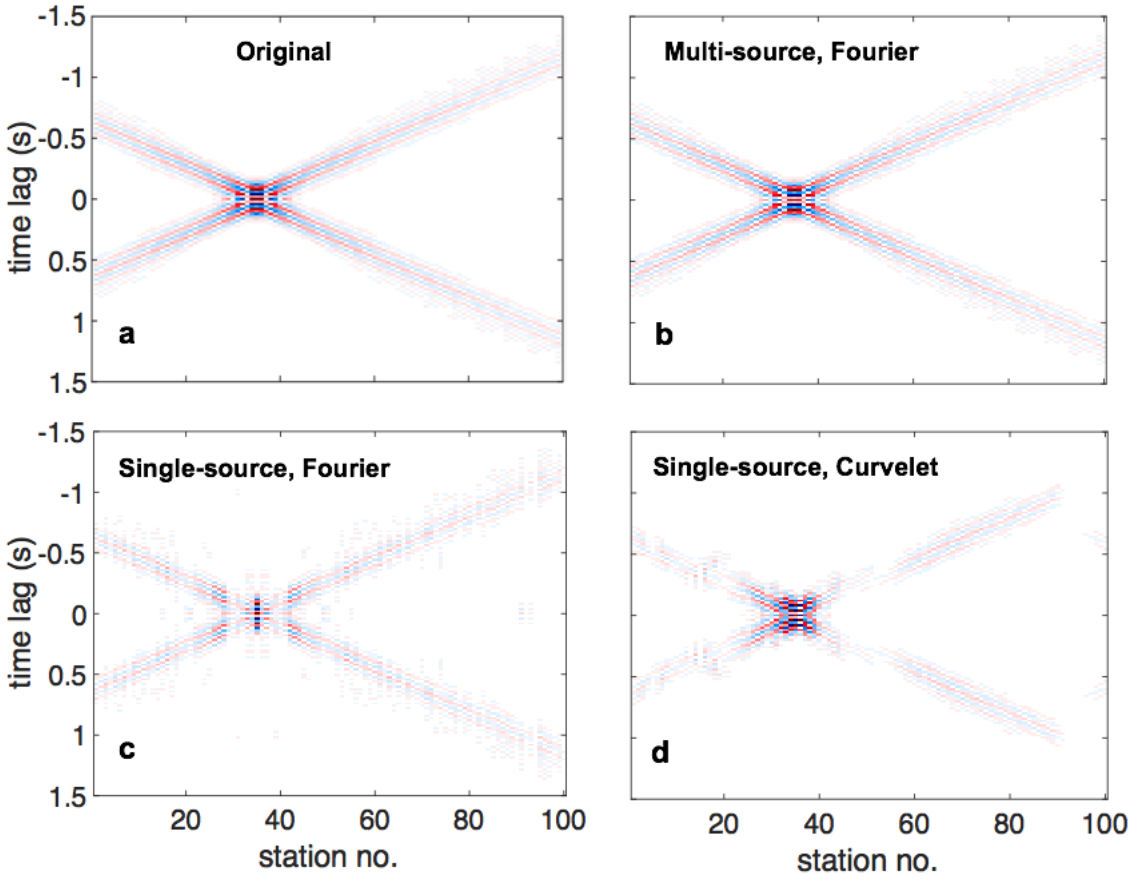


Figure 3.3 Comparison between (a) the original cross-correlated wavefields when the master sensor is the 35th receiver, and the wavefield images recovered by (b) Fourier multi-source reconstruction, (c) Fourier single-source reconstruction, and (d) Curvelet single-source reconstruction. In all reconstructions, 80% of the sensors are absent.

We compare the two schemes of wavefield reconstructions where CS reconstruction of interferometric wavefields is performed after cross-correlation interferometry: single- and multi- source wavefield reconstructions. We use signal-to-noise ratio (SNR) to determine the error of different wavefield reconstructions compared to the original wavefields. We then compare the SNR of the different CS reconstruction schemes. The SNR is defined as $20\log(||x_{ori}||_2/||x_{ori}-x_{rec}||_2)$, where x_{ori} is the original and x_{rec} is the reconstruction of the correlograms. Because multi-dimensional Fourier reconstruction is computationally faster than the Curvelet recovery, we use only the Fourier domain for multi-source reconstruction, while we use both the Fourier and Curvelet domains for single-source reconstruction. In our work, we apply a discrete Curvelet transform with wedge wrapping developed by Candès et al. (2006a), using arbitrary scale and direction parameters given by 4 and 16, respectively.

Figure 3.3 shows the original interferometric wavefield and wavefield reconstruction examples when the master sensor is the 35th receiver and 80% of the receivers are absent from our array; the locations of available and missing receivers are shown using yellow and blue colors in Figure 3.4, respectively. The Fourier and Curvelet single-source reconstructions give inaccurate wavefield reconstruction, in particular within 200 m (20 receivers) from the master sensor (Figure 3.3c and Figure 3.3d). In contrast, the Fourier multi-source reconstruction provides the interferometric wavefields (Figure 3.3b) recovered similarly to the original cross-correlated wavefields (Figure 3.3a).

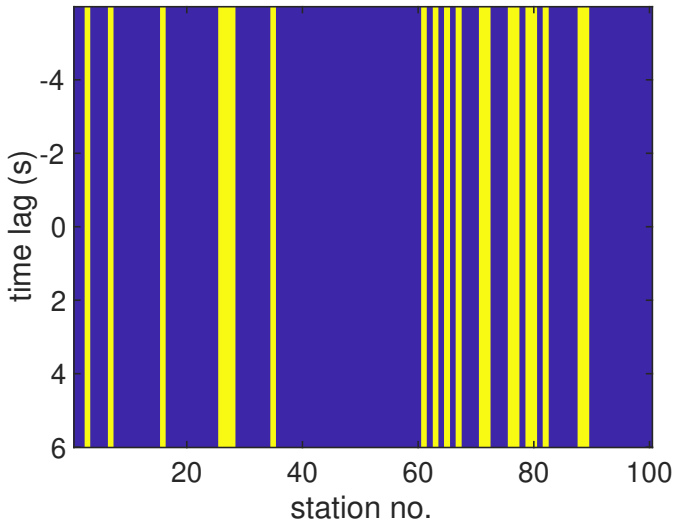


Figure 3.4 Available and missing seismic traces used in our wavefield reconstruction examples (Figure 3.3), shown in yellow and blue, respectively.

Figure 3.5 shows the error of different wavefield reconstructions, which is determined by the relative difference between each reconstruction image (Figure 3.3b, Figure 3.3c, and Figure 3.3d) and the original

image normalized with the maximum amplitude of the original image (Figure 3.3a). Note that the colorbar scale for the single- and multi- source methods are different (Figure 3.5). The relative error of the multi-source reconstruction is below 0.1% (Figure 3.5a) and is much smaller than the error of the single-source reconstruction that ranges between 0-50% (Figure 3.5b and Figure 3.5c). The reconstruction error of the Fourier and Curvelet single-source reconstruction ranges up to 10-40% at all locations along the receiver array (Figure 3.5b and Figure 3.5c). The Fourier and Curvelet single-source reconstructions produce a relative difference of approximately 40% at locations within 200 m from the master sensor. By contrast, the Fourier multi-source reconstruction gives a relative error of less than 0.1% at all locations along the array (Figure 3.5a).

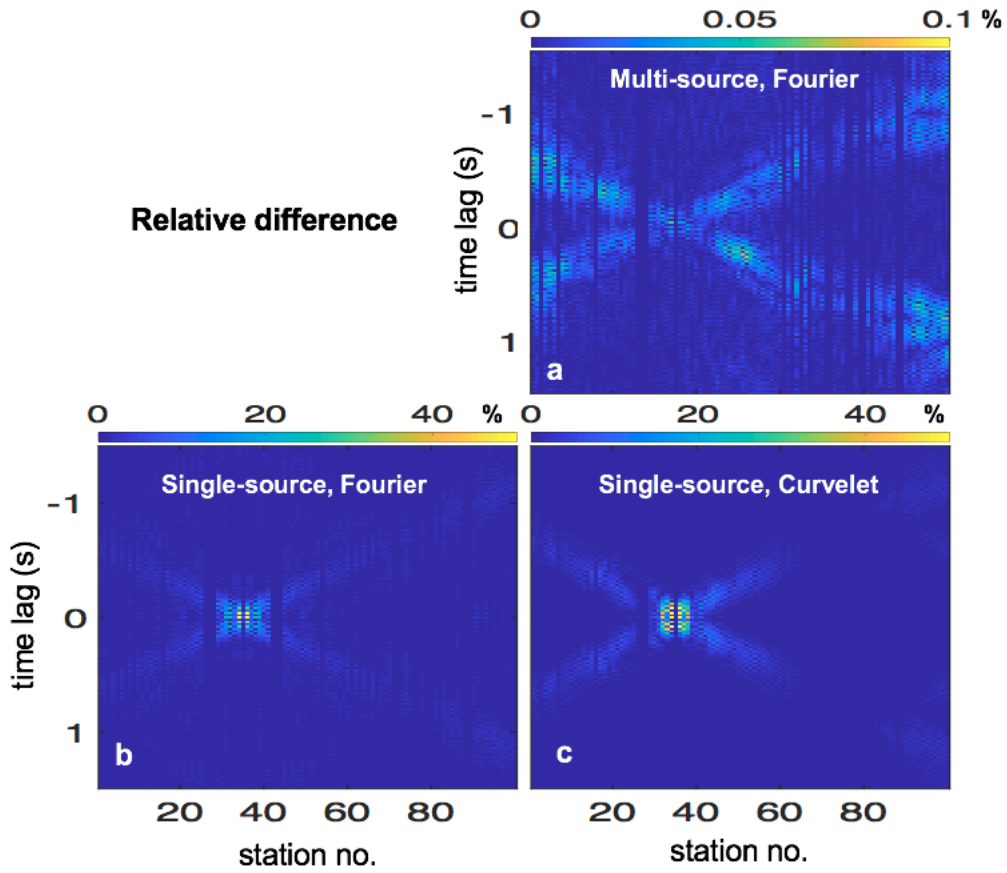


Figure 3.5 The relative differences (%) between the original wavefield image (Figure 3a) and the reconstructed images using (a) Fourier single-source reconstruction, (b) Curvelet single-source reconstruction, and (c) Fourier multi-source reconstruction. The colorbar scales of the two single-source reconstructions are different and much larger than the scale of the multi-source reconstruction.

To investigate the quality of wavefield reconstruction along the receiver array, we determine the SNR for the reconstruction using different master sensors. Figure 3.6 shows the SNR of different reconstruction

schemes averaged over different master sensors. In Figure 3.6, the multi-source reconstruction provides the SNR at every master sensor location, including the locations where the sensors are absent. By contrast, the single-source reconstructions do not provide the SNR when the master receivers are absent because as shown by Figure 3.2, one cannot reconstruct cross-correlated wavefields from missing master receivers using single-source wavefield reconstructions and thus one cannot provide SNR by comparing the reconstructed and original images of correlograms. Figure 3.6 shows that the SNR of the multi-source reconstruction is overall greater than the SNR of the single-source reconstruction by approximately 20-50 dB.

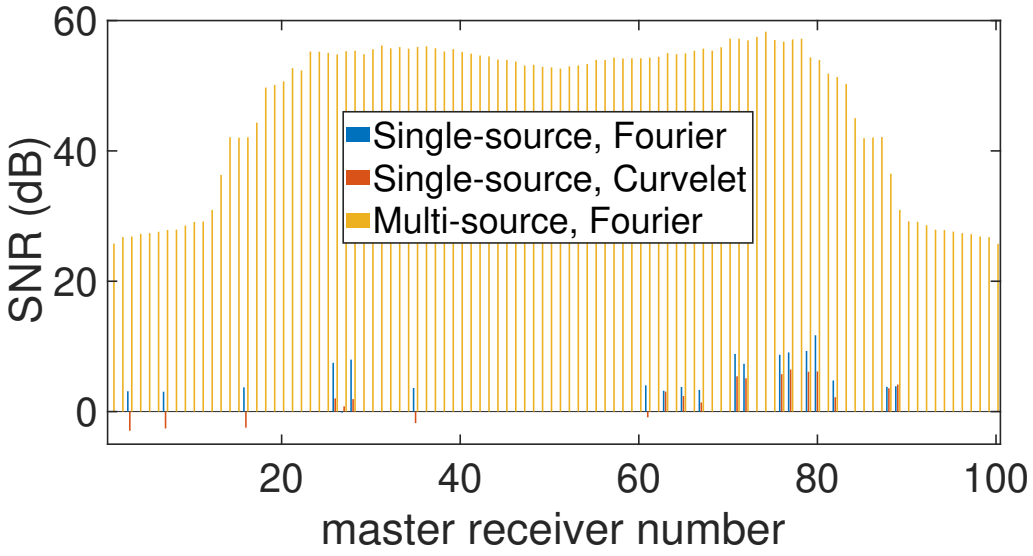


Figure 3.6 SNR in dB that is averaged over different master receivers for different reconstructions across the receiver array. Because the single-source reconstruction uses interferometric wavefields provided only from a VS, the single-source method cannot recover the wavefields when the master receivers are absent as shown in Figure 3.2. Thus, the single-source method cannot provide the correlation profiles and SNR by comparing the reconstructed and the original profiles for these missing master receivers. The multi-source reconstruction, by contrast, can recover the interferometric wavfields for all master receivers.

Using five different realizations of missing receiver locations, we estimate the mean and standard deviation of the SNR for the different reconstruction schemes. Figure 3.7 shows the SNR comparison for the different reconstructions for four different fractions of missing receivers. Overall, the SNR of the Fourier multi-source reconstruction is greater than the SNR of the Fourier and Curvelet single-source reconstructions by approximately 40-60 dB (Figure 3.7). The wavefield recovery using the Fourier multi-source reconstruction overall improves the reconstruction of interferometric wavefields compared to the single-source wavefield reconstructions at different fractions of remaining receivers.

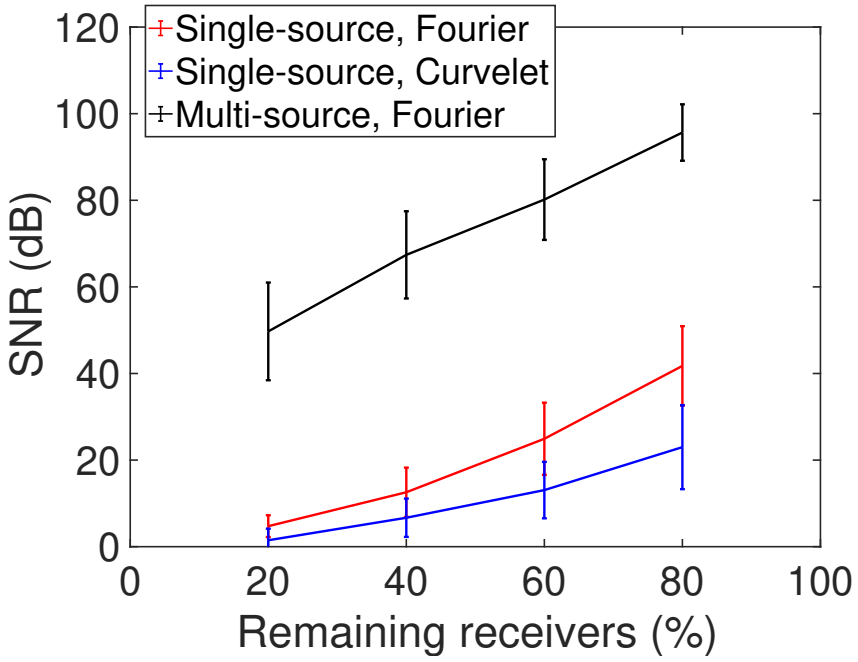


Figure 3.7 Mean and standard deviation of the SNR (dB) for different reconstructions using different fractions of remaining receivers and different missing receiver locations.

3.6 Discussion

We show that one can combine seismic interferometry and compressive sensing, which is a signal interpolation technique, to improve the reconstruction of interferometric wavefields retrieved from interferometry. Our reconstruction of interferometric wavefields takes 2 signal processing steps, where cross-correlation interferometry of seismic traces is performed prior to CS reconstruction of interferometric wavefields.

We propose to perform cross-correlation interferometry prior to CS wavefield recovery because our goal is not to recover noise but rather the Green's function itself. The phase spectrum of noisy signals is random (Figure 3.8a), while the phase spectrum of the correlated noisy signals is consistent as shown by the arrival at zero time lag for the auto-correlated signal in Figure 3.8b. In addition, the time series with the Green's function are short compared to the time series of the noise that is used to extract the Green's function. Thus, reconstruction of random noisy signals is more complicated than reconstruction of their correlations. In this way, performing wavefield reconstruction after interferometry is preferable. Using the Fourier transform, we compare the computational efficiency of the CS wavefield reconstruction performed before and after interferometry in the appendix by comparing the number of the Fourier coefficients associated with the CS reconstruction of noise and cross-correlated wavefields.

When one has two sets of signals A and B where signals in B are processed from signals in A, in some occasions, reconstructing the processed signals B rather than the direct measurements A may provide more computational efficiency and accurate signal recovery. For example, Chen et al. (2015b) and Arianandа and Leus (2012) directly reconstruct the covariance and power spectrum estimated from their original measurements by using less complex computation and minimal memory requirements than reconstructing the original signals before estimating the covariance and power spectrum. The benefits of reconstructing processed signals rather than the original signals are similar to our work, where the direct reconstruction of the cross-correlated signals is more efficient than recovering the original seismic profiles before performing the cross-correlation.

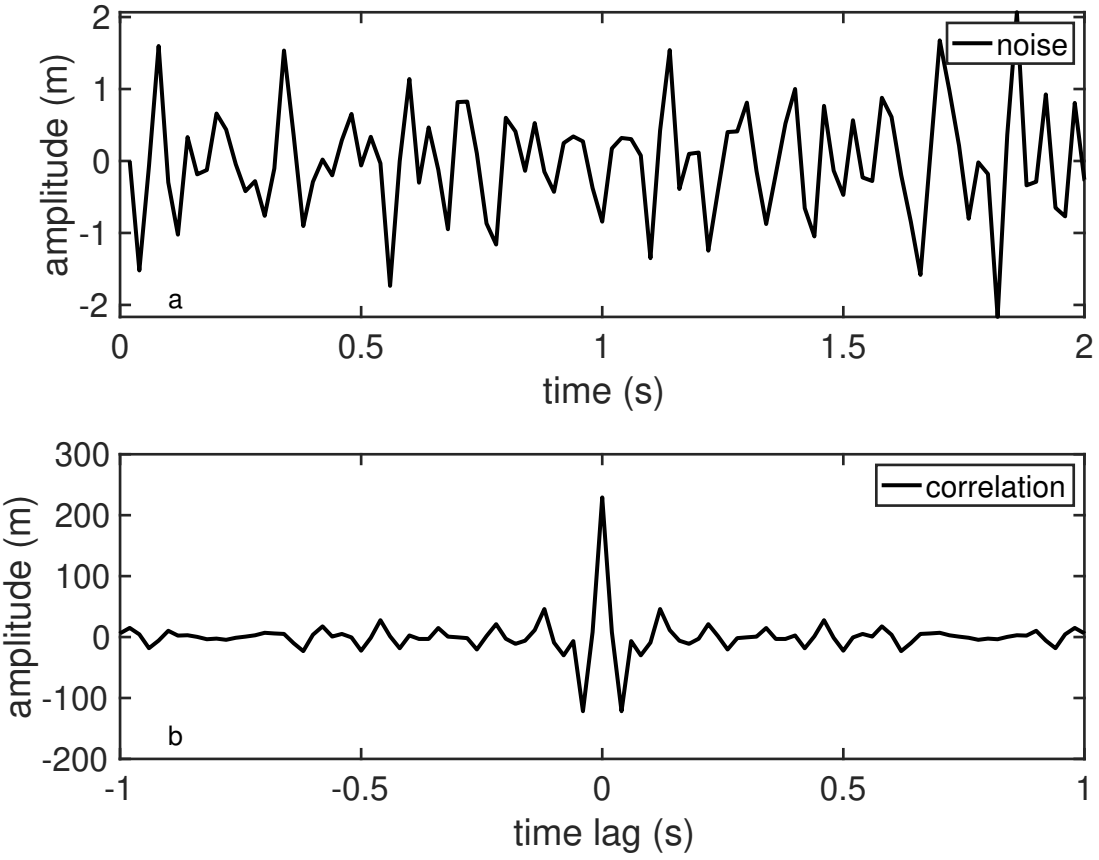


Figure 3.8 a) Noisy signal bandlimited between 5 and 20 Hz and b) auto-correlation of the noisy signal of a), representing seismic ambient noise and correlation of ambient noise, respectively.

In Section 3.5, we show two ways of reconstructing interferometric wavefields after cross-correlation: single- and multi- source reconstructions. One can use single-source reconstruction, which uses cross-correlated wavefields provided from a virtual source. Using SNR as a diagnostic for the quality of wavefield

reconstructions, we show in Section 3.5 that our proposed multi-source reconstruction, which uses cross-correlated wavefields provided from all available virtual sources, improves the recovery of interferometric wavefields compared to the single-source reconstruction. The Fourier multi-source reconstruction gives higher SNR by approximately 40-60 dB, compared to the SNR of the Fourier and Curvelet single-source reconstructions.

The CS signal reconstruction depends on the reconstruction method as well as the sparsity and the sampling of signals (Candès and Wakin, 2008; Wakin, 2017). Since we use the same randomly missing patterns of the receiver array for the single- and multi- source reconstructions, we use a consistent sampling of the signals. Thus, in the reconstruction comparison between the two reconstruction schemes, the wavefield recovery depends on the signal sparsity and the reconstruction method. Dabov et al. (2007) and Maggioni et al. (2013) show that by grouping similar data, which are image and volume fragments, into a volume and a tesseract (4D structure), respectively, one can enhance the sparsity of the data and improve the signal reconstruction and separation of signals and noise.

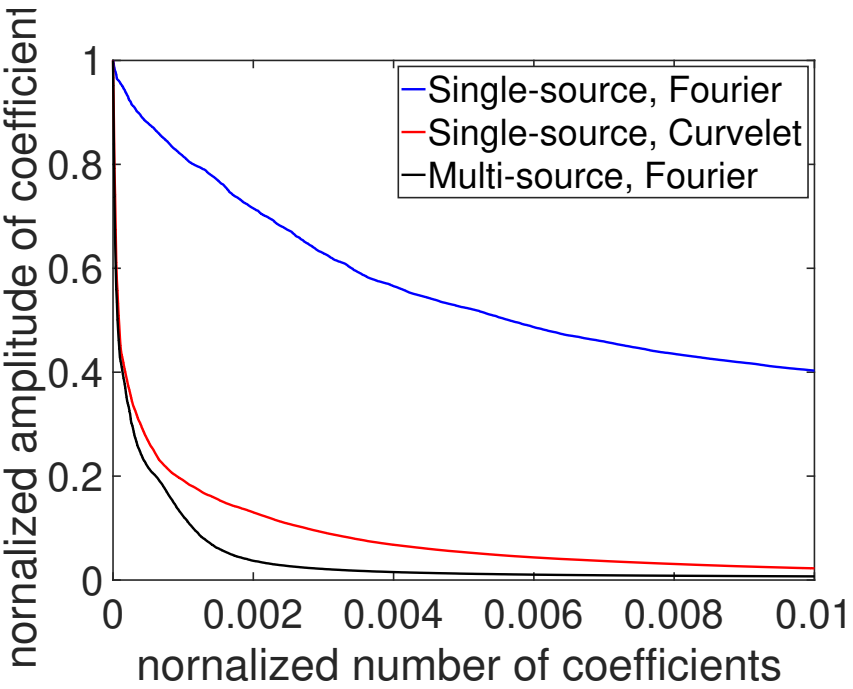


Figure 3.9 Decay of the Fourier and Curvelet coefficients. The graph shows the coefficient amplitudes normalized by the maximum amplitude of the coefficients across the number of the coefficients normalized by the maximum number of coefficients.

The enhancement of signal sparsity by grouping lower-dimensional fragments is similar to our work, where the sparsity of Fourier coefficients is enhanced when we group 2D images of the correlated wavefields from all virtual sources into a 3D volume, compared to the sparsity of the Fourier and Curvelet coefficients

of a 2D image of the wavefields only from a VS. Figure 3.9 shows the absolute value of the Fourier and Curvelet coefficients normalized with the maximum coefficient amplitude against the number of the coefficients normalized with the total number of coefficients. The normalized Fourier coefficients of the 3D volume grouping 2D images of the interferometric wavefields from all virtual sources decay more rapidly than the decay of the normalized Fourier and Curvelet coefficients of a 2D image from a VS. Thus, the coefficients of a 3D volume are sparser than those of a 2D image, improving the wavefield reconstruction. Apart from the sparsity enhancement, we speculate that the improvement of wavefield recovery using the multi-source reconstruction relies on exploiting the redundancy in the correlation that is due to the translational invariance of the interferometric waveforms. For example, the 5^{th} and 55^{th} master receivers give the same correlation response as the response provided from the 10^{th} and 60^{th} master receivers.

In addition, reconstructing a 3D volume rather than a 2D image of the interferometric wavefields requires fewer sensors. For a K-sparse signal of length N, one requires at least M measurements such that $M = O(Klog(N/K))$ to reconstruct the signal. Because the sparsity of a 3D volume is greater than the sparsity of a 2D image as shown by the different decay rates of coefficients in Figure 3.9, we can lower the requirements of M measurements (number of sensors). In our numerical examples of 20% remaining receivers, the single-source reconstruction recovers 100 virtual shots from 20 virtual shots, using 20% of the full data for the recovery. By contrast, the multi-source reconstruction recovers the 10000 virtual shots from 400 virtual shots, using 4% of the full data for the recovery. Thus, the multi-source reconstruction uses fewer relative measurements but gives more accurate reconstruction quality in term of SNR compared to the single-source reconstruction.

The improvement of wavefield recovery using multi-source reconstruction can be beneficial to applications in passive seismic explorations. One can apply the reconstruction technique to seismic profiles from existing linear receiver arrays, where some receivers are absent or malfunctioning. In addition, SI has been applied to seismic data collected from optical fiber networks using distributed acoustic sensing (DAS) technology to investigate and monitor the subsurface (Baird et al., 2020; Dou et al., 2017; Lellouch et al., 2019; Shragge et al., 2021; Zeng et al., 2017). Seismic data recorded using DAS fibers are highly sensitive to the arrival angles of wavefields and the fiber orientation (Martin et al., 2021; Shragge et al., 2021). Thus, the signals recorded from some parts of DAS fibers may be inadequate in signal quality and be discarded from subsequent signal processing steps. One can apply the multi-source reconstruction to mitigate the sensitivity limitation of DAS fibers by reconstructing the interferometric wavefields for the DAS sensors where inadequate-quality signals are recorded.

3.7 Conclusion

We implement a two-stage signal reconstruction, where compressive sensing is performed after cross-correlation interferometry of seismic wavefields. Using the Fourier and Curvelet transforms, we exploit the sparsity of interferometric wavefields for CS reconstruction of correlograms from a linear seismic array. We propose a technique called *multi-source wavefield reconstruction* to reconstruct interferometric wavefield by combining interferometry with compressive sensing. The multi-source method uses interferometric wavefields provided from all available virtual sources to fill the gaps of seismic correlation profiles. Our multi-source method is an alternative improvement of wavefield reconstruction, compared to the traditional single-source method that uses interferometric wavefields provided only from a VS to fill the gaps. The Fourier multi-source reconstruction improves the quality of recovered interferometric wavefields compared to the Fourier and Curvelet single-source reconstruction by a considerable SNR difference of approximately 50 dB. One can apply the multi-source method to recover missing seismic correlation profiles recorded from seismic receiver arrays where some receivers are absent, inoperative, or restricted for installation.

3.8 Acknowledgments

We thank colleagues from Center for Wave Phenomena (CWP) for useful discussions about seismic wavefield reconstructions. We thank the Development and Promotion of Science and Technology Talents Project (Royal Government of Thailand scholarship) for financially supporting P Saengduean to conduct his PhD research. The authors declare that they have no competing interests.

3.9 Data Availability

The numerical data and CS reconstruction used in this article are processed using steps and methods described in Section 3.3 and can be personally requested through P Saengduean. The code for the discrete Curvelet transform with wedge wrapping is available from the Curvelab group at http://curvelet.org/ (Candès et al., 2006a).

3.10 Appendix A: Fourier Coefficients Associated with SI and CS

In Section 3.5, we argue that directly reconstructing correlograms is more efficient than reconstructing original seismic waves prior to performing cross-correlation interferometry of seismic wavefields. Here, we use the Fourier transform to compare the number of the Fourier coefficients associated with CS when 1) one reconstructs seismic wavefields before performing SI (conventional method) and 2) one cross-correlates wavefields before performing CS reconstruction of the interferometric wavefields (our suggested method). These numbers of Fourier coefficients associated with CS reconstruction enable comparison of the computational

efficiency.

We use interferometric wavefields from ambient noise and active-source interferometry for our comparison of the Fourier coefficients. For ambient noise interferometry, the associated coefficients are the length of a noise window in samples (S), number of noise windows (Ws), and the number of receiver from the dense array (N). For active-source interferometry, the coefficients are the length of signals (S), number of sources (Ns), and the number of receivers from the dense array (N).

Table 3.1 shows the required number of Fourier coefficients for the conventional and our suggested methods. When the number of noise windows (Ws) in ambient noise interferometry and the number of sources (Ns) in active-source interferometry are much greater than the number of receivers from the dense array (N), the number of Fourier coefficients required for the CS reconstruction of our suggested method are smaller than the number of coefficients required for the conventional method. Note that the Fourier coefficients shown in Table 3.1 are the total number of coefficients associated with the wavefield reconstruction and are not the same as the number of computational operations such as the fast Fourier transform. In addition, rather than reconstructing the whole time windows of noise wavefields (Figure 3.8a), the window of the correlation of signals (Figure 3.8b) that we reconstruct can be shorter compared to the window of noise wavefields, where the length of this shorter time window is s and $s < S$ (Table 3.1). Thus, our method of directly reconstructing correlograms involves fewer Fourier coefficients and shorter time windows than reconstructing seismic wavefields before performing interferometry, allowing better computational efficiency.

Table 3.1 The number of the Fourier coefficients associated with CS wavefield reconstruction when the reconstruction of seismic wavefields is performed prior to interferometry (conventional method) and the reconstruction is performed after interferometry (our suggested methods). Both methods are carried out for ambient-noise and active-source examples.

Example	Method	Number of the Fourier coefficients associated with CS reconstruction
Ambient noise interferometry	Conventional	S x N x Ws
	Our method	s x N x N
Active-source interferometry	Conventional	S x N x Ns
	Our method	s x N x N

3.11 Appendix B: Multi-Source Reconstruction of 2D Wavefields from an Areal Array

Previous sections of Chapter 3 show that one can apply multi-source wavefield reconstruction to recover correlated wavefields from a linear receiver array. In this section, I show an application of the multi-source method to recover correlated wavefields from an areal seismic array.

3.11.1 Synthetic Model

I use a numerical model to simulate surface waves recorded at a 10×10 receiver array (blue) from 10000 sources located on a red rectangle surrounding the receivers (Figure 3.10). The spacing of the sources and receivers are 4 m and 100 m, respectively. Using the 2D Green's function of the Helmholz equation (Equation 3.6 and Equation 18.46 from Snieder and van Wijk (2015a)), I simulate synthetic surface waves recorded on each receiver. I use the spectrum of a delta function bandlimited between 5 and 20 Hz for the source-time function at angular frequency $W(\omega)$. The phase velocity $c(\omega)$, determined by Xia et al. (1999) for a laterally homogeneous medium, is used for the wavenumber k in Equation 5.3, where $k(\omega) = \omega/c(\omega)$. Applying seismic interferometry to the synthetic surface waves recorded at the receivers (blue array in Figure 3.10), where each receiver can act as a virtual source, I estimate the correlated wavefields which account for the wave propagation between each pair of receivers.

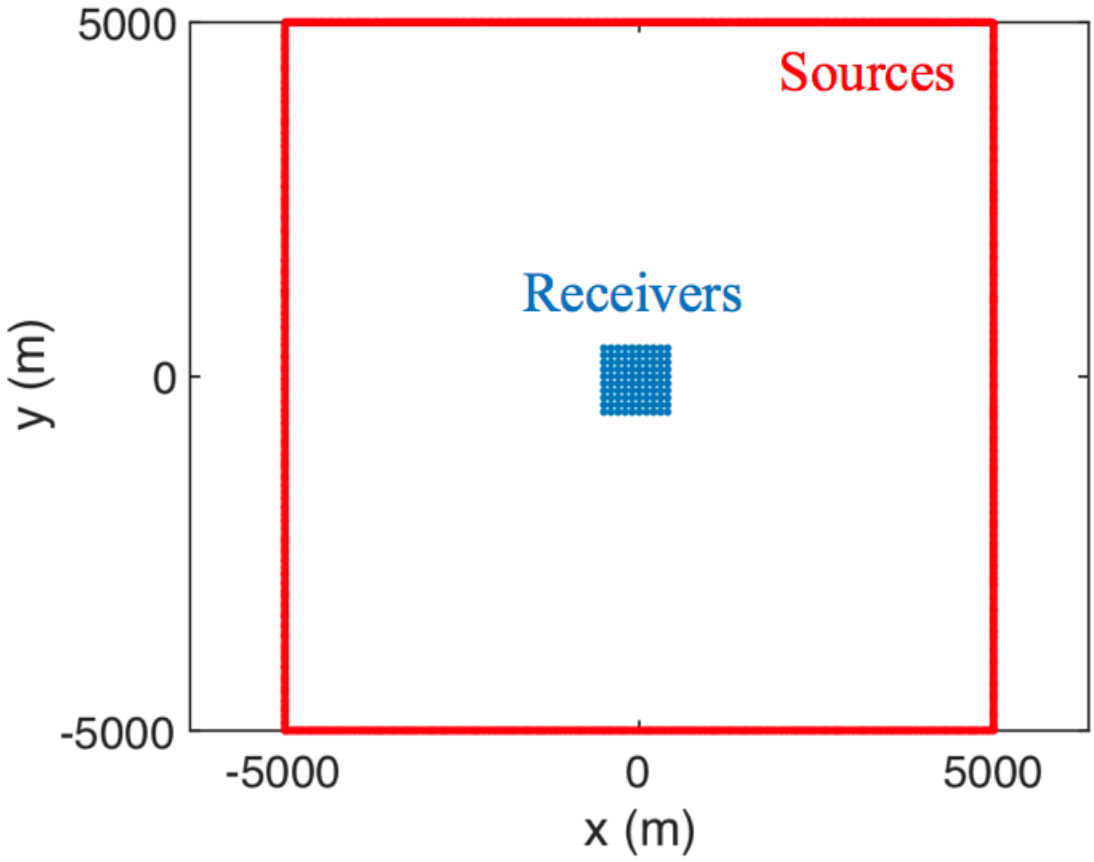


Figure 3.10 Geometry of sources and receivers used to generate synthetic surface waves. Sources are located on the red rectangle surrounding a 10×10 receiver array. Because the source spacing is 4 m which is much smaller than the scale of the figure, the sources appear as a solid rectangle.

3.11.2 Wavefield Reconstruction

In this section, I compare the single-source and multi-source wavefield reconstructions, described in previous sections of Chapter 3, for recovery of correlated wavefields recorded on an areal seismic array. Because the discrete Fourier transform is more computationally efficient than the discrete Curvelet transform for multiple dimensions, I use only the Fourier domain as a sparse domain for recovery of correlated wavefields for the single-source and multi-source methods.

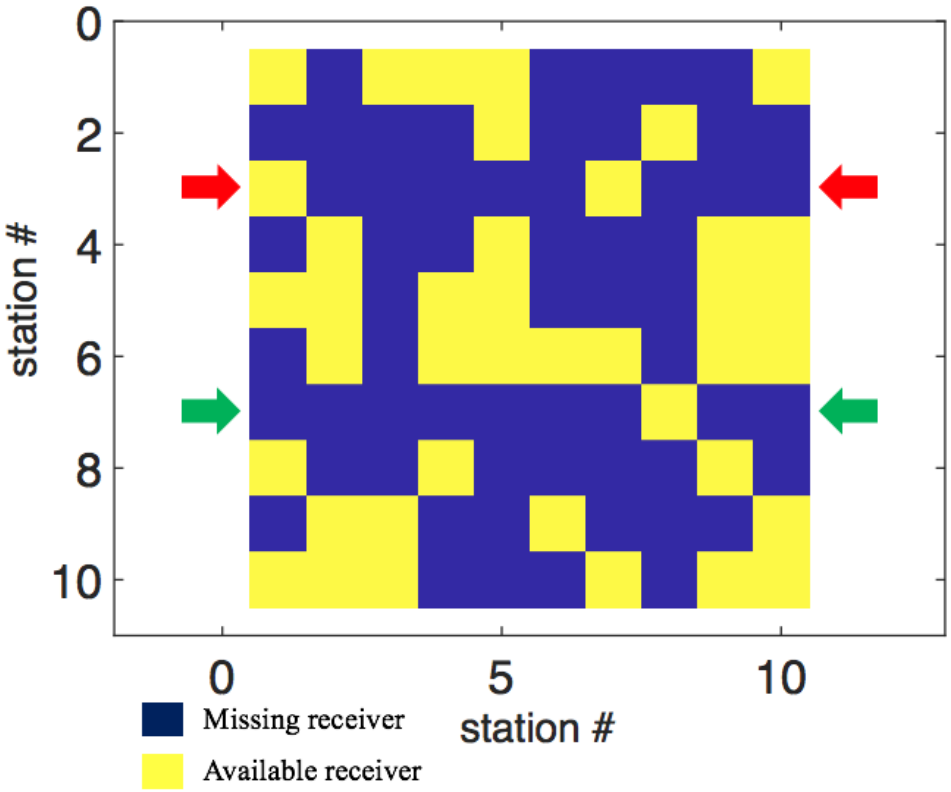


Figure 3.11 Locations of available (yellow) and missing (blue) receivers of the 10×10 receiver array. Because the graphical visualization of correlated wavefields recorded on an areal array is difficult, two linear profiles of correlated wavefields are chosen for the comparison of wavefield recovery using single-source and multi-source wavefield reconstruction (indicated by red and green arrows).

Using a 10×10 receiver array, where only 40% receivers are available, I show an example of wavefield recovery using single-source and multi-source wavefield reconstruction; the locations of missing and available receivers on the array are shown in blue and yellow colors, respectively (Figure 3.11). Because the graphical visualization of correlated wavefields recorded on the areal array is difficult to visualize, I select a linear profile (indicated by red arrows on Figure 3.11) of correlograms to show the comparison of the reconstruction of correlated wavefields using the single-source and multi-source methods.

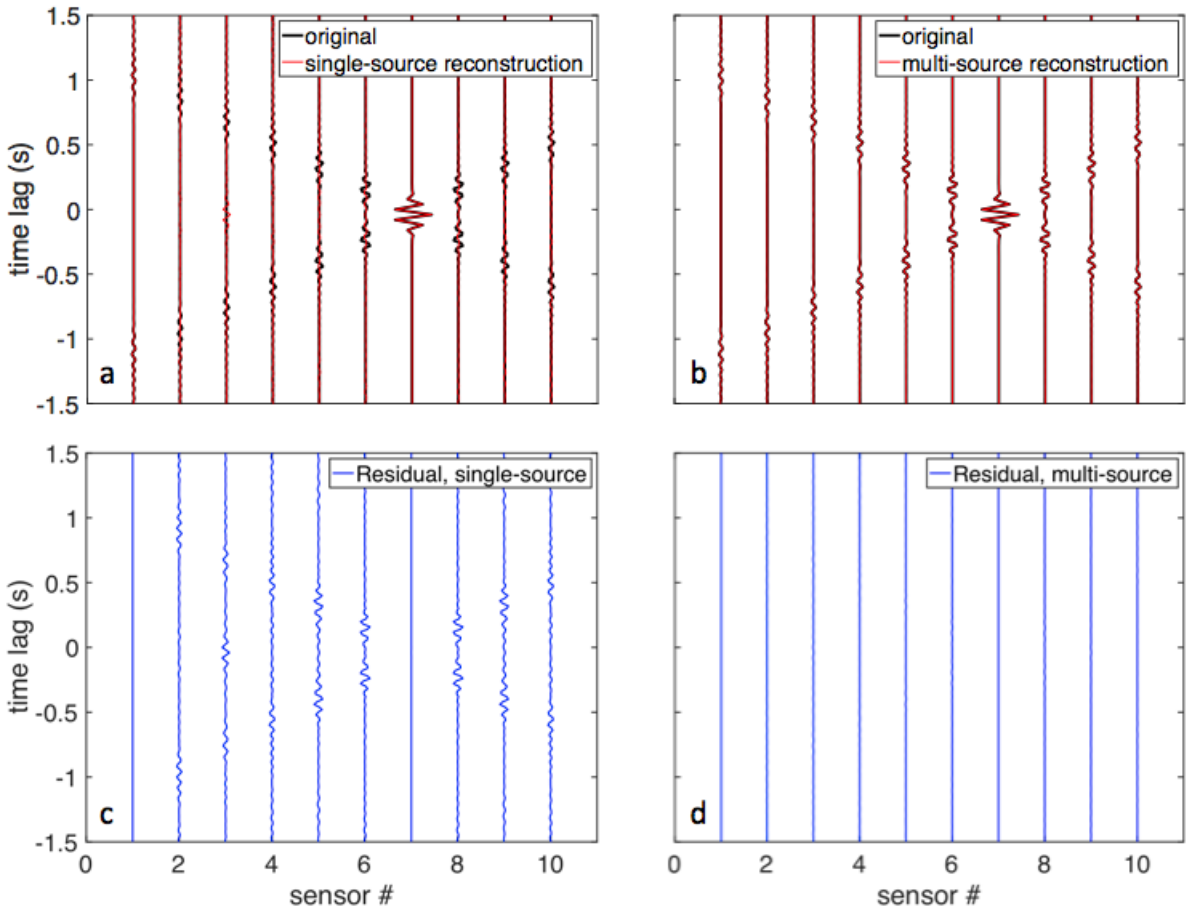


Figure 3.12 Profile of correlated traces, when the 7^{th} receiver is the master receiver, from the selected linear receiver array indicated by red arrows on Figure 3.11. The black original correlated traces are compared to the red traces recovered by (a) single-source and (b) multi-source reconstructions. Residuals (i.e., the difference between the original and the reconstructed traces) of (c) single-source and (d) multi-source reconstructions.

Figure 3.12 compares the correlated traces reconstructed by the single-source and multi-source methods to the original correlated traces recorded on the linear receiver array indicated by red arrows in Figure 3.11 and show the residual of wavefield reconstruction, which is the difference between the original and reconstructed traces, when the 7^{th} receiver is the master receiver. Single-source reconstruction provides accurate wavefield recovery only for the 1^{st} and 7^{th} receivers, which are the available receivers shown in Figure 3.11, but the single-source method recovers inaccurate correlated wavefields for the other missing receivers (Figure 3.12a), because the residuals of single-source reconstruction appear negligible for the 1^{st} and 7^{th} traces, compared to the residuals of the other traces (Figure 3.12c). By contrast, multi-source reconstruction recovers more accurate correlated wavefields than single-source reconstruction because the correlated traces reconstructed by the multi-source method overlap the original traces (Figure 3.12b) and the residuals of multi-source reconstruction appear negligible for all traces, compared to the residuals of single-source reconstruction (Figure 3.12d).

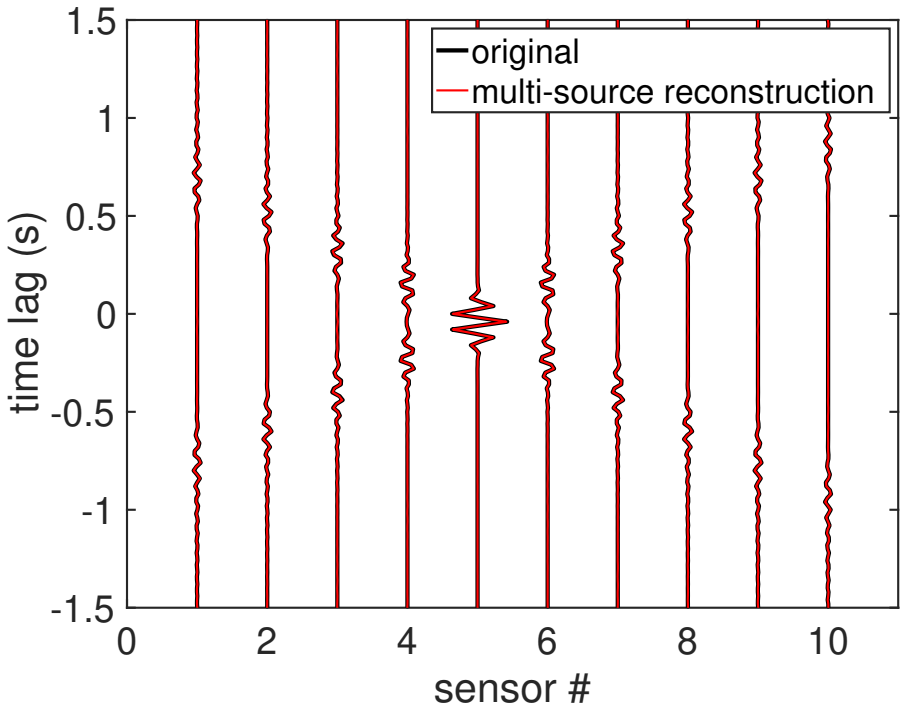


Figure 3.13 Comparison of the original traces (black) and the traces reconstructed by multi-source wavefield reconstruction (red) when the 5^{th} master receiver is missing on the linear profile indicated by green arrows in Figure 3.11.

In addition, as described in Section 3.5, the single-source method cannot reconstruct missing correlograms for the missing master receivers Figure 3.2. Thus, for the example of 10×10 receiver array where 60% receivers are missing, as shown in Figure 3.11, the single-source method can recover only 40 correlograms. In contrast,

the multi-source method can recover correlograms for all receivers including those from the missing master receivers. Figure 3.13 shows the black original correlated traces compared to the red traces reconstructed by the multi-source method for the linear profile, indicated by green arrows in Figure 3.11 when the 5^{th} master receiver is missing. Because the black and red traces overlap, the multi-source method recovers accurate correlated wavefields for the missing master receiver.

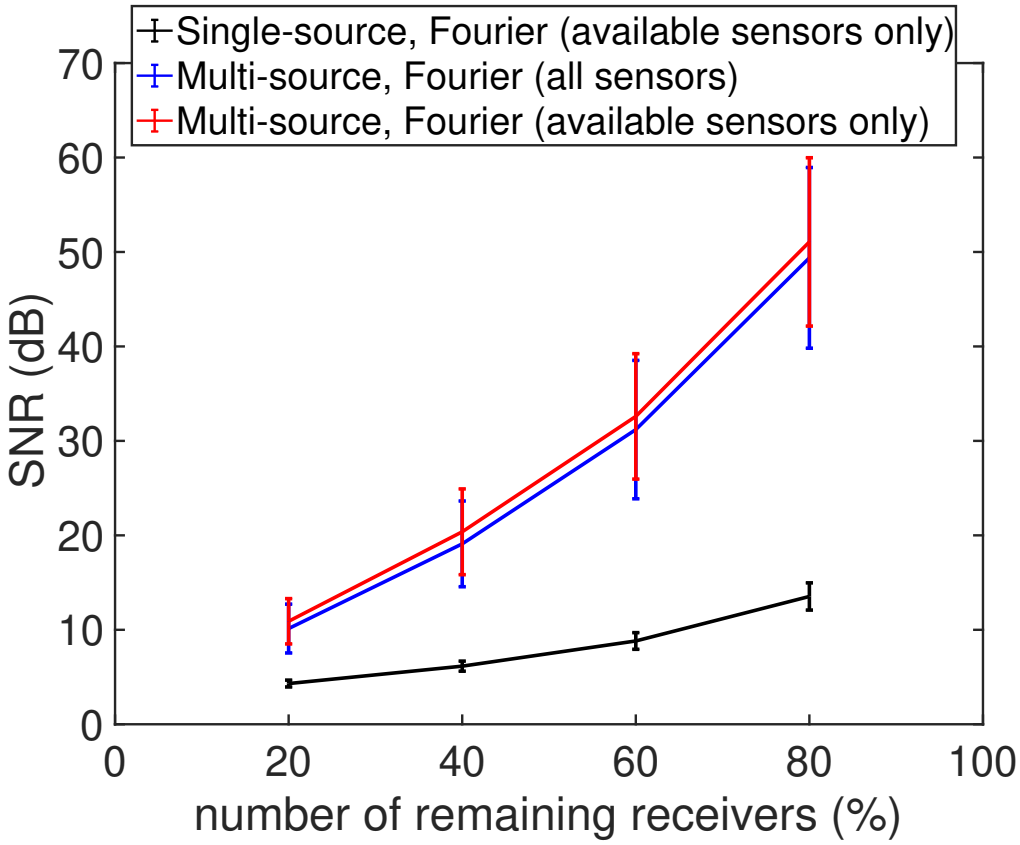


Figure 3.14 Mean and standard deviation of the SNR of Fourier single-source and multi-source reconstructions, using different realizations and fractions of available receivers in the 10×10 array.

I use SNR (described in Section 3.5) to quantitatively compare the reconstruction of correlated wavefields from the single-source and multi-source methods when one uses different realizations and fractions of missing receivers. Figure 3.14 shows the mean and standard deviation of the SNR of single-source and multi-source reconstructions when different realizations and fractions of available receivers in the 10×10 array are used. Because single-source reconstruction cannot recover correlograms of the missing master receivers, I estimate the mean and the standard deviation of the SNR for the single-source method only from available receivers (black line). By contrast, I estimate the mean and the standard deviation of the SNR of multi-source reconstruction from available receivers (red line) and all receivers (blue line). Figure 3.14 shows that

multi-source reconstruction provides more accurate wavefield recovery than single-source reconstruction by approximately 5-40 dB, depending on the fraction of available receivers.

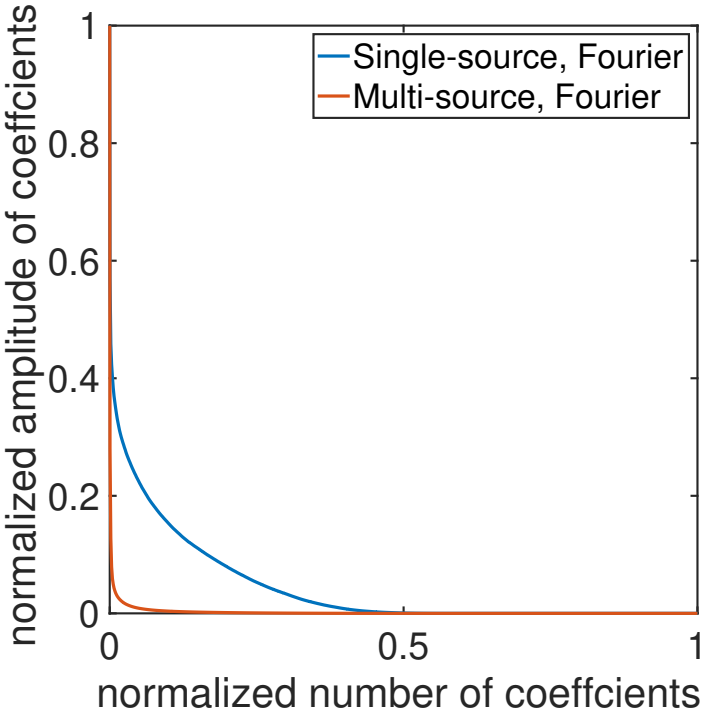


Figure 3.15 Comparison of the decay of the sorted Fourier coefficients between single-source and multi-source wavefield reconstruction. The graph shows the sorted Fourier coefficient amplitudes normalized by the maximum amplitude of the Fourier coefficients across the number of the coefficients normalized by the total number of the coefficients.

Section 3.6 explains that the multi-source reconstruction provides more accurate wavefield recovery because signal sparsity is enhanced by grouping 2D images of correlated wavefields into a 3D volume. For wavefields recorded from an areal array, one groups the 3D volumes of correlated wavefields into a 4D structure for multi-source wavefield reconstruction. Figure 3.15 shows the sorted absolute value of the Fourier coefficients normalized with the maximum amplitude of the coefficients against the number of the coefficients normalized with the maximum number of the coefficients. The decay of the Fourier coefficients of the 4D structure grouping 3D volumes of correlated wavefields from all virtual sources is faster than the decay of the coefficients of a 3D volume from a single virtual source. Thus, the sparsity of a 4D structure is higher than that of a 3D volume, improving the reconstruction of correlated wavefields. Additionally, in Section 3.6, I speculate that the multi-source method exploits the redundancy in the correlation from the translational invariance of correlated wavefields for the linear array example. For areal array cases, the multi-source reconstruction also exploits the redundancy of correlation from azimuthal invariance of correlated wavefields.

For example, the correlation of the 2^{nd} and the 7^{th} receivers on the west-east direction is the same as the correlation of the 4^{th} and 9^{th} receivers on the north-south direction.

CHAPTER 4

MULTI-SOURCE WAVEFIELD RECONSTRUCTION OF DAS DATA USING COMPRESSIVE SENSING AND SEISMIC INTERFEROMETRY

Modified from a paper to be submitted to the Journal of Acoustical Society of America.

Patipan Saengduean[8], Jihyun Yang[9], Jeffrey Shragge [10], Roel Snieder[11], and Michael B. Wakin[12]

4.1 Abstract

Seismic data recorded by distributed acoustic sensing (DAS) interrogator units on deployed optical fiber is increasingly being used for a variety of subsurface imaging and monitoring investigations. To reduce the costs of active-source surveying for DAS applications and to allow for quasi-continuous recording, seismic interferometry can be applied to estimate inter-sensor wavefields from DAS records. However, recording long-term DAS records for ambient interferometry requires considerable data storage and parts of DAS fibers may also be unusable because of broadside sensitivity considerations from the DAS fiber configuration and due to localized coherent energy sources with amplitudes significantly larger than the ambient signal of interest. Compressive sensing, a wavefield reconstruction technique exploiting wavefield sparsity, can mitigate the problems of large data storage and unusable data. We apply compressive-sensing-based multi-source wavefield reconstruction to estimate correlograms of ambient DAS records from an urban fiber array located in Perth, Australia. The multi-source method uses correlograms from all available virtual sources for simultaneous wavefield reconstruction and is different from the conventional single-source method that separately and serially reconstructs correlated wavefields from each virtual source. Using the Fourier and Curvelet transforms to sparsify correlated wavefields, we show that the multi-source method can be applied to the DAS data and that the Fourier multi-source reconstruction can improve the recovered wavefields by approximately 5-10 dB, compared to the Fourier and Curvelet single-source wavefield reconstructions.

4.2 Introduction

Distributed acoustic sensing (DAS) is a technology increasingly used in geophysics investigations to replace traditional geophones with an optical fiber to investigate the subsurface by recording strain along the fiber. DAS is widely applied to vertical seismic profiles (VSP) (Martinez et al., 2021; Nizkous et al., 2015;

[8]Graduate student, primary researcher, and author at Department of Geophysics, Colorado School of Mines
[9]Graduate student at Department of Geophysics, Colorado School of Mines
[10]Associate Professor at Department of Geophysics, Colorado School of Mines
[11]Supervisor at Department of Geophysics, Colorado School of Mines
[12]Supervisor at Department of Electrical Engineering, Colorado School of Mines

Young et al., 2022), micro-seismic monitoring (Karrenbach et al., 2017; Walter et al., 2020), and monitoring of reservoirs and hydraulic fracturing (Bakku et al., 2014; Mateeva et al., 2014).

To reduce costs for subsurface investigations using active-source profiling, seismic interferometry (SI), a technique for estimating inter-sensor wavefields, can be applied to DAS applications for passive surveys. Examples of DAS-based passive-source applications include near-surface monitoring from traffic-induced vibrations (Dou et al., 2017), time-lapse monitoring of geothermal fields using borehole signals (Chang and Nakata, 2022), monitoring of CO_2 sequestration using DAS-VSP interferometric imaging (Sidenko et al., 2021), subsurface investigation using surface-wave inversions from ambient waveform data (Shragge et al., 2021), and ocean-current observation using data from ocean-bottom fiber deployments (Williams et al., 2022).

One requirement for accurate estimation of surface waves is that SI requires long seismic ambient records to extract the waves that propagate between sensors (or sensing locations along a fiber). In practice, such records of long-term ambient waveforms require considerable storage and may not be fully available due to interference arising from strong but highly localized energy sources, and the insensitivity of DAS fibers to the angle of incoming wavefield energy with wave motion in the direction highly oblique to the fiber orientation (Bakku et al., 2014; Martin et al., 2021; Shragge et al., 2021). Thus, parts of a DAS fiber may not record usable wave-motion data, making the seismic record effectively irregular. The irregularization of DAS data can affect subsequent seismic analyses (e.g., full-wavefield imaging and inversion methods) which benefit from regular seismic profiles to enhance the signal-to-noise ratio (SNR) of subsurface images (Al-Gain et al., 2020; Chopra and Marfurt, 2013; Poole and Herrmann, 2007). However, one may perform signal reconstruction or regularization to mitigate the irregularization and data storage volume challenge.

Compressive sensing (CS) is a signal reconstruction technique that exploits signal sparsity to recover dense regularized wavefields from irregularly undersampled wavefields (Candès et al., 2006c). Mosher et al. (2014) and Zhan et al. (2018) show that one can reconstruct regularized seismic waves when seismometers are missing or unusable from regularly sampled arrays. CS is applied to recover wavefields from irregular marine and land seismic profiles (Jiang et al., 2019b; Mansour et al., 2012; Pawelec et al., 2019) and is used for wavefield interpolation and denoising for seismic exploration applications (Hennenfent et al., 2010; Herrmann et al., 2008). Muir and Zhan (2021a,b) use the Curvelet transform for the CS applications of field DAS data. Saengduean et al. (2022) use CS to recover regular interferometric wavefields and improve the SNR of the reconstruction of interferometric wavefields by using signals from multiple virtual sources; the multi-source method also allows source interpolation at unavailable locations.

We apply the multi-source CS reconstruction of interferometric wavefields proposed by Saengduean et al. (2022), which can recover regularly sampled correlated wavefields from an irregular seismic array, to help

alleviate the limitation of irregularization of interferometric DAS data due to unusable sensors and limited sensor sensitivity. Section 4.3 presents the acquisition and processing of the DAS data used for our wavefield reconstruction. In Section 4.4, we describe the basic theory of CS and the multi-source wavefield reconstruction proposed by Saengduean et al. (2022). We then show the reconstruction of interferometric DAS wavefields in Section 4.5. Finally, we respectively discuss the wavefield-recovery improvement from the application of multi-source wavefield reconstruction and the potential applications of the mulit-source technique in Sections 4.6 and 4.7.

4.3 Acquisition and Processing of DAS Data

To test the application of multi-source wavefield recovery on DAS records, we apply wavefield reconstruction to ambient virtual shot gathers (VSG) of DAS data generated by Shragge et al. (2021) following a seismic interferometry approach (Bensen et al., 2007; Shapiro et al., 2005; Snieder, 2004; Wapenaar et al., 2010). The seismic data are acquired from a DAS fiber deployed at approximately 0.5 m depth in the suburban area of Perth, Australia in a buried utility conduit. The usable deployed fiber length is 15 km and Figure 4.1 shows parts of the deployed DAS fiber (white line). The DAS data set was acquired in 19-20 December 2017 using an earlier prototype of the Terra15 Treble interrogator unit. The sampling rate of the acquired DAS data is 250 Hz and the waveform outputs are at 10 m receiver spacing with a 100 m gauge length applied as a post-acquisition filtering operation. We use the 4 km linear section of the DAS fiber (indicated by the red rectangle in Figure 4.1) that is perpendicular to the coast for seismic interferometry to minimize the sensitivity problem where the orientation of the DAS fiber is highly oblique to that of the microseism energy that largely propagates in an easterly direction.

The DAS data were acquired in 60 s windows roughly every 15 minutes from 7:30 pm and 1:00 am on a night where Perth had a winter storm with high ocean swell and significant wave activity. The lowpass filter between 0.04 and 1.80 Hz is applied to the acquired DAS data before smoothing this data with a 3×3 2D triangular-smoothing convolution operator (Shragge et al., 2021). We apply SI to the smoothed ambient noise DAS data to generate a VSG volume. Using the Fourier-domain representation, one can estimate a volume C of VSG by cross-correlating noise wavefields u recorded at two receivers, one of which acts as a virtual source excited at position x_a and the other as a receiver at position x_b. In theory, the ambient waveforms need to be generated on a closed surface that includes uncorrelated sources on the surface to generate the waves that propagate between x_a and x_b (Snieder et al., 2007; Wapenaar and Fokkema, 2006). For ambient seismic interferometry, one can generate the VSG by stacking the cross-correlation of ambient

waveform windows taken from N long ambient wavefield records,

$$C(\tau, x_b, x_a) = \mathcal{F}^{-1}\left[\frac{1}{N}\sum_{n=1}^{N} u(\omega, x_b, n)u^*(\omega, x_a, n)\right], \tag{4.1}$$

where τ is the correlation lag (s), ω is the angular frequency, n is the index of ambient waveform windows, $\mathcal{F}^{-1}$ denotes the inverse Fourier transform, and the asterisk denotes complex conjugation in the frequency domain. In our work, we stack 20 minutes of correlated 60 s ambient waveform windows when the storm-induced energy is the strongest.

Figure 4.1 Map of the Western Suburbs of Perth, Australia with a deployed DAS fiber (white line). The red box indicates the 4 km linear section of the DAS fiber that is used for seismic interferometry and our wavefield reconstruction example. Modified from Figure 1 in Shragge et al. (2021).

4.4 Compressive Sensing and Multi-Source Wavefield Reconstruction

We apply multi-source wavefield reconstruction developed by Saengduean et al. (2022) to the VSGs of the Perth DAS data set. The multi-source method uses CS, which is a wavefield reconstruction technique that recovers signals from few irregular samples (Candès et al., 2006c; Donoho, 2006; Wakin, 2017).

Using CS, one can recover a signal f by expanding the signal in its sparse or compressible domain Ψ:

$$f = \Psi\alpha, \tag{4.2}$$

where α denotes the sparse coefficients. For example, seismic signals can be represented using the Discrete Fourier Transform coefficients. An N-dimensional signal f is sparse when the number K of non-zero coefficients is much less than N, and f is compressible when the sorted coefficients decays rapidly to zero (Wakin, 2017). One can recover the sparse signal f without information loss if the sampling function Φ describing how the signals are sampled ensures the stability and energy conservation of sparse signals; examples of this

sampling function are Gaussian and sub-Gaussian matrices (Candès and Wakin, 2008; Donoho, 2006; Wakin, 2017). In our work, we aim to recover the the correlogram signal f, while the sampling function Φ prescribes the locations of receiver pair on a linear section of a DAS fiber where correlograms are available.

Recovering a K-sparse signal f using CS requires at least M compressive measurements d (Candès and Wakin, 2008; Wakin, 2017), where $d = \Phi f$ and

$$M = O(K \log(N/K)). \tag{4.3}$$

Rani et al. (2018) review different algorithms to solve the underdetermined equation $d = \Phi f$ for a solution f that is sparse in the desired transform domain Ψ. For our work using the multi-source method, we solve the following l_1 optimization problem to reconstruct the signal from noisy data measurements $d = \Phi f + n$:

$$\min ||\alpha||_1 \text{ subject to } ||\Phi\Psi\alpha - d||_2 \leq \sigma, \tag{4.4}$$

where σ is a constant bound of the added noise: $||n||_2 \leq \sigma$. One can recover the signal f once the sparse coefficients α are recovered, using $f = \Psi\alpha$.

For our CS wavefield reconstruction, we use the Fourier and Curvelet transforms for the sparse domains Ψ of our correlated DAS data. We use the 2D and 3D Fourier transforms across the time-space and time-space-space dimensions of correlated DAS data, respectively; reconstructing time-space-space wavefields using the 3D Fourier transforms ideally leads to the recovery of the band-limited Green's function $G(t, x_a, x_b)$ for time t at locations x_a and x_b. The discrete 2D and 3D Fourier transforms are equivalent to repeated discrete 1D Fourier transforms along each dimension (Nussbaumer, 1982). We also apply the 2D Curvelet transform for our CS wavefield recovery. The Curvelet transform uses scale, location, and direction parameters to represent a signal (Candès et al., 2006b; Candès and Demanet, 2005; Hennenfent et al., 2010; Stehly et al., 2011). We recover correlated DAS data (i.e., VSG) using the discrete Curvelet transform with wedge wrapping developed by Candès et al. (2006b).

One can interpolate or reconstruct the missing wavefield from a virtual source gather by using interpolation techniques (e.g., compressive sensing); this modality of wavefield recovery is called *single-source reconstruction*. The *multi-source wavefield reconstruction* developed by Saengduean et al. (2022) applies CS to recover correlated wavefields using correlograms provided from all available virtual sources. Figure 4.2 shows the difference between single-source and multi-source wavefield reconstructions. The panels illustrate examples of five receiver locations with three available and two unusable receivers for wavefield recovery; the second and fifth receivers (virtual sources) are unusable. The three available virtual sources produce three correlograms while two correlograms from VS2 and VS5 are missing because the master receivers for these VSG profiles are assumed to be unusable. The single-source method recovers the correlated traces from the two unusable receivers for each available correlogram but cannot recover the correlograms of VS2 and VS5

due to the unusable master receivers for these VGS profiles (left column of Figure 4.2). By contrast, using three available correlograms (VS1, VS3, and VS4), the multi-source reconstruction method simultaneously recovers the five correlograms of all receivers, including the unusable correlograms of VS2 and VS5 (right column of Figure 4.2).

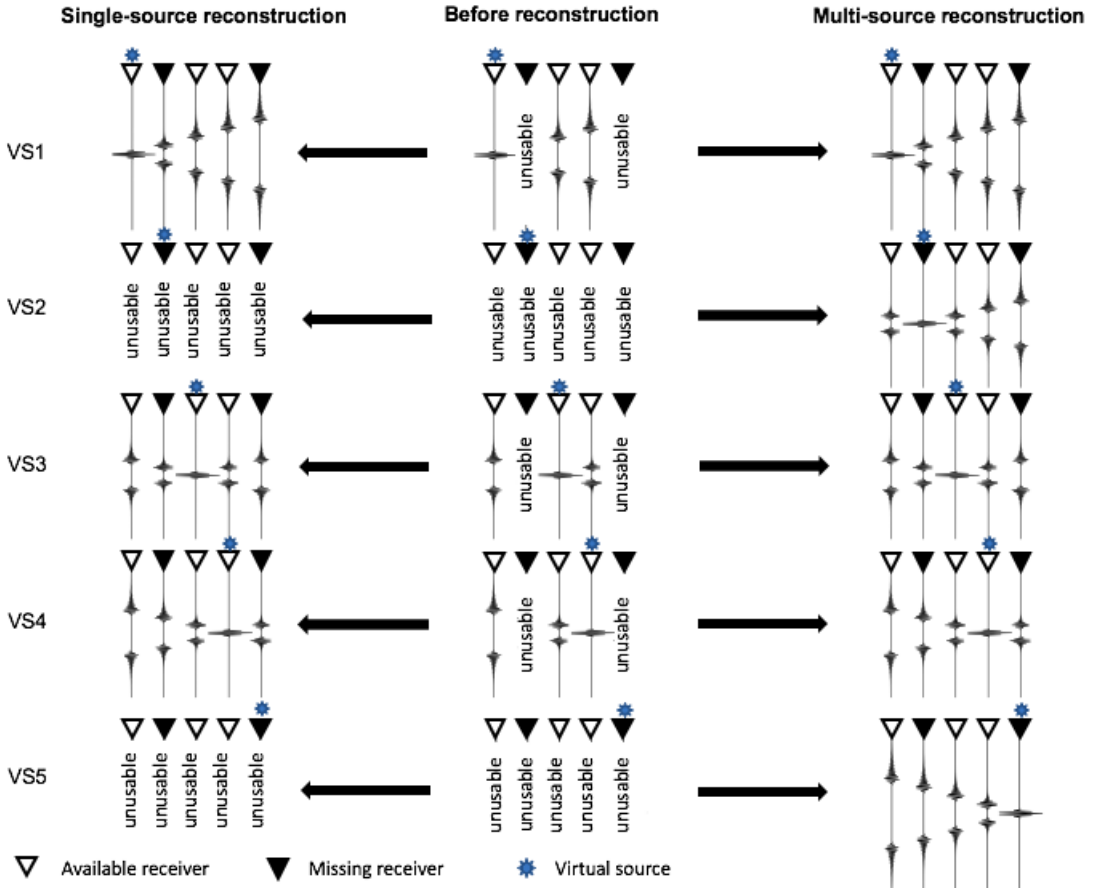


Figure 4.2 Single-source and multi-source reconstruction of correlated wavefields recorded on linear arrays of three available and two unusable receivers; the 2^{nd} and 5^{th} receivers are unusable. Single-source reconstruction (left column) is performed separately for three correlograms provided by three available receivers and recovers the two unusable correlated traces for each available correlogram. Note that the single-source method cannot reconstruct the correlograms of VS2 and VS5 because the master receivers for those VSGs are unusable. The multi-source method (right column) is performed simultaneously and uses correlograms with gaps provided by three available virtual sources to reconstruct correlograms of all virtual sources including the unusable VS2 and VS5 profiles. Adapted from Figure 2 in Saengduean et al. (2022).

4.5 Comparison of Different Wavefield Reconstructions

We apply the multi-source method to correlated DAS data, where the wavefield is missing at a number of recording points. We compare the reconstructed DAS correlograms using single-source and multi-source wavefield reconstructions to the original correlated wavefields. For the single-source reconstruction method,

we use the 2D Fourier and 2D Curvelet transforms as described in Section 4.3. Because the Fourier transform is more computationally efficient than the Curvelet transform for multiple dimensions, we use only the 3D Fourier transform for the multi-source method.

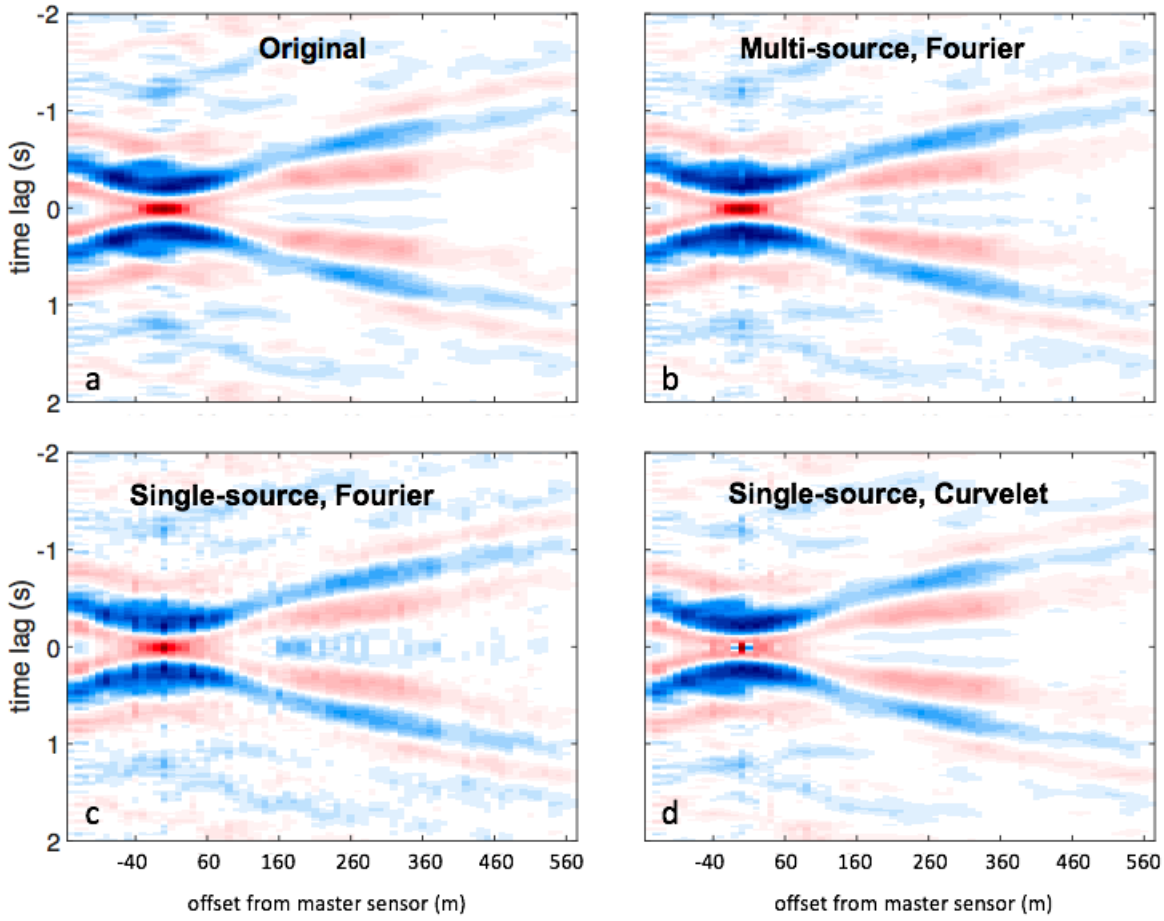


Figure 4.3 Images of (a) original correlograms when the 14^{th} sensor is the master receiver, and the correlograms recovered by (b) the Fourier multi-source reconstruction, (c) the Fourier single-source reconstruction, and (d) the Curvelet single-source reconstruction. 60% of the sensors are absent for each reconstruction in (b)-(d). The offset distance is from the master sensor along the linear segment of the DAS fiber.

We compare the recovery of wavefields using single-source and multi-source wavefield reconstructions with DAS data. Figure 4.3 shows the original cross-correlated wavefield and the reconstructions of cross-correlated wavefields using the single- and multi- source methods where the 14^{th} sensor is set as the master receiver. For each reconstruction in Figure 4.3, 60% of the sensors are assumed to be unusable, where the unusable and available sensors are indicated by blue and yellow colors, respectively, in Figure 4.4. The Fourier and Curvelet single-source reconstructions provide less accurate recovery of correlated wavefields, particularly within 200 m (20 sensors) from the location of the master sensor (Figure 4.3c and Figure 4.3d), compared

to the wavefield recovery by the Fourier multi-source method (Figure 4.3b).

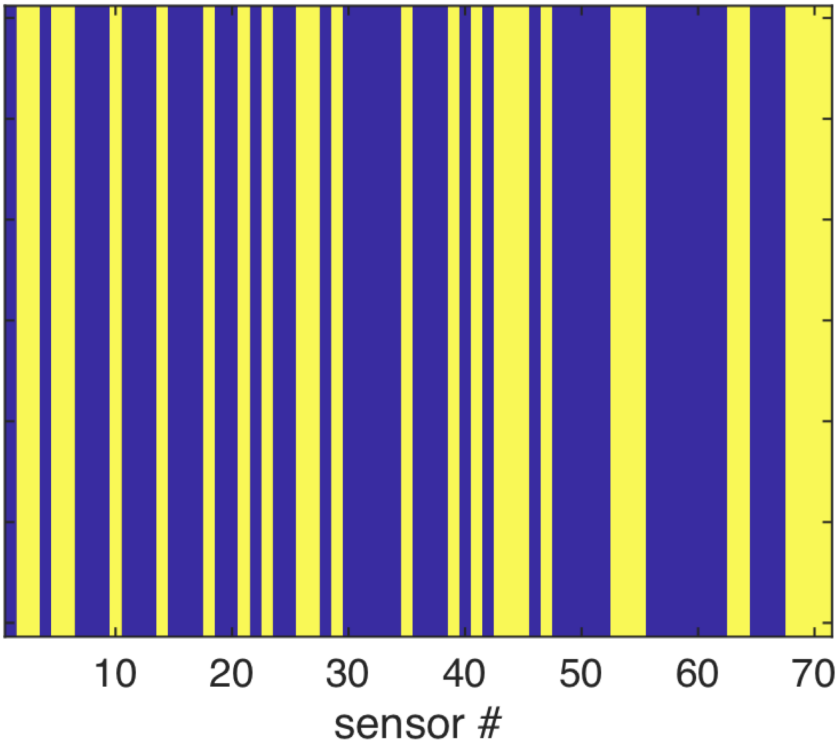


Figure 4.4 Locations of the assumed unusable (blue) and available (yellow) seismic traces in our wavefield recovery examples (Figure 4.3). The availability of traces does not depend on time lag, hence the vertical structure of the blue and yellow bands.

We use relative difference between each wavefield reconstruction (Figure 4.3b, Figure 4.3c, and Figure 4.3d) and the original image normalized to the maximum amplitude of the original image (Figure 4.3a) to quantify the error of wavefield recovery; Figure 4.5 shows the reconstruction errors of each wavefield recovery that range between 0% and 40%. The error in the single-source reconstruction of correlated wavefields (Figure 4.5b and Figure 4.5c) at locations within 200 m from the master sensor is approximately 30-40%. By contrast, the multi-source reconstruction produces a reconstruction error of less than 10% at all locations along the linear sensor arrays (Figure 4.5b).

Both the single-source and multi-source methods can recover the correlated wavefield for the 14^{th} master sensor because this virtual source is available in our reconstruction example as illustrated by Figure 4.2. Next, we test the reconstruction of correlograms for the 50^{th} master sensor, which is unusable as indicated by Figure 4.4. Because the virtual source is unusable, the single-source method cannot recover the correlogram for this virtual source whereas the multi-source method can recover this missing correlogram. Figure 4.6 compares the original correlated wavefields to those reconstructed by the multi-source method. These results show the reconstruction error shown by the relative difference between the original and reconstructed

correlogram. The multi-source method recovers a wavefield image (Figure 4.6b) that closely resemble that of the original wavefield (Figure 4.6a). The error of the multi-source wavefield recovery is below approximately 20% throughout the wavefield image (Figure 4.6c). Although this error is higher than the previously shown reconstruction error of less than 10% for the available 14^{th} master sensor (Figure 4.5a), the multi-source wavefield recovery can reconstruct the completely missing correlograms due to the missing virtual sources while the single-source method is precluded from recovering wavefield information at any missing virtual source location.

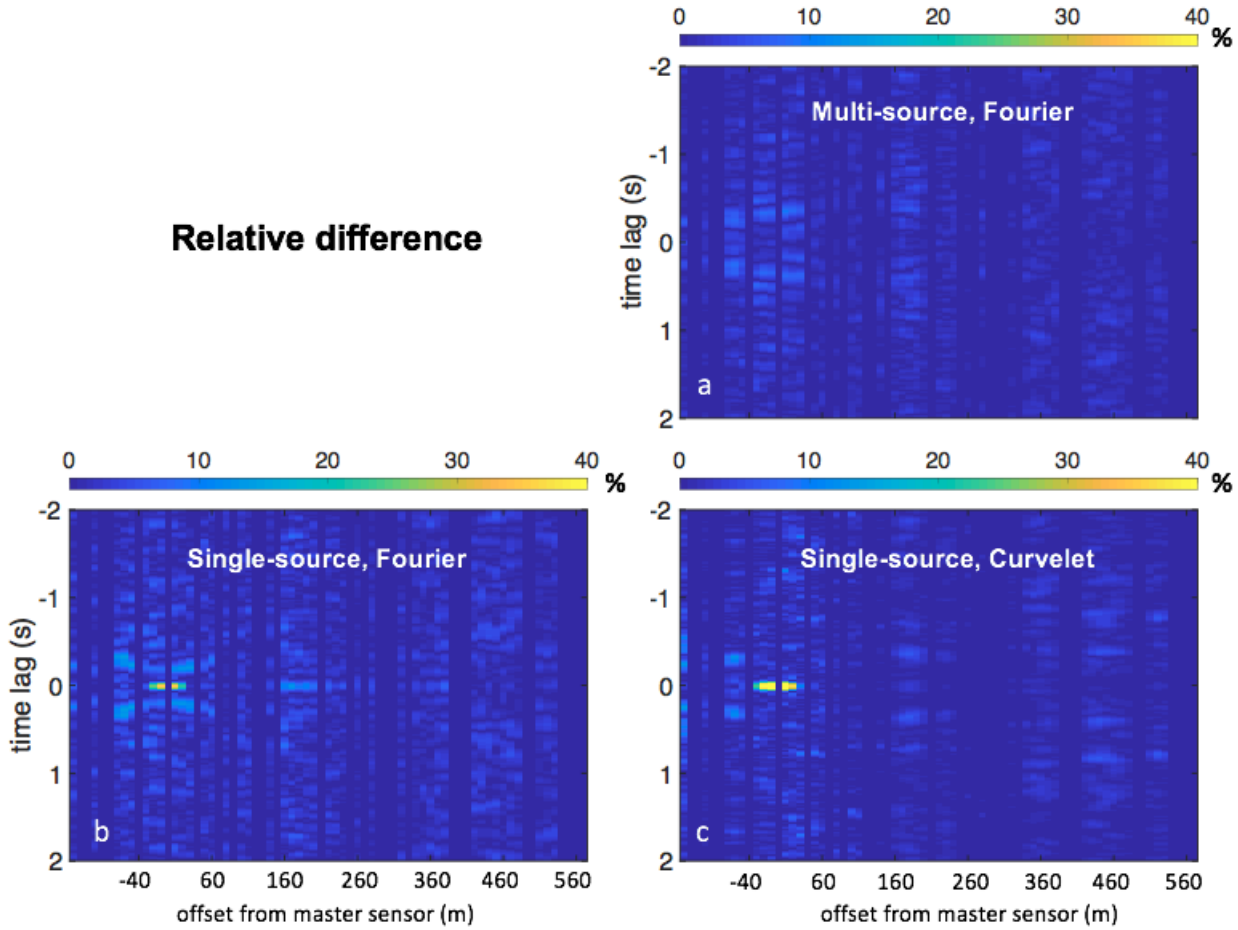


Figure 4.5 Relative difference (%) between the original correlogram (Figure 4.3a) and correlograms recovered by (a) the Fourier multi-source reconstruction, (b) the Fourier single-source reconstruction, and (c) the Curvelet single-source reconstruction. The offset distance is from the master sensor along the linear DAS fiber.

We use the signal-to-noise ratio (SNR) to quantitively determine the reconstruction error for each wavefield recovery; the SNR is defined by

$$SNR = 20\log(\frac{||f_{ori}||_2}{||f_{ori} - f_{rec}||_2}), \tag{4.5}$$

where f_{ori} and f_{rec} are the original and reconstructed correlograms, respectively. Using different realizations of missing DAS sensor locations, we determine the mean and standard deviation of the SNR for different wavefield reconstructions. Because single-source methods cannot recover correlograms for the missing virtual sources, we estimate the mean and the standard deviation only from all available virtual sources. For the multi-source method, the mean and the standard deviation are computed using 1) all available virtual sources and 2) all virtual sources (including the previously missing virtual sources). Figure 4.7 shows the comparison of the SNR of different wavefield reconstructions using four different fractions of remaining DAS sensors. The multi-source method overall achieves higher SNR than the single-source reconstruction by approximately 5-10 dB. The multi-source recovery improves the reconstruction of correlated wavefields for different fractions of remaining DAS sensors.

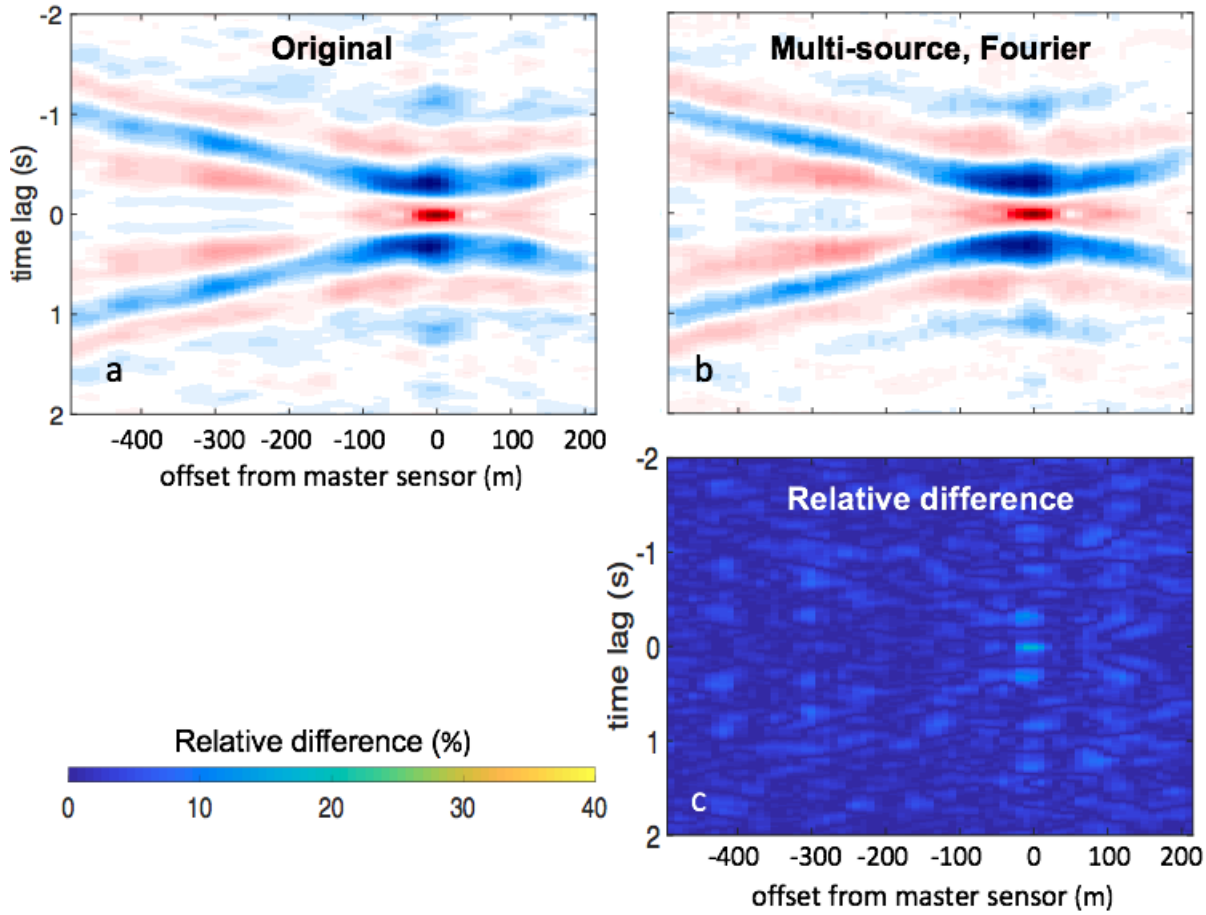


Figure 4.6 Comparison of (a) original correlogram and (b) the correlogram recovered by the Fourier multi-source reconstruction. (c) The relative difference (%) between the original and reconstructed wavefields in (a) and (b). The offset distance is from the 50^{th} master sensor along the linear DAS fiber.

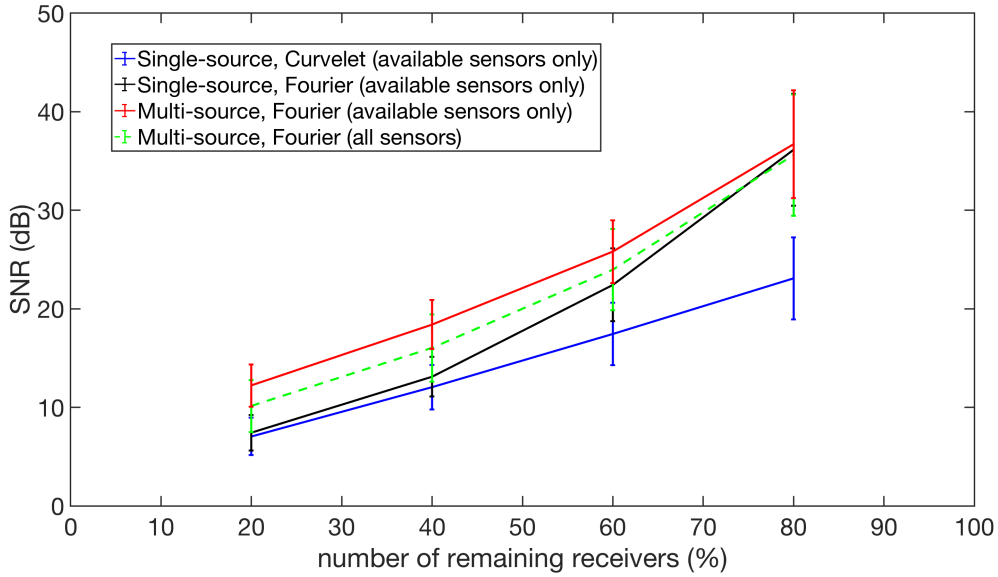


Figure 4.7 Mean and standard deviation of the SNR for different wavefield reconstructions using different fractions of remaining DAS sensors and different missing-sensor locations. For the single-source method, the mean and standard deviation are determined from only correlograms of all available virtual sources. The estimation of the mean and standard deviation for the multi-source method is presented both from correlograms of all available virtual sources (red) and from all virtual sources (green dash).

4.6 Discussion

Compressive sensing can be used to fill in gaps in interferometric waveforms obtained from DAS data. The improvement in wavefield reconstruction when some sensors are unusable is beneficial to DAS applications. Correlating DAS records has been used for Green's function retrieval (Dou et al., 2017; Lindsey et al., 2017; Martin et al., 2021). However, storing long-term DAS records for subsequent subsurface investigations and monitoring is expensive. In addition, during the deployment, DAS sensors can be inoperative and DAS fibers are highly sensitive to the orientation of incident wavefields, which can result in missing wavemode data from DAS fiber segments (Bakku et al., 2014; Martin et al., 2021; Shragge et al., 2021). Previous studies mitigate the expensive cost of data storage by storing seismic noise in a low-rank factorized form prior to interferometry (Martin, 2019) and mitigate the issue of missing data by applying wavefield reconstruction to fill the missing data (Saengduean et al., 2022; Zhan et al., 2018). With the improvement of wavefield recovery from DAS wavefield correlations, one may apply the multi-source method to reconstruct ambient correlograms of DAS data and the multi-source wavefield reconstruction can help mitigate the issue of missing data in DAS applications.

Similar improvement of wavefield reconstruction by the multi-source method over the single-source counterpart has been noted by Saengduean et al. (2022); the signal sparsity is enhanced when one groups images

of noise correlograms from multiple virtual sources into a signal volume for wavefield reconstruction, compared to reconstructing an individual correlogram image from a virtual source. The enhancement of signal sparsity may result from signal redundancy due to the translational invariance of correlated wavefields; for example, if the medium exhibits only weak lateral variations, the correlation of the wavefields from the 15^{th} and 35^{th} master sensors is the same as that of the wavefields of the 50^{th} and 70^{th} master sensors. Using a laterally homogeneous medium as a numerical example for the reconstruction of correlograms, Saengduean et al. (2022) show that the multi-source method improves the recovery of correlograms by approximately 40-60 dB, compared to the single-source method. The improvement of wavefield recovery shown by Saengduean et al. (2022) relies on a laterally invariant medium. We speculate that because the field DAS data sets is likely influenced by lateral variance of interferometric wavefields propagating along our linear DAS fiber, the improvement of our recovery of correlated DAS wavefields by approximately 5-10 dB is less than the improvement of wavefield recovery shown by Saengduean et al. (2022). Nevertheless, the present paper supports the use of the CS-based multi-source reconstruction on field data.

Because the multi-source method reconstructs correlated wavefields by relying on sparse signal representations, the multi-source wavefield reconstruction provides an alternative for mitigating the problem of expensive DAS data storage. One may require to store fewer significant sparse signal coefficients before recovering the correlated DAS wavefields for subsequent signal processing steps. In addition, although we show the application of multi-source wavefield reconstruct on DAS data and Saengduean et al. (2022) apply the multi-source method to seismic data recorded on a linear array, one may apply the multi-source technique to reconstruct wavefields recorded from an areal array using recorded VSG volumes.

4.7 Conclusion

We apply the multi-source wavefield reconstruction technique developed for recovering interferometric wavefields in the absence of some linear array receivers to DAS records from Perth, Australia. We exploit the sparsity of correlated DAS wavefields in the Curvelet and Fourier domains, and show that one can use the multi-source method for recovering noise correlograms of DAS data when some sensing locations are unusable. The Fourier multi-source method, which uses correlograms from multiple virtual sources, improves the reconstructed correlograms of DAS records by approximately 5-10 dB, compared to the Fourier and Curvelet single-source methods that separately use a correlogram from each virtual source for wavefield reconstruction. The multi-source method also allows recovery of correlograms at locations where master receivers are unavilable. The improvement of wavefield recovery from the multi-source method provides an alternative way to reconstruct wavefields for DAS data applications.

4.8 Acknowledgements

We thank the Development and Promotion of Science and Technology Talents Project (Royal Government of Thailand scholarship) for financially supporting P Saengduean to conduct his PhD research.

We thank colleagues from Center for Wave Phenomena who provide useful discussions about wavefield reconstruction and processing DAS data. We thank Terra15 Pty Ltd. for the DAS data from Perth, Australia which is used in this research. The authors declare that we have no competing interests.

CHAPTER 5

WEIGHTED COMPRESSIVE SENSING APPLIED TO SEISMIC INTERFEROMETRY: WAVEFIELD RECONSTRUCTION USING PRIOR INFORMATION

Modified from a paper to be submitted to Seismological Research Letters.

Patipan Saengduean[13], Roel Snieder[14], and Michael B. Wakin[15]

5.1 Abstract

Seismic interferometry is widely used for passive subsurface investigation using seismic noise. The technique requires much storage for long noise records to suppress interferometric noise which consists of spurious arrivals that do not correspond to the inter-receiver surface waves. Such long recordings may not be available in practice. Compressive sensing, which is a wavefield reconstruction technique operating on incomplete data, may reduce the unavailability and storage limitations of long noise time series. Using a numerical example of a linear array surrounded by sources and the Fourier basis for a sparse transform, we show that inter-receiver wavefields can be recovered at the locations where seismometers are unavailable, reducing the storage required for interferometry. We propose and develop a weighted compressive sensing algorithm that helps suppress the spurious arrivals by incorporating a-priori information about the arrivals of surface waves that can be expected.

5.2 Introduction

Signal acquisition is a significant step prior to the processes of analyzing acquired signals. The quality of acquired signals determines the complexity of the subsequent signal processing steps. Usually, acquired measurements or signals are contaminated by noise. In seismology, receiver arrays collecting seismic signals are often irregular and sparse because of missing or malfunctioning receivers. Some available receivers may record noisy signals or experience environmental changes (e.g. local temperature variations) that may fluctuate seismic signals recorded by the receivers (Wu et al., 2020; Yang et al., 2021). These noisy signals are usually unreliable and may be discarded from subsequent signal processing steps.

Seismic interferometry (SI) is a seismic processing technique using correlation of ambient seismic noise or diffusive wave coda to estimate wavefields that propagate between seismometers (Curtis et al., 2006; Larose et al., 2006; Snieder and Larose, 2013b; Wapenaar et al., 2010). SI has been applied for subsurface imaging and monitoring, using seismic noise (Draganov et al., 2007; Nakata et al., 2011; Shapiro et al.,

[13]Graduate student, primary researcher, and author at Department of Geophysics, Colorado School of Mines
[14]Supervisor at Department of Geophysics, Colorado School of Mines
[15]Supervisor at Department of Electrical Engineering, Colorado School of Mines

2005; Vasconcelos and Snieder, 2008a), active event sources (Matzel et al., 2016, 2017), and signals collected by distributed acoustic sensing (DAS) (Baird et al., 2020; Dou et al., 2017; Lellouch et al., 2019; Shragge et al., 2021). Some correlation traces used for SI applications may be contaminated with interferometric noise, which consists of wave arrivals that do not correspond to the surface wavefields retrieved by SI (Li et al., 2020; Snieder et al., 2008, 2006). These noisy traces may be unreliable and discarded from subsequent processing steps. One can alleviate interferometric noise by using the adequate source-distribution requirement for SI provided by Fan and Snieder (2009), or by stacking correlated windows of long noise records for noise interferometry. However, recording long seismic noise signals for correlation SI, as long as 3 months (Nakata et al., 2015) or 1 year (Lin et al., 2008) depending on the sampling frequency, requires much data storage. DAS increases the storage requirements for ambient records even more (Ben-Zion et al., 2015; Martin, 2019; Martinez et al., 2021; Sweet et al., 2018).

Recovery of correlation wavefields from fewer samples is an alternative to reduce the storage required for correlating long noise records. For example, Jayne et al. (2022) reduce the storage by reducing the frequency components that are used for cross-correlation in the Fourier domain. Exploiting the sparsity of seismic signals, Zhan et al. (2018) and Saengduean et al. (2022) reconstruct regular surface wavefields from irregular receiver arrays. One can improve seismic signal reconstruction using constraints from well-log or geological data (Gao et al., 2022; Wang et al., 2020). Additionally, when one expects varying reliability of samples in the signals, one can reconstruct signals by applying higher weights to signal samples with higher reliability (Anderson et al., 1997; Shin, 2013). For SI applications, although correlation traces with large interferometric noise can be discarded from subsequent signal processing, these traces may contain parts of surface wavefields which one aims to retrieve using SI. Thus, one may use these unreliable traces for signal reconstruction.

Rather than discarding unreliable signals before wavefield recovery, we propose and develop a weighted wavefield reconstruction technique to recover correlation wavefields and suppress interferometric noise. We formulate wavefield recovery as an optimization by minimizing an objective function which places higher weights on the signal samples with higher reliability. In Section 5.3, we describe the theory of traditional compressive sensing (CS), which is a wavefield reconstruction method exploiting the sparsity of signals, and explain the weighted CS algorithm that is modified from the traditional CS. In Section 5.4, we further explain the numerical model and synthetic wavefields used in our CS wavefield reconstruction. In Section 5.5, we show the comparison of the interferometric wavefields reconstructed using both traditional and weighted CS. Finally, we show that weighted wavefield recovery can reconstruct interferometric wavefields and reduce the spurious arrivals introduced by using inadequate source distribution for SI.

5.3 Sparse Wavefield Recovery using Compressive Sensing

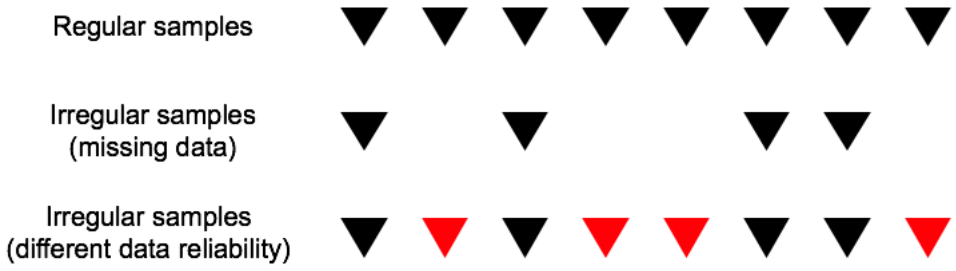


Figure 5.1 Regular samples of signals with an uniform interval (top). Irregular random samples with a non-uniform interval due to missing samples (missing triangles). Combination of two sets of irregular samples which have different degrees of signal reliability (black and red triangles).

The Nyquist sampling theorem states that one needs to sample signals at greater than twice the highest signal frequency. By explioting sparsity of signals, compressive sensing (CS) is a wavefield reconstruction technique that allows one to sample signals from fewer samples than the Nyquist theorem requires.

A signal f can be represented in a sparse domain Ψ: $f = \Psi\alpha$, where α is a vector of coefficients. A N-dimensional signal f is sparse when the number of non-zero coefficients K in α is less than N, and f is compressible when the sorted coefficients decay rapidly to zero (Candès and Wakin, 2008; Wakin, 2017). One can reconstruct a sparse signal when the sampling function Φ, which describes how signals are sampled, consists of random Gaussian or sub-Gaussian matrices; these matrices satisfy the restricted isometry property (RIP) that ensures energy conservation and stability of sparse signal reconstruction (Candès and Wakin, 2008; Donoho, 2006; Wakin, 2017). To reconstruct a K-sparse signal, one requires at least M measurements, where $M = O(Klog(N/K))$. Different algorithms can be used to seek sparse-signal reconstruction (Rani et al., 2018), but in our study, we minimize the following l_1 optimization to reconstruct sparse signals from noisy data d,

$$min||\alpha||_1 \text{ subject to } ||\Phi\Psi\alpha - d||_2 \leq \sigma, \tag{5.1}$$

where σ is the constant bound of the residual such that $||n||_2 \leq \sigma$, and $d = \Phi f + n$; n is signal residual which is the difference between the reconstructed and measured signals. After one recovers the coefficient vector α, the signal f can be reconstructed using $f = \Psi\alpha$.

When one regularly samples signals in space, some samples may be missing, which may render the signal acquisition irregular (Figure 5.1). Traditional CS (Equation 5.1) helps recover these samples from remaining samples. When one knows which signal samples are likely to be nonzero, one can perform weighted l_1 optimization to enhance signal sparsity and signal recovery (Khajehnejad et al., 2009). In addition, prior

information of signals (e.g. previously acquired signals or geology of investigated area) helps constrain traditional CS reconstruction of signals (Chen et al., 2008; Wang et al., 2020).

Rather than missing samples, one may acquire samples with different degrees of signal reliability (black and red triangles in Figure 5.1). In the spirit of weighted least-squares and prior-constrained CS, we propose to apply a weight matrix W to the l_1 optimization (Equation 5.1) for signal recovery from samples with different degrees of signal reliability. We generalize Equation 5.1 to

$$min||\alpha||_1 \text{ subject to } ||W(\Phi\Psi\alpha - d)||_2 \leq \sigma. \tag{5.2}$$

This weight matrix W is a diagonal matrix consisting of weights with values 0, 1, or w. Missing or completely unreliable signal samples have zero weight while fully reliable samples have weight 1 and $0 < w < 1$, depending on the reliability of the signal samples. The weighted CS (Equation 5.2) is equivalent to traditional CS (Equation 5.1) when the weight matrix consists of weights only with values 0 and 1.

The traditional and weighted CS algorithms are sparse-promoting recovery of wavefields. We use the Fourier transform for both unweighted and weighted CS wavefield reconstructions as a sparse transform for seismic signals (Abma and Kabir, 2006; Gao et al., 2013). The discrete 2D Fourier transform, associated with space-time dimension of seismic data, is equivalent to the 1D Fourier transform along each dimension of the signals (Nussbaumer, 1982). We avoid the wrap-around of the Fourier transform by padding zeroes at each end of our data matrix for every dimension.

5.4 Numerical Model and Synthetic Wavefields

Using SI for surface wavefield retrieval requires appropriate source distribution surrounding the receivers to mitigate interferometric noise. One can achieve the perfect source distribution (PSD) when sources are perfectly and appropriately distributed according to the requirement in Fan and Snieder (2009). Otherwise, sources are inappropriately and imperfectly distributed for SI. We call this case an imperfect source distribution (ISD) which introduces interferometric noise to the retrieved wavefields. In our work, we simulate surface wavefields using both PSD and ISD cases for wavefields with and without interferometric noise.

Figure 5.2 shows the source-receiver geometry used for our wavefield simulations. We use a linear seismic array consisting of 100 receivers that are regularly spaced at a separation of 10 m. For the PSD case, 8000 sources (shown in blue) are randomly distributed on the rectangle in Figure 5.2, where the sources are on average 5 m apart from each other. For ISD case, 400 red sources are randomly distributed around the array where the average source separation is approximately 100 m (Figure 5.2). To simulate surface waves, we use the 2D Green's function of the Helmholz equation (equation 18.46 from Snieder and van Wijk (2015a)),

$$u(x,t) = \int W(\omega) H_0^{(1)}(k(\omega)x) e^{-i\omega t} d\omega, \tag{5.3}$$

where x is the source-receiver distance, u is the wavefield at time t, $W(\omega)$ is the source-time function at angular frequency ω, $H_0^{(1)}$ is the first Hankel function of degree zero for wavenumber k; $k(\omega) = \omega/c(\omega)$, where c is the phase velocity.

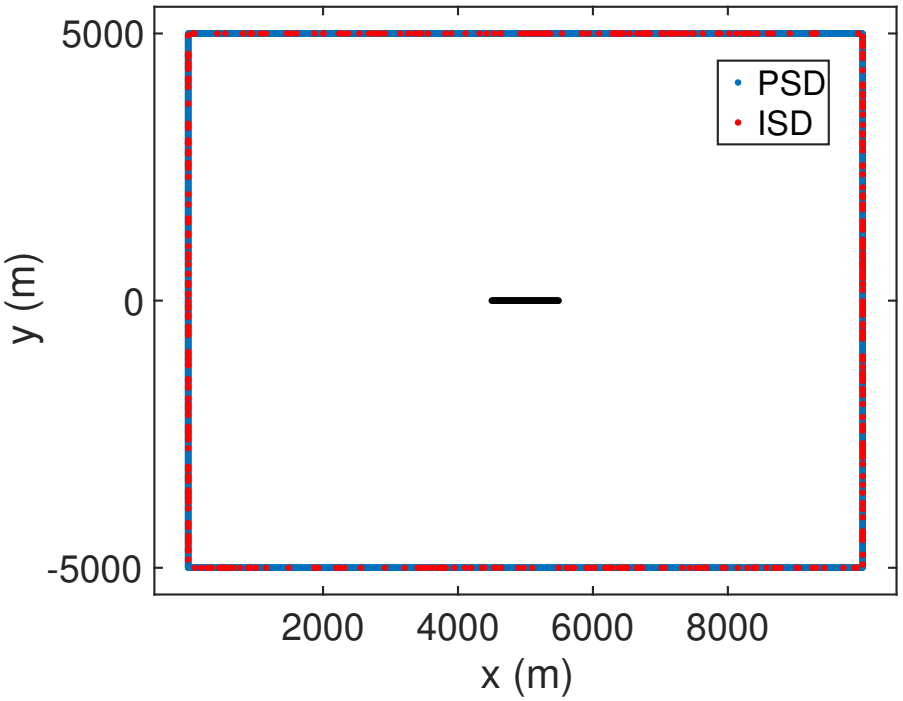


Figure 5.2 Source-receiver geometry. Black linear seismic array is located in the center of the geometry. Surrounding the receiver arrays, blue sources are perfectly distributed (PSD) while red sources are randomly and imperfectly distributed (ISD).

For $W(\omega)$, we use the spectrum of delta functions bandlimited between 5-20 Hz, and use the laterally layered homogeneous medium with phase velocity $c(\omega)$ determined by Xia et al. (1999) for the wavenumber in Equation 5.3. For SI, each receiver can act as a virtual source (master receiver). After we cross-correlate wavefields recorded at each receiver and at the master receiver, we stack the correlation wavefields for all sources for both PSD and ISD cases.

5.5 Comparison of Different CS Wavefield Reconstructions

Figure 5.3 shows the cross-correlated wavefields when we apply SI to seismic records on a linear array associated with the source-receiver geometry in Figure 5.2 when the 35^{th} receiver is used as master receiver. Figure 5.3a shows the cross-correlated wavefields using the PSD. Our aim is to reconstruct correlation wavefields that closely resemble the PSD wavefields (Figure 5.3a), when most receivers are missing or the seismic records of these receivers are unreliable.

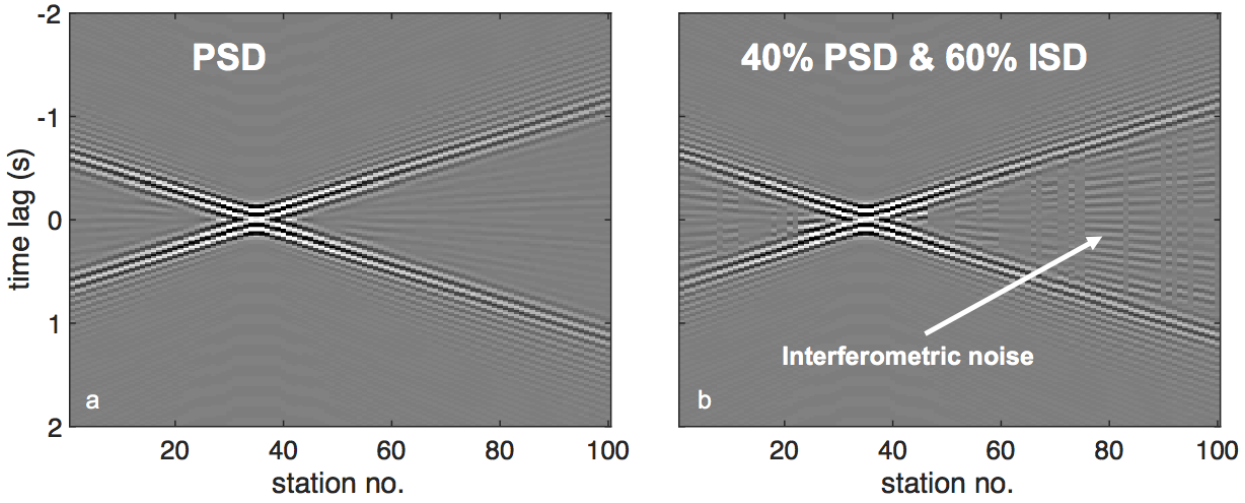


Figure 5.3 Cross-correlated wavefields recorded on a linear array surrounded by sources (Figure 5.2) when the master receiver is the 35^{th} and the source distribution is (a) fully PSD and (b) 40% PSD with 60% ISD.

When 60% receivers are absent from the array, one can apply the traditional CS to recover the missing correlation traces from the remaining traces; the missing and remaining traces (i.e. missing receivers) are shown using blue and yellow colors in Figure 5.4a, respectively. Figure 5.4d shows the correlation wavefields recovered by the traditional CS. For the case of seismic records of some receivers being unreliable, we use combined correlation wavefields of 40% PSD and 60% ISD traces (Figure 5.3b) as an example for wavefield reconstruction using our weighted CS; the PSD and ISD traces are shown by yellow and light blue colors in Figure 5.4b, respectively. Because we mix ISD traces, 60% of correlation wavefields exhibit interferometric noise which arrives in time between the acausal and causal surface waves (as indicated by the white arrow in Figure 5.3b). Using an arbitrary weight 0.4 for ISD traces and weight 1 for PSD traces, the correlation wavefields reconstructed by the weighted CS is shown in Figure 5.4e.

Since we know that interferometric noise arrives in time between the causal and acausal surface waves (Figure 5.3b), one can differently weigh the parts of interferometric noise and surface wave arrivals in the time dimension if one knows the expected arrivals of surface waves. Because parts of surface wave arrivals are more reliable than interferometric noise, one can give higher weight to the parts of surface waves than the parts of interferometric noise. We use the lowest and highest phase velocities of our numerical model to estimate the earliest and latest time arrivals of surface waves for each ISD traces, respectively. Once we estimate the time windows of the surface wave arrivals, we apply a weight 1 (yellow) to the PSD traces and we apply an arbitary weight 0.4 only to the expected time windows of surface wave arrivals (light blue cone) and a weight 0 (dark blue) to the other time windows for ISD traces with interferometric noise (Figure 5.4c). The reconstruction of correlation wavefields using weighted CS with known time arrivals of surface waves is

shown in Figure 5.4f.

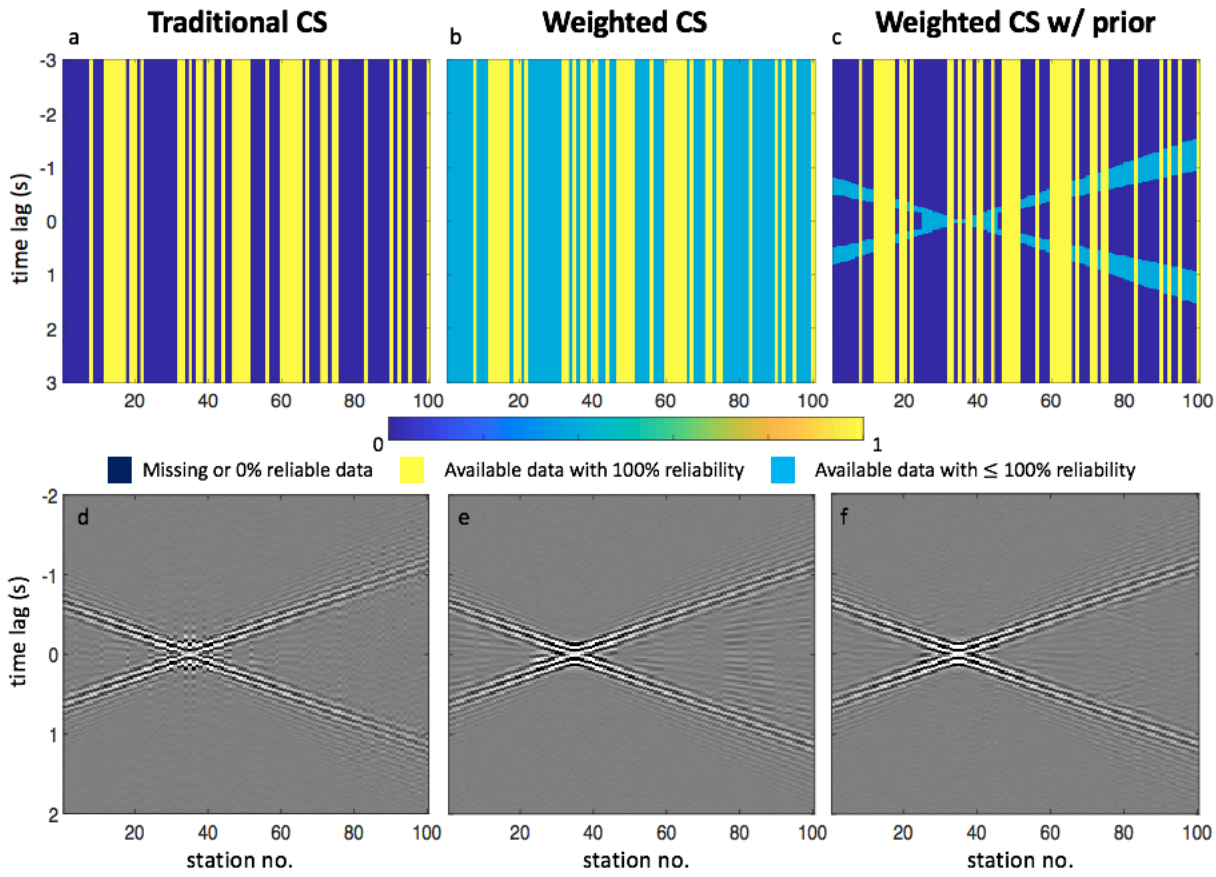


Figure 5.4 (a) traditional CS, (b) weighted CS, and (c) weighted CS with priori background velocity. Reconstruction of cross-correlated wavefields using (d) traditional CS, (e) weighted CS, and (f) weighted CS with priori background velocity.

Comparing reconstructed correlation wavefields to the PSD wavefields (Figure 5.3a), the traditional CS recovers a less accurate wavefield (Figure 5.4d) than the weighted CS wavefield reconstruction (Figure 5.4e and Figure 5.4f), in particular within 20 receivers (100 m) from the master receiver. We arbitrarily select the 88^{th} trace of the wavefields in Figure 5.4 to compare the reconstructed cross-correlations of different CS algorithms. Figure 5.5 shows the comparison of the PSD trace to the traces reconstructed using weighted and unweighted CS. Compared to the traditional CS (blue), the two weighted CS algorithms (red and green) reconstruct surface waves that are closer to the original PSD trace (black). Although the weighted CS can reconstruct the surface wave arrivals, without the prior known time arrivals of surface waves, the weighted CS also reconstructs the spurious arrivals (i.e. interferometric noise) of the cross-correlation between -0.6 and 0.6 s (red line in Figure 5.5). With the prior known time arrivals of surface waves, the weighted CS reconstructs surface waves that are comparable to the PSD signals but reduces the interferometric noise in

the wavefield reconstruction (green line in Figure 5.5).

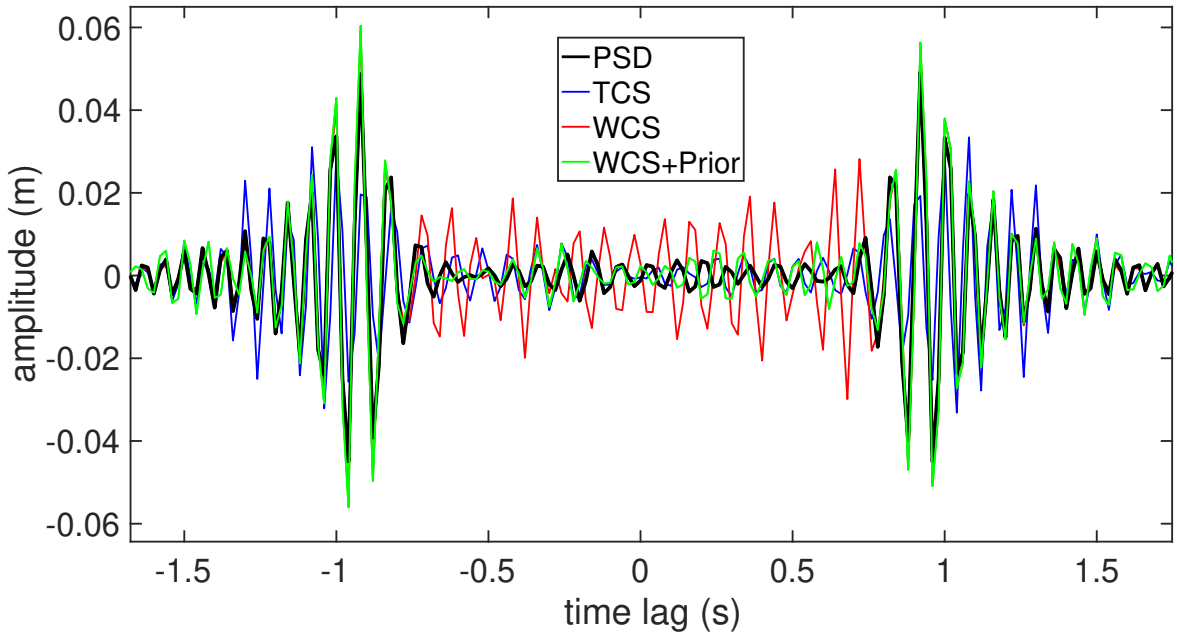


Figure 5.5 Comparison between the trace of the perfect-source distributed (PSD) wavefields and the traces recovered using traditional CS (TCS), weighted CS (WCS), and weighted CS with priori background-velocity knowledge. The traces are from the 88th receiver in Figure 5.4.

5.6 Discussion and Conclusions

We show that one can reconstruct the correlograms without interferometric noise from the combination of cross-correlations with and without interferometric noise by applying weights to the expected time windows of surface wave arrivals during the CS reconstruction. Compared to the traditional CS which uses only available traces without interferometric noise for wavefield recovery, the weighted CS gives accurate reconstruction of interferometric wavefields and reduces interferometric noise.

In our numerical example, we arbitrarily choose the weight ($w = 0.4$) for wavefield reconstruction but the weight can be any value from 0 to 1, depending on the reliability of the data. We show in the appendix that one can estimate the optimal weight w_{opt} that maximizes the signal-to-noise ratio (SNR) of wavefields; the SNR is defined by $= 20log(||x_{ori}||^2/||x_{ori} - x_{rec}||^2)$ where x_{ori} is the original and x_{rec} is the reconstructed wavefields. The optimal weight is given by

$$w_{opt} \leq \frac{\sigma}{\sqrt{\Sigma_{i \in cone} n_i^2}}, \tag{5.4}$$

where n is the signal error, which is the difference between the PSD and ISD signals, and the cone area is the light blue expected time windows of surface wave arrivals in Figure 5.4c. Our optimal weight expression

is similar to weighted least-squares inversion where the optimal weight for the weighted least-squares also depends on the signal error (Shin, 2013; Spilker and Vicini, 2001).

Although we use active-source interferometry as an example for weighted wavefield recovery, one can apply the weighted CS to noise interferometry that requires long seismic noise records to suppress interferometric noise. Previous studies mitigate storage problems by using records from sparse arrays for wavefield reconstruction (Zhan et al., 2018), using random subsampling in the frequency domain (Jayne et al., 2022), and storing seismic noise in a low-rank factorized form before performing interferometry in a compressed form (Martin, 2019). Rather than using subsampling or performing seismic interferometry in a compressed form, one may apply the weighted CS to reconstruct the correlograms from noise recorded by seismic receiver arrays, where majority of receivers records short duration of noise for interferometric contribution while the other few receivers records long duration of noise for noise interferometry. The weighted CS may help reduce storage requires for noise interferometry applications becuase one can store short noise records for most receivers and store long noise signals only from few receivers.

Although in our example, we apply weighted CS to interferometry, one can apply this weighted method to any digital measurements where parts of the signals are corrupted by noise or signal sensitivity. In these cases, one can down-weight the noisy parts of the data, depending on the reliability of the data.

5.7 Appendix

Using PSD and ISD cross-correlations, weighted CS reconstruction can recover PSD correlograms that are closer to the original PSD signals, compared to the traditional CS reconstruction. In the previous sections, we use an arbitrary weight for wavefield recovery. Here, we derive an optimal weight that provides the most accurate wavefield reconstruction by maximizing the SNR. Equation 5.2 shows that weights w and a constant bound of signal residual σ (i.e. the difference between the reconstructed and measured signals) are required for weighted CS reconstruction. Using the l_2 norm definition, we replace Equation 5.2 by

$$\Sigma_{i=1}^{N}[W(\Phi x - d)]_i^2 \leq \sigma^2, \tag{5.5}$$

where $x = \Psi\alpha$ is the reconstructed signals in the sparse domain and N is the signal length. Because the values w_i in W can only be 0, 1, and w, the summation reduces to

$$\begin{aligned} 0^2\Sigma_i(\Phi x - d)^2_{i,w_i=0} + 1^2\Sigma_i(\Phi x - d)^2_{i,w_i=1} + w^2\Sigma_i(\Phi x - d)^2_{i,w_i=w} &\leq \sigma^2 \\ \Sigma_i n^2_{i,w_i=1} + w^2\Sigma_i n^2_{i,w_i=w} &\leq \sigma^2, \end{aligned} \tag{5.6}$$

where $n = d - \Phi x$ and is the signal residual which is the difference between the ISD and PSD signals. Rearranging Equations 5.5 and 5.6 give the optimal weight,

$$w \leq \sqrt{\frac{\sigma^2 - \Sigma_i n_{i,w_i=1}^2}{\Sigma_i n_{i,w_i=w}^2}}. \tag{5.7}$$

When the signal error of fully-reliable data (at $w_i = 1$) is neglegible, the optimal weight expression reduces to

$$w \leq \frac{\sigma}{\sqrt{\Sigma_i n_{i,w_i=w}^2}}. \tag{5.8}$$

This optimal weight is scaled by a constant bound of noise and the signal error at parts of signals that are not fully reliable.

We use signal-to-noise ratio (SNR) to compare the accuracy of wavefield reconstruction. Using the numerical example in Sections 5.4 and 5.5, Figure 5.6 shows the SNR of weighted CS reconstruction at different weights and σ. We use Equation 5.8 to estimate the optimal weights for different σ (red line). We compare these expected optimal weights to the asterisk positions that gives the highest SNR for different σ from our numerical example. Figure 5.6 shows that the optimal weights estimated by Equation 5.8 are consistent with the highest-SNR wavefield reconstruction for different σ. Thus, one can use Equation 5.8 to estimate σ and weight that give the most accurate reconstruction of correlation wavefields.

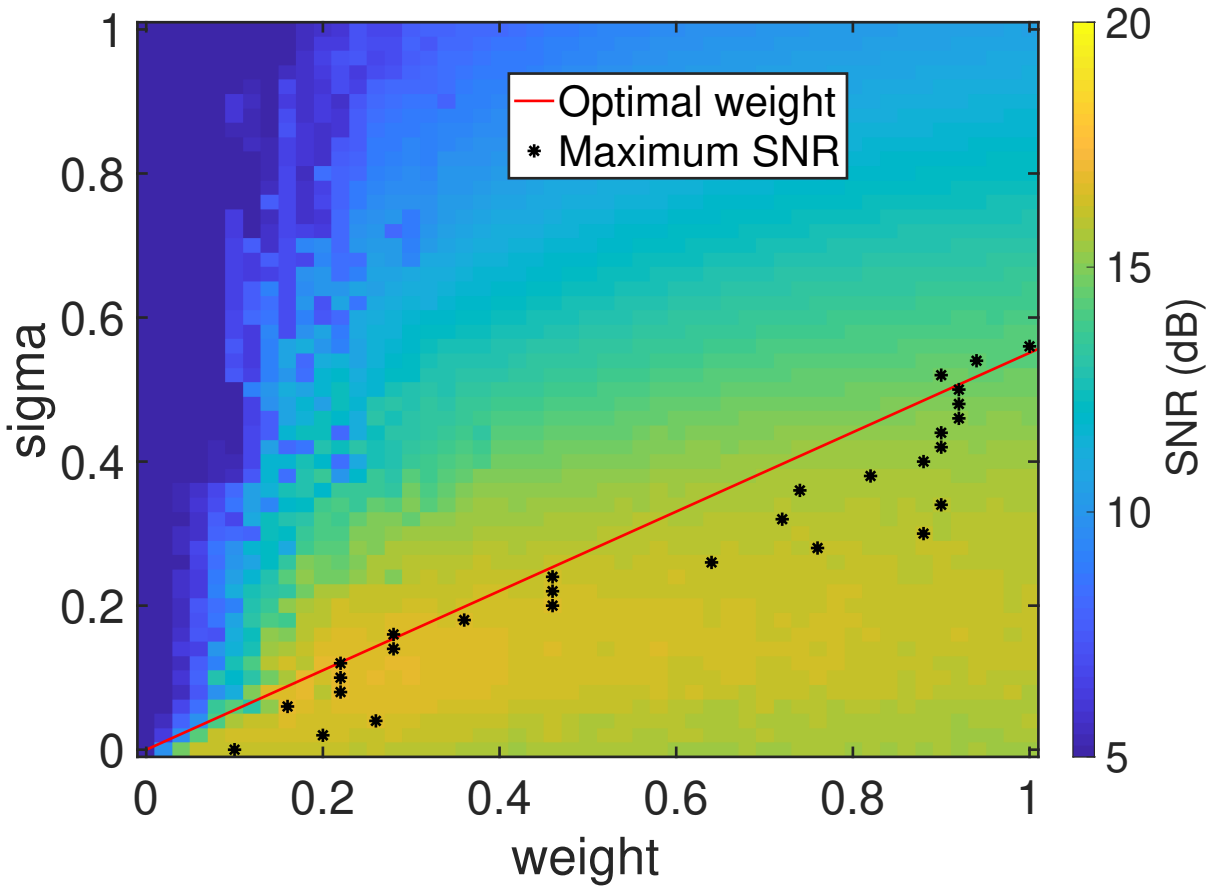


Figure 5.6 SNR (dB) of the weighted CS wavefield reconstruction, when the expected arrivals of surface waves are known, as a function of weights w and constant bounds of signal residual σ. The red line represents the optimal weight estimated by σ using Equation 5.8, while the black asterisks show the positions of highest SNR for each σ using our numerical example (Sections 5.4 and 5.5).

CHAPTER 6

CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

In this chapter, I provide a general summary of my thesis and recommend some future research ideas which can be built on the foundations that I have explored for the projects included in this thesis. The significant contribution of my thesis is that one can recover wavefields from the locations where either seismic receivers are missing or are impractical for receiver deployment, using inter-source interferometry and compressive sensing.

In Chapter 2, I show that one can extract body wavefields which account for the wave propagation between two seismic sources, using inter-source interferometry and wavefields recorded by surface receivers. The inter-source method requires four receivers per wavelength and the receiver array must cover the stationary phase region of the two sources to retrieve inter-source body wavefields. In practice, acquiring seismic data with adequate sampling is difficult. To achieve the sampling requirements, one may need dense seismic arrays that are usually deployed for industrial seismic exploration. Using wavefields recorded by surface receivers near the San Andreas Fault, I show that one can estimate the wavefields which account for the wave propagation between two earthquakes close to the fault. The inter-earthquake wavefields can be used to determine some properties of the fault. For example, I use the retrieved inter-source wavefields to determine the width of the damage zone of the San Andreas Fault which is approximately 4 km. The estimation of body wavefields using inter-source interferometry provides an alternative to receiver installation in the subsurface, which is expensive and may even be impossible in the mantle or inside volcanoes.

I show in Chapter 3 that one can use compressive sensing to reconstruct wavefields missing due to absent or inoperable receivers from a linear seismic receiver array. I propose a multi-source wavefield reconstruction, which is an application of compressive sensing, to reconstruct correlated wavefields using correlograms from multiple available virtual sources; the multi-source method is different from single-source wavefield reconstruction, which uses a correlogram for wavefield recovery from a single virtual source. The multi-source method can improve the reconstruction of correlograms obtained from interferometry by approximately 40-60 dB, compared to the single-source method, when the medium is laterally homogeneous. Although a linear receiver array is used as an example in Chapter 3, I show in the appendix that one can apply the multi-source method to 2D wavefields recorded on a 10×10 areal seismic receiver array. The multi-source method also reconstructs more accurate correlated wavefields than the single-source method because the fraction of significant sparse coefficients for 2D arrays is smaller than those of 1D arrays.

In Chapter 4, I apply the multi-source method proposed in Chapter 3 to DAS data recorded in Perth, Australia. I show that one can apply the multi-source method to reconstruct missing correlated wavefields due to the inoperable sections of DAS cables. The multi-source method improves the reconstruction of correlated DAS records by approximately 5-10 dB, compared to wavefield recovery using the single-source method. Although the wavefield-recovery improvement for DAS example is less than that of 40-60 dB from the numerical example shown in Chapter 3, the improvement supports the use of the multi-source method for DAS applications and for complicated (translation variant) wavefields.

In Chapter 5, I propose a weighted compressive sensing technique, which is a modified version of traditional compressive sensing, to reconstruct signals with variable signal reliability. Parts of the signals contain noise or are sensitive to unfavorable orientation of the fiber and incoming wavefields, and thus can be less reliable than the other parts of the signals and may be discarded or down-weighted from subsequent signal processing. I use correlated wavefields obtained from interferometry as an example for wavefield reconstruction. In general, interferometric noise can be produced if one cannot satisfy the source-reciver distribution required by Fan and Snieder (2009). My reconstruction example consists of correlated wavefields with and without interferometric noise. Using weighted compressive sensing for wavefield recovery, I show that one can reduce interferometric noise and reconstruct correlated wavefields from correlated wavefields with and without interferometric noise.

6.1 Future Research

A number of future innovations can be further implemented from the wavefield reconstruction methods shown in this thesis. These new directions include applications of inter-source interferometry to reconstruct inter-event wavefields for deep-earth investigations (e.g. inside subduction zones, the mantle, and the core), DAS data compression using the multi-source method and compressive sensing, and noise reduction using the weighted compressive sensing method. In the following sections, I discuss some possible future research directions for inter-source interferometry, multi-source wavefield reconstruction, and weighted compressive sensing.

6.1.1 Deep-Earth Investigation using Inter-Source Interferometry

I show that one can apply inter-source interferometry to retrieve body wavefields that account for wave propagation between two seismic events. One often uses surface or shallow subsurface receivers for deep-earth investigation because the installation of receivers in deep crust, mantle, or volcanoes is impossible. Successful retrieval of inter-earthquake body waves provides opportunities for deep-earth investigations by converting a deep earthquake into a virtual seismometer. One can recover inter-earthquake body wavefields

from deep earthquakes that occur due to rupture between a subduction slab and the mantle for deep-earth investigations. Such deep earthquakes include 2018 Fiji and Tonga earthquakes at a depth of approximately 600 km, which occurred due to subduction slabs (Jia et al., 2020; Lutikov et al., 2021).

Because subduction-zone earthquakes occur on the rupture surface of slabs and the mantle, the inter-event wavefields, which propagate along subduction slabs, can be used to investigate the property of the subduction zone or parts of the mantle. For example, Shen and Zhan (2020) uses inter-source wavefields to localize the meta olivine wedge inside a subduction slab beneath Japan. One can recover body waves traveling between subduction earthquakes to infer the properties (e.g. the thickness of the layers) of the subduction zone.

One may apply inter-source interferometry to S-waves generated from two deep earthquakes and are reflected from the core-mantle boundary for retrieval of inter-earthquake wavefields. This inter-event wavefields are sensitive to the structure and may help estimate the properties and locations of seismic structures (e.g. hot plumes and D" layers) located in the mantle, using modeling and imaging of travel-times and S-waves (Cottaar and Romanowicz, 2012; He et al., 2015; Stockmann et al., 2019; Yuan and Romanowicz, 2017). In addition, P-waves can travel through the outer core layers. One may retrieve inter-earthquake wavefields from P-waves generated by two earthquakes and passing through the core to infer the core structure, materials, and properties; the inter-event P-waves should be sensitive to the core structure. For example, modeling and analyzing the differential travel times and waveforms of P-waves can help estimate the structure and elastic properties of the outer cores (Irving et al., 2018; Yu et al., 2005).

6.1.2 DAS Data Compression using the Multi-Source CS Reconstruction of Wavefields

In this thesis, I show that using correlograms of multiple virtual sources provides higher-SNR recovery of correlated wavefields, compared to using a correlagram of a single virtual source. The multi-source method exploits sparser signal coefficients than the single-source method for the reconstruction of correlated wavefields.

Storing DAS data for subsequent data processing is often expensive. Thus, the idea of storing large amount of signals in a compressive format has recently become popular to alleviate the expensive data storage issue. For example, Martin (2019) stores DAS data into a low-rank factorized format before performing interferometry for noise correlation, and Rodriguez (2021b) uses compressive sensing and deep learning to record DAS data in a compressed form and facilitate fast data transmission and fast estimations of moment tensors and locations of events. Because the multi-source method reconstructs correlograms from sparser signal coefficients, the method allows a possibility that one may store only few significant sparse coefficients of a dense signal acquisition (e.g. DAS data) for subsequent signal recovery and processing.

6.1.3 Reasons of Wavefield-Recovery Improvement from the Multi-Source Reconstruction

In this thesis, I show that the multi-source method improves wavefield recovery, compared to the single-source method, by relying on sparser wavefield coefficients. I also speculate that the higher sparsity of coefficients is likely due to the redundancy of the translational invariance of correlated wavefields, which propagate along a linear seismic array.

The application of the multi-source method to field DAS data shows lower wavefield-recovery improvement than the reconstruction of wavefields numerically simulated using a laterally homogeneous medium, because the DAS wavefields may propagate through a laterally varying medium along a linear array. Thus, one needs a larger number of significant Fourier coefficients to represent the seismic signals. In the future, one may numerically test the effect of medium heterogeneity that affects the translation of correlated wavefields that propagate along a linear arrays, in terms of wavefield sparsity and recovery.

6.1.4 Wavefield Reconstruction for a Non-Uniform Sampling Grid

I show in this thesis that the multi-source wavefield reconstruction can be applied to wavefields recorded from a uniform-spaced seismic receiver array when some receivers are randomly missing or inoperative. However, arrays of uniformly spaced receivers are usually not available due to restrictions of receiver installation. For example, the USarray deployed by IRIS throughout the US is randomly spaced and irregular.

The idea of reconstructing signals for non-uniform sampling grid, which have been shown in non-geophysical disciplines, includes a variational approach based on spline spaces (Condat, 2013), neural network frameworks (Jiang et al., 2019a; Zhang et al., 2021), and compressive sensing using localized random samples (Barranca et al., 2016). Additionally, random subsampling for signal reconstruction provides fewer artifacts in the recovered signals, compared to regular subsamples (Snieder and Wakin, 2022). Thus, one may apply these techniques to incorporate the multi-source method using several virtual sources for wavefield recovery of a randomly non-uniform spaced receiver array such as the USarray.

6.1.5 Incorporating Machine Learning to Multi-Source Wavefield Reconstruction for Efficient Computation and Improved Wavefield Recovery

The multi-dimensional compressive sensing applications (e.g. the multi-source wavefield reconstruction) usually involve expensive computation. The fundamental idea of compressive sensing is to minimize l_1 optimization problem. One may shorten the computation of compressive sensing by using fast l_1 minimization algorithms. Examples of these algorithm include iterative-thresholding l_1 optimization (Drori, 2007; Yang et al., 2010), constrained l_1 optimization with an indicator function (Chen et al., 2015a), and $l_1 - l_2$ optimization using a proximal operator (Lou and Yan, 2018).

Incorporating machine learning to compressive sensing applications is an interesting alternative to improve the performance of the computation and wavefield recovery of compressive sensing. Previous studies show that integrating machine learning with compressive sensing can facilitate efficient computation and improve signal recovery, compared to conventional compressive sensing (Bright et al., 2013; Kravets and Stern, 2022; Zhang et al., 2021). In addition, the combination of compressive sensing and machine learning is applied to recovery of seismic wavefields, improving the wavefield recovery and efficiency of compressive-sensing computation for the reconstruction of missing seismic traces (Lu et al., 2019), moment tensor estimation (Rodriguez, 2021a), and passive seismic monitoring (Rodriguez, 2021b). One may incorporate machine learning to the multi-source method to facilitate the computation efficiency and wavefield recovery of multi-source wavefield reconstruction.

6.1.6 Applications of the Multi-Source CS Wavefield Reconstruction for Passive and Active Source Exploration

Although, in this thesis, I show the multi-source reconstruction for correlated wavefields from a linear array and a linear parts of a DAS fiber, the multi-source method can also be applied to correlated 2D wavefields recorded from an areal seismic array as shown in the appendix. I only show a numerical example of using the multi-source method to reconstruct correlated 2D wavefields. However, one may apply the multi-source method to field 2D seismic data recorded from areal arrays. Furthermore, the field data can be not only seismic noise but also active sources recorded on seismic arrays.

In this thesis, I use the multi-source method for the reconstruction of only correlated wavefields obtained by interferometry. However, one may apply the idea of using wavefield images from multiple sources for recovery of missing seismic profiles or shot gathers for exploration. Because the multi-source method relies on the translational invariance of wavefields propagating along a seismic array, the correlogram image from each virtual source contain similar structures of correlated waveforms, creating a linear translation structure for the data volume that consists of these correlograms. One may apply the multi-source method to seismic arrays deployed for nearly layered subsurface; an example of such surface is a geotechnical field in Sweden which consists of a bedrock overlain by sedimentary layers (Auken and Christiansen, 2004). In this case, the multi-source method can rely on the translational waveform structures from the nearly layered medium for wavefield reconstruction. The exploitation of the multi-source method for wavefield recovery is similar to that of 5D seismic interpolation which exploits the similar seismic structures from different inline, crossline, azimuth, offset, and frequency of seismic profiles (Al-Gain et al., 2020; Trad, 2009).

6.1.7 Non-Seismic Application of Weighted Compressive Sensing

Although using weighted compressive sensing, I show a numerical example of the successful reduction of interferometric noise from correlograms obtained by interferometry, one may apply the weighted compressive sensing method to any digital signals, where parts of the signals are corrupted or contaminated by some issues such as noise or instrument sensitivity.

For example, one may apply the weighted reconstruction method to separate singing voices and background music accompaniments. Singing voices and background music are usually harmonic and non-stationary. Because of the non-stationary behavior and variation of the frequency of human voice and background music, the voice separation problem is very interesting and challenging. Past studies use pitch-based or model-based method (Ozerov et al., 2007; Vembu and Baumann, 2005) and frequency factorization (Virtanen et al., 2008) for the separation of voice signals. Hsu et al. (2012), who exploit the temporal continuity and the hamonicity of different voice pitches from song recordings to detect the pitch range of singing voices for different time frames, outperform the previous studies for the extraction of singing voices. Because of the temporal continuity behavior of singing voices, one may apply the weighted compressive sensing for voice separation from music accompaniments, where one may give higher weight for the time frames of the sound signals during the pitches of singing voices than during the time frames of prioritized background music.

In addition, the weighted compressive sensing can be applied for noise reduction of digital noisy signals that have been computed using the weighted least squares method; examples of such application include fitting scattered noisy signals (Zhou and Han, 2008) and denoising data volumes (Zhang et al., 2019). Because one can set the weight depending on the size of the noise for the weighted least-squares method, one may use a similar idea to provide the weight for the weighted compressive sensing, depending on signal reliability.

REFERENCES

Abma, R. and Kabir, N. (2006). 3D Interpolation of irregular data with a POCS algorithm. *Geophysics*, 71(6):E91–E97.

Aki, K. and Richards, P. (2002). *Quantitative Seismology*. Univ. Science Books, second edition.

Al-Gain, M., Abdelrahman, K., Kahal, A., Al-Zahrani, S., Ibrahim, E., and Al-Otaibi, N. (2020). Impact of 5D regularization and interpolation on subsurface imaging: A case study of Stratton field, South Texas, United States of America. *Journal of King Saud University - Science*, 32(6):2733–2740.

Alwon, S. (2018). Generative adversarial networks in seismic data processing. *2018 SEG International Exposition and Annual Meeting*, pages 1991–1995.

Anderson, J. M., Mair, B. A., Rao, M., and Wu, C. H. (1997). Weighted least-squares reconstruction methods for positron emission tomography. *IEEE transactions on medical imaging*, 16(2):159–165.

Arianandа, D. D. and Leus, G. (2012). Compressive wideband power spectrum estimation. *IEEE Transactions on Signal Processing*, 60(9):4775–4789.

Asano, K., Iwata, T., Sekiguchi, H., Somei, K., Miyakoshi, K., Aoi, S., and Kunugi, T. (2017). Surface wave group velocity in the Osaka sedimentary basin, Japan, estimated using ambient noise cross-correlation functions. *Earth, Planets and Space*, 69(108):1–20.

Auken, E. and Christiansen, A. V. (2004). Layered and laterally constrained 2D inversion of resistivity data. *Geophysics*, 69(3):752–761.

Baird, A. F., Stork, A. L., Horne, S. A., Naldrett, G., Kendall, J. M., Wookey, J., Verdon, J. P., and Clarke, A. (2020). Characteristics of microseismic data recorded by distributed acoustic sensing systems in anisotropic media. *Geophysics*, 85(4):KS139–KS147.

Bakku, S. K., Fehler, M., and Wills, P. (2014). Monitoring hydraulic fracturing using distributed acoustic sensing in a treatment well. In *SEG Technical Program Expanded Abstracts 2014*, pages 5003–5008.

Bakulin, A. and Calvert, R. (2006). The virtual source method: Theory and case study. *Geophysics*, 71(4):SI139–SI150.

Barranca, V. J., Kovacic, G., Zhou, D., and Cai, D. (2016). Improved Compressive Sensing of Natural Scenes Using Localized Random Sampling. *Scientific Reports*, 6(31976):1–17.

Ben-Zion, Y., Vernon, F. L., Ozakin, Y., Zigone, D., Ross, Z. E., Meng, H., White, M., Reyes, J., Hollis, D., and Barklage, M. (2015). Basic data features and results from a spatially dense seismic array on the San Jacinto fault zone. *Geophysical Journal International*, 202(1):370–380.

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Shapiro, N. M., and Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, 169(3):1239–1260.

Boehm, C., Hanzich, M., de la Puente, J., and Fichtner, A. (2016). Wavefield compression for adjoint methods in full-waveform inversion. *Geophysics*, 81(6):R385–R397.

Bright, I., Lin, G., and Kutz, J. N. (2013). Compressive sensing based machine learning strategy for characterizing the flow around a cylinder with limited pressure measurements. *Physics of Fluids*, 25(12):7102.

Burdick, L. J. and Orcutt, J. A. (1979). A comparison of the generalized ray and reflectivity methods of waveform synthesis. *Geophysical Journal of the Royal Astronomical Society*, 58(2):261–278.

Candès, E., Demanet, L., Donoho, D., and Ying, L. (2006a). Fast discrete curvelet transforms. *Multiscale Modeling and Simulation*, 5(3):861–899.

Candès, E., Demanet, L., Donoho, D., and Ying, L. (2006b). Fast discrete curvelet transforms. *Multiscale Modeling and Simulation*, 5(3):861–899.

Candès, E. and Wakin, M. (2008). An Introduction To Compressive Sampling. *IEEE Signal Processing Magazine*, 25(2):21–30.

Candès, E. J. and Demanet, L. (2005). The curvelet representation of wave propagators is optimally sparse. *Communications on Pure and Applied Mathematics*, 58(11):1472–1528.

Candès, E. J., Romberg, J., and Tao, T. (2006c). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509.

Cao, A., Stump, B., and DeShon, H. (2018). High-resolution seismic data regularization and wavefield separation. *Geophysical Journal International*, 213(1):684–694.

Chang, H. and Nakata, N. (2022). Investigation of Time-Lapse Changes with DAS Borehole Data at the Brady Geothermal Field Using Deconvolution Interferometry. *Remote Sensing*, 14(1):1–18.

Chen, F., Shen, L., Suter, B. W., and Xu, Y. (2015a). A fast and accurate algorithm for Í 1 minimization problems in compressive sampling. *Eurasip Journal on Advances in Signal Processing*, 2015(1):1–12.

Chen, G.-H., Tang, J., and Leng, S. (2008). Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Medical Physics*, 35(2):660–663.

Chen, Y., Chi, Y., and Goldsmith, A. J. (2015b). Exact and Stable Covariance Estimation From Quadratic Sampling via Convex Programming. *IEEE Transactions on Information Theory*, 61(7):4034–4059.

Choi, J.-H., Edwards, P., Ko, K., and Kim, Y.-S. (2016). Definition and classification of fault damage zones: A review and a new methodological approach. *Earth-Science Reviews*, 152:70–87.

Chopra, S. and Marfurt, K. J. (2013). Preconditioning seismic data with 5D interpolation for computing geometric attributes. *The Leading Edge*, 32(12):1456–1460.

Condat, L. (2013). Reconstruction from non-uniform samples: A direct, variational approach in shift-invariant spaces. *Digital Signal Processing: A Review Journal*, 23(4):1277–1287.

Cottaar, S. and Romanowicz, B. (2012). An unsually large ULVZ at the base of the mantle near Hawaii. *Earth and Planetary Science Letters*, 355-356:213–222.

Curtis, A., Gerstoft, P., Sato, H., Snieder, R., and Wapenaar, K. (2006). Seismic interferometry – turning noise into signal. *The Leading Edge*, 25(9):1082–1092.

Curtis, A. and Halliday, D. (2010). Source-receiver wave field interferometry. *Physical Review E*, 81(4):046601.

Curtis, A., Nicolson, H., Halliday, D., Trampert, J., and Baptie, B. (2009). Virtual seismometers in the subsurface of the Earth from seismic interferometry. *Nature Geoscience*, 2(10):700–704.

Dabov, K., Foi, A., and Egiazarian, K. (2007). Image denoising by sparse 3D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095.

Delong, S. B., Hilley, G. E., Rymer, M. J., and Prentice, C. (2010). Fault zone structure from topography: Signatures of en echelon fault slip at Mustang Ridge on the San Andreas Fault, Monterey County, California. *Tectonics*, 29(5):1–16.

Diekmann, L. and Vasconcelos, I. (2021). Focusing and Green's function retrieval in three-dimensional inverse scattering revisited: A single-sided Marchenko integral for the full wave field. *Physical Review Research*, 3(1):13206.

Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306.

Dou, S., Lindsey, N., Wagner, A. M., Daley, T. M., Freifeld, B., Robertson, M., Peterson, J., Ulrich, C., Martin, E. R., and Ajo-Franklin, J. B. (2017). Distributed Acoustic Sensing for Seismic Monitoring of the Near Surface: A Traffic-Noise Interferometry Case Study. *Scientific Reports*, 7(1):1–12.

Draganov, D., Wapenaar, K., Mulder, W., Singer, J., and Verdel, A. (2007). Retrieval of reflections from seismic background-noise measurements. *Geophysical Research Letters*, 34(4):L04305.

Drori, I. (2007). Fast Í1 minimization by iterative thresholding for multidimensional NMR spectroscopy. *Eurasip Journal on Advances in Signal Processing*, 2007:1–10.

Entwistle, E., Curtis, A., Galetti, E., Baptie, B., and Meles, G. (2015). Constructing new seismograms from old earthquakes: Retrospective seismology at multiple length scales. *Journal of Geophysical Research: Solid Earth*, 120(4):2466–2490.

Eulenfeld, T. (2020). Towards source region tomography with inter-source interferometry: Shear wave velocity from 2018 West Bohemia swarm earthquakes. *Journal of Geophysical Research: Solid Earth [manuscript submitted for publication]*, pages 1–24.

Fan, Y. and Snieder, R. (2009). Required source distribution for interferometry of waves and diffusive fields. *Geophysical Journal International*, 179(2):1232–1244.

Forghani, F. and Snieder, R. (2010). Underestimation of body waves and feasibility of surface-wave reconstruction by seismic interferometry. *The Leading Edge*, 29(7):790–794.

Gan, S., Wang, S., Chen, Y., Zhang, Y., and Jin, Z. (2015). Dealiased Seismic Data Interpolation Using Seislet Transform with Low-Frequency Constraint. *IEEE Geoscience and Remote Sensing Letters*, 12(10):2150–2154.

Gao, J., Stanton, A., Naghizadeh, M., Sacchi, M. D., and Chen, X. (2013). Convergence improvement and noise attenuation considerations for beyond alias projection onto convex sets reconstruction. *Geophysical Prospecting*, 61(SUPPL.1):138–151.

Gao, Y., Zhang, J., Li, H., and Li, G. (2022). Incorporating Structural Constraint Into the Machine Learning High-Resolution Seismic Reconstruction. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–12.

Gold, R. D., Reitman, N. G., Briggs, R. W., Barnhart, W. D., Hayes, G. P., and Wilson, E. (2015). On- and off-fault deformation associated with the September 2013 Mw 7.7 Balochistan earthquake: Implications for geologic slip rate measurements. *Tectonophysics*, 660:65–78.

Gurbuz, A. C., McClellan, J. H., and Scott, W. R. (2009). Compressive sensing for subsurface imaging using ground penetrating radar. *Signal Processing*, 89(10):1959–1972.

Guy, M. R., Patton, J. M., Fee, J., Hearne, M., Martinez, E., Ketchum, D., Worden, C., Quitoriano, V., Hunter, E., Smoczyk, G., and Schwarz, S. (2015). National Earthquake Information Center systems overview and integration 2015-1120. *U.S. Geological Survey Open-File Report*, page 25.

Halliday, D. and Curtis, A. (2008). Seismic interferometry, surface waves and source distribution. *Geophysical Journal International*, 175(3):1067–1087.

Hanafy, S. M. and Schuster, G. T. (2014). Interferometric interpolation of sparse marine data. *Geophysical Prospecting*, 62(1):1–16.

He, Y., Wen, L., Capdeville, Y., and Zhao, L. (2015). Seismic evidence for an Iceland thermo-chemical plume in the Earth's lowermost mantle. *Earth and Planetary Science Letters*, 417:19–27.

Hennenfent, G., Fenelon, L., and Herrmann, F. J. (2010). Nonequispaced curvelet transform for seismic data reconstruction: A sparsity-promoting approach. *Geophysics*, 75(6):WB203–WB210.

Herrmann, F. J., Wang, D., Hennenfent, G., and Moghaddam, P. P. (2008). Curvelet-based seismic data processing: A multiscale and nonlinear approach. *Geophysics*, 73(1):A1–A5.

Holdsworth, R. E., van Diggelen, E. W., Spiers, C. J., de Bresser, J. H., Walker, R. J., and Bowen, L. (2011). Fault rocks from the SAFOD core samples: Implications for weakening at shallow depths along the San Andreas Fault, California. *Journal of Structural Geology*, 33(2):132–144.

Hsu, C. L., Wang, D., Jang, J. S. R., and Hu, K. (2012). A tandem algorithm for singing pitch extraction and voice separation from music accompaniment. *IEEE Transactions on Audio, Speech and Language Processing*, 20(5):1454–1463.

Innocent Oboué, Y. A. S., Chen, W., Wang, H., and Chen, Y. (2021). Robust damped rank-reduction method for simultaneous denoising and reconstruction of 5D seismic data. *Geophysics*, 86(1):V71–V89.

Irving, J. C. E., Cottaar, S., and Lekić, V. (2018). Seismically determined elastic parameters for Earth's outer core. *Science Advances*, 4(6):eaar2538.

Islam, S. R., Maity, S. P., and Ray, A. K. (2015). On compressed sensing image reconstruction using linear prediction in adaptive filtering. In *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 2317–2323.

Jayne, J., Wakin, M. B., and Snieder, R. (2022). Green's Function Estimation by Seismic Interferometry from Limited Frequency Samples. *Preprint (in preparation)*, pages 1–32.

Jeppson, T. N. and Tobin, H. J. (2015). San Andreas fault zone velocity structure at SAFOD at core, log, and seismic scales. *Journal of Geophysical Research: Solid Earth*, 120:4983–4997.

Jia, Z., Shen, Z., Zhan, Z., Li, C., Peng, Z., and Gurnis, M. (2020). The 2018 Fiji Mw 8.2 and 7.9 deep earthquakes: One doublet in two slabs. *Earth and Planetary Science Letters*, 531:115997.

Jiang, C. M., Wang, D., Huang, J., Marcus, P., and Niessner, M. (2019a). Convolutional Neural Networks on Non-uniform Geometrical Signals Using Euclidean Spectral Transformation. In *International Conference on Learning Representations*.

Jiang, T., Eick, P., Jiang, Y., Li, T., Hao, H., Chu, W., Holt, R., Blymyer, D., Koster, K., and Enns, D. (2019b). Compressive sensing seismic processing tests on a high density blended land data set. *SEG International Exposition and Annual Meeting 2019*, pages 4505–4509.

Kabir, M. M. and Verschuur, D. J. (1995). Restoration of missing offsets by parabolic Radon transform. *Geophysical Prospecting*, 43(3):347–368.

Karrenbach, M., Ridge, A., Cole, S., Boone, K., Kahn, D., Rich, J., Silver, K., and Langton, D. (2017). DAS Microseismic Monitoring and Integration With Strain Measurements in Hydraulic Fracture Profiling.

Khajehnejad, M. A., Xu, W., Avestimehr, A. S., and Hassibi, B. (2009). Weighted Í1 minimization for sparse recovery with prior information. In *2009 IEEE International Symposium on Information Theory*, pages 483–487.

Kim, B., Jeong, S., and Byun, J. (2015). Trace interpolation for irregularly sampled seismic data using curvelet-transform-based projection onto convex sets algorithm in the frequency-wavenumber domain. *Journal of Applied Geophysics*, 118:1–14.

Kiraz, M. S. R., Snieder, R., and Wapenaar, K. (2021). Focusing waves in an unknown medium without wavefield decomposition. *JASA Express Letters*, 1(5):55602.

Korneev, V., Nadeau, R., and Berkeley, L. (2004). Vibroseis Monitoring of San Andreas Fault in California. *University of California, Lawrence Berkeley National Laboratory*, (https://escholarship.org/uc/item/2t35c4x6).

Kravets, V. and Stern, A. (2022). Progressive compressive sensing of large images with multiscale deep learning reconstruction. *Scientific Reports*, 12(1):7228.

Larose, E., Margerin, L., Derode, A., van Tiggelen, B., Campillo, M., Shapiro, N., Paul, A., Stehly, L., and Tanter, M. (2006). Correlation of random wavefields: An interdisciplinary review. *Geophysics*, 71(4):SI11–SI21.

Lellouch, A., Yuan, S., Spica, Z., Biondi, B., and Ellsworth, W. L. (2019). Seismic Velocity Estimation Using Passive Downhole Distributed Acoustic Sensing Records: Examples From the San Andreas Fault Observatory at Depth. *Journal of Geophysical Research: Solid Earth*, 124(7):6931–6948.

Li, L., Boué, P., and Campillo, M. (2020). Observation and explanation of spurious seismic signals emerging in teleseismic noise correlations. *Solid Earth*, 11(1):173–184.

Li, S. and Yang, B. (2011). A new pan-sharpening method using a compressed sensing technique. *IEEE Transactions on Geoscience and Remote Sensing*, 49(2):738–746.

Li, X., Aravkin, A. Y., Van Leeuwen, T., and Herrmann, F. J. (2012). Fast randomized full-waveform inversion with compressive sensing. *Geophysics*, 77(3):A13–A17.

Li, Y. G., Leary, P., Aki, K., and Malin, P. (1990). Seismic trapped modes in the Oroville and San Andreas fault zones. *Science*, 249(4970):763–766.

Li, Y. G., Vidale, J. E., and Cochran, E. S. (2004). Low-velocity damaged structure of the San Andreas Fault at Parkfield from fault zone trapped waves. *Geophysical Research Letters*, 31(12):1–5.

Lin, F. C., Moschetti, M. P., and Ritzwoller, M. H. (2008). Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps. *Geophysical Journal International*, 173(1):281–298.

Lindsey, N., Dou, S., Martin, E. R., Wagner, A. M., and Ajo Franklin, J. B. (2017). 4-D permafrost thaw observations from ambient road traffic noise and a very dense distributed fiber optic sensing array. In *AGU Fall Meeting Abstracts*, volume 2017, pages S31A–0804.

Liu, Q., Fu, L., and Zhang, M. (2019). Deep-seismic-prior-based reconstruction of seismic data using convolutional neural networks. pages 1–5.

Liu, S., Liu, Y., Zhang, Y., and Tan, T. (2017). Compressive data gathering in wireless sensor networks via group sparse regularization. *IEEE International Conference on Communications*, pages 1–5.

Liu, W., Cao, S., Gan, S., Chen, Y., Zu, S., and Jin, Z. (2016a). One-Step Slope Estimation for Dealiased Seismic Data Reconstruction via Iterative Seislet Thresholding. *IEEE Geoscience and Remote Sensing Letters*, 13(10):1462–1466.

Liu, X., Zhu, H., Zhang, M., Richardson, A. G., Lucas, T. H., and Van Der Spiegel, J. (2015). Design of a low-noise, high power efficiency neural recording front-end with an integrated real-time compressed sensing unit. *Proceedings - IEEE International Symposium on Circuits and Systems*, 2015-July:2996–2999.

Liu, Y., Draganov, D., Wapenaar, K., and Arntsen, B. (2016b). Retrieving virtual reflection responses at drill-bit positions using seismic interferometry with drill-bit noise. *Geophysical Prospecting*, 64(2):348–360.

Lou, Y. and Yan, M. (2018). Fast L1–L2 Minimization via a Proximal Operator. *Journal of Scientific Computing*, 74(2):767–785.

Lu, P., Xiao, Y., Zhang, Y., and Mitsakos, N. (2019). Deep learning for 3D seismic compressive-sensing technique: A novel approach. *The Leading Edge*, 38(9):658–740.

Lutikov, A. I., Rogozhin, E. A., Dontsova, G. Y., and Zhukovets, V. N. (2021). The August 19, 2018 (MW = 8.2) Strong Deep Earthquake West off the Tonga Islands in the Context of Developing Our Concepts of Deep Earthquakes. *Izvestiya, Physics of the Solid Earth*, 57(2):203–216.

Ma, J. (2013). Three-dimensional irregular seismic data reconstruction via low-rank matrix completion. *Geophysics*, 78(5):V181–V192.

Maggioni, M., Katkovnik, V., Egiazarian, K., and Foi, A. (2013). Nonlocal transform-domain filter for volumetric data denoising and reconstruction. *IEEE Transactions on Image Processing*, 22(1):119–133.

Mann, M. E. and Emanuel, K. A. (2006). Atlantic hurricane trends linked to climate change. *Eos, Transactions American Geophysical Union*, 87(24):233–241.

Mansour, H., Wason, H., Lin, T. T., and Herrmann, F. J. (2012). Randomized marine acquisition with compressive sampling matrices. *Geophysical Prospecting*, 60(4):648–662.

Martin, E. R. (2019). A scalable algorithm for cross-correlations of compressed ambient seismic noise. In *SEG Technical Program Expanded Abstracts 2019*, pages 3006–3010.

Martin, E. R., Lindsey, N. J., Ajo-Franklin, J. B., and Biondi, B. L. (2021). Introduction to Interferometry of Fiber-Optic Strain Measurements. In *Distributed Acoustic Sensing in Geophysics*, chapter 9, pages 111–129. American Geophysical Union (AGU).

Martinez, A., Useche, M., Sayed, A., Duran, C., and Araujo, J. (2021). DAS: An efficient and effective solution for VSP acquisition. In *First International Meeting for Applied Geoscience & Energy Expanded Abstracts*, pages 3500–3504.

Mateeva, A., Lopez, J., Potters, H., Mestayer, J., Cox, B., Kiyashchenko, D., Wills, P., Grandi, S., Hornman, K., Kuvshinov, B., Berlang, W., Yang, Z., and Detomo, R. (2014). Distributed acoustic sensing for reservoir monitoring with vertical seismic profiling. *Geophysical Prospecting*, 62(4):679–692.

Matzel, E., Morency, C., Rhode, A., Templeton, D., Pyle, M., and Ave, E. (2016). Virtual Seismometers in Geothermal Systems: Looking Inside the Microseismic Cloud. *PROCEEDINGS, 41st Stanford Geothermal Workshop on Geothermal Reservoir Engineering*, pages 1–4.

Matzel, E., Zeng, X., Thurber, C., Luo, Y., and Morency, C. (2017). Seismic Interferometry Using the Dense Array at the Brady Geothermal Field. *PROCEEDINGS, 42nd Stanford Geothermal Workshop on Geothermal Reservoir Engineering*, (February):3–6.

Miyazawa, M., Snieder, R., and Venkataraman, A. (2008). Application of seismic interferometry to extract P- and S-wave propagation and observation of shear-wave splitting from noise data at Cold Lake, Alberta, Canada. *Geophysics*, 73(4):D35–D40.

Mordret, A., Jolly, A., Duputel, Z., and Fournier, N. (2010). Monitoring of phreatic eruptions using interferometry on retrieved cross-correlation function from ambient seismic noise: Results from Mt. Ruapehu, New Zealand. *Journal of Volcanology and Geothermal Research*, 191(1-2):46–59.

Morency, C. and Matzel, E. (2017). SGP-TR-212: Virtual Seismometers for induced seismicity monitoring and full moment tensor inversion. *PROCEEDINGS, 42nd Workshop on Geothermal Reservoir Engineering*, 5:1–5.

Mosher, C., Li, C., Morley, L., Ji, Y., Janiszewski, F., Olson, R., and Brewer, J. (2014). Increasing the efficiency of seismic data acquisition via compressive sensing. *Proceedings of the Annual Offshore Technology Conference*, 4:3273–3276.

Muir, J. B. and Zhan, Z. (2021a). Seismic wavefield reconstruction using a pre-conditioned wavelet–curvelet compressive sensing approach. *Geophysical Journal International*, 227(1):303–315.

Muir, J. B. and Zhan, Z. (2021b). Wavefield-based evaluation of DAS instrument response and array design. *Geophysical Journal International*, 229(1):21–34.

Naghizadeh, M. and Sacchi, M. D. (2007). Multistep autoregressive reconstruction of seismic records. *Geophysics*, 72(6):V111–V118.

Nakata, N., Chang, J. P., Lawrence, J. F., and Boué, P. (2015). Body wave extraction and tomography at Long Beach , California , with ambient-noise interferometry. *Journal of Geophysical Research: Solid Earth*, 120:1159–1173.

Nakata, N., Snieder, R., Kuroda, S., Ito, S., Aizawa, T., and Kunimi, T. (2013). Monitoring a building using deconvolution Interferometry. I: Earthquake-data analysis. *Bulletin of the Seismological Society of America*, 103(3):1662–1678.

Nakata, N., Snieder, R., Tsuji, T., Larner, K., and Matsuoka, T. (2011). Shear wave imaging from traffic noise using seismic interferometry by cross-coherence. *Geophysics*, 76(6):SA97–SA106.

Nizkous, I., Gerritsen, S., Kiyashchenko, D., and Joinson, D. (2015). Distributed Acoustic Sensing (DAS) VSP for Imaging and Velocity Model Building.

Nussbaumer, H. (1982). Fast Fourier Transform and Convolution Algorithms. *Springer-Verlag*, 1:80–111.

Oliveira, D. A. B., Ferreira, R. S., Silva, R., and Vital Brazil, E. (2018). Interpolating Seismic Data With Conditional Generative Adversarial Networks. *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, 15(12):1952–1956.

Otazo, R., Candès, E., and Sodickson, D. K. (2015). Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. *Magnetic Resonance in Medicine*, 73(3):1125–1136.

Ozerov, A., Philippe, P., Bimbot, F., and Gribonval, R. (2007). Adaptation of Bayesian Models for Single-Channel Source Separation and its Application to Voice/Music Separation in Popular Songs. *IEEE Transactions on Audio, Speech, and Language Processing*, 15(5):1564–1578.

Patel, V. M., Easley, G. R., Healy, D. M., and Chellappa, R. (2010). Compressed synthetic aperture radar. *IEEE Journal on Selected Topics in Signal Processing*, 4(2):244–254.

Pawelec, I., Sava, P., and Wakin, M. (2019). Wavefield reconstruction using wavelet transform. *SEG International Exposition and 89th Annual Meeting*, pages 147–151.

Peters, B., Herrmann, F., and van Leeuwen, T. (2014). Wave-equation Based Inversion with the Penalty Method - Adjoint-state Versus Wavefield-reconstruction Inversion. volume 2014, pages 1–5. European Association of Geoscientists & Engineers.

Pianese, G., Petrovic, B., Parolai, S., and Paolucci, R. (2018). Identification of the nonlinear seismic response of buildings by a combined Stockwell Transform and deconvolution interferometry approach. *Bulletin of Earthquake Engineering*, 16(7):3103–3126.

Poliannikov, O. V., Rondenay, S., and Chen, L. (2012). Interferometric imaging of the underside of a subducting crust. *Geophysical Journal International*, 189(1):681–690.

Poole, G. and Herrmann, P. (2007). Multidimensional data regularization for modern acquisition geometries. In *SEG Technical Program Expanded Abstracts 2007*, pages 2585–2589.

Prieto, G. A., Lawrence, J. F., and Beroza, G. C. (2009). Anelastic Earth structure from the coherency of the ambient seismic field. *Journal of Geophysical Research: Solid Earth*, 114(7):1–15.

Rani, M., Dhok, S. B., and Deshmukh, R. B. (2018). A Systematic Review of Compressive Sensing: Concepts, Implementations and Applications. *IEEE Access*, 6:4875–4894.

Rodriguez, I. V. (2021a). Compressive Sensing - Machine Learning combined for joint location and moment tensor estimation: a performance analysis. In *Proceedings, 82nd EAGE Annual Conference & Exhibition*, volume 2021, pages 1–5. European Association of Geoscientists & Engineers.

Rodriguez, I. V. (2021b). Towards fast DAS passive seismic monitoring combining Compressive Sensing with a deep learning decoder. *Conference Proceedings, EAGE GeoTech 2021 Second EAGE Workshop on Distributed Fibre Optic Sensing*, 2021(1):1–4.

Ronen, J. (1987). Wave-equation trace interpolation. *Geophysics*, 52(7):973–984.

Rose, J. H. (2001). "Single-sided" focusing of the time-dependent Schrodinger equation. *Physical Review A*, 65(1):12707.

Saengduean, P., Snieder, R., and Wakin, M. B. (2022). Multi-source wavefield reconstruction combining interferometry and compressive sensing: application to a linear receiver array. *Submitted to Geophys. J. Int.*, pages 1–26.

Sandoval-Hernandez, M., Vazquez-Leal, H., Hernandez-Martinez, L., Filobello-Nino, U. A., Jimenez-Fernandez, V. M., Herrera-May, A. L., Castaneda-Sheissa, R., Ambrosio-Lazaro, R. C., and Diaz-Arango, G. (2018). Approximation of Fresnel integrals with applications to diffraction problems. *Mathematical Problems in Engineering*, 2018:1–13.

Schuster, G. T. (2009). *Seismic Interferometry*. Cambridge University Press.

Shapiro, N. M., Campillo, M., Stehly, L., and Ritzwoller, M. H. (2005). High-resolution surface-wave tomography from ambient seismic noise. *Science*, 307(5715):1615–1618.

Sharma, S. K., Lagunas, E., Chatzinotas, S., and Ottersten, B. (2016). Application of compressive sensing in cognitive radio communications: A survey. *IEEE Communications Surveys and Tutorials*, 18(3):1838–1860.

Shen, Z. and Zhan, Z. (2020). Metastable olivine wedge beneath the Japan sea imaged by seismic interferometry. *Geophysical Research Letters*, 47(6):1–8.

Shin, H.-C. (2013). Weighted Least Squares Estimation with Sampling Weights. In *JSM 2013 - Survey Research Methods Section*, pages 1523–1530.

Shirzad, T., Riahi, M. A., and Assumpção, M. S. (2019). Crustal structure of the collision-subduction zone in South of Iran using virtual seismometers. *Scientific Reports*, 9(1):1–11.

Shragge, J., Yang, J., Issa, N., Roelens, M., Dentith, M., and Schediwy, S. (2021). Low-frequency ambient distributed acoustic sensing (DAS): Case study from Perth, Australia. *Geophysical Journal International*, 226(1):564–581.

Siahkoohi, A., Kumar, R., and Herrmann, F. (2018). Seismic Data Reconstruction with Generative Adversarial Networks. In *80th EAGE Conference and Exhibition*, volume 2018, pages 1–5. European Association of Geoscientists & Engineers.

Sidenko, E., Tertyshnikov, K., Bona, A., and Pevzner, R. (2021). DAS-VSP interferometric imaging: CO2CRC Otway Project feasibility study. *Interpretation*, 9(4):SJ1–SJ12.

Snieder, R. (2004). Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E*, 69(4):046610.

Snieder, R. and Larose, E. (2013a). Extracting Earth's elastic wave response from noise measurements. *Annual Review of Earth and Planetary Sciences*, 41(1):183–206.

Snieder, R. and Larose, E. (2013b). Extracting Earth's elastic wave response from noise measurements. *Annual Review of Earth and Planetary Sciences*, 41(1):183–206.

Snieder, R., Miyazawa, M., Slob, E., Vasconcelos, I., and Wapenaar, K. (2009). A comparison of strategies for seismic interferometry. *Surveys in Geophysics*, 30(4-5):503–523.

Snieder, R. and van Wijk, K. (2015a). *A guided tour of mathematical methods for physical sciences*. Cambridge University Press, 3rd edition.

Snieder, R. and van Wijk, K. (2015b). *A Guided Tour of Mathematical Methods for the Physical Sciences*. Cambridge University Press, 3rd edition.

Snieder, R., Van Wijk, K., Haney, M., and Calvert, R. (2008). Cancellation of spurious arrivals in Green's function extraction and the generalized optical theorem. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 78(3):036606.

Snieder, R. and Wakin, M. B. (2022). When randomness helps in undersampling. *SIAM Review*, 64(4):1062–1080.

Snieder, R., Wapenaar, K., and Larner, K. (2006). Spurious multiples in seismic interferometry of primaries. *GEOPHYSICS*, 71(4):SI111–SI124.

Snieder, R., Wapenaar, K., and Wegler, U. (2007). Unified Green's function retrieval by cross-correlation; Connection with energy principles. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 75(3):036103.

Spilker, M. E. and Vicini, P. (2001). An evaluation of extended vs weighted least squares for parameter estimation in physiological modeling. *Journal of Biomedical Informatics*, 34(5):348–364.

Spitz, S. (1991). Seismic trace interpolation in the F-X domain. *Geophysics*, 56(6):785–794.

Stehly, L., Cupillard, P., and Romanowicz, B. (2011). Towards improving ambient noise tomography using simultaneously curvelet denoising filters and SEM simulations of seismic ambient noise. *Comptes Rendus - Geoscience*, 343(8-9):591–599.

Stockmann, F., Cobden, L., Deschamps, F., Fichtner, A., and Thomas, C. (2019). Investigating the seismic structure and visibility of dynamic plume models with seismic array methods. *Geophysical Journal International*, 219:S167–S194.

Sweet, J. R., Anderson, K. R., Bilek, S., Brudzinski, M., Chen, X., DeShon, H., Hayward, C., Karplus, M., Keranen, K., Langston, C., Lin, F., Beatrice Magnani, M., and Woodward, R. L. (2018). A Community Experiment to Record the Full Seismic Wavefield in Oklahoma. *Seismological Research Letters*, 89(5):1923–1930.

Tegtmeier, S., Gisolf, A., and Verschuur, D. J. (2004). 3D sparse-data Kirchhoff redatuming. *Geophysical Prospecting*, 52(6):509–521.

Trad, D. (2009). Five-dimensional interpolation: Recovering from acquisition constraints. *GEOPHYSICS*, 74(6):V123–V132.

Trickett, S., Burroughs, L., Milton, A., Walton, L., and Dack, R. (2010). Rank-reduction-based trace interpolation. *Society of Exploration Geophysicists International Exposition and 80th Annual Meeting*, pages 3829–3833.

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

Tuck, E. O. (1967). A simple "Filon-Trapezoidal" rule. *Mathematics of Computation*, 21(98):239–241.

Unsworth, M. J., Malin, P. E., Egbert, G. D., and Booker, J. R. (1997). Internal structure of the San Andreas Fault at Parkfield, California. *Geology*, 25(4):359–362.

Vaidyanathan, P. P. (2007). *The theory of linear prediction*, volume 3.

Vallage, A., Klinger, Y., Grandin, R., Bhat, H. S., and Pierrot-Deseilligny, M. (2015). Inelastic surface deformation during the 2013 Mw 7.7 Balochistan, Pakistan, earthquake. *Geology*, 43(12):1079–1082.

van Dalen, K. N., Mikesell, T. D., Ruigrok, E. N., and Wapenaar, K. (2015). Retrieving surface waves from ambient seismic noise using seismic interferometry by multidimensional deconvolution. *Journal of Geophysical Research: Solid Earth*, 120(2):944–961.

Vasconcelos, I. and Snieder, R. (2008a). Interferometry by deconvolution, Part 1 - Theory for acoustic waves and numerical examples. *Geophysics*, 73(3):S115–S128.

Vasconcelos, I. and Snieder, R. (2008b). Interferometry by deconvolution: Part 2 — Theory for elastic waves and application to drill-bit seismic imaging. *Geophysics*, 73(3):S129–S141.

Vembu, S. and Baumann, S. (2005). Separation of vocals from polyphonic audio recordings. In *IN: PROC. ISMIR2005. (2005) 337–344*, pages 337–344.

Vijay Kartik, S., Carrillo, R. E., Thiran, J.-P., and Wiaux, Y. (2017). A Fourier dimensionality reduction model for big data interferometric imaging. *Monthly Notices of the Royal Astronomical Society*, 468(2):2382–2400.

Virtanen, T., Mesaros, A., and Ryynänen, M. (2008). Combining pitch-based inference and non-negative spectrogram factorization in separating vocals from polyphonic music. In *Proc. ITRW on Statistical and Perceptual Audio Processing (SAPA 2008)*, pages 17–22.

Wakin, M. B. (2017). Compressive sensing fundamentals. *Compressive Sensing for Urban Radar*, pages 1–47.

Walter, F., Gräff, D., Lindner, F., Paitz, P., Köpfli, M., Chmiel, M., and Fichtner, A. (2020). Distributed acoustic sensing of microseismic sources and wave propagation in glaciated terrain. *Nature Communications*, 11(1):2436.

Wang, B., Chen, X., Li, J., and Cao, J. (2016). An Improved Weighted Projection onto Convex Sets Method for Seismic Data Interpolation and Denoising. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(1):228–235.

Wang, B. and Lu, W. (2017). Accurate and efficient seismic data interpolation in the principal frequency wavenumber domain. *Journal of Geophysics and Engineering*, 14(6):1475–1483.

Wang, B., Wu, R. S., Geng, Y., and Chen, X. (2014). Dreamlet-based interpolation using POCS method. *Journal of Applied Geophysics*, 109:256–265.

Wang, B., Zhang, N., Lu, W., and Wang, J. (2019). Deep-learning-based seismic data interpolation: A preliminary result. *Geophysics*, 84(1):V11–V20.

Wang, B. F., Li, J. Y., Chen, X. H., and Liu, Y. (2015). Curvelet-based 3D reconstruction of digital cores using POCS method. *Chinese Journal of Geophysics*, 58(5):486–495.

Wang, J., Ng, M., and Perz, M. (2010). Seismic data interpolation by greedy local Radon t1. Yu, Z., McMechan, G. A., Ferguson, J. F. & Anno, P. D. Wavelet-Radon domain de-aliasing and interpolation of seismic data. 2004 SEG Annu. Meet. 72, (2004).ransform. *Geophysics*, 75(6):WB225–WB234.

Wang, Y., Ge, Q., Lu, W., and Yan, X. (2020). Well-Logging Constrained Seismic Inversion Based on Closed-Loop Convolutional Neural Network. *IEEE Transactions on Geoscience and Remote Sensing*, 58(8):5564–5574.

Wang, Y., Luo, Y., and Schuster, G. T. (2009). Interferometric interpolation of missing seismic data. *Geophysics*, 74(3):SI37–SI45.

Wapenaar, C. P. A. (1993). Kirchhoff-Helmholtz downward extrapolation in a layered medium with curved interfaces. *Geophysical Journal International*, 115(2):445–455.

Wapenaar, K., Draganov, D., Snieder, R., Campman, X., and Verdel, A. (2010). Tutorial on seismic interferometry: Part 1 — Basic principles and applications. *Geophysics*, 75(5):195–209.

Wapenaar, K. and Fokkema, J. (2006). Green's function representations for seismic interferometry. *Geophysics*, 71(4):SI33–SI46.

Wapenaar, K., Thorbecke, J., van der Neut, J., Broggini, F., Slob, E., and Snieder, R. (2014). Marchenko imaging. *GEOPHYSICS*, 79(3):WA39–WA57.

William A. Schneider, J., Phillip, L. D., and Paal, E. F. (1995). Wave-equation velocity replacement of the low-velocity layer for overthrust-belt data. *GEOPHYSICS*, 60(2):573–579.

Williams, E. F., Zhan, Z., Martins, H. F., Fernández-Ruiz, M. R., Mart\'\in-López, S., González-Herráez, M., and Callies, J. (2022). Surface Gravity Wave Interferometry and Ocean Current Monitoring With Ocean-Bottom DAS. *Journal of Geophysical Research (Oceans)*, 127(5):e18375.

Wu, J. and Bai, M. (2018). Adaptive rank-reduction method for seismic data reconstruction. *Journal of Geophysics and Engineering*, 15(4):1688–1703.

Wu, W., Zhan, Z., Peng, S., Ni, S., and Callies, J. (2020). Seismic ocean thermometry. *Science*, 369(6510):1510–1515.

Xia, J., Miller, R. D., and Park, C. B. (1999). Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves. *Geophysics*, 64(3):691–700.

Xiao, X. and Schuster, G. T. (2006). Redatuming CDP data below salt with VSP green's Function. In *SEG Technical Program Expanded Abstracts 2006*, pages 3511–3515.

Xu, Z., Sopher, D., Juhlin, C., Han, L., and Gong, X. (2018). Radon-domain interferometric interpolation for reconstruction of the near-offset gap in marine seismic data. *Journal of Applied Geophysics*, 151:125–141.

Yang, A. Y., Sastry, S. S., Ganesh, A., and Ma, Y. (2010). Fast Í1-minimization algorithms and an application in robust face recognition: A review. In *2010 IEEE International Conference on Image Processing*, pages 1849–1852.

Yang, J., Shragge, J., and Jin, G. (2021). 4D DAS fiber-coupling effects in freezing near-surface ground conditions. *First International Meeting for Applied Geoscience & Energy Expanded Abstracts*, pages 477–482.

Yang, P., Gao, J., and Chen, W. (2012). Curvelet-based POCS interpolation of nonuniformly sampled seismic records. *Journal of Applied Geophysics*, 79:90–99.

Young, C., Shragge, J., Schultz, W., Haines, S., Oren, C., Simmons, J., and Collett, T. S. (2022). Advanced Distributed Acoustic Sensing Vertical Seismic Profile Imaging of an Alaska North Slope Gas Hydrate Field. *Energy & Fuels*, 36(7):3481–3495.

Yu, W.-c., Wen, L., and Niu, F. (2005). Seismic velocity structure in the Earth's outer core. *Journal of Geophysical Research: Solid Earth*, 110:B02302.

Yu, Z., McMechan, G. A., Ferguson, J. F., and Anno, P. D. (2007). Wavelet-Radon domain dealiasing and interpolation of seismic data. *Geophysics*, 72(2):V41–V49.

Yuan, K. and Romanowicz, B. (2017). Seismic evidence for partial melting at the root of major hot spot plumes. *Science*, 357(6349):393–397.

Zeng, X., Lancelle, C., Thurber, C., Fratta, D., Wang, H., Lord, N., Chalari, A., and Clarke, A. (2017). Properties of noise cross-correlation functions obtained from a distributed acoustic sensing array at Garner Valley, California. *Bulletin of the Seismological Society of America*, 107(2):603–610.

Zhan, Z., Li, Q., and Huang, J. (2018). Application of wavefield compressive sensing in surface wave tomography. *Geophysical Journal International*, 213(3):1731–1743.

Zhang, H., Wang, Y., Ning, x., Lv, K., and Wang, N. (2019). Volume Data Denoising via Extended Weighted Least Squares. *IEEE Access*, 7:2750–2758.

Zhang, Z., Xu, K., and Ren, F. (2021). Selective Sensing: A Data-driven Nonuniform Subsampling Approach for Computation-free On-Sensor Data Dimensionality Reduction. In *ICLR 2021 Conference Blind Submission*.

Zhou, T. and Han, D. (2008). A weighted least squares method for scattered data fitting. *Journal of Computational and Applied Mathematics*, 217(1):56–63.

Zinke, R., Hollingsworth, J., and Dolan, J. F. (2014). Surface slip and off-fault deformation patterns in the 2013 MW 7.7 Balochistan, Pakistan earthquake: Implications for controls on the distribution of near-surface coseismic slip. *Geochemistry, Geophysics, Geosystems*, 15(12):5034–5050.

Zoback, M., Hickman, S., and Ellsworth, W. (2011). Scientific drilling into the San Andreas fault zone - An overview of SAFOD's first five years. *Scientific Drilling*, (1):14–28.

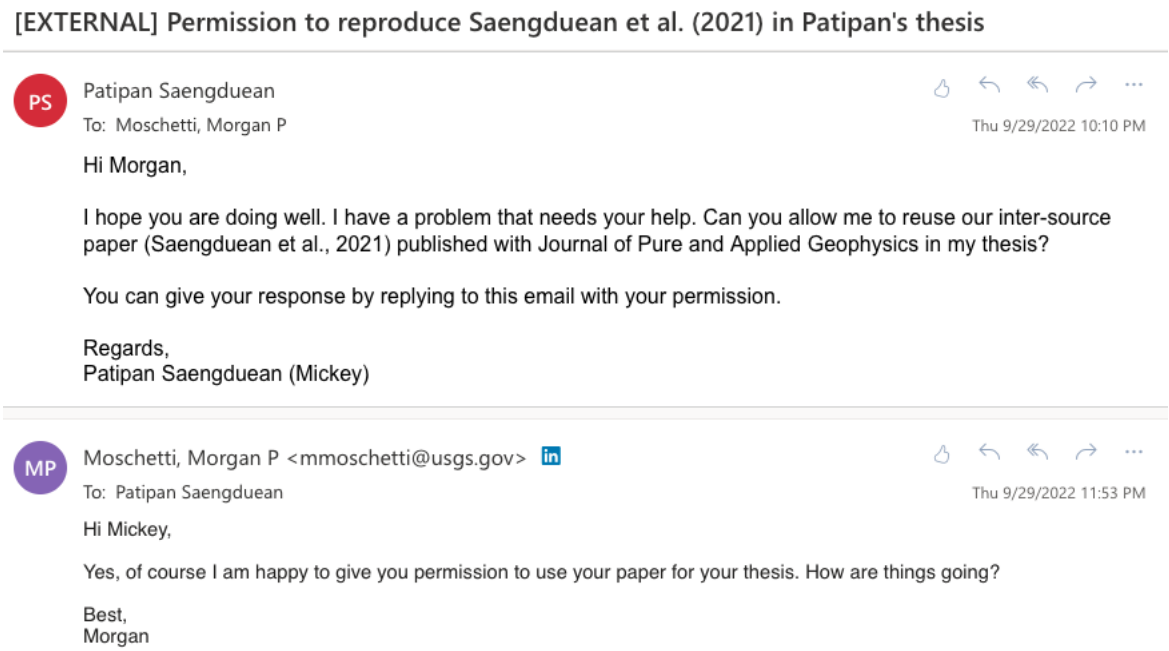
APPENDIX

COPYRIGHT PERMISSIONS

This appendix includes the permission from my co-authors and the Journal to modify and reuse my published paper as a chapter of my thesis.

A.1 Permission from Co-author

Chapter 2 is reproduced with a permission from a co-author, who works in different institute, for a paper published at Journal of Pure and Applied Geophysics. Figure A.1 shows the permission from the co-author for the published paper.


[EXTERNAL] Permission to reproduce Saengduean et al. (2021) in Patipan's thesis

PS Patipan Saengduean

To: Moschetti, Morgan P

Thu 9/29/2022 10:10 PM

Hi Morgan,

I hope you are doing well. I have a problem that needs your help. Can you allow me to reuse our inter-source paper (Saengduean et al., 2021) published with Journal of Pure and Applied Geophysics in my thesis?

You can give your response by replying to this email with your permission.

Regards,
Patipan Saengduean (Mickey)

MP Moschetti, Morgan P <mmoschetti@usgs.gov>

To: Patipan Saengduean

Thu 9/29/2022 11:53 PM

Hi Mickey,

Yes, of course I am happy to give you permission to use your paper for your thesis. How are things going?

Best,
Morgan

Figure A.1 Copyright permission to reuse the published paper from the co-author of Chapter 2.

A.2 Permission for Published Paper

I modify and reproduce a paper published at Journal of Pure and Applied Geophysics as Chapter 2 in my thesis. The copyright permission from the journal to modify and reuse the paper in my thesis are shown separately in Part 1 (Figure A.2), Part 2 (Figure A.3), Part 3 (Figure A.4), Part 4 (Figure A.5), Part 5 (Figure A.6), and Part 6 (Figure A.7).

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Oct 05, 2022

This Agreement between Colorado School of Mines -- Patipan Saengduean ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5377970128205
License date	Aug 28, 2022
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Pure and Applied Geophysics
Licensed Content Title	Inter-Source Interferometry of Seismic Body Waves: Required Conditions and Examples
Licensed Content Author	Patipan Saengduean et al
Licensed Content Date	Jul 19, 2021
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	full article/chapter
Will you be translating?	no

Figure A.2 Part 1 of the copyright permission to reuse my published paper in my thesis.

Circulation/distribution	50000 or greater
Author of this Springer Nature content	yes
Title	Wavefield reconstruction using seismic interferometry and compressive sensing
Institution name	Colorado School of Mines
Expected presentation date	Oct 2022
Requestor Location	Patipan Saengduean 13000 NW 119 St. MIAMI, FL 33186 United States Attn: Colorado School of Mines
Total	0.00 USD

Terms and Conditions

Springer Nature Customer Service Centre GmbH
Terms and Conditions

This agreement sets out the terms and conditions of the licence (the **Licence**) between you and **Springer Nature Customer Service Centre GmbH** (the **Licensor**). By clicking 'accept' and completing the transaction for the material (**Licensed Material**), you also confirm your acceptance of these terms and conditions.

1. Grant of License

1. 1. The Licensor grants you a personal, non-exclusive, non-transferable, world-wide licence to reproduce the Licensed Material for the purpose specified in your order only. Licences are granted for the specific use requested in the order and for no other use, subject to the conditions below.

1. 2. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of the Licensed Material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of

Figure A.3 Part 2 of the copyright permission to reuse my published paper in my thesis.

another entity (as credited in the published version).

1. 3. If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Scope of Licence

2. 1. You may only use the Licensed Content in the manner and to the extent permitted by these Ts&Cs and any applicable laws.

2. 2. A separate licence may be required for any additional use of the Licensed Material, e.g. where a licence has been purchased for print only use, separate permission must be obtained for electronic re-use. Similarly, a licence is only valid in the language selected and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence. Any content owned by third parties are expressly excluded from the licence.

2. 3. Similarly, rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.

2. 4. Where permission has been granted **free of charge** for material in print, permission may also be granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

2. 5. An alternative scope of licence may apply to signatories of the STM Permissions Guidelines, as amended from time to time.

3. Duration of Licence

3. 1. A licence for is valid from the date of purchase ('Licence Date') at the end of the relevant period in the below table:

Scope of Licence	Duration of Licence
Post on a website	12 months
Presentations	12 months
Books and journals	Lifetime of the edition in the language purchased

4. Acknowledgement

4. 1. The Licensor's permission must be acknowledged next to the Licenced Material in print. In electronic form, this acknowledgement must be visible at the same time as the

Figure A.4 Part 3 of the copyright permission to reuse my published paper in my thesis.

figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

5. Restrictions on use

5. 1. Use of the Licensed Material may be permitted for incidental promotional use and minor editing privileges e.g. minor adaptations of single figures, changes of format, colour and/or style where the adaptation is credited as set out in Appendix 1 below. Any other changes including but not limited to, cropping, adapting, omitting material that affect the meaning, intention or moral rights of the author are strictly prohibited.

5. 2. You must not use any Licensed Material as part of any design or trademark.

5. 3. Licensed Material may be used in Open Access Publications (OAP) before publication by Springer Nature, but any Licensed Material must be removed from OAP sites prior to final publication.

6. Ownership of Rights

6. 1. Licensed Material remains the property of either Licensor or the relevant third party and any rights not explicitly granted herein are expressly reserved.

7. Warranty

IN NO EVENT SHALL LICENSOR BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL OR INDIRECT DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND
WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

8. Limitations

8. 1. ***BOOKS ONLY:*** Where **'reuse in a dissertation/thesis'** has been selected the following terms apply: Print rights of the final author's accepted manuscript (for clarity, NOT the published version) for up to 100 copies, electronic rights for use only on a

Figure A.5 Part 4 of the copyright permission to reuse my published paper in my thesis.

personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

8. 2. For content reuse requests that qualify for permission under the STM Permissions Guidelines, which may be updated from time to time, the STM Permissions Guidelines supersede the terms and conditions contained in this licence.

9. Termination and Cancellation

9. 1. Licences will expire after the period shown in Clause 3 (above).

9. 2. Licensee reserves the right to terminate the Licence in the event that payment is not received in full or if there has been a breach of this agreement by you.

Appendix 1 — Acknowledgements:

For Journal Content:
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication papers:**
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)

For Adaptations/Translations:
Adapted/Translated by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

Note: For any republication from the British Journal of Cancer, the following credit line style applies:

Reprinted/adapted/translated by permission from [**the Licensor**]: on behalf of Cancer Research UK: : [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), **[COPYRIGHT]** (year of publication)

For **Advance Online Publication** papers:
Reprinted by permission from The [**the Licensor**]: on behalf of Cancer Research UK: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM])

Figure A.6 Part 5 of the copyright permission to reuse my published paper in my thesis.

For Book content:
Reprinted/adapted by permission from [**the Licensor**]: [**Book Publisher** (e.g. Palgrave Macmillan, Springer etc) [**Book Title**] by [**Book author**(s)] [**COPYRIGHT**] (year of publication)

Other Conditions:

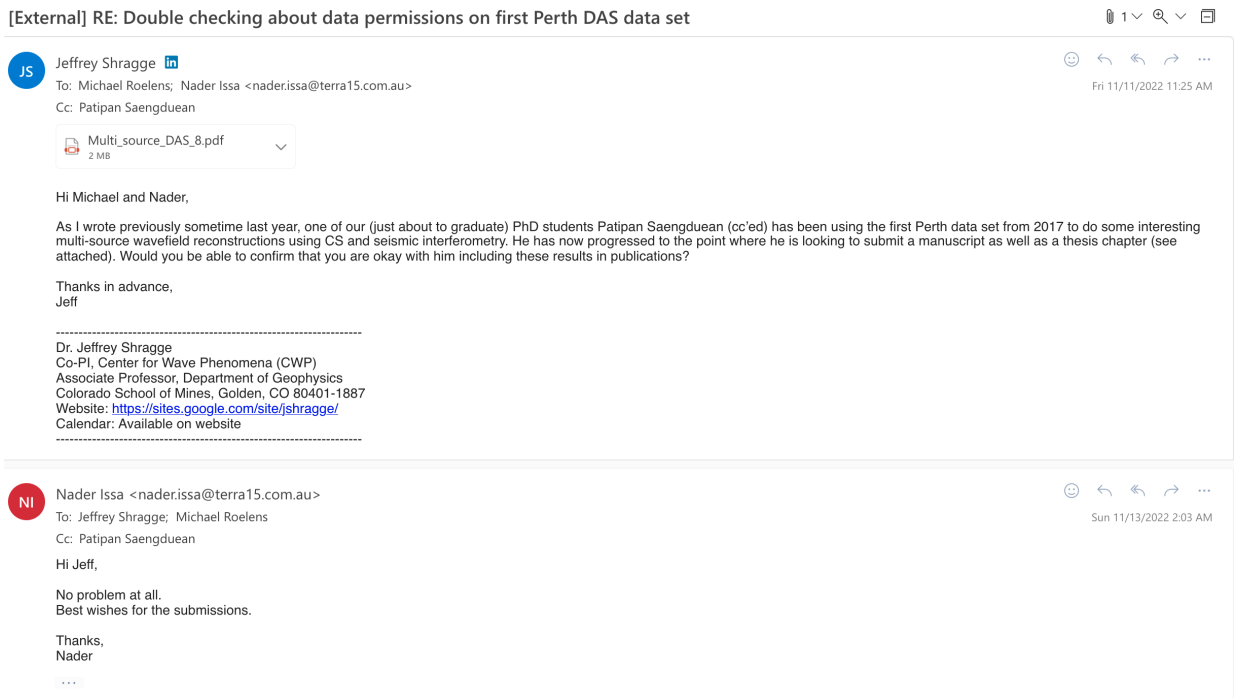
Version 1.3

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

Figure A.7 Part 6 of the copyright permission to reuse my published paper in my thesis.

A.3 Permission for DAS Data

In Chapter 4, I use the DAS data collected by Terra15 Pty Ltd. for the data analyses. Figure A.8 shows the permission from a representative of Terra15 Pty Ltd. to use the DAS data for my thesis.



[External] RE: Double checking about data permissions on first Perth DAS data set

JS Jeffrey Shragge

To: Michael Roelens; Nader Issa <nader.issa@terra15.com.au>

Cc: Patipan Saengduean

Fri 11/11/2022 11:25 AM

Multi_source_DAS_8.pdf
2 MB

Hi Michael and Nader,

As I wrote previously sometime last year, one of our (just about to graduate) PhD students Patipan Saengduean (cc'ed) has been using the first Perth data set from 2017 to do some interesting multi-source wavefield reconstructions using CS and seismic interferometry. He has now progressed to the point where he is looking to submit a manuscript as well as a thesis chapter (see attached). Would you be able to confirm that you are okay with him including these results in publications?

Thanks in advance,
Jeff

Dr. Jeffrey Shragge
Co-PI, Center for Wave Phenomena (CWP)
Associate Professor, Department of Geophysics
Colorado School of Mines, Golden, CO 80401-1887
Website: https://sites.google.com/site/jshragge/
Calendar: Available on website

NI Nader Issa <nader.issa@terra15.com.au>

To: Jeffrey Shragge; Michael Roelens

Cc: Patipan Saengduean

Sun 11/13/2022 2:03 AM

Hi Jeff,

No problem at all.
Best wishes for the submissions.

Thanks,
Nader

Figure A.8 Permission to use the DAS data collected by Terra15 Pty Ltd. from a representative of Terra15 Pty Ltd.