Multi-source wavefield reconstruction combining interferometry and compressive sensing: application to a linear receiver array

Patipan Saengduean[1], Roel Snieder[1], & Michael B. Wakin[2]

1. Center for Wave Phenomena and Dept. of Geophysics, Colorado School of Mines, Golden CO 80401
2. Dept. of Electrical Engineering, Colorado School of Mines, Golden CO 80401
Email psaengduean@mines.edu

ABSTRACT

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.

Key words: Seismic interferometry, compressive sensing, wavefield reconstruction.

1 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 review papers (Larose et al., 2006; Curtis et al., 2006; Wapenaar et al., 2010; Snieder and Larose, 2013). 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, 2008b,a), 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; Poole and Herrmann, 2007; Cao et al., 2018). 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 (Peters et al., 2014; Kim et al., 2015), linear prediction filter methods (Vaidyanathan, 2007; Islam et al., 2015), rank-reduction methods (Wu and Bai, 2018; Innocent Oboué et al., 2021), 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 (Herrmann et al., 2008; Hennenfent et al., 2010) 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 (Wang et al., 2009; Hanafy and Schuster, 2014; 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 2, 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, we show the different wavefield reconstruction techniques using different sparse transforms and compare these reconstructions to our proposed multi-source wavefield reconstruction.

2 THEORY AND SYNTHETIC EXAMPLE

2.1 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

$$u(\mathbf{x_r}, \mathbf{x}, \omega) = W(\mathbf{x}, \omega) G(\mathbf{x_r}, \mathbf{x}, \omega), \tag{1}$$

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

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

where the asterisk denotes complex conjugation. Integrating the cross-correlation equation (eq 2) over a closed surface that includes uncorrelated sources on the surface surrounding the receivers (Snieder et al., 2007) gives

$$\oint C_{ba} d^2\mathbf{x} = \langle |W(\mathbf{x},\omega)|^2 \rangle \oint G(\mathbf{x_b},\mathbf{x},\omega) G^*(\mathbf{x_a},\mathbf{x},\omega) d^2\mathbf{x}, \tag{3}$$

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 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

$$\big[G(\mathbf{x_a},\mathbf{x_b},\omega) - G^*(\mathbf{x_a},\mathbf{x_b},\omega)\big] \langle |W(\mathbf{x},\omega)|^2 \rangle \approx -\frac{2i\omega}{\rho c} \oint C_{ba} d^2\mathbf{x}, \tag{4}$$

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 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 4.

2.2 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 (Donoho, 2006; Wakin, 2017; Candès and Wakin, 2008). 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 (Wakin, 2017; Candès and Wakin, 2008). 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 $||n||_2 \leq \sigma$,

$$min||\alpha||_1 \text{ subject to } ||\Phi\Psi\alpha - d||_2 \leq \sigma. \tag{5}$$

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. (2006), to reconstruct interferometric wavefields.

2.3 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 5 m from each other and the sources are uniformly distributed on a rectangle

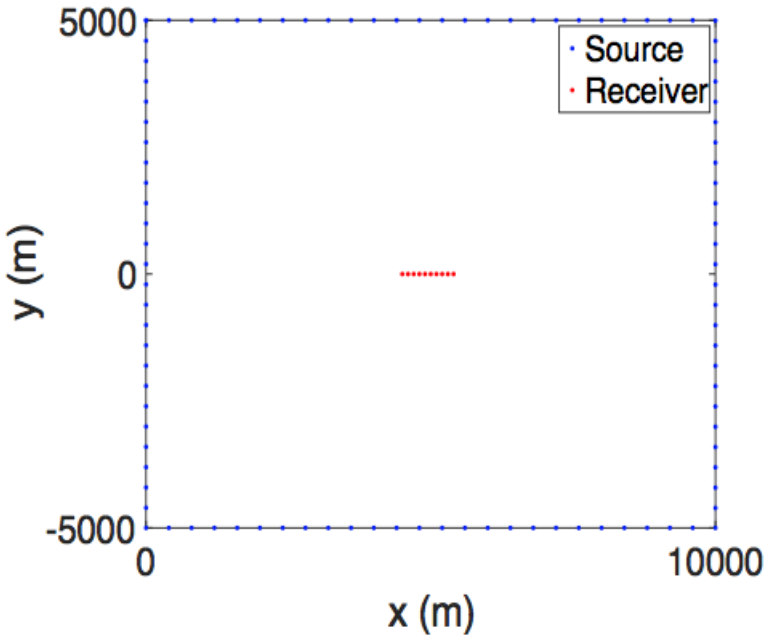

Figure 1. Geometry of sources and receivers used for the numerical examples. The sources (blue dots) are uniformly distributed on a rectangle surrounding a linear receiver array (red dots). For clarity the figure does not show all the sources and receivers. In the example, 100 receivers and 2000 sources are uniformly spaced.

surrounding the array (Figure 1). 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 (2015)) given by

$$u(x,t) = \int W(\omega) H_0^{(1)}(k(\omega)x) e^{-i\omega t} d\omega, \tag{6}$$

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 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 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.

3 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.

Using a linear array consisting of 3 available and 2 missing receivers as an example, Figure 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 2 shows 3 correlation profiles associated with the 3 available receivers (VS1, VS3, and VS4) before wavefield recovery.

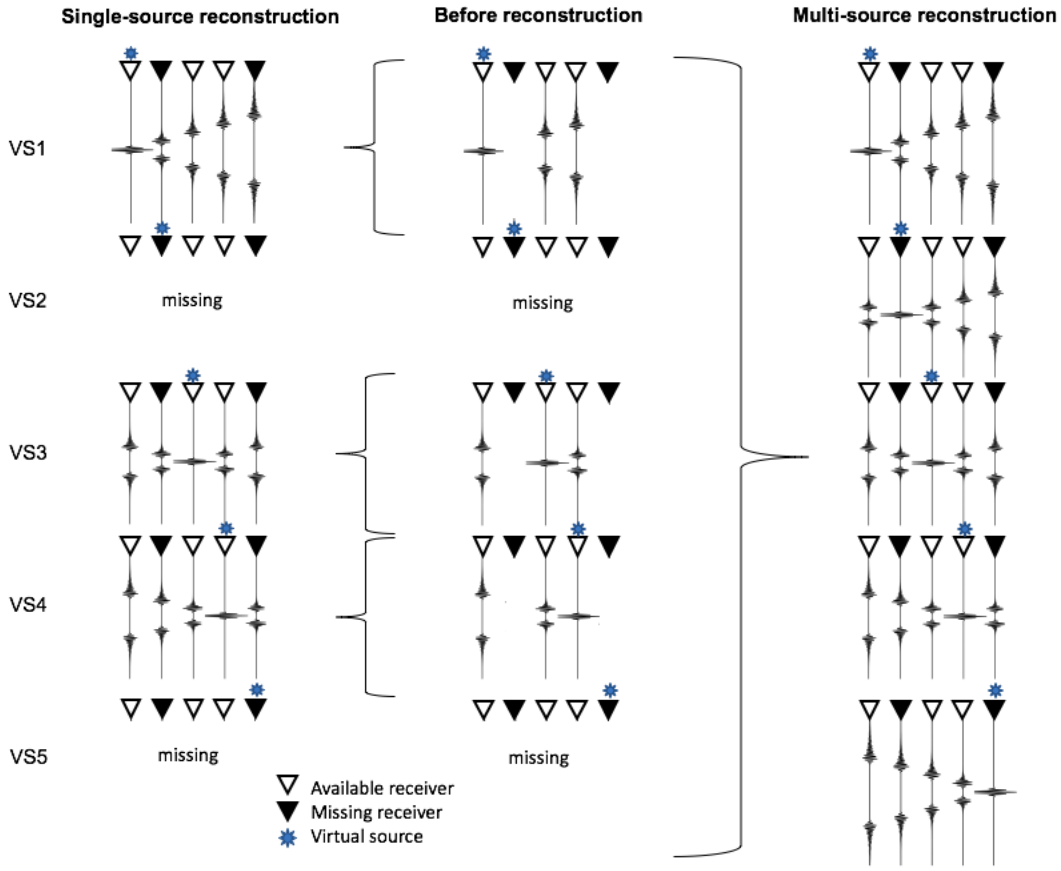

Figure 2. Difference between the two schemes of wavefield recovery of correlation profiles of a linear array. The middle column shows the correlation gathers before wavefield reconstruction. Single-source reconstruction (left column) of interferometric wavefields uses correlograms with gaps (i.e. missing 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.

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 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 2).

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 $20log(||x_{ori}||_2/||x_{ori} - x_{rec}||_2)$, 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. (2006), using arbitrary scale and direction parameters given by 4 and 16, respectively.

Figure 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 4, respectively. The Fourier and Curvelet single-source reconstructions give inaccurate

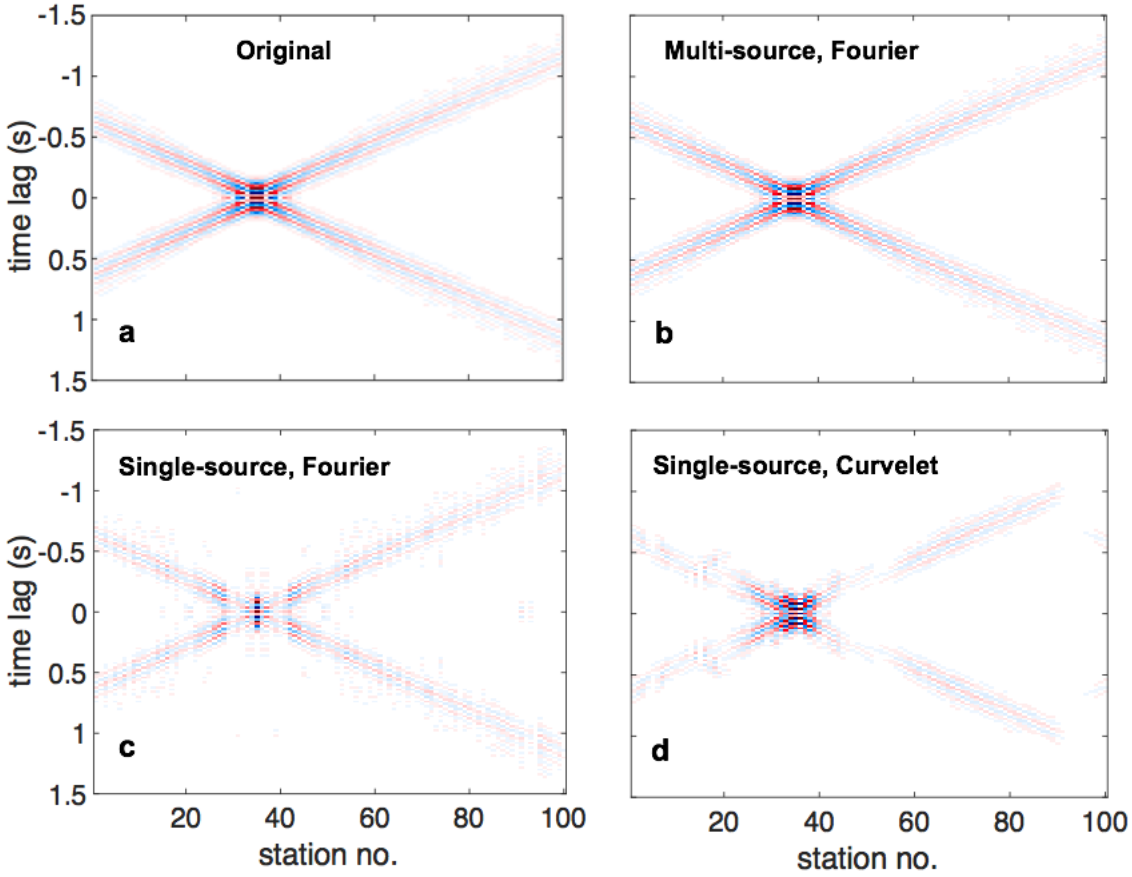

Figure 3. Comparison between (a) the original cross-correlated wavefields when the master sensor is the 35th receiver, and the wavefield images recovered by (b) Fourier multi-source reconstruction, (c) Fourier single-source reconstruction, and (d) Curvelet single-source reconstruction. In all reconstructions, 80% of the sensors are absent.

wavefield reconstruction, in particular within 100 m (20 receivers) from the master sensor (Figure 3c and 3d). In contrast, the Fourier multi-source reconstruction provides the interferometric wavefields (Figure 3b) recovered similarly to the original cross-correlated wavefields (Figure 3a).

Figure 5 shows the error of different wavefield reconstructions, which is determined by the relative difference between each reconstruction image (Figures 3b, 3c, and 3d) and the original image normalized with the maximum amplitude of the original image (Figure 3a). Note that the colorbar scale for the single- and multi- source methods are different (Figure 5). The relative error of the multi-source reconstruction is below 0.1% (Figure 5a) and is much smaller than the error of the single-source reconstruction that ranges between 0-50% (Figure 5b and 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 5b and 5c). The Fourier and Curvelet single-source reconstructions produce a relative difference of approximately 40% at locations within 100 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 5a).

To investigate the quality of wavefield reconstruction along the receiver array, we determine the SNR for the reconstruction using different master sensors. Figure 6 shows the SNR of different reconstruction schemes averaged over different master sensors. In Figure 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 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 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.

Using five different realizations of missing receiver locations, we estimate the mean and standard deviation of the SNR for the different reconstruction schemes. Figure 7 shows the SNR comparison for the different reconstructions for four different fractions

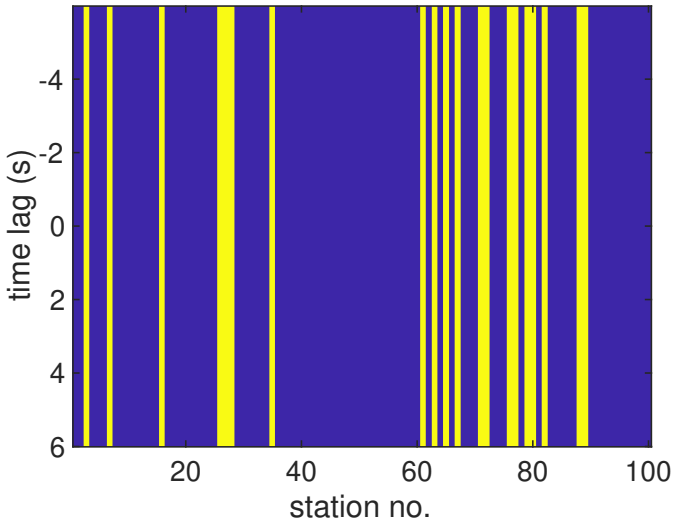

Figure 4. Available and missing seismic traces used in our wavefield reconstruction examples (Figure 3), shown in yellow and blue, respectively.

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 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.

4 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 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 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. (2015) and Arianandra 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.

In Section 3, 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 that our proposed multi-source reconstruction, which uses cross-correlated wavefields provided from all available virtual sources, improves the recovery of inter-

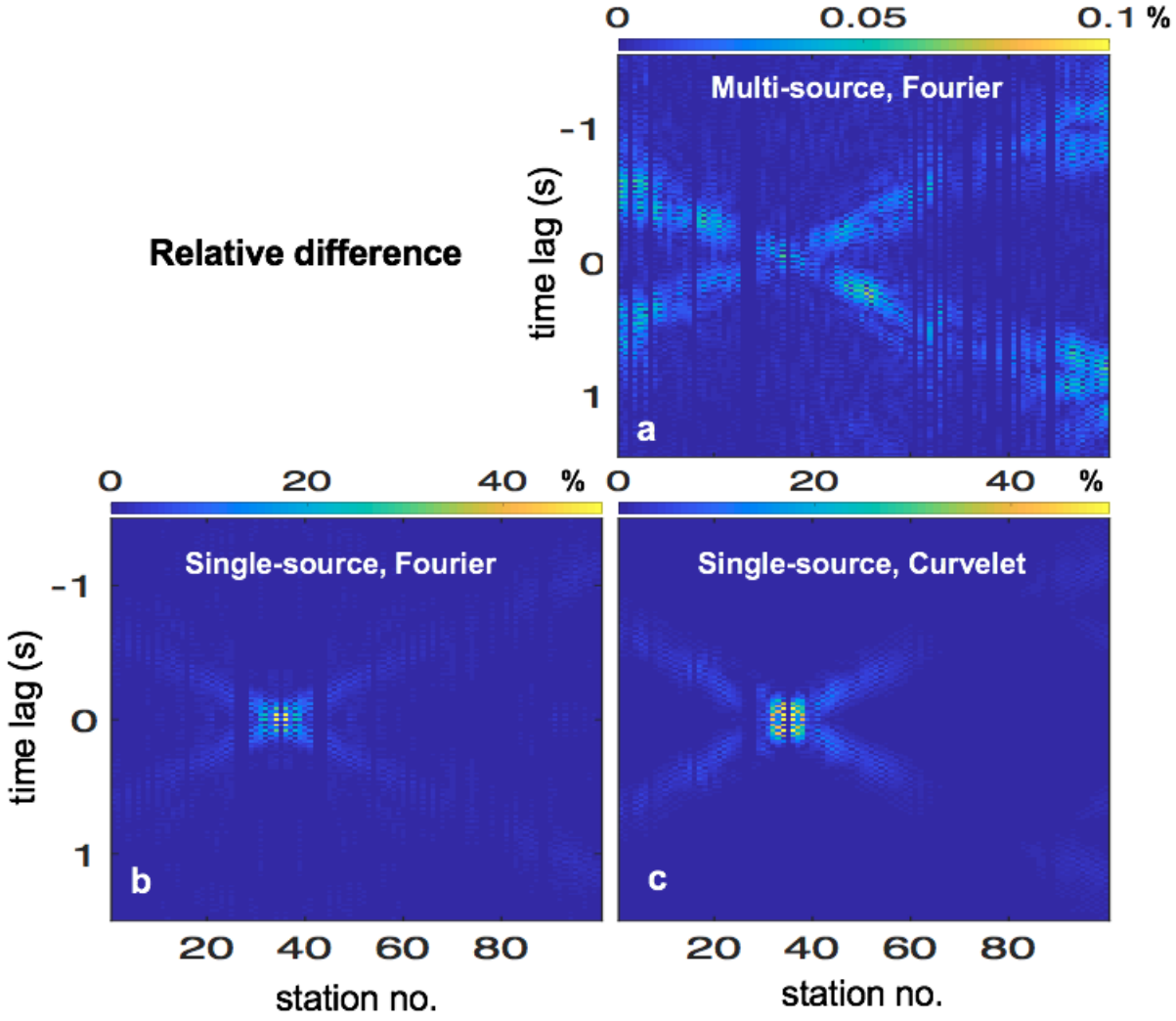

Figure 5. The relative differences (%) between the original wavefield image (Figure 3a) and the reconstructed images using (a) Fourier single-source reconstruction, (b) Curvelet single-source reconstruction, and (c) Fourier multi-source reconstruction. The colorbar scales of the two single-source reconstructions are different and much larger than the scale of the multi-source reconstruction.

ferometric 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 (Wakin, 2017; Candès and Wakin, 2008). 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.

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 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

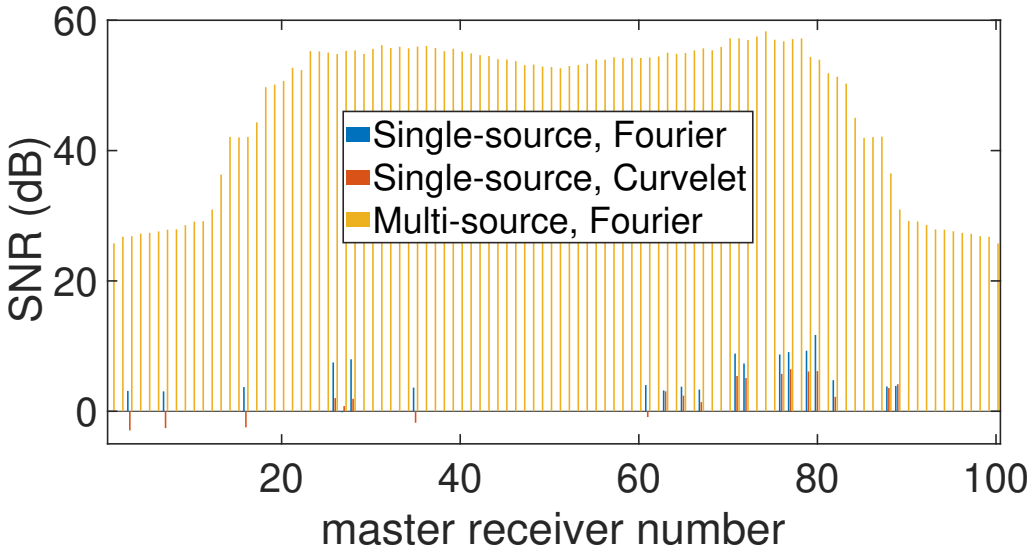

Figure 6. SNR in dB that is averaged over different master receivers for different reconstructions across the receiver array. Because the single-source reconstruction uses interferometric wavefields provided only from a VS, the single-source method cannot recover the wavefields when the master receivers are absent as shown in Figure 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.

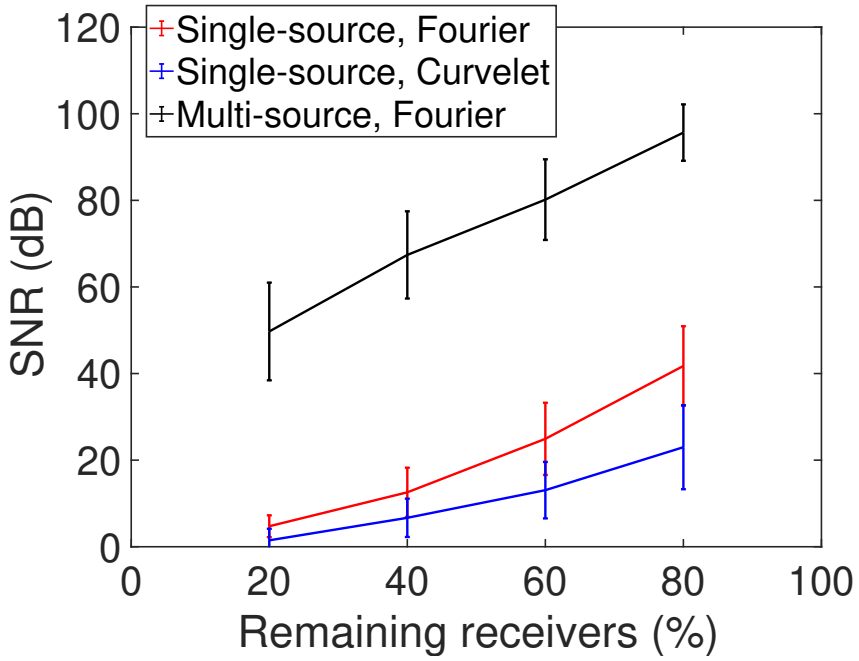

Figure 7. Mean and standard deviation of the SNR (dB) for different reconstructions using different fractions of remaining receivers and different missing receiver locations.

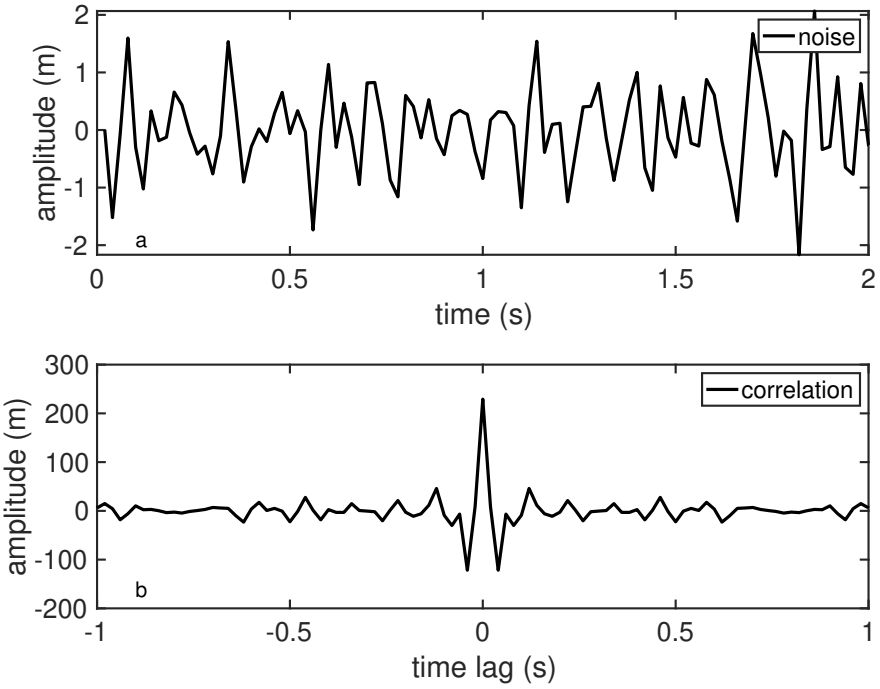

Figure 8. a) Noisy signal bandlimited between 5 and 20 Hz and b) auto-correlation of the noisy signal of a), representing seismic ambient noise and correlation of ambient noise, respectively.

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 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 (Dou et al., 2017; Zeng et al., 2017; Lellouch et al., 2019; Baird et al., 2020; Shragge et al., 2021). Seismic data recorded using DAS fibers are highly sensitive to the arrival angles of wavefields and the fiber orientation (Shragge et al., 2021; Martin 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.

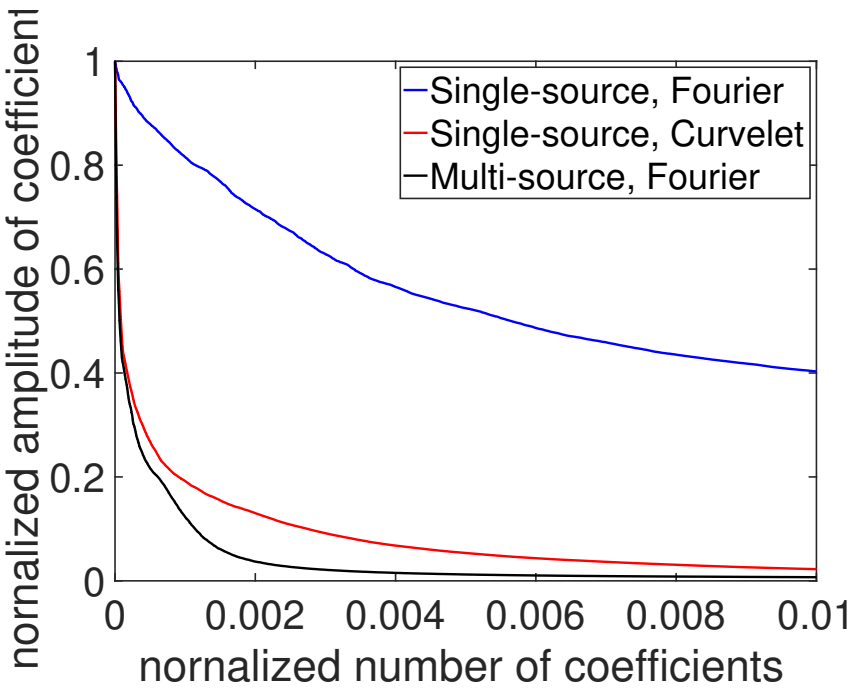

Figure 9. Decay of the Fourier and Curvelet coefficients. The graph shows the coefficient amplitudes normalized by the maximum amplitude of the coefficients across the number of the coefficients normalized by the maximum number of coefficients.

5 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.

6 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.

7 DATAAVAILABILITY

The numerical data and CS reconstruction used in this article are processed using steps and methods described in Section 2 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., 2006).

REFERENCES

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.

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.

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.

Candès, E., Demanet, L., Donoho, D., and Ying, L. (2006). 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.

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

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

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

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.

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.

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.

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.

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

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.

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.

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.

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.

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.

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.

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

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.

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).

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.

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., 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.

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

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.

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.

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.

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.

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

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

Snieder, R. and Wakin, M. B. (2022). When randomness helps in undersampling. *SIAM Review*, in press.

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.

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

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.

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

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

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.

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

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.

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.

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.

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

Table 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.

Appendix: Fourier coefficients associated with SI and CS

In Section 4, 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 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 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 8a), the window of the correlation of signals (Figure 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 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.