On beamforming of DAS ambient noise recorded in an urban environment and Rayleigh-to-Love waves ratio estimation

Yumin Zhao[1], Yunyue Elita Li[2,1], and Bei Li[1]

[1]Department of Civil and Environmental Engineering, National University of Singapore, Singapore
[2]Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, Indiana, U.S.A.

Key Points:

- We simulate geophone and DAS ambient noise data and analyze their differences and implications on beamforming results.
- The verified beamforming results confirm that strong urban traffic noise originated from the closest road intersection.
- We identify Rayleigh and Love waves from the DAS field data and resolve the Rayleigh-to-Love waves ratio.

Corresponding author: Y. E. Li, elitali@purdue.edu

Abstract

Ambient noise sources in urban environments are generally not perfectly uniformly distributed or travel along the axial direction of the receiver array. The horizontally deployed distributed acoustic sensing (DAS) fiber-optic cables record both Rayleigh and Love waves. Therefore, resolving the source propagation direction and the content of Rayleigh and Love waves from the DAS ambient noise data are essential for resolving a reliable near-surface shear-wave velocity model. However, a standard geophone-based beamforming procedure may not apply to DAS ambient noise data due to the complex measurements of DAS. We investigate the validity of beamforming results by analyzing the simulated seismic ambient noise recorded by a DAS array close to an L-shape under different source conditions. Synthetic examples show that correcting amplitude differences and polarity reversal of the DAS ambient noise data is crucial for a reliable beamforming analysis. In addition, we estimate the Rayleigh-to-Love waves (R/L) ratio from the DAS ambient noise data with the resolved source direction, the known array geometry, and the measured amplitude information. Finally, we apply the methods to the field data recorded by the Stanford DAS-1 array. The results suggest that the field DAS ambient noise was mainly generated from the local traffic and is dominated by Love waves.

Plain Language Summary

Human activities are the primary source that generates seismic ambient noise at higher frequencies (above 1 Hz). Resolving the main source propagation direction is essential to extract a reliable near-surface shear-wave velocity model from the seismic ambient noise recorded by the horizontal components of seismic sensors. Beamforming is a commonly used method to resolve the ambient noise source propagation directions. However, it is designed for seismic ambient noise data recorded by the vertical component of traditional seismic sensors. Due to the complex measurements of distributed acoustic sensing (DAS), beamforming may not apply to DAS ambient noise data. We investigate the validity of beamforming results resolved from DAS ambient noise data with several synthetic examples. We also improve the resolution of beamforming results with careful processing. Besides, the horizontally deployed DAS records Rayleigh and Love waves, both of which would be used to invert the shear-wave velocity model since they cannot be separated without enough information about the sources. We propose to estimate the Rayleigh-to-Love waves (R/L) ratio based on the resolved source direction, the known array geometry, and the measured amplitude information. Finally, we apply the method to the field DAS ambient noise data recorded by the Stanford DAS-1 array.

1 Introduction

Seismic ambient noise wavefield travels through Earth's layers and contains rich information about the subsurface at different scales. Compared with the costly vibroseis and environmentally harmful explosives, ambient noise sources are free and environmentally friendly. Therefore, seismic ambient noise becomes increasingly important for urban environment monitoring. The typical frequency of seismic ambient noise ranges from 0.001 Hz to 100 Hz (Nakata et al., 2019). It mainly results from natural vibrations (below 1 Hz) and anthropological activities (above 1 Hz) (Díaz et al., 2017). The single-station ambient noise data recorded in the urban environment appear random due to the random nature of the ambient noise sources. The assumptions in passive seismic interferometry are that ambient noise sources are uniformly distributed, and their generated wave modes are equipartitioned. Based on that, we can extract coherent signals from the seismic ambient noise recordings by cross-correlating the signals recorded by a specific sensor (virtual source) with those recorded by the array. The cross-correlation result is referred to as empirical Green's function, which approximates the response of each sensor to an impulse source at the virtual source (Claerbout, 1968; Wapenaar, 2004). Dis-

persion curve can be retrieved from the common virtual-shot gather, from which the near-surface shear-wave velocity model and soil properties can be inverted (e.g., Dou et al., 2017; Zhang et al., 2019; Spica et al., 2020).

However, ambient noise sources in urban environments are usually not perfectly uniformly distributed since the roads and traffic patterns are preferentially oriented according to geographic and urban development constraints. Besides, as the demand for real-time near-surface monitoring is rising, shorter periods of seismic ambient noise recordings are preferred to resolve the near-surface structural information with high temporal resolution. Although long-period (month- or year-long) seismic ambient noise recordings may show more randomness, such source averaging effect does not apply to short-period (minute-long) recordings. Seismic ambient noise recorded in the urban environment within a short period usually comes from a fixed location (e.g., Zhang et al., 2019; Nilot et al., 2019; Y. E. Li et al., 2020). The non-uniform distribution of anthropological noise sources leads to asymmetrical amplitude and travel times of the causal and anticausal parts of the extracted common virtual-shot gather (Stehly et al., 2006). The dispersion curves extracted from both parts of the common virtual-shot gather are apparent seismic wave phase velocities. Directional correction should be applied to the apparent phase velocities before inversion to obtain the true velocity information of seismic waves.

Distributed acoustic sensing (DAS) is a cutting-edge technology that enables continuous and long-term near-surface monitoring. The DAS system consists of an interrogator unit and a fiber-optic cable. The length of the fiber-optic cable can be tens of kilometers (Ajo-Franklin et al., 2019). Compared to conventional seismic sensors, DAS has a much higher spatial resolution and lower per-channel cost. It has been successfully applied to near-surface characterization in urban environments, where fiber-optic cables are usually horizontally and nonlinearly deployed (e.g., Dou et al., 2017; Fang et al., 2020; Spica et al., 2020). Dispersion curves or shear-wave velocity models were extracted from minute-long (Zhang et al., 2019; Nilot et al., 2019), hour-long (Zeng et al., 2017; Luo et al., 2020; Tribaldos et al., 2021), day-long (Dou et al., 2017; Martin et al., 2017; Ajo-Franklin et al., 2019), and month-long seismic ambient noise data (Martin & Biondi, 2018; Martin, 2018; Spica et al., 2020). The common virtual-shot gathers extracted from most of these studies are asymmetric, indicating insufficient source averaging and preferential source directionality. However, very few of these studies applied the directional correction that is essential for a reliable near-surface shear-wave velocity inversion, especially when dealing with DAS ambient noise.

The propagation direction and velocity of the seismic ambient noise wavefield recorded by the vertical component of the geophone are commonly resolved by beamforming(e.g., Brooks et al., 2009; Zhang et al., 2019). However, it is rarely applied to DAS ambient noise data due to the complexity of DAS measurements. One problem that hinders resolving a reliable source direction from the DAS ambient noise data is the amplitude differences between the data recorded by two nonparallel legs of the horizontally deployed DAS array (Lindsey et al., 2017). These differences are caused by the directionality of fiber-optic cables and geometry effect (Mateeva et al., 2012). Besides, polarity reversal may appear at the corner of the array when DAS records Love waves (e.g., Benioff, 1935; Lindsey et al., 2017; Yu et al., 2019). These two features of DAS ambient noise make beamforming analysis a challenging task. Lindsey et al. (2017) resolved the direction of propagation of the teleseismic signals recorded by an L-shape DAS array. Fang et al. (2020) detected the propagation direction of the quarry blast signals recorded by the Stanford DAS-1 array. The DAS ambient noise data in these two cases show coherent events instead of randomness since each was generated from a unique known far-field source from a fixed direction. However, it is more challenging when the source is unknown, moving, and close (within a few hundred meters) to the array.

The horizontal components of the receiver array record both Rayleigh and Love waves. As the physics of Rayleigh and Love waves are different, prior knowledge of the contents of each surface wave type in the common virtual-shot gather is essential for a reliable inversion of the near-surface velocity model. Rayleigh and Love waves recorded by a geophone can be separated by rotating the two horizontal components in the radial/transverse direction. However, those recorded by DAS cannot be separated if there is insufficient information about the sources. Previous studies show that when the virtual source is colinear with the DAS array, the extracted common virtual-shot gather mainly contains Rayleigh waves (Snieder, 2004; Martin & Biondi, 2017; Martin et al., 2017; Luo et al., 2020). The theory assumes that ambient noise sources are random and uniformly distributed, and the recording time is long (months or even years). However, the assumptions are unsuitable for real-time monitoring where a short period of data is utilized.

This paper aims to resolve the ambient noise source direction and estimate the content of Rayleigh and Love waves (i.e., Rayleigh-to-Love waves (R/L) ratio) from DAS ambient noise data recorded in an urban environment. Following the acquisition setup of the Stanford DAS-1 array, we simulate DAS ambient noise data under different source conditions with varying combinations of Rayleigh and Love waves. We normalize the amplitude of the DAS ambient noise data recorded by the two legs and extract the common virtual-shot gather from it. By observing the polarity of the common virtual-shot gather at the corner of the DAS array, we determine whether there is polarity reversal and apply a polarity flip on the common virtual-shot gather when polarity reversal occurs. After that, we apply multiple signal classification (MUSIC) beamforming on the extracted common virtual-shot gathers. We demonstrate that beamforming results obtained from the DAS ambient noise data after normalization and polarity correction are reliable when the source direction is unique. Based on the resolved source propagation direction, we calculate the R/L ratio from the noise amplitude measured by the fiber-optic cable deployed in two different directions. Additionally, when ambient noise arrives at the array from multiple directions, beamforming cannot provide universally reliable information for all scenarios, thus making the apparent velocity correction and R/L ratio inversion challenging. We apply the method to two 2-minute DAS ambient noise data recorded by Stanford DAS-1 array. The results suggest that the recorded seismic ambient noise is mainly generated from local traffic and dominated by Love waves.

2 Methodology

2.1 Urban seismic ambient noise simulation

We assume a receiver array is deployed close to the L-shape, and it records seismic ambient noise generated by ambient noise sources at fixed locations outside the array. To simulate DAS ambient noise, we first simulate the seismic ambient noise recorded by one leg of the three-component geophone array as displacement, and then we derive the DAS ambient noise as strain. We assume the recorded ambient noise is the superposition of plane surface waves, which are generated by multiple scattering between the source and receivers. The seismic ambient noise recorded by the selected leg of the vertical component of the geophone array is as follows:

$$d(x,\omega) = A(\omega)\sum_{i=1}^{N_s} a_{Ri} e^{-i\omega(t_i + x\cos\theta_i / v(\omega))}, \tag{1}$$

where x denotes the distance between the current receiver and the one at the corner of the L-shaped array, and ω is the angular frequency. $A(\omega)$ is the amplitude spectrum of the source wavelet, and N_s is the number of plane surface waves generated by the ambient noise sources during the recording period. a_{Ri} is the pulse amplitudes of Rayleigh waves generated by the i^{th} source, $i = 1, 2, ..., N_s$. t_i denotes the time at which the i^{th} plane surface wave is generated. a_{Ri} and t_i follow uniform distribution. θ_i is the intersection angle between the direction of propagation of the i^{th} plane surface wave and the

selected leg of the array (Figure 1). $v(\omega)$ is the surface wave phase velocity at angular frequency ω.

The ambient noise recorded by the horizontal components contains both Rayleigh and Love waves, where the Rayleigh waves components are:

$$d_R(x,\omega) = A(\omega)\sum_{i=1}^{N_s} a_{Ri}\cos\theta_i e^{-i\omega(t_i + x\cos\theta_i/v_R(\omega))}. \tag{2}$$

And the Love waves components can be expressed as:

$$d_L(x,\omega) = A(\omega)\sum_{i=1}^{N_s} a_{Li}\sin\theta_i e^{-i\omega(t_i + x\cos\theta_i/v_L(\omega))}, \tag{3}$$

where a_{Li} is the pulse amplitudes of Love waves generated by the i^{th} source, $i = 1, 2, ..., N_s$.

To approximate the DAS measurements, we first compute point-wise strain by taking the spatial derivative of the horizontal displacements of the Rayleigh and Love waves components:

$$s_R(x,\omega) = A(\omega)\sum_{i=1}^{N_s} -\frac{i\omega\cos^2\theta_i}{v(\omega)} a_{Ri} e^{-i\omega(t_i + x\cos\theta_i/v_R(\omega))}, \tag{4}$$

$$s_L(x,\omega) = A(\omega)\sum_{i=1}^{N_s} -\frac{i\omega\cos\theta_i\sin\theta_i}{v(\omega)} a_{Li} e^{-i\omega(t_i + x\cos\theta_i/v_L(\omega))}. \tag{5}$$

To include the gauge length effect of the DAS interrogator (Dean et al., 2017), we approximate the DAS measurements as the average of the point-wise strain over the gauge length L_g. Then DAS Rayleigh waves measurements can be expressed as:

$$\bar{s}_R(x,\omega) = \frac{1}{L_g}\int_{x-L_g/2}^{x+L_g/2} s_R(l,\omega)dl, \tag{6}$$

and DAS Love waves measurements can be expressed as:

$$\bar{s}_L(x,\omega) = \frac{1}{L_g}\int_{x-L_g/2}^{x+L_g/2} s_L(l,\omega)dl. \tag{7}$$

Similarly, DAS measurements of the other leg of the array can be derived. By changing angle θ_i, we can obtain the final DAS measurements of the whole close L-shaped receiver array. When the fiber-optic cable takes a 90° corner, Rayleigh waves strain can change significantly in their amplitudes without ever changing their polarity. In contrast, Love waves strain always flips its polarity without amplitude changes. These conclusions are consistent with those obtained in the literature (e.g., Benioff, 1935; Lindsey et al., 2017; Yu et al., 2019). Such measurement differences in Rayleigh and Love waves are exploited in our methodology to extract their relative amplitudes.

2.2 R/L ratio estimation

The horizontally deployed DAS array records both Rayleigh and Love waves. As we do not have sufficient knowledge about the sources, they cannot be separated. Thus, we estimate the R/L ratio in the data based on the resolved source propagation direction, array geometry, and amplitude of the seismic ambient noise data. When all ambient noise sources travel from one direction, we can rewrite equations 6 and 7 as:

$$\bar{s}_R(x,\omega) = -\frac{i\omega\cos^2\theta}{v_R(\omega)} A(\omega)R(x,\omega), \tag{8}$$

$$\bar{s}_L(x,\omega) = -\frac{i\omega\cos\theta\sin\theta}{v_L(\omega)} A(\omega)L(x,\omega), \tag{9}$$

where

$$R(x,\omega) = \int_{x-L_g/2}^{x+L_g/2} \sum_{i=1}^{N_s} a_{Ri} e^{-i\omega(t_i + y\cos\theta/v_R(\omega))} dy, \tag{10}$$

$$L(x,\omega) = \int_{x-L_g/2}^{x+L_g/2} \sum_{i=1}^{N_s} a_{Li} e^{-i\omega(t_i + y\cos\theta/v_L(\omega))} dy. \tag{11}$$

Then, the DAS measurements at frequency ω recorded at locations x_1 and x_2 (Figure 1) are:

$$D_1(x_1,\omega) = \bar{s}_R(x_1,\omega) + \bar{s}_L(x_1,\omega), \tag{12}$$

and

$$D_2(x_2,\omega) = \bar{s}_R(x_2,\omega) + \bar{s}_L(x_2,\omega). \tag{13}$$

Then, we can obtain the power spectrum ratio of $D_1(x_1,\omega)$ to $D_2(x_2,\omega)$:

$$\frac{|D_1|^2}{|D_2|^2} = \frac{D_1(x_1,\omega)D_1^*(x_1,\omega)}{D_2(x_2,\omega)D_2^*(x_2,\omega)} = \frac{(\bar{s}_R(x_1,\omega)+\bar{s}_L(x_1,\omega))(\bar{s}_R(x_1,\omega)+\bar{s}_L(x_1,\omega))^*}{(\bar{s}_R(x_2,\omega)+\bar{s}_L(x_2,\omega))(\bar{s}_R(x_2,\omega)+\bar{s}_L(x_2,\omega))^*}. \tag{14}$$

Assuming $\frac{v_L(\omega)}{v_R(\omega)} = \alpha(\omega)$, $\frac{|R|}{|L|} = m$, the power spectral ratio of $D_1(x_1,\omega)$ to $D_2(x_2,\omega)$ can be obtained according to equations 8-11 and 14:

$$\frac{|D_1|^2}{|D_2|^2} = \frac{a^2m^2\alpha(\omega)^2|L|^2 + b^2|L|^2 + 2abm\alpha(\omega)(R_rL_r + R_iL_i)}{c^2m^2\alpha(\omega)^2|L|^2 + d^2|L|^2 + 2cdm\alpha(\omega)(R_rL_r + R_iL_i)}, \tag{15}$$

where $a = -\cos^2\theta_1$, $b = -\cos\theta_1\sin\theta_1$, $c = -\cos^2\theta_2$, and $d = -\cos\theta_2\sin\theta_2$. θ_1 and θ_2 denote the intersection angle between the ambient noise source propagation direction and two legs of the DAS array, respectively. R_r and L_r denote the real parts of the Rayleigh and Love waves, respectively. R_i and L_i are the imaginary parts of the Rayleigh and Love waves, respectively.

If the phase of Rayleigh and Love waves are the same, $R_r/L_r = R_i/L_i$ holds. Then we have

$$R_rL_r + R_iL_i = m|L|^2, \tag{16}$$

and equation 15 can be simplified as

$$\frac{|D_1|^2}{|D_2|^2} = \frac{a^2m^2\alpha(\omega)^2 + b^2 + 2abm\alpha(\omega)}{c^2m^2\alpha(\omega)^2 + d^2 + 2cdm\alpha(\omega)}. \tag{17}$$

However, as their phase velocities are different, they do not share the same phase. Therefore, we approximate equation 15 as:

$$\frac{|D_1|^2}{|D_2|^2} = \frac{a^2m^2\alpha(\omega)^2 + b^2 + 2abm\alpha(\omega)\cos(\omega x_1 e_1 \delta_v(\omega))}{c^2m^2\alpha(\omega)^2 + d^2 + 2cdm\alpha(\omega)\cos(\omega x_2 e_2 \delta_v(\omega))}, \tag{18}$$

where $e_1 = \cos\theta_1$, $e_2 = \cos\theta_2$, and $\delta_v(\omega) = \frac{1}{v_R(\omega)} - \frac{1}{v_L(\omega)}$.

We only consider the fundamental mode surface waves in the DAS ambient noise data and assume $v_R(\omega)$ and $v_L(\omega)$ are constants at higher frequencies. We use the DAS ambient noise data at higher frequency bands to estimate the R/L ratio. The range of $v_R(\omega)$ at higher frequencies can be obtained from the dispersion curve extracted from the seismic ambient noise. Besides, $v_L(\omega)$ are normally larger than $v_R(\omega)$, and $\alpha(\omega) = \frac{v_L(\omega)}{v_R(\omega)}$ sits in the range $[1.0, 1.3]$ (Behm & Snieder, 2013). With the grid search method, we can solve the R/L ratio m.

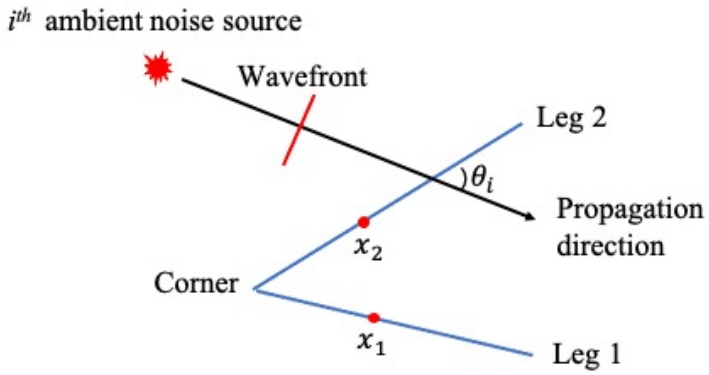

Figure 1. Simple illustration of the i^{th} plane wave generated by an ambient noise source and recorded by a 2D receiver array (two blue lines). Take Leg 2 as an example, θ_i is the intersection angle between the plane surface wave propagation direction and Leg 2. The two red dots represent the locations of two channels, one on each leg, which are x_1 and x_2, respectively. Leg 1 is connected to the DAS interrogator.

3 Synthetic data examples

We use the Multiple signal classification (MUSIC) beamforming technique (Goldstein & Archuleta, 1987; Kirlin, 1992; Godara, 1997; Gelius et al., 2013; Zhang et al., 2019) to estimate the ambient noise source propagation direction. The beamforming spectrum can be calculated from the seismic data or the corresponding common virtual-shot gather. This section presents the beamforming results resolved from common virtual-shot gathers extracted from the synthetic DAS ambient noise recordings with different R/L ratios when the ambient noise sources come from one direction, two directions (Figures S1-S2 in the supporting information), and multiple uniformly distributed directions (Figure S3 in the supporting information). For each scenario, We verify the reliability and identify the challenges of beamforming when applied to DAS measurements. We demonstrate the importance of identifying polarity reversal and compensating for the amplitude projection, although these are only effective when the sources mainly come from one direction.

The receiver array is designed according to the southwest corner of the Stanford DAS-1 array. The two legs of the array are about 85° from each other (Figure 2a). We use the published 1-D shear-wave velocity model (Spica et al., 2020) to calculate the phase velocities (Figure 2b). The velocity model was inverted from month-long DAS ambient noise data recorded by the same DAS array (Spica et al., 2020). The phase velocities are calculated using the Thomson-Haskell method (Thomson, 1950; Haskell, 1953; Buchen & Ben-Hador, 1996) with the open-source software *evodcinv* (Luu, 2019). Overall, the phase velocities of Love waves (red curve) are slightly higher than that of Rayleigh waves (blue curve). We use Rayleigh waves phase velocities (blue curve) to simulate both the displacement recorded by the vertical component of geophone and DAS Rayleigh waves measurements. We use Love waves phase velocities (red curve) to simulate DAS Love waves measurements.

The time interval of the synthetic seismic ambient noise data is 2 ms, the same as the field data to be discussed. During the 2-minute recording period, we assume 100 plane surface waves propagate from designated directions for each data. The gauge length is set as 7 meters based on the field settings (Martin et al., 2017). We define the direction of propagation towards the East as 0°, and the angle increases to 90° from East to North.

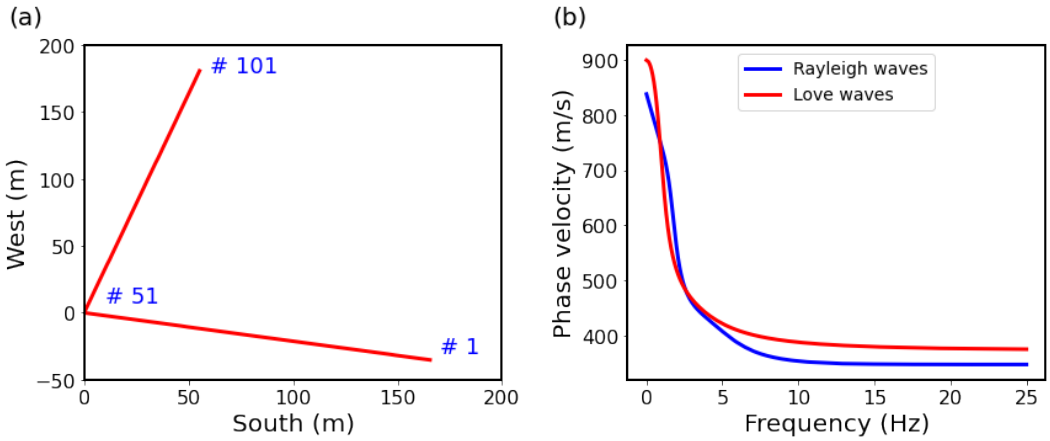

Figure 2. (a) Layout of the receiver array for the synthetic seismic ambient noise data examples, (b) phase velocities of Rayleigh (blue curve) and Love waves (red curve) calculated from the published 1-D shear-wave velocity model (Spica et al., 2020). They are used to simulate the seismic ambient noise data.

The data recorded at channel 55 is selected as the virtual source for the common virtual-shot gather calculation. We apply a 2 Hz to 10 Hz bandpass filter to all synthetic data.

3.1 Single surface wave type with one source direction

When all the plane waves generated by the motor vehicles propagate from one direction (110°), we simulate the vertical displacement (Figure 3a), horizontal DAS Rayleigh waves measurement (Figure 3d), and horizontal DAS Love waves measurement (Figure 3g), respectively. The DAS Rayleigh waves measurement (Figure 3d) shows the amplitude differences between the data recorded by different legs of the DAS array, which is caused by the unequal projection of Rayleigh waves. There is almost no amplitude difference in the DAS Love waves measurement (Figure 3g), which is consistent with the conclusions we have obtained from the equations in the methodology section.

Figures 3b, 3e, and 3h show the corresponding common virtual-shot gather of Figures 3a, 3d, and 3g, respectively. We can also observe the amplitude differences on the common virtual-shot gather obtained from DAS Rayleigh waves measurement (Figure 3e). It is difficult to determine if there is a polarity reversal on the raw ambient noise recordings. However, we can see from the common virtual-shot gathers, which maintain the polarity information, that the polarity is consistent at the corner for DAS Rayleigh waves measurement while there is a polarity reversal for DAS Love waves measurement. Therefore, we apply a polarity flip to the DAS Love waves measurement recorded by one leg of the array before beamforming. Figures 3c, 3f, and 3i display the corresponding beamforming results resolved from Figures 3a, 3d, and 3g, respectively. The estimated ambient noise source directions match well with the reference directions for all three beamforming spectra. However, the resolution of the beamforming result resolved from the DAS Rayleigh waves measurement is low (Figure 3f). This is due to the unequal amplitude projection of Rayleigh waves onto the two legs of the DAS array.

When the plane waves come from the direction of 185°, the source propagation direction is almost parallel to one leg of the array while almost perpendicular to the other. Thus, the amplitude projection is severe, and the amplitude differences are obvious in the DAS Rayleigh waves measurement (Figures 4e and 4f). As a result, the corresponding beamforming spectrum (Figure 4g) is blurred. To solve this problem, we apply am-

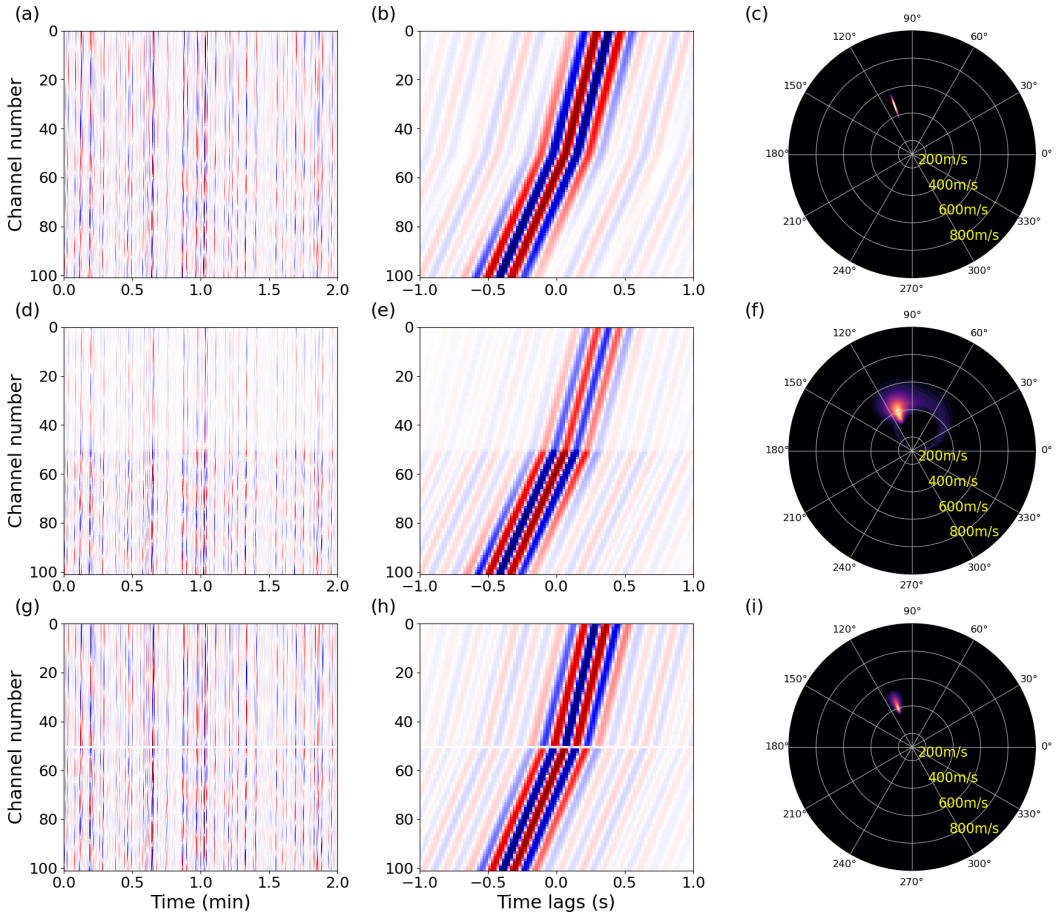

Figure 3. When the propagation direction of the ambient noise sources is 110°, (a) synthetic seismic ambient noise data recorded by the vertical component of the geophone array, (b) its corresponding common virtual-shot gather, (c) beamforming spectra calculated from (a). Note that if there is a polarity reversal in the data, a polarity flip will be applied to it before calculating the beamforming spectra. (d), (e), and (f) are the same as (a), (b), and (c), respectively, except that (d) is DAS Rayleigh waves measurement. (g), (h), and (i) are the same as (a), (b), and (c), respectively, except that (g) is DAS Love waves measurement.

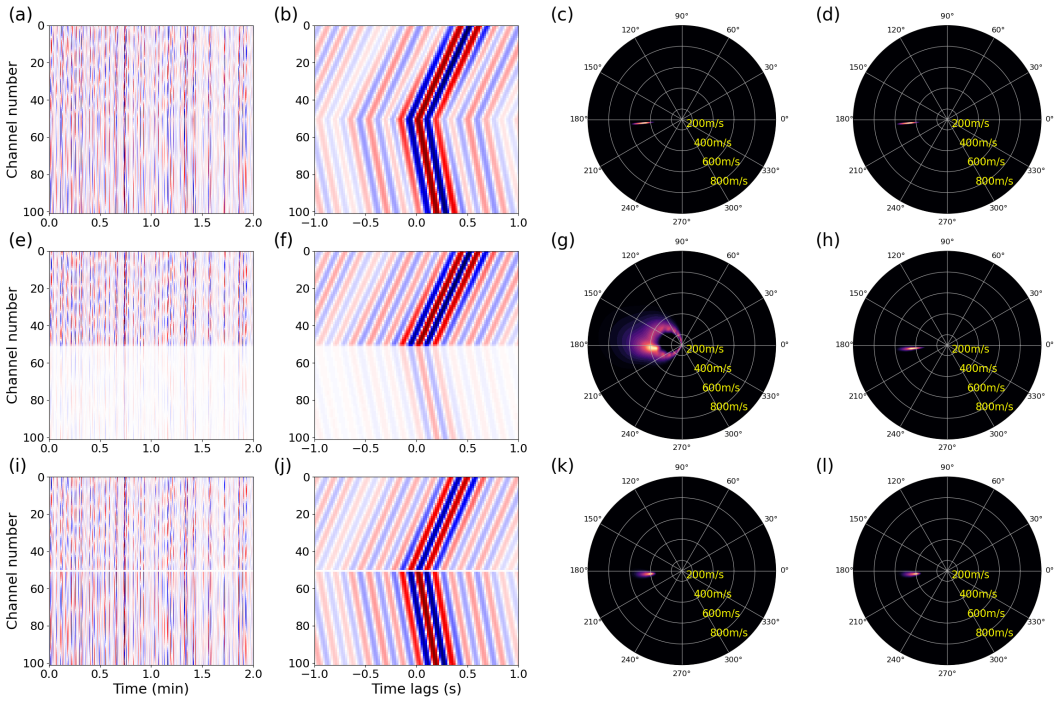

Figure 4. When the propagation direction of the ambient noise sources is 185°, (a) synthetic seismic ambient noise data recorded by the vertical component of the geophone array, (b) the corresponding common virtual-shot gather of (a), (c) beamforming spectra calculated from (b), and (d) beamforming spectra calculated from the common virtual-shot gather obtained from the normalized version of (a). (e), (f), (g), and (h) are the same as (a), (b), (c), and (d), respectively, except that the (e) is DAS Rayleigh waves measurement. (i), (j), (k), and (l) are the same as (a), (b), (c), and (d), respectively, except that (i) is DAS Love waves measurement.

plitude normalization to the DAS Rayleigh waves measurement using the root-mean-square ratio of the data recorded by the two legs of the array. Beamforming spectra obtained from the common virtual-shot gather calculated from the normalized data are shown in Figures 4d, 4h, and 4l, respectively. These results demonstrate that beamforming resolves the main source direction from each data. In addition, Figures 4k and 4i show that amplitude normalization is not critical for high-resolution beamforming of DAS Love waves measurement compared to DAS Rayleigh waves measurement. However, when the ambient noise sources come from multiple directions, beamforming may not resolve all the directions (Figures S1-S3 in the supporting information) due to the limitation of beamforming resolution (Johnson & Dudgeon, 1993) and the complexity of the data.

3.2 Two surface wave types with one source direction

Field DAS ambient noise recordings are often more complex when Rayleigh and Love waves mix up. When all plane waves come from one direction (185°), we test three cases where Rayleigh and Love waves occupy varying proportions in the ambient noise data: 75% Rayleigh waves and 25% Love waves, 50% Rayleigh waves and 50% Love waves, and 25% Rayleigh waves and 75% Love waves. Figures 5a, 5e, and 5 display the synthetic DAS ambient noise data for the three cases, respectively. We observe significant amplitude differences in the two legs of the seismic ambient noise data. The common virtual-shot gathers (Figures 5b, 5f, and 5j) show no polarity reversal when there are more Rayleigh

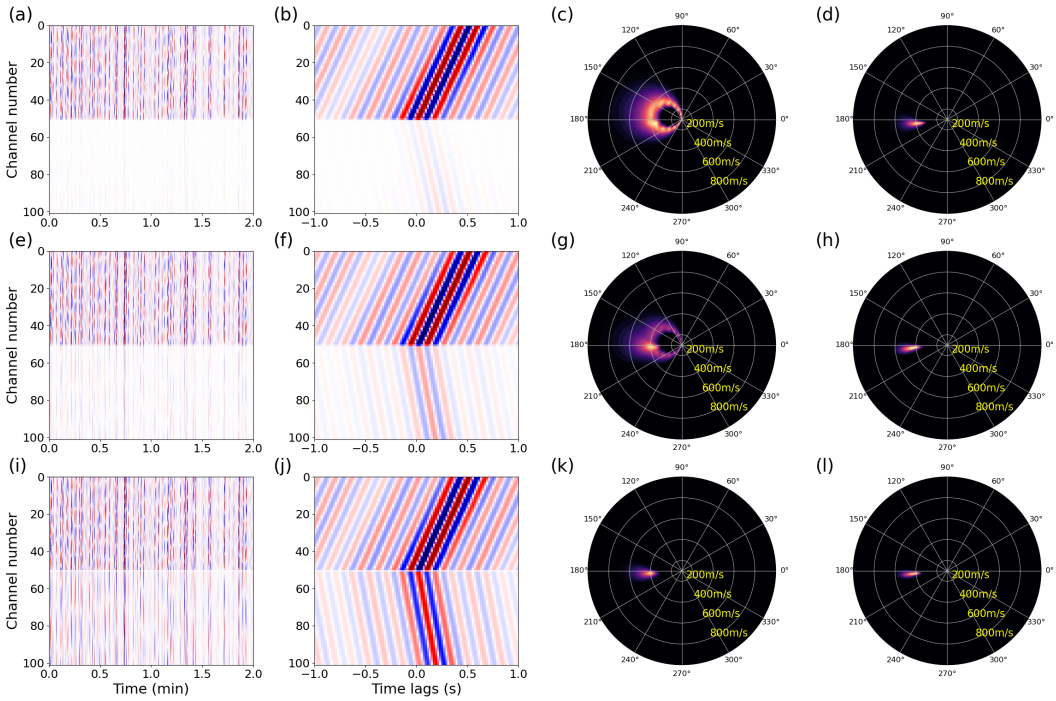

Figure 5. Same as Figure 4 except that the data are synthetic seismic ambient noise data recorded by the horizontally deployed DAS array, which contains (a) 75% Rayleigh waves and 25% Love waves, (e) 50% Rayleigh waves and 50% Love waves, and (i) 25% Rayleigh waves and 75% Love waves.

waves than Love waves and vice versa. The resolution of the beamforming spectra (Figures 5c and 5g) calculated from Figures 5b and 5f is very low. As a result, it is difficult to identify the source propagation direction from them. On the other hand, the high-resolution results displayed in Figures 5d, 5h, and 5i suggest that the beamforming spectra are reliable for all three cases where beamforming is applied to the data with amplitude normalization and polarity flip (only if there is polarity reversal).

Tables 1 and 2 summarize the beamforming results obtained from DAS ambient noise with varying portions of Rayleigh and Love waves when the ambient noise source propagation directions are 110° and 185°, respectively. Table 1 suggests that when the amplitude differences are slight, amplitude normalization and polarity correction are not critical for beamforming. Table 2 shows that the accuracy of the beamforming results increases with the decrease of the R/L ratio of the data. This is because the amplitude differences between the DAS ambient noise data recorded by the two legs of the array decrease with the decrease in the R/L ratio. It also suggests that when the amplitude differences are significant in the DAS ambient noise data, amplitude normalization and polarity correction could significantly improve the accuracy of the beamforming results. Thus, careful preprocessing of the data is essential. Although the resolved source propagation direction may not be the same as the reference, the errors are negligible. The results can be used for further studies such as apparent velocity correction and R/L ratio estimation.

Table 1. Beaforming results (reference source propagation direction: 110°)

R:L	10:1	5:1	2:1	1:1	1:2	1:5	1:10
Polarity reversal	No	No	Yes	Yes	Yes	Yes	Yes
Source direction (raw)	110°	111°	113°	108°	109°	110°	110°
Source direction (normalized)	110°	111°	111°	108°	109°	109°	110°

Table 2. Beaforming results (reference source propagation direction: 185°)

R:L	10:1	5:1	2:1	1:1	1:2	1:5	1:10
Polarity reversal	No	No	Yes	Yes	Yes	Yes	Yes
Source direction (raw)	–	–	–	–	184°	184°	185°
Source direction (normalized)	185°	186°	181°	184°	184°	184°	185°

3.3 R/L ratio estimation

We simulate DAS ambient noise data with different R/L ratios to investigate the reliability of the proposed method. We assume there are 100 sources, the source propagation direction is 110°, each source generates Rayleigh and Love waves simultaneously at random times, and the recording period is 2 minutes. We assume $a_L \sim U(-1, 1)$, which means the pulse amplitude of Love waves generated by each source is an independent sample from the uniform distribution. For the same source, the generated pulse amplitude of Rayleigh waves is $a_R = (1+\epsilon)ma_L$, where $\epsilon \sim U(-0.2, 0.2)$. These assumptions result in an averaged R/L ratio of m across all sources.

We first extract the dispersion curve from each data, from which we can obtain the range of Rayleigh waves phase velocities at higher frequencies (8-10 Hz). We assume the ratio between the Love and Rayleigh waves phase velocity is $\alpha \sim U[1.0, 1.3]$ (Behm & Snieder, 2013). Then, we estimate the R/L ratio based on the range of Rayleigh waves phase velocities, the ratio between the phase velocity of Love and Rayleigh waves α, and the amplitude ratio of the data recorded by the two legs with the grid search method. m can be constrained with the polarity information at the corner of the array and the amplitude ratio between the two legs of the array. For example, if there is polarity reversal at the corner and the amplitude ratio is close to 1, then the data must mainly contain Love waves and m should be a number close to 0.

Figure 6 shows the relationships between the true R/L ratios and those estimated from the synthetic DAS ambient noise data. Evidently, the estimated ratios are consistent with the references, demonstrating the applicability of the proposed inversion method for a wide range of Rayleigh and Love waves combinations.

4 Field data results

The layout map of the Stanford DAS-1 array (e.g., Biondi et al., 2017; Martin et al., 2017; Yuan et al., 2017; Fang et al., 2020) is displayed in Figure 7. We use two pieces of 2-min DAS ambient noise data to calculate the beamforming spectrum. These data were recorded by the DAS array leg marked by the yellow dashed line (Figure 7) and were less contaminated by the direct impact of road traffic. We filter out the low-frequency quarry blast signals (Fang et al., 2020) in the data with a 2-10 Hz bandpass filter and then remove the near-field traffic noise with the processing procedure published in Zhao et al. (2019). Figures 8a and 8b show two 2-minute preprocessed DAS data recorded in the evening on different days. Figure 8c and 8d are the corresponding common virtual-

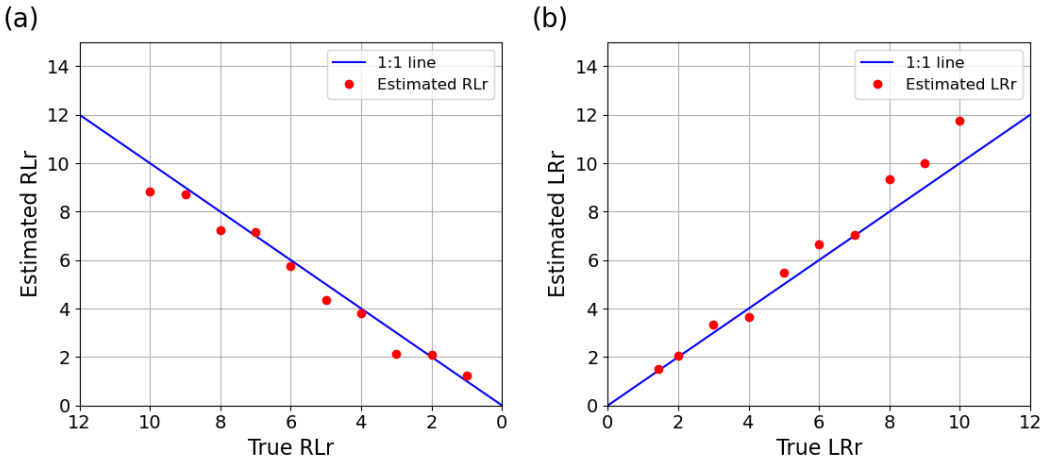

Figure 6. Comparison between the estimated R/L ratios and the references when the designated (a) R/L ratios are larger than or equal to 1, and (b) Love-to-Rayleigh waves (L/R) ratios are larger than 1.

shot gathers calculated from Figures 8a and 8b, respectively. The virtual source is at channel 93, about 15 meters to the corner. From the common virtual-shot gathers (Figures 8c and 8d), we observe clear polarity reversal around the corner of the DAS array. Thus, the data must contain Love waves. In addition, both the data (Figures 8a and 8b) and the common virtual-shot gathers (Figures 8c and 8d) show two types of amplitude differences. For each leg, we observe amplitude differences over different traces due to variations in fiber coupling. For example, the weaker amplitudes between Channel 70 and 90 are likely due to the presence of two manholes in this area (Martin, 2018). These channels are excluded from the R/L ratio estimation. Comparing overall amplitudes between the data recorded by the two legs, we observe systematic amplitude differences that suggest an uneven projection of Rayleigh waves energy to these two different angles.

Figures 9a and 9b show the beamforming results resolved from the two DAS data (Figures 8a and 8b), respectively. The energy of the beamforming spectra is concentrated. The resolved source propagation directions from the two data are 111° and 112°, respectively. The two directions are consistent, suggesting that the dominant traffic noise during the evening rush periods came from the NW direction. This is further validated by the travel time observations on the common virtual-shot gathers.

With the estimated source directions and the known array geometry, we estimate the R/L ratios from the amplitude measurements at higher frequencies (8-10 Hz) bounded by the red boxes in Figures 8a and 8b. The estimated R/L ratios for these two datasets are 0.68 and 0.47, respectively. These results suggest that the DAS array recorded more Love waves than Rayleigh waves propagating from the Campus Drive intersection in the northwest of the array.

Based on the estimated source propagation direction and R/L ratio, we simulate the corresponding DAS ambient noise data (Figures 10a and 10b) and extract the common virtual-shot gathers (Figures 10c and 10d) to verify their agreement with the two sets of field data. We observe the polarity reversal and amplitude differences between the data recorded by the two legs of the DAS array, which is consistent with those obtained from the field data. However, the time lags in Figures 10c and 10d deviate from those obtained from the field data. This is because the shear-wave velocity model (Spica

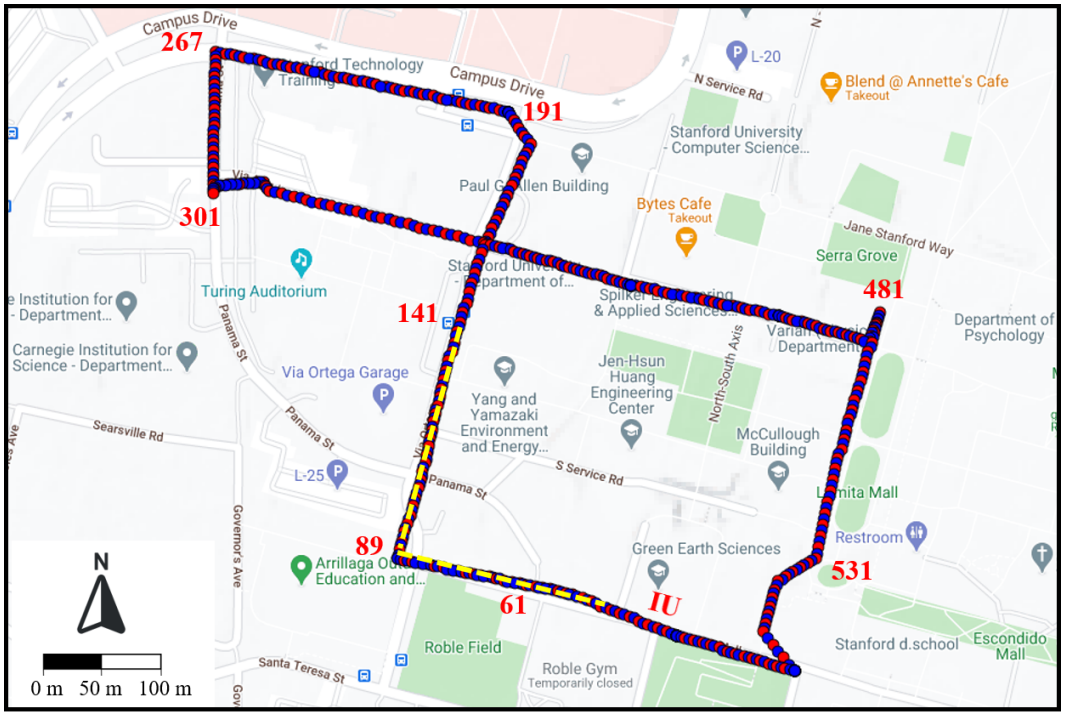

Figure 7. (a) Layout of the Stanford DAS-1 array. The red and blue points represent two loops of the fiber-optic cable (Martin et al., 2017). The data recorded by the fiber-optic cable marked by the yellow dashed line are used to estimate the source propagation direction.

et al., 2020) we utilized to simulate the dispersion curves may not precisely match the true one.

Figure 11a shows the smoothed average amplitude spectra of the data recorded on 3 Jan 2017, labeled by the red box in the upper panel (red curve) and lower panel (blue curve) in Figure 8a, respectively. Figure 11c shows the amplitude ratio between the two curves. Similarly, Figures 11b and 11d show the spectra and their ratio of the DAS data recorded on 12 April 2017. Figures 11c and 11d suggest that the average amplitude ratio between the two legs stays roughly constant at around 0.5 across the chosen frequency range. The amplitude ratios calculated from the two synthetic data within the same bandwidth are shown in Figure 12. The overall agreement between the field amplitude ratios (Figures 11c and 11d) and the synthetic amplitude ratios (Figures 12c and 12d) demonstrates the convergence of our inversion method.

5 Discussions

Ambient noise sources in urban environments are often generated from fixed anthropological activities instead of uniformly distributed. Resolving the ambient noise source propagation direction is essential to mitigate the overestimation of phase velocities of surface waves. When the ambient noise wavefield propagates from one direction, amplitude normalization and polarity flip (only when there is polarity reversal) improve the beamforming spectrum resolution significantly, especially when the amplitude differences are significant. However, when there are two directions of propagation, the resolution of the beamforming results is also influenced by the angle between the two directions and their relative angles to the array (Figures S1-S2 in the supporting information). When ambient noise sources are uniformly distributed, beamforming with an L-shape DAS array

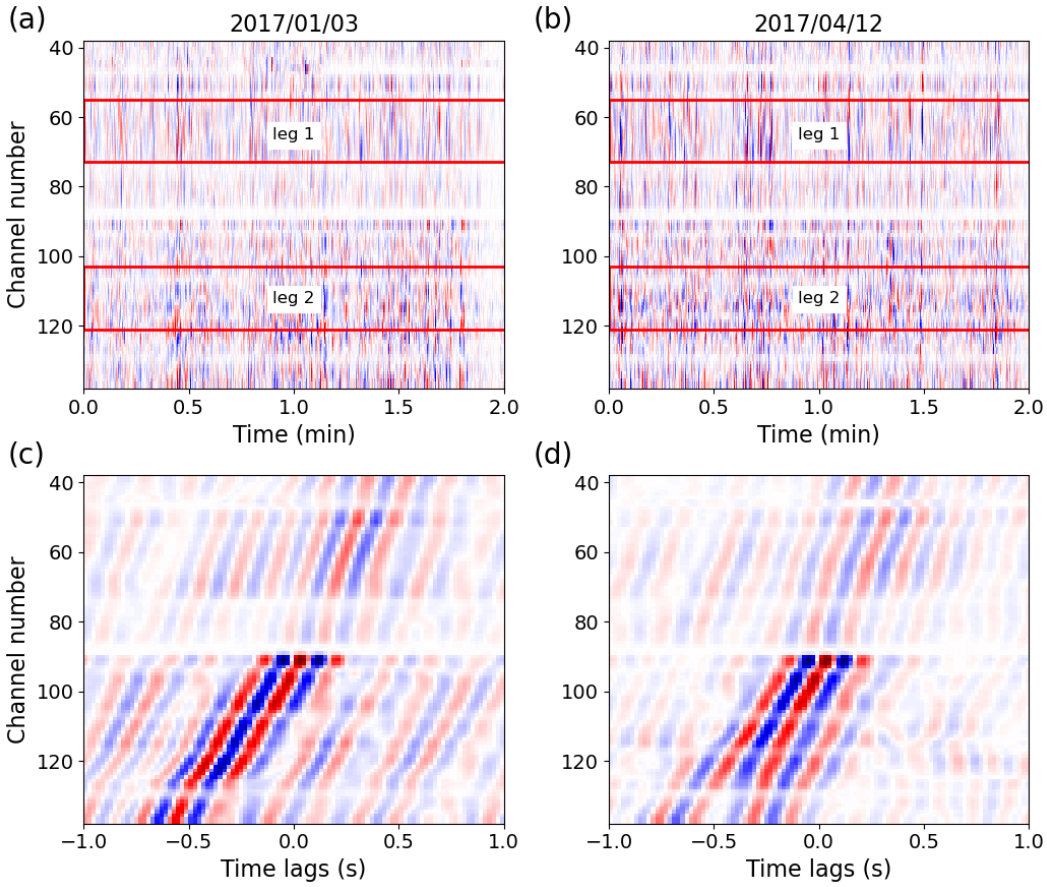

Figure 8. 2-min preprocessed DAS ambient noise data recorded by the Stanford DAS-1 array on (a) 3 Jan 2017 and (b) 12 April 2017. (c) and (d) are the common virtual-shot gathers extracted from (a) and (b), respectively. We use the data recorded at channel 93 as the virtual source. R/L ratios are estimated with the data labeled by the red boxes.

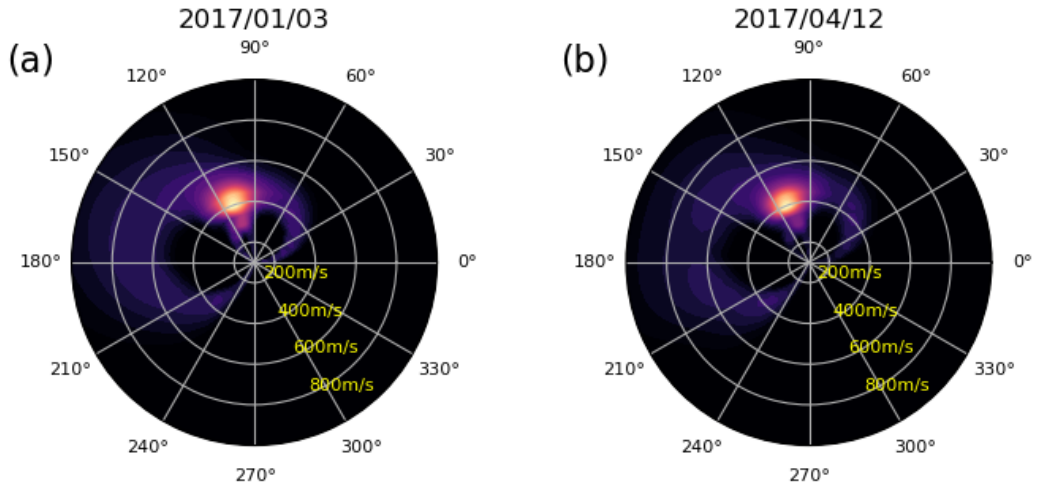

Figure 9. Beamforming spectra resolved from (a) Figure 8a, (b) Figure 8b.

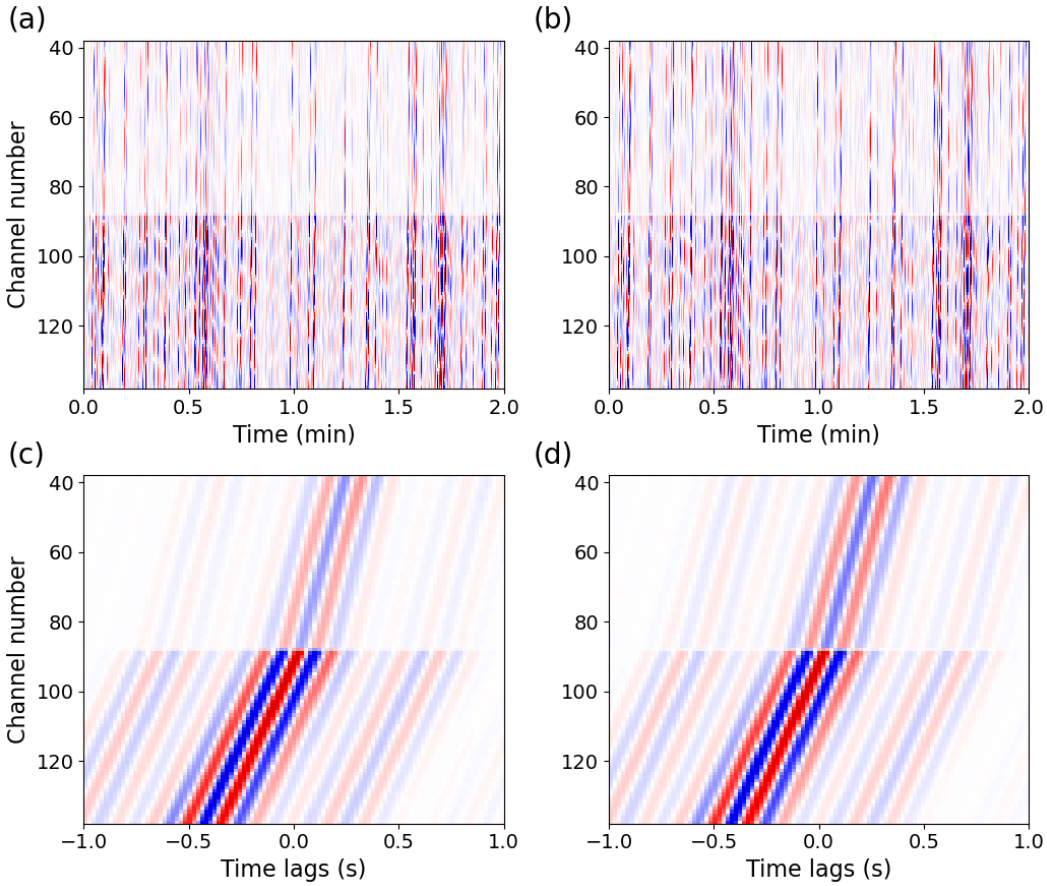

Figure 10. Synthetic DAS ambient noise data (a) when the source propagation direction is 111° and R/L ratio is 0.68, and (b) when the source propagation direction is 112° and R/L ratio is 0.47, (c) and (d) are the corresponding common virtual-shot gathers of (a) and (b), respectively.

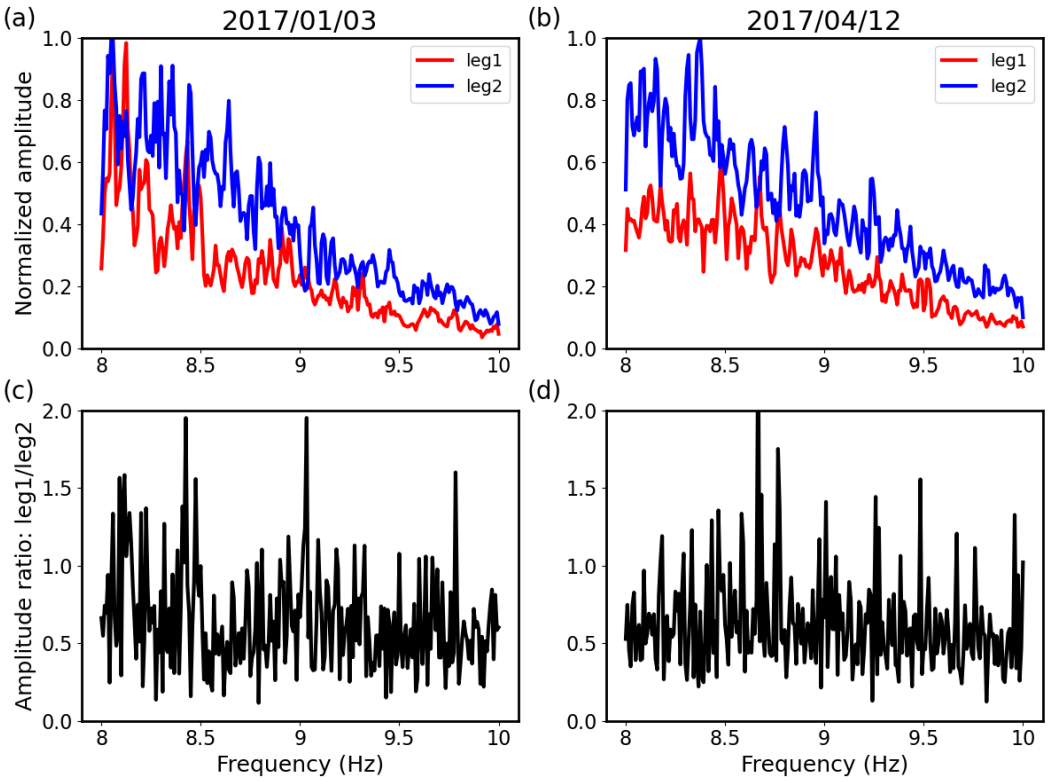

Figure 11. Averaged amplitude spectrum of the field data recorded by the two legs of the DAS array displayed in (a) Figure 8a and (b) Figure 8b, (c) and (d) are amplitude ratios calculated from (a) and (b), respectively.

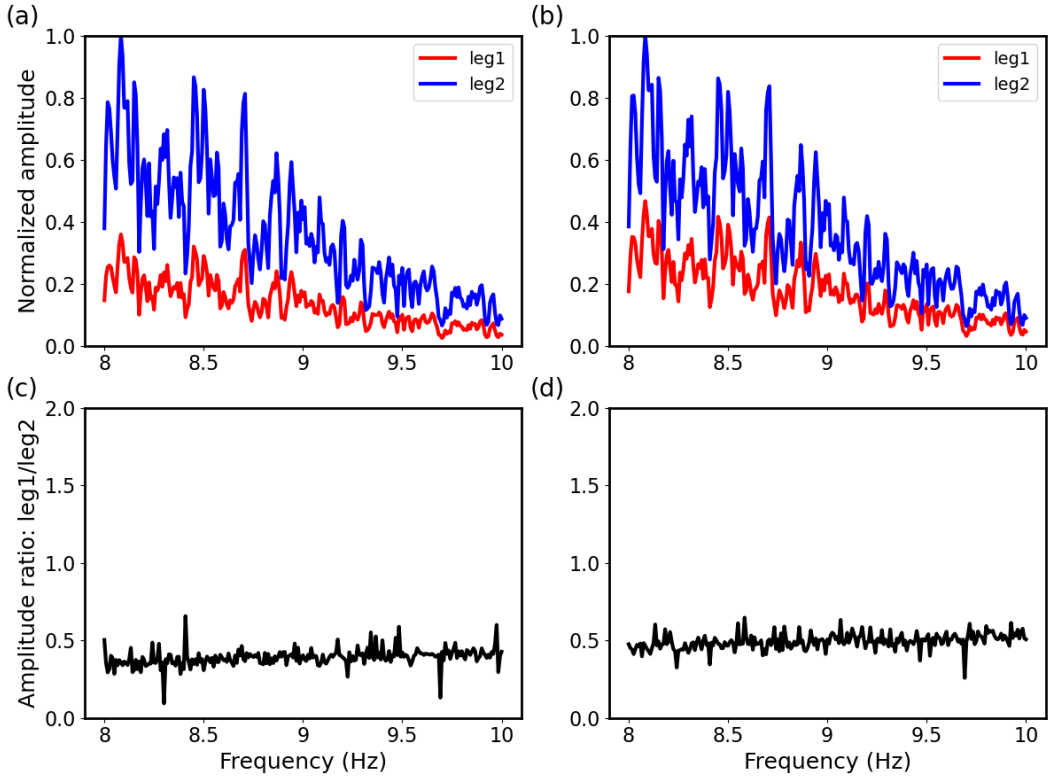

Figure 12. Same as Figure 11 except for the results are calculated from the synthetic DAS ambient noise data displayed in Figure 10a and 10b.

is strongly biased toward four parallel and orthogonal angles to the mid-line between the two legs of the array (Figure S3 in the supporting information).

We only focus on the cases when the receiver array is close to the "L" shape, which is most commonly seen for fiber-optic cable deployment in urban settings. However, when the array is in other shapes, more complex amplitude and polarity responses may be observed from DAS recordings, particularly when the source distribution is complex and unknown.

Although there are a few studies on Love waves tomography/imaging from 3C geophone/seismometer cross-correlations (e.g., Lin et al., 2008; H. Li et al., 2010; Galetti et al., 2017), these applications to DAS ambient noise data are rarely seen. Because DAS is a single-component receiver, Rayleigh and Love waves cannot be separated without enough information about the sources. Calculating the R/L ratio is a reasonable and easy-to-implement method to get information on the content of Rayleigh and Love waves. The R/L ratio can be further used to invert the near-surface shear-wave velocity model from Rayleigh and Love waves simultaneously (Zhao et al., 2023).

We assume uniform coupling at the selected locations of the fiber-optic cable and attribute the observed amplitude and polarity variations to the fiber's response to different types of surface waves. However, the accuracy of the R/L ratio may be compromised if this assumption is not satisfied. Such estimation becomes less reliable when the coupling effect dominates the amplitude response of DAS. Therefore, data quality control for fiber coupling is critical, particularly for real-time DAS monitoring applications.

6 Conclusions

Ambient noise sources generate both Rayleigh and Love waves. Unlike the common assumption that only Rayleigh waves can be extracted from a linear DAS array, we find the existence of Love waves in the extracted common virtual-shot gather even when the virtual source is co-linear with the DAS array. This is due to the non-uniform distribution of ambient noise sources in the urban environment. Thus, the contents of both Rayleigh and Love waves should be estimated and utilized for subsequent shear-wave velocity inversion.

Source propagation direction provides valuable information for estimating the content of Rayleigh and Love waves. However, beamforming results rely on source complexity when resolving the source propagation direction from the horizontally recorded DAS ambient noise data. We conclude that beamforming results are reliable when the noise sources mainly come from one direction and with careful data processing, such as amplitude normalization and polarity flip (only if there is polarity reversal). In contrast, beamforming may not accurately resolve all the directions when the noise sources are complex, such as coming from multiple directions.

Although Rayleigh and Love waves cannot be separated since we do not have sufficient information on the noise sources, the R/L ratio can be estimated with the estimated source propagation direction, the known array geometry, and the amplitude of the recorded DAS ambient noise data. The estimated R/L ratio could significantly contribute to a reliable near-surface shear-wave velocity model inverted from Rayleigh and Love waves simultaneously.

7 Open Research

Seismic data and the associate codes used in this study are available and can be downloaded from this link: `https://doi.org/10.5281/zenodo.7513783`.

Acknowledgments

We thank Professor Biondo Biondi for sharing the DAS data. We thank Dr. Gang Fang and Dr. Eileen Martin for the useful discussions. We acknowledge Cambridge Sensing Pte Ltd for its financial support. We thank the open-source software Madagascar (`http://www.ahay.org/`) and evodcinv (`https://pypi.org/project/evodcinv/1.0.0/`), with which we preprocess the data and simulate the dispersion curves.

References

Ajo-Franklin, J. B., Dou, S., Lindsey, N. J., Monga, I., Tracy, C., Robertson, M., ... others (2019). Distributed Acoustic Sensing Using Dark Fiber for Near-Surface Characterization and Broadband Seismic Event Detection. *Scientific Reports*, *9*(1), 1–14.

Behm, M., & Snieder, R. (2013). Love waves from local traffic noise interferometry. *The Leading Edge*, *32*(6), 628–632.

Benioff, H. (1935). A linear strain seismograph. *Bulletin of the Seismological Society of America*, *25*(4), 283–309.

Biondi, B., Martin, E., Cole, S., Karrenbach, M., & Lindsey, N. (2017). Earthquakes analysis using data recorded by the Stanford DAS array. In *Seg technical program expanded abstracts 2017* (pp. 2752–2756). Society of Exploration Geophysicists.

Brooks, L. A., Townend, J., Gerstoft, P., Bannister, S., & Carter, L. (2009). Fundamental and higher-mode Rayleigh wave characteristics of ambient seismic noise in New Zealand. *Geophysical Research Letters*, *36*(23).

Buchen, P., & Ben-Hador, R. (1996). Free-mode surface-wave computations. *Geophysical Journal International*, *124*(3), 869–887.

Claerbout, J. F. (1968). Synthesis of a layered medium from its acoustic transmission response. *Geophysics*, *33*(2), 264–269.

Dean, T., Cuny, T., & Hartog, A. H. (2017). The effect of gauge length on axially incident P-waves measured using fibre optic distributed vibration sensing. *Geophysical Prospecting*, *65*(1), 184–193.

Díaz, J., Ruiz, M., Sánchez-Pastor, P. S., & Romero, P. (2017). Urban seismology: On the origin of earth vibrations within a city. *Scientific reports*, *7*(1), 1–11.

Dou, S., Lindsey, N., Wagner, A. M., Daley, T. M., Freifeld, B., Robertson, M., ... 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.

Fang, G., Li, Y. E., Zhao, Y., & Martin, E. R. (2020). Urban Near-surface Seismic Monitoring using Distributed Acoustic Sensing. *Geophysical Research Letters*, *47*(6), e2019GL086115.

Galetti, E., Curtis, A., Baptie, B., Jenkins, D., & Nicolson, H. (2017). Transdimensional Love-wave tomography of the British Isles and shear-velocity structure of the East Irish Sea Basin from ambient-noise interferometry. *Geophysical Journal International*, *208*(1), 36–58.

Gelius, L.-J., Tygel, M., Takahata, A. K., Asgedom, E. G., & Serrano, D. R. (2013). High-resolution imaging of diffractions–A window-steered MUSIC approach. *Geophysics*, *78*(6), S255–S264.

Godara, L. C. (1997). Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations. *Proceedings of the IEEE*, *85*(8), 1195–1245.

Goldstein, P., & Archuleta, R. J. (1987). Array analysis of seismic signals. *Geophysical Research Letters*, *14*(1), 13–16.

Haskell, N. A. (1953). The dispersion of surface waves on multilayered media. *Bulletin of the seismological Society of America*, *43*(1), 17–34.

Johnson, D. H., & Dudgeon, D. E. (1993). *Array signal processing: concepts and techniques*. Prentice Hall, Englewood Cliffs, NJ.

Kirlin, R. L. (1992). The relationship between semblance and eigenstructure velocity estimators. *Geophysics*, *57*(8), 1027–1033.

Li, H., Su, W., Wang, C.-Y., Huang, Z., & Lv, Z. (2010). Ambient noise Love wave tomography in the eastern margin of the Tibetan plateau. *Tectonophysics*, *491*(1-4), 194–204.

Li, Y. E., Nilot, E., & Feng, X. (2020). Observation of guided and reflection P-waves in urban ambient noise cross-correlograms. In *Seg technical program expanded abstracts 2020* (pp. 2100–2104). Society of Exploration Geophysicists.

Lin, F.-C., Moschetti, M. P., & 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. J., Martin, E. R., Dreger, D. S., Freifeld, B., Cole, S., James, S. R., ... Ajo-Franklin, J. B. (2017). Fiber-optic network observations of earthquake wavefields. *Geophysical Research Letters*, *44*(23), 11–792.

Luo, B., Trainor-Guitton, W., Bozdağ, E., LaFlame, L., Cole, S., & Karrenbach, M. (2020). Horizontally orthogonal distributed acoustic sensing array for earthquake- and ambient-noise-based multichannel analysis of surface waves. *Geophysical Journal International*, *222*(3), 2147–2161.

Luu, K. (2019). *evodcinv: Inversion of dispersion curves using evolutionary algorithms*. Retrieved from `https://github.com/keurfonluu/evodcinv` doi: 10.5281/zenodo.5775193

Martin, E. R. (2018). *Passive imaging and characterization of the subsurface with distributed acoustic sensing*. PhD Thesis, Stanford University.

Martin, E. R., & Biondi, B. L. (2017). Ambient noise interferometry across two-dimensional DAS arrays. In *Seg technical program expanded abstracts 2017* (pp. 2642–2646). Society of Exploration Geophysicists.

Martin, E. R., & Biondi, B. L. (2018). Eighteen months of continuous near-surface monitoring with DAS data collected under Stanford University. In *Seg technical program expanded abstracts 2018* (pp. 4958–4962). Society of Exploration Geophysicists.

Martin, E. R., Castillo, C. M., Cole, S., Sawasdee, P. S., Yuan, S., Clapp, R., ... Biondi, B. L. (2017). Seismic monitoring leveraging existing telecom infrastructure at the SDASA: Active, passive, and ambient-noise analysis. *The Leading Edge*, *36*(12), 1025–1031.

Mateeva, A., Mestayer, J., Cox, B., Kiyashchenko, D., Wills, P., Lopez, J., ... others (2012). Advances in distributed acoustic sensing (DAS) for VSP. In *Seg technical program expanded abstracts 2012* (pp. 1–5). Society of Exploration Geophysicists.

Nakata, N., Nishida, K., Gualtieri, L., Fichtner, A., et al. (2019). Body wave exploration. *Seismic Ambient Noise*, *479*(239), 480.

Nilot, E., Zhang, Y., Li, Y. E., & Feng, X. (2019). Deep bedrock detection based on ambient noise recorded by a short geophone array: A Singapore case study. In *Seg technical program expanded abstracts 2019* (pp. 3121–3125). Society of Exploration Geophysicists.

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.

Spica, Z. J., Perton, M., Martin, E. R., Beroza, G. C., & Biondi, B. (2020). Urban seismic site characterization by fiber-optic seismology. *Journal of Geophysical Research: Solid Earth*, *125*(3), e2019JB018656.

Stehly, L., Campillo, M., & Shapiro, N. (2006). A study of the seismic noise from its long-range correlation properties. *Journal of Geophysical Research: Solid*

Earth, *111*(B10).

Thomson, W. T. (1950). Transmission of elastic waves through a stratified solid medium. *Journal of applied Physics*, *21*(2), 89–93.

Tribaldos, V. R., Ajo-Franklin, J. B., Dou, S., Lindsey, N. J., Ulrich, C., Robertson, M., ... Tracy, C. (2021). Surface Wave Imaging Using Distributed Acoustic Sensing Deployed on Dark Fiber: Moving Beyond High-Frequency Noise. *Distributed Acoustic Sensing in Geophysics: Methods and Applications*, 197–212.

Wapenaar, K. (2004). Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation. *Physical review letters*, *93*(25), 254301.

Yu, C., Zhan, Z., Lindsey, N. J., Ajo-Franklin, J. B., & Robertson, M. (2019). The potential of DAS in teleseismic studies: Insights from the Goldstone experiment. *Geophysical Research Letters*, *46*(3), 1320–1328.

Yuan, S., Martin, E., Chang, J., Cole, S., & Biondi, B. (2017). *Catalog of Northern California earthquakes recorded by DAS.* SEP-Report.

Zeng, X., Lancelle, C., Thurber, C., Fratta, D., Wang, H., Lord, N., ... 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.

Zhang, Y., Li, Y. E., Zhang, H., & Ku, T. (2019). Near-surface site investigation by seismic interferometry using urban traffic noise in Singapore. *Geophysics*, *84*(2), B169–B180.

Zhao, Y., Li, Y. E., & Fang, G. (2019). Extracting subsurface information based on extremely short period of DAS recordings. In *Seg technical program expanded abstracts 2019* (pp. 958–962). Society of Exploration Geophysicists.

Zhao, Y., Li, Y. E., & Li, B. (2023). Near-surface imaging with Rayleigh and Love waves extracted from DAS ambient noise data. *submitted to Journal of Geophysical Research*, *x*, x-xx.