Geophysical Journal International

Geophys. J. Int. (2021) **227,** 16–32
Advance Access publication 2021 May 24
GJI Seismology

https://doi.org/https://doi.org/10.1093/gji/ggab205

Effects of high-speed train traffic characteristics on seismic interferometry

Yujin Liu [1], Yubo Yue,[2] Yi Luo[3] and Youming Li[4]

[1] *Aramco Beijing Research Center, Aramco Asia, Beijing* 100102, *China*
[2] *Bureau of Geophysical Prospecting Inc., R&D Center, Zhuozhou* 72750, *China. E-mail: geoyyb@163.com*
[3] *EXPEC Advanced Research Center, Saudi Aramco, Dhahran 34465, Saudi Arabia*
[4] *Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing* 100029, *China*

Accepted 2021 May 21. Received 2021 April 19; in original form 2020 September 7

SUMMARY

Train traffic has been recognized as a powerful noise source for subsurface imaging and monitoring due to its properties of being strong, economic and repeatable. Compared with the traditional train, the high-speed train (HST) runs faster and generates vibrations that also manifest as periodical signals but with a more striking feature of sharp equidistant spectral lines in the frequency spectrum. Like the vibrations generated by the traditional train, the HST-induced vibrations can be utilized after the retrieval of interstation impulse responses using seismic interferometry. However, little attention has been paid to the characteristics of HST-induced vibrations during seismic interferometry. The aim of this study is to investigate the effects of train traffic characteristics on the seismic interferometry method. We first present an alternative derivation of seismic interferometry that is not based on the assumption of uncorrelation between sources at different locations. The derivation is valid for scalar waves in 3-D and for elastic surface waves. After analysing the effects of several HST-related factors, including HST moving direction, HST carriage number, rail structure, HST speed and the stacking number of HST records, on the seismic interferometry, we find that the seismic interferometry using HST traffic noise generates crosstalk that can be effectively attenuated by stacking the interferometric results of several HSTs with slightly different speeds. This finding has been validated by the seismic interferometry of both synthetic and field HST traffic noise. We further demonstrate that the retrieved surface waves can be used to estimate near-surface velocities.

Key words: Image processing; Seismic interferometry; Wave propagation.

1 INTRODUCTION

Ambient seismic noise has received much attention in recent years as a new probing signal for the solid earth (Sabra *et al.* 2005; Gerstoft *et al.* 2006; Bensen *et al.* 2007; Riahi *et al.* 2013; Weemstra *et al.* 2013; Zhang *et al.* 2019b) and can be used to predict earthquake ground motion (Denolle *et al.* 2014). Ambient seismic noise can be generated by natural processes, such as wind, volcano and ocean swell, or anthropogenic processes, such as machinery, traffic noise and construction activities. Among all these ambient noise types, train traffic noise is one of the most energetic anthropogenic seismic sources. In addition, compared with natural ambient noise like ocean swells, the train traffic noise has broad-band spectrum (Fuchs & Bokelmann 2018; Brenguier *et al.* 2019a; Wang *et al.* 2021), which is extremely important to image subsurface heterogeneities at different scales. However, due to the fact that the train traffic noise is a continuous (over a short time period) passive source, it is necessary to reconstruct impulsive responses of the earth from the passive seismic measurements prior to seismic imaging.

One way to turn passive seismic measurements into deterministic seismic responses is the seismic interferometry method (Campillo & Paul 2003; Schuster 2009). It involves the cross-correlation of noise measurements at two different sensors to obtain the impulse response of the earth between the two sensors as if a virtual source was emitted at each of the sensors. This approach has been proved successful for the reconstruction of low-frequency surface waves originating from ocean swell activity (Shapiro & Campillo 2004). Under specific virtual source and receiver configurations involving seismic arrays, it is possible to use cross-correlation of ambient noise to extract body waves traveling in the shallow crust (Roux *et al.* 2005; Nakata *et al.* 2015). However, the noise source for reconstructing these body waves, likely related to ocean shore break, wind or cultural activity, remains unclear. On the other hand, with a known seismic source generated by train traffic, both surface and body waves are successfully retrieved for shallow surface applications (Nakata *et al.* 2011; Quiros *et al.* 2016). Furthermore,

train traffic noise has been successfully utilized as a new seismic source to reconstruct direct body waves to monitor the San Jacinto Fault at 4 km depth (Brenguier *et al.* 2019a). Compared with other ambient noise, train traffic is broad-band and exhibits sharp equidistant spectral lines in the entire broad-band frequency range (Chen *et al.* 2004; Quiros *et al.* 2016; Brenguier *et al.* 2019a; Wang *et al.* 2021). Although this striking feature of train traffic noise was observed over ten years ago, little attention has been paid to the effects of such train traffic characteristics on the seismic interferometry method.

Compared with the traditional train, high-speed train (HST) often operates at a speed of 200 km hr^{-1} or more; it offers a faster and more robust travel option. Because of these advantages, high-speed rail (HSR) has been growing fast in recent decades across the world, especially in China. By 2015, China had constructed the longest HSR network (20 743 km) in the world (Wang *et al.* 2018). Like the traditional train, the HST generates strong ground vibrations when it moves on the railroad. However, since the HST demands a strict requirement for track regularity, bridge is a preferred construction type for HSR. As observed by He *et al.* (2017), the average bridge occupation ratio in China is greater than 50 per cent and the maximum bridge ratio is as high as 94.2 per cent on Guangzhou–Zhuhai intercity railway. Due to the regularity of HSR, the HST-generated vibrations are more regular and present more prominent feature of equidistant spectral lines in the entire broad-band frequency range than the traditional train (Fuchs & Bokelmann 2018). In addition, Liu *et al.* (2021b) demonstrate that the bridge pier plays an important role for HSTs in generating broad-band seismic waves that can be used for subsurface imaging and monitoring. When using seismic interferometry to extract the Green's function from the HST-induced vibration, Zhang *et al.* (2019a) find that the source crosstalk could generate artefacts due to the moving property of the HST source. Liu *et al.* (2021a) give a comprehensive analysis on the retrievability of different types of seismic waves when using train traffic noise with seismic interferometry. However, previous studies have not analysed the effects of HST-related properties, such as the HST structure and HST speed, on seismic interferometry and have not provided practical solutions to attenuate their effects to the retrieval of seismic waves.

The purpose of this paper is to investigate thoroughly how the characteristics of the HST traffic noise affect the retrieval of the Green's function using seismic interferometry. We first explicitly formulate the source function generated by HST and then present an alternative derivation of seismic interferometry without the assumption of uncorrelation between non-collocated sources. This derivation is valid for both scalar and surface waves. After the derivation, we analyse the effects of several HST-related factors, including HST moving direction, HST carriage number, rail structure, HST speed and the stacking number of HST records, on the seismic interferometry using both synthetic and field data. We further illustrate the potential of train traffic noise for near-surface characterization using the retrieved surface waves.

2 METHODOLOGY

In this section, we present an alternative derivation of seismic interferometry using HST-induced vibrations. Compared with the previous derivation of passive seismic interferometry (Snieder 2004), we explicitly formulate the source function generated by the HST and consider its effects on seismic interferometry deterministically, rather than assuming the source functions from different locations are uncorrelated statistically. Since the main difference between our derivation and the previous one comes from the treatment of source functions, we start this section by presenting the HST source function. Then we give the derivation of seismic interferometry using HST-induced vibrations for the simplest case of scalar waves in a homogeneous medium. Finally, we extent the derivation to elastic surface waves in 3-D.

2.1 HST source function

As described by Krylov (1994) and Hung & Yang (2001), the train-induced vibrations are caused by loading forces to the track from each wheel axle and then causing downward deflection of the track and underlying soils. Thus, the contact point between the track and underlying soils, such as the sleeper and the bridge pier, acts as a vibration source. Since, in the relevant frequency band, the dimensions of either sleeper or bridge pier are often negligible compared to the characteristic wavelengths of generated elastic waves, each of them can be considered as a point source when the HST is passing it. It is also reasonable to assume the high-speed railway to be a single line because the track gauge is negligible as well compared to the seismic wavelengths. Fig. 1(a) shows a two-carriage HST passing a sleeper or a bridge pier at x_s. The corresponding source function can be represented by a series of time-delayed signals, as shown in Fig. 1(b). The time delay is determined by the source location and the HST parameters (i.e. HST speed, carriage length and wheel spacing). Mathematically, the source function of a carriage can be expressed as

$$h_s\ (t) = \sum_{j=1}^{N_a} f_s \left(t - \frac{x_s + d_j}{v_s} \right), \tag{1}$$

where t is the time, N_a is the number of wheel axles in a carriage and is 4 in most HSTs, $f_s(t)$ is the wheel axle loading function that could change over source location x_s, d_j is the distance from the jth wheel axle to the head of this carriage, v_s is the HST speed when it is passing the location x_s and can reach 300–350 km hr^{-1} (or 83.3–97.2 m s^{-1}). Since the HST often consists of a series of equidistant carriages, the source function of a full-length HST that passes the location x_s can be represented as

$$w_s\ (t) = \sum_{i=1}^{N_c} h_s \left(t - \frac{i \times L}{v_s} \right), \tag{2}$$

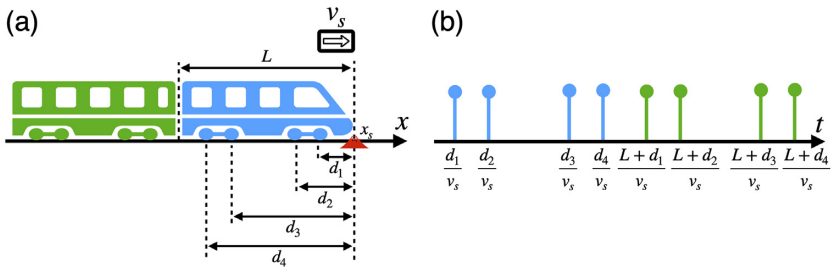

Figure 1. Schematic diagram of an HST source. (a) An HST with two carriages passing a sleeper or a bridge pier at x_s at a speed of v_s. The length of each carriage is L, and the distances from the carriage head to each pair of wheels within this carriage are d_1, d_2, d_3 and d_4, respectively. (b) The corresponding source function represented by a time series. The time delays are determined by the HST speed, the HST wheel distribution and the location of the sleeper or the bridge pier.

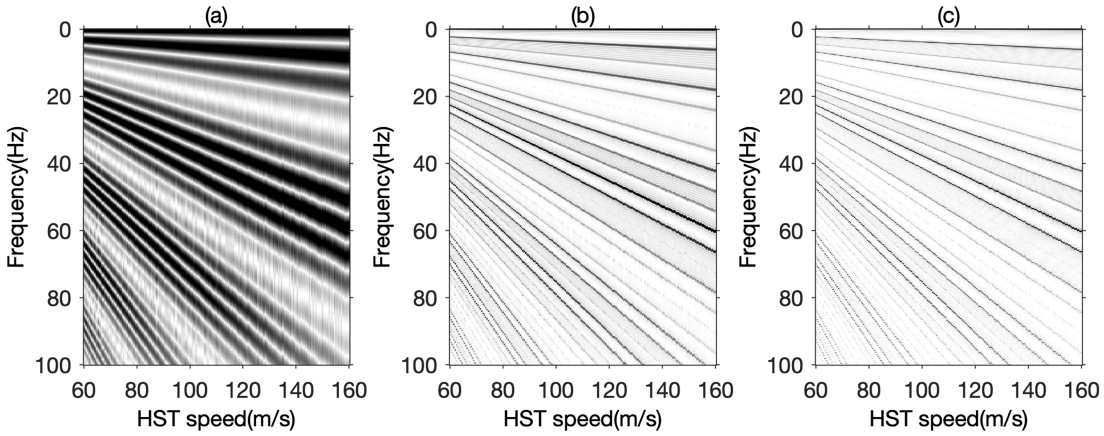

Figure 2. Train-speed-dependent frequency spectra of source functions for HSTs with (a) 1 carriage, (b) 8 carriages and (c) 16 carriages.

where N_c is the carriage number and equals 8 or 16 in most cases, and L is the carriage length. Eq. (2) represents a periodical signal repeating the same pattern of $h_s(t)$ over identical subsequent periods. In the Fourier domain, this signal is transformed to

$$W_s\ (\omega) = \sum_{k=1}^{N_\omega} H_s\,(\omega)\,\delta\,(\omega - k\omega_s)\,, \tag{3}$$

where N_ω is the sampling number with the sampling rate of ω_s for the entire frequency band, where $\omega_s = v_s/L$, and $H_s(\omega)$ is the Fourier transform of $h_s(t)$. As evident from eq. (3), the frequency spectrum of the HST source function is composed of a series of equidistant spikes and the spacing of these spikes ω_s depends on the HST speed and the carriage length. This spectral characteristic has been well observed in the field data (Chen *et al.* 2004; Quiros *et al.* 2016; Fuchs & Bokelmann 2018; Wang *et al.* 2021).

To further illustrate the properties of the HST source function, numerical experiments are conducted to analyse its dependence on the HST speed and structure. Fig. 2 shows the frequency spectra of HST source functions with different HST speeds and carriage numbers. The results not only validate that the spectrum of the HST source could be represented as a series of equidistant spikes, but also show that the higher the carriage number, the spikier the spectrum. In addition, the spacing of these spikes is proportional to the HST speed. This property could be used to determine the HST speed when the HST structure is known (X. Wang *et al.* 2021).

Before the derivation of seismic interferometry using HST traffic, we first analyse the cross-correlation of two HST sources at locations x_s and $x_{s'}$, respectively. The cross-correlation is

$$C\ (\omega) = W_s^*\ (\omega)\, W_{s'}\ (\omega) = \sum_{k=1}^{N_\omega} \sum_{k'=1}^{N_\omega} H_s^*\,(\omega)\, H_{s'}\,(\omega)\,\delta\,(\omega - k\omega_s)\ \delta\left(\omega - k'\omega_{s'}\right). \tag{4}$$

From eq. (4), we can conclude that the cross-correlation is zero, except when $k\ \omega_s = k'\ \omega_{s'}$. This means that the dominant contribution of the cross-correlation comes from the autocorrelation terms $v_s = v_{s'}$ and the cross terms $v_s \neq v_{s'}$ is small. This observation can be validated numerically. Fig. 3 shows the cross-correlations of HST source functions with different HST speeds and carriage numbers. As can be seen from Fig. 3, the dominant contribution of the cross-correlation between two HST sources comes from the autocorrelation terms, especially when the HSTs have more than one carriage.

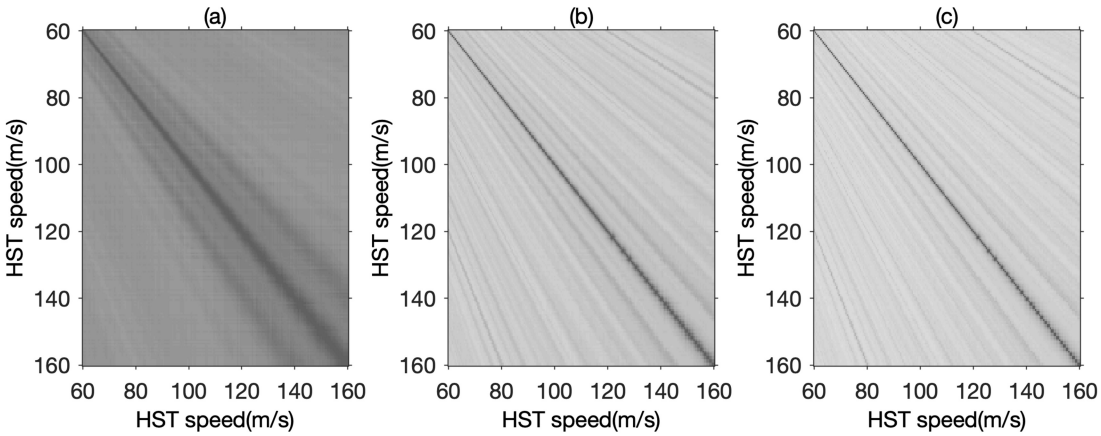

Figure 3. Train-speed-dependent cross-correlation of source functions for HSTs with (a) 1 carriage, (b) 8 carriages and (c) 16 carriages.

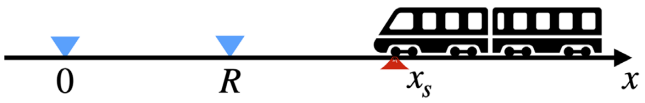

Figure 4. Definition of the geometric variables for the waves that travel from a source at location x_s to two receivers. Receivers 1 and 2 locate at $x = 0$ and $x = R$, respectively.

2.2 Seismic interferometry for scalar waves in a 3-D medium

With the previous analyses on HST source functions, we next derive the formula of seismic interferometry using the HST traffic for the simplest case of scalar waves in a homogeneous medium. Assuming that the propagation velocity of seismic waves is constant with value c, the wavefields at two receivers x_1 and x_2 can be written as a superposition of the waves radiated by all HST sources

$$p_{1,2} = \sum_s \frac{w_s\left(t - \frac{|x_{1,2} - x_s|}{c}\right)}{|x_{1,2} - x_s|}, \tag{5}$$

where w_s is the HST source function at location x_s and is given by eq. (2). In seismic interferometry, one correlates the wavefields recorded at two receivers over a time window of length T:

$$c\ (\tau) = \int_0^T p_2\,(t+\tau)\,p_1\,(t)\,\mathrm{d}t. \tag{6}$$

Considering that these two receivers are separated by a distance R, and receiver 1 locates at the origin, as shown in Fig. 4, inserting eq. (5) into eq. (6) gives a double sum over all sources

$$c\ (\tau) = \sum_{s,s'} \int \frac{w_s\,(t)\,w_{s'}\left(t + \frac{|x_{s'}| - |R - x_{s'}|}{c} + \tau\right)}{|x_s|\,|R - x_{s'}|}\mathrm{d}t. \tag{7}$$

The double sum in eq. (7) can be split into a sum over diagonal terms and a sum over cross terms. It has been demonstrated in the previous section that, for two HST sources with different train speeds, the ratio of the cross terms to the diagonal terms is small. By increasing the number of HST events to be averaged, the sum of the cross terms can be considered arbitrarily small.

In the following, we assume that sufficient event averaging is carried out so that the cross terms in the sum can be ignored, then

$$c\ (\tau) = \sum_{s} \int \frac{w_s\,(t)\,w_s\left(t + \frac{|x_s| - |R - x_s|}{c} + \tau\right)}{|x_s|\,|R - x_s|}\mathrm{d}t. \tag{8}$$

In Appendix A, we show that this integral can be solved following the stationary phase argument of Snieder (2004). The resulting cross-correlation is given in the frequency domain by

$$C\ (\omega) = \overline{|W_s\,(\omega)|^2}\left(\frac{e^{ikR}}{4\pi R}\int_{-\infty}^{0} \frac{\mathrm{R}}{x_s\,(x_s - R)} n dx_s + \frac{e^{-ikR}}{4\pi R}\int_R^{+\infty} \frac{R}{x_s\,(x_s - R)} n\mathrm{d}x_s\right), \tag{9}$$

where $W_s(\omega)$ is the Fourier transform of $w_s(t)$, k is the wavenumber and n is the source density defined as the number of sources per unit length. From eq. (9), we can conclude that the geometrical spreading of scalar waves cannot be correctly retrieved by seismic interferometry using HST traffic. However, the retrieved traveltime information is still reliable. More specifically, the first term in brackets $\frac{e^{ikR}}{4\pi R}$ is the Green's function that accounts for waves that propagate from receiver 1 to receiver 2. This term comes from the integration over $x_s < 0$; these waves correlate for a positive lag time $\tau > 0$. The second term $\frac{e^{-ikR}}{4\pi R}$, which comes from the integration over $x_s > R$, is the advanced Green's

function. The presence of the advanced Green's function is due to the waves that propagate from receiver 2 to receiver 1; these waves correlate for a negative lag time $\tau < 0$.

2.3 Seismic interferometry for surface waves in an elastic medium

In this section, we generalize the previous derivations to surface wave retrieval in a layered elastic 3-D medium. The surface wave Green's tensors of a layered medium, whose properties depend on the depth z only, can be written in the frequency domain (Snieder 2004) as

$$G_{ij}\ (\mathbf{r}, \mathbf{r}_s) = \sum_m G_{ij}^m\,(\mathbf{r}, \mathbf{r}_s)\ . \tag{10}$$

In eq. (10), the total surface wave Green's tensor $G_{ij}(\mathbf{r}, \mathbf{r}_s)$ is expressed as a sum over surface wave modes m that include both Rayleigh waves and Love waves.

The surface wave Green's tensor of mode m in the far field (Sneider 1986) is given by

$$G_{ij}^m(\mathbf{r}, \mathbf{r}_s) = p_i^m(z, \varphi)\, p_j^{m*}(z_0, \varphi)\frac{e^{i\left(k_m R+\frac{\pi}{4}\right)}}{\sqrt{\pi k_m R/2}}, \tag{11}$$

where $R = \sqrt{(x-x_s)^2+(y-y_s)^2}$ is the distance between the points measured in the horizontal plane, k_m is the horizontal wave number of mode m, and p_i^m and p_j^m are components of the polarization vector $\mathbf{p}^m(z, \varphi)$ that depends on the depth z and the azimuth φ of the path between point $\mathbf{r}_s$ and $\mathbf{r}$.

The displacement at two receivers in the frequency domain is given by a double sum over sources s and surface wave modes m (Snieder 2004) as

$$\mathbf{u}_{1,2} = \sum_s \sum_m \mathbf{p}^m\left(z_{1,2}, \varphi_{1,2}^{(s)}\right)\frac{e^{i\left(k_m|x_{1,2}-x_s|+\frac{\pi}{4}\right)}}{\sqrt{k_m\,|x_{1,2}-x_s|}}\,W_s^m\,(\omega)\,. \tag{12}$$

In this expression, $\varphi_{1,2}^{(s)}$ is the azimuth of the wave path from s to r_1 or r_2, and $W_s^m(\omega)$ is the frequency spectrum of the radiation of mode m from source s.

When the two receivers record the three-component wavefields of the ground motion induced by a moving HST, one can form the correlation tensor of all combinations of components

$$C_{ij}\ (\tau) = \int u_{2i}\,(t+\tau)\,u_{1j}\,(t)\,\mathrm{d}t, \tag{13}$$

where u_{2i}, for example, is the ith component of the displacement recorded at receiver 2. Inserting eq. (12) into the correlation eq. (13) gives a double sum over sources, that is,

$$C_{i,j}\ (\omega) = \sum_{s,s'}\sum_{m,m'} p_i^m\,(z_2, \varphi_2)\,p_j^{m'*}\,(z_1, \varphi_1)\,\frac{e^{i\left(k_m|R-x_s|-k_{m'}|x_{s'}|\right)}}{\sqrt{k_m k_{m'}\,|R-x_s|\,|x_{s'}|}}\,W_s^m\,(\omega)\,W_{s'}^{m'*}\,(\omega)\,. \tag{14}$$

Like eq. (7), the double sum over sources in eq. (14) can be split into a sum over diagonal terms and a sum over cross terms. For two HST sources with different train speeds, it has been demonstrated in the previous section that the ratio of the cross terms to the diagonal terms is small. By increasing the number of HST events to be averaged, the sum of the cross terms can be considered arbitrarily small.

In the following, we assume that sufficient event averaging is carried out so that the cross terms in the sum can be ignored, then eq. (14) reduces to

$$C_{i,j}\ (\omega) = \sum_{s}\sum_{m,m'} p_i^m\,(z_2, \varphi_2)\,p_j^{m*}\,(z_1, \varphi_1)\,\frac{e^{i\left(k_m|R-x_s|-k_{m'}|x_s|\right)}}{\sqrt{k_m k_{m'}\,|R-x_s|\,|x_s|}}\,W_s^m\,(\omega)\,W_s^{m'*}\,(\omega)\,. \tag{15}$$

In Appendix B, we show that eq. (15) can be solved following the stationary phase argument of Snieder (2004), resulting in

$$C_{ij}\ (\omega) = \sqrt{\frac{2}{\pi}}\sum_m c_m\left\{\frac{G_{ij}^m\,(\mathbf{r}_2, \mathbf{r}_1)}{i\omega}\int_{-\infty}^{0} n\mathrm{d}x + \left[\frac{G_{ij}^m\,(\mathbf{r}_1, \mathbf{r}_2)}{i\omega}\right]^*\int_R^{\infty} n\mathrm{d}x\right\}\overline{|W^m\,(\omega)|^2}, \tag{16}$$

where $\overline{|W^m(\omega)|^2}$ the average power spectrum of the radiated mode m. The first term is due to right-going waves that are generated in the region $x < 0$; the second term is due to waves propagated from the location $x > R$ that move toward the left. We thus demonstrate that although the geometrical spreading of surface waves is incorrectly retrieved by seismic interferometry using HST traffic, the retrieved traveltime information is still reliable. This indicates that the retrieved surface waves can be utilized for near-surface velocity inversion.

3 NUMERICAL EXAMPLES

In this section, both synthetic and field data are used to validate the theory of seismic interferometry using HST traffic. In the synthetic data tests, two models, including a homogeneous model and a layered model, are used to demonstrate the effectiveness of seismic interferometry

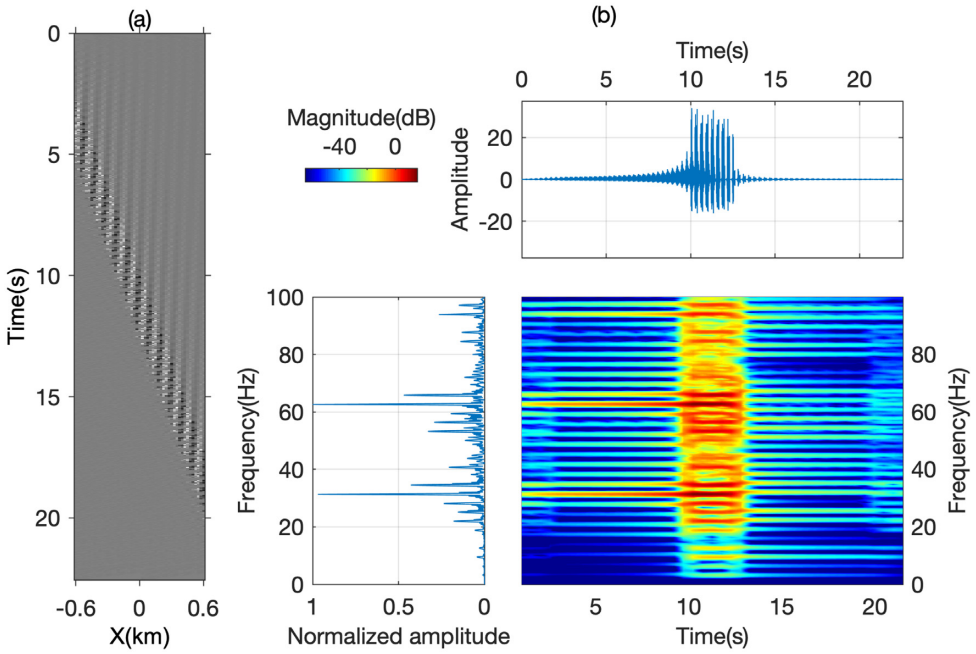

Figure 5. Simulated seismic records and their spectra. (a) Vibrations generated by 1 HST with 8 carriages passing with the speed of 300 km hr^{-1}. (b) Time–frequency analysis panels of the trace at $x = 0$ m in (a). The top-right, bottom-left and bottom-right panels in (b) show the seismogram, spectrum, and spectrogram, respectively.

in the retrieval of scalar and surface waves, respectively. Then a dispersion curve is extracted from the retrieved surfaces waves to estimate a shear wave velocity model. Finally, a field data set is used to validate the HST interferometry and demonstrate its potential in near-surface characterization.

3.1 Homogeneous model

We first use a homogeneous velocity model to validate the HST seismic interferometry theory for scalar-wave retrieval in a 3-D medium. The propagation velocity is 2 km s^{-1}. The HST has eight carriages, and each carriage has a length of 26.6 m and has four pairs of wheels. The distances from the four pairs of wheels to the carriage head within each carriage are 2.55, 5.25, 21.35 and 24.04 m, respectively. There are 81 bridge piers and 81 receivers along a straight high-speed railway. The bridge piers act as seismic sources when an HST is passing them. Both the source and receiver intervals are 25 m. Eq. (5) is implemented to simulate scalar wavefields in this numerical test. The one-wheel loading function is chosen to be a Ricker wavelet with a dominant frequency of 50 Hz. Fig. 5(a) shows a typical synthetic seismogram generated by an eight-carriage HST running from left to right at a constant speed of 83.33 m s^{-1} (that is 300 km hr^{-1}). The HST source signal could be easily identified from the record based on the vibration amplitude at each receiver and the apparent velocity of events across the receiver array. Fig. 5(b) shows the seismogram (top right), spectrum (bottom left) and spectrogram (bottom right) for the trace at $x = 0$ m in Fig. 5(a). It is obvious from the seismogram that the generated seismic waves are periodical. This is also evident from the spectrum, which is dominated by remarkably sharp spectrum lines with a constant frequency spacing of 3.13 Hz over the entire frequency band. The frequency spacing is determined by the HST speed and the carriage length, and could be well predicted by eq. (3). In addition, we can observe a clear shift of frequency components from the spectrogram. This could be explained by Doppler effect due to the moving property of the HST source. Even though the raw HST vibration records contain abundant seismic waves, utilizing it directly for subsurface imaging is challenging due to the complexity of the moving HST source. Seismic interferometry is therefore adopted to retrieve the impulse response from the vibrations induced by the HST source.

We first analyse the effect of HST moving direction on seismic interferometry. The interferometric results are shown as virtual shot gathers that can be generated by fixing the first receiver as a virtual shot location and then cross-correlating the observation at every receiver with that at the first one. It is worth noting that both time- and frequency-domain normalizations have been applied to the raw data before the cross-correlation. Figs 6(a) and (b) show the virtual shot gathers retrieved from the vibrations induced by an HST moving from left to right, and from right to left, respectively. Compared to the directly computed shot gather shown in Fig. 6(c), both of the virtual shot gathers are contaminated by strong periodical noise. As explained in Section 2, this noise comes from the crosstalk of non-collocated sources; the period of the crosstalk depends on the HST speed and the source spacing. Therefore, this crosstalk can be attenuated by stacking multiple records of vibrations with different HST speed. Another interesting aspect from Figs 6(a) and (b) is that the first one mostly retrieves the acausal part of Green's function, while the second one mostly retrieves the causal part. This ambiguity could be addressed by folding the

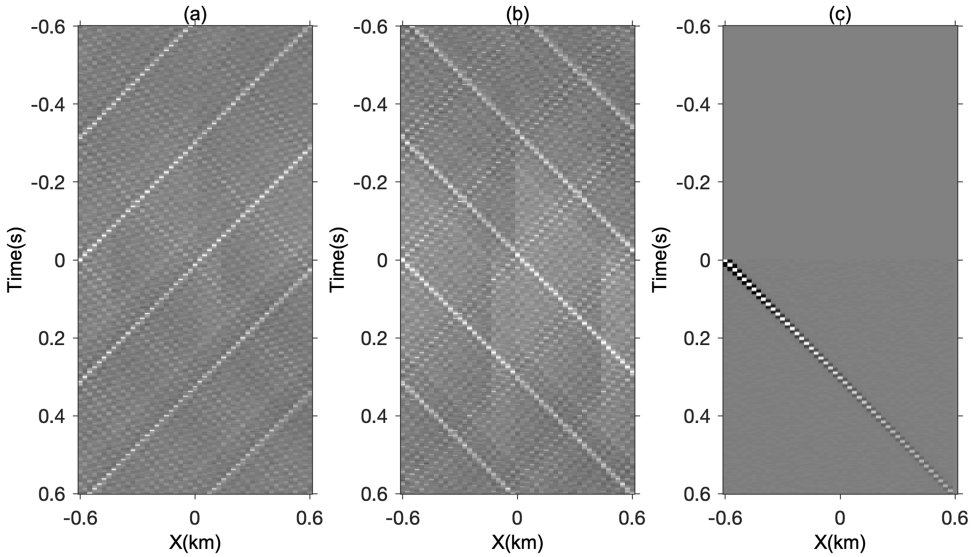

Figure 6. Scalar waves retrieved from seismic interferometry using vibrations induced by HSTs running from (a) left to right and (b) right to left. (c) The directly computed scalar wave.

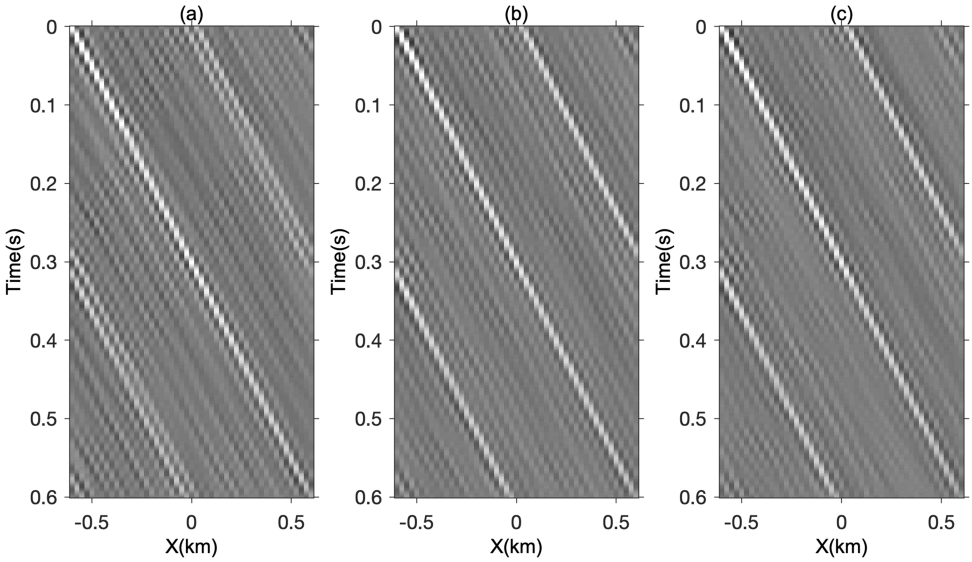

Figure 7. Scalar waves retrieved from seismic interferometry using vibrations induced by HSTs with (a) 1 carriage, (b) 8 carriages, and (c) 16 carriages. The train speed is 300 km hr^{-1}.

cross-correlations of negative lags with the positive lags, which can also help suppressing the crosstalk. It is worth noting that the geometric spreading information is lost during the seismic interferometry, whereas the traveltime information is retrieved correctly.

Then we analyse the effect of HST carriage number on seismic interferometry. The HST speed is fixed to be 300 km hr^{-1}. Figs 7(a)–(c) show the virtual short gathers retrieved from the vibrations induced by an HST with 1 carriage, 8 carriages and 16 carriages, respectively. After cross-correlation, the data in negative lags have been folded to the positive lags, followed by a dip filtering. Similar to the previous results shown in Fig. 6, all of the interferometric results contain useful scalar waves, but they are also contaminated by the crosstalk. In addition, these results reveal that the more carriages, the crosstalk is slightly better focused at the predicted location. This implies that the attenuation of the crosstalk could be more effective if there are more carriages in an HST. In practice, two types of HSTs, 8 carriages and 16 carriages, are common and the carriage number could be sufficient for the purpose of crosstalk attenuation.

We next analyse the effect of rail structure on seismic interferometry. Since the HSR is preferred to be built on a bridge, we mainly focus on the effect of the spacing of bridge piers on seismic interferometry. As mentioned in Section 2, the bridge piers act as seismic sources when the HST is passing through them. In this analysis, the HST speed is fixed to be 300 km hr^{-1} and the carriage number is 8. Figs 8(a)–(c) show the virtual shot gathers using HST-induced vibrations when the spacings of bridge piers are 12.5, 25 and 50 m, respectively. A comparison between these interferometric results shows that the periods of the crosstalk are proportional to the source spacing.

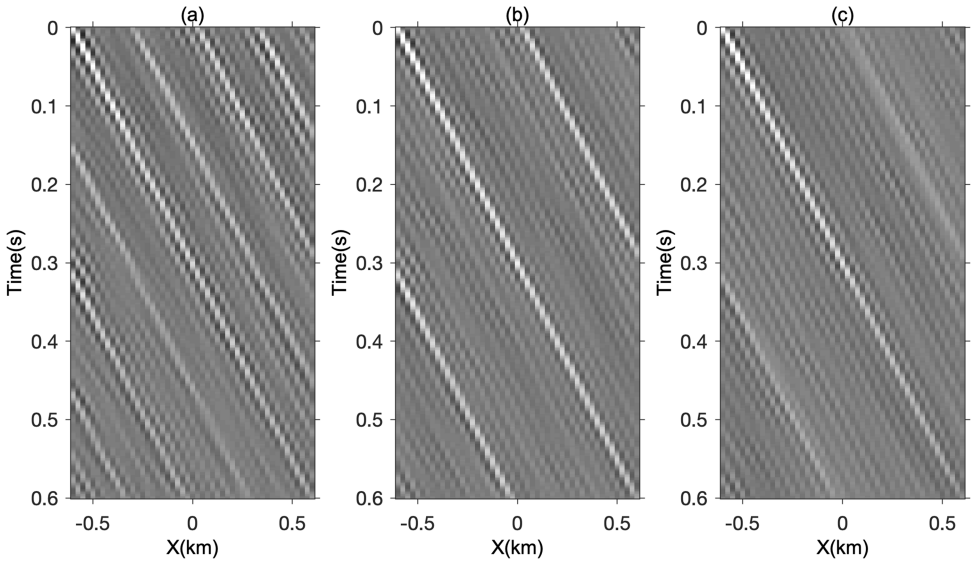

Figure 8. Scalar waves retrieved from HST-induced vibrations with source spacing of (a) 12.5 m, (b) 25 m, and (c) 50 m. The HST has eight carriages and runs at a speed of 300 km hr^{-1}.

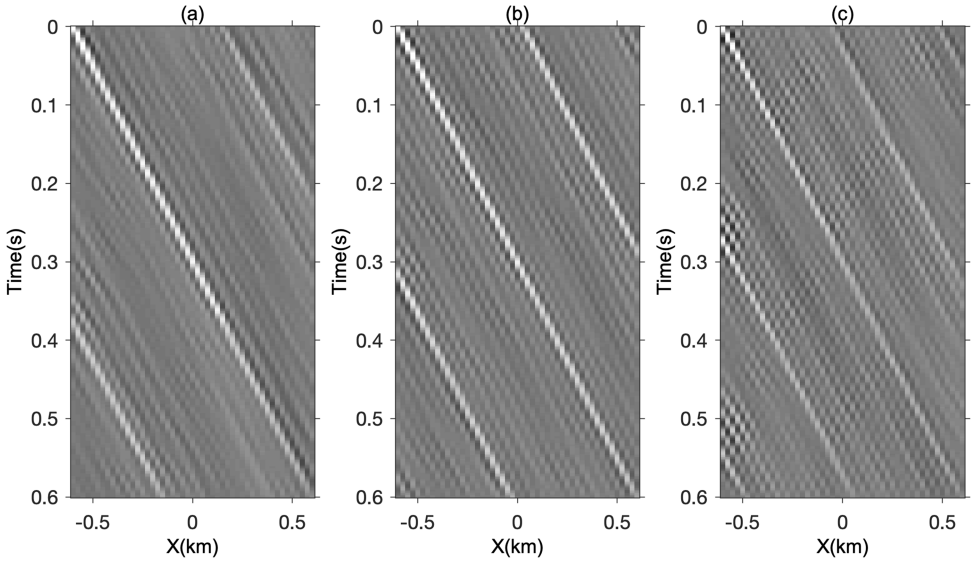

Figure 9. Scalar waves retrieved from seismic interferometry using vibrations of HSTs running at speeds of (a) 250 km hr^{-1}, (b) 300 km hr^{-1}, and (c) 350 km hr^{-1}, respectively.

Furthermore, we analyse the effect of HST speed on seismic interferometry. Figs 9(a)–(c) show the scalar waves retrieved from vibrations generated by HSTs at speeds of 250, 300 and 350 km hr^{-1}, respectively. We can observe that they are all contaminated by the crosstalk that is parallel to the desired scalar waves. In addition, the periods of the crosstalk are inversely proportional to the HST speeds, which indicates that the crosstalk can be attenuated by stacking all of these interferometric results.

Finally, we analyse the effects of the stacking number of HST records on the seismic interferometry. Figs 10(a)–(c) show the scalar waves retrieved from vibrations generated by 5, 10 and 20 HSTs, respectively. It is clear that the more HST records being used for stacking, the less crosstalk left in the final retrieved seismic waves. In addition, in this numerical example, 10 HSTs seem to be sufficient for the retrieval of the desired scalar waves.

3.2 Layered model

A three-layer model is used to verify the HST interferometry theory for surface wave retrieval in an elastic medium. Table 1 shows the model parameters. SPECFEM2D (Tromp *et al.* 2008) is used for the simulation of ground vibrations generated by the moving HST. Fig. 11(a) shows a typical seismic record with an 8-carriage HST running at a speed of 250 km hr^{-1} ($\approx$ 69.4 m s^{-1}). The one-wheel loading function is described by a Ricker wavelet with a dominant frequency of 20 Hz. The bridge piers are evenly distributed from 0 to 300 m with a spatial

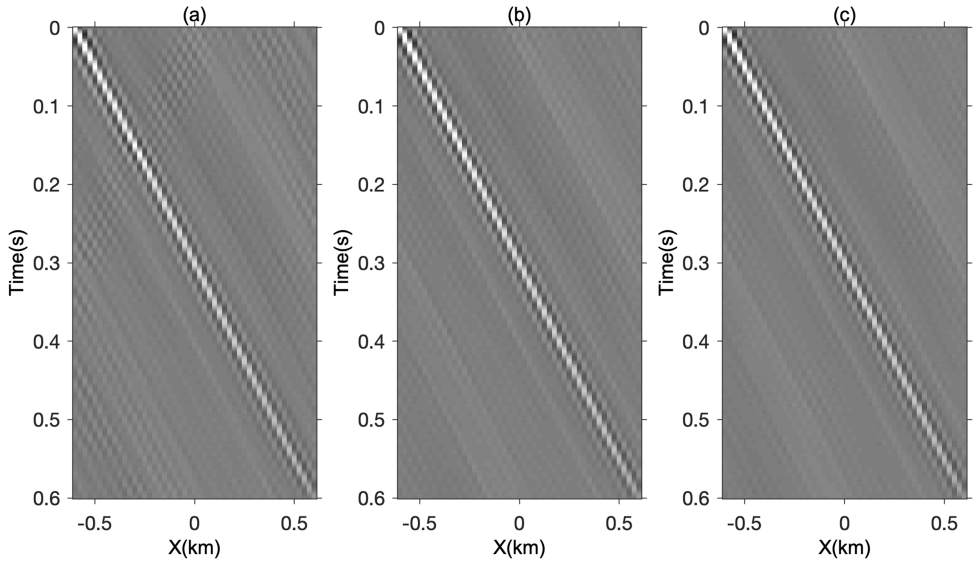

Figure 10. Scalar waves retrieved from seismic interferometry using vibrations of (a) 5 HSTs, (b) 10 HSTs, and (c) 20 HSTs.

Table 1. Parameters of a three-layer model.

Layer number	Shear wave velocity (m s^{-1})	Compressional wave velocity (m s^{-1})	Mass density (kg m^{-3})	Layer thickness (m)
1	100	680	2000	4.0
2	160	1300	2000	8.0
3 (half-space)	348	1400	2000	Infinite

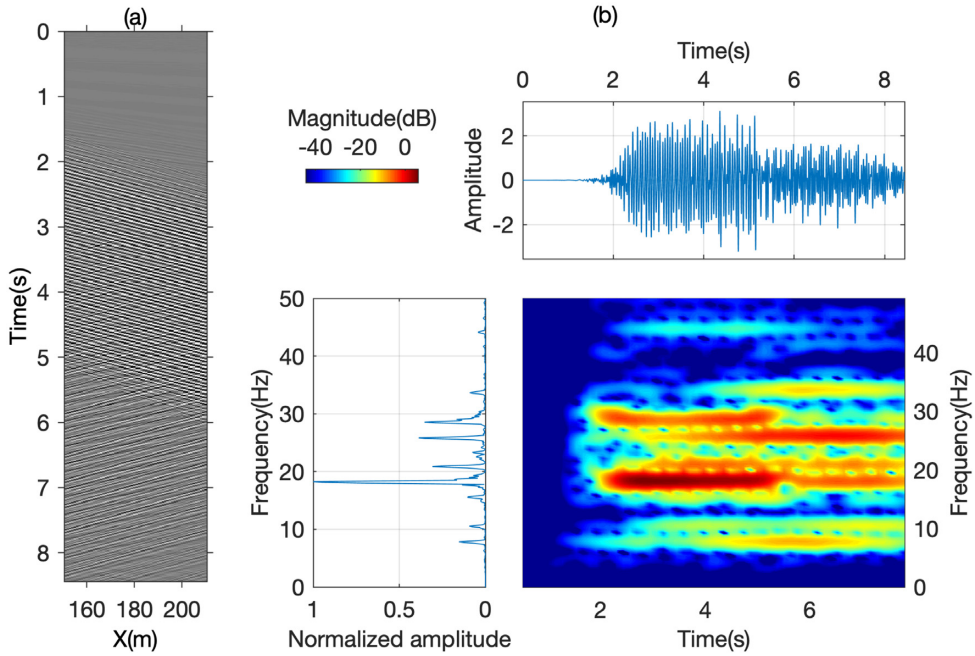

Figure 11. (a) Synthetic data stimulated by a moving HST. (b) Time–frequency analysis panels of the trace at $x = 180$ m in (a). The top-right, bottom-left and bottom-right panels in (b) show the seismogram, spectrum, and spectrogram, respectively.

interval of 25 m. 60 receivers are evenly deployed at the surface from 151 to 210 m with a spatial interval of 1 m. From the simulated data shown in Fig. 11(a), it is obvious that there are strong surface waves before and after the HST is passing the receivers. Fig. 11(b) shows the seismogram (top right), spectrum (bottom left) and spectrogram (bottom right) for the trace at $x = 180$ m in Fig. 11(a). It is clear from the seismogram that the generated seismic waves are periodical. This is also evident from the spectrum that is dominated by remarkably sharp spectrum lines with a constant frequency spacing that can be predicted by eq. (3). In addition, we can observe the Doppler effect from

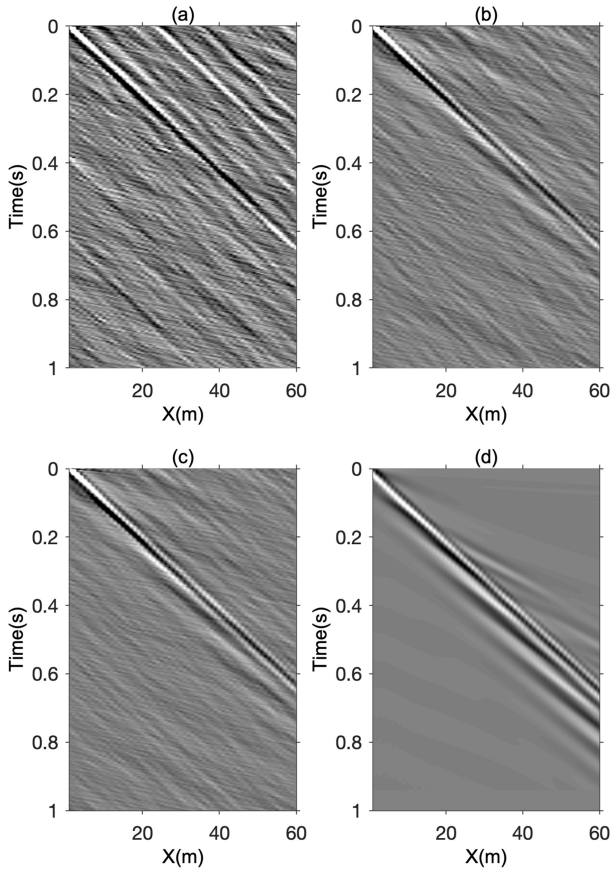

Figure 12. Surface waves retrieved from seismic interferometry using vibrations of (a) 1 HST, (b) 5 HSTs and (c) 10 HSTs. (d) The directly computed surface waves.

the spectrogram due to the moving property of the HST source. This makes direct utilization of HST-induced vibrations challenging for subsurface characterization.

Seismic interferometry is then used to extract surfaces waves. A total of 10 HSTs with speeds evenly ranging from 250 to 350 km s^{-1} are simulated. Figs 12(a)–(c) show the resulting virtual shot gathers by using vibrations of 1 HST, 5 HSTs and 10 HSTs, respectively. As expected, the more HSTs used for seismic interferometry, the less crosstalk is generated. Compared with the directly computed surface waves shown in Fig. 12(d), 5 HSTs seem to be sufficient to generate useful surface waves after seismic interferometry.

Then we use the extracted surface waves for dispersion imaging. Figs 13(a)–(c) show the phase-velocity maps calculated by the phase shift method (Park *et al.* 1998) using vibrations induced by 1 HST, 5 HSTs and 10 HSTs, respectively. As expected, the more HSTs used for seismic interferometry, the higher SNR in the final dispersion image. Fig. 13(d) shows the dispersion image using the directly computed surface waves. By overlaying with the theoretical dispersion curve computed by the stiffness matrix method (Kausel & Roësset 1981), we can observe that, for this numerical example, 5 HSTs seem to be sufficient to provide a reliable dispersion relationship that could be used for the dispersion inversion afterwards.

After extracting the dispersion curve from the dispersion image, the shear wave velocity model can be estimated using different methods, such as iterative algorithm (Xia *et al.* 1999), genetic algorithm (Dal Moro *et al.* 2007), and improved neighbourhood algorithm (Wathelet 2008). Fig. 14(a) shows the estimated 1-D shear wave velocity model using the improved neighbourhood algorithm. The estimated result is consistent with the true model described in Table 1. Fig. 14(b) compares the calculated and observed dispersion curves overlaid on the dispersion image. We can observe that the dispersion curve calculated from the inverted velocity model matches very well with the dispersion curve calculated from the true velocity model.

3.3 Field data tests

Finally, we use field data to demonstrate that seismic interferometry is able to retrieve surface waves from the HST-induced vibrations. 23 seismographs (indicated by red triangles in Fig. 15a) with an interval of 4 m are deployed parallelly to a busy section of high-speed railway near Xi'an, China. This section of railway is built on a viaduct with bridge piers that are evenly distributed at an interval of about 25 m, as indicated by yellow pentagrams in Fig. 15(a). The seismographs acquire data continuously for over half an hour. A sampling rate of 0.002 s is used, corresponding to a Nyquist frequency as high as 250 Hz.

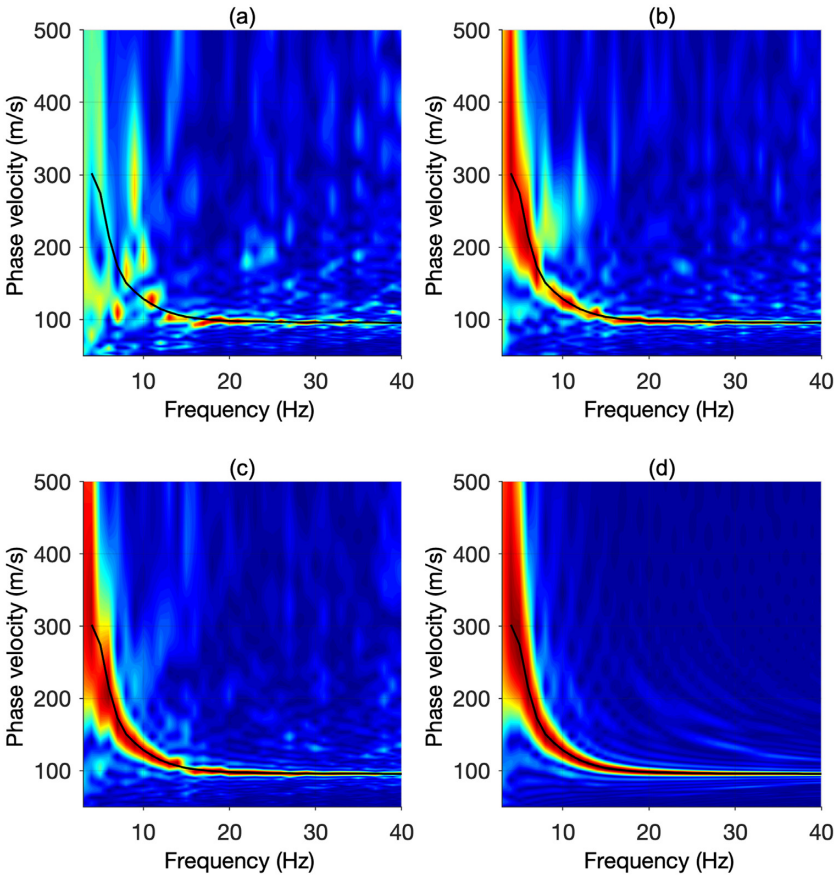

Figure 13. Dispersion images using vibrations of (a) 1 HST, (b) 5 HSTs and (c) 10 HSTs. (d) Dispersion image using directly computed surface waves. The black lines denote the dispersion curves calculated from the true model.

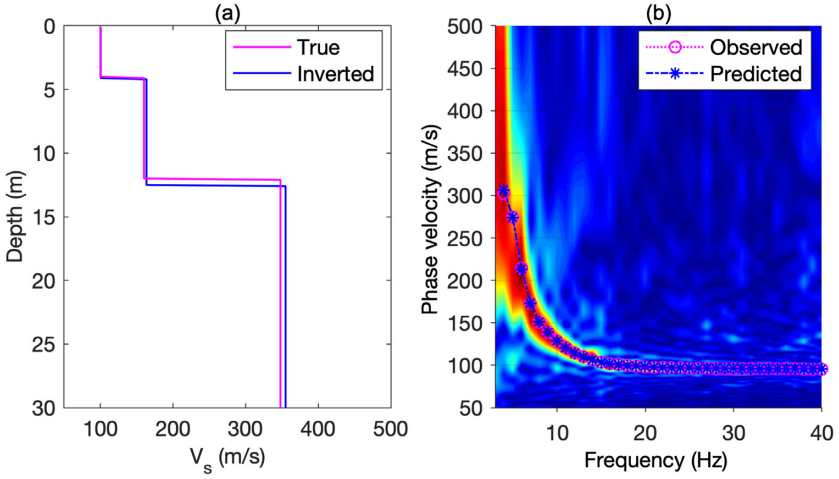

Figure 14. (a) 1-D shear wave velocity models. The pink and blue lines denote the true and inverted models, respectively. (b) Comparisons of the observed (pink) and predicted (blue) dispersion curves overlaid on the dispersion image.

While inspecting the raw records shown in Fig. 15(b), we can easily identify six HST-induced events based on the amplitude of the vibrations at each station and the velocity of the signal as it traverses the array. More specifically, the first, second, fourth and sixth events are generated by HSTs that run towards westbound at speeds of 71.8, 68.9, 69.2, and 72.6 m s^{-1}, respectively, and the third and fifth events are generated by HSTs that run towards eastbound at speeds of 58.3 and 77.1 m s^{-1}, respectively. After zooming in the third event, as shown in Fig. 15(c), we can observe that the HST-induced vibrations are characterized as being the strongest, long-durational and periodical. In addition to the waveforms with strong energy that are train vibrations, there are two series of parallel waveforms before and after the train vibrations. According to the slopes of these two series of events, the earlier series are mainly surface waves propagated to the seismographs when the HST is approaching; whereas the latter series are mainly surface waves generated when the HST is leaving the acquisition array.

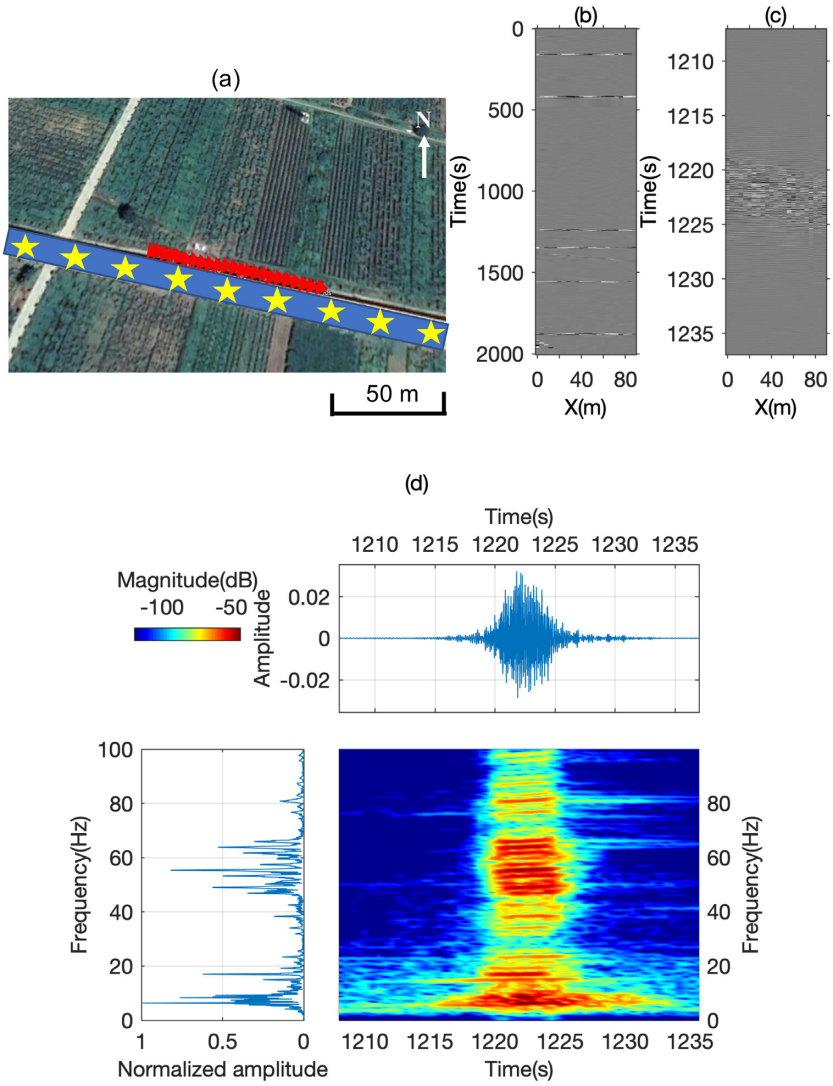

Figure 15. Field acquisition geometry and raw seismograms. (a) Locations of 23 seismographs (red triangles) with an interval of 4 m distributed paralleling the high-speed railway (blue line). The railway is built on a viaduct with bridge piers (yellow pentagrams) that are evenly distributed at an interval of about 25 m. (b) Full seismic records that contain 6 events. Two events (the third and fifth events) that dip down to the right are induced by trains that move toward the east, whereas four events (the first, second, fourth and sixth events) that dip down to the left are induced by trains that move towards the west. (c) A closeup of the third event. (d) Time–frequency analysis panels of the trace at $x = 56$ m in (c). The top-right, bottom-left and bottom-right panels in (d) show the seismogram, spectrum, and spectrogram, respectively.

Fig. 15(d) shows the seismogram (top right), spectrum (bottom left), and spectrogram (bottom right) for the trace at $x = 56$ m in Fig. 15(c). The spectrum shows substantial energy from 4 to 80 Hz and shows striking features of sharp equidistant spectral lines in the entire frequency band. As mentioned in Section 2, the spacing of these spectral lines mainly depends on the HST speed and the carriage length. In addition, we can observe the Doppler effect from the spectrogram due to the moving property of the HST source. Thus, it is challenging to utilize HST-induced vibrations directly for subsurface characterization.

Seismic interferometry is then applied to reconstruct the impulse responses between stations from the raw records. Figs 16(a) and (b) show the interferometric result after cross-correlating the first trace with other traces by using 1 HST and 6 HSTs, respectively. As expected, the more HSTs being used in seismic interferometry, the higher SNR the retrieved surface waves have. In particular, as indicated by the white arrows in Fig. 16, the crosstalk in the interferograms have been attenuated after stacking the interferometric results of several varying speed trains. In addition, we can observe that the surface waves present dominant energy propagating linearly with an apparent velocity of 240 m s^{-1}, which is consistent with the surface wave velocity measured from the raw records.

Then, we use the extracted surface waves for dispersion imaging. Figs 17(a) and (b) show the phase-velocity maps calculated by the phase shift method using vibrations induced by 1 HST and 6 HSTs, respectively. A comparison between these two phase-velocity maps, especially in the areas indicated by white circles in Fig. 17, suggests that the crosstalk generated during the surface wave retrieval is effectively attenuated by stacking the interferometric results using vibrations of several HSTs, resulting in a phase-velocity map with more coherent spectrum.

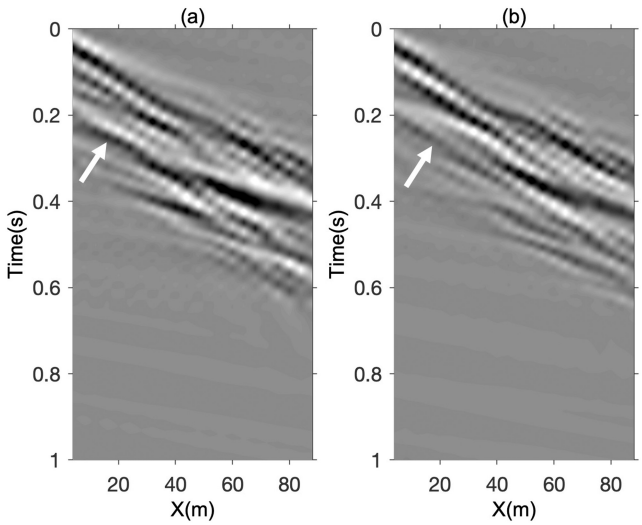

Figure 16. Surface waves retrieved using vibrations of (a) 1 HST and (b) 6 HSTs. The white arrows indicate the crosstalk that has been attenuated after stacking the interferometric results of several varying speed HSTs.

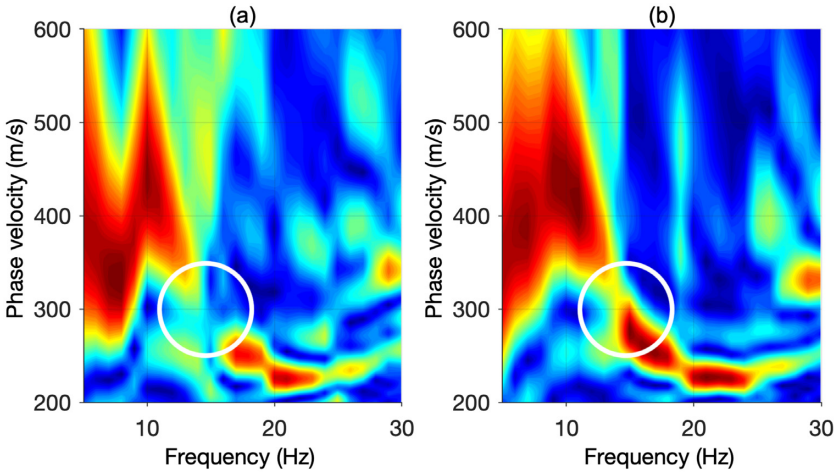

Figure 17. Phase-velocity maps of the retrieved surface waves using vibrations of (a) 1 HST and (c) 6 HSTs. The white circles indicate the area where the phase-velocity map has been improved after stacking the interferometric results of several varying speed HSTs.

Afterwards, dispersion inversion is used to estimate the near-surface shear wave velocity model from the phase-velocity map of the retrieved surface waves. The first step is picking a dispersion curve from the phase-velocity map. The picking result is denoted as the pink line in Fig. 18(a). Inversion of the picked dispersion curve within the frequency range 4–30 Hz results in a 1-D shear wave velocity model, as denoted by the blue line in Fig. 18(b). A comparison of the picked (pink) and predicted (blue) dispersion curves demonstrates that the inverted shear wave velocity model achieves small data mismatch after dispersion inversion, as shown in Fig. 18(a). It is worth noting that the surface wave dispersion inversion problem is highly nonlinear and is affected by solution non-uniqueness, the inversion conducted here is used to demonstrate that stacking interferograms of varying speed HSTs provides reliable dispersion image that could be used for the following dispersion inversion.

4 DISCUSSIONS

In this paper, we analyse the effects of HST traffic characteristics on seismic interferometry. This analysis starts from an explicit formula of HST source function that has deterministic properties relevant to the HST parameters. However, the HST source function in practice also has randomly oscillatory behaviour that should be taken into account as well during the derivations. This property will further reduce the cross terms, thus making the interferometric results less contaminated by the crosstalk. Since the effect of the random property on seismic interferometry is similar to its effect on conventional passive seismic interferometry, here we only investigate the influence of the deterministic properties of HST-induced vibrations on seismic interferometry.

In our derivations, we assume that the HST railway is straight. This is a reasonable assumption for the application of near-surface characterization in sites with weak lateral heterogeneities, because only a short receiver array is needed for the estimation of a 1-D shallow

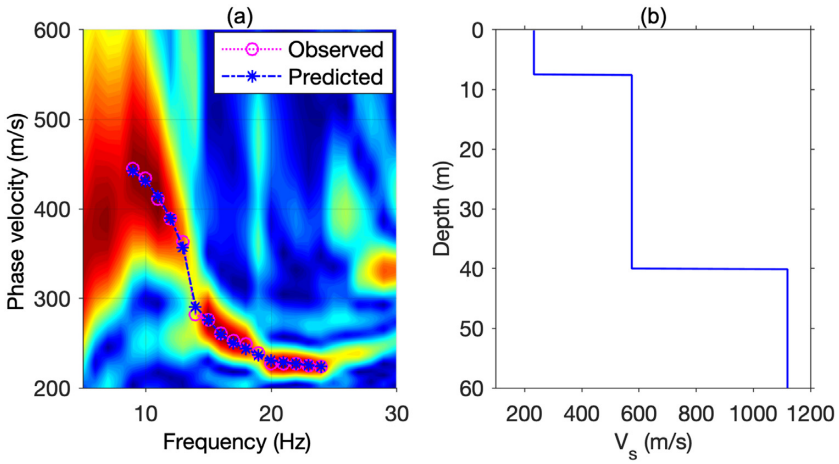

Figure 18. Surface wave dispersion inversion. (a) Phase-velocity map overlaid by the picked (pink) and predicted (blue) dispersion curves. (b) 1-D shear wave velocities obtained by dispersion inversion.

velocity profile. Whereas, for other applications in local scales with strong lateral heterogeneities, we should take the railway geometry into account during the derivation of seismic interferometry. Even though it has been demonstrated that seismic waves could be retrieved from train noise if the receiver array is not parallel to the railway (Brenguier *et al.* 2019b; Zhou & Paulssen 2020), little attention is paid on the effects of train traffic characteristics on seismic interferometry. Work is underway to fully investigate the influence of the deterministic properties of train-induced vibrations on seismic interferometry for a more general acquisition configuration.

5 CONCLUSIONS

Prior work has shown successful applications of train traffic noise for subsurface imaging with seismic interferometry. However, the influence of the train traffic characteristics on the seismic interferometry method has not been well studied. In this paper, we explicitly formulate the source function generated by a moving HST and present an alternative derivation of seismic interferometry using the HST-induced vibrations. In addition, we analyse the effects of several HST-related factors, including HST moving direction, HST carriage number, source locations, HST speed and stacking number of HST-induced vibration records, on the retrieval of seismic waves. We find that seismic interferometry using HST traffic noise generates strong crosstalk that could be effectively attenuated by the stacking of several vibration records generated by HSTs running with slightly different speeds. This is the first study, to our knowledge, that investigates the influence of train traffic characteristics on seismic interferometry. Most notably, our field data results provide compelling evidence for successful retrieval of surface waves from HST traffic noise. Moreover, the retrieved surface waves were used to estimate the near-surface velocity model. For long-time high-resolution near-surface monitoring, future work should be dedicated to recording dense, long-offset and continuous seismic data. This could be achieved by using distributed acoustic sensing acquisition method. HST traffic noise together with new advanced acquisition methods have a strong potential for monitoring temporal changes of the earth's near surface.

ACKNOWLEDGEMENTS

The authors thank Shoudong Huo from Chinese Academy of Science, Xiaokai Wang from Xi'an Jiaotong University and Jieyuan Ning from Peking University for their support on the field data acquisition. The authors also thank Lu Liu, Yi He, Tao Li and Yue Du from Aramco Beijing Research Center, Aramco Asia, for their help during the field data acquisition. In addition, the first author wants to thank his former colleagues Yan Wu and Bowen Guo for fruitful discussions about the theory of train-induced vibrations. Last but not least, the authors are grateful to the Editor Herve Chauris, Wen Zhou and an anonymous reviewer for their insightful criticism that helped to improve the manuscript.

DATA AVAILABILITY

The data underlying this paper will be shared on reasonable request to the corresponding author.

REFERENCES

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

Brenguier, F. *et al.*, 2019a. Train traffic as a powerful noise source for monitoring active faults with seismic interferometry, *Geophys. Res. Lett.,* **46**(16), 9529–9536.

Brenguier, F. *et al.*, 2019b. Noise-based ballistic wave passive seismic monitoring. Part 1: body waves, *Geophys. J. Int.,* **221,** 683–691.

Campillo, M. & Paul, A., 2003. Long range correlations in the diffuse seismic coda, *Science,* **299**(5606), 547–549.
Chen, Q., Chen, L. & Chen, Y., 2004. Seismic features of vibration induced by train, *Acta Seismol. Sin.,* **17**(6), 715–724.
Dal Moro, G., Pipan, M. & Gabrielli, P., 2007. Rayleigh wave dispersion curve inversion via genetic algorithms and Marginal Posterior Probability Density estimation, *J. Appl. Geophys.,* **61**(1), 39–55.
Denolle, M.A., Dunham, E.M., Prieto, G.A. & Beroza, G.C., 2014. Strong ground motion prediction using virtual earthquakes, *Science,* **343**(6169), 399–403.
Fuchs, F. & Bokelmann, G., 2018. Equidistant spectral lines in train vibrations, *Seismol. Res. Lett.,* **89**(1), 56–66.
Gerstoft, P., Fehler, M.C. & Sabra, K.G., 2006. When Katrina hit California, *Geophys. Res. Lett.,* **33**(17), L17308–L17308.
He, X., Wu, T., Zou, Y., Chen, Y.F., Guo, H. & Yu, Z., 2017. Recent developments of high-speed railway bridges in China, *Struct. Infrastruct. Eng.,* **13**(12), 1584–1595.
Hung, H.H. & Yang, Y.B., 2001. Elastic waves in visco-elastic half-space generated by various vehicle loads, *Soil Dyn. Earthq. Eng.,* **21**(1), 1–17.
Kausel, E. & Roësset, J.M., 1981. Stiffness matrices for layered soils, *Bull. seism. Soc. Am.,* **71**(6), 1743–1761.
Krylov, V.V., 1994. On the theory of railway-induced ground vibrations, *J. Phys.,* **4**(C5), 769–772.
Liu, Y., Yue, Y., Li, Y. & Luo, Y., 2021a. On the retrievability of seismic waves from high-speed-train-induced vibrations using seismic interferometry, *IEEE Geosci. Remote Sens. Lett.,* 1–5, doi:10.1109/LGRS.2021.3050205 .
Liu, Y., Yue, Y., Luo, Y. & Li, Y., 2021b. The seismic broad-band signature of high-speed trains running on viaducts, *Geophys. J. Int.,* **226**(2), 884–892.
Nakata, N., Snieder, R., Larner, K., Tsuji, T. & Matsuoka, T., 2011. Shear-wave imaging from traffic noise using seismic interferometry by cross-coherence, *Geophysics,* **76**(6), SA97–SA106.
Nakata, N., Chang, J.P., Lawrence, J.F. & Boué, P., 2015. Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry, *J. geophys. Res.,* **120**(2), 1159–1173.
Park, C.B., Miller, R.D., & Xia, J., 1998. Imaging dispersion curves of surface waves on multi-channel record, in *SEG Technical Program Expanded Abstracts,* SEG, pp.1377–1380.
Quiros, D.A., Brown, L.D. & Kim, D., 2016. Seismic interferometry of railroad induced ground motions: body and surface wave imaging, *Geophys. J. Int.,* **205**(1), 301–313.
Riahi, N., Bokelmann, G., Sala, P. & Saenger, E.H., 2013. Time-lapse analysis of ambient surface wave anisotropy: a three-component array study above an underground gas storage, *J. geophys. Res.,* **118**(10), 5339–5351.
Roux, P., Sabra, K.G., Gerstoft, P., Kuperman, W.A. & Fehler, M.C., 2005. P-waves from cross-correlation of seismic noise, *Geophys. Res. Lett.,* **32**(19), 1–4.
Sabra, K.G., Gerstoft, P., Roux, P., Kuperman, W.A. & Fehler, M.C., 2005. Extracting time-domain Green's function estimates from ambient seismic noise, *Geophys. Res. Lett.,* **32**(3), 1–5.
Schuster, G.T., 2009. *Seismic Interferometry,* Cambridge Univ. Press.
Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.,* **31**(7), 8–11.
Sneider, R., 1986. 3-D linearized scattering of surface waves and a formalism for surface wave holography, *Geophys. J. R. astr. Soc,* **84**(3), 581–605.
Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase, *Phys. Rev. E,* **69**(4), 8.
Tromp, J., Komatitsch, D. & Liu, Q., 2008. Spectral-element and adjoint methods in seismology, *Commun. Comput. Phys.,* **3**(1), 1–32.
Wang, L., Liu, Y., Mao, L., & Sun, C., 2018. Potential impacts of China 2030 high-speed rail network on ground transportation accessibility. *Sustainability,* **10**(4), 1–16.
Wang, X., Wang, B. & Chen, W., 2021. The second-order synchrosqueezing continuous wavelet transform and its application in the high-speed-train induced seismic signal, *IEEE Geosci. Remote Sens. Lett.,* **18**(6), 1109–1113.
Wathelet, M., 2008. An improved neighborhood algorithm: parameter conditions and dynamic scaling, *Geophys. Res. Lett.,* **35**(9), 1–5.
Weemstra, C., Boschi, L., Goertz, A. & Artman, B., 2013. Seismic attenuation from recordings of ambient noise, *Geophysics,* **78**(1), Q1–Q14.
Xia, J., Miller, R.D. & Park, C.B., 1999. Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves, *Geophysics,* **64**(3), 691–700.
Zhang, H., Wang, B., Ning, J. & Li, Y., 2019a. Interferometric imaging using high-speed-train induced seismic waves, *Chin. J. Geophys.,* **62**(6), 2321–2327. (in Chinses)
Zhang, Y., Li, Y.E., Zhang, H. & Ku, T., 2019b. Near-surface site investigation by seismic interferometry using urban traffic noise in Singapore, *Geophysics,* **84**(2), B169–B180.
Zhou, W. & Paulssen, H., 2020. Compaction of the Groningen gas reservoir investigated with train noise, *Geophys. J. Int.,* **223**(2), 1327–1337.

APPENDIX A. STATIONARY PHASE EVALUATION OF INTERFEROMETRIC INTEGRAL FOR SCALAR WAVES IN A 3-D MEDIUM

In this appendix, we show the details of the stationary phase analysis required to solve the interferometric integral for scalar waves. This analysis follows Snieder (2004).

Let the autocorrelation of the source signal $w_s(t)$ be denoted by

$$c_s\ (\tau) = \int_0^T w_s\,(t)\, w_s\,(t+\tau)\,\mathrm{d}t. \tag{A1}$$

Then eq. (8) in the main text can be rewritten as

$$c\ (\tau) = \sum_s \frac{c_s\left(\tau + \frac{|x_s|-|R-x_s|}{c}\right)}{|x_s\,||R-x_s|}\ . \tag{A2}$$

Since the Fourier transform of the cross-correlation is equal to the power spectrum, eq. (A2) can be expressed in the frequency domain as

$$C\ (\omega) = \sum_s \frac{|W_s\,(\omega)|^2 \exp\left[\frac{i\omega(|R-x_s|-|x_s|)}{c}\right]}{|x_s\,||R-x_s|}. \tag{A3}$$

The power spectrum $|W_s(\omega)|^2$ does not depend on the phase fluctuations of seismic waves but depends on the amplitude fluctuations. After applying spectral whitening to the HST-induced seismic data, the variations in the power spectrum could be considered to be uncorrelated

with the phase $\exp[\frac{i\omega(|R-x_s|-|x_s|)}{c}]$, then

$$C\ (\omega) = \overline{|W_s\,(\omega)|^2}\ \sum_s \frac{\exp\left[\frac{i\omega(|R-x_s|-|x_s|)}{c}\right]}{|x_s\,||R-x_s|}, \tag{A4}$$

with

$$\overline{|W_s\,(\omega)|^2} = \frac{1}{N_s}\ \sum_s |W_s\,(\omega)|^2, \tag{A5}$$

where N_s is the number of sources.

When there are many sources per wavelength, the summation over sources can be replaced by a linear integration that is weighted by the source density n defined as the number of sources per unit line. In this approximation, eq. (A4) is given by

$$C\ (\omega) = \overline{|W_s\,(\omega)|^2}\ \int_{-\infty}^{+\infty} \frac{\exp\left[\frac{i\omega(|R-x_s|-|x_s|)}{c}\right]}{|x_s\,||R-x_s|} n\mathrm{d}x_s. \tag{A6}$$

The above integral can be divided into three parts according to the region of the source location, that is

$$C\ (\omega) = \overline{|W_s\,(\omega)|^2}\ \left(e^{ikR}\int_{-\infty}^{0}\frac{1}{x_s\,(x_s-R)}n\mathrm{d}x_s + e^{-ikR}\int_{R}^{+\infty}\frac{1}{x_s\,(x_s-R)}n\mathrm{d}x_s + \int_0^R \frac{\exp\left[ik\,(R_s-2x_s)\right]}{x_s\,(R-x_s)}n\mathrm{d}x_s\right). \tag{A7}$$

Because the last summation is oscillatory, the region $0 < x_s < R$ gives a subdominant contribution to the final cross-correlation. Ignoring this contribution results in eq. (9) in the main text

$$C\ (\omega) = \overline{|W_s\,(\omega)|^2}\ \left(\frac{e^{ikR}}{4\pi\,R}\int_{-\infty}^{0}\frac{\mathrm{R}}{x_s\,(x_s-R)}n\mathrm{d}x_s + \frac{e^{-ikR}}{4\pi\,R}\int_{R}^{+\infty}\frac{R}{x_s\,(x_s-R)}n\mathrm{d}x_s\right). \tag{A8}$$

APPENDIX B. STATIONARY PHASE EVALUATION OF INTERFEROMETRIC INTEGRAL FOR SURFACE WAVES IN AN ELASTIC MEDIUM

In this appendix, we show the details of the stationary phase analysis required to solve the interferometric integral for surface waves. This analysis follows Snieder (2004) as well.

In the main text, we assume that sufficient event averaging is carried out so that the cross terms in the cross-correlation can be ignored, then the cross-correlation becomes

$$C_{i,j}\ (\omega) = \sum_s \sum_{m,m'} p_i^m\,(z_2,\varphi_2)\,p_j^{m*}\,(z_1,\varphi_1)\,\frac{e^{i\left(k_m|R-x_s|-k_{m'}|x_s|\right)}}{\sqrt{k_m k_{m'}\,|R-x_s|\,|x_s|}}\,W_s^m\,(\omega)\,W_s^{m'*}\,(\omega)\ . \tag{B1}$$

Similar to the derivation for scalar-wave interferometry, when there are many sources per wavelength, the summation over sources can be replaced by a linear integration that is weighted by the source density n, that is

$$C_{i,j}\ (\omega) = \sum_{m,m'}\int_{-\infty}^{+\infty}\mathrm{d}x_s n p_i^m\,(z_2,\varphi_2)\,p_j^{m'*}\,(z_1,\varphi_1)\,\frac{e^{i\left(k_m|R-x_s|-k_{m'}|x_s|\right)}}{\sqrt{k_m k_{m'}\,|R-x_s|\,|x_s|}}\,W_s^m\,(\omega)\,W_s^{m'*}\,(\omega)\ . \tag{B2}$$

The summation over surface wave modes in eq. (B2) is oscillatory, except when $k_m = k_{m'}$. This means that the dominant contribution comes from the terms $m = m'$, and for this reason, the mode coupling terms $m \neq m'$ can be ignored. Thus eq. (B2) reduces to

$$C_{i,j}\ (\omega) = \sum_{m}\int_{-\infty}^{+\infty}\mathrm{d}x_s n p_i^m\,(z_2,\varphi_2)\,p_j^{m*}\,(z_1,\varphi_1)\,\frac{e^{i(k_m|R-x_s|-k_m|x_s|)}}{\sqrt{k_m k_m\,|R-x_s|\,|x_s|}}\,\overline{|W^m\,(\omega)|^2}\ , \tag{B3}$$

with $\overline{|W^m(\omega)|^2}$ the average power spectrum of the radiated mode m. The above integral can be divided into three parts according to the region of the source location. Furthermore, the integrand is oscillatory over the range $0 < x < R$, and the dominant contribution comes from the regions $x < 0$ and $x > R$. Taking these approximations gives

$$C_{ij}\ (\omega) = \sqrt{\frac{2}{\pi}}\ \sum_m \left\{\frac{1}{ik_m}p_i^m\,(z_2,+)\,p_j^{m*}\,(z_1,+)\,\frac{e^{i\left(k_m R+\frac{\pi}{4}\right)}}{\sqrt{\frac{\pi}{2}k_m R}}\int_{-\infty}^{0} n\mathrm{d}x + \frac{1}{ik_m}p_i^m\,(z_2,-)\,p_j^{m*}\,(z_1,-)\,\frac{e^{-i\left(k_m R+\frac{\pi}{4}\right)}}{\sqrt{\frac{\pi}{2}k_m R}}\int_{R}^{\infty} n\mathrm{d}x\right\}\overline{|W^m\,(\omega)|^2}. \tag{B4}$$

In eq. (B4), the first term is due to right-going waves that are generated in the region $x < 0$; the polarization vectors correspond to the azimuth $\varphi = 0$, which is indicated by the plus signs. The second term is due to waves propagated from the location $x > R$ that move toward the left; their polarization vectors correspond to the azimuth $\varphi = \pi$, which is indicated by the minus signs.

A comparison with the surface wave Green's tensor of mode m in the far field (Sneider 1986)

$$G_{ij}^{m}(\mathbf{r},\mathbf{r}_s) = p_i^m(z,\varphi)\, p_j^{m^*}(z_0,\varphi)\, \frac{e^{i\left(k_m R+\frac{\pi}{4}\right)}}{\sqrt{\pi k_m R/2}} \tag{B5}$$

shows that, if we ignore the integral parts in eq. (B4), the first term is equal to $(c_m/i\omega)G_{ij}^{m}(\mathbf{r}_2,\mathbf{r}_1)$, while the second one equals $[(c_m/i\omega)G_{ij}^{m}(\mathbf{r}_2,\mathbf{r}_1)]^*$. The correlation tensor therefore becomes eq. (16) in the main text

$$C_{ij}\ (\omega) = \sqrt{\frac{2}{\pi}} \sum_m c_m \left\{ \frac{G_{ij}^{m}(\mathbf{r}_2,\mathbf{r}_1)}{i\omega} \int_{-\infty}^{0} n\mathrm{d}x + \left[\frac{G_{ij}^{m}(\mathbf{r}_1,\mathbf{r}_2)}{i\omega}\right]^* \int_{R}^{\infty} n\mathrm{d}x \right\} \overline{|W^m(\omega)|^2}. \tag{B6}$$