Article

Seismic Interferometry from Correlated Noise Sources

Daniella Ayala-Garcia [1,*], Andrew Curtis [2] and Michal Branicki [1,3]

1   School of Mathematics, The University of Edinburgh, Edinburgh EH9 3FE, UK; m.branicki@ed.ac.uk
2   School of GeoSciences, The University of Edinburgh, Edinburgh EH9 3FD, UK; Andrew.Curtis@ed.ac.uk
3   The Alan Turing Institute for Data Science, London NW1 2DB, UK
*   Correspondence: c.d.ayala-garcia@sms.ed.ac.uk

Abstract: It is a well-established principle that cross-correlating seismic observations at different receiver locations can yield estimates of band-limited inter-receiver Green's functions. This principle, known as Green's function retrieval or seismic interferometry, is a powerful technique that can transform noise into signals which enable remote interrogation and imaging of the Earth's subsurface. In practice it is often necessary and even desirable to rely on noise already present in the environment. Theory that underpins many applications of ambient noise interferometry assumes that the sources of noise are uncorrelated in time. However, many real-world noise sources such as trains, highway traffic and ocean waves are inherently correlated in space and time, in direct contradiction to the these theoretical foundations. Applying standard interferometric techniques to recordings from correlated energy sources makes the Green's function liable to estimation errors that so far have not been fully accounted for theoretically nor in practice. We show that these errors are significant for common noise sources, always perturbing or entirely obscuring the phase one wishes to retrieve. Our analysis explains why stacking may reduce the phase errors, but also shows that in commonly encountered circumstances stacking will not remediate the problem. This analytical insight allowed us to develop a novel workflow that significantly mitigates effects arising from the use of correlated noise sources. Our methodology can be used in conjunction with already existing approaches, and improves results from both correlated and uncorrelated ambient noise. Hence, we expect it to be widely applicable in ambient noise studies.

Keywords: ambient noise; interferometry; correlated noise; error quantification; seismology; traffic; trains; cars; processing

1. Introduction

The critical zone and Earth's crust are constantly monitored across ecological and geological disciplines due to their importance to terrestrial life, for industrial applications and for advancing science. Remote sensing methods are commonly deployed, in which physical energy fields are recorded on or above the Earth's surface, and are used to infer structure and properties of the subsurface. In particular, seismic interferometry is a powerful technique which transforms previously discarded data, such as seismic energy from earthquakes or from the ambient background noise field, into useful signals that remotely illuminate subsurface Earth structures [1–3]. The origin of seismic interferometry can be traced back to the seminal work of Claerbout [4] who showed that the reflection response of a horizontally layered medium could be estimated from the autocorrelation of its transmission response. Seismic, or more generally wavefield interferometry has since become a rapidly evolving field of research [5–8], leading to fundamental advances in our ability to image the Earth's crust at global [9,10], regional [11–13] and industrially relevant scales [14–16].

In general, interferometric methods rely on cross-correlating, convolving or deconvolving pairs of recorded signals to extract information about the medium. The goal is to estimate signals which would have been acquired if the receivers [17,18] or sources [1,2,19]

been deployed at different locations or at different times [20–23]. Theoretically, the response of a medium at a given location to an impulsive source at a different location is described by Green's functions which are associated with specific equations describing the wave dynamics. The purpose of seismic interferometry is usually to estimate various approximations to these Green's functions.

Two theoretical assumptions that underpin many interferometric methods are that recorded energy comes from sources that are distributed isotropically in space around the receivers, and that the time series emitted by the sources are statistically uncorrelated in time between pairs of sources. The latter assumption precludes sources that are spatio-temporally correlated. However, it is often necessary or even desirable to rely on noise sources that are already present in the environment. This is especially true in areas of the Earth where environmental constraints preclude the use of active artificial seismic sources, or when we wish to illuminate large volumes of the Earth that require more powerful sources than can be generated artificially. Sources of freely available ambient noise abound in the Earth and interferometry has been successfully performed by cross-correlating earthquake codas (the long tail of energy that is recorded after the initial impulses from first-arriving seismic waves) which are assumed to act as an approximately diffuse, reverberating wavefield approaching receivers from all directions [1], or using recorded wavefields assumed to come from isotropic noise fields [2,3,24,25]. Many known physical noise sources that contribute to ambient noise are in motion, e.g., storm sources, ocean waves or wind [26], in addition to seismic energy deriving from man-made activities such as shipping [27] or noise from traffic. Recently it has been recognized that highway and/or railway traffic can comprise a dominant component of the ambient noise field, and interferometry has been applied to highway traffic noise [28–30], railway noise [31–35], and noise generated from waves breaking along coastlines [36], as illustrated in Figure 1. However, all these sources of ambient noise are inherently correlated in time, which is in direct contradiction to the theoretical assumptions outlined above. Although different methodologies for mitigating non-ideal source distributions have been considered, for example in [37–40], the effects of statistical correlation in ambient noise sources were not considered, particularly in the important case of correlation induced by motion of the source. The impact that source motion has on cross-correlational interferometry was investigated in [41] but under the standard assumption of statistically uncorrelated ambient noise. Importantly, the extent of the error in the interferometric estimates of the Green's function due to the assumption of uncorrelated noise sources has so far not been quantified, and no general methods to reduce these errors have been published.
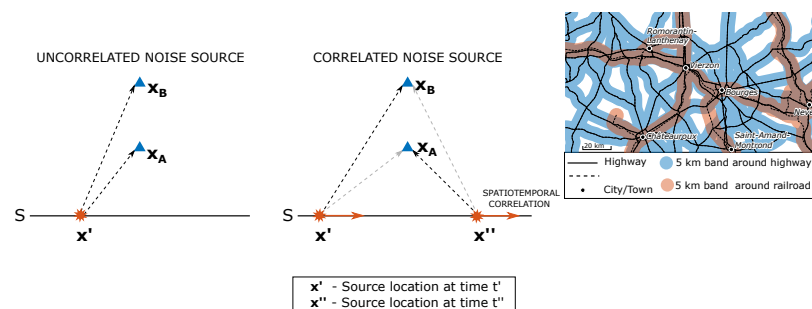
Figure 1. Geometry of a boundary *S* of uncorrelated, impulsive sources (see Equation (1)) versus a time-ordered train of impulsive sources (see Equation (4)). Blue triangles denote receivers and source locations are represented by orange stars. Orange arrows in the source boundary indicate the presence of a spatiotemporal correlation of sources along the boundary. The top-right panel shows the main roads (solid black) and railways (black dashes) within an area of about 9000 km^2 around Bourges, France. A zone of 5 km on either side of each road (blue) and railway (orange) highlights the prevalence of near-field traffic noise. The bottom panel shows oblique waves breaking against the coast of Skeleton Bay, Namibia. These are common examples of correlated noise sources.

In this paper, we present a unified theory of inter-receiver seismic interferometry that encompasses correlated and uncorrelated noise sources in both the near and far fields, including the case of moving sources. Our theoretical framework allows us to develop a novel workflow that mitigates the spurious effects arising from the use of correlated noise sources, leading to the retrieval of the Green's function from a short-time recording of a single correlated moving noise source. Moreover, our framework provides a systematic insight into the mechanisms through which the correlation in the sources induces errors in the estimates of the Green's function and its phase. We quantify these errors and show that they have a significant effect for commonly used noise sources, with spurious effects that obscure the estimates of phase (and hence of wave travel times). Our novel method is based on an appropriate randomization of the recorded traces which allows for an accurate interferometric retrieval even from a single moving energy source. Furthermore, our analysis explains why stacking multiple traces may reduce errors due to noise correlation in the interferometric estimates, but also highlights limitations of this approach, and identifies potentially commonly occurring circumstances in which it will fail. Our unified methodology is applicable to both correlated and uncorrelated ambient noise, and is particularly useful in cases where stacking does not improve the signal to noise ratio in the retrieval.

In what follows we begin by briefly presenting the theoretical principles of standard seismic interferometry for the case when individual recordings from static impulsive sources are available, as well as for an ambient noise scenario. Subsequently, we highlight the challenges in retrieving the interferometric inter-receiver Green's function from ambient noise. The important case of performing interferometric retrieval from wavefields generated by correlated moving sources is considered next, and the corrupting effect due to the Doppler spread on the quality of the interferometric retrieval is discussed. We then review the standard ensemble average approach in ambient noise interferometry, which is referred to as *stacking*. Potential pitfalls of the retrieval through the stacking average are discussed in terms of the correlation structure of the source signature. Moreover, an analytical decomposition of the ambient noise retrieval is derived which, based on the characteristic correlation length of the source signature, separates the retrieval into a term comprising predominantly coherent contributions that correctly approximate the inter-receiver phase, and the crosstalk term associated with spurious incoherent contributions. These analytical findings motivate the development of a novel and versatile technique, referred to as the *random windowing* approach. In contrast to the standard stacking average, the random windowing average does not rely on spatial localization of the correlation structure of the source signature, and it can be used for interferometric phase retrieval from short recordings of a correlated noise signal. We conclude by outlining a pseudo-algorithm for carrying out the random windowing retrieval and present an example with the resulting phases retrieved successfully from both correlated and uncorrelated ambient noise.

2. Theoretical Background

It is a well-established principle that cross-correlating seismic observations recorded at different receiver locations yields seismic responses that under certain conditions, provide useful estimates of the inter-receiver Green's function of linear wave equations. This principle has been expressed in a variety of ways, often via the acoustic or elastic representation theorems, and using either the convolution, deconvolution or cross-correlation of recorded seismic traces. For our purposes, a cross-correlational Green's function representation theorem is used in line with [2,42], in a regime of volume-injection rate impulsive sources.

Consider a medium of density ρ and sound speed c, and let $\hat{G}(\mathbf{x}, \mathbf{x}', \omega)$ denote the frequency-domain response of the medium to an impulsive source located at $\mathbf{x}'$ and recorded at location $\mathbf{x}$. Throughout, $\hat{\cdot}$ denotes quantities in the frequency domain, and $\cdot^*$ denotes complex conjugation. Given a pair of receiver locations $\mathbf{x_A}$ and $\mathbf{x_B}$ we assume an empirical estimate $\hat{g}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ of the inter-receiver Green's function $\hat{G}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ can be represented as

$$\hat{g}(\mathbf{x_A}, \mathbf{x_B}, \omega) = \frac{2}{\rho c} \int_S \hat{G}^*(\mathbf{x_A}, \mathbf{x}', \omega) \hat{G}(\mathbf{x_B}, \mathbf{x}', \omega) d\mathbf{x}', \tag{1}$$

as in Wapenaar and Fokkema [2], where each $\mathbf{x}'$ is an impulsive source location along the boundary S (see leftmost panel in Figure 1). The integrand in Equation (1) is equivalent to the convolution of the time-reversed trace recorded at $\mathbf{x_A}$ with the trace recorded at $\mathbf{x_B}$, and this operation corresponds to a cross-correlation in the time domain (this statement holds exactly for recordings $G(\mathbf{x}, \mathbf{x}', t)$ defined for all time). Please note that $\hat{g}$ in Equation (1) is proportional to the spatial average of the integrand in the right-hand side of Equation (1) over S; this interpretation will be helpful below.

The main theoretical requirements for Equation (1) to provide a good estimate of the inter-receiver Green's function are: (i) that the medium is lossless, (ii) that the source boundary S encloses the receivers and that energy is emitted equally from all directions, (iii) that S is sufficiently far from the receivers for the recorded energy flux to be emitted approximately perpendicular to the boundary, and (iv) that individual recordings are acquired independently for each impulsive source location $\mathbf{x}'$ within S. Under these idealized conditions, the empirical estimate $\hat{g}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ provides a good approximation to the homogeneous inter-receiver Green's function [2], i.e., $\hat{g}(\mathbf{x_A}, \mathbf{x_B}, \omega) = \hat{G}(\mathbf{x_A}, \mathbf{x_B}, \omega) - \hat{G}^*(\mathbf{x_A}, \mathbf{x_B}, \omega)$ or, equivalently $g(\mathbf{x_A}, \mathbf{x_B}, t) = G(\mathbf{x_A}, \mathbf{x_B}, t) - G(\mathbf{x_A}, \mathbf{x_B}, -t)$ in the time domain. Wapenaar et al. [24] discuss spurious effects that are introduced when the above theoretical assumptions are violated. However, in practice artefacts due to the violation of these theoretical assumptions are usually assumed to be sufficiently weak to allow for retrieval of useful empirical estimates of the Green's function. In particular, seismic interferometry has been shown empirically to be applicable in regimes where the medium illumination is one-sided [24,43–45]. Hence, while the boundary S in Equation (1) is typically assumed to be a closed contour which is sufficiently distant from the receivers, we assume S to be a straight source-line as depicted in the left insets of Figure 1. It is worth noting that while the amplitudes of empirical Green's function estimates are acknowledged to be unreliable, the phases of the Green's function recovered through such a procedure are assumed to be well approximated (see for example, [2,46], and remarks under Equation (9) below).

The main practical obstacle in the context of ambient noise interferometry stems from the fact that the medium's response to impulsive sources is not readily available. Hence, the information required for constructing the integrand in Equation (1) cannot be obtained directly. Instead, Wapenaar and Fokkema [2] propose the ambient noise interferometric relationship defined as

$$\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) := \frac{2}{\rho c} \hat{p}^*(\mathbf{x_A}, \omega) \hat{p}(\mathbf{x_B}, \omega), \tag{2}$$

where $\hat{p}(\mathbf{x}, \omega)$ denotes ambient noise recorded at a receiver location $\mathbf{x}$ for frequency ω. It is useful (and theoretically appropriate) to consider each recording $\hat{p}(\mathbf{x}, \omega)$ as a random field which is obtained via the Fourier transform of a realization of a stochastic process (a random mechanism) generating $p(\mathbf{x}, t)$ (these statements hold for $t, \omega \in \mathbb{R}$, we do not delve into the accuracy of finite-time approximations of the Fourier transform). In such a framework source locations $\mathbf{x}' \in S$ are no longer required to be known explicitly. Instead, the source characteristics are implicitly present in the ambient noise representation in Equation (2) and they are accounted for, in principle, in the ambient noise recordings represented via

$$\hat{p}(\mathbf{x}, \omega) = \int_S \hat{F}(\mathbf{x}', \omega) \hat{G}(\mathbf{x}, \mathbf{x}', \omega) d\mathbf{x}', \tag{3}$$

where $\hat{F}(\mathbf{x}, \omega)$ characterizes the unknown *source signature*. Wapenaar and Fokkema [2] showed that provided that $\hat{p}(\mathbf{x_A}, \omega)$ and $\hat{p}(\mathbf{x_B}, \omega)$ are obtained from recordings of uncorrelated noise sources, taking the empirical average over a large number ($N \gg 1$) of realizations of $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ in Equation (2), referred to hereafter as *stacking*, leads to a good

approximation (such an approximation is increasingly accurate for $N, t \rightarrow \infty$ provided that $h(\mathbf{x_A}, \mathbf{x_B}, t)$ is ergodic which seems to be a reasonable assumption in practice). of the phase of the inter-receiver Green's function $\hat{G}(\mathbf{x_A}, \mathbf{x_B}, \omega)$. However, as discussed in the subsequent sections, the stacking average is likely to produce unsatisfactory results in the presence of correlated noise sources even when long-time recordings are available.

In this paper, we delve deeper into the relationship between the retrievals $\hat{g}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ and $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ in order to identify conditions under which the ambient noise retrieval based on Equation (2) provides a satisfactory approximation of the retrieval Equation (1), and to develop techniques that allow us to achieve this for both uncorrelated and correlated noise sources.

To understand the need for the (statistical) averaging over multiple realizations of the ambient noise recordings in the retrieval (2) for some fixed boundary S, substitute the integral expression (3) for the noise recordings at receivers $\mathbf{x_A}$ and $\mathbf{x_B}$ into Equation (2) to obtain

$$\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) = \frac{2}{\rho c} \int_S \int_S \hat{F}^*(\mathbf{x}', \omega) \hat{F}(\mathbf{x}'', \omega) \, \hat{G}^*(\mathbf{x_A}, \mathbf{x}', \omega) \hat{G}(\mathbf{x_B}, \mathbf{x}'', \omega) d\mathbf{x}'' \, d\mathbf{x}' \, . \qquad (4)$$

Please note that the above formula represents a product of two independent integrals which are re-written in a form suitable for the subsequent analysis; to simplify notation, we skip the explicit dependence of $\hat{F}$ in Equation (4) on the particular recording of ambient noise, but this remains implicit. The geometry corresponding to the retrieval based on, respectively, Equations (1) and (4) is sketched on the left side of Figure 1. In contrast to $\hat{g}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ in Equation (1), where averaging over the source locations $\mathbf{x}' \in S$ leads to the cancellation of all phases except for the inter-receiver phase, the ambient noise retrieval $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ in Equation (4) does not represent a spatial average over S. Consequently, the retrieval in Equation (4) is dominated by acausal and spurious information contained in the integrand for $\mathbf{x}' \neq \mathbf{x}''$ (see middle panel of Figure 1).

It turns out that a reliable phase retrieval of the inter-receiver Green's function $\hat{G}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ from Equation (4) relies on an appropriate statistical average which couples the source characteristics $\hat{F}(\mathbf{x}', \omega)$ and $\hat{F}(\mathbf{x}'', \omega)$ for all $\mathbf{x}', \mathbf{x}'' \in S$ to generate an integral kernel which is localized in the neighborhood of $\mathbf{x}' = \mathbf{x}''$. As discussed in subsequent sections, the suitable statistical average depends on the nature of the noise. In particular, for correlated noise sources the (stacking) average over multiple realizations of Equation (2) is largely unsatisfactory regardless of the amount of available data, and a more versatile averaging approach must be used. Please note that while a constant medium density and propagation speed have been assumed, the analysis presented still holds when these two quantities are space-dependent.

In light of the above discussion, it is clear that the non-averaged relationship in Equation (4) is dominated by contributions from incongruous source locations, and these may lead to spurious contributions that completely obscure the phase of the inter-receiver Green's function. Consider, for example, the interferometric retrieval from a simulated sequence of sources which generate seismic impulses in a time-ordered fashion; hereafter, we refer to such a sequence as a *train of sources*. The train of sources is not to be confused with a *set of impulsive sources* recorded one-by-one and used to calculate the retrieval Equation (1); these two cases will be compared in what follows. In both cases we assume that the signal is recorded by a pair of receivers deployed in the vicinity of source-line S. To simulate the wavefield generated by trains of sources, we derived an analytical two-dimensional model which is described in Appendix A. Figure 2 illustrates the situation where the interferometric trace retrieved via Equation (4) from a train of sources differs significantly from the retrieval obtained from individually recorded impulsive sources, according to Equation (1). Both retrievals are based on sources on the same boundary S and recorded over the same time interval, except that the impulses in the train of sources are generated in a time-ordered sequence of physical locations which span S.
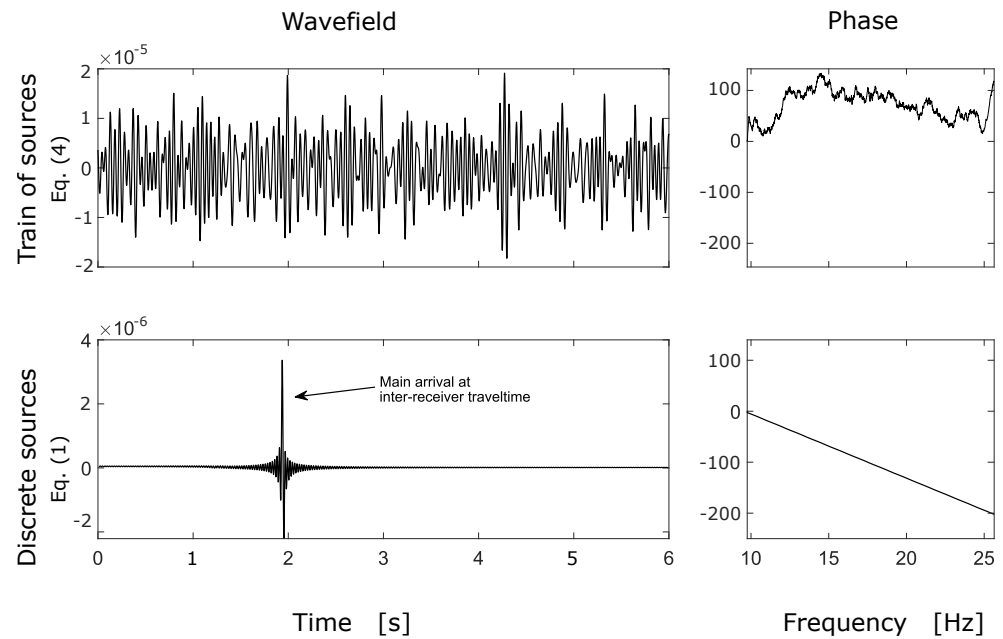
Figure 2. Comparison of the interferometric estimate of the Green's function from a simulated train of sources as per Equation (4) (**top**), versus the estimate from a boundary of static point sources (**bottom**) recorded individually, according to Equation (1). The left side panels show the waveforms in time, while the panels to the right show the corresponding estimated phases (unwrapped). The source train moves along the boundary line at 25 m/s (90 km/h) and emits sound in a frequency range between 10 and 25 Hertz. The receiver pair is oriented perpendicularly to the source boundary and 2000 m away from each other. Sound speed in the medium was set at 1000 m/s. The receiver closest to the source boundary was 400 m away.

When the sources are individually recorded (bottom of Figure 2) the inter-receiver Green's function is retrieved accurately with the main arrival found around 2 s, which corresponds to the inter-receiver travel time. In the top panels, the inter-receiver arrival is completely obscured by spurious arrivals, and the phase does not reproduce the reference estimate in the bottom panels.

In addition to the above issues, if the noise source is in motion, one must take into account both the spatiotemporal correlations of the noise sources, as well as the Doppler effect which shifts the energy emitted by the source at each frequency to a range of other recorded frequencies, even if the source is monochromatic. The instantaneous frequency $\breve{\omega}(\mathbf{x}, t')$ of the train of sources recorded at location $\mathbf{x}$ and time t' can be derived analytically from Equation (3) when it is used to represent a continuously moving source, resulting in the following expression (see the derivation of Equation (A23) in Appendix A)

$$\breve{\omega}(\mathbf{x}, t') = \frac{\tilde{\omega}}{1 - \frac{|\mathbf{v}|}{c} \cos \theta_{\mathbf{x}}(t')} , \tag{5}$$

where $\tilde{\omega}$ is the frequency emitted by the source, and $\theta_{\mathbf{x}}(t')$ is the angle between the line from the instantaneous source location at time t' to the receiver location $\mathbf{x}$, and the source-line S. Equation (5) agrees with the well-known expressions for the Doppler shift for the instantaneous frequency when the source is moving directly towards, or away from, the receiver. However, it is more general in that it quantifies the Doppler spread in the recorded frequency in terms of the relative source-receiver position angle $\theta_{\mathbf{x}}(t')$, when the source has an azimuthal velocity component. It is clear from Equation (5) that for a pair of distinct receivers $\mathbf{x}_{\mathbf{A}}$ and $\mathbf{x}_{\mathbf{B}}$, the corresponding relative position angles of the source $\theta_{\mathbf{x}_{\mathbf{A}}}(t')$ and $\theta_{\mathbf{x}_{\mathbf{B}}}(t')$ will be different; this means that the same source emitting the single frequency $\tilde{\omega}$ will be recorded with a different frequency at each location.

Figure 3 illustrates the relationship between the instantaneous location of a moving noise source and the recorded frequency $\breve{\omega}$ at each receiver $\mathbf{x_A}$ and $\mathbf{x_B}$. As expected, the recorded frequency is higher when the source is moving towards the receivers, and lower as it moves away. Clearly, as can be deduced from Equation (5) and observed in Figure 3, the only (instantaneous) location of the source that will lead to recording the true emitted frequency occurs when the source is collinear with the pair of receivers, and the line through $\mathbf{x_A}$ and $\mathbf{x_B}$ is perpendicular to the boundary S; depending on the geometry of the problem, such a location need not exist. Thus, in general, the frequency emitted by a moving noise source and recorded at the receivers will differ from the true frequency $\tilde{\omega}$, and the frequency bias will be different at $\mathbf{x_A}$ and $\mathbf{x_B}$ at all instantaneous locations except when the source is collinear with the receiver pair.
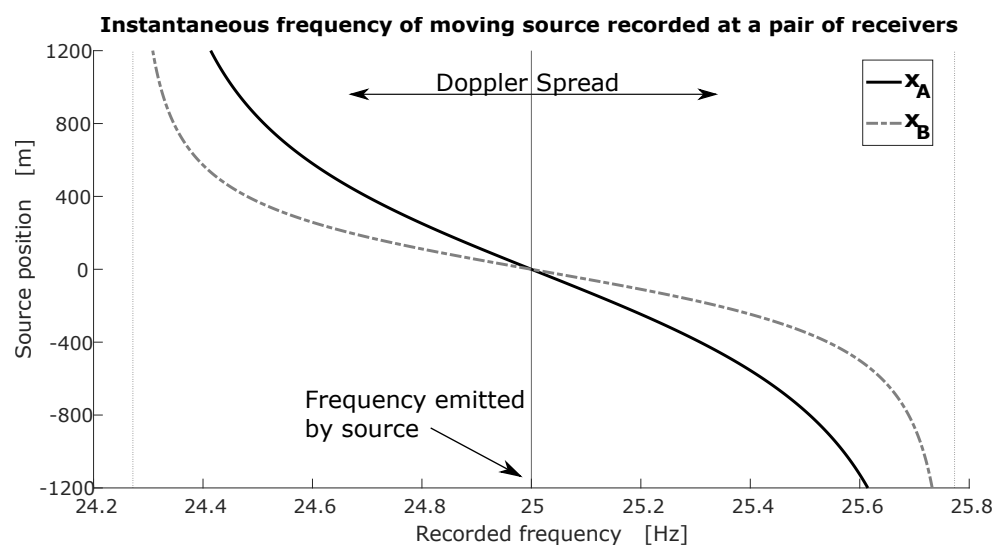
Figure 3. Relationship between source location and instantaneous frequency recorded at two different receiver locations $\mathbf{x_A}$, and $\mathbf{x_B}$. The simulated source was travelling at 30 m/s (108 km/h) and the medium speed was set at 1000 m/s. For illustration purposes the source was set to be monochromatic, emitting noise at a fixed frequency of 25 Hz. The limiting values for the Doppler spread, given by $25/(1 \pm |\mathbf{v}|/c)$ Hz, are marked with vertical dotted lines.

This has important implications for seismic interferometry, since recordings at any particular frequency will usually have been emitted at different source locations. For example, in Figure 3 the spectrum value recorded at 24.5 Hz would have been emitted by a source that was 800 m along the boundary for receiver $\mathbf{x_A}$, but only 400 m for receiver $\mathbf{x_B}$. Given that interferometry is performed frequency by frequency, this inconsistency in the recorded frequency is conceptually equivalent to introducing further spurious contributions from spatially incongruous source locations. Moreover, in the case of a more realistic broadband source the spread and consequent interaction among different frequencies becomes compounded, with multiple source locations along the boundary S introducing spatially incongruous contributions for every recorded frequency. Hence, the Doppler shift associated with moving noise sources further compounds the spurious effects inherent to ambient noise interferometry if we use the retrieval (2).

Equation (5) shows that the Doppler shift is amplified for higher frequencies, as well as increasing source train speeds. In addition, larger inter-receiver distances, as well as near-field recordings, will increase the discrepancies in frequencies recorded $\mathbf{x_A}$ and $\mathbf{x_B}$, resulting in further deterioration of the interferometric retrieval. Some of the potential obstacles for retrieving inter-receiver Green's functions from sources in motion have been considered theoretically by Sabra in [41], where the inconsistency induced by the differential Doppler effect at the receivers is mitigated via a stationary phase approximation, consistent with the discussion above, but under the assumption that the ambient noise sources are statistically uncorrelated in time. More recent work by Liu et al. [35], applies standard ambient noise

interferometry (2) to an expression for the wavefield generated by a train, and the potential retrievable wave types are discussed using intuitive ray-path arguments which are then tested on train recordings. However, the explicit role of the source speed in the retrieval, and systematic methods for mitigating these effects have not been addressed (and the Doppler effect is not considered in the latter study).

In summary, the crosstalk terms associated with the integrand in Equation (2) for $\mathbf{x}' \neq \mathbf{x}''$ emerge naturally in systematic ambient noise interferometry analysis but their effects are commonly removed from considerations by imposing the assumption that the ambient noise field is uncorrelated; such an assumption greatly simplifies the theoretical retrieval procedure but it is often unrealistic. This is the case, for example, when performing traffic or railway noise interferometry. Moreover, ambient noise interferometry (2) using moving point sources must contend with the Doppler effect in addition to the unavoidable correlations between the noise sources.

We discuss consequences of these issues in detail in the following section, where we point out that the delta-correlated assumption is unnecessarily restrictive, and that excluding near-diagonal terms reduces uncertainty in the retrieved amplitudes. We show in detail how the success of the standard stacking procedure depends on the properties of the source signature and on the source speed in cases where the noise sources are in motion. Moreover, we show how this standard stacking technique may fail, and propose an alternative workflow in Section 4 which proves to be successful at isolating terms from the incongruous crosstalk terms, leading to the extraction of the underlying inter-receiver phase regardless of the presence of correlation in the ambient noise field. Finally, for the scenario where the presence of correlation is caused by motion of the source, the workflow presented implicitly mitigates spurious contributions due the Doppler spread as explained above.

3. When and Why Stacking Works

A commonly used averaging procedure in ambient noise interferometry, referred to as stacking, relies on the arithmetic average over several interferometric retrievals $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ in Equation (2)—e.g., [2]. Here, we define *stacking* more formally as the operation consisting of taking the statistical average of functions evaluated on an ensemble of recordings of the ambient noise field over some time interval; we denote such an ensemble average by $\langle \, \cdot \, \rangle$. Below, we formally link the interferometric retrieval (2) to the retrieval (1) and determine conditions leading to the likely failure of the stacking average retrieval when the ambient noise field is correlated.

Application of the stacking average operator to the ambient noise retrieval represented via Equation (4) leads to

$$\langle \, \hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) \, \rangle = \frac{2}{\rho c} \int_S \int_S \langle \, \hat{F}^* \hat{F} \, \rangle (\mathbf{x}', \mathbf{x}'', \omega) \, \hat{G}^*(\mathbf{x_A}, \mathbf{x}', \omega) \hat{G}(\mathbf{x_B}, \mathbf{x}'', \omega) d\mathbf{x}' \, d\mathbf{x}'', \quad (6)$$

where

$$\langle \, \hat{F}^* \hat{F} \, \rangle (\mathbf{x}', \mathbf{x}'', \omega) := \langle \, \hat{F}^*(\mathbf{x}', \omega) \hat{F}(\mathbf{x}'', \omega) \, \rangle \quad (7)$$

is referred to as the *kernel of the stacking average* or simply as the *stacking average kernel*. This is determined by the correlation structure of the noise source characteristics $\hat{F}$. Recall that for brevity of notation we consistently skip the explicit dependence of $\hat{F}$ in Equation (6) on the particular recording of ambient noise, and that for any fixed $\mathbf{x}$ and ω, the source signature $\hat{F}(\mathbf{x}, \omega)$ should be thought of as a random variable, since it is associated with the Fourier transform of the ambient noise recording $p(\mathbf{x}, t)$. The stacking average $\langle \, \cdot \, \rangle$ is linear by construction and we assume that the order of integration is irrelevant (Fubini's theorem is satisfied). Please note that the averaging in $\langle \, \hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) \, \rangle$ effectively couples the two integrals which contribute to the single-recording ambient noise retrieval $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ in Equation (4). The formula in Equation (6) formally highlights the link between $\hat{g}(\mathbf{x_A}, \mathbf{x_B}, \omega)$

in Equation (1), the stacked ambient noise retrieval $\langle \hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) \rangle$, and the properties of the stacking average kernel $\langle \hat{F}^*\hat{F} \rangle$. Clearly, the properties of the kernel of the stacking average control the discrepancy between $\langle \hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) \rangle$ and $\hat{g}(\mathbf{x_A}, \mathbf{x_B}, \omega)$, which one is interested in minimizing.

In particular, the common interferometric assumption that the ambient noise field is uncorrelated leads to the stacking average kernel in Equation (6) of the form

$$\langle \hat{F}^*\hat{F} \rangle(\mathbf{x}', \mathbf{x}'', \omega) \propto \delta(\mathbf{x}' - \mathbf{x}'')|\hat{s}(\omega)|^2\,, \tag{8}$$

where $\hat{s}(\omega)$ is the power spectrum of the ambient noise source. There is a further implicit assumption in Equation (8) that the power spectrum of the source does not depend on the spatial variables. Clearly, if Equation (8) holds, evaluating the integrals in Equation (6) leads to

$$\langle \hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) \rangle \propto \hat{g}(\mathbf{x_A}, \mathbf{x_B}, \omega)|\hat{s}(\omega)|^2\,, \tag{9}$$

where $\hat{g}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ represents the retrieval from individually recorded discrete sources defined by Equation (1). The result in Equation (9) implies that even in the idealized scenario when Equation (8) holds, ambient noise interferometry does not generally allow for accurate retrieval of the amplitude of the Green's function due to the presence of the noise power spectrum $|\hat{s}(\omega)|^2$, and the lack of individual source recordings which could allow for a source-specific spectrum correction. On the other hand, the estimated phase remains unaffected by the presence of the real factor $|\hat{s}(\omega)|^2$ in Equation (9). Note also that in practice, one has to consider empirical averages in Equations (6)–(8) based on a finite number of ambient noise recordings instead of the abstract stacking in terms of the statistical average over an 'infinite' number of recordings implied by the operator $\langle \cdot \rangle$. In such a case, if the relationship in Equation (8) holds in the limit of averaging over several samples tending to infinity, it can be shown that the acausal contributions from off-diagonal, incongruent source locations (i.e., when $\mathbf{x}' \neq \mathbf{x}''$) decay with the number of samples and Equation (9) holds asymptotically; we postpone these more technical details to another publication. A related derivation of a bound on the signal to noise ratio can be found in Appendix B of [47].

The requirement that ambient noise be delta-correlated in space is restrictive and it is often unrealistic to assume such an approximation in applications. As pointed out in [48], most sources in practice are, unsurprisingly, neither fully correlated nor fully uncorrelated. Although some advances have been made in the context of simultaneously acting sources [40,49] based on deconvolution techniques to tackle the resulting underdetermined systems, further investigation of such a deconvolution setup is required. Nevertheless, the relationship in Equations (8) and (9) is desirable, as it allows for the retrieval of correct phases of the associated Green's functions from Equation (6). Therefore, in what follows we consider approaches which resemble the relationship in Equation (8) but without the requirement for the spatially uncorrelated noise sources. The remainder of this section is devoted to the analysis of the stacking average through the properties of the stacking average kernel $\langle \hat{F}^*\hat{F} \rangle$ in Equation (7). A new approach, which exploits a different averaging kernel and applies to both correlated and uncorrelated noise sources, is proposed and analyzed in Section 4.

To gain a theoretical insight into the ambient noise retrieval, we decompose Equation (4) as

$$\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) = \hat{g}_\epsilon(\mathbf{x_A}, \mathbf{x_B}, \omega) + \hat{x}_\epsilon(\mathbf{x_A}, \mathbf{x_B}, \omega)\,, \tag{10}$$

where $\epsilon \geqslant 0$ and

$$\hat{g}_\epsilon(\mathbf{x_A}, \mathbf{x_B}, \omega) = \frac{2}{\rho c} \int\limits_{\mathbf{x'} \in S} \int\limits_{\substack{\mathbf{x''} \in S \\ |\mathbf{x'} - \mathbf{x''}| < \epsilon}} \hat{F}^*(\mathbf{x'}, \omega) \hat{F}(\mathbf{x''}, \omega)\, \hat{G}^*(\mathbf{x_A}, \mathbf{x'}, \omega) \hat{G}(\mathbf{x_B}, \mathbf{x''}, \omega)\, d\mathbf{x''}\, d\mathbf{x'}, \tag{11}$$

and

$$\hat{x}_\epsilon(\mathbf{x_A}, \mathbf{x_B}, \omega) = \frac{2}{\rho c} \int\limits_{\mathbf{x'} \in S} \int\limits_{\substack{\mathbf{x''} \in S \\ |\mathbf{x'} - \mathbf{x''}| \geqslant \epsilon}} \hat{F}^*(\mathbf{x'}, \omega) \hat{F}(\mathbf{x''}, \omega)\, \hat{G}^*(\mathbf{x_A}, \mathbf{x'}, \omega) \hat{G}(\mathbf{x_B}, \mathbf{x''}, \omega)\, d\mathbf{x''}\, d\mathbf{x'} .$$

(12)

The term $\hat{g}_\epsilon(\mathbf{x_A}, \mathbf{x_B}, \omega)$, referred to hereafter as the *ϵ-diagonal term*, accounts for the contributions from sources such that $|\mathbf{x'} - \mathbf{x''}| < \epsilon$, and one might consider sufficiently small $\epsilon \geqslant 0$, in order to minimize the contribution from acausal arrivals from incongruous source locations in $\hat{g}_\epsilon$. The decomposition in Equation (10) holds for any choice of $\epsilon \geqslant 0$ but taking $\epsilon \to 0$ does not necessarily remove all acausal arrivals from $\hat{g}_\epsilon$ for moving noise sources, as explained in the discussion in Section 2. However, the aim of the above decomposition is to find an appropriate choice of the *spatial correlation length cut-off ϵ* so that $\hat{g}_\epsilon$ is dominated by the coherent terms, whereas $\hat{x}_\epsilon$ is dominated by the acausal contributions due to incongruous source locations. Please note that the decomposition in Equation (10) could be applied to the integral with respect to $\mathbf{x'}$ instead of $\mathbf{x''}$ without loss of generality.

Finally, we consider the stacking average of the decomposition in Equation (10) given by

$$\langle \hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) \rangle = \langle \hat{g}_\epsilon(\mathbf{x_A}, \mathbf{x_B}, \omega) \rangle + \langle \hat{x}_\epsilon(\mathbf{x_A}, \mathbf{x_B}, \omega) \rangle \tag{13}$$

with the aim of linking the characteristics of the stacking average kernel $\langle \hat{F}^* \hat{F} \rangle(\mathbf{x'}, \mathbf{x''}, \omega) = \langle \hat{F}^*(\mathbf{x'}, \omega) \hat{F}(\mathbf{x''}, \omega) \rangle$ to the spatial correlation length cut-off ϵ in the above decomposition. Please note that in this section no specific assumptions have been made about the source or its spectrum, as the representation (3) used to arrive at Equations (10) and (13) is general.

Interferometric Retrieval from Spatially Correlated Noise Due to a Moving Source

To show the utility of the ϵ-diagonal decomposition (10) in interferometric analysis, we consider a specific but common example of a broadband source train travelling at speed $\mathbf{v} \neq 0$ and emitting sound within the frequency band $[\tilde{\omega}_{\min}, \tilde{\omega}_{\max}]$, such that the phases for different emitted frequencies are random and independent from each other (we address correlation of the phase across different frequencies below), i.e., for each frequency $\tilde{\omega}$ emitted by the source, there is a corresponding random phase shift $\theta_{\tilde{\omega}}$ uniformly distributed between 0 and 2π, and we assume that the phase shifts $\theta_{\tilde{\omega}}$ and $\theta_{\tilde{\omega}'}$ are independent whenever $\tilde{\omega} \neq \tilde{\omega}'$. In such a case we find (see Equation (A34) for details) that the stacking average kernel in Equation (7) is given by

$$\langle \hat{F}^* \hat{F} \rangle(\mathbf{x'}, \mathbf{x''}, \omega) = \cos\left(\bar{\tilde{\omega}} \frac{|\mathbf{x''} - \mathbf{x'}|}{|\mathbf{v}|} \right) \mathrm{sinc}\left(\tfrac{1}{2} \Delta_{\tilde{\omega}} \frac{|\mathbf{x''} - \mathbf{x'}|}{|\mathbf{v}|} \right) |\hat{s}(\omega)|^2, \tag{14}$$

which depends on the train speed $\mathbf{v} \neq 0$, the mean emitted frequency $\bar{\tilde{\omega}} := \tfrac{1}{2}(\tilde{\omega}_{\max} + \tilde{\omega}_{\min})$ and the bandwidth $\Delta_{\tilde{\omega}} := \tilde{\omega}_{\max} - \tilde{\omega}_{\min}$, and where $\mathrm{sinc}(x) := \sin(x)/x$. The model and the resulting stacking average kernel (14) implicitly take account of the Doppler effect (see model derivations in Appendix A). Expression (14) for the kernel holds for non-monochromatic sources travelling at speeds $\mathbf{v} \neq 0$, and for sources whose spectrum can be represented in the form Equation (A27). In the limit of $\Delta_{\tilde{\omega}} \to 0$, i.e., when the source is monochromatic with frequency $\tilde{\omega}$, this kernel reduces to $\langle \hat{F}^* \hat{F} \rangle(\mathbf{x'}, \mathbf{x''}, \omega) = \cos(\tilde{\omega}(|\mathbf{x''} - \mathbf{x'}|)|\hat{s}(\omega)|^2)/|\mathbf{v}|$ which is no longer localized. See remarks under Equation (A24) in Appendix A concerning a more general model that allows for appropriate localization. For simplicity, we consider the

parametrization of the straight boundary S given by $\mathbf{x} = (x, 0)$, where $x = 0$ corresponds to the location where the line through the receiver pair intersects the boundary S, so that $\mathbf{x}' = (x', 0)$ and similarly $\mathbf{x}'' = (x'', 0)$. Using this parametrization we have that $|\mathbf{x}' - \mathbf{x}''| = |x' - x''|$ in Equation (14). For further details see Equation (A35) and accompanying remarks in Appendix B.

The movement of the noise source allows us to parameterize any locations $\mathbf{x}', \mathbf{x}'' \in S$ via the travel time so that $\mathbf{x}' = (|\mathbf{v}|t', 0)$ and $\mathbf{x}'' = (|\mathbf{v}|t'', 0)$. The analytical expression for the stacking average kernel associated with a broadband correlated train of sources allows one to identify the approximate correlation length beyond which the envelope of the kernel (14) decays rapidly. The unwanted long-range correlations contributing to the kernel away from the neighborhood of $\mathbf{x}' = \mathbf{x}''$, $\mathbf{x}', \mathbf{x}'' \in S$, represent the main contribution to $\langle \hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) \rangle$ in Equation (13) from the acausal arrivals. Based on Equation (14), we set the spatial correlation length cut-off ϵ to twice the value of the first zero of the envelope of the stacking average kernel (see the bottom panel of Figure 4), which is given by

$$\epsilon = \frac{|\mathbf{v}|}{\tilde{\omega}_{\max} - \tilde{\omega}_{\min}}, \tag{15}$$

for non-monochromatic sources travelling at speeds $\mathbf{v} \neq 0$. Figure 4 shows the numerically approximated stacking average kernel $\langle \hat{F}^*\hat{F} \rangle (\mathbf{x}', \mathbf{x}'', \omega)$ and its exact form (14). The rate of decay of the source kernel in terms of the spatial discrepancy $|\mathbf{x}' - \mathbf{x}''|$ is controlled by the envelope of the $\mathrm{sinc}(x)$ factor in Equation (14). As noted before, the spatial discrepancy can be represented in terms of the travel times so that $|\mathbf{x}' - \mathbf{x}''| = |t' - t''||\mathbf{v}|$, where $\mathbf{v}$ is the train speed. Although the choice of the spatial correlation cut-off length ϵ has so far been somewhat arbitrary, the main principle in its determination is to use the structure of the stacking average kernel $\langle \hat{F}^*\hat{F} \rangle$ in order to approximately identify the correlation length beyond which the crosstalk terms dominate the correlation structure represented by $\langle \hat{F}^*\hat{F} \rangle$.

Please note that while the decomposition (10) and its stacking average Equation (13) are given in the frequency domain to highlight the interaction between incongruous source locations, the corresponding stacking average decomposition in the time domain is given by

$$h(\mathbf{x_A}, \mathbf{x_B}, t) = g_\epsilon(\mathbf{x_A}, \mathbf{x_B}, t) + x_\epsilon(\mathbf{x_A}, \mathbf{x_B}, t), \tag{16}$$

and its stacking average is given by

$$\langle h(\mathbf{x_A}, \mathbf{x_B}, t) \rangle = \langle g_\epsilon(\mathbf{x_A}, \mathbf{x_B}, t) \rangle + \langle x_\epsilon(\mathbf{x_A}, \mathbf{x_B}, t) \rangle. \tag{17}$$

The above time-domain decompositions hold due to the linearity of the inverse Fourier transform; as is standard in the interferometric literature, we assume throughout that h is square-integrable in $t \in \mathbb{R}$ (so that the Fourier transform is well-defined on $h \in L^2(\mathbb{R})$) and that $\hat{F}$ and $\hat{G}$ are such that Fubini's theorem holds throughout. Thus, the diagonal band decompositions of h and $\hat{h}$ are dual to each other and they can be considered interchangeably.

The formulation of the ϵ-diagonal decomposition in the time domain (i.e., representation (16)) lends itself directly to the analysis based on time series recordings provided that a suitable spatial correlation length cut-off ϵ can be determined. When the noise source is in motion, as in the case considered here, the spatial correlation length cut-off ϵ has a temporal counterpart $T_\epsilon = \epsilon / |\mathbf{v}|$ for $\mathbf{v} \neq 0$. In Section 4, we will focus on a procedure for recovering the ϵ-diagonal band structure, analogous to that in $\langle g_\epsilon \rangle$ in Equation (17), from a single short recording when the value of the spatial correlation length cut-off ϵ cannot be estimated from the stacking average kernel. First, however, we outline the general properties of the stacking average and its potential pitfalls when retrieving the inter-receiver interferometric Green's function. To this end we consider two different classes of noise sources, and we compare the standard stacked retrieval $h(\mathbf{x_A}, \mathbf{x_B}, t)$ and its stacking average $\langle h(\mathbf{x_A}, \mathbf{x_B}, t) \rangle$ with retrievals based on isolating the ϵ-diagonal band terms $g_\epsilon(\mathbf{x_A}, \mathbf{x_B}, t)$

and $\left\langle g_\epsilon(\mathbf{x_A}, \mathbf{x_B}, t) \right\rangle$ in the decomposition Equations (16) and (17), respectively, using an 'informed' choice of the spatial correlation length cut-off ϵ given by Equation (15).
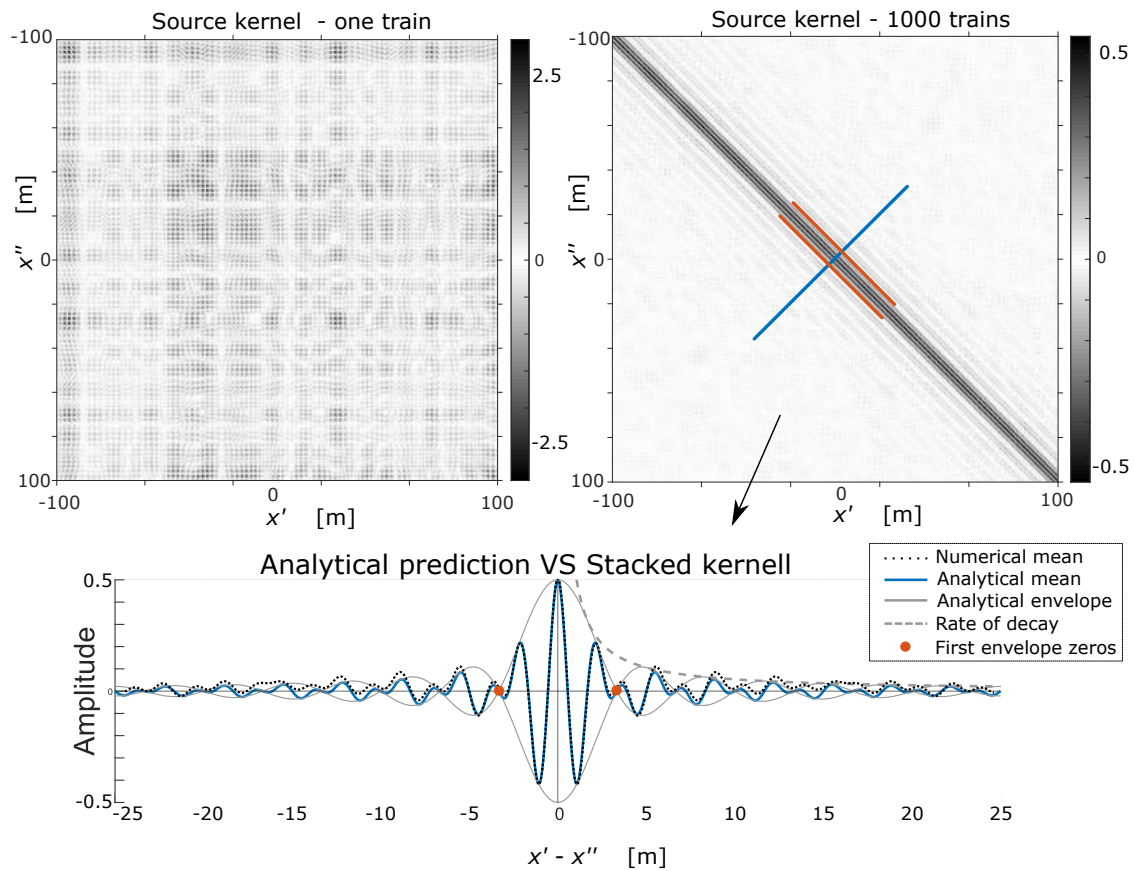
Figure 4. Example of the structure of the source kernel $\hat{F}^*(\mathbf{x'}, \omega)\hat{F}(\mathbf{x''}, \omega)$ and the stacking average kernel $\langle \hat{F}^*\hat{F} \rangle(\mathbf{x'}, \mathbf{x''}, \omega) = \langle \hat{F}^*(\mathbf{x'}, \omega)\hat{F}(\mathbf{x''}, \omega) \rangle$ derived in Equation (14). Top left panel shows the spatial part of source kernel (14) (that is, without including the frequency-dependent term $|\hat{s}(\omega)|^2$) as a function of the location of a single train of sources recorded at each receiver. The simulated train travels at 90 km/h and its spectrum ranges from 10 Hz to 25 Hz. The top-right panel shows the corresponding kernel after stacking 1000 source train recordings each five minutes long. The bottom panel compares the numerical anti-diagonal section in black dashes, extracted from the numerical mean in the top-right panel, with the prediction given by Equation (14) in blue, as well as its analytical envelope in solid gray and the rate of decay in dashed gray, in terms of the signed spatial discrepancy $x' - x''$, where the parametrization $\mathbf{x'} = (x', 0)$, $\mathbf{x''} = (x'', 0)$ has been used. The rate of decay is derived from the envelope of the sinc factor in Equation (14). The orange dots highlight the first pair of zeros of the analytical envelope, capturing the energy around the mean. The radius ϵ will be taken to be the distance between these zeros.

In Figure 5 we use the spatial cut-off radius ϵ determined in Equation (15) and apply the ϵ-diagonal decomposition procedure laid out in Equation (16). First, we assume that individual recordings are statistically independent. In particular, for a noise source with a broadband spectrum this constraint translates to the requirement that the frequencies associated with the random source signature $\hat{F}$ be independent. To generate this figure, a train of sources travelling at 90 km/h along a straight boundary and emitting sound between 10 Hz and 25 Hz was simulated to generate the ambient noise recordings $p(\mathbf{x_A}, t)$ and $p(\mathbf{x_B}, t)$. The receiver pair was oriented perpendicularly to the source-line, with an inter-receiver distance of 2000 m and the closest receiver 400 m from the source-line. The medium speed was set to 1000 m/s. The ambient noise recordings were transformed to the frequency domain and $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ was computed according to Equation (2). Then the ϵ-decomposition was calculated for a single run to generate the insets in the top row. The frequency-domain data were used to extract the phase in the top-right panel of Figure 5, as well as transformed back to the time domain to generate the time series $h(\mathbf{x_A}, \mathbf{x_B}, t)$,

$g_\epsilon(\mathbf{x_A}, \mathbf{x_B}, t)$, and $x_\epsilon(\mathbf{x_A}, \mathbf{x_B}, t)$ in the top left panel of the same Figure. This procedure was repeated 1000 times and averaged to calculate $\langle \hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega)\rangle$, as well as its ϵ-decomposition with the associated phases (right) and waveforms (left) in the bottom row of this figure.
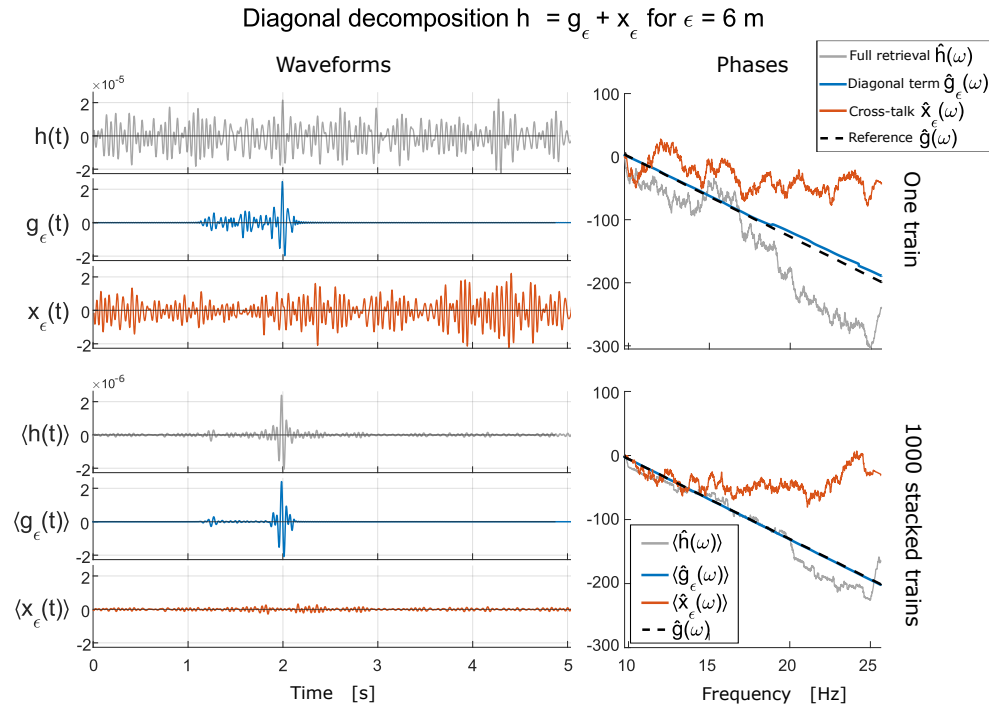
Figure 5. Example of the ϵ-diagonal decomposition procedure Equation (16) and its stacking average version Equation (17) for a train travelling at 90 km/h with the emitted spectrum between 10 Hz and 25 Hz. The receiver pair was placed perpendicularly to the road, with an inter-receiver distance of 2000 m and a distance of 400 m from the source-line to the closest receiver. The medium speed was set at 1000 m/s. The decomposition is performed for $\epsilon = 6$ m ($T_\epsilon = 0.25$ s) determined from Equation (15) for a single five-minute recorded source train (**top**) and for a stacked retrieval over 1000 five-minute train source recordings (**bottom**). The left panels show the waveforms of the components Equations (16) and (17), and the right panels show the corresponding phases in the frequency domain obtained from the frequency representation of the retrieval given by Equations (10) and (13). The retrieval $h(t)$ and its stacking average $\langle h(t)\rangle$ are shown in gray, and their components $g_\epsilon(t)$, $\langle g_\epsilon(t)\rangle$, and $x_\epsilon(t)$, $\langle x_\epsilon(t)\rangle$, are shown in blue and orange, respectively. The phase plots include the phase of the inter-receiver Green's function (dashed black) estimated from discrete sources, as prescribed by Equation (1) as a reference.

In this setup, and given the stacking average kernel in Equation (14), the retrieval of the inter-receiver Green's function is shown in gray in both the temporal and the spatial domains, with the ϵ-diagonal terms g_ϵ, $\hat{g}_\epsilon$ in blue and the crosstalk x_ϵ, $\hat{x}_\epsilon$ shown in orange. According to Equation (15), we chose $\epsilon = 6$ m which corresponds to a time window length $T_\epsilon \approx 0.25$ s. The resulting spectra were transformed to the temporal domain via inverse Fourier transform to show the effect of the decomposition on the retrieved waveforms; the components of Equation (16) are shown in the left panels of Figure 5. The decomposition was applied to a single five-minute recording of a source train (top), and to the averaged waveform resulting from stacking 1000 such recordings (bottom). The phases corresponding to each waveform are shown in the panels to the right in Figure 5.

This illustrates the improvement in the retrieval using the decomposition (10) which satisfactorily separates the causal contributions to the inter-receiver phase from the spurious crosstalk.

Please note that the phase of $\langle \hat{h}(\omega)\rangle$ in the bottom right panel of Figure 5 (in blue) typifies the phase retrieved from the standard ambient noise interferometry approach which relies on averaging over stacked independent recordings of ambient noise of the

form (2). Although stacking can clearly improve the quality of the retrieved phase of the inter-receiver Green's function, significant discrepancies remain especially in the higher frequency range of the spectrum due to the increased influence of the Doppler shift even for many stacked recordings. It should be noted that sources travelling at different speeds and overlapping with each other may help improve the rate of convergence of the standard stacking to the correct phase. On the other hand, the phase of the ϵ-diagonal term $\hat{g}_\epsilon(\omega)$ shows a much better agreement with the uncorrelated reference estimate even for a single short recording (blue phase in top-right panel of Figure 5). Application of the stacking average to the ϵ-diagonal term $\langle \hat{g}_\epsilon(\omega) \rangle$ leads to a further but slight improvement over the standard stacking applied directly to $\hat{h}$ which is evident in the higher frequency range of the spectrum. This increased uncertainty for higher frequencies is to be expected given the presence of the Doppler shift.

The above experiment indicates that although there is some flexibility in the choice of the spatial correlation length cut-off ϵ, it is possible to carry out the ϵ-diagonal decomposition (10) in a way which on average, largely separates the desired contributions from coherent sources and the crosstalk between incongruous source locations into two distinct terms $\hat{g}_\epsilon(t)$ and $\hat{x}_\epsilon(t)$ respectively. In the case of correlated moving noise sources, a sufficiently small correlation length cut-off ϵ results in attributing a significant proportion of the coherent noise contributions which should be captured by $\hat{g}_\epsilon$ relative to the 'crosstalk' term $\hat{x}_\epsilon$; the situation is reversed for a sufficiently large ϵ. If we were to evaluate these integrals explicitly in order to estimate the true phase, the challenge lies in determining an optimal value of the spatial correlation length cut-off ϵ that minimizes the error in the phase retrieval in the absence of knowledge of the spatial correlation structure contained in the stacking average kernel $\langle \hat{F}^* \hat{F} \rangle (\mathbf{x}', \mathbf{x}'', \omega)$ in Equation (7).

It is worth noting that a successful phase retrieval of the inter-receiver Green's function based on the empirical average of stacked recordings fundamentally relies, via the law of large numbers, on the statistical independence of the individual recordings. In particular, for a noise source with a broadband spectrum this constraint translates to the requirement that the frequencies associated with the random source signature $\hat{F}$ are independent; otherwise, the empirical average of the stacked retrieval $\frac{1}{N} \sum_{n=1}^{N} \hat{h}^{(n)}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ might still converge as $N \to \infty$ but there are no guarantees that the resulting limit $\langle \hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) \rangle$ will lead to the recovery of the true inter-receiver phase. Figure 6 shows one such example, where the setup and parameters are identical to those in Figure 5, except that the phases of the frequencies emitted by the sources are no longer uncorrelated. Such a scenario would be likely to occur, for example, due to an interaction between sleepers on a rail track and the wheels of the train [32]. Specifically, for each frequency $\tilde{\omega}$ emitted by the source, the random phases associated with each emitted frequency are such that for a pair of emitted frequencies $\tilde{\omega}$ and $\tilde{\omega}'$, $|\tilde{\omega} - \tilde{\omega}'| \ll 1$, the phase shifts satisfy $\theta_{\tilde{\omega}} = \theta'_{\tilde{\omega}} + r$, where r is uniformly distributed between 0 and 2π, so that the phase shifts $\theta_{\tilde{\omega}}$ and $\theta_{\tilde{\omega}'}$ are no longer uncorrelated.

The gray trace in the middle panel of Figure 6 shows the averaged waveform in the time domain obtained by stacking 1000 five-minute long recordings. In contrast to the result of the equivalent procedure shown in Figure 5 (bottom left), stacking the recordings does not converge (or does not converge sufficiently quickly) to the correct inter-receiver arrival, and the crosstalk term (orange) in the middle panel does not decay. The corresponding phases are shown in the right panel of Figure 6 with the phase estimated from stationary discrete sources (see Equation (1)) in dashed black as a reference. However, the phase of the inter-receiver signal is still successfully isolated by the ϵ-diagonal decomposition (10), as indicated by the blue line, showing a good agreement with the phase of the inter-receiver Green's function estimated from discrete sources, as per Equation (1).
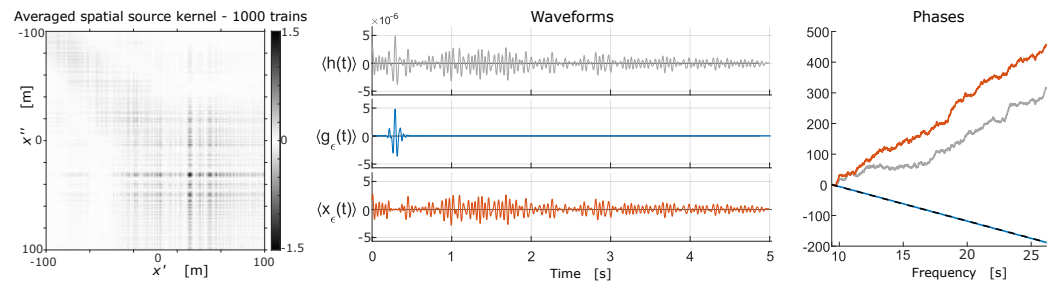
Figure 6. Example of a source signature from which the interferometric Green's function estimate cannot be retrieved by stacking. The left panel shows the numerical/empirical approximation of the stacking average kernel $\langle \hat{F}^* \hat{F} \rangle (\mathbf{x}', \mathbf{x}'', \omega)$ in Equation (7) estimated by stacking 1000 trains of sources travelling at 90 km/h with an emitted spectrum between 10 Hz and 25 Hz. The receivers were placed perpendicularly to the source-line, with an inter-receiver distance of 200 m and the receiver closest to the source-line at a distance of 400 m. Away from the diagonal, non-zero elements persist. The middle panel shows the stacked waveform $\langle h(t) \rangle$ and its ϵ-diagonal decomposition into $\langle g_\epsilon(t) \rangle$ and $\langle x_\epsilon(t) \rangle$. The right panel shows the corresponding unwrapped phases for the full term $\langle \hat{h}(\omega) \rangle$ (gray) and its decomposition into the ϵ-diagonal term $\langle \hat{g}_\epsilon(\omega) \rangle$ (blue) and the crosstalk term and $\langle \hat{x}_\epsilon(\omega) \rangle$ (orange), as well as the phase of the inter-receiver Green's function (dashed black) estimated from discrete sources as per Equation (1) as a reference. In this case, the stacking average kernel $\langle \hat{F}^* \hat{F} \rangle (\mathbf{x}', \mathbf{x}'', \omega)$ in Equation (7) is not localized and the value of the correlation length cut-off ϵ was chosen according to the approach described in Section 4. The phase retrieval from $\langle \hat{g}_\epsilon(\omega) \rangle$ and $\hat{g}_\epsilon(\omega)$ is similar and not shown.

Clearly, in most real-world applications it is not possible to gauge the retrieval against some reference benchmark, which leaves the retrieved results and the convergence trends open to subjective interpretation. Moreover, even if enough recorded data are available to allow for stacked averaging but the stacking average kernel $\langle \hat{F}^* \hat{F} \rangle (\mathbf{x}', \mathbf{x}'', \omega)$ in Equation (7) is not spatially localized, there is no a priori criteria for estimating the appropriate value for the spatial correlation length cut-off ϵ needed in the decomposition (10) (compare the example in Figure 6 with that illustrated in Figures 4 and 5). The key point of these considerations is that in principle it remains possible to estimate the correct phase of the underlying Green's function from a single short recording of a correlated ambient noise source, and that the cause of the noise correlation is not relevant in such considerations.

In summary, in this section we discussed a constructive example of the interferometric retrieval of the inter-receiver Green's function which was based on an abstract decomposition of the retrieval (2) as $\hat{h} = \hat{g}_\epsilon + \hat{x}_\epsilon$, as in Equations (10) or (16), and based on an 'informed' choice of the spatial correlation length cut-off ϵ so that the undesirable crosstalk terms are largely contained in the $\hat{x}_\epsilon$ term. This construct, driven by analytical insight and illustrated in Figures 4–6, suggests that it should be possible to carry out such a decomposition and subsequently obtain the phase estimate of the inter-receiver Green's function if a suitable correlation cut-off length ϵ can be estimated. Importantly, this approach would allow one to sidestep the classical stacking average approach and avoid its potential pitfalls. In the following section we propose a workflow that facilitates the extraction of the interferometric phase estimate from a single short recording of a passing train (i.e., an instance of a moving correlated noise source). In fact, the choice of the correlation length cut-off ϵ in the example in Figure 6 was determined via a procedure described in the next section, since in the setup of Figure 6 the stacking average kernel $\langle \hat{F}^* \hat{F} \rangle (\mathbf{x}', \mathbf{x}'', \omega)$ was not spatially localized due to implicit correlations between the noise sources.

4. The Random Windowing Technique

Based on the analytical insight outlined in the previous sections, we consider the problem of estimating the phase of the inter-receiver Green's function from a single short recording of a train of sources, without resorting to the standard averaging over a stacked ensemble of recordings. Nevertheless, the retrieval based on this procedure effec-

tively results in the extraction of the ϵ-diagonal term in the decomposition $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) = \hat{g}_\epsilon(\mathbf{x_A}, \mathbf{x_B}, \omega) + \hat{x}_\epsilon(\mathbf{x_A}, \mathbf{x_B}, \omega)$ introduced in Equation (10), where ϵ is determined without the explicit knowledge of the stacking average kernel $\langle \hat{F}^*\hat{F} \rangle (\mathbf{x}', \mathbf{x}'', \omega)$ and the need for it to be localized. Importantly, the technique applies to both correlated and uncorrelated noise sources.

Consider a single time-domain recording of a train of sources at a receiver pair $\mathbf{x_A}$ and $\mathbf{x_B}$ and denote these recordings by, respectively, $p(\mathbf{x_A}, t)$ and $p(\mathbf{x_B}, t)$ whose Fourier transform leads to the retrieval $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega) = \hat{p}^*(\mathbf{x_A}, \omega)\hat{p}(\mathbf{x_B}, \omega)$ in Equation (2), representing the interferometric approximation of the inter-receiver Green's function $\hat{G}(\mathbf{x_A}, \mathbf{x_B}, \omega)$. In the previous section the correlation length cut-off ϵ had to be estimated from the stacking average kernel $\langle \hat{F}^*\hat{F} \rangle (\mathbf{x}', \mathbf{x}'', \omega)$ in Equation (6), assuming that the stacking average had a localized kernel (e.g., Figure 4) which is not always the case for correlated noise sources (e.g., Figure 6). We now determine an optimal spatial/temporal window size that minimizes the effect of acausal arrivals in the interferometric retrieval from a single short recording. The approach relies on the use of an appropriate randomization technique, termed *random windowing*, in order to introduce incoherence into the acausal part of the phase in the retrieval $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ in the available recording and mitigate the influence of the crosstalk terms, largely contained in $\hat{x}_\epsilon(\mathbf{x_A}, \mathbf{x_B}, \omega)$ of the ϵ-diagonal decomposition of $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega)$, which hinder the accurate retrieval of the Green's function $\hat{G}(\mathbf{x_A}, \mathbf{x_B}, \omega)$.

4.1. General Setup

First, consider the interferometric retrieval $h(\mathbf{x_A}, \mathbf{x_B}, t)$ in Equation (2) obtained from recorded time series data within an interval/window of duration $T > 0$ and centered at some randomly drawn time t_n. For a pair of receiver locations $\mathbf{x_A}$ and $\mathbf{x_B}$ restrict the recorded time series $p(\mathbf{x_A}, t)$ and $p(\mathbf{x_B}, t)$ to that temporal window (see Figure 7). The random windowing method introduced below can be applied to recorded time series generated by arbitrary noise sources. However, in what follows we consider correlated noise generated by a source moving with speed $\mathbf{v} \neq 0$. This setup is relevant in applications, and it also aids understanding of the theoretical underpinning of the random windowing method, since it allows one to link the windowing of the time series to the spatial windowing, as illustrated in Figure 7. For time series generated by arbitrary noise sources, $\mathbf{v}$ does not correspond to a physical velocity, and stands for an abstract parameter that allows for the space time transformation. Next, for a given average source velocity, the window of duration T is linked to a spatial window which we define for a receiver location $\mathbf{x}$ as

$$S(\mathbf{x}, \mathbf{v}, T, t_n) = \left\{ \mathbf{x}' \in S : \mathbf{x}' = \mathbf{v}t' \text{ for } |t' - t_n| < \tfrac{1}{2}T \right\}, \tag{18}$$

and which corresponds to the segment of the (spatial) source boundary traversed by the source moving at speed $\mathbf{v}$ from the origin at time zero, and recorded at a receiver location $\mathbf{x}$ within the time window of duration T centered at t_n. As illustrated in the right panel of Figure 7, these boundary segments will differ for each receiver, depending on the inter-receiver distance, the distance to the source-line, and the speed of the source. The theoretical spatial location of $S(\mathbf{x}, \mathbf{v}, T, t_n)$ for each receiver can be calculated explicitly using Equation (A11) and scaling by the source speed (or by its estimated mean), as described in Appendix A. The frequency-domain representation of the time-windowed data at each receiver can be expressed as

$$\hat{p}(\mathbf{x_A}, \omega \, ; T, t_n) = \int_{S(\mathbf{x_A}, \mathbf{v}, T, t_n)} \hat{F}(\mathbf{x}', \omega)\hat{G}(\mathbf{x_A}, \mathbf{x}', \omega)d\mathbf{x}',$$
$$\hat{p}(\mathbf{x_B}, \omega \, ; T, t_n) = \int_{S(\mathbf{x_B}, \mathbf{v}, T, t_n)} \hat{F}(\mathbf{x}', \omega)\hat{G}(\mathbf{x_B}, \mathbf{x}', \omega)d\mathbf{x}', \tag{19}$$

where $S(\mathbf{x}, \mathbf{v}, T, t_n)$ is given by Equation (18). Please note that explicit knowledge of the signals integrated along boundary sections is not required in practice, as the windowing procedure is performed directly on the full recorded time series. A source travelling at speed $|\mathbf{v}|$ will span a distance no greater than $|\mathbf{v}|T$ during a time window of duration T, so that for any random time window location t_n the corresponding spatial location $\mathbf{x}_n = |\mathbf{v}|t_n$ satisfies $|\mathbf{x}' - \mathbf{x}_n| \leqslant |\mathbf{v}|T$. Restricting the time window size to T will implicitly restrict the spatial locations contributing to the cross-correlation terms, and it will limit the crosstalk between incongruous source locations. Choosing the set of source locations $\mathbf{x}'$ such that $|\mathbf{x}' - \mathbf{x}_n| \leqslant |\mathbf{v}|T$ allows us to write a recording Equation (19) in a manner consistent with representation (3); namely

$$\hat{p}(\mathbf{x}, \omega\,;\,T, t_n) := \int_S \hat{F}(\mathbf{x}', \omega)\hat{G}(\mathbf{x}, \mathbf{x}', \omega)\,\chi_{\{|\mathbf{x}' - \mathbf{x}_n| \leq T|\mathbf{v}|\}}(\mathbf{x}')\,d\mathbf{x}', \tag{20}$$

where $\chi_A(\mathbf{x})$ is the indicator function of a set A; i.e., $\chi_A(\mathbf{x}) = 1$ if $\mathbf{x} \in A$ and $\chi_A(\mathbf{x}) = 0$ if $\mathbf{x} \notin A$. Then, applying the ambient noise retrieval Equation (2) to the time-windowed recordings (see Figure 7) for a window of duration T centered at t_n results in the windowed retrieval.

$$\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega\,;\,T, t_n) := \frac{2}{\rho c}\,\hat{p}^*(\mathbf{x_A}, \omega\,;\,T, t_n)\,\hat{p}(\mathbf{x_B}, \omega\,;\,T, t_n)$$

$$= \frac{2}{\rho c} \int_S \int_S \hat{F}^*(\mathbf{x}', \omega)\hat{F}(\mathbf{x}'', \omega)\hat{G}^*(\mathbf{x}, \mathbf{x}', \omega)\hat{G}(\mathbf{x}, \mathbf{x}'', \omega)$$

$$\times \chi_{\{|\mathbf{x}' - \mathbf{x}_n| \leq T|\mathbf{v}|\}}(\mathbf{x}')\,\chi_{\{|\mathbf{x}'' - \mathbf{x}_n| \leq T|\mathbf{v}|\}}(\mathbf{x}'')\,d\mathbf{x}'d\mathbf{x}'', \tag{21}$$

where the second equality is derived by substituting Equation (20) for each recording, and recasting the resulting product of integrals in a manner consistent with the form of Equation (4). Please note that as in Equation (4), the term $\hat{F}^*(\mathbf{x}', \omega)\hat{F}(\mathbf{x}'', \omega)$ is still present in Equation (21), and one could consider applying the standard stacking to Equation (21). However, if the resulting averaged/stacked source kernel $\langle\,\hat{F}^*(\mathbf{x}', \omega)\hat{F}(\mathbf{x}'', \omega)\,\rangle$ is not sufficiently localized (as discussed in Section 3), there are no guarantees that we will retrieve the correct inter-receiver phase. Instead, we consider a new averaging operator which relies on averaging over randomized window locations t_n. We denote this operator by $\langle\,\cdot\,\rangle_T$ and define it for a scalar function $\gamma : \mathbb{R} \to \mathbb{R}$ as
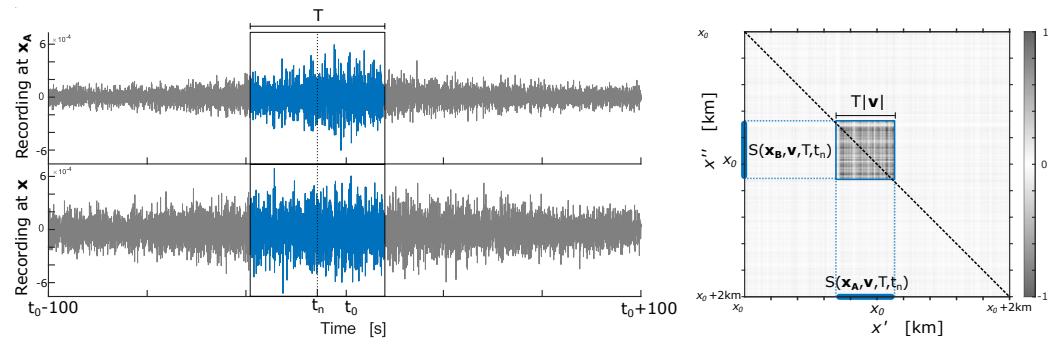
$$\langle\,\gamma\,\rangle_T := \int_{\mathbb{R}} \gamma(\tau)f_T(\tau)d\tau, \tag{22}$$

where f_T is the probability density function associated with the choice of locations t_n which is parameterized by T and assumed to be non-zero only on a bounded interval, and the function γ is integrable with respect to f_T. Next, we define the *averaged retrieval for a time window of duration T* as

$$\hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega) := \left\langle\,\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega\,;\,T, \cdot\,)\,\right\rangle_T$$

$$= \frac{2}{\rho c} \int_S \int_S \langle\chi\chi\rangle_T(\mathbf{x}', \mathbf{x}'')\hat{F}^*(\mathbf{x}', \omega)\hat{F}(\mathbf{x}'', \omega)\hat{G}^*(\mathbf{x}, \mathbf{x}', \omega)\hat{G}(\mathbf{x}, \mathbf{x}'', \omega)d\mathbf{x}'d\mathbf{x}, \tag{23}$$

where we assume that Fubini's theorem is satisfied and the *random windowing averaging kernel* is defined as

$$\langle\chi\chi\rangle_T(\mathbf{x}', \mathbf{x}'') := \left\langle\,\chi_{\{|\mathbf{x}' - \mathbf{x}_n| \leq T|\mathbf{v}|\}}(\mathbf{x}')\,\chi_{\{|\mathbf{x}'' - \mathbf{x}_n| \leq T|\mathbf{v}|\}}(\mathbf{x}'')\,\right\rangle_T. \tag{24}$$

Figure 7. Illustration of the random windowing procedure for a single realization of a window of size T in the temporal domain (**left**), and the corresponding window of length $T|\mathbf{v}|$ in the spatial domain (**right**) for $\mathbf{v} \neq 0$. The panels on the left show the waveforms recorded at a pair of receivers generated by a moving noise source. The approximate time at which the moving source is collinear with the receiver pair is denoted by t_0, and the location of a window of duration T (blue) is centered around the randomly drawn time t_n. The background in the inset on the right shows the spatial part of the source kernel $\langle \hat{F}^*(\mathbf{x}', \omega) \hat{F}(\mathbf{x}'', \omega) \rangle$ in Equation (6), where $\mathbf{x}_0$ corresponds to the spatial location of the moving source at time t_0. The boundary of the spatial window, centered at a random location $\mathbf{x}_n = t_n|\mathbf{v}|$, and its projections are highlighted in blue and denoted by $S(\mathbf{x}, \mathbf{v}, T, t_n)$ for each receiver; the resulting square window (of size $T|\mathbf{v}|$) is slightly off the diagonal. The parts of the source kernel excluded from this random window are grayed out.

Please note that given the above formulation and the ϵ-diagonal decomposition (10), we have

$$\hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega) \propto \hat{g}_\epsilon(\mathbf{x_A}, \mathbf{x_B}, \omega), \qquad \epsilon = 2T|\mathbf{v}|. \tag{25}$$

Furthermore, similarly to standard averaging via the stacking of recordings as in Equation (6), the structure of the random windowing averaging kernel determines the quality of the retrieval of the interferometric phase estimate from $\hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega)$. Note, in particular, that the random windowing averaging kernel $\langle \chi\chi \rangle_T(\mathbf{x}', \mathbf{x}'')$ is a symmetric function of the locations $\mathbf{x}', \mathbf{x}'' \in S$, and is concentrated in the neighborhood of $\mathbf{x}' = \mathbf{x}''$. Note also that the source speed $|\mathbf{v}|$ affects the properties of the random windowing averaging kernel. As suggested by the results illustrated in Figure 7 and confirmed by analytical estimates for a typical setup in Section 4.2, the random windowing procedure has an analogous effect to the ϵ-localization procedure in decomposition (13) in the sense that interactions between incongruous source locations are implicitly restricted to within the time/space window of the recordings before cross-correlating. Hence, the contribution of the acausal crosstalk is mitigated. Moreover, if sufficient data are available, one can additionally apply the standard stacking operation to $\hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega)$ which yields the general retrieval

$$\langle \hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega) \rangle = \frac{2}{\rho c} \int_S \int_S \langle \hat{F}^*(\mathbf{x}', \omega) \hat{F}(\mathbf{x}'', \omega) \rangle \langle \chi_{\{|\mathbf{x}' - \mathbf{x}_n| \leq T|\mathbf{v}|\}}(\mathbf{x}') \chi_{\{|\mathbf{x}'' - \mathbf{x}_n| \leq T|\mathbf{v}|\}}(\mathbf{x}'') \rangle_T$$

$$\times \hat{G}^*(\mathbf{x}, \mathbf{x}', \omega) \hat{G}(\mathbf{x}, \mathbf{x}'', \omega) \, d\mathbf{x}' d\mathbf{x}'', \tag{26}$$

since the stacking $\langle \cdot \rangle$ and the window randomization $\langle \cdot \rangle_T$ are independent by design, and therefore they commute. Importantly, the form of (26) implies that the random window average and the stacking average play, in principle, a similar role in the elimination of the crosstalk contributions in the ambient noise retrieval, in the spirit of the ϵ-diagonal decomposition (10). For $T \ll 1$ the random windowing kernel $\langle \chi\chi \rangle_T$ will concentrate the integrand in (26) to within the $\mathcal{O}(T|\mathbf{v}|)$ neighborhood of $\mathbf{x}' = \mathbf{x}''$. A similar effect is achieved if enough data are available, and the noise sources are such that the stacking average kernel $\langle \hat{F}^*\hat{F} \rangle$ localizes around $\mathbf{x}' = \mathbf{x}''$ (see Section 3 and Figure 4). In principle, both operations can be applied concurrently. However, in contrast to the stacking average,

the use of random windowing averaging relies on a single recording provided that a suitable value T of the time window can be determined; derivation of such a procedure described in Sections 4.3 and 4.4 is preceded by a simple example which is aimed at elucidating the main properties of the random windowing average.

4.2. Example: Uniformly Distributed Window Locations

Consider a default situation in which the temporal window location t_n is uniformly distributed within the interval $D = \left[t_0 - T,\ t_0 + T\right]$. We then have

$$f_T(t_n) = \frac{1}{2T}, \qquad \langle t_n \rangle_T = t_0, \qquad \mathrm{Var}(t_n) = \langle (t_n - t_0)^2 \rangle_T = T^2/3. \tag{27}$$

Then, if we define the spatial source location at time t_n as $\mathbf{x}_n = t_n\mathbf{v}$, we have $\langle \mathbf{x}_n \rangle_T = \mathbf{v}t_0 = \mathbf{x}_0$. Applying the operator $\langle \cdot \rangle_T$ with the uniform density f_T to the product of the indicator functions in Equation (24) corresponds to the convolution of two boxcar functions of constant height one and width $T|\mathbf{v}|$, parameterized in terms of the source location discrepancy $|\mathbf{x}' - \mathbf{x}''|$, $\mathbf{x}', \mathbf{x}'' \in S$.

A standard calculation shows that applying operator $\langle \cdot \rangle_T$ to the product of indicator functions in Equation (26) yields a symmetric kernel in the random windowing averaging of the form

$$\langle \chi_{\{|\mathbf{x}'-\mathbf{x}_n|\leq T|\mathbf{v}|\}}(\mathbf{x}')\,\chi_{\{|\mathbf{x}''-\mathbf{x}_n|\leq T|\mathbf{v}|\}}(\mathbf{x}'') \rangle_T = \left(|\mathbf{x}' - \mathbf{x}''| + T|\mathbf{v}| \right)\Theta\left(|\mathbf{x}' - \mathbf{x}''| + T|\mathbf{v}| \right)$$
$$+ \left(|\mathbf{x}' - \mathbf{x}''| - T|\mathbf{v}| \right)\Theta\left(|\mathbf{x}' - \mathbf{x}''| - T|\mathbf{v}| \right)$$
$$- 2|\mathbf{x}' - \mathbf{x}''|\,\Theta\left(|\mathbf{x}' - \mathbf{x}''| \right), \tag{28}$$

which represents the tent map supported on the interval $-2T|\mathbf{v}| \leqslant \mathbf{x}' - \mathbf{x}'' \leqslant 2T|\mathbf{v}|$ and centered at $\mathbf{x}' - \mathbf{x}'' = 0$, where Θ denotes the Heaviside function. An example of Equation (28) is plotted in the right inset in Figure 8, where the kernel (28) is plotted in blue, and its approximation obtained via direct simulations is indicated by the dotted black line.

It is informative to compare the kernel $\langle \chi\chi \rangle_T$ of the random windowing average in Equation (28) illustrated in Figure 8 with the stacking average kernel $\langle \hat{F}^*\hat{F} \rangle$ in Equation (7) which is illustrated in Figure 4. Recall that both these kernels are present in the retrieval (26) and that they play, in principle, a similar role aimed at reducing the contribution from the off-diagonal, crosstalk terms into the retrieval. One can see that the random windowing procedure is akin to extraction of the ϵ-diagonal term $\hat{g}_\epsilon$ in the decomposition (10), in the sense that for any pair of sources such that $|\mathbf{x}' - \mathbf{x}''| > 2T|\mathbf{v}|$ the random windowing kernel mutes the corresponding acausal contributions. For uniformly distributed locations of the random windows around t_0, only the contributions satisfying $|\mathbf{x}' - \mathbf{x}''| \leqslant 2T|\mathbf{v}|$ are included in the resulting retrieval (26). Moreover, the analytical expression (28) implies that the width of the diagonal band induced by the random windowing procedure is directly proportional to the source speed $|\mathbf{v}|$. This suggests that a progressively shorter time window size T should be chosen for increasing train speeds to mitigate acausal arrivals associated with the crosstalk terms in the ambient noise retrieval (10).

4.3. Estimation of the Retrieval $\hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega)$ for a Given Time Window Size T

In this section, we describe a procedure for the estimation of the random windowing average $\langle \cdot \rangle_T$ with some fixed time window length $T > 0$. A systematic way of choosing the optimal time window size T_{opt} for any given single recording by means of estimating the acausal energy present in each averaged retrieval $\hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega)$ in Equation (23) is discussed in Section 4.4.
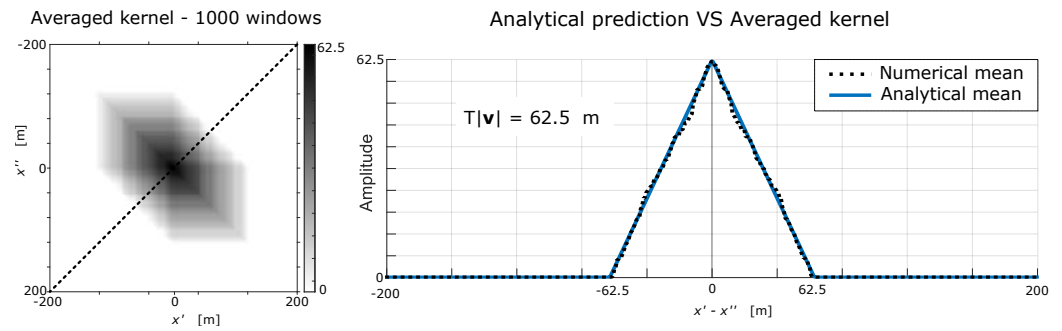
Figure 8. The structure of the kernel $\langle \chi\chi \rangle_T$ in Equation (24) of random windowing averaging (23) based on the uniform distribution of window locations. The kernel is given by a tent map (28) with height $T|\mathbf{v}|$ and width $2T|\mathbf{v}|$, centered at $\mathbf{x}' = \mathbf{x}''$. The right panel shows the kernel parameterized in terms of the distance from the source locations x' (for receiver $\mathbf{x_A}$) and x'' (for receiver $\mathbf{x_B}$) to the location of the source boundary with the receiver line (see the parametrization used in Figure 4); the analytical solution (28) is shown in blue and its approximation within the anti-diagonal section of the random windowing kernel (see the left panel) obtained via direct simulations is indicated by the dotted black line. The left panel shows the numerical approximation of the random window kernel based on 1000 windows for a single source train recording travelling at speed $|\mathbf{v}| = 25$ m/s (90 km/h), with $T = 2.5$ s. Compare the structure of the above random windowing kernel $\langle \chi\chi \rangle_T$ with the stacking average kernel $\langle \hat{F}^* \hat{F} \rangle$ illustrated in Figure 4. Both kernels are present in the general retrieval (26).

As outlined in Section 4.1 and illustrated in Section 4.2 the random windowing procedure will lead to an improved phase retrieval through damping the contributions from the cross-correlation terms in Equation (26) provided that a suitable temporal window length $T > 0$ is identified, so that the random windowing kernel $\langle \chi\chi \rangle_T$ in Equation (23) effectively enforces the recovery of the diagonal band term $\hat{g}_\epsilon$ in the decomposition (10). In the case of standard stacking the choice of the spatial correlation length cut-off ϵ can be inferred from the form of stacking average kernel $\langle \hat{F}^* \hat{F} \rangle$ in Equation (7), assuming that the kernel is sufficiently localized. Clearly, the structure of the stacking average kernel cannot be determined from a single short recording or in situations when the stacking average kernel can be estimated but it is not sufficiently localized (e.g., as in the example presented in Figure 6). On the other hand, the random windowing average with an appropriately chosen T can be applied to obtain the retrieval $\hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega)$ in the most challenging case where only a single (possibly short) recording of a source train is available. Moreover, as mentioned earlier and expressed via Equation (26), if enough recorded data are available, one can combine both averaging techniques to recover $\langle \hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega) \rangle$. It is clear that the use of such a doubly averaged approach is more versatile than using only standard stacking in the absence of knowledge about the structure of the noise sources, and it will lead to an improved retrieval, as long as at least one of the kernels is sufficiently concentrated around the set $\mathbf{x}' = \mathbf{x}''$, $\mathbf{x}', \mathbf{x}'' \in S$.

The first step in the procedure for the empirical approximation of the random windowing average $\langle \cdot \rangle_T$ with some fixed time window length T is to identify the approximate time in the recorded traces when the source train is collinear with the receiver pair (see for instance [31] for one example of a method that achieves this using plane-wave beamforming). We chose a configuration in which the line through the receiver pair is perpendicular to the source boundary, as depicted in Figure 2. In such a geometry the Fresnel zone of the receiver pair is symmetric about the stationary phase point (see [40]) on the source boundary but this is not essential to the random windowing method. To see this note that when the instantaneous source location is collinear with the receiver pair the signal recorded at the two receivers is generated at $\mathbf{x}' = \mathbf{x}''$, and the frequency recorded at each receiver, as well as the Doppler bias, are the same ($\theta_{\mathbf{x_A}}(t') = \theta_{\mathbf{x_B}}(t')$ in Equation (5)). Thus, in this instantaneous configuration the effect of spurious contributions to the retrieval is minimized. Finally, note that regardless of whether or not the noise source is in motion, the

source location contained in the line through the receiver pair coincides with the stationary phase point (on the source boundary) which is always contained in the Fresnel zone of the receiver pair, unless the line through the receivers is parallel to the source boundary.

The time tag at which the source is collinear with the receivers is denoted by t_0 (see left panels of Figure 7) and the corresponding source location is denoted by $\mathbf{x}_0 = \mathbf{v}t_0$. Please note that while t_0 is a known time tag, explicit knowledge of $\mathbf{x}_0$ is not necessary in practice. However, this notation will be useful when describing our method in the steps below. We note further that $\mathbf{x}_0$ will by construction be in the middle of the Fresnel zone of the receiver pair. Note also that for a given recording of a single source moving in one direction, the time tag t_0 remains fixed throughout the procedure.

To estimate $\hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega)$ consider the sequence $n = 1, \ldots, N$, $N \gg 1$, and perform the following steps for the given pair of recordings $p(\mathbf{x_A}, t)$ and $p(\mathbf{x_B}, t)$:

(i) *Generate a set of random window locations* $\{t_n\}_{n=1}^N$ *of a fixed duration* T *by sampling from the distribution associated with* f_T *in Equation* (22), *and such that* $|t_n - t_0| < T$. *There is flexibility in the choice of the probability distribution* f_T *of* t_n *provided that* $\langle t_{(\cdot)} \rangle_T = t_0$ *and that* f_T *is non-zero only on a bounded interval.*

(ii) *For each* t_n *extract the windowed recordings* $p(\mathbf{x_A}, t; T, t_n)$ *and* $p(\mathbf{x_B}, t; T, t_n)$ *of duration* T *and centered at* t_n *from* $p(\mathbf{x_A}, t)$ *and* $p(\mathbf{x_B}, t)$. *This procedure is illustrated in the left panels of Figure* 7 *and is performed on the recorded time series for each receiver.*

(iii) *Perform interferometry using the windowed data and derive a realization of* $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega; T, t_n)$ *according to Equation* (23), *using the Fourier transforms of the windowed recordings from* (ii).

(iv) *Repeat Steps* (i)–(iii) *for a large number* N *of random window locations* t_n *and take the arithmetic mean of* $\{\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega; T, t_n)\}_{n=1}^N$ *in order to approximate the random windowing kernel* $\langle \chi\chi \rangle_T$ *given by Equation* (24), *and to obtain the ambient noise retrieval* $\hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega)$ *in Equation* (23).

Due to the law of large numbers, one has

$$\frac{1}{N} \sum_{n=1}^N \hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega; T, t_n) \xrightarrow[N \to \infty]{} \hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega), \tag{29}$$

and the empirical mean of the sample $\{\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega; T, t_n)\}_{n=1}^N$ converges, in the limit of $N \to \infty$, to the expectation of $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega; T, \tau)$ as in Equation (22) provided that τ has a finite variance (i.e., $\int_{\mathbb{R}} \tau^2 f_T(\tau) d\tau < \infty$) and the locations $\{t_n\}_{n=1}^N$ are sampled independently from f_T. Finally, note that the steps outlined above do not require knowledge of the source speed $\mathbf{v}$, or for the source to be in motion, as long as it is possible to identify a suitable recorded time t_0 corresponding to the emission of energy from a source location collinear with the receiver pair.

4.4. Procedure for Choosing the Optimal Time Window Size T_{Opt} in the Random-Window Averaged Retrieval $\hat{h}_{T_{opt}}(\mathbf{x_A}, \mathbf{x_B}, \omega)$

The procedure described in the previous section can be applied to any time window of size $0 < T < \infty$. However, similarly to the general ϵ-diagonal decomposition defined in Equation (10), a successful approximation of the phase of the inter-receiver Green's function depends on the choice of an optimal support of the random windowing kernel $\langle \chi\chi \rangle_T$ in Equation (24). The steps outlined below provide a systematic way to choose an optimal time window size T_{opt} for any given single recording by estimating the acausal energy present in each averaged retrieval $\hat{h}_T(\mathbf{x_A}, \mathbf{x_B}, \omega)$ in Equation (23).

To determine T_{opt} for a given pair of recordings $p(\mathbf{x_A}, t)$ and $p(\mathbf{x_B}, t)$, consider first a sufficiently large range of test values $T \in (T_{\min}, T_{\max})$; this range will necessarily be limited by the length of the recordings available. Steps to find the optimal window size T_{opt} for any fixed pair of recordings $p(\mathbf{x_A}, t)$ and $p(\mathbf{x_B}, t)$ are:

(1) Choose an ordered collection of time window sizes $\{T^{(j)}\}_{j=1}^{J}$ contained in the test range $T \in (T_{\min}, T_{\max})$.

(2) Estimate the averaged retrieval $\hat{h}_{T^{(j)}}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ for each time window size $\{T^{(j)}\}_{j=1}^{J}$ using the procedure described in Section 4.3.

(3) Transform each estimate $\{\hat{h}_{T^{(j)}}(\mathbf{x_A}, \mathbf{x_B}, \omega)\}_{j=1}^{J}$ to the time domain to obtain the corresponding waveforms $\{h_{T^{(j)}}(\mathbf{x_A}, \mathbf{x_B}, t)\}_{j=1}^{J}$.

(4) Order the waveforms $h_{T^{(j)}}(\mathbf{x_A}, \mathbf{x_B}, t)$ according to the time window size $T^{(j)}$ (see bottom panel of Figure 9).

(5) Use the ordered family of waveforms $h_{T^{(j)}}(\mathbf{x_A}, \mathbf{x_B}, t)$ to identify the time tag of the *feature of interest*, which is normally the first-arriving dominant wave, denoted by t_I. In Figure 9, $t_I = 0.4$ s and is highlighted in orange.

(6) Calculate the waveform energy of each trace $h_{T^{(j)}}(\mathbf{x_A}, \mathbf{x_B}, t)$ from time $t = 0$ to a time slightly before t_I. This calculation quantifies the acausal energy present in each of the traces as a function of $T^{(j)}$. (In Figure 9 the waveform energy is calculated for each trace between $t = 0$ and $t = 0.3$ s. Please note that this curve would be different for different recordings.)

(7) Compute the spurious waveform energy calculated in Equation (4) as a function of the time window sizes $T^{(j)}$ (see the top panel of Figure 9).

(8) Minimize the spurious energy with respect to $T^{(j)}$; the minimizer is denoted by T_{opt}.

(9) Use the optimal window size T_{opt} determined in Equation (8) for the interferometric retrieval by following the steps described in Section 4.3. Compute the final/optimal ambient noise retrieval of the inter-receiver Greens function $\hat{h}_{T_{\mathrm{opt}}}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ given by Equation (23) or its time-domain counterpart $h_{T_{\mathrm{opt}}}(\mathbf{x_A}, \mathbf{x_B}, t)$.
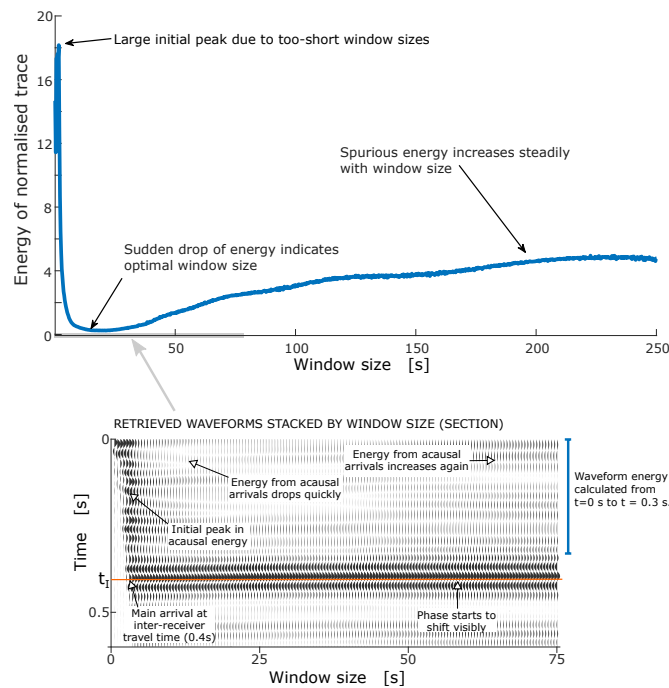
Figure 9. Calculation of acausal energy curve for a recording of a single moving source. Values of T tested ranged between 0.25 s and 250 s. The bottom panel shows the waveforms corresponding to the average random window retrieval for T only in the range between 0.25 s and 75 s, for clarity. The waveforms are arranged by window size T. The arrival of interest is highlighted in orange at time $t_I = 0.4$ s. The top panel shows the acausal energy curve in blue. The waveform energy was calculated for each waveform in the range from t = 0 to t = 0.3 s. The acausal energy curve exhibits a clear initial peak for small values of T, followed by a sharp drop and then a slow increase with T. The minimum is achieved at $T_{\mathrm{opt}} \approx 15$ s.

Please note that the above procedure can easily be automated. For multiple source train recordings Steps (1)–(8) can be performed for each recording, noting that the optimal window size T_{opt} will, in general, be different for each recording. Then, the standard stacking average can be applied to estimate $\langle \hat{h}_{T_{opt}}(\mathbf{x_A}, \mathbf{x_B}, \omega) \rangle$ and thus potentially improve the retrieval even further.

4.5. Results

We now present some examples of the interferometric retrieval obtained via the random windowing method described in Section 4.3 which is carried out for the optimal time window size T_{opt} determined according to the procedure described in Section 4.4. We present results for a receiver pair that is perpendicularly oriented with respect to the source-line, as well as results for a receiver pair that is oriented at 30 degrees with respect to the source-line. The results are of essentially identical quality.

The distribution of the window locations is chosen to be uniform, as in the example of Section 4.2. The random windowing method was applied to the two configurations discussed in Section 3. These configurations corresponded to (i) the case where the stacking average kernel $\langle \hat{F}^* \hat{F} \rangle$ of the ambient noise is localized, as in Figure 5 (left panels of Figure 10), and (ii) the challenging scenario when $\langle \hat{F}^* \hat{F} \rangle$ is not localized and the spurious crosstalk arrivals do not decay (right panels of Figure 10). The random windowing method leading to $\hat{h}_{T_{opt}}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ is applied in all cases to a single short recording of a train of sources, and the stacking average is not applied to $\hat{h}_{T_{opt}}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ unless otherwise stated. Figure 10 shows phase estimates produced by the random windowing average in blue and the phase estimate based on $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ in light gray for comparison. Receivers oriented perpendicularly with respect to the source-line were used in the top row, whereas the bottom row of Figure 10 shows the results for a receiver pair oriented at 30 degrees with respect to the source-line. The phase estimated from individually recorded sources, which allows us to calculate the reference retrieval using Equation (1), is indicated in dashed black. We also show the phase of the inter-receiver Green's function retrieved through the standard stacking average over six source train recordings restricted to a one-minute-long window in a similar fashion to [34]. For illustration, Figure 11 shows the waveforms corresponding to the phases depicted in the top-right panel of Figure 10 (i.e., a correlated noise source with correlated phases, and perpendicularly oriented receivers) with normalized amplitudes. As expected, the waveform corresponding to the random windowing estimate (in blue) exhibits the best agreement with the reference waveform (in black dashes). The phases estimated from the random windowing scheme agree well with the reference phase in all cases, showing that our method is effective in mitigating the undesired effects of correlation in the noise sources in ambient noise interferometry, especially where moving sources such as trains or traffic on a highway are concerned.

We note that the workflow outlined in Section 4.4 and illustrated above is not restricted to ambient noise interferometry from correlated noise. In fact, this methodology can be applied to short recordings of uncorrelated noise to find an optimal window size that minimizes spurious arrivals. In this scenario, the choice of t_0 as described in Section 4.3 is more flexible as the source no longer exhibits coherent spatiotemporal correlation or a Doppler spread differential, as would be the case for a source in motion. Therefore, the choice of time tag t_0 may be arbitrary and the random windowing method can still be applied as described in Section 4.4. Figure 12 shows the phase estimated by the random windowing method when the ambient noise recording is uncorrelated. The random windowing method was applied to a short (5 min) recording of uncorrelated ambient noise, producing satisfactory results. Our approach has the potential to speed up convergence to the inter-receiver phase when applied to uncorrelated ambient noise.
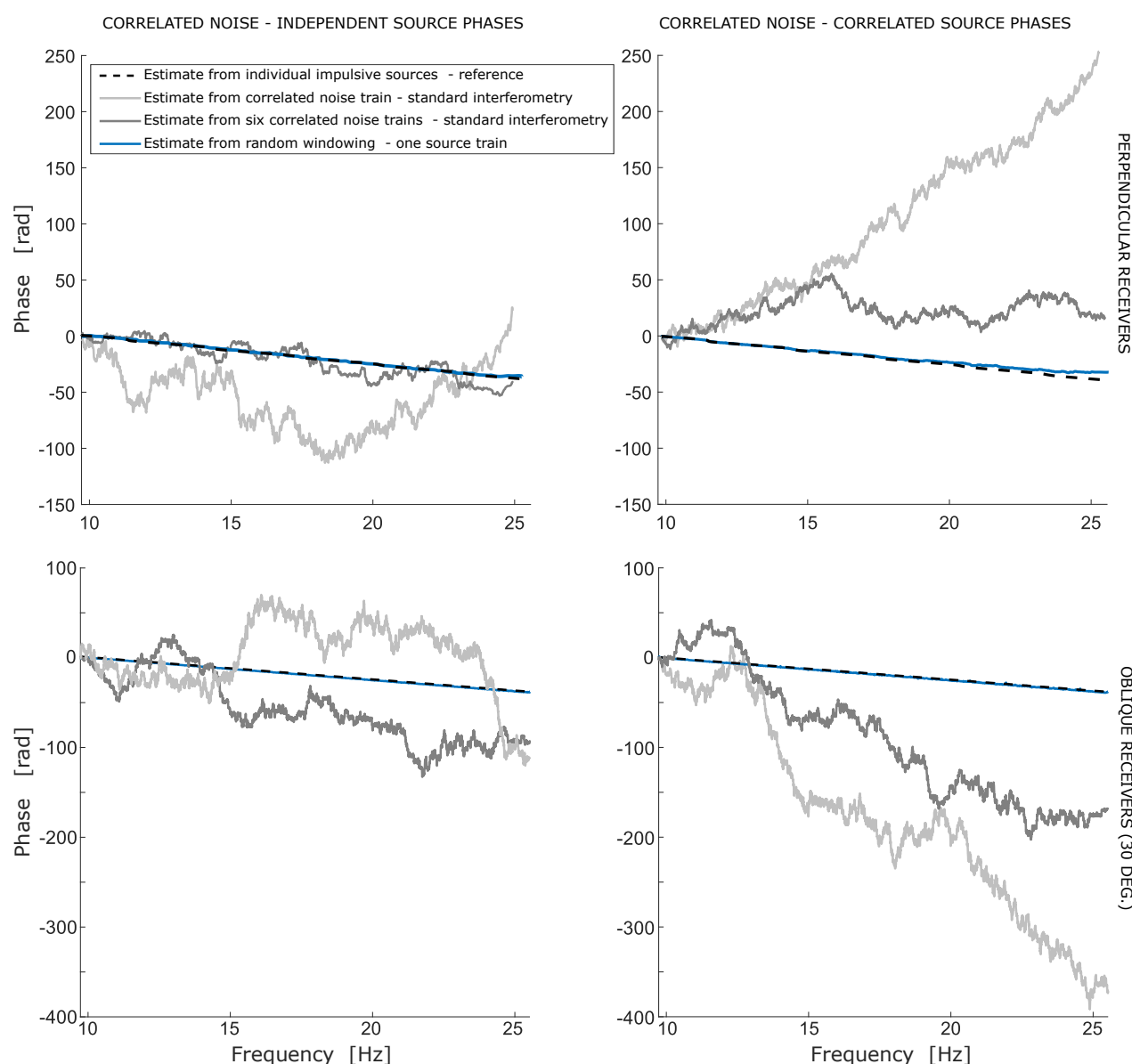
Figure 10. Phase estimates obtained from the random windowing scheme applied to a short (5 min) recording of a single source using receivers that are perpendicularly oriented with respect to the source-line (**top**) and oriented at 30 degrees with respect to the source-line (**bottom**). The left column shows the results for correlated noise where source phases at different frequencies are independent (as in Figure 4). The phase estimated from discrete sources, in black dashes, is used for reference. The phases produced by the random windowing scheme are shown in blue and exhibit very good agreement with the reference phase. The phase of the retrieval using the standard method and no stacking is shown in light gray, and the phase of six stacked sources is shown in dark gray for comparison. The same phase estimates are shown in the right column for correlated noise with correlated source phases (as shown in Figure 6). The phase estimates from the random windowing scheme, in blue, are better than the retrieval obtained using the standard method, with slightly reduced accuracy at higher frequencies.
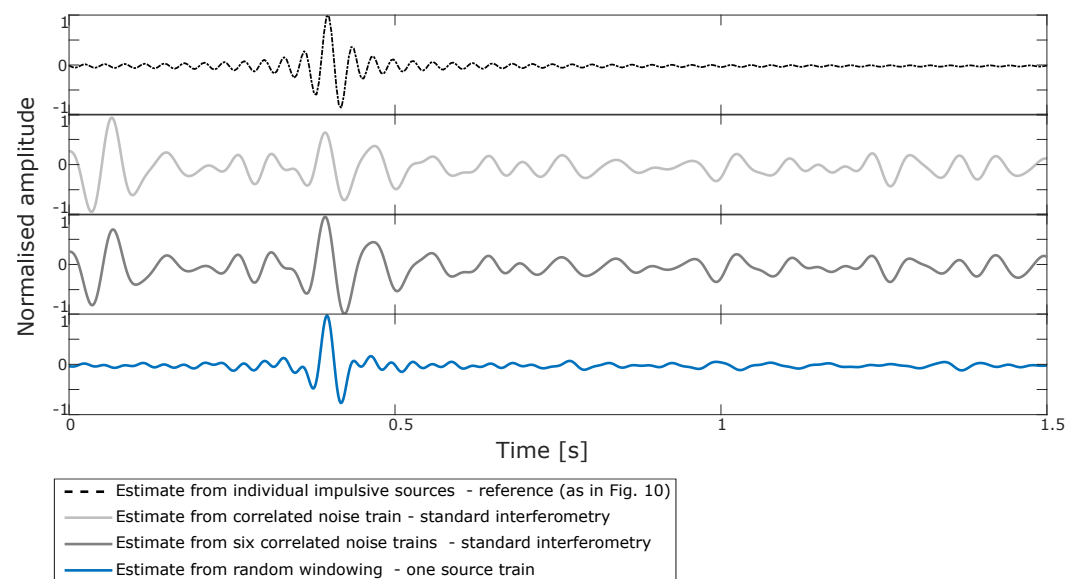
Figure 11. Waveforms corresponding to the estimated phases shown in the top-right panel of Figure 10, for correlated noise with source phases recorded at a perpendicularly oriented pair of receivers 400 m apart, with the medium speed set at 1000 m/s. For reference, the estimate calculated from individual point sources is depicted in black dashes (**top**), exhibiting the inter-receiver arrival at 0.4 s. The estimate produced by standard interferometry without stacking is shown in light gray (**second from top**). The waveform resulting from standard stacking of six estimates is depicted in dark gray (**third from top**). The estimate obtained from the random windowing method applied to a short recording of a single source is shown in blue (**bottom**).
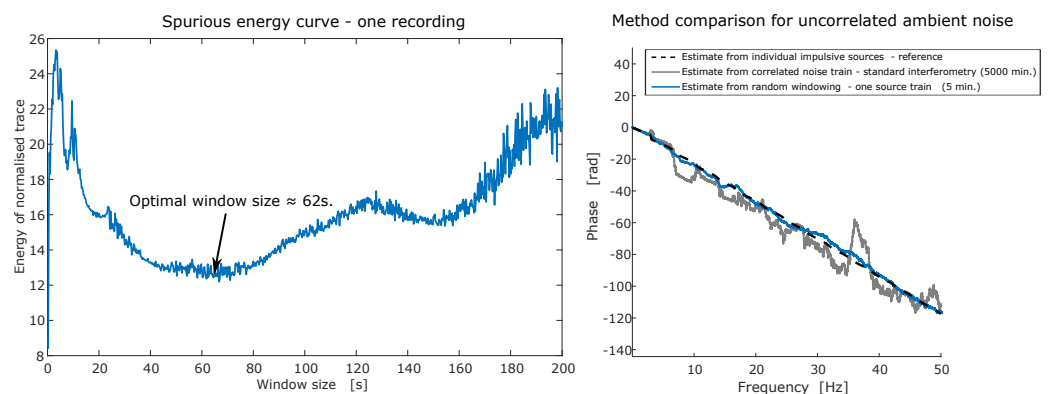
Figure 12. Random windowing method applied to uncorrelated ambient noise. The plot to the left shows the acausal energy curve estimated from one ambient noise recording. The plot to the right shows the resulting estimated phase in blue, which matches the reference phase in black dashes well. For comparison, the phase estimated from stacking 1000 simulated recordings of ambient noise five minutes long is shown in dark gray. We can see the stacking in this case will converge to the correct phase estimate as expected for truly uncorrelated ambient noise, but only after a very large number of ambient noise recordings are processed and stacked.

Finally, Figure 13 shows the mean optimal window sizes T_{opt} as a function of the distance from the closest receiver to the source-line; in the illustrated case the inter-receiver distance is 400 m, and three different source speeds are considered. Recall that optimal window sizes T_{opt} are, to some extent, recording-dependent due to the randomness in recorded signals, so their value will vary for different recordings even if the receiver geometry remains fixed; hence, 30 samples for each considered closest receiver distance and velocity were calculated and averaged to produce the curves in Figure 13; the fluctuations in the curves diminish with the number of samples used. The relationship illustrated in

Figure 13 indicates that for a fixed source speed and a fixed inter-receiver distance, the optimal window size decreases with the distance of the receivers to the source boundary. Moreover, the optimal window size decreases for increasing source speeds, which agrees with physical intuition and with the form of the random windowing average kernel (28) illustrated in Figure 8. Furthermore, as can be deduced from Equations (5) and (A11), the spatial discrepancy of the implicit source locations (i.e., the difference $x' - x''$) will be smaller for decreasing inter-receiver distances. Consequently, the error due to the presence of the residual crosstalk terms in the retrieval decreases with decreasing inter-receiver distance, while larger inter-receiver distances require smaller values of the time window size T_{opt} to mitigate this error.
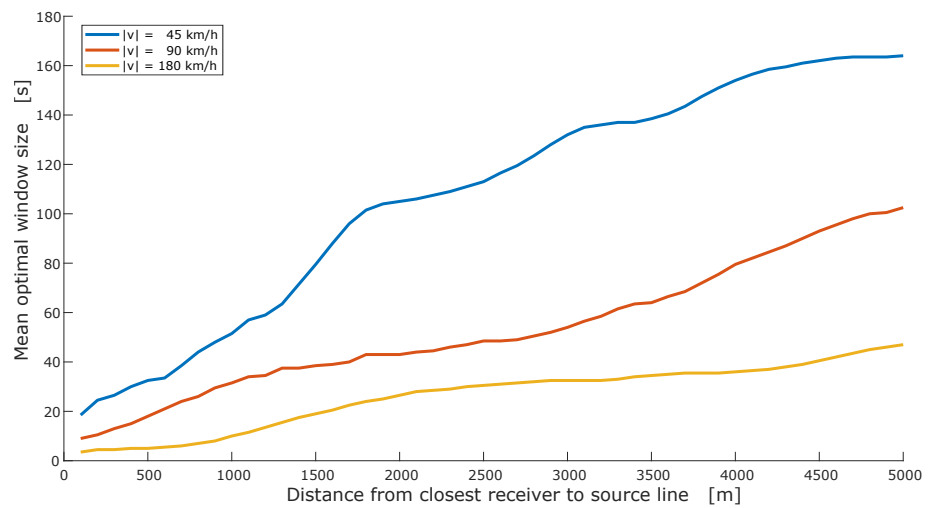
Figure 13. Estimated optimal window size for the random windowing scheme, calculated for trains travelling at 45 km/h, 90 km/h and 180 km/h, in terms of the distance between the receivers and the road (using the closest receiver as reference) and averaged over 30 samples each. The receivers were 400 m apart from each other.

5. Discussion

The ambient noise interferometric retrieval based on the random windowing average (Sections 4.3 and 4.4) is capable of accurately estimating the phase of the inter-receiver Green's function based on a short pair of recordings of a correlated noise source (Section 4.5). As pointed out in Equation (25), the random windowing average allows one to extract the ϵ-diagonal term in decomposition (10) of the ambient noise retrieval $\hat{h}(\mathbf{x_A}, \mathbf{x_B}, \omega)$ without the explicit knowledge of the stacking average kernel $\langle \hat{F}^* \hat{F} \rangle$ which need not be localized. For a given pair of recordings this new technique requires one to pre-process the spurious energy to determine the optimal time window size T_{opt}; then, the optimal spatial window length can be used in decomposition (10) to retrieve $\hat{g}_\epsilon$ with the optimal cut-off length $\epsilon = T_{opt}|\mathbf{v}|$. This procedure, illustrated in Figure 9, can be automated in practical applications. Based on the derivations and the discussion in Sections 3 and 4, the optimal value T_{opt} of the temporal window size decreases with the inter-receiver distance and with the distance between the source-line and the receiver pair, since the Doppler shift and the spatial incongruence become increasingly important. Moreover, the optimal value of the temporal window size decreases for increasing speeds of the moving noise source, as indicated in Equation (28). These trends are consistent with physical intuition, as faster source speeds have the compound effect of increasing the spatial coverage (and hence the prevalence of the crosstalk contributions to the ambient noise retrieval), and increasing the importance of the Doppler shift for the retrieval.

The random windowing approach can be used in conjunction with the standard stacking average procedure as discussed in Section 4. This is especially tractable when estimating inter-receiver Green's functions from a small ensemble of source train recordings, where the random windowing procedure can be applied to each individual train recording

before performing standard stacking to optimize the extraction of causal information. In Ref. [34], the distance from the receiver array to the rail track on which trains provide source energy is of the order of 5–10 km. This large distance combined with the stationary phase zone approach likely mitigates spurious effects arising from the Doppler effect induced by the trains' motion, since such a receiver array configuration and their localization methodology implicitly restrict the recordings to regions of the source boundary on which the train is roughly collinear with the receiver pairs. In fact, the choice of the optimal time window size in [34] in the stacked retrieval is consistent with our findings. Incorporating our methodology in such a setup could allow for the retrieval of inter-receiver estimates from near-field recordings (closer to the rail tracks) maintaining the optimality of extraction of coherent information from each recorded trace.

Sabra [41] concludes that under some circumstances the Doppler effect does not have a significant effect on the interferometric estimate of the inter-receiver Green's function. However, that work explicitly assumes that the ambient noise field is uncorrelated in time. To achieve this the sources considered were randomly and uniformly distributed in space around the receiver pair. The situation where the ambient noise field is statistically correlated is not considered. Moreover, to mitigate the differential Doppler effect, a stationary phase approximation is performed throughout, and the far-field assumption is made. These approaches are consistent with our findings, as well as with the random windowing method presented in this work.

In interferometric studies from traffic noise, such as the one performed by Behm et al. [30], wave velocities are retrieved from a linear array of receivers which crosses the main road perpendicularly. In this setup, the choice of the time window sizes could be further informed by our proposed framework, possibly using consistently shorter time windows for receiver pairs which are closer to the road. This interesting study points out that longer observation periods do not necessarily lead to better results, which indicates that correlations in the underlying ambient noise are too significant for the stacked-averaged retrieval to be appropriate, analogously to the case illustrated in Figure 6. Thus, integrating the random windowing approach could be beneficial by providing an informed method for selecting temporal window sizes so that the effect of spurious arrivals is minimized, or even averaging over different time window sizes. Liu et al. [35] conclude, using physically motivated ray-path arguments, that the crosstalk can be negligible when performing interferometry using seismic noise generated by high-speed trains in the case of direct, scattered and refraction waves. Our approach to interferometry from correlated noise sources links the quality of the retrieved phase of the inter-receiver Green's function from direct waves explicitly to the correlation structure of the noise source, as well as the receiver array geometry (inter-receiver distance and distance to the source-line) while further providing a methodology to mitigate the spurious effects induced by any of these causes. In the common case of sources in motion, we established a link between the spurious energy in the retrieved trace and the speed of the source, and the proposed methodology has the potential to mitigate spurious effects induced by the Doppler spread for this type of source. We further note that the receivers in [35] are very close to the location of the source train, confirming that correlated noise interferometry is applicable in the near-field of seismic noise generated by road traffic and trains.

We constrained our study to estimating the direct wave arrivals in two dimensions, which loosely resembles the case of surface waves in the three-dimensional setup. Similar derivations could be carried out for body waves that have been critically refracted in three dimensions along horizontal interfaces in the subsurface and then refract back up to the surface. Our methodology could be used to image and monitor the subsurface using these non-physical refracted arrivals, as suggested by King and Curtis [50] and Brenguier et al. [32], respectively.

Finally, our method is optimized when the location of the source in motion can be identified from a location tracker or from the acquired data. The latter approach has been shown to be possible for example in [32,34], but may not be feasible in an

urban environment with a more complicated road network and multiple simultaneous sources. In those situations, one could apply directional decomposition of the wavefield into constituent energy from its different individual sources prior to using these methods, similarly to the methods employed by Maranò et al. [51]. Although we have focused on the application of interferometry from noise sources in motion, representative of most known physical sources that contribute to ambient noise, the random windowing method can be used regardless of the underlying cause of statistical correlation in the ambient noise field.

6. Conclusions

We developed a novel, robust and general procedure for accurate retrieval of the phase of the inter-receiver Green's function from a short recording of a single source of ambient noise, without the need to rely on the commonly used stacking average procedure, and thus avoiding the potential pitfalls of standard ambient noise interferometry. The random windowing approach provides a versatile framework for ambient noise interferometry, and it should prove useful either as a standalone technique when only short recordings of correlated noise are available or in situations where the standard stacking operation fails due to significant correlations in the recorded signal and/or insufficiently long recordings of the seismic noise. The random windowing method mitigates the spurious crosstalk effects introduced by the presence of correlations in the ambient noise, regardless of the cause of the correlation. It includes the common case of correlation caused by a spatially migrating train of sources, and in that case it also potentially mitigates the bias introduced by the Doppler effect. This approach is general in the sense that it improves the quality of the phase of the inter-receiver Green's function without detriment to retrieval obtained using the standard stacking average methodology, and it applies to both correlated and uncorrelated ambient noise in both near and far-field configurations. Thus, this unified framework offers a novel workflow that can be deployed for robust estimation of inter-receiver phases from ambient noise interferometry, regardless of the characteristics of the ambient noise and, in general, using less data than are required for standard interferometry that uses the stacking average approach.

Appendix A. Analytical Derivation of Moving Point Source Model

Throughout this appendix, bold variables denote vectors whereas regular variables denote scalar quantities. We calculate the wavefield generated by a monochromatic point source that moves with constant speed $\mathbf{v}$ in a two-dimensional homogeneous medium, emitting sound at a constant angular frequency $\tilde{\omega}$. This wavefield is denoted by $p_{\tilde{\omega}}(t, \mathbf{x}, \mathbf{v})$

when measured at time t and at a receiver location $\mathbf{x}$. We may also refer to the constant source frequency as $\tilde{f}$ when we wish to express it in Hertz, so that $\tilde{\omega} = 2\pi\tilde{f}$. We characterize the moving point source through the function

$$s_{\tilde{\omega}}(t, \mathbf{x}, \mathbf{v}) = e^{-i\tilde{\omega}t}\delta(\mathbf{x} - \mathbf{v}t). \tag{A1}$$

It is assumed without loss of generality that at time $t = 0$ the monochromatic source is located at the origin. It is further assumed that the medium density ρ and sound speed c are constant. Then the wavefield, denoted by $p_{\tilde{\omega}}$, satisfies the linear wave equation

$$\frac{1}{c^2}\frac{\partial^2 p_{\tilde{\omega}}}{\partial t^2} - \nabla^2 p_{\tilde{\omega}} = \rho\frac{\partial s_{\tilde{\omega}}}{\partial t} \tag{A2}$$

where we restrict ourselves to real solutions. To solve Equation (A2) we introduce the abstract complex potential $\phi_{\tilde{\omega}}^{\mathcal{C}}(t, \mathbf{x}, \mathbf{v})$ of the particle velocity, so that

$$p_{\tilde{\omega}} = \rho\,\mathfrak{R}\left\{\frac{\partial\phi_{\tilde{\omega}}^{\mathcal{C}}}{\partial t}\right\}. \tag{A3}$$

This potential satisfies the wave equation

$$\frac{1}{c^2}\frac{\partial^2\phi_{\tilde{\omega}}^{\mathcal{C}}}{\partial t^2} - \nabla^2\phi_{\tilde{\omega}}^{\mathcal{C}} = s_{\tilde{\omega}}, \tag{A4}$$

which can be derived directly from Equation (A2) by integrating once with respect to time and cancelling a factor of ρ. Please note that if $\phi_{\tilde{\omega}}^{\mathcal{C}}$ is a solution to Equation (A4), then $p_{\tilde{\omega}}$ as defined by Equation (A3) will satisfy our original Equation (A2).

We can immediately write down the solution to Equation (A4) by convolving the source term Equation (A1) with the outgoing Green's function of the wave equation in time and space for a volume-injection source, given by

$$G(t, \mathbf{x}) = \frac{\Theta\left(t - \frac{|\mathbf{x}|}{c}\right)}{2\pi\sqrt{t^2 - \frac{|\mathbf{x}|^2}{c^2}}}, \tag{A5}$$

where Θ denotes the Heaviside function. Hence, the complex potential $\phi_{\tilde{\omega}}^{\mathcal{C}}(t, \mathbf{x}, \mathbf{v})$ can be expressed in integral form as

$$\phi_{\tilde{\omega}}^{\mathcal{C}}(t, \mathbf{x}, \mathbf{v}) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{\mathbb{R}^2} e^{-i\tilde{\omega}t'}\delta(\mathbf{x}' - \mathbf{v}t')\frac{\Theta\left(t - t' - \frac{|\mathbf{x}-\mathbf{x}'|}{c}\right)}{\sqrt{(t-t')^2 - \frac{|\mathbf{x}-\mathbf{x}'|^2}{c^2}}}\,d\mathbf{x}'\,dt'. \tag{A6}$$

Performing the integral in space is straightforward owing to the sifting property of the delta distribution, yielding

$$\phi_{\tilde{\omega}}^{\mathcal{C}}(t, \mathbf{x}, \mathbf{v}) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\tilde{\omega}t'}\frac{\Theta\left(t - t' - \frac{|\mathbf{x}-\mathbf{v}t'|}{c}\right)}{\sqrt{(t-t')^2 - \frac{|\mathbf{x}-\mathbf{v}t'|^2}{c^2}}}\,dt'. \tag{A7}$$

To perform the time integral in Equation (A7), define the auxiliary functions

$$q_1(t, \mathbf{x}, \mathbf{v}) = \frac{c^2 t - \mathbf{x}\cdot\mathbf{v}}{c^2 - |\mathbf{v}|^2}, \tag{A8}$$

and

$$q_2(t, \mathbf{x}, \mathbf{v}) = q_1(t, \mathbf{x}, \mathbf{v})^2 - \frac{c^2 t^2 - |\mathbf{x}|^2}{c^2 - |\mathbf{v}|^2}. \tag{A9}$$

Then, by the definition of the Heaviside function, we have

$$\Theta\left(t - t' - \frac{|\mathbf{x} - \mathbf{v}t'|}{c}\right) = \begin{cases} 1 & \text{if } t \le \mathring{t}, \\ 0 & \text{if } t > \mathring{t}, \end{cases} \tag{A10}$$

where

$$\mathring{t} = q_1(t, \mathbf{x}, \mathbf{v}) - \sqrt{q_2(t, \mathbf{x}, \mathbf{v})}. \tag{A11}$$

It is useful highlight the physical interpretation of time $\mathring{t}$. This value is such that the argument of the Heaviside function in the left-hand side becomes zero or, equivalently,

$$c(t - \mathring{t}) = |\mathbf{x} - \mathbf{v}\mathring{t}|.$$

Please note that $t - \mathring{t}$ is the time it takes for energy to travel from the location of the source at time $\mathring{t}$ (given by $\mathbf{v}\mathring{t}$) to the receiver location $\mathbf{x}$. In other words, $\mathring{t}$ represents the time at which the wave arriving at time t and location $\mathbf{x}$ was emitted. Equation (A11), scaled by the source speed $\mathbf{v}$, is useful to calculate the location of the spatial window induced by a time window in Section 4.

Next, performing the time integral in Equation (A7) yields the following solution for the potential:

$$\phi_{\tilde{\omega}}^{\mathcal{C}}(t, \mathbf{x}, \mathbf{v}) = \frac{ie^{-i\tilde{\omega}\, q_1(t, \mathbf{x}, \mathbf{v})}}{4c\sqrt{c^2 - |\mathbf{v}|^2}} \mathcal{H}_0^{(1)}\left(\tilde{\omega}\sqrt{q_2(t, \mathbf{x}, \mathbf{v})}\right). \tag{A12}$$

where $\mathcal{H}_0^{(1)}$ denotes the first-kind Hankel function of order zero and q_1 and q_2 are as defined by Equations (A8) and (A9) To simplify our expressions as possible, we define the non-dimensional parameter $M = \frac{|\mathbf{v}|}{c}$, which quantifies the ratio of the source speed with respect to the speed of sound in the medium. This parameter simplifies our analysis when considering the Doppler effect below. Furthermore, we place the origin of our reference system at the instantaneous location where the source is collinear with the perpendicular receiver pair. Hence we assume the receiver to be of the form $\mathbf{x} = (0, x_y)$, and that the source speed can be written as $\mathbf{v} = (v, 0)$. Finally, we define for simplicity

$$t_{\mathbf{x}} = \frac{|\mathbf{x}|}{c}, \tag{A13}$$

that is, the travel time from the reference origin described above to the receiver location $\mathbf{x}$. Under these assumptions we can write the wavefield generated by a monochromatic moving point source as

$$p_{\tilde{\omega}}(t, \mathbf{x}, M) = \Re\left\{ -\frac{i\rho\,\tilde{\omega}}{4c^2(1 - M^2)^{\frac{3}{2}}} \exp\left\{-i\frac{\tilde{\omega}}{1 - M^2}t\right\}\right.$$

$$\times \left[\frac{M^2 t}{\sqrt{M^2 t^2 + (1 - M^2)\,t_{\mathbf{x}}^2}} \mathcal{H}_1^{(1)}\left(\frac{\tilde{\omega}}{1 - M^2}\sqrt{M^2 t^2 + (1 - M^2)\,t_{\mathbf{x}}^2}\right)\right. \tag{A14}$$

$$\left.\left. + iH_0^{(1)}\left(\frac{\tilde{\omega}}{1 - M^2}\sqrt{M^2 t^2 + (1 - M^2)\,t_{\mathbf{x}}^2}\right)\right]\right\}.$$

Expressing the wavefield in this way allows us to see how its behavior depends on the source speed $|\mathbf{v}|$ and the distance to the road $|\mathbf{x}|$, albeit always scaled by the medium speed c. It also allows us to see how the distance from the receiver to the road affects its behavior. See Figure A1 for a direct implementation of this equation in the time domain, as well as its numerically calculated frequency spectrum.
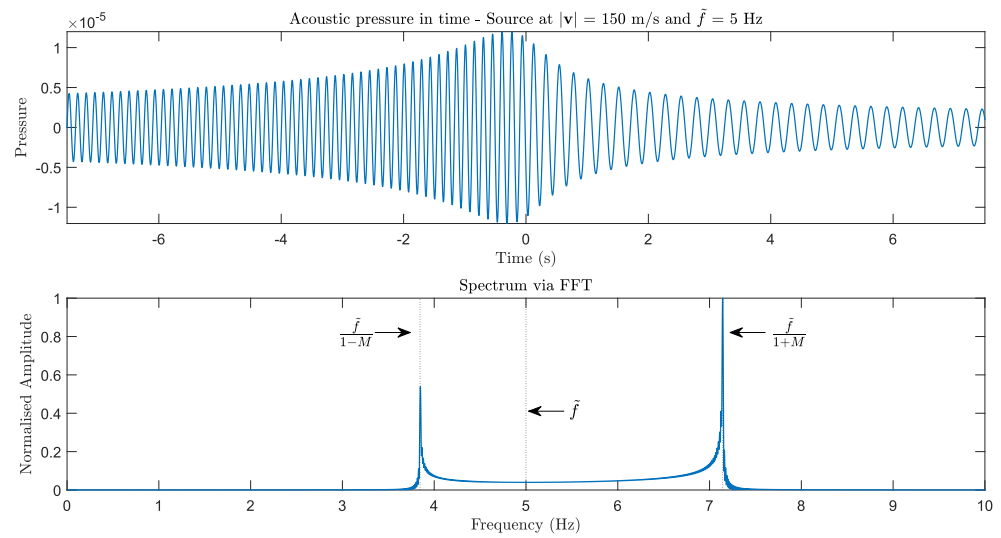
Figure A1. Example of simulated monochromatic wavefield $p_{\tilde{\omega}}$. Medium speed is $c = 500$ m/s and medium density is set to $\rho = 1$. The receiver location is $\mathbf{x} = (0\,\text{m}, 100\,\text{m})$. The monochromatic source travels with speed $|\mathbf{v}| = 150$ m/s along the positive direction of the x-axis, emitting sound at $\tilde{f} = 5$ Hz. It crosses the origin at time $t = 0$ according to our parametrization. For illustration purposes, source parameters were chosen so that the Doppler effect could be plainly appreciated in the waveform (low emitting frequency $\tilde{f}$, relatively high value of M at 0.3). The top panel shows the waveform in the time domain directly implemented from Equation (A14). We can see how the perceived frequency increases or decreases when the source is moving towards or away from the receiver, respectively. The bottom panel shows the spectrum of this source, calculated numerically via the fast Fourier transform (FFT) in Matlab. The spectrum spreads between the frequency values predicted by the standard Doppler formula, $\frac{\tilde{f}}{1 \pm M}$, highlighted in gray dotted lines.

To derive an exact frequency-domain expression for the interferometric retrieval we derive an analytical expression for the spectrum of Equation (A14). Consider the following Fourier transform of the real part of the potential (A12)

$$\hat{\phi}_{\tilde{\omega}}(\omega, \mathbf{x}, \mathbf{v}) = \int_{-\infty}^{\infty} e^{-i\omega t} \mathcal{R}\left\{ \phi_{\tilde{\omega}}^{\mathcal{C}}(t, \mathbf{x}, \mathbf{v}) \right\} dt. \tag{A15}$$

To perform this integration, we derive an integral representation for the spectrum of the potential by taking the real part of the integral representation (A7), where only the spatial integration has been performed, yielding

$$\mathcal{R}\left\{ \phi_{\tilde{\omega}}^{\mathcal{C}}(t, \mathbf{x}, \mathbf{v}) \right\} = \int_{-\infty}^{\infty} \cos(\tilde{\omega}t') \frac{\Theta\left(t - t' - \frac{|\mathbf{x} - \mathbf{v}t'|}{c} \right)}{\sqrt{(t - t')^2 - \frac{|\mathbf{x} - \mathbf{v}t'|^2}{c^2}}} dt' \tag{A16}$$

Substituting Equation (A16) into Equation (A15) results after a few manipulations in the following expression for the spectrum of the wavefield generated by a moving point source

$$\hat{\phi}_{\tilde{\omega}}(\omega, \mathbf{x}, \mathbf{v}) = -\frac{i}{4} \int_{-\infty}^{\infty} e^{-i\omega t'} \cos(\tilde{\omega}t') \mathcal{H}_0^{(2)}\left(\frac{\omega}{c} |\mathbf{x} - \mathbf{v}t'| \right) dt'. \tag{A17}$$

where $H_0^{(2)}$ denotes the second-kind Hankel function of order zero. Finally, to derive the spectrum of the wavefield $p_{\tilde{\omega}}$ Equation (A3) is Fourier transformed so that Equation (A17) is simply multiplied by the corresponding frequency factor $i\omega$ that corresponds to a differentiation in the time domain, that is

$$\hat{p}_{\tilde{\omega}}(\omega, \mathbf{x}, \mathbf{v}) = \frac{\omega\rho}{4} \int_{-\infty}^{\infty} e^{-i\omega t'} \cos(\tilde{\omega}t') \mathcal{H}_0^{(2)}\left(\frac{\omega}{c} |\mathbf{x} - \mathbf{v}t'| \right) dt'. \tag{A18}$$

Numerical implementation of the integral (A18) agrees with the spectrum of the time-domain wavefield (A14) calculated via FFT. Moreover, the spectrum (A18) can be written in the standard form of a recording given by Equation (3), where the source boundary S is parameterized by t' (via $\mathbf{x}'(t') = t'\mathbf{v}$) and the source kernel is given explicitly by

$$\hat{F}(\mathbf{x}'(t'), \omega) = e^{-i\omega t'} \cos(\tilde{\omega} t')$$

and the Green's function is indeed the Green's function of the Helmholtz operator for a volume-injection rate source, given by

$$\hat{G}(\mathbf{x}, \mathbf{x}'(t'), \omega) = \frac{\omega \rho}{4} \mathcal{H}_0^{(2)}\left(\frac{\omega}{c} |\mathbf{x} - \mathbf{x}'(t')|\right).$$

Appendix A.1. Doppler Effect

With frequency-domain analytical expressions available, the instantaneous frequency of the model can be evaluated. The complex model will be used. An analytical expression is derived for the instantaneous phase of the recorded source and then a derivative is taken to find the instantaneous frequency.

Consider the asymptotic approximation of the Hankel function.

$$\mathcal{H}_0^2(x) \sim e^{-i\left(x - \frac{\pi}{4}\right)} \sqrt{\frac{2}{\pi x}}. \tag{A19}$$

The frequency regime and the source-receiver distance are already required to be sufficiently large for the monopole interferometry equation to be applied, which makes approximation (A19) reasonable. Using this approximation, the spectrum of the theoretical complex wavefield $\hat{p}_{\tilde{\omega}}^{\mathcal{C}}(\omega, \mathbf{x}, \mathbf{v})$ (which is derived in a manner analogous to Equation (A18), taking the full complex potential) can be expressed as

$$\hat{p}_{\tilde{\omega}}^{\mathcal{C}}(\omega, \mathbf{x}, \mathbf{v}) \sim \frac{\omega \rho}{4} \int_{-\infty}^{\infty} e^{-i\left\{(\omega + \tilde{\omega})t' + \frac{\omega}{c}|\mathbf{x} - \mathbf{v}t'| - \frac{\pi}{4}\right\}} \sqrt{\frac{2c}{\pi \omega |\mathbf{x} - \mathbf{v}t'|}} \, dt'. \tag{A20}$$

The instantaneous phase recorded at a receiver location $\mathbf{x}$ will be denoted by $\mathrm{ph}_{\mathbf{x}}(t')$ and given by

$$\mathrm{ph}_{\mathbf{x}}(t') = (\omega + \tilde{\omega})t' + \frac{\omega}{c}|\mathbf{x} - \mathbf{v}t'| - \frac{\pi}{4}, \tag{A21}$$

and differentiating once in time yields

$$\frac{d}{dt'}\mathrm{ph}_{\mathbf{x}}(t') = \omega + \tilde{\omega} - \frac{\omega}{c}\left(\frac{\mathbf{x} - \mathbf{v}t'}{|\mathbf{x} - \mathbf{v}t'|}\right) \cdot \mathbf{v}. \tag{A22}$$

The integral (A20) is highly oscillatory and in accordance with the stationary phase approximation the main contributions correspond stationary points of the phase (A21). The spectrum will be otherwise negligibly small by the Riemann-Lebesgue lemma. Hence the spectrum will be largely negligible except at the zeros of Equation (A22). Solving the resulting equation and using the fact that $\mathbf{x} - \mathbf{v}t'$ is the position of the source with respect to the receiver at time t' results in the following expression for the instantaneous frequency recorded at location $\mathbf{x}$ and time t'

$$\breve{\omega}(\mathbf{x}, t') = \frac{\tilde{\omega}}{1 - \frac{|\mathbf{v}|}{c}\cos\theta_{\mathbf{x}}(t')}. \tag{A23}$$

The instantaneous recorded frequency (A23) exhibits the Doppler shift induced by the movement of the source with respect to the receiver. Please note that for any time t' the relative position angle $\cos\theta_{\mathbf{x}}(t')$ is bounded between $-\pi$ (when the source is infinitely far to the left of the receiver) and π (when the source is infinitely far to the right), corresponding to the frequencies $\tilde{\omega}/1 \pm M$ where $M = |\mathbf{v}|/c$. Equation (A23) agrees with the well-known

expressions for the recorded frequency when the receiver is directly in the path of the source, and is more general as the instantaneous frequency can be calculated for an arbitrary recording location $\mathbf{x}$. Moreover, it is straightforward to show that Equation (A23) is larger than $\tilde{\omega}$ when the source is moving towards the receiver and smaller when the source is moving away, with the true emitted frequency $\tilde{\omega}$ recorded the moment the source-receiver ray-path perpendicular to the source boundary.

Appendix A.2. Broadband Moving Source Model

We construct a broadband source emitting noise on a discrete frequency domain $\tilde{\omega} \in \tilde{\Omega}$ by applying the principle of superposition to the monochromatic wavefields (A18), and imposing a random phase shift for each emitted frequency, i.e., we define the broadband wavefield

$$\hat{p}(\mathbf{x}, \omega) = \sum_{\tilde{\omega} \in \tilde{\Omega}} c_{\tilde{\omega}} \hat{p}_{\tilde{\omega}}(\mathbf{x}, \omega), \tag{A24}$$

where $c_{\tilde{\omega}}$ is an appropriate scaling coefficient that depends on how the spectrum $\tilde{\Omega}$ is constructed, and the dependence on the source speed $\mathbf{v}$ is left implicit. In this paper, it is assumed that the source spectrum can be represented in the form (A24). A more general representation can be achieved by taking the Fourier transform of a wavefield of the form

$$p(\mathbf{x}, t) = \sum_{\tilde{\omega} \in \tilde{\Omega}} c_{\tilde{\omega}}(t) p_{\tilde{\omega}}(\mathbf{x}, t), \tag{A25}$$

where the coefficients $c_{\tilde{\omega}}(t)$ are time dependent and random with their own correlation structure. Although we do not deal with this case in detail here, we point out such a setup would be necessary to derive a kernel that becomes localized for the case of a monochromatic source, as remarked under Equation (14).

Assuming that the source spectrum can be represented in Equation (A24), we have

$$\hat{p}(\mathbf{x}, \omega) = \frac{\omega \rho}{4} \sum_{\tilde{\omega} \in \tilde{\Omega}} c_{\tilde{\omega}} \int_{-\infty}^{\infty} e^{-i\omega t'} \cos(\tilde{\omega} t' + \theta_{\tilde{\omega}}) \mathcal{H}_0^{(2)} \left(\omega \frac{|\mathbf{x} - \mathbf{v}t'|}{c} \right) dt', \tag{A26}$$

where $\theta_{\tilde{\omega}}$ are a collection of random variables indexed by $\tilde{\omega}$. The random quantity

$$F(t'; \theta_{\tilde{\Omega}}) = \sum_{\tilde{\omega} \in \tilde{\Omega}} c_{\tilde{\omega}} \cos(\tilde{\omega} t' + \theta_{\tilde{\omega}}) \tag{A27}$$

characterizes the spectrum emitted by the source, as well as how the component frequencies interact with each other depending on the distribution of $\theta_{\tilde{\omega}}$ and their correlation structure. Equation (A26) can then be written more compactly as

$$\hat{p}(\mathbf{x}, \omega) = \frac{\omega \rho}{4} \int_{-\infty}^{\infty} F(t'; \theta_{\tilde{\Omega}}) e^{-i\omega t'} \mathcal{H}_0^{(2)} \left(\omega \frac{|\mathbf{x} - \mathbf{v}t'|}{c} \right) dt', \tag{A28}$$

where it becomes clear that $\hat{p}$ is itself a random quantity.

Appendix B. Analytical Kernel for a Band-Limited Source Train with Random-Phase Emitted Frequency Components

The stacking average interferometric retrieval from the recordings of Equation (A28) at a receiver pair $\mathbf{x_A}$ and $\mathbf{x_B}$ can be written for a source travelling at speed $\mathbf{v}$ as

$$\langle \hat{h}(\mathbf{x_A}, \mathbf{x_B}, \mathbf{v}, \omega) \rangle = \frac{2}{\rho c} \langle \hat{p}^*(\mathbf{x_A}, \omega) \hat{p}(\mathbf{x_B}, \omega) \rangle$$

$$= \frac{\omega^2 \rho}{8c} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle F^*(t'; \theta_{\tilde{\Omega}}) F(t''; \theta_{\tilde{\Omega}}) \rangle e^{-i\omega(t''-t')}$$

$$\times \mathcal{H}_0^{(1)}\left(\omega \frac{|\mathbf{A} - \mathbf{v}t'|}{c} \right) \mathcal{H}_0^{(2)}\left(\omega \frac{|\mathbf{B} - \mathbf{v}t''|}{c} \right) dt' \, dt'' \tag{A29}$$

For an arbitrary discrete source spectrum $\tilde{\Omega}$ and associated source signature in the form of (A27), the generalized average source kernel in an ambient noise interferometric retrieval is given by

$$\langle F^*(t'; \theta_{\tilde{\Omega}}) F^*(t''; \theta_{\tilde{\Omega}}) \rangle = \sum_{\tilde{\omega} \in \tilde{\Omega}} c_{\tilde{\omega}}^2 \Big(\cos(\tilde{\omega} t') \cos(\tilde{\omega} t'') \langle \cos^2(\theta_{\tilde{\omega}}) \rangle$$
$$- \sin(\tilde{\omega}(t' + t'')) \langle \cos(\theta_{\tilde{\omega}}) \sin(\theta_{\tilde{\omega}}) \rangle$$
$$+ \sin(\tilde{\omega} t') \sin(\tilde{\omega} t'') \langle \sin^2(\theta_{\tilde{\omega}}) \rangle \Big)$$
$$+ \sum_{\tilde{\omega} \in \tilde{\Omega}} \sum_{\substack{\tilde{\omega}' \in \tilde{\Omega} \\ \tilde{\omega} \neq \tilde{\omega}'}} c_{\tilde{\omega}} c_{\tilde{\omega}'} \Big(\cos(\tilde{\omega} t') \cos(\tilde{\omega}' t'') \langle \cos(\theta_{\tilde{\omega}}) \cos(\theta_{\tilde{\omega}'}) \rangle$$
$$- \cos(\tilde{\omega} t') \sin(\tilde{\omega}' t'') \langle \cos(\theta_{\tilde{\omega}}) \sin(\theta_{\tilde{\omega}'}) \rangle$$
$$- \sin(\tilde{\omega} t') \cos(\tilde{\omega}' t'') \langle \sin(\theta_{\tilde{\omega}}) \cos(\theta_{\tilde{\omega}'}) \rangle$$
$$+ \sin(\tilde{\omega} t') \sin(\tilde{\omega}' t'') \langle \sin(\theta_{\tilde{\omega}}) \sin(\theta_{\tilde{\omega}'}) \rangle \Big), \tag{A30}$$

where the double summation is split into diagonal terms and cross-terms $\tilde{\omega} \neq \tilde{\omega}'$. Equation (A30) shows that the standard ensemble average operator depends on the phases of the emitted frequency components of the source, as well as their correlation.

For example, consider the case when the source spectrum is a regular partition of a bounded interval, that is $\tilde{\Omega} = \{\tilde{\omega}_j\}_{j=1}^{|\tilde{\Omega}|}$ for some finite spacing $\triangle \tilde{\omega} = \tilde{\omega}_{j+1} - \tilde{\omega}_j$, and associated random phase shifts θ_j for each frequency $\tilde{\omega}_j$. Moreover, assume that the random shifts are such that $\theta_j \sim \mathrm{Unif}(0, 2\pi)$ for all $j \in \{1, \dots, |\tilde{\Omega}|\}$. Furthermore, assume that the phase shifts θ_j and θ_k are independent whenever $\tilde{\omega}_j \neq \tilde{\omega}_k$. These assumptions result in the following probability distributions for the random variables in Equation (A30)

$$\sin(\theta), \cos(\theta) \sim \mathrm{ArcSin}(-1, 1),$$
$$\sin(\theta) \cos(\theta) \sim \mathrm{ArcSin}\left(-\tfrac{1}{2}, \tfrac{1}{2}\right),$$

and

$$\sin(\theta)^2, \cos^2(\theta) \sim \mathrm{ArcSin}(0, 1),$$

where ArcSin denotes the arcsine probability distribution with known means. Moreover, assume that the coefficients $c_{\tilde{\omega}}$ in Equation (A27) do not depend on the individual frequencies $\tilde{\omega}$, and let this coefficient be denoted henceforth simply as $c_{\tilde{\Omega}}$. Then the random source kernel (A30) reduces to the deterministic expression

$$\langle F^*(t'; \theta_{\tilde{\Omega}}) F(t''; \theta_{\tilde{\Omega}}) \rangle = c_{\tilde{\Omega}}^2 \frac{1}{|\tilde{\Omega}|} \sum_{j=0}^{|\tilde{\Omega}|} \cos(\tilde{\omega}_j(t'' - t')). \tag{A31}$$

Next, one can deduce in a standard fashion [52] that

$$\left\langle F^*(t';\theta_{\tilde{\Omega}})F(t'';\theta_{\tilde{\Omega}}) \right\rangle \xrightarrow[\substack{\triangle_{\tilde{\omega}}\to 0 \\ |\tilde{\Omega}|\triangle_{\tilde{\omega}}=const.}] \frac{c_{\tilde{\Omega}}^2}{\tilde{\omega}_{\max} - \tilde{\omega}_{\min}} \int_{\tilde{\omega}_{\min}}^{\tilde{\omega}_{\max}} \cos(\tilde{\omega}(t'' - t'))d\tilde{\omega}, \tag{A32}$$

where $\triangle_{\tilde{\omega}} = \tilde{\omega}_{j+1} - \tilde{\omega}_j$ and $\tilde{\omega}_{\max} - \tilde{\omega}_{\min} = |\tilde{\Omega}|\triangle_{\tilde{\omega}}$, which leads to

$$\left\langle F^*(t';\theta_{\tilde{\Omega}})F(t'';\theta_{\tilde{\Omega}}) \right\rangle \approx c_{\tilde{\Omega}}^2 \frac{\sin(\tilde{\omega}_{\max}(t'' - t)) - \sin(\tilde{\omega}_{\min}(t'' - t))}{(\tilde{\omega}_{\max} - \tilde{\omega}_{\min})(t'' - t')}. \tag{A33}$$

Finally, the right-hand side of Equation (A33) can be expanded using well-known trigonometric identities to derive

$$\left\langle F^*(t';\theta_{\tilde{\Omega}})F(t'';\theta_{\tilde{\Omega}}) \right\rangle \approx c_{\tilde{\omega}}^2 \cos\left(\frac{(\tilde{\omega}_{\max} + \tilde{\omega}_{\min})}{2}(t'' - t')\right)\text{sinc}\left(\frac{(\tilde{\omega}_{\max} - \tilde{\omega}_{\min})}{2}(t'' - t')\right) \tag{A34}$$

The kernel Equation (A34) is parameterized in time. Under the assumption that the source is in motion, i.e., $\mathbf{v} \neq 0$, Equation (A34) leads to expression (14). To see this, let $\mathbf{x}'$ be the source location at time t', and $\mathbf{x}''$ the source location at time t'', so that $\mathbf{x}' = (|\mathbf{v}|t', 0)$ and $\mathbf{x}'' = (|\mathbf{v}|t'', 0)$. Hence

$$t'' - t' = \frac{|\mathbf{v}|t'' - |\mathbf{v}|t'}{|\mathbf{v}|} = \frac{x'' - x'}{|\mathbf{v}|}. \tag{A35}$$

where $x' = |\mathbf{v}|t'$ and similarly $x'' = |\mathbf{v}|t''$. Combining Equations (A34) and (A35) yields an expression proportional to Equation (14), noting that the right-hand side of Equation (A34) is an even function (i.e., symmetric with respect to the sign of the argument).

References

1. Campillo, M.; Paul, A. Long-range correlations in the diffuse seismic coda. *Science* **2003**, *299*, 547–549.
2. Wapenaar, K.; Fokkema, J. Green's function representations for seismic interferometry. *Geophysics* **2006**, *71*, SI33–SI46.
3. Curtis, A.; Gerstoft, P.; Sato, H.; Snieder, R.; Wapenaar, K. Seismic interferometry—Turning noise into signal. *Lead. Edge* **2006**, *25*, 1082–1092.
4. Claerbout, J.F. Synthesis of a layered medium from its acoustic transmission response. *Geophysics* **1968**, *33*, 264–269.
5. Rickett, J.; Claerbout, J. Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring. *Lead. Edge* **1999**, *18*, 957–960.
6. Weaver, R.L.; Lobkis, O.I. Ultrasonics without a source: Thermal fluctuation correlations at MHz frequencies. *Phys. Rev. Lett.* **2001**, *87*, 134301.
7. Derode, A.; Larose, E.; Campillo, M.; Fink, M. How to estimate the Green's function of a heterogeneous medium between two passive sensors? Application to acoustic waves. *Appl. Phys. Lett.* **2003**, *83*, 3054–3056.
8. Slob, E.; Wapenaar, K. Electromagnetic Green's functions retrieval by cross-correlation and cross-convolution in media with losses. *Geophys. Res. Lett.* **2007**, *34*, 1–5.
9. Ruigrok, E.; Draganov, D.; Wapenaar, K. Global-scale seismic interferometry: Theory and numerical examples. *Geophys. Prospect.* **2008**, *56*, 395–417.
10. Nishida, K.; Montagner, J.P.; Kawakatsu, H. Global surface wave tomography using seismic hum. *Science* **2009**, *326*, 112.
11. Shapiro, N.M.; Campillo, M.; Stehly, L.; Ritzwoller, M.H. High-resolution surface-wave tomography from ambient seismic noise. *Science* **2005**, *307*, 1615–1618.
12. Nishida, K.; Kawakatsu, H.; Obara, K. Three-dimensional crustal S wave velocity structure in Japan using microseismic data recorded by Hi-net tiltmeters. *J. Geophys. Res. Solid Earth* **2008**, *113*, 1–22.
13. Arroucau, P.; Rawlinson, N.; Sambridge, M. New insight into Cainozoic sedimentary basins and Palaeozoic suture zones in southeast Australia from ambient noise surface wave tomography. *Geophys. Res. Lett.* **2010**, *37*, 1–6.
14. Bakulin, A.; Calvert, R. The virtual source method: Theory and case study. *Geophysics* **2006**, *71*, SI139–SI150.
15. Bakulin, A.; Mateeva, A.; Mehta, K.; Jorgensen, P.; Ferrandis, J.; Herhold, I.S.; Lopez, J. Virtual source applications to imaging and reservoir monitoring. *Lead. Edge* **2007**, *26*, 732–740.
16. Halliday, D.; Curtis, A. An interferometric theory of source-receiver scattering and imaging. *Geophysics* **2010**, *75*, SA95–SA103.
17. Hong, T.K.; Menke, W. Tomographic investigation of the wear along the San Jacinto fault, southern California. *Phys. Earth Planet. Inter.* **2006**, *155*, 236–248.

18. Curtis, A.; Nicolson, H.; Halliday, D.; Trampert, J.; Baptie, B. Virtual seismometers in the subsurface of the Earth from seismic interferometry. *Nat. Geosci.* **2009**, *2*, 700–704.
19. Wapenaar, K. Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation. *Phys. Rev. Lett.* **2004**, *93*, 254301.
20. Curtis, A.; Halliday, D. Source-receiver wave field interferometry. *Phys. Rev. E* **2010**, *81*, 046601.
21. Curtis, A.; Behr, Y.; Entwistle, E.; Galetti, E.; Townend, J.; Bannister, S. The benefit of hindsight in observational science: Retrospective seismological observations. *Earth Planet. Sci. Lett.* **2012**, *345*, 212–220.
22. Entwistle, E.; Curtis, A.; Galetti, E.; Baptie, B.; Meles, G. Constructing new seismograms from old earthquakes: Retrospective seismology at multiple length scales. *J. Geophys. Res. Solid Earth* **2015**, *120*, 2466–2490.
23. Chen, Y.; Saygin, E. Empirical Green's Function Retrieval Using Ambient Noise Source-Receiver Interferometry. *J. Geophys. Res. Solid Earth* **2020**, *125*, doi:10.1029/2019JB018261.
24. Wapenaar, K.; Van Der Neut, J.; Ruigrok, E.; Draganov, D.; Hunziker, J.; Slob, E.; Thorbecke, J.; Snieder, R. Seismic interferometry by crosscorrelation and by multidimensional deconvolution: A systematic comparison. *Geophys. J. Int.* **2011**, *185*, 1335–1364.
25. Nicolson, H.; Curtis, A.; Baptie, B.; Galetti, E. Seismic interferometry and ambient noise tomography in the British Isles. *Proc. Geol. Assoc.* **2012**, *123*, 74–86.
26. Ardhuin, F.; Stutzmann, E.; Schimmel, M.; Mangeney, A. Ocean wave sources of seismic noise. *J. Geophys. Res. Ocean.* **2011**, *116*, C006952.
27. Sabra, K.G.; Roux, P.; Kuperman, W. Arrival-time structure of the time-averaged ambient noise cross-correlation function in an oceanic waveguide. *J. Acoust. Soc. Am.* **2005**, *117*, 164–174.
28. Halliday, D.; Curtis, A.; Kragh, E. Seismic surface waves in a suburban environment: Active and passive interferometric methods. *Lead. Edge* **2008**, *27*, 210–218.
29. Nakata, N.; Snieder, R.; Tsuji, T.; Larner, K.; Matsuoka, T. Shear wave imaging from traffic noise using seismic interferometry by cross-coherence. *Geophysics* **2011**, *76*, SA97–SA106.
30. Behm, M.; Snieder, R. Love waves from local traffic noise interferometry. *Lead. Edge* **2013**, *32*, 628–632.
31. Dales, P.; Pinzon-Ricon, L.; Brenguier, F.; Boué, P.; Arndt, N.; McBride, J.; Lavoué, F.; Bean, C.J.; Beaupretre, S.; Fayjaloun, R.; et al. Virtual Sources of Body Waves from Noise Correlations in a Mineral Exploration Context. *Seismol. Res. Lett.* **2020**, *91*, 2278–2286.
32. Brenguier, F.; Boué, P.; Ben-Zion, Y.; Vernon, F.; Johnson, C.; Mordret, A.; Coutant, O.; Share, P.E.; Beaucé, E.; Hollis, D.; et al. Train traffic as a powerful noise source for monitoring active faults with seismic interferometry. *Geophys. Res. Lett.* **2019**, *46*, 9529–9536.
33. Quiros, D.A.; Brown, L.D.; Kim, D. Seismic interferometry of railroad induced ground motions: Body and surface wave imaging. *Geophys. Suppl. Mon. Not. R. Astron. Soc.* **2016**, *205*, 301–313.
34. Pinzon-Rincon, L.; Lavoué, F.; Mordret, A.; Boué, P.; Brenguier, F.; Dales, P.; Ben-Zion, Y.; Vernon, F.; Bean, C.J.; Hollis, D. Humming Trains in Seismology: An Opportune Source for Probing the Shallow Crust. *Seismol. Soc. Am.* **2021**, *92*, 623–635.
35. Liu, Y.; Yue, Y.; Li, Y.; Luo, Y. On the Retrievability of Seismic Waves From High-Speed-Train-Induced Vibrations Using Seismic Interferometry. *IEEE Geosci. Remote. Sens. Lett.* **2021**, doi:10.1109/LGRS.2021.3050205.
36. Gerstoft, P.; Sabra, K.G.; Roux, P.; Kuperman, W.; Fehler, M.C. Green's functions extraction and surface-wave tomography from microseisms in southern California. *Geophysics* **2006**, *71*, SI23–SI31.
37. Curtis, A.; Halliday, D. Directional balancing for seismic and general wavefield interferometry. *Geophysics* **2010**, *75*, SA1–SA14.
38. Fichtner, A.; Stehly, L.; Ermert, L.; Boehm, C. Generalised interferometry-I. Theory for inter-station correlations. *Geophys. J. Int.* **2017**, *208*, 603–638.
39. Van der Neut, J.R. Interferometric Redatuming by Multidimensional Deconvolution. 2012. Available online: http://homepage.tudelft.nl/t4n4v/9_Theses_students/Neut.pdf (accessed on 17th June 2021).
40. Van Dalen, K.N.; Mikesell, T.D.; Ruigrok, E.N.; Wapenaar, K. Retrieving surface waves from ambient seismic noise using seismic interferometry by multidimensional deconvolution. *J. Geophys. Res. Solid Earth* **2015**, *120*, 944–961.
41. Sabra, K.G. Influence of the noise sources motion on the estimated Green's functions from ambient noise cross-correlations. *J. Acoust. Soc. Am.* **2010**, *127*, 3577–3589.
42. Schuster, G. *Seismic Interferometry*; Cambridge University Press: Cambridge, UK, 2009.
43. Halliday, D.; Curtis, A. Seismic interferometry, surface waves and source distribution. *Geophys. J. Int.* **2008**, *175*, 1067–1087.
44. Mehta, K.; Bakulin, A.; Sheiman, J.; Calvert, R.; Snieder, R. Improving the virtual source method by wavefield separation. *Geophysics* **2007**, *72*, V79–V86.
45. Snieder, R. Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Phys. Rev. E* **2004**, *69*, 046610.
46. Thorbecke, J.; Wapenaar, K. Analysis of spurious events in seismic interferometry. In *SEG Technical Program Expanded Abstracts 2008*; Society of Exploration Geophysicists: Tulsa, OK, USA, 2008; pp. 1415–1420.
47. Van Manen, D.; Curtis, A.; Robertsson, J. Interferometric modeling of wave propagation in inhomogeneous elastic media using time reversal and reciprocity. *Geophysics* **2006**, *71*, SI47–SI60.
48. Van Dalen, K.N.; Wapenaar, K.; Halliday, D.F. Surface wave retrieval in layered media using seismic interferometry by multidimensional deconvolution. *Geophys. J. Int.* **2014**, *196*, 230–242.
49. Wapenaar, K.; van der Neut, J.; Thorbecke, J. On the relation between seismic interferometry and the simultaneous-source method. *Geophys. Prospect.* **2012**, *60*, 802–823.

50. King, S.; Curtis, A. Velocity analysis using both reflections and refractions in seismic interferometry. *Geophysics* **2011**, *76*, SA83–SA96.

51. Maranò, S.; Reller, C.; Loeliger, H.A.; Fäh, D. Seismic waves estimation and wavefield decomposition: Application to ambient vibrations. *Geophys. J. Int.* **2012**, *191*, 175–188.

52. Rudin, W. *Principles of Mathematical Analysis*; McGraw-Hill: New York, NY, USA, 1976; Volume 3.