On measuring surface wave phase velocity from station-station cross-correlation of ambient signal

Lapo Boschi, Cornelis Weemstra, Julie Verbeke, Göran Ekström, Andrea Zunino, Domenico Giardini

On measuring surface-wave phase velocity from station-station cross-correlation of ambient signal

Lapo Boschi (ETH Zürich), Cornelis Weemstra (Spectraseis, ETH Zürich), Julie Verbeke (ETH Zürich), Göran Ekström (LDEO), Andrea Zunino (Technical University of Denmark), Domenico Giardini (ETH Zürich)

October 11, 2012

Abstract

We apply two different algorithms to measure surface-wave phase velocity, as a function of frequency, from seismic ambient noise recorded at pairs of stations from a large European network. The two methods are based on consistent theoretical formulations, but differ in the implementation: one method involves the time-domain cross-correlation of signal recorded at different stations; the other is based on frequency-domain cross-correlation, and requires finding the zero-crossings of the real part of the cross-correlation spectrum. Furthermore, the time-domain method, as implemented here and in the literature, practically involves the important approximation that interstation distance be large compared to seismic wavelength. In both cases, cross-correlations are ensemble-averaged over a relatively long period of time (one year). We verify that the two algorithms give consistent results, and infer that phase velocity can be successfully measured through ensemble-averaging of seismic ambient noise, further validating earlier studies that had followed either approach. The description of our experiment and its results is accompanied by a detailed though simplifed derivation of ambient-noise theory, writing out explicitly the relationships between the surface-wave Green function, ambient-noise cross-correlation, and phase and group velocities.

1 Introduction

The ability to observe coherent surface-wave signal from the stacked cross-correlation of background noise recorded at different stations is essential to improve our resolution of Earth structure via seismic imaging. Surface waves generated by earthquakes are best observed at teleseismic distances, where the body- and surface-wave packets are well separated, and, owing to different geometrical spreading, surface waves are much more energetic than body waves; teleseismic surface waves, however, are dominated by intermediate to long periods ($\gtrsim$ 30s), and their speed of propagation is therefore related to mantle, rather than crustal structure [e.g., *Boschi and Ekström*, 2002]. The averaged cross-correlated ambient-noise signal is instead observed at periods roughly between 5 and 30 s [e.g., *Stehly et al.*, 2006, 2009], complementary to the period range of teleseismic surface waves, and allowing to extend imaging resolution upwards into the lithosphere-asthenosphere boundary region and the crust.

As first noted by *Shapiro and Campillo* [2004], the cross-correlation of seismic ambient signal recorded at two different stations approximates the Green function associated with a point source acting at one of the stations' location, and a receiver deployed at the other's. Such empirical Green function can then be analyzed in different ways, with the ultimate goal of obtaining information about Earth's structure at various depths between the two stations. Most authors either extract group velocity v_g from its envelope [e.g., *Shapiro et al.*, 2005; *Stehly et al.*, 2006, 2009], or isolate the phase velocity v [e.g., *Lin et al.*, 2008; *Nishida et al.*, 2008; *Yao and van der Hilst*, 2009; *Ekström et al.*, 2009]. Fewer authors [e.g., *Tromp et al.*, 2010; *Basini et al.*, 2012] attempt to explain (invert) the entire ambient-noise waveform.

Both v and v_g are useful expressions of shallow Earth properties between seismic source and receiver, or, in the present case, between two receivers. To measure v_g one must be able to identify the peak of the surface-wave envelope. This, as a general rule, is easier than isolating the carrying sinusoidal wave (i.e. measuring v) at a given frequency. There are, however, several properties of v_g that make phase-velocity observations useful and possibly preferable: (i) the envelope peak is less precisely defined than the phase of the carrying sinusoidal wave; (ii) at least so far as the surface-wave fundamental mode is concerned, v_g depends on, and is in turn used to image, structure over a narrower and shallower depth range than v [e.g., *Ritzwoller et al.*, 2001], so that v is particularly helpful to resolve larger depths; (iii) a v_g measurement needs to be made over a wider time window than a v measurement, and contamination by interfering phases is accordingly more likely.

While the validity of group-velocity estimates based on seismic ambient noise is widely recognized, phase velocity is more elusive. For instance, *Yao et al.* [2006] have noted an important discrepancy between two-station observations of phase velocity obtained from teleseismic vs. ambient signal. The systematic application of a far-field approximation, in the theoretical expression used to extract the phase from cross-correlation observations (see eqs. (41) and (35) below), results in a $\pi/4$ shift with respect to the cross-correlation of ballistic signal [e.g., *Harmon et al.*, 2008], which has caused some confusion as noted e.g. by *Tsai* [2009]. We apply here two different approaches to measure inter-station surface-wave phase velocity from one year of continuous recording at a dense, large array of European stations, first compiled by *Verbeke et al.* [2012b]. Both methods can be derived from the same basic theoretical formulation [*Tsai and Moschetti*, 2010]. One of them is based on time-domain cross-correlation, and is implemented, here and elsewhere, using a far-field approximation of the wavefield equation. The other is based on frequency domain cross-correlation, and on finding the roots of the real part of the cross-correlation spectrum; it involves no far-field approximation. The consistency between the two methods' results further validates earlier phase-velocity tomography studies conducted with either approach [e.g., *Lin et al.*, 2008; *Yao and van der Hilst*, 2009; *Ekström et al.*, 2009; *Fry et al.*, 2010; *Verbeke et al.*, 2012b].

2 Theory

We study the properties of the cross-correlation $C_{xy}(t, \omega)$, function of time t and frequency ω, of ambient surface-wave signal u recorded at two seismic instruments, located at positions

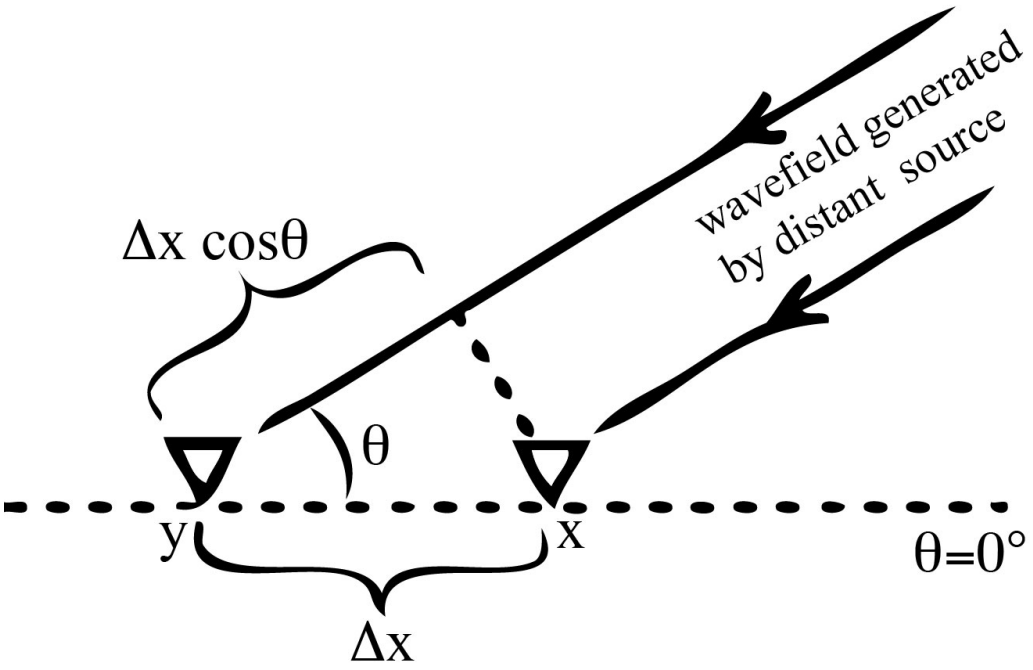

Figure 1: Modified from *Tsai* [2009]. Stations at x and y are separated by a distance $\Delta x > 0$. Noise sources are far enough that the azimuth θ of any given source is about the same with respect to either station.

x and y. By definition

$$C_{xy}(t,\omega) = \frac{1}{2T}\int_{-T}^{T} u(x,\tau,\omega)u(y,t+\tau,\omega)\mathrm{d}\tau, \tag{1}$$

with the parameter T defining the size of the window over which the cross-correlation is computed in practice. We limit our analysis to sources sufficiently far from both receivers for the source-receiver azimuth θ to be approximately the same. If we denote Δx the distance separating the two receivers, it then follows, as illustrated in Fig. 1, that the surface wave of frequency ω and phase velocity $v(\omega)$ generated by a plane-wave source at azimuth θ hits the receiver at y with an approximate delay

$$t_d = \Delta x \cos(\theta)/v(\omega) \tag{2}$$

with respect to the one at x.

Our treatment follows that of *Tsai* [2009] and *Tsai* [2011]; we review the formulation carried out in those works, confirming the theoretical consistency, and pointing out the practical differences between the data-analysis methods that we compare. The mathematical treatment leads to complete expressions for cross-correlation (section 2.4), and group, as well as phase-velocity of the ambient signal (section 2.5). Like *Tsai* [2009], we assume, as mentioned, that sources of ambient noise are far enough from our station pair for the source-receiver azimuth to be approximately the same at the two stations.

Another important assumption of our and most other formulations of ambient-noise theory is that the ambient signal be approximately "diffuse". In practice, this is not true at any

moment in time, but can be at least partially achieved if the ambient signal recorded over a very long time (e.g., one year) is subdivided into shorter (e.g., one-day-long) intervals, which are then whitened and (after station-station cross-correlation) stacked [*Yang and Ritzwoller*, 2008; *Mulargia*, 2012]. This procedure is described in detail by *Bensen et al.* [2007]; we refer to it as "ensemble-average", rather than time-average, since shorter time intervals can be chosen to overlap [e.g., *Seats et al.*, 2012; *Weemstra et al.*, 2012]. Over time, an array of seismic stations will record ambient signal generated over a wide range of azimuths and distances, and the process of stacking simulates the superposition of simultaneously acting sources. *Stehly et al.* [2006] show that, at least in the period range ∼5-15s, most ambient-noise signal is likely to be generated by the interaction between oceans and the solid Earth (i.e., ocean storms), and the source distribution of even the stacked ambient signal is accordingly nonuniform. Yet, there are both empirical [*Derode et al.*, 2003] and theoretical [*Snieder*, 2004] indications that as long as a significant fraction of ambient signal hits a receiver pair along the receiver-receiver azimuth, ensemble-averaging will result in successful applications of ambient-noise methods. In our formulation, we treat sources as uniformly distributed in azimuth with respect to the receiver pair.

2.1 Monochromatic signal from a single source

In the absence of strong lateral heterogeneity in elastic structure, the momentum equation for a Love or Rayleigh wave can be decoupled into a differential equation in the vertical, and another in the horizontal Cartesian coordinates. The latter coincides with the Helmholtz equation and is solved by sinusoidal functions [e.g., *Peter et al.*, 2007].

Seismic ambient noise can be thought of as the effect of a combination of sources more-or-less randomly distributed in space and time. It is however convenient to start our treatment, following *Tsai* [2009], from the simple case of a single source generating a monochromatic signal of frequency ω. The first receiver then records a signal

$$u(x,t) = S(x,\omega)\cos(\omega t + \phi), \tag{3}$$

where the constant phase delay ϕ is proportional to source-receiver distance, and the amplitude term $S(x,\omega)$ is inversely proportional, in the first approximation, to the square-root of source-receiver distance (geometrical spreading). The signal (3) is observed at y with a delay t_d, i.e.

$$u(y,t) = S(y,\omega)\cos\left[\omega(t+t_d)+\phi\right]. \tag{4}$$

(By virtue of eq. (2), t_d is negative when energy propagates from y to x ($0 < \theta < \pi/2$) and positive when energy propagates from x to y.)

Let us substitute (3) and (4) into (1), so that

$$C_{xy} = \frac{S(x)S(y)}{2T}\int_{-T}^{T}\cos(\omega\tau+\phi)\cos\left[\omega(\tau+t+t_d)+\phi\right]\mathrm{d}\tau. \tag{5}$$

It is convenient to substitute $z = \omega\tau$, to find

$$C_{xy} = \frac{S(x)S(y)}{2\omega T}\int_{-\omega T}^{\omega T}\cos(z+\phi)\cos\left[z+\phi+\omega(t+t_d)\right]\mathrm{d}z. \tag{6}$$

We next make use of the general trigonometric identity $\cos(A+B) = \cos A \cos B - \sin A \sin B$, valid for any A, B, and

$$\begin{aligned} C_{xy} = & \frac{S(x)S(y)}{2\omega T} \int_{-\omega T}^{\omega T} \Big\{ \cos^2(z) \cos\left[\phi + \omega(t+t_d)\right] \cos(\phi) \\ & + \sin^2(z) \sin\left[\phi + \omega(t+t_d)\right] \sin(\phi) \\ & - \sin(z)\cos(z)\cos(\phi)\sin\left[\phi + \omega(t+t_d)\right] \\ & - \sin(z)\cos(z)\sin(\phi)\cos\left[\phi + \omega(t+t_d)\right] \Big\} \, \mathrm{d}z, \end{aligned} \tag{7}$$

which can be simplified if one notices that

$$\int_{-\omega T}^{\omega T} \cos^2(z) \mathrm{d}z = \int_{-\omega T}^{\omega T} \frac{1+\cos(2z)}{2} \mathrm{d}z = \omega T + \frac{1}{2}\sin(2\omega T), \tag{8}$$

$$\int_{-\omega T}^{\omega T} \sin^2(z) \mathrm{d}z = \int_{-\omega T}^{\omega T} \frac{1-\cos(2z)}{2} \mathrm{d}z = \omega T - \frac{1}{2}\sin(2\omega T), \tag{9}$$

and finally

$$\int_{-\omega T}^{\omega T} \sin(z)\cos(z) \mathrm{d}z = \left[\frac{\sin^2(z)}{2}\right]_{-\omega T}^{\omega T} = 0, \tag{10}$$

where the notation $[f(z)]_A^B = f(B) - f(A)$.

After substituting the expressions (8), (9) and (10) into eq. (7),

$$\begin{aligned} C_{xy} = \frac{S(x)S(y)}{2} \Big\{ & \left[1 + \frac{\sin(2\omega T)}{2\omega T}\right] \cos(\phi)\cos\left[\phi + \omega(t+t_d)\right] \\ & + \left[1 - \frac{\sin(2\omega T)}{2\omega T}\right] \sin(\phi)\sin\left[\phi + \omega(t+t_d)\right] \Big\}. \end{aligned} \tag{11}$$

It then follows from simple trigonometric identities (cosine of the sum, sine of the sum) that

$$C_{xy} = \frac{S(x)S(y)}{2} \left\{ \cos\left[\omega(t+t_d)\right] + \frac{\sin(2\omega T)}{2\omega T} \cos\left[2\phi + \omega(t+t_d)\right] \right\}. \tag{12}$$

This expression can be simplified if one considers that the size $2T$ of the cross-correlation window should be large compared to the period of the surface waves in question, i.e. $T \gg 2\pi/\omega$, so that $2\omega T \gg 1$. Eq. (12) then reduces to

$$C_{xy} \approx \frac{S(x)S(y)}{2} \cos\left[\omega(t+t_d)\right] \tag{13}$$

(compare with eq. (1) of *Tsai* [2009]). From eq. (13) we infer that the station-station cross-correlation of a "ballistic" signal, i.e. generated by a single source localized in space, and not scattered, is only useful if the location of the source is known. It coincides (once amplitude is normalized) with the response, at one station, to a sinusoidal source located at the other, if and only if the two stations are aligned with the source, i.e. azimuth $\theta = 0$ or $\theta = \pi$, so that $t_d = \pm \Delta x / v$.

2.2 Monochromatic signal from a discrete set of sources

Recorded seismic ambient noise is believed to be the cumulative effect of numerous localized sources, distributed almost randomly all around our pair of recording instruments. The signal

generated by a discrete set of monochromatic sources can be written as a superposition of single-source signals, eqs. (3) and (4), resulting in

$$u(x,t) = \sum_i S_i(x,\omega)\cos(\omega t + \phi_i) \tag{14}$$

and

$$u(y,t) = \sum_i S_i(y,\omega)\cos\left[\omega(t + t_d^i) + \phi_i\right], \tag{15}$$

where the summation is over the sources, ϕ_i is the phase delay associated with source i, and the time delay t_d^i between stations x and y also changes with source azimuth, hence the superscript i. In analogy with sec. 2.1, we next substitute (14) and (15) into (1), and

$$C_{xy} = \frac{1}{2T}\sum_{i,k}\left\{S_i(x)S_k(y)\int_{-T}^{T}\cos(\omega\tau + \phi_i)\cos\left[\omega(\tau + t + t_d^k) + \phi_k\right]\mathrm{d}\tau\right\}. \tag{16}$$

Let us consider the "cross-terms" (cross-correlations between $\cos(\omega\tau+\phi_i)$ and $\cos\left[\omega(\tau + t + t_d^k) + \phi_k\right]$ with $i \neq k$) in eq. (16): they are sinusoidal with the same frequency ω but randomly out of phase, and therefore do not interfere constructively. The remaining ($i = k$) terms, on the other hand, interfere constructively, as we shall illustrate below, so that, after the contribution of a sufficient number of sources has been taken into account, the cross-term contribution becomes negligible relative to them. Following other derivations of noise-correlation properties, we thus neglect cross-terms from this point on [e.g., *Snieder*, 2004; *Tsai*, 2009]. We are left with a sum of integrals of the form (5), which we have proved in sec. 2.1 to be approximated by (13), so that

$$C_{xy} \approx \sum_i \frac{S_i(x)S_i(y)}{2}\cos\left[\omega(t + t_d^i)\right]. \tag{17}$$

2.3 Continuous distribution of sources

Eq. (17) can be further generalized to the case of a continuous distribution of sources,

$$C_{xy} \approx \int_{-\frac{\Delta x}{v}}^{\frac{\Delta x}{v}} \rho(t_d,\omega)\cos\left[\omega(t + t_d)\right]\mathrm{d}t_d, \tag{18}$$

where we have introduced the function $\rho(t_d,\omega)$, describing the density of sources as a function of inter-station delay t_d, or, which is the same (recall eq. (2)), azimuth θ. Integration is accordingly over t_d, and the integration limits correspond, through eq. (2), to the interval of possible azimuths, from 0 to π. ρ is also a function of ω, as signal generated by differently located sources generally has a different frequency content. To keep the notation compact, we have incorporated the continuous version of the source term $S_i(x,\omega)S_i(y,\omega)/2$ from eq. (17) in the source density function $\rho(t_d,\omega)$.

In analogy with earlier formulations of ambient-noise theory, we require the source distribution to be uniform with respect to azimuth θ. To find the corresponding (not constant) expression of ρ as a function of t_d, we note that, for azimuthally constant source density, $\rho(t_d)$ multiplied by a positive increment $|\mathrm{d}t_d|$ must coincide with the corresponding increment $|\mathrm{d}\theta|$ times a constant factor. Formally,

$$\frac{1}{2\pi}g(\omega)|\mathrm{d}\theta| = \rho(t_d,\omega)|\mathrm{d}t_d|, \tag{19}$$

where $g(\omega)/2\pi$ is the normalized value of uniform azimuthal source density, selected so that its integral between 0 and 2π is exactly $g(\omega)$. The factor $g(\omega)$ serves to remind us that source amplitude generally changes with frequency. After replacing $|\mathrm{d}t_d| = \Delta x \sin(\theta)|\mathrm{d}\theta|/v$,

$$\frac{g(\omega)}{2\pi}\,|\mathrm{d}\theta| = \rho(t_d,\omega)\frac{\Delta x \sin[\theta(t_d)]}{v}|\mathrm{d}\theta|, \tag{20}$$

or

$$\rho(t_d,\omega) = \frac{v(\omega)g(\omega)}{2\pi\Delta x \sin[\theta(t_d)]}, \tag{21}$$

which is the expression of $\rho = \rho(t_d,\omega)$ corresponding to azimuthally uniform source density.

2.4 Cross-correlation and Green function

It is convenient to separate the integral in eq. (18) into two integrals, one over positive, and the other over negative t_d,

$$C_{xy} \approx \int_{-\frac{\Delta x}{v}}^{0} \rho(t_d,\omega)\cos\left[\omega(t+t_d)\right]\mathrm{d}t_d \;+\; \int_{0}^{\frac{\Delta x}{v}} \rho(t_d,\omega)\cos\left[\omega(t+t_d)\right]\mathrm{d}t_d. \tag{22}$$

The negative- and positive-time contributions to C_{xy} are usually referred to as anticausal and causal, respectively.

2.4.1 Positive-time (causal) contribution to the cross-correlation

Let us first consider the second term ($t_d \geq 0$) at the right-hand side of (22), which, since $\rho(t_d,\omega)$ is real (see eq. (21)), can be rewritten

$$C_{xy}^{t_d>0} \approx \Re\left[\mathrm{e}^{\mathrm{i}\omega t}\int_{0}^{\frac{\Delta x}{v}} \rho(t_d,\omega)\mathrm{e}^{\mathrm{i}\omega t_d}\mathrm{d}t_d\right], \tag{23}$$

where $\Re(\ldots)$ equals the real part of its argument. It is convenient to replace $\rho(t_d,\omega)$ with its expression (21), and the integration variable t_d with θ. By differentiating eq. (2), $\mathrm{d}t_d = -\Delta x \sin(\theta)\mathrm{d}\theta/v$, while the limits of integration 0, $\Delta x/v$ correspond to azimuth $\theta = \pi/2$, 0, respectively, hence, using the symmetry of the cosine,

$$C_{xy}^{t_d>0} \approx \Re\left[\frac{g(\omega)\mathrm{e}^{\mathrm{i}\omega t}}{2\pi}\int_{0}^{\frac{\pi}{2}} \mathrm{e}^{\mathrm{i}\omega\Delta x\cos(\theta)/v}\mathrm{d}\theta\right]. \tag{24}$$

(Recall that positive t_d corresponds to azimuth $0 < \theta < \pi/2$, while the opposite holds for the $t_d \leq 0$ term corresponding to $\pi/2 < \theta < \pi$.)

We next rewrite the integral in terms of Bessel and Struve functions. Let us first consider the 0-order Bessel function of the first kind in its integral form

$$J_0(z) = \frac{1}{\pi}\int_{0}^{\pi}\cos(z\sin(\theta))\mathrm{d}\theta \tag{25}$$

(eq. (9.1.18) of *Abramowitz and Stegun* [1964]). The integral from 0 to π in (25) can be

196 transformed into an integral from 0 to $\pi/2$:

$$\begin{aligned} J_0(z) &= \frac{1}{\pi}\int_0^{\pi}\cos(z\sin(\theta))\mathrm{d}\theta \\ &= \frac{1}{\pi}\left[\int_0^{\frac{\pi}{2}}\cos(z\sin(\theta))\mathrm{d}\theta + \int_{\frac{\pi}{2}}^{\pi}\cos(z\sin(\theta))\mathrm{d}\theta\right] \\ &= \frac{1}{\pi}\left[\int_0^{\frac{\pi}{2}}\cos(z\sin(\theta))\mathrm{d}\theta - \int_{\frac{\pi}{2}}^{0}\cos(z\sin(\pi-\theta'))\mathrm{d}\theta'\right] \\ &= \frac{1}{\pi}\left[\int_0^{\frac{\pi}{2}}\cos(z\sin(\theta))\mathrm{d}\theta + \int_0^{\frac{\pi}{2}}\cos(z\sin(\theta'))\mathrm{d}\theta'\right] \\ &= \frac{2}{\pi}\int_0^{\frac{\pi}{2}}\cos(z\sin(\theta))\mathrm{d}\theta. \end{aligned} \tag{26}$$

197 We then replace $\sin(\theta) = \cos(\theta - \pi/2)$ and change the integration variable $\theta = \theta' + \pi/2$,

$$\begin{aligned} J_0(z) &= \frac{2}{\pi}\int_0^{\frac{\pi}{2}}\cos\left(z\cos\left(\theta-\frac{\pi}{2}\right)\right)\mathrm{d}\theta \\ &= \frac{2}{\pi}\int_{-\frac{\pi}{2}}^{0}\cos\left(z\cos\left(\theta'\right)\right)\mathrm{d}\theta' \\ &= \frac{2}{\pi}\int_0^{\frac{\pi}{2}}\cos\left(z\cos\left(\theta'\right)\right)\mathrm{d}\theta', \end{aligned} \tag{27}$$

198 and after substituting z with $\omega\Delta x/v$,

$$\begin{aligned} J_0\left(\frac{\omega\Delta x}{v}\right) &= \frac{2}{\pi}\int_0^{\frac{\pi}{2}}\cos\left(\frac{\omega\Delta x}{v}\cos\left(\theta\right)\right)\mathrm{d}\theta \\ &= \frac{2}{\pi}\,\Re\left[\int_0^{\frac{\pi}{2}}\mathrm{e}^{\mathrm{i}\omega\Delta x\cos(\theta)/v}\mathrm{d}\theta\right]. \end{aligned} \tag{28}$$

199 The 0-order Struve function also has an integral form

$$H_0(z) = \frac{2}{\pi}\int_0^{\frac{\pi}{2}}\sin(z\cos(\theta))\mathrm{d}\theta, \tag{29}$$

200 which coincides with eq. (12.1.7) of *Abramowitz and Stegun* [1964] at order 0 and substituting
201 $\Gamma(1/2) = \sqrt{\pi}$, with Γ denoting the Gamma function. We replace, again, z with $\omega\Delta x/v$, and

$$H_0\left(\frac{\omega\Delta x}{v}\right) = \frac{2}{\pi}\,\Im\left[\int_0^{\frac{\pi}{2}}\mathrm{e}^{\mathrm{i}\omega\Delta x\cos(\theta)/v}\mathrm{d}\theta\right], \tag{30}$$

202 with the operator $\Im$ mapping complex numbers to their imaginary part. It follows from (28)
203 and (30) that

$$\int_0^{\frac{\pi}{2}}\mathrm{e}^{\mathrm{i}\omega\Delta x\cos(\theta)/v}\mathrm{d}\theta = \frac{\pi}{2}\left[J_0\left(\frac{\omega\Delta x}{v}\right) + \mathrm{i}H_0\left(\frac{\omega\Delta x}{v}\right)\right], \tag{31}$$

204 and substituting into (24):

$$C_{xy}^{t_d>0} \approx \Re\left\{\frac{g(\omega)\mathrm{e}^{\mathrm{i}\omega t}}{4}\left[J_0\left(\frac{\omega\Delta x}{v}\right) + \mathrm{i}H_0\left(\frac{\omega\Delta x}{v}\right)\right]\right\}. \tag{32}$$

205 Following *Tsai* [2009], or all other authors conducting ambient-noise analysis in the time
206 domain, we next assume that inter-station distance be much larger than the wavelength of the

signal under consideration, i.e. $\omega\Delta x/v \gg 1$. It then follows from eq. (9.2.1) of *Abramowitz and Stegun* [1964] that

$$J_0\left(\frac{\omega\Delta x}{v}\right) \approx \sqrt{\frac{2\,v}{\omega\pi\Delta x}}\cos\left(\frac{\omega\Delta x}{v}-\frac{\pi}{4}\right), \tag{33}$$

and from eqs. (12.1.34) and (9.2.2) of *Abramowitz and Stegun* [1964],

$$H_0\left(\frac{\omega\Delta x}{v}\right) \approx Y_0\left(\frac{\omega\Delta x}{v}\right) \approx \sqrt{\frac{2\,v}{\omega\pi\Delta x}}\sin\left(\frac{\omega\Delta x}{v}-\frac{\pi}{4}\right), \tag{34}$$

with Y_0 denoting the 0-order Bessel function of the second kind.

Substituting equations (33) and (34) into (32),

$$\begin{aligned} C_{xy}^{t_d>0} &\approx \Re\left\{g(\omega)\mathrm{e}^{\mathrm{i}\omega t}\sqrt{\frac{v}{8\pi\omega\Delta x}}\left[\mathrm{e}^{\mathrm{i}(\omega\Delta x/v-\pi/4)}\right]\right\} \\ &= g(\omega)\sqrt{\frac{v}{8\pi\omega\Delta x}}\cos\left[\omega\left(\Delta x/v+t\right)-\pi/4\right]. \end{aligned} \tag{35}$$

Comparing eq. (35) to (13), we note a phase-shift $\pi/4$ between the cross-correlated signal generated by a teleseismic event aligned with the two stations, and that obtained from the ensemble-averaging of seismic ambient noise. $\pi/4$ is nothing but the phase-shift between a cosine and a Bessel function, for large values of the argument (i.e., in the far field). In our experimental set-up, a cosine describes the two-station cross-correlation of a plane wave hitting the receivers from a single azimuth; the Bessel function (and hence the $\pi/4$ shift) emerges from the combined effect of plane waves coming from all azimuths (i.e. focusing over the receiver array).

2.4.2 Negative-time (anticausal) contribution to the cross-correlation

An analogous treatment applies to the negative-time cross-correlation $C_{xy}^{t_d<0}$, i.e. the first term at the right hand side of eq. (22), which after the variable change from t_d to θ becomes

$$C_{xy}^{t_d<0} \approx \Re\left[\frac{g(\omega)\mathrm{e}^{\mathrm{i}\omega t}}{2\pi}\int_{\frac{\pi}{2}}^{\pi}\mathrm{e}^{\mathrm{i}\omega\Delta x\cos(\theta)/v}\mathrm{d}\theta\right]. \tag{36}$$

To express also this integral in terms of Bessel and Struve functions, we first notice that

$$\begin{aligned} \int_{\frac{\pi}{2}}^{\pi} f(\cos(\theta))\mathrm{d}\theta &= \int_0^{\frac{\pi}{2}} f\left(\cos\left(\theta'+\frac{\pi}{2}\right)\right)\mathrm{d}\theta' \\ &= \int_0^{\frac{\pi}{2}} f\left(\cos(\theta')\cos\left(\frac{\pi}{2}\right)-\sin(\theta')\sin\left(\frac{\pi}{2}\right)\right)\mathrm{d}\theta' \\ &= \int_0^{\frac{\pi}{2}} f(-\sin(\theta'))\mathrm{d}\theta', \end{aligned} \tag{37}$$

for an arbitrary function f. From eq. (36) it then follows that

$$C_{xy}^{t_d<0} \approx \Re\left[\frac{g(\omega)\mathrm{e}^{\mathrm{i}\omega t}}{2\pi}\int_0^{\frac{\pi}{2}}\mathrm{e}^{-\mathrm{i}\omega\Delta x\sin(\theta)/v}\mathrm{d}\theta\right]. \tag{38}$$

Similar to eq. (27) in section 2.4.1, we next replace $\cos(\theta) = \sin(\theta + \pi/2)$ in expression (29) for the Struve function, and change the integration variable $\theta' = \theta + \frac{\pi}{2}$,

$$\begin{aligned} H_0(z) &= \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \sin\left(z\cos\left(\theta\right)\right) \mathrm{d}\theta \\ &= \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \sin\left(z\sin\left(\theta + \frac{\pi}{2}\right)\right) \mathrm{d}\theta \\ &= \frac{2}{\pi}\int_{\frac{\pi}{2}}^{\pi} \sin\left(z\sin\left(\theta'\right)\right) \mathrm{d}\theta' \\ &= \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \sin\left(z\sin\left(-\theta'\right)\right) \mathrm{d}\theta' \\ &= -\frac{2}{\pi}\int_0^{\frac{\pi}{2}} \sin\left(z\sin\left(\theta'\right)\right) \mathrm{d}\theta'. \end{aligned} \tag{39}$$

Making use of eq. (39), and of expression (26) for the Bessel function J_0, with $z = \omega\Delta x/v$, in (38),

$$C_{xy}^{t_d<0} \approx \Re\left\{\frac{g(\omega)\mathrm{e}^{-\mathrm{i}\omega t}}{4}\left[J_0\left(\frac{\omega\Delta x}{v}\right) - \mathrm{i}H_0\left(\frac{\omega\Delta x}{v}\right)\right]\right\}, \tag{40}$$

where only the sign of H_0 at the right-hand side has changed with respect to eq. (32). We conclude that

$$C_{xy}^{t_d<0} \approx g(\omega)\sqrt{\frac{v}{8\pi\omega\Delta x}}\cos\left[\omega\left(-\Delta x/v + t\right) + \pi/4\right], \tag{41}$$

i.e. the negative-time phase-shift is symmetric to the positive-time one, in agreement with *Tsai* [2009].

Summing $C_{xy}^{t_d<0}$ (eq. (41)) and $C_{xy}^{t_d>0}$ (eq. (35)) one finds, according to eq. (22), an expression for C_{xy} valid at all, positive and negative times. To verify its validity, we implement it numerically and compare it in Fig. 2 to the result of eq. (17) applied to a very large set of sources, for the same frequency and inter-station distance. Confirming earlier findings, the two differently computed cross-correlations are practically coincident.

2.5 Group and phase velocity

We next consider the more general case of a seismogram formed by the superposition of surface waves with different frequencies. Let us start with our expression (35) for the cross-correlated signal, grouping the amplitude terms in a generic positive factor $S(\omega)$. We then find the mathematical expression of a surface-wave packet by (i) discretizing the frequency band of interest into a set of closely-spaced frequencies ω_i identified by the subscript i, and (ii) combining different-frequency contributions by integration around each frequency ω_i and summation over i, so that

$$u(x,t) = \sum_{i=1}^{\infty}\int_{\omega_i-\varepsilon}^{\omega_i+\varepsilon} S(x,\omega)\cos\left[\omega\left(\frac{\Delta x}{v(\omega)} + t\right) - \frac{\pi}{4}\right]\mathrm{d}\omega, \tag{42}$$

where $\varepsilon \ll \omega_i$. It is convenient to introduce the notation $\psi = \omega(\Delta x/v + t) - \pi/4$, and, since ε is small, replace it with its Taylor expansion around ω_i, i.e.

$$\psi(\omega) \approx \psi(\omega_i) + (\omega - \omega_i)\left[\frac{\mathrm{d}\psi}{\mathrm{d}\omega}\right]_{\omega_i}, \tag{43}$$

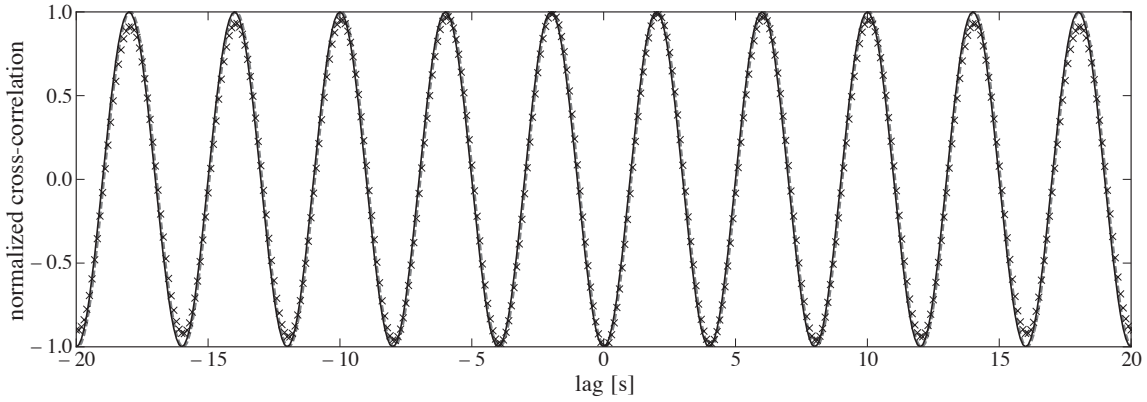

Figure 2: Numerical test of expression (41) + (35), with interstation distance of 500 km and wave speed of 3 km/s. C_{xy} resulting from the direct implementation of (41) + (35) is denoted by a solid line. We compare it with the result of applying eq. (17) to model C_{xy} from the combined effect of 1000, far, sinusoidal (with 4-s period) out-of-phase sources located at 200 different, uniformly distributed azimuths from the station couple. Finally, we also compute C_{xy} from eq. (16) (crosses), neglecting the cross-terms $i \neq k$; a slight decay, with increasing lag, in the latter estimate of C_{xy} is caused by the finite length of the time-integral in the implementation of (16). Amplitudes have been normalized. All modeled cross-correlations are perfectly in phase.

where $[f(\omega)]_{\omega_i}$ denotes the value of any function f evaluated at $\omega = \omega_i$. We rewrite eq. (42) accordingly, and find after some algebra that the integral at its right hand side

$$\int_{\omega_i-\varepsilon}^{\omega_i+\varepsilon} S(\omega) \cos\left[\omega\left(\frac{\Delta x}{v(\omega)}+t\right)-\frac{\pi}{4}\right] \mathrm{d}\omega \approx S(\omega_i)\cos\left[\psi(\omega_i)\right] \frac{2\sin\left\{\varepsilon\left[\frac{\mathrm{d}\psi}{\mathrm{d}\omega}\right]_{\omega_i}\right\}}{\left[\frac{\mathrm{d}\psi}{\mathrm{d}\omega}\right]_{\omega_i}} \tag{44}$$

(valid in the assumption that S be a smooth function of ω). If one introduces a function

$$v_g(\omega) = \frac{v(\omega)}{1-\frac{\omega}{v(\omega)}\frac{\mathrm{d}v}{\mathrm{d}\omega}}, \tag{45}$$

it follows that $\frac{\mathrm{d}\psi}{\mathrm{d}\omega}$ takes the compact form

$$\left[\frac{\mathrm{d}\psi}{\mathrm{d}\omega}\right]_{\omega_i} = \frac{\Delta x}{v_g(\omega_i)} + t; \tag{46}$$

we finally substitute it into (44) and substitute the resulting expression into (42), to find

$$u(x,t) = \sum_{i=1}^{\infty} S(\omega_i)\cos\left[\omega_i\left(\frac{\Delta x}{v(\omega_i)}+t\right)-\frac{\pi}{4}\right] \frac{2\sin\left[\varepsilon\left(\frac{\Delta x}{v_g(\omega_i)}+t\right)\right]}{\left[\frac{\Delta x}{v_g(\omega_i)}+t\right]}. \tag{47}$$

Each term at the right-hand side of eq. (47) is the product of a wave of frequency ω and speed $v(\omega)$ with one of frequency $\varepsilon \ll \omega$ and speed $v_g(\omega_i)$. The latter factor, with much lower frequency, modulates the signal, and we call "group velocity" its speed v_g, which coincides with the speed of the envelope of the signal. Eq. (45) shows that, in the absence of dispersion (i.e. $\frac{\mathrm{d}v}{\mathrm{d}\omega} = 0$) phase and group velocities coincide. In practice, the values of v and v_g are

always comparable, and the large difference in frequency results in a large difference in the wavelength of the phase and group terms.

Comparing eq. (47) to (42), it is important to notice that when phase velocity is measured from the station-station cross-correlation of ambient signal, a phase correction of $\pi/4$ must first be applied; the same is not true for group-velocity measurements. We have shown in sections 2.4.1 and 2.4.2 that ambient-noise cross-correlation coincides with a combination of Bessel functions, and that, for large values of their argument (corresponding to relatively large inter-station distance), Bessel functions can be replaced by sinusoidal functions, whose argument coincides with the argument of the Bessel functions minus $\pi/4$. The $\pi/4$-shift in (42) and (47) arises precisely from this far-field approximation.

3 How to measure phase velocity

To evaluate whether phase velocity can be accurately observed in the ensemble-averaged cross-correlation of ambient noise, we use two independent approaches to measure it from the same data. Consistency of the results is then an indication of their validity. The first approach (section 3.1) consists of cross-correlating and stacking the surface-wave signal (Δt-long records of ambient signal in our case) to find the empirical Green function (sec. 2.4), from which phase velocity can be measured [e.g., sec. 12.6.2 of *Udías*, 1999]. If, as is most often the case, one works in the far-field approximation, this requires that a $\pi/4$ correction be applied to the data as explained in section 2.5, eq. (47). The other approach we consider is based on the result of *Aki* [1957], confirmed by *Ekström et al.* [2009] for the frequency range of interest, that the spectrum of the two-station cross-correlation of seismic ambient noise should approximately coincide with a 0-order Bessel function of the first kind (section 3.2); in this case, no $\pi/4$ correction needs to be applied.

3.1 Time-domain cross-correlation

The procedure of ensemble-averaging ambient signal is described in detail, e.g., by *Bensen et al.* [2007]; a long (e.g., one year) continuous seismic record is subdivided into shorter Δt intervals. The records are whitened, and they are normalized in the time-domain so that the effects of possible ballistic signal (i.e., large earthquakes) present in the data are minimized. The cross-correlation between simultaneous Δt-long records from different stations is then computed for all available Δt intervals, and the results for each station pair are stacked over the entire year.

Bensen et al. [2007] measure group velocity from noise cross-correlations, and suggest that phase dispersion can be obtained by integration of group dispersion curves. This approach however is not sufficient to identify phase velocity uniquely. *Meier et al.* [2004] provide an algorithm to derive phase velocity from the cross-correlation of teleseismic signals recorded by stations aligned with the earthquake azimuth. *Fry et al.* [2010] and *Verbeke et al.* [2012a] show that the algorithm of *Meier et al.* [2004] can be successfully applied to the ambient signal recorded at a regional-scale array of broadband stations. In reference to the study of *Fry et al.* [2010] where it was first introduced, we shall dub this approach FRY. In the following we shall analyze a subset of the phase-dispersion database compiled by *Verbeke et al.* [2012a] via their own automated implementation of FRY.

The phase-velocity measurements of *Verbeke et al.* [2012a] are limited to the 0.02-0.1 Hz frequency range, where seismic ambient noise is known to be strong [*Stehly et al.*, 2009], most likely as an effect of ocean storms and the coupling between oceans and the solid Earth [*Stehly et al.*, 2006]. Frequency is discretized with increments whose length increases with increasing frequency (from 0.02 to 0.05 Hz). For each discrete frequency value, ensemble-averaged cross-correlations are (i) band-pass filtered around the frequency in question and (ii) windowed in the time-domain via a Gaussian window centered around the time of maximum amplitude of (filtered) cross-correlation. Causal and anticausal parts are folded together (i.e. stacked after reversing the time-dependence of the anticausal one). The resulting time series is Fourier-transformed, and its phase is identified as the arctangent of the ratio of the imaginary to real part of the Fourier spectrum, as explained by *Udías* [1999], section 12.6.1. Based on eq. (35), one must sum $\pi/4$ to the resulting folded ensemble-averaged cross-correlation phase before applying eq. (3) of *Meier et al.* [2004] (equivalent to eq. (12.56) of *Udías* [1999]). Importantly, this $\pi/4$ shift is specific to ambient-noise cross-correlation, and must not be applied in two-station analysis of ballistic surface-wave signal, as shown by eq. (13). Phase velocity is only known up to a $2\pi n$ "multiple cycle ambiguity", with $n = 0, \pm 1, \pm 2, ...$ After iterating over the entire frequency band, an array of dispersion curves is found, each corresponding to a value of n. *Verbeke et al.* [2012a] compare each curve (for all integer values of n between -5 and 5) with phase velocity as predicted by PREM [*Dziewonski and Anderson*, 1981], and pick the one closest to PREM, considering only the frequency range where the measurement is reliable (no large jumps for small variations in frequency). (More sophisticated procedures exist to resolve the ambiguity [e.g., *Lin and Ritzwoller*, 2011; *Gouedard et al.*, 2012], but here we stick to the simpler algorithm of *Verbeke et al.* [2012a].)

Ensemble-averaged cross-correlations for two Swiss stations (Fig. 3a) are shown in Fig. 3b. At long period (compared to interstation distance divided by wave speed) the causal and anti-causal parts of the cross-correlation overlap, complicating the time-domain analysis of cross-correlation, whose results are shown in Fig. 3c.

3.2 Frequency-domain cross-correlation and Bessel-function fitting

A different method, hereafter referred to as "AKI", to extrapolate phase velocity from the ambient signal recorded at two stations is proposed by *Ekström et al.* [2009], based on much earlier work by *Aki* [1957]. The theoretical basis of this method has been recently rederived by *Nakahara* [2006], *Yokoi and Margaryan* [2008] and *Tsai and Moschetti* [2010]. As pointed out by *Ekström et al.* [2009], this approach does not require that $\omega \Delta x/v \gg 1$, i.e. it will work for wavelengths comparable to interstation distance.

According to AKI, ambient signal recorded over a long time (e.g., one year) is, again, subdivided into shorter Δt intervals. Let us call $p_i(\omega)$ the frequency spectrum associated with a Δt-long record at station i (Fig. 4a, with Δt=2 hours). After whitening, this is multiplied with the simultaneous Δt-long recording made at another station j (Fig. 4b), resulting in the cross-spectrum, or spectrum of the cross-correlation between the two Δt-long records (Fig. 4c). This procedure is repeated for all available Δt-intervals in the year, which are then stacked together, i.e. ensemble-averaged (Fig. 4d). The resulting quantity is usually

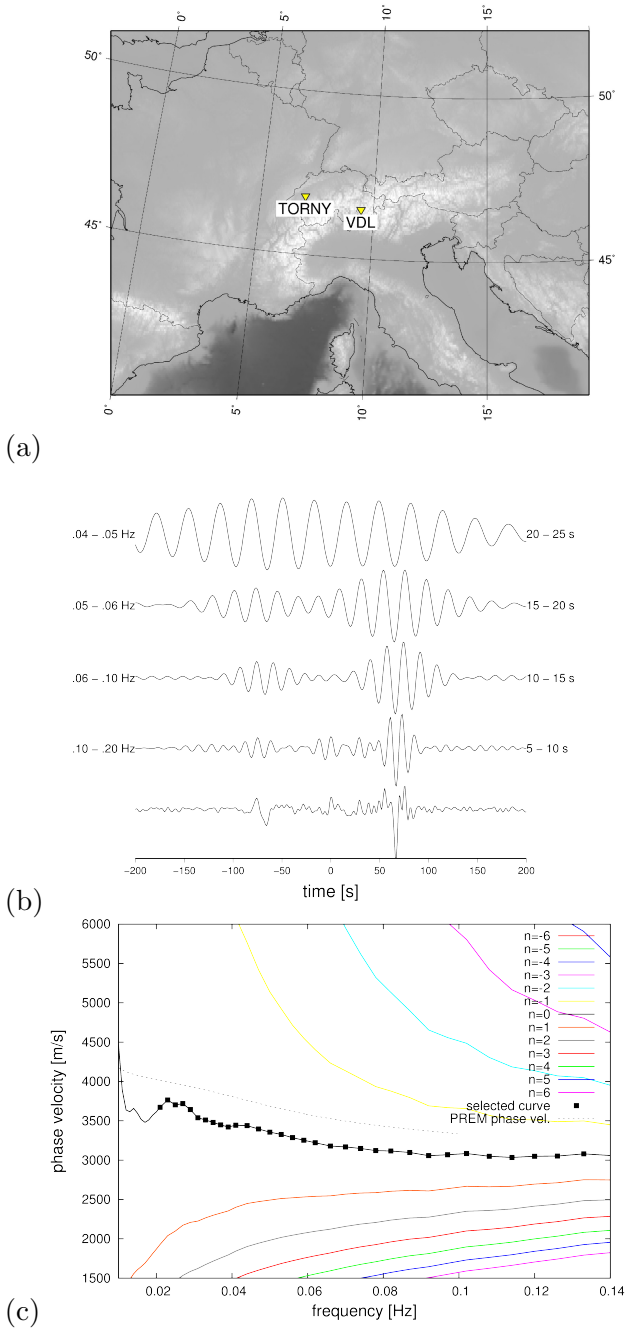

Figure 3: Illustration of the FRY method. (a) Locations (triangles) of stations TORNY and VDL, from the Swiss broadband network. (b) Ensemble-averaged cross-correlation of continuous signal recorded at TORNY and VDL, filtered over different frequency bands as indicated; the bottom trace is the "full" waveform. (c) Array of possible phase-velocity dispersion curves from cross-correlation of the continuous recordings made at TORNY and VDL; each curve corresponds to a different value of n, identified by the curve colour as indicated. The black curve, closest to our selected reference model (PREM), is our preferred one, but observations are only considered valid in the frequency range marked by black squares.

referred to as "coherency". Based on *Aki* [1957],

$$\left\langle \Re\left(\frac{p_i p_j^*}{|p_i|\,|p_j|}\right)\right\rangle \propto J_0\left(\frac{\omega\Delta x}{v(\omega)}\right), \tag{48}$$

where $< ... >$ denotes ensemble averaging, the left-hand side is precisely what we call coherency, and the superscript $*$ marks the complex conjugate of a complex number. The quantities at the right-hand side of (48) are defined as in section 2.4 above, with Δx distance between stations i and j. (The alert reader might notice at this point that the right-hand side of eq. (48) is proportional to C_{xy}: simply sum, according to eq. (22), its positive- and negative-time contributions (32) and (40), respectively [*Tsai and Moschetti*, 2010].) Again based on *Aki* [1957], the ensemble-averaged imaginary part

$$\left\langle \Im\left(\frac{p_i p_j^*}{|p_i|\,|p_j|}\right)\right\rangle = 0. \tag{49}$$

Importantly, both equations (48) and (49) are shown by *Aki* [1957] to be valid provided that the energy of ambient signal is approximately uniform with respect to azimuth. As anticipated at the beginning of section 2, this is typically not true at any moment in time, but can be achieved, at least to some extent, by ensemble-averaging [*Yang and Ritzwoller*, 2008].

Eq. (48) can be used to determine phase dispersion. In practice, observed coherency is first of all plotted as a function of frequency (i.e., the ensemble-averaged, whitened cross-spectrum is plotted). Values ω_i $(i = 1, 2, 3, ...)$ of frequency for which coherency is zero are identified. If $\omega = \omega_i$ for some i, the argument of (48) must coincide with one of the known zeros z_n $(n = 1, 2, ...)$ of the Bessel function J_0,

$$\frac{\omega_i \Delta x}{v(\omega_i)} = z_n. \tag{50}$$

Eq. (50) can be solved for v,

$$v(\omega_i) = \frac{\omega_i \Delta x}{z_n}, \tag{51}$$

and we now have an array of possible measurements of phase velocity at the frequency ω_i, each corresponding to a different value of n. Implementing (51) at all observed values of ω_i, an array of dispersion curves is found. Much like in the case of FRY (section 3.1), a criterion must then be established to select a unique curve.

Importantly, the observation of ω_i on ensemble-averaged cross-spectra like the one of Fig. (4d) is complicated by small oscillations that can be attributed to instrumental noise or inaccuracies related to nonuniformity in the source distribution. Before identifying ω_i, we determine the linear combination of cubic splines that best fits (in least-squares sense, via the LSQR algorithm of *Paige and Saunders* [1982]) observed coherency. Splines are equally spaced, and spacing must be selected so that "splined" coherency is sufficiently smooth (Fig. 4d).

Equations (48) and (49) are rarely satisfied by seismic ambient noise as observed in the real world. At a given time, the wavefield associated with ambient noise is not diffuse. The procedure of ensemble-averaging over a long time serves precisely to mimic a diffuse wavefield by combining non-diffuse ones. Yet, there are important systematic effects that ensemble-averaging does not remove: in Europe, for example, most of the recorded seismic

noise is generated in the Atlantic Ocean [*Stehly et al.*, 2006, 2009; *Verbeke et al.*, 2012b], and the requirement of an azimuthally uniform source distribution is accordingly not met. Presumably, scattering partly compensates for that, but the nonzero observed imaginary part of the coherency shown e.g. in Fig. 4d indicates that the problem remains [e.g., *Cox*, 1973]. The imaginary part should converge to zero if one ensemble-averages not only over time, but also over station-pair azimuth [e.g., *Weemstra et al.*, 2012], but then information on lateral Earth structure would be lost.

It is practical to focus the analysis on zero crossings, rather than measuring the overall fit between J_0 and measured coherency. The latter depends on the power spectrum of the noise sources, of which we know very little, and can be affected importantly by data processing [*Ekström et al.*, 2009].

4 Application to central European data and cross-validation of the two methods

Fig. 5 shows the set of ∼1000 randomly selected station pairs from *Verbeke et al.* [2012a] that we shall analyze here. The corresponding phase-velocity dispersion curves were measured by *Verbeke et al.* [2012a] following the procedure of section 3.1, after subdividing the entire year 2006 into day-long intervals and ensemble-averaging the resulting day-long cross-correlations.

We apply the AKI method of section 3.2 to continuous records associated with the station pairs of Fig. 5. Our implementation was originally designed for reservoir-scale application [*Weemstra et al.*, 2012], but could be applied to our continent-scale array of data after only minor modifications. For each station, continuous recording for the entire year 2006 is subdivided into intervals of $\Delta t = 2$ hours, with a very conservative 75% overlap between neighboring intervals to make sure that no coherent signal traveling from station to station is neglected [*Seats et al.*, 2012; *Weemstra et al.*, 2012]. This results in as many as 45 spectra per day.

In Fig. 6 we compare our new phase-velocity measurements with those of *Verbeke et al.* [2012a] for three example station pairs. A visual analysis (which we repeated on many more pairs) suggests that the two methods provide very similar results.

To evaluate quantitatively their level of consistency, we first expand FRY dispersion curves over a set of cubic splines, and apply spline interpolation to estimate FRY-based phase-velocity values at the exact frequencies (associated with zero-crossings of the Bessel function) where AKI measurements are available. We subtract the AKI phase velocities from the FRY ones interpolated at the same frequency, selecting at each frequency the AKI data point closest to the FRY one (we thus avoid the well known issue of multiple-cycle ambiguity, that equally affects both approaches). We count the number of discrepancy observations, independent of frequency, falling in each of a set of 50 m/s intervals, and plot the associated histogram in Fig. 7. Both mean and standard deviation of the FRY-AKI discrepancy are small (13 m/s and 151 m/s, respectively), and we conclude that, in our implementation, the two approaches provide consistent results when applied to the data. Outliers exist with misfit larger than ±1000 m/s, but they would not be visible in Fig. 7 even after extending the horizontal-axis range.

We next analyze the dependence of FRY-AKI discrepancy on interstation distance, through

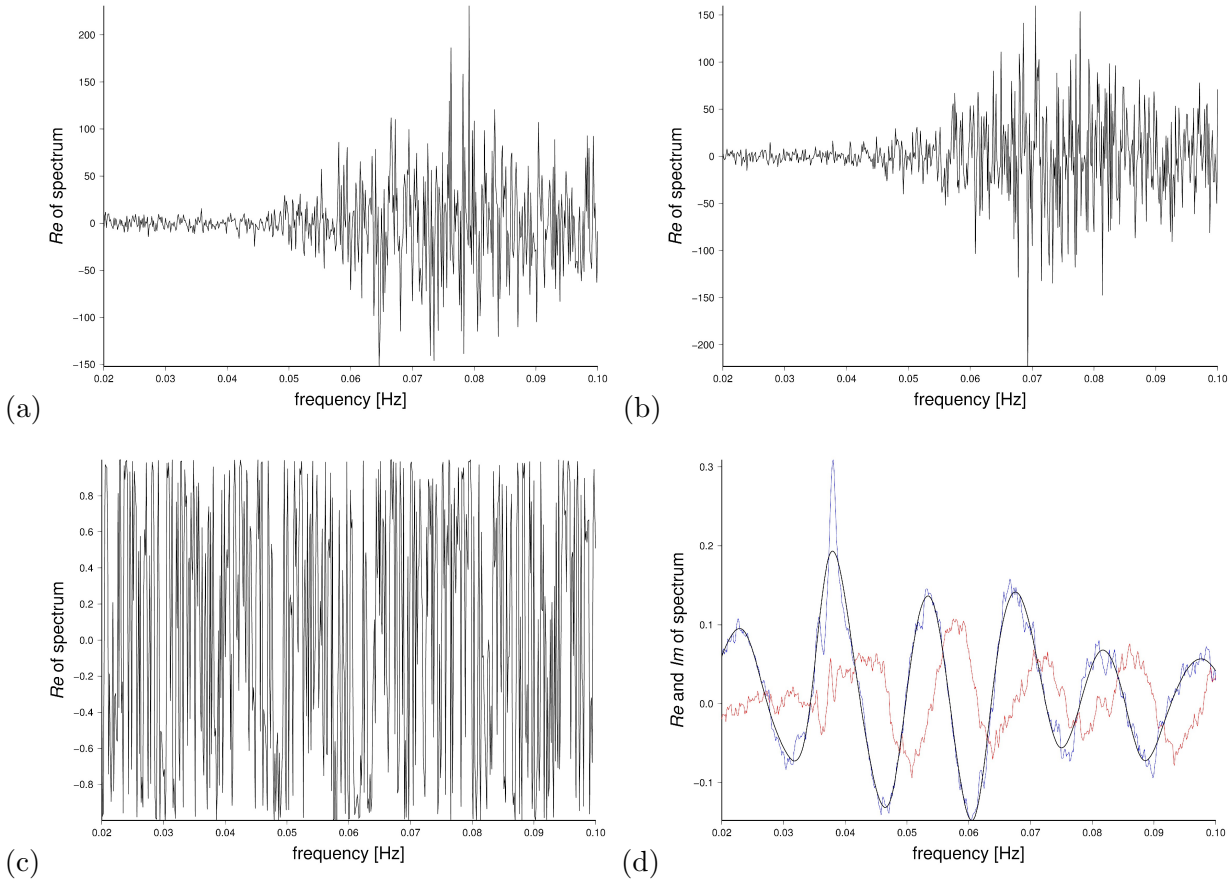

Figure 4: Illustration of the AKI approach. (a) Real part of the spectrum (m·s) obtained Fourier-transforming two hours of ambient recording at station TORNY. (b) same as (a), from the very same two hours of signal recorded at station VDL. (c) Product of (a) and (b) (coinciding with the real part of the spectrum of the cross-correlation of the two time-domain signals) obtained after whitening both. (d) Results of ensemble-averaging an entire year of spectra like the one at (c), for the same two stations: the blue and red lines identify values of real and imaginary parts found at various frequencies; the black solid line is the linear combination of cubic splines that best-fits the observed real part of the spectrum. The locations of stations TORNY and VDL are shown in Fig. 3a. Cross-spectra in both (c) and (d) are implicitly normalized and hence unitless.

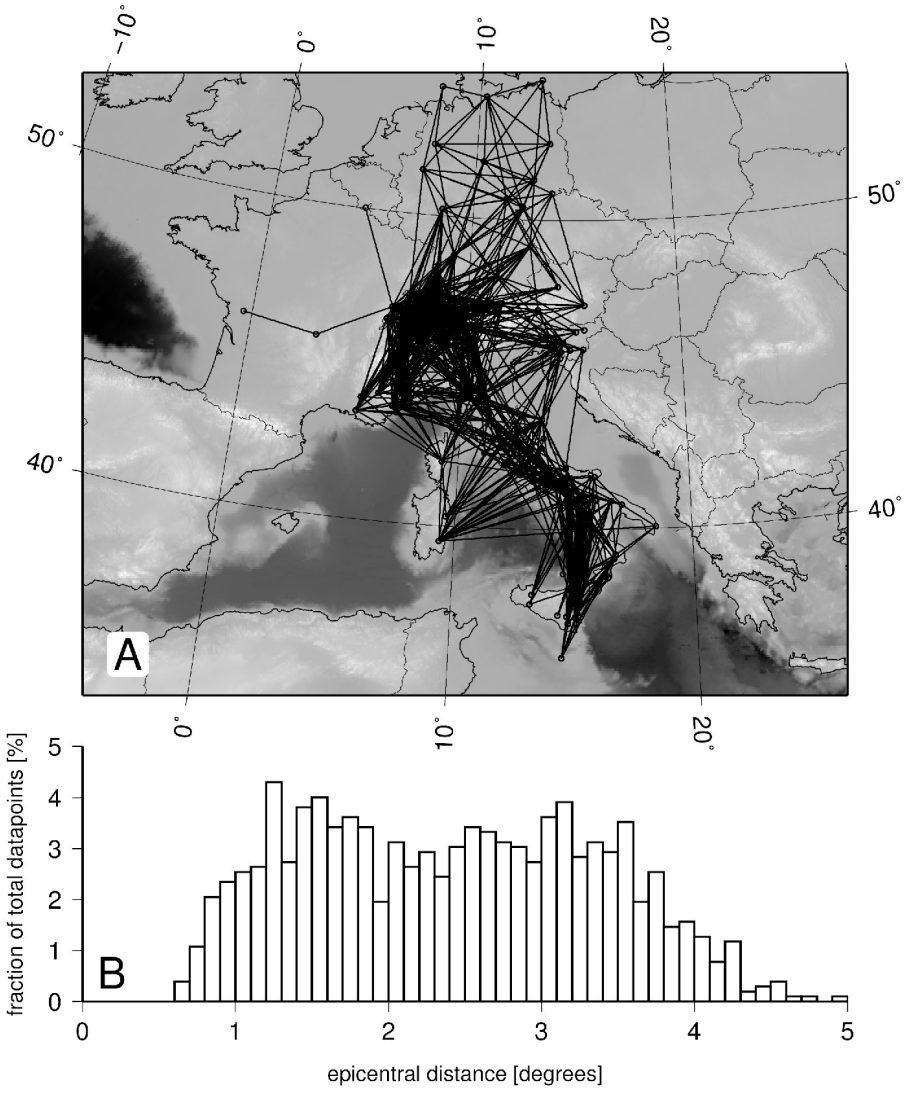

Figure 5: (A) Subset of European stations (circles) from *Verbeke et al.* [2012a] that are also included in our analysis. We only compare phase-velocity measurements associated with ∼1000 station pairs connected by solid lines. (B) Distribution of epicentral-distance values sampled by the data set at A.

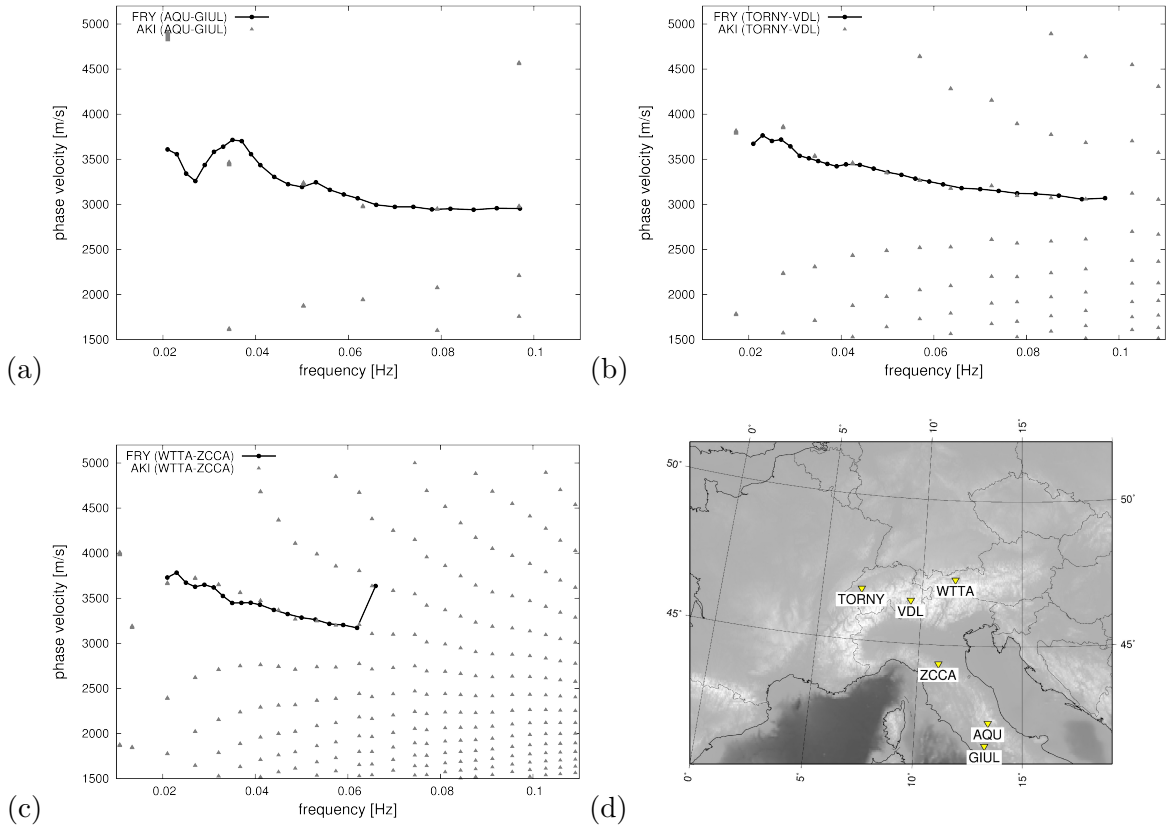

Figure 6: Selected FRY phase-velocity dispersion measurements (black circles, connected by a black line) compared with analogous frequency-domain (AKI) measurements (grey triangles), for three station pairs: (a) AQU and GIUL, in central Italy, only ∼90 km away from each other, with a North-South azimuth; (b) TORNY and VDL (see Fig. 3a), with interstation distance of ∼190 km; (c) WTTA in western Austria and ZCCA in northern Italy, ∼330 km to the south. Triangles in panel (d) mark the locations of all six stations considered here. We have not yet implemented an algorithm for automatic selection of a preferred AKI dispersion curve, but the FRY curves clearly fit a single branch of AKI datapoints. At low frequencies, and particularly at shorter epicentral distances, the match is less accurate. At longer epicentral distances and high frequencies, occasional one-cycle jumps as in (c) occur.

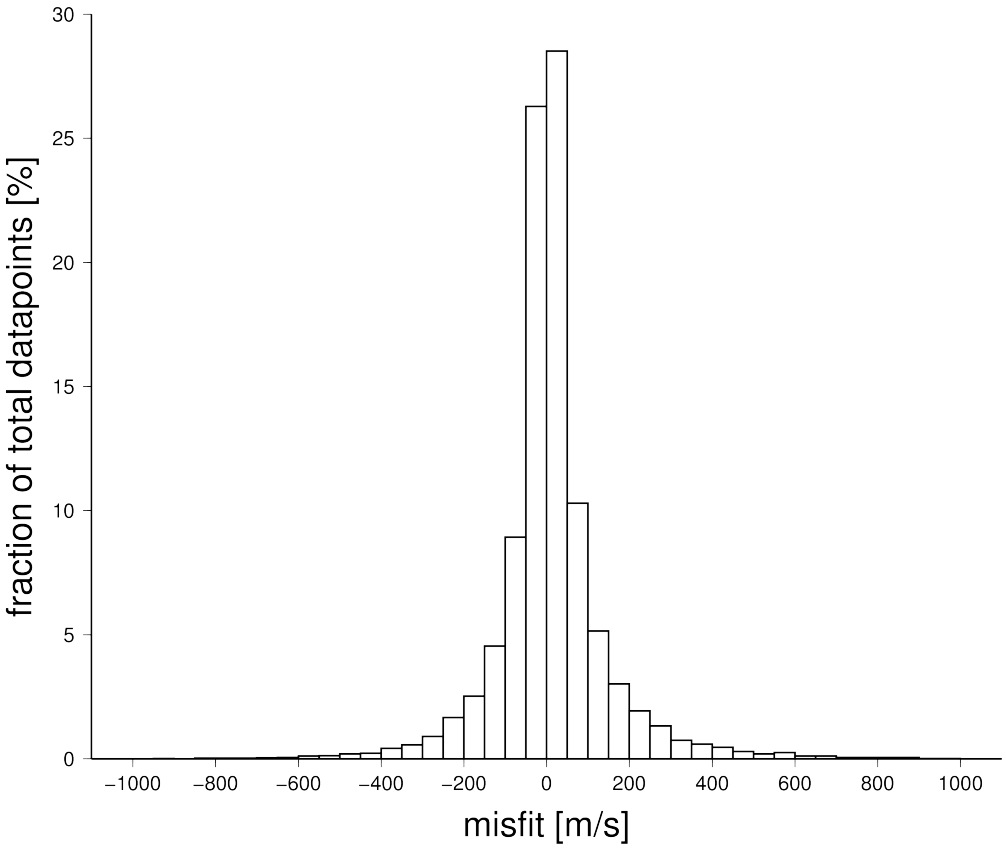

Figure 7: Frequency of observed phase-velocity misfit (AKI values subtracted from FRY ones) for the total set of ∼1000 analyzed station pairs. The mean is 13 m/s and the standard deviation is 151 m/s.

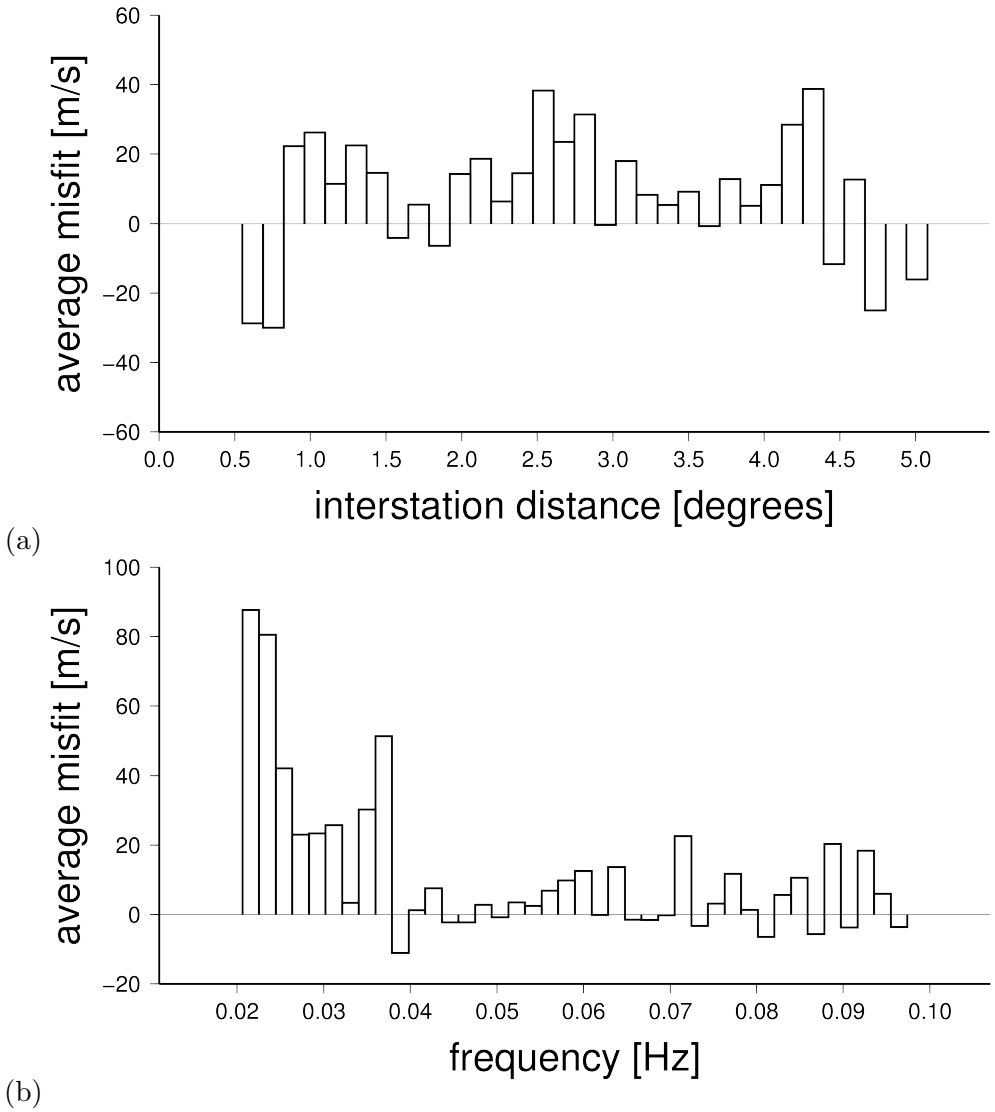

Figure 8: FRY-AKI phase-velocity misfit, for the total set of ∼1000 analyzed station pairs, averaged within (a) ∼0.3° interstation-distance bins, and (b) 2-mHz frequency bins.

a second histogram (Fig. 8a) where the misfit is averaged within ∼0.3° interstation-distance bins. In Fig. 8b the misfit is likewise averaged within 2-mHz increments spanning the whole frequency range of interest. Fig. 8a shows that FRY has a tendency to give slightly higher velocity estimates with respect to AKI; this effect is reversed at very small and very large interstation distances. The misfit remains low (∼30 m/s or less) at most interstation distances.

Fig. 8b shows clearly that misfit is systematically smaller ($\lesssim$20 m/s) at relatively high frequencies ($\gtrsim$0.04 Hz) than it is at low frequencies of ∼0.02-0.03 Hz. This is expected, as low frequency might result in relatively small $\omega\Delta x/v$, which would deteriorate the performance of FRY (but not of AKI) for short interstation distance Δx: in practice, the causal and anticausal parts tend to overlap in the short-Δx time-domain cross-correlations, making it difficult to measure phase via the FRY method [e.g., *Ekström et al.*, 2009].

The combined effect of short Δx and low frequency is perhaps better illustrated in Fig. 9a, where both frequency- and Δx-dependence of misfit are shown in a single, 2-D plot. It

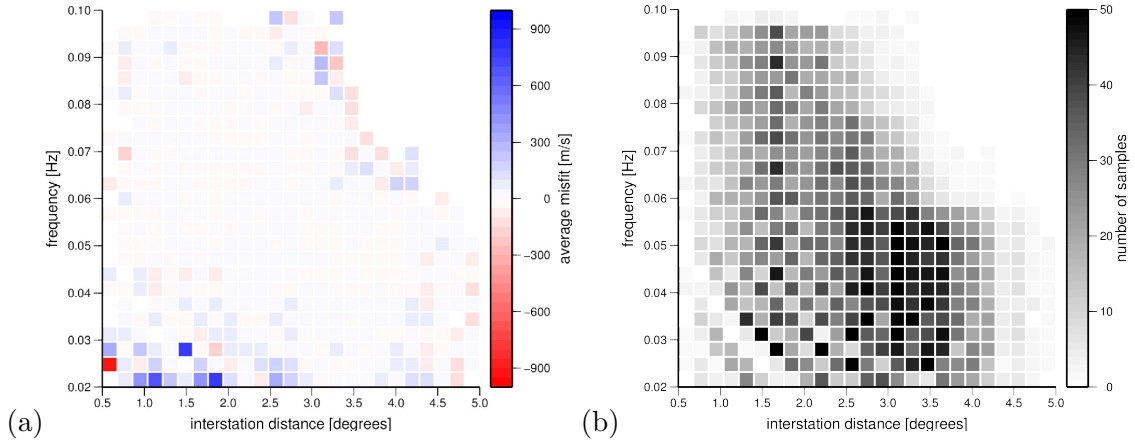

Figure 9: (a) FRY-AKI phase-velocity misfit, for the total set of ∼1000 analyzed station pairs, averaged within (a) ∼0.2° × 0.04-Hz distance/frequency bins; (b) number of pairs per distance/frequency bin.

emerges that, even at low frequency, AKI and FRY are in good agreement for sufficiently large interstation distance. Fig. 9b shows that, not surprisingly, sampling is not uniform with respect to frequency and Δx; most seismic-ambient-noise energy in our station array is found at frequencies around ∼0.05 Hz, and some of the discrepancy found at both higher and lower frequency (see in particular the top right of Fig. 8a) presubmably reflects the difficulty of finding coherent signal in the absence of a sufficiently strong ambient wavefield.

Overall, averaged discrepancies in Figs. 8 and 9 remain $\lesssim$50 m/s, with the exception of the lowest frequencies/shortest epicentral distances considered, where averaged values can exceed ∼100 m/s. A velocity difference of 50 m/s can be considered small if compared with the range of velocity heterogeneity in the frequency range and geographic area of interest, i.e. ∼1 km/s or more according to *Verbeke et al.* [2012b]. We take this as an indication that the AKI and FRY methods provide essentially consistent results, and we infer that such results can be considered reliable.

5 Conclusions

With this study we have conducted a detailed review of the theory of ensemble-averaged cross-correlation of surface waves generated by seismic ambient noise, as more tersely described by *Tsai* [2009], *Tsai and Moschetti* [2010] and *Tsai* [2011]. With our rederivation we attempt to focus the reader's attention on the potential discrepancy between the time-domain and frequency-domain approaches in phase-velocity measurements conducted on ambient-noise surface waves. The possibly most important difference between the two methods resides in the far-field approximation that is generally applied by time-domain practitioners [e.g., *Lin et al.*, 2008; *Yao and van der Hilst*, 2009; *Fry et al.*, 2010; *Verbeke et al.*, 2012b], and we have emphasized how this approximation is inadequate for interstation distances comparable to the seismic wavelengths. The frequency-domain approach of *Aki* [1957] and *Ekström et al.* [2009] does not suffer from this limitation: it is thus particularly useful for closely-spaced stations, provided that precursory noise caused by inhomogeneities in the source distribution is negligible [e.g., *Shapiro et al.*, 2006; *Villasenor et al.*, 2007; *Lin et al.*, 2007; *Zheng et al.*,

2011].

We have employed our own implementations of the non-asymptotic frequency-domain (AKI) and far-field asymptotic time-domain (FRY) approaches, to measure Rayleigh-wave phase dispersion from a year of seismic noise recorded at a dense array of European stations [*Verbeke et al.*, 2012b]. The two approaches provide overall consistent results. As shown in Fig. 9, discrepancies are limited to the lowest frequencies and shortest epicentral distances, where the far-field approximation on which the FRY method relies does not hold. We infer that Rayleigh-wave phase velocity can be successfully observed, via ensemble averaging, from continuous recordings of seismic ambient noise, at least within the frequency ($\sim$0.03-0.1 Hz) and inter-station distance ($\sim$0.5°-5°) ranges analyzed here. We further confirm the validity of published phase-velocity observations [e.g., *Verbeke et al.*, 2012b] obtained through the time-domain approach.

6 Acknowledgments

This study benefitted from our interactions with Michel Campillo, Bill Fry, Edi Kissling, Laurent Stehly, Victor Tsai, and Yang Zha. We are thankful to Eiichi Fukuyama, Mike Ritzwoller and one anonymous reviewer for their insightful comments. A. Z. wishes to thank Daniele Spallarossa for his advice and constant support.

References

Abramowitz, M., and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, National Bureau of Standards Applied Mathematics Series, 1964.

Aki, K., Space and time spectra of stationary waves with special reference to microtremors, *Bull. Earthquake Res. Inst. Univ. Tokyo*, *35*, 415–456, 1957.

Basini, P., T. Nissen-Meyer, L. Boschi, E. Casarotti, J. Verbeke, O. Schenk, and D. Giardini, Adjoint ambient-noise tomography of the European lithosphere, *submitted to Geochemistry, Geophysics, Geosystems*, 2012.

Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.*, *169*, 1239–1260, DOI:10.1111/j.1365–246X.2007.03,374.x, 2007.

Boschi, L., and G. Ekström, New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies, *J. Geophys. Res.*, *107*, doi:10.1029/2000JB000,059, 2002.

Cox, H., Line array performance when signal coherence is spatially dependent, *J. Acoust. Soc. Am.*, *54*, 1743–1746, DOI:10.1121/1.1914,473, 1973.

Derode, A., E. Larose, M. Campillo, and M. Fink, How to estimate the Green's function of a heterogeneous medium between two passive sensors? Application to acoustic waves, *Appl. Phys. Lett.*, *83*, 3054–3056, DOI:10.1063/1.1617,373, 2003.

Dziewonski, A. M., and D. L. Anderson, Preliminary reference earth model, *Phys. Earth Planet. Int.*, *25*, 297–356, 1981.

Ekström, G., G. A. Abers, and S. C. Webb, Determination of surface-wave phase velocities across USArray from noise and Aki's spectral formulation, *Geophys. Res. Lett.*, *36*, doi:10.1029/2009GL039,131, 2009.

Fry, B., F. Deschamps, E. Kissling, L. Stehly, and D. Giardini, Layered azimuthal anisotropy of Rayleigh wave phase velocities in the European Alpine lithosphere inferred from ambient noise, *Earth Planet. Sci. Lett.*, *297*, 95–102, DOI:10.1016/j.epsl.2010.06.008, 2010.

Gouedard, P., H. Yao, F. Ernst, and R. D. van der Hilst, Surface-wave eikonal tomography for dense geophysical arrays, *Geophys. J. Int.*, *34*, in press, 2012.

Harmon, N., P. Gerstoft, C. A. Rychert, G. A. Abers, M. S. de la Cruz, and K. M. Fischer, Phase velocities from seismic noise using beamforming and cross correlation in Costa Rica and Nicaragua, *Geophys. Res. Lett.*, *35*, L19,303, doi:10.1029/2008GL035,387, 2008.

Lin, F.-C., and M. H. Ritzwoller, Helmholtz surface wave tomography for isotropic and azimuthally anisotropic structure, *Geophys. J. Int.*, *186*, 1104–1120, doi:10.1111/j.1365–246X.2011.05,070.x, 2011.

Lin, F.-C., M. H. Ritzwoller, J. Townend, S. Bannister, and M. K. Savage, Ambient noise Rayleigh wave tomography of new Zealand, *Geophys. J. Int.*, *170*, 649–666,doi:10.1111/j.1365–246X.2007.03,414.x, 2007.

Lin, F.-C., M. P. Moschetti, and M. H. Ritzwoller, Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophys. J. Int.*, *173*, 281–298, DOI:10.1111/j.1365–246X.2008.03,720.x, 2008.

Meier, T., K. Dietrich, B. Stockhert, and H. Harjes, One-dimensional models of shear wave velocity for the eastern Mediterranean obtained from the inversion of Rayleigh-wave phase velocities and tectonic implications, *Geophys. J. Int.*, *156*, 45–58, DOI:10.1111/j.1365–246X.2004.02,121.x, 2004.

Mulargia, F., The seismic noise wavefield is not diffuse, *J. Acoust. Soc. Am.*, *131*, 2853–2858, doi:10.1121/1.3689,551, 2012.

Nakahara, H., A systematic study of theoretical relations between spatial correlation and Green's function in one-, two- and three-dimensional random scalar wavefields, *Geophys. J. Int.*, *167*, 1097–1105, 2006.

Nishida, K., H. Kawakatsu, and S. Obara, Three-dimensional crustal S wave velocity structure in Japan using microseismic data recorded by Hi-net tiltmeters, *J. Geophys. Res.*, *113*, B10,302, DOI:10.1029/2007JB005,395, 2008.

Paige, C. C., and M. A. Saunders, LSQR - an algorithm for sparse linear-equations and sparse least-squares, *ACM Transactions On Math. Software*, *8*, 43–71, 1982.

Peter, D., C. Tape, L. Boschi, and J. H. Woodhouse, Surface wave tomography: global membrane waves and adjoint methods, *Geophys. J. Int.*, *171*, 1098–1117, 2007.

Ritzwoller, N. M., M. H. amd Shapiro, A. L. Levshin, and L. G. M., Crustal and upper mantle structure beneath Antarctica and surrounding oceans, *J. Geophys. Res.*, *106*, 30,645–30,670, 2001.

Seats, K. J., J. F. Lawrence, and G. A. Prieto, Improved ambient noise correlation functions using Welch's method, *Geophys. J. Int.*, *188*, 513–523, doi:10.1111/j.1365–246X.2011.05,263.x, 2012.

Shapiro, N. M., and M. Campillo, Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.*, *31*, 2004.

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

Shapiro, N. M., M. H. Ritzwoller, and G. D. Bensen, Source location of the 26 sec microseism from cross-correlations of ambient seismic noise, *Geophys. Res. Lett.*, *33*, L18,310, doi:10.1029/2006GL027,010, 2006.

Snieder, R., Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase, *Phys. Rev. E*, *69*, doi:10.1103/PhysRevE.69.046,610, 2004.

Stehly, L., M. Campillo, and N. M. Shapiro, A study of the seismic noise from its long-range correlation properties, *J. Geophys. Res.*, *111*, 2006.

Stehly, L., B. Fry, M. Campillo, N. M. Shapiro, J. Guilbert, L. Boschi, and D. Giardini, Tomography of the Alpine region from observations of seismic ambient noise, *Geophys. J. Int.*, *178*, 338–350, 2009.

Tromp, J., Y. Luo, S. Hanasoge, and D. Peter, Noise cross-correlation sensitivity kernels, *Geophys. J. Int.*, *183*, 791–819,DOI:10.1111/j.1365–246X.2010.04,721.x, 2010.

Tsai, V. C., On establishing the accuracy of noise tomography travel-time measurements in a realistic medium, *Geophys. J. Int.*, *178*, 1555–1564, doi:10.1111/j.1365–246X.2009.04,239.x, 2009.

Tsai, V. C., Understanding the Amplitudes of Noise Correlation Measurements, *J. Geophys. Res.*, *116*, B09,311, doi:10.1029/2011JB008,483, 2011.

Tsai, V. C., and M. P. Moschetti, An explicit relationship between time-domain noise correlation and spatial autocorrelation (SPAC) results, *Geophys. J. Int.*, *182*, 454–460, DOI:10.1111/j.1365–246X.2010.04,633.x, 2010.

Udías, A., *Principles of Seismology*, Cambridge University Press, Cambridge, U.K., 1999.

Verbeke, J., L. Boschi, D. May, F. Deschamps, and E. Kissling, Anisotropic velocity structure of the crust and the upper mantle in central Europe from high resolution ambient noise database, *Geochemistry Geophysics Geosystems*, *to be submitted*, 2012a.

Verbeke, J., L. Boschi, L. Stehly, E. Kissling, and A. Michelini, High-resolution Rayleigh-wave velocity maps of central Europe from a dense ambient-noise data set, *Geophys. J. Int.*, *188*, 1173–1187, doi:10.1111/j.1365–246X.2011.05,308.x, 2012b.

Villasenor, A., Y. Yang, M. H. Ritzwoller, and J. Gallart, Ambient noise surface wave tomography of the Iberian Peninsula: Implications for shallow seismic structure, *Geophys. Res. Lett.*, *34*, L11,304, doi:10.1029/2007GL030,164, 2007.

Weemstra, C., L. Boschi, A. Goertz, and B. Artman, Constraining seismic attenuation from recordings of ambient noise, *Geophysics*, *submitted*, 2012.

Yang, Y., and M. H. Ritzwoller, Characteristics of ambient seismic noise as a source for surface wave tomography, *Geochemistry, Geophysics, Geosystems*, *9*, Q02,008, doi:10.1029/2007GC001,814, 2008.

Yao, H., and R. D. van der Hilst, Analysis of ambient noise energy distribution and phase velocity bias in ambient noise tomography, with application to SE Tibet, *Geophys. J. Int.*, *179*, 1113–1132, doi:10.1111/j.1365–246X.2009.04,329.x, 2009.

Yao, H., R. D. van der Hilst, and M. V. de Hoop, Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis - I. Phase velocity maps, *Geophys. J. Int.*, *166*, 732–744,DOI:10.1111/j.1365–246X.2006.03,028.x, 2006.

Yokoi, T., and S. Margaryan, Consistency of the spatial autocorrelation method with seismic interferometry and its consequence, *Geophysical Prospecting*, *56*, 435–451, 2008.

Zheng, Y., W. Shen, L. Zhou, Y. Yang, Z. Xie, and M. H. Ritzwoller, Crust and uppermost mantle beneath the North China Craton, northeastern China, and the Sea of Japan from ambient noise tomography, *J. Geophys. Res.*, *116*, B12,312, doi:10.1029/2011JB008,637, 2011.