Low-Frequency Ambient Noise Autocorrelations: Waveforms and Normal Modes

by M. Schimmel, E. Stutzmann, and S. Ventosa

ABSTRACT

Seismic interferometry by ambient noise autocorrelations is a special case of Green's function retrieval for single-station analysis. Although high-frequency noise autocorrelations are now used to extract the reflectivity beneath seismic stations, low-frequency autocorrelations are hardly applied. Here, we present the observation of the Earth orbiting surface waves from low-frequency noise autocorrelations which are used to extract normal-mode frequencies for the Hum. The performances of the classical and phase autocorrelations are analyzed using seismic data from GEOSCOPE station TAM in Algeria. Both approaches are independent and perform differently for data with large amplitude variability. We show that the phase autocorrelation can robustly extract Rayleigh waves and normal modes because it is not biased by large amplitude signals (e.g., earthquakes). This is convenient because no data preprocessing (data selection or amplitude clipping) is required as usually employed for the classical approaches. This implies that the phase correlation takes advantage of the full data set and waveform information to achieve a high signal extraction convergence. Single-station phase autocorrelations may become an important tool in planetary seismology where data are limited due to the expensive and difficult data acquisition and can consist of high-amplitude variability due to unknown conditions. The upcoming INSIGHT (Interior Exploration using Seismic Investigations, Geodesy and Heat Transport) Mars mission plans the deployment of one broadband seismometer and the successful measurement of normal-mode frequencies and surface-wave dispersion curves will constrain its reference structure. Although we present low-frequency autocorrelations, our findings remain valid for cross correlations, other applications, and other frequency bands.

Electronic Supplement: Amplitude spectra for the gravest modes of the 11 March 2011 Tohoku-Oki earthquake as recorded on GEOSCOPE station TAM in Algeria.

INTRODUCTION AND MOTIVATION

The whole Earth, as expected for finite bodies, resonates at discrete frequencies that are the elasto-gravitational modes or free oscillations of the Earth. The corresponding frequencies are functionals of the Earth structure, shape, and movement. The free oscillations are built up by the interference of propagating waves, which therefore relate to the modes by mode superposition (e.g., Woodhouse and Deuss, 2015). At the lowermost frequencies, these are mostly the Earth orbiting surface waves. These waves are excited by strong earthquakes or strong ambient noise sources.

The low-frequency noise sources cause continuous free background oscillations, called Hum, which are observed during the absence of strong earthquakes on land (Kobayashi and Nishida, 1998; Nawa *et al.*, 1998; Suda *et al.*, 1998; Tanimoto *et al.*, 1998) and since recently at the ocean bottom (Deen *et al.*, 2017). The more difficult to observe toroidal modes have also been reported (Kurrle and Widmer-Schnidrig, 2008; Nishida, 2014). The Hum is driven by the interaction of infragravity ocean waves with the sea floor (e.g., Rhie and Romanowicz, 2004; Webb, 2007; Ardhuin *et al.*, 2015; Nishida, 2017). Fundamental normal modes and corresponding surface-wave waveforms can be extracted from noise cross correlations (CCs; Ventosa *et al.*, 2017). Rayleigh waves in the Hum frequency band have been used in global ambient noise tomography studies (Nishida *et al.*, 2009; Haned *et al.*, 2016).

The signal extraction from ambient noise at any frequency band follows seismic interferometry principles (Lobkis and Weaver, 2001; Derode *et al.*, 2003; Shapiro and Campillo, 2004; Snieder, 2004; Wapenaar, 2004; Roux *et al.*, 2005; Curtis *et al.*, 2006; Sens-Schönfelder and Wegler, 2006; among others) and is based on the CC of noise recorded at two stations. Ideally, with the CC one retrieves the empirical Green's functions (EGFs) for equipartitioned wavefields and canceled noise cross terms. Equipartition is not warranted because noise sources and wavefield scattering are not random. Hence, the CCs are averaged over long time to assure a balanced source coverage and to decrease the impact of correlation cross terms (Medeiros *et al.*, 2015).

In the ideal case, the EGF is the impulse response recorded at one of the receivers with the other receiver being the virtual source and consisting of body waves and the Earth orbiting surface waves in the Hum frequency band. New data processing approaches based on analytic signal theory aid small-amplitude

doi: 10.1785/0220180027

signal extraction and now permit the efficient use of body waves and the Earth orbiting Rayleigh waves with less data and less computational effort (Schimmel *et al.*, 2011; Haned *et al.*, 2016; Ventosa *et al.*, 2017). Latter work uses the classical CC and focuses on the stacking approach through wavelet theory.

Here, we also work in the Hum frequency band but focus on different autocorrelations to perform single-station noise analyses. Autocorrelations are related to the spectral density of power or energy (e.g., Bornmann and Wielandt, 2013) through their Fourier transform and are therefore employed, implicitly or explicitly, when determining modes' strengths and frequencies from the spectral density. In seismic interferometry, noise autocorrelations have been used at high-frequency bands to extract the *P*-wave reflection response for the structure below individual stations (e.g., Claerbout, 1968; Tibuleac and von Seggern, 2012; Becker and Knapmeyer-Endrun, 2017). Ekström (2001) used a modified autocorrelation through computing the CC between the original and reverse-dispersed seismogram to construct a Rayleigh-wave detection algorithm with which he finds the Earth orbiting waves in the very low-frequency background noise. For low-frequency autocorrelations and orbiting surface waves, the station is at a caustic and the stationary phase region is wide (Snieder and Sens-Schönfelder, 2015), that is, sources for all azimuths contribute to extract orbiting waves with seismic interferometry.

Here, we analyze pure autocorrelations, that is, without employing a model and reverse-dispersion operators and demonstrate observations of the Earth orbiting Rayleigh waves extracted from noise autocorrelations at Hum frequencies. Different autocorrelation techniques are compared, and it is shown that the phase cross correlation (PCC) by Schimmel (1999) provides robust measurements which are not biased by energetic signals in the data. An interesting implication is that PCC can provide results for data acquired under difficult circumstances with a high-amplitude variability of signals and noise. This is especially interesting when working with a limited amount of data that would reduce to an insufficient database after removal of outlying amplitude segments. PCC may therefore become an important tool for the analysis of seismic data from future planetary missions. For instance, the deployment of a single broadband seismometer is planned for Mars with first data arriving on the Earth by the end of 2018 through the INSIGHT (Interior Exploration using Seismic Investigations, Geodesy and Heat Transport) mission (Banerdt *et al.*, 2013; Lognonné and Pike, 2015). Other future planetary missions, for example, to Venus and Jupiter moon Europa, may follow and revealing their interior structure may benefit from robust single-station approaches. In any case, here we are working with the Earth's Hum. That is, the low-frequency seismic background noise on extra-terrestrial planets and bodies, if present, has likely a different generation mechanism, especially in the absence of oceans. Although this work focuses on the low-frequency band, our findings are also valid at other frequencies and for CCs.

LOW-FREQUENCY AUTOCORRELATION

At very low frequencies ($f \leq 5$ mHz), the seismic motion is better represented by modes than by traveling waves. which justifies the frequency decomposition of seismograms to identify free oscillations and to measure their frequencies (and attenuation rates). With increasing frequencies, the number of normal modes becomes quickly too high and the resonant peaks are distorted mainly by the coupling of singlets due to heterogeneities. Consequently, it becomes more practical to analyze the waveforms of propagating waves, mainly surface waves and their dispersion characteristics. Both modes and waveforms are important to constrain the Earth structure and can be obtained with a single station through autocorrelations.

At very low frequencies, the identification of modes and the determination of their frequencies are based on Fourier analysis to decompose seismograms of single-station data. For this purpose, usually the power spectral density (PSD) and/or energy spectral density (ESD) are computed. The PSD and ESD can be obtained through the Fourier transform of differently defined autocorrelations, depending on whether one is dealing with infinite or finite length waveforms. The PSD has the advantage that it can be employed when the time series cannot be directly Fourier transformed, strictly due to infinite signal energy as happens for stationary (infinite duration) waveforms. The ESD adequately describes finite duration (transient) waveforms and equals the squared amplitude spectrum of the seismogram.

Different Autocorrelation Approaches

The autocorrelation measures the similarity between a time series and a delayed version of itself. It is the special case of the CC (equation 1) where both time series are the same, that is, $s_1(t) = s_2(t)$. Here, real-valued time series are considered, and the following equations are written for discrete time t and time delay or lag time τ:

$$c_{cc}(\tau) = \sum_{t=1}^{T} s_1(t) s_2(t+\tau). \tag{1}$$

If $s_1(t)$ is modulus square integrable, then the Fourier transform of the autocorrelation (as defined in equation 1 with $s_1(t) = s_2(t)$) is the ESD following the Wiener–Khinchin theorem (e.g., Alessio, 2015). The ESD describes the energy distribution of the time series with respect to frequency and equals the squared Fourier amplitude spectrum of $s_1(t)$. In the following, we use CC to indicate that a result is based on the correlation of equation (1).

CCs are often employed as similarity measures, for instance, to detect signals through their waveform similarity and are therefore normalized by the energy of the used sequences to limit the CC values between −1 and 1. Frequently used is the normalization by the geometric mean energy (denominator of equation 2). This geometrically normalized cross correlation (CCGN) reduces the energy sensitivity of the numerator and makes the correlation invariant to amplitude changes between the two data sets:

$$c_{\mathrm{CCGN}}(\tau) = \frac{\sum_{t=1}^{T} s_1(t)s_2(t+\tau)}{\sqrt{\sum_{t=1}^{T} s_1^2(t)}\sqrt{\sum_{t=1}^{T} s_2^2(t+\tau)}}. \quad (2)$$

With $s_1(t) = s_2(t)$, equation (2) measures also the similarity between a time series and a delayed version of itself. The autocorrelation based on equation (2) and its Fourier transform are therefore equally suited to detect multiple Rayleigh wavetrains and normal modes. However, in contrast to equation (1), no energy is measured. Along the same line, PCC and its Fourier transform can be used to detect multiple Rayleigh wavetrains and to measure the normal-mode frequencies. PCC (equation 3) has been designed by Schimmel (1999) in full analogy to a normalized CC:

$$c_{\mathrm{PCC}}(\tau) = \frac{1}{2T}\sum_{t=1}^{T} |e^{i\Phi(t)} + e^{i\Psi(t+\tau)}|^{\nu} - |e^{i\Phi(t)} - e^{i\Psi(t+\tau)}|^{\nu}. \quad (3)$$

It uses the instantaneous phases ($\Phi(t)$ and $\Psi(t)$) of the analytic signal for the time series $s_1(t)$ and $s_2(t)$, respectively. The analytic signal of a real-valued time series is a unique complex-valued representation of the time series, where the real and complex parts are orthogonal. The analytic signal is constructed from the real-valued time series and its Hilbert transform. Thanks to this representation, a real time series can be decomposed into an instantaneous phase and instantaneous amplitude. PCC (equation 3) uses this property, that is, it uses only the instantaneous phases and is therefore explicitly amplitude unbiased. Inherent to the analytic signal theory, the instantaneous phases contain information of the neighboring samples and therefore characterize a waveform as function of time (e.g., Taner *et al.*, 1979; Schimmel *et al.*, 2011). The phase autocorrelation is obtained using $\Phi(t) = \Psi(t)$.

With PCC, waveform similarity is measured through the amount of phase-coherent samples rather than the sum of amplitude products. PCC and CCGN are fully independent concepts, based on different philosophies, but which in full analogy determine similarity as function of lag time. Inherent to their different designs, the three definitions of cross/autocorrelations (equations 1–3) have a different data performance depending on signal and noise characteristics.

Example Using Earthquake Data

We now use the three autocorrelation approaches and show their ability to detect normal modes from an earthquake recording. We are analyzing the spectral composition as obtained through a Fourier transform of the previously defined autocorrelations based on equations (1)–(3). Our goal is to show that these different approaches can be used for mode detection after the occurrence of a strong earthquake.

Data Processing

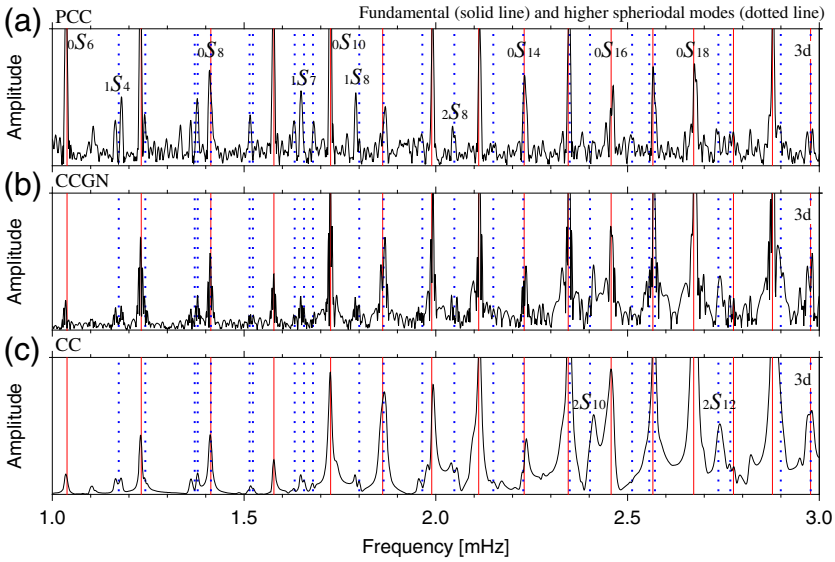
We use a vertical-component record of the Tohoku-Oki main earthquake (off coast of east Japan, moment magnitude 9.0, 11 March 2011) and its aftershocks recorded by the GEOSCOPE station TAM (Tamanrasset, Algeria). This station has been selected because it is installed at a calm place in the middle of the African continent far from oceans, working continuously for many years, and because its noise level is low at long periods. Our time series starts at midnight and is three days long. It has been converted to ground acceleration, decimated to 10 s sampling interval and zero-phase bandpassed with frequency plateau between 0.15 and 5 mHz.

Three autocorrelation functions have been computed using PCC, CCGN, and CC. Because the autocorrelations are symmetric with respect to zero lag time, we compute only the causal part to then construct the symmetric autocorrelation. The symmetric autocorrelations were Hanning tapered (width 0.5) and their spectra determined after zero padding of the time series to a power of 2 samples and applying a conventional fast Fourier transform algorithm. We plot the square root of the amplitude spectra. In the case of CC, it means that we determined the ESD, which is nonnegative for real-valued time series and where the ESD square root equals the Fourier amplitude spectrum of the seismogram.

Results

All spectra have been normalized for visual purposes and are shown in Figure 1. Solid and dotted lines mark fundamental and higher modes, respectively, as expected for the spherically symmetric preliminary reference Earth model (PREM) by Dziewonski and Anderson (1981). Mostly spheroidal modes S are expected on the vertical components. Toroidal modes may also appear on the vertical components due to Coriolis coupling between spheroidal and toroidal modes (Zürn *et al.*, 2000), but most of their energy is expected on the horizontal components, and therefore, in the following we only consider spheroidal modes. Spheroidal modes are denoted as ${}_nS_l^m$, in which n is the mode number and m and l label the azimuthal degree and order of the spherical harmonic functions. Each multiplet consists of $2l + 1$ singlets. For a spherical-symmetric nonrotating Earth model, the multiplets are degenerate, that is, have the same frequency. Index m, which marks each singlet has therefore been omitted. Some fundamental and higher modes have been marked in Figure 1a,c.

Although the same data have been used, the spectra for PCC (Fig. 1a), CCGN (Fig. 1b), and CC (Fig. 1c) are sufficiently different to catch attention. Some of the multiplets seem to be better detected by PCC (e.g., ${}_0S_6$, ${}_1S_4$, ${}_1S_7$, ${}_1S_8$, ${}_2S_8$) than by the other two approaches. Conversely, ${}_2S_{10}$ and ${}_2S_{12}$ are not detected with PCC while visible with CC. The differences are inherent to the different autocorrelation definitions. In particular, the PCC is amplitude unbiased and a large signal means that this mode is built up, thanks to its phase coherence and phase incoherence of other signals. The latter means that the PCC is cleaner due to the attenuation of phase incoherent signals that translates into the spectrum through the Fourier transform of PCC. The spectra with the gravest Earth modes ($f < 1$ mHz) are shown in Ⓔ Figure S1 (available in the electronic supplement to this article) for completeness, where some splitting of multiplets is visible. We did not further analyze singlets or earthquake data, which is out of the scope of this article.



▲ **Figure 1.** Amplitude spectra for the 11 March 2011 Tohoku-Oki earthquake as recorded at the GEOSCOPE station TAM in Algeria. Seismic recordings had a length of three days and spectra were obtained through the Fourier transform of the autocorrelations computed following equations (1) (cross correlation [CC]), (2) (geometrically normalized cross correlation [CCGN]), and (3) (phase cross correlation [PCC]). Plotted are the normalized amplitudes after taking the square root of the spectra amplitudes. Solid and dotted vertical lines mark the fundamental and higher spheroidal modes following preliminary reference Earth model (PREM). The color version of this figure is available only in the electronic edition.

Ambient Noise

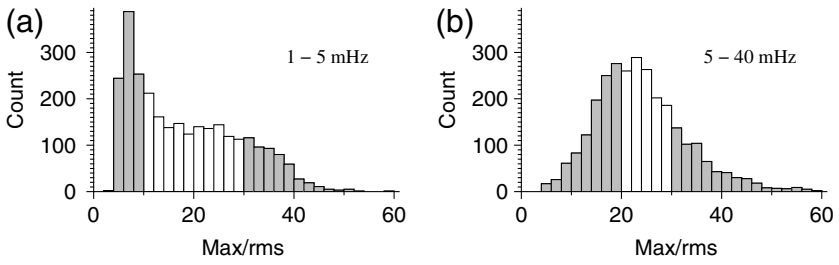
Here, our primary goal is to understand the performance of the three correlation methods on ambient noise data.

Main Data Processing

We use vertical-component seismograms for 8 yrs (2007–2014) of continuous recordings at the GEOSCOPE station TAM (Tamanrasset, Algeria). The data have been cut into three-day-long segments separated by one day, instrument corrected to ground acceleration, decimated to 10 s sampling interval, and zero-phase bandpassed with frequency plateau between 0.15 and 10 mHz. The three autocorrelations (CC, CCGN, and PCC) have been computed for each data segment with a maximum lag time of one day. No further preprocessing is performed for PCC, CC, and CCGN, but we also define CCGNN as follows. The band-passed seismograms have been amplitude balanced using the 1-bit time-domain normalization and spectral whitening and again bandpassed to then compute autocorrelations based on equation (2) which we label as CCGNN. Time- and frequency-domain normalizations are often used to reduce bias in the correlations from large amplitude signals and noise (e.g., Bensen *et al.*, 2007; Schimmel *et al.*, 2011). Alternative strategies are to identify data segments with larger amplitudes and to exclude them from the subsequent processing or to clip the amplitudes above a certain threshold.

The individual autocorrelograms are then linearly stacked for each of the employed approaches. The averaging is a common procedure to attenuate noise and signal variability and to enhance robust signals in the autocorrelations. No nonlinear stacking is shown here, as we are interested to focus on the performance of the correlation approaches and stacking can only enhance signals which are present in the correlations. Furthermore, the spectra of the stacked autocorrelograms are computed as described in the example with the earthquake data.

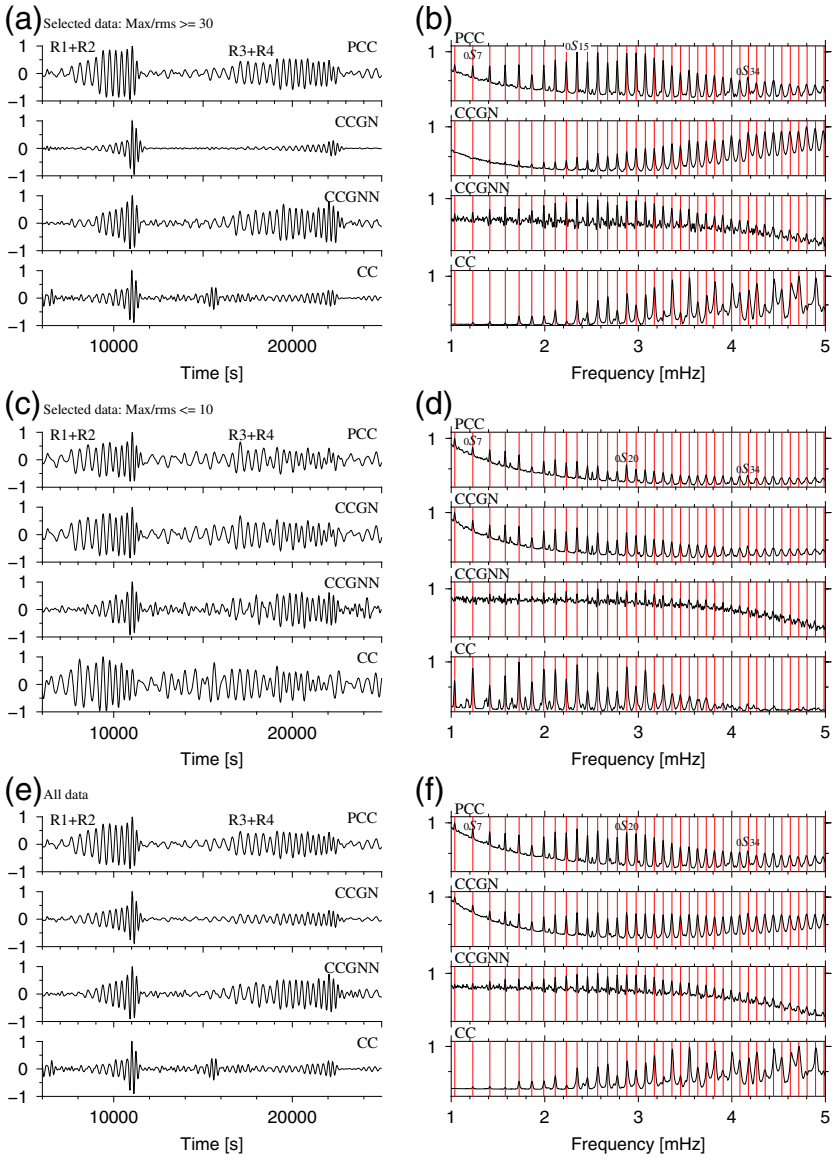
For the following analyses, we also determine the amplitude variability in our data. To that end, we compute for each data segment the ratio of the absolute maximum amplitude to the root mean square amplitude (Max/rms) in the 1–5 and 5–40 mHz frequency band. The corresponding Max/rms distribution of our data is shown in Figure 2. Data with a low Max/rms ratio should contain no strong earthquakes or other energetic signals that manifest through a large Max/rms ratio. The gray areas of Figure 2a mark the subsidiary data sets with Max/rms ≤ 10 and Max/rms ≥ 30 that are used to obtain the results of Figure 3a–d. In analogy, the gray areas of Figure 2b mark the subsidiary data sets with Max/rms ≤ 20 and Max/rms ≥ 30 for Figure 5a,b. For the considered frequencies, an event catalog can be used alternatively to select calm periods (Tanimoto and Um, 1999), nevertheless, we prefer to use the Max/rms ratio to also detect other outlying noise features.



▲ **Figure 2.** Distribution of the maximum amplitude to root mean square (rms) amplitude ratio (Max/rms) determined for the (a) three-day and (b) 30-hr-long seismogram sequences after band-passing the data in frequency band (a) 1–5 mHz and (b) 5–40 mHz. Gray areas mark subsidiary data sets with Max/rms ≤ 10 and Max/rms ≥ 30 in (a) and Max/rms ≤ 20 and Max/rms ≥ 30 in (b) as used for Figures 3a–d and 5a,b, respectively.

Results

Max/rms, determined from the 1–5 mHz bandpassed data, is now being used as selection criterion to stack only autocorrelations with corresponding Max/rms ≥ 30. The waveforms and spectra obtained from the autocorrelation



▲ **Figure 3.** (a,c,e) Waveforms and (b,d,f) amplitude spectra for the differently computed autocorrelations. (a,b) Based on the Max/rms ≤ 30 data set; (c,d) based on the Max/rms ≥ 10 data; and (e,f) based on the entire database. Solid vertical lines mark the fundamental modes following PREM. The color version of this figure is available only in the electronic edition.

stacks are shown in Figure 3a,b. The waveforms were filtered between 1.5 and 5 mHz and are shown for a zoom from 5.500 to 27.000 s lag time.

The extracted waveforms show two wavetrains which are multiple orbit Rayleigh waves. For earthquakes sources, Rayleigh waves that travel from the source on the minor and major arc path to the receiver are labeled R1 and R2, respectively. Accordingly, R3 and R4 are R1 and R2 waves that traveled in addition one entire Earth orbit. Because we are using autocorrelations, virtual source and receiver locations are the same and minor/major arc wavetrains merge to entire orbit wavetrains. We therefore mark with R1 + R2 and R3 + R4, one and two orbit Rayleigh wavetrains. It can be seen that the R3 + R4 wavetrain is much longer than the R1 + R2 wavetrain, which is inherent to the extra orbit and Rayleigh-wave dispersion. The dispersive character of the wavetrain is seen with the lowest frequency components traveling faster than the higher frequency waves.

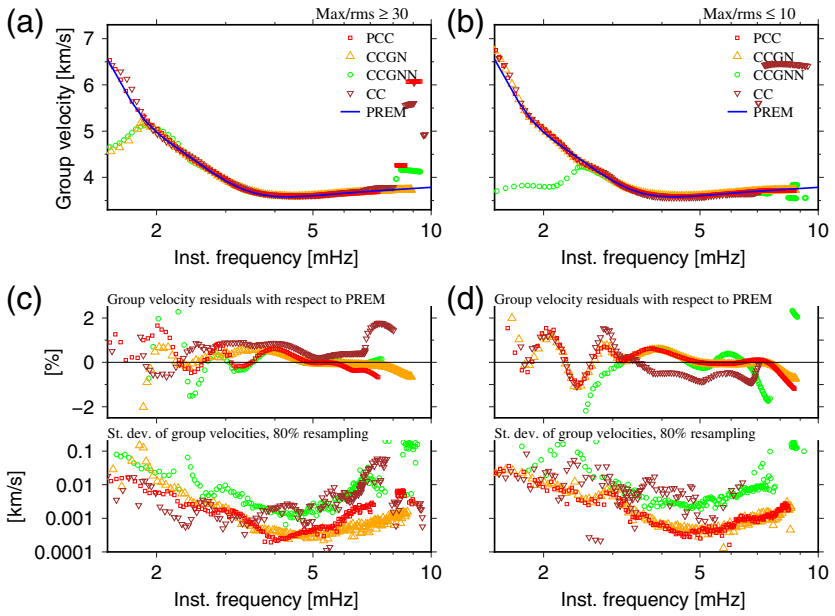
It is further observed that the different methods provide different waveforms in which CCGNN resembles most the PCC waveform. The differences between the different waveforms are due to the amplitude bias of the autocorrelation approaches. CCGNN uses time- and frequency-domain normalized data and resembles therefore most to the amplitude unbiased PCC.

The corresponding spectra are shown in Figure 3b. As expected from the waveforms (Fig. 3a), the lower frequency modes (e.g., ${}_0S_7$) are less represented in the spectra for CCGN and CC than in the spectra for PCC and CCGNN, where we can appreciate the detection of more modes.

Figure 3c,d shows in full analogy to Figure 3a,b the results for data with Max/rms ≤ 10. These data are expected to contain less amplitude variability due to the decreased amplitude ratio. Data sequences that contain strong earthquakes or other outlying signals and noise are therefore not included in this data set. Indeed, we now observe that CCGN resembles mostly to the PCC waveform. Also their spectra are more similar. The CCGN and CC waveform show now more dispersion than in Figure 3a, but R3 + R4 is hidden in the noise. CCGNN did not improve much because more data are needed for the EGF construction due to the waveform manipulation using the 1-bit normalization and spectral whitening. Indeed, both operations are nonunique and there exists an infinite number of waveforms that reduce to the same 1-bit and/or spectral whitened representation of the data, which decreases the signal extraction convergence.

Waveforms and spectra for the entire data set are shown in Figure 3e,f. The PCC and CCGNN results are similar to their respective results, obtained with the two subsidiary data sets. It further shows the invariance to the amplitude differences of these approaches and that these waveforms converged to the same result. Conversely, the CC waveform and spectrum resemble only to the results obtained with the Max/rms ≥ 30 data (Fig. 3a,b). This is expected due to the strong energy bias of CC. CCGN is less energy biased due to the normalization by the geometric mean energy and as a consequence the waveform and spectrum obtained with all data are less similar to the Max/rms ≥ 30 results. CC and CCGN waveforms did not converge to the same waveform for the three data sets.

In any case, Figure 3 shows that we can detect and extract propagating waves at the lowermost frequencies, that is, these frequencies are lower than the frequencies usually used in global Hum tomography studies (e.g., Nishida *et al.*, 2009; Haned *et al.*, 2016). These waveforms can be used to extract group velocities (e.g., Schimmel *et al.*, 2017) to further constrain the seismic structure. Figure 4a,b shows the group velocities for the two subsidiary data sets (Max/rms ≥ 30 and Max/rms ≤ 10) and the R1 + R2 waveform. The group velocities were obtained from the time–frequency representation of the waveforms shown in Figure 3 using 40,030.174 km for the propagation path and plotting the strongest amplitude maximum per frequency as group arrival. The solid line marks the theoretical group velocity curve for PREM, whereas the symbols correspond to the measured group velocities. It can be seen that PCC (squares) measures correctly the velocities down to the lowest frequencies, even for the data with large Max/rms (Fig. 4a). For large Max/rms, CCGNN (circles) and CCGN (triangles) fail to provide expected group velocities at frequencies below 2 mHz. Same happens for CCGNN for low Max/rms data (Fig. 4b). This is due to the spectral whitening and 1-bit normalization that deteriorate the waveforms and therefore decrease the detectability of the weak-amplitude signals, especially at the lowest frequencies. The other methods provide group velocity measurements close to the expected values and supposed for the data set without outlying amplitude events.



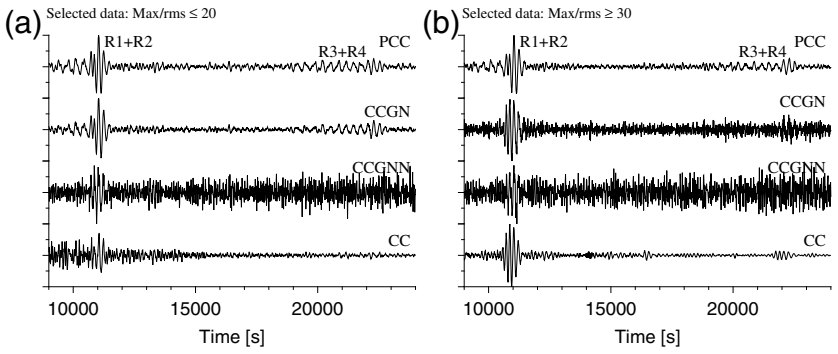
▲ **Figure 4.** (a,b) Group velocities extracted for the waveforms shown in Figure 3a, c using the Max/rms ≥ 30 (left) and Max/rms ≤ 10 (right) data sets. The solid line corresponds to the fundamental-mode group velocities expected for the spherical symmetric Earth model PREM. (c,d) The corresponding group velocity residuals with respect to the expected values for PREM (upper panels) and the standard deviations of measured group velocities for randomly resampled subsidiary data sets (lower panels). The subsidiary data sets use 80% of the available autocorrelations for corresponding Max/rms data. The color version of this figure is available only in the electronic edition.

The group velocity residuals with respect to PREM (top panels of Fig. 4c,d) are typically less than 1%. At frequencies below 2 mHz, PCC (squares) shows similar frequency-dependent deviations for Max/rms ≥ 30 (Fig. 4c) and Max/rms ≤ 10 (Fig. 4d). This trend is also seen for CCGN and CCGNN when Max/rms ≤ 10 (Fig. 4d). The lowermost panels of Figure 4c,d show the standard deviations of the measured group velocities for randomly down sampled subsidiary data sets. The resampled data sets use 80% of the autocorrelations. The lowest standard deviations are seen for frequencies of about 3–7 mHz. This means the dispersion curves at these frequencies are more robust with respect to data variability than at the frequencies outside this band, which suggests a better waveform convergence at 3–7 mHz. Further, the standard deviations for CCGNN (circles) for Max/rms ≥ 30 and Max/rms ≤ 10 show a similar frequency dependence as for PCC (squares) and CCGN (triangles), but displaced to higher values. We interpret this higher overall variability and standard deviations as an inferior EGF waveform convergence at all frequencies due to the additional preprocessing for CCGNN.

The waveforms have also been compared in a higher frequency band (5–40 mHz), to which end we reprocessed the continuous data as described before but with following differences: cutting the data into 30 hrs segments, decimating to 5 s sample interval, and band-passing with flat plateau from 5 to 40 mHz. The Max/rms distribution is shown in Figure 2b and the waveforms obtained with Max/rms ≤ 20 and ≥ 30 are shown in Figure 5a and 5b, respectively. As expected, PCC and CCGN resemble each other for low Max/rms (Fig. 5a). Further, PCC shows almost no waveform variations for the two different data sets. Only with CC, the R1 + R2 and R3 + R4 waveforms improve when there is more amplitude variability in the data (Fig. 5b). We believe this is due to energetic signals, for example, as caused by strong earthquakes, which are included in the data set for Max/rms ≥ 30. On contrary, CCGNN is less affected by energetic signals due to the applied amplitude balance (1-bit, spectral whitening) on each seismogram. This explains why the final CCGNN waveforms resemble each other for the two different data sets. Its signal-to-noise ratio (SNR) is expected to improve for more data. The waveform manipulation by the amplitude normalizations difficult the signal extraction and consequently slows down the convergence to a cleaner EGF.

DISCUSSION

We extracted spheroidal modes and Rayleigh waveforms from single-station low-frequency noise recordings using different



▲ **Figure 5.** Waveforms for the differently computed autocorrelations in the 5–40 mHz frequency band. (a,b) Based on the Max/rms ≤ 20 and Max/rms ≥ 30 data, respectively. The Max/rms distribution is shown in Figure 2b.

autocorrelation approaches. Although the high-frequency noise autocorrelations provide the zero-offset reflection response at the recording station (e.g., Claerbout, 1968; Wapenaar *et al.*, 2004; Galetti and Curtis, 2012), the low-frequency autocorrelations contain mainly the Earth orbiting surface waves. This is why long-period noise autocorrelations can be used to measure surface-wave dispersion and the Earth eigen frequencies, which give constraints on large-scale velocity models. To that end, autocorrelations can be computed using PCC or the classical CC (e.g., CC, CCGN, and CCGNN). Both correlation approaches are built on different and independent principles which cause that their performances depend on signal and noise characteristics (Schimmel, 1999). Using independent correlation methods is an advantage since jointly used they permit to determine the consistency of measurements and to fill measurement gaps whenever one of the methods fails to see the signal (Bonatto *et al.*, 2015). A joint use of both approaches may aid the search of less-observed modes or propagating waves.

The determination of ESDs can further be sophisticated, for example, using multitaper methods (Thomson, 1982; Prieto *et al.*, 2007) to reduce spectral leakage. This or other processing methods can be applied similarly on any of the presented autocorrelations and has not been used to avoid masking the performance of the correlations, which is the backbone for any further analysis and understanding. In addition, the here presented strategy can be used to extract orbiting Love waves and toroidal modes.

PCC is explicitly amplitude unbiased and finds the signals through their instantaneous phase coherence, while the signal detection with the classical correlation is due to the constructive summation of amplitude products, which makes this correlation technique amplitude biased. Because PCC is amplitude unbiased, no data preprocessing is required to balance the amplitudes through time- and/or frequency-domain normalizations or amplitude clipping (Sabra *et al.*, 2005; Bensen *et al.*, 2007) to reduce the influence of energetic features as strong earthquakes, sensor failures, local episodic noise, and so on. Normalization approaches as spectral whitening, 1-bit, and clipping are commonly used and remove the uniqueness of waveforms because an infinite number of waveforms reduces to the same clipped signal. In comparison to correlating the full signal waveforms, usually longer time series are needed for signal extraction (e.g., D'Hour *et al.*, 2016). We believe this slower convergence is to compensate the lost information through the nonunique normalization, which increases ambiguities between signals and noise. In other words, signals and noise may share same or similar waveforms after nonunique normalizations, which hinder signal extractions and noise attenuation. In that spirit, the minimum preprocessing required with PCC is an advantage over the other methods.

The amplitude bias of CC and CCGN explains why the waveforms differ for our data sets with large or small Max/rms. Conversely, the amplitude variations in these data sets were neutralized for CCGNN and the corresponding final waveforms are therefore more similar. Nevertheless, more data are needed to increase the SNR to obtain cleaner waveforms as often presented with PCC. That is, using the same limited amount of data may lead with CCGNN to signals that are not fully extracted as shown in figures 11a,b Schimmel *et al.* (2011), figures 3 and 4 of DHour *et al.* (2016), and figure 1 of Haned *et al.* (2016) for different frequency bands and wave types. It also explains the increased standard deviation of group velocity measurements shown in Figure 4c,d (bottom panel).

Rejecting data segments with outlying amplitudes (Pedersen and Krüger, 2007) is an excellent alternative to the normalization but reduces the database. In monitoring studies, this may imply a decrease in time resolution because one needs to compensate for the rejected length of data. Because PCC is less sensitive to amplitude variability it might be a good candidate to analyze data acquired under difficult conditions and limited length. However, one needs to keep in mind that PCC fails whenever the signals cannot be detected through their phase coherence, that is, whenever signals are less phase coherent than the surrounding noise.

PCC has been presented here for single-station analyses to also show its potential for extraterrestrial seismology, where the amount of stations might be restricted to one due to the high costs of stations and deployment. For example, the upcoming Mars mission (INSIGHT, e.g., Banerdt *et al.*, 2013; Lognonné and Pike, 2015; Panning *et al.*, 2017) is planned to place a single broadband seismometer on Mars. The corresponding data might be limited and of unknown characteristics. In any case, little is known on the seismic structure and seismic activity of extraterrestrial planets and bodies, and besides revealing local structure through the reflectivity obtained with autocorrelations at high frequencies (e.g., Tibuleac and von Seggern, 2012; Taylor *et al.*, 2016; Becker and Knapmeyer-Endrun, 2017), the low-frequency autocorrelations may provide constraints on the deep average structure through the identification of modes and orbiting surface waves.

CONCLUSION

Low-frequency autocorrelations are useful to extract normal modes and the Earth orbiting surface waves from single-station ambient noise recordings. As a single-station approach, it is important in planetary seismology owing high cost and consequently sparse sensor deployment. At low frequencies, the deep average structure can be constrained from the normal-mode frequencies and surface-wave dispersion measurements. We show that for that purpose different autocorrelation approaches can be employed and that their performance depends on the signal and noise characteristics. If the noise has a large amplitude variability and the signals are more phase coherent than the noise, then PCC outperforms the classical correlation (CC, CCGN, or CCGNN) approaches. This is mainly because the classical approaches are amplitude biased and may have a lower signal extraction convergence depending on the required data preprocessing steps to balance amplitudes in the data. The preprocessing to balance amplitudes reduces the data information contents through either discarding data with larger amplitudes or reducing waveform complexity to clipped waveforms. In that spirit, PCC is a good candidate for the upcoming INSIGHT Mars mission because data might be limited due to unfavorable deployment and detection conditions. However, if signals are detected through their energy rather than their coherence then the classical approaches are expected to perform better.

We further remark that the correlation is an important ingredient of seismic interferometry because any subsequent stacking approach can only enhance what has been detected in the correlation. We also advocate the joint usage of independent approaches because they may provide complementary constraints besides of being a means of testing the robustness or consistency of any measurement. All our findings can be extended to CCs whenever two or more stations are available and are also valid for different other applications and frequency bands.

DATA AND RESOURCES

The data used in this article are owned by the French GEOSCOPE observatory and are freely available through download from GEOSCOPE (http://geoscope.ipgp.fr/index.php/en/) or Incorporated Research Institutions for Seismology (IRIS; http://ds.iris.edu/ds/). The data processing was performed on a linux laptop without any special hardware requirements and the Fortran codes for the computation of correlations are freely distributed by the authors. The article was written using the free LaTex software and all plots were done with the Generic Mapping Tools v.5.2.1 (www.soest.hawaii.edu/gmt; Wessel and Smith, 1998). All websites were last accessed on May 15.

ACKNOWLEDGMENTS

The authors acknowledge Electronic Seismologist Editor Huajian Yao and Editor-in-Chief Zhigang Peng for their efforts in handling the article and two anonymous reviewers for their constructive and careful reviews which helped us improve our article. This work was supported by the projects CGL2013-48601-C2-1-R and ANR-14-CE01-0012. The authors also acknowledge support for discussions within TIDES (TIme DEpendent Seismology, COST Action ES1401). M. S. thanks IPGP for their invitations to several research stays. This is IPGP Contribution Number 3938.

REFERENCES

Alessio, S. M. (2015). *Digital Signal Processing and Spectral Analysis for Scientists: Concepts and Applications*, Springer International Publishing, Cham, Switzerland.

Ardhuin, F., L. Gualtieri, and E. Stutzmann (2015). How ocean waves rock the earth: Two mechanisms explain microseisms with periods 3 to 300 s, *Geophys. Res. Lett.* **42,** no. 3, 765–772.

Banerdt, W., S. Smrekar, P. Lognonné, T. Spohn, S. Asmar, D. Banfield, L. Boschi, U. Christensen, V. Dehant, W. Folkner, *et al.* (2013). Insight: A discovery mission to explore the interior of Mars, *Lunar and Planetary Science Conference*, Vol. 44, 1915.

Becker, G., and B. Knapmeyer-Endrun (2017). Crustal thickness across the trans-European suture zone from ambient noise autocorrelations, *Geophys.J. Int.* **212,** no. 2, 1237–1254.

Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.* **169,** no. 3, 1239–1260, doi: 10.1111/j.1365-246X.2007.03374.x.

Bonatto, L., M. Schimmel, J. Gallart, and J. Morales (2015). The upper-mantle transition zone beneath the Ibero-Maghrebian region as seen by teleseismic *Pds* phases, *Tectonophysics* **663,** 212–224, doi: 10.1016/j.tecto.2015.02.002.

Bornmann, P., and E. Wielandt (2013). Seismic signals and noise, in *New Manual of Seismological Observatory Practice 2*, P. Bormann (Editor), 1–62, doi: 10.2312/GFZ.NMSOP-2_ch4.

Claerbout, J. F. (1968). Synthesis of a layered medium from its acoustic transmission response, *Geophysics* **33,** no. 2, 264–269.

Curtis, A., P. Gerstoft, H. Sato, R. Snieder, and K. Wapenaar (2006). Seismic interferometry—Turning noise into signal, *The Leading Edge* **25,** no. 9, 1082–1092.

Deen, M., E. Wielandt, E. Stutzmann, W. Crawford, G. Barruol, and K. Sigloch (2017). First observation of the Earth's permanent free oscillations on ocean bottom seismometers, *Geophys. Res. Lett.* **44,** no. 21, 10,988–10,996.

Derode, A., E. Larose, M. Tanter, J. de Rosny, A. Tourin, M. Campillo, and M. Fink (2003). Recovering the Green's function from field–field correlations in an open scattering medium (*L*), *J. Acoust. Soc. Am.* **113,** no. 6, 2973, doi: 10.1121/1.1570436.

D'Hour, V., M. Schimmel, A. Do Nascimento, J. Ferreira, and H. L. Neto (2016). Detection of subtle hydromechanical medium changes caused by a small-magnitude earthquake swarm in NE Brazil, *Pure Appl. Geophys.* **173,** no. 4, 1097–1113, doi: 10.1007/s00024-015-1156-0.

Dziewonski, A. M., and D. L. Anderson (1981). Preliminary reference Earth model, *Phys. Earth Planet. In.* **25,** no. 4, 297–356.

Ekström, G. (2001). Time domain analysis of Earth's long-period background seismic radiation, *J. Geophys. Res.* **106,** no. B11, 26,483–26,493.

Galetti, E., and A. Curtis (2012). Generalised receiver functions and seismic interferometry, *Tectonophysics* **532,** 1–26.

Haned, A., E. Stutzmann, M. Schimmel, S. Kiselev, A. Devaille, and A. Yelles-Chaouche (2016). Global tomography from seismic hum, *Geophys. J. Int.* **204,** 1222–1236, doi: 10.1093/gji/ggv516.

Kobayashi, N., and K. Nishida (1998). Continuous excitation of planetary free oscillations by atmospheric disturbances, *Nature* **395,** no. 6700, 357–360.

Kurrle, D., and R. Widmer-Schnidrig (2008). The horizontal hum of the Earth: A global background of spheroidal and toroidal modes, *Geophys. Res. Lett.* **35,** no. 6, doi: 10.1029/2007GL033125.

Lobkis, O., and R. Weaver (2001). On the emergence of the Green's function in the correlations of a diffuse field, *J. Acoust. Soc. Am.* **110,** 3011–3017.

Lognonné, P., and W. Pike (2015). Planetary seismometry, *Extraterr. Seismol.* 36–48.

Medeiros, W. E., M. Schimmel, and A. F. do Nascimento (2015). How much averaging is necessary to cancel out cross-terms in noise correlation studies?, *Geophys. J. Int.* **203,** no. 2, 1096–1100, doi: 10.1093/gji/ggv336.

Nawa, K., N. Suda, Y. Fukao, T. Sato, Y. Aoyama, and K. Shibuya (1998). Incessant excitation of the Earth's free oscillations, *Earth Planets Space* **50,** no. 1, 3–8.

Nishida, K. (2014). Source spectra of seismic hum, *Geophys. J. Int.* **199,** no. 1, 416–429.

Nishida, K. (2017). Ambient seismic wave field, *Proc. Jpn. Acad. B Phys. Biol. Sci.* **93,** no. 7, 423–448.

Nishida, K., J.-P. Montagner, and H. Kawakatsu (2009). Global surface wave tomography using seismic hum, *Science* **326,** no. 5949, 112.

Panning, M. P., P. Lognonné, W. B. Banerdt, R. Garcia, M. Golombek, S. Kedar, B. Knapmeyer-Endrun, A. Mocquet, N. A. Teanby, J. Tromp, *et al.* (2017). Planned products of the Mars structure service for the insight mission to Mars, *Space Sci. Rev.* **211,** nos. 1/4, 611–650.

Pedersen, H. A., and F. Krüger (2007). Influence of the seismic noise characteristics on noise correlations in the Baltic shield, *Geophys. J. Int.* **168,** no. 1, 197–210.

Prieto, G. A., R. L. Parker, D. J. Thomson, F. L. Vernon, and R. L. Graham (2007). Reducing the bias of multitaper spectrum estimates, *Geophys. J. Int.* **171,** no. 3, 1269–1281, doi: 10.1111/j.1365-246X.2007.03592.x.

Rhie, J., and B. Romanowicz (2004). Excitation of Earth's continuous free oscillations by atmosphere–ocean–seafloor coupling, *Nature* **431,** no. 7008, 552–556.

Roux, P., K. G. Sabra, W. A. Kuperman, and A. Roux (2005). Ambient noise cross correlation in free space: Theoretical approach, *J. Acoust. Soc. Am.* **117,** no. 1, 79, doi: 10.1121/1.1830673.

Sabra, K. G., P. Gerstoft, P. Roux, W. Kuperman, and M. C. Fehler (2005). Surface wave tomography from microseisms in southern California, *Geophys. Res. Lett.* **32,** no. 14, doi: 10.1029/2005GL023155.

Schimmel, M. (1999). Phase cross-correlations: Design, comparisons, and applications, *Bull. Seismol. Soc. Am.* **89,** no. 5, 1366–1378.

Schimmel, M., E. Stutzmann, and J. Gallart (2011). Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale, *Geophys. J. Int.* **184,** no. 1, 494–506, doi: 10.1111/j.1365-246X.2010.04861.x.

Schimmel, M., E. Stutzmann, and S. Ventosa (2017). Measuring group velocity in seismic noise correlation studies based on phase coherence and resampling strategies, *IEEE Trans. Geosci. Remote Sens.* **55,** no. 4, 1928–1935, doi: 10.1109/TGRS.2016.2631445.

Sens-Schönfelder, C., and U. Wegler (2006). Passive image interferometry and seasonal variations of seismic velocities at Merapi volcano, Indonesia, *Geophys. Res. Lett.* **33,** no. 21, doi: 10.1029/2006GL027797.

Shapiro, N. M., and M. Campillo (2004). Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.* **31,** no. 7, doi: 10.1029/2004GL019491.

Snieder, R. (2004). Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase, *Phys. Rev. E* **69,** no. 4, doi: 10.1103/PhysRevE.69.046610.

Snieder, R., and C. Sens-Schönfelder (2015). Seismic interferometry and stationary phase at caustics, *J. Geophys. Res.* **120,** no. 6, 4333–4343.

Suda, N., K. Nawa, and Y. Fukao (1998). Earth's background free oscillations, *Science* **279,** no. 5359, 2089–2091.

Taner, M. T., F. Koehler, and R. Sheriff (1979). Complex seismic trace analysis, *Geophysics* **44,** no. 6, 1041–1063.

Tanimoto, T., and J. Um (1999). Cause of continuous oscillations of the Earth, *J. Geophys. Res.* **104,** no. B12, 28,723–28,739.

Tanimoto, T., J. Um, K. Nishida, and N. Kobayashi (1998). Earth's continuous oscillations observed on seismically quiet days, *Geophys. Res. Lett.* **25,** no. 10, 1553–1556.

Taylor, G., S. Rost, and G. Houseman (2016). Crustal imaging across the North Anatolian fault zone from the autocorrelation of ambient seismic noise, *Geophys. Res. Lett.* **43,** no. 6, 2502–2509.

Thomson, D. J. (1982). Spectrum estimation and harmonic analysis, *Proc. IEEE* **70,** no. 9, 1055–1096.

Tibuleac, I. M., and D. von Seggern (2012). Crust-mantle boundary reflectors in Nevada from ambient seismic noise autocorrelations, *Geophys. J. Int.* **189,** no. 1, 493–500.

Ventosa, S., M. Schimmel, and E. Stutzmann (2017). Extracting surface waves, hum and normal modes: Time-scale phase-weighted stack and beyond, *Geophys. J. Int.* **211,** no. 1, 30–44, doi: 10.1093/gji/ggx284.

Wapenaar, K. (2004). Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation, *Phys. Rev. Lett.* **93,** no. 25, 254301.

Wapenaar, K., J. Thorbecke, and D. Draganov (2004). Relations between reflection and transmission responses of three-dimensional inhomogeneous media, *Geophys. J. Int.* **156,** no. 2, 179–194.

Webb, S. C. (2007). The Earth's hum is driven by ocean waves over the continental shelves, *Nature* **445,** no. 7129, 754–756.

Wessel, P., and W. H. Smith (1998). New, improved version of generic mapping tools released *Eos Trans. AGU* **79,** no. 47, 579–579.

Woodhouse, J. H., and A. Deuss (2015). Theory and observations—Earth's free oscillations, in *Reference Module in Earth Systems and Environmental Sciences, Treatise on Geophysics*, Second Ed., Vol. 1, Elsevier, 79–115, ISBN: 978-0-12-409548-9.

Zürn, W., G. Laske, R. Widmer-Schnidrig, and F. Gilbert (2000). Observation of Coriolis coupled modes below 1 mHz, *Geophys. J. Int.* **143,** no. 1, 113–118.

M. Schimmel
S. Ventosa
Institute of Earth Sciences Jaume Almera – CSIC
Lluis Sole i Sabaris s/n
E-08028 Barcelona, Spain
schimmel@ictja.csic.es

E. Stutzmann
Institut de Physique du Globe de Paris
UMR 7154 CNRS
1 rue Jussieu
75005 Paris
France

Published Online 30 May 2018