Extracting surface waves, hum and normal modes: Time-scale phase-weighted stack and beyond

Sergi Ventosa[1], Martin Schimmel[1] and Eleonore Stutzmann[2]

[1] *Institute of Earth Sciences Jaume Almera - CSIC, Barcelona, Spain.*

[2] *Institut de Physique du Globe de Paris, Sorbonne Paris Cité, Univ Paris Diderot, UMR 7154 CNRS, F-75005 Paris, France.*

4 July 2017

SUMMARY

Stacks of ambient noise correlations are routinely used to extract empirical Green's functions (EGF) between station pairs. The time-frequency phase-weighted stack (tf-PWS) is a physically-intuitive non-linear denoising method that uses the phase coherence to improve EGF convergence when the performance of conventional linear averaging methods is not sufficient. The high computational cost of a continuous approach to the time-frequency transformation is currently a main limitation in ambient noise studies. We introduce the time-scale phase-weighted stack (ts-PWS) as an alternative extension of the phase-weighted stack that uses complex frames of wavelets to build a time-frequency representation that is much more efficient and fast to compute and that preserve the performance and flexibility of the tf-PWS. In addition, we propose two strategies: the unbiased phase coherence and the two-stage ts-PWS methods to further improve noise attenuation, quality of the extracted signals and convergence speed. We demonstrate that these approaches enable to extract minor- and major-arc Rayleigh waves (up to the sixth Rayleigh wave train) from many years of data from the GEOSCOPE global network. Finally we also show that fundamental spheroidal modes can be extracted from these EGF.

Key words: Surface waves and free oscillations; Seismic noise; Seismic interferometry; Wavelet transform.

1 INTRODUCTION

The ubiquity of noise is making seismic ambient noise an excellent complement to earthquakes in local to global tomography and monitoring studies. Seismic noise analyses are important from early studies (Gutenberg 1958) until today (e.g., GSN low-noise model (Berger et al. 2004)), since it defines the weakest signal an instrument can detect. Seismic imaging and monitoring using inter-station correlations became popular much later with works on the extraction of the empirical Green's function (EGF) (Campillo & Paul 2003; Shapiro & Campillo 2004; Snieder 2004) and the subsequent application to surface-wave tomography (Shapiro et al. 2005; Sabra et al. 2005), see, e.g., Wapenaar et al. (2010) and Campillo & Roux (2015) for a review. The main signals currently used are from wavefield fluctuations caused by scattered earthquake coda (Sens-Schönfelder et al. 2015) and ambient noise. The strongest and most commonly used ambient noise (of natural origin) is classified into (a) primary microseisms (periods of 10-20 s), generated by interactions of ocean gravity waves with the coast, (b) secondary microseisms (1-12 s), generated by wave-wave interactions between ocean gravity waves, and (c) hum (30-250 s), generated by interactions of infragravity waves with continental shelves (Ardhuin et al. 2015).

Equipartition of wavefields is fundamental to extract the exact Green's function. Direct correlation of raw seismic records produces an undesired predominance of earthquakes and other large amplitude signals over ambient-noise sources leading to low convergence rate towards the EGF. To improve the EGF convergence, a single inter-station correlation is replaced by many much shorter data sequences treated in processing flows that may include 1-bit amplitude normalization, spectral whitening, phase cross-correlation, adaptive filtering and array processing (Campillo & Paul 2003; Bensen et al. 2007; Baig et al. 2009; Schimmel et al. 2011). Ermert et al. (2016) use minimal preprocessing for the stack of correlograms in order to obtain correlations that can be easily forward modeled instead of EGFs. Later, Fichtner et al. (2017) also consider any linear and nonlinear processing.

In most applications, physically-intuitive signal-processing methods developed with other applications in mind (e.g., beamforming or phase-weighted stack) have proven useful for improving EGF convergence and observing weak signals in otherwise too noisy records. The search of more powerful methods is needed to tackle new imaging and monitoring challenges using current data volumes and to satisfy the need of better time and space resolution. Monitoring fast temporal variations (Grêt et al. 2005; Brenguier et al. 2008; Hadziioannou et al. 2011; D'Hour et al. 2016), demands short time windows to resolve corresponding variations at a local scale promoting the use of much less data to extract EGFs. Further, the extraction of body waves (e.g., at teleseismic distances, see Gerstoft et al. (e.g., 2008); Landès et al. (e.g., 2010); Poli et al. (e.g., 2012); Boué et al. (e.g., 2013, 2014); Poli et al. (e.g., 2015); Haned et al. (e.g., 2016); Farra et al. (e.g., 2016), among others), which is possible through stacking and/or beamforming methods, demands further SNR improvements to be able to constrain deep structure. In this context, stacking and denoising strategies introduced by the seismology, exploration and signal-processing communities become fundamental to boost the amount of high-quality observations.

Stacking (i.e., averaging) of a collection of traces (time series) is an ubiquitous method to improve signal-to-noise ratio (SNR) by combining coherent signals from many traces into a single trace, often at the expense of losing resolution. Seismic exploration has a long tradition introducing original stacking methods, see, e.g., Rashed (2014) for an historical review on common-mid-point stacking. A key problem still facing many modern linear stacking methods is assigning a weight to each trace in the stack (or locally to each time sample) and rejecting the anomalous ones (due to, e.g., coherent noise). Common criteria employed to estimate these weights involve measuring SNR (Neelamani et al. 2006) or correlation (Liu et al. 2009; Sanchis & Hanssen 2011; Deng et al. 2016) to a previously calculated reference trace. Non-linear methods such as the median stack (Claerbout & Muir 1973) and the alpha-trimmed average (Bednar & Watt 1984; Rashed 2008) help rejecting coherent noise. Other non-linear stacking methods used in seismology are the n-th root stack proposed by Muirhead (1968), a norm-like function carrying the sign of the signals that emphasizes small-amplitude signals at the expense of waveform distortion (Kanasewich et al. 1973), and the dual bootstrap stack proposed by Korenaga (2013), a linear stack weighted by signal

significance and coherence to improve the detection of weak amplitude signals that also introduces waveform distortion, particularly around zero crossings.

Denoising methods improve the SNR of a single trace in exchange of signal distortion. Likewise, they boost convergence in noise correlations by reducing the number of traces needed in the stacking to get a desired SNR. Generally applied after stacking, denoising preserves the main components of the signal and attenuates the rest. Most methods apply thresholding functions controlled by the level of noise to signal representations well adapted to the signal properties in order to boost noise attenuation and promote the most energetic signals (Mallat 2008), see, e.g., Galiana-Merino et al. (2003); Han & van der Baan (2015); Mousavi & Langston (2016) for some applications to seismic data. Alternatively, Schimmel & Paulssen (1997) propose the phase-weighted stack (PWS) to extract common coherent signals, regardless of their amplitude, instead of the most energetic signals. The PWS approach boosts the SNR of the linear stacked trace and reduces waveform distortion compared to more aggressive non-linear stacking methods such as the n-th root stack and dual bootstrap stack. The time-frequency PWS (tf-PWS), introduced by Schimmel & Gallart (2007) as an extension of PWS to the time-frequency domain, further improves the extraction of weak coherent signals and reduces signal distortion. First applied to clean lower crust and Moho reflections/refractions in wide-angle data (Schimmel & Gallart 2007; Garcia Cano et al. 2014), to detect coda phases, e.g., from upper mantle discontinuities (Schimmel & Gallart 2007) or from long period data tf-PWS (Ringler et al. 2016); it has been extensively used to extract empirical inter-station Green's functions (EGF) in seismic ambient-noise studies from regional to global scales (Baig et al. 2009; Schimmel et al. 2011; Kimman et al. 2012; Ren et al. 2013; Yang 2014; Dias et al. 2015; Chao et al. 2015; Cheng et al. 2015; Haned et al. 2016; Szanyi et al. 2016; Jiang et al. 2016; Pilia et al. 2016) and has proved useful for the analysis of low-frequency earthquakes (Thurber et al. 2014; Matoza et al. 2015; Lyons et al. 2016).

The high computational cost derived from a continuous approach to the time-frequency transformation is currently a limitation in main applications of noise correlations, tomography and monitoring that often demand big data volumes. A solution to this problem is developing better implementations, e.g., Zeng & Thurber (2016) use the graphic processing unit to accelerate

tf-PWS computation by a factor of 20 with respect to implementations using the FFTw3 library (Frigo & Johnson 2005). For our research, we reduce the operation complexity (see section 3) by building more efficient time-frequency representations instead.

This work is devoted to enhance the quality of stacked seismic data sequences, in particular from ambient noise correlations, by improving SNR and EGF convergence. Then, having computational efficient methods help us to do more with less resources. In section 2, we analyze tf-PWS as a non-linear denoising method which is related to approaches that use thresholding functions. In section 3, we introduce the time-scale PWS (ts-PWS), a method that obtains the same performance as tf-PWS with a much reduced computational cost and memory footprint. Further, in section 4, we propose and discuss two strategies: the unbiased ts-PWS and the two-stage ts-PWS to improve non-coherent noise attenuation, to reduce undesired attenuation of weak coherent signals, and to increase convergence speed. In order to validate our method, we compare in section 5 linear stack with alternative methods based on the ts-PWS using synthetic and field data examples. In these examples, we assume a diffuse wavefield, for instance, due to a balance of source distribution. We extract minor- and major-arc Rayleigh waves using many years of data, study and compare the quality of the extracted signals, and measure their convergence before and after denoising. Finally, we show that we can extract fundamental spheroidal modes from EGF.

2 UNDERSTANDING TF-PWS

Phase-weighted stack (PWS) is essentially the product of the linear stack multiplied by the phase stack (Schimmel & Paulssen 1997). The time-frequency PWS (tf-PWS) method improves SNR and reduces signal distortions compared to the PWS by being more data adaptive through the attenuation of incoherent signal components in the time-frequency domain, better adapted to seismic signals than the time domain. The tf-PWS implements the time-frequency expansion using the physically-intuitive S-transform (Stockwell et al. 1996; Schimmel & Gallart 2005) because it gives absolute phase information. Ventosa et al. (2008) further showed that the S-transform is a Morlet wavelet transform up to a phase correction.

In a linear stack, the main distinguishing feature between signal and noise is that the signal is

coherent across the prestacked traces while the noise is not. The actual physical sense of signal and noise changes depending on the application. Assuming x^k is the k-th (e.g., daily) correlation sequence between a pair of stations composed by an EGF, s, and w^k signals that do not contribute to the EGF:

$$x^k[m] = s[m] + w^k[m] \tag{1}$$

where m is the index of the coefficients in the time-frequency domain. Then, the estimation of the signal s, $\tilde{s}$, from the noisy measurements x^k using the linear stack is

$$\tilde{s}_{\mathrm{ls}}[m] = \frac{1}{K}\sum_{k=1}^{K} x^k[m] \tag{2}$$

where K is the number of data sequences in the stack and the sum symbol denotes a linear stack. The variance of the estimation of the EGF reduces in mean as the number of independent sequences increases and incoherent signal, e.g., due to isolated localized noise sources, average out.

Beyond using more and more data in the stack, non-linear denoising methods such as tf-PWS improve signal estimation in stacked sequences by defining non-linear attenuation functions that exploit known signal properties to further attenuate noise. Efficient non-linear estimators are conventionally done by applying simple attenuation functions in transformed domains (here the time-frequency domain) where signal and noise are better separated.

To improve the SNR of $\tilde{s}$ through denoising, regardless of the domain (e.g., time or time-frequency) used for data representation, we multiply the linear stacked trace by an attenuation function c that promotes the main components of the signal and attenuates the rest,

$$\tilde{s}_{\mathrm{d}}[m] = \tilde{s}_{\mathrm{ls}}[m]c[m], \tag{3}$$

where $\tilde{s}_{\mathrm{d}}$ is the estimation of the signal s after denoising. In the PWS method, c is the phase stack, an empirical noise attenuation function that measures the instantaneous coherence of a signal. The phase stack,

$$c^v_{\mathrm{ps}}[m] = \left|\frac{1}{K}\sum_{k=1}^{K}\frac{x^k[m]}{|x^k[m]|}\right|^v = \left|\frac{1}{K}\sum_{k=1}^{K} e^{i\theta^k[m]}\right|^v, \tag{4}$$

is a strictly-positive function that is one if and only if all the data sequences x at a coefficient m are equal in phase (full similarity) and close to zero if they are totally unrelated (full dissimilarity),

where the sharpness of the transition between these two extremes is controlled by v. Observe that the phase stack is a sum of phasors and consequently needs x^k to be analytic to assess the phase θ^k. For example, in the time domain, coefficient m is the time step t, and $x(t) = a(t)e^{i\theta(t)}$ where a is the envelope of x and θ the instantaneous phase.

3 TIME-SCALE PHASE-WEIGHED STACK

Improving the quality of extracted signals is one of the most demanding operations (in the computational sense) along the seismic processing flow. The computational cost of tf-PWS is mainly due to the regular sampling of the continuous time-frequency domain when using the S-transform. This leads to a very high redundancy (i.e., ratio of number of coefficients in the transformed domain to number of coefficients in the original domain) and, consequently, to many coherence measures to be computed. In this section we present more computational efficient time-frequency representations of the tf-PWS through the wavelet transform.

3.1 Seeking for an efficient representation of the time-frequency domain

The most suitable time-frequency resolution is often non-uniform and therefore Geophysical applications favor time resolutions proportional to period to account for the fast time variation of ballistic signals at all frequency bands. These applications also favor analytic signal transformations because they are more suitable to process oscillatory signals (here seismic signals) than real transformations which suffer from serious shift variance problems, i.e., small time shifts of the signal leads to strong variations of the transformed coefficients around singularities (e.g., Selesnick et al. 2005). Analytic signal representations enable new algorithms that exploit magnitude and phase information; in particular, to build weight functions using the phase coherency.

Traditionally, these time-frequency expansions use a set of band-pass Gaussian filters based on the assumption that time series are derived from a multidimensional normal distribution. These filters are defined as

$$X_{\mathrm{a}}(\omega, \omega_0) = X_{\mathrm{a}}(\omega) e^{-\frac{1}{2}\left(\frac{\sigma(\omega-\omega_0)}{\omega_0}\right)^2} \tag{5}$$

where $X_{\mathrm{a}}(\omega) = X(\omega)(1 + \mathrm{sign}(\omega))$ is the analytic signal following the Fourier transform con-

vention $X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-i\omega t}dt$, sign is the signum function, ω_0 is the central frequency of the filter, and σ the standard deviation of the Gaussian window. Observe that $X_{\mathrm{a}}(\omega, \omega_0)$ is only approximately analytic since band-pass Gaussian filters (the last term in eq. (5)) are strictly positive functions. The half-power bandwidth of these filters is proportional to their central frequency $\Delta\omega = 2\sqrt{\ln 2}\,\omega_0/\sigma$, equivalently, they are of constant Q quality factor (ratio of central frequency to bandwidth)

$$Q = \frac{\omega_0}{\Delta\omega} = \frac{\sigma}{2\sqrt{\ln 2}} \tag{6}$$

and their impulse responses have an equal number of cycles $N_{\mathrm{cycles}} = \sigma\sqrt{\ln 2}/\pi$. In seismology, this approach to the time-frequency expansion formalized by Dziewonski et al. (1969) is known as the multiple filter technique. Similarly, in the context of the short-time Fourier transform, Gambardella (1971) proposes the use of modulated windows of constant Q, instead of constant length, for audio analysis giving a "form invariant under time scaling", and Petersen & Boll (1983) show how to sample this time-frequency domain efficiently. Later on, the study of these ideas in physics, engineering and pure and applied mathematics converged to the wavelet transform (e.g., Daubechies 1992).

The Gaussian band-pass filters used in the multiple filter technique, eq. (5), are equal up to a scaling factor inversely proportional to their central frequency. In appendix A, we show how this scaling relationship and the close to zero mean (typically $\sigma \gg 1$) enable an efficient implementation of multiple filter technique using the continuous Morlet wavelet transform, by defining scale as $\lambda = \sigma/\omega_0$ and the central frequency of the mother wavelet as $\xi_0 = \omega_0\lambda = \sigma$.

The continuous wavelet transform (CWT) (Daubechies 1992; Sinha et al. 2005; Mallat 2008) of a data sequence x is given by its inner product with a wavelet collection $\psi_{\tau,\lambda}$,

$$x(\tau, \lambda) = \langle x, \psi_{\tau,\lambda}\rangle = \int_{-\infty}^{\infty} x(t)\overline{\psi_{\tau,\lambda}(t)}\,dt, \tag{7}$$

where τ is delay or lag-time, λ is scale and $\overline{\psi(t)}$ denotes the complex conjugate of $\psi(t)$. The mother wavelet $\psi \in \mathbf{L}^2(\mathbb{R})$ is a zero-mean function that generates a set of functions, the wavelet collection $\psi_{\tau,\lambda}$, through scaling and translating operations. Conventionally, this set of functions is

normalized so that the transform is unitary,

$$\psi_{\tau,\lambda}(t) = \lambda^{-1/2}\psi(\lambda^{-1}(t-\tau)). \tag{8}$$

Wavelets can be real or analytic. Real wavelets are good for detecting sharp signal transitions and thus often used in, e.g., image processing (Mallat 2008; Jacques et al. 2011). Complex analytic wavelets are attractive for analyzing oscillatory signals due to their ability to separate magnitude and phase and therefore useful for, e.g., studying seismic signals or measuring instantaneous frequency (Selesnick et al. 2005; Sinha et al. 2005; Ventosa et al. 2012; Bayram 2013).

To implement the multiple filter technique, eq. (5), we use analytic wavelets. Pseudo-analytic wavelets can be build with modulated windows. We use a modulated Gaussian window,

$$\psi(t) = \pi^{-1/4}e^{-t^2/2}e^{i\xi_0 t}, \tag{9}$$

this mother wavelet is conventionally known as the Morlet wavelet despite of not being exactly zero mean and has a $Q = \xi_0/2\sqrt{\ln 2}$. Zero mean is obtained in the exact Morlet wavelet with an additional term, $\psi(t) = \pi^{-1/4}e^{-t^2/2}(e^{i\xi_0 t} - e^{-\xi_0^2/2})$; still, eq. (9) is a good approximation for $\xi_0 \gg 1$ due to its extremely small mean. For low ξ_0 values both Morlets are not pseudo-analytic anymore.

3.2 Discretization of the CWT: frames of continuous wavelets

We can severely reduce redundancy by using known signal properties to design bases or frames (i.e., overcomplete bases) (Kovačević & Chebira 2007a,b) which capture efficiently the main features of the signal. The frame coefficients x^k_Ψ of the data sequence x^k are computed by inner products with a frame $\Psi = \{\psi_m\}_{0\le m<M}$,

$$x^k_\Psi[m] = \left\langle x^k, \psi_m \right\rangle = \sum_n x^k[n]\overline{\psi_m[n]}. \tag{10}$$

The sequences $\{\psi_m\}_{0\le m<M}$ constitute a frame if there are two real numbers $B \ge A > 0$ such that

$$A\left\|x\right\|^2 \le \sum_m \left|\langle x, \psi_m\rangle\right|^2 \le B\left\|x\right\|^2 \tag{11}$$

for any possible sequence x, where A and B are called the lower and upper frame bounds. When eq. (11) is satisfied, the signal x^k can be reconstructed from the x^k_Ψ coefficients with the pseudo-

inverse of the frame Ψ, implemented with a dual frame $\{\widetilde{\psi}_m\}_{0\leq m<M}$,

$$x^k[n] = \sum_m x^k_\Psi[m]\widetilde{\psi}_m[n]. \tag{12}$$

If $A = B$ the frame is called tight (satisfies the Parserval's identity) and the dual frame can be approximated by the "forward" frame up to a constant factor (Daubechies 1992).

Frames of wavelets are good approximating CWT (Daubechies 1992). The frame elements are distributed across the time-frequency domain according to their time and frequency resolutions in order to contain redundancy. A common choice for discretizing the wavelet collection $\psi_{\tau,\lambda}$, eq. (8), is $\lambda = 2^s$ and $\tau = u2^j b_0$, where u is the time index at the scale s and b_0 the sampling period at the scale 1, leading to a sampling period almost proportional to scale on the time axis and to its inverse on the frequency axis. The frame of wavelets is then

$$\psi_{u,s}[n] = \frac{1}{2^{s/2}}\psi\left(\frac{nT - u2^{\lfloor s\rfloor}b_0}{2^s}\right), \tag{13}$$

where T is the sampling period of the original time series and $\lfloor s\rfloor$ is the lower integer part of s. Scales are usually organized in voices $v \in [0, V-1]$ and octaves $j \in \mathbb{Z}$, $s = j + v/V$, where $V \in \mathbb{N}$ is the number of voices; and they are downsampled by $2^j b_0$ in order to keep roughly a constant number of samples per cycle in all frequency bands. For moderate redundancies, $R \geq 4$, the reconstruction error is typically much lower than 1% when the dual frame is approximated by the forward frame and reduces fast as V increases and b_0 decreases. For example, in the case of the exact Morlet wavelet, for the standard choice of $\xi_0 = \pi\sqrt{2/\ln 2}$, and with $V = 4$ and $b_0 = 1$ (that double when ξ_0 doubles), the redundancy is approximately $V/b_0 = 4$ and the normalized root-mean-squared error is below 3.61×10^{-4}.

The use of a sampling strategy adapted to the actual resolution has a major impact on operation complexity and memory footprint. The redundancy of the tf-PWS is about half the number of samples of each sequence, $R = N/2 + 1$, while the redundancy of the ts-PWS is typically $4 \leq R \leq 16$ and independent of N. For $R < 4$ the frame bounds diverge abruptly ($A \ll B$) and the wavelet collection stops being a frame. While from $R = 16$ the reconstruction error is extremely low and the filtering improvements are very minor. Likewise, the operation complexity of the ts-

PWS is proportional to the number of samples and scales, $\mathcal{O}(NS)$, and the tf-PWS, implemented using fast Fourier transforms, is $\mathcal{O}(N^2 \log_2 N)$.

3.3 Limits and alternatives to frames of Morlet Wavelets

The analytic approximation made in the Morlet wavelets degrades as ξ_0 (equivalently σ in eq. (5)) reduces and the negative frequency components becomes less and less negligible, which is problematic in applications requiring high temporal resolution. The complex Mexican hat wavelet (Addison et al. 2002) is an alternative with no negative frequency components and a $Q = \sqrt{3/2} \approx 1.2$. Equivalently to eq. (5), the band-pass filters of this wavelet are defined as

$$X_{\mathrm{a}}(\omega, \omega_0) = X_{\mathrm{a}}(\omega) \frac{2\omega^2}{\omega_0^2} e^{-\left(\frac{\omega}{\omega_0}\right)^2}. \tag{14}$$

A more flexible option is the Morse wavelet (Lilly & Olhede 2009, 2012), a family of analytic wavelets which permits the synthesis of a wide variety of wavelets from low to high Q quality factors.

Frames of continuous wavelets have a far lower redundancy than direct continuous implementations, still this redundancy is moderately high and the discretization leads to a non-perfect reconstruction. The discrete wavelet transform (DWT) is required to get the lowest redundancy and perfect reconstruction, e.g., Vetterli & Kovačević (1995). A certain degree of redundancy is necessary to construct an analytic DWT because discrete complex wavelets having a finite support cannot form an orthonormal or biorthogonal basis and be analytic or pseudo-analytic at the same time, e.g., Selesnick et al. (2005). For example, the dual-tree complex wavelet transform (Kingsbury 2001; Selesnick et al. 2005) uses wavelets with a low Q, has a redundancy of 2, and an operation complexity of $\mathcal{O}(N)$ which is lower than the fast Fourier transform. The dual-tree rational-dilation complex wavelet transform (Bayram & Selesnick 2011) and rational-dilation filter banks Bayram (2013) allow for a more flexible choice of Q, redundancy and time-frequency resolution. However, the filters obtained have to be constructed in the frequency domain to fulfill perfect reconstruction which increase operation complexity.

4 IMPROVING WEAK SIGNAL DETECTION

Most coherence measures can be used as empirical noise attenuation functions exploiting data from several realizations, being the main difference between them the actual definition of similarity (e.g., maximum coherence means amplitudes and waveforms being equal or just the waveforms) and the class of signals considered (e.g., real or analytic).

The phase stack is an instantaneous coherence estimator for analytic signals which gives more stable measurements than alternative estimators for real signals due to the near shift invariance of magnitude and phase. Other coherence measures such as semblance and geometrical-normalized cross-correlation (GNCC) allow both real and analytic signals (Taner et al. 1979). Note that the actual definition of similarity is not equal in these three estimators. The phase stack and GNCC define full similarity as equal waveforms (not amplitudes) and, for this reason, are said to be amplitude unbiased, whereas semblance defines full similarity as equal amplitude and waveforms and therefore it is amplitude biased.

The instantaneous coherence estimation of the phase stack is biased in the sense that it gives positive coherence values for full dissimilarity signals, whereas instantaneous GNCC coherence estimation is unbiased since coherence can become negative. Fig. 1 illustrates the impact of the coherence bias by showing the amplitude attenuation of the phase stack, eq. (4), as function of the SNR of full coherent signals contaminated by white Gaussian noise for the standard choice of $v = 2$. The signal attenuation is small for SNRs of the prestack higher than 1, and it increases on average to a maximum value of K for pure incoherent noise, see appendix B for the demonstration. In practice, this saturation of attenuation leads to lower than expected SNR increments for low SNR signals, but it guarantees a positive coherence.

We can improve weak signal detection by (i) promoting sequences with high SNR in a weighted phase stack, (ii) improving their SNR before tf/ts-PWS through two-stage stack, and (iii) correcting the coherence bias of the phase stack to increase noise attenuation with the unbiased phase coherence estimator. More specifically:

(i) *Weighted phase stack*: Introducing weights in the phase stack helps to promote top-quality data sequences and to reject the anomalous ones. Modern linear stacking methods used in explo-

ration can be adapted to estimate these weights, considering that the phase stack is a linear sum of phasors.

(ii) *Two-stage stack*: The phase stack attenuates low SNR signals much more than the stronger ones. For this reason, we obtain better results if we phase stack a few sequences with high SNR than if we stack many of much lower SNR, even when their linear stacks are identical. Consequently, if the number of sequences available is high, it seems reasonable to arrange them in few groups and apply linear stacking to each group to generate sequences with reduced signal variability (higher SNR) in a first stage, and then apply the phase stack in a second stage. In this manner, coherence is increased before tf/ts-PWS and therefore signals are less downweighted. Note that in ambient noise data the coherence can be increase before tf/ts-PWS using longer data windows in the cross-correlations. However, this reduces the adaptability to real data with gaps and the applicability of other processing steps, such as normalizations and weights to promote or reject sequences according their quality. For example, Jiang et al. (2016) linearly stack daily correlation in ten-days periods before using tf/ts-PWS to improve SNR of short-period ambient noise data.

(iii) *Unbiased phase coherence*: Noise attenuation can improve by using unbiased coherence measures, such as GNCC coherence, or by correcting the bias of the phase stack. In appendix B, we show that an unbiased phase coherence estimator, c_{upc} can be constructed using the phase stack with $v = 2$,

$$c_{\mathrm{upc}}^2 = \frac{K c_{\mathrm{ps}}^2 - 1}{K - 1}. \tag{15}$$

In Fig. 1 we show that, in mean, the attenuation of the unbiased phase coherence (solid lines) is much higher than the attenuation of the phase stack (dashed lines) when the SNR of the data sequences is very low and that both are similar when the SNR is higher than one. The akin PWS, constructed using c_{upc} rather than c_{ps}, is now called "unbiased" tf/ts-PWS.

In the next sections, we construct the two-stage ts-PWS using the linear stack in the first stage and the unbiased ts-PWS in the second.

5 EXAMPLES

In the following we use synthetic and real datasets to evaluate the performance of the linear stack, ts-PWS and two-stage ts-PWS. As measure of performance we use the similarity which we define as the cross-correlation coefficient CC between the original signal s and the estimated signal $\tilde{s}$,

$$\mathrm{CC} = \frac{|\langle s, \tilde{s} \rangle|}{\|s\| \, \|\tilde{s}\|}, \tag{16}$$

where $\langle s, \tilde{s} \rangle$ is the inner product and $\|s\|$ the norm. The corresponding misfit is defined as $1 - \mathrm{CC}$. Since time and frequency resolutions are often main constraints in distinct applications, the Q quality factor (defined in eq. 6) helps determining the best adapted wavelets. Generally, high Q wavelets are suitable for studies using surface waves due to their dispersion with frequency, while low Q wavelets are better adapted for analyzing body waves due to their short duration.

5.1 Synthetic data example: high Q wavelets

The test data, Fig. 2(a), consists of a chirp function,

$$x(t) = \begin{cases} \sin(2\pi f_s t) & \text{if } t_0 \leq t \leq t_1 \\ 0 & \text{otherwise} \end{cases}, \tag{17}$$

windowed with a 20% raised cosine taper, where $f_s = f_0 + (f_1/f_0)^{(t-t_0)/(t_1-t_0)}$ with $f_0 = 0.005$ Hz, $f_1 = 0.03$ Hz, $t_0 = 100$ s and $t_1 = 1001$ s, Fig. 2(b), embedded in white Gaussian noise of variance equal to 1 and sampled at 1 sample/s. Fig. 2(c) shows the convergence to the unperturbed chirp signal using three methods (black) linear stack, (blue) tf/ts-PWS, and (red) two-stage ts-PWS. Finally, Figs 2(d) and 2(e) show the extracted signals and their differences, respectively. The two-stage stack first arranges the original data sequences into 10 groups to linearly stack the sequences of each group, and then applies the unbiased ts-PWS. With 10 groups we seek a balance between increasing SNR before the ts-PWS and reducing the variance of the coherence estimation.

The main parameters to chose for the time-frequency transformation are the Q quality factor and the frequency band of interest. We opt for a Morlet wavelet with a relatively high Q of 5 ($\xi_0 \cong 8.33$ or about 2.2 cycles) considering the long signal duration and its high number of cycles. For extracting much shorter waveforms, such as body waves, a lower value is recommended to

increase time resolution. The center frequencies of the wavelet collection are ξ_0/λ. We opt for using 8 octaves starting at the scale $\lambda = 4$, which is more than enough to cover the frequencies of interest. Using a narrower frequency range would have further removed noise. We discretize the wavelet collection from the standard choice of $\xi_0 \cong 5.34$, and with $V = 4$ and $b_0 = 1$ by scaling the V and b_0 parameters in proportion to ξ_0. Then V is rounded to the closer integer value and b_0 to the largest power of two lower than b_0, leading to $V = 6$ and $b_0 = 1$ and a redundancy of $R \simeq 6$. The misfit ($1 -$ CC) of the ts-PWS using these parameters is 2.9×10^{-3} and the misfit of the tf-PWS using an equivalent configuration is 4.3×10^{-3}. This minor difference is mainly due to the different discretization used. Further increasing V or b_0 leads to very minor reduction of the misfit (Fig. S1).

Because tf-PWS and ts-PWS give identical waveforms, in the following we only show results for the ts-PWS which we call tf/ts-PWS. In the misfit ($1 -$ CC) results, Fig. 2(c), we see that the tf/ts-PWS reduces misfit much faster than linear stacking by removing noise across all frequency bands. Further, the linear stack needs more than 100 sequences to give results with the misfit that ts-PWS obtains using 10 sequences. For higher number of sequences, we observe that the tf/ts-PWS misfit starts saturating at about 20 sequences (and results no longer improve) while the linear stack keeps improving. When comparing the differences, Fig. 2(d), of the linear stack and the tf/ts-PWS, we can conclude that this saturation is due to the signal attenuation produced by the reduction of the phase coherency (the attenuation function used in the ts-PWS) on signals with low SNR, as predicted in Fig. 1.

We avoid the misfit saturation by introducing a two-stage stacking approach. The linear-stack stage delivers a few sequences with a much higher SNR than the original data sequences that are then stacked in the second stage using the unbiased ts-PWS. A moderate number of sequences (10 in this example) is sufficient to estimate coherence satisfactorily in practice. In exchange, we obtain a higher SNR that has a major impact on the misfit reduction (equivalently, decrement on waveform distortion) which now improves linearly with the number of sequences in parallel to the linear stack. The unbiased phase coherence helps to further improve misfit in the single-stage

ts-PWS (number of sequences lower than 10) and the two-stage ts-PWS as a result of the higher noise attenuation predicted in Fig. 1.

An implication of this result is that a faster convergence to a robust signal can be achieved through the introduced modifications. The fast convergence of the ts-PWS and the two-stage ts-PWS proposed here should be useful for, e.g., seismic monitoring as it allows us to use much shorter data windows leading to higher temporal resolution, key in these applications.

5.2 Field data example: extracting surface waves and normal modes from ambient noise

We extract minor- and major-arc Rayleigh waves from inter-station correlations using three different stacking strategies. For further information on the analysis of dispersion of Rayleigh waves see Levshin & Ritzwoller (2001); Herrmann & Ammon (2002); Schimmel et al. (2017) among others and, e.g., Haned et al. (2016) for a recent study deriving a global upper-mantle tomographic model and Zhao et al. (2017) building phase velocity maps across the USA. Fig. 3 shows the velocity EGF extracted from the vertical component (LHZ) of two pairs of broadband seismic stations from the GEOSCOPE network from the year 2006 to 2016 and convergences to the final waveforms. In Figs 5 and 6, we present the waveforms extracted from twenty station pairs using 500 days of data and more than 2000 days, respectively. See Medeiros et al. (2015) for some guidance on the choice of the window length and the number of correlations that ensures cross-terms canceling. In the pre-processing we remove the mean and trends, correct for the instrument response to produce ground velocity, apply a band-pass filter from 4 to 32 mHz, decimate to a sampling period of 4 s, reject seismograms having energies much higher than the average from the same time period (higher than 15 times the median of their standard deviations), and clip large-amplitude signals (higher than 4 times the standard deviation of the seismogram). Then, we compute GNCC of 1-day long velocity seismograms, reject correlations having energies much higher than average, and stack remaining correlations using linear stacking, ts-PWS and two-stage ts-PWS. The two-stage ts-PWS consists of a first stage formed by several linear stacks, which generate 10 correlation sequences having a SNR much higher than the original correlations, followed by the unbiased ts-PWS. Assigning an optimal weight to each inter-station correlation in the stack could potentially improve results. This

strategy is hindered in this example due to the very low SNR of the correlations and is also out of the scope of this paper. Therefore, for the sake of simplicity, we weight the correlations uniformly.

We use a power of $v = 2$ for the ts-PWS and the unbiased ts-PWS. For the time-frequency representation, we use the continuous Morlet wavelet transform with the standard choice of $\xi_0 = \pi\sqrt{2/\ln 2}$, equivalently $Q \cong 3.20$, discretized using a frame with a $b_0 = 1$ and 4 voices along 3 octaves, corresponding to a collection of 12 band-pass filters with central frequencies from 4 to about 27 mHz. Memory requirements reduce compared to the equivalent configuration of the tf-PWS due to the lower redundancy, $R = V/b_0 = 4$ versus 4126 (half the number of lags in Fig. 3), and the possibility of analyzing only the frequency bands of interest. Similarly, ts-PWS is much faster to compute than tf-PWS. The ts-PWS results shown in Fig. 6(b) take about 70 s to compute with a common desktop, and the results of the two-stage ts-PWS, Fig. 6(c), 18 s, mainly consumed in input/output operations.

The first minor-arc Rayleigh (R1) wave is the main signal used in ambient-noise studies from global to local tomography. In Fig. 3, we show up to the third major-arc Rayleigh (R6) wave train. These observations are important for global tomography studies to improve coverage in regions otherwise poorly sampled due to the lack of seismic stations, such as the oceans, or earthquakes. In Figs 3(a) and 3(b), we detect with the linear stack the first four Rayleigh waves quite clearly, partially due to the strict rejection of anomalous data before and after computing the correlations. We observe that with the linear stack, the level of noise is still significant even when using all the data available (about 11 years). With the ts-PWS, the level of noise reduces dramatically, but also the amplitude of the signals as a result of their low SNR. In contrast, the two-stage ts-PWS reduces the level of noise with a much lower signal attenuation allowing to observe up to R5 and R6, because (1) the linear stacks performed in the first stage provide correlations with higher SNR to the unbiased ts-PWS, and therefore reduce signal attenuation severely, and (2) the unbiased ts-PWS helps to further attenuate non-coherent noise. R3 to R6 wave trains are harder to observe when we limit the dataset to 500 days, Figs 3(c) and 3(d). All these signals are better observed using the two-stage ts-PWS, despite of the little signal attenuation.

Figs 3(e) and 3(f) show the convergence to the EGF using as a reference signal the waveforms

extracted by each method using the entire data set (Figs 3a and 3b). We see that the ts-PWS and the two-stage ts-PWS converge faster than the linear stack, the linear stack needs about 1000 days to get the similarity that the two-stage ts-PWS gets in about 250 days due to the lower non-coherent noise attenuation. As we have seen in the synthetic data example, Fig. 2(c), the misfit (equivalently similarity) saturates for ts-PWS while improves in parallel to the linear-stack for the two-stage ts-PWS.

The performance of the tf/ts-PWS and the two-stage ts-PWS is shown more in detail in Fig. 4. Here we zoom to the R1 and R4 wave trains from Fig. 3 with corresponding waveforms shown in Fig. 4(a) and 4(b). Figs 4(c) and 4(d) show the SNR which we define as the ratio of the maximum amplitude to the standard deviation of the noise ($\sigma = \text{median}\,|x|\,/0.6745$). Both methods, tf/ts-PWS and the two-stage ts-PWS provide better SNR than the linear stack. But this information does not evaluate the quality of the waveforms, e.g., a Rayleigh wave which is reduced to a signal pulse has a high SNR. Figs 4(e) and 4(f) do not contain information on the SNR, neither on the similarity to the unknown Rayleigh waveform. But, it does show the robustness and conversion of the extracted waveform to the best possible and expected waveform obtained by using all available data. It can be seen that the tf/ts-PWS and the two-stage ts-PWS converge faster to a stable waveform. This also means that we can obtain robust results with smaller amounts of data.

Fig. 5 shows extracted Rayleigh waves from 20 inter-station correlations normalized to better observe later-arriving phases using 500 days of data and, Fig. 6, using all the data available. When using linear stacking on 500 days of data, Fig. 5(a), we clearly see the first two Rayleigh waves (R1 and R2) but Rayleigh wave trains arriving later are difficult to observe due to the still high level of noise. We need to use all data, Fig. 6(a) in order to observe R3, R4 and eventually R5. When using ts-PWS, Figs 5(b) and 6(b), we get similar observations with an improved noise level but with a reduced amplitude due to the low SNR of the dataset. In contrast, we see the first four Rayleigh waves (R1-R4) clearly at most of the station pairs when using the two-stage ts-PWS, Figs 5(c) and 6(c), and the third minor- and major-arc Rayleigh waves (R5-R6) in Fig. 6(c) but with a lower SNR.

In Fig. 7, we show spheroidal modes extracted from EGF for the station pairs used in Fig. 6,

applying the linear stack and the two-stage ts-PWS. Changes in the preprocessing with respect to the previous example are that we now produce ground acceleration and we band-pass filter from 0.5 to 10 mHz. The time-frequency representation used in the ts-PWS changes accordingly, we now use 4 voices and 5 octaves corresponding to 20 band-pass filters with central frequencies from 0.5 to about 13.5 mHz. In Figs 7(a) and 7(b), we only detect fundamental modes, which is reasonable considering that EGF correspond to Green function for a surface vertical force at the locations of one of the stations recorded by the other station. The most energetic mode, detected at all station pairs, is the fundamental spheroidal mode ${}_0S_0$ at about 0.81 mHz. Much weaker, we detect other fundamental modes clearly at all inter-station distances, they are however better observed at frequencies lower than 4 mHz and short distances. We resolve the same modes in both stacking methods, the results are slightly cleaner when using the linear stack, Fig. 7(a). This is expected since the Fourier domain is the natural representation for normal modes. The contribution of earthquakes has not been analyzed since here we focus on the applicability of the modified PWS.

6 CONCLUSIONS

We improve the EGF convergence of the tf/ts-PWS by increasing non-coherent noise attenuation with the unbiased ts-PWS, and by reducing signal attenuation with the two-stage ts-PWS. These improvements are essential to increase the resolution in monitoring studies and to improve the quality of extracted signals. We show that using these methods one can obtain high-quality minor- and major-arc Rayleigh waves from R1 to R6 from inter-station EGF. We also clearly observe many fundamental spheroidal modes with eigenfrequencies lower than 4 mHz with the same dataset, being ${}_0S_0$ the most energetic. Constraints obtained from further analysis of high-order surface waves and normal modes extracted from ambient noise can complement earthquake data in upper-mantle studies because they sample regions otherwise poorly covered due to the lack of seismic stations or earthquakes.

The ts-PWS based methods are much faster and require much less memory than tf-PWS. The parameter controlling computational cost and memory requirements of the tf-PWS and ts-PWS is the redundancy of the signal representation in the time-frequency domain. Analytic discrete

wavelets enable the lowest redundancy and analytic continuous wavelets the highest. In between these extremes, a variety of solutions exist with a rich diversity of choices on the competing parameters of redundancy, time-frequency resolution and control on the wavelet function. In this paper, we opt for frames of continuous wavelets because they provide higher time-frequency resolutions and better wavelet control. For applications which demand perfect reconstruction or the lowest redundancy, alternative solutions based on analytic discrete wavelets can be of interest.

Non-linear stacking methods such as unbiased ts-PWS and two-stage ts-PWS help improving SNR in exchange of possible signal distortion due to varying signal coherence throughout its time and frequency components. These gains in signal quality are valuable for applications analyzing signals such as body or surface waves that benefit from conventional stacking methods. Other data, as earthquake recording could have been used in full analogy and without losing generality.

ACKNOWLEDGMENTS

The authors thank the editor and two anonymous reviewers for improving the manuscript. Sergi Ventosa thanks the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement de la Generalitat de Catalunya and the 7th European Framework Program Marie Curie COFUND - contract 600385; Beatriu de Pinós Fellowship. We acknowledge support from discussions within TIDES COST Actions ES1401. Martin Schimmel and Sergi Ventosa also thank the Spanish MISTERIOS project CGL2013-48601-C2-1-R. Eléonore Stutzmann was supported by the ANR MIMOSA project (ANR-16-CE33-0005)

REFERENCES

Addison, P. S., Watson, J. N., & Feng, T., 2002. Low-oscillation complex wavelets, *J. Sound Vib.*, **254**(4), 733–762.

Ardhuin, F., Gualtieri, L., & Stutzmann, E., 2015. How ocean waves rock the Earth: Two mechanisms explain microseisms with periods 3 to 300 s, *Geophys. Res. Lett.*, **42**(3), 765–772, 2014GL062782.

Baig, A., Campillo, M., & Brenguier, F., 2009. Denoising seismic noise cross correlations, *J. Geophys. Res.*, **114**(8).

Bayram, İ., 2013. An analytic wavelet transform with a flexible time-frequency covering, *IEEE Trans. Signal Process.*, **61**(5), 1131–1142.

Bayram, İ. & Selesnick, I. W., 2011. A dual-tree rational-dilation complex wavelet transform, *IEEE Trans. Signal Process.*, **59**(12), 6251–6256.

Bednar, J. & Watt, T., 1984. Alpha-trimmed means and their relationship to median filters, *IEEE Trans. Acoust., Speech, Signal Processing*, **32**(1), 145–153.

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

Berger, J., Davis, P., & Ekström, G., 2004. Ambient Earth noise: A survey of the global seismographic network, *J. Geophys. Res.*, **109**(B11), 1–10, B11307.

Boué, P., Poli, P., Campillo, M., Pedersen, H., Briand, X., & Roux, P., 2013. Teleseismic correlations of ambient seismic noise for deep global imaging of the earth, *Geophys. J. Int.*, **194**(2), 844–848.

Boué, P., Poli, P., Campillo, M., & Roux, P., 2014. Reverberations, coda waves and ambient noise: Correlations at the global scale and retrieval of the deep phases, *Earth Planet. Sci. Lett.*, **391**, 137–145.

Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N. M., Nadeau, R. M., & Larose, E., 2008. Postseismic relaxation along the San Andreas fault at Parkfield from continuous seismological observations, *Science*, **321**(5895), 1478–1481.

Campillo, M. & Paul, A., 2003. Long-range correlations in the diffuse seismic coda, *Science*, **299**(5606), 547–549.

Campillo, M. & Roux, P., 2015. 1.12 - crust and lithospheric structure - seismic imaging and monitoring with ambient noise correlations, in *Treatise on Geophysics (Second Edition)*, pp. 391–417, ed. Schubert, G., Elsevier, Oxford, second edition edn.

Chao, W.-A., Wu, Y.-M., Zhao, L., Tsai, V., & Chen, C.-H., 2015. Seismologically determined bedload flux during the typhoon season, *Scientific Reports*, **5**.

Cheng, C., Allen, R. M., Porritt, R. W., & Ballmer, M. D., 2015. Seismic constraints on a double-layered asymmetric whole-mantle plume beneath Hawaii, in *Hawaiian volcanoes: from source to surface*, vol. 208 of **Geophysical Monograph**, pp. 19–34, Amer Geophys Union, John Wiley & Sons Inc.

Claerbout, J. F. & Muir, F., 1973. Robust modeling with erratic data, *Geophysics*, **38**(5), 826–844.

Daubechies, I., 1992. *Ten Lectures on Wavelets*, CBMS-NSF, SIAM Lecture Series, Philadelphia, PA, USA.

Deng, P., Chen, Y., Zhang, Y., & Zhou, H.-W., 2016. Weighted stacking of seismic AVO data using hybrid AB semblance and local similarity, *Journal of Geophysics and Engineering*, **13**(2), 152.

D'Hour, V., Schimmel, M., Nascimento, A. F., Ferreira, J. M., & Lima Neto, H. C., 2016. Detection of subtle hydromechanical medium changes caused by a small-magnitude earthquake swarm in NE Brazil,

Pure Appl. Geophys., pp. 1–17.

Dias, R. C., Julia, J., & Schimmel, M., 2015. Rayleigh-wave, group-velocity tomography of the Borborema province, NE Brazil, from ambient seismic noise, *Pure Appl. Geophys.*, **172**(6), 1429–1449.

Dziewonski, A., Bloch, S., & Landisman, M., 1969. A technique for the analysis of transient seismic signals, *Bull. seism. Soc. Am.*, **59**(1), 427–444.

Dziewonski, A. M. & Anderson, D. L., 1981. Preliminary reference Earth model, *Phys. Earth Planet. In.*, **25**(4), 297–356.

Ermert, L., Villaseñor, A., & Fichtner, A., 2016. Cross-correlation imaging of ambient noise sources, *Geophys. J. Int.*, **204**(1), 347–364.

Farra, V., Stutzmann, E., Gualtieri, L., Schimmel, M., & Ardhuin, F., 2016. Ray-theoretical modeling of secondary microseism p-waves, *Geophys. J. Int.*, pp. 1730–1739.

Fichtner, A., Stehly, L., Ermert, L., & Boehm, C., 2017. Generalised interferometry - I. Theory for inter-station correlations, *Geophys. J. Int.*, **208**(2), 603–638.

Frigo, M. & Johnson, S. G., 2005. The design and implementation of FFTW3, *Proceedings of the IEEE*, **93**(2), 216–231, Special issue on “Program Generation, Optimization, and Platform Adaptation”.

Galiana-Merino, J. J., Rosa-Herranz, J., Giner, J., Molina, S., & Botella, F., 2003. De-noising of short-period seismograms by wavelet packet transform, *Bull. seism. Soc. Am.*, **93**(6), 2554–2562.

Gambardella, G., 1971. A contribution to the theory of short-time spectral analysis with nonuniform bandwidth filters, *IEEE Trans. Circuit Theory*, **18**(4), 455–460.

Garcia Cano, L., Galve, A., Charvis, P., & Marcaillou, B., 2014. Three-dimensional velocity structure of the outer fore arc of the Colombia-Ecuador subduction zone and implications for the 1958 megathrust earthquake rupture zone, *J. Geophys. Res.*, **119**(2), 1041–1060.

Gerstoft, P., Shearer, P. M., Harmon, N., & Zhang, J., 2008. Global p, pp, and pkp wave microseisms observed from distant storms, *Geophys. Res. Lett.*, **35**(23), 1–6, L23306.

Grêt, A., Snieder, R., Aster, R. C., & Kyle, P. R., 2005. Monitoring rapid temporal change in a volcano with coda wave interferometry, *Geophys. Res. Lett.*, **32**(6), 1–4, L06304.

Gutenberg, B., 1958. Microseisms, vol. 5 of **Advances in Geophysics**, pp. 53–92, eds Landsberg, H. & Mieghem, J., Elsevier.

Hadziioannou, C., Larose, E., Baig, A., Roux, P., & Campillo, M., 2011. Improving temporal resolution in ambient noise monitoring of seismic wave speed, *J. Geophys. Res.*, **116**(B7), 1–10, B07304.

Han, J. & van der Baan, M., 2015. Microseismic and seismic denoising via ensemble empirical mode decomposition and adaptive thresholding, *Geophysics*, **80**(6), KS69–KS80.

Haned, A., Stutzmann, E., Schimmel, M., Kiselev, S., Davaille, A., & Yelles-Chaouche, A., 2016. Global tomography using seismic hum, *Geophys. J. Int.*, **204**(2), 1222–1236.

Herrmann, R. & Ammon, C., 2002. Computer programs in seismology: Surface waves, receiver functions

and crustal structure: St. Louis, Missouri, St. Louis University, Tech. rep., Dept. Earth Atmos. Sci., Saint Louis Univ., St. Louis, MO, USA.

Jacques, L., Duval, L., Chaux, C., & Peyré, G., 2011. A panorama on multiscale geometric representations, intertwining spatial, directional and frequency selectivity, *Signal Processing*, **91**(12), 2699–2730, Advances in Multirate Filter Bank Structures and Multiscale Representations.

Jiang, C., Yang, Y., Rawlinson, N., & Griffin, W. L., 2016. Crustal structure of the newer volcanics province, SE Australia, from ambient noise tomography, *Tectonophysics*, **683**, 382–392.

Kanasewich, E. R., Hemmings, C. D., & Alpaslan, T., 1973. Nth-root stack nonlinear multichannel filter, *Geophysics*, **38**(2), 327–338.

Kimman, W., Campman, X., & Trampert, J., 2012. Characteristics of seismic noise: Fundamental and higher mode energy observed in the northeast of the Netherlands, *Bull. seism. Soc. Am.*, **102**(4), 1388–1399.

Kingsbury, N. G., 2001. Complex wavelets for shift invariant analysis and filtering of signals, *Appl. Comp. Harm. Analysis*, **10**(3), 234–253.

Korenaga, J., 2013. Stacking with dual bootstrap resampling, *Geophys. J. Int.*, **195**(3), 2023–2036.

Kovačević, J. & Chebira, A., 2007a. Life beyond bases: The advent of frames (part I), *IEEE Signal Process. Mag.*, **24**(4), 86–104.

Kovačević, J. & Chebira, A., 2007b. Life beyond bases: The advent of frames (part II), *IEEE Signal Process. Mag.*, **5**(24), 115–125.

Landès, M., Hubans, F., Shapiro, N. M., Paul, A., & Campillo, M., 2010. Origin of deep ocean microseisms by using teleseismic body waves, *J. Geophys. Res.*, **115**(B5), B05302.

Levshin, A. L. & Ritzwoller, M. H., 2001. *Automated Detection, Extraction, and Measurement of Regional Surface Waves*, pp. 1531–1545, Birkhäuser Basel, Basel.

Lilly, J. M. & Olhede, S. C., 2009. Higher-order properties of analytic wavelets, *IEEE Trans. Signal Process.*, **57**(1), 146–160.

Lilly, J. M. & Olhede, S. C., 2012. Generalized morse wavelets as a superfamily of analytic wavelets, *IEEE Trans. Signal Process.*, **60**(11), 6036–6041.

Liu, G., Fomel, S., Jin, L., & Chen, X., 2009. Stacking seismic data using local correlation, *Geophysics*, **74**(3), V43–V48.

Lyons, J., Haney, M., Werner, C., Kelly, P., Patrick, M., Kern, C., & Trusdell, F., 2016. Long period seismicity and very long period infrasound driven by shallow magmatic degassing at Mount Pagan, Mariana Islands, *J. Geophys. Res.*, **121**(1), 188–209.

Mallat, S., 2008. *A wavelet tour of signal processing: the sparse way*, Academic press.

Matoza, R., Chouet, B., Dawson, P., Shearer, P., Haney, M., Waite, G., Moran, S., & Mikesell, T., 2015. Source mechanism of small long-period events at Mount St. Helens in July 2005 using template matching,

phase-weighted stacking, and full-waveform inversion, *J. Geophys. Res.*, **120**(9), 6351–6364.

Medeiros, W. E., Schimmel, M., & do Nascimento, A. F., 2015. How much averaging is necessary to cancel out cross-terms in noise correlation studies?, *Geophys. J. Int.*, **203**(2), 1096–1100.

Mousavi, S. M. & Langston, C. A., 2016. Hybrid seismic denoising using higher-order statistics and improved wavelet block thresholding, *Bull. seism. Soc. Am.*, **106**(4).

Muirhead, K., 1968. Eliminating false alarms when detecting seismic events automatically, *Nature*, **217**, 533–534.

Neelamani, R., Dickens, T. A., & Deffenbaugh, M., 2006. *Stack-and-Denoise: A new method to stack seismic datasets*, pp. 2827–2831.

Petersen, T. L. & Boll, S. F., 1983. Critical band analysis-synthesis, *IEEE Trans. Acoust., Speech, Signal Processing*, **31**(3), 656–663.

Pilia, S., Arroucau, P., Rawlinson, N., Reading, A. M., & Cayley, R. A., 2016. Inherited crustal deformation along the east gondwana margin revealed by seismic anisotropy tomography, *Geophys. Res. Lett.*, **43**(23), 12,082–12,090, 2016GL071201.

Poli, P., Campillo, M., & Pedersen, H., 2012. Body-wave imaging of Earth's mantle discontinuities from ambient seismic noise, *Science*, **338**(6110), 1063–1065.

Poli, P., Thomas, C., Campillo, M., & Pedersen, H. A., 2015. Imaging the d" reflector with noise correlations, *Geophys. Res. Lett.*, **42**(1), 60–65, 2014GL062198.

Rashed, M., 2014. Fifty years of stacking, *Acta Geophysica*, **62**(3).

Rashed, M. A., 2008. Smart stacking: A new CMP stacking technique for seismic data, *The Leading Edge*, **27**(4), 462–467.

Ren, Y., Grecu, B., Stuart, G., Houseman, G., Hegedüs, E., & Group, S. C. P. W., 2013. Crustal structure of the carpathian-pannonian region from ambient noise tomography, *Geophys. J. Int.*, **195**(2), 1351–1369.

Ringler, A. T., Wilson, D. C., Storm, T., Marshall, B., Hutt, C. R., & Holland, A. A., 2016. Noise reduction in long-period seismograms by way of array summing, *Bull. seism. Soc. Am.*, **106**(5).

Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W. A., & Fehler, M. C., 2005. Surface wave tomography from microseisms in Southern California, *Geophys. Res. Lett.*, **32**(14), 1–4, L14311.

Sanchis, C. & Hanssen, A., 2011. Enhanced local correlation stacking method, *Geophysics*, **76**(3), V33–V45.

Schimmel, M. & Gallart, J., 2005. The inverse S-transform in filters with time-frequency localization, *IEEE Trans. Signal Process.*, **53**(11), 4417–4422.

Schimmel, M. & Gallart, J., 2007. Frequency-dependent phase coherence for noise suppression in seismic array data, *J. Geophys. Res.*, **112**(B4), B04303.

Schimmel, M. & Paulssen, H., 1997. Noise reduction and detection of weak, coherent signals through phase-weighted stacks, *Geophys. J. Int.*, **130**(2), 497–505.

Schimmel, M., Stutzmann, E., & Gallart, J., 2011. Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale, *Geophys. J. Int.*, **184**(1), 494–506.

Schimmel, M., Stutzmann, E., & Ventosa, S., 2017. Measuring group velocity in seismic noise correlation studies based on phase coherence and resampling strategies, *IEEE Trans. Geosci. Rem. Sens.*, **PP**(99), 1–8.

Selesnick, I. W., Baraniuk, R. G., & Kingsbury, N. G., 2005. The dual-tree complex wavelet transform, *IEEE Signal Process. Mag.*, **22**(6), 123–151.

Sens-Schönfelder, C., Snieder, R., & Stähler, S. C., 2015. The lack of equipartitioning in global body wave coda, *Geophys. Res. Lett.*, **42**(18), 7483–7489, 2015GL065108.

Shapiro, N. M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.*, **31**(7), 1–4, L07614.

Shapiro, N. M., Campillo, M., Stehly, L., & Ritzwoller, M. H., 2005. High-resolution surface-wave tomography from ambient seismic noise, *Science*, **307**(5715), 1615–1618.

Sinha, S., Routh, P. S., Anno, P. D., & Castagna, J. P., 2005. Spectral decomposition of seismic data with continuous-wavelet transform, *Geophysics*, **70**(6), P19–P25.

Snieder, R., 2004. Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase, *Phys. Rev. E*, **69**, 046610.

Stockwell, R. G., Mansinha, L., & Lowe, R. P., 1996. Localization of the complex spectrum: The S transform, *IEEE Trans. Signal Process.*, **44**(4), 998–1001.

Szanyi, G., Gráczer, Z., Győri, E., Kaláb, Z., & Lednická, M., 2016. Ambient seismic noise tomography of a high loess bank at Dunaszekcső (Hungary), *Pure Appl. Geophys.*, pp. 1–16.

Taner, M. T., Koehler, F., & Sheriff, R., 1979. Complex seismic trace analysis, *Geophysics*, **44**(6), 1041–1063.

Thurber, C., Zeng, X., Thomas, A., & Audet, P., 2014. Phase-weighted stacking applied to low-frequency earthquakes, *Bull. seism. Soc. Am.*, **104**(5), 2567–2572.

Ventosa, S., Simon, C., Schimmel, M., Danobeitia, J. J., & Manuel, A., 2008. The S-transform from a wavelet point of view, *IEEE Trans. Signal Process.*, **56**(7), 2771–2780.

Ventosa, S., Roy, S. L., Huard, I., Pica, A., Rabeson, H., Ricarte, P., & Duval, L., 2012. Adaptive multiple subtraction with wavelet-based complex unary wiener filters, *Geophysics*, **77**(6), V183–V192.

Vetterli, M. & Kovačević, J., 1995. *Wavelets and Subband Coding*, Prentice-Hall, Englewood Cliffs.

Wapenaar, K., Draganov, D., Snieder, R., Campman, X., & Verdel, A., 2010. Tutorial on seismic interferometry: Part 1 — basic principles and applications, *Geophysics*, **75**(5), 75A195–75A209.

Yang, Y., 2014. Application of teleseismic long-period surface waves from ambient noise in regional surface wave tomography: A case study in western USA, *Geophys. J. Int.*, **198**(3), 1644–1652.

Zeng, X. & Thurber, C., 2016. A graphics processing unit implementation for time-frequency phase-

weighted stacking, *Seismol. Res. Lett.*, **87**(2A), 358–362.

Zhao, K., Luo, Y., & Xie, J., 2017. Broad-band Rayleigh wave phase velocity maps (10–150 s) across the United States from ambient noise data, *Geophys. J. Int.*, **208**(2), 1265–1275.

LIST OF FIGURES

1 Attenuation of a linear-stacked signals as function of the SNR of the prestacked signal for (dashed lines) the phase stack with $v = 2$ and (solid lines) the unbiased phase coherence.

2 Extracted signal from 200 data sequences that contain a windowed chirp signal perturbed by white Gaussian noise using (black) linear stack, (blue) tf/ts-PWS ($Q = 5$) and (red) two-stage ts-PWS, firstly using the linear stack to produce 10 sequences and secondly using the unbiased ts-PWS. (a) Waveform of the test data and (b) of the chirp signal. (c) Misfit to the unperturbed chirp signal. (d) Extracted signals and (e) their differences (extracted minus original signal).

3 Convergence to the EGF of stacked ambient noise correlations (1-day long GNCC) of two pairs of GEOSCOPE stations with (black) linear stack, (blue) tf/ts-PWS and (red) two-stage ts-PWS (formed by a linear stack generating 10 sequences followed by the unbiased ts-PWS). Velocity waveforms extracted using, (a) and (b), all the data available (from about 2006 until 2016) and, (c) and (d), up to 500 days. All these figures use the same amplitude normalization, except ts-PWS results that are multiplied by 7. Convergence, (e) and (f), to the waveforms shown in (a) and (b), respectively.

4 (a) and (b) First and fourth Rayleigh wave trains extracted using all the data available shown in Fig. 3 with (black) linear stack, (blue) tf/ts-PWS and (red) two-stage ts-PWS. (c) and (d) Evolution of SNR with the number of correlations stacked. (e) and (f) Convergence to the waveforms extracted shown in (a) and (b), respectively.

5 Waveforms extracted by stacking ambient noise correlations (1-day long GNCC) from 20 pairs of GEOSCOPE stations using up to 500 days of data with (a) linear stack, (b) ts-PWS and (c) two-stage stack constructed using the unbiased ts-PWS. The normalization functions are applied to better observe later-arriving phases. Each station pairs is normalized to its maximum amplitude.

6 Waveforms extracted by stacking ambient noise correlations (1-day long GNCC) from 20 pairs of pairs of GEOSCOPE stations using all the data available (from about 2000 until 2016) with (a) linear stack, (b) ts-PWS and (c) two-stage stack constructed using the unbiased ts-PWS. The normalization functions are applied to better observe later-arriving phases. Each station pair is normalized to its maximum amplitude.

7 Normal modes using the same station pairs as Fig. 6 using (a) linear stack and (b) two-stage stack constructed using the unbiased ts-PWS. For each station, waveforms are normalized to the maximum amplitude on the both sides of the thick black line separately. In red, eigenfrequencies of the fundamental normal modes up to 5 mHz using the PREM model (Dziewonski & Anderson 1981).

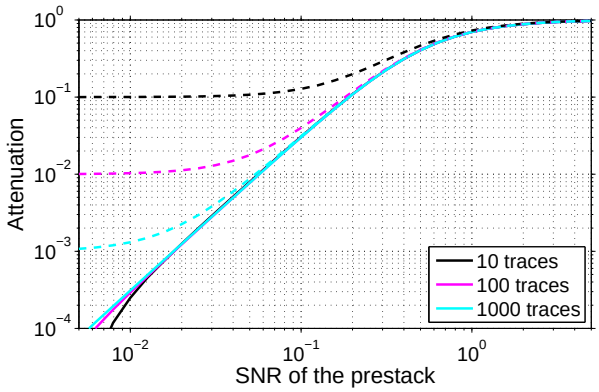

Figure 1. Attenuation of a linear-stacked signals as function of the SNR of the prestacked signal for (dashed lines) the phase stack with $v = 2$ and (solid lines) the unbiased phase coherence.

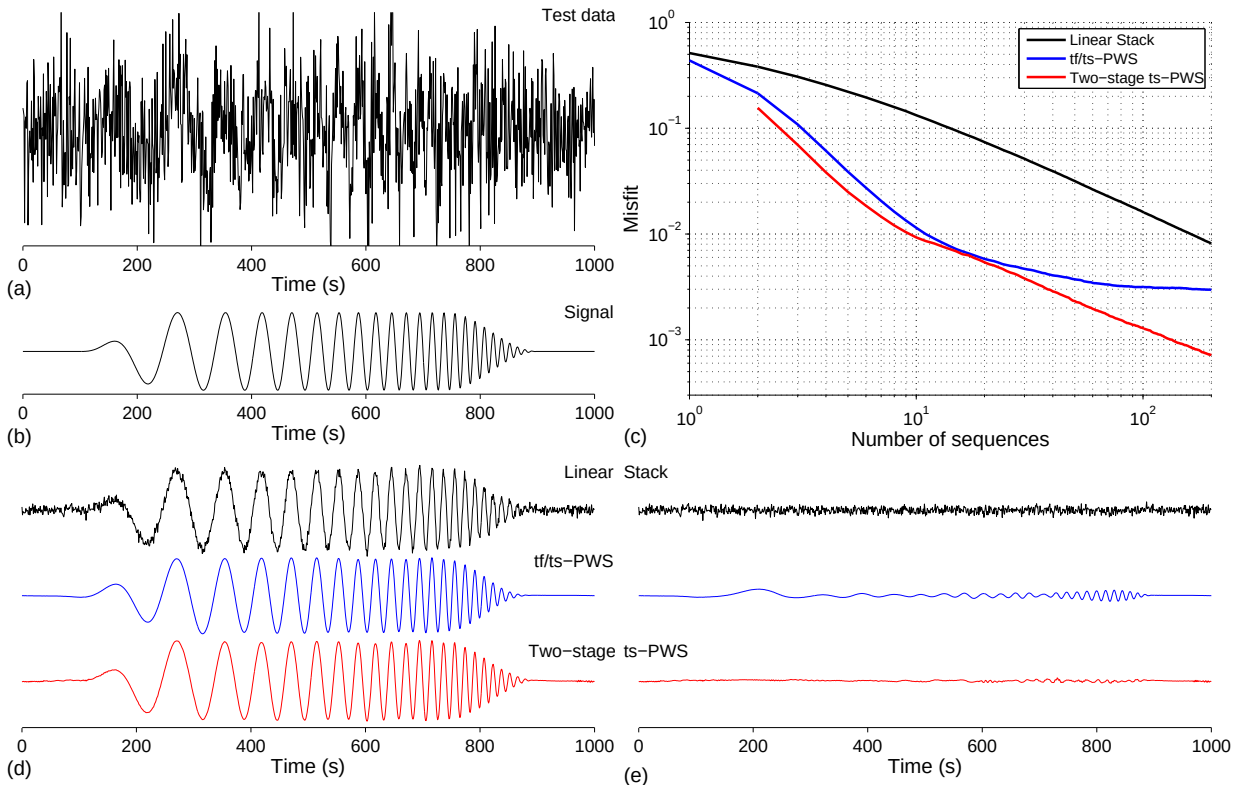

Figure 2. Extracted signal from 200 data sequences that contain a windowed chirp signal perturbed by white Gaussian noise using (black) linear stack, (blue) tf/ts-PWS ($Q = 5$) and (red) two-stage ts-PWS, firstly using the linear stack to produce 10 sequences and secondly using the unbiased ts-PWS. (a) Waveform of the test data and (b) of the chirp signal. (c) Misfit to the unperturbed chirp signal. (d) Extracted signals and (e) their differences (extracted minus original signal).

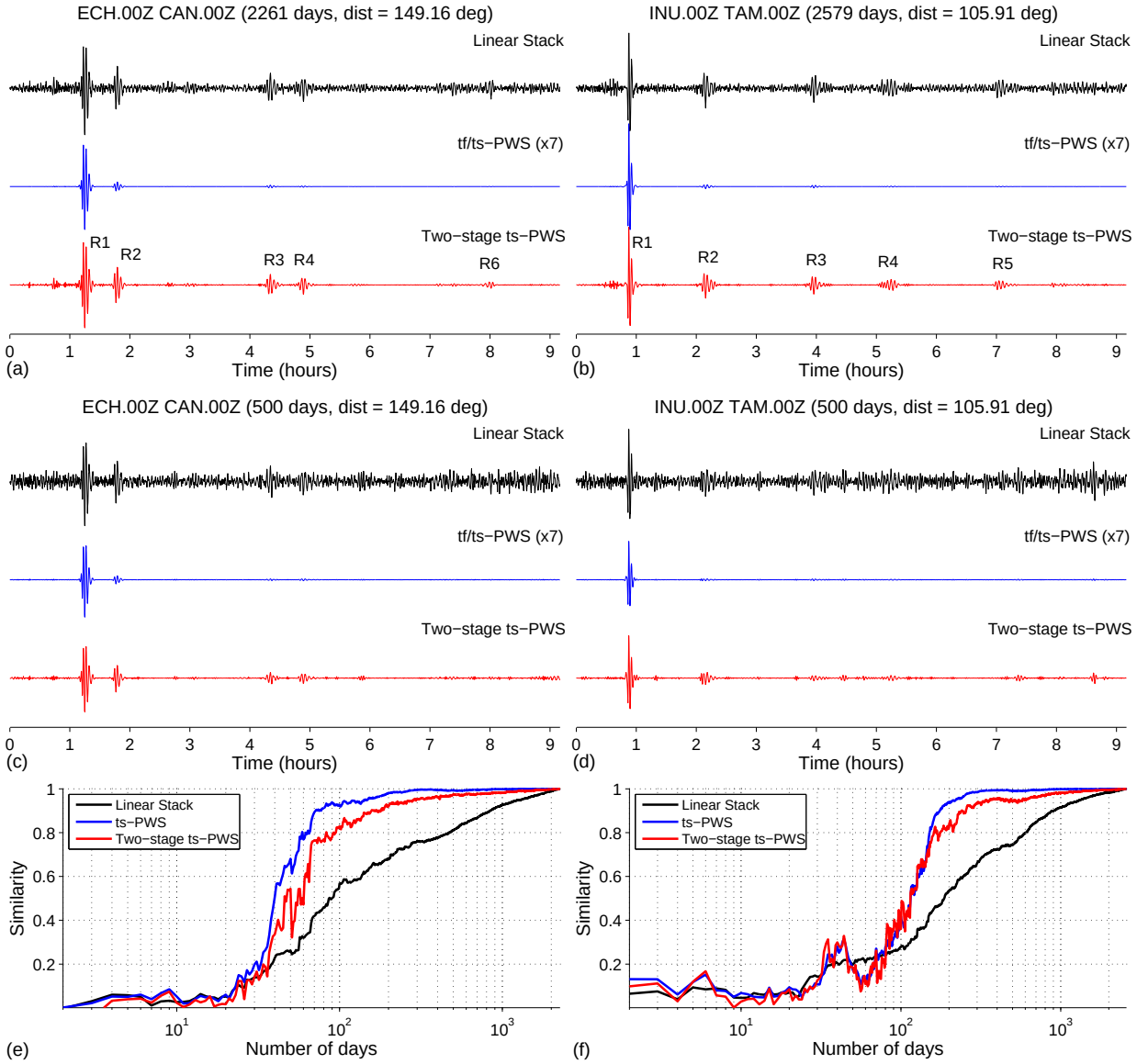

Figure 3. Convergence to the EGF of stacked ambient noise correlations (1-day long GNCC) of two pairs of GEOSCOPE stations with (black) linear stack, (blue) tf/ts-PWS and (red) two-stage ts-PWS (formed by a linear stack generating 10 sequences followed by the unbiased ts-PWS). Velocity waveforms extracted using, (a) and (b), all the data available (from about 2006 until 2016) and, (c) and (d), up to 500 days. All these figures use the same amplitude normalization, except ts-PWS results that are multiplied by 7. Convergence, (e) and (f), to the waveforms shown in (a) and (b), respectively.

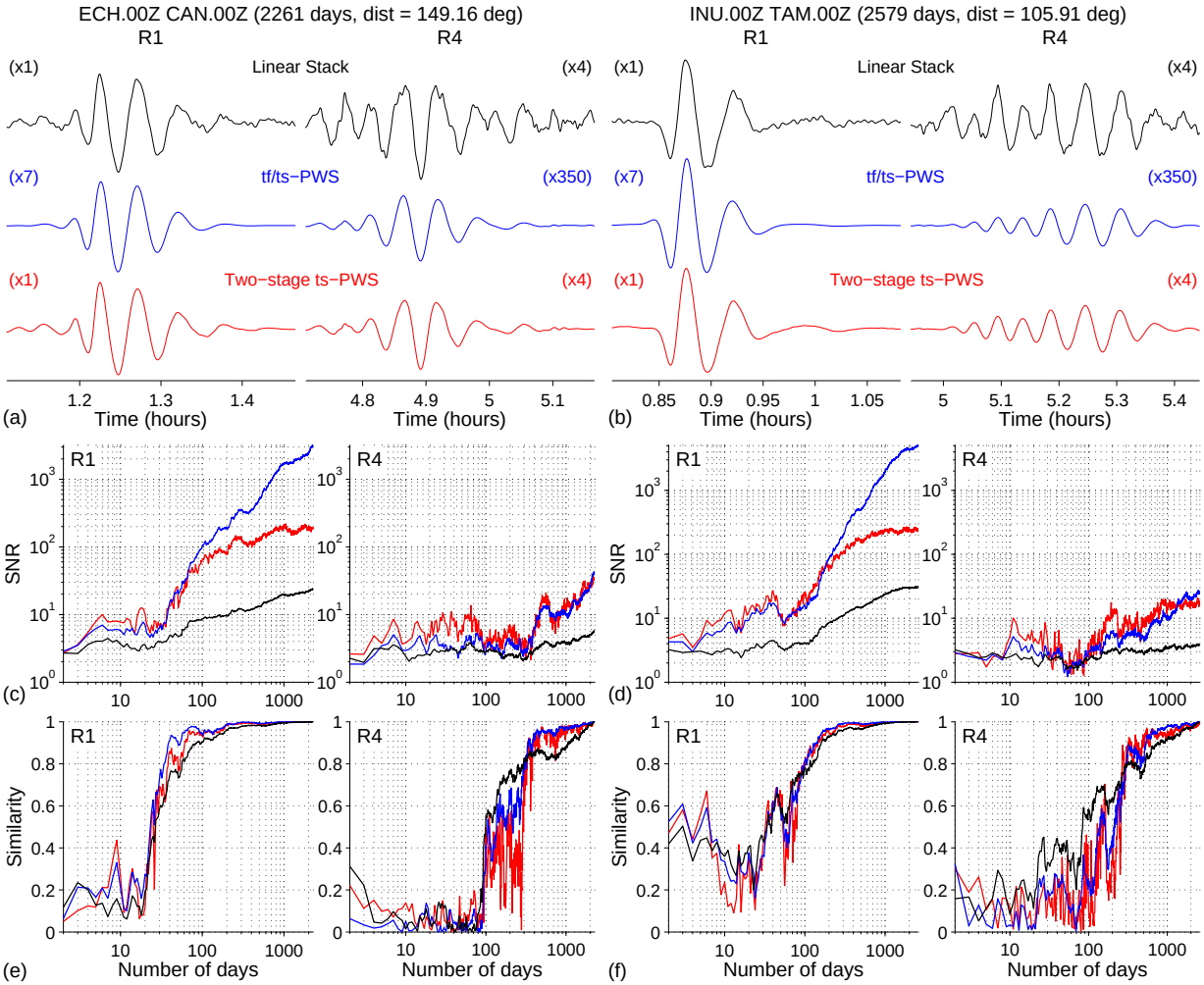

Figure 4. (a) and (b) First and fourth Rayleigh wave trains extracted using all the data available shown in Fig. 3 with (black) linear stack, (blue) tf/ts-PWS and (red) two-stage ts-PWS. (c) and (d) Evolution of SNR with the number of correlations stacked. (e) and (f) Convergence to the waveforms extracted shown in (a) and (b), respectively.

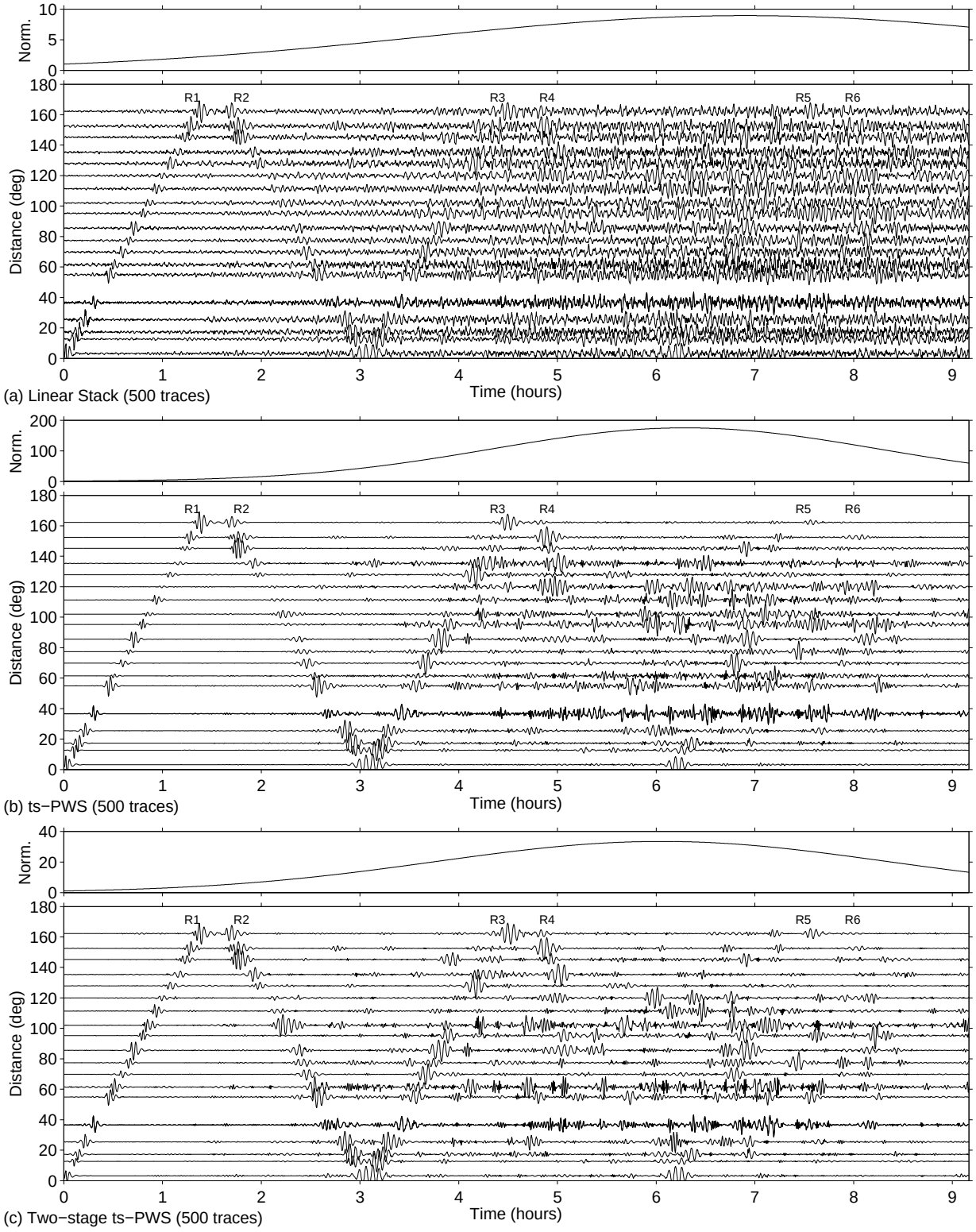

Figure 5. Waveforms extracted by stacking ambient noise correlations (1-day long GNCC) from 20 pairs of GEOSCOPE stations using up to 500 days of data with (a) linear stack, (b) ts-PWS and (c) two-stage stack constructed using the unbiased ts-PWS. The normalization functions are applied to better observe later-arriving phases. Each station pairs is normalized to its maximum amplitude.

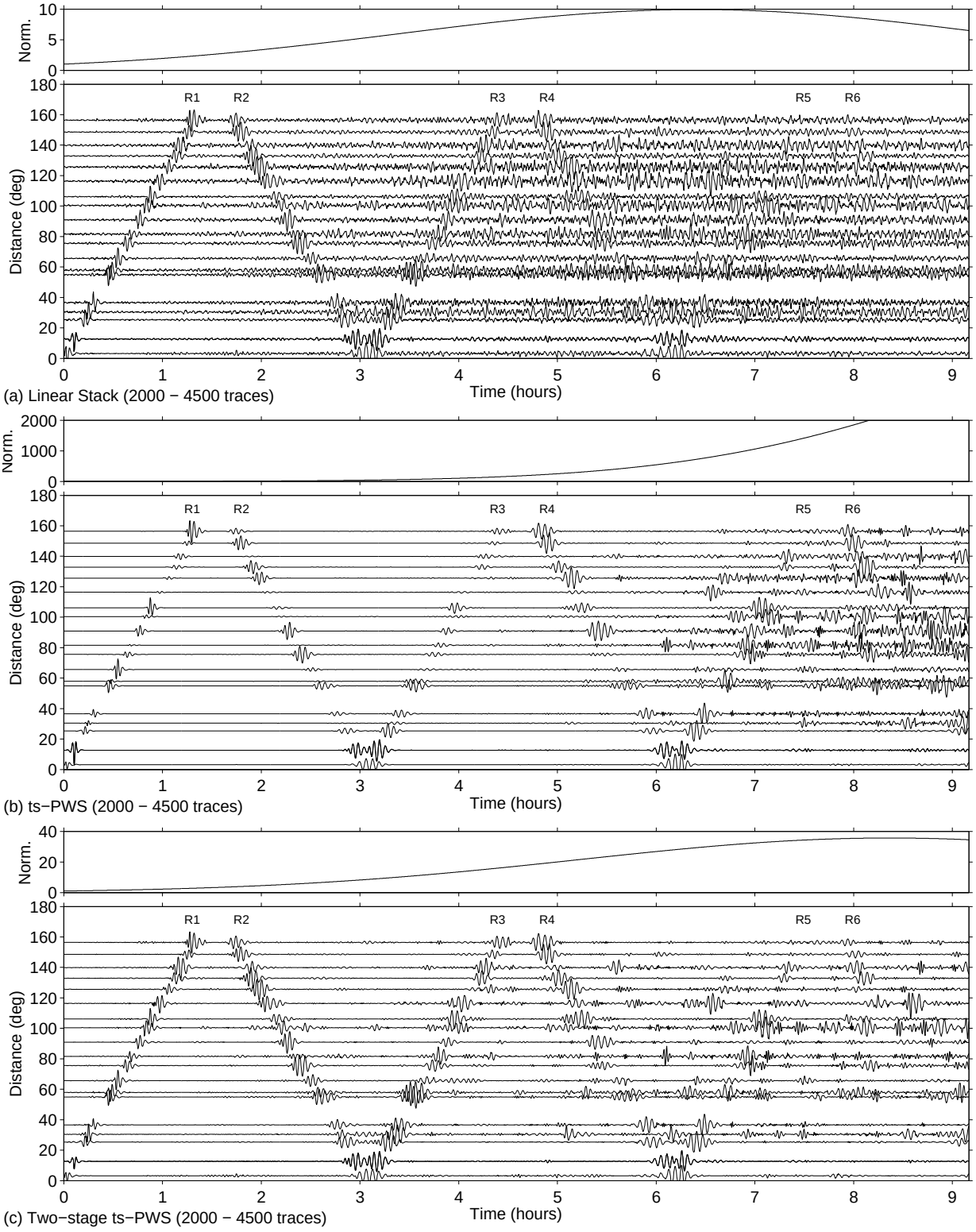

Figure 6. Waveforms extracted by stacking ambient noise correlations (1-day long GNCC) from 20 pairs of pairs of GEOSCOPE stations using all the data available (from about 2000 until 2016) with (a) linear stack, (b) ts-PWS and (c) two-stage stack constructed using the unbiased ts-PWS. The normalization functions are applied to better observe later-arriving phases. Each station pair is normalized to its maximum amplitude.

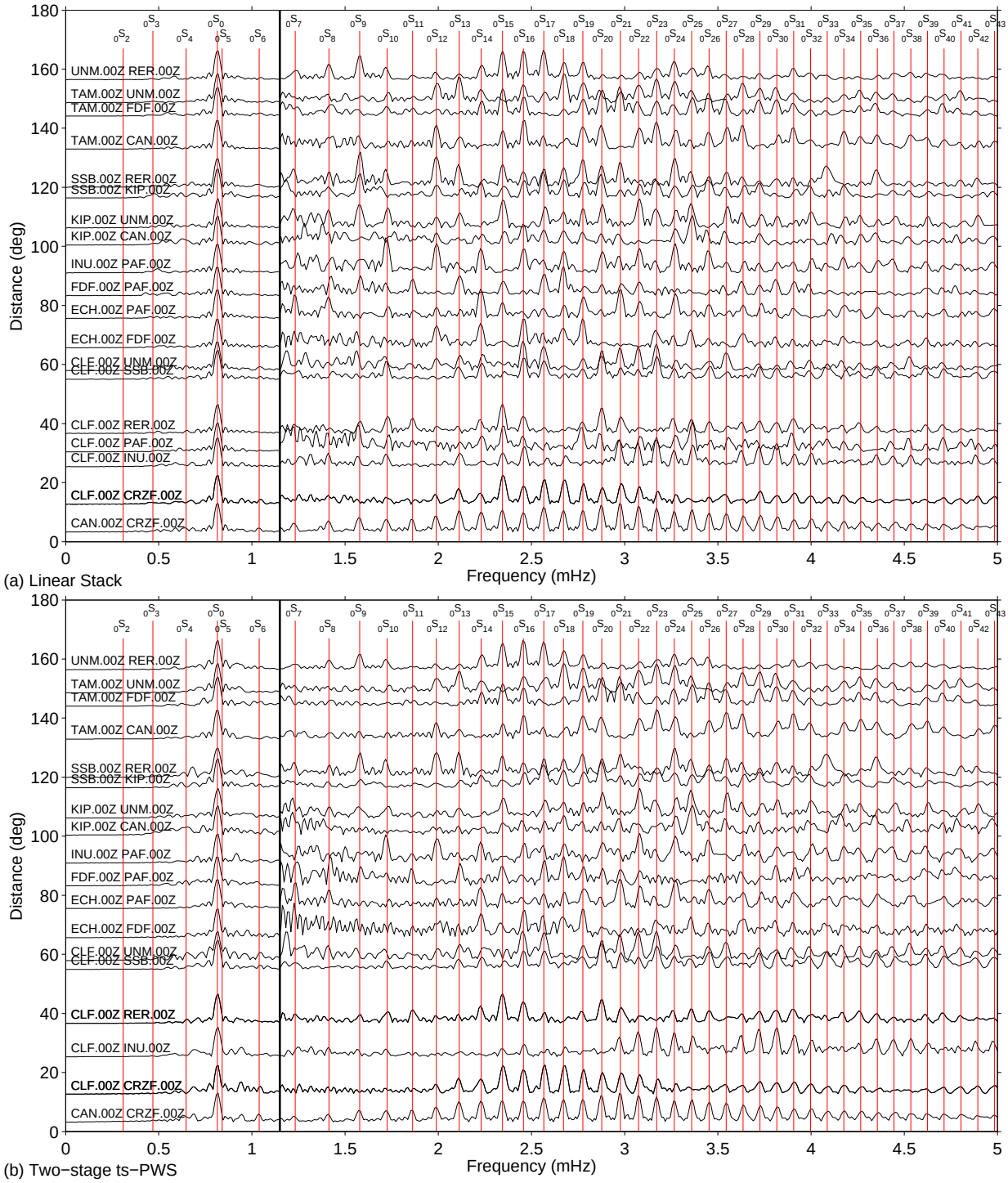

Figure 7. Normal modes using the same station pairs as Fig. 6 using (a) linear stack and (b) two-stage stack constructed using the unbiased ts-PWS. For each station, waveforms are normalized to the maximum amplitude on the both sides of the thick black line separately. In red, eigenfrequencies of the fundamental normal modes up to 5 mHz using the PREM model (Dziewonski & Anderson 1981).

APPENDIX A: IMPLEMENTING THE MULTIPLE FILTER TECHNIQUE WITH WAVELETS

We can identify the set of band-pass filters of constant Q quality factor in eq. (5) as a wavelet collection rewriting the product between the analytic signal $X_{\mathrm{a}}(\omega)$ and the band-pass filters in the frequency domain as a convolution of the real signal x with a modulated Gaussian window in the time domain,

$$x_{\mathrm{a}}(\tau,\omega_0) = \int_{-\infty}^{\infty} x(t) \frac{|\omega_0|}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\omega_0(\tau-t)}{\sigma}\right)^2} e^{i\omega_0(\tau-t)}\, dt. \tag{A.1}$$

When comparing the result with eq. (7), we obtain an "amplitude-normalized" Morlet wavelet defining the scale as $\lambda = \sigma/\omega_0$ and the central frequency of mother wavelet as $\xi_0 = \omega_0\lambda = \sigma$, that is,

$$\psi_{\tau,\lambda}(t) = \frac{1}{|\lambda|\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{t-\tau}{\lambda})^2} e^{i\xi_0(\frac{t-\tau}{\lambda})}, \tag{A.2}$$

or equivalently,

$$\psi_{\tau,\lambda}(t) = |\lambda|^{-1}\, \psi(\lambda^{-1}(t-\tau)) \tag{A.3}$$

with

$$\psi(t) = (2\pi)^{-1/2} e^{-t^2/2} e^{i\xi_0 t}. \tag{A.4}$$

Seismic studies traditionally use time-frequency transformations that preserve the amplitude of the signal in each frequency band, convenient for the analysis of oscillatory signals, instead of unitary transformations that conserve their energy. In this paper, we favor unitary transforms because they are better adapted for denoising (e.g., they lead to scale-independent thresholds for white Gaussian noise) and a default choice in the signal-processing literature. The difference between both is just cosmetic since equations (8) and (A.3) are equal up to a scaling factor of to $\lambda^{-1/2}$.

APPENDIX B: UNBIASED PHASE COHERENCE

To prove the coherence estimation of a phase stack with $v = 2$ and without time smoothing is biased, we rewrite the phase stack,

$$c^2_{\mathrm{ps}}[n] = \left| \frac{1}{K} \sum_{k=1}^{K} \frac{x_k[n]}{|x_k[n|} \right|^2 , \tag{B.1}$$

as

$$c^2_{\mathrm{ps}}[n] = \frac{1}{K^2} \sum_{i=1}^{K} \sum_{j=1}^{K} \frac{x_i[n]}{|x_i[n|} \frac{\overline{x_j[n]}}{|x_j[n]|} \tag{B.2}$$

by using the property that the square modulus of a sum equals the product of a sum by its conjugate. Then, we separate the elements where $i = j$ from the others,

$$c^2_{\mathrm{ps}}[n] = \frac{1}{K^2} \left(2 \sum_{i=1}^{K} \sum_{j>i}^{K} \frac{x_i[n]}{|x_i[n|} \frac{\overline{x_j[n]}}{|x_j[n]|} + \sum_{i=1}^{K} \frac{x_i[n]\overline{x_i[n]}}{|x_i[n|^2} \right) \tag{B.3}$$

and notice that the second term is constant,

$$c^2_{\mathrm{ps}}[n] = \frac{2}{K^2} \sum_{i=1}^{K} \sum_{j>i}^{K} \frac{x_i[n]}{|x_i[n|} \frac{\overline{x_j[n]}}{|x_j[n]|} + \frac{1}{K}. \tag{B.4}$$

In this form, it is clear that the first term is zero in mean if the signals $x_i[n]$ and $x_j[n]$ are fully dissimilar and $(K-1)/K$ if they are fully similar. The second term is consequently the bias of the phase stack with $v = 2$ for non-coherent signals.

Coherence is also measured using geometrically-normalized cross-correlations (GNCC),

$$c_{\mathrm{xcorr}}[n] = \frac{2}{K(K-1)} \sum_{i=1}^{K} \sum_{j>i}^{K} \frac{r_{ij}[n]}{\sqrt{r_{ii}[n] r_{jj}[n]}} \tag{B.5}$$

where $r_{ij}[n]$ is a local correlation at zero-lag between the data sequences x_i and x_j. These estimator is coherence unbiased but allows for negative values of coherence. In the particular case of using a single sample in this correlation, $r_{ij}[n] = x_i[n]\overline{x_j[n]}$, the GNCC and the phase-stack coherence estimators are related by

$$c^2_{\mathrm{ps}} = \frac{(K-1)c_{\mathrm{xcorr}} + 1}{K}. \tag{B.6}$$

Both estimators progressively diverge as the window length increases due to the different treatment of the amplitude of the time-sequences.

Using this relation we construct an unbiased phase coherence estimator as

$$c_{\mathrm{upc}}^2 = \frac{K c_{\mathrm{ps}}^2 - 1}{K - 1} \tag{B.7}$$

by correcting the bias of the phase stack with $v = 2$.