

Towards the Processing of Large Data Volumes with Phase Cross-Correlation

by Sergi Ventosa, Martin Schimmel, and Eleonore Stutzmann

ABSTRACT

Interstation correlation is the basic operation in seismic noise and coda-wave interferometry for signal extraction in imaging and monitoring applications. Conventional cross-correlations evaluate the similarity between two signals along lag time, and they are efficiently computed by the fast Fourier transform (FFT), valuable to manage the large data volumes that ambient noise applications demand. The phase cross-correlation (PCC) method contributes to increase convergence, a key issue in seismic ambient noise imaging and monitoring; however, it is much more computationally demanding. PCC evaluates similarity by subtracting the modulus of the sum and difference of the instantaneous phase of two signals. We introduce solutions to dramatically reduce the high-computational cost of PCC. We show that PCC can be rewritten as a complex cross-correlation and computed by the FFT when the moduli are raised to the power of 2, and we demonstrate PCC can improve waveform coherence and increase convergence compared with the default processing flow of 1-bit amplitude normalization and standard cross-correlation. Moreover, we develop a graphics processing unit implementation to accelerate computations when using powers other than 2 and particularly when using the power of 1. Finally, we extract Rayleigh- and body-wave signals from many years of data from seismic stations distributed worldwide using PCC without a significant increase in computational cost compared with conventional cross-correlation.

INTRODUCTION

Cross-correlation is a ubiquitous method in seismic data processing important for picking, empirical Green's function extraction, ambient noise monitoring, waveform comparison, and signal, event, and pattern detection or identification among others. Phase cross-correlation (PCC) was conceived to better detect weak P -to- S conversions at the upper-mantle discontinuities (Schimmel, 1999). In this work, PCC is introduced to estimate travel-time differences by searching for the lag time that maximizes an amplitude-unbiased misfit function between the pilot and converted waveforms, their phase

coherency. Phase coherency is useful for detecting weak signals in the coda of stronger waves because it effectively considers a larger portion of the waveforms than conventional correlation-based methods that consider mainly the most energetic components. Later, PCC was found useful to improve waveform coherence and convergence of interstation correlations, thanks to its amplitude-unbiased property (Schimmel *et al.*, 2011). Essentially, PCC can be useful for applications demanding more information on the waveform.

Interstation correlations of seismic ambient noise and scattered coda waves complement direct arrivals from earthquakes in surface-wave tomography, monitoring, and Earth's deep interior studies in regions otherwise poorly illuminated. Convergence to the empirical Green's function or a robust seismic signal extraction increases compared with the direct correlation of raw seismic records through customized processing flows that improve the balance of noise sources and reduce the influence of anomalous signals. These processing flows compute many correlations on relatively short-data sequences. The preprocessing of the raw data may include instrument response correction, anomalous signal rejection, spectral whitening, and time-domain normalization (Bensen *et al.*, 2007), and once the correlations are computed, weighted stacking and denoising (Baig *et al.*, 2009; Schimmel *et al.*, 2011; Cheng *et al.*, 2015; Moreau *et al.*, 2017). Finally, subsequent measurements of group and phase velocity dispersion are done in surface-wave tomography applications (e.g., Haned *et al.*, 2016; Obermann *et al.*, 2016; Schippkus *et al.*, 2018), of velocity and structure variations in monitoring (e.g., Hadziioannou *et al.*, 2011; D'Hour *et al.*, 2016; Takano *et al.*, 2017; Sánchez-Pastor *et al.*, 2018), or of travel times of body waves from ambient noise or earthquake coda in Earth's deep interior studies (e.g., Nishida, 2013; Boué *et al.*, 2014; Phạm *et al.*, 2018; Tkalčić and Phạm, 2018).

Spectral whitening or band-pass filtering to the frequency band of interest followed by 1-bit amplitude normalization and geometrical-normalized cross-correlation (GNCC) is the processing flow most often used, in which GNCC is the conventional cross-correlation divided by the standard deviation of the signals. The 1-bit cross-correlation method applies correlation on the sign of the input sequences (e.g., Bensen *et al.*, 2007; Cupillard *et al.*, 2011). This instantaneous amplitude normalization reduces the influence of poorly distributed high-energy features such as earthquakes, but it distorts

waveforms and slows down convergence (Schimmel *et al.*, 2018). The PCC method measures similarity between two analytic signals of unitary amplitude, that is, two instantaneous phases. Schimmel *et al.* (2011) show that PCC is a less aggressive alternative to 1-bit cross-correlation that may help to improve waveform coherence and increase convergence. This has been confirmed, for example, in D'Hour *et al.* (2016) for monitoring of hydromechanical changes, in Taylor *et al.* (2016) and Becker and Knapmeyer-Endrun (2018) for retrieving *P*-wave Moho reflections from noise autocorrelations, in Haned *et al.* (2016) for extracting group velocity observations for global tomography using seismic hum, in Schimmel *et al.* (2018) for extracting Rayleigh-wave group velocities using phase autocorrelation, and in Romero and Schimmel (2018) for mapping the basement of the Ebro basin in Spain using shallow subsurface reflections.

The computational cost of the signal processing methods conventionally used in the preprocessing of the raw time series is small compared with the correlation, weighted stacking, and denoising operations subsequently applied. This is mainly because preprocessing is done per station, and correlation and denoising operations are done per station pair. ObsPy (Krischer *et al.*, 2015) and MSNoise (Lecocq *et al.*, 2014) are open-source packages for Python that are excellent for proof-of-concept studies and often sufficient for preprocessing. Compiled programming language solutions using the central processing unit (CPU) or the graphics processing unit (GPU) are however necessary to apply the most demanding operations to large data volumes in a reasonable amount of time. Several tools have been introduced recently for improving the computational efficiency of key operations along interstation correlation processing flows. Zeng and Thurber (2016) present a GPU implementation of the time–frequency phase-weighted stack (tf-PWS; Schimmel *et al.*, 2011), which is up to 20 times faster than equivalent CPU implementations using the fast Fourier transform (FFT). Ventosa *et al.* (2017) reduce computational cost of tf-PWS by introducing the time-scale phase-weighted stack (ts-PWS) that uses the wavelet transform for the time–frequency representation, and propose unbiased phase coherence and two-stage ts-PWS to improve convergence speed and quality of the extracted signals. The unbiased phase coherence estimator contributes to increase noncoherent noise attenuation because it equals to 0 if the signals are totally incoherent. Further, the two-stage ts-PWS reduces signal attenuation by stacking linearly the sequences in a few groups to improve signal-to-noise ratio before applying the actual ts-PWS method. Fichtner *et al.* (2017) present Mirmex, a complete tool for the processing of seismic ambient noise data that uses GPUs to compute the most demanding operations including direct implementations of the PCC and GNCC methods.

Conventional cross-correlations, and in particular GNCC, are computed directly only if a very small number of lag times are required; otherwise, they are computed much faster using the FFT. The FFT is very efficiently implemented on the CPU by the fastest Fourier transform in the west (FFTW) library

(Frigo and Johnson, 2005) and can be further accelerated on the GPU using, for example, the compute unified device architecture (CUDA) platform. Despite the advantages of PCC compared with the 1-bit cross-correlation, the application of this method to the large data volumes that current seismic applications demand is limited by its high-computational cost.

In this article, we introduce two solutions to speed up the computation of PCC by a factor of 100 in conventional settings, so that PCC becomes only two to four times slower than conventional cross-correlation. We reduce operation complexity of the PCC when using modulus raised to the power of 2, and we present optimized CPU and GPU codes (see Data and Resources) to accelerate computations when using other powers, in particular, using the default power of 1 (PCC1) as recommended in Schimmel (1999) and Schimmel *et al.* (2011). These codes focus exclusively on the cross-correlation operation and allow for an easy integration with current processing flows.

CALCULATION OF THE PCC

The PCC methods evaluate similarity between two analytic signals s_1 and s_2 at the lag time τ by subtracting the modulus of the sum and difference of their unitary phasors. To obtain these unitary phasors, we represent real time-series u as analytic signals s using the Hilbert transform \mathcal{H} ,

$$s(t) = u(t) + \mathcal{H}\{u(t)\} = a(t)e^{i\theta(t)}, \quad (1)$$

in which a is the envelope, θ is the instantaneous phase, and t is the time. In this signal representation, we obtain unitary phasors containing the instantaneous phase information by dividing s by its modulus,

$$e^{i\theta(t)} = \frac{s(t)}{|s(t)|}. \quad (2)$$

Then, PCC uses the unitary phasors of the analytic signals $s_1(t) = a_1(t)e^{i\theta_1(t)}$ and $s_2(t) = a_2(t)e^{i\theta_2(t)}$ to compute the similarity between them.

The PCC of the discrete sequences $s_1[n]$ and $s_2[n]$ sampled at $t = t_0 + nT$, in which n is the sample number and T is the sampling period, can be written as

$$c_{\text{pcc}}[m] = \frac{1}{N} \sum_{n=0}^{N-1} \left| \frac{e^{i\theta_1[n+m]} + e^{i\theta_2[n]}}{2} \right|^\nu - \left| \frac{e^{i\theta_1[n+m]} - e^{i\theta_2[n]}}{2} \right|^\nu, \quad (3)$$

in which m is the lag time number (from the lag time $\tau = mT$), ν is the power of PCC, and N is the number of samples correlated from t_0 to $t_0 + (N - 1)T$.

Analyzing equation (3) in detail, we can see that if the signals are totally correlated, the first term in the sum equals to 1 and the second to 0, and thus $c_{\text{pcc}} = 1$; conversely, if they are totally anticorrelated, the first term is 0 and the second is 1,

and thus $c_{\text{pcc}} = -1$; and if they are totally not correlated, the two terms are equal and $c_{\text{pcc}} = 0$. In this context, the power factor ν controls the sharpness of the transition from totally correlated and anticorrelated to not correlated values. Default values for ν are 1 and 2. Setting $\nu = 1$ increases absolute correlation values and setting $\nu > 1$, and particularly $\nu = 2$, increases signal-to-noise ratio.

PCC with Power Factor 2

PCC as written in equation (3) simplifies for $\nu = 2$ by developing the square modulus. The key operation to be implemented when computing the PCC with $\nu = 2$ (PCC2) at one lag time is

$$c = \sum_i |a_i + b_i|^2 - |a_i - b_i|^2 = \|a + b\|^2 - \|a - b\|^2, \quad (4)$$

in which a_i and b_i are complex numbers, a and b are complex vectors, and $\|\cdot\|$ denotes the Euclidean norm. The squared norm of the sum and difference of two elements in a complex Hilbert space can be developed as functions of inner products as

$$\begin{aligned} \|a + b\|^2 &= \langle a + b, a + b \rangle \\ &= \langle a, a \rangle + \langle a, b \rangle + \langle b, a \rangle + \langle b, b \rangle, \end{aligned} \quad (5)$$

$$\begin{aligned} \|a - b\|^2 &= \langle a - b, a - b \rangle \\ &= \langle a, a \rangle - \langle a, b \rangle - \langle b, a \rangle + \langle b, b \rangle. \end{aligned} \quad (6)$$

Then, the difference of the two squared norms reduces to a single inner product,

$$\|a + b\|^2 - \|a - b\|^2 = 4\text{Re}\langle a, b \rangle, \quad (7)$$

which is the real part of the polarization identity for complex Hilbert spaces (Young, 1988). When we apply this result to equation (3), PCC2 simplifies to the real part of the complex cross-correlation between the phase vectors,

$$c_{\text{pcc2}}[m] = \text{Re} \left\{ \frac{1}{N} \sum_{n=0}^{N-1} e^{i\theta_1[n+m]} e^{-i\theta_2[n]} \right\}, \quad (8)$$

or equivalently, using $s_1[n]$ and $s_2[n]$,

$$c_{\text{pcc2}}[m] = \text{Re} \left\{ \frac{1}{N} \sum_{n=0}^{N-1} \frac{s_1[n+m] \overline{s_2[n]}}{|s_1[n+m]| |s_2[n]|} \right\}, \quad (9)$$

in which $\overline{s_2[n]}$ denotes the complex conjugate of $s_2[n]$.

Writing PCC2 as a conventional complex cross-correlation dramatically reduces its operation complexity and allows for further acceleration by the FFT. Overall, the computational cost of PCC2 and GNCC is now comparable, and the main differences are due to the required analytic signal

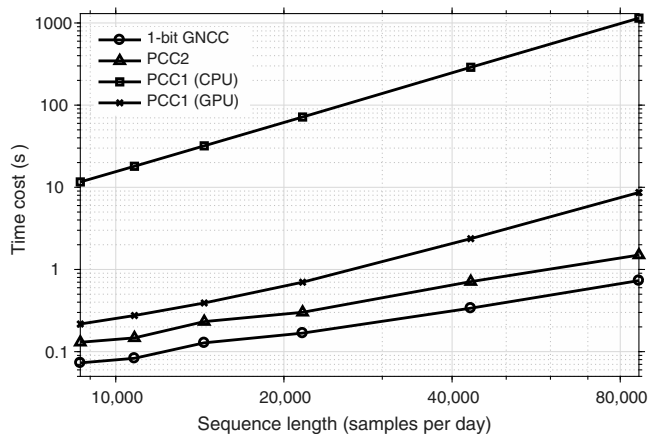
representation for the PCC. Finally, it is important to note that equation (9) defines PCC2 if s_1 and s_2 are analytic signals and defines 1-bit GNCC if s_1 and s_2 are real signals. Because the phase contains much more information about the signal than its sign, it is reasonable to say that PCC can improve waveform coherence and increase convergence compared with 1-bit GNCC in most applications.

IMPLEMENTATION

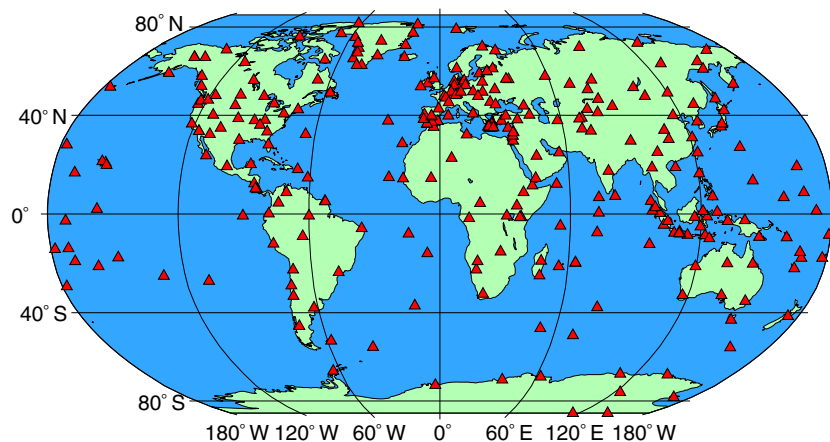
To our knowledge, the PCC equation with a power factor ν different from 2 does not simplify and the operation complexity cannot be reduced by algorithms such as the FFT; however, it can be accelerated on the GPU. GPUs are optimized for computing the same operations on large datasets in parallel and CPUs for long sequential series of operations. Thus, GPUs are optimal to solve simple but computationally intensive operations on relatively small datasets such as PCC, in which the wall-clock time reduction of the computations largely compensates for the time penalty of copying data to and from the GPU memory. Conversely, 1-bit GNCC and PCC2 are conventionally more efficiently implemented in the CPU using a few FFT operations than in the GPU, in which the acceleration of the FFT computation is not worth the extra time required to move data between the CPU and GPU memories. To compute the FFT in the CPU, we use the most recent iteration of the FFTW library (Frigo and Johnson, 2005), a very efficient implementation that generally outperforms direct implementations of the cross-correlation. Direct solutions are typically faster only if the number of lag times is very small, less than about 1% of the number of samples of the time series.

The computation of PCC between two analytic sequences of N samples at M lag times for the default value of $\nu = 1$ involves $2NM$ sums and modulus of complex numbers and $(2N - 1)M$ sums of real numbers. The most demanding operation among them is by far the computation of the modulus of complex numbers, mainly due to the square root involved. This operation is common in many physical problems and efficiently computed on the GPU.

We develop CPU only and GPU enhanced codes of 1-bit GNCC and PCC (see Data and Resources) that compute all the correlations of a station pair at once. We use OpenMP to compute these correlations in parallel on the CPU, instead of parallelizing each correlation individually, to reduce scheduling overhead and improve scaling with the number of cores. To accelerate the PCC1 calculation when a GPU is available, we compute the analytic representation of the data sequences on the CPU using OpenMP and, in parallel, the subsequent operations on the GPU. To speed up computations of the actual correlations, equation (3), we distribute the computations of one correlation across many GPU-thread blocks. Each thread in a block computes the correlation between 256 samples of the data sequences at one lag time using copies of the portions of the sequences needed on the small shared memory (much faster and with lower latency than the global memory) available to each block. Moreover, we use 16 streams



▲ **Figure 1.** Computational cost of cross-correlating 649 days of data from the GEOSCOPE stations CAN and ECH using 1-bit geometrical-normalized cross-correlation (GNCC), phase cross-correlation using the power of 2 (PCC2), and phase cross-correlation using 1 (PCC1) with a maximum correlation lag of $\pm 12,000$ s. CPU, central processing unit; GPU, graphics processing unit.



▲ **Figure 2.** Geographic distribution of the stations (red triangles) used in this test. The color version of this figure is available only in the electronic edition.

to compute several correlations in parallel, enough to fully exploit the cores available on modern GPUs. In this configuration, the GPU calculations limit wall-clock time and the computation of the analytic signals on the CPU reduces GPU usage and exploits resources that would be otherwise wasted.

EXAMPLE

We illustrate the computational performance of 1-bit GNCC and PCC using continuous data on the vertical component. The preprocessing flow consists of removing the mean and trends, correcting the instrument response to produce ground

velocity, band-pass filtering from 20 to 33 mHz (30–50 s period), dividing data into one day-long nonoverlapping sequences, and rejecting sequences with high-seismic activity either containing strong earthquakes or having a mean energy much higher than the average. We specifically reject raw-data sequences with absolute maxima greater than the median plus 100 times the median absolute deviation (MAD). Then, we reject preprocessed sequences with a standard deviation greater than the median of the standard deviation of the sequences recorded by a given station plus 500 times its MAD. The test environment is a workstation equipped with a 4-core 64-bit Intel Haswell CPU at 3.50 GHz (i5-4690K) with 16 GB of DDR3 at 1866 MHz, an Nvidia GTX 1060 graphics card, and a solid-state drive disk to store the data.

Figure 1 shows the wall-clock time cost of applying 1-bit GNCC, PCC2, and PCC1 with a maximum lag time of $\pm 12,000$ s to 649 day-long pairs of seismograms from the years 2016 and 2017 recorded by the GEOSCOPE stations CAN and ECH and downsampled to 1, 2, 4, 8, and 10 s per sample. We see that the time cost of 1-bit GNCC and

PCC2 has the same trend, mostly determined by FFT computations the operation complexity of which is $\mathcal{O}(N \log_2 N)$, independent of the number of lag times M . PCC2 is about two times slower than 1-bit GNCC because the PCC2 implementation uses seven single-precision FFT operations (four to compute the two analytic signals and three for the correlation) whereas 1-bit GNCC uses only three. In comparison, PCC1 on the CPU with an operation complexity of $\mathcal{O}(NM)$ is much slower and its computational cost increases faster. The GPU enhanced code of PCC1 substantially reduces the wall-clock time of the CPU-only code and makes PCC1 only about four times slower than 1-bit GNCC in conventional settings. For example, PCC1 takes 71.65 s on the CPU and 0.70 s on the GPU for the sequence length of 21,600 samples per day, which is 4.17 times slower than 1-bit GNCC and 1.79 times slower than PCC2, and 1.9 s on a single Nvidia Tesla K20 GPU

from the S-CAPAD cluster of the Institut de Physique du Globe de Paris (IPGP).

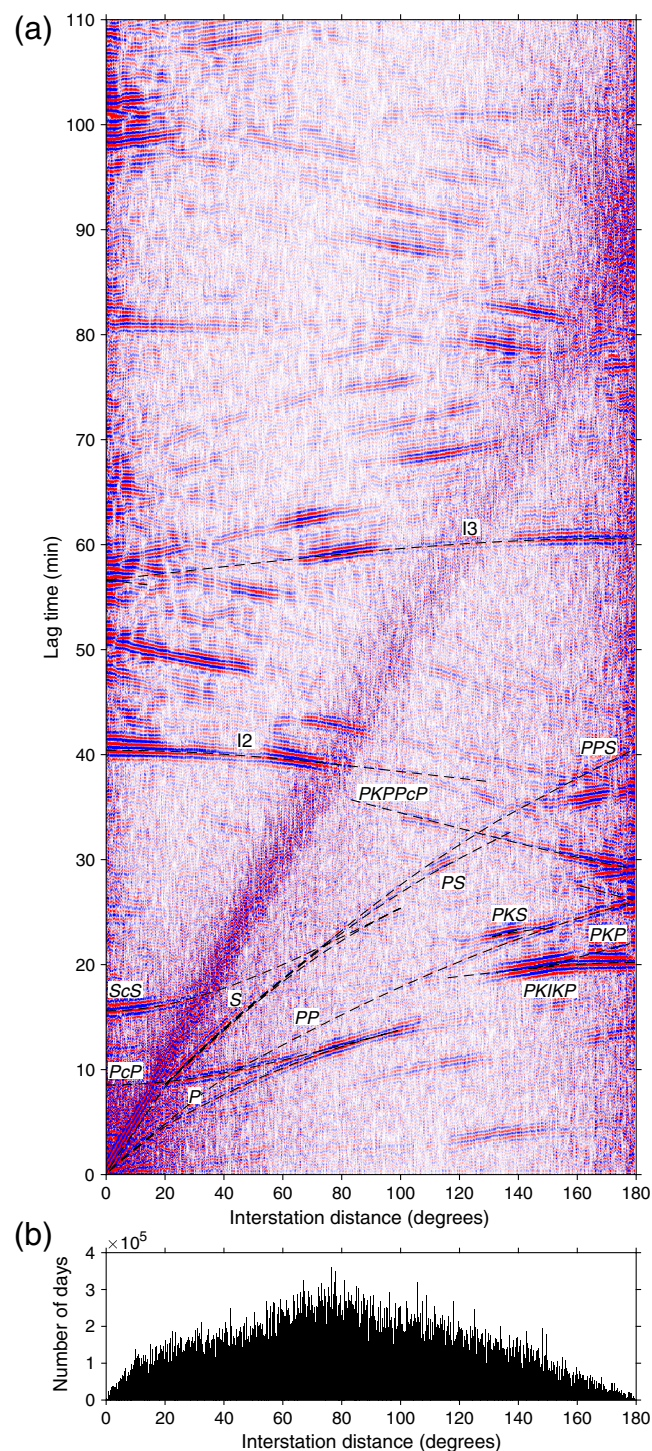
In the next example, we selected data from the years 1990 to 2017 recorded by 299 broadband seismic stations distributed worldwide as shown in Figure 2. In particular, we selected all stations from the global and regional networks of GEOFON (GE), GEOSCOPE (G), Global Seismographic Network (GSN-IRIS/IDA: II, GSN-IRIS/USGS: IU), and New China Digital Seismograph Network (IC). In addition, we included some stations from Geoscience Australia (AU), the Canadian National Seismograph Network (CN), the Danish Seismological Network (DK), the Pacific21 (PS) network, and the Canfranc Underground Laboratory (LC) to improve global Earth coverage.

We computed 1-bit GNCC, PCC2, and PCC1 of day-long sequences downsampled to sampling periods of 4 s per sample. Then, we rejected correlations with an anomalous standard deviation and stack the rest linearly. Under these settings, the faster algorithms would have spent most of the time on reading and writing operations. To accelerate these operations, we gathered all data from a given station in one file, and we used a ramdisk for temporal results such as the daily correlations. Using the aforementioned workstation, the correlation and stack operations require 939 min for 1-bit GNCC, 1291 min for PCC2, and 2380 min for PCC1 on the GPU.

Figure 3a shows 110 million day-long interstation correlations using PCC2 from 39,656 station pairs band-pass filtered from 20 to 33 mHz as a function of distance stacked on bins of 0.25° using only data from days with low-seismic activity. The large amount of day-long correlations stacked on each bin (Fig. 3b) is sufficient for the previous correlation methods to converge to the empirical Green's function. In applications in which few data are available, PCC usually contributes to improve signal quality and increases convergence (Schimmel *et al.*, 2011). The amplitude of the signals extracted is meaningful, in spite of the aggressive amplitude normalization of 1-bit GNCC or PCC but strongly influenced by the distribution of noise sources (e.g., Snieder, 2004). Hence, to extract accurate information from amplitudes, it is important to design processing flows that minimize this bias as, for example, Bowden *et al.* (2018) show in the construction of surface-wave attenuation maps across the United States.

The strongest signal we observe in Figure 3 is the Rayleigh wave starting at the bottom-left corner and progressively vanishing at larger distance due to heterogeneities in the crust and the top upper mantle. Much weaker, we see a great variety of body waves interacting with the mantle and the core such as the *P*, *PcP*, *S*, *ScS*, *PKP*, *PKIKP* (I), or *PKIKPPKIKP* (I2), and of noncausal phases (Pham *et al.*, 2018) such as *PKP-PcP* preceding the direct *P*-wave arrival in the range 120° – 140° by about 4 min. The number of station pairs per bin from 5° to 162° varies from about 20 to 100 pairs; however, it reduces to a few pairs close to 180° . At these interstation distances, some of these signals may be interpreted as spurious arrivals from dominant sources located far from the great circle from some station pairs (Retailleau *et al.*, 2017).

The signals we observe in this figure are produced by a mix of ambient noise sources and coda waves from moderate and smaller earthquakes, as can be seen from the anomalous high amplitude of the *ScS*, *PKP*, and *PKIKP* waves. Removing anomalously correlations and sequences with high-seismic activity removed contributions from large earthquakes the poor distribution of which in azimuth may introduce bias in, for example, velocity dispersion or arrival-time observations. More effective processing flows for removing coda waves can be found in Nishida (2013) and Boué *et al.* (2014). Pham *et al.* (2018) observe some of the signals we show in Figure 3 using coda from large earthquakes instead and explain the origin and generation mechanism of causal and noncausal phases. Further,



▲ **Figure 3.** (a) 110 million day-long interstation correlations from 39,656 station pairs sorted by distance and stacked in bins of 0.25° in the period band of 30–50 s using PCC2 and linear stacking on days with low-seismic activity. (b) Day-long correlations stacked per bin. The color version of this figure is available only in the electronic edition.

in Figure 3, we observe signals with low-apparent velocity, such as *P*, *S*, *PS*, and *PPS*, mainly produced by ambient noise sources.


CONCLUSIONS

We introduced approaches that enable the application of PCC on large data volumes using conventional workstations. The stand-alone code implementing PCC and 1-bit GNCC is designed to be easily integrated with current processing flows, whether it is executed directly or the main functions implementing each method are called from other codes.

We have shown that PCC simplifies to a conventional cross-correlation of unitary phasors when using the power factor 2. This cross-correlation equation is fully analogous to 1-bit GNCC with the single difference that 1-bit GNCC considers the sign of the real signals and PCC2 the phase factor of analytic signals. Because phase carries more information than sign, most applications using 1-bit GNCC should see a gain in waveform coherence and increased convergence when applying PCC2.

By writing PCC2 as a complex cross-correlation, we achieved a lower operation complexity and, in conventional settings, about 100 times faster code than direct implementations. This reduces the computational cost of PCC2 to about two times that of the 1-bit GNCC method using only the CPU. When using other powers, GPU acceleration allowed us to apply PCC1 to the large data volume needed to extract body-wave signals from interstation correlations using a conventional workstation and a low-cost GPU. In general, these computational cost reductions are important to efficiently find and tune the best adapted methods to a given problem and, especially, to process the large data volumes that modern seismic ambient noise imaging and monitoring applications demand with PCC.

DATA AND RESOURCES

The data used in this study were obtained from the Incorporated Research Institutions of Seismology (IRIS) and the European Integrated Data Archive (EIDA). The central processing unit (CPU) and graphics processing unit (GPU) accelerated versions of code implementing 1-bit geometrical-normalized cross-correlation (GNCC) and phase cross-correlation (PCC) are open source under the GNU Lesser General Public License (LGPL) version 3 (v.3) license and available at <https://github.com/sergiventosa/FastPCC>. The fastest Fourier transform in the west (FFTW) is available at <http://www.fftw.org> and the compute unified device architecture (CUDA) documentation and installation guides from <https://docs.nvidia.com/cuda/index.html>. All websites were last accessed on April 2019. 

ACKNOWLEDGMENTS

The authors thank the editors and the anonymous reviewers for improving the article. This work was supported by the MISTERIOS project CGL2013-48601-C2-1-R and the ANR MIMOSA project ANR-14-CE01-0012. The authors also acknowledge support from discussions within TIDES

COST Actions ES1401. Numerical computations were partly performed on the S-CAPAD platform, Institut du Physique du Globe de Paris (IPGP), France.

REFERENCES

- Baig, A., M. Campillo, and F. Brenguier (2009). Denoising seismic noise cross correlations, *J. Geophys. Res.* **114**, no. B8, 1–12, doi: [10.1029/2008JB006085](https://doi.org/10.1029/2008JB006085).
- Becker, G., and B. Knapmeyer-Endrun (2018). Crustal thickness across the Trans-European Suture Zone from ambient noise autocorrelations, *Geophys. J. Int.* **212**, no. 2, 1237–1254, doi: [10.1093/gji/ggx485](https://doi.org/10.1093/gji/ggx485).
- 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](https://doi.org/10.1111/j.1365-246x.2007.03374.x).
- Boué, P., P. Poli, M. Campillo, and P. Roux (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, doi: [10.1016/j.epsl.2014.01.047](https://doi.org/10.1016/j.epsl.2014.01.047).
- Bowden, D. C., V. C. Tsai, and F.-C. Lin (2018). Amplification and attenuation across USArray using ambient noise wavefront tracking, *J. Geophys. Res.* **122**, no. 12, 10,086–10,101, doi: [10.1002/2017JB014804](https://doi.org/10.1002/2017JB014804).
- Cheng, F., J. Xia, Y. Xu, Z. Xu, and Y. Pan (2015). A new passive seismic method based on seismic interferometry and multichannel analysis of surface waves, *J. Appl. Geophys.* **117**, 126–135, doi: [10.1016/j.jappgeo.2015.04.005](https://doi.org/10.1016/j.jappgeo.2015.04.005).
- Cupillard, P., L. Stehly, and B. Romanowicz (2011). The one-bit noise correlation: A theory based on the concepts of coherent and incoherent noise, *Geophys. J. Int.* **184**, no. 3, 1397–1414, doi: [10.1111/j.1365-246x.2010.04923.x](https://doi.org/10.1111/j.1365-246x.2010.04923.x).
- D'Hour, V., M. Schimmel, A. F. Nascimento, J. M. Ferreira, and H. C. Lima 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](https://doi.org/10.1007/s00024-015-1156-0).
- Fichtner, A., L. Ermert, and A. Gokhberg (2017). Seismic noise correlation on heterogeneous supercomputers, *Seismol. Res. Lett.* **88**, no. 4, 1141–1145, doi: [10.1785/0220170043](https://doi.org/10.1785/0220170043).
- Frigo, M., and S. G. Johnson (2005). The design and implementation of FFTW3, *Proc. IEEE* **93**, no. 2, 216–231, doi: [10.1109/JPROC.2004.840301](https://doi.org/10.1109/JPROC.2004.840301), special issue on “Program Generation, Optimization, and Platform Adaptation”.
- Hadziioannou, C., E. Larose, A. Baig, P. Roux, and M. Campillo (2011). Improving temporal resolution in ambient noise monitoring of seismic wave speed, *J. Geophys. Res.* **116**, no. B7, 1–10, doi: [10.1029/2011JB008200](https://doi.org/10.1029/2011JB008200).
- Haned, A., E. Stutzmann, M. Schimmel, S. Kiselev, A. Davaille, and A. Yelles-Chaouche (2016). Global tomography using seismic hum, *Geophys. J. Int.* **204**, no. 2, 1222–1236, doi: [10.1093/gji/ggv516](https://doi.org/10.1093/gji/ggv516).
- Krischer, L., T. Megies, R. Barsch, M. Beyreuther, T. Lecocq, C. Caudron, and J. Wassermann (2015). ObsPy: A bridge for seismology into the scientific Python ecosystem, *Comput. Sci. Disc.* **8**, no. 1, 014003, doi: [10.1088/1749-4699/8/1/014003](https://doi.org/10.1088/1749-4699/8/1/014003).
- Lecocq, T., C. Caudron, and F. Brenguier (2014). MSNoise, a Python package for monitoring seismic velocity changes using ambient seismic noise, *Seismol. Res. Lett.* **85**, no. 3, 715–726, doi: [10.1785/0220130073](https://doi.org/10.1785/0220130073).
- Moreau, L., L. Stehly, P. Boué, Y. Lu, E. Larose, and M. Campillo (2017). Improving ambient noise correlation functions with an SVD-based Wiener filter, *Geophys. J. Int.* **211**, no. 1, 418–426, doi: [10.1093/gji/ggx306](https://doi.org/10.1093/gji/ggx306).

- Nishida, K. (2013). Global propagation of body waves revealed by cross-correlation analysis of seismic hum, *Geophys. Res. Lett.* **40**, no. 9, 1691–1696, doi: [10.1002/grl.50269](https://doi.org/10.1002/grl.50269).
- Obermann, A., M. Lupi, A. Mordret, S. S. Jakobsdóttir, and S. A. Miller (2016). 3D-ambient noise Rayleigh wave tomography of Snæfellsjökull volcano, Iceland, *J. Volcanol. Geoth. Res.* **317**, 42–52, doi: [10.1016/j.jvolgeores.2016.02.013](https://doi.org/10.1016/j.jvolgeores.2016.02.013).
- Phạm, T.-S., H. Tkalčić, M. Sambridge, and B. L. N. Kennett (2018). Earth's correlation wavefield: Late coda correlation, *Geophys. Res. Lett.* **45**, no. 7, 3035–3042, doi: [10.1002/2018GL077244](https://doi.org/10.1002/2018GL077244).
- Retailleau, L., P. Boué, L. Stehly, and M. Campillo (2017). Locating microseism sources using spurious arrivals in intercontinental noise correlations, *J. Geophys. Res.* **122**, no. 10, 8107–8120, doi: [10.1002/2017JB014593](https://doi.org/10.1002/2017JB014593).
- Romero, P., and M. Schimmel (2018). Mapping the basement of the Ebro basin in Spain with seismic ambient noise autocorrelations, *J. Geophys. Res.* **123**, no. 6, 5052–5067, doi: [10.1029/2018JB015498](https://doi.org/10.1029/2018JB015498).
- Sánchez-Pastor, P., A. Obermann, and M. Schimmel (2018). Detecting and locating precursory signals during the 2011 El Hierro, Canary Islands, submarine eruption, *Geophys. Res. Lett.* **45**, no. 19, 10,288–10,297, doi: [10.1029/2018GL079550](https://doi.org/10.1029/2018GL079550).
- 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](https://doi.org/10.1111/j.1365-246X.2010.04861.x).
- Schimmel, M., E. Stutzmann, and S. Ventosa (2018). Low-frequency ambient noise autocorrelations: Waveforms and normal modes, *Seismol. Res. Lett.* **89**, no. 4, 1488–1496, doi: [10.1785/0220180027](https://doi.org/10.1785/0220180027).
- Schippkus, S., D. Zigone, G. Bokelmann, and the AlpArray Working Group (2018). Ambient-noise tomography of the wider Vienna basin region, *Geophys. J. Int.* **215**, no. 1, 102–117, doi: [10.1093/gji/ggy259](https://doi.org/10.1093/gji/ggy259).
- 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, doi: [10.1103/PhysRevE.69.046610](https://doi.org/10.1103/PhysRevE.69.046610).
- Takano, T., T. Nishimura, and H. Nakahara (2017). Seismic velocity changes concentrated at the shallow structure as inferred from correlation analyses of ambient noise during volcano deformation at Izu-Oshima, Japan, *J. Geophys. Res.* **122**, no. 8, 6721–6736, doi: [10.1002/2017JB014340](https://doi.org/10.1002/2017JB014340).
- Taylor, G., S. Rost, and G. Houseman (2016). Crustal imaging across the North Anatolian Fault Zone from the auto-correlation of ambient seismic noise, *Geophys. Res. Lett.* **43**, no. 6, 2502–2509, doi: [10.1002/2016GL067715](https://doi.org/10.1002/2016GL067715).
- Tkalčić, H., and T.-S. Phạm (2018). Shear properties of Earth's inner core constrained by a detection of J waves in global correlation wavefield, *Science* **362**, no. 6412, 329–332, doi: [10.1126/science.aau7649](https://doi.org/10.1126/science.aau7649).
- 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](https://doi.org/10.1093/gji/ggx284).
- Young, N. (1988). *An Introduction to Hilbert Space*, Cambridge University Press, Cambridge, United Kingdom.
- Zeng, X., and C. Thurber (2016). A graphics processing unit implementation for time-frequency phase-weighted stacking, *Seismol. Res. Lett.* **87**, no. 2A, 358–362, doi: [10.1785/0220150192](https://doi.org/10.1785/0220150192).

Sergi Ventosa

Martin Schimmel

Institute of Earth Sciences Jaume Almera–CSIC

C/Lluís Solé i Sabaris s/n

E-08028 Barcelona, Spain

sventosa@ictja.csic.es

sergiventosa@hotmail.com

Eleonore Stutzmann

Institut de Physique du Globe de Paris, Sorbonne Paris Cité

Université Paris Diderot

UMR 7154 CNRS

F-75005 Paris, France

Published Online 1 May 2019