

RESEARCH ARTICLE

10.1029/2018JB015937

Key Points:

- Symmetry of higher-order cross correlations arises primarily from signal processing rather than the diffuse character of correlation coda
- Prestack C3 accelerates convergence of the retrieved Green's functions and shows promise to extend the results to higher frequency

Supporting Information:

- Supporting Information S1

Correspondence to:

Y. Sheng,
yixiao2@stanford.edu

Citation:

Sheng, Y., Nakata, N., & Beroza, G. C. (2018). On the nature of higher-order ambient seismic field correlations. *Journal of Geophysical Research: Solid Earth*, 123. <https://doi.org/10.1029/2018JB015937>

Received 10 APR 2018

Accepted 2 SEP 2018

Accepted article online 5 SEP 2018

On the Nature of Higher-Order Ambient Seismic Field Correlations

Yixiao Sheng¹ , Nori Nakata² , and Gregory C. Beroza¹ 

¹Department of Geophysics, Stanford University, Stanford, CA, USA, ²ConocoPhillips School of Geology and Geophysics, University of Oklahoma, Norman, OK, USA

Abstract We find that higher-order ambient seismic field correlations, specifically C3 (correlation of the coda of correlations) of long-duration data sets from Southern California are symmetric despite the fact that the background wavefield is strongly directional. We demonstrate that this symmetry arises primarily from the structure of the C3 time series analysis rather than the diffuse character of the coda. For that reason, the resulting ambient-field Green's function symmetry may not be diagnostic of correlation quality. We propose an alternative prestack processing scheme that constructs C3 correlations from each C1 (cross correlation) coda time window, but stacks the results only at the final step of processing. Prestack C3 leads to faster convergence and extracts arrivals at higher frequencies than standard poststack processing.

1. Introduction

Multiple theoretical justifications connect time-averaged cross correlations of the ambient seismic field to receiver-to-receiver Green's functions (Lobkis & Weaver, 2001; Snieder, 2004; Snieder, Sheiman, & Calvert, 2006; Wapenaar, 2004; Wapenaar et al., 2006). Since then time-averaged noise correlation functions have been used for a series of applications ranging from inversion for Earth structure (e.g., Lin et al., 2008; Prieto et al., 2009; Sabra et al., 2005; Shapiro et al., 2005; Stehly et al., 2009), earthquake ground motion prediction (Denolle et al., 2013; Denolle et al., 2014; Sheng et al., 2017; Viens et al., 2014) to image time-dependent crustal changes (Brenuier et al., 2008; Sens-Schönfelder & Wegler, 2006). Despite these impressive results, unevenly distributed noise sources continue to bias travel time and amplitude measurements from correlation functions (Halliday & Curtis, 2008; Stehly & Boué, 2017; Tsai, 2009, 2011; Weaver et al., 2009). Uneven noise source distributions typically lead to highly asymmetric correlation functions, which contradict the physical constraint of Green's function's reciprocity. For this reason the symmetry of correlation functions has come to be used as a measure of their reliability.

The observation of a multiply scattered equipartitioned wavefield in earthquake coda (Hennino et al., 2001) led Campillo and Paul (2003) to the successful retrieval of the Green's function between seismic stations using the correlation of the coda from a group of earthquakes. Stehly et al. (2008) recognized that the coda of correlation-based Green's functions could itself be considered a diffuse wavefield and used the correlation of the coda of the ambient field correlations, which they referred to as C3, to retrieve Green's functions. Their interpretation of the improved symmetry relative to the C1 Green's functions was that it arose from the diffuse character of the waves comprising the C1 coda. Froment et al. (2011) further tested this idea and iterated correlation functions to higher orders. They demonstrated that the higher-order correlation of the coda of the C3 functions (C5) also yields symmetric Green's functions and suggested one could combine complementary information from different correlation functions to improve the results. The C3 method has important advantages and has seen increased application (Ma & Beroza, 2012; Sheng et al., 2017; Spica et al., 2016; Zhang & Yang, 2013). Though we refer higher-order correlation as C3, it is worth noting that there exists another branch of higher-order interferometry, named Source-Receiver Interferometry (SRI). SRI utilizes the energy injected from the surrounding boundaries to construct the Green's function between a source and a receiver (e.g., Curtis & Halliday, 2010; Halliday & Curtis, 2010).

In contrast to earlier studies, we find that the symmetry of C3 Green's functions does not arise primarily from the diffuse character of coda wavefield. We demonstrate through both numerical simulation and straightforward mathematical development that even for a nondiffuse wavefield, standard C3 signal

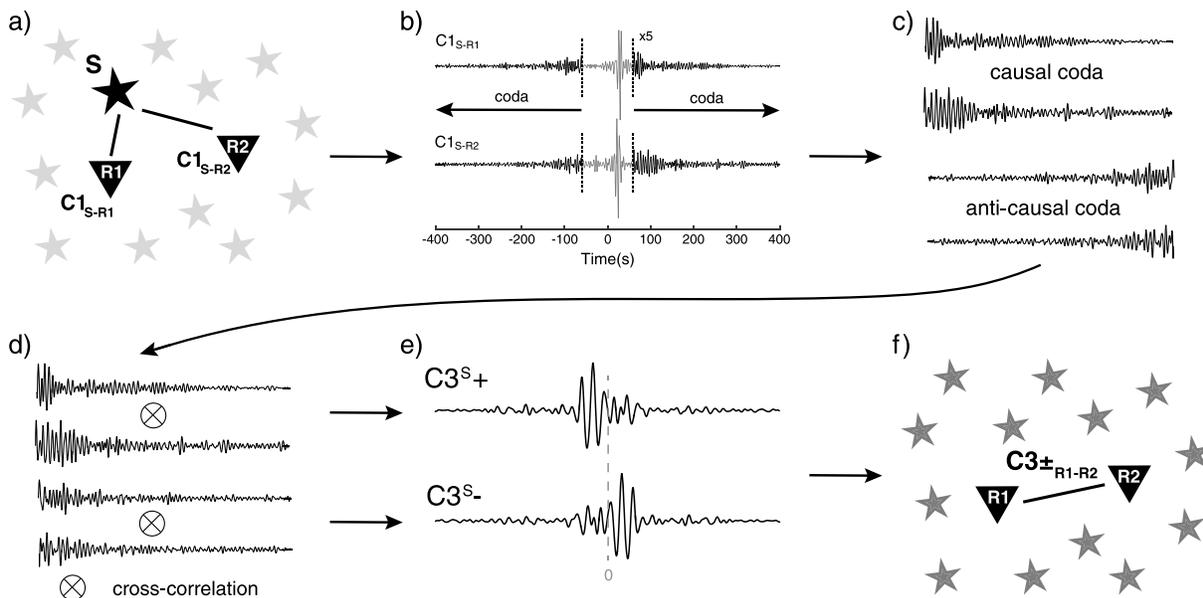


Figure 1. Workflow for constructing C3+ and C3-. (a) geometry. (b) C1 functions, which are first calculated between a virtual source S and target receivers R1 and R2. (c) Coda wave, starting at twice the Rayleigh wave travel time, is selected on both sides of the obtained C1 functions. (d) Anticausal coda is then time-reversed before performing the correlation of the correlation coda, which results in one ensemble of C3+ and C3- from the virtual source S, shown in (e). (f) Repeating the process and averaging the results over all the available virtual sources yield the final C3+ and C3- functions between the target receivers.

processing results in symmetric correlations. Although symmetric, such correlations can still be biased given an unfavorable source distribution. While we find that the C3 functions constructed only from the causal (C3+) or anticausal (C3-) part of the C1 coda leads to a more symmetric result than C1, this is a weak effect for the data we have studied. Because it better reflects the true degree to which the coda is equipartitioned, we suggest that the symmetry of the C3+ or C3- correlations individually provides a better measure of correlation function reliability.

Finally, we propose a modified C3 method, which we refer to as *prestack C3*, that improves the signal-to-noise ratio (SNR) of the correlation. As the name suggests, the technique constructs C3 correlations from each C1 coda time window but stacks the results at the end of the processing. This differs from previous C3 implementations that first stack over all time windows to construct stable C1 correlation functions and then correlates the coda of those results. Prestack C3 is straightforward to implement, and it significantly accelerates the convergence of the correlation functions, but at a cost of greater storage requirements for intermediate results. We also find that prestack C3 recovers coherent arrivals to higher frequencies than standard processing.

2. Symmetry of Higher-Order Correlation Functions

2.1. Asymmetry of C3+ and C3-

Figure 1 illustrates the C3 workflow, and typical results are displayed in Figure 2. By the standard C3 method, to construct the C3 function between stations CI.CLT and CI.BBR, we first calculate the long-term averaged C1 functions from each virtual source (broadband station in the CI network) to the two target stations. The cross correlation is performed with all the available data recorded in year 2015, with instrumental response removed. We divide the data into 30-min-long time series and discard those with spikes larger than 8 times the standard deviation of the window. We use coherence interferometry, so that each trace is normalized in the frequency domain during the calculation. Sheng et al. (2017) showed that using all three components of the virtual source station leads to improved results; however, in this paper we only use the vertical component. We select the coda by following previous studies (e.g., Froment et al., 2011; Sheng et al., 2017; Stehly et al., 2008), where a 1,200-s-long time window is picked, starting at twice the Rayleigh wave travel time. We then separately correlate the coda of the causal part (wave propagating toward either CI.CLT or

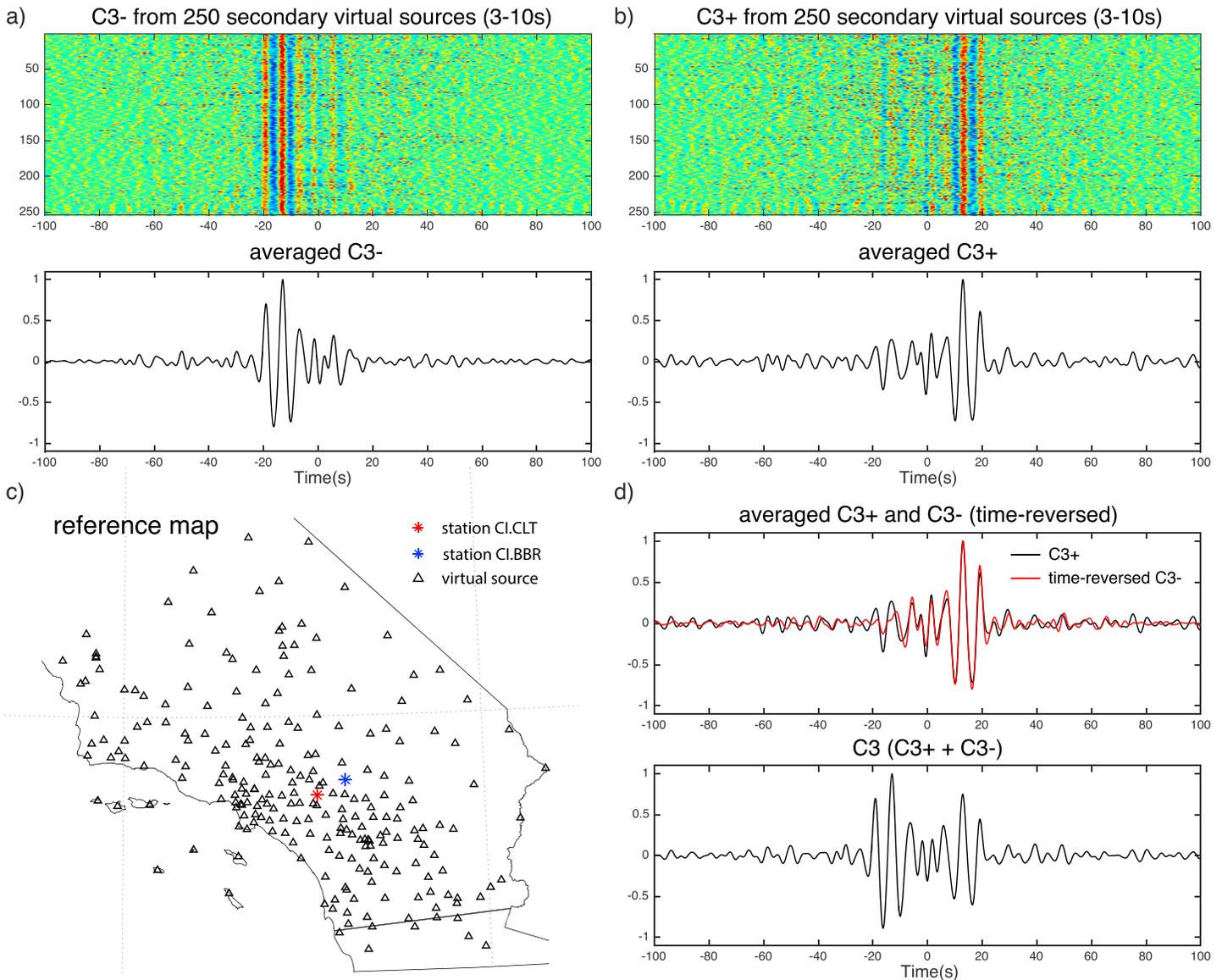


Figure 2. C3 functions between station pair CI.CLT and CI.BBR. (a) C3[−] functions from different virtual source stations are displayed in the upper panel with their average shown below. (b) Same as (a) but for C3⁺ functions. (c) Reference map shows location of target stations, indicated by stars, with virtual source stations marked by triangles. (d) The upper panel shows that the normalized time-reversed C3[−] function matches the C3⁺ function. Note that the sum is largely symmetric.

CI.BBR) and the time-reversed anticausal part (wave propagating outward from either CI.CLT or CI.BBR) of the C1 functions, such that for each virtual source, we have two measurements of the associated C3 functions: C3⁺ (using causal coda) and C3[−] (using time-reversed anticausal coda). To complete the procedure, we sum the C3⁺ and C3[−] functions over all the virtual sources.

Analogous to earthquake coda, the coda of the correlation functions should be more diffuse than the ballistic parts, and the symmetry of C3 has been assumed to result from the diffuse character of the coda wavefield (e.g., Froment et al., 2011; Sheng et al., 2017; Stehly et al., 2008). If symmetry arises from the diffuse wavefield, we would expect both C3⁺ and C3[−] to be symmetric; however, the observations in Figure 2 contradict these expectations. Neither the C3⁺ nor C3[−] function is nearly as symmetric as their sum. In particular, we find that the expected arrival does not depend on the direction of the corresponding virtual source. The averaged C3⁺ function over all the virtual sources acts like a mirror image of the averaged C3[−] function, as illustrated in Figure 2d. Not surprisingly, the sum of these two yields a largely symmetric C3 Green's function.

2.2. Numerical Simulations for a Homogeneous Medium

In the previous subsection we find that the symmetry of the Green's functions extracted by the C3 method arises from signal processing rather than a diffuse coda wavefield. We now examine this further through acoustic numerical simulations using the Madagascar package (Fomel et al., 2013). We start with a lossless homogeneous velocity model, generating only nondiffuse direct waves to examine whether symmetry is preserved in higher-order correlation without diffuse wavefield. Figure 3 gives two examples. Case I represents the situation where the source locates inside the stationary phase zone of two target stations, between which the Green's function is to be extracted. Since there is no scattered wave in such homogeneous velocity model, there is no coda either, and we use the ballistic part of C1 functions to perform the second-order correlation, for example, C2 (correlation of the correlation). C2+ denotes the C2 function constructed from the causal part of correlation functions, while C2− denotes that for the time-reversed anticausal part. In Figure 3a, the C2+ and C2− functions appear on the opposite sides of the time axis and their sum yields a symmetric correlation function, whose phase matches that of the C1 function between the target stations. We use this C1 function to mark the expected travel time throughout our numerical simulations, though it is known to suffer from a $\pi/4$ phase shift. The amplitude of the C2 function is stronger in the anticausal part because of the radiation pattern of the actual source and the locations of virtual sources contributing to either C2+ or C2− functions (Figure A1). Case II in Figure 3b is similar but with the source outside of the stationary phase zone. The overall C2 result is still symmetric but the travel time is biased, similar to biases in constructing C1 Green's functions discussed in other studies (e.g., Tsai, 2009). A series of derivations are given in Appendix A to support the observations.

2.3. Numerical Simulations for a Heterogeneous Medium

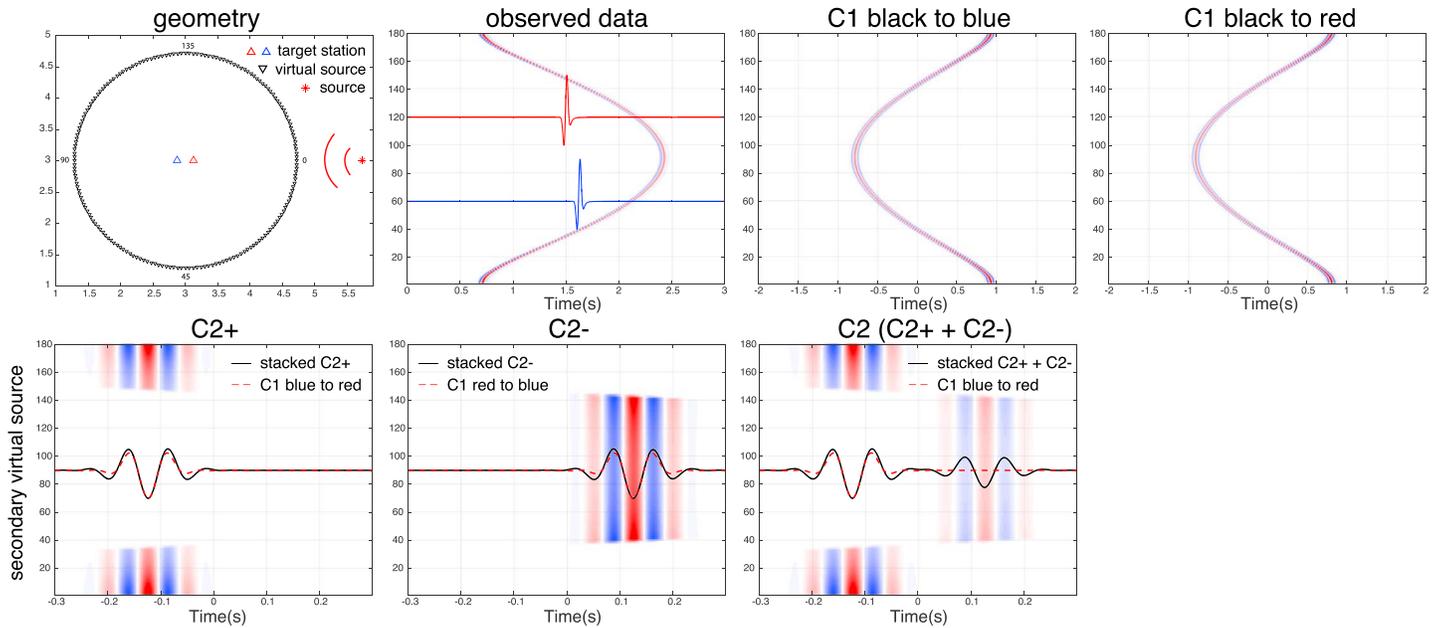
We create a 2-D heterogeneous medium by superimposing spatial velocity fluctuations onto a constant background velocity (Frankel & Clayton, 1986). The fluctuations are characterized by a von-Karman autocorrelation function (Nakata & Beroza, 2015). We set the autocorrelation length as 0.25 km and the fractal randomness is 0.05, with a maximum 40% velocity perturbation. We explore the results for the two cases shown above. To emphasize the correlation coda, we mute the first 1.5 s of the direct arrivals on the constructed C1 functions. We observe ringing signals on the positive time lag of the C1 functions, while the wavefield shown on the anticausal part appears to be more scattered (Figure 4). The ringing signals result from the cross correlation between the direct arrivals at the virtual sources and the scattered coda waves at the stations. We discuss this aspect of the results in Figure S1 in the supporting information. As shown in Figure 4b, cross-correlating the anticausal coda of the C1 functions, on the other hand, yields similar results to those for the homogeneous case: The correlation response is only observed on one side and the travel time is again biased given an unfavorable source location. Note the one-sided response and bias in travel time are still present even when the late coda is used (Figure S2).

These simple experiments indicate that the noise source distribution still has a strong effect on the C3 results and that the observed symmetry arises as a consequence of signal processing, rather than a diffuse coda wavefield. For this reason the increased symmetry of C3 functions may not be diagnostic of an improvement in the recovered Green's functions.

2.4. A Subtle C3 Improvement

Although signal processing may be the controlling factor for C3 function symmetry, there is evidence that using the correlation coda does mitigate the strong source irregularity to some extent. For the same station pair CI.CLT and CI.BBR, Figure 5 compares the causal and the anticausal parts of the retrieved Green's functions (several more examples are given in Figure S3). We find that the direct arrival on both sides of the C3+ Green's function is more consistent than those for the C1 Green's function. We also note that in this example, there is no significant improvement on the C3− Green's function, implying different wave contributions on the causal versus the anticausal C1 coda. This difference is probably due to the biased source distribution, so that the scatterers are illuminated differently by waves coming from different directions. We suggest that, rather than assessing reliability based on the symmetry of the overall C3 function, the symmetry of the C3+ (or C3−) function alone could serve as a more reliable indication of the reliability of the correlation function.

(a) case I



(b) case II

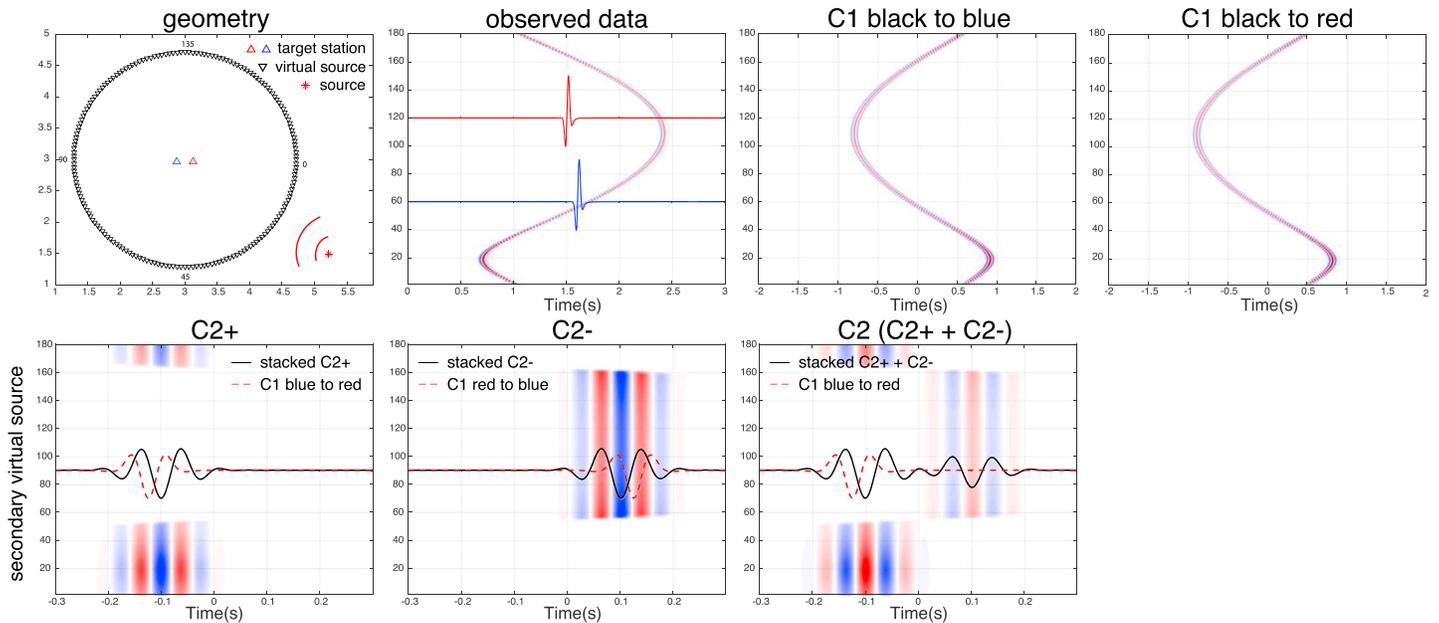
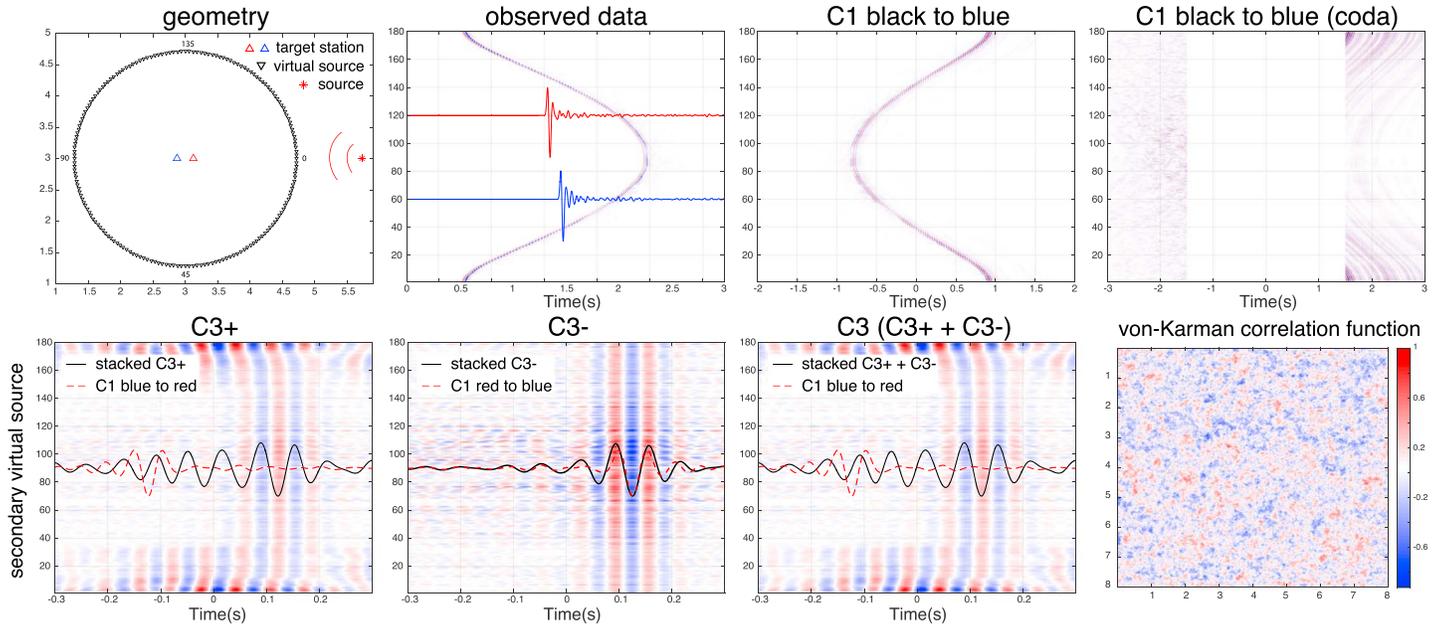


Figure 3. Numerical simulations for a homogeneous velocity model. (a) Top left shows the geometry of case I, for which two target receivers are surrounded by a circle of virtual source stations and the source is located on the extension of the receiver-receiver line. The following panel shows the simulated data from the source with target receiver recordings highlighted. The rest two display the C1 functions from virtual sources to each of the two target receivers. The first two panels on the second row illustrate the C2+ and C2− results. The C1 function (dashed curve) between the two target receivers approximates the expected travel time. Note the missing parts for certain virtual sources in both cases. This occurs because there is no available signal at those ranges on the C1 functions. The third panel shows the relatively symmetric C2 function that results from averaging the C2+ and C2− functions. (b) Similar to case I except that the source location moves outside of the stationary phase zone of the target receivers. The resulting C2 function, shown in the third figure on the second row, is still symmetric but has a notable bias in the phase shift away from the expected C1 functions in case I.

(a) case I



(b) case II

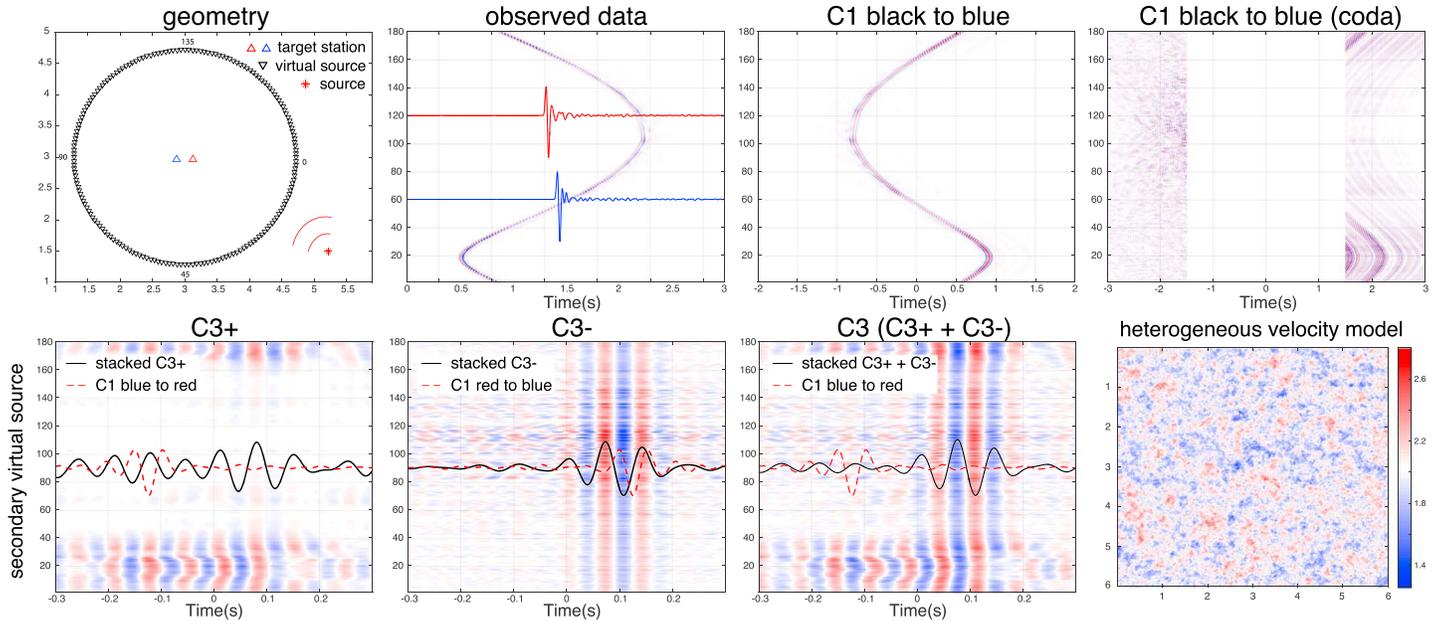


Figure 4. Numerical simulations with a heterogeneous velocity model. The setting of the two cases is the same as the one explained in Figure 3, except for the velocity model. We characterize the velocity model using a von-Karman spatial autocorrelation function, which is attached as the last plot in case I. The resulting heterogeneous velocity model is attached as the last plot in case II. C1 functions with first-arrivals muted are shown in the upper rightmost subplot in both cases. The coda parts used in constructing C3 functions are highlighted. Though only the C1 functions from the virtual sources to the blue receiver are displayed here, the results of the red receiver is very similar. Note the ringing patterns on the causal coda and more scattered anticausal coda in both cases. Cross correlation of the ringing patterns yield unexpected C3+ results.

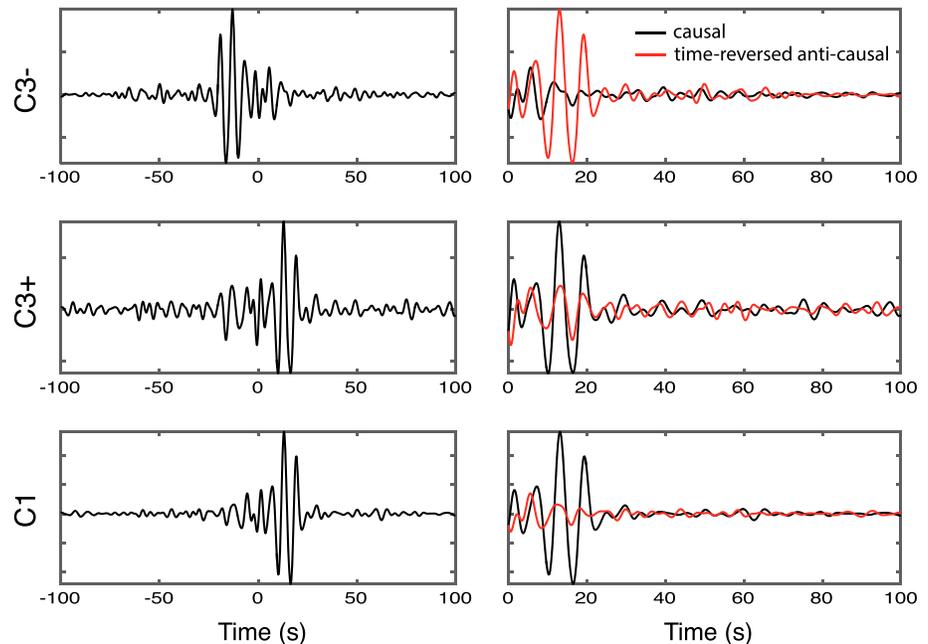


Figure 5. Correlation function between stations CI.CLT and CI.BBR. The three panels on the left, from top to bottom, show the averaged C3⁻, C3⁺, and C1 correlation functions, respectively. The panels on the right compare the causal and anticausal parts of the corresponding correlation functions. The match between the causal and the time-revised anticausal is improved for the C3⁺ correlation function.

3. A Cross-Talk Model for C3

Considering the above results, we propose an alternative explanation for how the C3 method works. It assumes noise sources are *correlated* to each other and that in constructing C1, we not only consider the correlation between the same signal recorded by different receivers, but also account for the correlation between different signals at those receivers.

We refer to the correlation between different signals as the cross talk of the associated sources. The cross-talk concept is not new. For higher-order interferometry, correlations (or convolutions) between multiple reflections (King & Curtis, 2012) or between the unperturbed background and scattered wavefield (Löer et al., 2014) have been used to explain the appearance of nonphysical energy in the SRI framework. Though we illustrate our model using scattering phases, repeating noise sources that are excited by the same mechanism should have a similar effect. For simplicity, we introduce the model using a 1-D case, where a source is excited from one side, scattered by several scatterers, and recorded on a linear array. Although the scatterers are considered secondary sources, the excited wavefields are not completely independent of each other. We simplify the signals with impulsive spikes in Figure 6a. Figure 6b reveals that the C1 functions with the first receiver as the virtual source consistently present a banded pattern across the array for both positive and negative time lags. Each banded pattern can be recognized as one virtual wavefield, excited at the first receiver and recorded by the rest of the array. If we take the C1 functions on two receivers and perform C3, adding the cross correlation within each virtual wavefield gives rise to a strong response, at the correct time lag, consistent with the travel time from one target receiver to another. Recall that in the C3 technique, we take the anticausal coda and time reverse it. By doing so, we mirror the banded pattern and therefore mirror the C3 result. This explains why the structure of signal processing is responsible for the observed symmetry. We offer this explanation for the C3 case, but the same logic applies to even higher-order correlations (e.g., C5) as well.

The cross-talk explanation predicts that the banding should be present in the data. We construct C1 functions among the stations in the XD array in Southern California. The geometry of the network (Figure 6c) is similar to the 1-D case illustrated above. The source of ambient seismic energy is mostly from the Pacific Ocean, to the west of the array. Figure 6d also shows the C1 functions constructed from the noise data in a 15-min-long

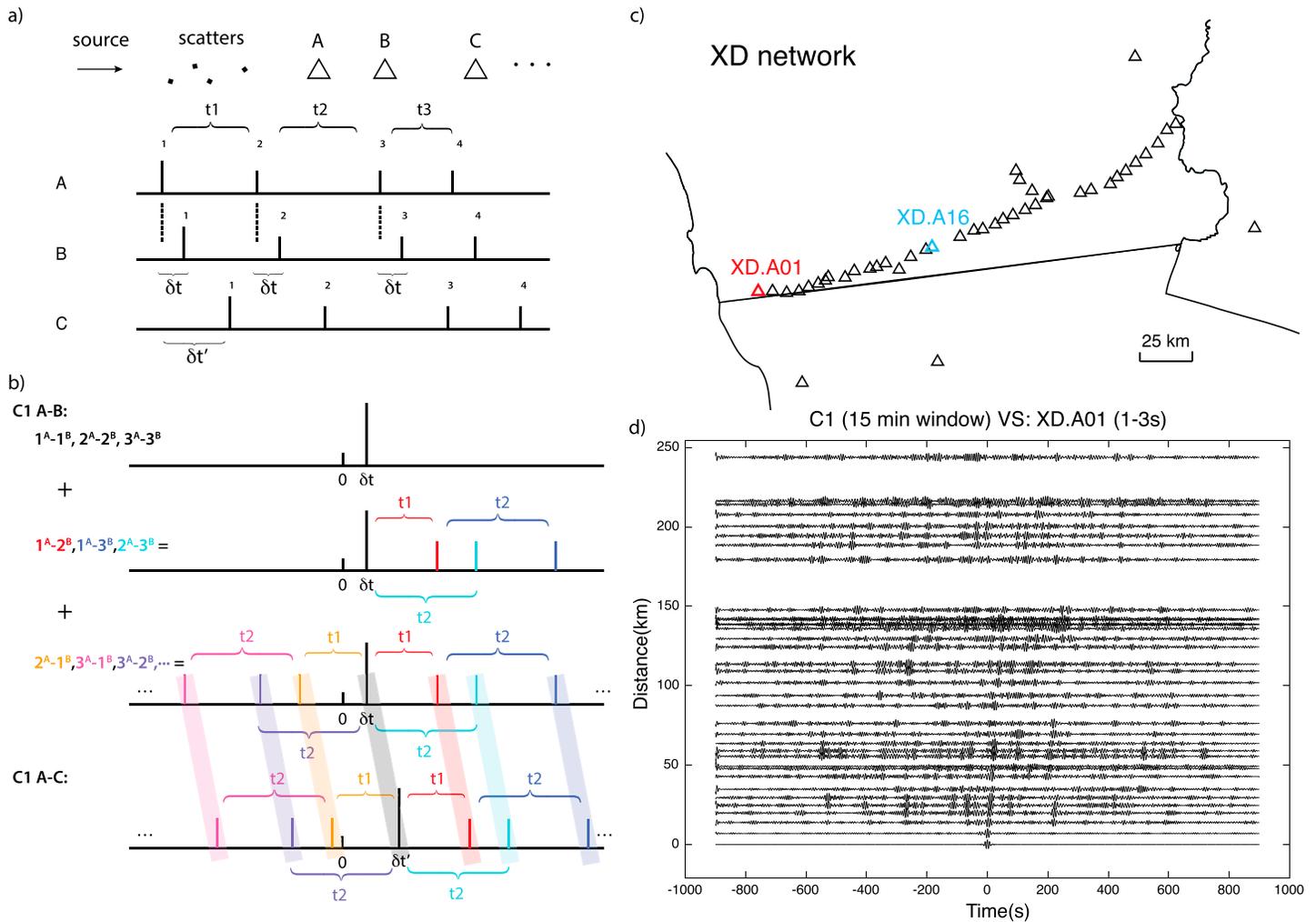


Figure 6. Schematic explanation for the 1-D cross-talk model. (a) The basic geometry of the 1-D model. A source is scattered by several scatterers and recorded by a linear array. The scattered signals are represented by numbered spikes. The time delay between each signal is marked by t_1, t_2 , and t_3 respectively; δt and $\delta t'$ indicate the travel time from the first receiver to the other two. (b) C1 function between the first receiver (a) and other receivers (b and c) in the linear array. The cross-talk model considers not only the cross correlation between the same signal recorded by both receivers, for example, 1A-1B (e.g., cross correlation between signal 1 on receivers A and B) but also the cross correlation between different signals recorded by these receivers, for example, 1A-2B (e.g., cross correlation between signal 1 on receiver A and signal 2 on receiver B) or 1B-2A. The cross talks between different pairs of signals are represented in different colors, which form banded patterns across the array. (c) Reference map showing the geometry of the XD network in the Salton trough of California. Locations of the station XD.A01 and XD.A16 are highlighted. (d) C1 functions, computed from ambient seismic field for a 15-min time window, between XD.A01 and other stations in the XD network. Localized arrivals, with consistent nearly linear moveout, can be observed on either side of the correlation functions. These arrivals correspond to the banded patterns in panel (b).

time window. The results are aligned with increasing distance from the virtual source, which is the leftmost station (XD.A01) in the array. It is clear that there are banded patterns showing up on both sides of the C1 functions. These localized patterns, with a consistent moveout, likely correspond to localized scatterers that are sensitive to the orientation of the incoming seismic waves. Though not specifically discussed, such banded patterns have also been observed in other studies, for example, Figure 2c in Lin et al. (2013) and Figures 11c and 11d in Nakata et al. (2015). Seismic noise sources vary spatially over time, suggesting that C1 functions from different time windows will have different cross-talk patterns.

The standard C3 method carries out ensemble averaging of the C1 functions over time. In constructing C3 functions, such time averaging may manifest another level of cross talk—the cross correlation between signals in different realizations. That is, it assumes intrinsic coherency between noise sources excited at different time periods. The existence of such higher-level cross talk would contaminate the C3 function, leading to

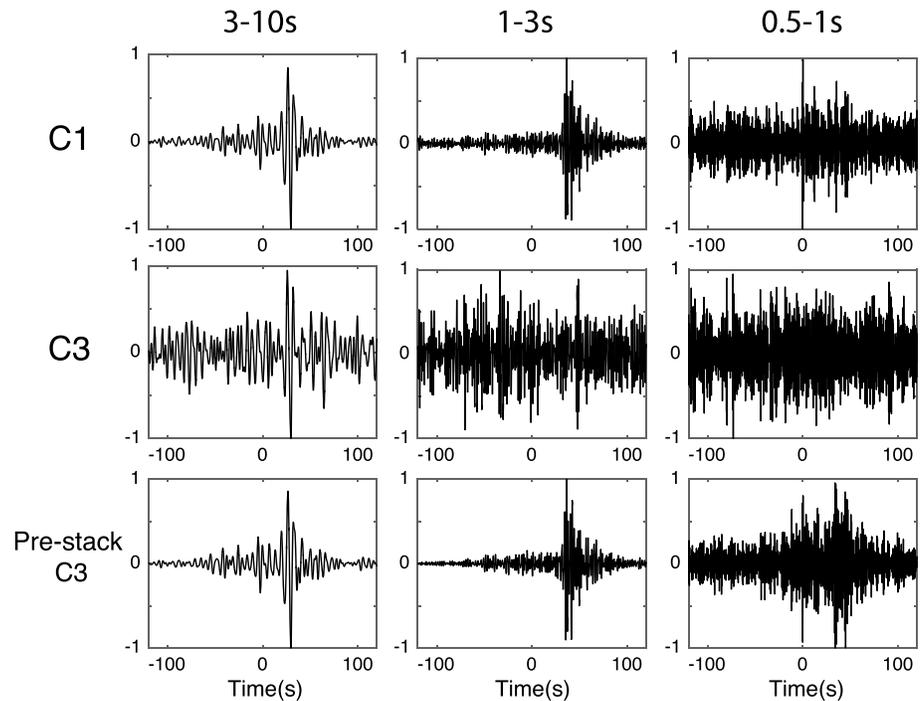


Figure 7. Normalized Green's function from XD.A01 to XD.A16 retrieved by using different techniques as labeled. From left to right the result is at 3–10 s, 1–3 s, and 0.5–1 s period. From top to bottom, the methods used are C1, standard C3, and prestack C3 methods.

relatively poor SNR, as has been observed in the previous studies (Froment et al., 2011; Sheng et al., 2017; Stehly et al., 2008). Figure S4 illustrates this with an example, in which we compare the cross correlation from different combinations of realizations. This observation motivates an approach to avoid such contamination by iterating cross correlation in each individual time window. We refer to this modified processing method as *prestack C3*.

4. The Pre-Stack C3 Method

As suggested previously, the difference between our prestack C3 and the standard C3 method lies in the order of forming and stacking correlation functions. Though in a different scenario, where a controlled source

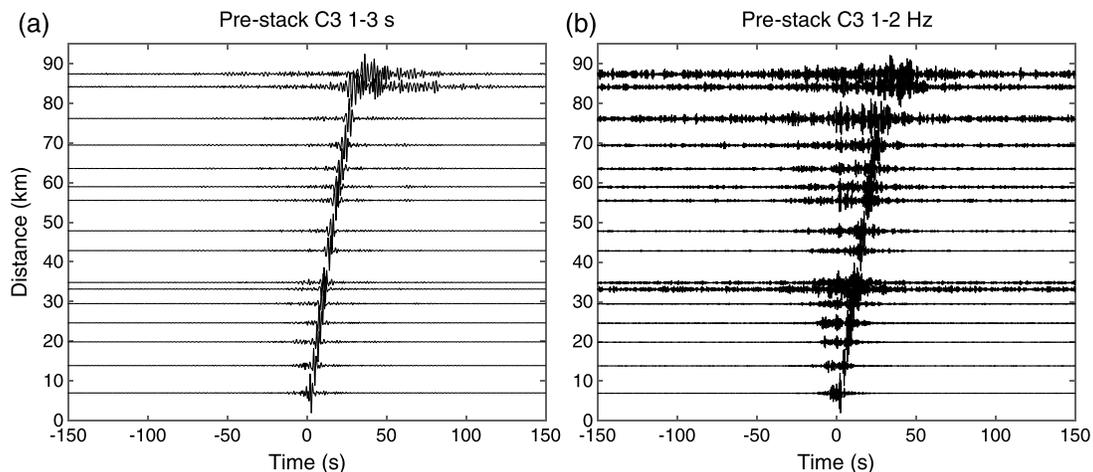


Figure 8. Green's functions from the westernmost station XD.A01 to the other stations through XD.A16. Each trace is normalized and plotted with increasing distance away from station XD.A01. (a) Green's functions in 1–3 s period and (b) Green's functions in 1–2 Hz frequency band.

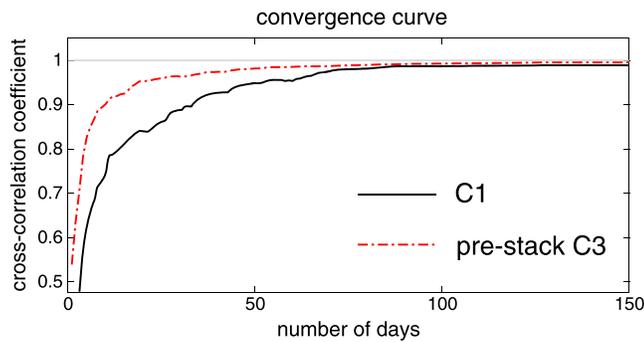


Figure 9. Convergence curve of C1 and prestack C3 method. The horizontal axis shows how many days of data are used to get the Green’s function. The vertical axis indicates the cross-correlation coefficient of the Green’s function relative to the long-term averaged result.

is used with a boundary of receivers and another boundary of sources, L er et al. (2014) demonstrated the idea of prestack by performing SRI source by source before the stacking. An analogy could be drawn between the SRI in L er et al. (2014) and our prestack C3 method, which performs C3 source subset by source subset before the stacking. We compare the effect of this order for correlation functions between the station pair XD.A01-XD.A16, with all available data recorded in 2011 (Figure 7). Each C1 function is constructed over a 15-min time window, and the coda starts at twice the Rayleigh wave travel time. The comparison of these estimated Green’s function serves as a performance measurement of different methods. We do not use symmetry for that comparison because we have shown that it is primarily attributable to signal processing. We time reverse the C3– functions and then add them to the C3+ functions to form the C3 result. We use all the stations, except XD.A01 and XD.A16, in the XD network (35 stations) as virtual sources in the standard C3 processing. While in the prestack C3 method, we use only the first five stations on

the left as virtual sources. Figure S5 illustrates the effect of the number of virtual sources in the prestack C3 method. It shows that adding more virtual sources will not always increase the SNR of the resulting correlation function; five virtual sources are enough for the data we use. SNR here is defined as the ratio between the peak amplitude of the first arrival and the root-mean-square of the late coda.

Compared with the other two methods, the prestack C3 technique consistently gives high-SNR Green’s functions over all frequency bands considered (Figure 7), while standard C3 fails to extract coherent signals at short periods (1–3 s and 0.5–1 s) and yields poor-SNR correlation functions even at the longest periods (3–10 s) considered. Note that neither C1 nor the standard C3 method is able to extract coherent signals at 0.5–1 s period, but prestack C3 appears to extend the recover of a coherent arrival in the correlation results to higher frequency.

To examine the reliability of the response, we plot a virtual-source gather of the Green’s functions constructed by the prestack C3 method. We take station XD.A01 as the virtual source. Figure 8 shows the results at two frequency bands. Although noisier than its counterpart, the moveout in Figure 8b can still be easily recognized. Moreover, at high frequency we observe weak arrivals showing up with approximately the correct time lag on the anticausal side of the Green’s functions. This suggests that energy propagating from the east side of the array is stronger at high frequency. Figure S6 shows the same moveout plots along with those constructed by the C1 and the standard C3 method.

We can also judge the performance of prestack C3 by examining how rapidly the correlation functions converge to the long-term correlations (Seats et al., 2012). We use the correlation coefficient between the short-term (accumulated daily) and the long-term correlation functions to assess the convergence rate. The correlation coefficients are calculated independently for each method, and the period band we focus on here is 1–3 s. Note that we window the correlation functions between –100 s and 100 s, so that the correlation coefficients are mainly determined by the ballistic wave contributions. Figure 9 shows the correlation coefficient curve, which indicates that the prestack C3 method significantly accelerates convergence.

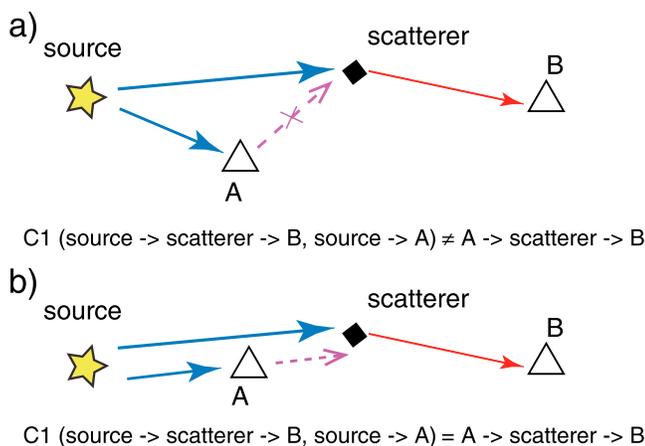


Figure 10. Illustration of the modulation in cross correlation. The source is situated in the stationary phase zone for (b) but not (a). solid lines with arrows indicate raypaths illuminated either by the source (blue) or the scatterer (red). (a) Source, station A and the scatterer are not collinear, which makes the cross correlation between source-scatterer-B and source-A not equal to A-scatterer-B. (b) Source, station A and the scatterer lie on the same line and cross-correlating source-scatterer-B with source-A gives the same travel time as A-scatterer-B.

5. Discussions and Conclusions

We propose an alternative explanation for higher-order cross correlations that explains the observations of asymmetry of the C3+ and C3– response and why their summation is symmetric. There remains a connection between our *cross-talk* model and the original conceptualization, which in effect states that the coda of a C1 function is composed of multiply-

scattered waves. In the cross-talk framework, these multiply-scattered waves can be understood as the cross-correlation between the signal from the original source and those illuminated by different scatterers. What we emphasize is that the contribution of the cross-talk to the virtual wavefield cannot be equivalent to an actual wave excited at one receiver, scattered and then recorded by another receiver. In cases when the actual source is not at the stationary point (Halliday & Curtis, 2009; Snieder, Wapenaar, & Lerner, 2006), the virtual ray path is modulated. The schematic representation in Figure 10 visualizes such modulation. It is worth pointing out that generally our interpretation of the correlation coda is not in conflict with its successful application, for example, on temporal variation analyses (e.g., Brenguier et al., 2008; Duputel et al., 2009; Mordret et al., 2016; Wegler & Sens-Schönfelder, 2007). Since the cross talks are primarily composed of scattered waves, they will be more sensitive to the medium than the original noise wavefield and therefore more suitable for time-dependent monitoring (Hadziioannou et al., 2009). However, if the correlation coda is dominated by cross talks, interpretations of the extracted temporal variations, especially their locations, could be affected.

Unlike seismic sources (or scatterers), virtual source stations cannot excite a real wavefield. Cross-correlating signals at stations is essentially a process of data manipulation, which reorganizes the original wavefield information without generating new information. From this point of view, the argument of the coexisting comparable wavefields (on C1 coda) between two stations in both directions, which has been used to explain the symmetry of the C3 functions, is misleading. However the above statement could still be true if there were equivalent energy flux from both directions, indicating a roughly homogeneous noise source distribution. In most cases including the one we have studied, a scattered wavefield from one station to another does not guarantee its counterpart. In other words, higher-order cross correlations, such as C3+ or C3−, will give a more symmetric Green's function only if the original wavefield allows it.

We present the cross-talk model in the 1-D case. The approach will work in 3-D; however, the geometric analogy is correspondingly more complex. A combination of an irregular array and varying seismic noise sources would present a complex C1 moveout distribution. An alternative understanding of the model would be to isolate the contribution from each individual signal at the virtual source. Therefore, a C3 function, modulated by the autocorrelation of each virtual source signal, could be recognized as a filtered version of the C1 function between the target stations. This interpretation is consistent with the simulations in section 2, which shows that wherever the virtual source is, the associated higher-order correlation function tends to mimic the C1 function, which itself is strongly influenced by the original noise wavefield. From this perspective, prestack C3 may not mitigate the bias due to nondiffuse sources. This argument may seem to be at odds with the SRI, which has been proven to suppress nonphysical arrivals by second-order correlations (e.g., King & Curtis, 2012). Though SRI is also suitable for extracting the Green's function between a virtual source and a receiver (Curtis & Halliday, 2010), it has been mostly used to bridge a real source and a receiver, which combines two independent wavefields: the ambient noise wavefield and the wavefield excited from the target source. While the Green's functions from the target source, either active (e.g., Duguid et al., 2011) or passive (e.g., Entwistle et al., 2015), are unbiased, it ensures that SRI sees fewer artifacts in the ultimate results.

A notable feature of cross correlation prior to stacking is that it accelerates convergence of the retrieved correlation functions. We demonstrate that cross-correlating the virtual sources discriminates the incoherent and coherent components of the recorded data and that the coherent signals improve the resolution of the second-order cross correlation. Faster convergence could be critical in situations where data are limited so that the C1 method fails to extract reliable Green's functions. It could also be useful for resolving time dependence, such as found by Hadziioannou et al. (2011) who have shown that high temporal resolution is necessary to observe small velocity change above the measurement precision; however, since prestack C3 may have time-variable sampling of the ambient field, it remains to be demonstrated that it would offer an advantage for this purpose.

Extending correlation results to higher frequency is crucial for many applications. Nakata et al. (2015) have demonstrated the feasibility of using dense large-N arrays with a phase isolation filter to achieve this purpose; however, despite the trend of deploying dense arrays, most available networks have much larger interstation distances. We show that at least in our experiments, owing to the rapid convergence, it is possible to use the prestack C3 method to extend the high-frequency limit substantially.

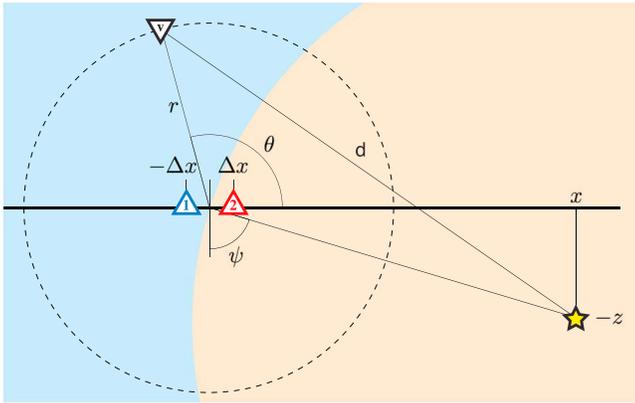


Figure A1. Geometry of source and receivers. Red and blue triangles are the target receivers. The black triangle represents one of the virtual source stations, which are on the dashed circle. The yellow star is the actual source, located at (x, z) . The background color indicates the region of virtual source stations that contribute to either causal ($C2^+$, marked by orange) or anticausal ($C2^-$, marked by blue) $C2$ functions, respectively.

Appendix A: Derivations of Higher-Order Cross Correlations for a Homogeneous Lossless Medium

The schematic of the derivation is shown in Figure A1. The location of each virtual source is $(r \cos \theta, r \sin \theta)$. The observed data u at each receiver is

$$u_1 = SJ(r_1) \exp \left\{ ik \left(\frac{z}{\cos \psi} + \Delta x \sin \psi \right) \right\} \quad (A1)$$

$$u_2 = SJ(r_2) \exp \left\{ ik \left(\frac{z}{\cos \psi} - \Delta x \sin \psi \right) \right\}$$

$$u_v = SJ(r_v) \exp \{ ikd \}.$$

where d is the distance between the actual source and the virtual source:

$$d = \sqrt{r^2 + x^2 + z^2 - 2r\sqrt{x^2 + z^2} \cos \left(\frac{\pi}{2} + \theta - \psi \right)} \quad (A2)$$

S is the source spectrum, k the wave number in the medium, which is assumed to be homogeneous, and J is the geometrical spreading function, which is: $J_{3D} = \frac{\rho}{4\pi R}$ and $J_{2D} = \frac{\rho \exp(i\pi/4)}{\sqrt{8\pi k R}}$, in 3-D and 2-D, respectively (Snieder

& Chapman, 1998). For simplicity, we assume that the source signature is a delta function ($S(\omega) = 1$) and ignore geometrical spreading J , since we are mostly interested in the phase of correlation functions. We also assume that Δx is small and that the plane wave approximation holds at stations 1 and 2.

The cross correlation of wavefields ($C1$) between a virtual source station and the two target stations is therefore given by

$$C1_{1v} = \exp \left\{ ik \left(\frac{z}{\cos \psi} + \Delta x \sin \psi - d \right) \right\} \quad C1_{2v} = \exp \left\{ ik \left(\frac{z}{\cos \psi} - \Delta x \sin \psi - d \right) \right\} \quad (A3)$$

Depending on the location of the virtual source stations, relative to the target stations, the energy could be on the causal or the anticausal part (Figure A1). Correlation of the correlation (with $C1_{1v}$ acting as the source) can be written as

$$C2^+ = (C1_{1v})^* (C1_{2v}) = \exp(-2ik\Delta x \sin \psi) \quad (A4)$$

$$C2^- = (C1_{1v}^*)^* (C1_{2v}^*) = \exp(2ik\Delta x \sin \psi)$$

where we obtain anticausal waves from $C2^+$ and causal waves from $C2^-$ as shown in Figure 3. We note that neither $C2^+$ nor $C2^-$ depends on d , which implies they are not controlled by the location of virtual source stations. This indicates that the extracted waves have the same phase for different virtual sources, although the amplitudes should vary as the geometrical spreading term J varies for different virtual source locations. Both the phase and amplitude match the simulation results in Figure 3. Averaging the higher-order cross correlation over virtual sources helps to increase the signal-to-noise ratio of target waves.

Equation (A4) also shows that the constructed waves have biases that vary with $\sin \psi$. Note that the correct Green's function should be proportional to $\exp(\pm 2ik\Delta x)$, and this is only true when $\psi = \pm 90^\circ$, that is, case I in section 2.2.

References

- 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. <https://doi.org/10.1126/science.1160943>
- Campillo, M., & Paul, A. (2003). Long-range correlations in the diffuse seismic coda. *Science*, 299(5606), 547–549. <https://doi.org/10.1126/science.1078551>
- Curtis, A., & Halliday, D. (2010). Source-receiver wave field interferometry. *Physical Review E*, 81(4), 046601. <https://doi.org/10.1103/PhysRevE.81.046601>
- Denolle, M., Dunham, E. M., Prieto, G. A., & Beroza, G. C. (2013). Ground motion prediction of realistic earthquake sources using the ambient seismic field. *Journal of Geophysical Research*, 118, 2102–2118. <https://doi.org/10.1029/2012JB009603>

Acknowledgments

Data used in this study were obtained from the Southern California Seismic Network CI (<https://doi.org/10.7914/SN/CI>), the San Jacinto Fault Zone Experiment YN (https://doi.org/10.7914/SN/YN_2010), and the Passive seismic study of a magma-dominated rift: the Salton Trough XD (https://doi.org/10.7914/SN/XD_2011). We are grateful to Fan-Chi Lin and Andrew Curtis for their constructive comments. We thank the Madagascar Development Team for providing reliable software package for numerical simulations. This work was supported through NSF award EAR-1520867.

- Denolle, M., Dunham, E. M., Prieto, G. A., & Beroza, G. C. (2014). Strong ground motion prediction using virtual earthquakes. *Science*, 343(6169), 399–403. <https://doi.org/10.1126/science.1245678>
- Duguid, C., Halliday, D., & Curtis, A. (2011). Source-receiver interferometry for seismic wavefield construction and ground roll removal. *The Leading Edge*, 30(8), 838–843. <https://doi.org/10.1190/1.3626489>
- Duputel, D., Ferrazzini, V., Brenguier, F., Shapiro, N., Campillo, M., & Nercessian, A. (2009). Real time monitoring of relative velocity changes using ambient seismic noise at the piton de la Fournaise volcano (La Réunion) from January 2006 to June 2007. *Journal of Volcanology and Geothermal Research*, 184(1-2), 164–173. <https://doi.org/10.1016/j.jvolgeores.2008.11.024>
- Entwistle, E., Curtis, A., Galetti, E., Baptie, B., & Meles, G. (2015). Constructing new seismograms from old earthquakes: Retrospective seismology at multiple length scales. *Journal of Geophysical Research: Solid Earth*, 120, 2466–2490. <https://doi.org/10.1002/2014jb011607>
- Fomel, S., Sava, P., Vlad, I., Liu, Y., & Bashkardin, V. (2013). Madagascar: Open-source software project for multidimensional data analysis and reproducible computational experiments. *Journal of Open Research Software*, 1, e8.
- Frankel, A., & Clayton, R. (1986). Finite difference simulations of seismic scattering: Implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity. *Journal of Geophysical Research*, 91(B6), 6465–6489. <https://doi.org/10.1029/JB091iB06p06465>
- Froment, B., Campillo, M., & Roux, P. (2011). Reconstructing the Green's function through iteration of correlations. *Comptes Rendus Geoscience*, 343(8-9), 623–632. <https://doi.org/10.1016/j.crte.2011.03.001>
- Hadziioannou, C., Larose, E., Baig, A., Roux, P., & Campillo, M. (2011). Improving temporal resolution in ambient noise monitoring of seismic wave speed. *Journal of Geophysical Research*, 116, B07304. <https://doi.org/10.1029/2011JB008200>
- Hadziioannou, C., Larose, E., Coutant, O., Roux, P., & Campillo, M. (2009). Stability of monitoring weak changes in multiply scattering media with ambient noise correlation: Laboratory experiments. *The Journal of the Acoustical Society of America*, 125(6), 3688–3695. <https://doi.org/10.1121/1.3125345>
- Halliday, D., & Curtis, A. (2008). Seismic interferometry, surface waves and source distribution. *Geophysical Journal International*, 175(3), 1067–1087. <https://doi.org/10.1111/j.1365-246X.2008.03918.x>
- Halliday, D., & Curtis, A. (2009). Seismic interferometry of scattered surface waves in attenuative media. *Geophysical Journal International*, 178(1), 419–446. <https://doi.org/10.1111/j.1365-246X.2009.04153.x>
- Halliday, D., & Curtis, A. (2010). An interferometric theory of source-receiver scattering and imaging. *Geophysics*, 75(6), SA95–SA103. <https://doi.org/10.1190/1.3486453>
- Hennino, R., Tégourès, N., Shapiro, N., Margerin, L., Campillo, M., van Tiggelen, B., & Weaver, R. (2001). Observation of equipartition of seismic waves. *Physical Review Letters*, 85(15), 3447–3450.
- King, S., & Curtis, A. (2012). Suppressing nonphysical reflections in Green's function estimates using source-receiver interferometry. *Geophysics*, 77(1), Q15–Q25. <https://doi.org/10.1190/geo2011-0300.1>
- Lin, F.-C., Li, D., Clayton, R. W., & Hollis, D. (2013). High-resolution 3D shallow crustal structure in Long Beach, California: Application of ambient noise tomography on a dense seismic array. *Geophysics*, 78(4), Q45–Q56. <https://doi.org/10.1190/geo2012-0453.1>
- Lin, F.-C., Moschetti, M. P., & Ritzwoller, M. H. (2008). Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and love wave phase velocity maps. *Geophysical Journal International*, 173(1), 281–298. <https://doi.org/10.1111/j.1365-246X.2008.03720.x>
- Lobkis, O. I., & Weaver, R. L. (2001). On the emergence of the Green's function in the correlations of a diffuse field. *The Journal of the Acoustical Society of America*, 110(6), 3011–3017. <https://doi.org/10.1121/1.1417528>
- Löer, K., Meles, G. A., Curtis, A., & Vasconcelos, I. (2014). Diffracted and pseudo-physical waves from spatially limited arrays using source-receiver interferometry (SRI). *Geophysical Journal International*, 196(2), 1043–1059. <https://doi.org/10.1093/gji/ggt435>
- Ma, S., & Beroza, G. C. (2012). Ambient-field Green's functions from asynchronous seismic observations. *Geophysical Research Letters*, 39, L06301. <https://doi.org/10.1029/2011GL050755>
- Mordret, A., Mikesell, T. D., Harij, C., Lipovsky, B. P., & Prieto, G. A. (2016). Monitoring Southwest Greenland's ice sheet melt with ambient seismic noise. *Science Advances*, 2(5), e1501538. <https://doi.org/10.1126/sciadv.1501538>
- Nakata, N., & Beroza, G. C. (2015). Stochastic characterization of mesoscale seismic velocity heterogeneity in Long Beach, California. *Geophysical Journal International*, 203, 2049–2054.
- Nakata, N., Chang, J. P., Lawrence, J. F., & Boué, P. (2015). Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry. *Journal of Geophysical Research: Solid Earth*, 120, 1159–1173. <https://doi.org/10.1002/2015JB011870>
- Prieto, G. A., Lawrence, J. F., & Beroza, G. C. (2009). Anelastic earth structure from the coherency of the ambient seismic field. *Journal of Geophysical Research*, 114, B07303. <https://doi.org/10.1029/2008JB006067>
- Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W. A., & Fehler, M. C. (2005). Surface wave tomography from microseisms in southern California. *Geophysical Research Letters*, 32, L14311. <https://doi.org/10.1029/2005GL023155>
- Seats, K. J., Lawrence, J. F., & Prieto, G. A. (2012). Improved ambient noise correlation functions using Welch's method. *Geophysical Journal International*, 188(2), 513–523. <https://doi.org/10.1111/j.1365-246X.2011.05263.x>
- Sens-Schönfelder, C., & Wegler, U. (2006). Passive image interferometry and seasonal variations of seismic velocities at Merapi volcano, Indonesia. *Geophysical Research Letters*, 33, L21302. <https://doi.org/10.1029/2006GL027797>
- 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. <https://doi.org/10.1126/science.1108339>
- Sheng, Y., Denolle, M., & Beroza, G. C. (2017). Multicomponent C3 Green's functions for improved long period ground motion prediction. *Bulletin of the Seismological Society of America*, 107(6), 2836–2845. <https://doi.org/10.1785/0120170053>
- Snieder, R. (2004). Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E*, 69(4), 046610. <https://doi.org/10.1103/PhysRevE.69.046610>
- Snieder, R., & Chapman, C. (1998). The reciprocity properties of geometrical spreading. *Geophysical Journal International*, 132, 89–95.
- Snieder, R., Sheiman, J., & Calvert, R. (2006). Equivalence of the virtual source method and wavefield deconvolution in seismic interferometry. *Physical Review E*, 73(6), 066620. <https://doi.org/10.1103/PhysRevE.73.066620>
- Snieder, R., Wapenaar, K., & Larner, K. (2006). Spurious multiples in seismic interferometry of primaries. *Geophysics*, 71(4), S1111–S1124. <https://doi.org/10.1190/1.2211507>
- Spica, Z., Perton, M., Calò, M., Legrand, D., Córdoba-Montiel, F., & Iglesias, A. (2016). 3-D shear wave velocity model of Mexico and south US: Bridging seismic networks with ambient noise cross-correlations (C1) and correlation of coda of correlations (C3). *Geophysical Journal International*, 206(3), 1795–1813. <https://doi.org/10.1093/gji/ggw240>
- Stehly, L., & Boué, P. (2017). On the interpretation of the amplitude decay of noise correlations computed along a line of receivers. *Geophysical Journal International*, 209(1), 358–372.

- Stehly, L., Campillo, M., Froment, B., & Weaver, R. L. (2008). Reconstructing Green's function by correlation of the coda of the correlation (C3) of ambient seismic noise. *Journal of Geophysical Research*, *113*, B11306. <https://doi.org/10.1029/2008JB005693>
- Stehly, L., Campillo, M., Shapiro, N. M., Fry, B., & Boschi, L. (2009). Tomography of the alpine region from observations of seismic ambient noise. *Geophysical Journal International*, *178*(1), 338–350. <https://doi.org/10.1111/j.1365-246X.2009.04132.x>
- Tsai, V. C. (2009). On establishing the accuracy of noise tomography travel-time measurements in a realistic medium. *Geophysical Journal International*, *178*(3), 1555–1564. <https://doi.org/10.1111/j.1365-246X.2009.04239.x>
- Tsai, V. C. (2011). Understanding the amplitudes of noise correlation measurements. *Journal of Geophysical Research*, *116*, B09311. <https://doi.org/10.1029/2011JB008483>
- Viens, L., Laurendeau, A., Bonilla, L. F., & Shapiro, N. M. (2014). Broad-band acceleration time histories synthesis by coupling low-frequency ambient seismic field and high frequency stochastic modeling. *Geophysical Journal International*, *199*(3), 1784–1797. <https://doi.org/10.1093/gji/ggu362>
- Wapenaar, K. (2004). Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation. *Physical Review Letters*, *93*(25), 254301. <https://doi.org/10.1103/PhysRevLett.93.254301>
- Wapenaar, K., Slob, E., & Snieder, R. (2006). Unified Green's function retrieval by cross correlation. *Physical Review Letters*, *97*(23), 234301. <https://doi.org/10.1103/PhysRevLett.97.234301>
- Weaver, R. L., Froment, B., & Campillo, M. (2009). On the correlation of non-isotropically distributed ballistic scalar diffuse waves. *The Journal of the Acoustical Society of America*, *126*(4), 1817–1826. <https://doi.org/10.1121/1.3203359>
- Wegler, U., & Sens-Schönfelder, C. (2007). Fault zone monitoring with passive image interferometry. *Geophysical Journal International*, *168*(3), 1029–1033. <https://doi.org/10.1111/j.1365-246X.2006.03284.x>
- Zhang, J., & Yang, X. (2013). Extracting the surface wave attenuation from seismic noise using correlation of the coda of correlation. *Journal of Geophysical Research: Solid Earth*, *118*, 2191–2205. <https://doi.org/10.1002/jgrb.50186>