Extracting time-domain Green's function estimates from ambient seismic noise

Karim G. Sabra, Peter Gerstoft, Philippe Roux, and W. A. Kuperman

Marine Physical Laboratory, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA

Michael C. Fehler

Los Alamos National Laboratory, Los Alamos, New Mexico, USA

Received 27 October 2004; revised 5 January 2005; accepted 11 January 2005; published 15 February 2005.

[1] It has been demonstrated experimentally and theoretically that an estimate of the Green's function between two seismic stations can be obtained from the time-derivative of the long-time average cross correlation of ambient noise between these two stations. This TDGF estimate from just the noise field includes all tensor components of the Green's function and these Green's function estimates can be used to infer Earth structure. We have computed cross correlations using 1 to 30 continuous days of ambient noise recorded by over 150 broadband seismic stations located in Southern California. The data processing yielded thousands of cross-correlation pairs, for receiver separations from 4–500 km, which clearly exhibit coherent broadband propagating dispersive wavetrains across frequency band 0.1–2 Hz. **Citation:** Sabra, K. G., P. Gerstoft, P. Roux, W. A. Kuperman, and M. C. Fehler (2005), Extracting time-domain Green's function estimates from ambient seismic noise, *Geophys. Res. Lett.*, *32*, L03310, doi:10.1029/2004GL021862.

1. Introduction

[2] The seismic Time Domain Green's Function (TDGF) between two points is usually estimated from the measured response from an active controlled source or using seismic sources of opportunity. Thus multiple point-to-point measurements over a large number of seismic stations separated by large distances can be cumbersome. On the other hand, coherent deterministic arrival-times can be extracted from the time-derivative of the time-averaged ambient noise cross-correlation function (NCF) between two receivers. These coherent arrival-times are related to the TDGF between these two receivers. Large and dense networks of broadband seismic sensors are commonly used for earthquake monitoring worldwide and they measure ambient noise continuously. These ambient noise recordings can be readily processed to provide an estimate of the TDGFs based on the NCFs between all station pairs of the network.

[3] The TDGF estimate emerges from those correlations that contain noise sources whose field passes through both receivers (i.e., coherent noise). Experimental evidence of this result has been presented in helioseismology [*Rickett and Claerbout*, 1999], ultrasonics [*Lobkis and Weaver*, 2001; *Weaver and Lobkis*, 2003; *Larose et al.*, 2004; *Malcolm et al.*, 2004; *Weaver and Lobkis*, 2004], underwater acoustics [*Roux et al.*, 2004; *Sabra et al.*, 2005], and in seismology [*Shapiro and Campillo*, 2004]. Furthermore, this non-intuitive result was demonstrated theoretically under the hypothesis of normal-mode equipartition [*Lobkis and Weaver*, 2001] and more generally for the case of a homogeneous medium with attenuation [*Snieder*, 2004; *Roux et al.*, 2005] and in a waveguide [*Sabra et al.*, 2005]. Relationships between the cross correlation of the transmission response with the reflection response were also derived for inhomogeneous elastic media [*Wapenaar et al.*, 2004].

0094-8276/05/2004GL021862$05.00

2. Ambient Noise Cross Correlations

[4] For a spatially uniform broadband noise-distribution in a uniform medium with sound speed c, it can be argued [*Snieder*, 2004; *Roux et al.*, 2005] that the NCF C of the traces at two receivers separated by distance L is zero for $|t| > L/c$ (no wave can travel slower than c), non-continuous at $t = \pm L/c$ and continuous for $|t| < L/c$ (waves originating from broadside will appear to travel faster than c). Thus the derivative of the NCF is

$$\frac{dC}{dt} \approx -\frac{\delta(t+L/c)}{L} + \frac{\delta(t+L/c)}{L} \tag{1}$$

It is recognized that the RHS is proportional to the free space TDGF and the time reversed TDGF.

[5] Based on analytic derivations for specific propagation models [*Snieder*, 2004; *Roux et al.*, 2005] and discussions by *Lobkis and Weaver* [2001] and *Weaver and Lobkis* [2003], we construct the following relationship for the time-derivative of the ambient noise cross correlation C_{ij} (1, 2, t) between two seismic stations 1 (located at $\mathbf{r}_1$ recording component i) and 2 (located at $\mathbf{r}_2$ recording component j) to the TDGF G_{ij} ($\mathbf{r}_1$; $\mathbf{r}_2$, t)

$$\frac{dC_{ij}}{dt} \approx -G_{ij}(\mathbf{r}_1;\mathbf{r}_2,t) + G_{ji}(\mathbf{r}_2;\mathbf{r}_1,-t) \tag{2}$$

The TDGF G_{ij} ($\mathbf{r}_1$; $\mathbf{r}_2$, t) relates a unit concentrated impulse displacement in direction i at $\mathbf{r}_1$ to the displacement response in direction j at receiver $\mathbf{r}_2$.

[6] In equation (1), the terms on the RHS are respectively: (1) the TDGF which comes from noise events that propagate from station 1 to 2 and yields a positive correlation time-delay t and (2) the time-reversed TDGF which comes from noise events that propagate from station 2 to 1 and yields a negative correlation time-delay $-t$. Thus, for a uniform noise

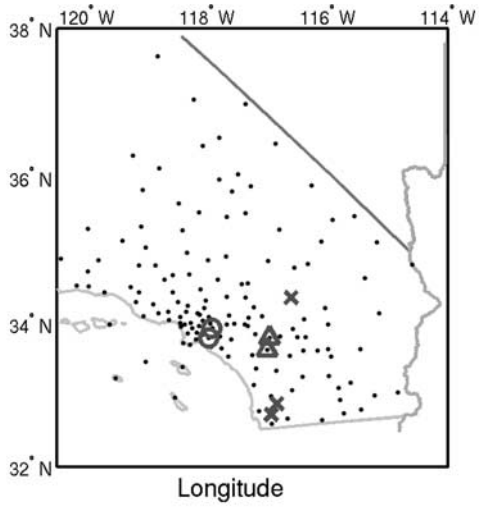

Figure 1. Map of the 150 online stations in the Southern California Seismic network, circle: BRE, OLI; triangle: DGR, MSJ; cross: JVA, SDG and SDL, EDL.

source distribution surrounding the two stations the derivative of the NCF will be a symmetric function with respect to the arrival time because ambient noise sources are distributed on both sides of the station pair. However, in the case of a predominant directional noise source (e.g., noise originating from ocean microseism, see next section) the NCF and its time-derivative likely will be one-sided.

[7] The cross correlation C_{ij} is computed from the observed fields $v_i(\mathbf{r}_1, t)$ and $v_j(\mathbf{r}_2, t)$ by integration over the whole observation period T

$$C_{ij}(1,2,t) = \int_0^T v_i(\mathbf{r}_1,\tau)v_j(\mathbf{r}_2,\tau+t)\mathrm{d}\tau. \quad (3)$$

Several researchers have used a form similar to equation (1), but instead of the derivative they used the cross correlation [e.g., *Campillo and Paul*, 2003; *Shapiro and Campillo*, 2004]. The ambient noise recorded on broadband seismic stations typically has a relatively small bandwidth. Hence the time series of the NCF and its time-derivative will have similar arrival structure but with a phase shift ($\pi/2$ for narrowband signals). Using the derivative also enhances the higher frequencies in the derived TDGF usually enhance noise. If undesired noise becomes an experimental issue when performing the time derivative of the NCF, it is not a bad approximation to estimate the TDGF as the NCF itself.

3. Data Processing Procedure and Experimental Results

[8] The cross-correlation technique is applied to ambient noise recordings collected at the 150 available stations of the Southern California network (Figure 1) for a continuous recording period (1–18 July 2004) using a 20-Hz sampling frequency. The NCFs were computed for all station pairs separated by a 4.3–500 km distance. It is computationally advantageous to obtain the 18-day NCF by cross correlating using continuous data for each day and then stacking the daily segments.

[9] This cross-correlation technique works best when the noise distribution is uniform in space and time [*Larose et al.*, 2004; *Snieder*, 2004; *Shapiro and Campillo*, 2004; *Roux et al.*, 2005]. Hence the effects of large seismic events should be minimized since they would otherwise dominate the arrival-time structure of the NCF. One straightforward way is to disregard the amplitude completely by keeping only the sign of the signals [*Campillo and Paul*, 2003; *Shapiro and Campillo*, 2004, *Larose et al.*, 2004]. However, this rough clipping creates high frequency noise and artificially modifies the ambient noise spectrum. To minimize the effect of clipping, one possibility is using a clipping threshold for each station. The threshold is determined as the minimum of the standard deviations measured over each day, hoping that few events appear in that time interval so that the threshold is only determined by the noise. This way the effect of large events is reduced, but the high frequency content of the ambient noise is less distorted.

[10] The selected frequency band for computing the NCFs was 0.1–2 Hz, which contained most of the coherent ambient noise between these stations pairs (as can be seen from Figure 6 which will be discussed later). The main part of energy was below 0.4 Hz. Typical ambient noise spectra exhibit a maximum in the low-frequency band [0.1–0.2 Hz] [*Aki and Richards*, 1980, chap. 10]. The lower bound was chosen to mitigate the effects of larger earthquakes in the data.

[11] Given the network's vicinity to the Pacific coast, it is expected that the NCF for station pairs oriented perpendicular and close to the coast will emerge more reliably than for station pairs oriented parallel to the coastline and located further inland. Furthermore, the directionality of the ambient noise originating mostly from the ocean microseisms should yield one-sided NCFs.

[12] NCFs (Figure 2) and their derivatives (Figure 3) for the vertical component of the ambient noise recordings were computed for all receiver pairs separated by 4.3–500 km having a bearing angle with respect to the north direction of

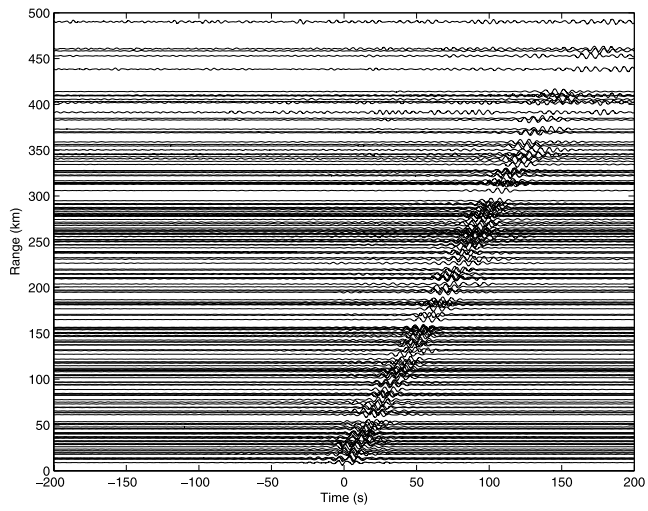

Figure 2. NCF Seismic record section up to 500 km. Station pairs having a bearing between 1–21 deg to N and a high SNR are shown.

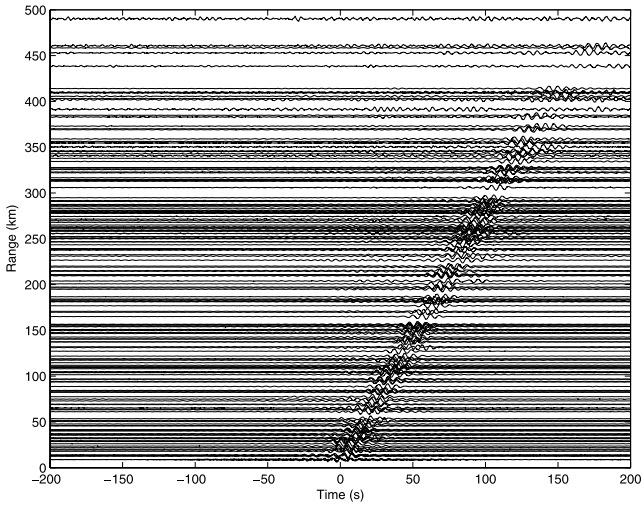

Figure 3. Time-derivative NCF Seismic record section up to 500 km. Station pairs having a bearing between 1 and 21 deg to N and a high SNR are shown.

1–21 deg and pointing away from the ocean, roughly a 45 deg angle to the coast. A record section of the NCFs and their derivatives as a function of increasing receiver separation shows clearly that the recovered signals are propagating wavetrains (Figures 2 and 3) with an average group velocity of 2.8 km/s. NCFs were averaged across the whole 18 day recording interval. We estimated the group velocity of Rayleigh waves within the region containing the stations used in calculating the NCF record section. We began by estimating an average layered P and S wave velocity model for the region from the 3D model of *Kohler et al.* [2003]. Group velocity as a function of period was calculated for the model structure. The estimated group velocity ranges from 2.8 km/s at 5-s period to 3.0 km/s at 10-s period, which agrees well with the estimated group velocity of 2.8 km/s at a period of about 6 s from our NCF record section.

[13] For this particular bearing of station pairs, the NCFs are typically not symmetric in time because the ambient noise is dominated by ocean microseisms propagating away from the coast. Station pairs oriented parallel to the coast still reveals propagating wavetrains and often one sided NCFs but the overall SNR diminishes because ocean generated noise contribute less to the NCF (not shown).

[14] To investigate the important issue of the rate at which the TDGF emerges as a function of averaging time and separation distance (range) we define the SNR as the ratio (in dB) of the peak amplitude in a 50-s time-window centered around the main arrival and the standard deviation of a time-window, −250 to −200 s, of the time derivative of the NCF. Figure 4a shows SNR for the station pairs from Figure 2 with the highest SNR. Due to our choice of data, one cannot use Figure 4a to draw conclusions about the decay of SNR versus range. In Figures 4b and 4c, we focus on stations within a specific range, 200–220 km. The SNR for each day (Figure 4b) shows moderate variation as the noise and seismicity varies. Based on the cross-correlation time series for each day, we sum these to obtain the time series for N number of days and compute the SNR. Each day is treated as an independent observation and therefore for each N the sum is computed for all combinations of the 18 days. The accumulated SNR based on the summed traces

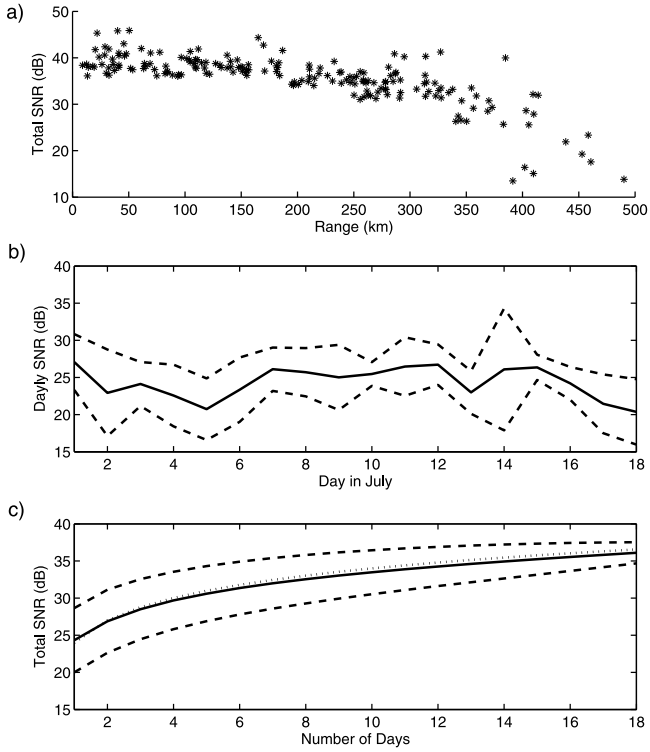

Figure 4. a) SNR for each station pair used in plotting the record sections, Figures 2 and 3. b) Mean and standard deviation of SNR for each day for stations used in a) within a range of 200–220 km. c) Mean and standard deviation of SNR computed by summing all combinations of traces for a number of days, a square root dependence is shown as a dotted line.

(Figure 4c) shows clearly that the SNR increases proportional to the square root of recording time in agreement with *Roux et al.* [2005] and *Snieder* [2004].

[15] The estimated waveforms are quite stable over time as shown in Figure 5 for each day of July 2004 for a single

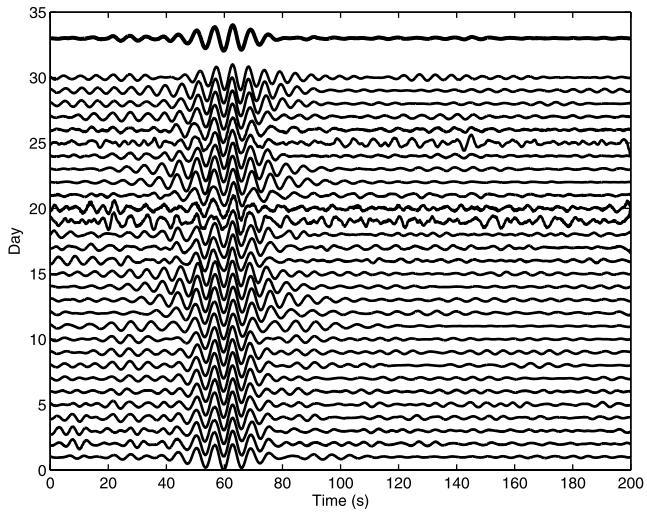

Figure 5. Estimated NCF for vertical component JVA-SDG for each day of July. The summed trace for the whole month of July is shown as heavy solid.

station pair JVA-SDG (cross in Figure 2, separation 182 km). The waveforms are shown in Figure 5 for each day of July 2004. While there are some fluctuations in the small amplitude waveforms, the main arrivals are stable. For the more noisy days, large earthquakes were present in the data.

[16] At long range, the NCF is dominated by the Rayleigh mode. However, for short separations (less than 50 km) spatial variations of the arrival-time of the NCF are observed that are likely due to local inhomogeneities and differences in seismic phases at short ranges. Figure 6 displays time-derivatives and spectrograms of three NCFs at comparable distances (respectively 16, 18 and 19 km). Since the waveforms are band-pass-filtered TDGFs, the peak of the arrival should be used for determining arrival-time. The time series in Figure 6a is for a station pair located in the Los Angeles sedimentary basin (circle in Figure 2) and has a slow group velocity, 0.6 km/s, in agreement with shear speed estimates in the area [*Kohler et al.*, 2003]. The dispersive structure of this NCF is clearly visible in the spectrogram with a high frequency slow arrival. For a station pair located more inland (Figure 6c, triangle in Figure 2) the directionality of the ambient noise is less pronounced. Consequently, the time symmetry is more obvious with arrival-times ±7 s (group velocity 2.6 km/s). The last time series (Figure 6e, cross in Figure 2) is typical of most waveforms in Figure 2 with a 3.0-km/s group velocity and a one-sided arrival structure.

[17] For one station pair (JVA and SDG, separation 182 km, cross in Figure 2) the cross correlation between all three components for each station were computed giving a total of 9 components, Figure 7. The arrival and polarization plots show strong correlations between the vertical and radial components indicating that Rayleigh wave propagation is dominating, in agreement with *Campillo and Paul* [2003].

4. Conclusion

[18] Estimates of the surface wave components of the time domain Green's function can be extracted from the time-derivative of the ambient noise cross correlation between two stations. Broadband noise cross correlations computed over a dense network of station pairs exhibit frequency dispersion and three-dimensional variations of the arrival-time structure. Short-period content is more likely to be recovered for short propagation distances. These waveforms can be used to construct a geophysical model. The observed emergence of the coherent waveforms confirms that they build up proportional to the square root of the recording time.

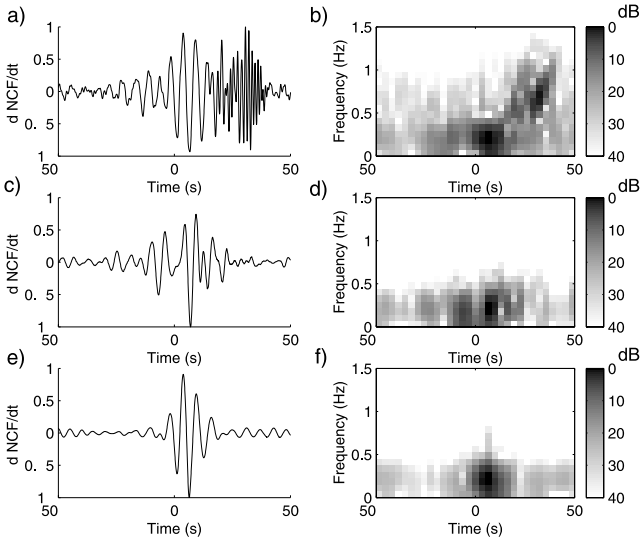

Figure 6. Time series (a, c and e) and spectrogram (b, d and f) for station pairs BRE-OLI (a and b), DGR-MSJ (c and d), and EDL-SDR (e and f).

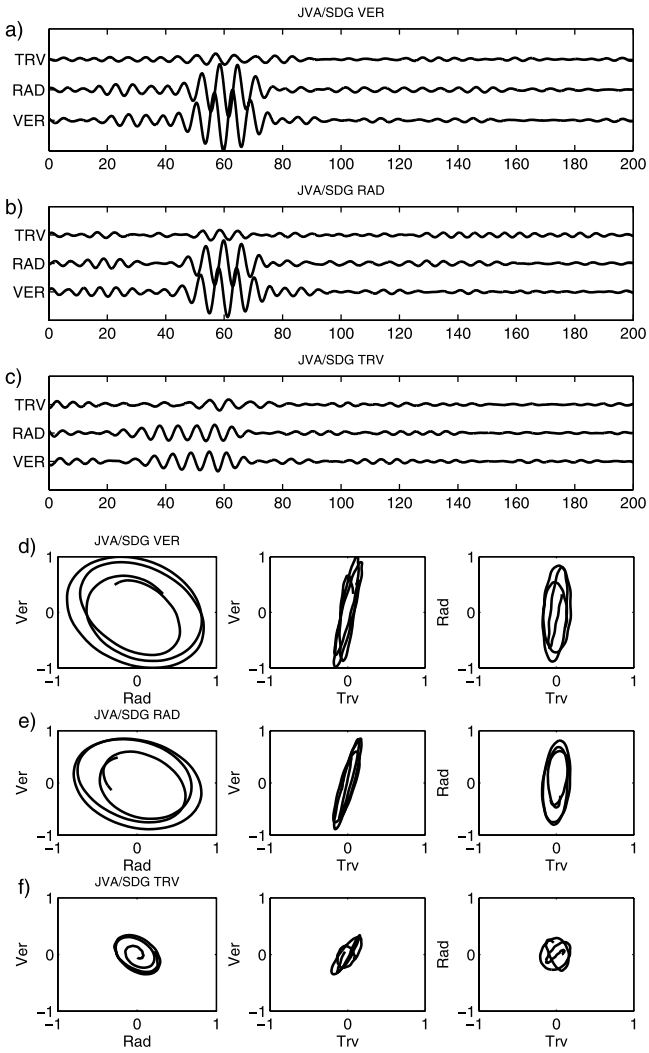

Figure 7. (a, b, c) Time series and d, e, f: Polarization diagram for the JVA-SDG station pair. Cross correlating with JVA vertical component (a, d), radial component (b, e), and transversal component (c, f) with the vertical, radial, and transversal component of SDG.

[19] **Acknowledgments.** Funding was provided by the UCSD/LANL CARE-program. Data came from the Southern California Earthquake Center. We thank George Randall, LANL, for assistance with group velocity calculations and Kees Wapenaar and an anonymous reviewer for constructive comments on the manuscript.

References

Aki, K., and P. G. Richards (1980), *Qualitative Seismology*, W. H. Freeman, New York.

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

Kohler, M. D., H. Magistrale, and R. W. Clayton (2003), Mantle heterogeneities and the SCEC Reference Three-Dimensional Seismic Velocity Model Version 3, *Bull. Seismol. Soc. Am.*, *93*, 757–774.

Larose, E., A. Derode, M. Campillo, and M. Fink (2004), Imaging from one-bit correlations of wideband diffuse wavefields, *J. Appl. Phys.*, *95*, 8393–8399.

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

Malcolm, A. E., J. A. Scales, and B. A. van der Tiggelen (2004), Retrieving the Green function from diffuse equipartitioned waves, *Phys. Rev. E*, *70*, doi:10.1103/PhysRevE.70.015601.

Rickett, J., and J. Claerbout (1999), Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring, *Leading Edge*, *18*, 957–960.

Roux, P., W. A. Kuperman, and the NPAL Group (2004), Extracting coherent wavefronts from acoustic ambient noise in the ocean, *J. Acoust. Soc. Am.*, *116*, 1995–2003.

Roux, P., K. G. Sabra, W. A. Kuperman, and A. Roux (2005), Ambient noise cross correlation in free space: Theoretical approach, *J. Acoust. Soc. Am.*, *117*, 79–84.

Sabra, K. G., P. Roux, and W. A. Kuperman (2005), Arrival-time structure of the time-averaged ambient noise cross-correlation function in an oceanic waveguide, *J. Acoust. Soc. Am.*, *117*, 164–174.

Shapiro, N. M., and M. Campillo (2004), Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.*, *31*, L07614, doi:10.1029/2004GL019491.

Snieder, R. (2004), Extracting the Green's function from the correlation of coda waves: A derivation based on stationary phase, *Phys. Rev. E*, *69*, doi:10.1103/PhysRevE.69.046610.

Wapenaar, K., J. Thorbecke, and D. Draganov (2004), Relations between reflection and transmission responses of three-dimensional inhomogeneous medias, *Geophys. J. Int.*, *156*, 179–194.

Weaver, R. L., and O. I. Lobkis (2003), Elastic wave thermal fluctuations, ultrasonic waveforms by correlation of thermal photons, *J. Acoust. Soc. Am.*, *113*, 2611–2621.

Weaver, R. L., and O. I. Lobkis (2004), Diffuse fields in open systems and the emergence of the Green's function, *J. Acoust. Soc. Am.*, *116*, 2731–2734.

M. C. Fehler, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.

P. Gerstoft, W. A. Kuperman, P. Roux, and K. G. Sabra, Marine Physical Laboratory, University of California, San Diego, La Jolla, CA 92093-0238, USA. (gerstoft@ucsd.edu)