

CONSTRUCTION OF EMPIRICAL GREEN'S FUNCTIONS FROM DIRECT WAVES, CODA WAVES, AND AMBIENT NOISE IN SE TIBET

HUAIJIAN YAO*, XANDER CAMPMAN[†], MAARTEN V. DE HOOP[‡], AND ROBERT D. VAN DER HILST[§]

Abstract. Empirical Green's functions (EGFs) between receivers can be obtained from seismic interferometry through cross-correlation of pairs of ground motion records. Full reconstruction of the Green's function requires diffuse wavefields or a uniform distribution of (noise) sources. In practice, EGFs differ from actual Green's functions because wavefields are not diffuse and the source-distribution not uniform. This difference, which may depend on medium heterogeneity, complicates (stochastic) medium characterization as well as imaging and tomographic velocity analysis with EGFs. We investigate how source distribution and scale lengths of medium heterogeneity influence Green's function reconstruction in the period band of primary microseisms ($t = 10 - 20$ s). With data from a broadband seismograph array in SE Tibet we analyze the symmetry and travel-time properties of EGFs from correlation of data in different windows: ambient noise, direct minor or major arc surface waves, and surface wave coda. The EGFs from these different windows show similar dispersion characteristics, which demonstrates that the Green's function can be recovered from direct wavefields (e.g., ambient noise or earthquakes) or from wavefields scattered by heterogeneity on a regional scale. Late surface wave coda is more diffuse than the early surface wave coda and is generally expected to produce a more symmetric EGF. We show, however, that directional bias is also manifest in EGFs from late coda and that this bias is similar to that in EGFs from ambient noise. This suggests that (in the period band studied) late coda is dominated not by signal from multiple (local) scattering of surface waves but by ambient noise (for instance from oceans). Directional bias and signal-to-noise ratio of EGFs can be understood better with (plane wave) beamforming of the energy contributing to EGF construction. Beamforming also demonstrates that seasonal variations in cross-correlation functions correlate with changes in ocean activity.

Key words. Empirical Green's function; seismic interferometry; ambient noise; surface waves; coda; beamforming.

1. Introduction. Traditional seismic imaging and tomographic velocity analysis of Earth's interior relies on data associated with ballistic (source-receiver) wave propagation. However, over the past few years one has also started to use information contained in seismic coda waves and ambient noise to image the Earth's structure from regional scale to continental scale (Campillo & Paul, 2003, Shapiro & Campillo, 2004; Shapiro et al., 2005; Bakulin & Calvert, 2006; Willis et al., 2006, Yao et al, 2006, 2008; Yang et al., 2007). Modal representation of diffuse wavefields, elastodynamic representation theorems, and stationary phase arguments (Weaver & Lobkis, 2004; Wapenaar, 2004; Snieder, 2004; Paul et al., 2005; Roux et al., 2005; Nakahara, 2006) have been used to argue that the Green's function between the two stations can be estimated from the summation of cross correlations of continuous records of ground motion at these stations. These studies make different assumptions about noise characteristics and (stochastic) properties of the medium. The results of ambient noise cross correlation are analyzed by Colin de Verdière (2006a, 2006b), Bardos et al. (2008), and De Hoop and Solna (2008).

Continuous records of ground motion typically contain seismic energy in several regimes. For example, earthquakes generate deterministic, transient energy that can be registered as distinct phase arrivals by seismometers. Non-smooth medium heterogeneity can, however, complicate waveforms in such a way that they can no longer be described deterministically. After multiple scattering the wave field may become diffuse. This regime is often called the seismic coda, mostly arriving after the ballistic waves (see, for instance, Sato and Fehler, 1998). Outside the time windows containing direct and coda waves from earthquakes continuous records contain energy that is mainly due to continuous processes near and below Earth's surface. This regime is often referred to as ambient

*Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, USA (hjyao@mit.edu)

[†]Shell International E&P B.V., Kessler Park 1, 2288 GS, Rijswijk, The Netherlands

[‡]Center for Computational and Applied Mathematics, and Geo-Mathematical Imaging Group, Purdue University, West Lafayette, IN 47907, USA (mdehoop@purdue.edu).

[§]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, USA

seismic noise. In theory, the cross-correlation-and-summation approach can be applied to each of these regimes to obtain an empirical Green's function (EGF), as long as energy arrives at the two seismic stations from all directions and in all possible modes (assuming equipartitioning).

For simple media cross correlation of the ballistic responses due to sources surrounding two receivers gives the exact Green's function between the receivers (De Hoop & De Hoop, 2000; Wapenaar, 2004). In practice, seismic energy is neither uniformly distributed nor equipartitioned (Malcolm et al., 2004; Sánchez-Sesma et al., 2008; Paul et al., 2005). In field experiments, equipartitioning is generally not achieved because the mode structure of the wave field depends on the mechanism and the location of the noise sources. Moreover, equipartitioned waves are weak and their contribution to the wavefield can easily be overwhelmed by (directional) waves and noise, as shown below. As a consequence, Green's functions are not fully reconstructed, and the accuracy of reconstruction is generally unknown. How well the Green's function is estimated depends on the mechanism and spatial distribution of the noise sources as well as the properties of the medium beneath the receiver arrays. On the positive side, one could exploit this dependence to constrain (stochastic) medium properties (e.g., Scales et al., 2004) if the effects of noise distribution can be accounted for. In this context, the length scale of heterogeneity, the frequency content of the wave fields, and the spatial and temporal spectra of noise sources are all important (De Hoop and Solna, 2008). On the negative side, the (unknown) uncertainty in Green's function construction complicates imaging and, in particular, multi-scale (tomographic) velocity analysis with EGFs.

The problem of incomplete Green's function reconstruction has been recognized before – see, for instance, Yao et al. (2006) for cases of incomplete reconstruction of EGFs for Rayleigh wave propagation) – and practical solutions have been proposed. For active source applications of seismic interferometry, source distributions can be designed with the objective to optimize the retrieval of the Green's function (Metha et al., 2008). In earthquake seismology, where the source configuration cannot be manipulated, one can enhance the illumination of receiver arrays by ballistic waves either by waiting long enough for contributions from a large range of source areas to accumulate or one can make better use of the (continuously) recorded wavefield. For example, as we will show here, the estimation of EGFs for surface wave propagation can be improved by considering not only minor-arc source-receiver propagation (associated with minimum travel time stationarity) but also the time windows relevant for major-arc (maximum travel time) propagation.

To improve the inference of medium properties from EGFs or the imaging or velocity analysis of complex media with EGFs we need a more comprehensive understanding of the relationships between EGFs and medium heterogeneity and properties of noise sources. De Hoop & Solna (2008) present a theoretical framework for the estimation of Green's functions in medium with random fluctuations; and show that EGFs are related to the actual Green's function through a convolution with a statistically stable filter that depends on the medium fluctuations.

Using field observations (from an array in SW China) we investigate here the different contributions of the wavefield to the construction of EGFs through cross correlation. For this purpose we analyze EGFs obtained from windows of ambient noise, direct surface waves, or surface-wave coda. Cross correlation of (direct) surface windows yield EGFs (only) for direct surface wave propagation, but by changing the data window we can manipulate the parts of the wavefield that contribute to the construction of the EGF. Cross correlation of coda waves should yield EGFs that include scattered waves. The latter can also be obtained by correlation of long records of ambient noise. In principle, coda wave and (pure) ambient noise correlation should produce similar EGFs and differences between them can give information about the energy distribution and heterogeneity under and near the array. We complement our analysis with plane-wave beam forming (in the frequency-wavenumber domain), which quantifies the directional energy distribution of the signals that contribute to the EGF. This beamforming analysis reveals (temporal) variations in source regions of ambient noise, which – in turn – help understand the (changes in) symmetry and signal-to-noise ratio (SNR) of the EGFs.

2. Data and Processing. We use 10 months (November 2003 to August 2004) of continuously

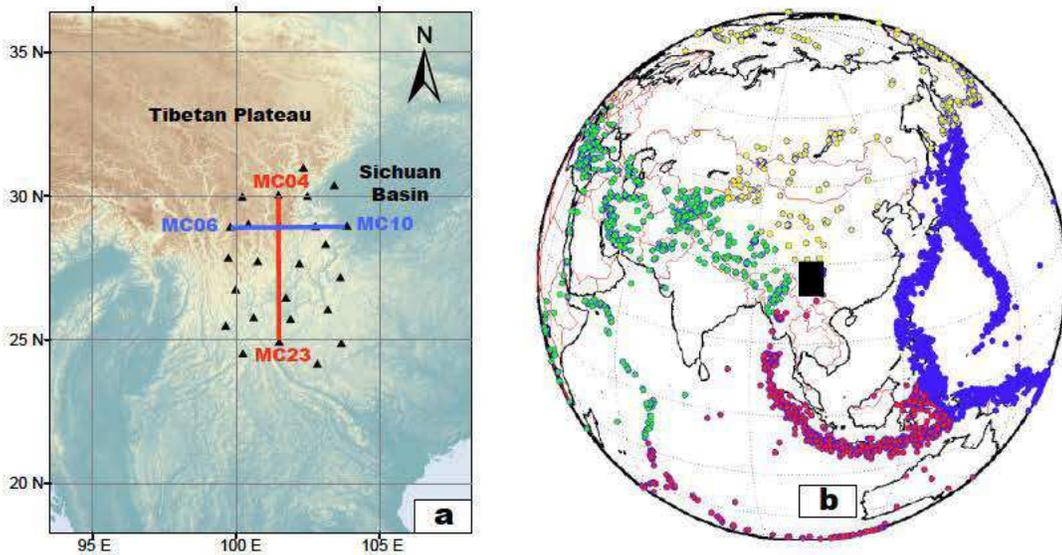


FIG. 1. (a) The location of 25 stations (black triangles) of the MIT-CIGMR array in SE Tibet. The red line and the blue line show the two-station paths for the S-N directional station pair MC04-MC23 and the E-W directional station pair MC06-MC10, respectively. (b) Epicenters of earthquakes with $m_b \geq 4$ that occurred in the 10 months from November 2003 to August 2004 (from EHB catalogue by Engdahl et al., 1998). The black rectangle shows the location of the MIT-CIGMR array. Earthquakes from east (E), south (S), west (W), and north (N) of the array are shown as blue, red, green, and yellow dots. The azimuth angle θ from the center of the array to the earthquake satisfies $-45^\circ \leq \theta \leq 45^\circ$, $45^\circ \leq \theta \leq 135^\circ$, $225^\circ \leq \theta \leq 315^\circ$, $-45^\circ \leq \theta \leq 45^\circ$, for the earthquakes in S, W, N, and E quadrant, respectively.

recorded, vertical component broadband data from a temporary seismograph array in southeastern (SE) Tibet (see Figure 1a). The 25 station array, with average station spacing ~ 100 km, was deployed by MIT and the Chengdu Institute of Geology and Mineral Resources (CIGMR). For more detailed descriptions of the array data and the preliminary results from surface wave array tomography (ambient noise and traditional two-station analysis), which reveal strong heterogeneity in the crust, we refer to Yao et al. (2006, 2008). Figure 1b depicts epicenters of earthquakes with $m_b \geq 4$ that occurred between November 2003 and August 2004. Earthquakes occur predominantly along tectonic plate margins East and South of the MIT-CIGMR array (e.g., near Japan, Philippine, Mariana and Indonesian trenches), while seismicity in the west (Himalayan boundary) and north is much weaker.

Basic data pre-processing includes the removal of the mean and compensation for the instrument response. For the analyses presented here we band-passed the data between periods of 10 – 20 s. Next, we select particular parts of the data (direct surface waves, surface wave coda, and ambient noise, as shown in Figure 2) from the continuous recordings using a group velocity window procedure

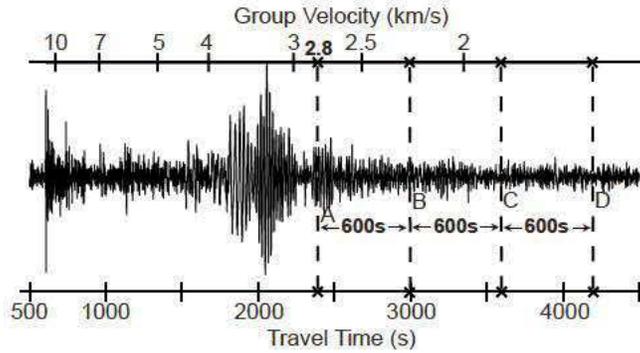


FIG. 2. Illustration of group velocity windows and coda windows used for the EGF retrieval. The seismogram is recorded by the station MC06 shown in Figure 1a and the earthquake is located at $(-5.4^\circ S, 151.5^\circ E)$. The epicenter distance is 6697 km. The bottom and top horizontal axis shows the travel time (with respect to the source time) and the corresponding group velocity (or horizontal propagation speed) of the records, respectively.

(Yao et al., 2006).

Consider seismic waves released by an earthquake with source time, at t_s , recorded by a seismograph station at epicentral distance Δ (km). For any time t after t_s (i.e., $t > t_s$) the corresponding average group velocity (or horizontal propagation speed) for 2-D surface waves $v_g = \Delta/(t - t_s)$, as shown in Figure 2. The example shows a main surface wave within window $v_g = (2.5 - 5.0)$ km/s. By muting (setting the amplitude of the seismic trace to zero) outside or inside a specific group velocity window (e.g., Figure 2), we select specific data windows associated with (known) earthquakes or (unknown) ambient noise.

We apply one-bit cross correlation to the data band-pass-filtered in these windows to obtain the cross correlation function. EGFs are then obtained from the time-derivative of the cross correlation function by $-G_{AB}(t) + G_{BA}(-t) = \frac{dC_{AB}(t)}{dt}$, where $G_{AB}(t)$ ($t \geq 0$) is the causal part EGF at station B for a fictitious (point) source located at A, $G_{BA}(-t)$ ($t \leq 0$) is the anti-causal part EGF at A for a fictitious (point) source at B, and $C_{AB}(t)$ is the one-bit cross correlation between the two stations (Yao et al., 2006). Since for this analysis we use vertical component data we recover predominantly the Green's function for (fundamental mode) Rayleigh wave propagation. Similarly, Love waves can be recovered from transverse component data (Campillo & Paul, 2003; Paul et al., 2005; Lin et al., 2008).

3. EGFs from different data windows. In a heterogeneous medium, the Green's function for wave propagation between two points contains contributions from scattering anywhere in the medium – not just from structure located between these points. EGFs are estimates of the Green function obtained from correlation and summation of the diffuse wavefields recorded at two receivers. How well the EGF reconstructs the actual Green function strongly depends on the characteristics of the energy in the wavefields used. EGFs from cross correlation of field data usually show a strong dependence on (non uniform) energy distribution (Yao et al., 2006; Yao and Van der Hilst, in preparation).

In this section we evaluate EGFs extracted from one-bit cross correlation of data in different time windows. From the continuous records, we extract data associated with ambient noise, direct surface waves, and surface wave coda using the group velocity window procedure described above. The group velocities that successfully separate these windows depend on factors such as the magnitude of the earthquake, epicentral distance, heterogeneity, etc., and cannot be defined uniquely (Turner, 1998, Malcolm et al, 2004). Our choice is, arguably, ad hoc but the analysis presented here is not sensitive to the precise values of the bounding group velocities. We illustrate our analysis with data from two station pairs (Figure 1a): MC04-MC23 is a N-S (north-south) pair with inter-station

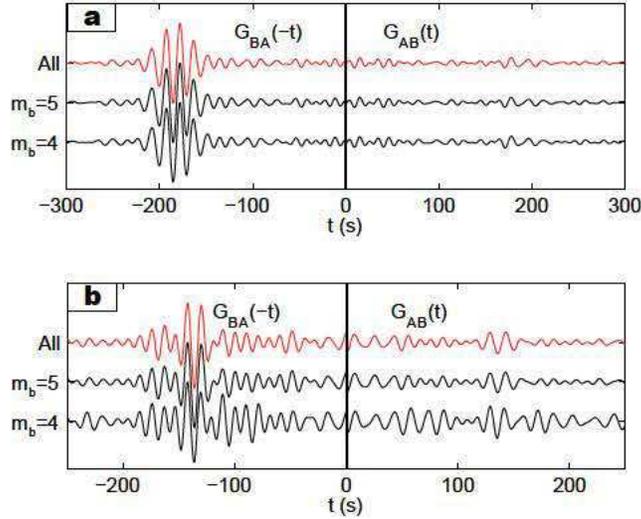


FIG. 3. EGFs (black traces) for the station pair (a) MC04-MC23 and (b) MC06-MC10 in the period band 10 – 20 s from approximately ambient noise after muting the wave trains in the group velocity window 2–10 km/s from earthquakes larger than the cut-off magnitude ($m_b = 5$ or 4, shown at the left side of each black trace) for all the 10 months data. The red traces in (a) and (b) are the EGFs from all the 10 months continuous data. The causal part shows for the EGF recorded at the station MC23 (or MC10) generated by a fictitious source at the station MC04 (or MC06) while the anti-causal part for recordings at MC04 (or MC06) with the source at MC23 (or MC10), same as shown in Figures 4, 5, 6, 7, and 10a.

distance ~ 570 km; MC06-MC10 (E-W, ~ 400 km).

3.1. EGFs from all continuous data. For reference, we first calculate EGFs for the two station pairs from one-bit cross correlation of the entire 10-month record, shown as the red trace in Figures 3a,b for MC04-MC23 and MC06-MC10, respectively. Like other normalized cross correlation methods (e.g., Bensen et al., 2007) one-bit cross correlation normalizes the energy of all sources contributing to the construction of the EGF, so that the average energy flux is an indicator of the number (or normalized strength) of these sources, not their real magnitude. For both station pairs, the EGFs reveal Rayleigh wave arrivals with group speed around 3 km/s. Neither EGF is time-symmetrical, however, and the amplitude (or SNR) in the anti-causal part is much larger than in the causal part. For MC04-MC23 the 10-month average energy flux seems much higher from S to N (which contributes to the recovery of the anti-causal part of the EGF) than from N to S (the causal part). For MC06-MC10 the average energy flux in the 10 months is larger from E to W than from W to E.

3.2. EGFs from ambient noise. In the previous section we used continuous 10-month records. In this section and the next, we partition the data in specific energy propagation regimes (direct waves, coda waves, and ambient noise). The group velocity window procedure described above allows us to obtain EGFs (mostly) from ambient noise by suppressing signals associated with large earthquakes (see also Yao et al., 2006). Most direct body waves and surface waves, as well as their coda, appear in the group-velocity window 2 – 10 km/s (see Figure 2). Using earthquake origin times t_s from the EHB catalog by Engdahl et al., (1998) we suppress the amplitude of signal within the 2 – 10 km/s group velocity windows for earthquakes larger than certain magnitude. One-

bit cross-correlation to the remaining signals is then used to extract EGFs (approximately) from ambient noise.

Note that ambient noise is here defined as all seismic energy unrelated to earthquakes larger than the cut-off magnitude. Thus defined, ambient noise contains contributions from small earthquakes, but the smaller the cut-off magnitude the closer the remaining seismograms are to ambient seismic noise proper. The energy from such a source distribution approximately corresponds to the diffuse wave field theoretically required for accurate Green’s function construction. In this study we set the smallest cut-off magnitude to $m_b = 4$, because many earthquakes smaller than $m_b = 4$ are not listed in the EHB catalogue and recorded signals from those small earthquakes are usually below the ambient noise level due to the attenuation and geometrical spreading over a few thousand kilometers.

EGFs obtained from 10-month records of ambient noise, as defined above, are shown as the black traces in Figure 3 for two cut-off magnitudes. These EGFs are almost identical to the EGFs from the continuous 10-month records (red traces in Figure 3). This implies that in the period band considered (10 – 20 s) the contributions from large earthquakes is small compared to that from ambient noise, as expected from one-bit cross correlation (see also Bensen et al., 2007). This also implies that the asymmetry of the EGFs is not caused by non-uniform distribution of large earthquakes but (for the time period considered) by ambient noise directionality, with most noise sources in South and East. Furthermore, tests (not shown here) with 1-month records showed that variations of EGFs over time are not related to the temporal variations in earthquake activity. In fact, (plane wave) beam forming with the EGFs (see Section 4 below) demonstrates that the temporal changes in EGF symmetry and amplitude are related to seasonal variations of ocean microseisms (see also Stehly et al. 2006, Pedersen et al., 2007). Together, these results suggest that for $t = 10 - 20$ s ambient noise is dominated by primary microseisms, which are usually attributed to coupling of oceanic wave energy into seismic energy in the Earth in shallow waters (Cessaro, 1994; Bromirski et al., 2005).

3.3. EGFs from direct surface waves. In earthquake seismology, sources are non-uniformly distributed along plate boundaries (Figure 1b) and Green’s function reconstruction from direct waves is often incomplete. To study the symmetry properties of the EGFs from direct surface waves the data selection is opposite of that of the previous section. Here we suppress signal outside and keep the data inside the 2.5 – 5 km/s group velocity window (calculated for earthquakes with $M_w \geq 5$ anywhere in the world). This window contains mainly the (dispersive) fundamental surface wave mode (Figure 2). From stationary phase analysis it is easily understood that the strongest contribution for a particular station pair comes mainly from sources located on or near the line connecting the stations (Snieder, 2004). We can, therefore, choose the direction from which we want sources to contribute for a given seismic station pair. To this end, we divide the earthquake source regions into East, South, West and North quadrants (Figure 1b). As before, we applied one-bit normalization to the records before cross correlations.

For both station pairs, the EGFs from all earthquake data (Figure 4, black traces labeled ‘ESWN’) show a similar time-asymmetry as EGFs from the 10-month continuous data (Figures 3 and 4, red trace). For MC04-MC23 the anti-causal part of the EGF from earthquake data in each quadrant is similar to the anti-causal part from all data (Figure 4a). However, the causal part (that is, surface waves propagating from N to S) can be only recovered from the earthquakes located north of the array (yellow circles in Figure 1b). Seismicity in the North is relatively low but we still observe a causal phase around the same time as the reference phase (Figure 4a, blue trace). The causal EGF is, however, much noisier than the anti-causal part due to the sparse event distribution in the north. For the E-W station pair we can make similar observations (Figure 4b). The anti-causal EGF from earthquakes in the East, South, and North are, again, similar to that from all data. Data from events in the west produce both a causal and anti-causal part (Figure 4b, black trace labeled ‘W’), even though the latter is substantially weaker. This demonstrates that we can indeed recover the (anti-) causal parts of the surface wave Green’s function by using earthquake

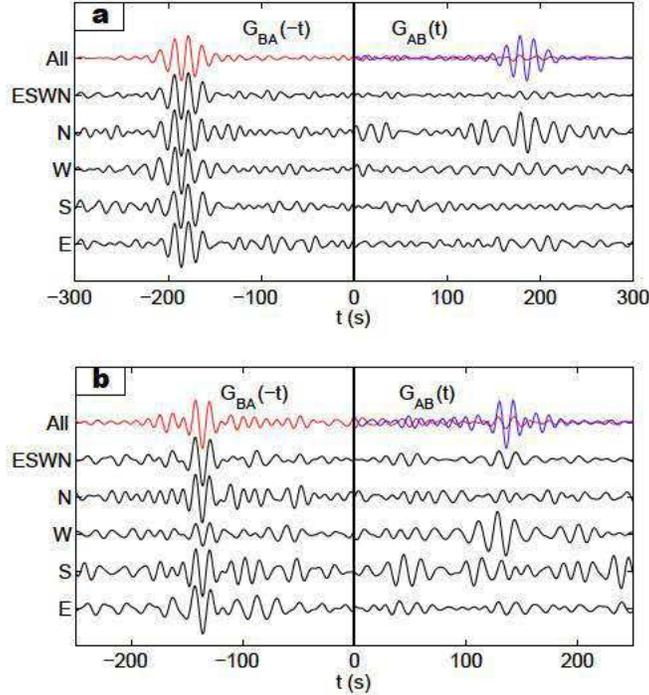


FIG. 4. EGFs (black traces) in the period band 10 – 20s for the station pair (a) MC04-MC23 and (b) MC06-MC10 only from direct (minor-arc) surface waves in the group velocity window 2.5-5 km/s from all the earthquakes with $M_w \geq 5$ in the 10 months everywhere in the world (labeled ‘ESWN’ at the left side), and from four different directions (labeled ‘E’, ‘S’, ‘W’, and ‘N’ at the left side; for location of each quadrant, see Figure 1b). The red trace in (a) and (b) is the same as that shown in Figure 3a and b, respectively. The blue trace in (a) and (b) is the time reversal of the anti-causal EGF of the red trace, shown for comparison with the causal EGF.

data from a specific direction.

The fact that for both the N-S and E-W station pair we can recover (anti-causal) surface wave EGFs for all seismicity quadrants is surprising. In principle, energy from directions perpendicular to the geographical orientation of the receiver pair contributes little to the Green’s function of (surface) wave propagation between them. We speculate that the successful recovery of EGFs is due to presence of ambient noise energy in the surface wave records (which, as mentioned above, seems to originate mainly in South (Indian Ocean) and East (Pacific Ocean)).

Major-arc surface waves (associated with maximum time paths) sample receiver arrays from the opposite direction as the more commonly used minor-arc (minimum time) waves, and they have been used to improve the resolution of tomography with ballistic surface waves (Levshin et al., 2005). Symmetry in EGFs of surface wave propagation between two receivers can be restored by including major-arc surface wave windows in the correlation. Consider for example Figure 4b, that is, for the station pair MC06-MC10 in the E-W direction. The trace labeled ‘E’, obtained by correlation of the responses from earthquakes with $M_w \geq 5.5$ in the East, reveals a one-sided (anti-causal) EGF. However, correlation of data windows that contain major-arc surface waves (traveling in opposite direction along the station pair) improves the retrieval of the causal Green’s function. We remark that since the major-arc surface waves are much weaker than the minor-arc surface waves, the contribution from ambient noise energy is substantial. This explains why, for this example, also the anti-causal part is recovered well (even better than the causal part).

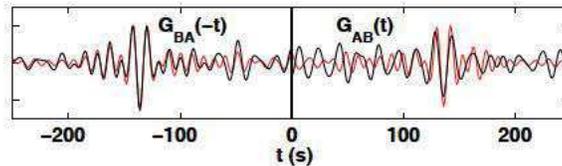


FIG. 5. EGF (black trace) in the period band 10 – 20 s for the station pair MC06-MC10 from the major-arc surface waves in the group velocity window 2.5-5 km/s of earthquakes only located in the east quadrant (Figure 1b) with $M_w \geq 5.5$ in the 10 months. Both the anti-causal and causal parts of the red trace are the anti-causal EGF using all the 10 months data as shown in Figure 3b.

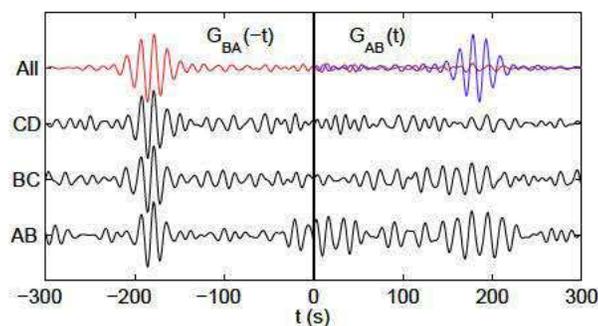


FIG. 6. EGFs (black traces) in the period band 10 – 20 s for the station pair MC04-MC23 from surface wave coda of earthquakes with $M_w \geq 5$ in the 10 months in three different time windows (AB, BC, and CD) shown at the left side of each black trace. The surface wave coda window starts at the point A with group velocity 2.8 km/s (Figure 1b) for each earthquake. Three time windows AB, BC, and CD all have 600 seconds in length. The red and blue traces are the same as those shown in Figure 4a.

3.4. EGFs from coda waves. Independent of the source distribution, one can improve conditions for Green’s function construction by exploiting wavefield scattering due to medium heterogeneity (Paul et al., 2005; Campillo & Paul, 2003). Coda waves are due to (multiple) scattering in the shallow subsurface (Sato & Fehler, 1998) and can be divided into two regimes (Malcolm et al., 2004): an earlier diffusion regime and a later equipartitioning regime. The equipartitioning regime is theoretically the optimal regime for interferometric Green’s function reconstruction because no preferred direction and mode of propagation exists (Van Tiggelen, 2004).

For surface wave applications in solid Earth seismology the diffusion regime is usually found in the (late) coda of direct S (Campillo & Paul, 2003; Paul et al., 2005) or Rayleigh waves (Langston, 1989). Equipartitioning has indeed been observed in late coda waves from short-period S waves (Hennino et al., 2001), but the associated energy usually falls below the ambient noise level because it arrives many mean-free times after the direct waves. As a consequence, EGFs from late coda may show the same directional bias as EGFs from ambient noise (e.g., Paul et al., 2005).

Using the S-N station pair MC04-MC23 as example, we investigate the correlations of coda from all earthquakes with $M_w \geq 5$ in three 600 s long coda windows (Figure 2). We define the early, middle, and late coda by, respectively, the first, second, and third 600-s time window (AB, BC, and CD) after the recording time with group velocity of 2.8 km/s (Figure 2).

The trace obtained by correlating data from the third window (CD) in Figure 6 shows the EGF from late coda for the S-N station pair MC04-MC23. Clearly, the energy propagates predominantly from S to N in the late coda rather than in all directions, as we would expect for an equipartitioned, isotropic wave field. In fact, it is similar to the EGF from all 10 months (ambient noise) data, marked as ‘All’ in Figure 6, with a dominant anti-causal part. In Section 3.2, we attributed this

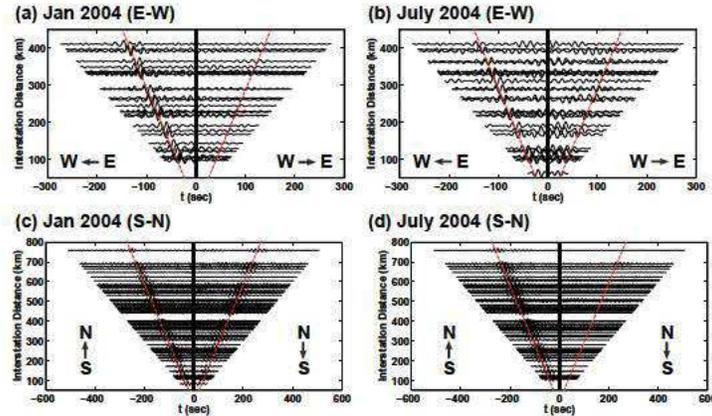


FIG. 7. Comparison of cross correlation functions in the period band 10 – 20 s from one month data of (a) January 2004 for E-W directional two-station pairs, (b) July 2004 for E-W directional two-station pairs, (c) January 2004 for S-N directional two-station pairs, and (d) July 2004 for S-N directional two-station pairs. Here the E-W or S-N directional two-station pairs are station pairs directed roughly from East to West or North to South with a maximum of 15° deviation. ‘E→W’ means the fictitious noise sources approximately to the east (E) of the array generate waves propagating to the west (W) for the retrieval of the anti-causal EGF as shown in (a) or (b), similarly for ‘W→E’ but for the retrieval of the causal EGF in (a) and (b), ‘S→N’ for the anti-causal EGF in (c) and (d), and ‘N→S’ for the causal EGF in (c) and (d).

directional bias to the dominating energy flux from the South Indian Ocean, as expected from the observations of Paul et al. (2005).

Traces from the early coda window (AB) and middle coda window (BC) are also shown in Figure 6. Similarly, we construct the anti-causal part of the EGF, that is, the surface wave propagating from S to N, well from these windows. However, we are now able to construct the anti-causal (surface wave) EGF propagating from N to S from these windows, especially for the early coda window (AB) when we observe significant scattered energy above the ambient noise level as observed in the trace for late coda window (CD). This implies scattered energy has not attenuated yet and does contribute to the EGF construction. But, the low SNR in the causal EGF for the coda windows AB and BC suggests scattering has not produced a sufficiently diffuse wave field necessary to accurately construct the Rayleigh-wave phase.

4. Seasonal variability and origin of ambient noise energy. The energy density and distribution of ambient noise – and as a consequence, the reliability of EGFs from wavefield cross-correlation – varies with frequency and time. In this section we investigate the temporal changes in the directional distribution and origin of ambient noise energy (in the period band 10 – 20 s) with respect to the MIT-CIGMR array in SE Tibet. We first analyze the variations of the amplitude of one-bit cross-correlation functions (CFs) over time (Figures 7 and 8). Subsequently we perform a (frequency-wavenumber) beamforming analysis in order to constrain the temporal variations in the geographical origin of the ambient noise energy (Figure 9).

As in Stehly et al. (2006), we analyze the symmetry and amplitude of CFs using data band-passed between 10 – 20 s (the frequency band of the primary microseisms) during different seasons. We correlate one month of continuous records during the northern hemisphere summer (July 2004) and northern hemisphere winter (January 2004) for station pairs directed roughly from north to south and east to west (with 15° deviation). In the winter, the CFs for the E-W station pairs are dominated by energy traveling from the east, as is evident from the one-sided CFs (Figure 7a). For the E-W pairs, the summer CFs (Figure 7b) have lower SNR than in the winter but not seem to have a preferred direction, and (weak) very early arrivals become apparent. The CFs calculated for the N-S station pairs show fairly good symmetry in winter (see Figure 7c) indicating a similar energy flux into the array from the south or north. In summer time (Figure 7d). The apparent

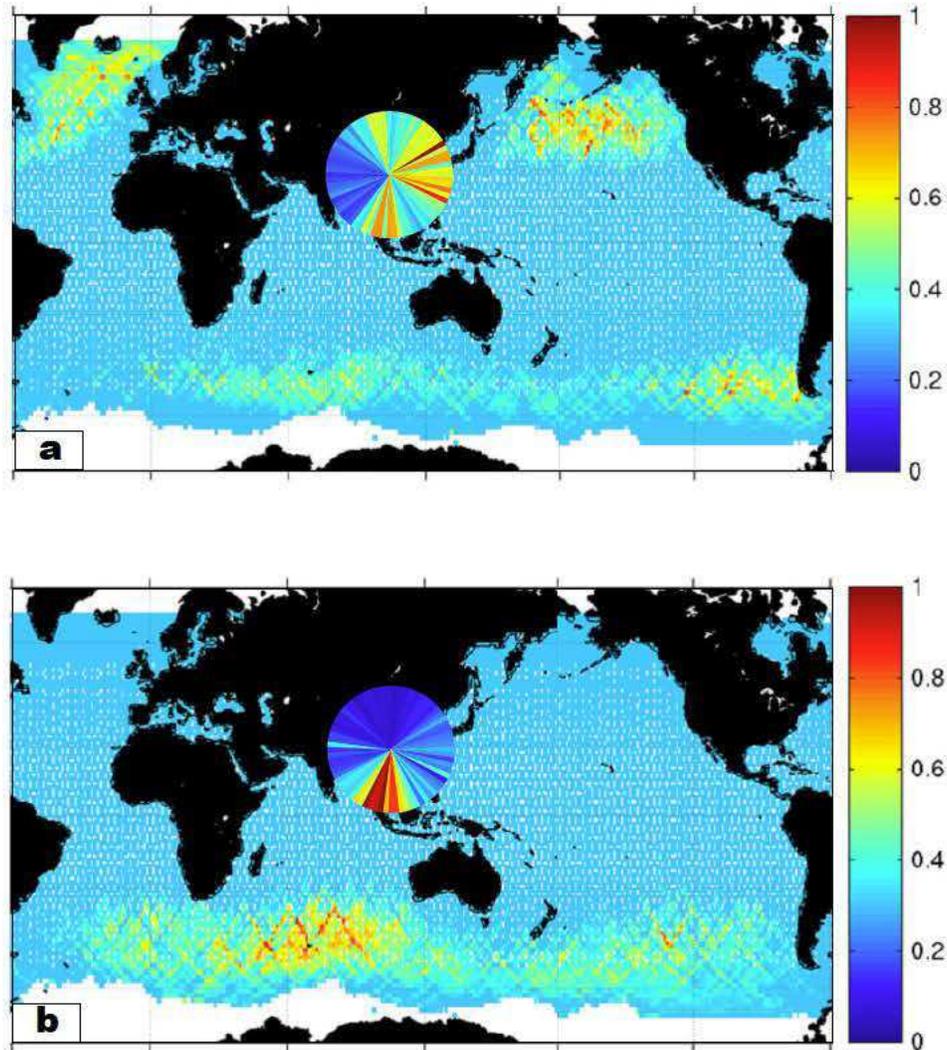


FIG. 8. Seasonal variation of the azimuthal dependence of the normalized amplitude of the cross correlation functions (shown as the pie chart) for all possible MIT-CIGMR array station pairs: (a) in northern hemisphere winter time (January 2004) and (b) in northern hemisphere summer time (July 2007). The pie charts are constructed using the procedure from Stehly et al. (2006) by averaging the amplitude of all CFs in each azimuthal sector (5° width here) with a geometrical spreading amplitude correction considering the difference in interstation distance. The background image shows the distribution of the normalized global ocean wave height in winter time (a) and in summer time (b) (modified after Stehly et al., 2006). The color bar in the right gives the value of normalized amplitude for both cross correlation functions and the ocean wave heights. In the pie chart, the red sector at certain azimuth angle approximately implies that more energy is coming from that azimuth angle and propagating into the array (center of the pie chart).

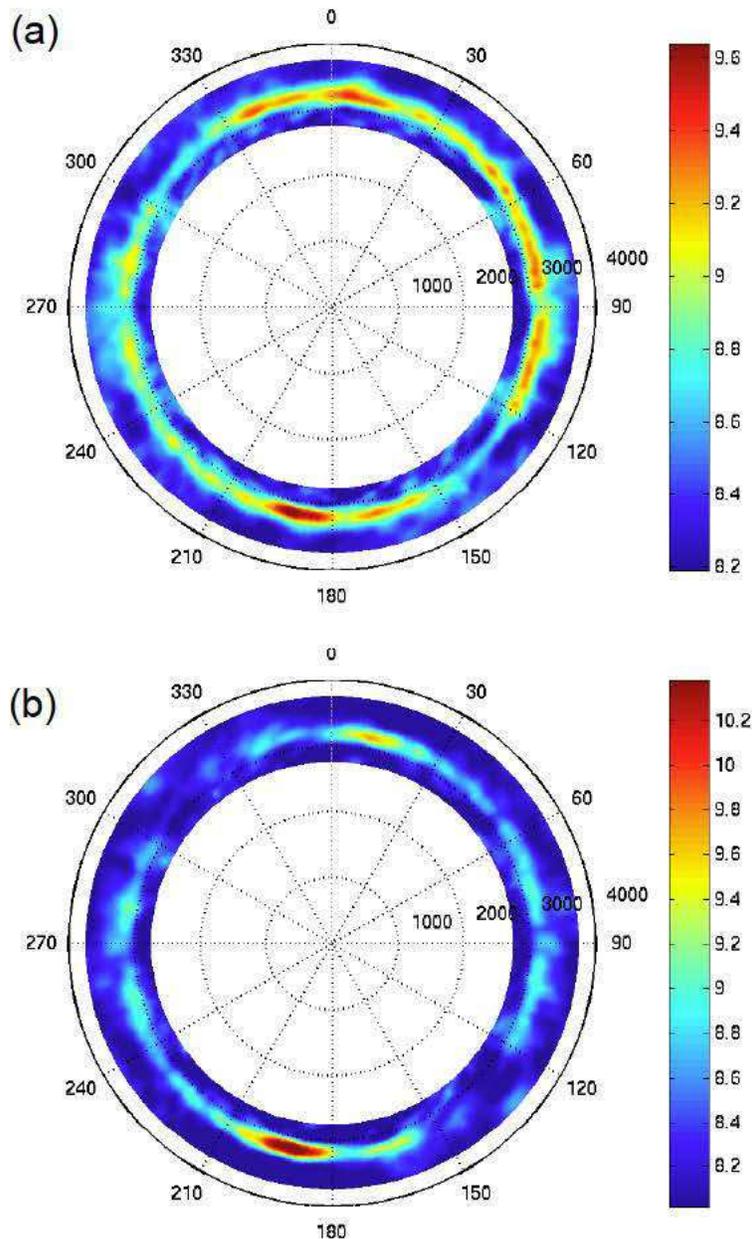


FIG. 9. (a) Noise power from beamforming analysis in January 2004, for the period band 10 – 20 s. The noise mainly arrives from the south-southwest and from between the north-northwest and east-southeast. The apparent velocity is around 3200 m/s. b) Noise power from beamforming analysis in July 2004, for the period band 10 – 20 s. The noise mainly arrives from the south-southwest. The apparent velocity is around 3200 m/s.

asymmetry of the CFs indicates that energy coming from the south is much larger than from the north.

The traces in Figure 7 correspond to E-W and N-S station pair orientations, but pie charts illustrate the azimuthal dependence of the normalized amplitude of the CFs (or ambient noise energy flux) for all station pairs, both for winter (Figure 8a) and summer (Figure 8b). The background

image in Figure 8 shows the distribution of the normalized global ocean wave height, modified after Stehly et al. (2006). The pie charts show that the ambient noise energy in the winter (Figure 8a) is more uniformly distributed than in the summer (Figure 8b). In the winter, noise energy is dominant in the east and north-east directions (possibly related to enhanced wave power in the Northern Pacific) and also from the south (Indian Ocean) and the north (Northern Atlantic). In the summer, the main direction of the ambient noise energy is from the south-south west, pointing to an origin in the Indian Ocean. These results are consistent with the observations of Stehly et al. (2006) and Yang & Ritzwoller (2008).

To confirm, quantify, and interpret the above illustration of seasonal CF amplitude variations, we perform a wavenumber-frequency analysis of the same data. Wavenumber-frequency analysis of random noise fields decomposes the wave field into plane waves, which allows one to characterize the noise wave field – or the wavenumber-frequency power-spectral density – by an azimuth and apparent slowness (or velocity) (Lacoss et al. 1969, Aki & Richards, 1980, Johnson & Dudgeon, 1993). We divide approximately one month of data (January 2004 or July 2004) into 512 s windows with an overlap of 100 s. Using the algorithm due to Lacoss et al. (1969) we beamform the data in these windows for 20 central periods between 10 and 20 s using a narrow band-pass filter of about 0.002 s. The angle resolution is 2 degrees, while the velocity resolution is 20 m/s. The beamforming results in all time windows and frequency bands are then normalized and stacked to produce the final images of the power of the noise wave field in the period band 10 – 20 s in terms of velocity in m/s along the radial axis and azimuth in degrees, along the angle, shown in Figure 9.

Figures 9a and 9b show the noise power during January 2004 and July 2004, respectively. The wave field is dominated by energy coming from the south-south west during the July 2004 (Figure 9b), in excellent agreement with results of the above analysis of CF amplitudes (Figure 8b). The apparent velocity is around 3200 m/s, which agrees very well with the velocities obtained from dispersion analysis (see Figure 10b). The noise power during January 2004 has less obvious directionality (Figure 9a). The same direction in the south-south east causes arrivals with velocities around 3200 m/s, but significant energy also arrives from the north and east with approximately equal amounts and much weaker energy flux from the west. This is also similar to the result from the above CF analysis (Figure 8a). Overall, the noise power in the January is less than during July.

The above observations that the CFs for E-W station pairs have a lower SNR in the summer (Figure 7b) than in the winter (Figure 7a) and that early arrivals appear in the summer time CFs may both be explained by the overall dominance of energy from the south in the summer, as established by the beamforming. If plane waves arrive from the south-south west at an E-W station pair, the result will be an arrival with very high apparent velocity (and thus early arrival time).

5. Discussion. In Section 3 we evaluated the recovery of (surface wave) Green’s functions from ambient noise, direct surface waves, and surface wave coda (for $t = 10 - 20$ s). Figure 10a shows the EGFs from these different data windows for the S-N station pair MC04-MC23. The EGFs for the different data windows give similar phase arrival times for both the causal and anti-causal part of the EGFs (less than 1 s difference if measuring the peak travel time for each trace around ± 178 s in Figure 10a). In Figure 10b we give the dispersion curves for each of these EGFs for $t = 12 - 18$ s. For the anti-causal part of the EGFs, the variation in phase velocities over the 12 – 18 s band is about 0.5% or less. For the causal part (the 2nd and 4th traces in Figure 10a), the phase velocity difference is still quite small (within 1%) at most periods with respect to that of the anti-causal part. The difference among the phase velocities from different causal or anti-causal part of the EGFs reflects the difference of source distribution and energy for the construction of surface wave Green’s functions through cross correlation.

For S to N wave propagation along MC04-MC23 the anti-causal part of the EGF can be reconstructed equally well from each of the different data windows. The causal part of the EGF (representing N-S wave propagation) can be reconstructed from earthquake-generated direct surface waves (second trace, Figure 10a), but due to the much sparser distribution of the earthquakes to the north of the array (Figure 1b) its recovery is worse than that of the anti-causal EGF. Scattered

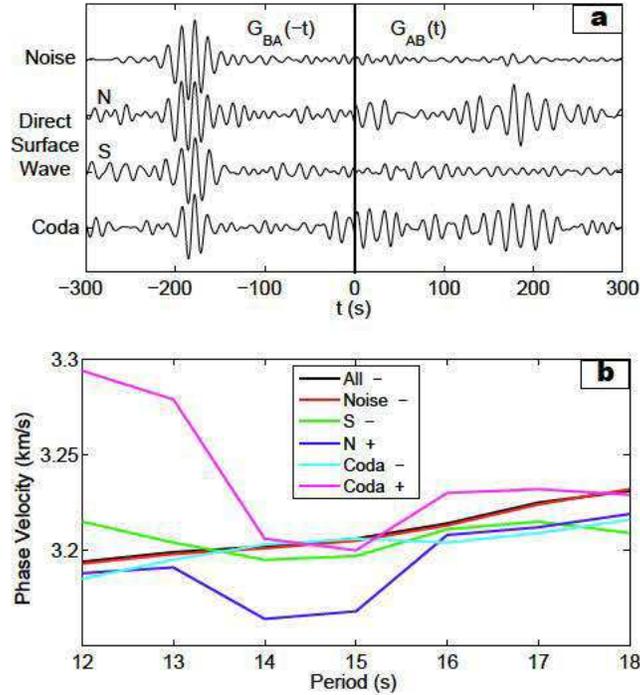


FIG. 10. (a) Comparison of the EGFs of the S-N directional two-station pair MC04-MC23 constructed from the one-bit cross correlation of three different data windows, i.e., ambient noise (top trace labeled 'Noise', same as the bottom trace in Figure 3a labeled ' $m_b = 4$ '), direct surface waves (middle two traces labeled 'Direct Surface Wave', same as the two traces in Figure 4a labeled 'N' and 'S'), and surface wave coda (bottom trace labeled 'Coda', same as the bottom trace in Figure 6 labeled 'AB'). (b) Comparison of phase velocity dispersion in the period band 12-18 s of the EGFs in (a) from different data windows for the station pair MC04-MC23. In the legend box, 'All' means the EGF from all the 10 months continuous data, shown as the red trace in Figure 3a; 'Noise' means the EGF shown as the top trace in (a); 'N' (or 'S') means the EGF shown as the middle two traces in (a); 'Coda' means the EGF shown as the bottom trace in (a); '+' and '-' denote the causal and anti-causal part of the EGF, respectively.

surface wave (coda) wavefields propagating from N to S give better recovery of the causal part for periods ≥ 14 s (Figure 10b) but is less reliable for periods less than 14 s as inferred from the dispersion (Figure 10b) of the causal part EGF from coda (Figure 10a). This frequency dependence of Green's function recovery may be due to frequency-dependent attenuation of the scattered waves.

By changing the group velocities that define the data windows we effectively manipulate the character of seismic energy that contributes to the construction of the EGF. This, in turn, also alters the type of information that can be retrieved about the medium. The early coda is expected to be dominated by single scattering, whereas in the late coda, multiple scattering contributes to the diffusion of energy. In theory a diffuse wavefield should produce a more symmetric EGF (Sanchez-Sesma et al., 2008, Malcolm et al, 2004). However, in reality, multiply scattered energy decays faster and can quickly fall below the noise level, especially for the range of interstation distances considered in our study. This explains our observation that even the EGFs from data in the late coda are not symmetric but show similar effects of directionality (manifest in differences between causal and anti-causal parts) as EGFs retrieved from ambient noise.

The strength of scattering determines the level of spreading of energy and, as a consequence, the symmetry and SNR of the EGF. The strength of scattering is determined by the scale length

and strength of the medium heterogeneity. A comparison of EGFs from different coda windows should therefore allow one to estimate the heterogeneity of the medium (e.g., Malcolm et al, 2004) by observing the emergence of the causal and anti-causal EGF (that is, the symmetry of the EGF).

Our study illustrates that a comparison of EGFs extracted from different regimes in the seismic trace is complicated by various factors. Much depends on the frequency band one uses for the correlations. For periods between 10 and 20 s ambient noise is dominated by the primary microseism and effects of scattering are relatively weak. For shorter periods, scattering is stronger (due to the shorter wavelength compared to heterogeneity) and ocean generated ambient noise may be weaker if the array is far from the coastline. For shorter periods we may, therefore, expect to retrieve more symmetric EGFs from late coda data for station pairs with shorter distance considering high attenuation at shorter periods. At longer periods, say, from 20 to 120 s, the effect of scattering (Langston, 1989) is less and ambient noise energy generally shows much weaker directionality (Yang & Ritzwoller, 2008) or even without directivity (Pederson et al, 2007). Therefore, in this period band one would mainly rely on direct waves and noise to retrieve Green's functions.

The quality of Green's function recovery relies on the azimuthal distribution of coherent sources (both direct sources or scatters) and their (normalized) strength. Generally, wavefields propagating in or near the orientation of the two-station pair will help EGF construction but contributions from directions perpendicular to the station pair will generate bias or noise in the EGFs if the source distribution is not sufficiently uniform to cause destructive interference. In practice, one can steer the known sources (e.g., larger earthquakes) within the regime of constructive interference to recover the Green's function better. The steering process may include both the selection of sources and compensation of source energy to enable the perfect recovery.

6. Conclusions. We demonstrated that the surface wave empirical Green's function can be retrieved from cross-correlation of different data windows (ambient noise, direct surface waves, or surface wave coda) using array data from SE Tibet. Phase velocity dispersion also reveals very similar dispersion characteristics of these empirical Green's functions. By examining the symmetry and amplitude of the cross-correlation functions and performing a frequency-wavenumber beamforming analysis, we conclude that the dominant ambient noise field in the period band 10 – 20 s is from the ocean activities and shows clear seasonal dependence. The average phase velocity between 10 – 20 s of the study area from beam forming analysis is very similar to what we obtained from dispersion analysis. The directionality of ambient noise energy distribution seems to have a large effect on the recovery of the Green's function, especially when one tries to recover the Green's function from late coda which tends to be more diffuse to recover the symmetric Green's functions but is easily overwhelmed by ambient noise fields in reality. Wavenumber-frequency analysis of the noise wave-field has the potential to help in interpreting the Green's function obtained from cross-correlation.

REFERENCES

- [1] AKI, K. & RICHARDS, P.G., 1980. *Quantitative Seismology, Theory and methods*. Vol.1, W.H. Freeman. San Francisco, CA.
- [2] BAKULIN, A., & CALVERT, R., 2006. *The virtual source method: theory and case study*, *Geophys.*, 71(4), SI139–SI150.
- [3] BARDOS, C., GARNIER, J. & PAPANICOLAOU, G., 2008. *Identification of Green's's functions singularities by cross correlation of noisy signals*, *Inverse Problems*, 24, 015011.
- [4] BENSON, G. D., RITZWOLLER, M. H., BARMIN, M. P., LEVSHIN, A. L., LIN, F., MOSCHETTI, M. P., SHAPIRO, N. M., YANG, Y., 2007. *Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements*, *Geophys. J. Int.*, 169, 1239–1260.
- [5] BROMIRSKI, P. D., DUENNEBIER, F. K. & STEPHEN, R. A., 2005. *Mid-ocean microseisms*, *Geochem. Geophys. Geosys.*, 6, Q04009, doi:10.1029/2004GC000768.
- [6] COLIN DE VERDIÈRE, Y., 2006a. *Mathematical models for passive imaging I: general background*. <http://fr.arxiv.org/abs/math-ph/0610043/>
- [7] COLIN DE VERDIÈRE, Y., 2006b. *Mathematical models for passive imaging II: effective Hamiltonians associated to surface waves*. <http://fr.arxiv.org/abs/math-ph/0610044/>
- [8] CAMPILLO, M., PAUL, A., 2003. *Long-Range correlations in the diffuse seismic coda*, *Science*, 299, 547–549.
- [9] CESSARO, R. K., 1994. *Sources of primary and secondary microseisms*, *Bull. Seism. Soc. Am.*, 84, 142–148.

- [10] CORREIG, A. M. & URQUIZÚ, M., 2002. *Some dynamical characteristics of microseismic time-series*, Geophys. J. Int., 149, 589–598.
- [11] DE HOOP, M.V. & DE HOOP, A.T., 2000. *Wave-field reciprocity and optimization in remoting sensing*, Proc. R. Soc. Lond. A (Mathematical, Physical and Engineering Sciences), 456, 641–682.
- [12] DE HOOP, M. V. & SOLNA, K., 2008. *Estimating a Green's function from field-field correlations in a random medium*, SIAM J. Appl. Math., in press.
- [13] ENGDahl, E.R., VAN DER HILST, R.D. & BULAND, R.P., 1998. *Global teleseismic earthquake relocation from improved travel times and procedures for depth determination*, Bull. Seism. Soc. Am., 88, 722–743.
- [14] GU, Y. J., DUBLANKO, C., LERNER-LAM, A., BRZAK, K. & STECKLER, M. 2007. *Probing the sources of ambient seismic noise near the coasts of Southern Italy*, Geophys. Res. Lett., 34, L22315, doi:10.1029/2007GL031967.
- [15] HENNINO, R., TRÉGOURÈS, N., SHAPIRO, N., MARGERIN, L., CAMPILLO, M., VAN TIGGELEN, B. & WEAVER, R. L., 2001, it Observation of equipartition of seismic waves in Mexico, Phys. Rev. Lett., 86, 3447–3450.
- [16] LANGSTON, C. A., 1989. *Scattering of long-period Rayleigh waves in Western North America and the interpretation of coda Q measurements*, Bull. Seism. Soc. Am., 79, 774–789.
- [17] LEVSHIN, A. L., BARMIN, M. P., RITZWOLLER, M. H., TRAMPERT, J., 2005. *Minor-arc and major-arc global surface wave diffraction tomography*, Geophys. J. Int., 149, 205–223.
- [18] LIN, F.-C., M.P. MOSCHETTI, AND M.H. RITZWOLLER, 2008. *Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps*, Geophys. J. Int., doi:10.1111/j1365-246X.2008.03720.x.
- [19] MALCOLM, A. E., SCALES, J. A. & VAN TIGGELEN, B. A., 2004. *Extracting the Green's function from diffuse, equipartitioned waves*, Phys. Rev. E, 70, 015601.
- [20] MARGERIN, L., CAMPILLO, M. & VAN TIGGELEN, B. A., 2001. *Coherent backscattering of acoustic waves in the near-field*, Geophys. J. Int., 145, 593–603.
- [21] MCNAMARA, N.M. & BULAND, R. P., 2004. *Ambient noise levels in the Continental United States*, Bull. Seism. Soc. Am., 94, 1517–1527.
- [22] METHA, K, SNIEDER, R., CALVERT, R. & SHEIMAN, J. 2008. *Acquisition geometry requirements for generating virtual-source data*, The Leading Edge, 27, 620–629.
- [23] PEDERSEN, H. A., KRUGER, F. AND THE SVEKALAPKO SEISMIC TOMOGRAPHY WORKING GROUP, 2007. *Influence of the seismic noise characteristics on noise correlations in the Baltic shield*, Geophys. J. Int., 168, 197–210.
- [24] PAUL, A., CAMPILLO, M., MARGERIN, L., LAROSE, E., DERODE, A., 2005. *Empirical synthesis of time-asymmetrical Green's functions from the correlation of coda waves*, J. Geophys. Res., 110, B08302, doi:10.1029/2004JB003521.
- [25] ROUX, P., SABRA, K. G., KUPERMAN, W. A. & ROUX, A., 2005. *Ambient noise cross correlation in free space: Theoretical approach*, J. Acoust. Soc. Am., 117(1), 79–84.
- [26] POLLITZ, F., 1999. *Regional velocity structure in northern California from inversion of scattered seismic surface waves*, J. Geophys. Res., 104, 15043–15072.
- [27] SANCHEZ-SESMA, F. J., PEREZ-RUIZ, J. A., LUZON, F., CAMPILLO, M., RODRIGUEZ-CASTELLANOS, A., 2008. *Diffuse fields in dynamic elasticity*, Wave Motion, 45(5), 641–654.
- [28] SATO, H., AND M. FEHLER, 1998. *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*, American Institute of Physics Press.
- [29] SCALES, J. A., MALCOLM, A. E. & VAN TIGGELEN, B. A., 2004. *Estimating scattering strength from the transition to equipartitioning*, AGU Fall Meeting Abstracts, B1053.
- [30] 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.
- [31] SHAPIRO, N. M., CAMPILLO, 2004. *Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise*, Geophys. Res. Lett., 31, L07614, doi:10.1029/2004GL019491.
- [32] SNIEDER, R., 1986. *The influence of topography on the propagation and scattering of surface waves*, Phys. Earth Planet. Inter., 44, 226–241.
- [33] 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.
- [34] STEHLY, L., CAMPILLO, M., SHAPIRO, N. M., 2006. *A study of the seismic noise from its long-range correlation properties*, J. Geophys. Res., 111, B10306, doi:10.1029/2005JB004237.
- [35] TREGOURES, N., HENNINO, R., LACOMBE, C., SHAPIRO, N. M., MARGERIN, L., CAMPILLO, M. & VAN TIGGELEN, B. A., 2002. *Multiple scattering of seismic waves*, Ultrasonics, 40, 269–274.
- [36] TURNER, J. A., 1998. *Scattering and diffusion of seismic waves*, Bull. Seism. Soc. Am., 88, 1, 276–283.
- [37] VAN TIGGELEN, B. A., 2003. *Green's function retrieval and time-reversal in a disordered world*, Phys. Rev. Lett., 91, 243904.
- [38] WAPENAAR, K., 2004. *Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation*, Phys. Rev. Lett., 93, 254301.
- [39] WAPENAAR, C.P.A., FOKKEMA, J. T. & SNIEDER, R., 2005. *Retrieving the Green's function in an open system by crosscorrelation: a comparison of approaches*, J. Acoust. Soc. Am., 118, 2783–2786.
- [40] WAPENAAR, K., 2006. *Green's function retrieval by cross-correlation in case of one-sided illumination*, Geophys. Res. Lett., 33, L19304, doi:10.1029/2006GL027747.

- [41] WEAVER, R. & LOBKIS, O. I., 2004. *Diffuse fields in open systems and the emergence of the Green's function*, J. Acoust. Soc. Am., 116, 2731–2734.
- [42] WEAVER, R. & LOBKIS, O. I., 2005. *Fluctuations in diffuse field-field correlations and the emergence of the Green's's function in open systems*, J. Acoust. Soc. Am., 117(6), 3432–3439.
- [43] WILLIS, M. E., LU, R., CAMPMAN, X., TOKSÖZ, M. N., ZHANG, Y. & DE HOOP, M., 2006. *A novel application of time reverse acoustics: salt dome flank imaging using walk away VSP surveys*, Geophys., 71(2), A7–A11.
- [44] YANG, Y., RITZWOLLER, M.H., LEVSHIN, A.L., & SHAPIRO, N.M., 2007. *Ambient noise Rayleigh wave tomography across Europe*, Geophys. J. Int., 168, 259–274.
- [45] YANG, Y. AND RITZWOLLER, M.H., 2008. *Characteristics of ambient seismic noise as a source for surface wave tomography*, Geochem. Geophys. Geosyst., 9, doi:10.1029/2007GC001814.
- [46] YAO, H., VAN DER HILST, R. D., DE HOOP, M. V., 2006. *Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis – I Phase velocity maps*, Geophys. J. Int., 166, 732–744.
- [47] YAO, H., BEGHEIN, C., VAN DER HILST, R. D., 2008. *Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis – II Crustal and upper mantle structure*, Geophys. J. Int., 173, 205–219.