

Application of temporal reweighting to ambient noise cross-correlation for improved seismic Green's function

Jiangtao Li¹, Richard L. Weaver,² John Y. Yoritomo² and Xiaodong Song^{1,3}

¹Department of Geology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. E-mail: jiangtaoli.whu@gmail.com

²Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

³Institute of Theoretical and Applied Geophysics, School of Earth and Space Sciences, Peking University, Beijing 100871, China. E-mail: xiao.d.song@gmail.com

Accepted 2020 January 1. Received 2019 December 19; in original form 2019 June 27

SUMMARY

Due to the partly diffuse character of ambient noise, the retrieval of amplitude information and attenuation from noise cross-correlations has been difficult. Here, we apply the temporal reweighting method proposed by Weaver & Yoritomo to seismic data from the USArray in the central-midwest US. The results show considerable improvements in retrieved Green's functions in both symmetry and causality. The reweighting is able to make the effective incident noise field more isotropic (though not yet truly isotropic). It produces more robust amplitude measurements and also makes both the causal and anticausal parts usable. This suggests that it could be widely applicable for retrieval of Green's functions from ambient noise for attenuation study. The results also suggest an alternative measure of signal-to-noise ratio that complements the conventional one.

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

1 INTRODUCTION

The retrieval of Green's function (Lobkis & Weaver 2001; Weaver & Lobkis 2001) between two seismic stations from ambient noise cross-correlation has become a routine technique in seismology (Shapiro & Campillo 2004; Shapiro *et al.* 2005). The technique is now widely applied to perform surface wave tomography (e.g. Shapiro *et al.* 2005; Yao *et al.* 2006; Lin *et al.* 2008; Zheng *et al.* 2008; Bao *et al.* 2015) and to monitor changes in Earth's elastic properties (e.g. Brenguier *et al.* 2008, 2014; Xu & Song 2009). These studies utilized the traveltime information from the retrieved surface waves. However, due to the partly diffuse character of ambient noise, retrieval of amplitude information and attenuation, which are strongly affected by the distribution of noise sources, has proved difficult (Zhang & Yang 2013). Moreover, non-fully diffuse ambient noise, even when smoothly non-isotropic, would generate spurious arrivals in the cross-correlations when scatterers exist (Snieder *et al.* 2008; Snieder & Fleury 2010), which distort the main arrivals and thus deteriorate amplitude measurements (Yoritomo & Weaver 2016).

The standard seismic signal processing procedure (Bensen *et al.* 2007) normalizes the data by a running absolute mean on each station's record. This suppresses earthquake signals before cross-correlation. However, such normalization rescales each station differently, which makes it difficult to interpret the relative amplitudes between station pairs. Weaver (2011, 2013) proposed 'temporal

flattening' which chooses weights that are inversely proportional to each day's total energy on the array; this keeps the relative amplitudes while minimizing residual noise due to averaging over a finite time record. But this does not remove the influence of a non-isotropic noise field. Thus, using simulated noise wavefield, Weaver & Yoritomo (2018) proposed an alternative temporal reweighting method that seeks the retrieved Green's function that is symmetric or causal (no energy before the main arrivals). The method takes advantage of noise fields that, while non-isotropic, have their non-isotropy varying in time. By minimizing an antisymmetry energy or acausality energy as measures of noise, they constructed optimal reweightings of the cross-correlations from different blocks of time ('days') before stacking. Note that they defined 'acausal' differently from Bensen *et al.* (2007). In Weaver & Yoritomo (2018), 'acausal' or 'acausality' refers to the situation where there are spurious arrivals/energy before the main arrivals, which should not exist in an ideal fully diffuse noise field. 'Causal' or 'causality' means the opposite, and 'causal' also refers to the positive lag in the Green's function. They use 'anticausal' instead of 'acausal' to describe the negative lag. The acausality energy is the energy before the main arrivals, and antisymmetry energy measures the difference between the positive and negative lags. Weaver & Yoritomo (2018) tested the method with numerical simulations for a 2-D medium with scatterers, found significant improvement in the retrieved Green's functions and showed that the reweighted noise field had been made more isotropic, and thus closer to a fully diffuse field, by the reweighting.

In this study, we apply the temporal reweighting method proposed by Weaver & Yoritomo (2018) to seismic data from the USArray and find that it works well. The results show considerable improvements in the retrieved Green's functions in both symmetry and causality. The new method makes the effective noise field more isotropic. It produces more robust amplitude measurements and also makes both the causal (positive lag) and anticausal (negative lag) parts of the Green's function usable. This suggests that it could be widely applicable for retrieval of Green's functions from ambient noise. The results also suggest an alternative measure of signal-to-noise ratio (SNR), one that complements the conventional SNR.

2 METHOD AND DATA

Weaver & Yoritomo (2018) proposed eight distinct schemes of temporal reweighting based on minimization of different quantities, and preferred their schemes V–VIII that minimize certain noise-to-signal ratios. In all their simulations, the four schemes produced very similar results. Regardless of choice of signal energy definition, and regardless of whether the noise is measured by antisymmetry or acausality, they retrieved Green's functions that were both more symmetric and more causal than the raw correlations or with temporal flattening. Here we adopt their schemes V and VI for their simpler definition of signal energy, which is the square of the sum of all auto-correlation amplitudes. Scheme V measures a noise defined as the antisymmetry energy (noise), which is the difference (L2 norm) between causal and anticausal parts over a defined window (e.g. from zero to long after the main arrival). Scheme VI measures a noise defined as the acausality energy (noise), which is the integral of amplitude squared over pre-causal times (between the two main arrivals in the positive or negative lags, respectively). The total noise is the sum of these over all station pairs. For each scheme, the minimization generates optimal weights with which the cross-correlations from different blocks of time ('days') are combined before stacking to get the empirical Green's functions (EGFs). See eqs (1)–(21) in Weaver & Yoritomo (2018) for more details.

In this study, we prefer scheme V (minimizing antisymmetry) over VI (minimizing acausality) for two reasons: (1) real data could have body wave signals in pre-causal times; (2) it is more straightforward to define an antisymmetry window than an acausality window. Thus we focus on their scheme V. We also show the results from scheme I (simple stack, stack of raw correlations after pre-processing) and scheme II (temporal flattening; Weaver 2011, 2013) for comparison. We did not show the results from running absolute mean normalization (e.g. Bensen *et al.* 2007) since it does not preserve the relative amplitudes, although the waveforms would look somewhat similar to the ones from temporal flattening.

We apply the temporal reweighting method (Weaver & Yoritomo 2018) to continuous data (with about one year overlap) recorded at USArray stations in the central-midwest US (Fig. 1). We selected this region because it is in a stable craton with less structural complexity. For the pre-processing, we removed the instrument response, band-passed within 2–150 s, removed the average and trend, and removed records with very large amplitudes ($>10^4$ nm, presumably due to instrument glitches or strong earthquakes). Then we performed strict data selection that requires every selected station to have a full-day data at each selected day, which ended up with 290 d (~ 24 d in each month of year 2012) of data at 60 stations. Afterwards, we followed the procedure in Weaver & Yoritomo

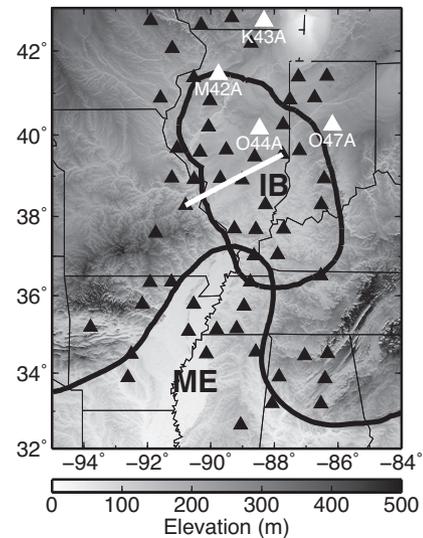


Figure 1. Distribution of seismic stations used in this study with surface topography. The triangles represent 60 USArray stations, and the white ones mark the 4 example stations in Fig. 5. The thick white line represents the example pair shown in Fig. 2. The thick black lines outline the approximate locations of two major sedimentary basins in this region [modified from Chen *et al.* (2016) and Buschbach & Kolata (1991)], the Illinois Basin (IB) and the Mississippi Embayment (ME). The thin black lines are state boundaries.

(2018) to perform cross-correlations in frequency domain and filter by a narrow-band filter around 0.1 Hz (a cosine-bell-squared filter), and then used their proposed schemes to calculate optimal weights for each day and stacked the weighted cross-correlations to get the EGFs for all station pairs. We focus on the period of 10 s (0.1 Hz) as it is a typical period in surface wave studies, and it appears to have relatively stronger scattering compared to longer periods.

To define the windows for different measurements, we estimated the average group velocity in the study region to be around 3.1 km s^{-1} (e.g. Shen & Ritzwoller 2016). Then we defined the antisymmetry window to be from zero to $D/3.1 + 100$ s, and the acausality window to be within $\pm(D/3.1 - 20)$ s, where D is the distance between two stations. For calculation of SNR, we defined the signal to be the maximum amplitude within $D/3.1 - 20$ to $D/3.1 + 50$ s, conventional noise to be the root mean square (rms) after $D/3.1 + 200$ s, and an alternative measure of noise to be the rms within the acausality window (before the signal). We checked these windows on record sections like Fig. 3 and confirmed they are reasonable.

With the data set of 290 d, we have tested the method with different lengths of time. As noise sources vary with seasons (e.g. Stehly *et al.* 2006; Stutzmann *et al.* 2012; Traer *et al.* 2012), we started with moving windows of 60-d data (with 30-d overlap, i.e. days 1–60, 31–90, ...) to test the applicability of the method at different time of the year. We noted that the tests which contain data in summer months (some June, July, August and some September) always resulted in some spurious arrivals very close to zero time in the correlations, regardless of the scheme used. We speculate that this additional residual noise may be related to summer thunderstorms (Supporting Information Fig. S1). Thus we removed 70 d in the summer, forming a new data set of 220 d. We also ran a test with 120 d of data formed by the even months (February, April, half June, October and December) to see how it compares with data covering most of the year (the case with 220 d).

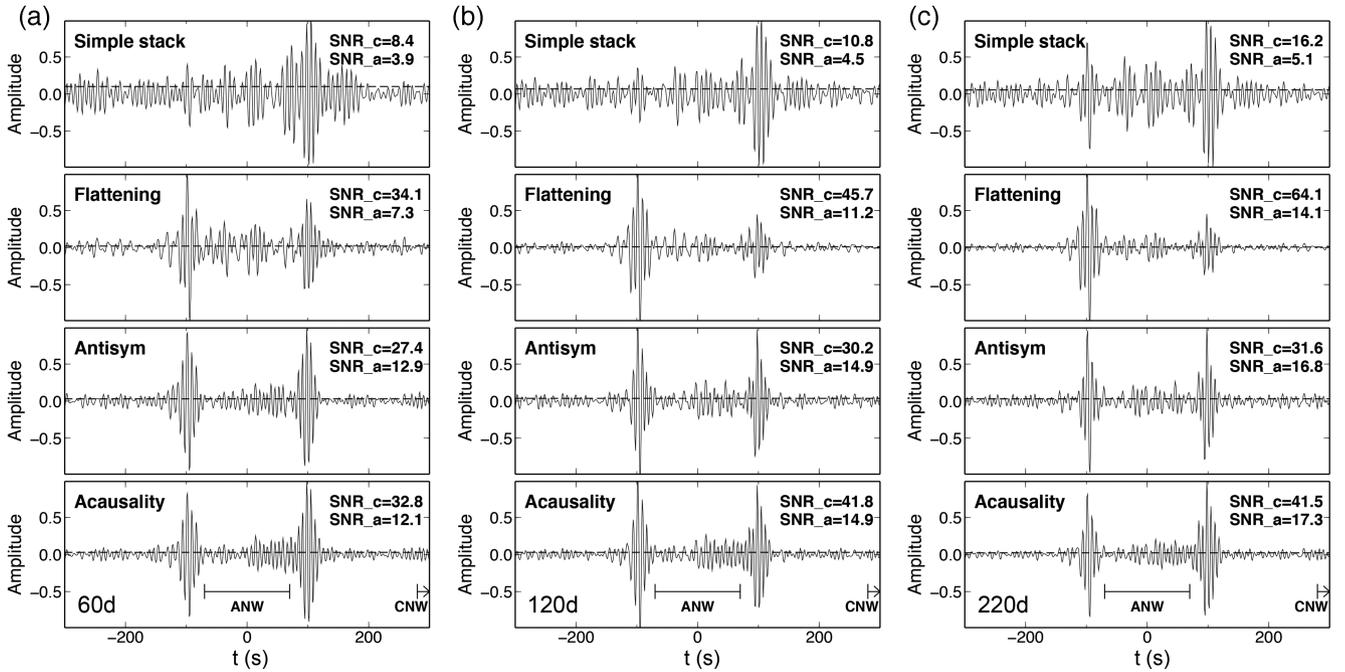


Figure 2. Retrieved EGFs at 10 s for an example pair (P45A to R42A) from different schemes with (a) 60 d, (b) 120 d or (c) 220 d of data, respectively. From top to bottom are results from simple stack, temporal flattening (Weaver 2011, 2013), minimizing antisymmetry and minimizing acausality. The waveforms are normalized by their maximum amplitudes. The acausality noise window (ANW) and conventional noise window (CNW) are marked in the bottom panels. The conventional SNR (SNR_c) and alternative SNR (SNR_a), calculated using the symmetric part of causal and anticausal EGFs, are also labelled. The dashed lines mark the levels of finite-T noise (Weaver & Yoritomo 2018).

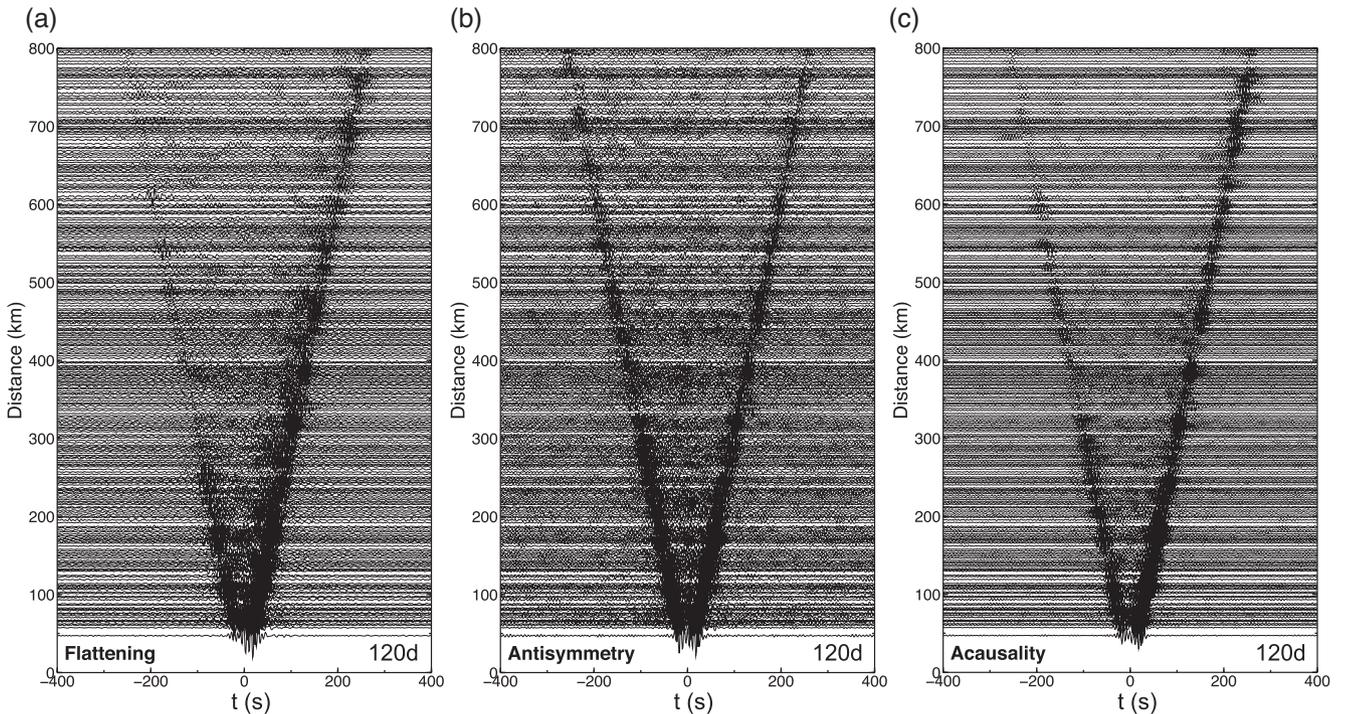


Figure 3. Retrieved EGFs at 10 s for 290 selected pairs with 120 d of data, from (a) flattening, (b) minimizing antisymmetry and (c) minimizing acausality. The pairs with the lowest SNR_a (alternative SNR) from flattening within every 2-km distance range are selected (until 800 km). The average alternative SNRs for the 290 pairs are 4.0, 6.3 and 6.9, respectively.

Since seismic data contain a lot of earthquake signals, we have also tested different time-domain normalizations in the pre-processing to suppress them. Those include normalizing each day's records by the average of running absolute mean of all records at

that day (also able to keep the relative amplitudes; similar to temporal flattening), or automated event detection and removal for all stations at the same day (−10 to 50 min for amplitudes larger than 10 times of the rms). The results show small difference compared to

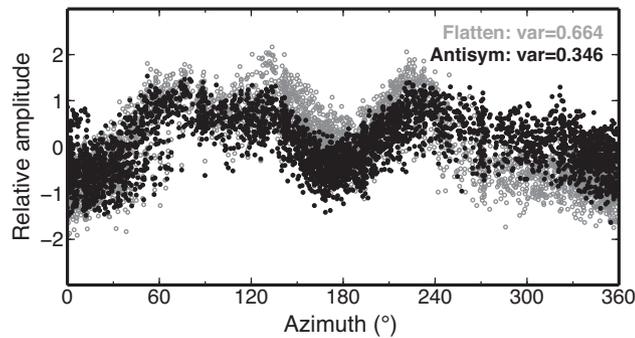


Figure 4. Variation of amplitudes (two measurements for each pair in both causal and anticausal parts of EGFs: the maxima in the signal windows, corrected for geometric spreading, in natural logarithm) with respect to azimuth (or backazimuth for anticausal part) for 120 d of data. The grey open circles show the measurements from temporal flattening (Weaver 2011, 2013), and the black dots show the measurements from minimizing antisymmetry. The averages from both schemes are removed, and the variances are labelled.

those without time-domain normalization (Supporting Information Fig. S2), suggesting that the reweighting method is able to suppress earthquake signals by lowering the weights of those days. Therefore, the reweighting method enables one to not worry or choose a time-domain normalization.

3 RESULTS

The results show considerable improvements in retrieved EGFs compared to both simple stack and temporal flattening (Weaver 2011, 2013), in both symmetry and causality, as shown in Figs 2 and 3. The case with 60 d shown here (Fig. 2a) is the one with the largest improvement by the new schemes among all the 60-d moving window tests (see description below). Our results suggest that 60 d of data is sufficient to obtain a high enough SNR. At the same time, we find that the schemes which minimize the ratio of either antisymmetry energy (noise) or acausality energy (noise) to signal energy would result in similar EGFs, which improve in both symmetry and causality (Figs 2 and 3). The reason is that both schemes reduce their measure of noise by effectively making the incident noise field more isotropic (Weaver & Yoritomo 2018). Note that although the results from minimizing acausality (Fig. 3c) are not as symmetric as from minimizing antisymmetry (Fig. 3b), they are still more symmetric than the ones from temporal flattening (Fig. 3a). As mentioned above, considering that it is easier to minimize the antisymmetry (for acausality a proper window needs to be defined carefully), and also that real data could have body wave arrivals before the surface waves, we prefer to use the scheme that minimizes the antisymmetry noise-to-signal ratio (scheme V; for Figs 4 and 5).

It may be noted that this data set shows little coda (e.g. Figs 2 and 3). Scattering is therefore presumably weak. This may be contrasted with the system studied by Weaver & Yoritomo (2018), for which scattering attenuation was not negligible, and for which there was a very notable coda. Thus the reweighting schemes proposed appear to be applicable regardless of the strength of scattering.

Fig. 4 shows the variation of the corrected amplitudes (corrected for geometric spreading, in natural logarithm) of all EGFs with respect to azimuth. We choose to show the case with 120 d for its smallest variance with scheme of minimizing antisymmetry, compared to 60 d or 220 d (Supporting Information Fig. S3). Note that we measure the amplitudes (the maximum in the signal window)

for both casual and anticausal parts of the EGFs, thus each pair has two measurements. The result suggests that scheme V (minimizing antisymmetry) is able to make the effective noise field more isotropic compared to temporal flattening. Fig. 4 and Supporting Information Fig. S3 also imply that the length of data is not critical as long as the time period contains noise sources from different directions. Note that the 60 d (late September to early December; Supporting Information Fig. S3a, also for Fig. 2a) are picked for the largest variance reduction of ~ 43 per cent among all the tests with 60-d moving windows, which show an average variance reduction of ~ 30 per cent. It turns out that this time period has very similar azimuthal pattern with the cases covering most of the year (Fig. 4 and Supporting Information Fig. S3b). The cases with 120 or 220 d show a variance reduction of 48 per cent or 51 per cent, respectively. On the other hand, it is notable that even with scheme V the reweighted field is not truly isotropic, with minima around azimuths of 0° and 180° , the reason for which is discussed in Section 4.1.

Fig. 5 shows the variation of the corrected amplitudes of the EGFs with respect to distance between two stations, also with 120 d of data. Four typical examples of measurements from one source station to all other stations are displayed, with either small (Fig. 5a), moderate (Fig. 5b), or large improvements (Figs 5c and d). The results show not only much larger R^2 , but also larger and more consistent (presumably more reasonable) attenuation (slope). There are even some unphysical negative attenuations from temporal flattening (e.g. Fig. 5d), but that is not a problem for scheme V (minimizing antisymmetry). We have a statistic of such measurements (slope and R^2) for all the 60 stations, as shown in Fig. 6. The average R^2 increases from 0.0456 (temporal flattening) to 0.0989 (minimizing antisymmetry; Fig. 6b). The average slope (a) increases from 3.78×10^{-4} to 7.03×10^{-4} , and the standard deviation decreases from 4.44×10^{-4} to 2.55×10^{-4} (Fig. 6a). The results from the new scheme are not only more robust but also more consistent and reasonable. The larger attenuation (slope) correlates better with the local geology of sedimentary basin at shallow depth (e.g. Mitchell 1975, 1995).

We also compare the results with measurements from an earthquake (Fig. 7). Throughout the time range of the 60 stations (2011.1 to 2013.6), the largest earthquake in the study region that was recorded by at least 3 stations was only $M_{\text{W}} 3.9$ (M_{Lg}). The surface wave signal around 10 s can hardly be seen. Therefore, we could not make a direct comparison. Instead, we looked for earthquakes with $M > 5$ since 2000, and found just one event on 2008 April 18 with an M_{W} of 5.2, which was recorded by 36 stations (not USArray) in the study region (Fig. 7a). It is close to a USArray station R45A in year 2012 (Fig. 7a). Fig. 7(b) shows the measurements of relative amplitudes (corrected for geometric spreading, in natural logarithm) from the M_{W} 5.2 earthquake at 10 s, which has a slope of 11.5×10^{-4} . We noted that some measurements (the grey dots) behave like a separate group and found that they are all within the Mississippi Embayment (ME), where the sediment is the thickest in the region (Fig. 7a). After removing these measurements (more than 1/3 of the stations), the new slope is 9.1×10^{-4} , which is close to the slope measured at station R45A from minimizing antisymmetry (8.5×10^{-4} ; Fig. 7c). Even the original slope of 11.5×10^{-4} is still comparable, and it is well within the range of the USArray measurements (Fig. 6a). This suggests the slope (attenuation) measurements from the new scheme are reliable. Note that we did not remove the measurements in ME for our ambient noise study, because there are only 1/6 of the stations inside ME and the effect would not make a difference.

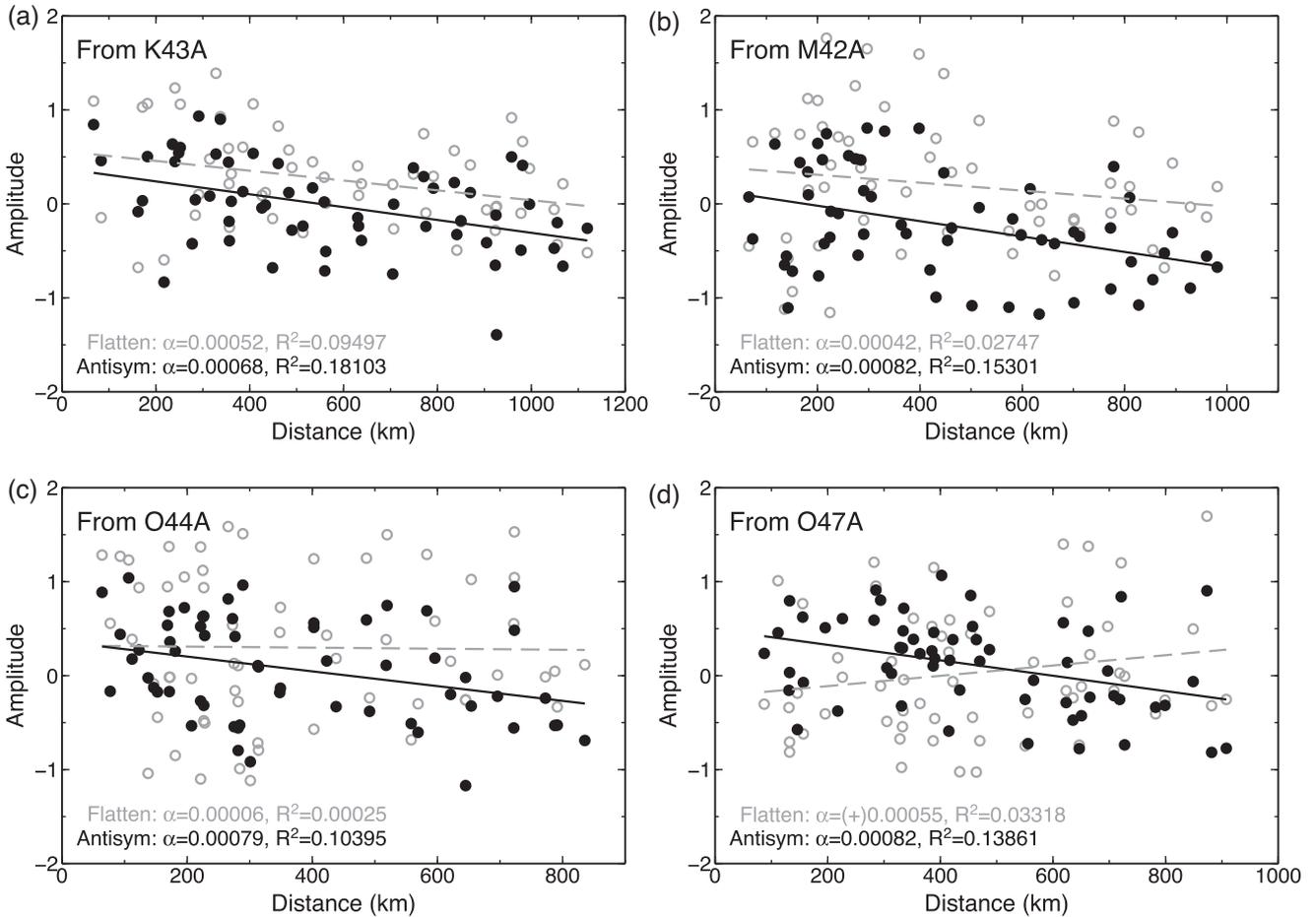


Figure 5. Four typical examples for variation of amplitudes (same measurements as in Fig. 4) with distance from one station to all other stations, for 120 d of data. The results are from temporal flattening (grey) and minimizing antisymmetry (black), respectively. The linear fit, slope (α), and R -squared to the values are also shown. (a) With slightly larger slope and small improvement in R -squared. (b) With larger slope and large improvement in R -squared. (c) With much larger and more reasonable slope, and significant improvement in R -squared. (d) With unphysical negative attenuation from temporal flattening but reasonable slope from minimizing antisymmetry, and large improvement in R -squared.

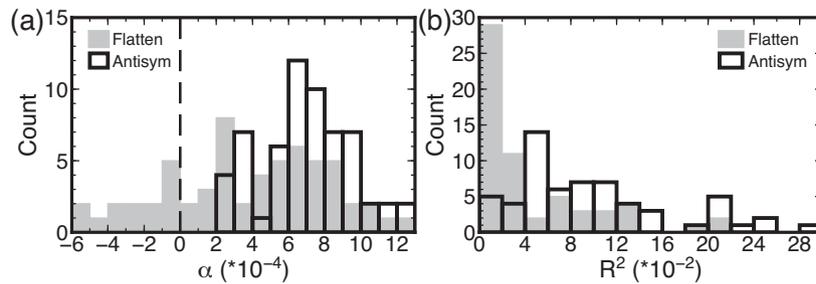


Figure 6. Histograms of slope (a) and R -squared measurements for all the 60 stations (same measurements as in Fig. 5), for 120 d of data. The results are from temporal flattening (grey) and minimizing antisymmetry (black), respectively. (a) Histogram of slope measurements. The dashed line marks a slope of zero. Values smaller than that are unphysical. (b) Histogram of R -squared measurements.

4 DISCUSSION

4.1 Improvements in retrieved EGFs and reweighted field

The results (Figs 2–7) suggest that the temporal reweighting method works for the real Earth and can be widely applied in the retrieval of EGFs. The reweighting is able to make the effective noise field more isotropic, though not yet truly isotropic (Fig. 4). The method is able to extract more symmetric and causal EGFs, which reduces the spurious arrivals from scattering or earthquakes. By minimizing

antisymmetry or acausality, it can make the EGFs more causal and symmetric at the same time (Figs 2 and 3), as each attempts to minimize their own measure of noise, and can only do that by making the incident noise field more isotropic.

Although we are focusing on the amplitude information but not traveltimes, we still show an enlarged view of Fig. 2 to compare the traveltimes (Fig. 8). It can be seen that the new scheme has negligible effect on traveltimes and thereby dispersion measurements. On the other hand, it is well known that arrival times are only

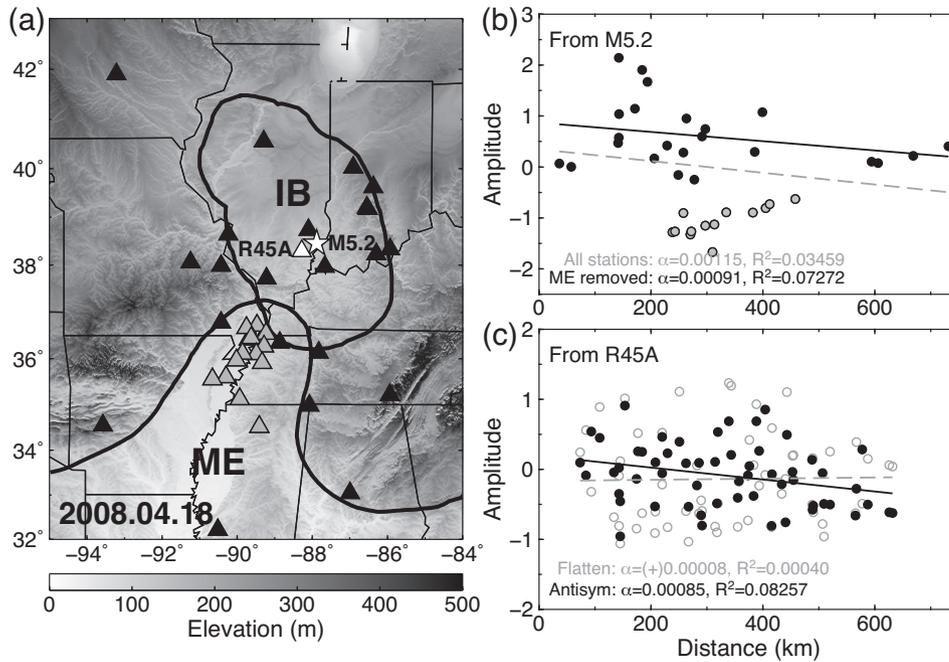


Figure 7. Comparison of amplitude slopes measured from ambient noise methods and from an earthquake. (a) The location of the M_W 5.2 earthquake (white star) in 2008 April 18 and the recording stations (black and grey triangles). The grey triangles correspond to the grey dots with lower amplitudes in (b). The white triangle marks the closest USArray station R45A in year 2012 (not yet deployed in 2008). The topography and boundaries are the same as in Fig. 1. (b) The measurements of relative amplitudes (the maxima in the signal windows, corrected for geometric spreading, in natural logarithm, and with the average removed) with respect to distance from the M_W 5.2 earthquake. The grey dots are the ones with lower amplitudes and behaving like a separate group, which are found to be all within the Mississippi Embayment. The dashed line and text show the fit for all stations, while the solid line and text for the black dots only (without the grey stations). (c) Same plot as Fig. 5 but for station R45A.

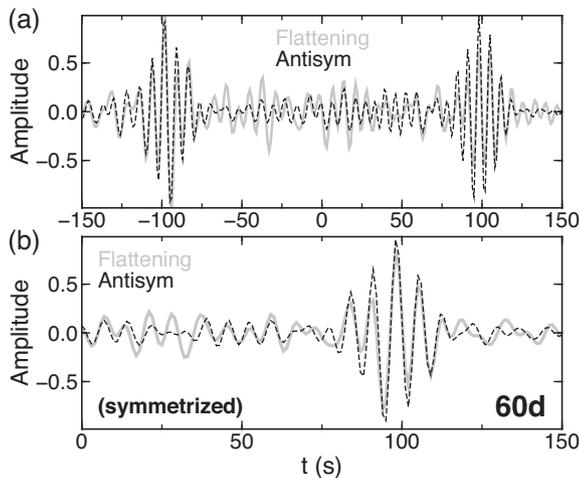


Figure 8. Comparison of waveform and traveltimes between the schemes of temporal flattening (grey) and minimizing antisymmetry (black), for the same pair as Fig. 2. (a) An enlarged view of Fig. 2(a) (60 d, the second and third rows). (b) The symmetrized waveforms of (a). The situation for 120 or 220 d of data is very similar.

weakly sensitive to anisotropy in the ambient noise field (Weaver *et al.* 2009). That is why practical applications have been successful in recovering speeds and dispersions. Therefore, by making the incident field more isotropic would not make much difference on velocity dispersion measurements.

As mentioned above, it seems that even with the new method (scheme V) the reweighted noise field is not truly isotropic, with minima around azimuths of 0° and 180° . The reason is that at most times there is little noise energy coming from the north or the south of the study region (e.g. Stehly *et al.* 2006; Stutzmann *et al.* 2012; Traer *et al.* 2012). Therefore, there is no linear combination of daily weights to make the noise field isotropic. This case was examined as synthetic case ‘B’ by Weaver & Yoritomo (2018). Ideally, if in some years there were also sufficient noise energy from the Gulf of Mexico or the Arctic for at least some of the time, it might be possible for the reweighted noise field to get closer to being isotropic, like the other synthetic cases in Weaver & Yoritomo (2018). On the other hand, in other places with better noise distribution, the schemes might work better.

4.2 Improvements in amplitude measurements

As shown by Figs 5 and 6, we are able to extract more robust amplitude information with the temporal reweighting method; this could be useful for attenuation studies. With a more isotropic noise field and more causal EGFs, the amplitude measurements are less affected by the spurious arrivals at short times (Snieder *et al.* 2008; Yoritomo & Weaver 2016). In comparison, the temporal flattening (Weaver 2011, 2013) can also produce relatively clean EGFs and acceptable SNRs, but not as robust amplitude measurements (Figs 5 and 6). Moreover, with the new schemes, now both the causal and anticausal parts of the EGFs can be utilized, which avoids the need to artificially symmetrize the EGFs as conventional method do, and

thus keeps more information. This is another advantage of the new method. We did not show the results along linear arrays, as there would be too few measurements for the current data set, which could be dominated by a few local extremes.

From the measurements of slope (a ; Fig. 6), the average slope is 7.03×10^{-4} , with a standard deviation of 2.55×10^{-4} . It corresponds to a relatively small average Q value of ~ 144 ($Q = \pi f/aV$, $f = 0.1$ Hz, $a = 7.03 \times 10^{-4}$ km $^{-1}$, $V = 3.1$ km s $^{-1}$) and a range of ~ 106 – 226 . More than 1/3 of the values are consistently around 6 – 8×10^{-4} , which corresponds to a Q range of ~ 127 – 169 . The relatively strong attenuation (small Q) correlates well with the local geology of sedimentary basin (the Illinois Basin and the ME; Fig. 1) at shallow depth (e.g. Mitchell 1975, 1995), considering that 10 s surface waves are primarily sensitive to shallow upper crust. The comparison with earthquake measurements (Fig. 7) further suggests that the attenuation measurements with the temporal reweighting method are reasonable.

4.3 Alternative measure of SNR

Using the definition of noise-to-signal ratio in Weaver & Yoritomo (2018), by reversing it we define an alternative measure of SNR. The signal is the same as conventional definition, while the ‘noise’ refers to antisymmetry energy or acausality energy. The conventional SNR (e.g. Bensen *et al.* 2007) measured noise as the rms in a certain time window (e.g. of 500 s) hundreds of seconds after the signal window. However, such ‘noise’ window may contain true signal, that is, multiply scattered coda (Weaver & Yoritomo 2018), and furthermore is dominated by the noise component that Weaver & Yoritomo (2018) call ‘finite-T noise’ (noise from averaging over finite times). A more practical measure of ‘noise’ would be the energy within the acausal window before the signal, or the energy in the antisymmetric parts of the EGF, as defined in Weaver & Yoritomo (2018) and used elsewhere (e.g. Larose *et al.* 2008). It is that energy that can contaminate assessment of the main arrival. Thus, we propose an alternative SNR (as labelled in Fig. 2) defined as the maximum in the signal window over the rms within the acausality window. Note that here we still measure the SNR of the symmetric part to be consistent with conventional measurements. It can be seen that although schemes V (minimizing antisymmetry) and VI (minimizing acausality) increase in the alternative SNR, they actually have smaller conventional SNR compared to temporal flattening (Fig. 2). That is because temporal flattening (Weaver 2011, 2013) attempts to minimize the finite-T noise at later times, while the new schemes focus on the antisymmetry or acausality noise.

5 CONCLUSIONS

We apply the temporal reweighting method proposed by Weaver & Yoritomo (2018) to seismic data from the USArray in the central-midwest US. The results show considerable improvements in the retrieved EGFs in both symmetry and causality. The reweighting is able to make the effective noise field more isotropic (though not yet truly isotropic). It produces more robust amplitude measurements, and also makes both the causal and anticausal parts of the EGFs usable, which avoids the need to artificially symmetrize the EGFs and thus keeps more information. These improvements suggest that it could be widely applicable for the retrieval of EGFs from ambient noise. The extent of improvement also depends on whether there

is sufficient noise energy coming from most of the directions. The results also suggest an alternative measure of SNR, which complements the conventional one. It would be meaningful to test the method in other locations, for larger and/or denser data set, and at other frequencies of interest. It would also be intriguing to see how it works in case of stronger coda.

ACKNOWLEDGEMENTS

The continuous seismic data in this study are downloaded from the Data Management Center of the Incorporated Research Institutions for Seismology (<https://ds.iris.edu/ds/nodes/dmc/>). Most figures were prepared using Generic Mapping Tools (Wessel & Smith 1998). Detailed constructive comments from Morgan Moschetti and an anonymous reviewer helped improve the manuscript greatly. This research was supported by the National Natural Science Foundation of China (41774056) and the National Science Foundation of the United States (EAR 1620595).

REFERENCES

- Bao, X., Song, X. & Li, J., 2015. High-resolution lithospheric structure beneath Mainland China from ambient noise and earthquake surface-wave tomography, *Earth planet. Sci. Lett.*, **417**, 132–141.
- Bensen, G.D., Ritzwoller, M.H., Barmin, M.P., Levshin, A.L., Lin, F., Moschetti, M.P., Shapiro, N.M. & Yang, Y., 2007. Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophys. J. Int.*, **169**(3), 1239–1260.
- 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.
- Brenguier, F., Campillo, M., Takeda, T., Aoki, Y., Shapiro, N.M., Briand, X., Emoto, K. & Miyake, H., 2014. Mapping pressurized volcanic fluids from induced crustal seismic velocity drops, *Science*, **345**(6192), 80–82.
- Buschbach, T.C. & Kolata, D.R., 1991. Regional setting of Illinois Basin, in *Interior Cratonic Basins*, Vol. 51, pp. 29–55, eds Leighton, M.W. *et al.*, Am. Assoc. Petr. Geol. Mem.
- Chen, C., Gilbert, H., Andronicos, C., Hamburger, M.W., Larson, T., Marshak, S., Pavlis, G.L. & Yang, X., 2016. Shear velocity structure beneath the central United States: implications for the origin of the Illinois Basin and intraplate seismicity, *Geochem. Geophys. Geosyst.*, **17**, 1020–1041.
- Larose, E., Roux, P., Campillo, M. & Derode, A., 2008. Fluctuations of correlations and Green’s function reconstruction: role of scattering, *J. Appl. Phys.*, **103**(11), 114907, doi.org/10.1063/1.2939267.
- 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, *Geophys. J. Int.*, **173**(1), 281–298.
- Lobkis, O.I. & Weaver, R.L., 2001. On the emergence of the Green’s function in the correlations of a diffuse field, *J. acoust. Soc. Am.*, **110**, 3011–3017.
- Mitchell, B.J., 1975. Regional Rayleigh wave attenuation in North America, *J. geophys. Res.*, **80**, 4904–4916.
- Mitchell, B.J., 1995. Anelastic structure and evolution of the continental crust and upper mantle from seismic surface wave attenuation, *Rev. Geophys.*, **33**(4), 441–462.
- Shapiro, N.M. & Campillo, M., 2004. Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophys. Res. Lett.*, **31**, L07614, doi:10.1029/2004GL019491.
- 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.
- Shen, W. & Ritzwoller, M.H., 2016. Crustal and uppermost mantle structure beneath the United States, *J. geophys. Res.*, **121**(6), 4306–4342.

- Snieder, R. & Fleury, C., 2010. Cancellation of spurious arrivals in Green's function retrieval of multiple scattered waves, *J. acoust. Soc. Am.*, **128**, 1598–1605.
- Snieder, R., van Wijk, K., Haney, M. & Calvert, R., 2008. Cancellation of spurious arrivals in Green's function extraction and the generalized optical theorem, *Phys. Rev. E*, **78**, 036606, doi:10.1103/PhysRevE.78.036606.
- Stehly, L., Campillo, M. & Shapiro, N.M., 2006. A study of the seismic noise from its long-range correlation properties, *J. geophys. Res.*, **111**(B10), 1–12.
- Stutzmann, E., Arduin, F., Schimmel, M., Mangeny, A. & Patau, G., 2012. Modelling long-term seismic noise in various environments, *Geophys. J. Int.*, **191**(2), 707–722.
- Traer, J., Gerstoft, P., Bromirski, P.D. & Shearer, P.M., 2012. Microseisms and hum from ocean surface gravity waves, *J. geophys. Res.*, **117**, B11307, doi:10.1029/2012JB009550.
- Weaver, R.L., 2011. On the amplitudes of correlations and the inference of attenuations, specific intensities and site factors from ambient noise, *C. R. Geosci.*, **343**(8–9), 615–622.
- Weaver, R.L., 2013. On the retrieval of attenuation and site amplifications from ambient noise on linear arrays: further numerical simulations, *Geophys. J. Int.*, **193**(3), 1644–1657.
- Weaver, R.L. & Lobkis, O.I., 2001. Ultrasonics without a source, thermal fluctuation correlations at MHz frequencies, *Phys. Rev. Lett.*, **87**(13), 134301, doi:10.1103/PhysRevLett.87.134301.
- Weaver, R.L. & Yoritomo, J.Y., 2018. Temporally weighting a time varying noise field to improve Green function retrieval, *J. acoust. Soc. Am.*, **143**, 3706–3719.
- Weaver, R.L., Froment, B. & Campillo, M., 2009. On the correlation of non-isotropically distributed ballistic scalar diffuse waves, *J. acoust. Soc. Am.*, **126**, 1817–1826.
- Wessel, P. & Smith, W.H.F., 1998. New, improved version of generic mapping tools released, *EOS, Trans. Am. geophys. Un.*, **79**(47), 579, doi:10.1029/98EO00426.
- Xu, Z.J. & Song, X., 2009. Temporal changes of surface wave velocity associated with major Sumatra earthquakes from ambient noise correlation, *Proc. Natl. Acad. Sci. USA*, **106**, 14 207–14 212.
- Yao, Y., van der Hilst, R. & de Hoop, M., 2006. Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis—I. Phase velocity maps, *Geophys. J. Int.*, **166**(2), 732–744.
- Yoritomo, J.Y. & Weaver, R.L., 2016. Fluctuations in the cross-correlation for fields lacking full diffusivity: the statistics of spurious features, *J. acoust. Soc. Am.*, **140**(1), 702–713.
- Zhang, J. & Yang, X., 2013. Extracting surface wave attenuation from seismic noise using correlation of the coda of correlation, *J. geophys. Res.*, **118**, 2191–2205.
- Zheng, S., Sun, X., Song, X., Yang, Y. & Ritzwoller, M.H., 2008. Surface wave tomography of China from ambient seismic noise correlation, *Geochem. Geophys. Geosyst.*, **9**, Q05020, doi:10.1029/2008GC001981.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1. Plots for why removing the data in the summer. (a) EGFs of the same 290 selected pairs as in Figs 2(d)–(f) but with all 290 d of data, from the scheme of minimizing antisymmetry. There are some arrivals very close to zero time. (b) EGFs of the 290 pairs with 220 d of data (with summer data removed), from minimizing antisymmetry. (c) Day-long records at station M42A in different months of year 2012. Each line is the record for one day. The summer months (June, July and August) have many irregular signals (checked to be not from earthquakes), which might be related with summer thunderstorms, while the other months (e.g. December) are more noise-like.

Figure S2. Results with different time-domain normalizations in pre-processing for pair M46A to O44A with 60 d of data, showing not much difference. (a) No time-domain normalization (adopted in this study). (b) Normalization by the average of running absolute mean of all records at the same day. (c) Automated event detection and removal for all stations at the same day.

Figure S3. Same as Fig. 3, but for (a) 60 d and (b) 220 d of data.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.