Referee 1 (Cornelis Wapenaar):

It is not clear why the correlation function is first introduced in terms of integrals (eqs 2 and 4) and later discretized (eq 7). Since in this paper the correlation function is not used as a representation of a Green's function between x_1 and x_2, it seems that one doesn't need the integral form of eq 4. Wouldn't it make more sense to define the correlation function as a summation (like in eq 7) right from the start? (honoring the fact that sources are in general sparsely and irregularly distributed, such as in Figures 3a and 6a). If there is a good reason to start with the integral representation then this should be clearly explained.

We considered this question in detail and discussed about it. We agree with the reviewer that the choice to start from an integral representation is somewhat arbitrary because we later approximate by summation.

The reason why we would prefer to continue using the integral representation is that for the most general case, we assume that noise excitation can be represented by sources acting and varying continuously over extended areas. While we don't assume that their occurrence is homogeneous, we believe that for example the secondary microseism is excited by sources distributed over areas where strong, opposing wave trains occur, for example through storms, and which are not always strongly localized.

In this sense, the example in Figure 3 is not the most representative case. Sources in Figure 6 are actually continuously distributed, although most of them have low amplitudes compared to the Gaussian-shaped, more localized ones.

We added the following sentence to make this thought process explicit:

"We adopt an integral description here, as we assume that the noise sources N(xi_1, w) and N(xi_2, w) are generally extended and vary continuously over more or less extended source areas."

Figure 5 shows sensitivity kernels. This needs more explanation. I assume the left panel shows eq 6 as a function of xi, for fixed x_1 and x_2 (shown by the triangles) and fixed n and m. If this is correct, please state this explicitly (or if it is not correct, explain what it shows instead). How are omega_0 and omega_1 chosen?

Indeed the left panels shows equation 6. We have added the suggested information, which indeed helps clarifying the figure caption, namely that station locations x_1, x_2 are marked by triangles and that nm are fixed to zz. Integration of omega is over the entire frequency band that was used in the simulation, but the source spectrum, which peaks at 0.05 Hz in this case, acts as a filter. (See also the point below).

The right panel, which shows the function A of eq 8, needs even more explanation. There are no spatial variables in eq 8. I assume C in eq 8 is implicitly a function of x_1 and x_2, but not of xi. It remains unclear to me which of these variables are taken along the axes in the right panel of Figure 5.

What is shown is actually not A, but the sensitivity kernel that one obtains after measuring A on both synthetic and observed cross-correlations. This was apparently wholly unclear from our description and the figure caption, since the usage of the measurement also led to questions from referee 2. Therefore, we have adapted the respective paragraph (lines 310 ff in the original manuscript) and mentioned which functions are plotted in both panels. The figure caption now reads:

"Illustration of sensitivity kernels. Left panel: Normalized vertical-component sensitivity kernel K zz (x 1 , x 2 , ξ) of full waveform L2-misfit χ fwi (equation 9). The station locations x 1 , x 2 are marked by red triangles. Frequency integration runs from 0 Hz to the Nyquist frequency, but the source spectrum peaks at dominant frequency 0.05 Hz and filters out everything above 0.1 Hz. Right: Normalized sensitivity kernel $K_{zz}(x_1, x_2, \xi)$ of windowed asymmetry measurement χ_A (equation 11). Similar figures can be obtained by adapting the Jupyter notebook tutorial for noisi."

(Note that equation numbers have shifted with respect to the original manuscript). Please also consider the answer to the last question of Referee 2, where we incorporate changes of the mentioned paragraph.

Lines 15-20: I suggest to add some references to the pioneering papers in each of the fields of application mentioned here.

We understand this suggestion well. Since there is such a large number of high-quality publications in the field, our selection here is somewhat arbitrary, and citing the pioneering papers is more systematic. We have added references, and the paragraph now reads:

"Cross-correlations of ambient seismic noise form the basis of many applications in seismology, from site effects studies (e.g., Aki, 1957; Roten et al., 2006; Bard et al., 2010; Denolle et al., 2013; Bowden et al., 2015) to ambient noise tomography (e.g., Shapiro et al., 2005; Yang et al., 2007; Nishida et al., 2009; Haned et al., 2016; de Ridder et al., 2014; Fang et al., 2015; Singer et al., 2017) and coda wave interferometry (e.g., Sens-Schönfelder and Wegler, 2006; Brenguier et al., 2008; Obermann et al., 2013; Sánchez-Pastor et al.,2019). Auto-correlations of the ambient noise are also increasingly used to study seismic interfaces as suggested by Claerbout (1968) (e.g., Taylor et al., 2016; Saygin et al., 2017; Romero and Schimmel, 2018) and to monitor subsurface properties (Viens et al., 2018; Clements and Denolle, 2018)."

Line 51: 'It presents a . . . alternative for cross-correlation modeling.' Why do you say 'an alternative'? Isn't cross-correlation modeling what the paper is about?

What we mean here is that the tool is an alternative approach to modeling cross-correlations compared to the implementations in specfem3d and salvus, which are more versatile, but also more computationally expensive. To make this clearer, we adapted the paragraph as:

"The noise cross-correlation implementations of Tromp et al. (2010) and Sager et al. (2018a) honor the physics of wave propagation to the greatest possible extent, but require substantial HPC resources for inversion (Sager et al., 2020). The noisi tool uses databases of pre-calculated seismic wavefields instead to compute cross-correlations and sensitivity kernels. It therefore presents a computationally inexpensive alternative for cross-correlation modeling and noise source inversion when updates to the structure model (i.e., seismic velocities, density, and attenuation) are not required. Owing to the reuse of Green's functions, computation is quick and inexpensive. However, storage resources, typically on the order of 1 GB per station, are needed to hold the Green's function database."

Line 93: The star instead of the bracket should be set as a superscript.

This has been corrected.