Dear editor and reviewers,

We thank you for carefully reading our manuscript tc-2019-225 entitled "On the Green's function emergence from interferometry of seismic wavefields generated in high-melt glaciers: implications for passive imaging and monitoring". On the following pages, we provide a point-by-point response to the reviewers' comments.

In our answers, we use the following color code:

- The reviewer comments are highlighted in yellow.
- Our answer is in blue.
- When applicable, our answers contain line numbers which refer to the revised manuscript.
- When applicable, our answers often quote the modified text in the revised manuscript: old and not modified text is in black, changes are in red.

In addition, enclosed you will find a list of substantial changes as well as an annotated PDF with all changes highlighted (new modified text in red, ~~old deleted text in blue~~).

If you have any questions, we would be happy to answer them. We are looking forward to hearing from you about your decision.

Best regards,

Amandine Sergeant,
Malgorzata Chmiel and the co-authors

List of substantial changes:

- We re-read carefully the whole manuscript and cut out repetitive statements, try to simplify the text and homogenize the technical terms (examples: empirical Green's function → Green's function estimate, uneven/even → an/isotropic, anti-causal → acausal, …). Lots of changes appear in the annotated pdf which do not necessarily refer to one reviewer comment. These changes were made to facilitate the reading, and remove any redundant statement. According to both reviewers.

- We shortened the general explanations we give on theoretical assumptions for the Green's function estimation (introduction) and the origin of diffusive coda waves (section 5, 5.1). We try to address more accurate statements and refer to the previous studies on theoretical developments. According to both reviewers.

- Introduction: We add explanations to better explain why seismic interferometry is difficult on glaciers as we cannot use microseism which travel through the crust given the small arrays deployed on small glaciers. According to reviewer 2.

- Section 2.1: We shortened the section.

- Section 2.2: We added a paragraph explaining why we only use vertical components. According to reviewer 1.

- Section 3.1: We rephrased one paragraph about our analysis of NCC with station pair azimuth and modify our statement on noise source distributions. According to reviewer 2.

- Section 3.2: We added explanations on how the misfit values are computed in Geopsy and what they represent for the inversion of dispersion curves and ground models. We provide additional explanations on the error expected from the depth inversion due to the uncertainties on seismic velocity measurements which arise from the array geometry. According to both reviewers.

- Section 4.2: We clarify why we use only 1 day for the MFP analysis and the noise source localizations at the Greenland ice-sheet. We provide additional details for the need of automatic criteria to recompute the same analysis on other days and for longer time range. According to reviewer 1.

- Section 5.2: We rearranged and rephrased the discussion on CWCC, scattered codas and effects of anisotropic wavefields. The discussion appears clearer. According to both revoewers.

- Figures are now referenced correctly.
- Figures: we increased the fonts on all figures.
- Figure 5c: we indicate the low P-velocity surface layer on top of ice to better illustrate the text as it was not clear to one of the reviewer.
- Figure 12: we modified the figure as it appeared that there was a filtering artefact in the acausal part.

Comments of reviewer 1

Sergeant et al. presents multiple approaches to extract Green's functions from the seismic record on the surface of glaciers. Such an effort is challenging and this study looks technically correct, but the manuscript needs several revisions. One potential major problem is the length of the paper. I found many paragraphs are not necessary after spending time to read through the article. Considering that this is not a dissertation nor textbook, reducing unnecessary texts makes the points of this study clearer. In addition to this, I provide specific comments below.

Thank you for your careful reading and comments. We shortened the text by cutting out some sections and paragraphs on general seismology. We integrate your comments into the new manuscript.

Introduction:

- L36–37: The "crustal" and "local scales" do not look contrary. Probably "local to regional" or "regional to global"?
  We modified to « from regional to local ». Line 36

- L39: What are the original observations? Please rephrase.
  We now give additional details on the original observations: « Ambient noise studies have so far led to original observations such as thermal variations of the subsoil, spatio-temporal evolution of the water content, stress changes along fault zones with applications to geomechanics, hydrology and natural hazard ». Lines 39-40

- L60: Rather than theoretically, experimentally? If authors want to claim the condition is met THEORETICALLY, please explain which theory.
  We rephrased the entire paragraph (Lines 65-68) being more precise on theoretical conditions (i.e. equipartitioned source wavefield) and the limitations of their applicability to the Earth which then ends up with the simplified assumption of a diffuse wavefield in practice. We replaced the specific statement you are referring to by: "The latter condition (ii) is sufficiently met only at high- frequency (> 0.5 Hz) in the inhomogeneous Earth's crust. Even if the noise wavefield is not generally diffuse (Mulargia, 2012), the presence of scatterers in the Earth's crust and the generation of oceanic ambient noise all around Earth make ambient noise interferometry applications generally successful. »

- L64: The statement looks the opposite. Considering that the next sentence is mentioning the source distribution, it sounds better to say: the condition (i) can compensate for the lack of condition (ii).
  Thank you for reporting this. Indeed, the statement was incorrect and has been modified.

- L80–81: Brief explanation is expected for "virtual reflector" if authors want to mention it here.
  Instead of speaking of « virtual reflector seismology », we specifically refer to the process that was used in the referred study as « multidimensional deconvolution on a contour of receivers » and refer to section 6.2. for additional explanation.

Section 2: Material and data

- L102: waves that => whose seismic waves
  We modified as suggested.

- L116: Moment magnitude? Local magnitude? Surface-wave magnitude?

We specified local magnitude.

- Figure 1: Please label all panels: one panel below b and above d.
  We changed the label accordingly.

- L123: brief => short
  We modified as suggested

- L151: Please explain why the authors use only the vertical component, while the sensor is three-component?
  We add a paragraph to explain that using horizontal components from on-ice recordings require additional processing step to reorient the sensors (Lines 158-162).

- L164: How long was the experiment? It should be explained beforehand.
  The experiment was 5 week-long as stated in Line 164.

- L166: Please make the sentence complete. GPS stations were deployed, and the DEM is used for later analyses?
  Thank you for reporting this. We rephrased the sentence. Line 174-175

- L167: What "GPR" stands for?
  We modified GPR to Ground-Penetrating Radar.

- L170: Insert space before: from
  Space added

Section 3: Glacier d'Argentière dense array

- L187–188: because of instrumental sensitivity => because of low instrumental sensitivity
  We modified as suggested.

- Figure3: ranging from 150 m to 250 m => either 150 or 250 m
  We modified as suggested.

- L213: anticausal => acausal *Please use the same word consistently throughout the Manuscript
  Modified to « acausal ». We checked for the whole manuscript.

- L223: Why is it claimed to be correct?

We want to highlight here the differences in ICC and NCC arising from different source contributions. When compared to NCC, ICC quality is very much less sensitive to the orientation of the station pair because we control the icequake source distribution. We then obtain « correct » GF estimates for ICC, in the sense that spurious arrivals vanish when controlling the icequake source aperture. On the contrary, NCC yield to poor Green's function estimates at some station paths which are not aligned with the most abundant noise sources down/upstream of the array. We rephrased the sentence as: « The control of the icequake source aperture enables to minimize the spurious arrivals which are observed on some NCC (Fig. 3b) and obtain more accurate Rayleigh wave traveltimes at most station paths (Fig. B2b) .» Lines 230-231.

- L236: Fig. 3b => 3c?
  We modified to Fig 3c

- L240–241: ICC has limited energy at low frequencies just because spectral whitening is not applied?

ICC and NCC are both computed on spectrally whitened seismograms (for same frequency corners). The difference in ICC and NCC spectral energy mainly result from the spectral content of icequake seismograms. Indeed, icequakes have short duration of about 0.2 s and then do not carry much energy below 5 Hz. We specify this: « Reconstruction of Rayleigh waves and resolution of their phase velocities using f-k processing are differently sensitive for NCC and ICC at frequencies below 5 Hz (Fig. 3c versus Fig. B2c) as ICC have limited energy at low frequency (Fig. 4a) due to the short and impulsive nature of icequake seismograms (Fig. 1d-e). » Lines 245-247

- Figure 5: What is "all" generated misfit values? Isn't it 2500, as explained in the main text? To calculate deviation around each node line, how long horizontal distance did authors consider?
  Yes, it is right and we modified the text in the caption accordingly to the main text. Ice thickness uncertainty comes from inversion results reached for misfit values that are below one standard deviation of the 2500 best misfit values. This consideration generally gives a misfit value threshold of 0.02. We added in the main text additional details on how is calculated the misfit (normalized RMS error), Lines 266-267. A maximum misfit value of 0.02 corresponds to ground models which reproduce the data dispersion curve with an approximate error of 2%. The normalizing term for the misfit function is the standard deviation of the uncertainties of the data dispersion curve measured at each frequency. These uncertainties are larger at lower frequencies below 5 Hz because we have less redundant measurements at these wavelengths considering 450m-long profiles. We specify this in Line 273-274.

- L282: What is "thickness absolute values"? Probably no need to say "absolute values" here.
  « absolute » was removed.

- L284: Why errors and uncertainties are linked to bedrock velocities? Is there any theory ? or is it just implied based on the result? Please explain.
  Errors and then uncertainties in ice thickness (as derived here from the deviation of best-fitting ground models) are linked to bedrock velocities as stated few lines before (Lines 270-272: « Walter et al. (2015) explored the sensitivity of the basal layer depth to the other model parameters and report a trade-off leading to an increase in inverted ice thickness when increasing both ice and bedrock velocities. »). Indeed, using the Greenland ice-sheet array, Walter at al (2015) find a best match for the ice-thickness of 540 m by fitting the Rayleigh wave dispersion curve obtained from match-field processing. They test the sensitivity of the result to the values of S-wave velocities when fixing them in the ice and in the bedrock. However, in our case, the velocities in the ice are very well constrained as all misfit values generally lower than 5% yield to Vs around 1700m/s. We did not perform any further tests like Walter et al, 2015 to assess the ice thickness dependency on the bedrock velocity solely. But from our inversions, we see that the ice thickness estimate slightly depends on the rock S-velocity (Fig 5b). We say on Lines 272-273: « Here the ice thickness estimation is most influenced by the rock velocities as we notice that a 100 m/s increase of basal S-velocity results in an increase in ice thickness up to 15 m. » So our statement purely comes from observations and refers to the analysis of Walter et al, 2015. We add additional details Lines 273-374 about the errors and constraints on the bedrock velocities that also related to 3D effects and larger uncertainties on the dispersion curves at lower frequencies (see also answer above): « These results are moreover influenced by larger errors at lower frequencies (Fig. 5a) which comes from less redundant measurements at large distances. »

  Walter, F., Roux, P., Röösli, C., Lecointre, A., Kilb, D., and Roux, P.: Using glacier seismicity for phase velocity measurements and Green's function retrieval, Geophysical Journal International, 201, 1722–1737, 2015.

- L315: Taking derivatives should not increase the number of cycles of sinusoids in the theory of math. Please rephrase the sentence.
  This is a misleading from the terms we used. We modified « derivation » to « the formulation

of equation 1 also gives rise to […] ».

- L319: It should be explained earlier that they calculate "within 250 m of the target point".
  We moved this sentence earlier to line 327.

- L323: Please add explain "transversal crevasses" as "perpendicular to flow" here.
  We now specify « transversal crevasses, i.e. perpendicular to the ice flow ».

- L331–332: The sentence of "However we do not exclude ..." is unclear. It sounds like some excuse for something but needs modifications to make it logical, definitely.
  We rephrased the entire paragraph. What we want to point out is that our anisotropy measurements are not punctual but are averaged over the ice column given the consistency of the align-flow fast-axis pattern with frequency. At frequencies 10-15 Hz, the sensitivity kernels are not zero near the base of the glacier (Fig 4b). We then attribute the along-flow fast axis pattern to features at depth, likely a water conduit. We rewrote the paragraph as: « Alignment of the fast-axis directions with that of ice flow appears along the central lines of the glacier (receiver lines 4-5) with anisotropy degrees of 0.5% to 1.5%. This feature is only observed along the deepest part of the glacier where it flows over a basal depression. Results are here computed for seismic measurements at 25 Hz and maps of anisotropy do not change significantly with frequency over the 15-30 Hz range. If we extend our analysis down to 7 Hz, we notice that the aligned-flow fast-axis pattern starts to become visible at 10 Hz. At frequencies lower than 10 Hz, the fast-axis generally tend to align perpendicular to the glacier flow because lateral topographic gradients introduce 3D effects and non-physical anisotropy. The results presented here are not punctual measurements but are rather averaged over the entire ice column. The vertical sensitivity kernels for Rayleigh waves (Fig. 4b) are not zero in the basal ice layers at the considered frequencies. The align-flow anisotropic pattern is likely attributed to a thin water-filled conduit as also suggested by locations of seismic hydraulic tremors at the study site (Nanni et al, 2019). » Lines 335-340

- L333: aligned-flow => flow-parallel
  We did not modify the wording.

Section 4: MFP at the Greenland ice-sheet

- L350: Short explanation of MFP is needed here. Especially, providing the direct purpose of the processing helps readers. Here, "to locate specific localize sources" would be appropriate.
  We added the following short explanation: "One of the approaches we apply here is Matched-Field Processing (MFP) (Kuperman and Turek, 1997), which is an array processing technique allowing to locate low-amplitude sources. MFP is similar to a traditional beamforming that is based on phase-delay measurements". Lines 356

- L361: "localized" is better than just dominant.
  We removed "dominant" from the title of the subsection: "Location of noise sources at the GIS via matched-field processing".

- L380: frequency band of => frequency band between
  Modified

- Figure 7: frequency band of => frequency band between
  Modified

- L403: SVD itself is not a process to decrease the number of information, and therefore it is inappropriate to mention "as few coefficients as possible" in this sentence.

We corrected the sentence to: "SVD is a decomposition of the CSDM that projects the maximum signal energy into independent coefficients […]."

- L408: Are eigenvectors normalized to be unit?
  U and V are unitary matrices, so their columns (eigenvectors) are unit vectors (e.g., |eigenvector| = 1) forming an orthonormal basis. We replace "orthogonal matrices" with "unitary matrices".

- L409: Please choose either "singular value" or "eigenvalues" to be consistent throughout the manuscript.
  We choose "eigenvalues" and modified the text accordingly.

- L414: threshold in "singular values / eigenvalues"?
  As above.

- L415: Eigenvector is not a scalar, so it is inappropriate to say "eigenvector" above the threshold.
  We changed the phrase to "the index of eigenvectors".

- L426–427: Why it is thought to be related to frequency content differences? Different eigenvector does not necessarily mean they consist of different frequency contents.
  We think that the dominant source (the moulin) is located either in the first eigenvector or the second eigenvector depending on the frequency. This can be related to the seismic signature of the tremor and the distinctive frequency bands of either elevated or suppressed seismic energy that can be also noted on the spectrogram of the hydraulic tremor (Figure 1c).
  We changed the following sentence: "This might be related to the change in the distribution of the dominant sources depending on the frequency related to the seismic signature of the hydraulic tremor and the distinctive frequency bands generated by the moulin activity (Figure 1c)."

- L427: SVD => eigenvalue?
  We mean the "eigenvalue distribution". We corrected the phrase to: "Moreover, the eigenvalue distribution decays steadily […]."

- L427–428: I do not understand why the number of ambient noise sources is more than the number of receivers based on this observation. Please explain this more logically.
  We do not claim that the number of noise sources is higher than the number of receivers, but that the number of degrees of freedom (number of parameters that can be used to describe the seismic wavefield, as explained earlier in the manuscript L:400-405) is higher than the number of receivers, as defined in Seydoux, 2017. This confirms that the wavefield is undersampled by the seismic array (see Seydoux et al, 2017 for details).
  We changed the phrase into the following: "The latter confirms that the wavefield is undersampled by the seimic array (see Seydoux et al., 2017) for details)."

- L429: reconstructed => decomposed
  Here we mean "reconstructed", as we talked about the reconstruction of the CSDM by using individual eigenvectors.

- L430: Please use a different symbol for K_i, which is different from the previous K_i.
  Thank you for this comment, we now use the symbol $\widetilde{K}_i$ for the reconstructed CSDM.

- L444–446: This explanation is too redundant: i.e., L431–432.
  We removed the last two repeating phrases.

- L461: wavenumber vectors are normalized => wavenumber is normalized

We change the phrase to the following: "The wavenumbers k_x and k_y are normalized by the wavenumber k_0 corresponding to Rayleigh wave slowness s=1/1680 s/m."

- L481: Why only one-day used in this study?

  Please see our reply below.

- L483: Again, it is too straightforward to apply the same analyses to the other day, but why authors are not interested in doing so as a part of this study?

  We agree that it is straightforward to apply the same analysis to another day. For example, Figure R1 shows the same analysis for the 28th of July (the following day after the day presented in our study). Moreover, we can take it even further, and stack the Green's function estimates for these two days (Figure R2). However, one can see that the moulin seismic signature is located in the second eigenvector for the 27th of July and the first eigenvector for the 28th of July. Therefore, in order to extend this analysis to other days, one should find an automatic criterion to find the index of eigenvectors that corresponds to the moulin. Selecting the eigenvectors manually in order to extend the study to more days would be a tedious task and it would not change the interpretation of the current results.

  We added the following paragraph in the manuscript, although we do not think that adding the Figures R1, and R2 would enhance the manuscript scientific impact. However, if the Reviewer thinks that this is important, we can add the Figures R1, and R2 in the Appendix.

  "For example, a similar procedure could be performed on other days and the eigenormalized NCF could be stacked over a few days to increase the SNR. However, we verified that the index of eigenvectors corresponding to the moulin changes over days (the moulin can be located in the first, second, third etc. eigenvector). This is the reason why it would be useful to find an automatic criterion for the eigenvalue selection based on the MFP output. However, this is beyond of the scope of this paper."

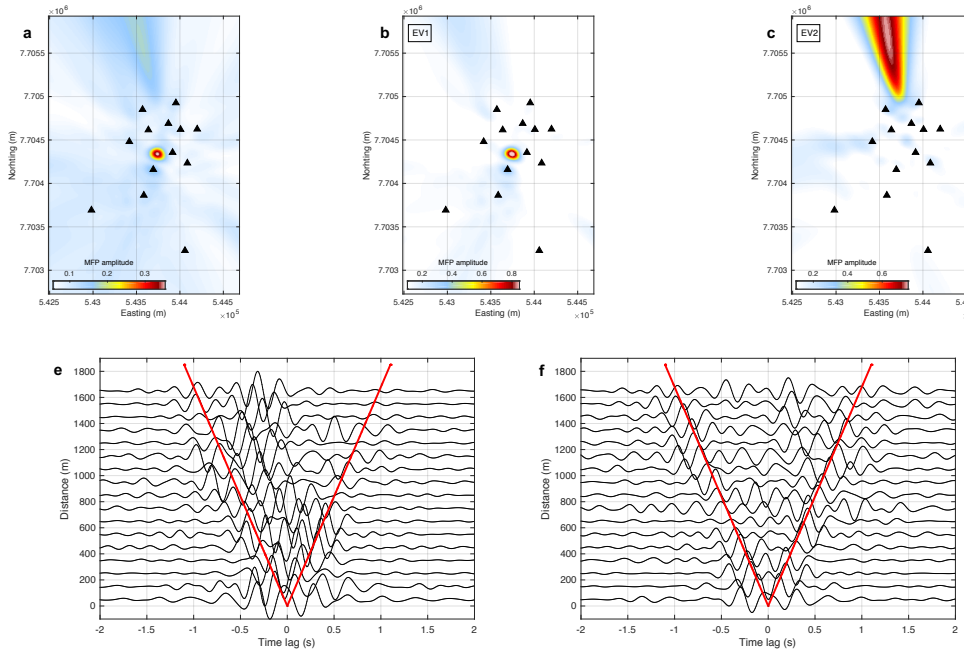

Figure R1: (a) Location of the dominant noise sources using MFP in the frequency band of 2.5 Hz and 6 Hz (the MFP output is averaged over 30 discrete frequencies). (b), (c) Reconstruction of the CSDM for the 28th of July by using first (in (b)) and second (in (c)) eigenvectors that are related to different noise sources. Each figure represents the MFP gridsearch output calculate for the corresponding eigenvectors. (e) Stacked sections of NCC in the frequency band 2.5-6 Hz. The red line shows the

propagation of the Rayleigh waves with velocity of 1680 m/s (also in f). (f) Stacked sections of NCC reconstructed in the frequency band 2.5-6 Hz from the CSDM eigenspectrum equalization.

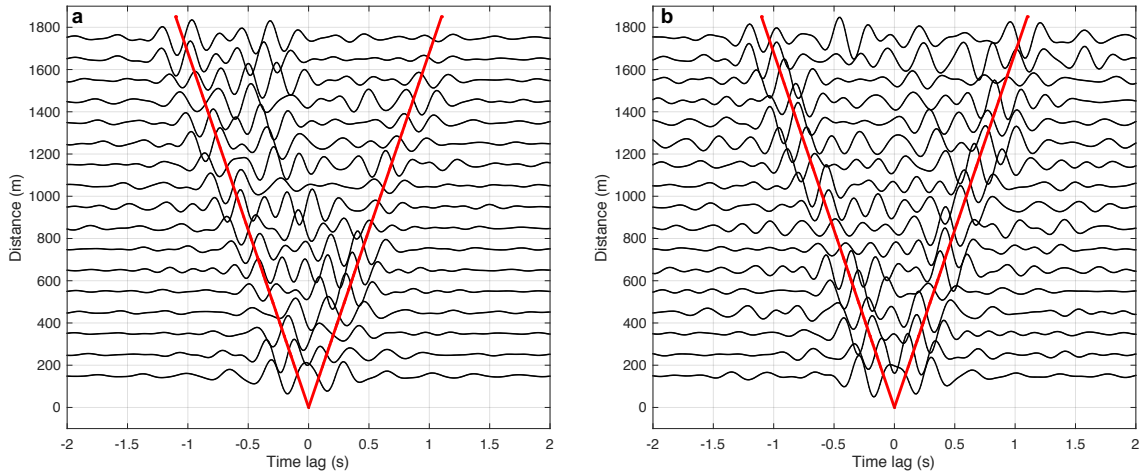

Figure R2: (a) Stacked sections of NCC in the frequency band 2.5-6 Hz averaged over two days (the 27th and the 28th of July 2016). The red line shows the propagation of the Rayleigh waves with velocity of 1680 m/s (also in f). (b) Stacked sections of NCC reconstructed in the frequency band from 2.5 to 6 Hz from the CSDM eigenspectrum equalization averaged over two days (the 27th and the 28th of July 2016).

Section 4: CWI at Gornergletscher

- L493: Subtitle should represent what is done in each section. In this sense, the study site "Gornergletscher" is not necessary and instead making it confusing.
  We removed « using the Gornergletscher array » from the title. To be consistent all through the manuscript, we also removed « using the GIS array » from the title of section 4. We now refer to the study sites in the following first section (i.e. « 5.1. Icequake coda waves at Gornergletscher »).

- L525: Fig 1c => 1d ?
  We modified to Fig 1d.

- L532: Fig 1c => 1d ?
  We modified to Fig 1d.

- L540: Do authors see the potential effect of early aftershock in the coda wave?

Yes, we must carefully select event codas to avoid anisotropic fields generated by incident waves from aftershocks. The event selection is a first important step as stated in the manuscript. The window-optimization scheme further helps to reduce their potential effects. Aftershock affect and cause of spurious arrivals are explained in Lines 598-599 "spurious arrivals at times 0 or later could result from seismic reflections on the glacier bed beneath the stations, cross-correlations of early aftershocks or other noise sources. » and lines 617-620 « the abundance of seismic sources in glaciers often pollutes coda wave seismograms. We often find the situation where ballistic body and surface waves generated by early aftershocks from repetitive and subsequent events (or bed reflections) arrive at the seismic sensor only a few milliseconds after the onset of the first event of interest and therefore fall in its coda window. This typically introduces anisotropic wavefields. »

- L542: Please show the coda window to be used in the figure 1d as an example.
  We indicated the coda window by the gray horizontal bar in Fig 1d.

- L542: Fig 1c => 1d
  We modified to Fig 1d.

- L546: Does it mean the authors are applying spectral whitening and 1-bit normalization together?
  The waveforms are first spectrally whitened as stated in Line 534 and then one-bit normalized.

- L572–586: Please explain this as a regular paragraph. In addition, many words are redundant (e.g., N or M is already explained).
  We rephrase this as a regular paragraph and removed the redundant information.

- L613: Indeed, the arrivals could result from the reflection at the bed. Isn't it easy to calculate and validate?
  It is not easy to calculate as at the study site, we are very close to the glacier margins. We are sensitive to 3D effects from the lateral margins and also strong basal topographic gradients with about 200 m of depth difference around the array (Walter et al. 2009). Such spurious arrivals may result from the correlation cross-terms from very different reflection arrivals in the coda.
  Walter, F., Clinton, J., Deichmann, N., Dreger, D. S., Minson, S., and Funk, M.: Moment tensor inversions of icequakes on Gornergletscher, Switzerland, Bulletin of the Seismological Society of America, 99, 852–870, 2009.

- L629: Surprisingly, there are no further analyses of the GF in this section.
  We improved the discussion and re-organized it. We already discussed the obtained GF in terms of (1) symmetry, (2) spurious arrivals, (3) (non-)diffuse character of coda waves which arise from single scattering rather than multiply-scattering (medium effect) and non-homogeneously distributed sources and (4) our ability to compute reliable Rayleigh wave dispersion curves. We expose the advantage of CWI to obtain Green's function at seismic arrays where uneven icequake or noise source distributions prevent the Green's function estimation.

Section 5: Discussion

- L643: allow => allows
  Modified

- L695: Does SH refer to SH wave?
  Yes. Specified

Comments of reviewer 2

Sergeant et al. present an overview of different methods used to estimate structural information about glaciers and an ice sheet using seismic waves. The main focus of the study is on estimating the direct Rayleigh wave between seismic station using passive recordings of ambient noise and (near-surface) icequakes. The authors present multiple methods (traditional correlation, MFP, MDD, etc.) and highlight the benefits and limitations of each method. The overall goal is to show the usefulness of high-density seismic arrays deployed on ice to image the structure directly beneath the ice, as well as monitoring changes in this structure through time.

The authors demonstrate each of the proposed methods, albeit using dataset from different glaciers. All of the results are quite convincing, but I do question some of the interpretations about why certain features exist in the recovered wavefields. I have listed those below in the comments. In general, the paper is well written and clear. In some aspects the text needs to be tightened up though. Statement are not entirely complete or not entirely accurately described.
The figure fonts can all be enlarged as well. In general, the actual results in all methods are very nice and I congratulate the authors on recovering such nice GFs from icequake and glacier noise data. This is not easy.
Thank you for your comment and appreciation. We improved the discussions you were referring to in the comments below (i.e. distributed noise sources in Argentière, anisotropic wavefields from coda waves at Gornergletscher). We shortened some parts, revised our writing and used fewer technical terms to smoothen the reading. We revised the general statements on seismic interferometry with more accurate arguments. We enlarge the fonts on every figure.

I have listed major comments/concerns below and provide an annotated PDF with many more minor comments.

Introduction:

- Paragraph around line 50: Interferometry recovers an approximation to the Green's function/impulse response. There are many assumptions that influence the accuracy of this approximation. This should be better explained or at least noted. Also, the description of the causal and acausal parts of the GF estimate should not be limited to only the direct wave (as is currently done). The more accurate way to characterize what is happening is to describe virtual sources. The direct wave is commonly observed because not all of the assumptions in SI are valid in most field data studies. This paragraph as written is too simplistic and not an accurate depiction of the theory. Please revise to be more complete.
  We rephrased the paragraph with more accurate statements (Lines 48-55). We use the terms "approximate" rather than "reconstruct" the Green's function. We always say "under specific conditions" or "simplified approximations" that refer to the next paragraph which further details the theoretical requirement for an equipartitioned wavefield (which is often not met in reality) and is replaced in practice by the diffusive condition of the seismic wavefield (see comment below). For the assumptions in seismic interferometry, we also refer later to the study of Fichtner et al (2017) which give an accurate overview. We removed the term "direct wave".

- Paragraph starting at line 55: a diffuse or equipartitioned wavefield is not the same thing. Equipartition means that all modes are excited (P,S,Rayleigh,Love,etc.). Diffuse means waves propagating in all directions. This distinction is commonly neglected by most people that write about SI. The current description in this manuscript again confuses these two distinct properties of the wavefield. Please revise throughout the manuscript.
  We corrected this. We give a correct definition of equipartion. We rewrote the paragraph to say

that equipartition is a theoretical requirement for Green's function estimate from interstation correlation, but is not met in practice in the Earth. We then work in simplified approximations of a diffuse wavefield. We modify the paragraph as (Lines 57-63): "In theory, the GF estimate is obtained in media capable of hosting an equipartitioned wavefield, that is random and uncorrelated modes of seismic propagation with same amount of energy. In practice, the equipartition argument has limited applicability to the Earth because non-homogeneously distributed sources, in the forms of ambient noise sources, earthquakes and/or scatterers, prevent the ambient wavefield from being equipartitioned across the entire seismic scale (Fichtner et al, 2017, and references therein). The GF estimation from interstation correlation therefore usually relies on simplified approximations of diffusive wavefields which can be reached in (i) the presence of equally-distributed sources around the recording network (Wapenaar, 2004; Gouédard et al., 2008b) and/or (ii) in strong-scattering settings as scatterers act like secondary seismic sources and likely create a diffuse homogenized wavefield in all propagation directions (e.g. Hennino et al., 2001; Malcolm et al., 2004; Larose et al., 2008). »

- Line 60: The statement that strong scattering exists in the crust is not true at the frequencies commonly used for ambient noise tomography. This statement is very much untrue. At high frequencies (>0.5 Hz) this statement is true, but ambient noise imaging often works because oceans generate the microseismic wavefield all around earth so it is more accurate that condition (i) is met. See the following reference and the papers that have since referenced this paper. Mulargia, F. (2012). The seismic noise wavefield is not diffuse. The Journal of the Acoustical Society of America, 131(4), 2853. https://doi.org/10.1121/1.3689551
  Thank you for bringing this up. We modified the paragraph and include your suggestions as (Lines 65-68): "The latter condition (ii) is sufficiently met only at high-frequency (> 0.5 Hz) in the inhomogeneous Earth's crust. Even if the noise wavefield is not generally diffuse (Mulargia, 2012), the presence of scatterers in the Earth's crust and the generation of oceanic ambient noise all around Earth make ambient noise interferometry applications generally successful. »

- The sentence beginning on line 61 is also not entirely accurate. Most studies on glaciers have been unable to reconstruct GFs not because of the lack of scattering but because of the dominant frequency of the background noise. Seismic arrays on glaciers are tiny compared to regional or continental arrays. In order to recover a usable GF in the microseism band you need stations that are more than 1 wavelength apart (neglecting methods like SPAC). When we correlate signals on glaciers in the microseism band the resulting correlations look like autocorrelations because the sensors are pretty much in the same location at the wavelengths of the microseism band. It is more appropriate to state that the noise field lacks the high frequencies needed to generate GFs that contain useful information at the scale of the glacier. If you wish to use icequakes with frequencies above 0.5 Hz, then yes, your statement is accurate, but you should explicitly state this. Everything depends on the frequencies considered and you are neglecting this point in the way that you are writing these statements.
  That is correct. We add two sentences about these details (Lines 69-72): "In glaciers, the oceanic ambient noise field commonly used in crustal studies lacks the high frequencies needed to generate GFs that contain useful information at the scale of the glacier. To target shallower glaciers and their bed, we must work with other sources such as nearby icequakes and flowing water which excite higher-frequency (> 1 Hz) seismic modes (Sect. 2.1). In this context, the lack of seismic scattering […] »

- Line 75: You say on line 73 that they do obtain accurate GFs, but then on line 76 you say they don't obtain accurate GFs. Which is it?
  Indeed, Preiswerk and Walter (2018) were able to compute accurate Green's functions on some glacier settings which presented at that time efficient drainage systems. However, due to changes in the drainage system (localized noise sources which sometimes appeared, or initiation of lake drainage), they could not compute accurate GF at every time scale. We add this late info without going into such details: "However, due to localized noise sources in the drainage system that also change positions over time over the course of the melting season, they could not

systematically obtain accurate coherent GF when computed for different times, limiting the applications for glacier monitoring."

Section 2: Material and data

- Figure 1: fonts are way too small. I also cannot tell which color is "this event" or the "1000 event average". The colors look identical to me. I am assuming the smoother line is the average.
  We increased the fonts and modified the color of the red line to blue in Fig 1d.

- Line 143: Figure 1b –> Figure 1c
  We modified to Fig. 1c.

- Line 166: What is the reason for the partial statement about the 20m resolution DEM? It does not make sense in this sentence. Please read out loud to yourself to see the mistake.
  We rephrased the sentence which had indeed no meaning. The 20 m resolution of the DEM comes from the spatial interpolation of the GPR tracks.

- Figure 2: Fonts on axes are again very small.
  We increased the fonts.

- Appendix A: Line 783: What does "network vs. array" mean?
  We distinguish an array and a network mainly from the processing involved and the size of the array with respect to the considered seismic wavelength. We explain this in a first paragraph in appendix A3. We removed this statement here as it is explained a few lines later.

- Line 793: Do you really mean to reference Fig. 1c here? This is a figure of the spectrogram of a GIS signal.
  We modified to Fig.1a.

Section 3: Glacier d'Argentière dense array

- Line 194-195: plane wave approximation –> stationary phase approximation. I do not understand why plane wave is used here. The proper interpretation of the sinusoidal shape is the stationary-phase. See Snieder, R., Van Wijk, K., Haney, M., & Calvert, R. (2008). Cancellation of spurious arrivals in Green's function extraction and the generalized optical theorem. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics, 78(3), 1–8. https://doi.org/10.1103/PhysRevE.78.036606
  We refer to the plane wave approximation as we restrict ourselves to far-field events. The Green function is recovered when the incident plane waves arrive with equal strength from all azimuths (and with randomly distributed and statistically independent amplitudes, i.e. with homogeneous source distributions). You need same azimuthal angle for same cross-correlation arrivals at both stations for individual sources and this is induced by the plane wave approximation (Gouedard et al, 2008). In contrast, to get accurate GF, you only need stationary phase sources. We did not modify our statement, although we revised the paragraph for better clarity.

- Line 194: Why are you referencing Fig. B4 before Fig B1, B2, or B3? Please fix the referencing so that things are referenced in order of appearance. It makes reading easier.
  Figure B4 is here to illustrate two things: (1) the source-azimuthal dependency of the reconstructed arrival times in the cross-correlation functions when computed on plane waves (as explained in the main text, Lines 200-210), and (2) the method introduced to measure Rayleigh wave phase velocity from the sinusoidal fit to the azimuthal-dependent arrival times

(as described in appendix B3). Figure B4 is then an integral part of the appendices and is more cited here as a supplementary figure for readers who are not familiar with this. Figure B4 is difficult to be renamed without completely rearranging the appendix which would not completely make sense as the appendix is ordered in the same way as the referencing to the different processing techniques that are used in the main text. We did not modify the referencing to figure B4.

- Figure 3 caption: Can you please explain why you think the GF converge better in the along-flow direction based on Fig. 3b? I wonder if you are seeing anisotropy in the ice velocities, rather than some sort of convergence related to the strongest noise sources. That reasoning is somewhat counter to your argument for sign-bitting the data. It can be that the density, not the amplitude, of sources is larger in the along flow direction. That would explain differences in convergence, but what you are stated here is not quite correct. Please revise. Also, note that in line 215 you state that the sources are located homogeneously around the array, which implies the density of sources is even with azimuth. Is this true? Did you do beamforming to look at the azimuthal amplitude of incident waves on the array?

  You are right about this. We used misleading terms when speaking about "stronger sources" rather than "more sources". We modified the text and caption accordingly. We modified our statements about distributed noise sources which are, according to our analysis, not distributed homogeneously. We rewrite the paragraph (Lines 219-229) to better explain this.

  Now in our discussion, we separate spurious arrivals at time 0 at station pairs perpendicular to the flow line, with spurious arrivals at other pairs constituted by faster waves which maybe arise from non-aligned sources and/or anisotropy as you point out. We modified the text as: "More sources downstream are likely generated by faster water flow running into subglacial conduits toward the glacier ice fall (Gimbert et al, 2016, Nanni et al., 2019b). Indeed, looking closer at NCC for individual receiver pairs, we sometimes observe spurious arrivals around time 0 (marked as green dots in Fig. 3b), mostly at stations pairs which are oriented perpendicular to the glacier flow (i.e. azimuth $0° \leq \psi \leq 50°$), indicating that dominant noise sources are located along the flow line. At other station pairs (i.e. azimuth $\psi \sim 90°$), the reconstructed arrival times are slightly faster than expected. This could be an effect of non-distributed noise sources, or anisotropy introduced by englacial features (Sect. 3.3). This analysis shows that even if the noise sources are not equally distributed in space, averaging the NCC in regular distance intervals on a dense array deployment helps the GF convergence. »

  Furthermore, we located the noise sources by taking advantage of the causal/acausal amplitude asymmetry and the spurious arrivals (approach similar to Stehly et al, 2009 and Retailleau et al 2017). We find that dominating sources are near the glacier ice fall (downstream) and upstream of the array as stated in the main text, but also along the glacier flow line at the center of the array. This observation also satisfies our expectation for an englacial water conduit along the flow line given the along-flow anisotropic patterns we measure at the array center (Fig.6b and modified text in section 3.3). Nanni et al (AGU Fall Meeting abstract 2019) also looked at the locations of seismic hydraulic tremors with beamforming and their observations are well correlated with our noise source locations. Noise source locations are beyond the topic of the paper as our take-home message is on the advantages of using dense arrays which allow to stack and average the GF estimates (see modified text above). We do not further discuss the noise source locations as it will be part of a future study.

  Stehly et al, 2009 : A study of the seismic noise from its long-rangecorrelation properties, JGR.

  Retailleau et al, 2017 : Locating microseism sources using spurious arrivals in intercontinental noise correlations, JGR

  Nanni et al, 2019 : Mapping the Subglacial Drainage System from Dense Array Seismology: a Multi-method Approach, AGU Fall Meeting abstract.

- Figure 3: Text is again very hard to read.
  We increased the fonts.

- Figure 3: The dashed blue line is not the array response. That is the frequency-dependent resolution limit. You actually correctly state this in line 238.
  We modified the caption accordingly to the main text.

- Line 226: Don't you mean 1b, not 1a?
  We modified to Fig.1b

- Line 227: See annotated PDF. This first sentence can be stated more accurately because not all phases are dispersive. Instead, you are using the f-k domain to identify phases. We just happen to that particular transform a lot for surface wave dispersion, but as you show in your dispersion image, the P wave is not dispersive.
  Thank you for reporting this. We modified our statement accordingly (Line 235): « Seismic phases and their velocities can be identified on the frequency-velocity diagram (Fig. 3c, black dots) that is obtained from frequency-wavenumber (f-k) analysis […] ».

- Line 236: Fig. 3b –> Fig. 3c (You should really pay attention to not mislabeling your figures in the future.)
  We modified to Fig.3c.

- Line 240: Fig. B2b –> Fig. B2c.
  We modified to Fig B2c.

- Figure 4: axes fonts could be larger
  We increased the fonts.

- Figure 5: What are the units on the misfit values? Are the misfits the same in (a) and (b)?
  The misfit values are the same in (a) and (b) and are calculated as the normalized RMS of the residuals between dispersion curves. More precisely, the misfit can be defined in Geopsy as the RMS error normalized by the standard deviation provided by the error estimate on the dispersion curve to be fitted. When the maximum misfit value is reached (i.e. 0.1), the dispersion curve is reproduced with an approximate error of 10%. We explain this more accurately in Lines 266-270: « Misfit values correspond here to the root-mean square error on the dispersion curve residuals, normalized by the uncertainty average we obtained from the seismic data extraction (error bars in Fig. 5a). The inversion well resolves the S-wave velocity in the ice layer as all best matching models yield to V_S = 1707 m/s for misfit values below 0.05 meaning that the data dispersion curve is adjusted with an approximate error below 5%. »

- Line 258: Fig. 3b –> 3c
  We modified to Fig 3c.

- Table 1 states that Vs in the granite can be as low as 1000 m/s, but there are not gray lines in Figure 5b that show you tested this velocity. Can you please explain why? I think it would be easy to change the range in Table 1 and not influence the results of the inversion. It actually appears that the lower layer velocity never goes below the upper layer velocity. Is there something in Geopsy that imposes increasing depth and prevents low velocity layers?
  This is right, thank you for bringing this up. Indeed, we force Geopsy to explore only increasing velocities with depth, so the lower layer velocities are always higher than the ice velocities. We then have modified the velocity range for the low layer in Table 1, and add this condition to the text, Line 264.

- Line 271: I do not follow the statement that the ice thickness is 7 to 15 meters thick on the edges.

Figure 5c shows ice on the edges more than 100m thick. Can you please explain this discrepancy between the text and the figure? Am I missing something here?
Indeed, the previous text was maybe unclear. Seismic inversions for the lines on the array edges yield to the determination of three layers: (1) a thin layer of 7-15 meters where we find decreased P-velocities with respect to the ice velocities, (2) the ice layer and (3) the bedrock. We discuss the origin of the low P-velocity zones which could be attributed to snow and more likely to the presence of transversal crevasses which introduce anisotropy that can then be modelled by a slow top layer of a few dozens of meters as shown by Lindner et al, 2018a.
We also modified Figure 5c to represent this low velocity layer on the top of the ice. We now better explain this in the main text: « For the receiver lines near the array edges (Lines 1-3 and 8), the inversion yields to a low P-velocity surface layer of thickness 15 m and 7 m respectively, above thicker ice (dashed blue zone in Fig. 5c). […] This low-velocity surface layer could also at least partially be attributed to the presence of pronounced transversal crevasses (i.e. perpendicular to the receive lines) near the array edges, which do not extend deeper than a few dozens of meters (Van der Veen, 1998) and can be modelled as a slow layer above faster underlying ice (Lindner et al., 2018a). »

Lindner, F., Laske, G., Walter, F., and Doran, A. K.: Crevasse-induced Rayleigh-wave azimuthal anisotropy on Glacier de la Plaine Morte, Switzerland, Annals of Glaciology, pp. 1–16, 2018a.

- Appendix B: Line 915: SPAC works for single stations when you have an isotropic incident wavefield, otherwise you need an array and averaging. You even state this on line 919 with "azimuthally averaged". You should be careful with your wording in line 915. You are not telling the whole story.
  Thank you for reporting this. We added the condition of the isotropic wavefield in Line 915: « this technique does not require specific array geometries to compute phase velocities and can be used on single pairs of stations as long as you are in the presence of an isotropic incident wavefield. »

Section 4: CWI at Gornergletscher

- Section 5.1: Do you really need an "origin of coda waves" section? This is already explained with a references in the introduction of the paper. To me this paper is unnecessarily long because everything is explained rather than simply cited.
  We shortened this section and moved some general descriptions with appropriate referencing on the diffusive character of coda waves to the introduction of section 5. We renamed section 5.1 as « Icequake coda waves at Gornergletscher ». This section now focuses on the description of icequake coda seismogram.

- Line 533: What is your reason to state that the energy is back-scattered? Rayleigh waves have significant forward scattering. See Snieder, R. (1986). 3D linearized scattering of surface waves and a formalism for surface wave holography. Geophysical Journal of the Royal Astronomical, 581–605, in particular Figures 6 and 7 for example.
  We did not want to imply anything on the back or forward scattering of the Rayleigh wave propagation mode. We were referring to single versus multiply scattering and removed the term "back-scattered" accordingly.

- Figure 10: Why are the azimuth ranges in (a) and (b) not the same? Are you using difference sources for each station? Or is the azimuth relative to the interstation path, rather than absolute azimuth? I would think the two matrices should be missing the same azimuths if the icequakes used were the same in the two cases. (It is a very nice result by the way!!)
  We are essentially using the same sources for the two station paths, except that we exclude events that lie in the vicinity of the stations, i.e. at distances shorter than approximately half of the interstation distance given the considered seismic wavelength (as described in the general section 3). As the first station pair in Fig 10a is closer than the one in Fig 10b, we excluded more

events for the computation in Fig 10a. Furthermore, the azimuth is here defined as the event azimuth relative to the station path (and not absolute azimuth). As the two station paths are oriented differently with respect to the source distribution, it results in a different azimuth range. The definition of the azimuth was well stated in the main text (Line 563). We define it correctly in the figure caption.

- Line 556: Why not beamform the coda? Take the average beam over all time windows. This would highlight illumination problems.

  We did not try beamforming for the following reasons:
  - The short coda duration and the fast drop of coda coherency across the array prevents for time-lapse beamforming. This is also why CWI fails at some station paths as in here.
  - Low beamformer resolution given the seismic array and coda coherency
  - Given sensitivity kernels of seismic coda, seismic coda are strongly sensitive right beneath the stations, and we do not expect to see any convergence to the illuminated zones given the two points above.

- Line 616: What is an anisotropic diffuse wavefield?

With "anisotropic diffuse wavefield" we were referring to the anisotropy of the energy flux which is still observed in the late coda as it is still contains information on the source incidence. To overcome this effect in weak scattering medium, numerical simulations of Paul et al (2005) suggest to work with equally-distributed sources. We modified "anisotropic diffused wavefield" to "the long-lasting anisotropy of the diffuse energy flux ». We add these explanations (Lines 608-610): « Indeed, in weak (or homogeneous) media, the incident energy flux from earthquakes can still dominate the late coda resulting in GF time-asymmetry, provided the sources are located in the same distant region. The CWCC asymmetry is expected to disappear with an isotropic distribution of sources or scatterers around the seismic network. ». In general we rearranged and rephrased the whole discussion for more accurate and clearer statements.

- Please also note the supplement to this comment: https://www.the-cryosphere-discuss.net/tc-2019-225/tc-2019-225-RC2-supplement.pdf

  Thank you for such a careful reading. We took into account all of your suggestions.

On the Green's function emergence from interferometry of seismic wavefields generated in high-melt glaciers: implications for passive imaging and monitoring

Amandine Sergeant[1,2], Małgorzata Chmiel[1], Fabian Lindner[1], Fabian Walter[1], Philippe Roux[3], Julien Chaput[4], Florent Gimbert[3], and Aurélien Mordret[3,5]

[1]Laboratory of Hydraulics, Hydrology and Glaciology, ETH Zürich, Zürich, Switzerland
[2]Now at Aix Marseille Univ, CNRS, Centrale Marseille, LMA, France
[3]University of Grenoble Alpes, CNRS, IRD, ISTerre, Institut des Géosciences de l'Environnement (IGE), Grenoble, France
[4]Department of Geological Sciences, University of Texas El Paso, El Paso, USA
[5]Massachusets Institute of Technology, Boston, USA

Correspondence: Amandine Sergeant (sergeant@lma.cnrs-mrs.fr)

Abstract. Ambient noise seismology has revolutionized seismic characterization of the Earth's crust from local to global scales. The elastic Green's function (GF) between two receivers can be reconstructed via cross-correlation of the ambient noise seismograms. An homogenized wavefield illuminating the propagation medium in all directions is a pre-requesite for obtaining accurate GF. For seismic data recorded on glaciers, this condition imposes strong limitations on GF convergence, because of minimal seismic scattering in homogeneous ice. We address this difficulty by investigating three patterns of seismic wavefields: a favourable distribution of icequakes and noise sources recorded on a dense array of 98 sensors on Glacier d'Argentière (France), a dominant noise source constituted by a moulin within a smaller seismic array on the Greenland ice-sheet, and crevasse-generated scattering at Gornergletscher (Switzerland). In Glacier d'Argentière, surface melt routing through englacial channels produces turbulent water flow creating sustained ambient seismic sources and thus favorable conditions for GF estimates. From the velocity measurements of reconstructed Rayleigh waves, we invert bed properties and depth profiles, and map seismic anisotropy, which is likely introduced by crevassing. In Greenland, we employ an advanced pre-processing scheme which include match-field processing and eigenspectral equalization of the cross-spectra to remove the moulin source signature and reduce the effect of inhomogeneous wavefields on the GF. At Gornergletscher, cross-correlations of icequake coda waves show evidence for homogenized wavefields. Optimization of coda correlation windows further promotes the GF convergence. This study presents new processing schemes on suitable array geometries for passive seismic imaging and monitoring of glaciers.

1 Introduction

Passive seismic techniques have proven efficient to better understand and monitor glacier processes on a wide range of time and spatial scales. Improvements in portable instrumentation have allowed rapid deployments of seismic networks in remote terrain and harsh polar conditions (Podolskiy and Walter, 2016; Aster and Winberry, 2017). Studies on seismic source processes have revealed unprecedented details about englacial fracture propagation (e.g. Walter et al., 2009; Mikesell et al., 2012), basal processes (e.g. Winberry et al., 2013; Röösli et al., 2016a; Lipovsky et al., 2019), glacier hydrology (Bartholomaus et al., 2015; Gimbert et al., 2016) and iceberg calving (e.g. Walter et al., 2010; Bartholomaus et al., 2012; Sergeant et al., 2016, 2018).

The subsurface structure of ice sheets and glaciers has been characterized by analysis of seismic wave propagation in ice bodies. As a few examples, Harland et al. (2013) and Smith et al. (2017) used records of basal seismicity to measure elastic anisotropy in two Antarctic ice streams. Lindner et al. (2019) identified crevasse-induced anisotropy in an Alpine glacier from velocity anomalies by analyzing icequake seismograms at seismic arrays. Walter et al. (2015) used transient seismic signals generated in moulins to compute frequency-dependent seismic velocities through matched-field processing and estimate the depth of the ice-to-bedrock transition beneath a seismic network deployed on the Greenland ice-sheet (GIS).

At the same time, a new approach appeared in seismology ~~about a decade ago~~ which explores not only earthquakes but also ambient noise sources generated by climate and ocean activity (Ekström, 2001; Rhie and Romanowicz, 2004; Webb, 1998; Bonnefoy-Claudet et al., 2006). Shapiro and Campillo (2004) and Shapiro et al. (2005) pointed out the possibility of using continuous noise recordings to reconstruct propagating surface waves across a seismic array and to use them for crustal tomography in California. Other studies followed, shaping the ambient noise background into a powerful tool to constrain the elastic properties of the illuminated medium, making it possible to image the Earth's interior from ~~crustal~~ regional (Yang et al., 2007; Lin et al., 2008) to local scales (e.g. Lin et al., 2013; Nakata et al., 2015) and monitor e.g., seismic fault (e.g. Brenguier et al., 2008b; Olivier et al., 2015) and volcanic processes (Sens-Schönfelder and Wegler, 2006; Brenguier et al., 2011). Moreover, ambient noise studies have so far led to original observations such as thermal variations of the subsoil, spatio-temporal evolution of the water content, stress changes along fault zones ~~of the subsurface structure and evolution~~ with applications to geomechanics, hydrology and natural hazard (Larose et al., 2015, and references therein).

For the cryosphere, few studies have successfully used oceanic ambient noise records from permanent broadband stations deployed on the rocky margins of glaciers or 10 to 500 km away to monitor the subsurface processes. Mordret et al. (2016) and Toyokuni et al. (2018) tracked the strain evolution in the upper 5 km of the Earth's crust beneath the GIS due to seasonal loading and unloading of the overlaying melting ice mass. More recently, Zhan (2019) detected slowing down of surface wave velocities up to 2% in the basal till layer of the largest North American glacier (Bering Glacier, 20 km wide) during surge likely due to the switch of the subglacial drainage from channelized to distributed.

The underlying seismic interferometry techniques used in ambient noise studies are rooted in the fact that the elastic impulse response between two receivers, the Green's function (GF), can be ~~reconstructed~~ approximated via cross-correlation of ~~ambient noise seismograms~~ a diffuse wavefield recorded at the two sites (Lobkis and Weaver, 2001; Campillo et al., 2014). Seismic interferometry consists in turning ~~one~~ each of the two receivers ~~R_T~~ into a virtual source and retrieving the estimated elastic

response of the medium at the ~~second~~ other receiver R_2. ~~The retrieved GF is composed of a causal part (being the direct wavefield travelling from the virtual source R_1 to R_2) and an acausal part (being the time-reversed wavefield from the virtual source R_2 to R_1).~~ Under specific assumptions on the source wavefield (see below), the ~~constructed~~ GF estimate is thus expected to be symmetric in its causal and acausal portions (referred to as "causal-acausal symmetry"). ~~Seismic propagation properties can then be estimated by analyzing the reconstructed wavefield which travels between R_1 and R_2.~~

In ~~principle~~ theory, the GF estimate is obtained in media capable of hosting an equipartitioned wavefield, ~~which provides homogeneous (i.e. isotropic) seismic illumination of the medium at the two-receiver site~~ that is random modes of seismic propagation with same amount of energy. In practice, the equipartition argument has limited applicability to the Earth because non-homogeneously distributed sources, in the forms of ambient noise sources, earthquakes and/or scatterers, prevent the ambient wavefield from being equipartitioned across the entire seismic scale (Fichtner et al., 2017, and references therein). The GF estimation from interstation correlation therefore usually relies on simplified approximations of diffusive wavefields which can be reached ~~An equipartioned seismic wavefield can be reached~~ in (i) the presence of equally-distributed sources around the recording network (Wapenaar, 2004; Gouédard et al., 2008b) and/or (ii) in strong-scattering settings as scatterers act like secondary seismic sources and likely create a diffuse homogenized wavefield in all propagation directions (e.g. Hennino et al., 2001; Malcolm et al., 2004; Larose et al., 2008). Even if the noise wavefield is not generally diffuse (Mulargia, 2012), the presence of scatterers in the Earth's crust and the generation of oceanic ambient noise all around Earth make ambient noise interferometry applications generally successful.

In glaciers, the oceanic ambient noise field commonly used in crustal studies lacks the high frequencies needed to generate GFs that contain useful information at the scale of the glacier. To target shallower glaciers and their bed, we must work with other sources such as nearby icequakes and flowing water which excite higher-frequency (> 1 Hz) seismic modes (Sect. 2.1). In this context, ~~In the ice,~~ the lack of seismic scattering in homogeneous ice (Podolskiy and Walter, 2016) renders challenging the reconstruction of the GF from on-ice recordings. ~~Indeed, typical glacier seismic events lack sustained scattered coda waves (Walter et al, 2015) which normally arise from seismic wave deviations on multiscale heterogeneities in the medium (see Fig. 1a and further details in Sect. 5.1).~~ Condition ~~(ii)~~ (i) can compensate for lack of condition ~~(i)~~ (ii). However, microseismicity generated on glaciers is often confined to narrow regions such as crevasse margin-zones (Roux et al., 2008; Mikesell et al., 2012) or other water-filled englacial conduits (Röösli et al., 2014; Walter et al., 2015; Preiswerk and Walter, 2018; Lindner et al., 2020). This often prevents ~~for~~ the occurrence of homogeneous source distributions on glaciers. Nevertheless, the abundance of local seismicity ~~in glaciers such as near-surface crevasse icequakes and high-frequency ambient noise from glacier melt (see Sect. 2.1)~~ indicates a considerable potential for glacier imaging and monitoring with interferometry.

Few attempts have been conducted on glaciers to obtain GF estimates from on-ice seismic recordings. Zhan et al. (2013) first calculated ambient noise cross-correlations on the Amery ice-shelf (Antarctica) but could not compute accurate GF at frequencies below 5 Hz due to the low-velocity water layer below the floating ice-shelf, which causes resonance effects and a significantly nondiffusive inhomogeneous noise field. Preiswerk and Walter (2018) successfully retrieved accurate GF on two Alpine glaciers from the cross-correlation of high-frequency (≥ 2 Hz) ambient noise seismograms, generated by meltwater flow. However, due to localized noise sources in the drainage system that also change positions over time over the course of

the melting season, they could not systematically obtain accurate coherent GF when computed for different times, limiting the applications for glacier monitoring.

As an alternative to continuous ambient noise, Walter et al. (2015) used crevassing icequakes recorded during a one-month seismic deployment at Gornergletscher (Switzerland). They recorded thousands of point source events which offered an idealized spatial source distribution around one pair of seismic sensors and could obtain accurate GF estimates. To overcome the situation of a skewed illumination pattern often arising from icequake locations, Lindner et al. (2018) used ~~virtual reflector seismology~~ multi-dimensional deconvolution (Wapenaar et al., 2011; Weemstra et al., 2017) that relies on a contour of receivers enclosing the region of interest (see also Sect. 6.2). This technique proved to be efficient to suppress spurious arrivals in the cross-correlation function which emerge in the presence of heterogeneously distributed sources. However, this method was applied on active sources and synthetic seismograms and its viability still needs to be addressed for passive recordings.

In this study, we provide a catalogue of methods to tackle the challenge of applying passive seismic interferometry on glaciers in the absence of significant scattering and/or ~~even~~ an isotropic source distribution. After a review on glacier seismic sources (Sect. 2.1), we investigate the GF retrieval on three glacier settings with different patterns of seismic wavefields. In a first ideal case (case A, Glacier d'Argentière in the French Alps, Sect. 3), we take advantage of a favorable distribution of noise sources and icequakes recorded on a dense array. In case B (GIS, Sect. 4), a dominant persistent noise source constituted by a moulin prevents the accurate estimate of the GF across the array. We use a recently proposed scheme (Corciulo et al., 2012; Seydoux et al., 2017) that involves matched-field processing to remove the moulin signature and improve the GF estimate. In case C (Gornergletscher in Swiss Alps, Sect. 5), the limited distribution of icequakes is overcome by the use of crevasse-generated scattered coda waves to obtain homogenized diffuse wavefields before conducting cross-correlations. In order to serve as a practical scheme for future studies, the three above sections are nearly independent from each other. They focus on the processing schemes to compute or improve the GF estimate. We refer the reader who is not familiar with more ambient noise seismic processing to the appendix sections providing details on seismic detection methods, seismic array processing and seismic velocity measurements. Finally, in the light of our analysis, we discuss suitable array geometries and measurement types for future applications of passive seismic imaging and monitoring studies on glaciers.

2 Material and data

2.1 Glacier seismic sources

Glaciogenic seismic waves couple with the bulk Earth, and can be recorded by seismometers deployed at local (Podolskiy and Walter, 2016) to global ranges (Ekström et al., 2003). In this study, we focus on three classes of local sources ~~waves that whose seismic waves are typically recorded in the ablating and melting environment of active grounded glaciers and generate a mix of high-frequency seismic waves, including surface waves which are of interest for the analysis which follows~~. For an exhaustive inventory of glacier seismicity and associated source mechanisms, we refer to the review papers of Podolskiy and Walter (2016) and Aster and Winberry (2017).

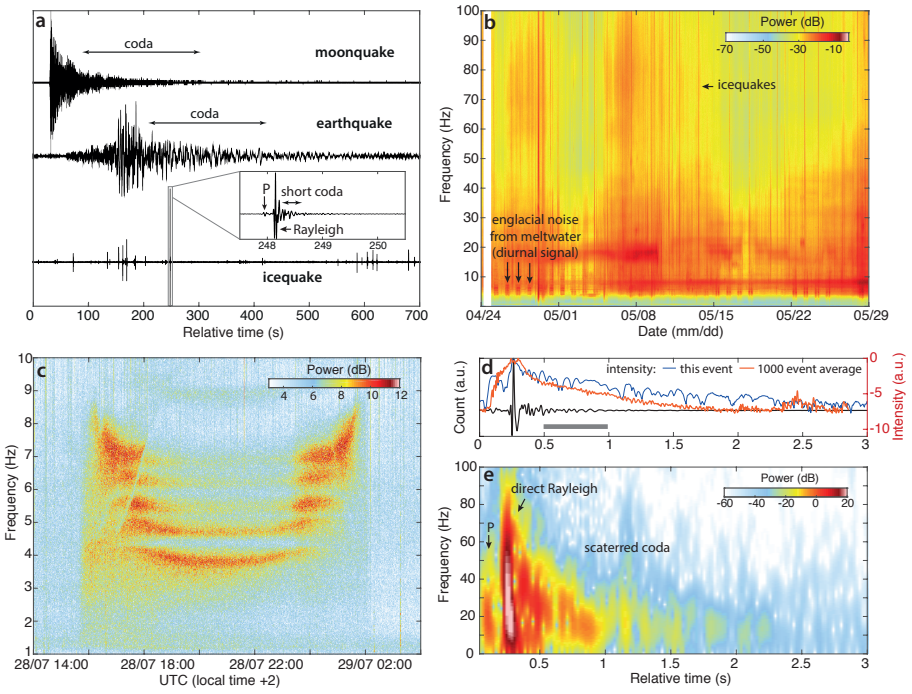

Figure 1. (a) Seismograms of a moonquake, regional earthquake and typical Alpine glacier seismicity. Moonquake seismogram was recorded during the 1969-1977 Apollo passive seismic experiment (Nunn, 2017). Zoom on icequake waveform shows the lack of sustained coda in homogeneous ice when compared to other signals propagating in ~~rocky grounds~~ crustal rocks. (b) Spectrogram of one month of continuous recording at Glacier d'Argentière (French Alps) showing abundance of icequakes (5-100 Hz) and englacial noise (2-30 Hz) produced by turbulent meltwater flow. (c) Spectrogram for a 10 h-long hydraulic tremor produced by the water moulin activity within the Greenland ice-sheet network (Fig.2b). (d) Seismic waveform and associated spectrogram (e) for one icequake recorded at Gornergletscher (Swiss Alps). Color lines in (d) are the signal intensity (see main text, Sect. 5.1) for this event in blue, and averaged over 1000 events in orange (right y-axis, note the logarithmic scale). The horizontal gray bar indicate the coda window which is used to generate first estimations of Green's functions (Sect. 5.2).

Typically on Alpine glaciers and ablation zones, the most abundant class of recorded seismicity is related to brittle ice failure which leads to the formation of near-surface crevasse (e.g. Neave and Savage, 1970; Mikesell et al., 2012; Röösli et al., 2014; Podolskiy et al., 2019) and the generation of $10^2 - 10^3$ daily recorded icequakes (Fig. 1a-b).. ~~On grounded ice, near-surface crevasse formation and expansion in response to glacial stress generate the vast majority of small-scale impulsive seismic events with about $10^2 - 10^3$ daily recorded icequakes (Fig. 1a-b). Daily fluctuations in surface temperature and meltwater production modulate melt-enhanced sliding and seismicity rate (ref) by introducing thermal-stress and strain rate variations (ref) or through hydrofracturing at intermediate depth (ref).Subsurface fractures have a predominantly isotropic (volumetric) source mechanism which describes the crack opening, although a few studies report that a shear failure can also be a significant~~

~~fracture component (references).~~ Near-surface icequakes have local magnitudes -1 to 1 and seismic waves propagate a few hundred meters before falling below the background noise level (Neave and Savage, 1970). Icequake waveforms have durations of 0.1–0.2 s and ~~energy is concentrated in the 5–50 Hz range~~ thus do not carry much energy at frequency below 5 Hz (Fig. 1d-e). With its maximum amplitude on the vertical component, Rayleigh waves dominate the seismogram. In contrast, the prior P-wave arrival is substantially weaker and for distant events often below noise level. Rayleigh waves propagate along the surface and are not excited by a source at depth exceeding one wavelength (Deichmann et al., 2000). In addition, the crevasse zone is mostly confined to the surface (≤ 30 m) since ice-overburden pressure inhibits tensile fracturing at greater depths (Van der Veen, 1998). That is why such icequakes are usually considered to originate at shallow depth (Walter, 2009; Roux et al., 2010; Mikesell et al., 2012). The ~~brief~~ short duration and weak seismic coda after the Rayleigh wave arrival (compared to earthquake coda ~~travelling in rocky ground~~ propagating in the crust, Fig. 1a, see also further details in Sect. 5.1) are the result of limited englacial scattering. This typically allows seismologists to approximate the glacier's seismic velocity model by a homogeneous ice layer on top of a rock half-space when locating events or modelling seismic waveforms (e.g. Walter et al., 2008, 2015).

From spring to the end of summer, another seismic source superimposes on icequake records and takes its origin in fluvial processes. Ice melting and glacier runoff create turbulent water flow at the ice surface that interacts with englacial and subglacial ~~channels and~~ linked conduits. Gravity-driven transport of meltwater creates transient forces on the bulk Earth (e.g. Schmandt et al., 2013; Gimbert et al., 2014) and surrounding ice (Gimbert et al., 2016) that generate a mix of ~~seismic~~ body and surface waves (Lindner et al., 2020; Vore et al., 2019) ~~including Rayleigh surface waves (Lindner et al, 2019) that constitute a large component of the ambient noise recorded in melting glaciers~~. ~~Analytic expressions for expected relative change of seismic power spectral density have been derived for turbulent flow in englacial channels, and allow to monitor subglacial discharge by analyzing the evolution of the channel radius and water pressure gradient from recorded seismic power over the melting season (Gimbert et al, 2016).~~ ~~It~~ Meltwater flow noise is recorded continuously at frequencies 1-20 Hz as shown in the one-month spectrogram of ~~seismic recording~~ ground velocity at Glacier d'Argentière (Fig. 1b). Seismic noise power shows diurnal variations of ambient seismicity correlated with higher discharge during daytime and reduced water pressure at night (Preiswerk and Walter, 2018; Nanni et al., 2019b).

Englacial and subglacial conduits can also generate accoustic (Gräff et al., 2019) and seismic wave resonances (Röösli et al., 2014) known as hydraulic ~~seismic~~ tremors. ~~For example, moulin tremors are produced by surface water routed through channels and moulins to the glacier base and~~ In the presence of moulin, water flowing to the glacier base creates seismic tremors (Fig. 1c) which often dominate the ambient noise during peak melt hours. ~~Spectrograms of moulin tremors exhibit energy between 2 and 10 Hz as shown in Fig. 1b.~~ Frequency bands of either elevated or suppressed seismic energy reflect the geometry of the englacial conduit as they act as a resonanting semi-open pipe, modulated by the moulin water level (Röösli et al., 2016b).

Finally, in Alpine environments, seismic signatures of anthropogenic activity generally overlap with glacier ambient noise at frequencies > 1 Hz. Whereas anthropogenic monochromatic sources can usually be distinguished by their temporal pattern (Preiswerk and Walter, 2018), separation of all active sources recorded on glacier seismograms can prove difficult. Never-

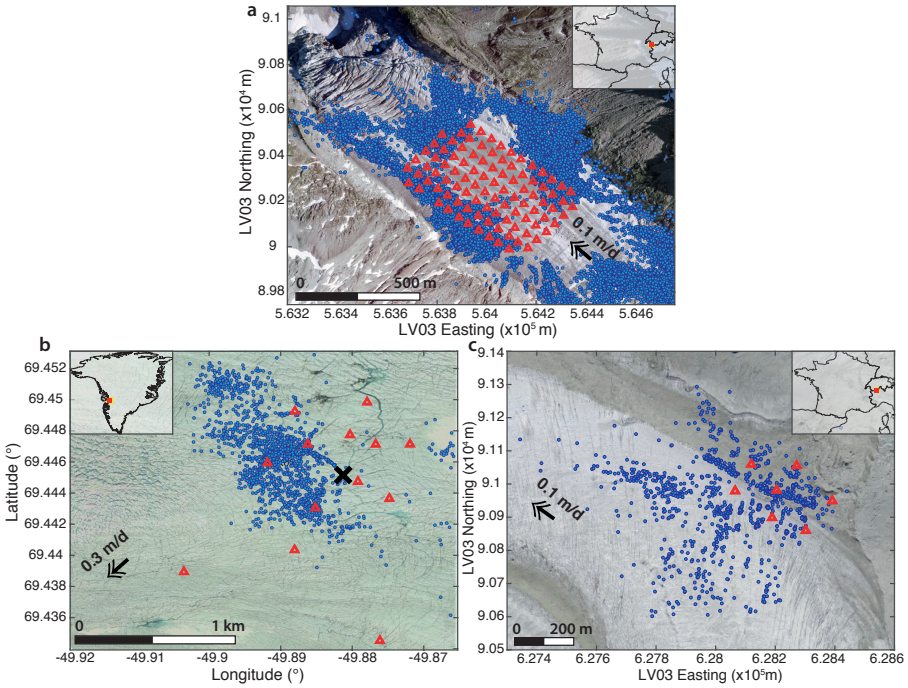

Figure 2. Icequake locations (blue dots) and seismic stations (red triangles) superimposed on aerial photographs of (a) Argentière (photo source: IGN France), (b) Greenland ice-sheet (photo source: Google, Mixar Technologies) and (c) Gornergletscher (Photo source: Federal Office of Topography swisstopo). Black arrows indicate ice flow direction. Black cross in (b) indicates the ~~presence~~ location of the moulin within the array.

theless, locating the source regions through matched-field processing (Corciulo et al., 2012; Chmiel et al., 2015) can help to identify the noise source processes in glaciated environments (Fig. 1c).

2.2 Study sites and seismic experiments

We use ~~vertical component data of ground motion recorded on~~ seismic recordings from three seasonally-deployed networks in the ablation zones of two temperate Alpine glaciers and of the GIS. Each of the acquired datasets presents different patterns of seismic wavefields corresponding to the three configurations investigated for GF estimate retrieval, as defined in the introduction. All networks recorded varying amounts of near-surface icequakes (blue dots in Fig. 2a-c). Different processing schemes were used to constitute the icequake catalogues and are detailed in Appendix A. In this study we only use vertical component data of ground motion to generate ZZ cross-correlation functions which primarily contain the Rayleigh wave fundamental mode (Shapiro and Campillo, 2004). Some of the datasets involve surface seismometers whose horizontal components are regularly shifted over the course of the melting season. Obtaining GF estimates from horizontal component data require additional pre-processing to obtain accurate orientations of the seismic sensors.

2.2.1 Glacier d'Argentière array

The Argentière seismic array (Fig. 2a) was deployed in late April 2018 and was recording for five weeks. It consists of 98 three-component surface sensors regularly spaced on a grid with a 350 m$\times$480 m aperture and a station-to-station spacing of ~ 40 m for the along-flow profiles and ~ 50 m for the across-flow profiles. This large N-array experiment used the technology of nodes (Fairfield Nodal Z-Land 3C) that combine a geophone, digitizer, battery, data storage, and GPS in a single box (Hand, 2014) and allowed a rapid deployment within a few hours. Z-Land ~~nodal sensors~~ geophones have natural frequency of 5 Hz and recorded continuously at a sampling rate of 500 Hz. Besides seismic sensors, four on-ice GPS were deployed. At the array site, the ice is 80-260 m thick (Hantz, 1981) and flows at an approximate rate of 0.1 m/d. The sensors were placed about 30 cm into the snow and accumulated about 4 m of downstream displacement at the end of the experiment. Because of snow melt, we had to level and reorient the instruments twice during the experiment. A Digital Elevation Model (DEM) for the glacier bed was obtained using 14 Ground-Penetrating Radar tracks over the area covered by the seismic array, and a glacier surface DEM was acquired from drone survey.

2.2.2 Greenland ice-sheet array

The GIS network (Fig. 2b) was deployed 30 km North of the calving front of the Jakobshavn Isbræ from 2011 July 2 to August 17. The details of the study site and the seismic network can be found in Röösli et al. (2014); Ryser et al. (2014) and Andrews et al. (2014). We use seismic recordings from 13 stations: 12 seismometers (1 Hz Lennartz) installed on the surface or shallow boreholes (2-3 m deep), and one surface broadband seismometer (Trillium Compact 120 s corner period). Seismometers recorded continuously with a sampling frequency of 500 Hz. The array has a 1.8 km aperture. It is located around a prominent moulin with an average intake of 2.5 m^3/s of meltwater. At the study site, the ice is approximately 600 m-thick and flows at ~ 0.3 m/d (Röösli et al., 2016a).

2.2.3 Gornergletscher array

The Gornergletscher network (Fig. 2c) operated between 2007 May 28 and July 22. It consists of 7 seismometers (six 8 Hz Geospace 11D and one 28 Hz Geospace 20D) installed in shallow boreholes (2-3 m deep). They recorded continuously with a sampling frequency of 1000 Hz. The array has a 320 m aperture. At the study site, the ice is approximately 160 m-thick and flows at ~ 0.1 m/d (Walter, 2009).

3 Passive interferometry at the Glacier d'Argentière dense array

We use here a standardized processing scheme for computing ~~cross-correlation functions~~ GF estimates ~~at the dense array of Glacier d'Argentière~~. We either cross-correlate seismogram time windows which encompass ~~icequake~~ ballistic seismic waves of the icequake catalogue, or cross-correlate continuous seismograms ~~that encompass both ambient noise and icequake signals~~ as traditionally done in ambient noise studies. Prior to any calculation, ~~vertical components of~~ seismic records are corrected

for instrumental response and converted to ground velocity. Seismograms are then spectrally whitened between 1 and 50 Hz because of low instrumental sensitivity at lower frequency.

For icequake cross-correlation (ICC), we follow the method of Gouédard et al. (2008b) and Walter et al. (2015)on $\sim 11 \times 10^3$ events. The correlation time window length T is adjusted to the nature of the seismic source and the array aperture ~~to be able to encompass the icequake Rayleigh wave and~~ to reconstruct propagating waves across the array. Here we use $T = 0.5$ s given the short icequake duration and the maximum station separation of 690 m. To avoid near-field source effects and to account for near-planar wave fronts, we ~~only~~ select events that lie outside a circle centered at the midpoint between the two considered stations and with a radius equal to the inter-station distance (Fig. B4a). ~~For individual ICC (i.e. ICC computed on single events), t~~ The plane wave approximation implies a sinusoidal dependence of ~~surface wave~~ the arrival times with respect to event azimuth (Fig. B4b). ~~In addition, the surface wave arrival time depends on the propagation velocity and the sensor spacing of the cross-correlated pairs. Final ICC are obtained by stacking all individual correlation functions.~~ When stacking the individual ICC on all events, only the sources that lie in the stationary phase zones, i.e. aligned with the receiver direction, actually contribute to the GF (Gouédard et al., 2008b). The aperture of the stationary phase zones, also called "the endfire lobes" (Fig. B4a), depends on the considered seismic wavelength (Roux et al., 2004). In the case of anisotropic source distribution, the contribution of non-stationary sources eventually do not vanish and give rise to spurious arrivals in the final GF estimate. Prior stacking, we assign all cross-correlations to event azimuth bins of 5^o to attribute equal weights to all incident directions. To reduce eventual spurious arrivals, we compute the GF on selected sources in the endfire lobes whose aperture is calculated for maximum wavelengths corresponding to 3 Hz (Appendix B3).

For noise cross-correlation (NCC), we use a similar protocol as the one of Preiswerk and Walter (2018). To reduce the effects of teleseismic events or strongest icequakes, we disregard the seismic amplitudes completely and consider only one-bit normalized seismograms (Bensen et al., 2007). By doing so, we attribute a similar weight to ambient noise and icequake source contributions to the GF. The traces are cross-correlated in nonoverlapping 30-min long windows. Resulting NCCs are stacked daily and then averaged over the five weeks of recording. We finally obtain a set of 4371 NCC that correspond to the GF estimates for all combinations of sensor pairs.

3.1 Green's function estimates

Figure 3a shows the stacked section of NCC averaged in 1 m-binned distance intervals ~~and sorted by increasing inter-station distances~~. Coherent Rayleigh waves with propagation velocity of 1600 m/s are well reconstructed across the array. We also observe emergence of weak but faster waves identified at higher frequencies as P-waves traveling in the ice.

Slight disparities in amplitudes of the causal and ~~anti~~ acausal parts of the GF estimates (positive versus negative times) are related to the noise source density and distribution. ~~The symmetry of NCC stacked in close distance intervals indicates that noise sources are homogeneously distributed around the network, and h~~ Higher acausal amplitudes observed at larger distances are evidence for ~~strongest~~ a higher density of sources located downstream of the array, according to our cross-correlation definition. ~~Strongest~~ More sources downstream ~~could be related to higher discharge near the glacier tongue~~ are likely generated by faster water flow running into subglacial conduits toward the glacier ice fall (Gimbert et al., 2016; Nanni

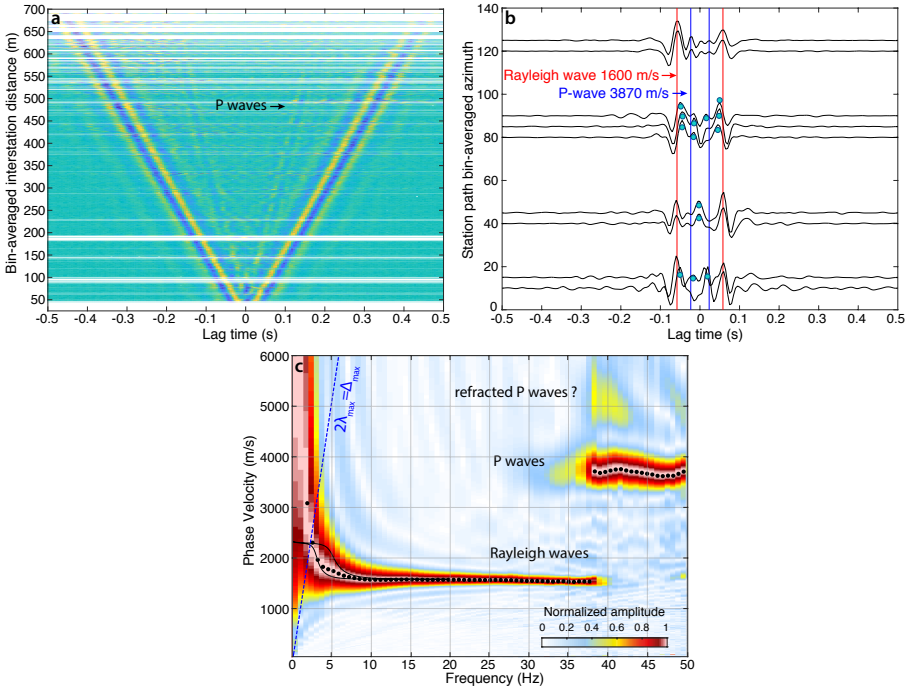

Figure 3. (a) Noise cross-correlations (NCC) sorted by increasing inter-station distances at Glacier d'Argentière. For the representation, correlation functions are averaged in 1 m distance bins and bandpass filtered between 10-50 Hz to highlight the presence of high-frequency P-waves. (b) Azimuthal dependence of GF estimates for pairs of stations 100 m apart. Accurate GF estimates are obtained at station paths roughly aligned with the glacier flow (azimuth $\sim 120^o$), indicating ~~strongest~~ more noise sources likely located downstream and upstream of the array. For other station paths, we observe spurious arrivals (indicated by green dots) before the expected arrival times for Rayleigh (red bars) and P waves (blue bars) which primarily arise from ~~the non-alignment of noise sources in the stationary phase zones of the stations~~ non-distributed noise sources that lie outside the stationary phase zones of these stations (see main text). (c) Frequency-velocity diagram obtained from f-k analysis of NCC in (a). The dispersion curve of phase velocity for Rayleigh waves and P-waves are plotted in black dots. The dashed blue line shows ~~the array response~~ the frequency-dependent resolution limit, given the maximum wavelength and sensor spacing $\lambda_{\max} = \Delta_{\max}/2$. Black lines are theoretical dispersion curves for fundamental mode Rayleigh wave velocity computed for ice thickness either 150 or 250 m with Geopsy software. We used the elastic parameters for the ice and bedrock as given in (Preiswerk and Walter, 2018, Sect. 6.1). Same figure for icequake cross-correlations is available in Appendix (Fig. B2).

et al., 2019b). Looking closer at NCC for individual receiver pairs, we sometimes observe spurious arrivals around time 0 (marked as green dots in Fig. 3b) ~~before the expected Rayleigh wave arrival time~~, mostly at stations pairs which are oriented perpendicular to the glacier flow (i.e. azimuth $0^o \leq \psi \leq 50^o$), indicating that dominant noise sources are located along the flow line. At other station pairs (i.e. azimuth $\psi \sim 90^o$), the reconstructed arrival times are slightly faster than expected. This could be an effect of non-distributed noise sources, and/or anisotropy introduced by englacial features (Sect. 3.3). This analysis shows

that even if the noise sources are not equally distributed in space, averaging the NCC in regular distance intervals on a dense array deployment helps the GF convergence.

The stacked section of ICC (Fig. B2a) yields similar results to those of the NCC (Fig. 3a). ~~Due to the presence of icequake sources in every azimuthal direction, the ICC yield correct GF estimates at most station pairs.~~ The control of the icequake source aperture enables to minimize the spurious arrivals which are observed on some NCC (Fig. 3b) and obtain more accurate Rayleigh wave traveltimes at most station paths (Fig. B2b). The differences in ICC and NCC support that NCC are more sensitive to the ~~distributed~~ noise sources rather than icequake sources. Icequake contributions certainly enable to widen the spectral content of the NCC to frequency higher than 20 Hz, as the most energetic ambient noise is recorded in the 1-20 Hz frequency band (Fig. 1~~a~~ b).

Seismic phases and their velocities can be identified ~~from dispersion curves~~ on the frequency-velocity diagram (Fig. 3c, black dots) that is obtained from frequency-wavenumber (f-k) analysis of the NCC computed on a line of receivers (Appendix B1). ~~Seismic velocities and dispersion curves (black dots) are extracted from the maximum intensity of the f-k diagram at each frequency.~~ As identified above, the correlation functions reconstruct well P-waves traveling in the ice with an average velocity $V_p = 3870$ m/s. We also observe weak intensity but fast seismic phases at frequencies above 35 Hz, which could correspond to refracted P-waves traveling along the basal interface with velocity around 5000 m/s.

Surface waves are dispersive meaning that their velocity is frequency-dependent with higher frequencies being sensitive to surface layers and conversely, lower frequencies being sensitive to basal layers. Theoretical dispersion curves for Rayleigh wave fundamental mode are indicated in black solid lines in Fig. 3~~b~~ c. They correspond to a two-layer model with the top ice layer of thickness $H = 150$ m and $H = 250$ m over a semi-half space representing the bedrock. The dashed blue line indicates the array resolution capability that corresponds to the maximum wavelength limit $\lambda_{\max} = \Delta_{\max}/2$ (Wathelet et al., 2008), with $\Delta_{\max}$ being the maximum sensor spacing. Reconstruction of Rayleigh waves and resolution of their phase velocities using f-k processing are differently sensitive for NCC and ICC at frequencies below 5 Hz (Fig. 3c versus Fig. B2~~b~~ c) as ICC have limited energy at low frequency (Fig. 4a) due to the short and impulsive nature of icequake seismograms (Fig. 1d-e). Given the vertical sensitivity kernels for Rayleigh wave phase velocity (Fig. 4b) and the dispersion curves obtained from the cross-correlation sections (Fig. 3c and B2c), Rayleigh waves that are reconstructed with NCC are capable of sampling basal ice layers and bedrock while ICC are more accurately sensitive to the ice surface. These results reflect the S-wave velocity dependence on depth.

3.2 Dispersion curve inversion and glacier thickness estimation

~~As discussed above, Rayleigh waves are well suited to determine shear-wave velocities and thicknesses of shallow layers as they are primarily sensitive to the S-wave velocity at a depth of approximately $1/3$ of the wavelength (Song et al, 1989).~~ Sensitivity of Rayleigh waves obtained on NCC to frequencies below 5 Hz enable us to explore the subsurface structure of the glacier with inversions of velocity dispersion curves. Due to the general noise source locations up-flow and down-flow of the network, we limit our analysis to receiver pairs whose accurate GF could be obtained. We thus compute the dispersion curves on eight along-flow receiver lines which constitute the array (inset map in Fig. 5). For each line, we invert the 1D ground

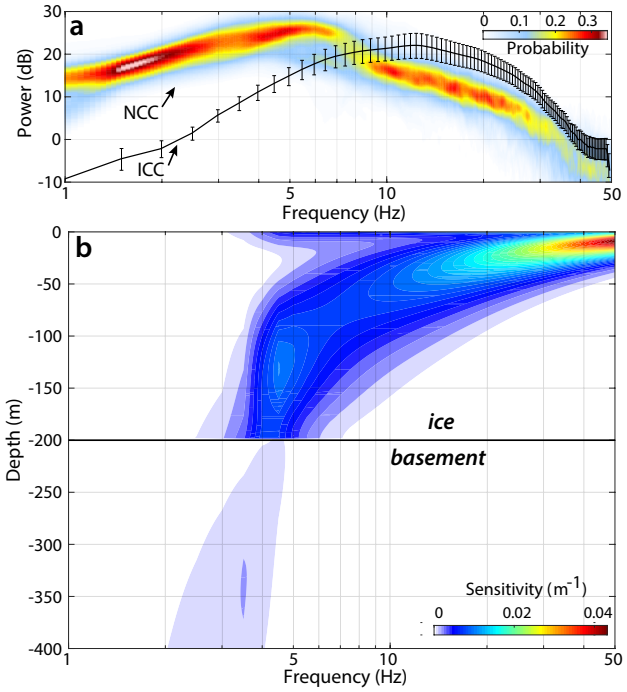

Figure 4. (a) Probability density function of noise cross-correlation (NCC) spectra (colors) and median average of icequake cross-correlation (ICC) spectra (black line). Note that raw data (i.e continuous noise or icequake waveforms) were spectrally whitened between 1-50 Hz prior cross-correlation. Due to spectral content of englacial noise and icequakes, NCC and ICC have different depth sensitivity due to spectral response. (b) Vertical sensitivity kernels for phase velocities of the Rayleigh wave fundamental mode for an ice thickness of 200 m over a semi-half space representing the bedrock. The kernels were computed using the freely-available code of Haney and Tsai (2017).

profile which best matches seismic velocity measurements in the 3-20 Hz frequency range, using the neighbourhood algorithm encoded in the Geopsy software (Wathelet, 2008).

Following Walter et al. (2015), we assume a two-layer medium consisting of ice and underlying granite bedrock. This is a simplified approximation and does not include ~~lateral or longitudinal variations in basal topography along the receiver line, nor allow for the presence of snow and basal till layers as well as other heterogeneities such as cavities, crevasses and other water-filled channels~~ 2D and 3D effects, and anisotropy introduced by englacial features (Sect. 3.3). The grid search boundaries for seismic velocity, ice thickness and density are given in Table 1. We fix the seismic P-velocity in ice to 3870 m/s as measured in Fig. 3~~b~~ c, and couple all varying parameters to the S-wave velocity structure with the imposed condition of increasing velocity with depth.

Figure 5 a-b show the inversion results for the receiver line at the center of the array labelled "4". Velocity measurements are indicated by yellow squares and dispersion curves corresponding to explored velocity models are in colors sorted by misfit values. Misfit values correspond here to the root-mean square error on the dispersion curve residuals, normalized by

Table 1. Parameter ranges and fixed parameters for grid search to invert the dispersion curves in Fig. 5 for ice thickness. Poisson's ratios of ice and granite were varied between 0.2 and 0.5. Poisson's ratio, ice thickness and P-wave velocity V_p were coupled to the S-wave velocity V_s.

Material	Thickness (m)	V_p (m/s)	V_s (m/s)	Density (kg/m^3)
Ice	50-500	3870 (fix)	1500-2100	917 (fixed)
Granite	∞	3870-6000	1500-3500	2750 (fixed)

the uncertainty average we obtained from the seismic data extraction (error bars in Fig. 5a). The inversion well resolves the S-wave velocity in the ice layer as all best matching models yield to $V_s = 1707$ m/s for misfit values below 0.05 meaning that the data dispersion curve is adjusted with an approximate error below 5%. The best fitting model gives a 236 m-thick ice layer and bedrock S-velocity of 2517 m/s. Walter et al. (2015) explored the sensitivity of the basal layer depth to the other model parameters and report a trade-off leading to an increase in inverted ice thickness when increasing both ice and bedrock velocities. Here the ice thickness estimation is most influenced by the rock velocities as we notice that a 100 m/s increase of basal S-velocity results in an increase in ice thickness up to 15 m. These results are moreover influenced by larger errors at lower frequencies (Fig. 5a) which comes from less redundant measurements at large distances. Furthermore 3D effects could lead to some errors in the depth inversion results which need to be further investigated.

From the eight receiver line inversions, we find average S-wave velocities of 1710 m/s for the ice, 2570 m/s for the granite and a P-wave velocity of 4850 m/s in the basement, which is consistent with our measurement for refracted P-waves in Fig. 3c. V_p/V_s ratios are found to be 2.2 and 1.9 for ice and granite, giving Poisson's ratios of 0.37 and 0.3, respectively. For the receiver lines near the array edges (Lines 1-3 and 8), the inversion yields to a ~~top~~ low P-velocity surface layer of thickness 15 m and 7 m respectively, above thicker ice (dashed blue zone in Fig. 5c). ~~where P-wave velocity is slightly decreased while S-wave velocity is maintained at ~ 1710 m/s as for the ice.~~ In this thin top layer, the matching S-velocity corresponds to the one for the ice (i.e. 1710 m/s). The V_p/V_s ratio is around 1.6 and corresponds to a Poisson's ratio of 0.2. This is what is expected for snow although only a ~ 5 m snow cover was present in the area at the time of the experiment. This low-velocity surface layer could also at least partially be attributed to the presence of pronounced transversal crevasses (i.e. perpendicular to the receive lines, see Sect. 3.3) near the array edges, which do not extend deeper than a few dozens of meters (Van der Veen, 1998) and can be modelled as a slow layer above faster ice (Lindner et al., 2019).

Inversion results for the ice thickness are plotted in red in Fig. 5c. Associated uncertainties (red error bars) are given by the models which fit the dispersion curves with misfit values below one standard deviation of the 2500 best-fitting models. The errors on the basal interface depth generally correspond to a maximum misfit value of 0.02. The black solid line shows the across-flow profile of the glacier baseline, that was extracted from the DEM of Glacier d'Argentière (Sect. 2.2.1) and averaged over the nodal sensor positions. Results show that ambient noise interferometry determines the depth of the basal interface with a 10 m resolution as we are able to reproduce the transverse variations of the ice thickness. Differences in ice thickness

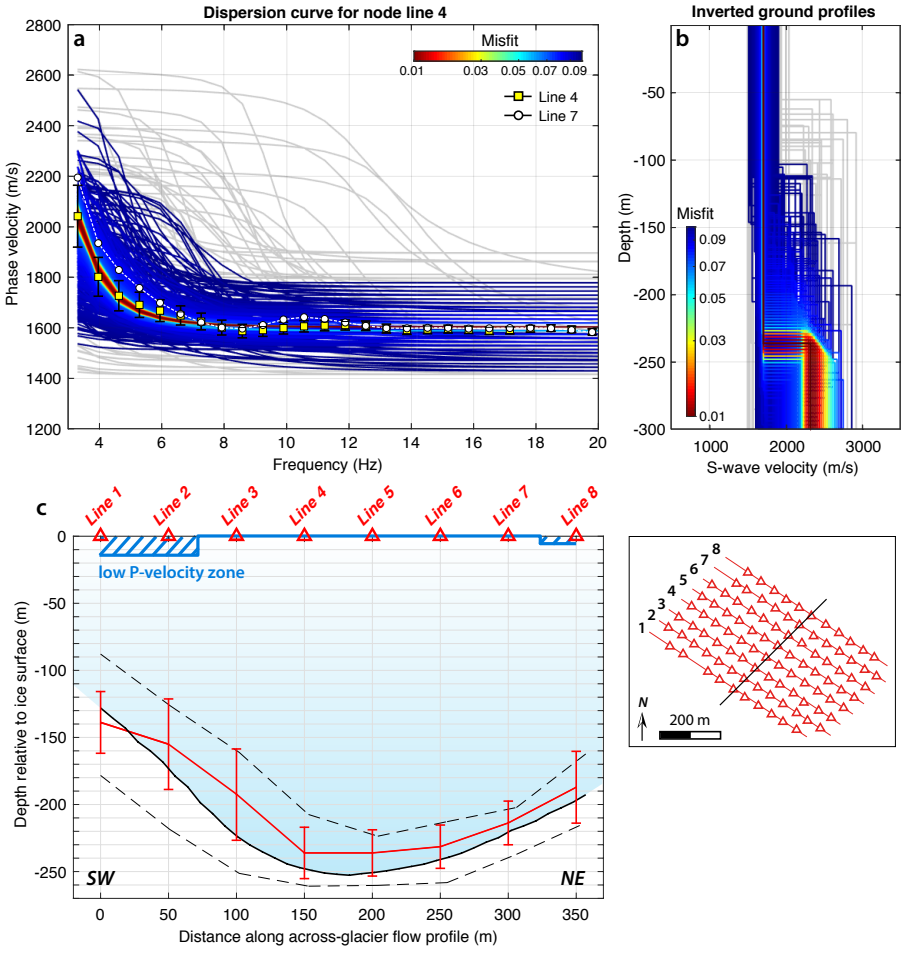

Figure 5. Inversion of glacier thickness using velocity dispersion curves of Rayleigh waves uand the Geopsy neighborhood algorithm. Dispersion curve measurements are obtained from f-k analysis of noise cross-correlations on eight receiver lines whose geometry is described in the bottom-right panel. (a-b) Colour-coded population of (a) dispersion curve fits and (b) S-wave velocity profiles for the node along-flow line labelled 4 in (c). Warmer colours correspond to smaller misfit, gray lines correspond to models with misfit values higher than 0.1. In (a) the dispersion curve and uncertainties obtained from seismic measurements are overlaid in yellow squares. For comparison, the dispersion curve computed for the node line 7 and associated to a thinner ice layer is plotted in white dots. (c) Across-flow profile of (red line) ice thickness estimates from Rayleigh wave velocities obtained at eight along-flow node lines and (black line) average basal topography from a DEM. Dashed-blue zones indicate the presence of a low P-velocity top layer from seismic inversions. Uncertainties in ice thickness estimates (red error bars) correspond to seismic inversion results which yield to a misfit lower than one standard deviation of ~~all~~ misfit values from the 2500 best-fitting models. Black dashed lines indicate deviations to the glacier baseline around each node line due to longitudinal topography gradients.

~~absolute~~ values between our measurements and the DEM are generally less than 20 m (being the DEM resolution), and the maximum error is 35 m for Line 3.

Errors and uncertainties on mapping the basal interface are primarily linked to bedrock velocities, as discussed above. Potentially, the bed properties can be refined using additional measurements from refracted P-waves that should be reconstructed on NCC obtained on such a dense and large array and stacked over longer times. Ice thickness estimation is also affected by ~~any lateral and longitudinal variations in the subsurface~~ 2D and 3D effects as phase velocities are here averaged over multiple receiver pairs. The confidence interval we obtain for basal depth is similar order to the actual variations in glacier thickness along the receiver lines ~~due to longitudinal basal topography gradients~~ (black dashed lines). More accurate 3D seismic models of the glacier subsurface could be obtained using additional station pairs as discussed in section 6.

3.3 Azimuthal anisotropy from average phase velocities

Smith and Dahlen (1973) show that for a slightly anisotropic medium the velocity of surface waves varies in 2ϕ-azimuthal dependence according to

$$c(\phi) = c_0 + A\cos\left[2(\phi - \psi)\right] \tag{1}$$

where c_0 is the isotropic component of the phase velocity, A is the amplitude of anisotropy and ψ defines the orientation of the anisotropic fast-axis. ~~Several studies have been able to measure azimuthally-dependent phase speed anisotropy from ambient noise correlations computed on dense arrays, and relate them to the presence of geological features and cracks in the Earth's crust or temperature variations in the upper mantle (references).~~ On glaciers, azimuthal anisotropy can be induced by englacial crevasses, with fast direction for Rayleigh wave propagation being expected to be oriented parallel to the crack alignment (Lindner et al., 2019). ~~Preferred orientations of ice crystals can also introduce anisotropy whose axis of symmetry depend on the type of ice fabrics Anisotropy in glaciers and ice-sheets can also result from preferred orientations of ice crystals. There exists different ice fabrics and each one represents a transversely isotropic medium with different types of symmetry (Diez et al, 2015; Maurel et al, 2015). In particular, girdle fabrics can have horizontal axes of symmetry and also induce azimuthal anisotropy as observed by (Smith et al, 2017) in polar ice.~~ Glacier and ice-sheets are also represented as transversely isotropic media whose type of symmetry depends on the ice fabric (e.g. Diez and Eisen, 2015; Horgan et al., 2011; Smith et al., 2017; Picotti et al., 2015).

The dense array experiment of Glacier d'Argentière covers a wide range of azimuths ϕ defined by the orientation of the station pairs, which allows us to investigate azimuthal variation in Rayleigh wave velocities at any given sensor. In order to cover a maximum range of ϕ-azimuth, we compute velocity dispersion curves for Rayleigh waves obtained on the correlation functions computed on icequake signals (ICC), since ~~correct~~ accurate GF estimates from ambient noise are limited to station pair directions with azimuth ϕ roughly aligned with ice flow ($\phi \sim 120^o$, Sect. 3.1). To measure phase velocities, we use a slant-stack technique similar to Walter et al. (2015), to octave-wide frequency ranges by bandpass filtering the individual ICC, at each station pair (Appendix B3).

For each sensor position, we obtain c velocity measurements as a function of the ϕ-azimuth of the receiver pair that includes the target station. To reduce the effect of spatial averaging, we compute anisotropy parameters ψ and A considering subarrays of stations that lie within 250 m of the target point (inset map in Fig. 6a). ψ and A values are found at each station cell by fitting $c(\phi)$ with equation 1 using a Monte-Carlo inversion scheme. Note that the ~~derivation~~ formulation of Equation 1 also gives rise to an additional 4ϕ dependence of velocities. Lindner et al. (2019) used a beamforming approach on icequake records at 100 m-aperture arrays and found that adding the 4ϕ component to describe the azimuthal variations of phase velocities induced by glacier crevasses yields similar ψ and A. We therefore neglect the 4ϕ term in the present analysis. ~~Finally, to reduce the effect of spatial averaging, we compute anisotropy parameters considering subarrays of stations that lie within 250 m of the target point (inset map in Fig. 6a).~~

Anisotropy is observed to be more pronounced near the glacier margins (lines 1-2 and 8 as labelled in Fig. 5c), where the anisotropy strength varies between 2% and 8% (Fig. 6b). There, fast-axis directions of Rayleigh wave propagation coincide with the observed surface strike of the ice-marginal narrow crevasses that are also responsible for the generation of icequakes indicated by red dots. At other locations, fast-axis directions indicate the presence of transversal crevasses (i.e. perpendicular to the ice flow) with weaker degrees of anisotropy up to 4%. While the near-surface crevasses observed at the array edges result from shear stress from the margin of the glacier, the transversal crevasses are formed by longitudinal compressing stress from lateral extension of the ice away from the valley side walls, which is typical for glacier flow dynamics in ablating areas (Nye, 1952).

Alignment of the fast-axis directions with that of ice flow appears along the central lines of the glacier (receiver lines 4-5) with anisotropy degrees of 0.5% to 1.5%. This feature is only observed along the deepest part of the glacier where it flows over a basal depression. Results are here computed for seismic measurements at 25 Hz and maps of anisotropy do not change significantly with frequency over the 15-30 Hz range. If we extend our analysis down to 7 Hz, we notice that the aligned-flow fast-axis pattern starts to become visible at 10 Hz. At frequencies lower than 10 Hz, the fast-axis generally tend to align perpendicular to the glacier flow because lateral topographic gradients introduce 3D effects and non-physical anisotropy. The results presented here are not punctual measurements but are rather averaged over the entire ice column. The vertical sensitivity kernels for Rayleigh waves (Fig. 4b) are not zero in the basal ice layers at the considered frequencies. The align-flow anisotropic pattern is likely attributed to a thin water-filled conduit as also suggested by locations of seismic hydraulic tremors at the study site (Nanni et al., 2019a). ~~Given sensitivity kernels for Rayleigh waves (Fig. 4b), frequencies above 15 Hz are most sensitive to the top 100 m layer. However we do not exclude the possibility that such seismic measurements could be averaged over the entire ice column due to the large spectral band used for velocity estimates. Indeed, we extend our analysis down to 7 Hz and notice that the aligned-flow fast-axis pattern starts to become visible at 10 Hz. The octave-wide frequency range used for computing the velocity at 10 Hz gives corner frequencies of 6-13 Hz for ICC filtering and then include Rayleigh waves that are sensitive to basal layers.~~

Generally, we observe an increase in the degree of anisotropy with frequency, which is evident for a shallow anisotropic layer. Conversely increase in anisotropy strength at lower frequency would indicate deeper anisotropic layer. At the Alpine plateau Glacier de la Plaine Morte, Lindner et al. (2019) find azimuthal anisotropy at frequencies 15-30 Hz with strength up to

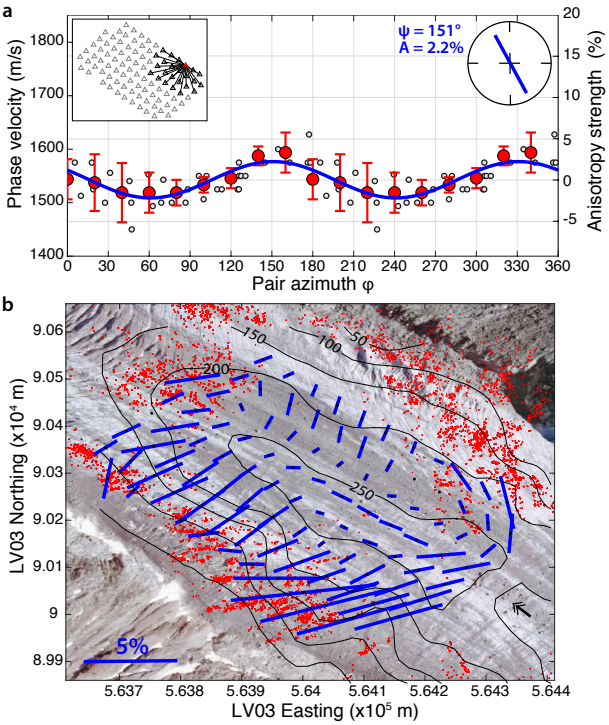

Figure 6. (a) Azimuthal variation of phase velocities measured at one node (red triangle in the inset map). White dots are the phase velocity measurements obtained for different azimuths ϕ that are defined by the station pair orientation. To avoid spatial averaging, we only consider subarrays of 250 m aperture around the target node at the center as described by black triangles in the inset map. Red dots are phase velocities averaged in 20° azimuth bins. The thick blue curve is the best fit for the 2ϕ azimuthal variation of the averaged velocity measurements in red. Fast-axis angle and anisotropy strength are indicated in the top right corner circle. (b) Map of fast-axis direction and amplitude of anisotropy measured at 25 Hz. Locations of icequakes active for 7 days are plotted in red dots to highlight the orientations of surface crevasses. Basal topography contour lines are indicated every 50 m. Black arrow indicates ice flow direction. Photo source: Federal Office of Topography swisstopo.

8%. They also find that constraining the depth of the anisotropy layer is not straightforward as there exists a trade-off between its thickness and the degree of anisotropy. Without any further modelling effort, we refrain from further interpreting our results in terms of crevasse extent and depth of the anisotropic layer or any other cause for the observed anisotropic patterns.

4 Matched-field processing of englacial ambient seismic noise ~~using the GIS array~~

As pointed out earlier, localized englacial noise sources related to water drainage can prevent the reconstruction of stable GF estimates. In addition, noise sources change over time preventing stable GF estimate reconstruction when averaged over short-time scale (Preiswerk and Walter, 2018). These kind of strong, localized seismic noise sources within a seismic array create

spurious arrivals in the NCC, preventing GF estimate convergence (i.e. Walter, 2009; Zhan et al., 2013; Preiswerk and Walter, 2018). In this case, the workflow processing traditionally used in NCC procedure as presented in Bensen et al. (2007) and Sect. 3 is not sufficient. Accordingly, we need to apply more advanced processing methods that can reduce the influence of localized sources, and enhance a more ~~even~~ isotropic distribution of the noise sources around receiver pairs.

One of the approaches we apply here is Matched-Field Processing (MFP) (Kuperman and Turek, 1997) which is an array processing technique allowing to locate low-amplitude sources. MFP is similar to a traditional beamforming that is based on phase-delay measurements. MFP was used for location and separation of different noise sources in various applications, i.e., to monitor geyser activity (Cros et al., 2011; Vandemeulebrouck et al., 2013), in an exploration context (Chmiel et al., 2016), in geothermal field (Wang et al., 2012) and fault zone (Gradon et al., 2019) event detection. MFP was also used by Walter et al. (2015) to measure phase velocities of moulin tremor signals on the GIS.

Moreover, joint use of MFP and the Singular Value Decomposition (SVD) of the cross-spectral density matrix allows separation of different noise source contributions, as in Multi-Rate Adaptive Beamforming (MRABF: Cox (2000)). The SVD approach was explored by i.e. Corciulo et al. (2012) to locate weak amplitude subsurface sources, and Chmiel et al. (2015) for microseismic data denoising. Also, Seydoux et al. (2017) and Moreau et al. (2017) showed that the SVD-based approach improves the convergence of NCC towards ~~empirical GF~~ the GF estimate. Here, we combine MFP and SVD in order to remove spurious arrivals in NCC caused by the moulin located within the GIS array, and thus improve the GF estimate emergence.

4.1 Location of ~~dominant~~ noise sources at the GIS via matched-field processing

Röösli et al. (2014) and Walter et al. (2015) documented the presence of hour-long tremor signals in GIS seismic records, typically starting in the afternoon hours. These events occurred on 29 days out of the 45-day total monitoring period. Signal intensity and duration depended on the days of observations, and the energy was mostly concentrated in the 2-10 Hz range within distinct frequency bands (Fig. 1b). Röösli et al. (2014) and Röösli et al. (2016b) showed a clear correlation between water level in the moulin, and start and end times of the tremor, therefore the tremor signal is referred to as "moulin tremor". Figure 1c shows a spectrogram of a moulin tremor lasting for 10 hours on the night of 28th/29th July 2011 and recorded at one station located 600 m away from the moulin. This signal is generated by the water resonance in the moulin, is coherent over the entire array, and dominates the ambient noise wavefield during peak melt hours (Röösli et al., 2014; Walter et al., 2015).

We briefly summarise the basics of MFP and the details of the method can be found in Cros et al. (2011); Walter et al. (2015) and Chmiel et al. (2016). MFP exploits the phase coherence of seismic signals recorded across an array. It is based on the match between the Cross-Spectral Density Matrix (CSDM) and a modelled GF. The CSDM captures the relative phase difference between the sensors, as it is the frequency-domain equivalent of the time-domain correlation of the recorded data. The CSDM is a square matrix with a size equivalent to the number ($N = 13$ for our GIS array) of stations (N-by-N matrix). MFP is a forward propagation process. It places a "trial source" at each point of a search grid, computing the model-based GF on the receiver array, and then calculating the phase match between the frequency-domain-modeled GF and the Fourier transform of time-windowed data. The optimal noise source location is revealed by the grid points with maximum signal coherence across the array.

In order to calculate the MFP output we use 24 h data of continuous recordings on the 27$^{\text{th}}$ of July which encompass the moulin tremor. We calculate a daily estimate of the CSDM by using 5 min-long time segments in the frequency band ~~of~~ between 2.5 and 6 Hz, which gives in total M = 288 of segments for one day. This ensures a robust, full-rank estimation of the CSDM ($M >> N$). The modelled GFs are computed over the two horizontal spatial components (Easting and Northing) using a previously optimized Rayleigh wave velocity of c = 1680 m/s corresponding to the propagation of Rayleigh waves within the array obtained by Walter et al. (2015). The MFP output is averaged over 30 discrete frequencies in the 2.5-6 Hz range.

The lower frequency bound (2.5 Hz) ensures a higher rank regime of the seismic wavefield, as defined in Seydoux et al. (2017). It means that the degree of freedom of the seismic wavefield is higher than the number of stations. The degree of freedom of the seismic wavefield is defined as a number of independent parameters that can be used to describe the wavefield in the chosen basis of functions. This number depends on the analyzed frequency, slowness of the medium (inverse of velocity), and the average interstation spacing of the array (here 736 m). The higher frequency bound (6 Hz) ensures no spatial aliasing in the beamformer output, given the minimum sensor spacing of 156 m.

Figure 7a shows the grid search for MFP performed over Easting and Northing positions. In order to reveal the location of the source, we use the ~~conventional~~ Bartlett processor (Baggeroer et al., 1993) to measure the match between the recorded and modelled wavefield. The MFP output reveals two dominant noise sources: a well constrained focal spot corresponding to the moulin position inside the GrIS array and another source located North of the array. The latter source is revealed by a hyperbolic shape. This shape is related to a poor radial resolution of the beamformer for sources located outside of an array. Walter et al. (2015) suggested that this dominant source might correspond to another moulin as satellite imagery shows the presence of several drainage features North of the array. Both noise source signals contribute to the NCC. However, while the source located outside of the array contributes to the stationary-phase zone ("end-fire lobes") of certain receiver pairs, the moulin located within the array will mostly cause spurious arrivals on NCC. In order to separate the contribution of these noise sources we first perform SVD of the CSDM, and then use a selection of eigen~~components~~ vectors and eigenspectral equalization (Seydoux et al., 2017) to improve the convergence of NCC towards ~~empirical~~ an estimate of the GF.

4.2 Green's function ~~emergence~~ estimate from eigenspectral equalization

SVD is a decomposition of the CSDM that projects the maximum signal energy into ~~as few coefficients as possible~~ independent coefficients (i.e. Moonen et al., 1992; Konda and Nakamura, 2009; Sadek, 2012). It allows to split the recorded wavefield into a set of linearly independent eigencomponents, each of them corresponding to the principal direction of incoming coherent energy and bearing their own seismic energy contribution:

$$K = USV^T = \sum_{i=1}^{N} K_i = \sum_{i=1}^{N} U_i \Gamma_i V_i^T \tag{2}$$

where, K is the CSDM, N is the number of receivers, U and V are ~~orthogonal~~ unitary matrices containing the eigenvectors, and S is a diagonal matrix representing the eigenvalues Γ, and ^{T} denotes the transpose of the matrix. The total number of eigenvalues corresponds to the number N of receivers. The CSDM can be represented as the arithmetic mean of individual CSDMs (K_i), where each K_i is a CSDM corresponding to a given singular value Γ_i.

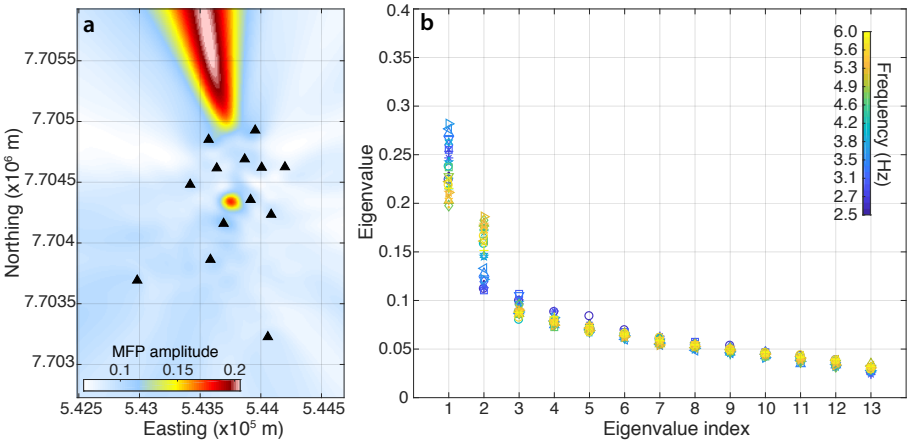

Figure 7. (a) Location of the dominant noise sources using MFP in the frequency band ~~of~~ between 2.5 Hz and 6 Hz (the MFP output is averaged over 30 discrete frequencies). The MFP was calculated using daily data recorded on the 27th of July at the 13 presented stations (black triangles).(b) Eigenvalue distribution of the CSDM for 30 discrete frequencies in the analyzed frequency band.

The SVD separates the wavefield into dominant (coherent) and subdominant (incoherent) subspaces. It has been shown that the incoherent sources correspond to the smallest ~~singular~~ eigenvalues (Bienvenu and Kopp, 1980; Wax and Kailath, 1985; Gerstoft et al., 2012; Seydoux et al., 2016, 2017). Therefore, a common noise removal method consists of setting a threshold that distinguishes between coherent signal and noise, and to keep only the index of eigenvectors that are above the threshold before reconstructing the CSDM (Moreau et al., 2017). The CSDM reconstruction consists of eigenspectral normalization (as explained in the following) and summing a selection of individual CSDMs (K_i). The "denoised" NCC in the time domain are obtained with the inverse Fourier transform of the reconstructed CSDM.

Here, we follow the approach of Seydoux et al. (2016) for choosing the threshold. In the 2.5-6 Hz frequency band, the wavefield is undersampled by the seismic array (which means that the typical radius of GrIS seismic array is ~~smaller~~ larger than half a wavelength of the analyzed Rayleigh waves). Seydoux et al. (2016) showed that in this case, the eigenvalue index cut-off threshold should be set to $N/2$ in order to maximize the reconstruction of the CSDM. This means that we reject the last eigenvectors (from 7^{th} to 13^{th}) as they do not contain coherent phase information.

Figure 7b shows the eigenvalue distribution for 30 discrete frequencies in the analyzed frequency band. The first two eigenvalues correspond to the two dominant noise sources visible on Fig. 7b, and show larger value variation with frequency in comparison with the rest of the distribution. This might be related to the change in the distribution of the dominant sources depending on the frequency related to the seismic signature of the hydraulic tremor and the distinctive frequency bands generated by the moulin activity (Fig. 1c). Moreover, the ~~SVD~~ eigenvalue distribution decays steadily, and does not vanish with high eigenvalue indexes. ~~This confirms that the number of the degrees of freedom of the ambient wavefield is higher than the number of receivers.~~ The latter confirms that the wavefield is undersampled by the seismic array (see Seydoux et al. (2017) for details).

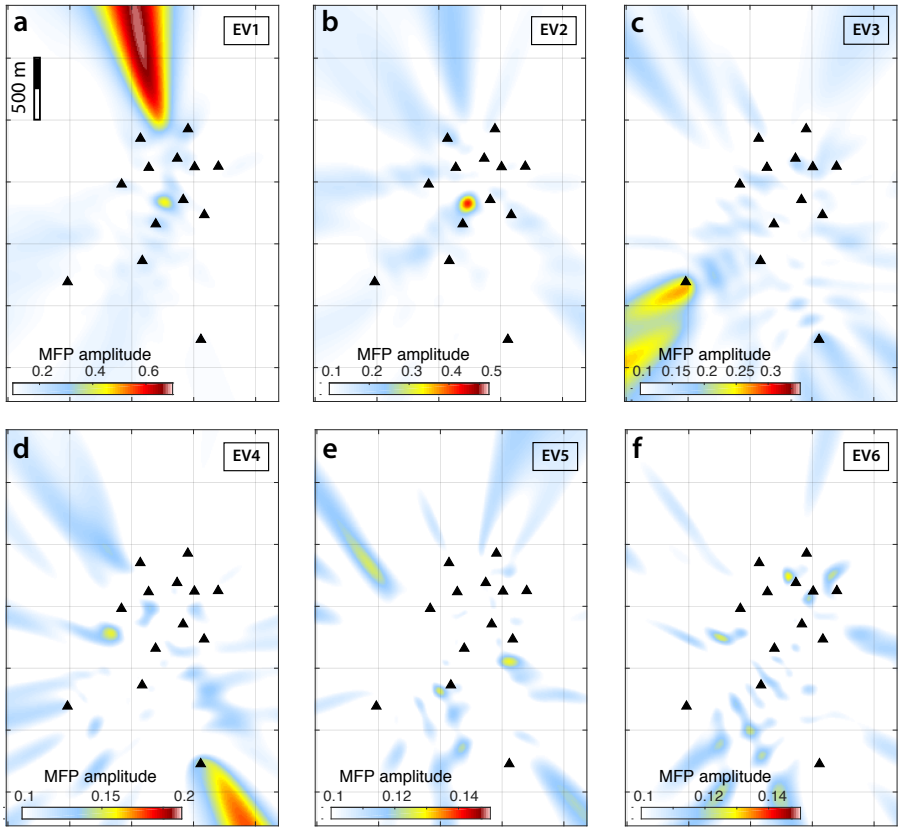

Figure 8. Reconstruction of the CSDM by using individual eigenvectors (EV) that are related to different noise sources. Each plot shows MFP output computed using the reconstructed CSDM with individual eigenvectors as in equation 3. Each figure represents the MFP gridsearch output calculated for the first eigenvectors: (a) 1, (b) 2, (c) 3, (d) 4, (e) 5, and (f) 6. The spatial coordinates are the same as on Fig. 7a.

The CSDM can be then reconstructed by using only individual eigenvectors as in:

$$\widetilde{K_i} = U_i V_i^T \tag{3}$$

Note that we do not include the eigenvalues Γ in the CSDM reconstruction, which is equivalent to equalizing them to 1. That is why we refer to the reconstructed CSDM as "equalized" (Seydoux et al., 2016).

Figure 8 shows the equalized and reconstructed CSDM by using individual eigenvectors that correspond to the principal directions of incoming coherent energy and are related to different noise sources. Each plot represents MFP grid-search output computed on reconstructed CSDM. Figure 8b shows that the second eigenvector corresponds to the moulin source located inside the array. However, we note that the first eigen~~component~~ vector also reveals a weaker focal spot corresponding to the moulin location. This indicates, similarly to the hydraulic tremor spectrum (Fig. 1c) and the singular value distribution, that the spatial distribution of dominant noise sources varies ~~with frequency~~ within the analyzed frequency band. Furthermore, higher

485 eigenvectors do not reveal any strong noise sources localized within the array, and their MFP output points towards sources located outside of the array.

This MFP-based analysis of spatial noise source distribution allows us to select the eigenvectors of CSDM that contribute to noise sources located in the stationary phase zone (i.e. in the endfire lobes of each station path). We now reconstruct the NCC in the frequency band of 2.5-6 Hz with a step equivalent to the frequency sampling divided by the number of samples
490 in the time window (here 0.0981Hz, so 1019 individual frequencies in total). We perform the inverse Fourier transform of the equalized CSDM reconstructed using the 1^{st}, 3^{rd}, 4^{th}, 5^{th}, and 6^{th} eigenvectors. ~~Note that we do not include the eigenvalues Γ in the CSDM reconstruction, which is equivalent to equalizing them to 1. This is why we refer to the reconstructed CSDM as "equalized".~~

Figures 9a and 9b compare the NCC before (in a) and after (in b) the eigenvector selection and eigenspectrum equalization
495 procedure. The displayed NCC are bin-averaged in fixed distance intervals (every 100 m) in order to improve the signal-to-noise ratio (SNR). The blue line shows the propagation of the Rayleigh waves with the velocity of 1680 m/s. On Figure 9a we observe spurious arrivals (marked with green dots) that dominates the NCC together with a non-symmetrical shape. On average, the CSDM equalization process (Fig. 9b) enhances the symmetry of NCC by 40%. To quantify the symmetry of NCC we used the correlation asymmetry as proposed in Ermert et al. (2015), equation 11.

500 Unfortunately, we notice that the equalization process reduces the overall SNR ratio and does not eliminate all spurious arrivals. This might be related to the imperfect separation of different noise sources. For example, we still keep some contribution of the central moulin in the 1^{st} eigen~~component~~ vector. Moreover, by removing the second eigen~~value~~ vector we remove not only the seismic signature of the moulin, but also contribution of coherent far-field sources.

To further assess the isotropy of the reconstructed noise field, we use the conventional plane wave beamformer (e.g. Veen and
505 Buckley, 1988). The plane wave beamforming technique estimates the ~~isotropie~~ isotropy and coherence of the ambient seismic noise wavefield with respect to the slowness and back-azimuth. For the plane wave beamforming calculation, we use the original (Fig. 9c) and the previously equalized CSDM (Fig. 9d). Figure 9c, d shows the beamforming output before (in Fig. 9c) and after the selection and equalization of eigen~~components~~ vectors (in Fig. 9d). The ~~wavenumber vectors~~ wavenumbers k_x and k_y are normalized by the wavenumber k_0 corresponding to Rayleigh wave slowness of $s = 1/1680$ s.m^{-1}. A perfectly isotropic
510 noise wavefield consisting of Rayleigh waves would locate energy near the slowness circle of radius 1. After the removal of the second eigenvalue and the equalization of the strongest eigenvectors, we observe a more isotropic wavefield, meaning other noise sources are enhanced. This quasi-circular shape reflects the energy that arrives from different azimuths. The difference in beamformer amplitude can be caused by the non-regular shape of the GrIS array and different energy contributions of the noise sources. The results show not only the strong source of noise coming from the North, but also energy incident from the
515 South-West that might be related to oceanic ambient noise in the Labrador Sea (Sergeant et al., 2013) or other continuous noise generated by calving and ice-mélange dynamics in the proglacial fjord of nearly Jakobshavn Isbræ (Amundson et al., 2010), one of Greenland's largest ice-streams. Finally, it seems that not much seismic energy is incident from the inland of the East GIS.

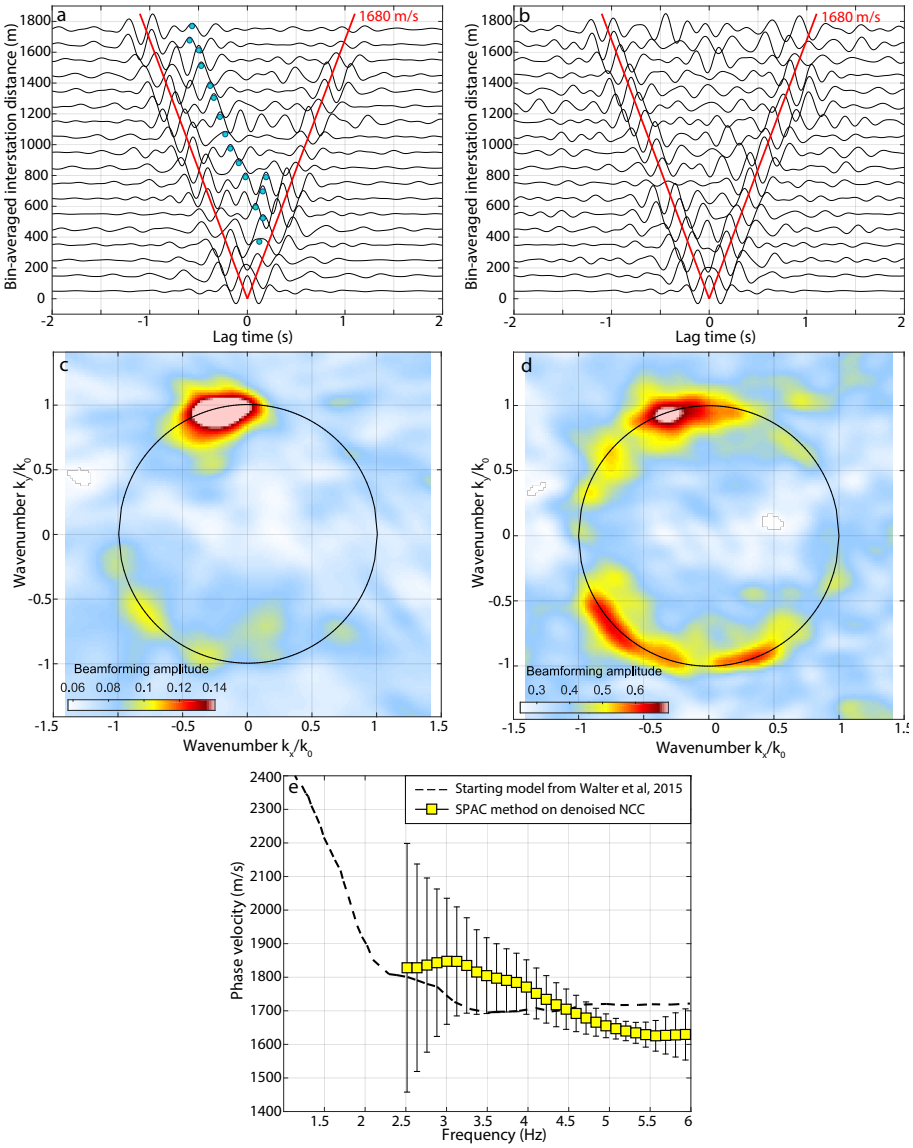

Figure 9. (a) Stacked sections of NCC in the frequency band from 2.5 to 6 Hz. The red line shows the propagation of the Rayleigh waves with velocity of 1680 m/s (also in b). Spurious arrivals are marked with green dots. (b) Stacked sections of NCC reconstructed in the frequency band from 2.5 to 6 Hz from the CSDM eigenspectrum equalization. (c) Plane wave beamforming before the eigenspectrum separation and normalization, and (d) afterwards. (e) Rayleigh wave phase velocity dispersion curves (yellow squares) calculated over the averaged seismic section in (d) with errorbars indicating discrepancies in velocity measurements at different station paths. The dispersion curve is obtained with the Aki's spectral method. Dashed black line is the dispersion curve obtained by Walter et al. (2015) with MFP.

After the eigenspectrum equalization, we are able to extract Rayleigh waves dispersion curve from the averaged seismic section obtained in Fig. 9b. For calculating the averaged dispersion curve we use a version of the Aki's spectral method (Aki, 1957) which consists of fitting a Bessel function to the real part of the cross-correlation spectrum. This method is referred to as SPAC and is described in Appendix B2. The Rayleigh wave phase velocity dispersion curve averaged over all station measurements is shown in Fig. 9e with yellow squares and errorbars representing the measurement discrepancies for individual NCC. The dotted line presents an apriori Rayleigh velocity dispersion curve extracted from Walter et al. (2015). High discrepancies observed at lower frequencies mainly arise from the limited frequency band for computing the NCC (see the appendix method section B). The slight differences between the two dispersion curves might be related to the different approaches used for phase velocity dispersion curve extraction (MFP in Walter et al. (2015), and SPAC in the current work). Moreover, Walter et al. (2015) worked on a wider frequency band and averaged their dispersion measurements over 46 days and in our study we use only one-day of data.

Several additional tests could be used to further improve the SNR of the NCC and their convergence to GF. ~~A similar procedure could be performed for other days. It would be useful to find an automatic criterion of the eigenvalue selection based on the near-field beamforming (calculated over spatial positions with no plane wave approximation, as in Fig. 7a) and far-field beamforming (calculated over wavenumbers using the the plane wave approximation, as in Fig. 9c, d).~~ For example, a similar procedure could be performed on other days, and the eigenormalized NCF could be stacked over a few days to increase the SNR. However, we verified that the index of eigenvectors corresponding to the moulin changes over days (the moulin can be located in the first, second, third etc. eigenvector). This is the reason why it would be useful to find an automatic criterion for the eigenvalue selection based on the MFP output. However, this is beyond of the scope of this paper. Another improvement could consist of azimuthal stacking the NCC according to the direction of the noise sources, although the GrIS array does not have sufficient azimuthal and spatial coverage to implement this. Moreover, we could envisage calculating a projector based on the SVD (as in MRABF) only for the time period when the moulin is active and then project-out the moulin signature from the continuous seismic data.

In summary, we conclude that the CSDM eigenspectrum equalization together with beamforming-based selection of eigenvectors is an useful method to separate localized seismic sources in a glaciated environment. It can further improve the GF emergence from ambient seismic noise in the presence of strong, localized englacial noise sources for imaging applications.

5 Cross-correlation of icequake coda waves ~~using the Gornergletscher array~~: a window-optimization approach

~~Inhomogeneities in the Earth's crust give rise to coda waves (following the ballistic seismic arrivals after an earthquake) consisting of reflections and scattered waves (references).~~ Contrary to ballistic waves, the likely diffuse coda arises from multiple seismic scattering (Aki and Chouet, 1975; Shapiro et al., 2000; Hennino et al., 2001) and is expected to contain all possible modes and propagation directions following an equipartition principle (Paul et al., 2005; Colombi et al., 2014). ~~Diffusion of seismic coda waves~~ Scattered coda waves after an earthquake ~~can thus provide an isotropic wavefield to allow for~~ favors isotropy of the incident wavefield and GF estimate retrieval GF estimate retrieval via the cross-correlation of a coda

window at two sensors (Campillo and Paul, 2003; Malcolm et al., 2004; Paul et al., 2005; Gouédard et al., 2008b; Chaput et al., 2015a, b).

In contrast, glacier ice is highly homogeneous with scattering crevasses mostly confined to the surface crevasse zone, typically 30 m-thick. As a result, icequakes typically lack a strong coda (when compared to earthquakes, Fig.1a), usually limited to a few milliseconds after the ballistic waves, with a fast decay of seismic amplitudes toward the noise level. However, focusing on few strongest events recorded in Gornergletscher (Fig. 2c), we still observe sustained coda (1.5 s-long) above the background noise (Fig. 1c) whose origin is discussed below. In the following, we explore the application of coda wave interferometry (CWI) on selected near-surface icequakes to estimate the GF, which could not be obtained from traditional processing of icequake cross-correlations because of lacking sources in the stationary phase zones of the seismic array. The use of icequakes here is fundamentally different than in section 3, in that the ballistic arrivals are specifically avoided (whereas in the other case, the ballistic component was the primary source of energy in the cross-correlation functions).

5.1 Origin of Icequake coda waves at Gornergletscher

The coda of a local earthquake is a mix of multipath arrivals and is formed by waves that are deviated from their original path by multiscale heterogeneities (i.e. velocity or density anomalies as cracks, rock pores, fluid bubbles, etc) present in the propagation medium. The seismic coda can also consist of reverberations in horizontally layered structure under the receiver, reverberations in layered structures between the source and receiver , and surface waves scattered by lateral heterogeneities. In addition, coda can be generated by the conversion of body waves into surface waves by topography at the free surface or at buried interfaces at depth (refererences).

In a strong scattering medium, ballistic and singly-scattered waves will follow a random walk between the scatterers until forming a diffusive halo which consists of multiply-scattered waves, and whose propagation velocity is influenced by the scattering strength often measured by the scattering mean free path (refererences). The diffuse character of coda waves has been demonstrated by the observation of the stabilization of the S-to-P energy ratio. Contrary to ballistic waves, fully diffuse wavefields are expected to contain all possible modes and propagation directions following an equipartition principle (refererences). Diffusion of seismic coda waves is the pre-requesite condition for a homogenized seismic wavefield and the correct estimation of the GF through the application of CWI (refererences).

The strongest 720 events chosen out of more than 24000 icequakes detected at Gornergletscher exhibit a sustained coda with approximate duration of 1.5 s (Fig. 1e d-e). The propagation regime of seismic waves can be identified by the evolution of the elastic intensity ("coda power spectrum") being the squared seismic amplitudes (color lines in Fig. 1e d). Before the source energy has reached the receiver, the elastic intensity is equal to some background or ambient level. Once the source pulse arrives at the receiver, the intensity rises up and then begins to decay exponentially. This is the ballistic regime. After several mean-free times which are to be related to the scattering strength (De Rosny and Roux, 2001), the intensity begins to decay diffusively with time as multiple scattering slows the transport of energy out of the scan region (Malcolm et al., 2004). This is the diffusion regime and it is characterized by a linear decay of the coda intensity (Aki and Chouet, 1975). Eventually, intrinsic attenuation (anelastic loss) dominates and the energy falls to the noise level.

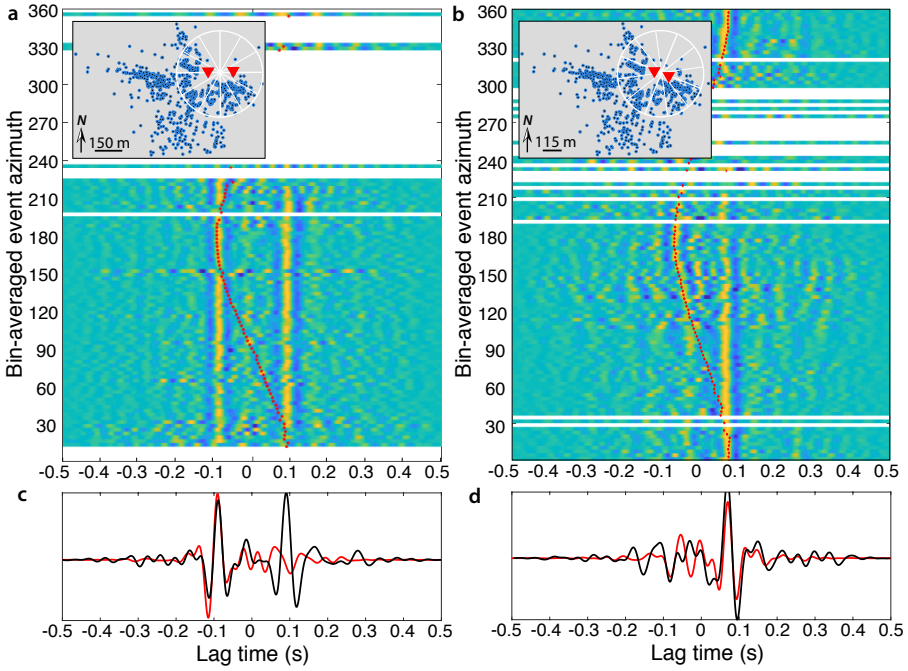

Figure 10. (a, b) cross-correlation functions obtained at two station pairs from each source and plotted as a function of event azimuth relative to the station path, for a time window in the coda part. Rayleigh wave arrival times of correlation functions obtained for a time window which encompasses ballistic waves of icequake records are plotted in red dots. Stations, station separation and source geometry are plotted in the inset maps. (c, d) Stack of the correlation functions obtained from icequake ballistic waves (red) and coda (black).

Figure 1e d shows such linear decay of the coda power spectrum starting at ~ 0.5 s indicating that icequake seismogram signals contain single or multiply ~~back-~~scattered energy that may approach from a wide range of directions assuming that the scatterers are homogeneous around the network site (Chaput et al., 2015a). In the present study we do not investigate further the cause of wave scattering in glaciers and particularly in Gornergletscher but suggest a relation to the presence of pervasive conspicuous near-surface crevasses (Fig. 2c) and deeper fractures as intermediate-depth and basal fault planes have been reported at the study site (Walter et al., 2008, 2009, 2010), as well as topography gradients, reflections at the glacier margins and/or rock and air inclusions.

5.2 Coda wave interferometry and Green's function ~~emergence~~ estimate

We first apply a standardized CWI processing scheme following Gouédard et al. (2008b). The cross-correlations are computed on 10-30 Hz spectrally whitened seismograms to reduce the influence of background noise. As a first guess, coda waves are arbitrarily time windowed around 0.5-1 s (gray horizontal bar in Fig. 1e d) by looking at the decay of the waveform amplitudes. The first sample of the coda correlation window corresponds to the two station average of the time when the seismogram

envelope falls below 5% of the ballistic wave maximum amplitudes. Because of the decrease of coda amplitude with time, we cannot perform a simple cross-correlation between the coda signals without strongly overweighing the earliest part of the coda. To avoid this problem, we follow Paul et al. (2005) by disregarding the amplitudes and considering only one-bit signals.

Figure 10a shows the individual coda wave cross-correlation functions (CWCC) sorted by the azimuth of the source event relative to the station path. In contrast to conventional ICC (correlations of icequake ballistic waves) whose computed arrival times (see also Fig. B4 and Sect. 3) are plotted in red, the coherent arrival times in the CWCC no longer depend on the event azimuth. The CWCC arrival times correspond to stationary Rayleigh waves travelling between the two stations. The causal and acausal parts of the individual CWCC tend to symmetrize as we are in the scattering regime. This results in a symmetric correlation stack (Fig. 10c) whereas only the acausal part of the GF is reconstructed from the ICC due to missing sources behind one of the two stations.

For the pair of closer stations (Fig. 10b), the reconstructed acausal times still depend on the source position while the CWCC causal times are stable with the event azimuth. The source position signature on one side of the correlation function could be an effect of heterogenous scatterers which cause single-~~back~~ scattering and then skew the illumination pattern at one side of the receiver pair. Another explanation could be that the correlation window used here for CWI is still influenced by the incoming energy flux from ballistic waves which then create an anisotropic incident wavefield as we are in the presence of limited energy diffusion (Paul et al., 2005).

Focusing on a complex scattering medium at the glaciated Erebus volcano (Antarctica), Chaput et al. (2015b) showed that symmetric GF could be recovered when optimizing the icequake coda correlation window over the sources. In the case of a weak scattering medium as glacial ice, the coda time window for the diffusion regime should notably depend on the distance of the scatterers to the recording seismic sensor. We therefore use a similar optimizing-window processing scheme for improving the GF convergence at each station pair.

The overall processing and technical details of coda window optimization are described in Chaput et al. (2015b). We refer to the method as MCMC processing as it involves a Markov Chain Monte Carlo scheme. A Bayesian inversion determines the best coda window to generate a set of CWCC that are the most coherent and symmetric across the source events. We first construct a matrix of CWCC that are bin-averaged over N events and then iterate this correlation matrix by randomly shifting the coda window along a certain number M of random traces. At each iteration, a misfit function is constructed based on the coherency of the N-binned CWCC matrix and the causal-acausal symmetry of the CWCC stack. In the end, the best optimized models consisting of the cross-correlation matrices computed for different sets of individual event coda windows are stored and used to generate an average stack of CWCC which is our estimate of the final GF.

MCMC processing involves several parameters that need to be tuned such as: the frequency band of analysis (here 10-30 Hz); the number of selected traces $N = 40$ and $M = 10$ at each iteration; the coda correlation window length T, we use $T = 0.5$ s and we force the algorithm to shift the correlation window to no later than 1.5 s in order to stay within the icequake coda and to not correlate noise (see Fig. 1c); the portion of the cross-correlation stack where we want to optimize the causal-acausal symmetry (i.e. ballistic waves of the constructed CWCC stack and/or the cross-correlation coda part), we choose here to optimize the CWCC symmetry for both reconstructed ballistic and coda waves, ie. at later times than what expected for a

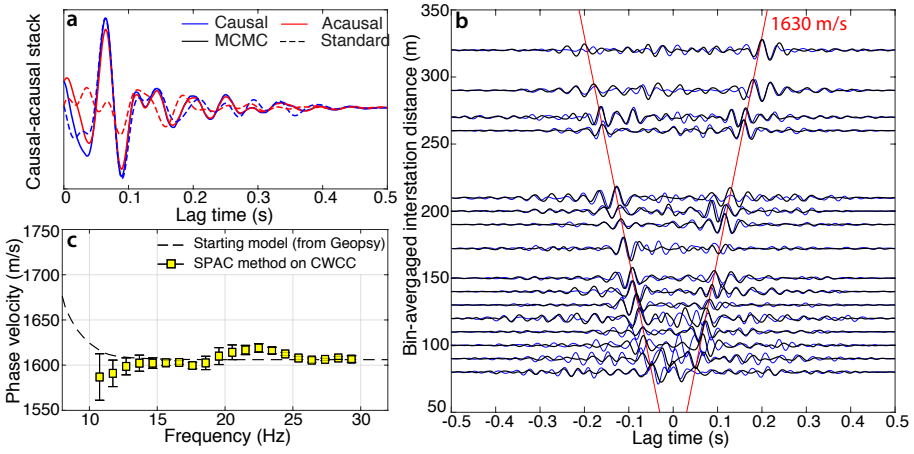

Figure 11. (a) Causal-acausal symmetry of the CWCC obtained one station pair using standard coda interferometry processing as in Fig. 10d (dashed lines) and MCMC processing (solid lines). (b) Correlation gather sorted by increasing distance and averaged in 10 m intervals. Black and blue lines result from MCMC and standard processing, respectively. Red lines show the propagation of Rayleigh waves with the velocity of 1630 m/s. (c) Rayleigh wave phase velocity dispersion curve (yellow squares) averaged over measurements for all station pairs. Error bars indicate discrepancies in velocity measurements at different station paths. The dispersion curve is obtained with the Aki's spectral method. Dashed black line is the theoretical dispersion curve for a 160 m-thick ice layer computed with Geopsy.

wave propagating at velocity 1700 m/s; the relative importance of the causal-acausal symmetry and the cross-correlation matrix coherency that is used to optimize the misfit function (coefficient factors A and B in Chaput et al., 2015b, , equation 2), we here weight the symmetry and coherency evenly; the maximum number of iterations we allow, which naturally depend on the number of available sources, N and M (we use up to 2×10^4 iterations).

Figure 11a shows the MCMC optimization of the CWCC symmetry at the station pair already presented in Fig. 10d. Blue and red lines are the causal and acausal parts of the correlation stack, respectively. Solid and dashed lines are the resulting CWCC obtained from MCMC processing and standard processing (i.e. first iteration of the MCMC inversion), respectively. While the ballistic Rayleigh waves could not be reconstructed in the acausal part of the correlation function using the first coda wave windowing, the MCMC output approaches the symmetrical GF as we see a Rayleigh wave propagation in both directions, and also the emergence of a coherent coda in both parts of the correlation function. The MCMC inversion gives optimized coda window in the range 0.7-1.2 s for the majority of events, i.e. at times larger than the initially used coda window. Enhanced symmetry of the CWCC computed in later time windows is consistent with the expectation to have a more isotropic diffuse wavefield as all directions of propagation are closer to be equally represented after several mean free paths (Paul et al., 2005).

Figure 11b shows the final stack of CWCC gather sorted by increasing inter-station distance and averaged in 10 m-binned intervals. For comparison, CWCC computed with the standard processing are in blue. CWCC are noisier that ICC obtained from correlations of icequake ballistic waves when computed on homogeneous source distributions (Gouédard et al., 2008b). Nevertheless, at most station pairs, the MCMC processing managed to extract Rayleigh waves with consistent traveltimes in

both causal and acausal parts. We extract a dispersion curve of phase velocities averaged over the station components (Fig. 11c) using the SPAC method already used in section 4.2 (Appendix B2). We find an average Rayleigh wave velocity of 1600 m/s which is in the estimate range of what Walter et al. (2015) find at Gornergletscher using slant-stacking of ICC arrival times (Appendix B3). Errors introduced at lower frequencies arise from the limited frequency band used for computing the CWCC and filtering effects (see the appendix method section B).

Note that MCMC processing coherently increases the presence of energy at zero lag-time (Fig. 11a-b). Such spurious arrival likely arises because scattered coda also contains a strong vertically trapped body wave that correlates at 0 across relatively near receivers, even if it is not part of the "true" GF. Obtained CWCC may contain spurious arrivals and seismic modes that are not purely the result of an isotropic point source GF. We point out two reasons for this.

On the one hand, as stated before, spurious arrivals at times 0 or later could result from seismic reflections on the glacier bed beneath the stations, cross-correlations of early aftershocks or other noise sources. A certain portion of the icequake coda may be still influenced by background noise sources especially at distant stations from the source as the coda time window of one station may fall in the noise window of the further one.

On the other hand, spurious arrival contributions will not vanish in case of localized scatterers around the seismic array if the incident waves do not illuminate the scatterers with equal power from all directions because of limited source aperture (Snieder et al., 2008). This second argument is also supported by the observations of non-symmetric CWCC. At some locations, there still exist differences in the amplitudes of the Rayleigh waves in the causal and acausal parts of the final GF estimate (Fig. 11a), meaning that the icequake coda is not entirely diffuse and may result from single reflections on preferred scatterers. Paul et al. (2005) could not obtain symmetric CWCC from regional earthquake coda seismograms and attribute this to ~~anisotropic diffuse wavefield~~ the long-lasting anisotropy of the diffuse energy flux. Indeed, in weak (or homogeneous) media, the incident energy flux from earthquakes can still dominate the late coda resulting in GF time-asymmetry, provided the sources are located in the same distant region. The CWCC asymmetry is expected to disappear with an isotropic distribution of sources or scatterers around the seismic network. In the case of Gornergletscher icequakes, we still see the influence of the energy flux approaching from the direction of the source at a few station pairs and for some events as depicted in Fig. 10b, supporting the argument for single scattering rather than multiply scattering as the cause of icequake coda.

In general, CWI must be processed on carefully selected events which show sustained coda above the background noise. We try coda wave correlations on the Argentière node grid and notice an influence of the source position on the retrieved CWCC likely because we did not select strong enough signals with sustained coda that is coherent enough across the sensors. Similarly, GF convergence does not work for weak Gornergletscher icequakes. Moreover, the abundance of seismic sources in glaciers often pollutes coda wave seismograms. We often find the situation where ballistic body and surface waves generated by early aftershocks from repetitive and subsequent events (or bed reflections) arrive at the seismic sensor only a few milliseconds after the onset of the first event of interest and therefore fall in its coda window. This typically ~~biases CWCC computation~~ introduces anisotropic wavefields. The brief icequake coda duration, the interevent time distribution and the weak scattering in glacial ice impose limitations of CWI on large arrays.

To conclude, even if limited, the extraction of GF from icequake coda waves allows imaging a glacier subsurface between station pairs. In principle, this can be done even in cases when (skewed) distribution of icequake sources or sustained noise sources does not allow for GF estimation.

6 Discussion

~~In this study, we show the possibility to extract informations about the glacier subsurface structure and elastic properties from cross-correlations of typical seismic wavefields recorded in ablating ice-covered areas.~~ The three methods proposed in this study could theoretically be applied to each one of the explored dataset but were not further tested here. The standard processing schemes proposed in section 3 for computations of NCC and ICC ~~due to distributed sources~~ as we benefit from favorable seismic illumination patterns and redundancy of the measurements on the dense array. ~~We are able to obtain a precise image of the glacier subsurface (with uncertainties of the order of 10 m) thanks to the dense measurements points which allows to improve the SNR of the cross-correlation functions by stacking them in regular distance intervals. Spatial averaging~~ The need of spatial averaging of the correlation responses in regular distance intervals could potentially be overcome when (1) stacking the GF over longer time range of typically weeks or months (e.g. Sabra et al., 2005; Larose et al., 2008) or (2) applying more advance processing schemes as the one proposed for GIS and Gornergletscher studies, to overcome ~~uneven~~ anisotropic source distribution effects.

~~MPF and c~~ Eigenspectral equalization of the cross-spectral matrix of ambient noise seismograms (Sect. 4) enable to distinguish propagation directions of incoming seismic energy. ~~This method coupled to an adequate selection of the eigencomponents allows to remove or at least attenuate the signatures from dominant and localized noise sources as in the case of the GIS moulin.~~ This method improves the spatial homogenization of the incident wavefield in order to reveal weaker sources that could eventually be used for extracting the GF at sensor pairs which initially lacked stationary phase contributions. Same technique could then be applied to Glacier d'Argentière in order to improve the GF estimates ~~from NCC at station pairs not aligned with the strongest noise sources that lie down and upstream of the array~~ which show spurious arrivals arising from the dominant sources near the glacier tongue (Fig. 3b). By doing so, ~~dispersion curves of Rayleigh wave velocities could be obtained at more station pairs and used to~~ we would improve our spatial coverage of seismic velocity measurements and could invert a more accurate 3D model of the subsurface. ~~The eigenvalue distribution of the cross-spectral density matrix depends on the number of receivers.~~ The eigenspectral equalization method is then particularly well suited for large sensor arrays, also the geometry of the glacier discharge system (channelized versus distributed) is the primary controlling factor for successful applications.

In the absence of distributed noise sources, the use of icequakes is a good alternative ~~for GF estimates~~ especially in winter when the glacier freezes ~~and no melt occurs~~ preventing for generation of coherent water flow noise. As icequakes propagate to few hundreds meters, ICC studies are well suited for medium-size arrays with aperture of typical order of 500 m. In the case of ~~uneven~~ anisotropic distributions of icequakes which map the crevassed ice, GF estimates can be optimized with CWI and the coda-window optimization approach used in section 5. CWI successfully works on the strongest selected events at Gornergletscher because we could record a coherent coda at adjacent sensors, with seismic amplitudes above background

noise level. T CWI should be appropriate for smaller icequake datasets (here we employ the method on a loop of 700 events, i.e. less than 10% of the icequake catalogue used in Argentière) and smaller size arrays to be able to record coherent coda across the network (typically 250 m-wide). Such arrays can be deployed in targeted regions where pervasive crevasses are present, i.e. typically near the glacier margins. As an example, CWI could help to estimate the GF at smaller subarrays at the edges of a larger array as in the configuration of Glacier d'Argentière experiment.

In the following we discuss the type of array deployment, geometry and measurements suitable for structural and monitoring studies using the GF obtained by either one of the processing described above.

6.1 Implications for glacier imaging

~~The capability of extracting accurate seismic velocities on a line of receivers can be substantially improved when stacking the correlation functions on a large number of receiver pairs. Dense array deployment of sensors regularly spaced on a grid enables to densify the measurement points and will thus improves the results. Technological improvements in the portability of seismic instruments and storage capacity make today such kind of experiment feasible to deploy, as it was done on Glacier d'Argentière.~~

The performance of an array for ~~deriving phase velocity values in a wavenumber or frequency range depends on~~ imaging the structure at depth first relies on its geometry, ~~and~~ secondly on the wavefield characteristics as discussed in section 3.1. Wathelet et al. (2008) recommend that the array diameter should be at least as large as the longest wavelength of interest (we conventionally take two to three wavelengths). The minimum station spacing for any direction should be less than half the shortest wavelength of interest to avoid spatial aliasing. ~~To be able to sample small-scale glacier heterogeneities (e.g. individual crevasses, cavities, short-wavelength basal topography, etc), the investigated seismic wavelength should be of the same order of the anomaly, giving typical sensor spacing of the order of 10 m.~~ To be able to sample the basal interface and target elastic parameters of sediments constituting the till layer or the bedrock, one should design suitable array geometry. For a glacier thickness of 200 m, we need an array aperture of at least 600 m to measure propagating surface waves with a one-wavelength cycle. For a 500 m-thick glacier, we need sensors that are at least 1500 m apart, although prior knowledge on the basal interface allows to better constrain the inversion of depth with lower-size arrays.

~~Finally, as outlined in Fig. 4, Rayleigh waves reconstructed on the cross-correlations of continuous ambient englacial noise (NCC) and of icequake records (ICC) are differently sensitive to the subsurface layers, primarily due to the recorded wavefield characteristics and spectral content. Given typical glacier noise sources, NCC can reconstruct propagating waves down to frequencies of 1 Hz and thus well sample the basal layer for glaciers thinner than $\sim$ 600 m. In contrast, surface crevasse icequakes have dominant energy at frequencies above 5 Hz. ICC are thus more appropriate to sample the surface layers. To obtain elastic properties of basal layers in thicker glacier settings and ice-sheets, cross-correlations of oceanic ambient noise recorded on broadband seismometers should be considered as such sources sample lower frequencies below 1 Hz (Zhan et al, 2013).~~

Seismic velocity measurements can additionally be complemented by other types of observations computed on the horizontal and vertical components of the GF, such as the horizontal-to-vertical (H/V) ratio. ~~This technique can be implemented on~~

~~cross-correlation functions (ref) or on raw continuous seismograms (ref) as long as the seismic sensors are well oriented and coupled to the ice. The H/V measurements exploit the ellipticity of the propagating Rayleigh wave which is strongly sensitive to the vertical-layer structure beneath the recording site. H/V ratio can be computed at different frequencies, we then call it H/V spectral ratio. The spectrum of the H/V ratio reveals the fundamental SH-~~wave ~~resonance frequency of a site which is related to the basal impedance contrast. 1D resonance of S-waves is particularly strong for a "soft" layer (such as ice) on top of a more rigid layer (such as bedrock).~~ Assuming a horizontally homogeneous medium, the resonance frequency of the H/V ratio spectrum can be used to constrain the layered structure (Zhan et al., 2013) and the ice thickness (Picotti et al., 2017) or to investigate 3D effects of the recorded wavefield (Preiswerk et al., 2018). H/V ratios and dispersion curves of surface wave velocities can be jointly inverted (Lin et al., 2014) for an even more accurate 3D model of the glacier subsurface.

6.2 Implications for glacier monitoring

~~Changes in subsurface properties affect seismic wave propagation.~~ Repeated analysis of cross-correlation analysis allows to detect changes in the subsurface properties. ~~The accuracy for detecting and quantifying seismic velocity changes $\delta v/v$ from time shifts in the ballistic wave traveltime measurements depends on (1) the quality of the GF estimate which is primarily controlled by the stability of sources over time and (2) the magnitude of the velocity variation.~~Seismic velocity monitoring is usually performed in the coda part of the cross-correlation function through the application of CWI, as multiple scattered coda waves are less sensitive to the source distribution and travel larger distances accumulating time delays (Hadziioannou et al., 2009). CWI enables to detect relative velocity changes as small as ~~$\delta v/v \sim 10^{-3}$~~ 0.1% (e.g. Sens-Schönfelder and Wegler, 2006; Brenguier et al., 2008a, b; Mainsant et al., 2012).

CWI computed on ~~on-ice seismic recordings~~ the cross-correlation functions could lead to the monitoring of englacial crevasses, failure of calving icebergs, glacial lake outburst flood, break-off of hanging glaciers, surface mass balance and such bed conditions as the evolution of the glacier hydraulic system and subglacial till properties. Such topics are currently investigated with active seismics or through the spatio-temporal evolution of passive seismicity and associated source mechanisms (e.g. Walter et al., 2008; Bartholomaus et al., 2015; Preiswerk et al., 2016; Podolskiy et al., 2017; Lipovsky et al., 2019). Passive seismic monitoring of glaciers could lead to the detection and understanding of processes related to climate conditions, glacier hydraulics and ice flow dynamics, which today are labor-intensive to investigate with active geophysical measurements or satellite imagery.

Unfortunately, the weak ice scattering limits the emergence of coherent coda in the correlation functions which appear still to be affected by the changing nature of primary ambient noise sources (Walter, 2009; Preiswerk and Walter, 2018). To overcome source effects and enhance the coda SNR, it is possible to design arrays for multidimensional deconvolution (MDD) of the time-averaged cross-correlations. Seismic interferometry by MDD measures the illumination pattern (e.g. Wapenaar et al., 2011) using recordings by a set of additional receivers along a contour which goes through the virtual-source's location (Fig. 12a). MDD processing then enables to remove the imprint of the (non-uniform) source distribution from the correlation responses and improves the quality of the retrieved GF. ~~By employing virtual boundary conditions in seismic interferometry by MDD (Weemstra et al, 2017) on active-shot signals, Lindner et al (2018b) could create artificial coda following the ballistic~~

wave. This coda consists of multiple seismic reflections within the (closed) receiver contour trapping the waves. Applied to icequake signals recorded at the Glacier d'Argentière array (Fig. 12a), this technique is successful in retrieving stable GF estimates over time with coherent coda energy (Fig. 12c). We follow Lindner et al. (2018) and use the outer receiver contour of the Glacier d'Argentière array (black triangles in Fig. 12a) as virtual boundaries for MDD of the icequake correlations computed at one pair of sensors and sorted in two-day overlapping windows. MDD processing appears successful in retrieving accurate GF estimates which feature a coda that is coherent over time (Fig. 12c) and does not appear on more classical ICC as defined in section 3 (Fig. 12b). The technique based on virtual reflectors allows to create artificial coda which consists of multiple reflections of the ballistic wavefield trapped within the receiver contour. This approach is a promising candidate for a monitoring technique of any changes in seismic velocities in glaciers even in the absence of scattering coda. Long-term installations directly on or in the ice have increased in recent years due to technological improvements (Aster and Winberry, 2017). MDD could be applied to glacier seismic sources recorded over time scales longer than a month, using borehole sensors which ensure a solid coupling to the ice.

7 Summary and concluding remarks

This study explores the application of seismic interferometry on on-ice recordings to extract the elastic response of the glacier subsurface beneath one array deployment. In contrast to ambient noise studies focusing on the Earth's crust, the GF retrieval from the cross-correlation of glacier seismicity at sensor pairs is notoriously known as difficult due the limited spatial coverage of glacier point sources and the lack of seismic scattering in homogeneous ice. We investigate the GF emergence on three particular cases: (A) a favourable distribution of icequakes and noise sources recorded on a dense array (Sect. 3), (B) a spatially limited distribution of icequakes and a dominant noise source constituted by a moulin (Sect. 4), and (C) crevasse-generated scattering (Sect. 5) . We design processing schemes suitable for each configuration of seismic deployment, wavefield characteristics and glacier setting.

In Glacier d'Argentière (Sect. 3), cross-correlations of water flow ambient noise and icequake recordings result in accurate GF estimates, when averaged in regular inter-station distance intervals as (1) we face the situation of a favourable source distribution and (2) the large number of sensors and their dense spatial sampling allows to stack redundant measurements on a line of receivers.. The averaged GF well reconstruct P and dispersive Rayleigh waves propagating across the array. Seismic velocity inversions enable to conduct structural studies, map the englacial crevasses and the glacier bed with vertical resolution of ~ 10 m.

On the GIS (Sect. 4), spurious arrival times for Rayleigh waves are obtained from ambient noise cross-correlation and are attributed to the moulin seismic signature which acts like a secondary dominating source. They can be removed by using matched-field processing on the cross-spectral matrix of seismic recordings and equalization of eigenspectral values. The processing involves decomposition of the wavefield into independent components of the seismic energy contribution. This method allows to locate and separate different noise sources and improve the GF convergence by selecting appropriate directional components. cross-correlations of ambient noise seismograms give rise to spurious arrivals which are not part of

the true GF. MFP identifies a localized point source constituted by a moulin within the seismic array, out of the stationary phase zones of most of the receiver paths. Eigenspectral equalization of the cross-spectral matrix coupled to an adequate selection of its eigencomponents enables to remove or at least attenuate the imprint of the moulin on the correlation functions. Such technique allows to separate and weight the contributions of different noise sources shaping optimized conditions for GF convergence.

In Gornergletscher (Sect. 5), cross-correlations of icequake coda waves show evidence for a quasi-homogenized incident wavefield ~~coming from a wide range of directions as a result of seismic back-scattering likely attributed to pervasive and penetrating englacial crevasses~~ as a result of seismic scattering by the crevassed ice. ~~By optimizing the correlation window in the coda part, we compute symmetric GF which could not be obtained otherwise by cross-correlation of icequake ballistic waves due to a skewed source distribution which creates a non-homogeneous illumination pattern. Coda wave interferometry is a technique usually employed in crustal seismology but works fine on accurately selected strong icequakes with sustained 1.5 s-long coda waves that can be recorded above the background noise level.~~ An optimization of the coda window based on the coherency and the causal-acausal symmetry of the cross-correlation functions is used to improve the GF convergence. CWI allows to retrieve GFs at seismic arrays where icequakes are not recorded evenly around the study site.

The capability of extracting accurate seismic velocities on a line of receivers can be substantially improved when stacking the correlation functions on a large number of receiver pairs. However, MPF and CWI allow for new kinds of measurements on sparse seismic networks and enable to speedup the GF convergence for non-idealized seismic illumination patterns which commonly arise in glacier settings.

Finally, the use of nodal sensor technology enables fast deployment of large N-arrays suitable for seismic interferometry studies. This opens new ways for characterizing and monitoring glacial systems using continuous passive seismic recordings.

Code and data availability. All data except for Argentière are available by request on the server of the Swiss Seismological Service (http://seismo.ethz.ch). Obspy Python routines (www.obspy.org) were used to download waveforms and process icequake catalogues. NCC of Argentière dataset were computed using the MSNoise Python package (www.msnoise.org, Lecocq et al., 2014).

Author contributions. PR and FG designed the Argentière experiment in the scope of the RESOLVE project (https://resolve.osug.fr/). FL, PR, FW and AS processed the icequake catalogues. AS processed the correlation functions and analyzed the results at Argentière. MFP of Greenland data was processed by MC. CWI at Gornergletscher was processed by AS and FW with the input of JC. Icequake correlations by MDD were processed by FL. AM provided the code for computing phase velocities by fitting the cross-spectrum with a Bessel function. AS and MC prepared the manuscript with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was funded by the Swiss National Science Foundation (SNSF) project Glacial Hazard Monitoring with Seismology (GlaHMSeis, grant PP00P2_157551). Additional financial support was provided to AS by the Swiss Federal Institute of Technology (ETH Zürich). The Observatory of Sciences of the Universe of Grenoble funded the project RESOLVE for the data acquisition at Argentière. SNSF and ETH Zürich participated to the data collection in Greenland (grants 200021_127197 SNE-ETH and 201 ETH-27 10-3) and on Gornergletscher (grants 200021-103882/1, 200020-111892/1). We gratefully acknowledge all the people involved in the RESOLVE project and who participated to the array deployment in Argentière, collected and processed raw formats of seismic data. We thank Olivier Laarman, Bruno Jourdain, Christian Vincent and Stéphane Garambois for having constructed bed and surface DEMs using ground penetrating radar for the bed and drone data for the surface. We also thank Léonard Seydoux for insightful discussions on MPF and SVD.

Appendix A: Construction of icequake catalogues

There exists a wide range of seismic sources in glaciers as well as detection schemes (see Podolskiy and Walter, 2016, for a review on the methodology). The processing employed for icequake detections and localizations must be adapted to the type of network (~~network versus array,~~ number of sensors, sensor spacing and array aperture) and to the type of the events of interest which involve various waveforms (e.g. Walter, 2009; Röösli et al., 2014; Podolskiy and Walter, 2016). Dispersive Rayleigh waves are well suited for investigating the glacier subsurface including the basal interface as they are primarily sensitive to the S-wave velocity structure and can sample depths up to approximately $1/3$ of the wavelength. We then focus our study on one class of glacier seismic events being surface icequakes generated by ice crevassing. Another advantage of using such events is their high rate of time occurence ($\sim 10^2 - 10^3$ recorded events per day) and their potentially wide spatial coverage which is optimal for the application of seismic interferometry techniques (Walter et al., 2015). We here introduce the methods used to compute the icequake catalogues at Gornergletscher (Sect. A1 and A2) and Glacier d'Argentière (Sect. A3).

A1 Icequake detection

Seismic waveforms of surface icequakes generally exhibit a first low-amplitude P-wave followed by impulsive Rayleigh waves (Fig. 1~~e~~ a). Such events can be detected using a template matching on continuous seismograms (Mikesell et al., 2012). This cross-correlation method exploits the signal coherency with a reference waveform and can be used on single stations or across a network. Nevertheless, the most common and straigthforward detection approach is to implement an amplitude threshold trigger.

The most broadly used algorithm in weak-motion seismology and on glaciers for detecting impulsive events (e.g. Walter, 2009; Canassy et al., 2012; Barruol et al., 2013) is the "short-time-average long-time-average ratio trigger" referred to as STA/LTA (Allen, 1978). It continuously calculates the average values of the absolute amplitude of a seismic signal in two consecutive moving-time windows. The short time window (STA) is sensitive to seismic events while the long time window (LTA) provides information about the temporal amplitude of seismic noise at the recording site. When the ratio of both exceeds a preset value at a single station or coherently across a network, an event is declared.

As icequakes usually propagate to a few hundred meter distances before attenuating to the background noise level, the amount of identified events varies with the network configuration and the minimum number of stations to required trigger

concurrently. For the Gornergletscher study, we work with events that have been detected by running a STA/LTA trigger over 5-15 Hz bandpass filtered continuous seismograms, using an STA windows of 0.3 s (i.e. typical icequake duration), and an LTA windows 10 times longer with a threshold value of 8. To declare an event, we require at least half of the network stations to trigger.

A2 Icequake location

The vast majority of icequakes recorded on glaciers are localized near crevasses that extend no deeper than $\sim$ 30 m (Walter et al., 2008; Lindner et al., 2019). To locate the events at Gornergletscher, we fix the source depth to the surface and invert for the epicenter distance following the automated approach of Roux et al. (2010) and Walter et al. (2015), also similar to Mikesell et al. (2012). This method employs cross-correlations to automatically measure differences in Rayleigh wave arrival times across the network.

To be able to record coherent icequake waveforms with high enough signal-to-noise ratios (SNR) at couples of sensors, the network aperture should be less than 1 km, or at least consist of several pairs of stations whose separation is shorter than the distance at which surface waves start to be strongly attenuated in the ice.

In the same spectral band that is used for event detection, icequake signals are first windowed around the Rayleigh wave and cross-correlated for each pair of stations to obtain time delays. Time delay measurements are then refined to subsample precision by fitting a quadratic function to the cross-correlation function centered on its discrete maximum. The best Easting and Northing coordinates of the source are concurrently inverted with the apparent propagation velocity to match the time delay catalogue for pre-selected pairs of stations. We only consider time shift measurements at pairs of stations whose cross-correlation maximum is above 0.8. This allows us to minimize complex source and/or propagation effects on seismic waveforms and the observed arrival times that can not be fit with the oversimplified velocity model.

The inversion process is an iterative procedure using a quasi-Newton scheme (Roux et al., 2010, equation 3). Reliability of icequake locations varies as a function of the events being inside or outside the network. Using a seismic network similar to the one of Gornergletscher used in the present study, Walter et al. (2015) estimated that, in the azimuthal direction the error remains below 2° for average apparent velocities in the range of 1600-1650 m/s.

A3 Array processing: matched-field processing using beamforming

A seismic network is called an array if the network aperture is shorter than the correlation radius of the signals, which is the maximum distance between stations at which time series are coherent, i.e. typically less than 1 km for glacier sources recorded by on-ice deployments (Podolskiy and Walter, 2016). A seismic array differs from a local network mainly by the techniques used for data analysis.

Dense sensor arrays have many advantages as the SNR can be improved by summing the individual recordings of the array stations. Compared to a single or couples of sensors, array processing techniques, such as beamforming, allow for time domain stacking which constructively sums coherent signals over the sensors and cancels out incoherent random noise, enhancing the signal detection capability.

Continuous data are scanned through matched-field processing (MFP) which involves time domain beamforming (Kuperman and Turek, 1997). Beamforming uses the differential travel times of the plane wave front due to a specific apparent slowness (inverse of velocity) and back azimuth to individual array stations (Rost and Thomas, 2002). If the single-station recordings are appropriately shifted in time for a certain back azimuth and velocity, all signals with the matching back azimuth and slowness will sum constructively. MFP can be processed using a decomposition of seismic signals in frequency components. To be declared as an event and furthermore, with accurate location, the beampower of aligned seismic waveforms at a given frequency (i.e. the norm of the cross-spectral density matrix and the array response, see Lindner et al., 2019, equation 3) must pass a preset trigger threshold. The final event location can be averaged from the beamforming solutions obtained at several successive discrete frequencies (assuming that slowness and back azimuth are close to constant in the considered frequency band).

MFP was successfully used by Corciulo et al. (2012) to localize microseismic sources at the exploration scale using ambient-noise data. Moreover, recent studies focused on developing an automatic, optimization-based MFP approach that does not require grid-search to localize thousands of weak seismic events in a complex fault zone (Gradon et al., 2019) and hydraulically fractured area (Chmiel et al., 2019).

The MFP method of Chmiel et al. (2019) was used on Argentière array to detect and locate about 4000 events each day, with a beampower threshold averaged over frequencies between 5 and 30 Hz set to 0.5. Locations of icequakes recorded over the 5 week deployment are presented in Fig. 2a. For computing the cross-correlation functions from icequake waveform described in Sect. 3, we restrict ourselves to 11100 events (red stars in Fig. 12). Such events have been identified following Lindner et al. (2019) with a STA/LTA trigger on 8-16 Hz bandpass filtered continuous seismograms (STA$=0.3$ s, LTA$=3.6$ s, trigger threshold$=11$ for events detected concurrently at the four corner stations of the array). Locations are the beamforming solutions using a grid search over Easting and Northing positions in 25 m$\times$25 m steps. All events with beampower lower than 0.5 were discarded.

Appendix B: Computation of phase velocity dispersion curves

Because dispersive surface waves of different frequencies propagate at different speeds, computation of seismic velocities generally involves Fourier analysis to decompose the wave into frequencies that compose it. One can distinguish two types of wave speeds.

The phase velocity c is the speed at which the phase of a wave propagates in space and is related to the angular frequency ω and the wavenumber k as

$$c(\omega) = \frac{\omega}{k(\omega)} \tag{B1}$$

The angular frequency is related to the time periodicity of the signal of frequency f as $\omega = 2\pi f$. The ground displacement is also periodic in space over a distance equal to the wavelength λ that is used to describe how the wave oscillation repeats in space via the wavenumber $k = 2\pi/\lambda$.

If the harmonic waves of different frequencies propagate with different phase velocities, the velocity at which a wave group propagates differs from the phase velocity at which individual harmonic waves travel (Stein and Wysession, 2009). The group velocity u of a wave is the velocity with which the overall energy of the wave propagates through space. If the signal has energy over a wide range of frequencies, $u = \mathrm{d}\omega/\mathrm{d}k$ and the group velocity is related to the phase velocity as

$$u(\omega) = c(\omega) + \frac{\mathrm{d}c}{\mathrm{d}k} \tag{B2}$$

In ambient noise tomography, it is common to measure the group velocity of dispersive Rayleigh waves travelling in the Earth's crust and upper mantle (e.g. Shapiro et al., 2005; Mordret et al., 2013). Dispersion curves of group velocities are usually computed using the Frequency Time Analysis (FTAN) of the noise cross-correlation time-series (Levshin et al., 1992). FTAN employs a system of narrow-band Gaussian filters, with varying central frequency, that do not introduce phase distortion and give a good resolution in the time-frequency domain. For each filter band the envelope of the inverse Fourier transform of the filtered signal is the energy carried by the central frequency component of the original signal. Since the arrival time is inversely proportional to group velocity, for a known distance, the maximum energy of the time-frequency diagram is obtained as a function of group velocity with frequency.

In glaciers, due to homogeneous ice, only weakly dispersive surface waves are recorded at on-ice seismometers. It is then difficult to use FTAN to measure Rayleigh wave group velocity dispersion. We choose here to compute the phase velocity dispersion curve for Rayleigh waves using different approaches. Obtaining the group velocity from the phase velocity is then straightforward while the reverse is not possible due to unknown additive constants which arise from the integration of the phase velocity over frequency (equation B2).

B1 Array processing: frequency-wavenumber analysis

Frequency-wavenumber analysis (f-k) was used to compute the velocity dispersion curves at Glacier d'Argentière (Sect. 3.1 and 3.2). The f-k analysis is a standard array processing for computing phase velocities from seismic time-series recorded on a line of receivers (Capon, 1969). It enables to identify and separate wave types and wave modes and also design appropriate f-k filters to remove any seismic energy in the original signal time-series.

The most basic f-k processing employs a 2D Fourier transform on both time and spatial components to construct the f-k diagram (Fig. B1b). We then need to select the dispersion curve of the phase of interest by picking the energy maxima of the 2D Fourier transform output. The absolute value of the f-k space is then transformed into the velocity $c(f)$ via the equation B1 (Fig. B1c).

There exist multiple array techniques for computing frequency-velocity diagrams using spectral analysis in time and space domains. Some of them are described in Rost and Thomas (2002), Gouédard et al. (2008a) and referenced in Ohrnberger et al. (2004). Concerning the cross-correlation functions obtained at Argentière array, we employ the phase-shift method of Park et al. (1998) which allows to construct a frequency-velocity diagram where dispersion trends are identified from the pattern of energy accumulation in this space. Then, necessary dispersion curves are extracted by following the diagram amplitude trends (Fig. 3c and Fig. B2c).

The performance of an array for deriving phase velocity values in a wavenumber or frequency range depends on its geometry and on the wavefield characteristics (i.e. frequency range and magnitude of seismic energy with respect to attenuation). The capability to resolve phase velocity at a given frequency depends on the array aperture (described by the array diameter $\Delta_{\max}$) and minimum sensor spacing ($\Delta_{\min}$) so that at least two wavelengths are sampled between adjacent receivers to avoid aliasing in the wavenumber domain (e.g. Wathelet et al., 2008). Phase velocities should then be computed for frequencies which satisfy $\Delta_{\min} \leq n\lambda \leq \Delta_{\max}$, with n usually taken as 2 or 3. This relationship depends on the expected phase velocity as $\lambda = c/f$.

B2 Aki's spectral method

This method was used to compute the Rayleigh wave velocity at the GIS and Gornergletscher arrays (Sect. 4.2 and 5.2). Whereas the f-k techniques are based on the assumption of a plane wave arriving at the array, the Aki's spectral method (also referred to SPAC method, Tsai and Moschetti, 2010) bases its theoretical foundation on the precondition of a scalar wavefield which is stationary in both space and time. As detailed below, this technique does not require specific array geometries to compute phase velocities and can be used on single pairs of stations as long as you are in the presence of an isotropic incident wavefield. Another advantage concerns the capability to resolve discrete frequencies on a potential wider range than what is possible using f-k methods as the Aki method produces robust and unbiased measurements at distances smaller than two wavelengths (Ekström et al., 2009). The major limit is set by the seismic wavefield characteristics.

Aki (1957) states that the azimuthally averaged normalized cross-spectrum $S(\Delta,\omega_0)$ for a receiver separation Δ and frequency ω_0 varies as J_0, the zero-order Bessel function of the first kind

$$S(\Delta,\omega_0) = J_0\left(\frac{\omega_0}{c(\omega_0)}\Delta\right) \tag{B3}$$

This relation suggests that the dispersion curve of phase velocities can be obtained from the fit of a J_0 Bessel function to the cross-spectrum obtained on a loop of receivers of same radius Δ, or also, as demonstrated by Cox (1973), to the cross-spectrum obtained for a single station pair if computed on an azimuthally isotropic wavefield.

Ekström et al. (2009) successfully obtain phase velocity estimates at discrete frequencies from ambient noise cross-correlations, by associating the zero-crossing of the real part of the data cross-spectrum with zeros of a Bessel function following equation B3. Preiswerk and Walter (2018) and Lindner et al. (2018) both use this method to obtain dispersion curves of Rayleigh wave speeds from cross-correlations of on-ice seismic records.

Application of this method is presented in Fig. B3 for one cross-correlation function obtained at Argentière. Because of possible noise contained in the correlation time-serie, the cross-spectrum is first smoothed to avoid any extra zero-crossing measurement. As there is a possibility of having missed one or several zero-crossings (indicated by black squares in b), several dispersion curves are generated (black dashed lines in c). The correct one still needs to be identified by judging the plausibility of the results given the expected velocities of the propagation medium (Ekström et al., 2009).

The dispersion curve estimation can be refined by fitting the entire cross-spectrum with a Bessel function, instead of fitting the zero-crossings only. We develop an approach similar to Menke and Jin (2015) who employ a grid search to generate

an initial estimate of the phase velocity that matches the observed cross-spectrum and then use a generalized least-squares procedure to refine this initial estimate.

The prior dispersion curve has to be as close as possible to the measured dispersion curve to avoid cycle skipping during the fitting. In our procedure, we take as a starting model, the average dispersion curve obtained from f-k analysis of the cross-correlation section computed at Argentière array. For sparse network configuration (as in Greenland and Gornergletscher, i.e. Fig. 9d and 11c), we take the theoretical dispersion curve computed with Geopsy software and based on the prior knowledge we have on the subsurface. The dispersion curves are then modelled as polynomial functions to enforce their smoothness and to reduce the number of fitting parameters and help the convergence of the fitting process. The order of the polynomial can be varied. Usually a polynomial order of 5 is a good compromise between smoothness and complexity of the dispersion curve. We then use a least-square inversion procedure of the polynomial coefficients to estimate the dispersion curve which best reproduces the observed cross-spectrum.

Fig. B3 shows the output of this procedure which yields to the same dispersion curve (red line in c) as the most probable one computed by fitting the zero-crossings. The overall-fitting method is particularly efficient for estimating more accurate phase velocities than with the zero-crossing fit, when considering correlation functions with low SNR (Menke and Jin, 2015). However it is particularly sensitive to the frequency range in which the least-square inversion is performed. When considering cross-correlation functions computed in narrow frequency bands as in Sect. 4 and 5, the method introduces strong side effects near the frequency corner limits due to filtering (Fig. 9d and 11c). The cross-spectrum must then be fitted considering carefully selected frequency components.

In the example shown in Fig. B3, the cross-correlation function is computed for frequencies 1-25 Hz. The Bessel fitting method is applied to frequencies above 2 Hz and enables to widen the velocity estimates down to 3 Hz (that is the frequency at which the inter-station distance is approximately equal to one wavelength) when compared to the zero-crossing output.

B3 Slant-stack technique on discrete sources

This method employed by Walter et al. (2015) can only be applied on cross-correlation functions that are computed on discrete sources (i.e. icequakes). It was used to obtain the velocity from ICC at Glacier d'Argentière in section 3.3. We here exploit the phase time-difference in the arrival times of Rayleigh waves with respect to the source position, and reproduce the azimuthal variations of phase times assuming a constant velocity.

The plane wave approximation implies a sinusoidal dependence of the arrival times which depend on the source azimuth and propagation velocity c as $\Delta \cos(\theta)/c$ with θ here defined as the source azimuth relative to the station pair axis (Fig. B4 a-b). We call the endfire lobes the two areas aligned with the receiver direction, in which the phase of the correlation function is stationary with respect to azimuth. The angular aperture of the endfire lobes depends on the ratio between the seismic wavelength λ and the station separation as $\delta\theta = \sqrt{\lambda/\Delta}$ (Roux et al., 2004; Gouédard et al., 2008b).

To measure seismic velocities at one station path and at different frequencies, we filter the individual correlation functions computed for each event to octave-wide frequency ranges. The lower frequency we can resolve is determined by the icequake

spectral content and is most importantly related to the station separation as we require that at least two wavelengths are sampled. We then restrict the analysis to 15-30 Hz.

We assign all cross-correlations to event azimuth bins of 5° to minimize the effects of location errors. For each trace, the arrival time of the Rayleigh wave is measured as the maximum of the correlation function computed at each frequency. We then invert the best sinusoide fit to the times of the maxima (in the least-square sense). The best solution gives the velocity estimate at the central frequency of the spectral band.

The velocity solution estimated by this method is naturally averaged over the azimuth range and can only be considered as the average velocity in the presence of strong azimuthal anisotropy which implies azimuthal variations of propagation velocities (Sect. 3.3). Nevertheless, the least-square solution fits very well the Rayleigh wave arrival times in the azimuth range of the stationary phase zones (Fig. B4b), and then is considered to represent well the propagation medium between the two stations.

To minimize the effects of location errors or low SNR of the correlation components, we perform a jackknife test on randomly selected events to fit the sinusoide. We require that the maximum of the cross-correlation stacked over bootstrap samples (i.e. selected correlation functions that have been shifted by the inverted arrival times prior stacking) exceeds 0.7. We require that at least 10 azimuth bins including the endfire lobes are considered in the sinusoidal fit. The final velocity at each frequency is then averaged over 200 jackknife tests.

References

Aki, K.: Space and Time Spectra of Stationary Stochastic Waves, with Special Reference to Microtremors, Bulletin of the Earthquake Research Institute, 35, 415–456, 1957.

Aki, K. and Chouet, B.: Origin of coda waves: source, attenuation, and scattering effects, Journal of geophysical research, 80, 3322–3342, 1975.

Allen, R. V.: Automatic earthquake recognition and timing from single traces, Bulletin of the Seismological Society of America, 68, 1521–1532, 1978.

Amundson, J. M., Fahnestock, M., Truffer, M., Brown, J., Lüthi, M. P., and Motyka, R. J.: Ice mélange dynamics and implications for terminus stability, Jakobshavn Isbræ, Greenland, Journal of Geophysical Research: Earth Surface, 115, 2010.

Andrews, L. C., Catania, G. A., Hoffman, M. J., Gulley, J. D., Lüthi, M. P., Ryser, C., Hawley, R. L., and Neumann, T. A.: Direct observations of evolving subglacial drainage beneath the Greenland Ice Sheet, Nautre, 514, 80–83, 2014.

Aster, R. and Winberry, J.: Glacial seismology, Reports on Progress in Physics, 80, 126 801, 2017.

Baggeroer, A. B., Kuperman, W. A., and Mikhalevsky, P. N.: An overview of matched field methods in ocean acoustics, IEEE J. Oceanic Eng., 18, 401–424, 1993.

Barruol, G., Cordier, E., Bascou, J., Fontaine, F. R., Legrésy, B., and Lescarmontier, L.: Tide-induced microseismicity in the Mertz glacier grounding area, East Antarctica, Geophysical Research Letters, 40, 5412–5416, 2013.

Bartholomaus, T., Larsen, C., O'Neel, S., and West, M.: Calving seismicity from iceberg–sea surface interactions, Journal of Geophysical Research: Earth Surface (2003–2012), 117, 2012.

Bartholomaus, T. C., Amundson, J. M., Walter, J. I., O'Neel, S., West, M. E., and Larsen, C. F.: Subglacial discharge at tidewater glaciers revealed by seismic tremor, Geophysical research letters, 42, 6391–6398, 2015.

Bensen, G. D., Ritzwoller, M. H., Barmin, M. P., Levshin, A. L., Lin, F., Moschetti, M. P., Shapiro, N. M., and Yang, Y.: Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, Geophysical Journal International, 169, 1239–1260, 2007.

Bienvenu, G. and Kopp, L.: Adaptivity to background noise spatial coherence for high resolution passive methods, ICASSP '80. IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 307–310, 1980.

Bonnefoy-Claudet, S., Cotton, F., and Bard, P.-Y.: The nature of noise wavefield and its applications for site effects studies: A literature review, Earth-Science Reviews, 79, 205–227, 2006.

Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N., Nadeau, R., and Larose, E.: Postseismic relaxation along the San Andreas fault at Parkfield from continuous seismological observations, Science, 321, 1478–1481, 2008a.

Brenguier, F., Shapiro, N., Campillo, M., Ferrazzini, V., Duputel, Z., Coutant, O., and Nercessian, A.: Towards forecasting volcanic eruptions using seismic noise, Nature Geoscience, 1, 126–130, 2008b.

Brenguier, F., Clarke, D., Aoki, Y., Shapiro, N. M., Campillo, M., and Ferrazzini, V.: Monitoring volcanoes using seismic noise correlations, Comptes Rendus Geoscience, 343, 633–638, 2011.

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

Campillo, M., Roux, P., Romanowicz, B., and Dziewonski, A.: Seismic imaging and monitoring with ambient noise correlations, Treatise on Geophysics, 1, 256–271, 2014.

Canassy, P. D., Faillettaz, J., Walter, F., and Huss, M.: Seismic activity and surface motion of a steep temperate glacier: a study on Triftgletscher, Switzerland, Journal of Glaciology, 58, 513–528, 2012.

Capon, J.: High-resolution frequency-wavenumber spectrum analysis, Proceedings of the IEEE, 57, 1408–1418, 1969.

Chaput, J., Campillo, M., Aster, R., Roux, P., Kyle, P., Knox, H., and Czoski, P.: Multiple scattering from icequakes at Erebus volcano, Antarctica: Implications for imaging at glaciated volcanoes, Journal of Geophysical Research: Solid Earth, 120, 1129–1141, 2015a.

Chaput, J., Clerc, V., Campillo, M., Roux, P., and Knox, H.: On the practical convergence of coda-based correlations: a window optimization approach, Geophysical Journal International, 204, 736–747, 2015b.

Chmiel, M., Roux, P., and Bardainne, T.: Attenuation of Seismic Noise in Microseismic Monitoring from Surface Acquisition, 77th EAGE Conference and Exhibition, 2015.

Chmiel, M., Roux, P., and Bardainne, T.: Extraction of phase and group velocities from ambient surface noise in a patch-array configuration, Geophysics, 81, KS231–KS240, 2016.

Chmiel, M., Roux, P., and Bardainne, T.: High-sensitivity microseismic monitoring: automatic detection and localization of subsurface noise sources using matched-field processing and dense path arrays, Geophysics, 2019.

Colombi, A., Boschi, L., Roux, P., and Campillo, M.: Green's function retrieval through cross-correlations in a two-dimensional complex reverberating medium, The Journal of the Acoustical Society of America, 135, 1034–1043, 2014.

Corciulo, M., Roux, P., Campillo, M., Dubucq, D., and Kuperman, W. A.: Multiscale matched-field processing for noise-source localization in exploration geophysics, Geophysics, 77, KS33–KS41, 2012.

Cox, H.: Spatial correlation in arbitrary noise fields with application to ambient sea noise, The Journal of the Acoustical Society of America, 54, 1289–1301, 1973.

Cox, H.: Multi-rate adaptive beamforming (MRABF), Sensor Array and Multichannel Signal Processing Workshop. Proceedings of the 2000 IEEE, pp. 306–309, 2000.

Cros, E., P., R., Vandemeulebrouck, J., and S., K.: Locating hydrothermal acoustic sources at Old Faithful Geyser using matched field processing, Geophysical Journal International, 187, 385–393, 2011.

De Rosny, J. and Roux, P.: Multiple scattering in a reflecting cavity: Application to fish counting in a tank, The Journal of the Acoustical Society of America, 109, 2587–2597, 2001.

Deichmann, N., Ansorge, J., Scherbaum, F., Aschwanden, A., Bernard, F., and Gudmundsson, G.: Evidence for deep icequakes in an Alpine glacier, Annals of Glaciology, 31, 85–90, 2000.

Diez, A. and Eisen, O.: Seismic wave propagation in anisotropic ice-Part 1: Elasticity tensor and derived quantities from ice-core properties, The Cryosphere, 9, 367–384, 2015.

Ekström, G.: Time domain analysis of Earth's long-period background seismic radiation, Journal of Geophysical Research: Solid Earth, 106, 26 483–26 493, 2001.

Ekström, G., Nettles, M., and Abers, G. A.: Glacial earthquakes, Science, 302, 622–624, 2003.

Ekström, G., Abers, G. A., and Webb, S. C.: Determination of surface-wave phase velocities across USArray from noise and Aki's spectral formulation, Geophysical Research Letters, 36, 2009.

Ermert, L., Villaseñor, A., and Fichtner, A.: Cross-correlation imaging of ambient noise sources, Geophysical Journal International, 204, 347–364, 2015.

Fichtner, A., Stehly, L., Ermert, L., and Boehm, C.: Generalised interferometry-I: Theory for inter-station correlations, Geophysical Journal International, 208, 603–638, 2017.

Gerstoft, P., Menon, R., Hodgkiss, W., and Mecklenbräuker, C.: Eigenvalues of the sample covariance matrix for a towed array, J. acoust. Soc. Am., 132, 2388–2396, 2012.

Gimbert, F., Tsai, V. C., and Lamb, M. P.: A physical model for seismic noise generation by turbulent flow in rivers, Journal of Geophysical Research: Earth Surface, 119, 2209–2238, 2014.

Gimbert, F., Tsai, V. C., Amundson, J. M., Bartholomaus, T. C., and Walter, J. I.: Subseasonal changes observed in subglacial channel pressure, size, and sediment transport, Geophysical Research Letters, 43, 3786–3794, 2016.

Gouédard, P., Cornou, C., and Roux, P.: Phase-velocity dispersion curves and small-scale geophysics using noise correlation slantstack technique, Geophysical Journal International, 172, 971–981, 2008a.

Gouédard, P., Roux, P., Campillo, M., and Verdel, A.: Convergence of the two-point correlation function toward the Green's function in the context of a seismic-prospecting data set, Geophysics, 73, V47–V53, 2008b.

Gradon, C., Moreau, L., Roux, P., and Ben-Zion, Y.: Analysis of surface and seismic sources in dense array data with match field processing and Markov chain Monte Carlo sampling, Geophysical Journal International, 218, 1044–1056, 2019.

Gräff, D., Walter, F., and Lipovsky, B. P.: Crack wave resonances within the basal water layer, Annals of Glaciology, pp. 1–9, 2019.

Hadziioannou, C., Larose, E., Coutant, O., Roux, P., and Campillo, M.: Stability of monitoring weak changes in multiply scattering media with ambient noise correlation: Laboratory experiments, The Journal of the Acoustical Society of America, 125, 3688–3695, 2009.

Hand, E.: A boom in boomless seismology, 2014.

Haney, M. M. and Tsai, V. C.: Perturbational and nonperturbational inversion of Rayleigh-wave velocities, Geophysics, 82, F15–F28, 2017.

Hantz, D.: Dynamique et hydrologie du glacier d'Argentière (Alpes françaises), Ph.D. thesis, Université Scientifique et Médicale de Grenoble, 1981.

Harland, S., Kendall, J.-M., Stuart, G., Lloyd, G., Baird, A., Smith, A., Pritchard, H., and Brisbourne, A.: Deformation in Rutford Ice Stream, West Antarctica: measuring shear-wave anisotropy from icequakes, Annals of Glaciology, 54, 105–114, 2013.

Hennino, R., Trégourès, N., Shapiro, N., Margerin, L., Campillo, M., Van Tiggelen, B., and Weaver, R.: Observation of equipartition of seismic waves, Physical review letters, 86, 3447, 2001.

Horgan, H. J., Anandakrishnan, S., Alley, R. B., Burkett, P. G., and Peters, L. E.: Englacial seismic reflectivity: Imaging crystal-orientation fabric in West Antarctica, Journal of Glaciology, 57, 639–650, 2011.

Konda, T. and Nakamura, Y.: A new algorithm for singular value decomposition and its parallelization, Parallel Computing, 2009.

Kuperman, W. A. and Turek, G.: Matched field acoustics, Mechanical Systems and Signal Processing, 11, 141–148, 1997.

Larose, E., Roux, P., Campillo, M., and Derode, A.: Fluctuations of correlations and Green's function reconstruction: role of scattering, Journal of Applied Physics, 103, 114 907, 2008.

Larose, E., Carrière, S., Voisin, C., Bottelin, P., Baillet, L., Guéguen, P., Walter, F., Jongmans, D., Guillier, B., Garambois, S., et al.: Environmental seismology: What can we learn on earth surface processes with ambient noise?, Journal of Applied Geophysics, 116, 62–74, 2015.

Lecocq, T., Caudron, C., and Brenguier, F.: MSNoise, a python package for monitoring seismic velocity changes using ambient seismic noise, Seismological Research Letters, 85, 715–726, 2014.

Levshin, A., Ratnikova, L., and Berger, J.: Peculiarities of surface-wave propagation across central Eurasia, Bulletin of the Seismological Society of America, 82, 2464–2493, 1992.

Lin, F.-C., Moschetti, M. P., and Ritzwoller, M. H.: Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, Geophysical Journal International, 173, 281–298, 2008.

Lin, F.-C., Li, D., Clayton, R. W., and Hollis, D.: High-resolution 3D shallow crustal structure in Long Beach, California: Application of ambient noise tomography on a dense seismic array, Geophysics, 78, Q45–Q56, 2013.

Lin, F.-C., Tsai, V. C., and Schmandt, B.: 3-D crustal structure of the western United States: application of Rayleigh-wave ellipticity extracted from noise cross-correlations, Geophysical Journal International, 198, 656–670, 2014.

Lindner, F., Weemstra, C., Walter, F., and Hadziioannou, C.: Towards monitoring the englacial fracture state using virtual-reflector seismology, Geophysical Journal International, 214, 825–844, 2018.

Lindner, F., Laske, G., Walter, F., and Doran, A. K.: Crevasse-induced Rayleigh-wave azimuthal anisotropy on Glacier de la Plaine Morte, Switzerland, Annals of Glaciology, 60, 96–111, 2019.

Lindner, F., Walter, F., Laske, G., and Gimbert, F.: Glaciohydraulic seismic tremors on an Alpine glacier, The Cryosphere Discussion, 14, 28–308, https://doi.org/https://doi.org/10.5194/tc-2019-155, 2020.

Lipovsky, B. P., Meyer, C. R., Zoet, L. K., McCarthy, C., Hansen, D. D., Rempel, A. W., and Gimbert, F.: Glacier sliding, seismicity and sediment entrainment, Annals of Glaciology, p. 1–11, https://doi.org/10.1017/aog.2019.24, 2019.

Lobkis, O. I. and Weaver, R. L.: On the emergence of the Green's function in the correlations of a diffuse field, The Journal of the Acoustical Society of America, 110, 3011–3017, 2001.

Mainsant, G., Larose, E., Brönnimann, C., Jongmans, D., Michoud, C., and Jaboyedoff, M.: Ambient seismic noise monitoring of a clay landslide: Toward failure prediction, Journal of Geophysical Research: Earth Surface, 117, 2012.

Malcolm, A. E., Scales, J. A., and van Tiggelen, B. A.: Extracting the Green function from diffuse, equipartitioned waves, Physical Review E, 70, 015 601, 2004.

Menke, W. and Jin, G.: Waveform fitting of cross spectra to determine phase velocity using Aki's formula, Bulletin of the Seismological Society of America, 105, 1619–1627, 2015.

Mikesell, T., Wijk, K., Haney, M. M., Bradford, J. H., Marshall, H.-P., and Harper, J. T.: Monitoring glacier surface seismicity in time and space using Rayleigh waves, Journal of Geophysical Research: Earth Surface, 117, 2012.

Moonen, M., van Dooren, P., and Vandewalle, J.: Singular value decomposition updating algorithm for subspace tracking, SIAM Journal on Matrix Analysis and Applications, 1992.

Mordret, A., Shapiro, N. M., Singh, S., Roux, P., Montagner, J.-P., and Barkved, O. I.: Azimuthal anisotropy at Valhall: The Helmholtz equation approach, Geophysical Research Letters, 40, 2636–2641, 2013.

Mordret, A., Mikesell, T., Harig, C., Lipovsky, B., and Prieto, G.: Monitoring southwest Greenland's ice sheet melt with ambient seismic noise, Science advances, 2, e1501 538, 2016.

Moreau, L., Stehly, L., Boué, P., Lu, Y., Larose, E., and Campillo, M.: Improving ambient noise correlation functions with an SVD-based Wiener filter, Geophysical Journal International, 211, 418–426, 2017.

Mulargia, F.: The seismic noise wavefield is not diffuse, The Journal of the Acoustical Society of America, 131, 2853–2858, 2012.

Nakata, N., Chang, J. P., Lawrence, J. F., and Boué, P.: Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry, Journal of Geophysical Research: Solid Earth, 120, 1159–1173, 2015.

Nanni, U., Gimbert, F., Roux, P., Urruty, B., and Lecointre, A.: Mapping the Subglacial Drainage System from Dense Array Seismology: a Multi-method Approach, AGU Fall Meeting, 2019a.

Nanni, U., Gimbert, F., Vincent, C., Gräff, D., Walter, F., Piard, L., and Moreau, L.: Seasonal and Diurnal Dynamics of Subglacial Channels: Observations Beneath an Alpine Glacier, The Cryosphere Discussions, https://doi.org/https://doi.org/10.5194/tc-2019-243, 2019b.

Neave, K. and Savage, J.: Icequakes on the Athabasca glacier, Journal of Geophysical Research, 75, 1351–1362, 1970.

Nunn, C.: Apollo Passive Seismic Experiments, International Federation of Digital Seismograph Networks, https://doi.org/10.7914/SN/XA_1969, 2017.

Nye, J.: The mechanics of glacier flow, Journal of Glaciology, 2, 82–93, 1952.

Ohrnberger, M., Schissele, E., Cornou, C., Bonnefoy-Claudet, S., Wathelet, M., Savvaidis, A., Scherbaum, F., and Jongmans, D.: Frequency wavenumber and spatial autocorrelation methods for dispersion curve determination from ambient vibration recordings, in: Proceedings of the 13th World Conference on Earthquake Engineering, vol. 946, Vancouver Canada, 2004.

Olivier, G., Brenguier, F., Campillo, M., Roux, P., Shapiro, N., and Lynch, R.: Investigation of coseismic and postseismic processes using in situ measurements of seismic velocity variations in an underground mine, Geophysical Research Letters, 42, 9261–9269, 2015.

Park, C. B., Miller, R. D., and Xia, J.: Imaging dispersion curves of surface waves on multi-channel record, in: SEG Technical Program Expanded Abstracts 1998, pp. 1377–1380, Society of Exploration Geophysicists, 1998.

Paul, A., Campillo, M., Margerin, L., Larose, E., and Derode, A.: Empirical synthesis of time-asymmetrical Green functions from the correlation of coda waves, Journal of Geophysical Research: Solid Earth, 110, 2005.

Picotti, S., Vuan, A., Carcione, J. M., Horgan, H. J., and Anandakrishnan, S.: Anisotropy and crystalline fabric of Whillans Ice Stream (West Antarctica) inferred from multicomponent seismic data, Journal of Geophysical Research: Solid Earth, 120, 4237–4262, 2015.

Picotti, S., Francese, R., Giorgi, M., Pettenati, F., and Carcione, J.: Estimation of glacier thicknesses and basal properties using the horizontal-to-vertical component spectral ratio (HVSR) technique from passive seismic data, Journal of Glaciology, 63, 229–248, 2017.

Podolskiy, E. and Walter, F.: Cryo-seismology, Reviews of Geophysics, 2016.

Podolskiy, E. A., Genco, R., Sugiyama, S., Walter, F., Funk, M., Minowa, M., Tsutaki, S., and Ripepe, M.: Seismic and infrasound monitoring of Bowdoin Glacier, Greenland, Low Temperature Science, 75, 15–36, 2017.

Podolskiy, E. A., Fujita, K., Sunako, S., and Sato, Y.: Viscoelastic modeling of nocturnal thermal fracturing in a Himalayan debris-covered glacier, Journal of Geophysical Research: Earth Surface, https://doi.org/10.1029/2018JF004848, 2019.

Preiswerk, L., Walter, F., Anandakrishnan, S., Barfucci, G., and others.: Monitoring unstable parts in the ice-covered Weissmies northwest face, in: 13th Congress Interpraevent 2016, ETH-Zürich, 2016.

Preiswerk, L. E. and Walter, F.: High-Frequency (>2 Hz) Ambient Seismic Noise on High-Melt Glaciers: Green's Function Estimation and Source Characterization, Journal of Geophysical Research: Earth Surface, https://doi.org/10.1029/2017JF004498, 2018.

Preiswerk, L. E., Michel, C., Walter, F., and Fäh, D.: Effects of geometry on the seismic wavefield of Alpine glaciers, Annals of Glaciology, pp. 1–13, 2018.

Rhie, J. and Romanowicz, B.: Excitation of Earth's continuous free oscillations by atmosphere–ocean–seafloor coupling, Nature, 431, 552, 2004.

Röösli, C., Walter, F., Husen, S., Andrews, L., Lüthi, M., Catania, G., and Kissling, E.: Sustained seismic tremors and icequakes detected in the ablation zone of the Greenland ice sheet, Journal of Glaciology, 60, 563–575, 2014.

Röösli, C., Helmstetter, A., Walter, F., and Kissling, E.: Meltwater influences on deep stick-slip icequakes near the base of the Greenland Ice Sheet, Journal of Geophysical Research: Earth Surface, 121, 223–240, 2016a.

Röösli, C., Walter, F., Ampuero, J.-P., and Kissling, E.: Seismic moulin tremor, Journal of Geophysical Research: Solid Earth, 121, 5838–5858, 2016b.

Rost, S. and Thomas, C.: Array seismology: Methods and applications, Reviews of geophysics, 40, 2–1, 2002.

Roux, P., Kuperman, W., and Group, N.: Extracting coherent wave fronts from acoustic ambient noise in the ocean, The Journal of the Acoustical Society of America, 116, 1995–2003, 2004.

Roux, P.-F., Marsan, D., Métaxian, J.-P., O'Brien, G., and Moreau, L.: Microseismic activity within a serac zone in an alpine glacier (Glacier d'Argentiere, Mont Blanc, France), Journal of Glaciology, 54, 157–168, 2008.

Roux, P.-F., Walter, F., Riesen, P., Sugiyama, S., and Funk, M.: Observation of surface seismic activity changes of an Alpine glacier during a glacier-dammed lake outburst, Journal of Geophysical Research: Earth Surface, 115, 2010.

Ryser, C., Lüthi, M., Andrews, L., Hoffman, M., Catania, G., Hawley, R., Neumann, T., and Kristensen, S.: Sustained high basal motion of the Greenland ice sheet revealed by borehole deformation, Journal of Glaciology, 60, 647, 2014.

Sabra, K. G., Gerstoft, P., Roux, P., Kuperman, W., and Fehler, M. C.: Surface wave tomography from microseisms in Southern California, Geophysical Research Letters, 32, 2005.

Sadek, R. A.: SVD Based Image Processing Applications: State of The Art, Contributions and Research Challenges, International Journal of Advanced Computer Science and Applications, 3, 2012.

Schmandt, B., Aster, R. C., Scherler, D., Tsai, V. C., and Karlstrom, K.: Multiple fluvial processes detected by riverside seismic and infrasound monitoring of a controlled flood in the Grand Canyon, Geophysical Research Letters, 40, 4858–4863, 2013.

Sens-Schönfelder, C. and Wegler, U.: Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia, Geophysical research letters, 33, 2006.

Sergeant, A., Stutzmann, E., Maggi, A., Schimmel, M., Ardhuin, F., and Obrebski, M.: Frequency-dependent noise sources in the North Atlantic Ocean, Geochemistry, Geophysics, Geosystems, 14, 5341–5353, 2013.

Sergeant, A., Mangeney, A., Stutzmann, E., Montagner, J., Walter, F., Moretti, L., and Castelnau, O.: Complex force history of a calving-generated glacial earthquake derived from broadband seismic inversion, Geophysical Research Letters, 43, 1055–1065, 2016.

Sergeant, A., Yastrebov, V. A., Mangeney, A., Castelnau, O., Montagner, J.-P., and Stutzmann, E.: Numerical modeling of iceberg capsizing responsible for glacial earthquakes, Journal of Geophysical Research: Earth Surface, 123, 3013–3033, 2018.

Seydoux, L., Shapiro, N., de Rosny, J., Brenguier, F., and Landès, M.: Detecting seismic activity with a covariance matrix analysis of data recorded on seismic arrays, Geophysical Journal International, 204, 1430–1442, 2016.

Seydoux, L., Rosny, J., and Shapiro, N. M.: Pre-processing ambient noise cross-correlations with equalizing the covariance matrix eigenspectrum, Geophysical Journal International, 210, 1432–1449, 2017.

Shapiro, N., Campillo, M., Margerin, L., Singh, S., Kostoglodov, V., and Pacheco, J.: The energy partitioning and the diffusive character of the seismic coda, Bulletin of the Seismological Society of America, 90, 655–665, 2000.

Shapiro, N. M. and Campillo, M.: Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, Geophysical Research Letters, 31, 2004.

Shapiro, N. M., Campillo, M., Stehly, L., and Ritzwoller, M. H.: High-resolution surface-wave tomography from ambient seismic noise, Science, 307, 1615–1618, 2005.

Smith, E. C., Baird, A. F., Kendall, J. M., Martín, C., White, R. S., Brisbourne, A. M., and Smith, A. M.: Ice fabric in an Antarctic ice stream interpreted from seismic anisotropy, Geophysical Research Letters, 44, 3710–3718, 2017.

Smith, M. L. and Dahlen, F.: The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium, Journal of Geophysical Research, 78, 3321–3333, 1973.

Snieder, R., Van Wijk, K., Haney, M., and Calvert, R.: Cancellation of spurious arrivals in Green's function extraction and the generalized optical theorem, Physical Review E, 78, 036 606, 2008.

Stein, S. and Wysession, M.: An introduction to seismology, earthquakes, and earth structure, John Wiley & Sons, 2009.

Toyokuni, G., Takenaka, H., Takagi, R., Kanao, M., Tsuboi, S., Tono, Y., Childs, D., and Zhao, D.: Changes in Greenland ice bed conditions inferred from seismology, Physics of the Earth and Planetary Interiors, 277, 81–98, 2018.

Tsai, V. C. and Moschetti, M. P.: An explicit relationship between time-domain noise correlation and spatial autocorrelation (SPAC) results, Geophysical Journal International, 182, 454–460, 2010.

Van der Veen, C.: Fracture mechanics approach to penetration of surface crevasses on glaciers, Cold Regions Science and Technology, 27, 31–47, 1998.

Vandemeulebrouck, J., Roux, P., and E., C.: The plumbing of Old Faithful Geyser revealed by hydrothermal tremor, Geophysical Research Letters, 40, 1989–1993, 2013.

Veen, B. V. and Buckley, K.: Beamforming: a versatile approach to spatial filtering, IEEE ASSP Mag., 5, 4–24, 1988.

Vore, M. E., Bartholomaus, T. C., Winberry, J. P., Walter, J. I., and Amundson, J. M.: Seismic tremor reveals spatial organization and temporal changes of subglacial water system, Journal of Geophysical Research: Earth Surface, 124, 427–446, 2019.

Walter, F.: Seismic activity on Gornergletscher during Gornersee outburst floods, Ph.D. thesis, ETH Zurich, 2009.

Walter, F., Deichmann, N., and Funk, M.: Basal icequakes during changing subglacial water pressures beneath Gornergletscher, Switzerland, Journal of Glaciology, 54, 511–521, 2008.

Walter, F., Clinton, J., Deichmann, N., Dreger, D. S., Minson, S., and Funk, M.: Moment tensor inversions of icequakes on Gornergletscher, Switzerland, Bulletin of the Seismological Society of America, 99, 852–870, 2009.

Walter, F., Dreger, D., Clinton, J., Deichmann, N., and Funk, M.: Evidence for near-horizontal tensile faulting at the base of Gornergletscher, a Swiss alpine glacier, Bulletin of the Seismological Society of America, 100, 458–472, 2010.

Walter, F., Roux, P., Röösli, C., Lecointre, A., Kilb, D., and Roux, P.: Using glacier seismicity for phase velocity measurements and Green's function retrieval, Geophysical Journal International, 201, 1722–1737, 2015.

Wang, J., Templeton, D., and Harris, D.: A New Method for Event Detection and Location - Matched Field Processing Application to the Salton Sea Geothermal Field, Search and Discovery Article #40946, 2012.

Wapenaar, K.: Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation, Physical review letters, 93, 254 301, 2004.

Wapenaar, K., Van Der Neut, J., Ruigrok, E., Draganov, D., Hunziker, J., Slob, E., Thorbecke, J., and Snieder, R.: Seismic interferometry by crosscorrelation and by multidimensional deconvolution: A systematic comparison, Geophysical Journal International, 185, 1335–1364, 2011.

Wathelet, M.: An improved neighborhood algorithm: parameter conditions and dynamic scaling, Geophysical Research Letters, 35, 2008.

Wathelet, M., Jongmans, D., Ohrnberger, M., and Bonnefoy-Claudet, S.: Array performances for ambient vibrations on a shallow structure and consequences over V s inversion, Journal of Seismology, 12, 1–19, 2008.

Wax, M. and Kailath, T.: Detection of signals by information theoretic criteria, IEEE Transactions on Acoustics, Speech, and Signal Processing, 33, 387–392, 1985.

Webb, S. C.: Broadband seismology and noise under the ocean, Reviews of Geophysics, 36, 105–142, 1998.

Weemstra, C., Wapenaar, K., and Van Dalen, K. N.: Reflecting boundary conditions for interferometry by multidimensional deconvolution, The Journal of the Acoustical Society of America, 142, 2242–2257, 2017.

Winberry, P., Anandakrishnan, S., Wiens, D., and Alley, R.: Nucleation and seismic tremor associated with the glacial earthquakes of Whillans Ice Stream, Antarctica, Geophysical Research Letters, 40, 312–315, 2013.

Yang, Y., Ritzwoller, M. H., Levshin, A. L., and Shapiro, N. M.: Ambient noise Rayleigh wave tomography across Europe, Geophysical Journal International, 168, 259–274, 2007.

Zhan, Z.: Seismic noise interferometry reveals transverse drainage configuration beneath the surging Bering Glacier, Geophysical Research Letters, 46, 4747–4756, 2019.

Zhan, Z., Tsai, V., Jackson, J., and Helmberger, D.: Ambient noise correlation on the Amery Ice Shelf, east Antarctica, Geophysical Journal International, 196, 1796–1802, 2013.

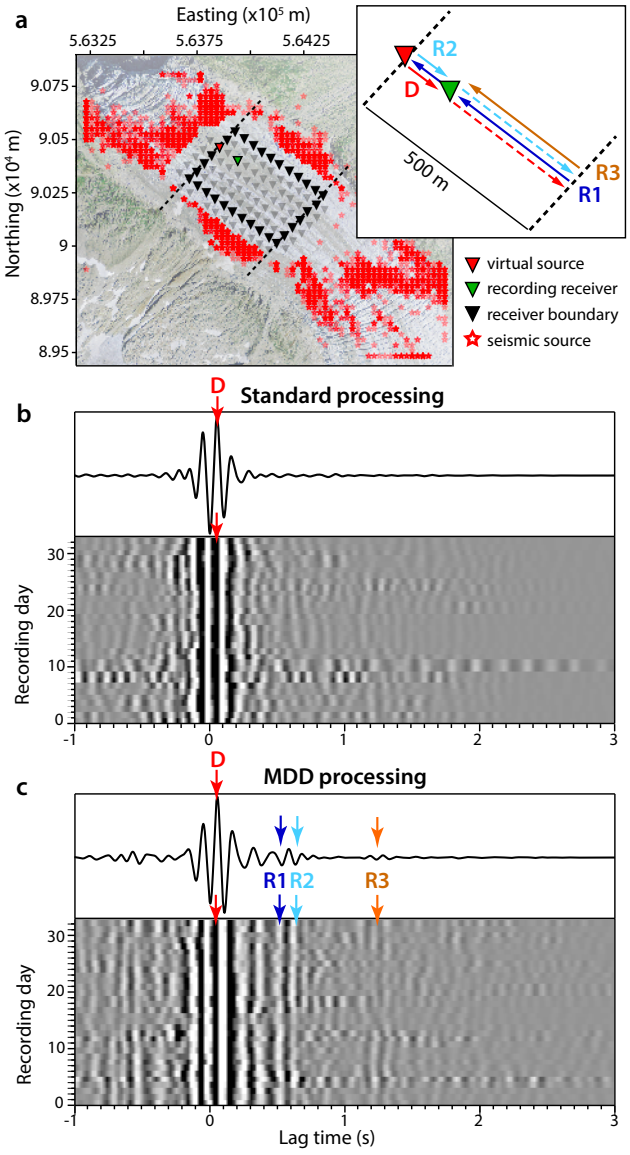

Figure 12. (a) Source–receiver geometry used for the computation of icequake virtual-source responses through the application of MDD (c) to the cross-correlations in (b). In this particular case, a receiver contour (consisting of 32 regularly-spaced receivers) and a single receiver between these receiver lines (green triangle) are illuminated by 4282 sources (i.e. icequakes) on either side of the receiver cavity. The receiver colored in red is here turned into a virtual source, whose response is recorded at the green receiver. Due to the reflecting boundary conditions and the receiver geometry, the emitted wave travels back and forth between the two receiver lines indicated with the black dashed lines. We obtain multiple reflections noted R_i (i indicates the number of virtual reflection), which are visible on the MDD correlation gather and averaged stack in (c). Virtual reflections R_i create an artificial coda after the ballistic Rayleigh wave reconstructed on the GF estimate. This coda is observed to be coherent through time (here we show cross-correlation stacks computed on two-day sliding window with an overlap of one day) and is suitable for CWI studies. (b) is the same as (c) but for standard processing of icequake cross-correlations at the two sensors in green and red.

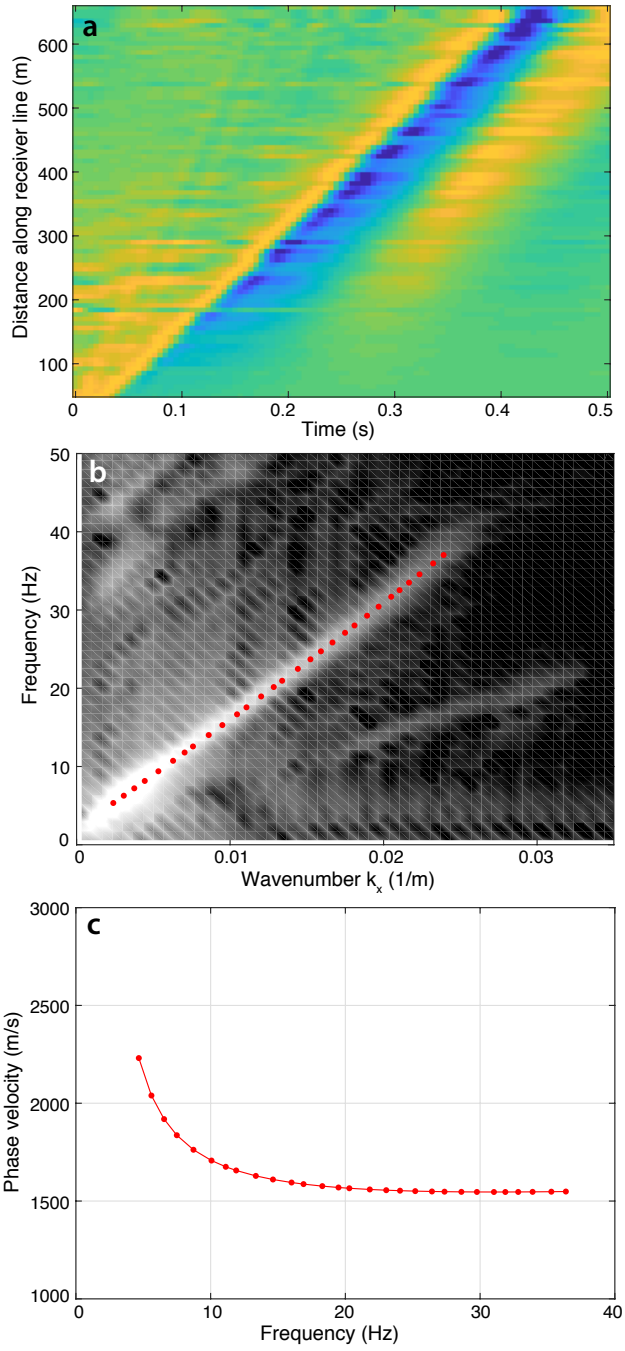

Figure B1. Computation of Rayleigh wave phase velocity from f-k analysis of the noise cross-correlation section obtained at Argentière array (a). (b) Frequency-wavenumber diagram obtained from the 2D Fourier transform of (a). Red dots show the peaks of energy maximum at each frequency, and correspond to the Rayleigh wave fundamental mode. (c) Phase velocity dispersion curve obtained from the interpolated peaks in (b).

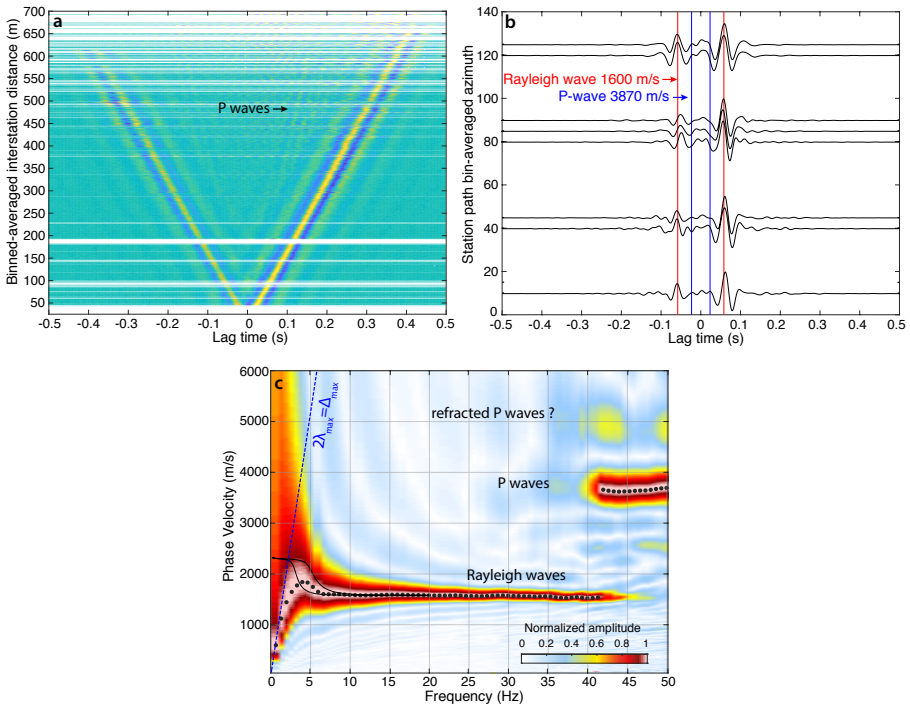

Figure B2. Same as Fig. 3 but for icequake-cross-correlations (a) computed in Glacier d'Argentière. (b) Frequency-velocity diagram obtained from f-k analysis of the correlation functions in (a) using the phase-shift method of Park et al. (1998). The extracted dispersion curve of Rayleigh and P-waves are plotted in black dots.

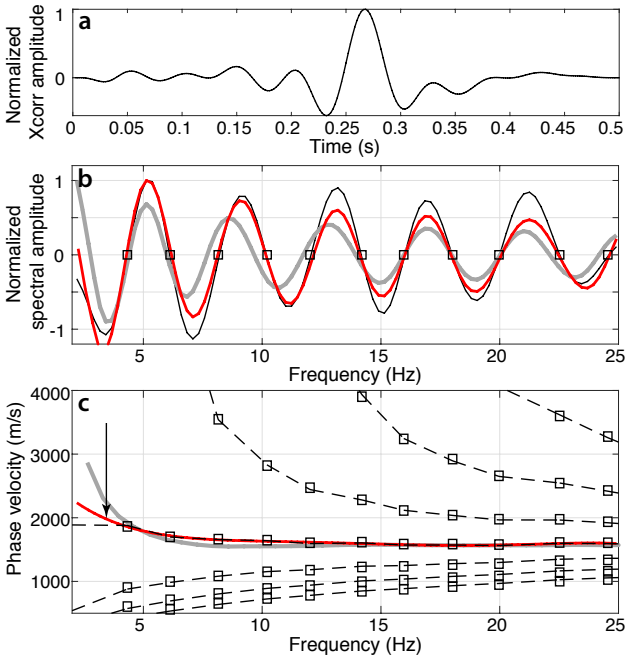

Figure B3. Computation of Rayleigh wave phase velocity using Aki's spectral method. (a) Symmetric icequake cross-correlation function obtained for one receiver pair in Glacier d'Argentière, with 450 m apart stations. (b) The real part of the spectrum of (a) is in black and associated zero-crossings are marked by squares. The gray line indicates the real part of the spectrum associated with a priori phase velocity dispersion curve which serves as a starting model for the least-square fit (in red) of the observations. (c) Corresponding phase velocities estimated by zero-crossings (black dashed lines) and least-square fit (red). The prior dispersion curve used for the Bessel fit is plotted in gray. The black arrow indicates the minimum frequency above which we can trust the velocity measurements and corresponds to approximately one wavelength.

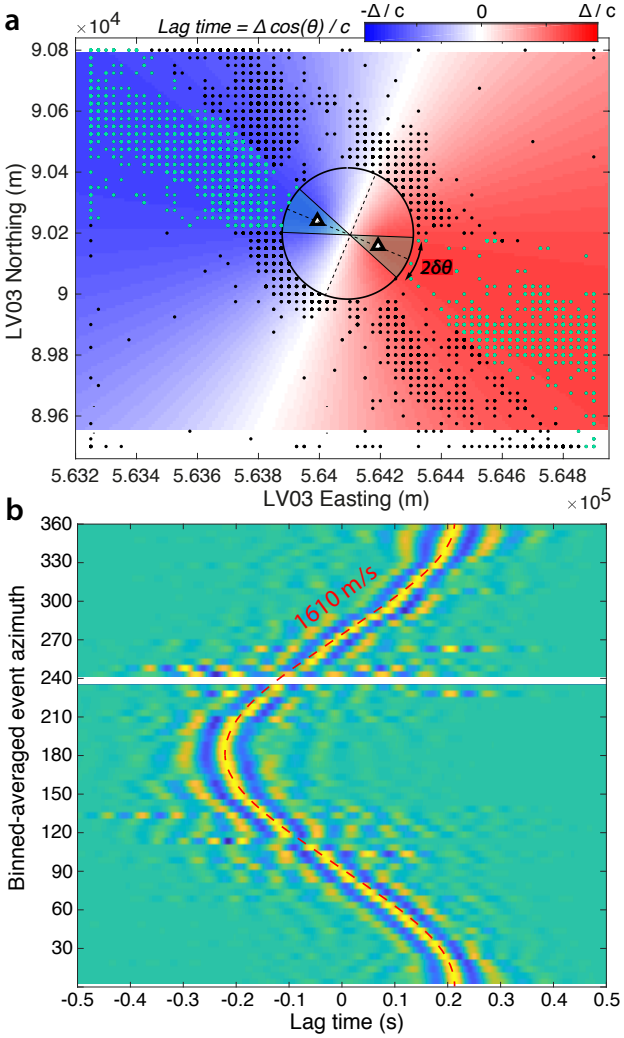

Figure B4. (a) Icequake locations (dots) in Glacier d'Argentière whose waveforms are cross-correlated at the two stations (triangles) to obtain (b). Green dots in (a) show the events that lie in the endfire lobes of aperture $2\delta\theta$ (see main text). (b) The icequake cross-correlation functions (here averaged in 5° azimuth bins) give coherent arrival times for Rayleigh waves traveling in the ice between the two stations, which are a function of the inter-station distance Δ and the event azimuth relative to the station path θ, as plotted in background colors in (a). The red dashed line shows the least-square fit of the arrival times with a sinusoidal function of phase velocity $c = 1610$ m/s for central frequency 15 Hz. Correlation functions are here filtered between 10-20 Hz.