Supplementary materials for

Numerical and observational study of *Sn*-to-*Lg* conversion due to crustal-thickening: implications for identification of continental mantle earthquakes

*Shiqi Wang**

Simon L. Klemperer

*Corresponding author: axelwang@stanford.edu

Department of Geophysics, Stanford University, Stanford, CA 94305, USA.

Table of contents

S1. Reconstruction of undulated geometry with AxiSEM3D

It is easy to use SalvusMeshLite (https://gitlab.com/swp_ethz/public/SalvusMeshLite) to generate a mesh to use with AxiSEM3D. However, these meshes can only be regular, meaning that they do not contain desired geometrical undulations on element boundaries (e.g. for a Cartesian domain, all element boundaries will be flat and for a spherical domain all element boundaries will be concentrically spherical). It also can be easy to generate an undulation file, called a geometric model in AxiSEM3D. For our case, the geometric model is the Moho ramp, and we only needed to specify it (distance to ramp start d, width of the ramp w, height of the ramp h, and method of undulation, a sigmoid function) as a function of simulation-domain latitude, assuming the source is at 90°. However, the geometrical model is quite disconnected with the finite element mesh. For example, even though the mesh is spherical, the geometrical model is represented linearly in simulation-domain latitude and the sampling of this geometrical model in simulation-domain latitude is usually different from the element nodes in the mesh (i.e. the Gauss-Lobatto-Legendre, GLL, points). Our geometrical model can be specified using many fewer points than there are element nodes as our mesh is necessarily extremely dense in order to sample waves up to 5Hz.

When AxiSEM3D is given the mesh and geometrical model files, the program requires the user to specify on which element boundary to deform the mesh; the surrounding area must also be specified to prevent over-deforming individual elements which may cause numerical instabilities. This is done for the GLL points within the deformation region by specifying a local perturbation value (the amount the GLL points need to be moved in the radial direction) as a Fourier series as a function of azimuthal angle to enable the representation of fully 3D models (Al-Attar &Crawford, 2016). The program only works with Fourier series coefficients that represent the local perturbations, but never actually calculates the perturbations. This approach is also used to represent all other properties in the model, even when the simulation domain is axisymmetric, because at least five Fourier coefficients are needed to represent the 3D point source.

Because a user only has access to the mesh and geometrical model files, but not the intermediate results from the internal computations, it is hard for a user to analyze the computed wavefield on a correctly deformed mesh and to check whether the internal program has represented the

geometrical model correctly with the GLL points, since the sampling of the two can be quite different.

To add these important capabilities, we save and output intermediate results from the program, and calculate the mesh deformation externally with Fourier interpolation. The intermediate results we output are the element tags for elements within the undulation region (from *Undulation.cpp*), the Fourier coefficients associated with these elements for undulations (from *Undulation.cpp*) and the coordinates of all GLL points (from *SE_MODEL.cpp* and *Quad.cpp*). We also output the element tag for the first element whose bottom edge is the start of the undulation surface (from *SE_MODEL.cpp*), and this allows us to extract the coordinates on this undulated surface exactly as AxiSEM3D sees it internally. This then allows plotting such a surface as shown in Figs. 1&6a-d and allows checking against the original geometrical model for accuracy of representation by AxiSEM3D (Fig. S1-1).

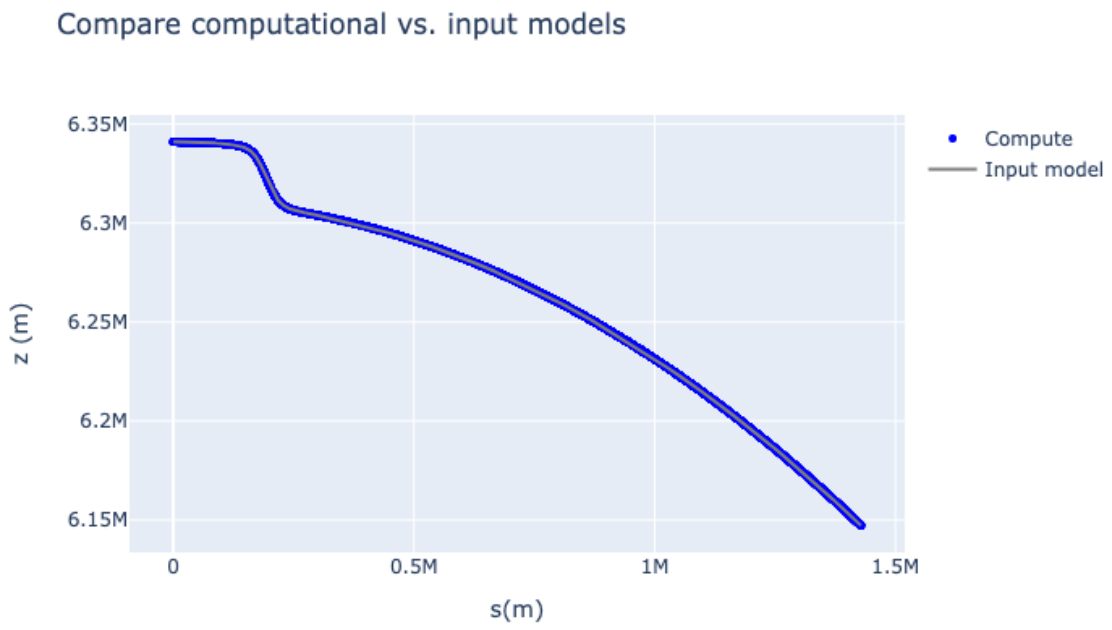

Fig. S1-1. Checking representation of undulation by AxiSEM3D. The input geometrical model is shown as a grey line, and its representation by GLL points used for computation is shown as blue dots. The match is exact. Vertical and horizontal axes are the z, vertical, and s, radial, cylindrical coordinates measured respectively from Earth's center and from the source. M represents 10^6. This example is shown for a Moho ramp with d =100, w =200 and h = 30 km.

S2. Mesh coarsening

We require the mesher to use a minimum of two elements (~10 GLL points) per wavelength in our simulation domain. By default, the mesher ensures this first with the smallest wavelength, which occurs in the upper crust when the shear wave velocity is the smallest. The rest of the mesh is then built honoring the position of discontinuities and the background geometry and velocity models. Only when the velocity at a deeper layer allows doubling the element size while still satisfying the two element-per-wavelength requirement does the mesher automatically coarsen the mesh through a two-refinement transition template (Fig. S2-1). For our specific case with the PREM background model, this doesn't automatically happen. The consequence is that although the upper crust is being sampled with 2 elements per wavelength, the deeper, faster layers are being sampled with more elements per wavelength, reaching a maximum of around 2.9 below the LVZ (Fig. S2-2). This number is still much lower than 4, which is needed to trigger automatic mesh coarsening.

The default mesh is not optimal in terms of computational cost, because a rather small area in the computational domain (i.e. upper crust) is being sufficiently sampled, whereas the rest of the domain is being over-sampled (Fig. S2-2). To combat this, we modified the shear wave velocities in the background model in the upper crust and at and below the depth where we wish the element size to become doubled. This depth is chosen at 70 km, i.e. 10 km below the deepest Moho in our models to avoid conflict with the Moho undulation. These velocities are modified so that the mesher will automatically introduce mesh coarsening at our desired depth, in such a way that when the real shear velocities are replaced back, the crust, a smaller portion of the domain, will now be over-sampled, whereas the majority of the simulation domain, below the coarsening depth, will be sufficiently sampled with ~2 elements per wavelength (Fig. S2-2). Implementing this strategy enabled us to use ~16.8% less elements for our simulations. We compare seismograms computed with both coarsened and uncoarsened meshes at short and long distances to verify that our mesh-coarsening technique didn't influence our results (Fig. S2-3).

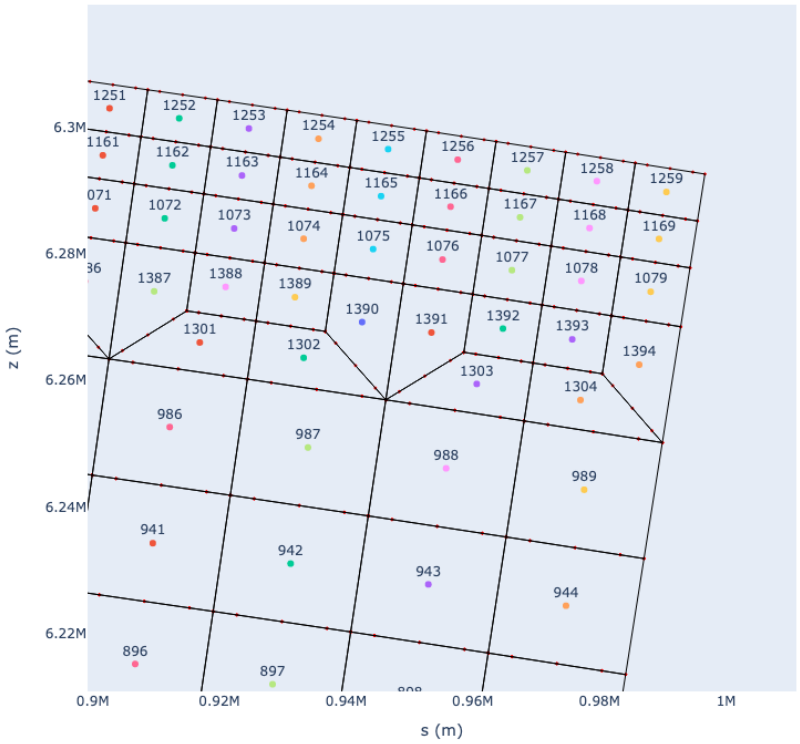

Fig. S2-1. Two-refinement transition templates. Example showing a low-resolution mesh (not used in our simulations) at an offset to demonstrate the way a mesh can be coarsened with SalvusMeshLite. The horizontal and vertical axes are the same as in Fig. S1-1. Numbers on each element are its tag number. A 2-refinement transition template consists of three elements such as elements 1387, 1388 and 1301. Faint dots on element edges are GLL points on element edges (those inside an element are not shown).

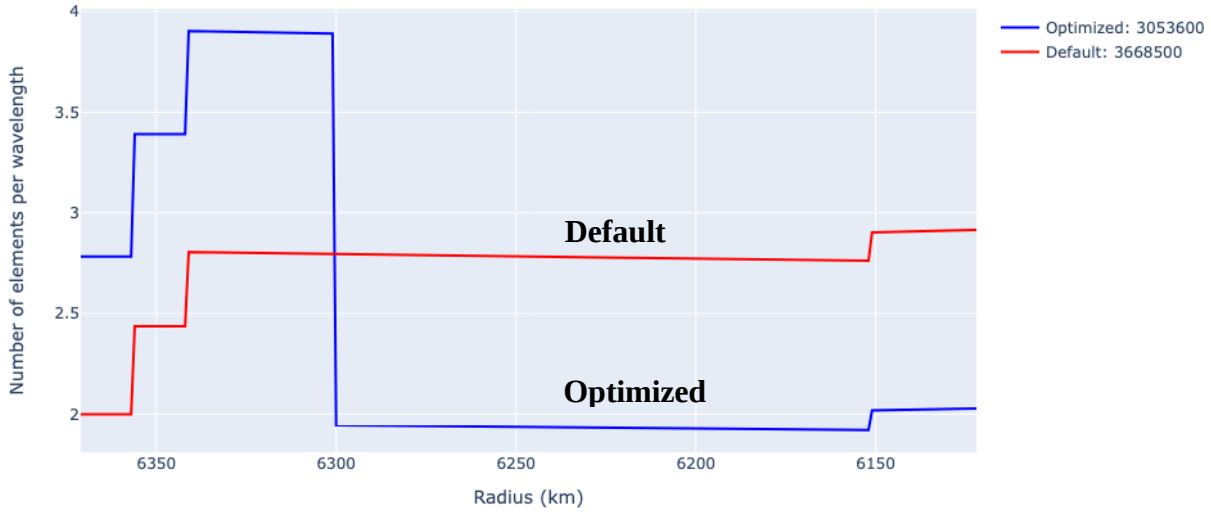

Fig. S2-2. Number of elements per wavelength for the coarsened and uncoarsened meshes. Two elements per wavelength is maintained by the default, uncoarsened mesh (red line), near the surface of the earth, with radius close to 6371 km, whereas this is honored by the optimized, coarsened mesh (blue line), at 6301 km radius, i.e. 70 km below below the surface. The latter is achieved at an expense of increasing the number of elements per wavelength above this depth. The total number of elements used in the optimized and default meshes are indicated in the legend, with the optimized mesh achieving a total element number reduction of ~16.8%.

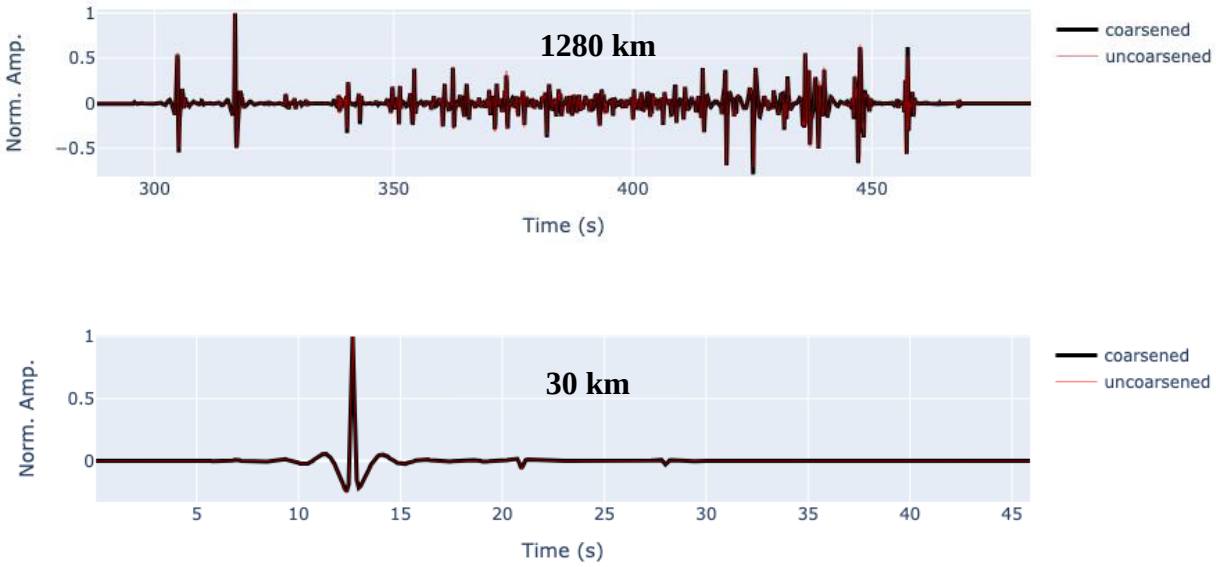

Fig. S2-3. Representative seismograms computed with the coarsened and uncoarsened meshes. Top: at epicentral distance 1,280 km. Bottom: at epicentral distance 30 km. The seismograms shown are the transverse component and are computed for the model with a Moho ramp with d =100 and w= 200 km. Horizontal axis represents time in seconds and vertical axis is normalized amplitudes from the coarsened model (thick black line) and uncoarsened mesh (thin red line).

S3. *Sn* and *Lg* windows

We follow essentially the same windowing strategy as in Wang and Klemperer (2021), in that we use the group-velocity ranges reported from regional observations and use a simple, layer-over-halfspace model to calculate a zero-offset intercept for *Sn*, following Barron and Priestley (2008). However, Wang and Klemperer (2021) maximized the *Sn* window by picking its end time using the latest possible arrival time from a 10-km earthquake even though, for such a shallow crustal earthquake, extending the *Sn* window this late will more likely include the earliest *Lg* energy than any meaningful additional *Sn*. Instead, here we simply use the earliest and latest *Sn* arrival time from an earthquake at the Moho as the *Sn* window. We suggest picking *Sn* and *Lg* windows should be approached from an empirical rather than an analytical perspective, as discussed more in the main text. In contrast to Song and Klemperer (2024) who used published earthquake depths to define *Lg* and *Sn* windows as a function of hypocentral depth, we process all our data with identical windows to avoid any pre-judgment and to robustly test our *Sn/Lg* discriminant.

S4. Additional simulation results with fixed ramp widths (*w*)

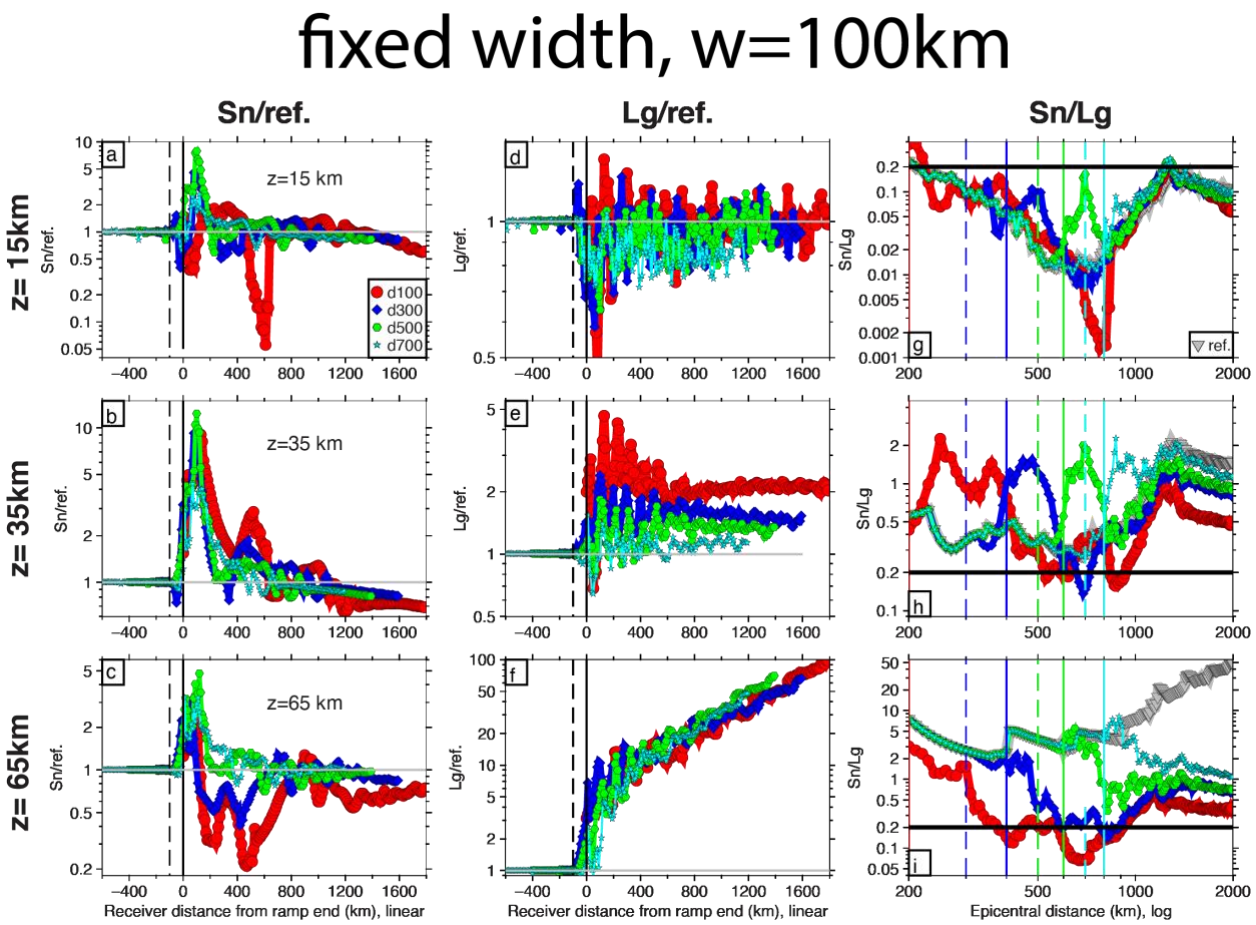

Fig. S4-1. *Sn*, *Lg* amplitudes and *Sn/Lg* with varying distance to ramp start *d* but fixed width *w* = 100 km. The layout is the same as Fig. 6 except for right columns, where colored dashed lines represent the start of a ramp and colored solid lines represent the end of a ramp, for the correspondingly colored ramp models. The start of the ramp for the *d*=100 km case is not shown and only its end is shown at 200–km.

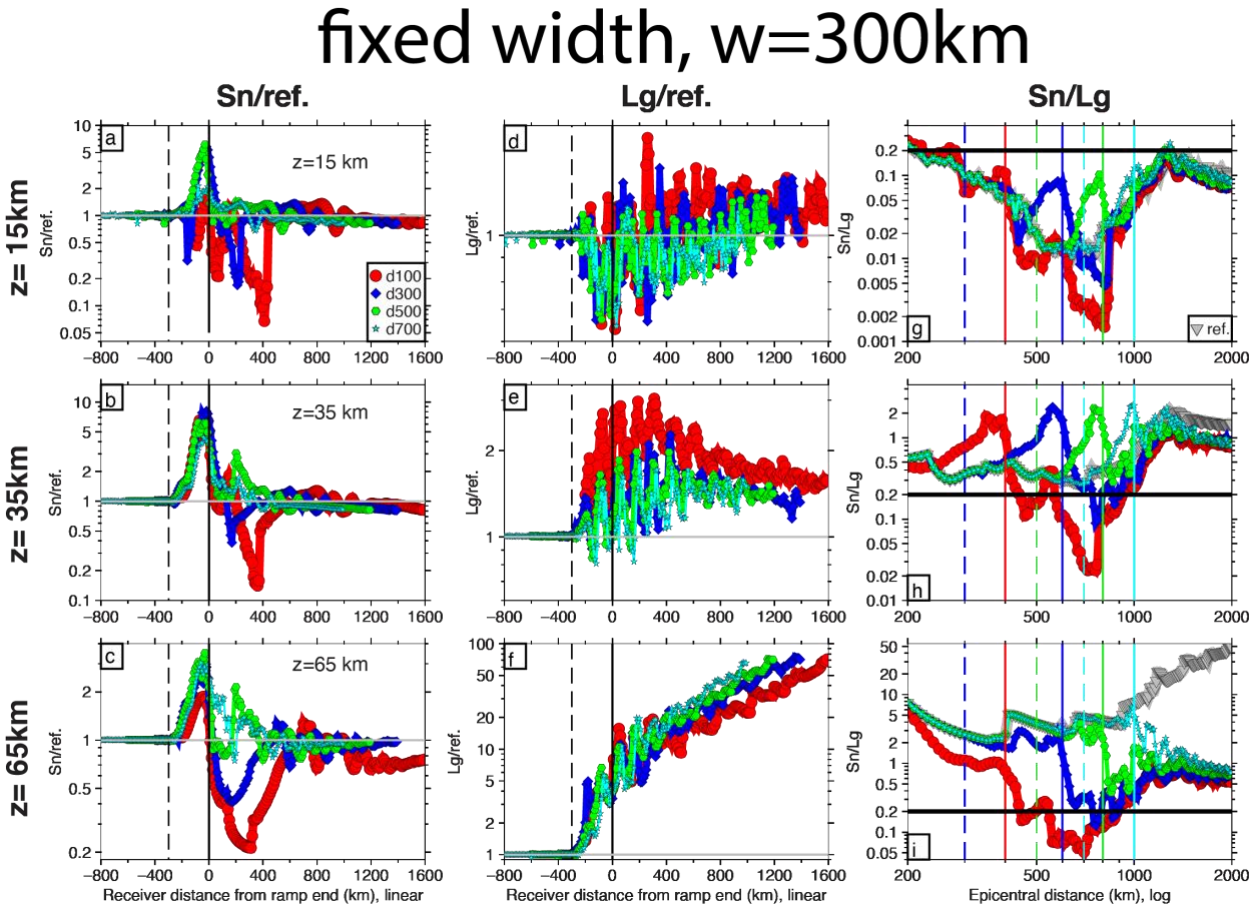

Fig. S4-2. *Sn*, *Lg* amplitudes and *Sn/Lg* with varying distance to ramp start *d* but fixed width *w* = 300 km. The layout is the same as Fig. S4-1.

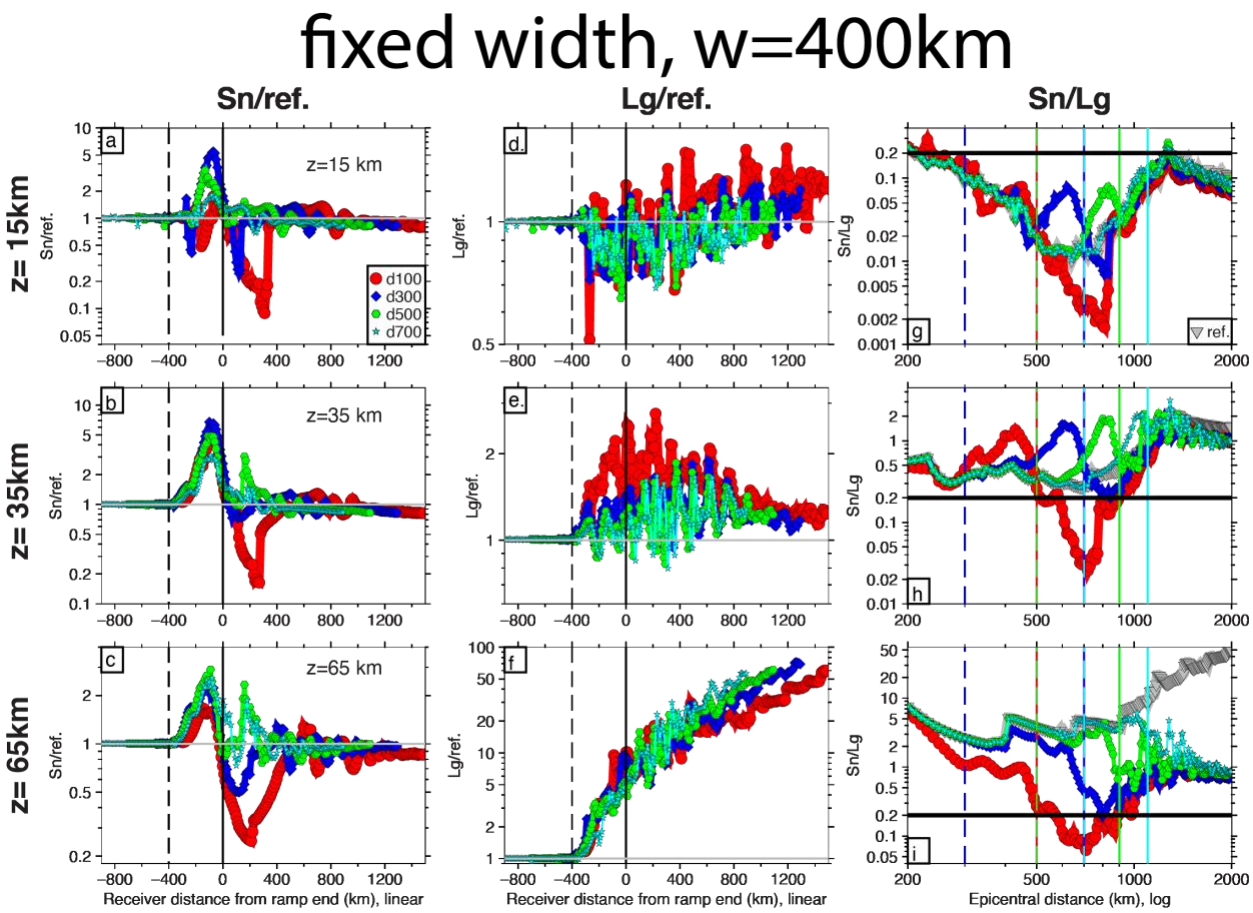

Fig. S4-3. *Sn*, *Lg* amplitudes and *Sn/Lg* with varying distance to ramp start *d* but fixed width *w* = 400 km. The layout is the same as Fig. S4-1.

S5. Additional simulation results with fixed distance to ramp (*d*)

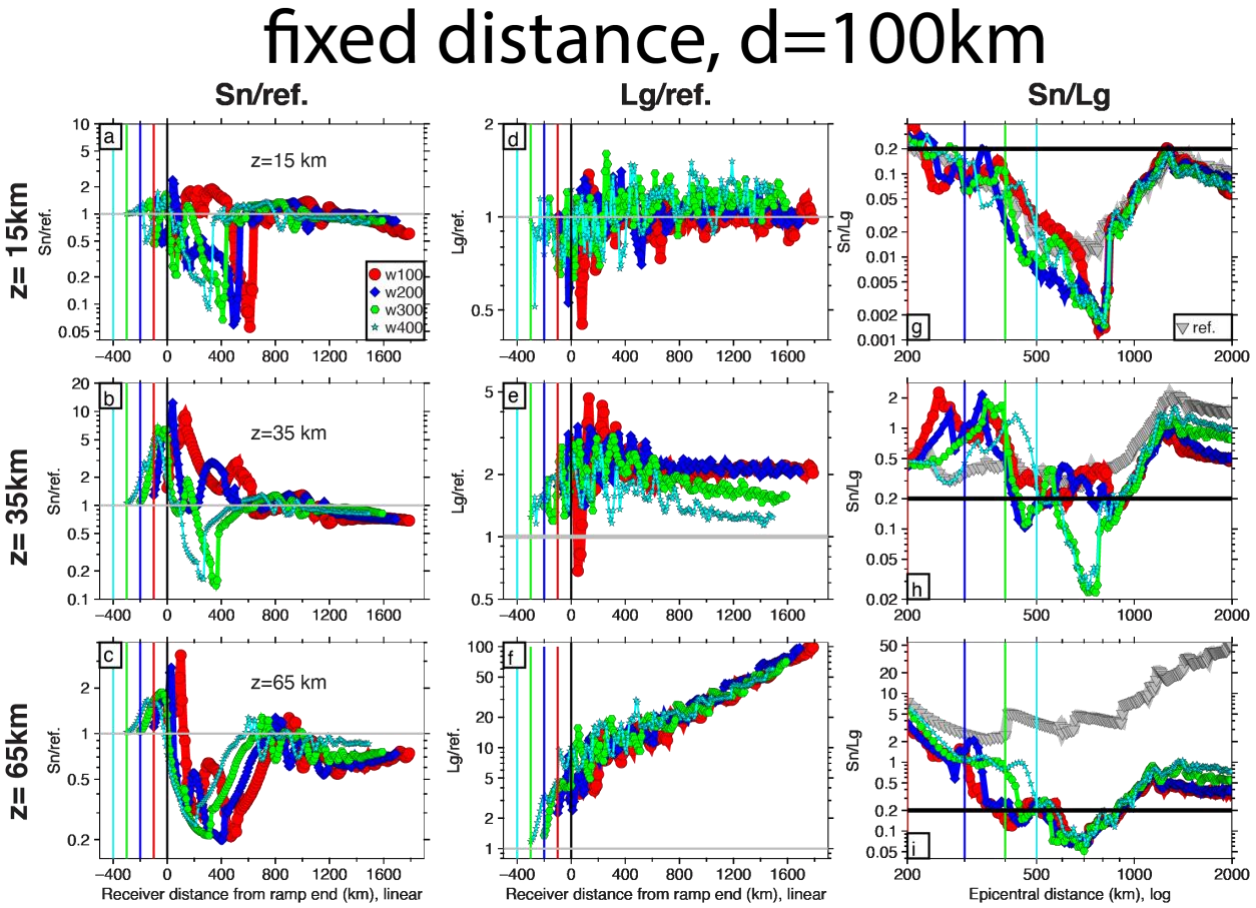

Fig. S5-1. *Sn, Lg* amplitudes and *Sn/Lg* with varying ramp width *w* but fixed ramp distance *d* = 100 km. The layout is the same as Fig. 9, except the start of the ramp is not shown the right column.

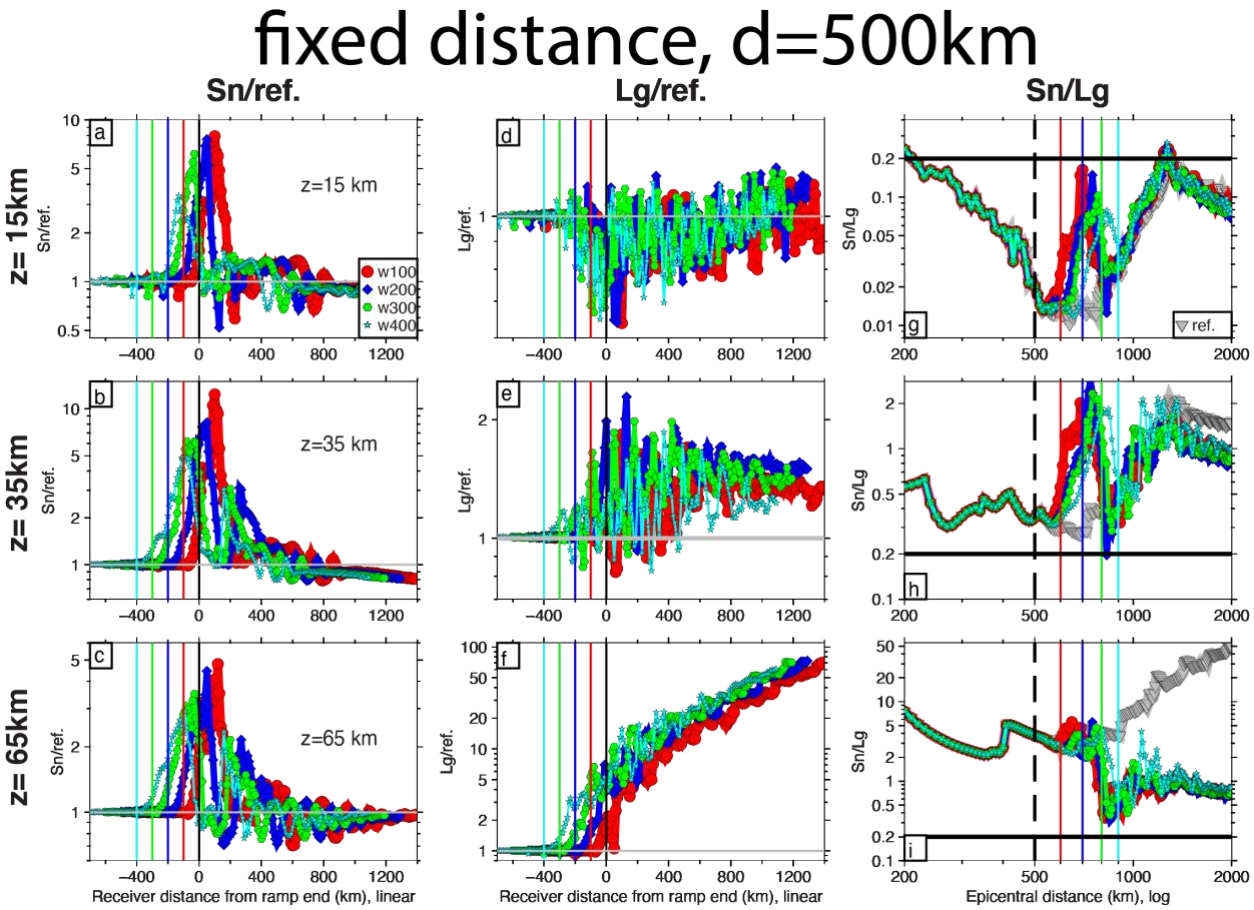

Fig. S5-2. ***Sn, Lg*** **amplitudes and** ***Sn/Lg*** **with varying ramp width** ***w*** **but fixed ramp distance** ***d*** **= 500 km.** The layout is the same as Fig. 9.

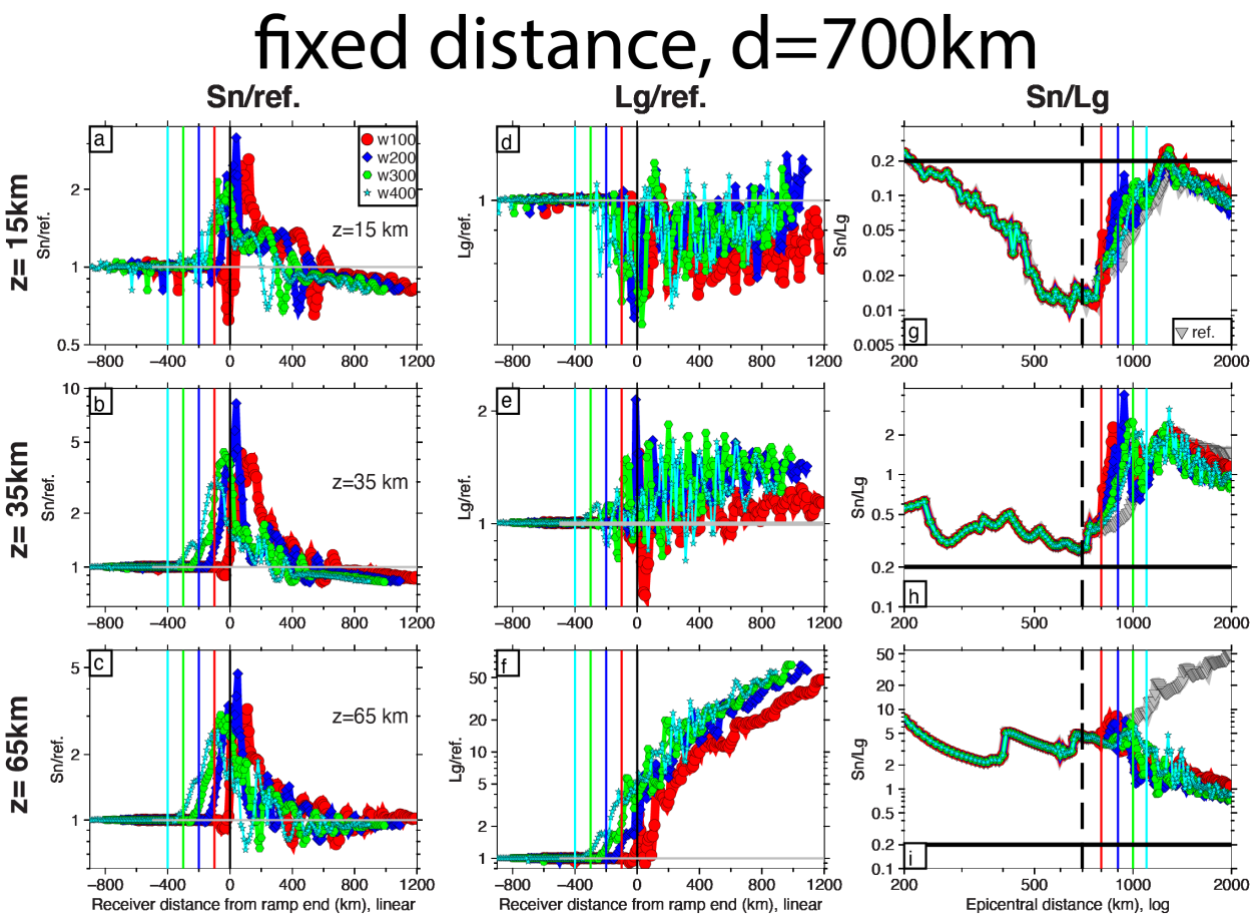

Fig. S5-3. ***Sn, Lg*** **amplitudes and** ***Sn/Lg*** **with varying ramp width** ***w*** **but fixed ramp distance** ***d*** **= 700 km.** The layout is the same as Fig. 9.

S6. HiCLIMB events plotted against epicentral distance

The southern events recorded on the HiCLIMB array are chosen so that they cover roughly the same epicentral distance ranges compared to the previously studied northern events, to facilitate their direct comparison. Except for a few closer stations recording event S4 and more-distant stations recording S1, almost all of our recordings, both from the northern and southern events, have an epicentral distance range of ~580–1100 km (Fig. S6-1). A separation of *Sn/Lg*, for the mantle (solid symbols) and crustal (open symbols) events, can also be observed in this plotting scheme (Fig. S6-1c).

This plotting scheme is ideal to investigate geometrical spreading effects in our raw data. We first plot a representative *Lg* geometrical spreading curve with $\gamma = -1$ (Fig. S6-1c&d, grey line). It is only the trend of this curve that matters but not its absolute values, i.e. it can be moved up or down freely. The southern events (Fig. S6-1c, colored symbols) have a rise of *Lg* amplitudes at around 700-800 km epicentral distances, followed by a smooth amplitude decrease, i.e. the effect of the attenuation region (Fig. 2a) is not visible. However, for the northern events (Fig. S6-1d, black symbols), the effect of the attenuation region is visible because *Lg* amplitudes rise first, then followed by a smooth decrease. In any case, the smooth or overall decreases of both groups of events are too steep compared with the geometrical spreading model (for reference, $\gamma = -8$ is plotted as a purple line in Fig. S6-1c&d, fitting the decreasing trend much better, but such a rate of amplitude decrease is unreasonable for *Lg* geometrical spreading), and the apparent deviations as discussed above are clearly correlated with station locations (Fig. 10 a-d). Hence, apparent *Lg* geometrical spreading in real data is severely influenced by other factors such as site effects.

We then filter the *Sn* wavetrain at 1 Hz and at 5 Hz (with a bandpass filter from $f/\sqrt{2}$ to $\sqrt{2}\,f$, the same way as in Yang et al. (2007), where f if the frequency of interest) and overlay the geometrical spreading model at these frequencies (Fig. S6-2). All of the data start to lose amplitude at around 800-900 km, and the changes in amplitude are not clearly related with the interference head wave phenomenon but are clearly correlated with station location and location of the attenuation zone (Fig.10 a–d). Hence, we reach the same conclusion for *Sn* geometrical spreading as for *Lg*: apparent spreading effects are severely influenced by local factors.

These comparisons of real data and geometrical spreading models shed some lights on the apparent deviations of real data from our synthetic predictions on individual amplitudes. If the basic geometrical spreading patterns are obscured in the real data, then it is unreasonable to expect to observe the superimposed perturbations due to crossing a Moho ramp.

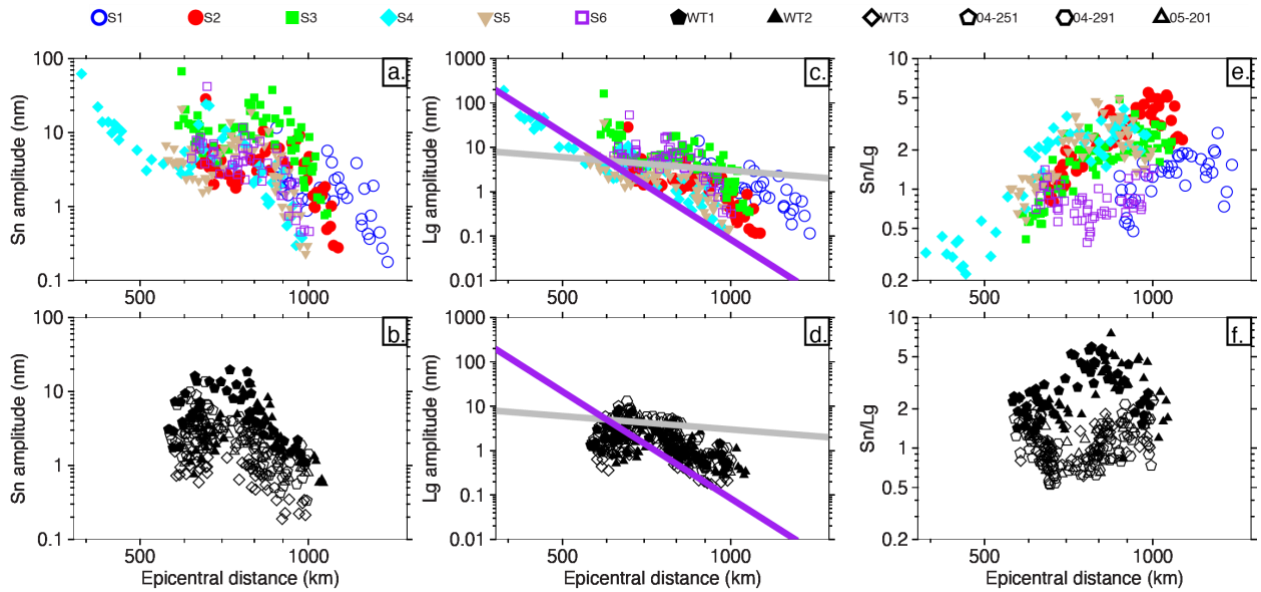

Fig. S6-1. HiCLIMB data plotted against epicentral distance (log-log plots). All symbols are the same as in the main paper for HiCLIMB events. Amplitudes are absolute without any normalization. Data are filtered from 1–5 Hz. (a)&(b) *Sn* amplitudes for northern and southern events, respectively. (c)&(d) *Lg* amplitudes for northern and southern events, respectively. Geometrical spreading model with $\gamma = -1$ (grey line) and $\gamma = -8$ (purple line) are shown. (e)&(f) *Sn/Lg* for northern and southern events, respectively. Note clear separation of solid (mantle earthquake) and open (crustal earthquake) symbols from ~600–1000 km. Horizontal axes for the top and bottom rows are identical and aligned, for easy comparison of distance ranges for our events.

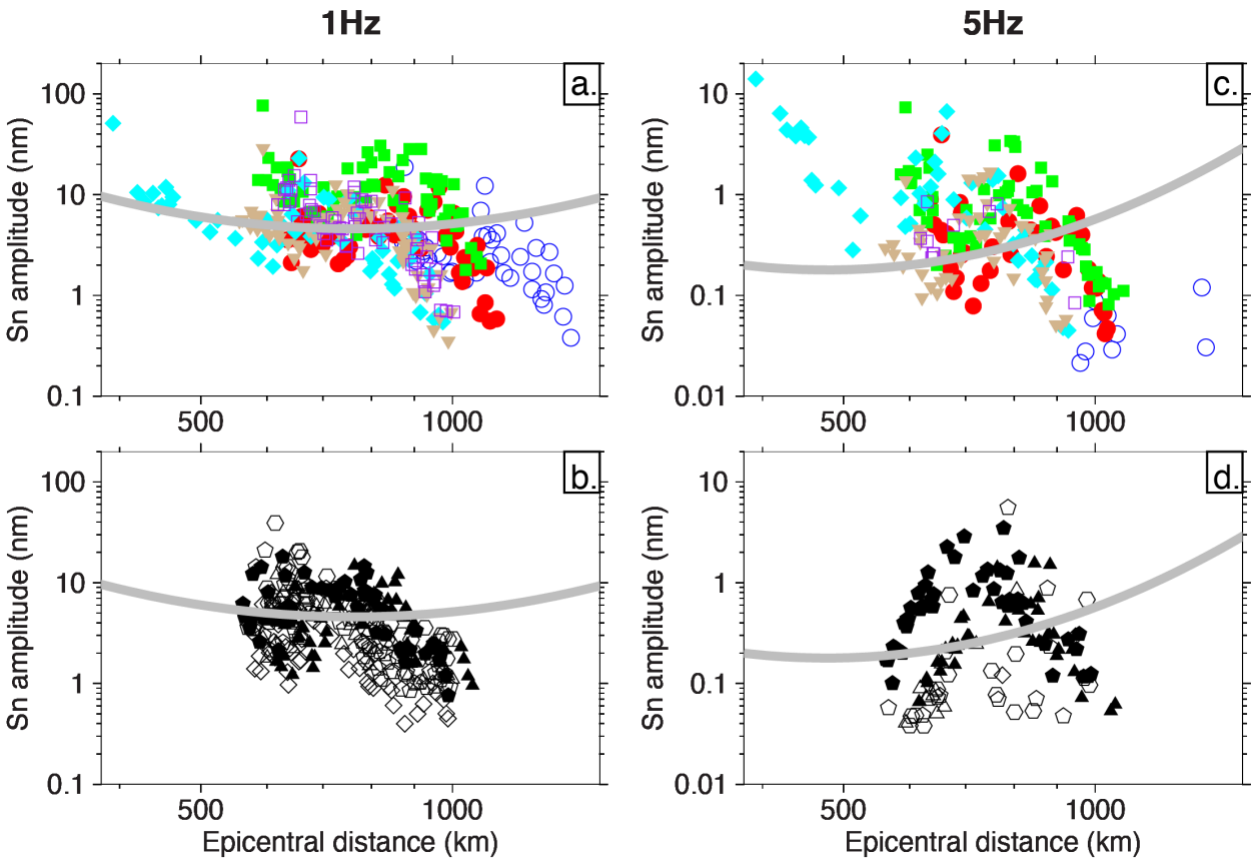

Fig. S6-2. *Sn* filtered to two frequencies. All symbols are the same as in the main paper. Amplitudes are absolute without any normalization. (a)&(b) *Sn* at 1 Hz for northern and southern events, respectively. (c)&(d). *Sn* at 5 Hz for northern and southern events, respectively. Corresponding geometrical spreading models from Yang et al. (2007) are shown as grey lines.

S7. Data from the Gangdese-92 array

Code name	Date, time	Location (N°, E°)	Magnitude (m_b)	Catalog depth (km)	Distance to ramp start d (km)
G1	2012-01-01, 02:35:21	23.472, 91.834 *23.503, 91.804*	4.6	27.8±10.4 *15*	450
G2	2012-09-06, 18:27:11	25.455, 91.208 *25.600, 91.049*	4.5	45.1±8.5 *15*	245
G3	2011-09-18, 19:20:54	25.759, 91.178 *25.576, 91.113*	4.0	37.1±9.5 *14.2*	214
G4	2012-05-11, 12:41:35	26.175, 92.889 *26.246, 92.882*	5.4	43.4±? *43.0*	193

Table S7-1. Events G1–G4 recorded on the Gangdese-92 array. Event details are from USGS except Distance to Ramp, which is measured according to the description in the main paper. These events are essentially aligned with the array, so a single distance to ramp start is reported, instead of a range. The ramp width they cross is ~constant at ~160 km. Italicized hypocentral locations and depths are from the Seismological Bulletins of the Indian National Center for Seismology.

The *Sn* and *Lg* amplitudes and *Sn/Lg* amplitude ratios for events G1–G4 are shown in Fig. S7-1. Apart from *Sn* from G4, all other *Sn* and *Lg* amplitudes vary coherently with station locations (Fig. S7-1a&b), demonstrating likely site effects. *Lg* for these four events has a plateau until ~100 km beyond the end of the ramp, then at greater distances *Lg* amplitudes start to fall off faster than r^{-1} where r is epicentral distance. This could be a representation of enhanced *Lg* near the ramp end. *Sn/Lg* for these four events largely overlaps (as does their *Lg* HF/LF, Fig. 12c), and is always <2, a working *Sn/Lg* threshold established at the nearby PASSCAL array (Wang & Klemperer, 2021) for mantle earthquakes. Without a clear *Sn/Lg* baseline for collocated crustal earthquakes, we are unable to rigorously establish the depths of these events with respect to Moho, but they are very likely deep-crustal events given their catalog depths. Our inability to perform a full analysis on the Gangdese array is the reason we show its results only in supplementary materials (apart from its *Lg* HF/LF).

Record sections for these four events are shown in Fig. S7-2, indicating good signal-to-noise ratio.

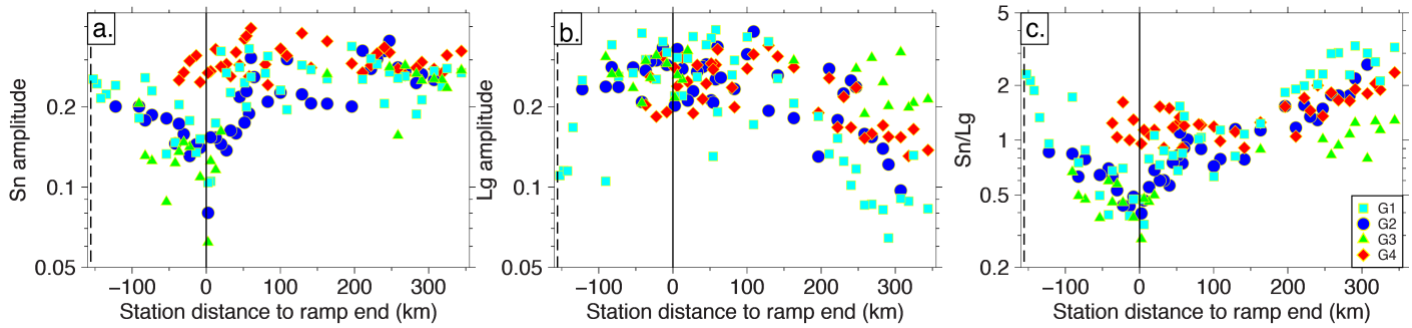

Fig. S7-1. *Sn, Lg* amplitudes and *Sn/Lg* as a function of distance to ramp end for Gangdese-92 events. All annotations are as in Fig. 10. This array did not extend into the *Sn* attenuation region.

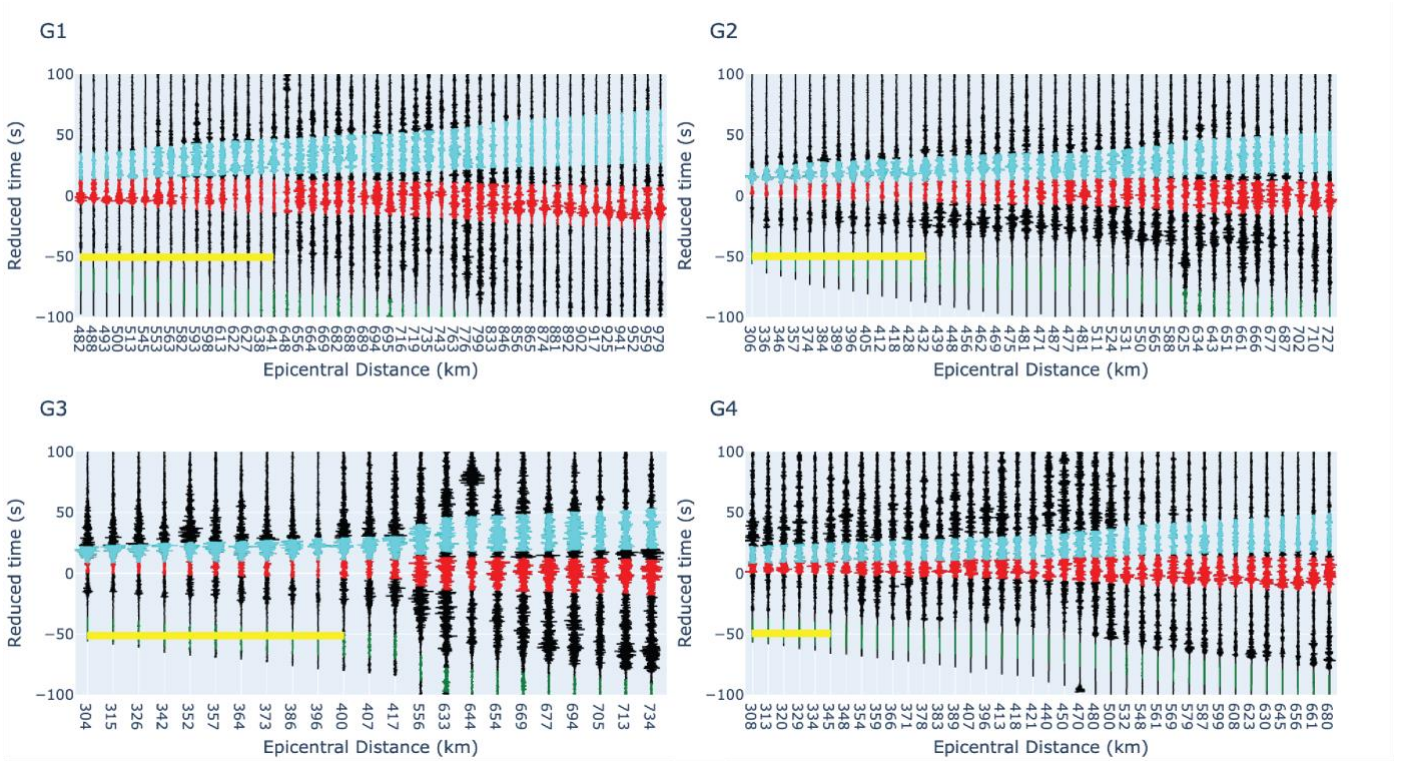

Fig. S7-2. Record sections for Gangdese-92 events. Record sections are shown with reduction velocity of 4 km/s. Green section: noise window. Red section: *Sn* window. Cyan section: *Lg* window. Horizontal yellow bar indicates stations south of YZS, i.e. within the Moho ramp region.

S8. Measuring *Sn/Lg* 100 km north of the end of the ramp

We remeasure the percentage of *Sn/Lg* > 2 for all of the HiCLIMB events, but only use stations that are at least 100 km north of YZS. This is performed to evaluate whether *Sn/Lg* signature is stronger if we move further away from the ramp, as predicted by our synthetics and qualitatively shown in Fig. 10e.

The four southern events interpreted as mantle earthquakes (S2, S3, S4, S5) show a significant increase in the percentage of stations that record *Sn/Lg* >2 if we only use stations > 100 km north of the end of the Moho ramp (Fig. S8-1). However, for the northern events that see no ramp, and for the southern events interpreted as crustal earthquakes (S1 & S6), using only the stations > 100 km north of the end of the Moho ramp did not significantly change the percentage of stations surpassing our threshold (Fig. S8-1). Hence we demonstrate that, using only stations >100 km distant from the ramp does increase the contrast of crustal and mantle earthquakes as measured by *Sn/Lg*.

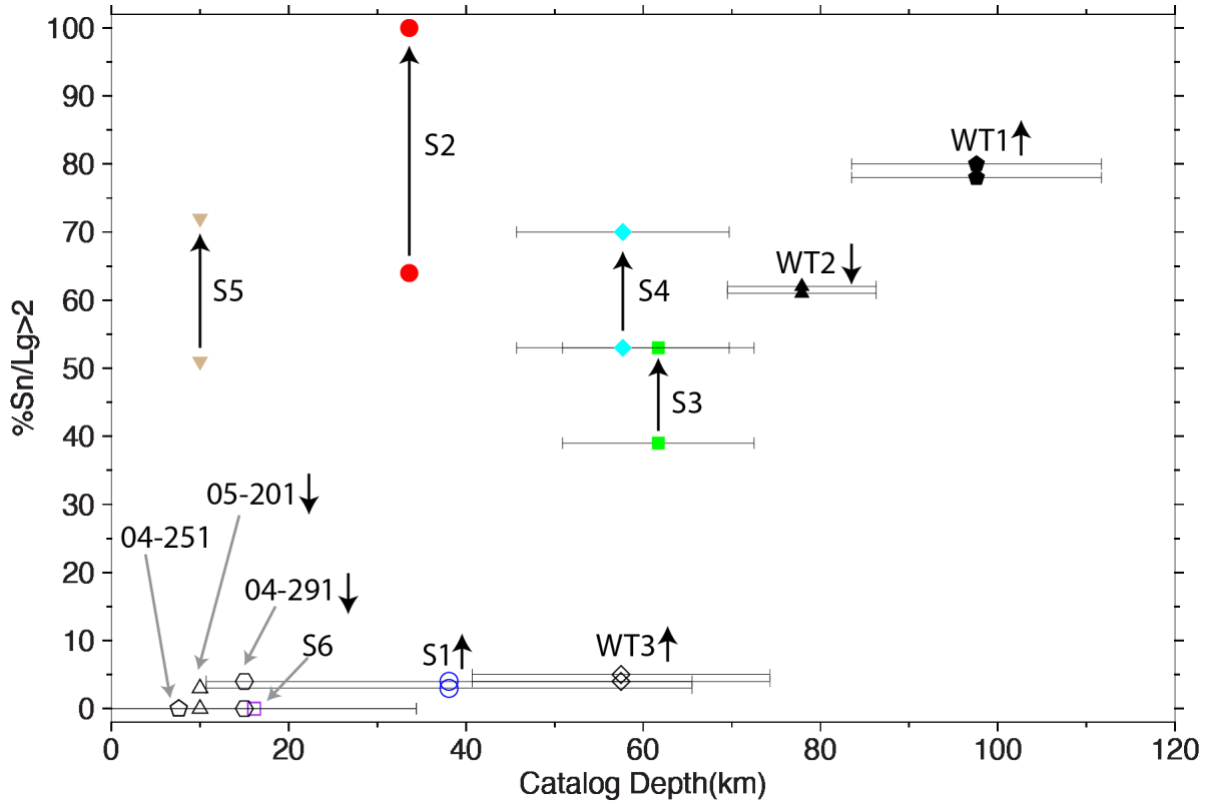

Fig. S8-1. Percentage of stations with *Sn/Lg*>2 for HiCLIMB events, *vs.* catalog depth. Symbols are the same as in Fig. 10; colored symbols S1–S6 are earthquakes in the Himalayan foreland with Moho depth likely < 45 km (Singh et al., 2015; Mitra et al., 2018), whereas black symbols are from northwest Tibet with Moho depth likely > 80 km. Horizontal axis shows catalog depths from USGS with uncertainties when available. Black arrows next to event names correspond to whether percentage of stations with *Sn/Lg*>2 has increased or decreased from using all stations to using only stations at least 100 km north of YZS. When there is no change, no black arrow is shown. Grey arrows lead event names to event symbols.