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

Shiqi Wang[1] and Simon L. Klemperer[1]

[1]Stanford University

April 19, 2024

Abstract

We study Sn-to-Lg conversion at regional distances due to significant crustal thickening, particularly in the context of using Sn and Lg amplitude ratios (Sn/Lg) to identify continental mantle earthquakes. We further enhance recent developments in computational seismology to perform 2.5D simulations up to 5 Hz and 2,000 km. Our simulations compare propagation in a reference, constant-thickness crust from a source at three depths straddling the Moho, to 48 models of the same three sources propagating through Moho ramps of four different widths (dips) at four different distances from the source. We compare our synthetics to data from 12 earthquakes recorded on the HiCLIMB array across Tibet, of which six events from northwestern Tibet traverse no major crustal-thickness variation, and six located south of the Himalaya cross a major Moho ramp. Our observations on real data show that amplitude perturbations on individual Sn and Lg waves are smooth and mostly limited to near the ramp end. Even the more-pronounced amplitude variations seen in our simulations show that Sn/Lg for mid-crustal earthquakes is consistently lower than those for mantle earthquakes. Hence we can directly compare Sn/Lg for ramp-crossing and non-ramp-crossing earthquakes and identify new mantle earthquakes in northern India. Sn-to-Lg converted waves may be readily detected near the Moho ramp end through an enhancement in high-frequency content. In addition, we observe higher frequency content in Lg from crustal than from mantle earthquakes, which offers a new discriminant for continental mantle earthquakes based on frequency content of Lg waves alone.

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

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

*Corresponding author. Email address: axelwang@stanford.edu

Key points

- Synthetics and data show the *Sn/Lg* method successfully identifies mantle earthquakes with thickening crust across the Himalaya.
- *Sn*-to-*Lg* conversions can be recognized by enhanced high frequency content of *Lg*
- *Lg* frequency content discriminates between crustal and mantle near-Moho earthquakes

Abstract

We study *Sn*-to-*Lg* conversion at regional distances due to significant crustal thickening, particularly in the context of using *Sn* and *Lg* amplitude ratios (*Sn/Lg*) to identify continental mantle earthquakes. We further enhance recent developments in computational seismology to perform 2.5D simulations up to 5 Hz and 2,000 km. Our simulations compare propagation in a reference, constant-thickness crust from a source at three depths straddling the Moho, to 48 models of the same three sources propagating through Moho ramps of four different widths (dips) at four different distances from the source. We compare our synthetics to data from 12 earthquakes recorded on the HiCLIMB array across Tibet, of which six events from northwestern Tibet traverse no major crustal-thickness variation, and six located south of the Himalaya cross a major Moho ramp. Our observations on real data show that amplitude perturbations on individual *Sn* and *Lg* waves are smooth and mostly limited to near the ramp end. Even the more-pronounced amplitude variations seen in our simulations show that *Sn/Lg* for mid-crustal earthquakes is consistently lower than those for mantle earthquakes. Hence we can directly compare *Sn/Lg* for ramp-crossing and non-ramp-crossing earthquakes and identify new mantle earthquakes in northern India. *Sn*-to-*Lg* converted waves may be readily detected near the Moho ramp end through an enhancement in high-frequency content. In addition, we observe higher frequency content in *Lg* from crustal than from mantle earthquakes, which offers a new discriminant for continental mantle earthquakes based on frequency content of *Lg* waves alone.

Plain language summary

Seismic waves *Sn* and *Lg* respectively propagate largely below and above the Moho. Previous work showing that *Sn* and *Lg* amplitudes can distinguish whether near-Moho continental earthquakes nucleated in the crust or mantle (the '*Sn*/*Lg* method') used only 1D (flat-Moho) theory and synthetics, and data from areas with little Moho topography. Here we extend this work with synthetic seismograms across large Moho ramps and with data recorded across the Himalaya from India to Tibet. By comparing earthquakes with source-receiver raypaths that do and do not cross a Moho ramp we show the *Sn*/*Lg* method can still identify mantle earthquakes provided multiple recorders are used. We also show that the frequency content of *Lg* contains information about *Sn*-to-*Lg* conversions, and can by itself be used to identify mantle earthquakes. Traditionally, *Sn* and *Lg* waves have not been modeled at high-frequencies (>1 Hz) and long-distances (>1000 km) due to high computing costs. Here, we take advantage of and enhance recent developments in computational seismology to model *Sn* and *Lg* propagation up to 5 Hz and for 2000 km through a 2D lithosphere, paying special attention to their amplitude ratio and its application to distinguish exotic continental mantle earthquakes from commonplace crustal earthquakes.

Keywords: *Sn*, *Lg*, crustal thickening, continental mantle earthquakes, Himalaya, Tibet

1. Introduction

Seismic waves *Sn* and *Lg* are the most prominent arrivals on high-frequency (~1–5 Hz) seismograms recorded at regional distances (~200–2,000 km). They are guided shear waves within the entire crust (*Lg*) or the entire lithosphere (*Sn*), and can be represented equivalently either by Airy phases from surface-wave normal modes (Stephens and Isacks,1977, Knopoff, 1973) or by interference patterns of waves multiply reflected between the surface and the Moho top-side (Oliver and Ewing, 1958) or under-side (Červený and Ravindra, 1971; Menke and Richards, 1980), respectively (Fig. 1). Their excitation and propagation characteristics derived from the above representations are directly related to the wave amplitudes that have been useful for a variety of purposes such as determining focal depths for crustal earthquakes from amplitude spectra (Baker et al., 2004), serving as the dominant measure for regional earthquake magnitude (e.g. Patton and Walter, 1993), monitoring nuclear tests based on *Pg* and *Lg* amplitude ratios (e.g. Zhang and Wen, 2013), as well as estimating local properties relating to the attenuation (e.g. Mousavi et al., 2014) and amplification (i.e. seismic hazards, e.g. Kebeasy and Husebye, 2003, Rodgers et al., 2019, 2020) of these waves.

Recently, addressing a half-century-long controversy regarding whether earthquakes can nucleate in the continental mantle (Chen and Molnar, 1983; Maggi et al., 2000; Chen and Yang, 2004; Schulte-Pelkum et al., 2019; Priestley et al., 2008; Craig et al., 2011; Prieto et al., 2017), we demonstrated the use of *Sn* and *Lg* amplitude ratios (hereafter "*Sn/Lg*") to discriminate continental mantle earthquakes from crustal ones using Tibetan earthquakes recorded on the Tibetan plateau (Wang and Klemperer, 2021) (Fig. 2). The signature of a mantle origin is a higher *Sn/Lg* compared with nearby crustal earthquakes recorded on a common array. For a group of earthquakes in NW Tibet, *Sn/Lg* ratios > 2 (averaged over many stations) were found to identify sub-Moho earthquakes. This method has the advantages of making the discrimination by relying on prominent waveform features of the earthquakes themselves (as opposed to Zhu and Helmberger, 1996 and Yang and Chen, 2010, who relied on more subtle waveform features), thus avoiding comparing independently derived earthquake and Moho depths at different locations, which has been a popular method (Chen and Yang, 2004, Priestley et al., 2008), and also can be performed using any stations/arrays that lie within regional distances of an earthquake (as opposed to Schulte-Pelkum et al., 2019 who relied on stations essentially on top of earthquakes).

Our *Sn/Lg* method is based on predictions from 1D surface-wave normal-mode theory, but given the *Sn* and *Lg* interconversions (*Lg* blockage or leakage) created when waves are incident on a dipping Moho, it is far from certain how the method will perform if there exists a large-scale structural variation between the earthquakes and recording stations (e.g. earthquakes in India recorded by stations in Tibet). Necessary corrections may be small – Song and Klemperer (2023) show general agreement between the catalog depths and *Sn/Lg* of hundreds of earthquakes with paths crossing the boundaries of Tibetan Plateau recorded on either of two permanent stations (KBL and LSA) – or may be significant, as where *Lg* blockage is used to study large-scale geologic features (e.g. North Sea: Mendi et al.,1997, Japan: Furumura et al., 2014, Pyrenees: Sens-Schönfelder et al., 2009).

Given the number of seismological applications utilizing regional-wave amplitudes, and that large-scale Moho topography is often well-known, many attempts have been made to quantify *Lg* blockage as *Lg* propagates through a suddenly thinned crust. Coupled-mode theory builds on the 1D surface-wave eigenproblem which can synthesize regional waves with only vertical (1D) heterogeneity. Coupled-mode theory represents the wavefield as a sum of basis functions (motion-stress vectors for a 1D problem, e.g. Aki and Richards, 2002) with laterally-varying amplitude coefficients obtained through the orthogonality principle of the normal modes (Maupin, 1988). Most relevant here are to consider the width across which Moho depth varies (Kennett, 1972; Drake, 1972; Kennett 1984; Maupin, 1988) and to incorporate undulating structural boundaries using local modes (i.e. motion-stress vectors corresponding to a flat (1D) model locally identical to a small section of the laterally varying 2D model; Odom, 1986) and representing the continuity conditions on the tilted surfaces as a volume force in both 2D (Maupin, 1988) and 3D (Tromp, 1994). A 2D coupled-local-mode method, incorporating all these ideas, was applied to *Lg* propagation in the North Sea (Maupin, 1989) to model transmitted and reflected wavefields for incident waves both perpendicular and at a sub-critical angle to the strike of the Moho topography. Maupin (1989) reported little difference between perpendicular and oblique incidences; the reflected wavefield is negligible and the strong *Lg* attenuation seen in the North Sea cannot be fully explained simply by structural effects, a conclusion that has been corroborated by later studies (Cao and Muirhead, 1993; Mendi et al., 1997) using 2D finite-difference simulations. An important observation is that mode-coupling occurs most strongly between neighboring modes. In Maupin (1989)'s North Sea model at a fixed frequency of 1 Hz, *Lg* mostly leaks into the mantle as *Sn* waves from the first (lowest) few *Sn*-forming normal modes (Fig. 3), as predicted by Kennett (1984). This means that only the lowest few *Sn*-forming normal modes, or the highest few *Lg*-forming normal modes, get enhanced by *Sn* and *Lg* interconversion, and if these enhanced modes do meaningfully contribute to either the *Sn* or *Lg* wavetrain then they contribute more to the low-frequency content of *Sn* or the high-frequency content of the *Lg* wavetrain. An alternative to the coupled-mode method is the ray-diagram method (Kennett, 1986 for *Lg*; Xie, 1996 for *Pn*), whose results are mostly graphical and do not account for interference between different rays once their initial coherent pattern is broken (Kennett, 1986). Nonetheless, for an initial bundle of rays with the same inclination (i.e. apparent velocity), focusing and de-focusing effects due to the lateral structure can be clearly seen (e.g. Kennett, 1986, his figures 2 and 3). These methods study the interactions of different modes (i.e. different dispersion relations: frequency vs. wavenumber) by either fixing the frequency (the coupled-mode methods) or the wavenumber (proxy to apparent velocity, the ray-diagram method). These methods yield valuable insights, but cannot represent the full broadband wavefield, which for regional waves is dominated by interference patterns.

Fully-numerical simulations can calculate the full broadband wavefield for any arbitrary structure, however, significant computational challenges exist given the frequency and range requirements for simulating regional waves. We are not aware of any 3D simulations that simultaneously reach frequencies up to 5 Hz and distance ranges up to 2,000 km, common observational parameters for *Sn* and *Lg* waves. Furumura et al. (2014) simulated regional wave propagation around Japan up to 1.5 Hz; and Rodgers et al. (2019, 2020) simulated ground motion in the San Francisco Bay Area covering an area of 120 km x 80 km up to 10 Hz. More importantly, these 3D simulations are run with very specific models, so are hard to generalize to other cases. On the other hand, 2D simulations, which recently focused on *Pn* propagation (Bakir and Nowack, 2012; Xie and Lay, 2017a&b; Wang et al., 2017), are attractive as they are much cheaper, so may simultaneously satisfy the frequency and range requirements, and may be more generalizable. However, these simulations, if performed in a Cartesian grid, require the earth-flattening transformation to produce physical sphericity which is vital for simulating interference head waves such as *Pn* and *Sn*. More importantly, these 2D simulations require a non-straightforward correction from their 2D line sources to 3D point sources (Li et al., 2014), and this correction cannot be exact if lateral heterogeneities exist (Li et al., 2014).

Here we establish first-order features of *Sn* and *Lg* transmission and inter-conversion with a set of 2.5D axisymmetric simulations allowing exact representations of Earth's sphericity and of 3D point sources. Our simulations have a maximum range of 2,000 km and frequency of 5 Hz, typical values used in observations. We view our synthetic results as building on those of Yang (2002) and Yang et al. (2007) who investigated *Lg* and *Sn* geometrical spreading for simple 1D models, and so we do not include effects such as intrinsic attenuation or random scatterers. The only factor that should make our synthetic results deviate from the 1D studies is the laterally varying crustal thickness, which is also typically well-known, thereby allowing our results to be quickly adapted to multiple regions of the world. We restrict our structural models to a Moho ramp leading to crustal thickening. Crustal thickening has been less explored, perhaps due to its subtler influence compared to crustal thinning, but this limited scope allows us to discuss comparisons with real data (Fig. 2).

After discussing the computational setup of our model, we explore individual *Sn* and *Lg* amplitudes, and their amplitude ratios (*Sn/Lg*) in a 1D reference model (Figs. 4&5) and in Moho-thickening models (Figs. 6–9), in which we establish the effectiveness of using *Sn/Lg* to identify continental mantle earthquakes in the presence of significant Moho thickening. We next examine real data from Tibet (Figs. 10&11) by directly comparing ramp-crossing (S events, Table 1& Fig. 2) and non-ramp crossing events (non-S events, Table 1 & Fig. 2), and show that *Sn/Lg* is a valid criterion for separating mantle from crustal earthquakes for the ramp crossing events just as for the non-ramp-crossing events. Hence the *Sn/Lg* method, if used rigorously with local shallow comparison events and multiple recording stations, can recognize the signature of a mantle earthquake even with stations in a region of crust much thicker (Tibet) than the source region (Indian Shield). Although we do not reliably detect effects of the ramp on individual *Sn* and *Lg* amplitudes, we are able to confirm enhancement of high-frequency *Lg* across the ramp (Fig. 12) due to neighboring mode-coupling during *Sn*-to-*Lg* conversion (Figs. 3, 13). Indeed, *Lg* frequency content is another powerful discriminant for continental mantle earthquakes.

2. Computational aspects

We use the AxiSEM3D software package, whose main advances compared to previous 2.5D axisymmetric methods (Bottero et al., 2016; van Driel et al., 2015) are that it can account for fully 3D variations in terms of volumetric perturbations (Leng et al., 2016) as well as through undulating surfaces (i.e. structural variations to either internal surfaces such as a Moho ramp or external surface such as the ellipticity of the earth or topography) that break the spherical geometry necessary for an axisymmetric method (Leng et al., 2019). We first briefly discuss these new features from a user's perspective and introduce two necessary modifications made to the source code in order to enable simulations with our desired scale and output.

2.1 Computational method and its enhancements

Without considering undulating surfaces, the azimuthal component of the 3D wavefield (from 0 to 2π in the plane perpendicular to the source-receiver direction) can be conveniently represented by a Fourier series, which localizes the equations to a single meridian plane not associated with any physical location, and then can be solved with a 2D spectral-element method (Leng et al., 2016). Recognizing that lateral heterogeneities in earth are much smaller than vertical ones, this hybrid scheme essentially uses "one line-shaped element" and high-order Fourier series in the azimuthal direction, and in the 2D meridian plane uses 4th-order Lagrange polynomials on a mesh with the quad-shaped elements that are necessary for a conventional spectral-element method. The cost of 3D simulations in AxiSEM3D depends not on the length of the 3rd dimension which in AxiSEM3D is always 0 to 2π, but on the 2D model size and highest wave frequency, since these determine the number of elements that each have an associated Fourier series. Even though the AxiSEM3D hybrid scheme is much more efficient for a global 3D model than a fully 3D scheme, we note that a small 3D model needs the same Fourier orders as the global model with the same level of lateral heterogeneity, and so a conventional 3D method might be more desirable in this case. Testing for this specific threshold is beyond the scope of this study. We estimated the Fourier orders needed if we were to extend our 2.5D simulation to 3D based on Equation (5) in Szenicer et al. (2020), and found 3D simulations are well out of reach given our available computational resources.

Even for our 2.5D simulations, the spatially and temporally down-sampled wavefield (discussed in detail in the next section) on one meridian plane on one wavefield component is about 1.5 terabytes, and AxiSEM3D by default directs all parallel processes (MPI ranks) to output the data to the same location. For our output size, the bandwidth to one location in a filesystem is overloaded, greatly reducing the performance (since MPI ranks spend most of their time waiting for I/O instead of computing) and more importantly, causing frequent filesystem crashes. We take advantage of local hard drives physically connected to each computing node on Stanford University's Sherlock HPC cluster (https://www.sherlock.stanford.edu/) through infiniband, and we modified AxiSEM3D so that each MPI rank can identify its own computing node at runtime and output its results to that node's physically-connected hard drive. This resolved the problem of crashing the filesystem, and increased the performance of AxiSEM3D by at least one order of magnitude for our problem size.

AxiSEM3D accounts for structural boundary variations diffeomorphic to a spherical or flat (for Cartesian mesh) boundary through the use of a "particle-relabeling transformation" (Leng et al. 2019; Al-Attar &Crawford, 2016), which finds the change in radial coordinate for each collocation point inside an element necessary to represent the structural boundary (and a finite thickness transition zone around it). The input mesh is a 2D mesh without any geometric variation (undulations on structural boundaries) or volumetric variation (perturbations of material properties such as density and elastic constants), so is a 1D vertically-layered model compatible with the axisymmetric requirement. The 2D and 3D variations are added on later with separate files (for geometric and volumetric variations) and then described as Fourier coefficients for each element in the 2D mesh. The 2D mesh can be related to a specific physical location and its properties only by association with a specific azimuthal angle ϕ. The current version of AxiSEM3D (Leng et al., 2019) does not output any of the built 2D or 3D models, and computed wavefields can only be plotted on the coordinates of the spherical mesh with no undulating surfaces. This causes a distortion of the wavefield visualizations that is too small to see on a global scale (Nissen-Meyer, pers. comm.), but unacceptable for our regional-scale lithospheric simulations. Further, although users can define models for undulating surfaces, there is currently no way to check if this is being represented accurately inside the program. We modified the source code to output the Fourier coefficients related to structural boundary variations for affected elements, and then deform the input mesh to obtain 2D variations at any azimuthal angle ϕ (constant in our 2.5D simulation) specified in our input geometric model. This enables us to visualize our regional wavefield without distortion and to confirm that our Moho undulation is exactly represented by the Fourier series (Supplementary material S1).

Lastly, we note that it is still not a straightforward and cheap task to perform regional wave simulations even with the proliferation of computing resources and advancements in efficient computational methods. Each of 51 simulations presented here required ~2 days with 500 cores on Stanford's Sherlock supercomputer. Since most regional wave applications focus on amplitudes, a much cheaper method based on radiative transport theory (a form of advanced ray theory), that can only calculate absolute amplitude but can easily account for 3D structures and random scatters (Sanborn et al., 2017) is potentially attractive. We did not use radiative transport because we prefer a fully-numeric method, and because radiative transport has been shown to underestimate shear energy both in 2D (Pryzbilla et al., 2006) and in 3D (Pryzbilla et al., 2008).

2.2 Model design

Our simulation domain is shown in Fig. 1. It has an effective size of ~2,000 x 230 km within the absorbing boundary conditions. The vertical properties are based on PREM (Dziewonski and Anderson, 1981) with the Moho depth adjusted to 30 km to better represent continental areas. Our 230-km depth includes 10 km of the positive velocity gradient below the mantle low velocity zone (LVZ) in order to include the LVZ trapped waves as well as *Sa*: the shear wave trapped between the free-surface and the bottom of LVZ (Schwab et al., 1974; Wang and Klemperer, 2023) that are important for the energy partitioning of surface waves.

Our reference model has Moho depth fixed at 30 km. For our other simulations, we introduce a Moho ramp, in the shape of a sigmoid function at distance d from the source and with width w that smoothly transitions the Moho depth from 30 km on the source side to 60 km beyond the ramp (Fig. 1). The distance-to-ramp parameter d controls the portion of the wavefield that will interact with the ramp and, with our fixed ramp height of 30 km, the ramp width parameter w controls the steepness of the ramp. We vary d from 100 to 700 km with a 200 km interval, and w from 100 to 400 km with a 100 km interval, values chosen to capture a wide range of realistic scenarios. Our steepest ramp, that thickens by 30 km over a distance of 100 km (>16° dip), is analogous to the steepest part of the Himalayan Moho ramp, typically 20-25 km vertical change over a 100-km width (e.g. Gao et al., 2016; Nabelek et al., 2009; Shi et al., 2016). The lowest slopes we model, ~4° dip (here, a 30-km ramp spanning 400 km), is more characteristic of eroded mountain belts in which total Moho relief of ~15 km within ~200 km across strike is typical (e.g. Cook et al., 2010). However, it is commonplace for earthquakes to be recorded along raypaths that are oblique, not perpendicular to orogens, and the first-order influence of oblique incidence can be approximated through an increase of ramp width (Bostock & Kennett, 1990). Our gentlest ramp (w = 400 km) is analogous to that seen by an earthquake recorded at 45° obliquity to the Himalayan ramp. We use as our source a thrust earthquake with moment magnitude $M_w = 6$, dip $\delta = 45°$, rake $\lambda = 90°$, and a Gaussian source time function with a half-width of 0.2s. For each combination of ramp parameters w and d, as well as for the reference model, we calculate the SH wavefield due to placement of this source at 3 different depths: 15 km (mid-crust), 35 km (shallow-lid, just below the Moho at the source but within the depth range of the ramp), and 65 km (deeper-lid, below the Moho everywhere in the model).

We constructed our 2D finite-element mesh to balance accuracy and efficiency. Our absorbing boundary conditions, including the thickness of the sponge layers, are set such that 97% of reflections for waves with >0.5 Hz frequency are eliminated (Haindl et al., 2020). We use two elements (10 collocation points) per wavelength in our simulations, and confirmed there was no visible numerical dispersion. Using just less than two elements per wavelength – as needed for the mesh coarsening – also produced no visible difference. We coarsened our mesh using two-refinement transition templates (Anderson et al., 2009) (by "tricking" the built-in mesher) at a depth of 70 km, 10 km below the deepest Moho, to ensure that mesh coarsening and deformation (due to the Moho ramp) do not conflict. Our coarsening strategy enforced two elements per wavelength at the coarsening depth, and more (or slightly fewer) elements per wavelength above (or below) this depth, which resulted in a ~16.8% reduction in the number of elements needed (Supplementary materials S2).

3. Numerical results from the reference model

As a check of our computational setup, and to ensure that *Sn* and *Lg* amplitude variations we discuss later are due only to the presence of the Moho ramp, we first calculate wavefields in a reference model with constant crustal thickness (no ramp) and compare our results to geometrical-spreading results calculated with a full-waveform method for a similar but non-identical 1D earth model for both *Lg* (Yang, 2002) and *Sn* (Yang et al., 2007) (Fig. 4). Both our source depths and distance range are larger than those explored by previous studies, so some discrepancy is expected apart from differences in earth models. We measure the *Sn* and *Lg* amplitudes at each offset as the RMS value over a time window defined by the expected range of group velocities, 4.0 to 4.7 km/s for *Sn* and 3.0 to 3.8 km/s for *Lg*. Our *Sn* windows are picked slightly differently compared to Wang and Klemperer (2021) to minimize overlaps with *Lg* windows at short distances and the mis-categorization of fast *Lg* waves at long distances, and also to account for non-zero source depths while not pre-judging whether an earthquake has a mantle or crustal hypocenter (Supplementary materials S3). All amplitudes are reported as displacements.

Fig. 4a&b show synthetic absolute *Lg* and *Sn* amplitudes, respectively, for the three source depths (15, 35 and 65 km) in the reference model (Moho depth fixed at 30 km), filtered from 1–5 Hz for *Lg* and around 3 Hz (from $3/\sqrt{2}$ to $3\sqrt{2}$ Hz, following Yang et al., 2007) for *Sn* and plotted at 10-km intervals from 200 to 2,000 km epicentral distance. Our frequency filters are 8th-order Butterworth filters. We also plot the relative amplitude decay provided by geometrical spreading models. For *Lg*

$$G_{Lg}(r) = r^{-\gamma}$$

where r is the epicentral distance and $\gamma = 1$ is an empirical constant (Yang, 2002). Yang (2002) modelled a variety of parameters such as source depth (but only tested crustal sources, above the Moho), frequency content, and amplitude-measurement technique, and found γ remained close to 1. *Sn* geometrical spreading is more complicated due to its propagation path (whispering gallery or interference head waves; cf. Avants et al., 2011), and has been modeled with both frequency (f) and distance dependence:

$$G_{Sn}(r,f) = \frac{10^{n_3(f)}}{r_0}\left(\frac{r_0}{r}\right)^{n_1(f)\log(r_0/r)+n_2(f)}$$

where $r_0 = 1$ km and $n_i(f)$ are fixed parameters for a specific two-layer Earth model, calculated by Yang et al. (2007) for a source at 15 km depth in a uniform 40-km thick crust. The frequency dependence of *Sn* geometrical spreading is the reason why we show single-frequency *Sn* in Fig. 4b.

We first note the preferential *Lg* excitation by crustal sources (amplitudes for the crustal source are an order of magnitude greater than for the mantle sources: Fig. 4a) and preferential *Sn* excitation by mantle sources (amplitudes for the 15-km source are an order of magnitude less than for the 65-km source: Fig. 4b). This is reflected by the amplitude ratios *Sn/Lg* for these three source depths, where although they vary as a function of epicentral distance, at each location, *Sn/Lg* for mantle earthquakes is always higher than *Sn/Lg* for crustal earthquakes (Fig. 5). In an ideal 1D model (i.e. the reference model), *Sn/Lg* varies according to epicentral distances in different ways for different source depths, but the common feature is the increase of *Sn/Lg* at long distances, i.e. about 600-1300 km, as noted before (Wang and Klemperer, 2021). The amount of increase is the highest for the 15-km event. For this same 15-km event, there is also a rather large decrease of *Sn/Lg* at shorter distances, whereas the values remain relatively constant for the 35- and 65-km events. The reasons behind this is due to an over-estimation of *Sn* following conventional observation methods, to which we return below while discussing individual *Sn* and *Lg* amplitudes.

The higher-amplitude models (15 and 35 km for *Lg*, 65 and 35 km for *Sn*) vary fairly smoothly with distance (Fig. 4 a&b). The abrupt discontinuities present in the bottom traces of Fig. 4a (65-km trace) and Fig. 4b (15-km trace) are artifacts due to mis-categorizations of the waves in our windowing process. For example, Fig. 4c,d&e show the synthetic seismograms for the 65-km source at 400, 410 and 460 km (triangles in Fig. 4a&b). For this deep source, we expect essentially no *Lg* excitation. However, at 400 km, there are two prominent *Sn* peaks included in the *Lg* window (Fig. 4c), which explains why measured *Lg* amplitudes are unexpectedly high for distances from 200-400 km (Fig. 4a). Measured *Sn* amplitudes are correspondingly lower than the total amplitude within the *Sn* phase but by a smaller proportional amount because the later *Sn* peaks are lower amplitude than the first-arriving *Sn* peak. At 410 km (Fig. 4d), the second *Sn* peak is no longer within the *Lg* window, resulting in a sharp drop of measured *Lg* amplitudes. At 460 km (Fig. 4e) the second *Sn* peak moves into the *Sn* window, resulting in a small proportional increase in measured *Sn* amplitudes for the 65-km source at that distance (Fig. 4b). This artifact (peaks moving in and out of a window) is present for all cases in Fig. 4, but is small for the major phase from each source depth, e.g. the small sinusoidal oscillations in measured *Lg* amplitudes for the 15- and 35-km sources. These small variations were also shown but not explained in previous synthetic studies (Yang, 2002). Although it is important that we fully understand our synthetics, these phenomena have no relevance for real data for which small-scale scatterers will always act to smooth out the strong amplitude peaks seen in our synthetics (and those of Yang, 2002 and Yang et al., 2007). Our observations offer insight into the relationship between physical *Sn* and *Lg* (as defined by propagation waveguides) and observational *Sn* and *Lg* (as defined by group-velocity windows), and into the fundamental inaccuracy of using the same *Sn* and *Lg* velocity windows for different events because these window-bounding velocities are dependent on epicentral distance even for the same source depth in a 1D model (a point also touched on by Aki and Richards, 2002, their Box 7.1). For example, all three wavelets in Fig. 4c are physical *Sn* waves, ordered by their relative mantle and crustal path lengths, and the *Sn* velocity window would need to be extended down to 3.4 km/s to capture all three in the *Sn* window. At just 400-km range, the second and third wavelet appear in the *Lg* window (and also a small wavelet around 127s, Fig. 4c), but with increasing offset all these peaks would have travelled proportionally greater distances in the mantle, gaining higher apparent horizontal group velocity, and be captured in the conventional *Sn* velocity window (Fig. 4d&e).

Working with synthetics it would be possible to measure the apparent group velocity of each arrival and thereby correctly separate *Sn* from *Lg*; but in real data such an approach is likely difficult or impossible. Hence we do not seek to change the conventional observation method, but rather we acknowledge the prevalence of this issue and highlight the irrelevance of fine-tuning velocity windows and the care needed to avoid over-interpreting amplitude measurements. Fortunately, the conventional and tractable *Sn*- and *Lg*-velocity windows method are historically proven to be adequate, especially when only crustal sources are considered. Since our *Sn/Lg* method for identifying mantle earthquakes fundamentally relies on comparisons between potentially crustal and mantle earthquakes, rather than absolute-value *Sn/Lg* thresholds, it is more important to use a simple and unified approach for a group of events (to enable comparisons) than to strive for picking the most accurate windows for individual events, which in practice is also hard to achieve.

Sn synthetic amplitudes (Fig. 4b) exhibit the classic interference head-wave behavior at distances <~1,300 km, in that amplitudes first decrease then increase with distance due to the spherical focusing effect, because at larger distances more energy from waves multiply-reflected at the Moho underside will contribute to the amplitudes, in addition to the direct arrival. The distance at which the *Sn* amplitudes begin to increase and the amount of the increase depends on source depth, and is closest/strongest for the shallowest source (note this is not captured by the geometrical spreading model, black lines in Fig. 4b, as that is an empirical fit based on a crustal source only). For the 15-km source, we expect mostly *Lg* excitation. While our *Sn* window is already shortened compared to Wang and Klemperer (2021), the earliest *Lg* waves could appear within the *Sn* window, which results in the artificial amplitude jump at 820- and 830-km distances (circles in Fig. 4b; Fig. 4f&g). As before, these sudden amplitude changes are due to mis-categorizations, but are likely much smaller in real data due to presence of smoothing effects. However, if these smoothing effects are not accounted for and if the conventional windowing method is followed (Yang et al., 2007) (completely justified if the intent is to study geometrical spreading alone), the amount of *Sn* amplitude increase might be over-estimated at these long offsets due to incorporation of *Lg* waves. This incorporation of *Lg* waves at long offsets potentially explains the earlier rise to larger *Sn* amplitudes, leading to a larger increase of *Sn/Lg*, for the crustal source compared with the mantle sources (Fig. 4b, Fig. 5). For a 35-km source, the smooth amplitude increases from e.g. 900- to 1,000-km distances (squares in Fig. 4b) are due to the spherical focusing effect of interference head waves, as the number of peaks within the *Sn* window is not changed, yet their amplitudes (most notably the third peak) grow larger (Fig. 4h, i). For distances >1,300 km, all three sources have about the same amount of *Sn* energy (Fig. 4b), but their *Lg* energy is vastly different (Fig. 4a) so our method is still very effective at these long offsets. Beyond ~1,300 km our measured amplitudes start to drop, a phenomenon not previously noted because the Yang et al. (2007) study was limited to shorter offsets, but completely reasonable because the spherical focusing effect must eventually wear off, i.e. the multiply-reflected waves at the Moho underside eventually become too small to meaningfully contribute. In all other respects, our 3-Hz *Sn* amplitudes exhibit the same interference head-wave behavior as in Yang et al. (2007). In fact, the fit to our 15-km source (the same source depth simulated by Yang et al. (2007), but in a slightly different earth model) is good (Fig. 4b). The misfit to the deeper sources beyond ~500 km clearly originates from the fact that only a crustal source was considered by Yang et al. (2007), which, combined with other reasons discussed above, led them to an over-estimation of the spherical focusing effect.

We have an almost exact match between our 15- and 35-km sources' synthetic *Lg* amplitudes and the simple *Lg* geometrical spreading model (Yang, 2002), and between our 15-km source's *Sn* synthetic amplitudes and the more complicated *Sn* geometrical spreading model (Yang et al., 2007). The mismatches between our synthetic amplitudes and the previous models can all be understood. This gives us confidence in our modelling approach. Our results cover a larger parameter space and exhibit a greater range of features than previous studies, so already provide useful new information as well as serving as a benchmark against which to test our simulations with a Moho ramp.

4. Numerical results from Moho ramp models

Our parameter-space study includes 48 2.5D crustal-thickening simulations (plus reference simulations). Here, we present a selection of these results (Figs. 6–9) as two groups by first fixing the ramp width $w = 200\ km$ and varying the distance to the start of the ramp d =100, 300, 500 and 700 km; and then by fixing $d = 300\ km$ but varying w = 100 km ~17°, 200 km ~9°, 300 km ~6° and 400 km ~4°. The rest of our numerical results can be found in Supplementary materials S4&5. The amplitudes are measured as discussed in Section 3, except that now our *Sn* amplitudes are measured using the same broader frequency band we use for *Lg*, i.e. 1-5 Hz. We present our results for *Sn* and *Lg* amplitudes (Figs. 6, 9, left and middle columns) as ratios to our reference-model results, aligned by distance relative to the end of the ramps to highlight deviations relative to ramp locations. We show results for *Sn/Lg* (Figs. 6, 9, right columns) relative to epicentral distance and overlaid on reference-model results to highlight the ramp effects and to illustrate the absolute values of *Sn/Lg* for different source depths.

Our interest is in phenomena that have the potential to be recognized and measured in real data. Measurements on synthetics of *Sn* alone, or *Lg* alone, coupled with inspection of synthetic seismograms and compared to the reference (flat Moho) model reveal the physics of wave propagation across Moho ramps. However, in the real world no reference data are available and it is the *Sn/Lg* ratios that, by removing source and receiver dependencies, may allow recognition of source depth with respect to Moho (Wang & Klemperer, 2021).

4.1 Fixed ramp width, $w = 200\ km$

This fixed ramp width, combined with our constant 30-km ramp height, produces a Moho ramp with fixed gradient that is among the steepest in nature (though not the steepest Moho ramp we test, see Supplementary materials S4&5), especially considering that many earthquake-receiver geometries involve oblique incidence onto the ramp, effectively increasing the width of the ramp. We highlight this example, *w* = 200 km, because incidence onto a steep Moho ramp produces the clearest effect on *Sn* and *Lg* amplitudes. For all our simulations (variable *w*, d and source depth *z*) the measured amplitudes (and the *Sn/Lg* ratios) coincide with those from the reference model until the waves reach the ramp apart from tiny numerical errors (e.g. Figs. 6a,d show some symbols at distances -700–0 km slightly below the grey line that represents equality with reference model results) .

4.1.1 Sn amplitudes

For relative *Sn* amplitudes (Fig. 6, left column, a.-c.), one of the most striking features is the focusing of *Sn* waves that starts close to the middle of the ramp and peaks slightly beyond the ramp, for most source depths *z* and distances to ramp d. This phenomenon corresponds to the breaking of the *Sn* waveguide by the Moho ramp (Fig. 7). Before the leading wavefront in the mantle reaches the ramp (at time = 73 s, Fig. 7a), the wavefield is the same as in the reference model, with the same first reflected wave as the leading wavefront and the same first arrival at the surface (Fig. 7a&b). At time = 100s the leading wavefront is at about the middle of the ramp and the first arrival is just beyond the start of the ramp (Fig. 7c). At 100s, the Moho underside transmitted wave has a much larger amplitude in the ramp model than in the reference model (Fig. 7d), and this increased amplitude extends to the surface, representing the start of the *Sn* peak just beyond the start of the ramp. The reasons behind the focusing are twofold. First, the Moho ramp increases the local curvature of the Moho so that deeper energy on the leading wavefront, which in the reference model would refract up at a greater distance (Fig. 7e, blue arrows), refracts up to the surface from the ramp (Fig. 7e, yellow dashed arrows), locally increasing the amount of energy being transmitted into the crust. Second, the incidence angle of this deeper energy changes from almost grazing to a smaller angle ($i_{ramp} < i_{flat}$) (Fig. 7e), which could flip the energy partitioning of the reflected and transmitted waves (Fig. 7f) in favor of transmission (calculated using plane-wave transmission and reflection coefficients, e.g. von Seggern, 2012, which are a good approximation at long distances from the source). The large reflected energy at large incidence angles (Fig. 7f) enables multiple reflections at the Moho underside, and essentially gives rise to the whispering-gallery waveguide. The increase in transmitted energy at decreased incidence angles shows how this waveguide is broken by a Moho ramp. On a seismogram (Fig. 7g), other than the prominently increased amplitude of the first arrival of the ramp model, the effect of the ramp shows up as delays in individual arrivals due to the transmitted waves travelling a longer distance in the crust and travelling at a steeper angle ($r_{ramp} < r_{flat}$, Fig. 7e) leading to a smaller horizontal apparent velocity. In general, the magnitude of this focusing (~2 to 10 times stronger than reference, Fig. 6, left column) is a proxy for how much of the original wavefield interacts with the ramp, which is inversely proportional to d and is largest for the 35-km event (that lies within the vertical extent of the ramp), followed by the 15- and 65-km events. This explains why the 65-km deeper-lid event has about half the focusing strength of the shallow-lid earthquake (35-km) and the mid-crustal earthquake (15-km), which both have similar degrees of focusing. For the 35-km event, the focusing strength is strictly inversely proportional to d, but this is not the case for the 15- and 65-km events for which, for shorter distances to the ramp, more complicated interferences occur that decrease the strength of the *Sn* peak ($d = 100$ for 15-km event, Fig. 6a, and $d = 100, 300$ km for 65-km event, Fig. 6c).

The only case without an *Sn* peak at ramp exit is z=15 km, d=100 km (Fig. 6a). Regardless of ramp width, no crustal sources at this short distance to ramp show an Sn peak, but instead have a large decrease of Sn for longer ranges beyond the ramp exit, and all eventually recover back to close to reference values (Fig. S5-1a). Similarly, *Sn* de-focusing is present for the 65-km events when *d* is short (i.e. 100 and 300 km) (Fig. S5-1c). Unlike the *Sn* focusing peak that is just outside (<200 km) of the ramp exit (left column of Figs. 6, S4-1), or even completely contained inside the ramp region for wider ramps (left column of Figs. S4-2, S4-3), these *Sn* de-focusing regions can extend up to 600 km (Fig. 6a) to 800 km (Fig. 6c), having a broad influence on *Sn/Lg* (e.g. Fig. 6i, d=100 &300 km). It would seem that e.g. if an earthquake with z = 65 km, d = 100 km is measured at ~700-800 km epicentral distance (Fig. 6i), its *Sn/Lg* could be confused with that of an earthquake with z = 15 km and d = 500 km measured at the same distance (Fig. 6g). However, such a confusion requires a careful orchestration of a broad de-focusing zone and a localized *Sn* focusing peak, as well as potentially very different back-azimuths to a particular station to produce the 400-km difference in effective ramp width, and although this could occur in the real data, array-based measurements with varying source-station geometry should be able to mitigate, if not completely avoid, this effect.

The behavior of *Sn* is quite complex in the presence of a crustal-thickening Moho ramp due to its interference head-wave nature. For example, there are also secondary focusing peaks for the 35-km events for all d's (Fig. 6b). However, such specific observations on synthetics are likely too detailed to observe in real data, so we do not further discuss or decipher these phenomena.

Lastly, in contrast to the above-mentioned deviations from the reference model, the other most striking feature, common to all our simulations (Figs. 6, 9, left columns, Supplementary materials, S4&5) is that, at long offsets, *Sn* amplitude always returns to about the same level as the reference model. The mechanism for unperturbed amplitude at long offset is that deeper energy (below the black dot on the wavefront, Fig. 7e) may never interact with the ramp, and hence at long distances this energy is transmitted into the crust as if the Moho had always been flat beneath a thickened 65-km crust. However, arrival delays (Fig. 7g) persist for the ramp model even at long distances beyond the ramp, simply due to the increased travel path in the crust.

4.1.2 Lg amplitudes

Clear *Sn*-to-*Lg* conversion due to Moho thickening can be seen in Fig. 6, middle column, especially for the mantle sources. To first order, when the ramp shape is fixed as in the present case (i.e. height = 30 km and w = 200 km), we expect the degree of conversion to be controlled by the amount of *Sn* excited (positively correlates with source depth z) and the subset of this amount that interacts with the ramp (negatively correlates with d and source depth relative to Moho). Our results indicate that the source depth plays a far more important role. For the 65-km event, all d's share a similar growth pattern for relative Lg amplitude, which increases by >5 times between the ramp start and end, due to *Sn*-to-Lg conversion. At longer offsets, beyond the ramp, the continued gradual increase in relative Lg is due to the decrease in Lg for the reference model for this sub-crustal source (Fig. 4a, magenta triangles) rather than to Lg growth in the ramp models. For the 35-km event, even though the source depth is within the vertical extent of the ramp, *Sn*-to-Lg conversion is relatively modest (Fig. 6e). Except for the source closest to the ramp ($d = 100km$), the relative Lg amplitude increase from the reference model is less than a factor of 2, which means real-world observations perturbed by scatterers and noise could be difficult. Beyond the ramp, after the initial oscillations in relative Lg amplitudes, we see relative amplitudes decrease. This is especially prominent for d = 300 and 500 km, and is subtle for $d = 100$ km perhaps due to more initial wavefield interaction with the ramp, and is not shown for $d = 700km$ because our 2000-km maximum simulation range does not include distances sufficiently far beyond the ramp (Fig. 6e). As ramp width increases, drop-off of amplified Lg becomes even more pronounced (e.g. Figs. S4-2e, S4-3e). The crustal 15-km source (Fig. 6d) does not show this same behavior of relative-amplitude decrease, even though in the reference model Lg decays at the same rate for both 15-km and 35-km sources. The relative-amplitude drops for the 35-km below-Moho source may therefore indicate that the crustal waveguide cannot sustain the increased Lg frequencies that are created by *Sn*-to-Lg conversion, a topic we return to below (see Fig. 8).

Relative Lg amplitude from the 15-km source decreases by factor <2 as we cross the ramp (Fig. 6d). The magnitude of decrease is inversely proportional to ramp distance d because disruption of the Lg waveguide (i.e. the crust) causes Lg de-focusing as illustrated by Kennett (1986). Though this de-focusing must also occur for the 35-km and 65-km sources it is more than compensated for by the strong *Sn*-to-Lg conversion from these deeper sources. For the 15-km mid-crustal source, *Sn*-to-Lg conversion is hard to observe as there is much less initial *Sn* energy that can be potentially converted to Lg (Fig. 4b). However, *Sn*-to-Lg conversion must still be occurring because at larger distances beyond from the ramp relative *Lg* amplitude for all d's increases above 1, indicating extra *Lg* energy than expected if the Moho was uniform.

When *Sn* converts to *Lg* we expect not only amplitude but also frequency effects: the higher-frequency portion of the *Lg* becomes enriched because mode coupling tends to happen at neighboring modes (Fig. 3) (Maupin et al., 1989), thus higher-mode *Sn* tends to excite higher *Lg*-forming modes, which contribute to *Lg* at higher frequencies. We test this with our full-waveform results by comparing the *Lg* wavetrain filtered from 1–5 Hz, as shown thus far, to the same wavetrain filtered 0.1–0.8 Hz. We plot the high-frequency (HF) to low-frequency (LF) ratio (*Lg* HF/LF) (Fig. 8) of the ramp model divided by the reference model, and confirm that *Lg* has a higher frequency component that develops across the ramp due to *Sn*-to-*Lg* conversion. For the 15-km source, although subtle, there is a slight increase of high-frequency content further away from the ramp, which suggests some *Sn*-to-*Lg* conversion for the mid-crustal source. Comparing Fig. 8b&c to Fig. 6e&f we see similar trends, implying that a large part of the increase in *Lg* beyond the ramp is due to the increased HF component from *Sn*-to-*Lg* conversion.

4.1.3 Sn/Lg amplitude ratios

The changes in *Sn* or *Lg* amplitude or frequency content, relative to the reference model, are by a factor typically <2, so can be hard to recognize on real, noisy, data (except for the deepest source at the largest offset, Figs. 6f, 9c). In contrast, the amplitude ratio *Sn/Lg* is a direct measure of the relative strengths of *Sn* and *Lg* amplitude perturbations. *Sn/Lg* in the reference model increases linearly in log amplitude–log distance space for distances from ~600–1,400 km (Figs. 5, 6&9 g-i), confirming our earlier conclusion (Wang and Klemperer, 2021) from analysis of the empirical geometrical spreading models that differ between *Sn* and *Lg*. When the source is at 15 km, *Sn/Lg* largely follows the shape of *Sn* variations (Fig. 6g), because the *Sn*-to-*Lg* conversion is rather weak. Since any ramp only locally perturbs *Sn* amplitude perturbations, if amplitude ratios are measured far enough beyond the ramp, there is virtually no difference between the ramp and the reference models. Hence, crossing a significant Moho ramp (as in the present example) does not affect *Sn/Lg* observations for a crustal earthquake provided the measurements are made sufficiently far beyond the ramp. Exactly how far is sufficient is related to d, and ranges from ~800 km beyond the ramp for $d = 100\ km$ (measured from the red vertical line in Fig.6g) to

~100 km beyond the ramp for $d = 700\ km$ (measured from the cyan vertical line in Fig.6g). More simply, for source depth = 15 km for all d studied here (100 km ≤ *d* ≤ 700 km), a propagation distance of ~1100 km is sufficient to erase most of the ramp effect on *Sn/Lg*: at this distance all symbols coincide with the reference model (grey inverted triangles, Fig. 6g). Lastly, we note that *Sn/Lg* almost nowhere exceeds 0.2 for the 15-km event, marked by a black fiducial line in Fig. 6g,h,i.

In contrast, for the 35-km and 65-km sources, *Sn/Lg* is typically an order-of-magnitude larger than for the crustal source and rarely drops below 0.2 regardless of their variations, except sometimes just beyond the ramp. For the 35-km source (Fig. 6h) we see a combined effect from *Sn* and *Lg* variations, with *Sn* controlled by the local, transient focusing behavior, primarily in the ramp region, and *Sn/Lg* determined largely by *Lg* amplitudes further beyond the ramp. The biggest decrease of *Sn/Lg*, to ~3 times lower than the reference model at epicentral distance ~450 km is for the source that is closest to the ramp start and is due to increased *Lg* amplitudes (Fig. 6h, red circles). For the 65-km source, *Sn/Lg* is primarily controlled by the amplitude variations of *Lg*, and in some cases with small *d*, *Sn* de-focusing. Further away from the ramp there are significant decreases from the reference model by more than an order of magnitude, with the source closest to the ramp again exhibiting the largest decrease. Even though *Sn/Lg* for the two sub-Moho sources (Figs. 6h,i) far beyond the ramp can be smaller than had there been no Moho ramp, *Sn/Lg* remains 5-10 times larger than for the mid-crustal, 15-km, event. Visually, crustal and mantle earthquakes can be largely separated by the black fiducial line (Fig. 6g,h,i). In consequence, the *Sn/Lg* method (Wang & Klemperer, 2021) is robust for all cases tested here, especially if the recording stations are not limited to the ramp region and some measurements are made far beyond the end of the ramp.

4.2 Fixed distance to ramp, $d = 300\ km$

Here, d is fixed, so all ramps start at the same location but they extend out by different distances from $w = 100$, 200, 300 and 400 km (Fig. 9). Since our ramp height is fixed, by increasing w, we are decreasing the steepness of the ramp. As the dip of the Moho ramp decreases, the amplitudes will converge to the reference model.

Many key observations remain the same as in the previous section, including the *Sn* focusing peaks being localized close to the ramp and then recovering to the reference model at long distances (Fig. 9, left column), and the increase of the *Lg* amplitudes (quite subtle for the 15-km mid-crustal event) beyond the ramp that persists to greater distances in the crustal waveguide (Fig. 9, middle column). Most important, even though the *Sn/Lg* ratio for each source depth varies by a factor of ~20 with offset (Fig. 9, right column), *Sn/Lg* for the 35- and 65-km events exceeds *Sn/Lg* for the 15-km event at all distances, typically by a factor of 5–10. Thus – as seen also from Section 4.1 and Fig. 6, right column – *Sn/Lg* ratios are a robust metric for interpreting source depth above or below the Moho even in the presence of a Moho ramp, but particularly beyond the end of the ramp (Fig. 9, right column).

Increases in steepness of the Moho ramp can be thought of as an effective increase in the local Moho curvature. The tighter the curvature, the more intense the upward focusing of *Sn* from the Moho underside (Fig.7, e&f), leading to the most prominent feature in our simulations, the factor of 2–10 increase in *Sn* compared to the reference model vertically above and immediately beyond the ramp (Fig. 9a-c). Fig. 9a-c shows the steepest Moho ramp (w=100km) leads to the strongest and earliest (with respect to epicentral distance) *Sn* amplitude increase, reaching peak *Sn* amplitude just beyond the ramp region (for ramp widths <200 km). For wider ramps ($w = 300$ and 400 km), the peaks are completely within the ramp region. Just as for the models in Fig. 6, these results imply the need for observations across the ramp region if they are to be relevant for real data. The *Lg* energy increase beyond the ramps is again well-aligned with the start of the ramps (Fig.9, d-f) and is most prominent for the 65-km source, and decreases as the source depths decreases. As expected, the gentlest ramp (w=400 km) has the smallest increase in *Lg*. Ramp width seems to have less influence on *Sn* de-focusing. For the 15- and 65-km events, *Sn* de-focusing occurs except for w=100 km when z=15 km (Fig. 9a,c), and for d=100 km, even this exception doesn't hold anymore (Fig. S5-1a,c). However, at d=300 km, Sn de-focusing barely influences *Sn/Lg* (Fig. 9g&i) and the de-focusing becomes completely absent for larger d's (Figs. S5-2, S5-3, left column).

5. Observational results

We study earthquakes recorded by the HiCLIMB array not only for its high data quality, but also because the Moho structure beneath this array is well-studied (Nabelek et al., 2009), providing good definition of the Moho ramp structure (Fig. 2). Along the HiCLIMB profile (IRIS data code XF), the Moho is relatively flat beneath northern India and the Main Frontal Thrust (MFT). The Moho ramp begins about 100 km further north, beneath the Main Central Thrust (MCT). Because the surface trace of the MCT is tortuous due to laterally varying exhumation of a low-angle structure (Martin, 2017), we use a line 100-km north of the MFT as our proxy for the start of the ramp (Fig. 2a). North of the MCT, the Moho deepens from ~45 km to ~65 km over a distance of 150 km to the Yarlung-Zangpo Suture (YZS), a geometry present all along the Himalayan arc (e.g. Gao et al., 2016; Shi et al., 2016). Our data set includes very few southern stations over the northern part of the ramp (usually ≤7 due to limited operating time and noisier data), with most available stations lying further north beyond the YZS. This station distribution offers only a glimpse close to the ramp region to investigate effects on individual *Sn* and *Lg* amplitudes but gives ample opportunity to observe *Sn/Lg* away from the end of the ramp, where synthetics predict its effectiveness. In addition to this main dataset, we also analyzed 4 events from the Gangdese-92 array (Fig. 2). These 4 earthquakes are directly due south of the stations, offering an opportunity to evaluate the influence of oblique incidence for the HiCLIMB events. We use the same *Sn* and *Lg* velocity windows as used in Wang & Klemperer (2021), based on regional observations in our study area, i.e. 4.3-4.8 km/s for *Sn* (McNamara et al., 1995) and 3.1-3.6 km/s for *Lg* (McNamara et al., 1996). The data is bandpass filtered from 1-5 Hz with an 8th order Butterworth filter. We select traces only if either or both *Sn* or *Lg* has a root mean square (RMS) amplitude at least twice as high as that of a noise window, defined to start 30s and end 5s before the *Pn* arrival that we calculate using a constant velocity of 8.1 km/s.

There have been reports of sub-crustal earthquakes beneath southern Nepal and northern India (e.g. Chen & Molnar, 1983; Chen & Yang, 2004; Baur, 2007; Song & Klemperer, 2024) but none are confirmed and counter-claims exist that all earthquakes in these regions are likely intra-crustal (Maggi et al., 2000; Mitra et al., 2005; Priestley et al., 2008). Here, we look at six earthquakes, our 'southern events' (S1–S6 from south to north) with catalog depths 10–62 km (Table 1), recorded on the HiCLIMB array (Nabelek et al., 2009; Fig. 2). The travel paths of S1–

S6 traverse a Moho ramp ~20 km high at distances $50 \leq d \leq 660$ km from the source, and spanning widths measured obliquely along the path $160 \leq w \leq 475$ km. This ramp is smaller than used for some of our synthetics, and the obliquity of raypaths to both the ramp and the HiCLIMB profile means that different recording stations have a different d and w. This varied geometry is beneficial to our method as it helps to avoid systematic errors. HiCLIMB also allows us to compare our six southern events that do traverse the ramp with six earthquakes in northwestern Tibet (Wang & Klemperer, 2021) (Fig. 2) that do not traverse a ramp or indeed any major Moho topography. These six 'northern events', spanning upper-crustal to upper-mantle hypocentral depths (Wang and Klemperer, 2021), were recorded on the same array as the southern events, over roughly the same distance ranges (Supplementary materials S6). The southern and northern events are also similar in magnitudes, ranging from m_b 3.5-4.3, with S4 being the smallest event studied (Table 1). The HiCLIMB array operated only from mid-2004 to end-2005 and seismicity in Northern India is not nearly as prolific as on northwestern Tibet, so our six southern events have a much larger spatial spread than the six northern events.

We look at our data from three perspectives to illustrate the effect of regional waves traversing through a Moho ramp: gross amplitude measurements (Fig. 10), individual seismogram changes across a record section for a given event (Fig. 11), and *Lg* HF/LF, i.e. ratios of *Lg* amplitudes in 1-5Hz (HF) and 0.1-0.8Hz (LF) frequency ranges (Fig. 12).

5.1 *Sn, Lg* amplitudes and *Sn/Lg* ratios

We plot *Sn*, *Lg* and *Sn/Lg* for all our events against station distance north of the end of the ramp (Fig. 10), YZS (Fig. 2), so that the horizontal axis is also a proxy for station locations allowing evaluation of site effects along the array. We normalize the individual *Lg* and *Sn* amplitudes to the first recording station, i.e. the southernmost and northernmost stations for southern and northern events respectively, to remove first-order differences between earthquakes, e.g. their different magnitudes. As we move north from the end of the ramp, epicentral distances increase for the southern events, but they decrease for the northern events (Fig. 10). Alternatively, plotting both groups of events against epicentral distance shows that amplitude generally decreases as epicentral distance increases (Supplementary Fig. S6-1).

Individual *Sn* and *Lg* amplitudes for both southern (Fig. 10 a&b) and northern (Fig.10 c&d) events show remarkable coherence as a function of station location, despite the many differences within and between the two groups. At about 200-300 km beyond the end of the ramp, all measurements (Fig. 10 a-d) are amplified. For the southern events (Fig. 10 a&b), this increase of *Sn* and *Lg* amplitudes superficially resembles the *Sn* peak and increased *Lg* due to *Sn*-to-*Lg* conversion predicted by our modelling (Fig. 6&9, left and middle columns). However, our synthetics show both increases should occur closer to the end of the ramp, reaching their maxima within 100–200 km beyond the end of the ramp. The distance of these maxima from the end of the ramp decreases as ramp width increases, and the obliquity of our source-receiver azimuths to the ramp creates very large effective ramp widths (Table 1). Hence the location of the *Sn* and *Lg* maxima moves even closer to the ramp (Fig. 9, left & middle columns) so the amplitude increases at 200–300 km (Fig. 10a&b) are most unlikely related to traversing the ramp. Indeed, the northern events also show *Sn* and *Lg* amplification at the same stations, implying the peaks at 200–300 km are likely due to local variation in crustal and mantle seismic attenuation (Fig. 2).

Another potential candidate for an *Sn* focusing peak is shown by the few stations that recorded the southern events within and closely adjacent to the ramp (~-50 – +100 km) (Fig. 10a). This is promising because the northern events (Fig. 10c) do not seem to show this peak, and for our southern events that traverse wider (less-steep) ramps the *Sn* peak should occur within the ramp region (Fig. 9, left column). However, we are not confident that this is a true observation of an *Sn* focusing peak because our secondary dataset from the Gangdese-92 array does not show the same feature (Supplementary materials S7). The lack of an *Sn* focusing peak on the Gangdese-92 array, that recorded events with almost perpendicular incidence to the Moho ramp, implies that obliquity of ray-paths to the ramp is likely not the cause of our inability to observe amplitude variations due to the ramp. Observations of individual amplitudes in real data are subject to many variables such as site effects (which likely is strong in the HiCLIMB data based on the coherence seen in Fig.10 a-d), anelastic attenuation, and small-scale heterogeneities that could completely erase the *Sn* and *Lg* ramp-traversal signatures in our synthetics.

The lack of unequivocal observations of ramp effects in the *Sn* and *Lg* amplitude data is disappointing in that we cannot confirm the predictions of our synthetics from an amplitude perspective, but the negligible influence of the ramp on amplitudes is a *positive* result for the ability of the *Sn/Lg* method to distinguish below-Moho from above-Moho earthquakes. The *Sn/Lg* method is robust because it is largely immune to site effects, due to ratioing of the two portions of the same waveform recorded at the same location. Hence *Sn/Lg* ratios (Fig. 10e), unlike individual *Sn* and *Lg* amplitudes (Fig. 10 a-d), do not show any strong correlation with station locations and *Sn/Lg* ratios span similar values for both the southern and the northern earthquakes. We can separate our events into two groups either visually (Fig. 10e), or more quantitatively according to whether at least half of station *Sn/Lg* values are above or below our previous experimental threshold for this region (Wang and Klemperer, 2021), *Sn/Lg* =2. Southern event S1 has a single station and S6 has no station recording *Sn/Lg* >2 (Figs.10, 11): we believe both are crustal earthquakes. In contrast, events S2, S4, and S5 have >50% stations reporting *Sn/Lg* >2 (Figs.10, 11), and visually they behave like northern events WT1 and WT2 (Fig. 10e), which have previously been identified as upper-mantle events (Wang and Klemperer, 2021). This distinction is particularly clear >100 km north of the end of the ramp, and remains clear across most of the northern attenuation zone. Measured across all the stations, southern event S3 has just 39% of measurements with *Sn/Lg* > 2 (Fig. 11), but this rises to 52% if we only consider stations >100 km north of the ramp end (Supplementary materials S8). If the catalog depths for S3 and S4 are correct (~60 km) they are certainly below-Moho events. A full-waveform inversion put S3 at 53 km (Baur, 2007), clearly below the local Moho (Singh et al., 2015; Mitra et al., 2018), a conclusion (weakly) supported by our *Sn/Lg* results. S5 has an arbitrarily assigned depth of 10 km, which is not a useful determinant of the real depth, and based on the *Sn/Lg* data we believe it is in fact a sub-Moho event. Events S1 and S2 have depths ~35 km, around Moho depth (Singh et al., 2015; Mitra et al., 2018) yet our method suggests S2 occurred below the Moho and S1 above it. S6, with a relatively reliable catalog depth of 16.1 km, in the upper crust, is also suggested by our *Sn/Lg* criterion to be a crustal earthquake. These results show that although there is in general a positive correlation of *Sn/Lg* measurements with catalog depth (Song and Klemperer, 2024), there could also be inconsistencies particularly for the case of S5. Because comparison between different Himalayan catalogs shows numerous large depth discrepancies (Song and Klemperer, 2024), and dedicated re-location efforts have found some egregious catalog mis-locations (Craig et al., 2023), we suggest that our determination of

S5 as a sub-Moho earthquake from its *Sn/Lg* character may be more reliable than the assigned catalog depth.

5.2 Record sections

To further investigate the excitation of *Sn* and *Lg* for the southern events, we turn to their record sections (Fig. 11). normalized to the maximum value on each trace to highlight relative amplitude changes within a trace. For our current dataset, the *Sn* and *Lg* windows do not overlap, making their amplitude measurements distinct.

For the four events that we believe are of mantle origin (S2, S3, S4 and S5), clear *Sn* excitation can be observed in the middle part of the record section, at distances >~100 km north of the YZS (the Moho ramp end, labelled as 0 on the upper x-axes of the record sections, Fig. 11). At distances >~400 km beyond YZS there is some diminution of *Sn*, as waves reaching these stations have propagated partly within the region of high *Sn* attenuation (Fig. 2) (Barron & Priestley, 2009). Although *Sn* is clearly strongly excited for S3, the *Sn* energy arrives towards the end of the *Sn* window (Fig. 11). This likely represents a delayed *Sn* arrival rather than incorporation of early *Lg* into the *Sn* window, because early *Lg* should be followed by stronger subsequent *Lg* waves (Fig. 4, f&g) yet the energy in the *Sn* window is already the strongest in the entire record. Because our standard *Sn* window does not capture much of the *Sn* wavetrain for S3, inevitably *Sn/Lg* – calculated as the ratio of the RMS amplitudes of the respective windows – is lower than expected, explaining why only 39% of stations record *Sn/Lg* > 2. This analysis, and the clear increase in *Sn/Lg* for stations ~100 km north of the ramp (Supplementary materials S8) persuade us that S3 is indeed a mantle earthquake. The S1 and S6 record sections are quite different from S2, S3, S4 and S5. Neither S1 nor S6 shows significant *Sn* excitation relative to *Lg* excitation, and they do not show increase in *Sn/Lg* for stations ~100 km north of the ramp (Supplementary materials S8), further corroborating their crustal origin.

The *Lg* wavetrains for shallow events S1 and S6 have rather uniform amplitudes across the HiCLIMB array, but *Lg* varies dramatically for likely below-Moho events S2–S5. A common pattern for S2–S5 is that the southernmost few traces (<~15 km beyond YZS for S2 & S3, and <~100 km beyond YZS for S4 & S5) have *Lg* wavetrains comparable to, or even larger than (S3 and S4) their respective *Sn* wavetrains; then the *Lg* wavetrain becomes uniformly low amplitude further north. We believe this pattern may be a signature of enhanced *Lg* due to *Sn*-to-*Lg* conversion at the ramp. If true, it means *Sn*-to-*Lg* conversions waves may not persist in the crust for long distances, and may attenuate much faster than predicted by our modelling (which uses a scatterer-free crust). Note that the relative change of *Lg* amplitudes across the array that is obvious for events S2–S5 in their record sections, i.e. by comparison within traces (Fig. 11), is not obvious when looking only at the array-normalized *Lg* amplitudes (Fig. 10b), which are essentially the same as absolute amplitudes.

5.3 *Lg* HF/LF, ratio of *Lg* amplitudes at higher and lower frequencies

Another possibility to identify *Sn*-to-*Lg* conversion in real data, instead of relying on observing an increase of *Lg* amplitudes that can be strongly influenced by factors such as site effects (Fig. 10b), is the enrichment of high-frequency (HF) *Lg*. We analyze our twelve HiCLIMB earthquakes and four Gangdese-92 earthquakes exactly as we processed our synthetics. We have no measurements from within the *Sn*-attenuation region (Fig. 2): Gangdese-92 did not extend into this area, and the HiCLIMB stations here all lack high-quality low-frequency (LF) data. For *Lg* from the six southern events (Fig. 12a), we see the southern few stations, in particular those within the ramp region (negative distances), do have a much larger high-frequency component compared to the more northern stations, where *Lg* HF/LF ratio is more uniform. The peaking of *Lg* HF/LF may be smaller for the crustal events (open symbols) than for the mantle events colored symbols, Fig. 12a), as predicted by synthetics (Fig. 8a). For the four events recorded on the Gangdese-92 array, we more clearly see the rise of *Lg* HF/LF associated with the end of the ramp (Fig. 12c) because there are more stations vertically above the Moho ramp. However, we do not see an *Lg* HF/LF peak associated with the end of the Moho ramp for the six northern events (Fig. 12b), because these events have not traversed the ramp.

6. Discussion

We now bring together our numerical and observational results, to address our three main results: the ability to use *Sn/Lg* to recognize below-Moho earthquakes even in the presence of significant crustal thickening, our identification of *Sn*-to-*Lg* conversion in real data, and the value of *Lg* frequency content as another discriminant for continental mantle earthquakes.

Our numerical results (Figs. 6–9) show that significant Moho topography, that locally enhances *Sn* amplitudes and more regionally enhances *Lg* amplitudes, does not strongly influence *Sn/Lg* ratios which remain useful as a comparative measure to separate mantle and crustal earthquakes. The resilience of the *Sn/Lg* method to crustal thickening is clear because *Sn/Lg* ratios for the deeper-lid (65-km) and shallow-lid (35-km) events are always above the *Sn/Lg* ratios for the mid-crustal earthquake (15-km) at the same distance (Figs. 6g,h,i, 9g,h,i). The best separation, an order of magnitude, occurs between our shallow-lid earthquake and our mid-crustal earthquake at stations far beyond the end of the ramp, because of the ramp-transient nature of *Sn* amplitude perturbations and modest *Sn*-to-*Lg* conversion for shallow-lid earthquakes.

Thus our simulation results show we can apply *Sn/Lg* criteria to identify mantle earthquakes regardless of the presence of a Moho-thickening ramp. Observations of *Sn* and *Lg* on the HiCLIMB (Fig. 10 a&b) and Gangdese-92 arrays (Supplementary S7) show less significant effects than our simulation results (Figs. 6&9, right columns) that therefore likely represent the strongest possible scenarios for ramp effects on *Sn/Lg* signatures. Our HiCLIMB events are strongly influenced by site effects and are obliquely incident on the array (though as noted above, this obliquity is likely unimportant), whereas our Gangdese-92 events do not exhibit strong site effects and are nearly in-line with the array. Nonetheless, neither set of events shows either the predicted strong focusing of *Sn* near the end of the ramp nor the predicted sustained increase of *Lg* energy beyond the end of the ramp (Fig.10, Supplementary materials S7). We believe these inconsistencies between data and simulations originate from the absence in our models of small-scale features such as inhomogeneities in the crust or less-smooth Moho topography. Additional small-scale features should spatially smooth a localized feature such as the *Sn* peak (Figs. 6&9, left column), and selectively attenuate the higher-frequency *Lg* in the crust produced by *Sn*-to-*Lg* conversion (Fig. 8), which we discuss more below. Our observations on individual *Sn* and *Lg* waves agree with findings in the North Sea (Mendi et al., 1997) that regional waves are more influenced by small-scale scatterers than large-scale features. Because the largest perturbations from the reference model due to a Moho ramp are the *Sn* peak above and the increased *Lg* beyond the ramp, smoothing out these effects in the real data likely means *Sn/Lg* in the real world is even more robust than predicted by our simulations.

We can directly compare *Sn/Lg* for events traversing one of Earth's largest Moho ramps with *Sn/Lg* for events traversing relatively uniform Moho topography (Figs. 2, 9e). Using a previously established *Sn/Lg* threshold that identified two new below-Moho earthquakes in NW Tibet (Wang and Klemperer, 2021), we can identify four earthquakes (S2, S3, S4, S5) south of the MCT that nucleated below the Moho, including one previously tentatively identified as such (S3=H82 of Baur, 2007) and one that has a nominal (assigned) catalog depth of 10 km (S5). We can similarly show that a different event with a catalog depth close to the Moho (S1) is in fact a crustal event. We emphasize that these conclusions are quite reliable, as they are based on measurements on multiple stations that show *Sn/Lg* significantly larger than the regional low-*Sn/Lg* baseline established for multiple nominally shallow earthquakes in both northern India and in northwestern Tibet.

Sn-to-*Lg* converted waves maybe most easily identified in the frequency domain (Fig. 12), rather than in the amplitude domain (Figs. 10&11), through *Lg* HF/LF. This diagnostic is motivated by early mode-coupling studies (Maupin, 1989) (Fig. 3) and verified with our full-waveform synthetics (Fig. 8). Two groups of events with significant Moho ramp crossing recorded on two separate arrays both exhibit increase of *Lg* HF/LF (Fig. 12a&c) associated with the end of the ramp, but another group of non-Moho-ramp crossing events recorded on one of the same arrays does not show this (Fig. 12b). Hence, we believe *Lg* HF/LF is a rather robust signature of *Sn*-converted-*Lg* waves. This implies that the enhanced *Lg* above and close to the ramp on the record sections of the southern mantle events S2–S5 (Fig. 11) represents *Sn*-to-*Lg* conversions, enriched in high-frequencies. The enriched HF content for *Lg* close to the ramp corroborates our suspicion that small-scale crustal scatterers are the reason we do not see persistent high *Lg* energy after conversion in real data, unlike in the numerical results.

A prominent feature of *Lg* HF/LF is the clear separation of mantle and crustal earthquakes recorded on HiCLIMB (Fig. 12a&b) following our interpretations based on *Sn/Lg* (Fig. 10e), whereas the overlapping of *Lg* HF/LF for the Gangdese events (Fig. 12c) matches their overlapping *Sn/Lg* values (Supplementary materials S7). This can be understood from a normal-mode perspective in that the only *Lg* energy excitable by a mantle earthquake is associated with lower-frequency Airy phases (Knopoff et al., 1973) that could have a displacement/strain eigenfunction sampling the mantle lid to some depths (Wang and Klemperer, 2023, their Fig. 3a), whereas the higher-frequency *Lg* Airy phases have displacement/strain eigenfunctions much more tightly bounded within the crust (Wang and Klemperer, 2023, their Fig. 3b). In our reference model with flat Moho at 30 km, a source 5 km above Moho (z=25 km) has essentially the same *Lg* HF/LF as a source 15 km above Moho (z= 15 km), and both are clearly distinct from *Lg* HF/LF for a source that is 5 km below the Moho (z=35 km). The deeper-lid earthquakes at z = 65 and 95 km have similar *Lg* HF/LF as the crustal earthquakes at short offsets because of the artificial inclusion of *Sn* in our measurement windows (Fig. 4 a,c,d,e), but their *Lg* HF/LF quickly drops beyond ~400-500 km epicentral distance as *Sn* exits the *Lg* window. Hence, like *Sn/Lg*, *Lg* HF/LF is not particularly sensitive to absolute source depths, but rather to their relative position with respect to the Moho as predicted by the normal-mode explanation, so can also be used as a discriminant for sub-Moho earthquakes. This frequency discriminant *Lg* HF/LF is even simpler than *Sn/Lg* because *Lg* HF/LF remains almost constant with epicentral distance (Fig. 13a) (apart from the artificial sinusoidal oscillations due to overlapping *Lg* and *Sn* windows, see Section 3) in contrast to *Sn/Lg* (Fig. 5), and because *Lg* in general is a much simpler wave than *Sn* (i.e. crustal waveguide for *Lg* vs. whispering-gallery waveguide for *Sn*).

We further selected crustal thickening models with small d's and w's in order to capture the effects a Moho thickening ramp can produce on *Lg* HF/LF (Fig. 13 b-e). The difference between *Lg* HF/LF for a crustal and an upper-mantle earthquake is present for all our selected models (beyond ~400–500 km where some *Sn* is present in the *Lg* window), representing among the strongest effects a Moho ramp can produce. We note d has a stronger effect than w in terms of increasing the upper-mantle event's high-frequency *Lg* thereby raising its *Lg* HF/LF, and when $d = 100$ km, the separation with the mid-crustal event is quite small (Fig. 13 b,d&e). In addition, our synthetics show overlapping *Lg* HF/LF for the deeper-lid event and the mid-crustal earthquake in these extreme models (Fig. 13 b–e). However, these are a worst-case because the high-frequency enriched *Lg* due to *Sn*-to-*Lg* conversion, that is persistent at large distances in our synthetics (Fig. 8), in real data fades away quickly after the end of the ramp as observed in data

(Figs. 11&12). Hence, it is unlikely *Lg* HF/LF will be undistinguishable for upper-mantle and mid-crustal earthquakes nor will it mis-classify a deeper-lid earthquake as a crustal earthquake if measurements are made on sufficient stations beyond the end of the ramp. In ongoing work, we are exploring the correlation between *Lg* HF/LF and *Sn/Lg* amplitude-ratio discriminants, and their joint potential to resolve relative location of earthquakes above and below the Moho.

7. Conclusion

We enhanced the code AxiSEM3D to perform 2.5D regional wave simulations across a Moho ramp and achieved a combination of higher frequency ranges and longer propagation distances than other recent studies. Most notably, our modifications enabled checking the representation of an undulated geometry within AxiSEM3D and using this technique to stretch a uniform mesh so that the computed wavefield can be shown at the correct positions, avoiding wavefield distortions that will be visible for simulations at our scale (i.e. regional, vs. global).

We compare our numerical results in a 1D reference model, with flat Moho, with previous studies on *Sn* and *Lg* geometrical spreading to confirm the accuracy of our numerical approach. In addition, with this benchmarking exercise we emphasize the fact that regional-wave arrival windows, as defined by group velocities, cannot be fine-tuned in real data. The windows will always overlap leading to artificial abrupt or oscillatory changes in measured amplitudes and frequencies whenever an *Sn* or *Lg* phase moves in or out of its window (a phenomenon previously noted by Yang (2002)). It is likely that mischaracterization of phases contributed to an over-estimation of *Sn* amplitude increase at ~700–1,300 km by Yang et al. (2007) leading to an inaccuracy in their *Sn* geometrical spreading model (Fig. 4b).

As we vary distance to ramp start d and ramp width w in our crustal thickening model (Fig. 1), the synthetics for *Lg* absolute amplitudes are relatively simple and consistently display sustained increase *Lg* for amplitudes as well as *Lg* HF/LF across the ramp, though to the smallest degree for the crustal source. On the other hand, synthetic *Sn* absolute amplitudes are much more complicated due to its complex propagation path as an interference head wave. Nonetheless, commonalities are present, including the *Sn* focusing peak around the ramp end, and the return to *Sn* amplitudes similar to the reference model at larger distances for almost all parameter ranges tested. These phenomena are closely related to the shape of the *Sn* waveguide (Fig. 7). Even with the presence of these perturbations on individual amplitudes, among all cases tested in our simulations, *Sn/Lg* ratios for mid-crustal earthquakes are persistently lower than for mantle earthquakes on noise-free synthetics, and potential confusions are unlikely when using a recording array with varying source-station geometry.

There are substantial differences between real-world data and synthetics for individual *Sn* and *Lg* absolute (or array-normalized) amplitudes, as in addition to factors like site effects, the real world contains many finer-scale details, such as crustal scatters and irregular Moho/ramp surfaces that tend to average the sharp *Sn* focusing peak and sustained *Lg* amplitude increase seen in our synthetics. We therefore believe our ramp models provide a worst-case scenario for the utility of *Sn/Lg* in the real world as waveforms are smoother in the real world. We verified the effectiveness of *Sn/Lg* through direct comparison with ramp-crossing and non-ramp-crossing events from southern and northwestern Tibet, recorded on the same array with roughly the same epicentral distances (Fig. 10e; Supplementary materials S6), providing strong evidence for four mantle earthquakes in northern India.

Sn-to-*Lg* converted waves are generally hard to recognize from their amplitudes, though not impossible (Fig. 11), but can more easily be identified by the shift of *Lg* frequency content, as shown here with our full-waveform synthetics (Fig. 8) and demonstrated with non-ramp-crossing events (Fig. 12b) and two sets of ramp-crossing events recorded on two different arrays (Fig. 12a,c). *Lg* HF/LF is a promising new discriminant to identify continental mantle earthquakes from their decreased *Lg* HF/LF as predicted by normal-mode theory and verified in both our reference and ramp models (Fig. 13).

Acknowledgements

The authors wish to thank Tarje Nissen-Meyer and Kuangdai Leng for insightful discussions regarding the software package AxiSEM3D. Gangdese-92 data collection was funded by the National Natural Science Foundation of China Grant No. 41374109. This study was funded by Stanford University, and in part by NSF-EAR-1628282.

CREDIT statement

Shiqi Wang:

Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing-Original draft, Writing-Review and editing, Visualization.

Simon Klemperer:

Methodology, Validation, Formal analysis, Investigation, Writing-Review and editing, Visualization, Funding acquisition.

Data Availability Statement

All seismic data analyzed in this paper are available via https://www.fdsn.org/networks/detail/XF_2002 (HiCLIMB data) and at https://doi.org/10.5281/zenodo.10971752 (Gangdese-92 data). Our custom-version AxiSEM3D can be found at https://github.com/axelwang/AxiSEM3D_Modified.

Competing interest

The authors declare no competing interest.

References

Aki, K. and Richards, P.G., 2002. Quantitative seismology. University Science Books.

Al-Attar, D. and Crawford, O., 2016. Particle relabelling transformations in elastodynamics. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, *205*(1), pp.575-593.

Anderson, B.D., Benzley, S.E. and Owen, S.J., 2009. Automatic all quadrilateral mesh adaption through refinement and coarsening. In *Proceedings of the 18th international meshing roundtable* (pp. 557-574). Springer Berlin Heidelberg.

Avants, M., Lay, T., Xie, X.B. and Yang, X., 2011. Effects of 2D Random Velocity Heterogeneities in the Mantle Lid and Moho Topography on P n Geometric Spreading. *Bulletin of the Seismological Society of America*, *101*(1), pp.126-140.

Baker, G.E., Stevens, J. and Xu, H., 2004. Lg group velocity: A depth discriminant revisited. *Bulletin of the Seismological Society of America*, *94*(2), pp.722-739.

Bakir, A.C. and Nowack, R.L., 2012. Modeling seismic attributes of Pn waves using the spectral-element method. *Pure and applied geophysics*, *169*, pp.1539-1556.

Barron, J. and Priestley, K., 2009. Observations of frequency-dependent S n propagation in Northern Tibet. *Geophysical Journal International*, *179*(1), pp.475-488.

Baur, Jan, 2007. Seismotectonics of the Himalayas and the Tibetan Plateau: moment tensor analysis of regional seismograms. MA thesis. Oregon State University.

Bostock, M.G. and Kennett, B.L.N., 1990. The effect of 3-D structure on Lg propagation patterns. *Geophysical Journal International*, *101*(2), pp.355-365.

Bottero, A., Cristini, P., Komatitsch, D. and Asch, M., 2016. An axisymmetric time-domain spectral-element method for full-wave simulations: Application to ocean acoustics. *The journal of the acoustical society of America*, *140*(5), pp.3520-3530.

Cao, S. and Muirhead, K.J., 1993. Finite difference modelling of Lg blockage. *Geophysical Journal International*, *115*(1), pp.85-96.

Červený, V., and R. Ravindra (1971). Theory of Seismic Head Waves, University of Toronto Press, Toronto.

Chen, W.P. and Molnar, P., 1983. Focal depths of intracontinental and intraplate earthquakes and their implications for the thermal and mechanical properties of the lithosphere. Journal of Geophysical Research: Solid Earth, 88(B5), pp.4183-4214.

Chen, W.P. and Yang, Z., 2004. Earthquakes beneath the Himalayas and Tibet: Evidence for strong lithospheric mantle. *Science*, *304*(5679), pp.1949-1952.

Cook, F.A., White, D.J., Jones, A.G., Eaton, D.W., Hall, J. and Clowes, R.M., 2010. How the crust meets the mantle: Lithoprobe perspectives on the Mohorovičić discontinuity and crust–mantle transitionThis article is one of a series of papers published in this Special Issue on the theme Lithoprobe—parameters, processes, and the evolution of a continent. *Canadian Journal of Earth Sciences*, *47*(4), pp.315-351.

Craig, T.J., Jackson, J.A., Priestley, K. and McKenzie, D., 2011. Earthquake distribution patterns in Africa: their relationship to variations in lithospheric and geological structure, and their rheological implications. *Geophysical Journal International*, *185*(1), pp.403-434.

Craig, T.J., Jackson, J., Priestley, K. and Ekström, G., 2023. A Cautionary Tale: examples of the mis-location of small earthquakes beneath the Tibetan plateau by routine approaches. *Geophysical Journal International*, *233*(3), pp.2021-2038.

Drake, L.A., 1972. Love and Rayleigh waves in nonhorizontally layered media. *Bulletin of the Seismological Society of America*, *62*(5), pp.1241-1258.

Dziewonski, A.M. and Anderson, D.L., 1981. Preliminary reference Earth model. *Physics of the earth and planetary interiors*, *25*(4), pp.297-356.

Furumura, T., Hong, T.K. and Kennett, B.L., 2014. Lg wave propagation in the area around Japan: observations and simulations. *Progress in Earth and Planetary Science*, *1*, pp.1-20.

Gao, R., Lu, Z., Klemperer, S.L., Wang, H., Dong, S., Li, W. and Li, H., 2016. Crustal-scale duplexing beneath the Yarlung Zangbo suture in the western Himalaya. *Nature Geoscience*, *9*(7), pp.555-560.

Haindl, C., Leng, K. and Nissen-Meyer, T., 2021. A 3D complexity-adaptive approach to explore sparsity in elastic wave propagation. *Geophysics*, *86*(5), pp.T321-T335.

Kebeasy, T.R.M. and Husebye, E.S., 2003. A finite-difference approach for simulating ground responses in sedimentary basins: quantitative modelling of the Nile Valley, Egypt. *Geophysical Journal International*, *154*(3), pp.913-924.

Kennett, B.L.N., 1972. Seismic waves in laterally inhomogeneous media. *Geophysical Journal International*, *27*(3), pp.301-325.

Kennett, B.L.N., 1984. Guided wave propagation in laterally varying media—I. Theoretical development. *Geophysical Journal International*, *79*(1), pp.235-255.

Kennett, B.L.N., 1986. Lg waves and structural boundaries. *Bulletin of the Seismological Society of America*, *76*(4), pp.1133-1141.

Knopoff, L., Schwab, F. and Kauselt, E., 1973. Interpretation of Lg. *Geophysical Journal International*, *33*(4), pp.389-404.

Laske, G., Masters, G., Ma, Z. and Pasyanos, M., 2013, April. Update on CRUST1. 0—A 1-degree global model of Earth's crust. In *Geophysical research abstracts* (Vol. 15, No. 15, p. 2658). Vienna, Austria: EGU General Assembly.

Leng, K., Nissen-Meyer, T. and van Driel, M., 2016. Efficient global wave propagation adapted to 3-D structural complexity: a pseudospectral/spectral-element approach. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, *207*(3), pp.1700-1721.

Leng, K., Nissen-Meyer, T., Van Driel, M., Hosseini, K. and Al-Attar, D., 2019. AxiSEM3D: broad-band seismic wavefields in 3-D global earth models with undulating discontinuities. *Geophysical Journal International*, *217*(3), pp.2125-2146.

Li, D., Helmberger, D., Clayton, R.W. and Sun, D., 2014. Global synthetic seismograms using a 2-D finite-difference method. *Geophysical Journal International*, *197*(2), pp.1166-1183.

Maggi, A., Jackson, J.A., Priestley, K. and Baker, C., 2000. A re-assessment of focal depth distributions in southern Iran, the Tien Shan and northern India: Do earthquakes really occur in the continental mantle?. *Geophysical Journal International, 143*(3), pp.629-661.

Martin, A.J., 2017. A review of definitions of the Himalayan Main Central Thrust. *International Journal of Earth Sciences*, *106*, pp.2131-2145.

Maupin, V., 1988. Surface waves across 2-D structures: a method based on coupled local modes. *Geophysical Journal International*, *93*(1), pp.173-185.

Maupin, V., 1989. Numerical modelling of Lg wave propagation across the North Sea Central Graben. *Geophysical Journal International*, *99*(2), pp.273-283.

McNamara, D.E., Owens, T.J. and Walter, W.R., 1995. Observations of regional phase propagation across the Tibetan Plateau. *Journal of Geophysical Research: Solid Earth*, *100*(B11), pp.22215-22229.

McNamara, D.E., Owens, T.J. and Walter, W.R., 1996. Propagation characteristics of Lg across the Tibetan Plateau. *Bulletin of the Seismological Society of America*, *86*(2), pp.457-469.

Mendi, C.D., Ruud, B.O. and Husebye, E.S., 1997. The north sea Lg-blockage puzzle. *Geophysical Journal International*, *130*(3), pp.669-680.

Menke, W.H. and Richards, P.G., 1980. Crust-mantle whispering gallery phases: A deterministic model of teleseismic Pn wave propagation. *Journal of Geophysical Research: Solid Earth*, *85*(B10), pp.5416-5422.

Mitra, S., Priestley, K., Bhattacharyya, A.K. and Gaur, V.K., 2005. Crustal structure and earthquake focal depths beneath northeastern India and southern Tibet. *Geophysical Journal International*, *160*(1), pp.227-248.

Mitra, S., Priestley, K.F., Borah, K. and Gaur, V.K., 2018. Crustal structure and evolution of the Eastern Himalayan plate boundary system, Northeast India. *Journal of Geophysical Research: Solid Earth*, *123*(1), pp.621-640.

Mousavi, S.M., Cramer, C.H. and Langston, C.A., 2014. Average QLg, QSn, and observation of Lg blockage in the continental margin of Nova Scotia. *Journal of Geophysical Research: Solid Earth*, *119*(10), pp.7722-7744.

Nábělek, J., Hetényi, G., Vergne, J., Sapkota, S., Kafle, B., Jiang, M., Su, H., Chen, J., Huang, B.S. and Team, T.H.C., 2009. Underplating in the Himalaya-Tibet collision zone revealed by the Hi-CLIMB experiment. *Science*, *325*(5946), pp.1371-1374.

Nissen-Meyer, T., 2022, *personal communications*.

Odom, R.I., 1986. A coupled mode examination of irregular waveguides including the continuum spectrum. *Geophysical Journal International*, *86*(2), pp.425-453.

Oliver, J. and Ewing, M., 1957. Higher modes of continental Rayleigh waves. *Bulletin of the Seismological Society of America*, *47*(3), pp.187-204.

Patton, H.J. and Walter, W.R., 1993. Regional moment: magnitude relations for earthquakes and explosions. *Geophysical Research Letters*, *20*(4), pp.277-280.

PDE Bulletin, 2024. U.S. Geological Survey National Earthquake Information Center (NEIC) Preliminary Determination of Epicenters (PDE) Bulletin, last accessed 03/01/2024 at URL *http://earthquake.usgs.gov/data/comcat/catalog/us/*

Priestley, K., Jackson, J. and McKenzie, D., 2008. Lithospheric structure and deep earthquakes beneath India, the Himalaya and southern Tibet. *Geophysical Journal International*, *172*(1), pp.345-362.

Prieto, G.A., Froment, B., Yu, C., Poli, P. and Abercrombie, R., 2017. Earthquake rupture below the brittle-ductile transition in continental lithospheric mantle. *Science Advances*, *3*(3), p.e1602642.

Przybilla, J. and Korn, M., 2008. Monte Carlo simulation of radiative energy transfer in continuous elastic random media—three-component envelopes and numerical validation. *Geophysical Journal International*, *173*(2), pp.566-576.

Przybilla, J., Korn, M. and Wegler, U., 2006. Radiative transfer of elastic waves versus finite difference simulations in two-dimensional random media. *Journal of Geophysical Research: Solid Earth*, *111*(B4).

Rodgers, A.J., Petersson, N.A., Pitarka, A., McCallen, D.B., Sjogreen, B. and Abrahamson, N., 2019. Broadband (0–5 Hz) fully deterministic 3D ground-motion simulations of a magnitude 7.0 Hayward fault earthquake: Comparison with empirical ground-motion models and 3D path and site effects from source normalized intensities. *Seismological Research Letters*, *90*(3), pp.1268-1284.

Rodgers, A.J., Pitarka, A., Pankajakshan, R., Sjögreen, B. and Petersson, N.A., 2020. Regional-scale 3D ground-motion simulations of Mw 7 earthquakes on the Hayward fault, northern California resolving frequencies 0–10 Hz and including site-response corrections. *Bulletin of the Seismological Society of America*, *110*(6), pp.2862-2881.

Sanborn, C.J., Cormier, V.F. and Fitzpatrick, M., 2017. Combined effects of deterministic and statistical structure on high-frequency regional seismograms. *Geophysical Journal International*, *210*(2), pp.1143-1159.

Schulte-Pelkum, V., Monsalve, G., Sheehan, A.F., Shearer, P., Wu, F. and Rajaure, S., 2019. Mantle earthquakes in the Himalayan collision zone. *Geology*, *47*(9), pp.815-819.

Schwab, F., Kausel, E. and Knopoff, L., 1974. Interpretation of Sa for a shield structure. *Geophysical Journal International*, *36*(3), pp.737-742.

Schulte-Pelkum, V., Monsalve, G., Sheehan, A.F., Shearer, P., Wu, F. and Rajaure, S., 2019. Mantle earthquakes in the Himalayan collision zone. *Geology*, *47*(9), pp.815-819.

Seismological Bulletins of the Indian National Center for Seismology, 2024. Last accessed 04/08/2024 at URL *https://seismo.gov.in/seismological-bulletins*.

Sens-Schönfelder, C., Margerin, L. and Campillo, M., 2009. Laterally heterogeneous scattering explains Lg blockage in the Pyrenees. *Journal of Geophysical Research: Solid Earth*, *114*(B7).

Shi, D., Zhao, W., Klemperer, S.L., Wu, Z., Mechie, J., Shi, J., Xue, G. and Su, H., 2016. West–east transition from underplating to steep subduction in the India–Tibet collision zone revealed by receiver-function profiles. *Earth and Planetary Science Letters*, *452*, pp.171-177.

Singh, A., Singh, C. and Kennett, B.L.N., 2015. A review of crust and upper mantle structure beneath the Indian subcontinent. *Tectonophysics*, *644*, pp.1-21.

Stephens, C. and Isacks, B.L., 1977. Toward an understanding of Sn: Normal modes of Love waves in an oceanic structure. *Bulletin of the Seismological Society of America*, *67*(1), pp.69-78.

Szenicer, A., Leng, K. and Nissen-Meyer, T., 2020. A complexity-driven framework for waveform tomography with discrete adjoints. *Geophysical Journal International*, *223*(2), pp.1247-1264.

Tolstoy, I., 1956. Resonant frequencies and high modes in layered wave guides. *The Journal of the Acoustical Society of America*, *28*(6), pp.1182-1192.

Tromp, J., 1994. A coupled local-mode analysis of surface-wave propagation in a laterally heterogeneous waveguide. *Geophysical Journal International*, *117*(1), pp.153-161.

van Driel, M., Krischer, L., Stähler, S.C., Hosseini, K. and Nissen-Meyer, T., 2015. Instaseis: instant global seismograms based on a broadband waveform database. *Solid Earth*, *6*(2), pp.701-717.

von Seggern, D.H., 2012. Exploring reflection and transmission coefficients in elastic media. *Math. J.*, 14, p.50.

Wang, H., Ni, S., Jin, P., Liu, W., Xu, X., Pan, C., Zhu, H. and Xu, H., 2017. Anomalous Pn Amplitudes through the Southeastern Tarim Basin and Western Tien Shan along Two Profiles: Observations and Interpretations. *Bulletin of the Seismological Society of America*, *107*(2), pp.760-769.

Wang, S. and Klemperer, S.L., 2021. Love-wave normal modes discriminate between upper-mantle and crustal earthquakes: Simulation and demonstration in Tibet. *Earth and Planetary Science Letters*, *571*, p.117089. https://doi.org/10.1016/j.epsl.2021.117089.

Wang, S. and Klemperer, S.L., 2024. Gangdese-92 Data [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.10971752.

Xie, J., 1996. *Synthetic and Observational Study of Pn Excitation and Propagation in Central Asia*. Saint Louis Univ MO Dept of Earth and Atmospheric Sciences.

Xie, X.B. and Lay, T., 2017a. Effects of laterally varying mantle lid velocity gradient and crustal thickness on Pn geometric spreading with application to the north Korean test site. *Bulletin of the Seismological Society of America*, *107*(1), pp.22-33.

Xie, X.B. and Lay, T., 2017b. Frequency-dependent effects of 2D random velocity heterogeneities in the mantle lid on Pn geometric spreading. *Bulletin of the Seismological Society of America*, *107*(1), pp.482-488.

Yang, X., 2002. A numerical investigation of Lg geometrical spreading. *Bulletin of the Seismological Society of America*, *92*(8), pp.3067-3079.

Yang, X., Lay, T., Xie, X.B. and Thorne, M.S., 2007. Geometric spreading of Pn and Sn in a spherical Earth model. *Bulletin of the Seismological Society of America*, *97*(6), pp.2053-2065.

Yang, Z. and Chen, W.P., 2010. Earthquakes along the East African Rift System: A multiscale, system-wide perspective. Journal of Geophysical Research: Solid Earth, 115(B12).

Zhang, M. and Wen, L., 2013. High-precision location and yield of North Korea's 2013 nuclear test. *Geophysical Research Letters*, *40*(12), pp.2941-2946.

Zhu, L. and Helmberger, D.V., 1996. Intermediate depth earthquakes beneath the India-Tibet collision zone. *Geophysical Research Letters*, *23*(5), pp.435-438.

Figure and table legends

Table 1. Earthquakes recorded on the HiCLIMB array. Southern events are named S1–S6. These events nucleated in northern India and cross a significant Moho ramp before being recorded by HiCLIMB (Fig. 2). The distance to the ramp (*d*) and ramp-width (*w*) are shown as ranges because of the different azimuth (hence obliquity to the ramp) from each earthquake to the southern and northern limits of the HiCLIMB stations (Fig. 2). The six events with no values for *d* and *w* comprise our 'northern' events that do not cross significant Moho topography before reaching the stations (Wang and Klemperer, 2021). Magnitude and depth data are from PDE, 2024. Values in parentheses from Baur (2007). Italicized hypocentral locations and depths are from the Seismological Bulletins of the Indian National Center for Seismology.

Fig. 1. Computational model and a representative wavefield. Computational region extends to 2000 km in range and 230 km in depth. Thick black line shows the Moho, which is at 30 km on the left side, and transitions smoothly to 60 km through a 30-km high ramp whose width (*w*) and distance from source (*d*) are labelled. Small red stars represent the 3 source depths (*z*) we study, 15, 35 and 65 km, respectively. Thin black lines represent the top and bottom of the mantle low-velocity zone (LVZ) at 80 and 220 km. In this example the source is at 65 km depth. A snapshot wavefield (transverse component, filtered 1-5 Hz) is plotted at time 235.5s with amplitude shown in the color bar on lower left, showing multiply-reflected and interfering regional wave trains. The wavefield in the crust is complex as it is a combination of multiple reflections from the Moho top-side (*Lg*), under-side (*Sn*), as well as from just below the LVZ (*Sa*). The absence of visible reflections from the bottom of the computational domain, despite the clearly visible reflections from the bottom of the LVZ, demonstrates the performance of our absorbing boundary condition.

Fig. 2. Earthquakes and stations. (a) Earthquakes in India south of the Main Frontal Thrust (MFT) and in the Bhutan Himalaya (red stars, red labels S1–S6; Table 1) recorded on the HiCLIMB array (Nabelek et al., 2009) (purple triangles) after crossing a ~15–20 km high Moho ramp from thinner to thicker crust. Earthquakes in northwestern Tibet (black stars) (Wang and Klemperer, 2021) are recorded on the same array but their paths do not cross significant Moho undulations. Moho depths are interpolated from CRUST1.0 to 0.05° (Laske et al., 2013). Gangdese-92 array (Shi et al., 2015) (yellow triangles) recorded four nominally deep earthquakes ~ due south of the array (yellow stars, yellow labels G1-G4) (Supplementary materials table S7-1). YZS: Yarlung-Zangpo, BNS: Banggong-Nujiang, JRS: Jinsha River sutures. H: Himalaya, L: Lhasa, Q: Qiangtang, SG: Songpan-Ganzi terranes. MFT: Main Frontal thrust, KKF: Karakoram fault, KXF: Karakax fault. White dashed lines border a well-known attenuation zone for *Sn* (e.g. Barron and Priestley, 2009). Black double arrow indicates approximate distance from YZS to the attenuation zone. YZS represents the ending of the Moho ramp, while two thick green lines represent the Moho ramp beginning directly south of the arrays at the approximate location for

Main Central Thrust (MCT), which is too tortuous to show on our map (see main text). (b). Cartoon crustal and Moho cross-section along white solid line shown in (a), redrawn after Nabelek et al., 2009 based on their receiver function analysis on the HiCLIMB array.

Fig. 3. Normal-mode-coupling results. Transmission amplitude coefficients for Rayleigh waves due to perpendicular incidence onto a North Sea graben-type model, visualized from Figure 4 of Maupin (1989). The matrix is symmetric; for convenience we label each row as representing the incidence of a pure mode and each column as the converted amplitude with the amplitude coefficients representing the degree of partitioning of energy due to incidence onto a large-scale Moho depth variation. The calculations are done for 1 Hz, at which a strict separation (shown as black dashed lines) can be made between *Lg* (mode numbers ≤ 11) and *Sn* (mode numbers ≥ 12). Note the amplitude coefficients are typically large along the diagonal (no mode conversion), and are very small in the upper right and lower left sections of the figure as separated by the dashed lines. Because *Lg*-to-*Sn* coupling is strongest into the lowest *Sn* modes (12, 13, 14) *Lg*-to-*Sn* coupling preferentially excites the lower frequencies of *Sn*. For example, looking at the row for mode 9, the squares in columns 1-11 represent mode coupling to other *Lg*-forming normal modes, though most of the energy remains as mode 9 (highlighted with thick black border). Across the dashed line, squares in columns 12-25 represent mode coupling into *Sn*-forming normal modes, leading to *Lg*-to-*Sn* conversion with the strongest coupling to mode 13 (highlighted with dashed border), a low mode number for *Sn* normal modes. Similarly, *Sn*-to-*Lg* coupling is dominantly from the lower *Sn* modes (e.g. 12, 13, 14) to the higher *Lg*-forming-modes (e.g. 8, 9, 10, highlighted with dotted lines, contributing dominantly to higher-frequency *Lg* Airy phases.

Fig. 4. *Sn* and *Lg* in reference model (no ramp). Transverse-component displacements are shown. (a) *Lg* amplitude filtered 1–5 Hz, for three source depths, “mid-crustal” (15 km), “shallow-lid” (35 km) and “deeper-lid” (65 km). (b) *Sn* amplitude at 3 Hz for the same three source depths. Black lines in (a) and (b) are the best-fit models of Yang (2002) (*Lg*) and Yang et al. (2007) (*Sn*) that only predict relative amplitudes as a function of distance, so are set to be equal to our results at 200-km distance for *Lg*, and 300 km for *Sn* (the starting modelling distance in Yang et al. (2007). Our extrapolation of the Yang (2002) and Yang et al. (2007) formulae beyond the distance range they studied leads to large misfits at large offsets. Both (a) and (b) are log-log, amplitude vs. distance. Seismograms for symbols with black border are shown in (c)-(i) with corresponding labels. Two red vertical lines bound the *Sn* windows, and cyan lines bound the *Lg* windows. Full *Lg* windows are not shown for (f)-(i) as the focus there is on the *Sn* window. Seismogram amplitudes shown are absolute values without normalizations, in units of nanometers (nm).

Fig. 5. *Sn/Lg* in reference model (no ramp). *Sn/Lg* for three source depths, 15-km (mid-crust), 35-km (shallow-lid) and 65-km (deeper-lid) are clearly separated at all epicentral distances,

despite their individual variations with offset. A black fiducial line at *Sn/Lg* =0.2 further illustrates separation of crustal and mantle earthquakes.

Fig. 6. *Sn, Lg* amplitudes and *Sn/Lg* with varying distance to ramp start d but fixed ramp width w = 200 km. Rows, top to bottom, display results when the source is mid-crustal (15 km), shallow-lid (35 km) and deeper-lid (65 km). Columns, left to right, show amplitudes of *Sn* and *Lg* relative to the reference model, and *Sn/Lg*. *Left and middle columns* are plotted with data aligned at the ramp, with its beginning marked as a vertical dashed black line and end marked as a vertical solid black line (0 on the horizontal axis). A grey horizontal line at 1 marks no deviation from reference-model results. The vertical axis is plotted in $\log_{10}$ scale while the horizontal axis is linear. Note because of the ramp alignment and a fixed total simulation range, larger d has a shorter distance covered beyond the end of the ramp. *Right column* plots *Sn/Lg* against epicentral distance and superimposed on the reference-model results (grey inverted triangles). Each colored bar represents the end of the ramp for the correspondingly colored symbol (e.g. the red bar marks the end of the ramp at 300 km epicentral distance for d = 100 km (red circles), and its ramp starts outside the range of the plots). The total ramp ranges for the other cases are shown between the vertical lines (e.g. for d = 300 km (blue diamonds) the ramp range is between the red and blue bars). Both the vertical and horizontal axes are in $\log_{10}$ scale. A solid black line at *Sn/Lg* = 0.2 (as in Fig. 5) in all rows in the right column shows that despite the variability within each plot, *Sn/Lg* for our mantle earthquakes (h) and (i) is greater than *Sn/Lg* for our crustal earthquake (g) at every common offset.

Fig. 7. Disruption of *Sn* waveguide by the Moho ramp. (a)-(d) Representative wavefield snapshots (for z = 35 km, corresponding to the blue diamonds in Fig. 6 b,e,h) bandpass filtered 1-5 Hz and shown with same color scale for the amplitudes. (a) and (c): d = 300 km, w = 200 km; (b) and (d): reference model (flat Moho). UT: displacement on the transverse component, in meters. Moho and top of LVZ are marked by thick and thin black lines. At time = 73 s, before the wavefield interacts with the ramp, the ramp model (a) and reference model (b) show the same wavefield, with the first arrival due to a rather weak transmitted wave through the Moho from the leading strong sub-Moho wavefront, showing an effective *Sn* waveguide. At time = 100 s, after the wavefield in the ramp model starts to interact with the ramp, the transmitted wave becomes much stronger in the ramp model (c) than in the reference model (d) (red arrows) corresponding to the onset of the *Sn* peak right after the vertical black line (0 km, end of the ramp) in Fig. 6b. (e) Schematics of sub-Moho wavefront interacting with a flat Moho (black line/blue raypath) and with a Moho with a thickening ramp (grey line/dashed yellow raypath). (f) calculated energy partitioning for a transverse S-wave incident on the Moho from below. In (e), black dot on the red wavefront represents a point slightly below the flat Moho that will contribute to *Sn* for the reference model where the purple arrow intercepts the Moho with incidence angle, i_{flat} close to 90°, suggesting most of the energy is reflected back below the Moho (dashed black curve in f), representing the *Sn* waveguide, while a smaller amount is transmitted into the crust at a smaller angle r_{flat}. The introduction of a Moho ramp reduces these angles to i_{ramp} and r_{ramp}, as shown by the yellow dashed arrows, and sharply increases the amount of energy transmitted into the crust (black curve in f). Because r_{ramp} is smaller than

r_{flat}, the horizontal velocity (apparent velocity) of *Sn* is reduced in the ramp region. Points on the wavefront deeper than the black dot will not interact with the ramp, but will enter the thickened crust beyond the end of the ramp, thus explaining the recovery of *Sn* amplitude at distances further away from the end of the ramp. (g) seismogram at 510 km (10 km beyond the ramp and approximately corresponding to the largest amplitude peak in Fig. 6b) for the reference model (black line) and ramp model (blue line) with *Sn* and *Lg* windows marked by red and cyan lines, showing the large growth of the first-arrival *Sn* wave and the phase delays experienced by the ramp model.

Fig. 8. Change of *Lg* frequency content as a result of *Sn*-to-*Lg* conversion. Same as in the middle column of Fig. 6, but the vertical axis '*Lg* HF/LF' is the ratio of high-frequency (HF, 1–5 Hz) *Lg* to low-frequency (LF, 0.1–0.8 Hz) *Lg* of the ramp model divided by the equivalent ratio for the reference model. For all panels, the horizontal axis is linear while the vertical axis is in log scale.

Fig. 9. *Sn*, *Lg* amplitudes and *Sn/Lg* with varying ramp width *w* but fixed ramp distance *d* = 300 km. Figure organization as for Fig. 6, except for the left and middle columns the vertical colored bars represent the start of the ramps for the correspondingly colored symbols. The end of the ramp is aligned for all of these cases at 0 km and marked by a solid black line, as in Fig. 6. For the right column, the beginning of the ramp is marked by a dashed black line and the end of the ramp is marked by a solid colored line for the corresponding colored symbol. For example, the ramp region for the red circles is within the dashed black line and the solid red line. For all panels, the vertical axis is in log scale. Horizontal axes are linear for parts (a)–(f), and log scale for (g), (h), (i).

Fig. 10. *Sn*, *Lg* amplitudes and *Sn/Lg* as a function of south-north distance from ramp end. Southern events are shown with colored symbols and northern events with black symbols. Events interpreted as mantle earthquakes are shown with solid symbols and crustal events with open symbols. The beginning of the ramp, perpendicular to the array, is shown with a dashed black line and the end of the ramp is shown with a solid black line. The southern limit of the *Sn* attenuation zone is marked with a magenta dashed line. (a)-(d) Individual *Sn* and *Lg* amplitudes for the southern and northern events, respectively. Data points are aligned vertically for each individual station location. (e) Comparison of *Sn/Lg* for the northern and southern events. This very different combined group of events can be clearly separated by high and low *Sn/Lg*, especially >~100 km beyond the end of the ramp (Supplementary materials S7). For all panels, the horizontal axis is linear while the vertical axis is in log scale.

Fig. 11. Record sections of the southern events. Top four events (S2–S5) are interpreted here as below-Moho earthquakes, and bottom two events (S1, S6) as crustal events. For each event, an upper panel shows *Sn/Lg* measured at each station, on a linear scale from 0-8, with our

arbitrary threshold *Sn/Lg* = 2 shown as a grey line. The event code along with the percentage of stations that registered a *Sn/Lg* > 2 are labelled in the upper panel. For each event, between the two panels is each station's distance north of the YZS, recognized as the end of the Moho ramp. The bottom panels show trace-normalized amplitudes for each event, with *Sn* window colored red, *Lg* window cyan, and the noise window green. The traces are shown with a reduction velocity of 4 km/s. Traces are displayed south to north with epicentral distance shown beneath each record section. Thick yellow line marks stations south of YZS (i.e. within the ramp), and thick magenta line marks stations within the *Sn* attenuation zone.

Fig. 12. High-frequency (HF) to low-frequency (LF) ratio of *Lg* waves in data. Moho ramp beginning is marked by a black dashed line while its ending is marked by a black solid line. The start of the *Sn* attenuation zone is shown by a magenta dashed line. The horizontal axis are distances aligned at the end of the Moho ramp. The vertical axis shows the *Lg* HF/LF in a log scale. (a) southern events. (b) northern events. Symbol styles are as in Fig. 10 (open symbols: crustal events; closed symbols: mantle earthquakes). (c) events recorded on the Gangdese-92 array (not categorized as crustal or mantle because we lack comparison events). Note different vertical scale compared with (a)&(b).

Fig. 13. High-frequency (HF) to low-frequency (LF) ratio of *Lg* waves in reference and select ramp models. (a) *Lg* HF/LF for five source depths with the reference model with a flat Moho at 30 km. Two of the sources are located within the crust, one in the shallow-lid, and two deeper within the mantle. (b) & (c) *Lg* HF/LF for ramp models with fixed w=100 km, testing the effect of increasing d. (d)&(e) *Lg* HF/LF for ramp models with fixed d=100 km, testing the effect of increasing w. All panels are log-log. Dashed and solid black lines indicate the start and end of the ramp, when located beyond 200 km. A fiducial line at *Lg* HF/LF = 0.7 is drawn to emphasize the separation of crustal and mantle events for all panels.

Tables

Code name	Date, time	Location (N°, E°)	Magnitude (m_b)	Catalog depth (km)	Distance to ramp, d (km)	Effective ramp width, w (km)
S1	2005-07-26, 18:27:05	23.27, 91.41 *23.281, 91.516*	4.0	38.1$\pm$ 27.4 *10*	560-660	200-280
S2	2005-05-03, 00:38:57	25.76, 91.06 *26.078, 91.033*	4.3	33.6$\pm$? *33*	285-385	210-280
S3 (H82)	2004-08-04, 02:09:21	25.92, 90.26 *25.865, 90.333*	4.2 (4.1, M_w)	61.7$\pm$ 10.8 *20* (53$\pm$?)	250-330	180-260
S4	2005-05-27, 22:12:20	26.14, 87.21 *26.170, 87.685*	3.5	57.7$\pm$ 12 *15*	270-280	130-160
S5	2004-11-24, 22:35:42	27.33, 90.94 *27.337, 90.875*	4.0	10$\pm$? *10*	100-140	160-380
S6	2004-08-09, 08:18:18	27.58, 91.80 *27.547, 91.718*	4.1	16.1$\pm$? *14.9*	50-120	195-475
WT1	2005-05-19, 05:43:30	35.63, 78.38	4.2	97.6 $\pm$ 14.1	-	-
WT2	2005-06-20, 22:52:26	36.23, 77.92	3.9	77.9 $\pm$ 8.4	-	-
WT3	2005-03-03, 15:07:39	35.65, 77.85	3.7	57.5 $\pm$ 16.8	-	-
04-251	2004-09-07, 04:01:05	35.72, 78.25	4.2	7.6 $\pm$ 26.8	-	-
04-291	2004-10-17, 15:35:45	35.20, 77.67	4.3	15$\pm$?	-	-
05-201	2005-07-20, 10:54:49	35.34, 77.79	4.2	10$\pm$?	-	-

Table 1

Figures

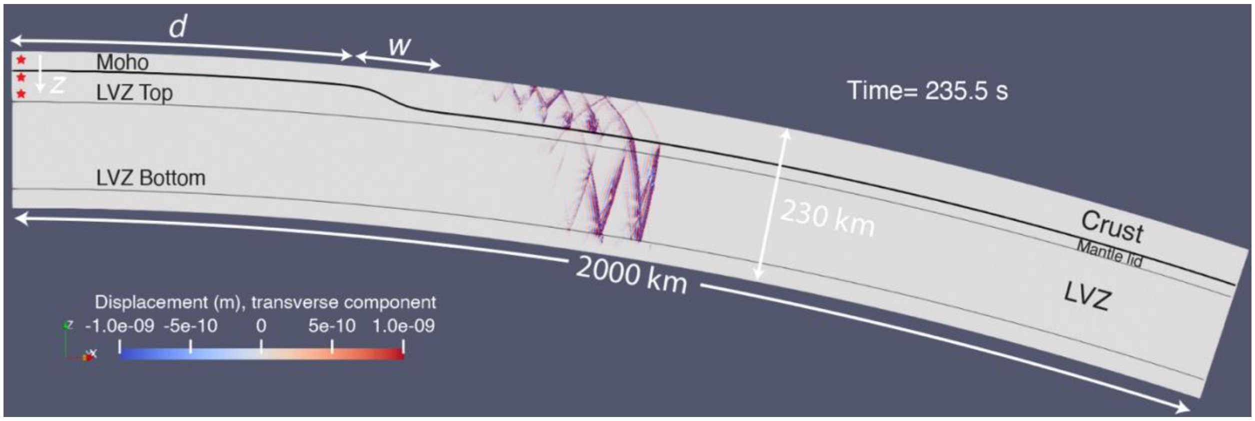


Fig. 1

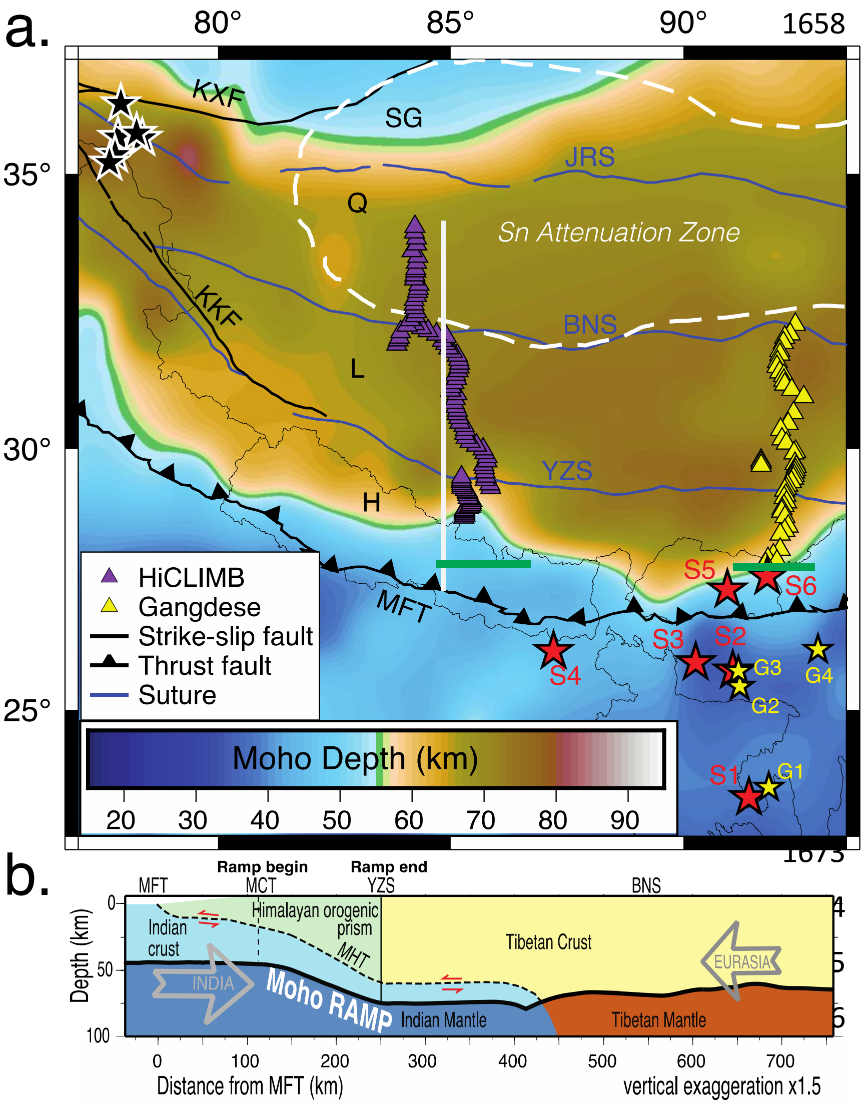
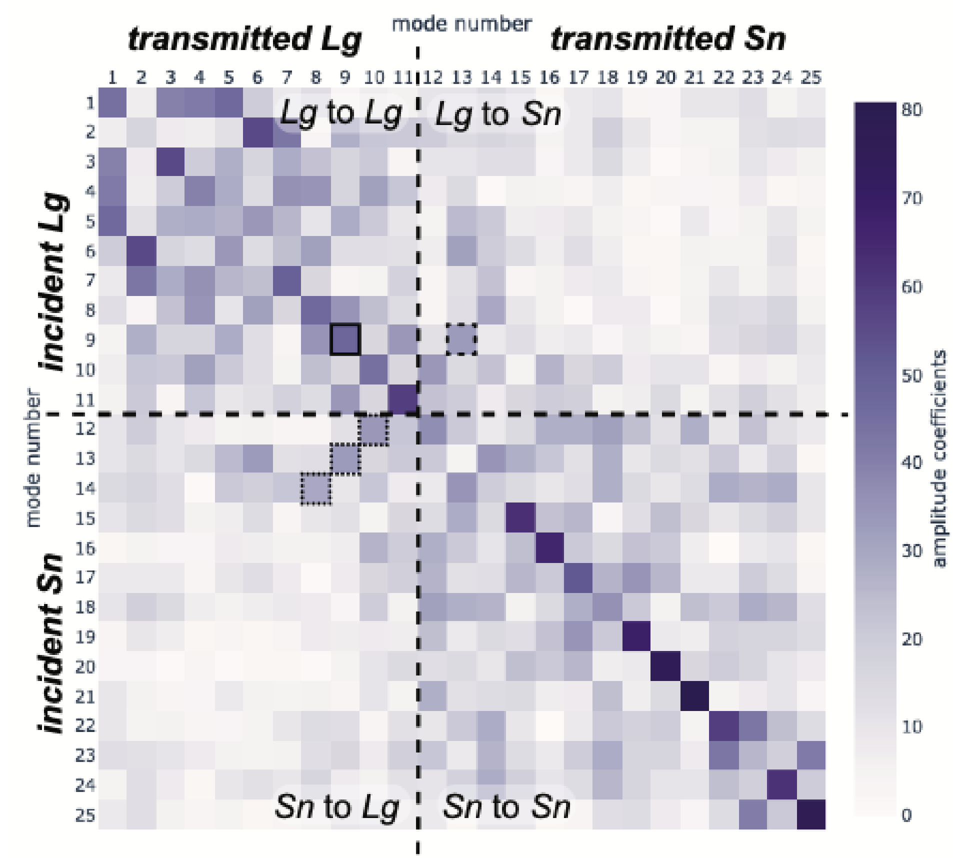
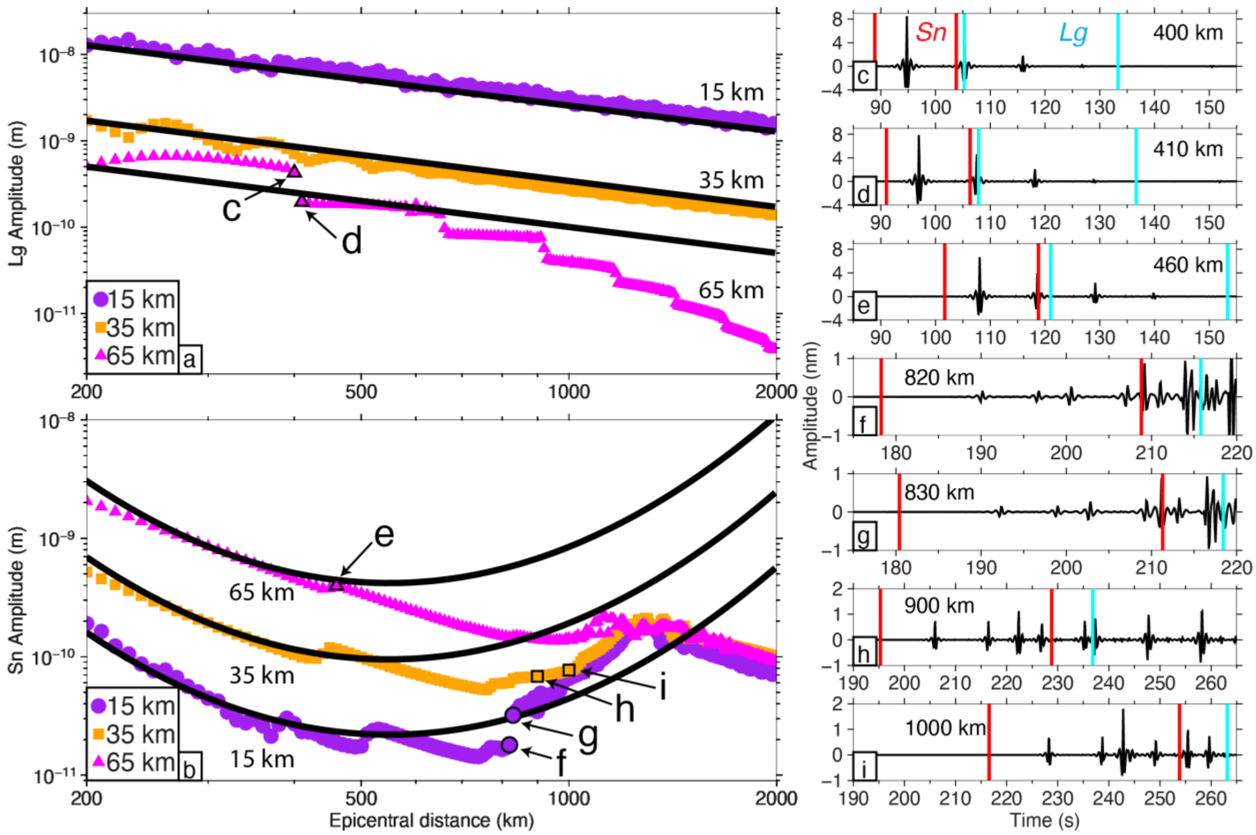

a.
80°
85°
90°
35°
30°
25°
KXF
SG
JRS
Q
Sn Attenuation Zone
KKF
BNS
L
YZS
H
MFT
HiCLIMB
Gangdese
Strike-slip fault
Thrust fault
Suture
S1
S2
S3
S4
S5
S6
G1
G2
G3
G4
Moho Depth (km)
20 30 40 50 60 70 80 90
b.
MFT
Ramp begin
MCT
Ramp end
YZS
BNS
Depth (km)
Indian crust
Himalayan orogenic prism
MHT
Tibetan Crust
INDIA
EURASIA
Moho RAMP
Indian Mantle
Tibetan Mantle
Distance from MFT (km)
vertical exaggeration x1.5

Fig. 2



transmitted Lg
mode number
transmitted Sn
Lg to Lg
Lg to Sn
incident Lg
mode number
incident Sn
Sn to Lg
Sn to Sn
amplitude coefficients

Fig. 3



Lg Amplitude (m)
15 km
35 km
65 km
c
d
a
Sn Amplitude (m)
e
f
g
h
i
b
Epicentral distance (km)
Sn
Lg
400 km
410 km
460 km
820 km
830 km
900 km
1000 km
Amplitude (nm)
Time (s)

Fig. 4

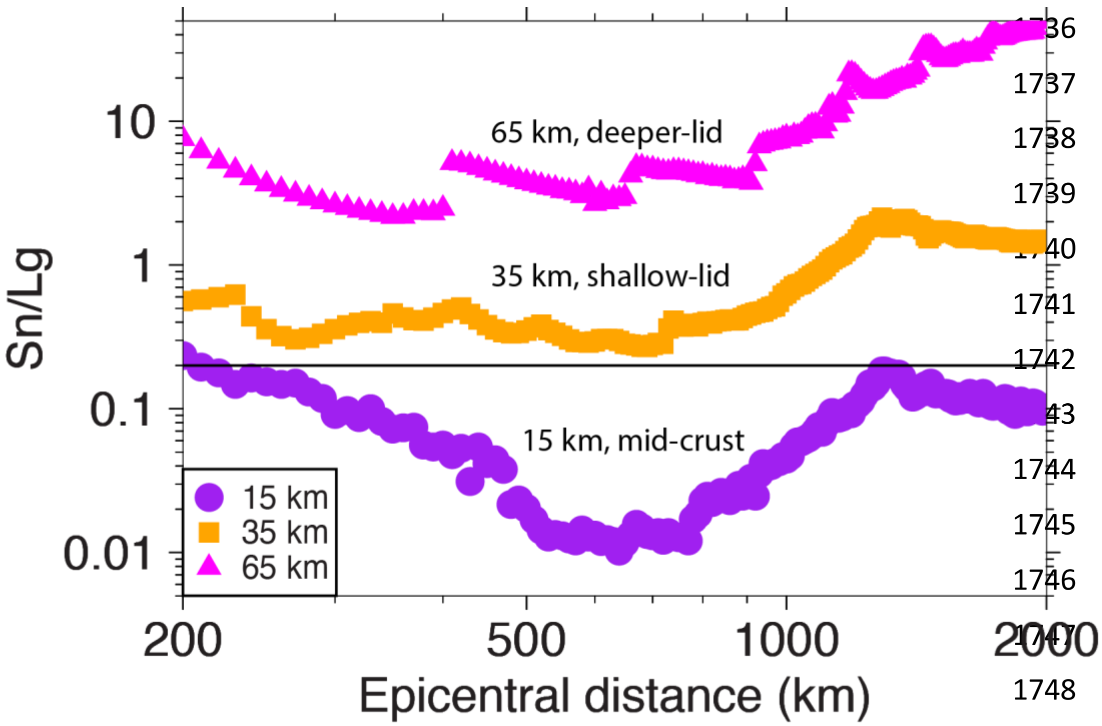

65 km, deeper-lid
35 km, shallow-lid
15 km, mid-crust
15 km
35 km
65 km
10
1
0.1
0.01
Sn/Lg
200
500
1000
2000
Epicentral distance (km)

Fig. 5

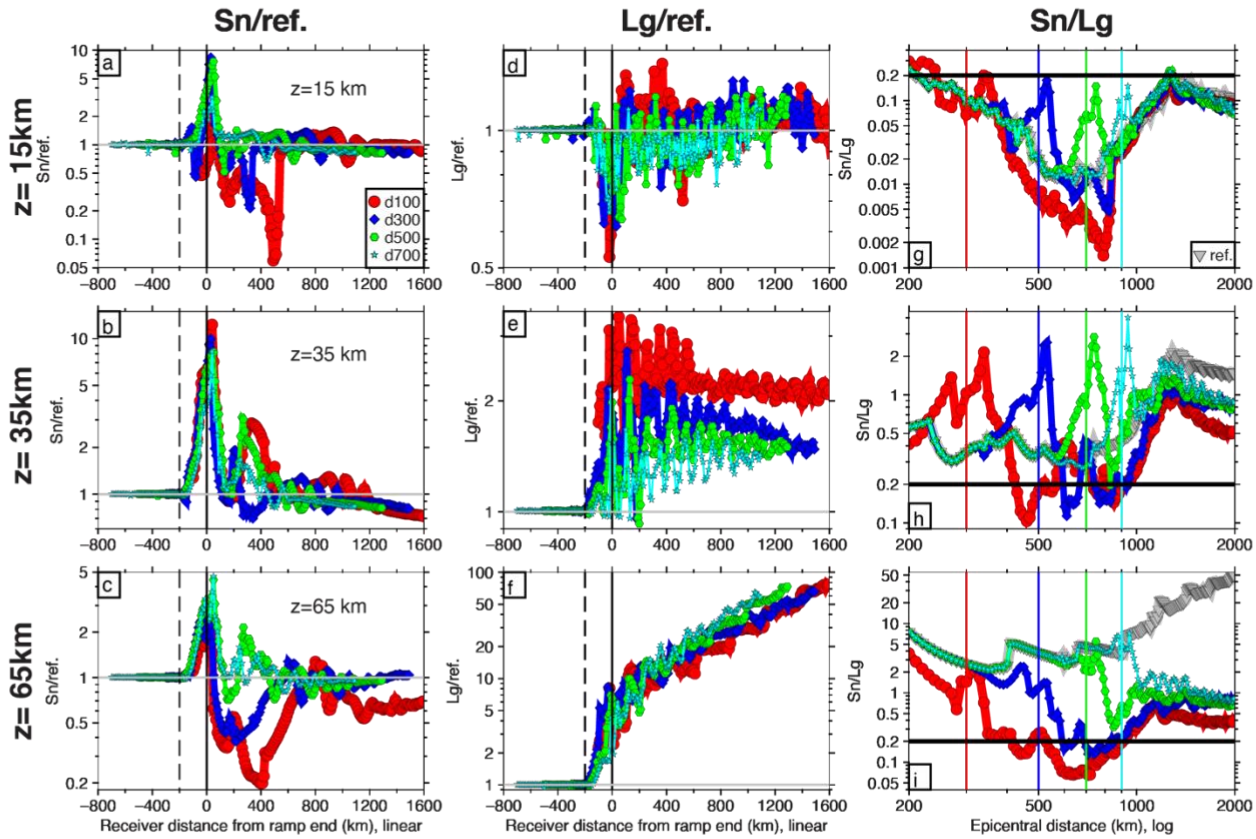

Sn/ref.
Lg/ref.
Sn/Lg
z= 15km
z= 35km
z= 65km
z=15 km
z=35 km
z=65 km
a
b
c
d
e
f
g
h
i
d100
d300
d500
d700
ref.
Receiver distance from ramp end (km), linear
Epicentral distance (km), log

Fig. 6

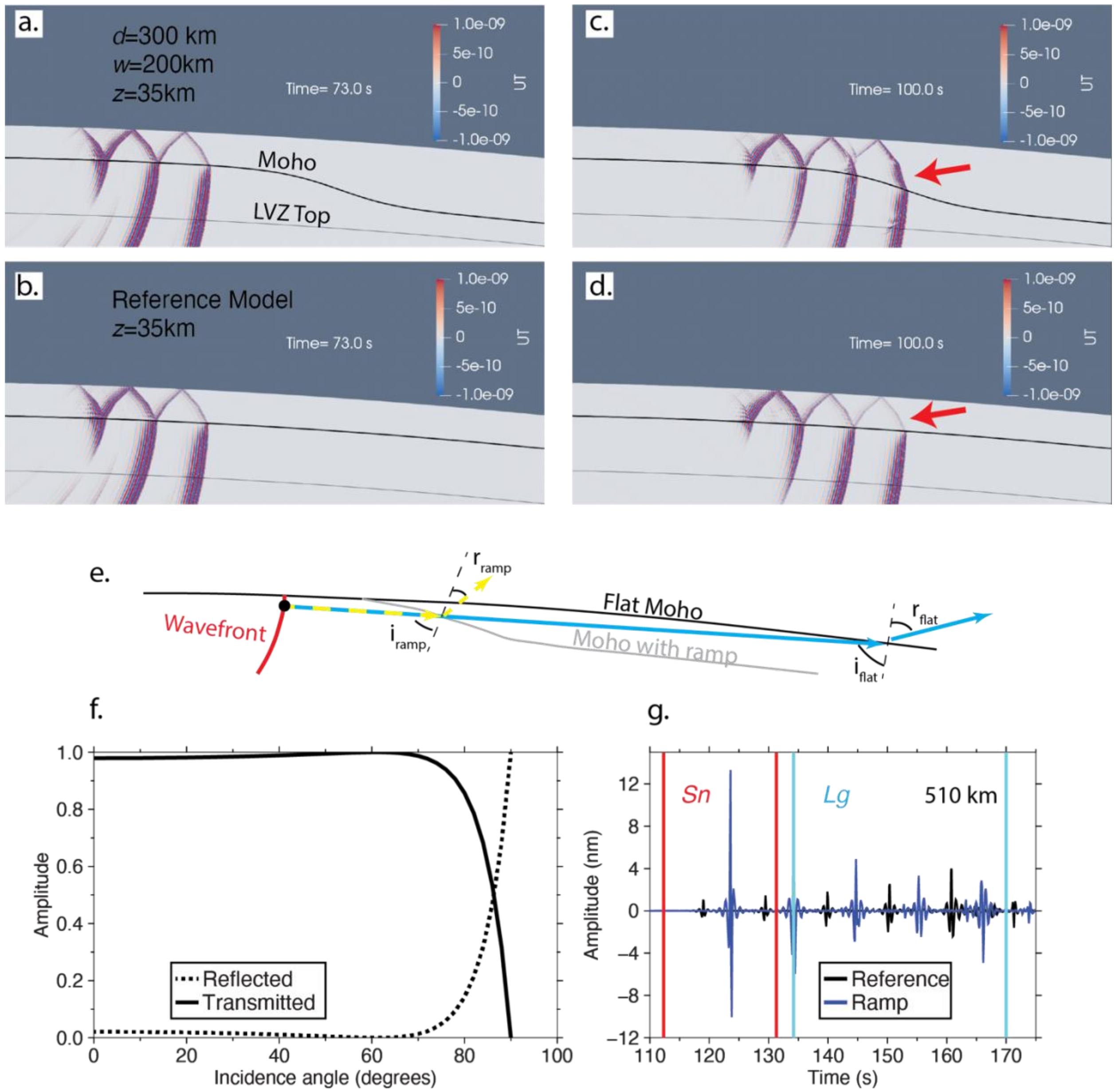

a.
d=300 km
w=200km
z=35km
Time= 73.0 s
Moho
LVZ Top
b.
Reference Model
z=35km
Time= 73.0 s
c.
Time= 100.0 s
d.
Time= 100.0 s
1.0e-09
5e-10
0
-5e-10
-1.0e-09
e.
Wavefront
Flat Moho
Moho with ramp
r ramp
i ramp
r flat
i flat
f.
Amplitude
Incidence angle (degrees)
Reflected
Transmitted
g.
Sn
Lg
510 km
Amplitude (nm)
Time (s)
Reference
Ramp

Fig. 7

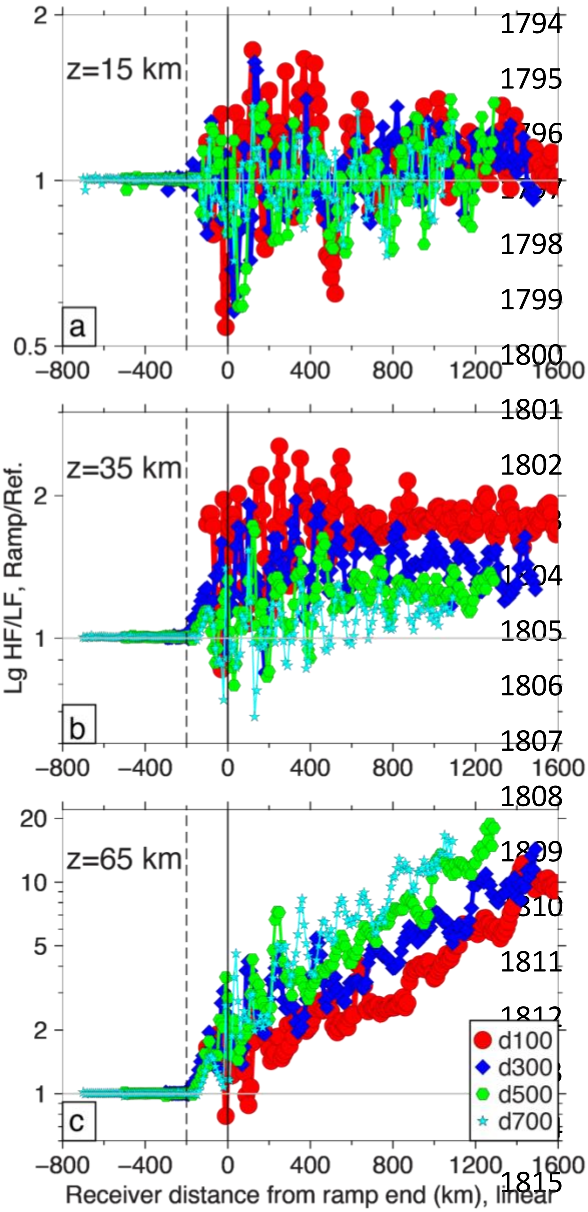

z=15 km
a
z=35 km
b
z=65 km
c
Lg HF/LF, Ramp/Ref.
Receiver distance from ramp end (km), linear
d100
d300
d500
d700
−800
−400
0
400
800
1200
1600
0.5
1
2
5
10
20

Fig. 8

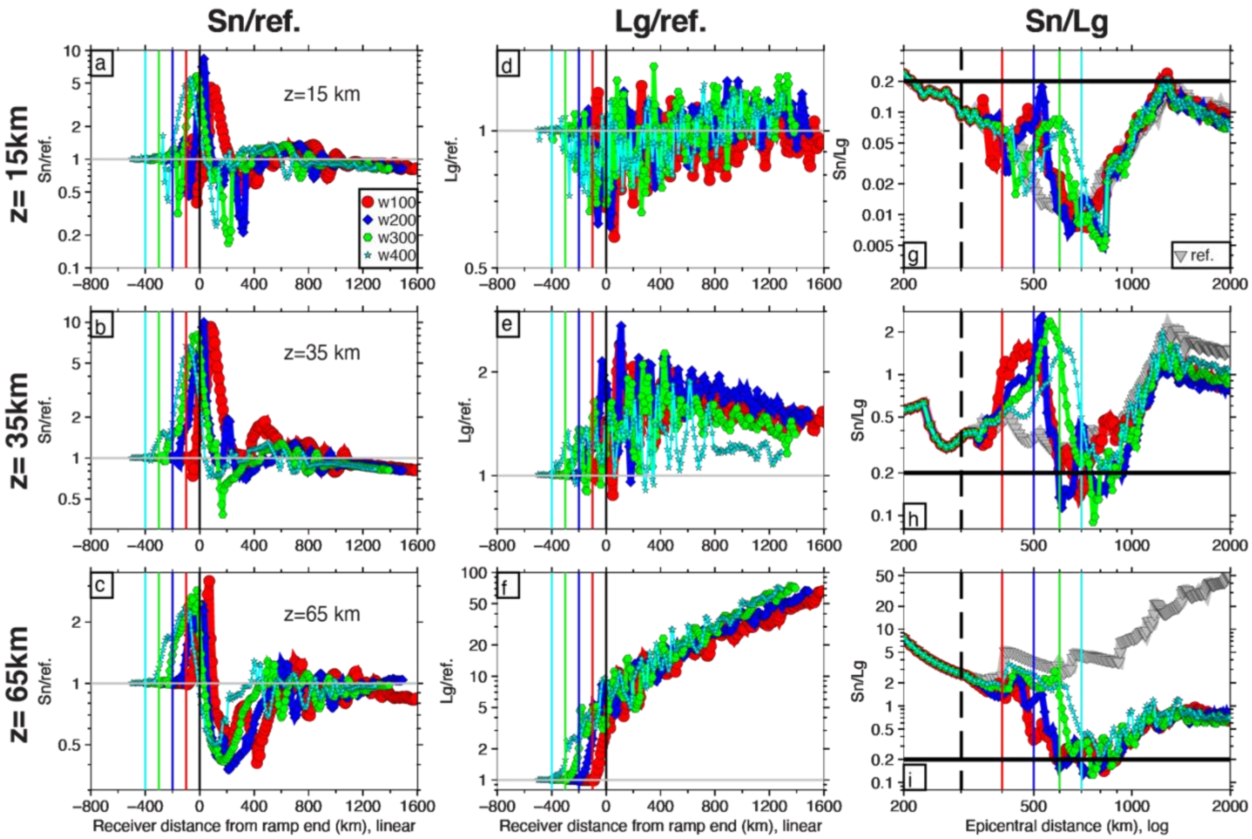
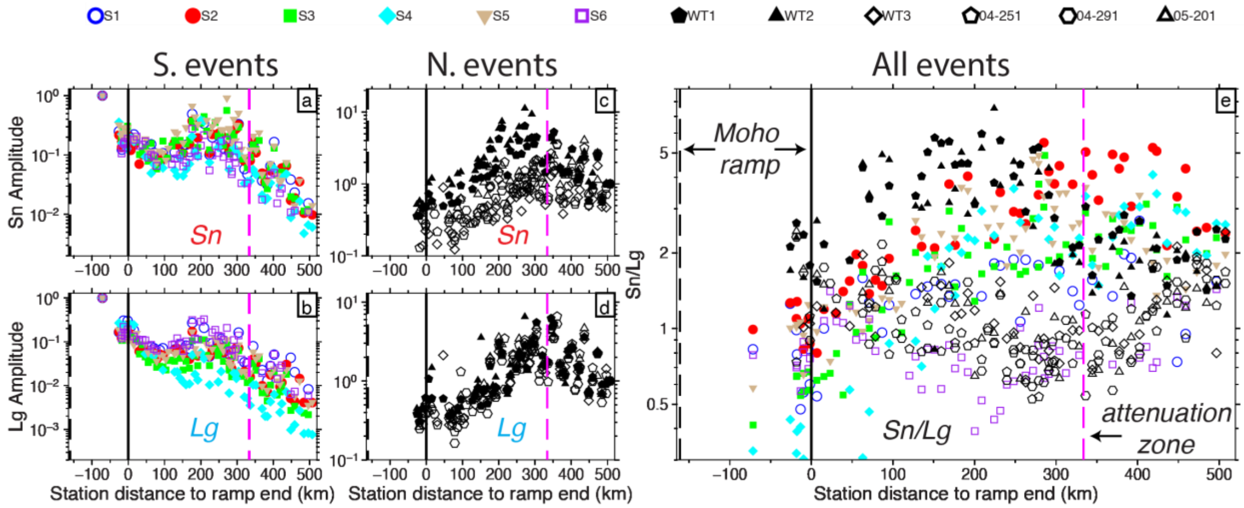

Sn/ref.
Lg/ref.
Sn/Lg
z= 15km
z= 35km
z= 65km
a
b
c
d
e
f
g
h
i
z=15 km
z=35 km
z=65 km
w100
w200
w300
w400
ref.
Sn/ref.
Lg/ref.
Sn/Lg
Receiver distance from ramp end (km), linear
Receiver distance from ramp end (km), linear
Epicentral distance (km), log

Fig. 9



S1
S2
S3
S4
S5
S6
WT1
WT2
WT3
04-251
04-291
05-201
S. events
N. events
All events
a
b
c
d
e
Sn Amplitude
Lg Amplitude
Sn
Lg
Sn/Lg
Moho ramp
attenuation zone
Station distance to ramp end (km)

Fig. 10

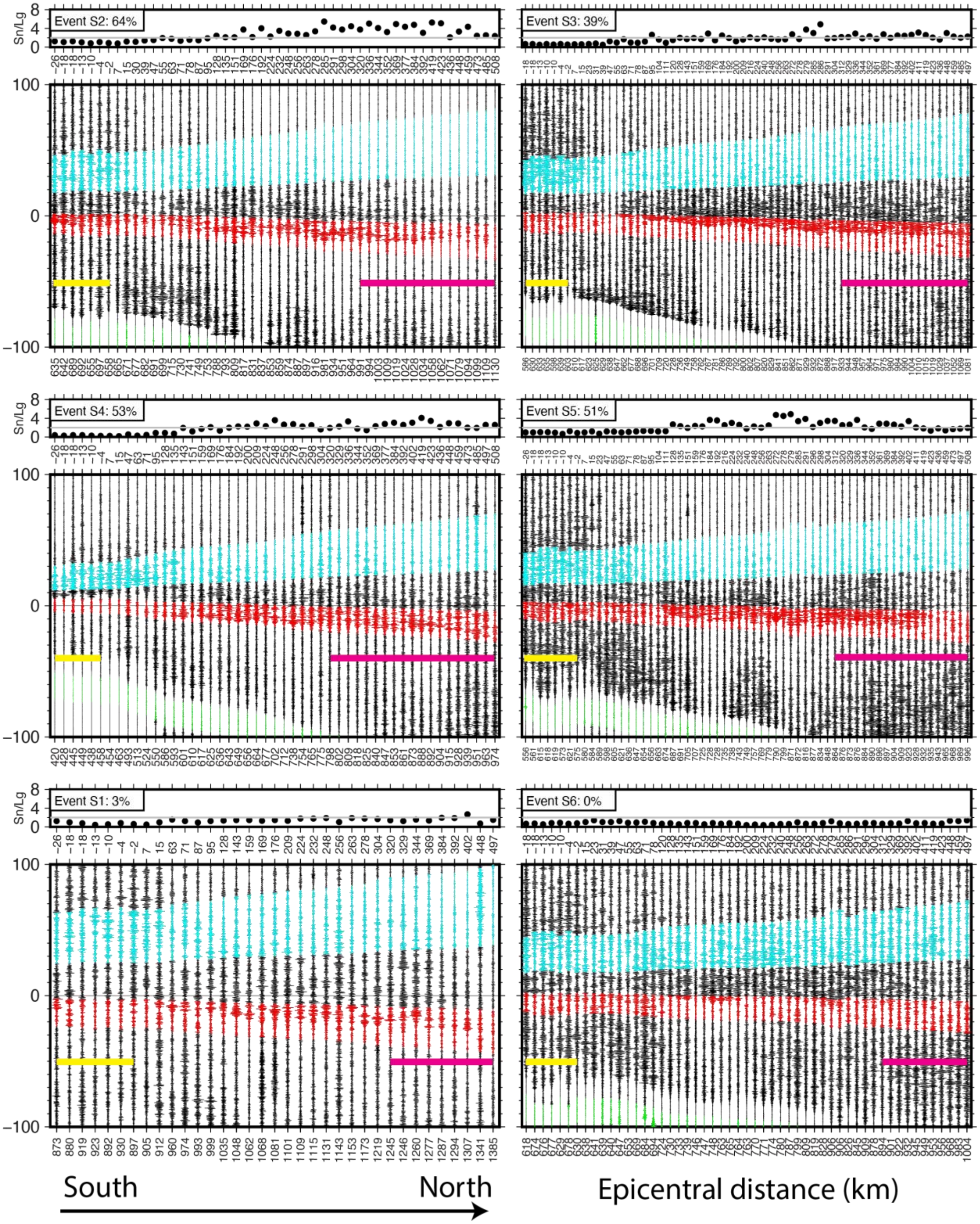

Event S2: 64%
Event S3: 39%
Event S4: 53%
Event S5: 51%
Event S1: 3%
Event S6: 0%
Sn/Lg
South
North
Epicentral distance (km)

Fig. 11

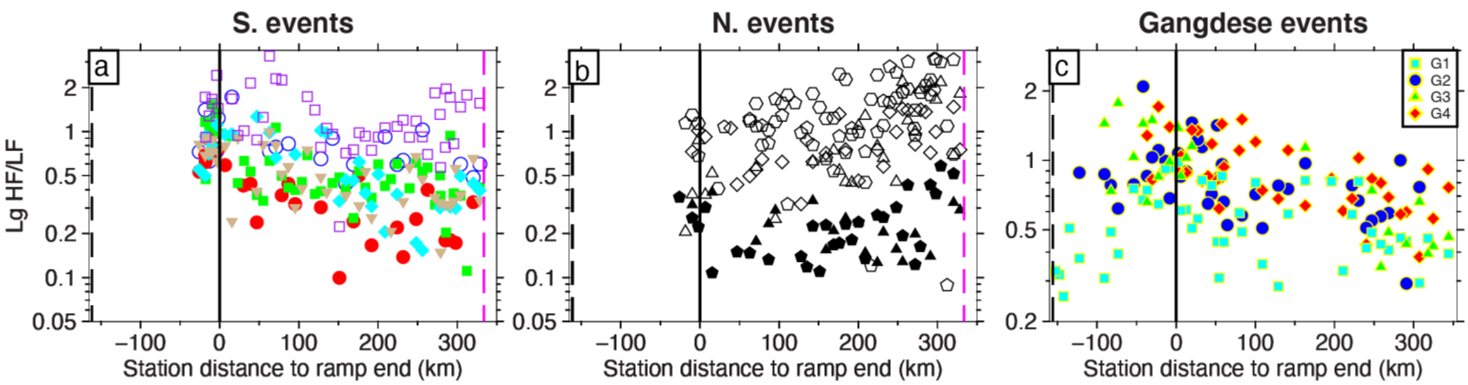

S. events
N. events
Gangdese events
a
b
c
G1
G2
G3
G4
Lg HF/LF
0.05
0.1
0.2
0.5
1
2
−100
0
100
200
300
Station distance to ramp end (km)

Fig. 12

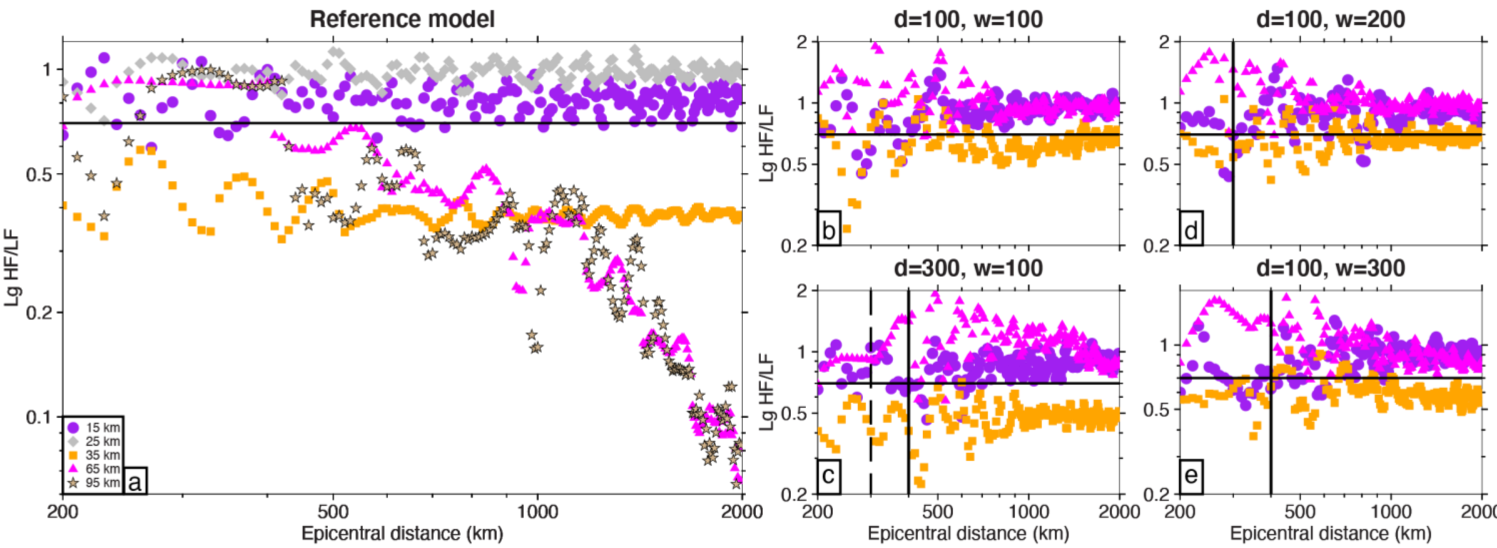

Reference model
d=100, w=100
d=100, w=200
d=300, w=100
d=100, w=300
Lg HF/LF
Epicentral distance (km)
15 km
25 km
35 km
65 km
95 km
a
b
c
d
e

Fig. 13

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

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

*Corresponding author. Email address: axelwang@stanford.edu

Key points

- Synthetics and data show the *Sn/Lg* method successfully identifies mantle earthquakes with thickening crust across the Himalaya.
- *Sn*-to-*Lg* conversions can be recognized by enhanced high frequency content of *Lg*
- *Lg* frequency content discriminates between crustal and mantle near-Moho earthquakes

Abstract

We study *Sn*-to-*Lg* conversion at regional distances due to significant crustal thickening, particularly in the context of using *Sn* and *Lg* amplitude ratios (*Sn/Lg*) to identify continental mantle earthquakes. We further enhance recent developments in computational seismology to perform 2.5D simulations up to 5 Hz and 2,000 km. Our simulations compare propagation in a reference, constant-thickness crust from a source at three depths straddling the Moho, to 48 models of the same three sources propagating through Moho ramps of four different widths (dips) at four different distances from the source. We compare our synthetics to data from 12 earthquakes recorded on the HiCLIMB array across Tibet, of which six events from northwestern Tibet traverse no major crustal-thickness variation, and six located south of the Himalaya cross a major Moho ramp. Our observations on real data show that amplitude perturbations on individual *Sn* and *Lg* waves are smooth and mostly limited to near the ramp end. Even the more-pronounced amplitude variations seen in our simulations show that *Sn/Lg* for mid-crustal earthquakes is consistently lower than those for mantle earthquakes. Hence we can directly compare *Sn/Lg* for ramp-crossing and non-ramp-crossing earthquakes and identify new mantle earthquakes in northern India. *Sn*-to-*Lg* converted waves may be readily detected near the Moho ramp end through an enhancement in high-frequency content. In addition, we observe higher frequency content in *Lg* from crustal than from mantle earthquakes, which offers a new discriminant for continental mantle earthquakes based on frequency content of *Lg* waves alone.

Plain language summary

Seismic waves *Sn* and *Lg* respectively propagate largely below and above the Moho. Previous work showing that *Sn* and *Lg* amplitudes can distinguish whether near-Moho continental earthquakes nucleated in the crust or mantle (the '*Sn*/*Lg* method') used only 1D (flat-Moho) theory and synthetics, and data from areas with little Moho topography. Here we extend this work with synthetic seismograms across large Moho ramps and with data recorded across the Himalaya from India to Tibet. By comparing earthquakes with source-receiver raypaths that do and do not cross a Moho ramp we show the *Sn*/*Lg* method can still identify mantle earthquakes provided multiple recorders are used. We also show that the frequency content of *Lg* contains information about *Sn*-to-*Lg* conversions, and can by itself be used to identify mantle earthquakes. Traditionally, *Sn* and *Lg* waves have not been modeled at high-frequencies (>1 Hz) and long-distances (>1000 km) due to high computing costs. Here, we take advantage of and enhance recent developments in computational seismology to model *Sn* and *Lg* propagation up to 5 Hz and for 2000 km through a 2D lithosphere, paying special attention to their amplitude ratio and its application to distinguish exotic continental mantle earthquakes from commonplace crustal earthquakes.

Keywords: *Sn*, *Lg*, crustal thickening, continental mantle earthquakes, Himalaya, Tibet

1. Introduction

Seismic waves *Sn* and *Lg* are the most prominent arrivals on high-frequency (~1–5 Hz) seismograms recorded at regional distances (~200–2,000 km). They are guided shear waves within the entire crust (*Lg*) or the entire lithosphere (*Sn*), and can be represented equivalently either by Airy phases from surface-wave normal modes (Stephens and Isacks,1977, Knopoff, 1973) or by interference patterns of waves multiply reflected between the surface and the Moho top-side (Oliver and Ewing, 1958) or under-side (Červený and Ravindra, 1971; Menke and Richards, 1980), respectively (Fig. 1). Their excitation and propagation characteristics derived from the above representations are directly related to the wave amplitudes that have been useful for a variety of purposes such as determining focal depths for crustal earthquakes from amplitude spectra (Baker et al., 2004), serving as the dominant measure for regional earthquake magnitude (e.g. Patton and Walter, 1993), monitoring nuclear tests based on *Pg* and *Lg* amplitude ratios (e.g. Zhang and Wen, 2013), as well as estimating local properties relating to the attenuation (e.g. Mousavi et al., 2014) and amplification (i.e. seismic hazards, e.g. Kebeasy and Husebye, 2003, Rodgers et al., 2019, 2020) of these waves.

Recently, addressing a half-century-long controversy regarding whether earthquakes can nucleate in the continental mantle (Chen and Molnar, 1983; Maggi et al., 2000; Chen and Yang, 2004; Schulte-Pelkum et al., 2019; Priestley et al., 2008; Craig et al., 2011; Prieto et al., 2017), we demonstrated the use of *Sn* and *Lg* amplitude ratios (hereafter "*Sn/Lg*") to discriminate continental mantle earthquakes from crustal ones using Tibetan earthquakes recorded on the Tibetan plateau (Wang and Klemperer, 2021) (Fig. 2). The signature of a mantle origin is a higher *Sn/Lg* compared with nearby crustal earthquakes recorded on a common array. For a group of earthquakes in NW Tibet, *Sn/Lg* ratios > 2 (averaged over many stations) were found to identify sub-Moho earthquakes. This method has the advantages of making the discrimination by relying on prominent waveform features of the earthquakes themselves (as opposed to Zhu and Helmberger, 1996 and Yang and Chen, 2010, who relied on more subtle waveform features), thus avoiding comparing independently derived earthquake and Moho depths at different locations, which has been a popular method (Chen and Yang, 2004, Priestley et al., 2008), and also can be performed using any stations/arrays that lie within regional distances of an earthquake (as opposed to Schulte-Pelkum et al., 2019 who relied on stations essentially on top of earthquakes).

Our *Sn/Lg* method is based on predictions from 1D surface-wave normal-mode theory, but given the *Sn* and *Lg* interconversions (*Lg* blockage or leakage) created when waves are incident on a dipping Moho, it is far from certain how the method will perform if there exists a large-scale structural variation between the earthquakes and recording stations (e.g. earthquakes in India recorded by stations in Tibet). Necessary corrections may be small – Song and Klemperer (2023) show general agreement between the catalog depths and *Sn/Lg* of hundreds of earthquakes with paths crossing the boundaries of Tibetan Plateau recorded on either of two permanent stations (KBL and LSA) – or may be significant, as where *Lg* blockage is used to study large-scale geologic features (e.g. North Sea: Mendi et al.,1997, Japan: Furumura et al., 2014, Pyrenees: Sens-Schönfelder et al., 2009).

Given the number of seismological applications utilizing regional-wave amplitudes, and that large-scale Moho topography is often well-known, many attempts have been made to quantify *Lg* blockage as *Lg* propagates through a suddenly thinned crust. Coupled-mode theory builds on the 1D surface-wave eigenproblem which can synthesize regional waves with only vertical (1D) heterogeneity. Coupled-mode theory represents the wavefield as a sum of basis functions (motion-stress vectors for a 1D problem, e.g. Aki and Richards, 2002) with laterally-varying amplitude coefficients obtained through the orthogonality principle of the normal modes (Maupin, 1988). Most relevant here are to consider the width across which Moho depth varies (Kennett, 1972; Drake, 1972; Kennett 1984; Maupin, 1988) and to incorporate undulating structural boundaries using local modes (i.e. motion-stress vectors corresponding to a flat (1D) model locally identical to a small section of the laterally varying 2D model; Odom, 1986) and representing the continuity conditions on the tilted surfaces as a volume force in both 2D (Maupin, 1988) and 3D (Tromp, 1994). A 2D coupled-local-mode method, incorporating all these ideas, was applied to *Lg* propagation in the North Sea (Maupin, 1989) to model transmitted and reflected wavefields for incident waves both perpendicular and at a sub-critical angle to the strike of the Moho topography. Maupin (1989) reported little difference between perpendicular and oblique incidences; the reflected wavefield is negligible and the strong *Lg* attenuation seen in the North Sea cannot be fully explained simply by structural effects, a conclusion that has been corroborated by later studies (Cao and Muirhead, 1993; Mendi et al., 1997) using 2D finite-difference simulations. An important observation is that mode-coupling occurs most strongly between neighboring modes. In Maupin (1989)'s North Sea model at a fixed frequency of 1 Hz, *Lg* mostly leaks into the mantle as *Sn* waves from the first (lowest) few *Sn*-forming normal modes (Fig. 3), as predicted by Kennett (1984). This means that only the lowest few *Sn*-forming normal modes, or the highest few *Lg*-forming normal modes, get enhanced by *Sn* and *Lg* interconversion, and if these enhanced modes do meaningfully contribute to either the *Sn* or *Lg* wavetrain then they contribute more to the low-frequency content of *Sn* or the high-frequency content of the *Lg* wavetrain. An alternative to the coupled-mode method is the ray-diagram method (Kennett, 1986 for *Lg*; Xie, 1996 for *Pn*), whose results are mostly graphical and do not account for interference between different rays once their initial coherent pattern is broken (Kennett, 1986). Nonetheless, for an initial bundle of rays with the same inclination (i.e. apparent velocity), focusing and de-focusing effects due to the lateral structure can be clearly seen (e.g. Kennett, 1986, his figures 2 and 3). These methods study the interactions of different modes (i.e. different dispersion relations: frequency vs. wavenumber) by either fixing the frequency (the coupled-mode methods) or the wavenumber (proxy to apparent velocity, the ray-diagram method). These methods yield valuable insights, but cannot represent the full broadband wavefield, which for regional waves is dominated by interference patterns.

Fully-numerical simulations can calculate the full broadband wavefield for any arbitrary structure, however, significant computational challenges exist given the frequency and range requirements for simulating regional waves. We are not aware of any 3D simulations that simultaneously reach frequencies up to 5 Hz and distance ranges up to 2,000 km, common observational parameters for *Sn* and *Lg* waves. Furumura et al. (2014) simulated regional wave propagation around Japan up to 1.5 Hz; and Rodgers et al. (2019, 2020) simulated ground motion in the San Francisco Bay Area covering an area of 120 km x 80 km up to 10 Hz. More importantly, these 3D simulations are run with very specific models, so are hard to generalize to other cases. On the other hand, 2D simulations, which recently focused on *Pn* propagation (Bakir and Nowack, 2012; Xie and Lay, 2017a&b; Wang et al., 2017), are attractive as they are much cheaper, so may simultaneously satisfy the frequency and range requirements, and may be more generalizable. However, these simulations, if performed in a Cartesian grid, require the earth-flattening transformation to produce physical sphericity which is vital for simulating interference head waves such as *Pn* and *Sn*. More importantly, these 2D simulations require a non-straightforward correction from their 2D line sources to 3D point sources (Li et al., 2014), and this correction cannot be exact if lateral heterogeneities exist (Li et al., 2014).

Here we establish first-order features of *Sn* and *Lg* transmission and inter-conversion with a set of 2.5D axisymmetric simulations allowing exact representations of Earth's sphericity and of 3D point sources. Our simulations have a maximum range of 2,000 km and frequency of 5 Hz, typical values used in observations. We view our synthetic results as building on those of Yang (2002) and Yang et al. (2007) who investigated *Lg* and *Sn* geometrical spreading for simple 1D models, and so we do not include effects such as intrinsic attenuation or random scatterers. The only factor that should make our synthetic results deviate from the 1D studies is the laterally varying crustal thickness, which is also typically well-known, thereby allowing our results to be quickly adapted to multiple regions of the world. We restrict our structural models to a Moho ramp leading to crustal thickening. Crustal thickening has been less explored, perhaps due to its subtler influence compared to crustal thinning, but this limited scope allows us to discuss comparisons with real data (Fig. 2).

After discussing the computational setup of our model, we explore individual *Sn* and *Lg* amplitudes, and their amplitude ratios (*Sn/Lg*) in a 1D reference model (Figs. 4&5) and in Moho-thickening models (Figs. 6–9), in which we establish the effectiveness of using *Sn/Lg* to identify continental mantle earthquakes in the presence of significant Moho thickening. We next examine real data from Tibet (Figs. 10&11) by directly comparing ramp-crossing (S events, Table 1& Fig. 2) and non-ramp crossing events (non-S events, Table 1 & Fig. 2), and show that *Sn/Lg* is a valid criterion for separating mantle from crustal earthquakes for the ramp crossing events just as for the non-ramp-crossing events. Hence the *Sn/Lg* method, if used rigorously with local shallow comparison events and multiple recording stations, can recognize the signature of a mantle earthquake even with stations in a region of crust much thicker (Tibet) than the source region (Indian Shield). Although we do not reliably detect effects of the ramp on individual *Sn* and *Lg* amplitudes, we are able to confirm enhancement of high-frequency *Lg* across the ramp (Fig. 12) due to neighboring mode-coupling during *Sn*-to-*Lg* conversion (Figs. 3, 13). Indeed, *Lg* frequency content is another powerful discriminant for continental mantle earthquakes.

2. Computational aspects

We use the AxiSEM3D software package, whose main advances compared to previous 2.5D axisymmetric methods (Bottero et al., 2016; van Driel et al., 2015) are that it can account for fully 3D variations in terms of volumetric perturbations (Leng et al., 2016) as well as through undulating surfaces (i.e. structural variations to either internal surfaces such as a Moho ramp or external surface such as the ellipticity of the earth or topography) that break the spherical geometry necessary for an axisymmetric method (Leng et al., 2019). We first briefly discuss these new features from a user's perspective and introduce two necessary modifications made to the source code in order to enable simulations with our desired scale and output.

2.1 Computational method and its enhancements

Without considering undulating surfaces, the azimuthal component of the 3D wavefield (from 0 to 2π in the plane perpendicular to the source-receiver direction) can be conveniently represented by a Fourier series, which localizes the equations to a single meridian plane not associated with any physical location, and then can be solved with a 2D spectral-element method (Leng et al., 2016). Recognizing that lateral heterogeneities in earth are much smaller than vertical ones, this hybrid scheme essentially uses "one line-shaped element" and high-order Fourier series in the azimuthal direction, and in the 2D meridian plane uses 4th-order Lagrange polynomials on a mesh with the quad-shaped elements that are necessary for a conventional spectral-element method. The cost of 3D simulations in AxiSEM3D depends not on the length of the 3rd dimension which in AxiSEM3D is always 0 to 2π, but on the 2D model size and highest wave frequency, since these determine the number of elements that each have an associated Fourier series. Even though the AxiSEM3D hybrid scheme is much more efficient for a global 3D model than a fully 3D scheme, we note that a small 3D model needs the same Fourier orders as the global model with the same level of lateral heterogeneity, and so a conventional 3D method might be more desirable in this case. Testing for this specific threshold is beyond the scope of this study. We estimated the Fourier orders needed if we were to extend our 2.5D simulation to 3D based on Equation (5) in Szenicer et al. (2020), and found 3D simulations are well out of reach given our available computational resources.

Even for our 2.5D simulations, the spatially and temporally down-sampled wavefield (discussed in detail in the next section) on one meridian plane on one wavefield component is about 1.5 terabytes, and AxiSEM3D by default directs all parallel processes (MPI ranks) to output the data to the same location. For our output size, the bandwidth to one location in a filesystem is overloaded, greatly reducing the performance (since MPI ranks spend most of their time waiting for I/O instead of computing) and more importantly, causing frequent filesystem crashes. We take advantage of local hard drives physically connected to each computing node on Stanford University's Sherlock HPC cluster (https://www.sherlock.stanford.edu/) through infiniband, and we modified AxiSEM3D so that each MPI rank can identify its own computing node at runtime and output its results to that node's physically-connected hard drive. This resolved the problem of crashing the filesystem, and increased the performance of AxiSEM3D by at least one order of magnitude for our problem size.

AxiSEM3D accounts for structural boundary variations diffeomorphic to a spherical or flat (for Cartesian mesh) boundary through the use of a "particle-relabeling transformation" (Leng et al. 2019; Al-Attar &Crawford, 2016), which finds the change in radial coordinate for each collocation point inside an element necessary to represent the structural boundary (and a finite thickness transition zone around it). The input mesh is a 2D mesh without any geometric variation (undulations on structural boundaries) or volumetric variation (perturbations of material properties such as density and elastic constants), so is a 1D vertically-layered model compatible with the axisymmetric requirement. The 2D and 3D variations are added on later with separate files (for geometric and volumetric variations) and then described as Fourier coefficients for each element in the 2D mesh. The 2D mesh can be related to a specific physical location and its properties only by association with a specific azimuthal angle ϕ. The current version of AxiSEM3D (Leng et al., 2019) does not output any of the built 2D or 3D models, and computed wavefields can only be plotted on the coordinates of the spherical mesh with no undulating surfaces. This causes a distortion of the wavefield visualizations that is too small to see on a global scale (Nissen-Meyer, pers. comm.), but unacceptable for our regional-scale lithospheric simulations. Further, although users can define models for undulating surfaces, there is currently no way to check if this is being represented accurately inside the program. We modified the source code to output the Fourier coefficients related to structural boundary variations for affected elements, and then deform the input mesh to obtain 2D variations at any azimuthal angle ϕ (constant in our 2.5D simulation) specified in our input geometric model. This enables us to visualize our regional wavefield without distortion and to confirm that our Moho undulation is exactly represented by the Fourier series (Supplementary material S1).

Lastly, we note that it is still not a straightforward and cheap task to perform regional wave simulations even with the proliferation of computing resources and advancements in efficient computational methods. Each of 51 simulations presented here required ~2 days with 500 cores on Stanford's Sherlock supercomputer. Since most regional wave applications focus on amplitudes, a much cheaper method based on radiative transport theory (a form of advanced ray theory), that can only calculate absolute amplitude but can easily account for 3D structures and random scatters (Sanborn et al., 2017) is potentially attractive. We did not use radiative transport because we prefer a fully-numeric method, and because radiative transport has been shown to underestimate shear energy both in 2D (Pryzbilla et al., 2006) and in 3D (Pryzbilla et al., 2008).

2.2 Model design

Our simulation domain is shown in Fig. 1. It has an effective size of ~2,000 x 230 km within the absorbing boundary conditions. The vertical properties are based on PREM (Dziewonski and Anderson, 1981) with the Moho depth adjusted to 30 km to better represent continental areas. Our 230-km depth includes 10 km of the positive velocity gradient below the mantle low velocity zone (LVZ) in order to include the LVZ trapped waves as well as *Sa*: the shear wave trapped between the free-surface and the bottom of LVZ (Schwab et al., 1974; Wang and Klemperer, 2023) that are important for the energy partitioning of surface waves.

Our reference model has Moho depth fixed at 30 km. For our other simulations, we introduce a Moho ramp, in the shape of a sigmoid function at distance d from the source and with width w that smoothly transitions the Moho depth from 30 km on the source side to 60 km beyond the ramp (Fig. 1). The distance-to-ramp parameter d controls the portion of the wavefield that will interact with the ramp and, with our fixed ramp height of 30 km, the ramp width parameter w controls the steepness of the ramp. We vary d from 100 to 700 km with a 200 km interval, and w from 100 to 400 km with a 100 km interval, values chosen to capture a wide range of realistic scenarios. Our steepest ramp, that thickens by 30 km over a distance of 100 km (>16° dip), is analogous to the steepest part of the Himalayan Moho ramp, typically 20-25 km vertical change over a 100-km width (e.g. Gao et al., 2016; Nabelek et al., 2009; Shi et al., 2016). The lowest slopes we model, ~4° dip (here, a 30-km ramp spanning 400 km), is more characteristic of eroded mountain belts in which total Moho relief of ~15 km within ~200 km across strike is typical (e.g. Cook et al., 2010). However, it is commonplace for earthquakes to be recorded along raypaths that are oblique, not perpendicular to orogens, and the first-order influence of oblique incidence can be approximated through an increase of ramp width (Bostock & Kennett, 1990). Our gentlest ramp (w = 400 km) is analogous to that seen by an earthquake recorded at 45° obliquity to the Himalayan ramp. We use as our source a thrust earthquake with moment magnitude $M_w = 6$, dip $\delta = 45°$, rake $\lambda = 90°$, and a Gaussian source time function with a half-width of 0.2s. For each combination of ramp parameters w and d, as well as for the reference model, we calculate the SH wavefield due to placement of this source at 3 different depths: 15 km (mid-crust), 35 km (shallow-lid, just below the Moho at the source but within the depth range of the ramp), and 65 km (deeper-lid, below the Moho everywhere in the model).

We constructed our 2D finite-element mesh to balance accuracy and efficiency. Our absorbing boundary conditions, including the thickness of the sponge layers, are set such that 97% of reflections for waves with >0.5 Hz frequency are eliminated (Haindl et al., 2020). We use two elements (10 collocation points) per wavelength in our simulations, and confirmed there was no visible numerical dispersion. Using just less than two elements per wavelength – as needed for the mesh coarsening – also produced no visible difference. We coarsened our mesh using two-refinement transition templates (Anderson et al., 2009) (by “tricking” the built-in mesher) at a depth of 70 km, 10 km below the deepest Moho, to ensure that mesh coarsening and deformation (due to the Moho ramp) do not conflict. Our coarsening strategy enforced two elements per wavelength at the coarsening depth, and more (or slightly fewer) elements per wavelength above (or below) this depth, which resulted in a ~16.8% reduction in the number of elements needed (Supplementary materials S2).

3. Numerical results from the reference model

As a check of our computational setup, and to ensure that *Sn* and *Lg* amplitude variations we discuss later are due only to the presence of the Moho ramp, we first calculate wavefields in a reference model with constant crustal thickness (no ramp) and compare our results to geometrical-spreading results calculated with a full-waveform method for a similar but non-identical 1D earth model for both *Lg* (Yang, 2002) and *Sn* (Yang et al., 2007) (Fig. 4). Both our source depths and distance range are larger than those explored by previous studies, so some discrepancy is expected apart from differences in earth models. We measure the *Sn* and *Lg* amplitudes at each offset as the RMS value over a time window defined by the expected range of group velocities, 4.0 to 4.7 km/s for *Sn* and 3.0 to 3.8 km/s for *Lg*. Our *Sn* windows are picked slightly differently compared to Wang and Klemperer (2021) to minimize overlaps with *Lg* windows at short distances and the mis-categorization of fast *Lg* waves at long distances, and also to account for non-zero source depths while not pre-judging whether an earthquake has a mantle or crustal hypocenter (Supplementary materials S3). All amplitudes are reported as displacements.

Fig. 4a&b show synthetic absolute *Lg* and *Sn* amplitudes, respectively, for the three source depths (15, 35 and 65 km) in the reference model (Moho depth fixed at 30 km), filtered from 1–5 Hz for *Lg* and around 3 Hz (from $3/\sqrt{2}$ to $3\sqrt{2}$ Hz, following Yang et al., 2007) for *Sn* and plotted at 10-km intervals from 200 to 2,000 km epicentral distance. Our frequency filters are 8th-order Butterworth filters. We also plot the relative amplitude decay provided by geometrical spreading models. For *Lg*

$$G_{Lg}(r) = r^{-\gamma}$$

where r is the epicentral distance and $\gamma = 1$ is an empirical constant (Yang, 2002). Yang (2002) modelled a variety of parameters such as source depth (but only tested crustal sources, above the Moho), frequency content, and amplitude-measurement technique, and found γ remained close to 1. *Sn* geometrical spreading is more complicated due to its propagation path (whispering gallery or interference head waves; cf. Avants et al., 2011), and has been modeled with both frequency (f) and distance dependence:

$$G_{Sn}(r,f) = \frac{10^{n_3(f)}}{r_0}\left(\frac{r_0}{r}\right)^{n_1(f)\log(r_0/r)+n_2(f)}$$

where $r_0 = 1$ km and $n_i(f)$ are fixed parameters for a specific two-layer Earth model, calculated by Yang et al. (2007) for a source at 15 km depth in a uniform 40-km thick crust. The frequency dependence of *Sn* geometrical spreading is the reason why we show single-frequency *Sn* in Fig. 4b.

We first note the preferential *Lg* excitation by crustal sources (amplitudes for the crustal source are an order of magnitude greater than for the mantle sources: Fig. 4a) and preferential *Sn* excitation by mantle sources (amplitudes for the 15-km source are an order of magnitude less than for the 65-km source: Fig. 4b). This is reflected by the amplitude ratios *Sn/Lg* for these three source depths, where although they vary as a function of epicentral distance, at each location, *Sn/Lg* for mantle earthquakes is always higher than *Sn/Lg* for crustal earthquakes (Fig. 5). In an ideal 1D model (i.e. the reference model), *Sn/Lg* varies according to epicentral distances in different ways for different source depths, but the common feature is the increase of *Sn/Lg* at long distances, i.e. about 600-1300 km, as noted before (Wang and Klemperer, 2021). The amount of increase is the highest for the 15-km event. For this same 15-km event, there is also a rather large decrease of *Sn/Lg* at shorter distances, whereas the values remain relatively constant for the 35- and 65-km events. The reasons behind this is due to an over-estimation of *Sn* following conventional observation methods, to which we return below while discussing individual *Sn* and *Lg* amplitudes.

The higher-amplitude models (15 and 35 km for *Lg*, 65 and 35 km for *Sn*) vary fairly smoothly with distance (Fig. 4 a&b). The abrupt discontinuities present in the bottom traces of Fig. 4a (65-km trace) and Fig. 4b (15-km trace) are artifacts due to mis-categorizations of the waves in our windowing process. For example, Fig. 4c,d&e show the synthetic seismograms for the 65-km source at 400, 410 and 460 km (triangles in Fig. 4a&b). For this deep source, we expect essentially no *Lg* excitation. However, at 400 km, there are two prominent *Sn* peaks included in the *Lg* window (Fig. 4c), which explains why measured *Lg* amplitudes are unexpectedly high for distances from 200-400 km (Fig. 4a). Measured *Sn* amplitudes are correspondingly lower than the total amplitude within the *Sn* phase but by a smaller proportional amount because the later *Sn* peaks are lower amplitude than the first-arriving *Sn* peak. At 410 km (Fig. 4d), the second *Sn* peak is no longer within the *Lg* window, resulting in a sharp drop of measured *Lg* amplitudes. At 460 km (Fig. 4e) the second *Sn* peak moves into the *Sn* window, resulting in a small proportional increase in measured *Sn* amplitudes for the 65-km source at that distance (Fig. 4b). This artifact (peaks moving in and out of a window) is present for all cases in Fig. 4, but is small for the major phase from each source depth, e.g. the small sinusoidal oscillations in measured *Lg* amplitudes for the 15- and 35-km sources. These small variations were also shown but not explained in previous synthetic studies (Yang, 2002). Although it is important that we fully understand our synthetics, these phenomena have no relevance for real data for which small-scale scatterers will always act to smooth out the strong amplitude peaks seen in our synthetics (and those of Yang, 2002 and Yang et al., 2007). Our observations offer insight into the relationship between physical *Sn* and *Lg* (as defined by propagation waveguides) and observational *Sn* and *Lg* (as defined by group-velocity windows), and into the fundamental inaccuracy of using the same *Sn* and *Lg* velocity windows for different events because these window-bounding velocities are dependent on epicentral distance even for the same source depth in a 1D model (a point also touched on by Aki and Richards, 2002, their Box 7.1). For example, all three wavelets in Fig. 4c are physical *Sn* waves, ordered by their relative mantle and crustal path lengths, and the *Sn* velocity window would need to be extended down to 3.4 km/s to capture all three in the *Sn* window. At just 400-km range, the second and third wavelet appear in the *Lg* window (and also a small wavelet around 127s, Fig. 4c), but with increasing offset all these peaks would have travelled proportionally greater distances in the mantle, gaining higher apparent horizontal group velocity, and be captured in the conventional *Sn* velocity window (Fig. 4d&e).

Working with synthetics it would be possible to measure the apparent group velocity of each arrival and thereby correctly separate *Sn* from *Lg*; but in real data such an approach is likely difficult or impossible. Hence we do not seek to change the conventional observation method, but rather we acknowledge the prevalence of this issue and highlight the irrelevance of fine-tuning velocity windows and the care needed to avoid over-interpreting amplitude measurements. Fortunately, the conventional and tractable *Sn*- and *Lg*-velocity windows method are historically proven to be adequate, especially when only crustal sources are considered. Since our *Sn/Lg* method for identifying mantle earthquakes fundamentally relies on comparisons between potentially crustal and mantle earthquakes, rather than absolute-value *Sn/Lg* thresholds, it is more important to use a simple and unified approach for a group of events (to enable comparisons) than to strive for picking the most accurate windows for individual events, which in practice is also hard to achieve.

Sn synthetic amplitudes (Fig. 4b) exhibit the classic interference head-wave behavior at distances <~1,300 km, in that amplitudes first decrease then increase with distance due to the spherical focusing effect, because at larger distances more energy from waves multiply-reflected at the Moho underside will contribute to the amplitudes, in addition to the direct arrival. The distance at which the *Sn* amplitudes begin to increase and the amount of the increase depends on source depth, and is closest/strongest for the shallowest source (note this is not captured by the geometrical spreading model, black lines in Fig. 4b, as that is an empirical fit based on a crustal source only). For the 15-km source, we expect mostly *Lg* excitation. While our *Sn* window is already shortened compared to Wang and Klemperer (2021), the earliest *Lg* waves could appear within the *Sn* window, which results in the artificial amplitude jump at 820- and 830-km distances (circles in Fig. 4b; Fig. 4f&g). As before, these sudden amplitude changes are due to mis-categorizations, but are likely much smaller in real data due to presence of smoothing effects. However, if these smoothing effects are not accounted for and if the conventional windowing method is followed (Yang et al., 2007) (completely justified if the intent is to study geometrical spreading alone), the amount of *Sn* amplitude increase might be over-estimated at these long offsets due to incorporation of *Lg* waves. This incorporation of *Lg* waves at long offsets potentially explains the earlier rise to larger *Sn* amplitudes, leading to a larger increase of *Sn/Lg*, for the crustal source compared with the mantle sources (Fig. 4b, Fig. 5). For a 35-km source, the smooth amplitude increases from e.g. 900- to 1,000-km distances (squares in Fig. 4b) are due to the spherical focusing effect of interference head waves, as the number of peaks within the *Sn* window is not changed, yet their amplitudes (most notably the third peak) grow larger (Fig. 4h, i). For distances >1,300 km, all three sources have about the same amount of *Sn* energy (Fig. 4b), but their *Lg* energy is vastly different (Fig. 4a) so our method is still very effective at these long offsets. Beyond ~1,300 km our measured amplitudes start to drop, a phenomenon not previously noted because the Yang et al. (2007) study was limited to shorter offsets, but completely reasonable because the spherical focusing effect must eventually wear off, i.e. the multiply-reflected waves at the Moho underside eventually become too small to meaningfully contribute. In all other respects, our 3-Hz *Sn* amplitudes exhibit the same interference head-wave behavior as in Yang et al. (2007). In fact, the fit to our 15-km source (the same source depth simulated by Yang et al. (2007), but in a slightly different earth model) is good (Fig. 4b). The misfit to the deeper sources beyond ~500 km clearly originates from the fact that only a crustal source was considered by Yang et al. (2007), which, combined with other reasons discussed above, led them to an over-estimation of the spherical focusing effect.

We have an almost exact match between our 15- and 35-km sources' synthetic *Lg* amplitudes and the simple *Lg* geometrical spreading model (Yang, 2002), and between our 15-km source's *Sn* synthetic amplitudes and the more complicated *Sn* geometrical spreading model (Yang et al., 2007). The mismatches between our synthetic amplitudes and the previous models can all be understood. This gives us confidence in our modelling approach. Our results cover a larger parameter space and exhibit a greater range of features than previous studies, so already provide useful new information as well as serving as a benchmark against which to test our simulations with a Moho ramp.

4. Numerical results from Moho ramp models

Our parameter-space study includes 48 2.5D crustal-thickening simulations (plus reference simulations). Here, we present a selection of these results (Figs. 6–9) as two groups by first fixing the ramp width $w = 200\ km$ and varying the distance to the start of the ramp d =100, 300, 500 and 700 km; and then by fixing $d = 300\ km$ but varying w = 100 km ~17°, 200 km ~9°, 300 km ~6° and 400 km ~4°. The rest of our numerical results can be found in Supplementary materials S4&5. The amplitudes are measured as discussed in Section 3, except that now our *Sn* amplitudes are measured using the same broader frequency band we use for *Lg*, i.e. 1-5 Hz. We present our results for *Sn* and *Lg* amplitudes (Figs. 6, 9, left and middle columns) as ratios to our reference-model results, aligned by distance relative to the end of the ramps to highlight deviations relative to ramp locations. We show results for *Sn/Lg* (Figs. 6, 9, right columns) relative to epicentral distance and overlaid on reference-model results to highlight the ramp effects and to illustrate the absolute values of *Sn/Lg* for different source depths.

Our interest is in phenomena that have the potential to be recognized and measured in real data. Measurements on synthetics of *Sn* alone, or *Lg* alone, coupled with inspection of synthetic seismograms and compared to the reference (flat Moho) model reveal the physics of wave propagation across Moho ramps. However, in the real world no reference data are available and it is the *Sn/Lg* ratios that, by removing source and receiver dependencies, may allow recognition of source depth with respect to Moho (Wang & Klemperer, 2021).

4.1 Fixed ramp width, $w = 200\ km$

This fixed ramp width, combined with our constant 30-km ramp height, produces a Moho ramp with fixed gradient that is among the steepest in nature (though not the steepest Moho ramp we test, see Supplementary materials S4&5), especially considering that many earthquake-receiver geometries involve oblique incidence onto the ramp, effectively increasing the width of the ramp. We highlight this example, w = 200 km, because incidence onto a steep Moho ramp produces the clearest effect on *Sn* and *Lg* amplitudes. For all our simulations (variable w, d and source depth z) the measured amplitudes (and the *Sn/Lg* ratios) coincide with those from the reference model until the waves reach the ramp apart from tiny numerical errors (e.g. Figs. 6a,d show some symbols at distances -700–0 km slightly below the grey line that represents equality with reference model results) .

4.1.1 Sn amplitudes

For relative *Sn* amplitudes (Fig. 6, left column, a.-c.), one of the most striking features is the focusing of *Sn* waves that starts close to the middle of the ramp and peaks slightly beyond the ramp, for most source depths z and distances to ramp d. This phenomenon corresponds to the breaking of the *Sn* waveguide by the Moho ramp (Fig. 7). Before the leading wavefront in the mantle reaches the ramp (at time = 73 s, Fig. 7a), the wavefield is the same as in the reference model, with the same first reflected wave as the leading wavefront and the same first arrival at the surface (Fig. 7a&b). At time = 100s the leading wavefront is at about the middle of the ramp and the first arrival is just beyond the start of the ramp (Fig. 7c). At 100s, the Moho underside transmitted wave has a much larger amplitude in the ramp model than in the reference model (Fig. 7d), and this increased amplitude extends to the surface, representing the start of the *Sn* peak just beyond the start of the ramp. The reasons behind the focusing are twofold. First, the Moho ramp increases the local curvature of the Moho so that deeper energy on the leading wavefront, which in the reference model would refract up at a greater distance (Fig. 7e, blue arrows), refracts up to the surface from the ramp (Fig. 7e, yellow dashed arrows), locally increasing the amount of energy being transmitted into the crust. Second, the incidence angle of this deeper energy changes from almost grazing to a smaller angle ($i_{ramp} < i_{flat}$) (Fig. 7e), which could flip the energy partitioning of the reflected and transmitted waves (Fig. 7f) in favor of transmission (calculated using plane-wave transmission and reflection coefficients, e.g. von Seggern, 2012, which are a good approximation at long distances from the source). The large reflected energy at large incidence angles (Fig. 7f) enables multiple reflections at the Moho underside, and essentially gives rise to the whispering-gallery waveguide. The increase in transmitted energy at decreased incidence angles shows how this waveguide is broken by a Moho ramp. On a seismogram (Fig. 7g), other than the prominently increased amplitude of the first arrival of the ramp model, the effect of the ramp shows up as delays in individual arrivals due to the transmitted waves travelling a longer distance in the crust and travelling at a steeper angle ($r_{ramp} < r_{flat}$, Fig. 7e) leading to a smaller horizontal apparent velocity. In general, the magnitude of this focusing (~2 to 10 times stronger than reference, Fig. 6, left column) is a proxy for how much of the original wavefield interacts with the ramp, which is inversely proportional to d and is largest for the 35-km event (that lies within the vertical extent of the ramp), followed by the 15- and 65-km events. This explains why the 65-km deeper-lid event has about half the focusing strength of the shallow-lid earthquake (35-km) and the mid-crustal earthquake (15-km), which both have similar degrees of focusing. For the 35-km event, the focusing strength is strictly inversely proportional to d, but this is not the case for the 15- and 65-km events for which, for shorter distances to the ramp, more complicated interferences occur that decrease the strength of the *Sn* peak ($d = 100$ for 15-km event, Fig. 6a, and $d = 100, 300$ km for 65-km event, Fig. 6c).

The only case without an *Sn* peak at ramp exit is z=15 km, d=100 km (Fig. 6a). Regardless of ramp width, no crustal sources at this short distance to ramp show an Sn peak, but instead have a large decrease of Sn for longer ranges beyond the ramp exit, and all eventually recover back to close to reference values (Fig. S5-1a). Similarly, *Sn* de-focusing is present for the 65-km events when *d* is short (i.e. 100 and 300 km) (Fig. S5-1c). Unlike the *Sn* focusing peak that is just outside (<200 km) of the ramp exit (left column of Figs. 6, S4-1), or even completely contained inside the ramp region for wider ramps (left column of Figs. S4-2, S4-3), these *Sn* de-focusing regions can extend up to 600 km (Fig. 6a) to 800 km (Fig. 6c), having a broad influence on *Sn/Lg* (e.g. Fig. 6i, d=100 &300 km). It would seem that e.g. if an earthquake with z = 65 km, d = 100 km is measured at ~700-800 km epicentral distance (Fig. 6i), its *Sn/Lg* could be confused with that of an earthquake with z = 15 km and d = 500 km measured at the same distance (Fig. 6g). However, such a confusion requires a careful orchestration of a broad de-focusing zone and a localized *Sn* focusing peak, as well as potentially very different back-azimuths to a particular station to produce the 400-km difference in effective ramp width, and although this could occur in the real data, array-based measurements with varying source-station geometry should be able to mitigate, if not completely avoid, this effect.

The behavior of *Sn* is quite complex in the presence of a crustal-thickening Moho ramp due to its interference head-wave nature. For example, there are also secondary focusing peaks for the 35-km events for all d's (Fig. 6b). However, such specific observations on synthetics are likely too detailed to observe in real data, so we do not further discuss or decipher these phenomena.

Lastly, in contrast to the above-mentioned deviations from the reference model, the other most striking feature, common to all our simulations (Figs. 6, 9, left columns, Supplementary materials, S4&5) is that, at long offsets, *Sn* amplitude always returns to about the same level as the reference model. The mechanism for unperturbed amplitude at long offset is that deeper energy (below the black dot on the wavefront, Fig. 7e) may never interact with the ramp, and hence at long distances this energy is transmitted into the crust as if the Moho had always been flat beneath a thickened 65-km crust. However, arrival delays (Fig. 7g) persist for the ramp model even at long distances beyond the ramp, simply due to the increased travel path in the crust.

4.1.2 Lg amplitudes

Clear *Sn*-to-*Lg* conversion due to Moho thickening can be seen in Fig. 6, middle column, especially for the mantle sources. To first order, when the ramp shape is fixed as in the present case (i.e. height = 30 km and w = 200 km), we expect the degree of conversion to be controlled by the amount of *Sn* excited (positively correlates with source depth z) and the subset of this amount that interacts with the ramp (negatively correlates with d and source depth relative to Moho). Our results indicate that the source depth plays a far more important role. For the 65-km event, all d's share a similar growth pattern for relative Lg amplitude, which increases by >5 times between the ramp start and end, due to *Sn*-to-Lg conversion. At longer offsets, beyond the ramp, the continued gradual increase in relative Lg is due to the decrease in Lg for the reference model for this sub-crustal source (Fig. 4a, magenta triangles) rather than to Lg growth in the ramp models. For the 35-km event, even though the source depth is within the vertical extent of the ramp, *Sn*-to-Lg conversion is relatively modest (Fig. 6e). Except for the source closest to the ramp ($d = 100km$), the relative Lg amplitude increase from the reference model is less than a factor of 2, which means real-world observations perturbed by scatterers and noise could be difficult. Beyond the ramp, after the initial oscillations in relative Lg amplitudes, we see relative amplitudes decrease. This is especially prominent for d = 300 and 500 km, and is subtle for $d = 100$ km perhaps due to more initial wavefield interaction with the ramp, and is not shown for $d = 700km$ because our 2000-km maximum simulation range does not include distances sufficiently far beyond the ramp (Fig. 6e). As ramp width increases, drop-off of amplified Lg becomes even more pronounced (e.g. Figs. S4-2e, S4-3e). The crustal 15-km source (Fig. 6d) does not show this same behavior of relative-amplitude decrease, even though in the reference model Lg decays at the same rate for both 15-km and 35-km sources. The relative-amplitude drops for the 35-km below-Moho source may therefore indicate that the crustal waveguide cannot sustain the increased Lg frequencies that are created by *Sn*-to-Lg conversion, a topic we return to below (see Fig. 8).

Relative Lg amplitude from the 15-km source decreases by factor <2 as we cross the ramp (Fig. 6d). The magnitude of decrease is inversely proportional to ramp distance d because disruption of the Lg waveguide (i.e. the crust) causes Lg de-focusing as illustrated by Kennett (1986). Though this de-focusing must also occur for the 35-km and 65-km sources it is more than compensated for by the strong *Sn*-to-Lg conversion from these deeper sources. For the 15-km mid-crustal source, *Sn*-to-Lg conversion is hard to observe as there is much less initial *Sn* energy that can be potentially converted to Lg (Fig. 4b). However, *Sn*-to-Lg conversion must still be occurring because at larger distances beyond from the ramp relative Lg amplitude for all d's increases above 1, indicating extra Lg energy than expected if the Moho was uniform.

When *Sn* converts to *Lg* we expect not only amplitude but also frequency effects: the higher-frequency portion of the *Lg* becomes enriched because mode coupling tends to happen at neighboring modes (Fig. 3) (Maupin et al., 1989), thus higher-mode *Sn* tends to excite higher *Lg*-forming modes, which contribute to *Lg* at higher frequencies. We test this with our full-waveform results by comparing the *Lg* wavetrain filtered from 1–5 Hz, as shown thus far, to the same wavetrain filtered 0.1–0.8 Hz. We plot the high-frequency (HF) to low-frequency (LF) ratio (*Lg* HF/LF) (Fig. 8) of the ramp model divided by the reference model, and confirm that *Lg* has a higher frequency component that develops across the ramp due to *Sn*-to-*Lg* conversion. For the 15-km source, although subtle, there is a slight increase of high-frequency content further away from the ramp, which suggests some *Sn*-to-*Lg* conversion for the mid-crustal source. Comparing Fig. 8b&c to Fig. 6e&f we see similar trends, implying that a large part of the increase in *Lg* beyond the ramp is due to the increased HF component from *Sn*-to-*Lg* conversion.

4.1.3 Sn/Lg amplitude ratios

The changes in *Sn* or *Lg* amplitude or frequency content, relative to the reference model, are by a factor typically <2, so can be hard to recognize on real, noisy, data (except for the deepest source at the largest offset, Figs. 6f, 9c). In contrast, the amplitude ratio *Sn/Lg* is a direct measure of the relative strengths of *Sn* and *Lg* amplitude perturbations. *Sn/Lg* in the reference model increases linearly in log amplitude–log distance space for distances from ~600–1,400 km (Figs. 5, 6&9 g-i), confirming our earlier conclusion (Wang and Klemperer, 2021) from analysis of the empirical geometrical spreading models that differ between *Sn* and *Lg*. When the source is at 15 km, *Sn/Lg* largely follows the shape of *Sn* variations (Fig. 6g), because the *Sn*-to-*Lg* conversion is rather weak. Since any ramp only locally perturbs *Sn* amplitude perturbations, if amplitude ratios are measured far enough beyond the ramp, there is virtually no difference between the ramp and the reference models. Hence, crossing a significant Moho ramp (as in the present example) does not affect *Sn/Lg* observations for a crustal earthquake provided the measurements are made sufficiently far beyond the ramp. Exactly how far is sufficient is related to d, and ranges from ~800 km beyond the ramp for $d = 100\ km$ (measured from the red vertical line in Fig.6g) to

~100 km beyond the ramp for $d = 700\ km$ (measured from the cyan vertical line in Fig.6g). More simply, for source depth = 15 km for all d studied here (100 km ≤ d ≤ 700 km), a propagation distance of ~1100 km is sufficient to erase most of the ramp effect on *Sn/Lg*: at this distance all symbols coincide with the reference model (grey inverted triangles, Fig. 6g). Lastly, we note that *Sn/Lg* almost nowhere exceeds 0.2 for the 15-km event, marked by a black fiducial line in Fig. 6g,h,i.

In contrast, for the 35-km and 65-km sources, *Sn/Lg* is typically an order-of-magnitude larger than for the crustal source and rarely drops below 0.2 regardless of their variations, except sometimes just beyond the ramp. For the 35-km source (Fig. 6h) we see a combined effect from *Sn* and *Lg* variations, with *Sn* controlled by the local, transient focusing behavior, primarily in the ramp region, and *Sn/Lg* determined largely by *Lg* amplitudes further beyond the ramp. The biggest decrease of *Sn/Lg*, to ~3 times lower than the reference model at epicentral distance ~450 km is for the source that is closest to the ramp start and is due to increased *Lg* amplitudes (Fig. 6h, red circles). For the 65-km source, *Sn/Lg* is primarily controlled by the amplitude variations of *Lg*, and in some cases with small *d*, *Sn* de-focusing. Further away from the ramp there are significant decreases from the reference model by more than an order of magnitude, with the source closest to the ramp again exhibiting the largest decrease. Even though *Sn/Lg* for the two sub-Moho sources (Figs. 6h,i) far beyond the ramp can be smaller than had there been no Moho ramp, *Sn/Lg* remains 5-10 times larger than for the mid-crustal, 15-km, event. Visually, crustal and mantle earthquakes can be largely separated by the black fiducial line (Fig. 6g,h,i). In consequence, the *Sn/Lg* method (Wang & Klemperer, 2021) is robust for all cases tested here, especially if the recording stations are not limited to the ramp region and some measurements are made far beyond the end of the ramp.

4.2 Fixed distance to ramp, $d = 300\ km$

Here, d is fixed, so all ramps start at the same location but they extend out by different distances from $w = 100$, 200, 300 and 400 km (Fig. 9). Since our ramp height is fixed, by increasing w, we are decreasing the steepness of the ramp. As the dip of the Moho ramp decreases, the amplitudes will converge to the reference model.

Many key observations remain the same as in the previous section, including the *Sn* focusing peaks being localized close to the ramp and then recovering to the reference model at long distances (Fig. 9, left column), and the increase of the *Lg* amplitudes (quite subtle for the 15-km mid-crustal event) beyond the ramp that persists to greater distances in the crustal waveguide (Fig. 9, middle column). Most important, even though the *Sn/Lg* ratio for each source depth varies by a factor of ~20 with offset (Fig. 9, right column), *Sn/Lg* for the 35- and 65-km events exceeds *Sn/Lg* for the 15-km event at all distances, typically by a factor of 5–10. Thus – as seen also from Section 4.1 and Fig. 6, right column – *Sn/Lg* ratios are a robust metric for interpreting source depth above or below the Moho even in the presence of a Moho ramp, but particularly beyond the end of the ramp (Fig. 9, right column).

Increases in steepness of the Moho ramp can be thought of as an effective increase in the local Moho curvature. The tighter the curvature, the more intense the upward focusing of *Sn* from the Moho underside (Fig.7, e&f), leading to the most prominent feature in our simulations, the factor of 2–10 increase in *Sn* compared to the reference model vertically above and immediately beyond the ramp (Fig. 9a-c). Fig. 9a-c shows the steepest Moho ramp (w=100km) leads to the strongest and earliest (with respect to epicentral distance) *Sn* amplitude increase, reaching peak *Sn* amplitude just beyond the ramp region (for ramp widths <200 km). For wider ramps ($w = 300$ and 400 km), the peaks are completely within the ramp region. Just as for the models in Fig. 6, these results imply the need for observations across the ramp region if they are to be relevant for real data. The *Lg* energy increase beyond the ramps is again well-aligned with the start of the ramps (Fig.9, d-f) and is most prominent for the 65-km source, and decreases as the source depths decreases. As expected, the gentlest ramp (w=400 km) has the smallest increase in *Lg*. Ramp width seems to have less influence on Sn de-focusing. For the 15- and 65-km events, *Sn* de-focusing occurs except for w=100 km when z=15 km (Fig. 9a,c), and for d=100 km, even this exception doesn't hold anymore (Fig. S5-1a,c). However, at d=300 km, Sn de-focusing barely influences *Sn/Lg* (Fig. 9g&i) and the de-focusing becomes completely absent for larger d's (Figs. S5-2, S5-3, left column).

5. Observational results

We study earthquakes recorded by the HiCLIMB array not only for its high data quality, but also because the Moho structure beneath this array is well-studied (Nabelek et al., 2009), providing good definition of the Moho ramp structure (Fig. 2). Along the HiCLIMB profile (IRIS data code XF), the Moho is relatively flat beneath northern India and the Main Frontal Thrust (MFT). The Moho ramp begins about 100 km further north, beneath the Main Central Thrust (MCT). Because the surface trace of the MCT is tortuous due to laterally varying exhumation of a low-angle structure (Martin, 2017), we use a line 100-km north of the MFT as our proxy for the start of the ramp (Fig. 2a). North of the MCT, the Moho deepens from ~45 km to ~65 km over a distance of 150 km to the Yarlung-Zangpo Suture (YZS), a geometry present all along the Himalayan arc (e.g. Gao et al., 2016; Shi et al., 2016). Our data set includes very few southern stations over the northern part of the ramp (usually ≤7 due to limited operating time and noisier data), with most available stations lying further north beyond the YZS. This station distribution offers only a glimpse close to the ramp region to investigate effects on individual *Sn* and *Lg* amplitudes but gives ample opportunity to observe *Sn/Lg* away from the end of the ramp, where synthetics predict its effectiveness. In addition to this main dataset, we also analyzed 4 events from the Gangdese-92 array (Fig. 2). These 4 earthquakes are directly due south of the stations, offering an opportunity to evaluate the influence of oblique incidence for the HiCLIMB events. We use the same *Sn* and *Lg* velocity windows as used in Wang & Klemperer (2021), based on regional observations in our study area, i.e. 4.3-4.8 km/s for *Sn* (McNamara et al., 1995) and 3.1-3.6 km/s for *Lg* (McNamara et al., 1996). The data is bandpass filtered from 1-5 Hz with an 8th order Butterworth filter. We select traces only if either or both *Sn* or *Lg* has a root mean square (RMS) amplitude at least twice as high as that of a noise window, defined to start 30s and end 5s before the *Pn* arrival that we calculate using a constant velocity of 8.1 km/s.

There have been reports of sub-crustal earthquakes beneath southern Nepal and northern India (e.g. Chen & Molnar, 1983; Chen & Yang, 2004; Baur, 2007; Song & Klemperer, 2024) but none are confirmed and counter-claims exist that all earthquakes in these regions are likely intra-crustal (Maggi et al., 2000; Mitra et al., 2005; Priestley et al., 2008). Here, we look at six earthquakes, our ‘southern events’ (S1–S6 from south to north) with catalog depths 10–62 km (Table 1), recorded on the HiCLIMB array (Nabelek et al., 2009; Fig. 2). The travel paths of S1–

S6 traverse a Moho ramp ~20 km high at distances $50 \le d \le 660$ km from the source, and spanning widths measured obliquely along the path $160 \le w \le 475$ km. This ramp is smaller than used for some of our synthetics, and the obliquity of raypaths to both the ramp and the HiCLIMB profile means that different recording stations have a different d and w. This varied geometry is beneficial to our method as it helps to avoid systematic errors. HiCLIMB also allows us to compare our six southern events that do traverse the ramp with six earthquakes in northwestern Tibet (Wang & Klemperer, 2021) (Fig. 2) that do not traverse a ramp or indeed any major Moho topography. These six 'northern events', spanning upper-crustal to upper-mantle hypocentral depths (Wang and Klemperer, 2021), were recorded on the same array as the southern events, over roughly the same distance ranges (Supplementary materials S6). The southern and northern events are also similar in magnitudes, ranging from m_b 3.5-4.3, with S4 being the smallest event studied (Table 1). The HiCLIMB array operated only from mid-2004 to end-2005 and seismicity in Northern India is not nearly as prolific as on northwestern Tibet, so our six southern events have a much larger spatial spread than the six northern events.

We look at our data from three perspectives to illustrate the effect of regional waves traversing through a Moho ramp: gross amplitude measurements (Fig. 10), individual seismogram changes across a record section for a given event (Fig. 11), and *Lg* HF/LF, i.e. ratios of *Lg* amplitudes in 1-5Hz (HF) and 0.1-0.8Hz (LF) frequency ranges (Fig. 12).

5.1 *Sn, Lg* amplitudes and *Sn/Lg* ratios

We plot *Sn*, *Lg* and *Sn/Lg* for all our events against station distance north of the end of the ramp (Fig. 10), YZS (Fig. 2), so that the horizontal axis is also a proxy for station locations allowing evaluation of site effects along the array. We normalize the individual *Lg* and *Sn* amplitudes to the first recording station, i.e. the southernmost and northernmost stations for southern and northern events respectively, to remove first-order differences between earthquakes, e.g. their different magnitudes. As we move north from the end of the ramp, epicentral distances increase for the southern events, but they decrease for the northern events (Fig. 10). Alternatively, plotting both groups of events against epicentral distance shows that amplitude generally decreases as epicentral distance increases (Supplementary Fig. S6-1).

Individual *Sn* and *Lg* amplitudes for both southern (Fig. 10 a&b) and northern (Fig.10 c&d) events show remarkable coherence as a function of station location, despite the many differences within and between the two groups. At about 200-300 km beyond the end of the ramp, all measurements (Fig. 10 a-d) are amplified. For the southern events (Fig. 10 a&b), this increase of *Sn* and *Lg* amplitudes superficially resembles the *Sn* peak and increased *Lg* due to *Sn*-to-*Lg* conversion predicted by our modelling (Fig. 6&9, left and middle columns). However, our synthetics show both increases should occur closer to the end of the ramp, reaching their maxima within 100–200 km beyond the end of the ramp. The distance of these maxima from the end of the ramp decreases as ramp width increases, and the obliquity of our source-receiver azimuths to the ramp creates very large effective ramp widths (Table 1). Hence the location of the *Sn* and *Lg* maxima moves even closer to the ramp (Fig. 9, left & middle columns) so the amplitude increases at 200–300 km (Fig. 10a&b) are most unlikely related to traversing the ramp. Indeed, the northern events also show *Sn* and *Lg* amplification at the same stations, implying the peaks at 200–300 km are likely due to local variation in crustal and mantle seismic attenuation (Fig. 2).

Another potential candidate for an *Sn* focusing peak is shown by the few stations that recorded the southern events within and closely adjacent to the ramp (~-50 – +100 km) (Fig. 10a). This is promising because the northern events (Fig. 10c) do not seem to show this peak, and for our southern events that traverse wider (less-steep) ramps the *Sn* peak should occur within the ramp region (Fig. 9, left column). However, we are not confident that this is a true observation of an *Sn* focusing peak because our secondary dataset from the Gangdese-92 array does not show the same feature (Supplementary materials S7). The lack of an *Sn* focusing peak on the Gangdese-92 array, that recorded events with almost perpendicular incidence to the Moho ramp, implies that obliquity of ray-paths to the ramp is likely not the cause of our inability to observe amplitude variations due to the ramp. Observations of individual amplitudes in real data are subject to many variables such as site effects (which likely is strong in the HiCLIMB data based on the coherence seen in Fig.10 a-d), anelastic attenuation, and small-scale heterogeneities that could completely erase the *Sn* and *Lg* ramp-traversal signatures in our synthetics.

The lack of unequivocal observations of ramp effects in the *Sn* and *Lg* amplitude data is disappointing in that we cannot confirm the predictions of our synthetics from an amplitude perspective, but the negligible influence of the ramp on amplitudes is a *positive* result for the ability of the *Sn/Lg* method to distinguish below-Moho from above-Moho earthquakes. The *Sn/Lg* method is robust because it is largely immune to site effects, due to ratioing of the two portions of the same waveform recorded at the same location. Hence *Sn/Lg* ratios (Fig. 10e), unlike individual *Sn* and *Lg* amplitudes (Fig. 10 a-d), do not show any strong correlation with station locations and *Sn/Lg* ratios span similar values for both the southern and the northern earthquakes. We can separate our events into two groups either visually (Fig. 10e), or more quantitatively according to whether at least half of station *Sn/Lg* values are above or below our previous experimental threshold for this region (Wang and Klemperer, 2021), *Sn/Lg* =2. Southern event S1 has a single station and S6 has no station recording *Sn/Lg* >2 (Figs.10, 11): we believe both are crustal earthquakes. In contrast, events S2, S4, and S5 have >50% stations reporting *Sn/Lg* >2 (Figs.10, 11), and visually they behave like northern events WT1 and WT2 (Fig. 10e), which have previously been identified as upper-mantle events (Wang and Klemperer, 2021). This distinction is particularly clear >100 km north of the end of the ramp, and remains clear across most of the northern attenuation zone. Measured across all the stations, southern event S3 has just 39% of measurements with *Sn/Lg* > 2 (Fig. 11), but this rises to 52% if we only consider stations >100 km north of the ramp end (Supplementary materials S8). If the catalog depths for S3 and S4 are correct (~60 km) they are certainly below-Moho events. A full-waveform inversion put S3 at 53 km (Baur, 2007), clearly below the local Moho (Singh et al., 2015; Mitra et al., 2018), a conclusion (weakly) supported by our *Sn/Lg* results. S5 has an arbitrarily assigned depth of 10 km, which is not a useful determinant of the real depth, and based on the *Sn/Lg* data we believe it is in fact a sub-Moho event. Events S1 and S2 have depths ~35 km, around Moho depth (Singh et al., 2015; Mitra et al., 2018) yet our method suggests S2 occurred below the Moho and S1 above it. S6, with a relatively reliable catalog depth of 16.1 km, in the upper crust, is also suggested by our *Sn/Lg* criterion to be a crustal earthquake. These results show that although there is in general a positive correlation of *Sn/Lg* measurements with catalog depth (Song and Klemperer, 2024), there could also be inconsistencies particularly for the case of S5. Because comparison between different Himalayan catalogs shows numerous large depth discrepancies (Song and Klemperer, 2024), and dedicated re-location efforts have found some egregious catalog mis-locations (Craig et al., 2023), we suggest that our determination of

S5 as a sub-Moho earthquake from its *Sn/Lg* character may be more reliable than the assigned catalog depth.

5.2 Record sections

To further investigate the excitation of *Sn* and *Lg* for the southern events, we turn to their record sections (Fig. 11). normalized to the maximum value on each trace to highlight relative amplitude changes within a trace. For our current dataset, the *Sn* and *Lg* windows do not overlap, making their amplitude measurements distinct.

For the four events that we believe are of mantle origin (S2, S3, S4 and S5), clear *Sn* excitation can be observed in the middle part of the record section, at distances >~100 km north of the YZS (the Moho ramp end, labelled as 0 on the upper x-axes of the record sections, Fig. 11). At distances >~400 km beyond YZS there is some diminution of *Sn*, as waves reaching these stations have propagated partly within the region of high *Sn* attenuation (Fig. 2) (Barron & Priestley, 2009). Although *Sn* is clearly strongly excited for S3, the *Sn* energy arrives towards the end of the *Sn* window (Fig. 11). This likely represents a delayed *Sn* arrival rather than incorporation of early *Lg* into the *Sn* window, because early *Lg* should be followed by stronger subsequent *Lg* waves (Fig. 4, f&g) yet the energy in the *Sn* window is already the strongest in the entire record. Because our standard *Sn* window does not capture much of the *Sn* wavetrain for S3, inevitably *Sn/Lg* – calculated as the ratio of the RMS amplitudes of the respective windows – is lower than expected, explaining why only 39% of stations record *Sn/Lg* > 2. This analysis, and the clear increase in *Sn/Lg* for stations ~100 km north of the ramp (Supplementary materials S8) persuade us that S3 is indeed a mantle earthquake. The S1 and S6 record sections are quite different from S2, S3, S4 and S5. Neither S1 nor S6 shows significant *Sn* excitation relative to *Lg* excitation, and they do not show increase in *Sn/Lg* for stations ~100 km north of the ramp (Supplementary materials S8), further corroborating their crustal origin.

The *Lg* wavetrains for shallow events S1 and S6 have rather uniform amplitudes across the HiCLIMB array, but *Lg* varies dramatically for likely below-Moho events S2–S5. A common pattern for S2–S5 is that the southernmost few traces (<~15 km beyond YZS for S2 & S3, and <~100 km beyond YZS for S4 & S5) have *Lg* wavetrains comparable to, or even larger than (S3 and S4) their respective *Sn* wavetrains; then the *Lg* wavetrain becomes uniformly low amplitude further north. We believe this pattern may be a signature of enhanced *Lg* due to *Sn*-to-*Lg* conversion at the ramp. If true, it means *Sn*-to-*Lg* conversions waves may not persist in the crust for long distances, and may attenuate much faster than predicted by our modelling (which uses a scatterer-free crust). Note that the relative change of *Lg* amplitudes across the array that is obvious for events S2–S5 in their record sections, i.e. by comparison within traces (Fig. 11), is not obvious when looking only at the array-normalized *Lg* amplitudes (Fig. 10b), which are essentially the same as absolute amplitudes.

5.3 *Lg* HF/LF, ratio of *Lg* amplitudes at higher and lower frequencies

Another possibility to identify *Sn*-to-*Lg* conversion in real data, instead of relying on observing an increase of *Lg* amplitudes that can be strongly influenced by factors such as site effects (Fig. 10b), is the enrichment of high-frequency (HF) *Lg*. We analyze our twelve HiCLIMB earthquakes and four Gangdese-92 earthquakes exactly as we processed our synthetics. We have no measurements from within the *Sn*-attenuation region (Fig. 2): Gangdese-92 did not extend into this area, and the HiCLIMB stations here all lack high-quality low-frequency (LF) data. For *Lg* from the six southern events (Fig. 12a), we see the southern few stations, in particular those within the ramp region (negative distances), do have a much larger high-frequency component compared to the more northern stations, where *Lg* HF/LF ratio is more uniform. The peaking of *Lg* HF/LF may be smaller for the crustal events (open symbols) than for the mantle events colored symbols, Fig. 12a), as predicted by synthetics (Fig. 8a). For the four events recorded on the Gangdese-92 array, we more clearly see the rise of *Lg* HF/LF associated with the end of the ramp (Fig. 12c) because there are more stations vertically above the Moho ramp. However, we do not see an *Lg* HF/LF peak associated with the end of the Moho ramp for the six northern events (Fig. 12b), because these events have not traversed the ramp.

6. Discussion

We now bring together our numerical and observational results, to address our three main results: the ability to use *Sn/Lg* to recognize below-Moho earthquakes even in the presence of significant crustal thickening, our identification of *Sn*-to-*Lg* conversion in real data, and the value of *Lg* frequency content as another discriminant for continental mantle earthquakes.

Our numerical results (Figs. 6–9) show that significant Moho topography, that locally enhances *Sn* amplitudes and more regionally enhances *Lg* amplitudes, does not strongly influence *Sn/Lg* ratios which remain useful as a comparative measure to separate mantle and crustal earthquakes. The resilience of the *Sn/Lg* method to crustal thickening is clear because *Sn/Lg* ratios for the deeper-lid (65-km) and shallow-lid (35-km) events are always above the *Sn/Lg* ratios for the mid-crustal earthquake (15-km) at the same distance (Figs. 6g,h,i, 9g,h,i). The best separation, an order of magnitude, occurs between our shallow-lid earthquake and our mid-crustal earthquake at stations far beyond the end of the ramp, because of the ramp-transient nature of *Sn* amplitude perturbations and modest *Sn*-to-*Lg* conversion for shallow-lid earthquakes.

Thus our simulation results show we can apply *Sn/Lg* criteria to identify mantle earthquakes regardless of the presence of a Moho-thickening ramp. Observations of *Sn* and *Lg* on the HiCLIMB (Fig. 10 a&b) and Gangdese-92 arrays (Supplementary S7) show less significant effects than our simulation results (Figs. 6&9, right columns) that therefore likely represent the strongest possible scenarios for ramp effects on *Sn/Lg* signatures. Our HiCLIMB events are strongly influenced by site effects and are obliquely incident on the array (though as noted above, this obliquity is likely unimportant), whereas our Gangdese-92 events do not exhibit strong site effects and are nearly in-line with the array. Nonetheless, neither set of events shows either the predicted strong focusing of *Sn* near the end of the ramp nor the predicted sustained increase of *Lg* energy beyond the end of the ramp (Fig.10, Supplementary materials S7). We believe these inconsistencies between data and simulations originate from the absence in our models of small-scale features such as inhomogeneities in the crust or less-smooth Moho topography. Additional small-scale features should spatially smooth a localized feature such as the *Sn* peak (Figs. 6&9, left column), and selectively attenuate the higher-frequency *Lg* in the crust produced by *Sn*-to-*Lg* conversion (Fig. 8), which we discuss more below. Our observations on individual *Sn* and *Lg* waves agree with findings in the North Sea (Mendi et al., 1997) that regional waves are more influenced by small-scale scatterers than large-scale features. Because the largest perturbations from the reference model due to a Moho ramp are the *Sn* peak above and the increased *Lg* beyond the ramp, smoothing out these effects in the real data likely means *Sn/Lg* in the real world is even more robust than predicted by our simulations.

We can directly compare *Sn/Lg* for events traversing one of Earth's largest Moho ramps with *Sn/Lg* for events traversing relatively uniform Moho topography (Figs. 2, 9e). Using a previously established *Sn/Lg* threshold that identified two new below-Moho earthquakes in NW Tibet (Wang and Klemperer, 2021), we can identify four earthquakes (S2, S3, S4, S5) south of the MCT that nucleated below the Moho, including one previously tentatively identified as such (S3=H82 of Baur, 2007) and one that has a nominal (assigned) catalog depth of 10 km (S5). We can similarly show that a different event with a catalog depth close to the Moho (S1) is in fact a crustal event. We emphasize that these conclusions are quite reliable, as they are based on measurements on multiple stations that show *Sn/Lg* significantly larger than the regional low-*Sn/Lg* baseline established for multiple nominally shallow earthquakes in both northern India and in northwestern Tibet.

Sn-to-*Lg* converted waves maybe most easily identified in the frequency domain (Fig. 12), rather than in the amplitude domain (Figs. 10&11), through *Lg* HF/LF. This diagnostic is motivated by early mode-coupling studies (Maupin, 1989) (Fig. 3) and verified with our full-waveform synthetics (Fig. 8). Two groups of events with significant Moho ramp crossing recorded on two separate arrays both exhibit increase of *Lg* HF/LF (Fig. 12a&c) associated with the end of the ramp, but another group of non-Moho-ramp crossing events recorded on one of the same arrays does not show this (Fig. 12b). Hence, we believe *Lg* HF/LF is a rather robust signature of *Sn*-converted-*Lg* waves. This implies that the enhanced *Lg* above and close to the ramp on the record sections of the southern mantle events S2–S5 (Fig. 11) represents *Sn*-to-*Lg* conversions, enriched in high-frequencies. The enriched HF content for *Lg* close to the ramp corroborates our suspicion that small-scale crustal scatterers are the reason we do not see persistent high *Lg* energy after conversion in real data, unlike in the numerical results.

A prominent feature of *Lg* HF/LF is the clear separation of mantle and crustal earthquakes recorded on HiCLIMB (Fig. 12a&b) following our interpretations based on *Sn/Lg* (Fig. 10e), whereas the overlapping of *Lg* HF/LF for the Gangdese events (Fig. 12c) matches their overlapping *Sn/Lg* values (Supplementary materials S7). This can be understood from a normal-mode perspective in that the only *Lg* energy excitable by a mantle earthquake is associated with lower-frequency Airy phases (Knopoff et al., 1973) that could have a displacement/strain eigenfunction sampling the mantle lid to some depths (Wang and Klemperer, 2023, their Fig. 3a), whereas the higher-frequency *Lg* Airy phases have displacement/strain eigenfunctions much more tightly bounded within the crust (Wang and Klemperer, 2023, their Fig. 3b). In our reference model with flat Moho at 30 km, a source 5 km above Moho (z=25 km) has essentially the same *Lg* HF/LF as a source 15 km above Moho (z= 15 km), and both are clearly distinct from *Lg* HF/LF for a source that is 5 km below the Moho (z=35 km). The deeper-lid earthquakes at z = 65 and 95 km have similar *Lg* HF/LF as the crustal earthquakes at short offsets because of the artificial inclusion of *Sn* in our measurement windows (Fig. 4 a,c,d,e), but their *Lg* HF/LF quickly drops beyond ~400-500 km epicentral distance as *Sn* exits the *Lg* window. Hence, like *Sn/Lg*, *Lg* HF/LF is not particularly sensitive to absolute source depths, but rather to their relative position with respect to the Moho as predicted by the normal-mode explanation, so can also be used as a discriminant for sub-Moho earthquakes. This frequency discriminant *Lg* HF/LF is even simpler than *Sn/Lg* because *Lg* HF/LF remains almost constant with epicentral distance (Fig. 13a) (apart from the artificial sinusoidal oscillations due to overlapping *Lg* and *Sn* windows, see Section 3) in contrast to *Sn/Lg* (Fig. 5), and because *Lg* in general is a much simpler wave than *Sn* (i.e. crustal waveguide for *Lg* vs. whispering-gallery waveguide for *Sn*).

We further selected crustal thickening models with small d's and w's in order to capture the effects a Moho thickening ramp can produce on *Lg* HF/LF (Fig. 13 b-e). The difference between *Lg* HF/LF for a crustal and an upper-mantle earthquake is present for all our selected models (beyond ~400–500 km where some *Sn* is present in the *Lg* window), representing among the strongest effects a Moho ramp can produce. We note d has a stronger effect than w in terms of increasing the upper-mantle event's high-frequency *Lg* thereby raising its *Lg* HF/LF, and when $d = 100$ km, the separation with the mid-crustal event is quite small (Fig. 13 b,d&e). In addition, our synthetics show overlapping *Lg* HF/LF for the deeper-lid event and the mid-crustal earthquake in these extreme models (Fig. 13 b–e). However, these are a worst-case because the high-frequency enriched *Lg* due to *Sn*-to-*Lg* conversion, that is persistent at large distances in our synthetics (Fig. 8), in real data fades away quickly after the end of the ramp as observed in data

(Figs. 11&12). Hence, it is unlikely *Lg* HF/LF will be undistinguishable for upper-mantle and mid-crustal earthquakes nor will it mis-classify a deeper-lid earthquake as a crustal earthquake if measurements are made on sufficient stations beyond the end of the ramp. In ongoing work, we are exploring the correlation between *Lg* HF/LF and *Sn/Lg* amplitude-ratio discriminants, and their joint potential to resolve relative location of earthquakes above and below the Moho.

7. Conclusion

We enhanced the code AxiSEM3D to perform 2.5D regional wave simulations across a Moho ramp and achieved a combination of higher frequency ranges and longer propagation distances than other recent studies. Most notably, our modifications enabled checking the representation of an undulated geometry within AxiSEM3D and using this technique to stretch a uniform mesh so that the computed wavefield can be shown at the correct positions, avoiding wavefield distortions that will be visible for simulations at our scale (i.e. regional, vs. global).

We compare our numerical results in a 1D reference model, with flat Moho, with previous studies on *Sn* and *Lg* geometrical spreading to confirm the accuracy of our numerical approach. In addition, with this benchmarking exercise we emphasize the fact that regional-wave arrival windows, as defined by group velocities, cannot be fine-tuned in real data. The windows will always overlap leading to artificial abrupt or oscillatory changes in measured amplitudes and frequencies whenever an *Sn* or *Lg* phase moves in or out of its window (a phenomenon previously noted by Yang (2002)). It is likely that mischaracterization of phases contributed to an over-estimation of *Sn* amplitude increase at ~700–1,300 km by Yang et al. (2007) leading to an inaccuracy in their *Sn* geometrical spreading model (Fig. 4b).

As we vary distance to ramp start *d* and ramp width *w* in our crustal thickening model (Fig. 1), the synthetics for *Lg* absolute amplitudes are relatively simple and consistently display sustained increase *Lg* for amplitudes as well as *Lg* HF/LF across the ramp, though to the smallest degree for the crustal source. On the other hand, synthetic *Sn* absolute amplitudes are much more complicated due to its complex propagation path as an interference head wave. Nonetheless, commonalities are present, including the *Sn* focusing peak around the ramp end, and the return to *Sn* amplitudes similar to the reference model at larger distances for almost all parameter ranges tested. These phenomena are closely related to the shape of the *Sn* waveguide (Fig. 7). Even with the presence of these perturbations on individual amplitudes, among all cases tested in our simulations, *Sn/Lg* ratios for mid-crustal earthquakes are persistently lower than for mantle earthquakes on noise-free synthetics, and potential confusions are unlikely when using a recording array with varying source-station geometry.

There are substantial differences between real-world data and synthetics for individual *Sn* and *Lg* absolute (or array-normalized) amplitudes, as in addition to factors like site effects, the real world contains many finer-scale details, such as crustal scatters and irregular Moho/ramp surfaces that tend to average the sharp *Sn* focusing peak and sustained *Lg* amplitude increase seen in our synthetics. We therefore believe our ramp models provide a worst-case scenario for the utility of *Sn/Lg* in the real world as waveforms are smoother in the real world. We verified the effectiveness of *Sn/Lg* through direct comparison with ramp-crossing and non-ramp-crossing events from southern and northwestern Tibet, recorded on the same array with roughly the same epicentral distances (Fig. 10e; Supplementary materials S6), providing strong evidence for four mantle earthquakes in northern India.

Sn-to-*Lg* converted waves are generally hard to recognize from their amplitudes, though not impossible (Fig. 11), but can more easily be identified by the shift of *Lg* frequency content, as shown here with our full-waveform synthetics (Fig. 8) and demonstrated with non-ramp-crossing events (Fig. 12b) and two sets of ramp-crossing events recorded on two different arrays (Fig. 12a,c). *Lg* HF/LF is a promising new discriminant to identify continental mantle earthquakes from their decreased *Lg* HF/LF as predicted by normal-mode theory and verified in both our reference and ramp models (Fig. 13).

Acknowledgements

The authors wish to thank Tarje Nissen-Meyer and Kuangdai Leng for insightful discussions regarding the software package AxiSEM3D. Gangdese-92 data collection was funded by the National Natural Science Foundation of China Grant No. 41374109. This study was funded by Stanford University, and in part by NSF-EAR-1628282.

CREDIT statement

Shiqi Wang:

Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing-Original draft, Writing-Review and editing, Visualization.

Simon Klemperer:

Methodology, Validation, Formal analysis, Investigation, Writing-Review and editing, Visualization, Funding acquisition.

Data Availability Statement

All seismic data analyzed in this paper are available via https://www.fdsn.org/networks/detail/XF_2002 (HiCLIMB data) and at https://doi.org/10.5281/zenodo.10971752 (Gangdese-92 data). Our custom-version AxiSEM3D can be found at https://github.com/axelwang/AxiSEM3D_Modified.

Competing interest

The authors declare no competing interest.

References

Aki, K. and Richards, P.G., 2002. Quantitative seismology. University Science Books.

Al-Attar, D. and Crawford, O., 2016. Particle relabelling transformations in elastodynamics. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, *205*(1), pp.575-593.

Anderson, B.D., Benzley, S.E. and Owen, S.J., 2009. Automatic all quadrilateral mesh adaption through refinement and coarsening. In *Proceedings of the 18th international meshing roundtable* (pp. 557-574). Springer Berlin Heidelberg.

Avants, M., Lay, T., Xie, X.B. and Yang, X., 2011. Effects of 2D Random Velocity Heterogeneities in the Mantle Lid and Moho Topography on P n Geometric Spreading. *Bulletin of the Seismological Society of America*, *101*(1), pp.126-140.

Baker, G.E., Stevens, J. and Xu, H., 2004. Lg group velocity: A depth discriminant revisited. *Bulletin of the Seismological Society of America*, *94*(2), pp.722-739.

Bakir, A.C. and Nowack, R.L., 2012. Modeling seismic attributes of Pn waves using the spectral-element method. *Pure and applied geophysics*, *169*, pp.1539-1556.

Barron, J. and Priestley, K., 2009. Observations of frequency-dependent S n propagation in Northern Tibet. *Geophysical Journal International*, *179*(1), pp.475-488.

Baur, Jan, 2007. Seismotectonics of the Himalayas and the Tibetan Plateau: moment tensor analysis of regional seismograms. MA thesis. Oregon State University.

Bostock, M.G. and Kennett, B.L.N., 1990. The effect of 3-D structure on Lg propagation patterns. *Geophysical Journal International*, *101*(2), pp.355-365.

Bottero, A., Cristini, P., Komatitsch, D. and Asch, M., 2016. An axisymmetric time-domain spectral-element method for full-wave simulations: Application to ocean acoustics. *The journal of the acoustical society of America*, *140*(5), pp.3520-3530.

Cao, S. and Muirhead, K.J., 1993. Finite difference modelling of Lg blockage. *Geophysical Journal International*, *115*(1), pp.85-96.

Červený, V., and R. Ravindra (1971). Theory of Seismic Head Waves, University of Toronto Press, Toronto.

Chen, W.P. and Molnar, P., 1983. Focal depths of intracontinental and intraplate earthquakes and their implications for the thermal and mechanical properties of the lithosphere. Journal of Geophysical Research: Solid Earth, 88(B5), pp.4183-4214.

Chen, W.P. and Yang, Z., 2004. Earthquakes beneath the Himalayas and Tibet: Evidence for strong lithospheric mantle. *Science*, *304*(5679), pp.1949-1952.

Cook, F.A., White, D.J., Jones, A.G., Eaton, D.W., Hall, J. and Clowes, R.M., 2010. How the crust meets the mantle: Lithoprobe perspectives on the Mohorovičić discontinuity and crust–mantle transitionThis article is one of a series of papers published in this Special Issue on the theme Lithoprobe—parameters, processes, and the evolution of a continent. *Canadian Journal of Earth Sciences*, *47*(4), pp.315-351.

Craig, T.J., Jackson, J.A., Priestley, K. and McKenzie, D., 2011. Earthquake distribution patterns in Africa: their relationship to variations in lithospheric and geological structure, and their rheological implications. *Geophysical Journal International*, *185*(1), pp.403-434.

Craig, T.J., Jackson, J., Priestley, K. and Ekström, G., 2023. A Cautionary Tale: examples of the mis-location of small earthquakes beneath the Tibetan plateau by routine approaches. *Geophysical Journal International*, *233*(3), pp.2021-2038.

Drake, L.A., 1972. Love and Rayleigh waves in nonhorizontally layered media. *Bulletin of the Seismological Society of America*, *62*(5), pp.1241-1258.

Dziewonski, A.M. and Anderson, D.L., 1981. Preliminary reference Earth model. *Physics of the earth and planetary interiors*, *25*(4), pp.297-356.

Furumura, T., Hong, T.K. and Kennett, B.L., 2014. Lg wave propagation in the area around Japan: observations and simulations. *Progress in Earth and Planetary Science*, *1*, pp.1-20.

Gao, R., Lu, Z., Klemperer, S.L., Wang, H., Dong, S., Li, W. and Li, H., 2016. Crustal-scale duplexing beneath the Yarlung Zangbo suture in the western Himalaya. *Nature Geoscience*, *9*(7), pp.555-560.

Haindl, C., Leng, K. and Nissen-Meyer, T., 2021. A 3D complexity-adaptive approach to explore sparsity in elastic wave propagation. *Geophysics*, *86*(5), pp.T321-T335.

Kebeasy, T.R.M. and Husebye, E.S., 2003. A finite-difference approach for simulating ground responses in sedimentary basins: quantitative modelling of the Nile Valley, Egypt. *Geophysical Journal International*, *154*(3), pp.913-924.

Kennett, B.L.N., 1972. Seismic waves in laterally inhomogeneous media. *Geophysical Journal International*, *27*(3), pp.301-325.

Kennett, B.L.N., 1984. Guided wave propagation in laterally varying media—I. Theoretical development. *Geophysical Journal International*, *79*(1), pp.235-255.

Kennett, B.L.N., 1986. Lg waves and structural boundaries. *Bulletin of the Seismological Society of America*, *76*(4), pp.1133-1141.

Knopoff, L., Schwab, F. and Kauselt, E., 1973. Interpretation of Lg. *Geophysical Journal International*, *33*(4), pp.389-404.

Laske, G., Masters, G., Ma, Z. and Pasyanos, M., 2013, April. Update on CRUST1. 0—A 1-degree global model of Earth's crust. In *Geophysical research abstracts* (Vol. 15, No. 15, p. 2658). Vienna, Austria: EGU General Assembly.

Leng, K., Nissen-Meyer, T. and van Driel, M., 2016. Efficient global wave propagation adapted to 3-D structural complexity: a pseudospectral/spectral-element approach. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, *207*(3), pp.1700-1721.

Leng, K., Nissen-Meyer, T., Van Driel, M., Hosseini, K. and Al-Attar, D., 2019. AxiSEM3D: broad-band seismic wavefields in 3-D global earth models with undulating discontinuities. *Geophysical Journal International*, *217*(3), pp.2125-2146.

Li, D., Helmberger, D., Clayton, R.W. and Sun, D., 2014. Global synthetic seismograms using a 2-D finite-difference method. *Geophysical Journal International*, *197*(2), pp.1166-1183.

Maggi, A., Jackson, J.A., Priestley, K. and Baker, C., 2000. A re-assessment of focal depth distributions in southern Iran, the Tien Shan and northern India: Do earthquakes really occur in the continental mantle?. *Geophysical Journal International, 143*(3), pp.629-661.

Martin, A.J., 2017. A review of definitions of the Himalayan Main Central Thrust. *International Journal of Earth Sciences*, *106*, pp.2131-2145.

Maupin, V., 1988. Surface waves across 2-D structures: a method based on coupled local modes. *Geophysical Journal International*, *93*(1), pp.173-185.

Maupin, V., 1989. Numerical modelling of Lg wave propagation across the North Sea Central Graben. *Geophysical Journal International*, *99*(2), pp.273-283.

McNamara, D.E., Owens, T.J. and Walter, W.R., 1995. Observations of regional phase propagation across the Tibetan Plateau. *Journal of Geophysical Research: Solid Earth*, *100*(B11), pp.22215-22229.

McNamara, D.E., Owens, T.J. and Walter, W.R., 1996. Propagation characteristics of Lg across the Tibetan Plateau. *Bulletin of the Seismological Society of America*, *86*(2), pp.457-469.

Mendi, C.D., Ruud, B.O. and Husebye, E.S., 1997. The north sea Lg-blockage puzzle. *Geophysical Journal International*, *130*(3), pp.669-680.

Menke, W.H. and Richards, P.G., 1980. Crust-mantle whispering gallery phases: A deterministic model of teleseismic Pn wave propagation. *Journal of Geophysical Research: Solid Earth*, *85*(B10), pp.5416-5422.

Mitra, S., Priestley, K., Bhattacharyya, A.K. and Gaur, V.K., 2005. Crustal structure and earthquake focal depths beneath northeastern India and southern Tibet. *Geophysical Journal International*, *160*(1), pp.227-248.

Mitra, S., Priestley, K.F., Borah, K. and Gaur, V.K., 2018. Crustal structure and evolution of the Eastern Himalayan plate boundary system, Northeast India. *Journal of Geophysical Research: Solid Earth*, *123*(1), pp.621-640.

Mousavi, S.M., Cramer, C.H. and Langston, C.A., 2014. Average QLg, QSn, and observation of Lg blockage in the continental margin of Nova Scotia. *Journal of Geophysical Research: Solid Earth*, *119*(10), pp.7722-7744.

Nábělek, J., Hetényi, G., Vergne, J., Sapkota, S., Kafle, B., Jiang, M., Su, H., Chen, J., Huang, B.S. and Team, T.H.C., 2009. Underplating in the Himalaya-Tibet collision zone revealed by the Hi-CLIMB experiment. *Science*, *325*(5946), pp.1371-1374.

Nissen-Meyer, T., 2022, *personal communications*.

Odom, R.I., 1986. A coupled mode examination of irregular waveguides including the continuum spectrum. *Geophysical Journal International*, *86*(2), pp.425-453.

Oliver, J. and Ewing, M., 1957. Higher modes of continental Rayleigh waves. *Bulletin of the Seismological Society of America*, *47*(3), pp.187-204.

Patton, H.J. and Walter, W.R., 1993. Regional moment: magnitude relations for earthquakes and explosions. *Geophysical Research Letters*, *20*(4), pp.277-280.

PDE Bulletin, 2024. U.S. Geological Survey National Earthquake Information Center (NEIC) Preliminary Determination of Epicenters (PDE) Bulletin, last accessed 03/01/2024 at URL *http://earthquake.usgs.gov/data/comcat/catalog/us/*

Priestley, K., Jackson, J. and McKenzie, D., 2008. Lithospheric structure and deep earthquakes beneath India, the Himalaya and southern Tibet. *Geophysical Journal International*, *172*(1), pp.345-362.

Prieto, G.A., Froment, B., Yu, C., Poli, P. and Abercrombie, R., 2017. Earthquake rupture below the brittle-ductile transition in continental lithospheric mantle. *Science Advances*, *3*(3), p.e1602642.

Przybilla, J. and Korn, M., 2008. Monte Carlo simulation of radiative energy transfer in continuous elastic random media—three-component envelopes and numerical validation. *Geophysical Journal International*, *173*(2), pp.566-576.

Przybilla, J., Korn, M. and Wegler, U., 2006. Radiative transfer of elastic waves versus finite difference simulations in two-dimensional random media. *Journal of Geophysical Research: Solid Earth*, *111*(B4).

Rodgers, A.J., Petersson, N.A., Pitarka, A., McCallen, D.B., Sjogreen, B. and Abrahamson, N., 2019. Broadband (0–5 Hz) fully deterministic 3D ground-motion simulations of a magnitude 7.0 Hayward fault earthquake: Comparison with empirical ground-motion models and 3D path and site effects from source normalized intensities. *Seismological Research Letters*, *90*(3), pp.1268-1284.

Rodgers, A.J., Pitarka, A., Pankajakshan, R., Sjögreen, B. and Petersson, N.A., 2020. Regional-scale 3D ground-motion simulations of Mw 7 earthquakes on the Hayward fault, northern California resolving frequencies 0–10 Hz and including site-response corrections. *Bulletin of the Seismological Society of America*, *110*(6), pp.2862-2881.

Sanborn, C.J., Cormier, V.F. and Fitzpatrick, M., 2017. Combined effects of deterministic and statistical structure on high-frequency regional seismograms. *Geophysical Journal International*, *210*(2), pp.1143-1159.

Schulte-Pelkum, V., Monsalve, G., Sheehan, A.F., Shearer, P., Wu, F. and Rajaure, S., 2019. Mantle earthquakes in the Himalayan collision zone. *Geology*, *47*(9), pp.815-819.

Schwab, F., Kausel, E. and Knopoff, L., 1974. Interpretation of Sa for a shield structure. *Geophysical Journal International*, *36*(3), pp.737-742.

Schulte-Pelkum, V., Monsalve, G., Sheehan, A.F., Shearer, P., Wu, F. and Rajaure, S., 2019. Mantle earthquakes in the Himalayan collision zone. *Geology*, *47*(9), pp.815-819.

Seismological Bulletins of the Indian National Center for Seismology, 2024. Last accessed 04/08/2024 at URL *https://seismo.gov.in/seismological-bulletins*.

Sens-Schönfelder, C., Margerin, L. and Campillo, M., 2009. Laterally heterogeneous scattering explains Lg blockage in the Pyrenees. *Journal of Geophysical Research: Solid Earth*, *114*(B7).

Shi, D., Zhao, W., Klemperer, S.L., Wu, Z., Mechie, J., Shi, J., Xue, G. and Su, H., 2016. West–east transition from underplating to steep subduction in the India–Tibet collision zone revealed by receiver-function profiles. *Earth and Planetary Science Letters*, *452*, pp.171-177.

Singh, A., Singh, C. and Kennett, B.L.N., 2015. A review of crust and upper mantle structure beneath the Indian subcontinent. *Tectonophysics*, *644*, pp.1-21.

Stephens, C. and Isacks, B.L., 1977. Toward an understanding of Sn: Normal modes of Love waves in an oceanic structure. *Bulletin of the Seismological Society of America*, *67*(1), pp.69-78.

Szenicer, A., Leng, K. and Nissen-Meyer, T., 2020. A complexity-driven framework for waveform tomography with discrete adjoints. *Geophysical Journal International*, *223*(2), pp.1247-1264.

Tolstoy, I., 1956. Resonant frequencies and high modes in layered wave guides. *The Journal of the Acoustical Society of America*, *28*(6), pp.1182-1192.

Tromp, J., 1994. A coupled local-mode analysis of surface-wave propagation in a laterally heterogeneous waveguide. *Geophysical Journal International*, *117*(1), pp.153-161.

van Driel, M., Krischer, L., Stähler, S.C., Hosseini, K. and Nissen-Meyer, T., 2015. Instaseis: instant global seismograms based on a broadband waveform database. *Solid Earth*, *6*(2), pp.701-717.

von Seggern, D.H., 2012. Exploring reflection and transmission coefficients in elastic media. *Math. J.*, 14, p.50.

Wang, H., Ni, S., Jin, P., Liu, W., Xu, X., Pan, C., Zhu, H. and Xu, H., 2017. Anomalous Pn Amplitudes through the Southeastern Tarim Basin and Western Tien Shan along Two Profiles: Observations and Interpretations. *Bulletin of the Seismological Society of America*, *107*(2), pp.760-769.

Wang, S. and Klemperer, S.L., 2021. Love-wave normal modes discriminate between upper-mantle and crustal earthquakes: Simulation and demonstration in Tibet. *Earth and Planetary Science Letters*, *571*, p.117089. https://doi.org/10.1016/j.epsl.2021.117089.

Wang, S. and Klemperer, S.L., 2024. Gangdese-92 Data [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.10971752.

Xie, J., 1996. *Synthetic and Observational Study of Pn Excitation and Propagation in Central Asia*. Saint Louis Univ MO Dept of Earth and Atmospheric Sciences.

Xie, X.B. and Lay, T., 2017a. Effects of laterally varying mantle lid velocity gradient and crustal thickness on Pn geometric spreading with application to the north Korean test site. *Bulletin of the Seismological Society of America*, *107*(1), pp.22-33.

Xie, X.B. and Lay, T., 2017b. Frequency-dependent effects of 2D random velocity heterogeneities in the mantle lid on Pn geometric spreading. *Bulletin of the Seismological Society of America*, *107*(1), pp.482-488.

Yang, X., 2002. A numerical investigation of Lg geometrical spreading. *Bulletin of the Seismological Society of America*, *92*(8), pp.3067-3079.

Yang, X., Lay, T., Xie, X.B. and Thorne, M.S., 2007. Geometric spreading of Pn and Sn in a spherical Earth model. *Bulletin of the Seismological Society of America*, *97*(6), pp.2053-2065.

Yang, Z. and Chen, W.P., 2010. Earthquakes along the East African Rift System: A multiscale, system-wide perspective. Journal of Geophysical Research: Solid Earth, 115(B12).

Zhang, M. and Wen, L., 2013. High-precision location and yield of North Korea's 2013 nuclear test. *Geophysical Research Letters*, *40*(12), pp.2941-2946.

Zhu, L. and Helmberger, D.V., 1996. Intermediate depth earthquakes beneath the India-Tibet collision zone. *Geophysical Research Letters*, *23*(5), pp.435-438.

Figure and table legends

Table 1. Earthquakes recorded on the HiCLIMB array. Southern events are named S1–S6. These events nucleated in northern India and cross a significant Moho ramp before being recorded by HiCLIMB (Fig. 2). The distance to the ramp (*d*) and ramp-width (*w*) are shown as ranges because of the different azimuth (hence obliquity to the ramp) from each earthquake to the southern and northern limits of the HiCLIMB stations (Fig. 2). The six events with no values for *d* and *w* comprise our 'northern' events that do not cross significant Moho topography before reaching the stations (Wang and Klemperer, 2021). Magnitude and depth data are from PDE, 2024. Values in parentheses from Baur (2007). Italicized hypocentral locations and depths are from the Seismological Bulletins of the Indian National Center for Seismology.

Fig. 1. Computational model and a representative wavefield. Computational region extends to 2000 km in range and 230 km in depth. Thick black line shows the Moho, which is at 30 km on the left side, and transitions smoothly to 60 km through a 30-km high ramp whose width (*w*) and distance from source (*d*) are labelled. Small red stars represent the 3 source depths (*z*) we study, 15, 35 and 65 km, respectively. Thin black lines represent the top and bottom of the mantle low-velocity zone (LVZ) at 80 and 220 km. In this example the source is at 65 km depth. A snapshot wavefield (transverse component, filtered 1-5 Hz) is plotted at time 235.5s with amplitude shown in the color bar on lower left, showing multiply-reflected and interfering regional wave trains. The wavefield in the crust is complex as it is a combination of multiple reflections from the Moho top-side (*Lg*), under-side (*Sn*), as well as from just below the LVZ (*Sa*). The absence of visible reflections from the bottom of the computational domain, despite the clearly visible reflections from the bottom of the LVZ, demonstrates the performance of our absorbing boundary condition.

Fig. 2. Earthquakes and stations. (a) Earthquakes in India south of the Main Frontal Thrust (MFT) and in the Bhutan Himalaya (red stars, red labels S1–S6; Table 1) recorded on the HiCLIMB array (Nabelek et al., 2009) (purple triangles) after crossing a ~15–20 km high Moho ramp from thinner to thicker crust. Earthquakes in northwestern Tibet (black stars) (Wang and Klemperer, 2021) are recorded on the same array but their paths do not cross significant Moho undulations. Moho depths are interpolated from CRUST1.0 to 0.05° (Laske et al., 2013). Gangdese-92 array (Shi et al., 2015) (yellow triangles) recorded four nominally deep earthquakes ~ due south of the array (yellow stars, yellow labels G1-G4) (Supplementary materials table S7-1). YZS: Yarlung-Zangpo, BNS: Banggong-Nujiang, JRS: Jinsha River sutures. H: Himalaya, L: Lhasa, Q: Qiangtang, SG: Songpan-Ganzi terranes. MFT: Main Frontal thrust, KKF: Karakoram fault, KXF: Karakax fault. White dashed lines border a well-known attenuation zone for *Sn* (e.g. Barron and Priestley, 2009). Black double arrow indicates approximate distance from YZS to the attenuation zone. YZS represents the ending of the Moho ramp, while two thick green lines represent the Moho ramp beginning directly south of the arrays at the approximate location for

Main Central Thrust (MCT), which is too tortuous to show on our map (see main text). (b). Cartoon crustal and Moho cross-section along white solid line shown in (a), redrawn after Nabelek et al., 2009 based on their receiver function analysis on the HiCLIMB array.

Fig. 3. Normal-mode-coupling results. Transmission amplitude coefficients for Rayleigh waves due to perpendicular incidence onto a North Sea graben-type model, visualized from Figure 4 of Maupin (1989). The matrix is symmetric; for convenience we label each row as representing the incidence of a pure mode and each column as the converted amplitude with the amplitude coefficients representing the degree of partitioning of energy due to incidence onto a large-scale Moho depth variation. The calculations are done for 1 Hz, at which a strict separation (shown as black dashed lines) can be made between *Lg* (mode numbers ≤ 11) and *Sn* (mode numbers ≥ 12). Note the amplitude coefficients are typically large along the diagonal (no mode conversion), and are very small in the upper right and lower left sections of the figure as separated by the dashed lines. Because *Lg*-to-*Sn* coupling is strongest into the lowest *Sn* modes (12, 13, 14) *Lg*-to-*Sn* coupling preferentially excites the lower frequencies of *Sn*. For example, looking at the row for mode 9, the squares in columns 1-11 represent mode coupling to other *Lg*-forming normal modes, though most of the energy remains as mode 9 (highlighted with thick black border). Across the dashed line, squares in columns 12-25 represent mode coupling into *Sn*-forming normal modes, leading to *Lg*-to-*Sn* conversion with the strongest coupling to mode 13 (highlighted with dashed border), a low mode number for *Sn* normal modes. Similarly, *Sn*-to-*Lg* coupling is dominantly from the lower *Sn* modes (e.g. 12, 13, 14) to the higher *Lg*-forming-modes (e.g. 8, 9, 10, highlighted with dotted lines, contributing dominantly to higher-frequency *Lg* Airy phases.

Fig. 4. *Sn* and *Lg* in reference model (no ramp). Transverse-component displacements are shown. (a) *Lg* amplitude filtered 1–5 Hz, for three source depths, "mid-crustal" (15 km), "shallow-lid" (35 km) and "deeper-lid" (65 km). (b) *Sn* amplitude at 3 Hz for the same three source depths. Black lines in (a) and (b) are the best-fit models of Yang (2002) (*Lg*) and Yang et al. (2007) (*Sn*) that only predict relative amplitudes as a function of distance, so are set to be equal to our results at 200-km distance for *Lg*, and 300 km for *Sn* (the starting modelling distance in Yang et al. (2007). Our extrapolation of the Yang (2002) and Yang et al. (2007) formulae beyond the distance range they studied leads to large misfits at large offsets. Both (a) and (b) are log-log, amplitude vs. distance. Seismograms for symbols with black border are shown in (c)-(i) with corresponding labels. Two red vertical lines bound the *Sn* windows, and cyan lines bound the *Lg* windows. Full *Lg* windows are not shown for (f)-(i) as the focus there is on the *Sn* window. Seismogram amplitudes shown are absolute values without normalizations, in units of nanometers (nm).

Fig. 5. *Sn/Lg* in reference model (no ramp). *Sn/Lg* for three source depths, 15-km (mid-crust), 35-km (shallow-lid) and 65-km (deeper-lid) are clearly separated at all epicentral distances,

despite their individual variations with offset. A black fiducial line at *Sn/Lg* =0.2 further illustrates separation of crustal and mantle earthquakes.

Fig. 6. *Sn, Lg* amplitudes and *Sn/Lg* with varying distance to ramp start d but fixed ramp width *w* = 200 km. Rows, top to bottom, display results when the source is mid-crustal (15 km), shallow-lid (35 km) and deeper-lid (65 km). Columns, left to right, show amplitudes of *Sn* and *Lg* relative to the reference model, and *Sn/Lg*. *Left and middle columns* are plotted with data aligned at the ramp, with its beginning marked as a vertical dashed black line and end marked as a vertical solid black line (0 on the horizontal axis). A grey horizontal line at 1 marks no deviation from reference-model results. The vertical axis is plotted in $\log_{10}$ scale while the horizontal axis is linear. Note because of the ramp alignment and a fixed total simulation range, larger d has a shorter distance covered beyond the end of the ramp. *Right column* plots *Sn/Lg* against epicentral distance and superimposed on the reference-model results (grey inverted triangles). Each colored bar represents the end of the ramp for the correspondingly colored symbol (e.g. the red bar marks the end of the ramp at 300 km epicentral distance for d = 100 km (red circles), and its ramp starts outside the range of the plots). The total ramp ranges for the other cases are shown between the vertical lines (e.g. for d = 300 km (blue diamonds) the ramp range is between the red and blue bars). Both the vertical and horizontal axes are in $\log_{10}$ scale. A solid black line at *Sn/Lg* = 0.2 (as in Fig. 5) in all rows in the right column shows that despite the variability within each plot, *Sn/Lg* for our mantle earthquakes (h) and (i) is greater than *Sn/Lg* for our crustal earthquake (g) at every common offset.

Fig. 7. Disruption of *Sn* waveguide by the Moho ramp. (a)-(d) Representative wavefield snapshots (for z = 35 km, corresponding to the blue diamonds in Fig. 6 b,e,h) bandpass filtered 1-5 Hz and shown with same color scale for the amplitudes. (a) and (c): d = 300 km, w = 200 km; (b) and (d): reference model (flat Moho). UT: displacement on the transverse component, in meters. Moho and top of LVZ are marked by thick and thin black lines. At time = 73 s, before the wavefield interacts with the ramp, the ramp model (a) and reference model (b) show the same wavefield, with the first arrival due to a rather weak transmitted wave through the Moho from the leading strong sub-Moho wavefront, showing an effective *Sn* waveguide. At time = 100 s, after the wavefield in the ramp model starts to interact with the ramp, the transmitted wave becomes much stronger in the ramp model (c) than in the reference model (d) (red arrows) corresponding to the onset of the *Sn* peak right after the vertical black line (0 km, end of the ramp) in Fig. 6b. (e) Schematics of sub-Moho wavefront interacting with a flat Moho (black line/blue raypath) and with a Moho with a thickening ramp (grey line/dashed yellow raypath). (f) calculated energy partitioning for a transverse S-wave incident on the Moho from below. In (e), black dot on the red wavefront represents a point slightly below the flat Moho that will contribute to *Sn* for the reference model where the purple arrow intercepts the Moho with incidence angle, i_{flat} close to 90°, suggesting most of the energy is reflected back below the Moho (dashed black curve in f), representing the *Sn* waveguide, while a smaller amount is transmitted into the crust at a smaller angle r_{flat}. The introduction of a Moho ramp reduces these angles to i_{ramp} and r_{ramp}, as shown by the yellow dashed arrows, and sharply increases the amount of energy transmitted into the crust (black curve in f). Because r_{ramp} is smaller than

r_{flat}, the horizontal velocity (apparent velocity) of *Sn* is reduced in the ramp region. Points on the wavefront deeper than the black dot will not interact with the ramp, but will enter the thickened crust beyond the end of the ramp, thus explaining the recovery of *Sn* amplitude at distances further away from the end of the ramp. (g) seismogram at 510 km (10 km beyond the ramp and approximately corresponding to the largest amplitude peak in Fig. 6b) for the reference model (black line) and ramp model (blue line) with *Sn* and *Lg* windows marked by red and cyan lines, showing the large growth of the first-arrival *Sn* wave and the phase delays experienced by the ramp model.

Fig. 8. Change of *Lg* frequency content as a result of *Sn*-to-*Lg* conversion. Same as in the middle column of Fig. 6, but the vertical axis '*Lg* HF/LF' is the ratio of high-frequency (HF, 1–5 Hz) *Lg* to low-frequency (LF, 0.1–0.8 Hz) *Lg* of the ramp model divided by the equivalent ratio for the reference model. For all panels, the horizontal axis is linear while the vertical axis is in log scale.

Fig. 9. *Sn*, *Lg* amplitudes and *Sn/Lg* with varying ramp width *w* but fixed ramp distance *d* = 300 km. Figure organization as for Fig. 6, except for the left and middle columns the vertical colored bars represent the start of the ramps for the correspondingly colored symbols. The end of the ramp is aligned for all of these cases at 0 km and marked by a solid black line, as in Fig. 6. For the right column, the beginning of the ramp is marked by a dashed black line and the end of the ramp is marked by a solid colored line for the corresponding colored symbol. For example, the ramp region for the red circles is within the dashed black line and the solid red line. For all panels, the vertical axis is in log scale. Horizontal axes are linear for parts (a)–(f), and log scale for (g), (h), (i).

Fig. 10. *Sn*, *Lg* amplitudes and *Sn/Lg* as a function of south-north distance from ramp end. Southern events are shown with colored symbols and northern events with black symbols. Events interpreted as mantle earthquakes are shown with solid symbols and crustal events with open symbols. The beginning of the ramp, perpendicular to the array, is shown with a dashed black line and the end of the ramp is shown with a solid black line. The southern limit of the *Sn* attenuation zone is marked with a magenta dashed line. (a)-(d) Individual *Sn* and *Lg* amplitudes for the southern and northern events, respectively. Data points are aligned vertically for each individual station location. (e) Comparison of *Sn/Lg* for the northern and southern events. This very different combined group of events can be clearly separated by high and low *Sn/Lg*, especially >~100 km beyond the end of the ramp (Supplementary materials S7). For all panels, the horizontal axis is linear while the vertical axis is in log scale.

Fig. 11. Record sections of the southern events. Top four events (S2–S5) are interpreted here as below-Moho earthquakes, and bottom two events (S1, S6) as crustal events. For each event, an upper panel shows *Sn/Lg* measured at each station, on a linear scale from 0-8, with our

arbitrary threshold *Sn/Lg* = 2 shown as a grey line. The event code along with the percentage of stations that registered a *Sn/Lg* > 2 are labelled in the upper panel. For each event, between the two panels is each station's distance north of the YZS, recognized as the end of the Moho ramp. The bottom panels show trace-normalized amplitudes for each event, with *Sn* window colored red, *Lg* window cyan, and the noise window green. The traces are shown with a reduction velocity of 4 km/s. Traces are displayed south to north with epicentral distance shown beneath each record section. Thick yellow line marks stations south of YZS (i.e. within the ramp), and thick magenta line marks stations within the *Sn* attenuation zone.

Fig. 12. High-frequency (HF) to low-frequency (LF) ratio of *Lg* waves in data. Moho ramp beginning is marked by a black dashed line while its ending is marked by a black solid line. The start of the *Sn* attenuation zone is shown by a magenta dashed line. The horizontal axis are distances aligned at the end of the Moho ramp. The vertical axis shows the *Lg* HF/LF in a log scale. (a) southern events. (b) northern events. Symbol styles are as in Fig. 10 (open symbols: crustal events; closed symbols: mantle earthquakes). (c) events recorded on the Gangdese-92 array (not categorized as crustal or mantle because we lack comparison events). Note different vertical scale compared with (a)&(b).

Fig. 13. High-frequency (HF) to low-frequency (LF) ratio of *Lg* waves in reference and select ramp models. (a) *Lg* HF/LF for five source depths with the reference model with a flat Moho at 30 km. Two of the sources are located within the crust, one in the shallow-lid, and two deeper within the mantle. (b) & (c) *Lg* HF/LF for ramp models with fixed w=100 km, testing the effect of increasing d. (d)&(e) *Lg* HF/LF for ramp models with fixed d=100 km, testing the effect of increasing w. All panels are log-log. Dashed and solid black lines indicate the start and end of the ramp, when located beyond 200 km. A fiducial line at *Lg* HF/LF = 0.7 is drawn to emphasize the separation of crustal and mantle events for all panels.

Tables

Code name	Date, time	Location (N°, E°)	Magnitude (m_b)	Catalog depth (km)	Distance to ramp, d (km)	Effective ramp width, w (km)
S1	2005-07-26, 18:27:05	23.27, 91.41 *23.281, 91.516*	4.0	38.1$\pm$ 27.4 *10*	560-660	200-280
S2	2005-05-03, 00:38:57	25.76, 91.06 *26.078, 91.033*	4.3	33.6$\pm$? *33*	285-385	210-280
S3 (H82)	2004-08-04, 02:09:21	25.92, 90.26 *25.865, 90.333*	4.2 (4.1, M_w)	61.7$\pm$ 10.8 *20* (53$\pm$?)	250-330	180-260
S4	2005-05-27, 22:12:20	26.14, 87.21 *26.170, 87.685*	3.5	57.7$\pm$ 12 *15*	270-280	130-160
S5	2004-11-24, 22:35:42	27.33, 90.94 *27.337, 90.875*	4.0	10$\pm$? *10*	100-140	160-380
S6	2004-08-09, 08:18:18	27.58, 91.80 *27.547, 91.718*	4.1	16.1$\pm$? *14.9*	50-120	195-475
WT1	2005-05-19, 05:43:30	35.63, 78.38	4.2	97.6 $\pm$ 14.1	-	-
WT2	2005-06-20, 22:52:26	36.23, 77.92	3.9	77.9 $\pm$ 8.4	-	-
WT3	2005-03-03, 15:07:39	35.65, 77.85	3.7	57.5 $\pm$ 16.8	-	-
04-251	2004-09-07, 04:01:05	35.72, 78.25	4.2	7.6 $\pm$ 26.8	-	-
04-291	2004-10-17, 15:35:45	35.20, 77.67	4.3	15$\pm$?	-	-
05-201	2005-07-20, 10:54:49	35.34, 77.79	4.2	10$\pm$?	-	-

Table 1

Figures

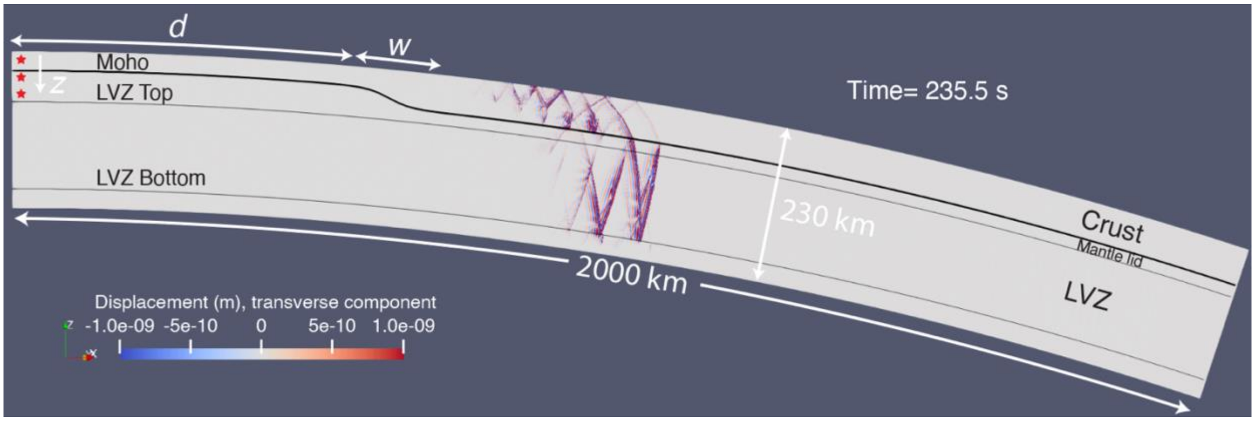


Fig. 1

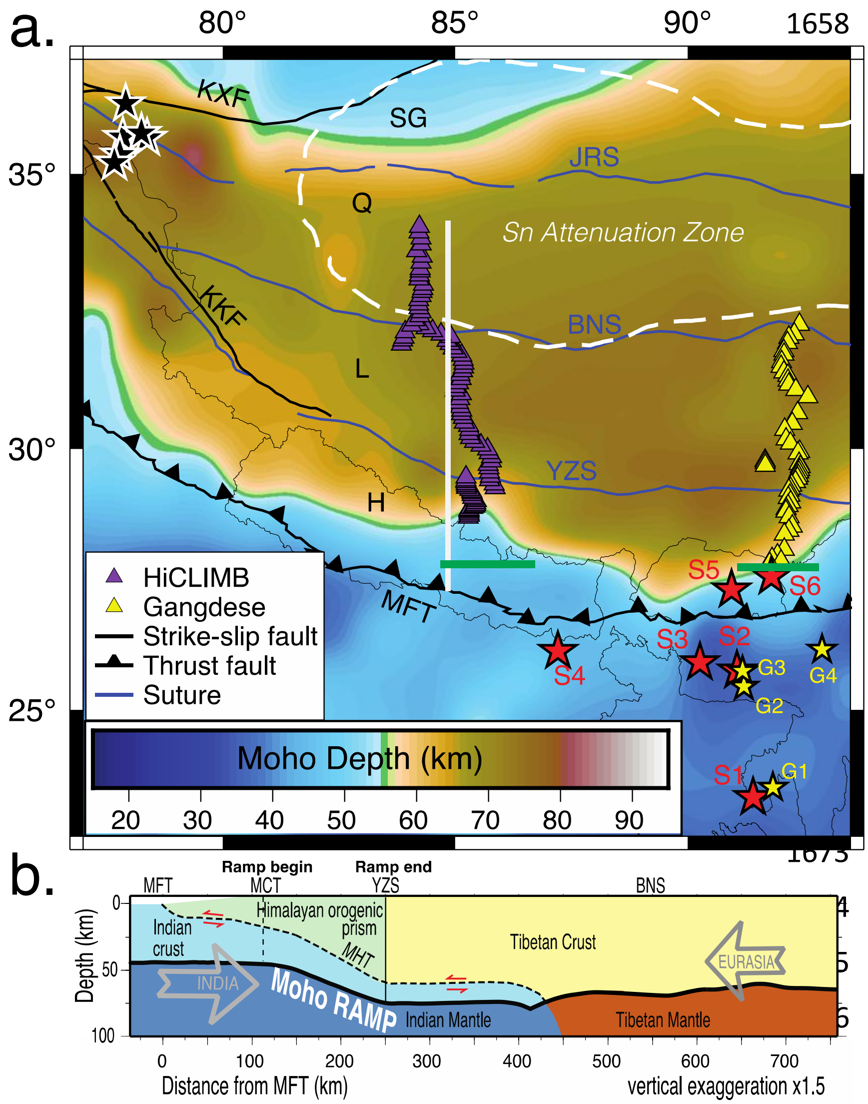
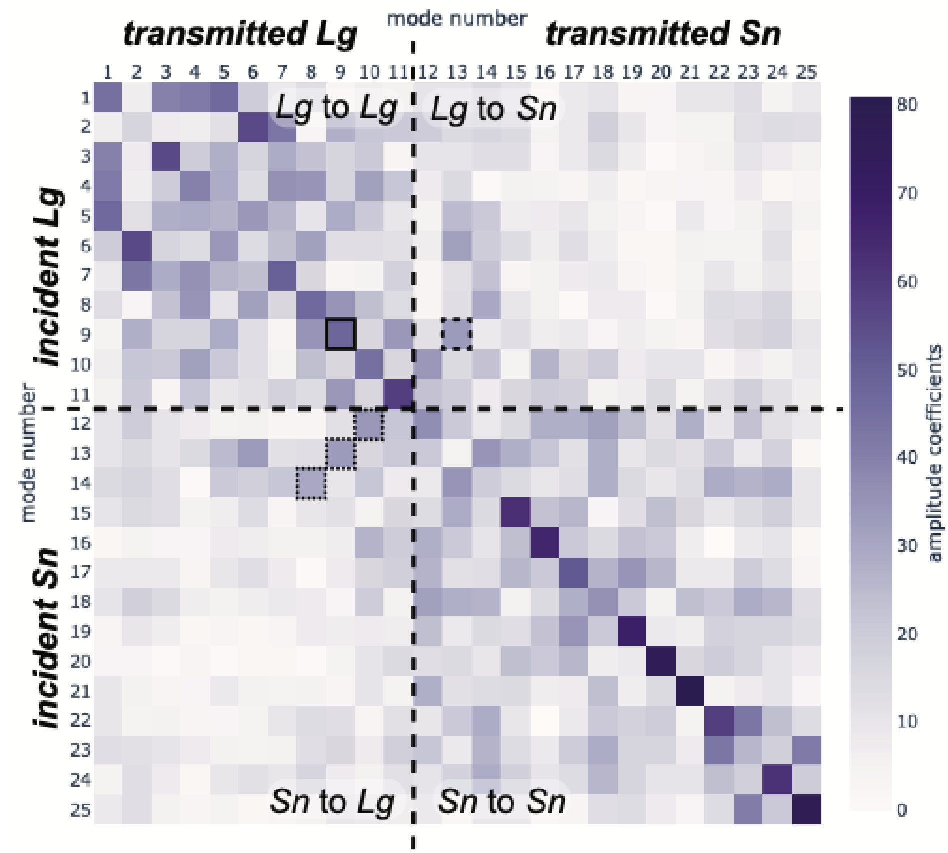

a.
80°
85°
90°
35°
30°
25°
KXF
SG
JRS
Q
Sn Attenuation Zone
KKF
BNS
L
YZS
H
MFT
S1
S2
S3
S4
S5
S6
G1
G2
G3
G4
HiCLIMB
Gangdese
Strike-slip fault
Thrust fault
Suture
Moho Depth (km)
20 30 40 50 60 70 80 90
b.
MFT
Ramp begin
MCT
Ramp end
YZS
BNS
Himalayan orogenic prism
Indian crust
MHT
Tibetan Crust
INDIA
EURASIA
Moho RAMP
Indian Mantle
Tibetan Mantle
Depth (km)
0
50
100
0 100 200 300 400 500 600 700
Distance from MFT (km)
vertical exaggeration x1.5

Fig. 2



transmitted Lg
mode number
transmitted Sn
Lg to Lg
Lg to Sn
incident Lg
mode number
incident Sn
Sn to Lg
Sn to Sn
amplitude coefficients

Fig. 3

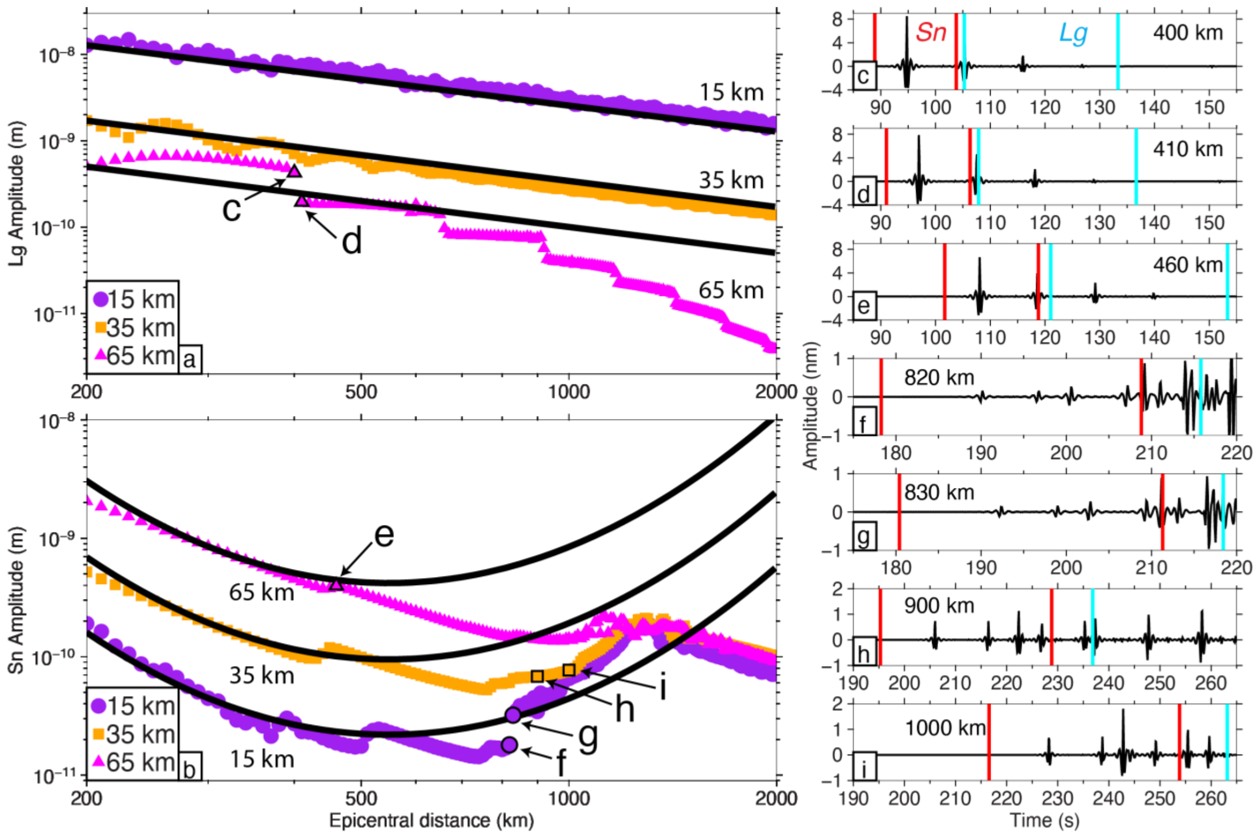
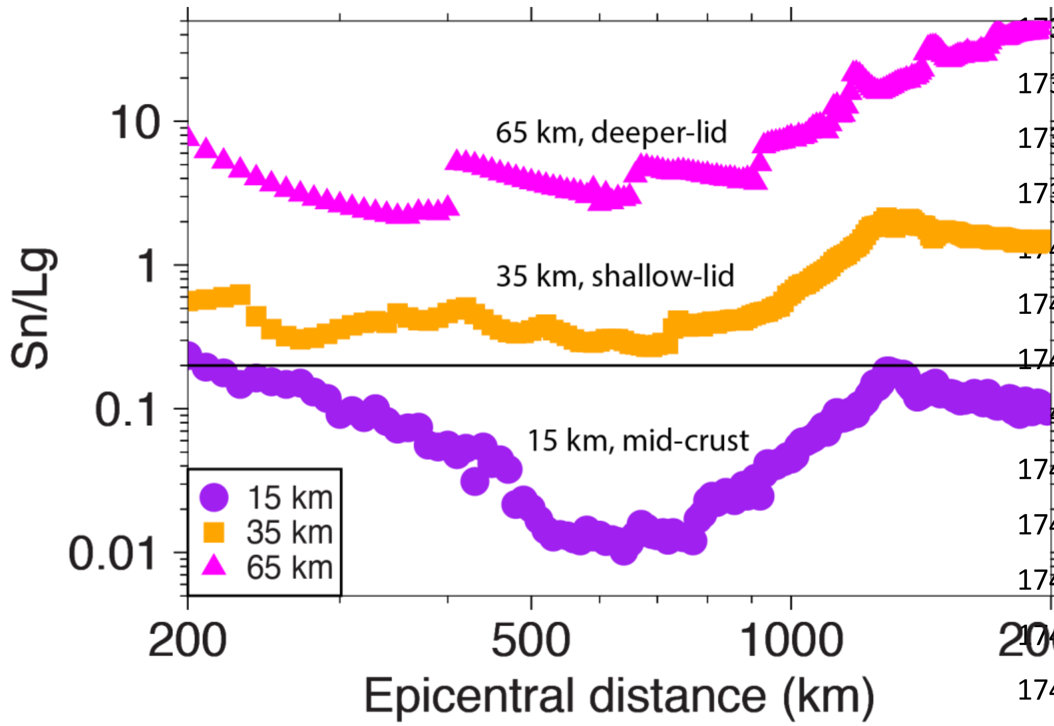
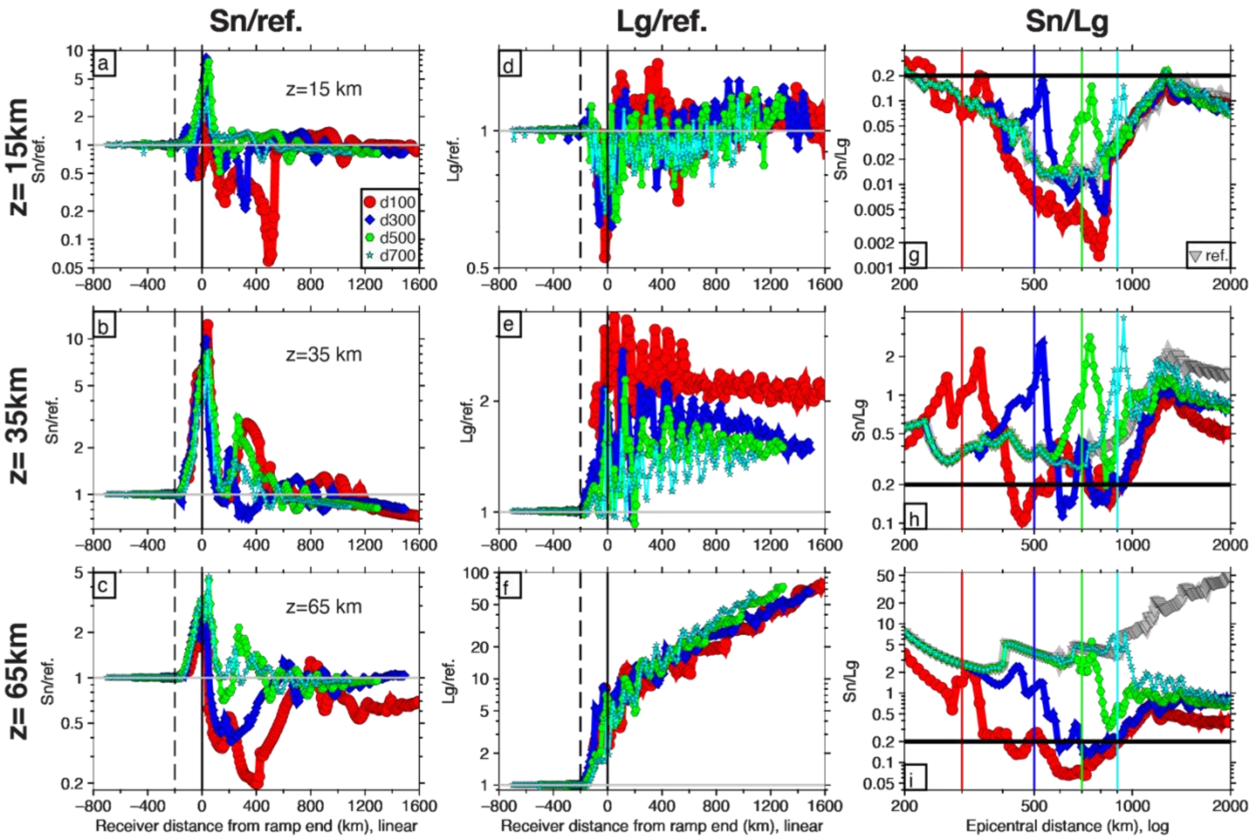

Lg Amplitude (m)
15 km
35 km
65 km
c
d
a
Sn Amplitude (m)
e
f
g
h
i
b
Epicentral distance (km)
Sn
Lg
400 km
410 km
460 km
820 km
830 km
900 km
1000 km
Amplitude (nm)
Time (s)

Fig. 4



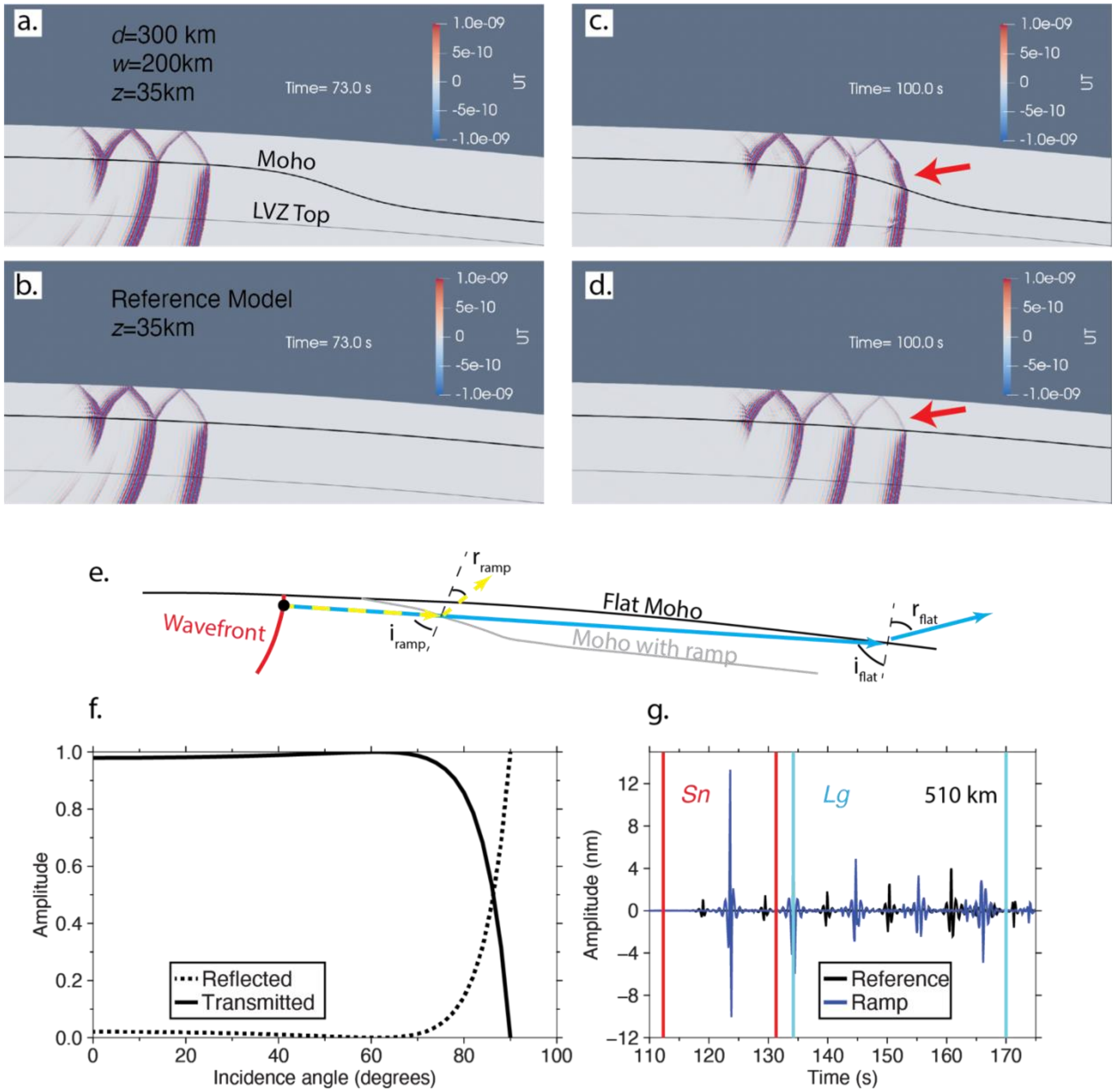
Sn/Lg
10
1
0.1
0.01
65 km, deeper-lid
35 km, shallow-lid
15 km, mid-crust
15 km
35 km
65 km
200
500
1000
2000
Epicentral distance (km)

Fig. 5



Sn/ref.
Lg/ref.
Sn/Lg
z= 15km
z= 35km
z= 65km
z=15 km
z=35 km
z=65 km
d100
d300
d500
d700
ref.
Receiver distance from ramp end (km), linear
Receiver distance from ramp end (km), linear
Epicentral distance (km), log

Fig. 6



a.
d=300 km
w=200km
z=35km
Time= 73.0 s
Moho
LVZ Top
b.
Reference Model
z=35km
Time= 73.0 s
c.
Time= 100.0 s
d.
Time= 100.0 s
1.0e-09
5e-10
0
-5e-10
-1.0e-09
e.
Wavefront
Flat Moho
Moho with ramp
r_{ramp}
i_{ramp}
r_{flat}
i_{flat}
f.
Amplitude
Incidence angle (degrees)
Reflected
Transmitted
g.
Sn
Lg
510 km
Amplitude (nm)
Time (s)
Reference
Ramp

Fig. 7

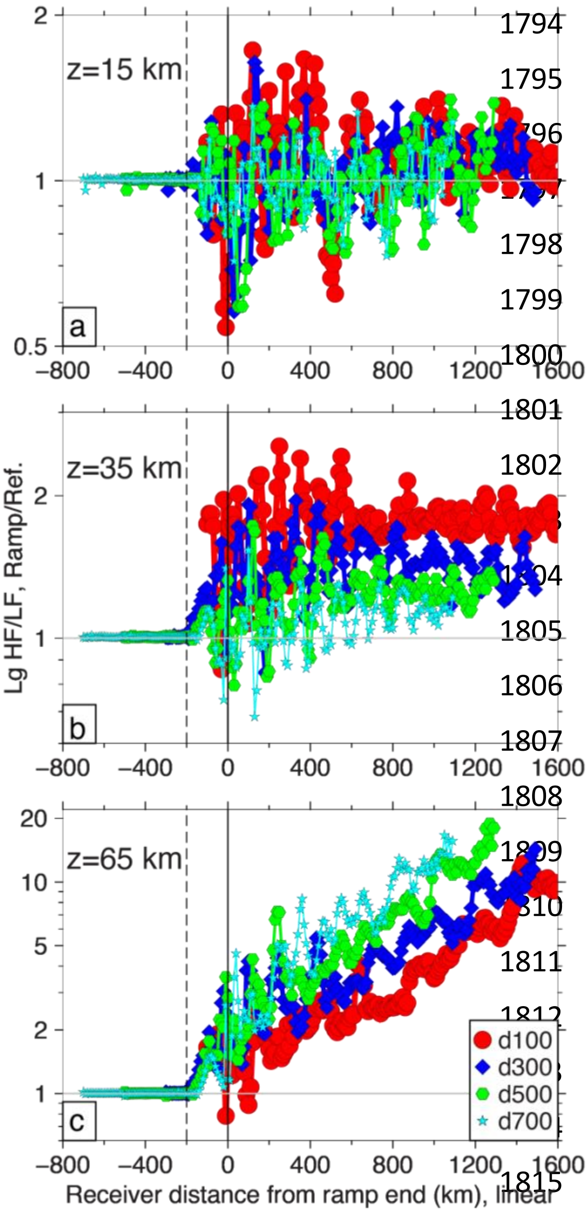

z=15 km
a
z=35 km
b
z=65 km
c
Lg HF/LF, Ramp/Ref.
Receiver distance from ramp end (km), linear
d100
d300
d500
d700

Fig. 8

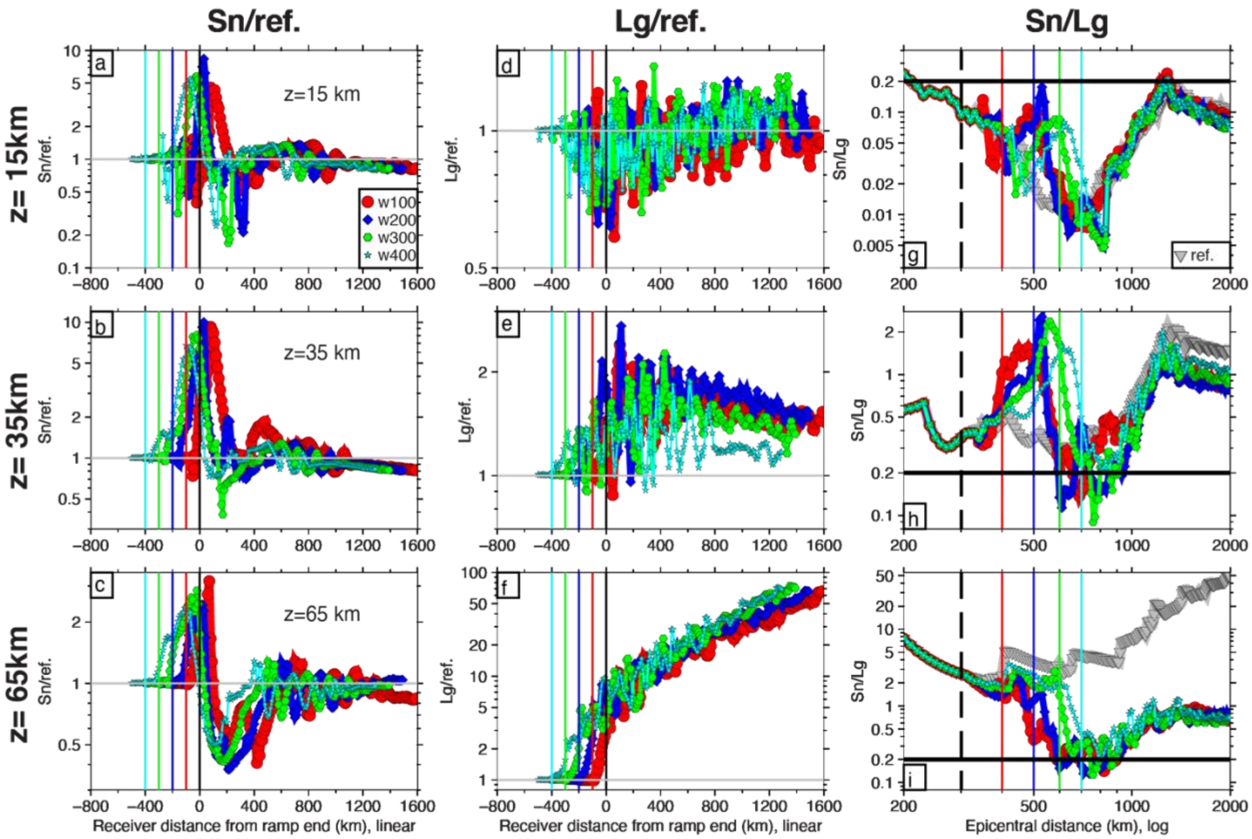

Sn/ref.
Lg/ref.
Sn/Lg
z= 15km
z= 35km
z= 65km
z=15 km
z=35 km
z=65 km
a
b
c
d
e
f
g
h
i
w100
w200
w300
w400
ref.
Sn/ref.
Lg/ref.
Sn/Lg
Receiver distance from ramp end (km), linear
Epicentral distance (km), log

Fig. 9

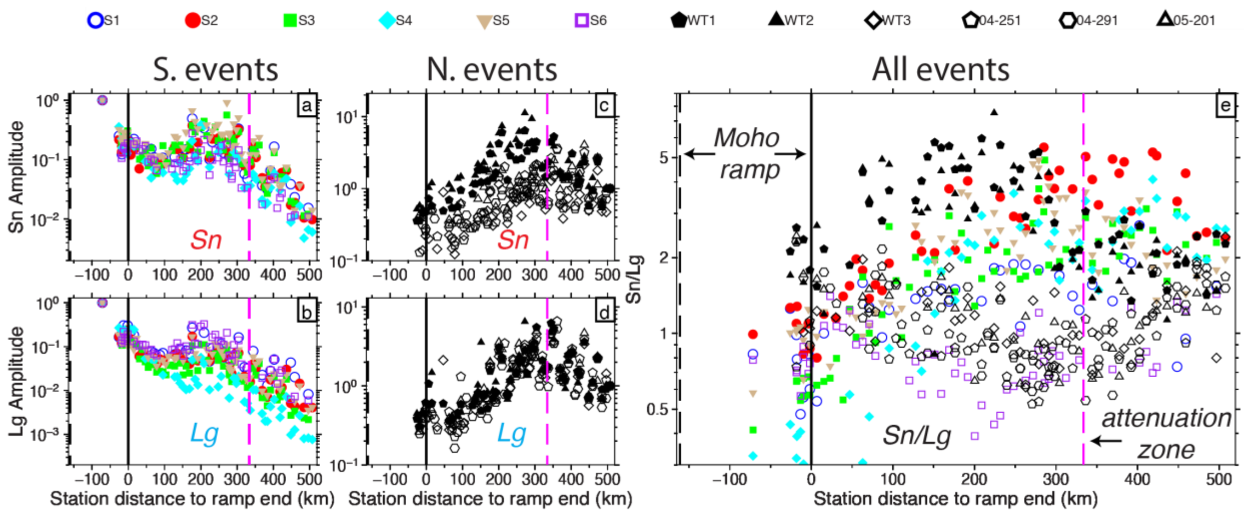

S1
S2
S3
S4
S5
S6
WT1
WT2
WT3
04-251
04-291
05-201
S. events
N. events
All events
Sn Amplitude
Lg Amplitude
Sn
Lg
Sn/Lg
Moho ramp
attenuation zone
Station distance to ramp end (km)
a
b
c
d
e

Fig. 10

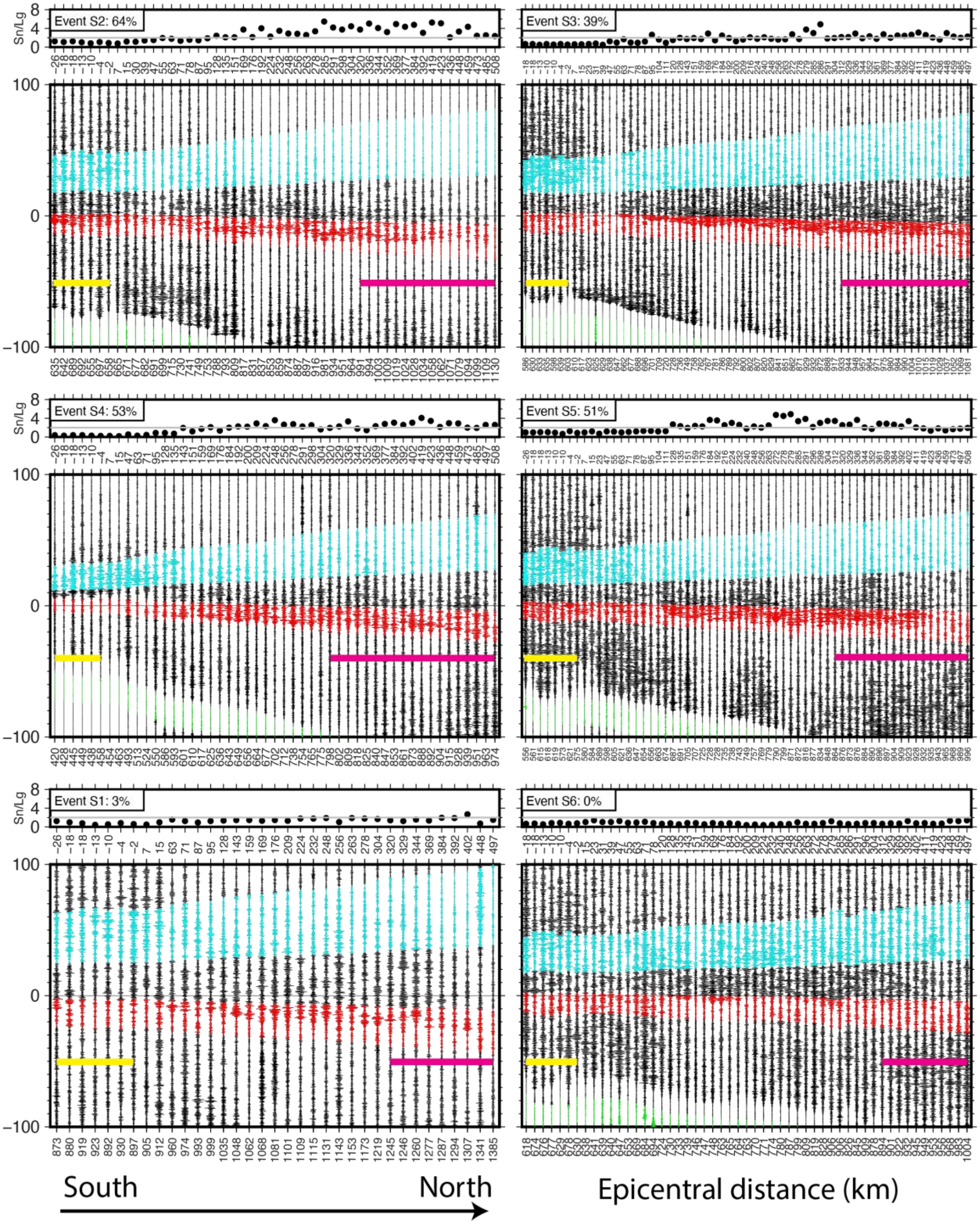
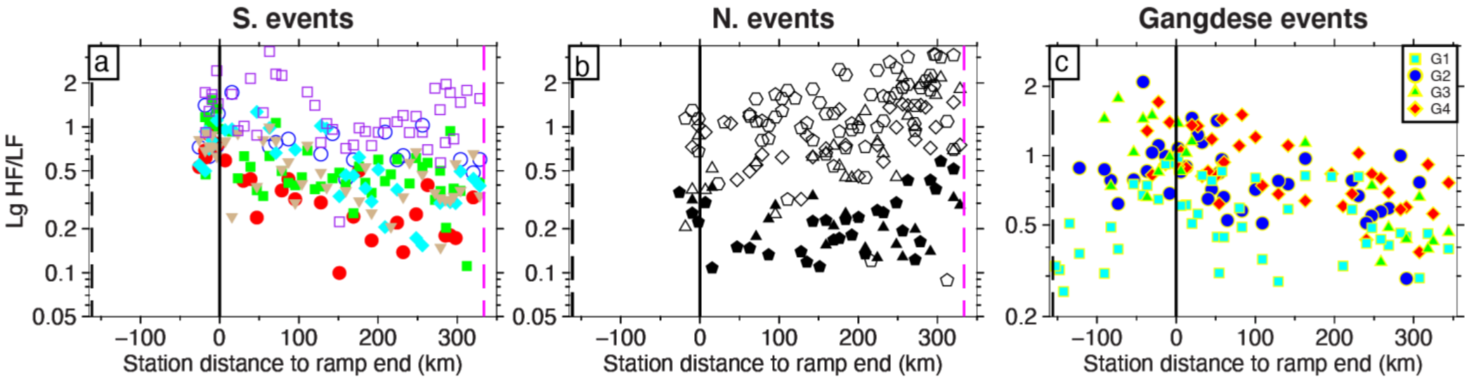

Event S2: 64%
Event S3: 39%
Event S4: 53%
Event S5: 51%
Event S1: 3%
Event S6: 0%
Sn/Lg
South
North
Epicentral distance (km)

Fig. 11



S. events
N. events
Gangdese events
a
b
c
G1
G2
G3
G4
Lg HF/LF
Station distance to ramp end (km)
Station distance to ramp end (km)
Station distance to ramp end (km)

Fig. 12

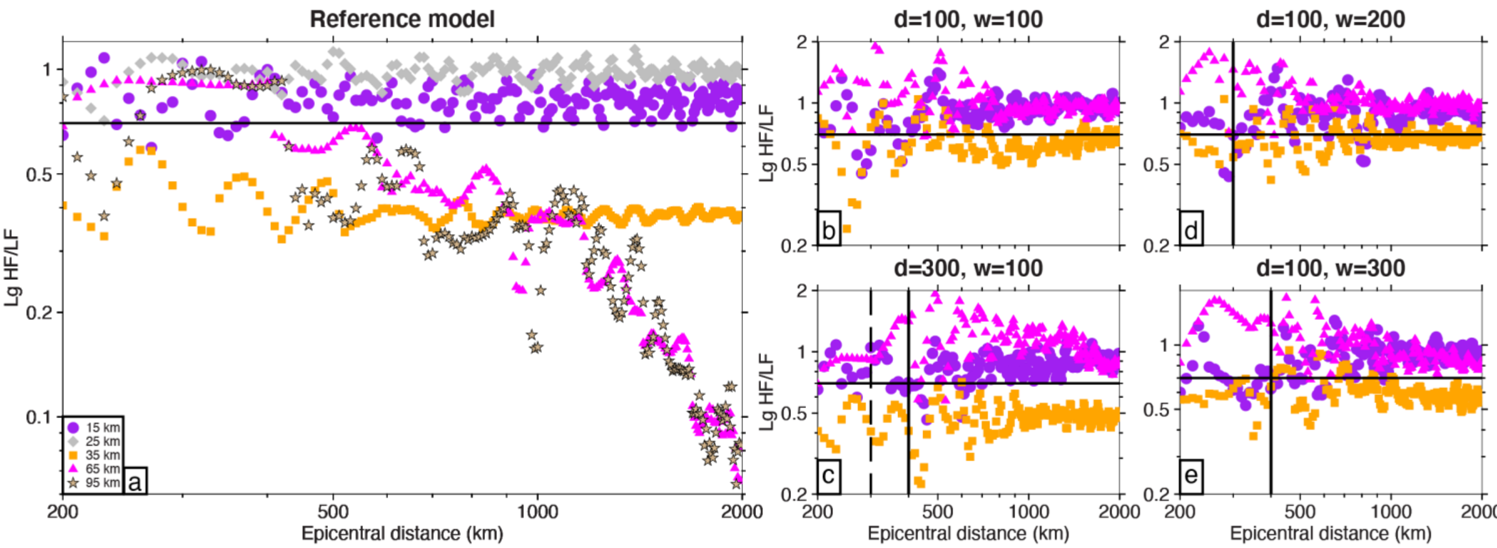

Reference model
d=100, w=100
d=100, w=200
d=300, w=100
d=100, w=300
Lg HF/LF
Epicentral distance (km)
15 km
25 km
35 km
65 km
95 km
a
b
c
d
e

Fig. 13

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).

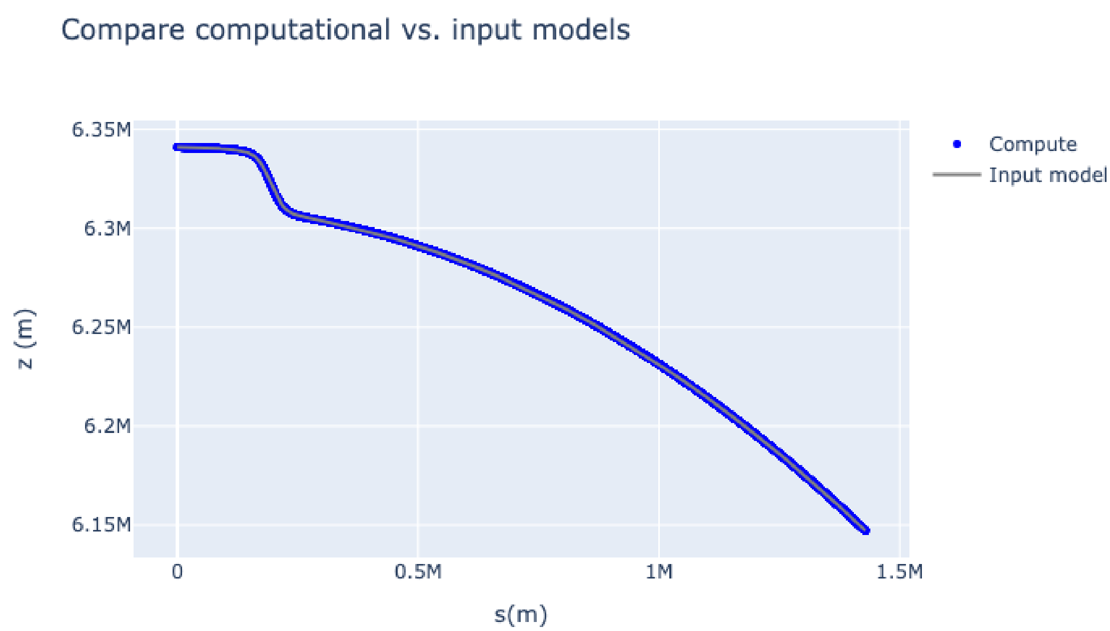


Fig. S1-1. Checking representation of undulation by AxiSEM3D. The input geometrical model is shown as a grey line, and its representation by GLL points used for computation is shown as blue dots. The match is exact. Vertical and horizontal axes are the z, vertical, and s, radial, cylindrical coordinates 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).

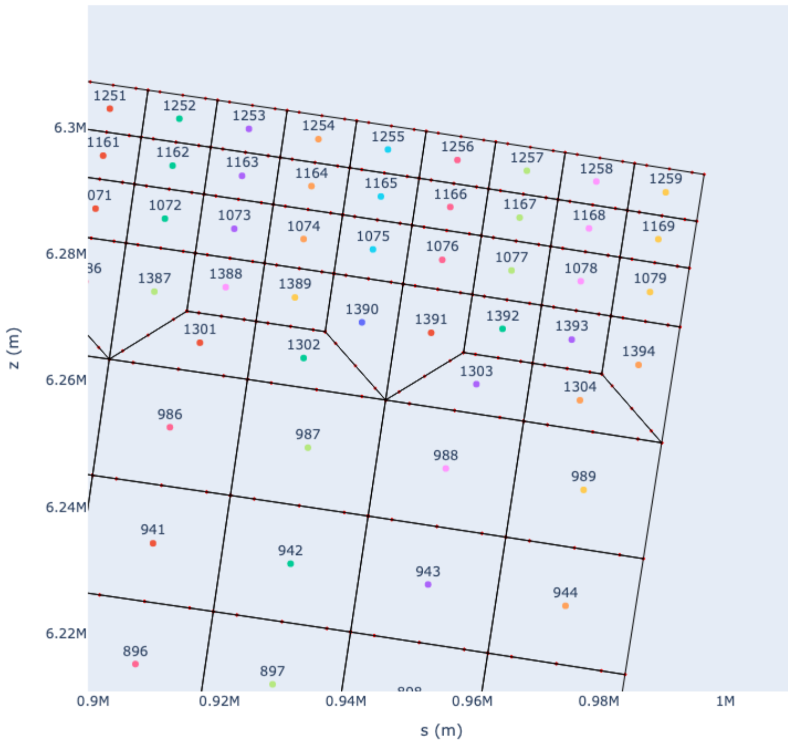


Fig. S2-1. Two-refinement transition templates. Example showing a low-resolution mesh (not used in our simulations) at an offset to demonstrate the way a mesh can be coarsened with SalvusMeshLite. The horizontal and vertical axes are the same as in Fig. S1-1. Numbers on each element are its tag number. 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).

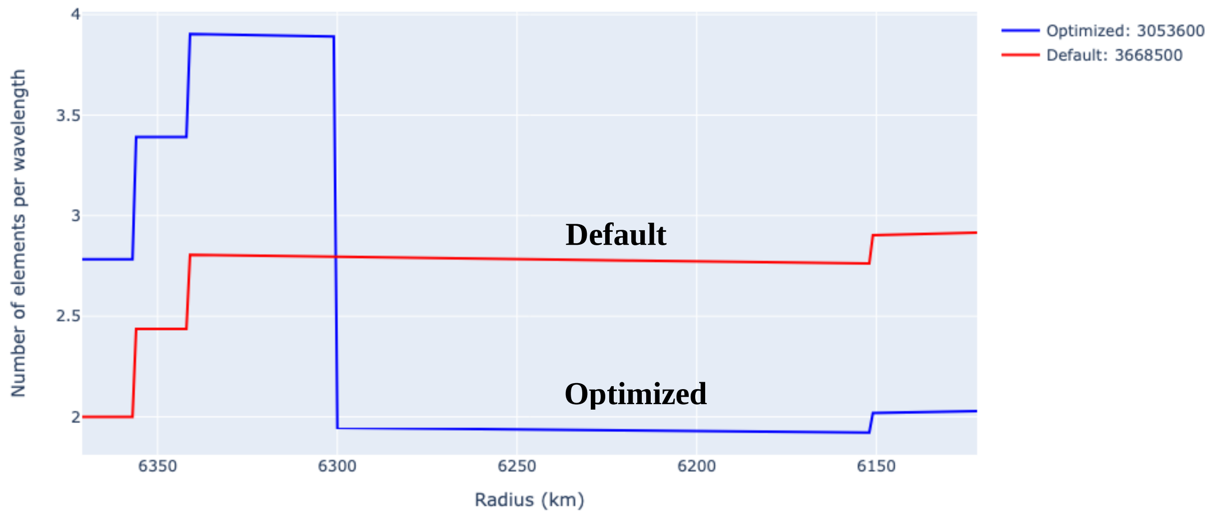


Fig. S2-2. Number of elements per wavelength for the coarsened and uncoarsened meshes. Two elements per wavelength is maintained by the default, uncoarsened mesh (red line), near the surface of the earth, with radius close to 6371 km, whereas this is honored by the optimized, coarsened mesh (blue line), 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%.

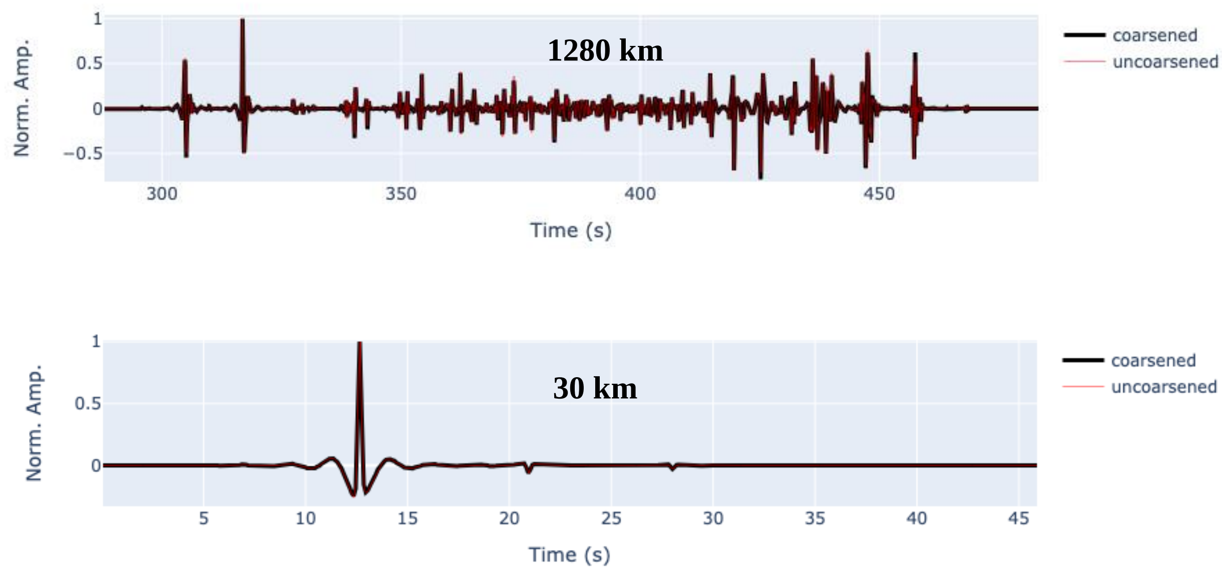


Fig. S2-3. Representative seismograms computed with the coarsened and uncoarsened meshes. Top: at epicentral distance 1,280 km. Bottom: at epicentral distance 30 km. The seismograms shown are the transverse component and are computed for the model with a Moho ramp with d =100 and w= 200 km. Horizontal 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*)

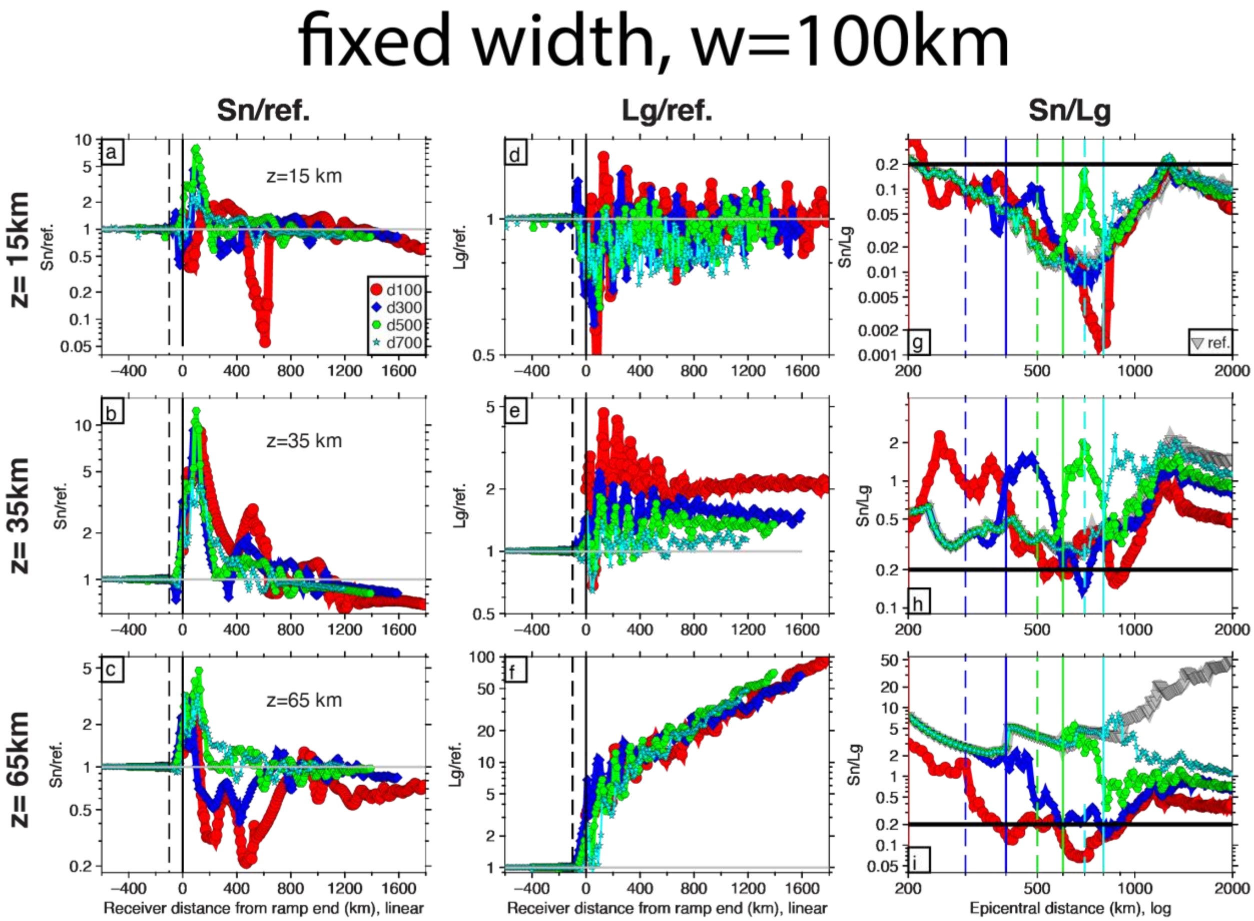


Fig. S4-1. *Sn*, *Lg* amplitudes and *Sn/Lg* with varying distance to ramp start *d* but fixed width *w* = 100 km. The layout is the same as Fig. 6 except for right columns, where colored dashed lines represent the start of a ramp and colored solid lines represent the end of a ramp, for the correspondingly 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.

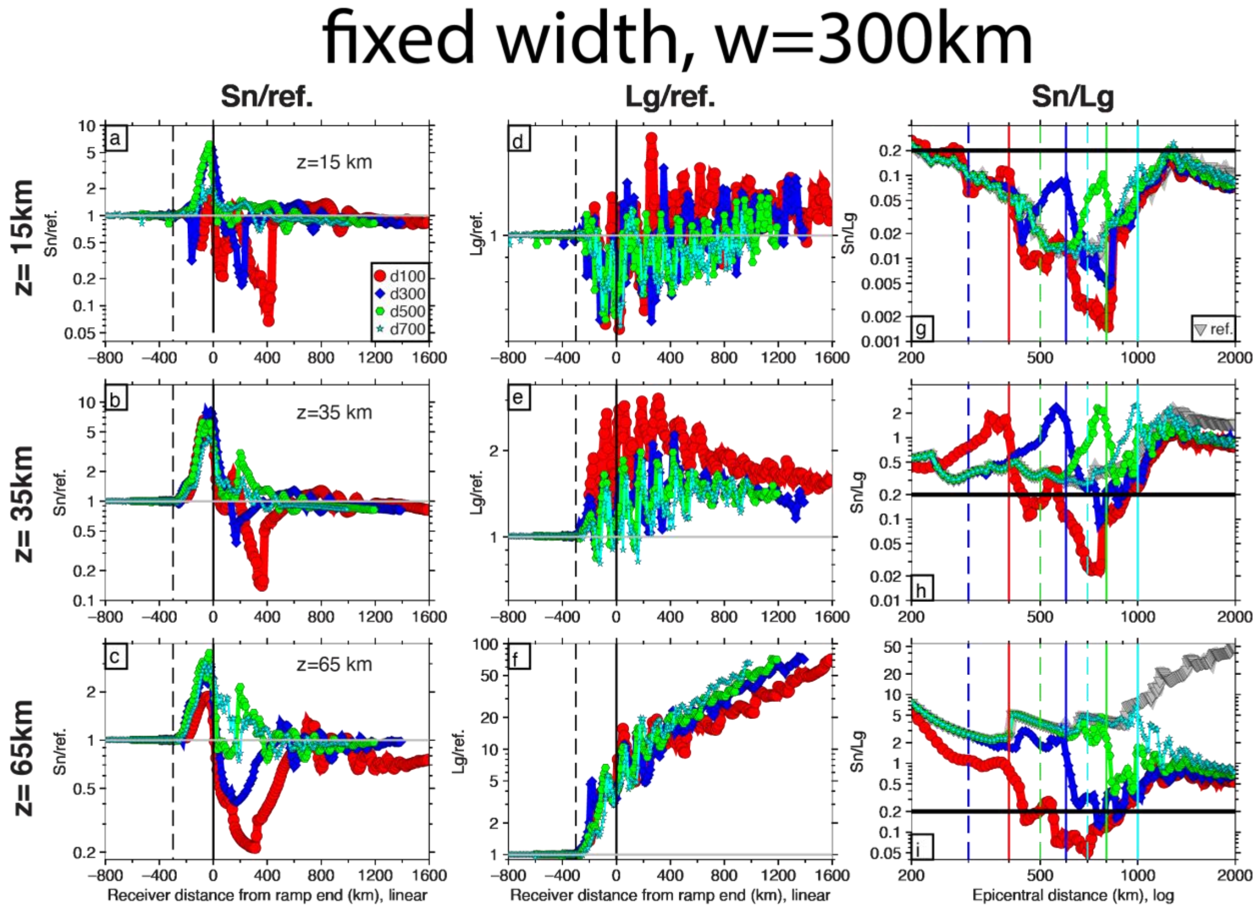


Fig. S4-2. *Sn, Lg* amplitudes and *Sn/Lg* with varying distance to ramp start *d* but fixed width *w* = 300 km. The layout is the same as Fig. S4-1.

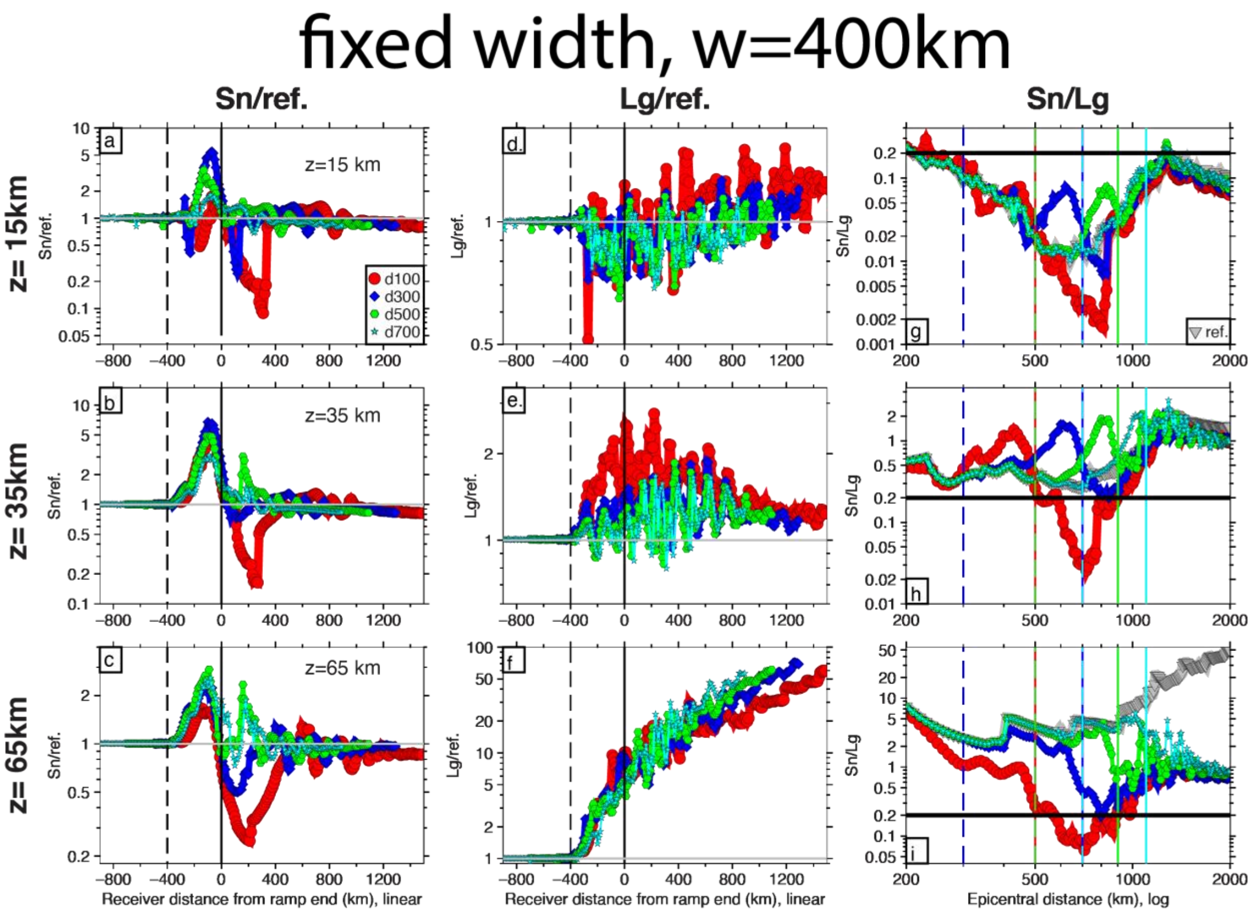


Fig. S4-3. *Sn*, *Lg* amplitudes and *Sn/Lg* with varying distance to ramp start *d* but fixed width *w* = 400 km. The layout is the same as Fig. S4-1.

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

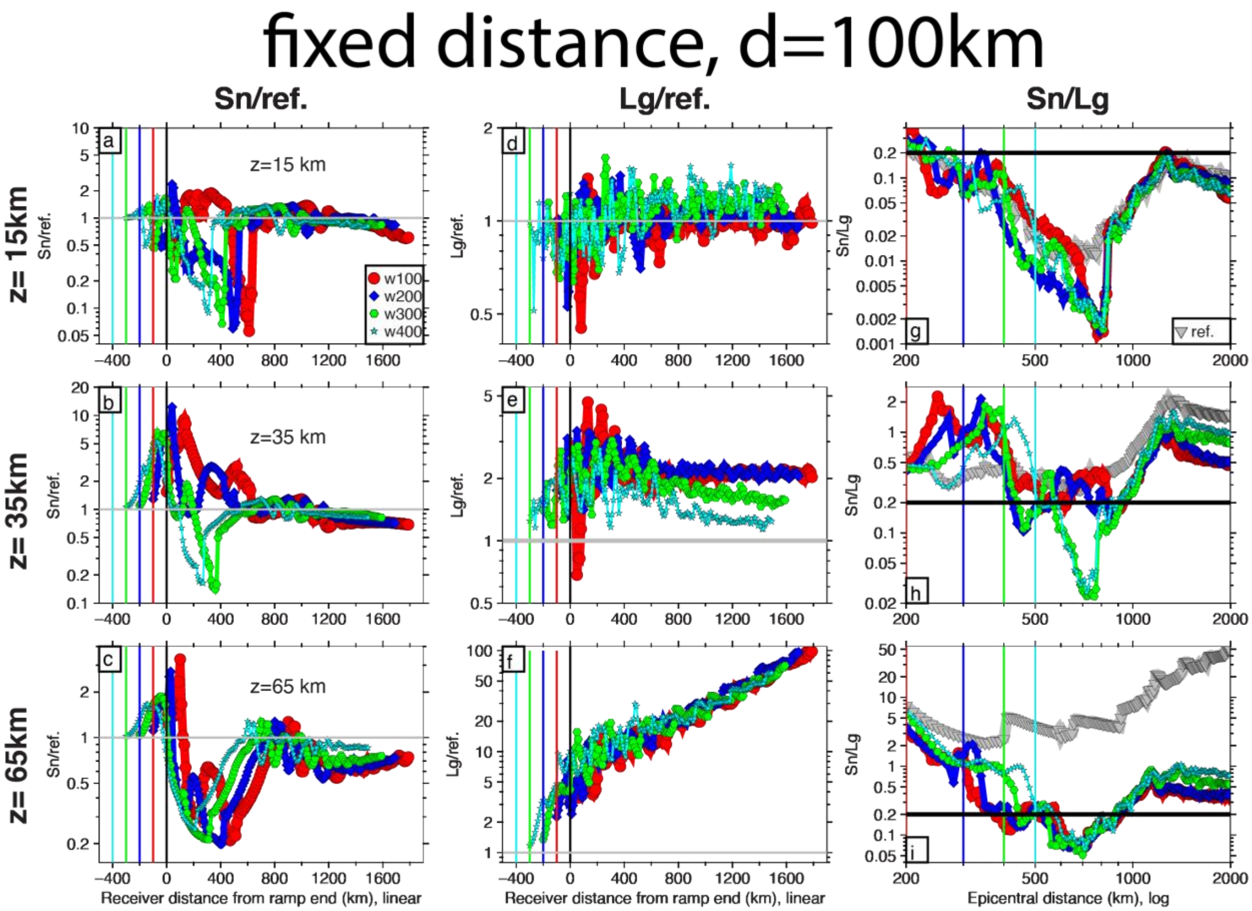


Fig. S5-1. ***Sn*, *Lg* amplitudes and *Sn/Lg* with varying ramp width *w* but fixed ramp distance *d* = 100 km.** The layout is the same as Fig. 9, except the start of the ramp is not shown the right column.

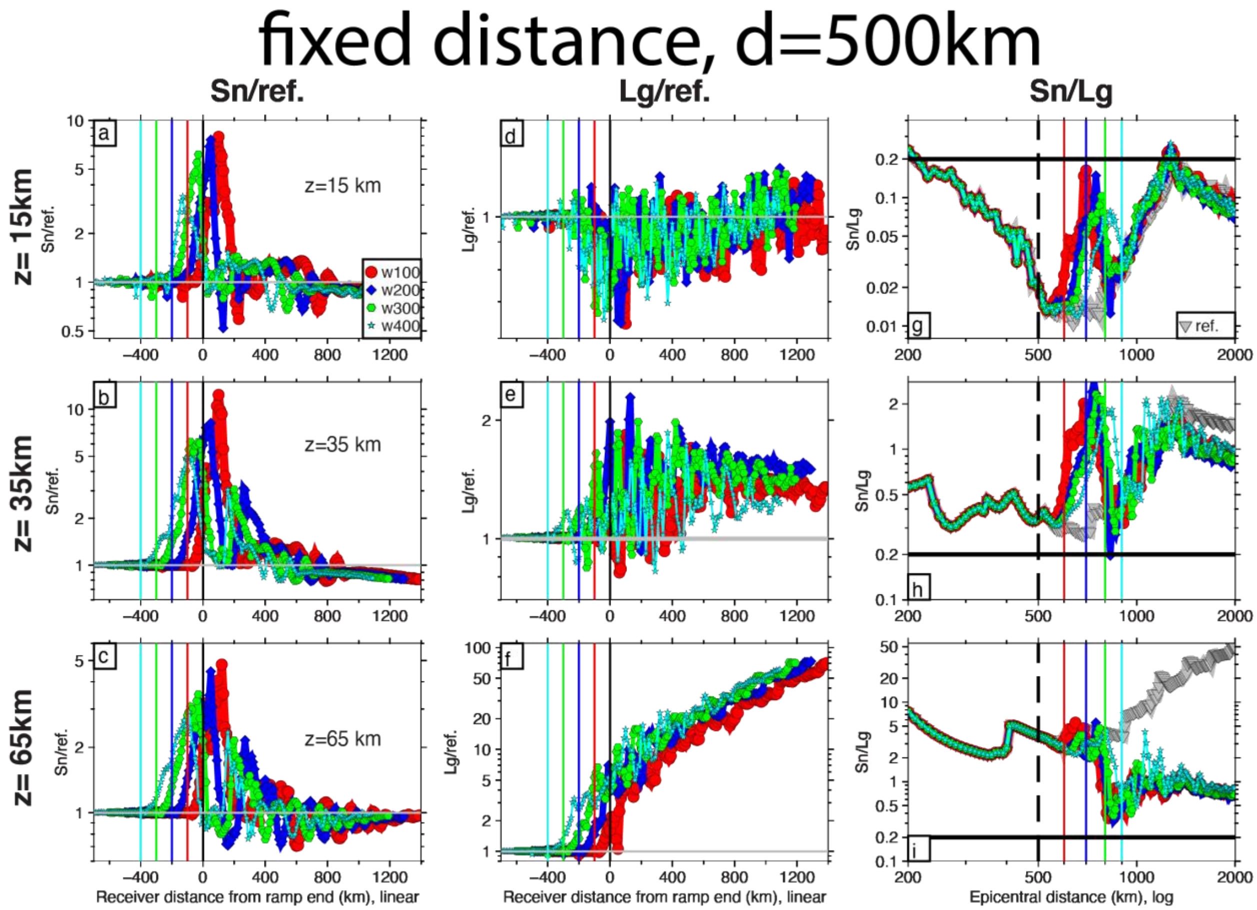


Fig. S5-2. ***Sn*, *Lg* amplitudes and *Sn/Lg* with varying ramp width *w* but fixed ramp distance *d* = 500 km.** The layout is the same as Fig. 9.

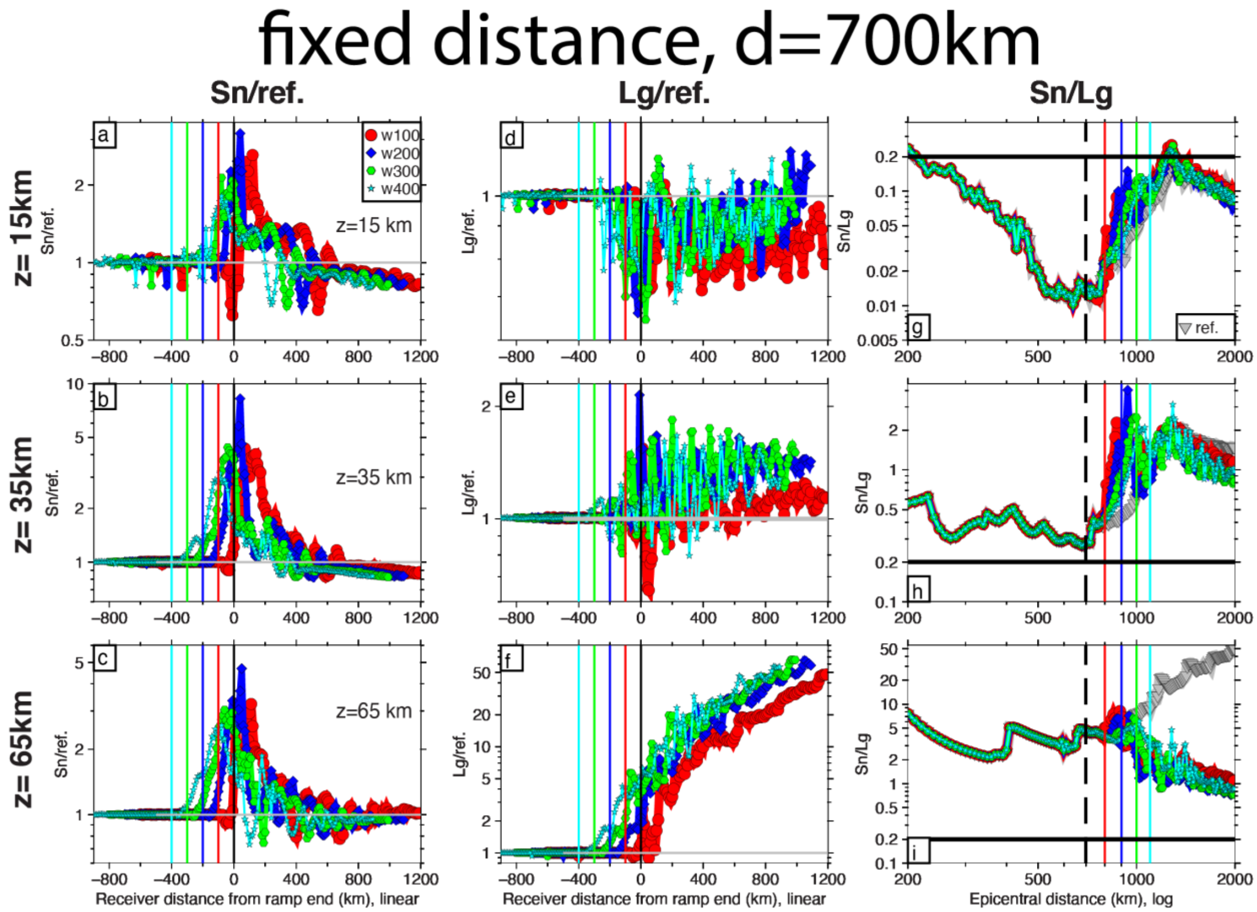


Fig. S5-3. ***Sn*, *Lg* amplitudes and *Sn/Lg* with varying ramp width *w* but fixed ramp distance *d* = 700 km.** The layout is the same as Fig. 9.

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.

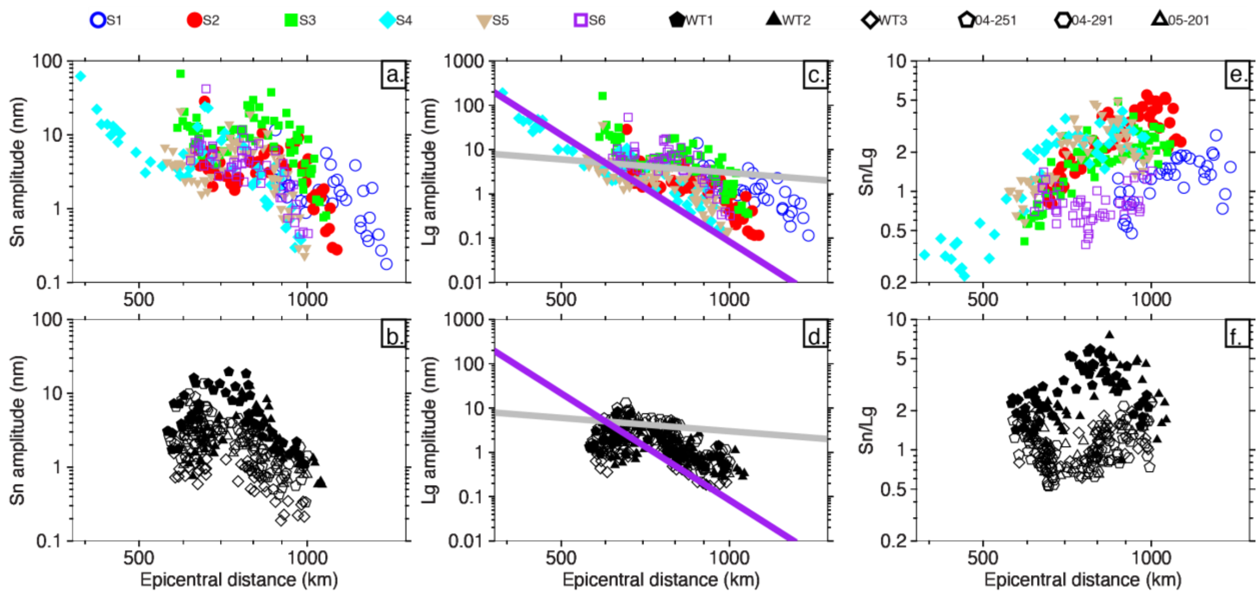


Fig. S6-1. HiCLIMB data plotted against epicentral distance (log-log plots). All symbols are the same as in the main paper for HiCLIMB events. Amplitudes are absolute without any normalization. Data are filtered from 1–5 Hz. (a)&(b) *Sn* amplitudes for northern and southern events, respectively. (c)&(d) *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.

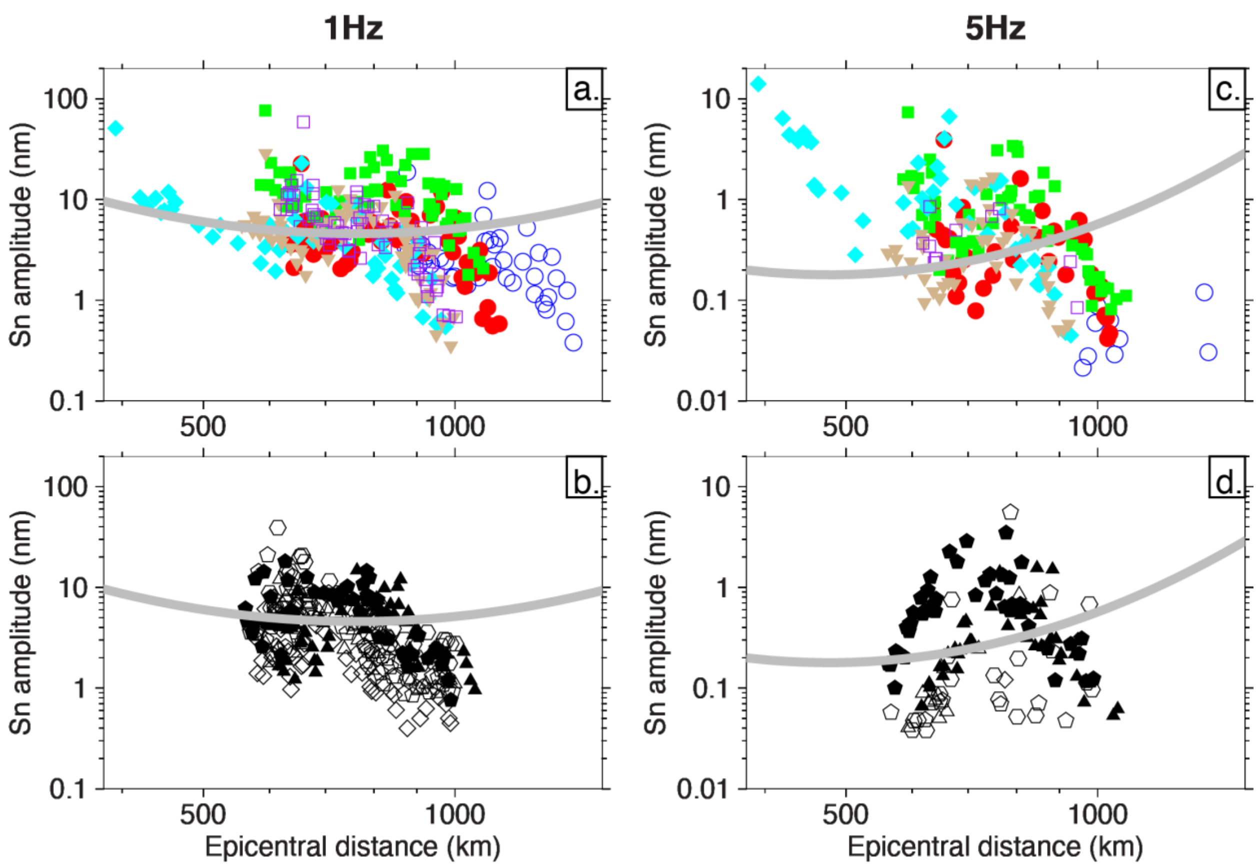


Fig. S6-2. *Sn* filtered to two frequencies. All symbols are the same as in the main paper. Amplitudes are absolute without any normalization. (a)&(b) *Sn* at 1 Hz for northern and southern events, respectively. (c)&(d). *Sn* at 5 Hz for northern and southern events, respectively. Corresponding geometrical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

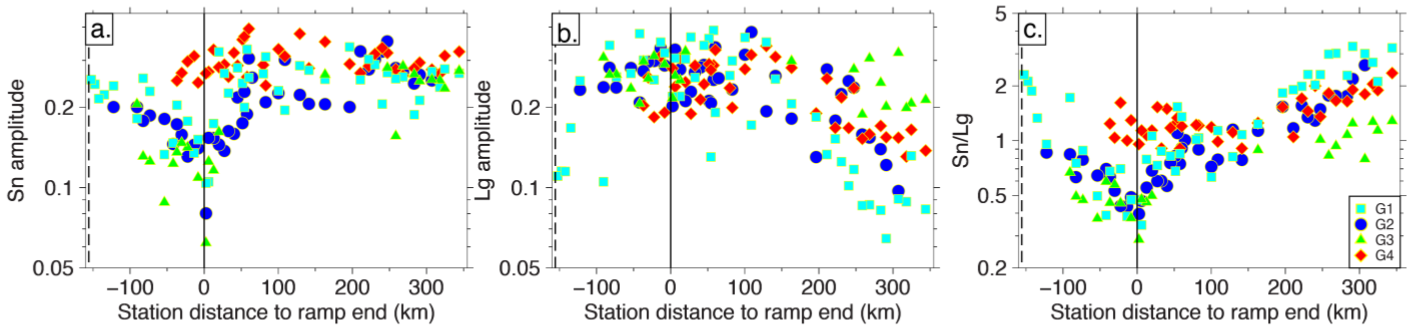


Fig. S7-1. *Sn, Lg* amplitudes and *Sn/Lg* as a function of distance to ramp end for Gangdese-92 events. All annotations are as in Fig. 10. This array did not extend into the *Sn* attenuation region.

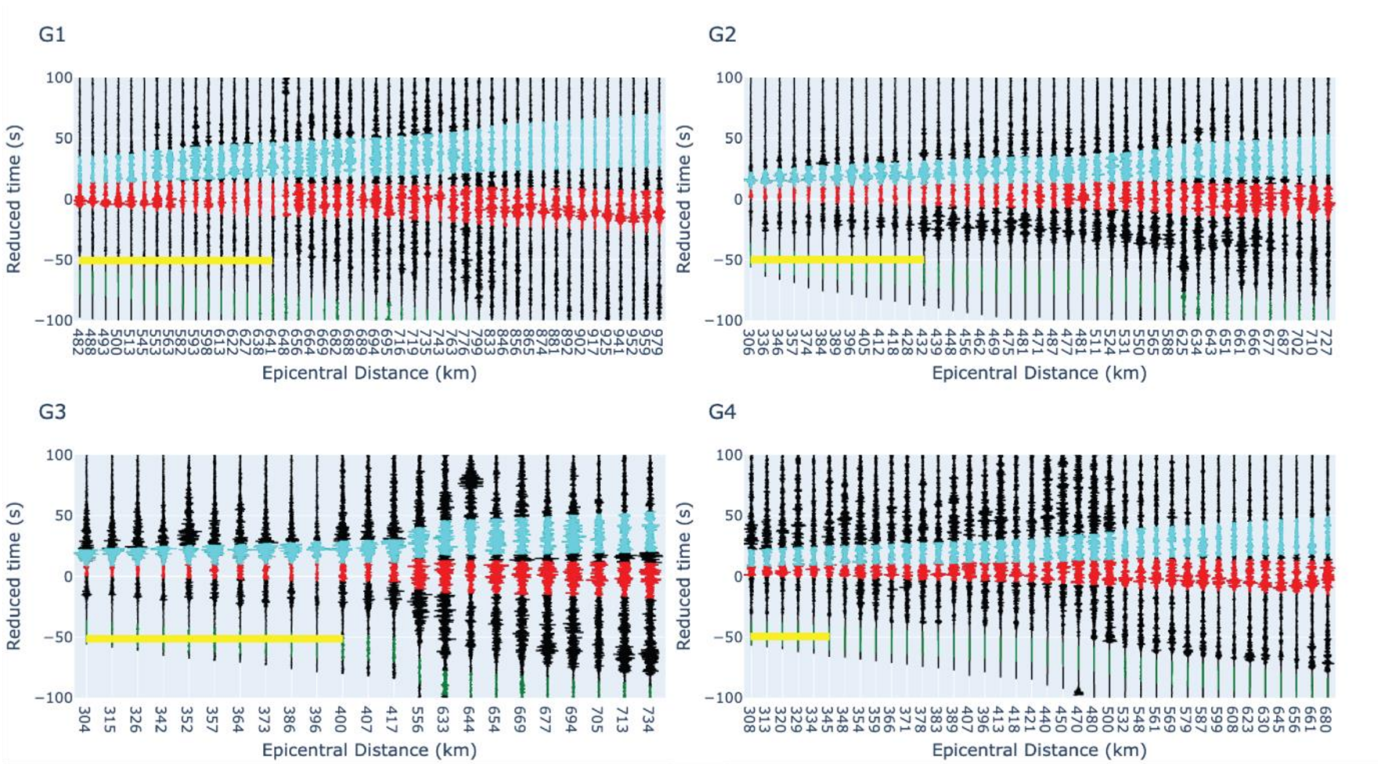


Fig. S7-2. Record sections for Gangdese-92 events. Record sections are shown with reduction velocity of 4 km/s. Green section: noise window. Red section: *Sn* window. Cyan section: *Lg* window. Horizontal yellow bar indicates stations south of YZS, i.e. within the Moho ramp region.

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*.

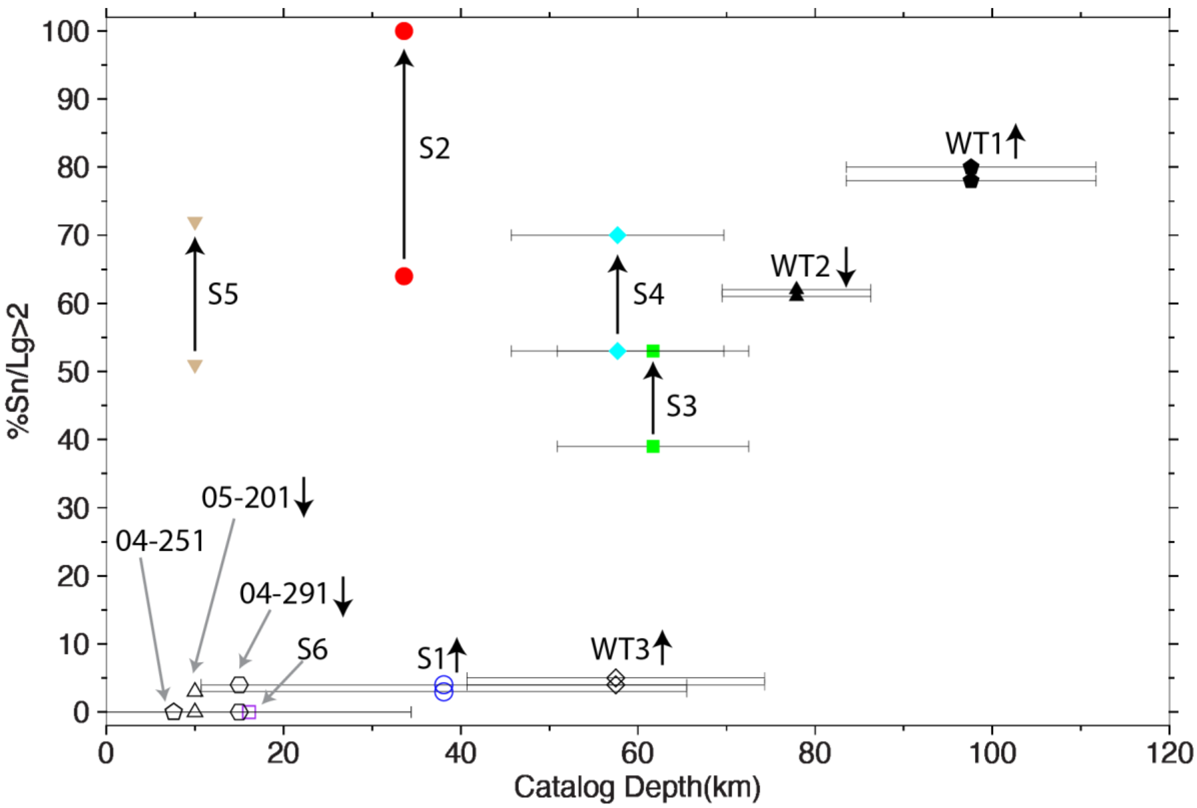


Fig. S8-1. Percentage of stations with *Sn/Lg*>2 for HiCLIMB events, *vs.* catalog depth. Symbols are the same as in Fig. 10; colored symbols S1–S6 are earthquakes in the Himalayan foreland with Moho depth likely < 45 km (Singh et al., 2015; Mitra et al., 2018), whereas black symbols are from northwest 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.