On the Origin of Small-Scale Seismic Scatters at 660-km Depth

Wei Mao1, Michael Gurnis1, and Wenbo Wu1,2

1Seismological Laboratory, California Institute of Technology, Pasadena, CA, USA, 2Woods Hole Oceanographic Institution, Falmouth, MA, USA

Abstract Strong small-scale seismic scatters (<10 km) have been recently observed at 660 km depth, but their origin remains uncertain. We systematically conduct both high-resolution 2-D geodynamic computations that include realistic thermodynamic properties, synthetic seismic waveforms, and insight from shallow seismic observations to explore their origin. We demonstrate that neither short-term subduction, nor long-term mechanical mantle mixing processes can produce sufficiently strong heterogeneities to explain the origin of such small-scale seismic scatters. Instead, the intrinsic heterogeneities inside the oceanic lithosphere which subducts into the mantle transition zone and the uppermost lower mantle can explain the observed short-wavelength scatter waves.

Plain Language Summary Similar to the Earth's bumpy surface separating the top atmosphere and hydrosphere from the underlying solid Earth, roughness could also exist on the Earth's other deeper discontinuities. Strong small-scale roughness with a laterally scale of about 10 km has been recently seen at the 660 km depth from seismic wave records, but their origin remains uncertain. We conduct computations which incorporate realistic rock properties to simulate the recorded seismic waves. We demonstrate that neither short-term (millions of years) subduction of oceanic lithosphere from the surface into the mantle, nor long-term (billions of years) mechanical mantle mixing processes inside the Earth are able to explain the origin of the seismic observations. Instead, the intrinsic heterogeneities inside the oceanic lithosphere which subducts into the mantle are capable of explaining the observed short-wavelength topography.

1. Introduction

Seismic tomography models reveal broad long-wavelength structures associated with mantle convection. Both global and regional tomographic models resolve ubiquitous high velocity seismic anomalies in subduction zones, indicative of subduction of oceanic lithosphere into the deep mantle (e.g., Grand et al., 1997; van der Hilst et al., 1997). The seismic inversions show that diverse slab morphologies are imaged in the mantle transition zone (e.g., French & Romanowicz, 2014; Fukao & Obayashi, 2013; Goes et al., 2017; Ritsema et al., 2011). Additionally, two Large Low Shear Velocity Provinces (LLSVPs) generally below African and the Pacific, which extend several thousand kilometers horizontally and (at least) several hundred kilometers radially, are observed above the core-mantle boundary (CMB) (e.g., French & Romanowicz, 2015; He & Wen, 2012; Ishii & Tromp, 1999; Masters et al., 2000; Ni et al., 2002; Ritsema et al., 2011; Su & Dziewonski, 1997; Wen et al., 2001). By incorporating surface plate motion history, geodynamic models suggest these long-wavelength seismic structures can be explained by subduction history (e.g., Bower et al., 2013; Lithgow-Bertelloni & Richards, 1998; Liu et al., 2008; McNamara & Zhong, 2005).

In addition, significant progress has been made on constraining the character and distribution of fine-scale structures in the deep mantle by studies of seismic scatters (Kaneshima, 2016; Shearer, 2015; Yu & Garnero, 2018). Seismic scatters are usually observed as relatively incoherent energy which arrive after (the coda) a direct seismic arrival (e.g., P, S, and PKP) or before the direct arrival (a precursor). The seismic scatters could be associated with volumetric structural heterogeneities (Shearer, 2015; Zhang et al., 2018) or roughness and/or topography at specific reflection depths (Benz & Vidale, 1993; Earle & Shearer, 1997). Unlike the long wavelength (>100 km) topography at 410 and 660 depths, which has been extensively studied (e.g., Deuss et al., 2006; Schmerr & Garnero, 2007), short wavelength topography (<100 km) at these depths has been poorly constrained.

Recently, by analyzing high-frequency short-distance PKPKP (P’P’) precursors, Wu et al. (2019) reported the existence of strong (2%) small-scale scatters at 660 km depth. The observations have been interpreted as scattering waves due to short wavelength (1–10 km) topography at 660 km depth. The potential sampling by the
observed data cover regions not only near Japan, Central America, South America, and the Hellenic subduction zones, but also around the Antarctic Plate where there is no present-day subduction (Figure 1). In contrast, short wavelength topography at the 410 km depth (P’•410•P’) could be hardly detected (Wu et al., 2019).

The underlying origin of these short-wavelength topography/scatters remain unclear. One hypothesis is the accumulation of subducted oceanic crust in the mantle transition zone (Irifune & Tsuchiya, 1994). Previous geodynamic models have demonstrated that the phase change at 660 km depth could act as an efficient filter that prevents the lower density material (e.g., harzburgite) from descending to greater depths and denser material (e.g., basalt) below the spinel-to-post-spinel phase change (e.g., Ballmer et al., 1999; Mambole & Fleitout, 2002; Nakagawa & Buffett, 2005; Nakagawa et al., 2010; Ogawa, 2000, 2003; van Keken et al., 1996; Weinstein, 2010). The principal controlling parameters in these models are the Clapeyron slope of the spinel-to-post-spinel phase change and the density contrast of the chemical heterogeneities relative to ambient mantle. Some chemical layering could occur if the Clapeyron slope or the density contrast of the chemical heterogeneities, either individually or together, are sufficiently large. Some previous studies have suggested the possibility of buoyant oceanic crust accumulating at the bottom of the mantle transition zone (Ballmer et al., 2015; Davies, 2008; Irifune & Tsuchiya, 2015; Mambole & Fleitout, 2002; Nakagawa & Buffett, 2005; Nakagawa et al., 2010; Ogawa, 2000, 2003; Ringwood, 1994; van Keken et al., 1996; Yan et al., 2020). However, those studies only explore structures at relatively longer wavelength (>20 km; e.g., Haugland et al., 2018; Jones et al., 2020; Waszek et al., 2021), which are incapable of explaining structures smaller than 10 km as observed by Wu et al. (2019). Consequently, high resolution dynamic models are required to explore the origin of short-wavelength structures at 660 km depth.

In this study, we perform high-resolution geodynamic experiments to explore the possible origin of short-wavelength roughness at 660-km depth. Specifically, we compute the influence of phase transitions, long-term mantle mixing, detachment of oceanic crust and internal heterogeneities of subducted oceanic lithosphere.

2. Geodynamic Model Formulation

We solve the mass, momentum, and energy conservation equations governing mantle convection using the finite element method Citcom (Moresi & Gurnis, 1996; Moresi et al., 1996) in a 2-D Cartesian domain. The dimensionless conservation equations are:

\[ \mathbf{V} \cdot \mathbf{u} = 0, \] (1)

\[ -\nabla P + \nabla \cdot (\eta \dot{\varepsilon}) + \text{Ra} \ddot{\zeta} = 0, \] (2)

\[ \frac{\partial T}{\partial t} + (\mathbf{u} \cdot \nabla) T = \nabla^2 T, \] (3)

where \( \mathbf{u} \) is the flow velocity vector, \( P \) is the dynamic pressure, \( \eta \) is the viscosity, \( \dot{\varepsilon} \) is the strain rate tensor, \( T \) is temperature, Ra is the Rayleigh number, and \( \ddot{\zeta} \) is the normalized horizontal density anomaly (usually called the
density anomaly $R$ in other papers), respectively. $\hat{e}_i$ is the unit vector in vertical direction, and $t$ is the time. The reference parameters are listed in Table 1.

Ra and $\hat{p}$ are defined as:

$$ Ra = \frac{\rho g a \Delta T d^3}{\kappa \eta_0}, $$

$$ \hat{p} = \frac{\Delta \rho_{P,T,Ci} \kappa}{\rho a \Delta T}, $$

where $\rho$ is the mantle reference density, $g$ is the gravitational acceleration, $a$ is the thermal expansivity, $\Delta T$ is the temperature difference between the surface and CMB, $\kappa$ is the thermal diffusivity, $\eta_0$ is the reference viscosity, $\Delta \rho_{P,T,Ci}$ is the horizontal density anomaly, and $d$ is the thickness of the mantle.

Instead of assuming a simplified linear parameterization of thermodynamic properties of mantle rocks, we apply a new and more realistic thermodynamic data set (Stixrude & Lithgow-Bertelloni, 2022) computed from the code HeFESTo (Stixrude & Lithgow-Bertelloni, 2005, 2011). We use tracers to track the composition field from which we calculate the density $\rho_{P,T,C}$ and seismic velocities ($V_p$ and $V_s$) of three end-member rocks, that is, basalt, harzburgite, and pyrolite (Baker & Beckett, 1999; Workman & Hart, 2005). We also apply a pre-calculated look-up table of the thermodynamic data set from the code Perple_X (Connolly, 2005) to accelerate the calculation.

Mantle viscosity is depth-, temperature-, and composition-dependent with the non-dimensional form

$$ \eta(T, r, c) = \eta_c(r) \eta_c \exp[E(0.5 - T)], $$

where $\eta_c(r)$ is the depth-dependent prefactor as a function of radius $r$, $\eta_c$ is the composition dependent prefactor, and $E$ is the non-dimensional activation energy. The dimensional viscosity depends on reference viscosity $\eta_0$ and hence Ra. Ra in most cases is $2.3 \times 10^6$ with a corresponding lower mantle viscosity of $\sim 2.5 \times 10^{22}$ Pa·s (parameters in Table 1). The non-dimensional activation energy is defined as:

$$ E = \frac{Q_c \Delta T}{RT_{\text{CMB}}^2}, $$

Its value is 18.4 (Table 1) in most cases and results in a total viscosity contrast of $10^8$ for dimensionless temperature varying from 0 to 1. For most cases, the viscosity of basalt is 100 times smaller than those of harzburgite and pyrolite (Immoor et al., 2022; van Keken et al., 1996).

The computational domain is 8,610 km across and 2,870 km deep (i.e., aspect ratio of 3). Generally, solutions are computed on a mesh with 1921 horizontal and 513 node points and refined near the surface, subduction zone and mantle transition zone so that in these regions the elements are 2–6 km in size. The averaged tracers per element is 64 and the total number of tracers is $\sim 63 \times 10^6$. We also perform resolution test for regional subduction models with a resolution of 1 $\times$ 1 km and the evolution, the temperature anomaly field, and the amplitude and travel time of synthetic waveforms of the model are nearly identical to the standard resolution.

The non-dimensional surface temperature is fixed at 0, while the CMB is fixed at 1. For regional subduction models, the initial temperature for oceanic lithosphere is calculated from a half-space cooling model (Turcotte & Schubert, 2014) based on the age determined by the ridge spreading rate (Figure S7a in Supporting Information S1). The initial dimensionless temperature for overriding lithosphere increases linearly from 0 at the surface to 0.52 at 125 km depth. Below the lithosphere, the initial temperature is 0.52 except near the CMB.

We also conducted long-term mixing models (with a total integration time of 2 Gyr), which was suggested to be the origin of small-scale mantle heterogeneities (e.g., Gurnis, 1986; Olson et al., 1984). The main differences between the long-term mantle mixing model and the short-term subduction model are: (a) While the velocity boundary condition at all sides is free-slip in long-term models, a prescribed kinematic surface boundary condition is imposed at the surface to control trench motions for regional subduction models (Christensen, 1996; Han

### Table 1

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mantle thickness, $d$</td>
<td>2,870 km</td>
</tr>
<tr>
<td>Mantle density, $\rho$</td>
<td>3,300 kg/m³</td>
</tr>
<tr>
<td>Gravitational acceleration, $g$</td>
<td>9.8 m/s²</td>
</tr>
<tr>
<td>Thermal expansivity, $a$</td>
<td>$3 \times 10^{-5}$/K</td>
</tr>
<tr>
<td>Surface temperature, $T_0$</td>
<td>273 K</td>
</tr>
<tr>
<td>CMB temperature, $T_{\text{CMB}}$</td>
<td>2773 K</td>
</tr>
<tr>
<td>Reference temperature difference, $\Delta T$</td>
<td>2500 K</td>
</tr>
<tr>
<td>Thermal diffusivity, $\kappa$</td>
<td>$10^{-6}$ m²/s</td>
</tr>
<tr>
<td>Gas constant, $R_p$</td>
<td>8.314 J/(K·mol)</td>
</tr>
<tr>
<td>Specific heat, $C_p$</td>
<td>1250 J/(kg·K)</td>
</tr>
</tbody>
</table>
& Gurnis, 1999; H. Liu et al., 2021). Note that since previous studies with different treatments of surface boundary condition gave rise to similar characteristics of long-duration mantle mixing structures (e.g., Brandenburg et al., 2008; Nakagawa et al., 2010; Yan et al., 2020), we did not apply complex surface boundary conditions in long-term mixing models. (b) Near the surface, a 7-km basalt layer and a 23-km harzburgite layer (Ringwood, 1982) is continuously generated at regions where there is upwelling flow in long-term mixing models. While no extra basalt or harzburgite is generated in short-term subduction models.

3. Statistical Characteristics of Seismic Heterogeneities and Synthetic Seismic Waveform Calculations

3.1. Statistical Characteristics of Seismic Heterogeneity

The spatial distribution of seismic heterogeneities can be quantitatively described by statistical models. Different types of statistical models are commonly classified by their auto-correlation function (ACF), which characterizes the spatial variation and intensity of heterogeneities statistically. The ACF can be transformed into its Power Spectral Density (PSD) of anomalies in the frequency domain. In 2-D, this becomes:

$$P_{2D}(k_x, k_y) = \iint_{-\infty}^{\infty} ACF(x, y)e^{-(k_x x + k_y y)} dxdy,$$

where $k_x$ and $k_y$ are wavenumbers in $x$ and $y$ directions.

We choose exponential ACF models to characterize seismic heterogeneities (Wu et al., 2019), in which:

$$ACF(r) = \sigma^2 e^{-\lambda r},$$

The corresponding PSDs are:

$$P_{1D}(k) = \frac{2\sigma^2 A}{[1 + (Ak)^2]^{\frac{3}{2}}} \text{ in } 1D, \quad (10)$$

$$P_{2D}(k) = \frac{2\pi\sigma^2 A^2}{[1 + (Ak)^2]^{\frac{3}{2}}} \text{ in } 2D, \quad (11)$$

where $\sigma$ is the root-mean-square of the media property (i.e., $P$-wave speed, $S$-wave speed and density) perturbation, which describes the magnitude of the seismic velocity anomaly, $r$ is the distance, and $A$ is the auto-correlation length which determines how PSD or energy is distributed in the wavenumber domain. A larger correlation wavelength $A$ indicates a larger portion of energy at a small wavenumber. In 2-D, $A$ can be decomposed into two directions, that is, $A = \sqrt{(A_x^2 + A_y^2)}$. In the study of Wu et al. (2019), a combination of $A = 4$ km and $\sigma = 2\%$ could fit the observation reasonably well (Figures 1b and 1c).

3.2. Synthetic Seismic Wave Calculation

Synthetic waveforms are computed using the Spectral Element Method (SEM) SPECFEM3D Cartesian package tool (Komatitsch & Tromp, 1999; Tromp et al., 2008; Wu et al., 2018), and are compared with observed seismic waveforms (Figures 1a and 1b, Wu et al., 2019). Performing calculations in a whole Earth’s domain is computationally expensive and we substantially improved efficiency by applying a hybrid method (Wu et al., 2018). The idea of this hybrid method is limiting the expensive 3-D numerical simulation to a small target region, where the structural heterogeneities are present, and computing the wavefield propagation outside of the target region using a faster 1-D Direct-Solution-Method (DSM) (Kawai et al., 2006; Wu et al., 2018). In this deep Earth coupling scheme, we need to conduct the DSM computations twice—the first computes the waves that propagate from the source to the boundary elements of the SEM box, and then additionally propagating the 3-D SEM scattering wavefield to the remote receiver. We set up a SPECFEM model with dimensions of $200 \times 200 \times 100$ km$^3$ at the 660 km depth and incorporate small-scale heterogeneities to simulate the 3-D scattering waves. For the target region, the distributions of density and seismic velocities are either taken from stochastic models or geodynamic computations. Because the geodynamic computations are time-dependent, we computed the synthetic waveform...
from when there the slab arrives at the mantle transition zone until 10–30 Myr later when either the majority part of the slab which is in the mantle transition zone and has reached the lower mantle or when the main characteristic of slab does not change.

4. Results

4.1. Short-Term Regional Subduction Model

We first present models of short-term (tens of millions of years [Myr]) regional subduction models. Similar to previous computations of subduction zone dynamics, we varied trench retreat velocity to explore slab dynamics in the mantle transition zone. Our short-term regional subduction models show similar dynamic evolution as in previous models with more simplified density (e.g., Christensen, 1996; Mao & Zhong, 2021; Zhong & Gurnis, 1995).

Case A1 with relatively large trench retreat velocity (5 cm/yr), as the slab subducts to the bottom of the mantle transition zone, both the viscosity jump across the interface between the upper and lower mantle and the buoyancy force due to density contrast between the slab and ambient mantle resist the slab from sinking into the lower mantle vertically. The slab is horizontally deflected in the mantle transition zone before sinking into the lower mantle (Figures 2a–2d, Figure S7 in Supporting Information S1). Because the phase transition from ringwoodite to bridgmanite and magnesiowüstite of the cold slab occurs at larger depth compared with ambient mantle, a localized low density and low seismic velocity layer of 30–60 km thick forms beneath 660 km depth (Figures S8e and S8k in Supporting Information S1). At the top of the slab, the basaltic crust layer also shows relatively small seismic velocity compared with underlying slab lithosphere. These two low seismic velocity layers are at relatively long wavelength of several hundreds of kilometers.

The evolution of slab changes as the trench retreat velocity varies, as demonstrated with Case A2 which is identical to Case A1 except that the trench is fixed (i.e., trench retreat velocity is 0 cm/yr; Table 2). In this case, as the slab sinks deeper into the lower mantle, it buckles resulting in an “S” shape (Figure 2e). The basaltic crust layer is 100 times lower in viscosity and ~4% denser than the slab lithosphere, such that crustal segregation occurs during subduction (Figure 2g, Figure S7 in Supporting Information S1). A more complex seismic velocity structure results at ~660 km depth region through the combination of temperature, composition, and slab geometry (Figure 2h, Figure S9 in Supporting Information S1). However, based on our synthetic waveform computations, those complex velocity structures still fail to reproduce observed short-wavelength P*660*P scatter waves (Figure 3d) which require strong short-wavelength heterogeneities (A = 0.5–10 km, σ = 2%; Figures 3a–3c, Figures S4 and S5 in Supporting Information S1).

In summary, four end-member scenarios of slab morphology in the mantle transition zone as widely reported in previous geodynamic studies are found: (a) Vertical penetration, (b) buckling, (c) horizontal deflection in the mantle transition zone, and (d) horizontal deflection below the transition zone. However, despite diverse slab morphologies in the mantle transition zone and uppermost lower mantle, all of these models show long-wavelength structures and are inconsistent with the observed short-wavelength P*660*P scatter waves (e.g., Figure 3d for Case A2).

<table>
<thead>
<tr>
<th>Case number</th>
<th>Duration</th>
<th>Trench retreat velocity (cm/yr)</th>
<th>Ra</th>
<th>Basalt viscosity reduction ratio</th>
<th>Initial imposed heterogeneities in oceanic lithosphere</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Short-term</td>
<td>5</td>
<td>$2.3 \times 10^6$</td>
<td>100</td>
<td>–</td>
</tr>
<tr>
<td>A2</td>
<td>Short-term</td>
<td>0</td>
<td>$2.3 \times 10^6$</td>
<td>100</td>
<td>–</td>
</tr>
<tr>
<td>B1</td>
<td>Long-term</td>
<td>–</td>
<td>$2.3 \times 10^6$</td>
<td>100</td>
<td>–</td>
</tr>
<tr>
<td>B2</td>
<td>Long-term</td>
<td>–</td>
<td>$4.6 \times 10^6$</td>
<td>100</td>
<td>–</td>
</tr>
<tr>
<td>B3</td>
<td>Long-term</td>
<td>–</td>
<td>$2.3 \times 10^6$</td>
<td>10</td>
<td>–</td>
</tr>
<tr>
<td>B4</td>
<td>Long-term</td>
<td>–</td>
<td>$2.3 \times 10^6$</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>C1</td>
<td>Short-term</td>
<td>5</td>
<td>$2.3 \times 10^6$</td>
<td>100</td>
<td>Extra intrinsic perturbation</td>
</tr>
<tr>
<td>D1</td>
<td>Short-term</td>
<td>5</td>
<td>$2.3 \times 10^6$</td>
<td>100</td>
<td>Basalt distribution</td>
</tr>
</tbody>
</table>
4.2. Long-Term Mantle Mixing Model

The short-term regional subduction models shown above cannot capture the effects of long-term mantle mixing, which leads to crustal segregation and a reduction in the wavelength of heterogeneities (Brandenburg & Van...
Figure 3.
Therefore, we also compute a series of long-term mantle mixing models from which we can assess their influence on the seismic wavefield. These series of models differ from those in Section 4.1 in that the surface velocity boundary condition is free-slip instead of pre-imposed. The initial temperature and composition are in layers (Figures 4a and 4e). Case B1 has the same reference Rayleigh number as in Case A1 (Ra = 2.3 \times 10^6, Table 2). The models exhibit a diverse range of behaviors when cold downwelling encounters resistance at 660 km depth from both the viscosity and density (Figures 4b, 4d, and 4f). Similar to previous studies (e.g., Brandenburg & Van Keken, 2007; Nakagawa et al., 2010), we observe accumulation of basalt piles both around 660 km depth and above the CMB due to density contrast of basalt compared with pyrolite (Figure 4f).

We also varied Rayleigh number and the viscosity reduction ratio of basalt, both of which have been implied to have a strong influence on mixing (Brandenburg & Van Keken, 2007; Manga, 1996). Case B2 has a larger Ra of 4.6 \times 10^6 and exhibits thinner slabs and plumes as expected (Figure 5, Figure S10 in Supporting Information S1). Cases B3 and B4 with smaller viscosity reduction ratio of 10 and 1 shows similar behaviors as in Case B1.

The characteristics of seismic anomalies within the mantle in the long-term mixing models, are quantified by calculating the velocity perturbation and auto-correlation wavelength (Equation 10) as follows. First, we calculate the horizontal P-wave velocity anomaly $\delta V_p$ from the convection models between 640 and 680 km depths (Figure S11 in Supporting Information S1). Then, we calculated the 1-D PSD of P-wave velocity anomaly at those depths and found their average. Finally, we applied a least square method to recover the velocity perturbation $\sigma$ and auto-correlation length $\text{A}_{\sigma}$ based on the averaged PSD profile. From the time evolution of $\sigma$ and $\text{A}_{\sigma}$ (Figure 6), a sharp increase of $\sigma$ and $\text{A}_{\sigma}$ occurs with the arrival of slabs in the transition zone. The auto-correlation wavelength, $\text{A}_{\sigma}$ decreases until 600 Myr of forward evolution and indicates the development of small-scale structures. Even after 2.2 Gyr, the auto-correlation wavelength $\text{A}_{\sigma}$ is still above 100 km and $\sigma$ is around 1% (red solid line in Figure 6).

We calculate the minimum auto-correlation length $\text{A}_{\sigma}^{\text{mix}}$ based on a 500 km horizontal sampling window between 640 and 680 km depths. The changing rate of $\text{A}_{\sigma}^{\text{mix}}$ and $\sigma$ increases with increasing viscosity reduction and Rayleigh number. For most cases, $\text{A}_{\sigma}^{\text{mix}}$ fluctuates between 5 and 80 km and $\sigma$ fluctuates between 0.1% and 1.2%. The characteristic wavelength is significantly longer, and the perturbation magnitude is smaller than what is required to explain the observed P'•660•P' waves (e.g., Figure 3e for Case B2). We also conduct a series of synthetic waveform computations and confirm that these models cannot reproduce the observed P'•660•P' waves (e.g., Figure 3e for Case B2).

**4.3. Intrinsic Chemical Heterogeneities of the Subducted Plate**

Seismic studies suggest strong heterogeneities are present inside subducted oceanic plates (Furumura & Bennett, 2005; Shito et al., 2014; Keken, 2007; Christensen & Hofmann, 1994; Gurnis, 1986; Manga, 1996).
et al., 2013, 2015; Sun et al., 2014), and so we propose an alternative mechanism that the observed seismic scatters are intrinsic heterogeneities subducted into the mantle transition zone. For example, strong internal small-scale heterogeneities ($A_x = 10–50$, $A_z = 0.5$ km, $\sigma = 2–3\%$) are proposed from Po/So guiding waves in the Pacific and Philippine Sea oceanic lithosphere across varying plate ages (Furumura & Kennett, 2005; Kennett et al., 2014; Shiito et al., 2013, 2015). Therefore, we perform computations accounting for such intrinsic heterogeneities, which are generated inside the subducted oceanic lithosphere during the formation and evolution of the lithosphere before subduction. Appropriate initial conditions for the convection models are generated by first conducting a synthetic test for a stochastic model with $A_x = 10$, $A_z = 0.5$ km, and $\sigma = 2\%$ as suggested from guiding wave observations. This stochastic model also generates similar amplitude of P•660•P' waveform as the model with $A_x = A_z = 4$ km and $\sigma = 2\%$ (Figures 3a–3c). The result indicates the possibility of intrinsic heterogeneities in the subducted oceanic lithosphere for the origin of the short-wavelength scatters at 660 km depth. Finally, such models with internal heterogeneities are tested based on the inferred distribution from the seismic studies.

As the origin of seismic heterogeneities inside the oceanic lithosphere is not well-constrained (Kennett & Furumura, 2015), we explored two types of models that treat the origin of such heterogeneities differently. For both kinds of models, the harzburgite layer thickness is increased from 23 to 63 km from those in Section 1. Those values are more consistent with those observed in oceanic lithosphere (Kennett et al., 2014; Shiito et al., 2013, 2015). The first model type (Case C1) is identical to Case A1 in Section 4.1, except that we add extra density and seismic perturbations as intrinsic heterogeneities to the harzburgite layer (Figures S12c and S12d in Supporting Information S1). The extra perturbations follow the distribution of stochastic model with $A_x = 10$, $A_z = 0.5$ km, and $\sigma = 2\%$. The other type of model Case D1 treats the heterogeneities as distributed regions of basalt within the harzburgite layer (Figures S12e and S12f in Supporting Information S1). Following the distribution of anomalies of the same stochastic model, basalt is placed into the harzburgite layer where the anomaly is smaller than a certain threshold value ($\sim 1\%$).

The overall dynamic behavior of these models is generally similar as those in Section 4.1 without extra intrinsic chemical heterogeneities as the long-wavelength buoyancy structure changes little (e.g., Figure S13 in Supporting Information S1 for Case C1 and Figure S14 in Supporting Information S1 for Case D1), but the predicted seismic signal is not. With the addition of intrinsic chemical heterogeneities within the subducted lithosphere, the amplitudes of synthetic scatter waves of P•660•P' in both types of models are significantly higher than those in Sections 4.1 and 4.2 and comparable with those observed (Figure 3). In the first type of model (Case C1) with extra anomaly shows clear P•660•P' phase when the anomaly is either above or below 660 km depth. But in the second type of model with basaltic distribution inside the harzburgite layer (Case D1) clear P•660•P' phase only shows when the harzburgite layer is below 660 km depth (Figures 3g and 3h). These results arise because basalt shows stronger seismic contrast between harzburgite and pyrolite below a depth of 660 km depth (Figure S15 in Supporting Information S1).

5. Discussion

Global long-wavelength mantle structures have been well-recognized from both global tomography and geodynamic studies. While the relatively short wavelength structures are detected from various studies of seismic scatters (Kaneshima, 2016; Shearer, 2015; Yu & Garnero, 2018), their underlying origin has received little attention. In
our dynamic models, we include the most recent thermodynamic data set (Stixrude & Lithgow-Bertelloni, 2022) to compute both density and seismic velocity. These types of models with realistic thermodynamic properties are more physically self-consistent compared with previous models with either simplified thermodynamic properties or purely statistical models (Nakagawa et al., 2010). For the short-term (tens of Myrs) subduction models, these new models show similar dynamic patterns as previous studies (Christensen, 1996; Mao & Zhong, 2021; Zhong & Gurnis, 1995). One of the largest differences is the complex density and seismic velocity pattern in the mantle transition zone (Figures S8–S10 in Supporting Information S1), which previous models could not capture. However, those structures are over tens and hundreds of kilometers and cannot explain observed short-wavelength seismic scatters. Such seismic structures might be helpful in explaining observed long-wavelength multiple seismic interfaces in the mantle transition zone (Wang et al., 2020).

Long-term mechanical mixing is often considered as the dominant mechanism in explaining the origin of small-scale structures and geochemical heterogeneities inside Earth's mantle (Ferrachat & Ricard, 1998, 2001; Gurnis, 1986; Gurnis & Davies, 1986; Haugland et al., 2018; Hunt & Kellogg, 2001; Li, 2021; Manga, 1996; Olson et al., 1984; van Keken & Zhong, 1999; van Keken et al., 2002; Yan et al., 2020). However, it remains uncertain how the small-scale heterogeneities are distributed inside the mantle, especially those in the mantle transition zone where complex dynamics occur, but studies of OIBs and MORBs have been interpreted in terms of pervasive small-scale heterogeneity in the mantle (e.g., Davies, 1984; Helffrich & Wood, 2001; Hofmann, 1997; Jackson & Dasgupta, 2008; Stixrude & Lithgow-Bertelloni, 2012). By conducting long-term high-resolution numerical computations (2 x 2 km) of mantle mixing simulations, our results indicates that even after 2 Gyr of mechanical mixing, the heterogeneities are not strong enough to explain the observed P’•660•P’ (Figure 3e and Figure S10 in Supporting Information S1).

Our dynamic models indicate that intrinsic heterogeneities inside the oceanic lithosphere can explain the small-scale scatters when subducted slabs reach the mantle transition zone and the uppermost lower mantle. The detailed distribution and thickness of such heterogeneity varies at different subduction zones. For example,
Sun et al. (2014) argued that heterogeneities in the top 40 km of the slab are required to fit the high frequency coda wave at the Calabria subduction zone above 300 km depth. Shen et al. (2021) argued that the strength of scatters, especially those within the top of the slab, weakens below 410 km in the Japan subduction zone. We also perform calculations with heterogeneities only in the bottom 50 km (Kennett & Furumura, 2015) to test the influence of distribution of heterogeneities in the slab. For models with extra density and seismic perturbations as the intrinsic heterogeneities to the harzburgite layer, they show similar scattering signals as Case C1 (Figure S16a in Supporting Information S1). While models with basalts distributed inside the harzburgite layer, they do not exhibit as strong scatter signals as Case D1 (Figure S16b in Supporting Information S1). Considering that seismic velocity contrast between basalt and harzburgite at 300–600 km depth in cold slab is significantly smaller than those at ~660 km (Figure S15 in Supporting Information S1), a thicker layer of harzburgite with basalt heterogeneities inside the slab might explain the decrease in the strength of scattering signals in the slab observed by Shen et al. (2021).

The origin of such intrinsic heterogeneities remains uncertain. The origin of the elongated heterogeneity inside the oceanic lithosphere is interpreted as frozen melts as the oceanic plate moves away from the mid-ocean ridge (Furumura & Kennett, 2005) or intraslab hydrous minerals (Shen et al., 2021). The origin of the observed seismic scatters at 660 km depth might not just come from the subducted oceanic lithosphere, but also from the deaminated continental lithosphere which might also be stagnant in the mantle transition zone (Peng et al., 2022), as the averaged auto-correlation wavelength $\lambda$ is 4 km and the perturbation $\sigma$ is 4% from surface to 200 km depth (Shearer & Earle, 2004), which is similar to those at 660 km depth. Better constraints on global distribution of such heterogeneities could strengthen our understanding of their origin. Other mechanisms, such as partial melting, fluids migration and strain localization during subduction (Tommasi & Vauchez, 2015) might also play an important role.

In this study, we only perform 2-D dynamic simulations. While in 3-D, both less (Schmalzl et al., 1996) and more efficient mixing are reported (Ferrachat & Ricard, 1998; van Keken & Zhong, 1999) compared to 2-D. The difference among those models is whether chaotic mixing, which is generated by toroidal flow in complex geometric settings (e.g., transform fault and ridge), is considered. When no chaotic mixing is considered (Schmalzl et al., 1996), the wavelength of heterogeneities is larger than those in 2-D model, which makes it more difficult to explain high frequency seismic scattering. In contrast, when chaotic mixing is considered, the scaling law of chaotic mixing indicates that a 7 km thick oceanic crust can be reduced to 10 cm in less than 4.2 Gyr (Ferrachat & Ricard, 1998), which is 4–5 orders smaller than the wavelength of the interested wavelength (1–10 km). For our 2-D dynamic models, we extrapolate the same 2D slice into 3-D domain, which might result in smaller strength of heterogeneities as lateral variations are neglected. Low viscosity basalt could enhance the crustal segregation process (Manga, 1996; van Keken et al., 1996), while the basalt viscosity in the deep mantle remains poorly constrained (Immoor et al., 2022). Furthermore, while we have applied the latest thermodynamic data set (Stixrude & Lithgow-Bertelloni, 2022), large uncertainties remain in the density and seismic velocity, especially those in the lower mantle. Future studies with better model configurations can enhance our understanding in the origin and distribution of small-scale scatters in the mantle.

We note that scatters due to short-term slab subducting and long-term mantle mixing are much weaker than the seismic estimates, but the perturbations estimated by Wu et al. (2019) have uncertainties. To estimate the perturbation from heterogeneities, Wu et al. (2019) used $P'$Surf$' (P'P' scattered near the surface) and PKiKP (inner core reflected P-wave) as reference phases to fit the amplitude of 660-km precursor. Although the ray paths of the two reference phases show some similarities as the 660-km precursors, large portions of their ray paths remaining different and would bias scattering estimations. In addition, Wu et al. (2019) analyzed seismic records from only 12 earthquakes, which sample only a few regions in the world. Future studies using more records and better quantification of uncertainties would provide an improved picture of global 660-km scatter distribution.

6. Conclusion

We perform systematic analyses of the origin of short-wavelength (<10 km) scatters at 660 km depth observed by Wu et al. (2019). By conducting both high-resolution geodynamic and synthetic waveform computations, we find that the origin of the short-wavelength scatters are neither likely from the complex phase transition in the subduction zone, nor from long-term mechanical mixing in the mantle. Instead, the seismic signals might arise from intrinsic heterogeneity with subducted oceanic lithosphere.
Data Availability Statement

Figures are drawn using the Generic Mapping Tools (GMT, www.soest.hawaii.edu/gmt/). The hybrid SEM-DSM code is available at https://github.com/wenbowu-geo/SEM_DSMM_hybrid. The thermodynamic computation code HeFesto is available at https://github.com/stixrude/HeFestoRepository. Perple_X is available at https://www.perplex.ethz.ch. The mantle convection code Citcom is available at https://geodynamics.org/cig. All model input parameters are given in Table 1, and all data are available at https://doi.org/10.6084/m9.figshare.19839763.v1.

References


The authors are grateful to two anonymous reviewers for careful reviews of the paper. This work is partially supported by the National Science Foundation through EAR-2009935. The authors acknowledge the Texas Advanced Computing Center (TACC) at the University of Texas at Austin for providing HPC resources that have contributed to the research results reported within this paper. The authors are grateful to discussions during the Seismolab Deep Earth Seminar in Caltech, especially comments from Jennifer Jackson and Zhongwen Zhan.


