Constraining Jumps in Density and Elastic Properties at the 660 km discontinuity Using Normal Mode Data via the Backus-Gilbert Method

Harriet C.P. Lau[1], and Barbara Romanowicz[1,2,3]

[1]Department of Earth and Planetary Science, University of California, Berkeley, CA 94720, USA
[2]College de France, Paris, France
[3]Institut de Physique du Globe de Paris, Paris, France

Key Points:

- We use recent normal mode center frequency data to constrain the elastic/density properties of the mantle 660 km discontinuity
- We find that acceptable range of jumps in P wave-speed and density fall outside that of standard seismic reference models
- Our data preclude the global discontinuity being as shallow as 650 km depth

Corresponding author: Harriet C.P. Lau, hcplau@berkeley.edu

Abstract

We apply the Backus-Gilbert approach to normal mode center frequency data, to constrain jumps in P, S, bulk-sound speed and density at the "660" discontinuity in the earth's mantle (∼650-670 km depth). Different 1D models are considered to compute sensitivity kernels. When using model PREM (Dziewonski and Anderson, 1981) as reference, with a "660" at 670 km depth , the best-fitting jumps in density, P and S wave-speeds range from (5.1-8.2)%, (5.3-8.0)%, (5.0-7.0)%, respectively, so the PREM values lie outside the ranges of acceptable density and P wave-speed jumps. When shifting the depth of "660" to 660 km, the density and S wave-speed jumps increase while the P wave-speed jump decreases. Normal mode data do not support a global transition at 650 km depth. The density jumps are closer to those of pyrolite than PREM while our bulk-sound wave-speed jumps suggest a larger garnet proportion at "660".

1 Introduction

Phase transitions that occur throughout the mantle greatly affect mantle dynamics and their precise location can provide information about the thermal and compositional variations within the earth. In this study we focus on the so-called "660" discontinuity (hereafter, 660) which has been observed to occur at depths between ∼650–670 km, attributed, at least in part, to the transition between the mineral phases ringwoodite/spinel at lower pressures to bridgmanite and oxides at higher pressures (Birch, 1952; Ringwood, 1991; Shearer, 2000; Shim et al., 2001; Stixrude & Lithgow-Bertelloni, 2005; Frost, 2008). Two approaches have provided important insights into the nature of 660.

Such a phase transition will produce sharp jumps in seismic wave-speed, manifested by seismological observations of reflected phases such as precursors to short period P'P' phases (Xu et al., 2003), precursors to long period SS phases (Shearer, 2000), and converted phases as detected in receiver function studies (Andrews & Deuss, 2008). The depth of the sharp jumps in wave speed listed in seismic reference, spherically symmetric (1D) models vary from 670 km (for PREM, Dziewonski & Anderson, 1981) to 650 km (for STW105, Kustowski et al., 2008), with a currently preferred value of 660 km. At the global scale, the topography of this discontinuity reaches up to ±30 km (Andrews & Deuss, 2008), with somewhat larger excursions locally, e.g., in subduction zones (e.g., Niu & Kawakatsu, 1995). Observed jumps in S-wave speed, Δv_{s}, P wave-speed, Δv_{p}, and in density, $\Delta\rho$, range from 4.5–10.1%, 2.5–5.6%, and 4.2–10.2%, respectively (Montagner & Anderson, 1989; Kennett & Engdahl, 1991; Morelli & Dziewonski, 1993; Estabrook & Kind, 1996; Shearer & Flanagan, 1999; Castle & Creager, 2000). Along with inherent trade-offs between the different physical parameters, the complicated nature of seismic signals observed across the boundary itself must contribute to the wide range of seismically observed jumps (Andrews & Deuss, 2008).

Efforts to combine the mineral physics and seismological approaches aim to tie physical causes to observed seismic properties. For example, by applying equations of state derived from mineral physics, assuming a mantle of adiabatic pyrolite composition,Cammarano et al. (2005) showed that wave-speeds jumps that satisfy seismic reference models lay towards the higher end of permissible values from mineral physics constraints. Indeed, in order for wave-speed jumps to be of larger magnitude, an increased proportion of garnet compared to that of pyrolite is usually invoked (Vacher et al., 1998; Wang et al., 2006; Ishii et al., 2018).

In this study, we revisit the estimation of globally averaged Δv_{s}, Δv_{p}, Δv_{b}, and $\Delta\rho$ across 660 by applying Backus-Gilbert based methods (Backus & Gilbert, 1970; Pijpers & Thompson, 1992; Masters & Gubbins, 2003) to an extensive recent normal mode catalogue (Roult et al., 2010; Deuss et al., 2013).

2 Data and Methodology

2.1 Data

We use the normal mode center frequencies (and uncertainties) compiled by Robson and Romanowicz (2019) which is based on a combination of the Reference Earth Model catalogue (Laske & Masters, n.d.), observations from Deuss et al. (2013) and radial modes from Roult et al. (2010). The data are provided in Supplementary Table 1.

2.2 Methodology

In a spherical elastic and isotropic earth model, the eigenfrequency ω_k of any isolated normal mode multiplet (denoted by the index k) has distinct sensitivity kernels to v_s, v_p, and ρ structure across the mantle and to topography of any discontinuity, d_i. In the framework of first order perturbation theory, the fractional change in eigenfrequency may be expressed as (e.g. (Woodhouse & Dahlen, 1978)):

$$\frac{\delta\omega_k}{\omega_k} = \int_0^a \left[M^k_{v_\mathrm{p}}(r)\frac{\delta v_\mathrm{p}}{v_\mathrm{p}}(r) + M^k_{v_\mathrm{s}}(r)\frac{\delta v_\mathrm{s}}{v_\mathrm{s}}(r) + M^k_{\rho}(r)\frac{\delta v_\rho}{v_\rho}(r) \right] \mathrm{d}r + \sum_i M^k_{\mathrm{d},i}\delta\mathrm{d}_i\,, \tag{1}$$

where r is the radius and $r = a$ is the surface, $M^k_{v_\mathrm{p}}$, $M^k_{v_\mathrm{s}}$, M^k_ρ are the sensitivity kernels of mode k to perturbations in v_p, v_s, and ρ, respectively. $M^k_{\mathrm{d},i}$ is the sensitivity to topography, d, on the i-th discontinuity.

By considering a linear combination of equation (1) over a set of modes k, we obtain:

$$\sum_k c_k \frac{\delta\omega_k}{\omega_k} = \int_0^a \left[\mathcal{K}_{v_\mathrm{p}}(r)\frac{\delta v_\mathrm{p}}{v_\mathrm{p}}(r) + \mathcal{K}_{v_\mathrm{s}}(r)\frac{\delta v_\mathrm{S}}{v_\mathrm{S}}(r) + \mathcal{K}_{\rho}(r)\frac{\delta v_\rho}{v_\rho}(r) \right] \mathrm{d}r + \sum_i \mathcal{K}_{\mathrm{d},i}\delta\mathrm{d}_i \tag{2}$$

where $\mathcal{K}_\mathrm{X} = \sum_k c_k M^k_\mathrm{X}$ for parameter X which, in this study, X $= v_\mathrm{p}$, v_s or ρ. The coefficients c_k may be determined such that $\mathcal{K}_\mathrm{X}$ is designed to enhance the sensitivity of the weighted observations (left-hand side of eq. 2) to a specific region within the mantle and a specific parameter X, while simultaneously reducing the sensitivity to other parameters, Y, Z, and d_i. If the weights $\mathbf{c}$ are successfully determined, in the most ideal case $\mathcal{K}_\mathrm{X}$ will be only non-zero across the region of interest, and $\mathcal{K}_\mathrm{Y}$, $\mathcal{K}_\mathrm{Z}$ and $\mathcal{K}_{\mathrm{d},i}$ will be zero everywhere. We will refer to both the weighted kernels and data as *composite* kernels and data. An additional condition required of the *composite* kernel is that it should be unimodular:

$$\int_0^a \mathcal{K}_X(r)\mathrm{d}r = 1. \tag{3}$$

This is the essence of the Backus-Gilbert methodology. Finding the best combination of data, i.e., finding $\mathbf{c}$, requires solving an inverse problem and thus carries with it the same regularization issues as in typical geophysical inverse problems.

To expand upon this, we introduce the concept of a *target kernel*, $\mathcal{T}$, as introduced by Pijpers and Thompson (1992), whose methodology we closely follow (though they considered only one free parameter). $\mathcal{T}$ will be designed such that it follows the shape of the desired sensitivity. Here, we will explore three kernels: (1) a narrow Gaussian centered at 660 (solid black line, Fig. 1a) which is defined as:

$$\mathcal{T}_\mathrm{full} = \frac{1}{\Lambda}\exp\left(-\left(\frac{r-r_0}{\Delta}\right)^2\right) \tag{4}$$

where Λ is chosen so that the area under $\mathcal{T}_{\text{full}}$ is 1, Δ is the characteristic width of the Gaussian centered at $r = r_0$. The remaining two target kernels are truncated versions of this Gaussian, one where the kernel is identical to the full Gaussian above 660, but is zero below 660 , $\mathcal{T}_+$ (orange kernel, Fig. 1a) and the other where the kernel is identical to the full Gaussian below 660, but zero above, $\mathcal{T}_-$ (blue kernel, Fig. 1a). The truncated kernels,$\mathcal{T}_-$ and $\mathcal{T}_+$, will provide estimates on either side of the 660, which we will use to determine new jump constraints for ρ, v_{p}, and v_{s}. The full Gaussian, $\mathcal{T}_{\text{full}}$ will provide an overall constraint across the 660 boundary when testing synthetic models in Section 3.2.

In order to determine $\mathbf{c}$ such that the resulting $\mathcal{K}_{\text{X}}$ is as similar to $\mathcal{T}$ as possible, we use the method of Lagrange multipliers to minimize the following expression:

$$\Phi = \int_0^a \left[(\mathcal{K}_{\text{X}} - \mathcal{T})^2 + \mathcal{K}_{\text{Y}}^2 + \mathcal{K}_{\text{Z}}^2\right] \text{d}r + \sum_i \mathcal{K}_{\text{d},i}^2 + \mu \sum_{ij} E_{ij} c_i c_j + \lambda \int_0^a \mathcal{K}_{\text{X}}(r)\, \text{d}r \qquad (5)$$

where λ is a Lagrange multiplier, $\mathbf{E}$ is the covariance matrix of data errors and μ is its corresponding trade-off parameter. Minimizing Φ with respect to the $(N + 1)$ unknowns, c_j (where $j = 1, 2, ..N$ and N is the number of normal mode center frequencies considered) and λ, yields $N + 1$ linear equations. The first N equations have the form:

$$\begin{aligned}\sum_j \quad & \left[\int_0^a M_{\text{X}}^i M_{\text{X}}^j + M_{\text{Y}}^i M_{\text{Y}}^j + M_{\text{Z}}^i M_{\text{Z}}^j \ \text{d}r + \sum_d M_{\text{d},i} M_{\text{d},j} + \mu E_{ij}\right] c_j - \\ & \int_0^a M_{\text{X}}^i \mathcal{T} \text{d}r = 0, \qquad (6)\end{aligned}$$

and the $(N + 1)$-th equation, resulting from (3) is:

$$\sum_j c_j \int_0^a M_{\text{X}}^j(r) - 1 = 0. \qquad (7)$$

This system of $N+1$ linear equations with $N+1$ unknowns $c_j, j = 1, N$ and λ can be written in matrix form as:

$$\mathbf{Ac} = \mathbf{v} \qquad (8)$$

where $\mathbf{c} = [c_1, c_2, c_3, ..., c_N, \lambda]$ and vector $\mathbf{v}$ is

$$v_i = \int_0^a M_{\text{X}}^i \mathcal{T} \text{d}r \qquad (9)$$

for $i \leq N$ and $v_{N+1} = 1$. The elements of the $(N + 1) \cdot (N + 1)$ symmetric matrix $\mathbf{A}$ are:

$$\begin{aligned} A_{ij} &= \int_0^a M_{\text{X}}^i M_{\text{X}}^j + M_{\text{Y}}^i M_{\text{Y}}^j + M_{\text{Z}}^i M_{\text{Z}}^j \ \text{d}r + \sum_d M_{\text{d},i} M_{\text{d},j} + \mu E_{ij} \text{ when } i, j \leq N, \\ &= \int_0^a M_{\text{X}}^i \text{d}r \text{ when } i \leq N, j = N + 1, \\ &= \int_0^a M_{\text{X}}^j \text{d}r \text{ when } j \leq N, i = N + 1, \\ &= 0 \text{ when } i = j = N + 1. \qquad (10)\end{aligned}$$

Since the constraint (3) is satisfied by the inclusion of the Lagrange multiplier, an estimate of the quantity of interest, $\tilde{\text{X}}$, may be obtained as follows. (Note that in the following expressions we make explicit any dependence on r and that the "$\sim$" symbol denotes any value integrated over r.)

$$\sum_k c_k \frac{\delta\omega_k}{\omega_k} = \int_0^a \frac{\delta \text{X}}{\text{X}_0}(r)\, \mathcal{K}_{\text{X}}(r)\, \text{d}r = \widetilde{\left(\frac{\delta \text{X}}{\text{X}_0}\right)}, \qquad (11)$$

where X_0 is the unperturbed depth profile of parameter X from which the kernels $\mathcal{K}_X$ were determined. Taking $\delta X = X - X_0$, we may write

$$1 + \sum_k c_k \frac{\delta\omega_k}{\omega_k} = \int_0^a \frac{X}{X_0}(r)\, \mathcal{K}_X(r) \mathrm{d}r \tag{12}$$

(the value of 1 arising from eq. 3). The radially dependent solution is thus:

$$X(r) = X_0(r) \left[1 + \frac{\delta X}{X_0}(r)\right]. \tag{13}$$

However, as eq. (11) does not yield radially dependent $\delta X/X_0$, our solution, $\tilde{X}$, is an approximation of this (i.e., $\tilde{X} \approx X$) over the radial range for which the kernel is nonzero (or non-negligible):

$$\tilde{X}(r) = X_0(r) \left[1 + \widetilde{\left(\frac{\delta X}{X_0}\right)}\right], \tag{14}$$

where r is the radius of interest.

The uncertainty in the estimate of $\tilde{X}$, ε, is due to two sources: data error, ε_{obs}, and contamination from imperfections in the composite kernels, ε_{con}, since in practice, they will not be fully zero where desired. The contribution from errors in observation is given by

$$\varepsilon_{\text{obs}}^2 = \mathbf{c} \cdot \mathbf{E} \cdot \mathbf{c}, \tag{15}$$

and ε_{con} is due to non-zero contributions from $\mathcal{K}_Y$, $\mathcal{K}_Z$ and $\mathcal{K}_{d,i}$ (Masters & Gubbins, 2003) but also differences between $\mathcal{T}$ and $\mathcal{K}_X$. This may be estimated by the following expression

$$\varepsilon_{\text{con}}^2 = \int_0^a |\mathcal{K}_X - \mathcal{T}||\varepsilon_X| + |\mathcal{K}_Y||\varepsilon_Y| + |\mathcal{K}_Z||\varepsilon_Z| \; \mathrm{d}r + \sum_i |\mathcal{K}_{d,i}||\varepsilon_d|, \tag{16}$$

where ε_X, ε_Y, ε_Z, and ε_d are uncertainties in the parameters X, Y, Z, and d, respectively, and $|\cdot|$ denotes taking the absolute value.

3 Results

3.1 Composite Kernels

We calculated the sensitivity kernels for each mode to each parameter M according to Dahlen and Tromp (1998), using the software package MINEOS (https://geodynamics.org/cig/software/mineos/), adopting the widely used seismic reference model PREM (Dziewonski & Anderson, 1981) as reference. In Section 4.1 we explore the effect of the choice of 1-D reference model.

We solve for three separate sets of three kernels $\mathcal{T}_{\text{full}}$, $\mathcal{T}_-$, $\mathcal{T}_+$ (Fig. 1a). In the first set, we enhance sensitivity to ρ and suppress sensitivities to v_p, v_s, and d_i (Fig. 1b). In Figs 1(c,d) we show the analogous resulting composite kernels for enhancing sensitivity to v_p and v_s, respectively.

While not perfect, the overall shape of the composite kernels $\mathcal{K}_X$ capture $\mathcal{T}$ very well, with small amounts of noise in all composite kernels (i.e., $\mathcal{K}_Y$ and $\mathcal{K}_Z$), which will add to the uncertainty in the estimate.

3.2 New Estimates on Jumps

In Figs 1(e-g) we show the results obtained by solving eq. (8) successively for each set of composite kernels. In each figure, it is important to focus on the shift

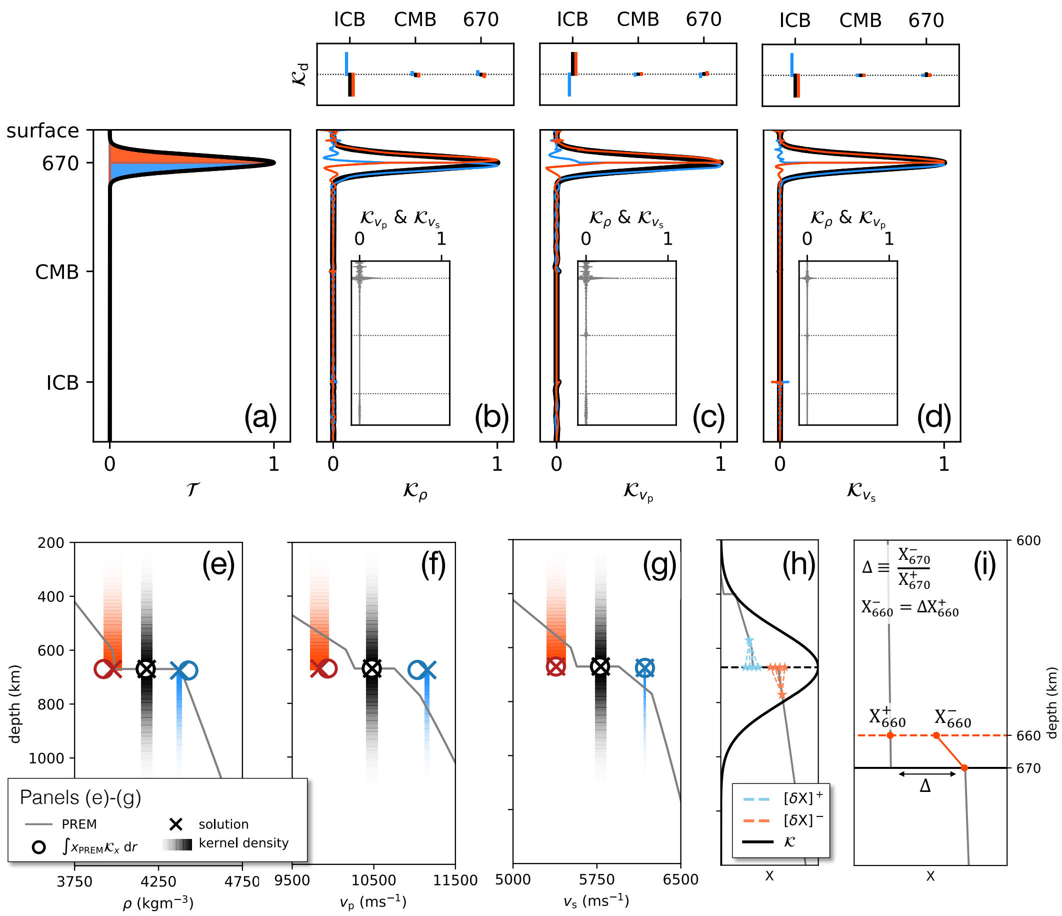


Figure 1. **Composite Kernels and Inversion Results.** (a) Target kernels for study (blue, $\mathcal{T}_-$, a half-Gaussian beneath 660; orange, $\mathcal{T}_+$, a half-Gaussian above, and black, $\mathcal{T}_{\text{full}}$, a full-Gaussian straddling both beneath and above 660). (b-d) The resulting kernels to enhance sensitivity to ρ, v_p, and v_s, respectively. (In the inset panels, the resulting kernels for the other parameters chosen to be muted). The top panels show relative sensitivity to topography on each discontinuity. (e)-(g) Resulting perturbations in parameters X (where X is ρ, v_p, and v_s in panels (a,b,c), respectively. The circles denote how each kernel samples the background PREM model, while the crosses are the resulting perturbations when applying the composite data. The color intensity of the bar represents the kernel density (see panels b-d). The width of each bar is the uncertainty in the result (eqs 15-16). The gray lines display the PREM profile. (h) A schematic depiction of the test models we consider. The perturbations above and below 660 are guided by solutions shown in panels (e)-(g), and we consider these to increase or decrease linearly from the background PREM model in two cases: 100 km above and below the discontinuity (dashed ines) and 200 km above and below the discontinuity (not depicted). The thick black line is the scaled kernel that is associated with the black bars in panels (e)-(g). This is used as a criterion to exclude test models that do not fit the constraint from the black crosses in (e)-(g). (i) Schematic diagram of how we adjust PREM for a shallower discontinuity at both 660 km and 650 km depth, preserving the percent jump value to that of the standard PREM model. Panels (e)-(h) share the same vertical axes.

between the circle and cross, rather than the shift from the PREM profile. Focusing on the truncated-Gaussian solutions (orange and blue) and the width of the kernel density bar, it can be seen that for all parameters, the available normal mode data better constrain the parameter beneath 660 (which is at 670 km in PREM) than above. For v_s, the PREM values satisfy those of the composite data well (crosses overlap the circles in Fig. 1g). However, the composite data call for a smaller density jump (Fig. 1e),and for a larger jump in v_p (Fig. 1f) than in PREM, with shifts both above and below 660.

When one considers averaged values of PREM both above and below 660, PREM satisfies the composite data for all parameters. This is indicated by the black circles falling on the black crosses in all panels (Figs 1(e)-(g)).

To explore this further, we tested how these new estimates perform in reproducing the composite data. We perturbed PREM both above and below its 670 km discontinuity with our new estimates. The values of v_s, v_p and ρ were linearly interpolated to the PREM background model across two length scales: 100 km and 200 km. A schematic depiction of this perturbation is shown in Fig. 1h where we show test models for the 100 km length scale.

We produced many perturbed PREM models ($\mathrm{X} + \delta\mathrm{X}$), choosing values of parameters above and below the discontinuity within the uncertainty shown in Figs 1(e-g), applying all possible combinations of ρ, v_p, and v_s. This resulted in 46656 models for each length scale tested. To further scrutinize these models, and before confronting them with the data, we tested whether these models met the constraint provided by $\mathcal{T}_{\rm full}$ (shown in a scaled version in Fig. 1h). This condition simply requires that the value $\bar{\mathrm{X}}$ must lie within the horizontal span of the associated black colored bar where

$$\bar{\mathrm{X}} = \int_0^a (\mathrm{X} + \delta\mathrm{X})\mathcal{T}_{\rm full}\ \mathrm{d}r \tag{17}$$

and ($\mathrm{X} + \delta\mathrm{X}$) is the perturbed model tested. That is,

$$(\tilde{\mathrm{X}}_{\rm full} - \varepsilon) \leq \bar{\mathrm{X}} \leq (\tilde{\mathrm{X}}_{\rm full} + \varepsilon). \tag{18}$$

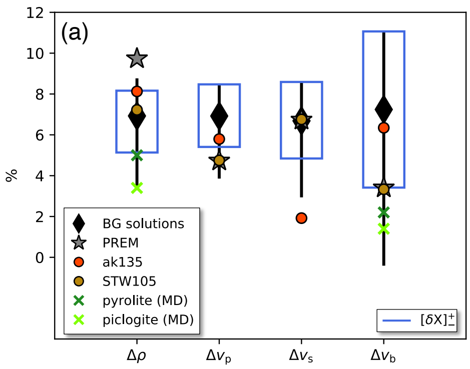
The result of this additional condition is that none of the models perturbed across a length scale of 200 km were able to meet the constraint, whereas for the 100 km length scale, 10,000 models satisfied the constraint. Suggesting that any perturbation from PREM cannot be too wide.

For all these results,listed in Fig. 2b, the associated jumps are displayed as percentages in Fig. 2a, where the final models tested are shown by the blue boxes. We note that, after the initial culling of models, $\Delta\rho$ and Δv_p are significantly different from their respective PREM values. We subject this culled subset of models to an additional test, as follows. For all models, we predict the set of composite data and define the chi-squared misfit, χ^2, as

$$\chi^2 = \sum_i \frac{(\Omega_i^{\rm mod} - \Omega_i^{\rm obs})^2}{\sigma_i^2}. \tag{19}$$

We note that Ω_i is the composite datum from each inversion performed where $i = [\rho^-, \rho^+, v_p^-, v_p^+, v_s^-, v_s^+]$ and $\Omega_i = \sum_{k=1}^N c_k^i \omega_k$. Indeed, for each parameter i enhanced, a different set of coefficients c_k^i is obtained, where k is the index of a mode. Each inversion is accompanied by a composite uncertainty, σ_i, weighted in the same manner.

We present the misfit reduction, γ, in Figs (3a-d), where γ is the ratio of χ^2 calculated from the test models to χ^2 calculated by PREM. These models correspond only to the perturbed models over a length scale of 100 km, since the 200 km



(b)

Model	$\Delta\rho$ (%)	Δv_p (%)	Δv_s (%)	Δv_b (%)
PREM[1]	9.7	4.7	6.7	3.4
ak135[2]	8.1	5.8	1.9	6.3
STW105[3]	7.2	4.7	6.8	3.3
pyrolite[4]	5.0	-	-	2.2
piclogite[4]	3.4	-	-	1.4
This study, $[\delta X]^{\pm}$	5.1-8.2	5.3-8.0	5.0-7.0	4.0-11.0

Figure 2. Resulting Jumps Across 660. (a) Black bars (Backus-Gilbert solutions) depict the jumps based on the results Figs 1(e)-(g). The length of the vertical gray lines here corresponds to the width of the colored bars in the latter figure. The blue boxes denote synthetically produced models tested against the data, perturbing PREM above and below the discontinuity over a length scale of 100 km (Fig. 1h). These ranges are less than the gray bars as models that did not satisfy the constraint imposed by $\mathcal{T}_{\text{full}}$ (condition 18) were culled. Symbols represent values from different seismic reference models and from molecular dynamics (MD) calculations. (b) Table of jumps across 660 all listed as percentages. The references for each are as follows: 1: Dziewonski and Anderson (1981); 2: Kennett et al. (1995); 3: Kustowski et al. (2008); 4: Matsui (2000).

models did not satisfy the $\mathcal{T}_{\text{full}}$ constraint (eq. 18). They span the blue boxes shown in Fig. 2a. Trade-offs between wave-speeds Δv_s and Δv_p are shown in Figs 3(a,c), while trade-offs between wave-speeds Δv_s and Δv_b are shown in Figs 3(b,d). Each row corresponds to models at two fixed density jumps, where $\Delta\rho$ is 5.1% and 8.2% for panels (a,b) and (c,d), respectively. These values correspond to the minimum and maximum $\Delta\rho$ values in the blue boxes of Figure 3.

In the models tested, the misfit was reduced (i.e., $\gamma < 1$) for a significant portion of the models. We see a general preference for lower Δv_s across the range tested and preference for higher Δv_p across the range we test, though these two parameters display some covariance. A trade-off also exists between v_s and v_b, whereby larger values of Δv_s are paired with smaller values of Δv_b.

We performed F-tests for all cases at the 99% level of significance (solid red line). For Δv_s, we see that we reach levels of 99% significance in the region of $\sim$(4.5–7.5)% for a $\Delta\rho$ of 5.1% and $\sim$(5.0–7.5)% for $\Delta\rho$ of 8.2%. As such, we report that these composite data provide revised estimates of these jumps at a 99% significance value of: $\Delta\rho$ = (5.1–8.2)%, Δv_s = (5.0–7.0)%, Δv_p = (5.3–8.0)% and Δv_b = (4.0–9.5)%, though these ranges are all correlated and should not be taken at face value. The original PREM values are 9.7%, 6.7%, 4.7%, and 3.4%, respectively (see also Fig. 2b).

4 Discussion

4.1 Effect of Background Model

In both Figs 2a and 3 we have overlain our results with the corresponding jumps in PREM (gray stars) and the seismological reference models “ak135” (orange circles, Kennett et al., 1995) and “STW105”(yellow circles, Kustowski et al., 2008).

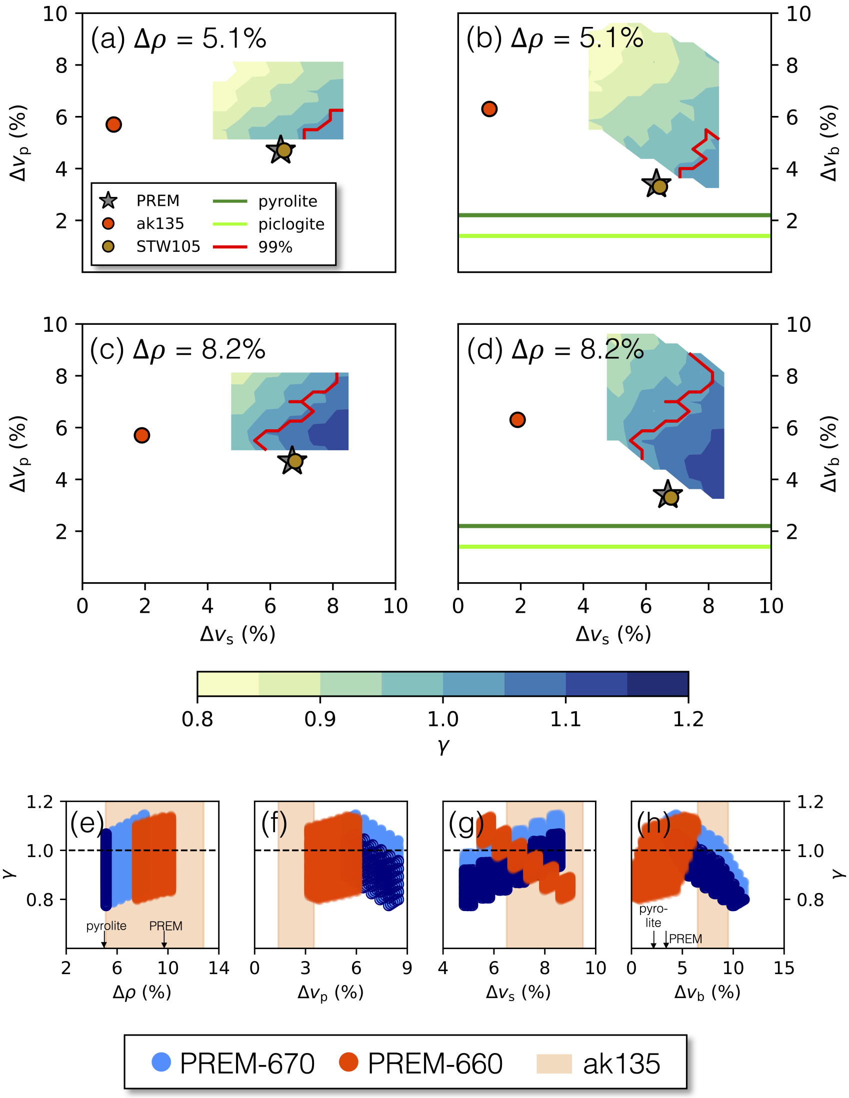


Figure 3. Misfits from Synthetic Tests and Their Trade-Offs. Panels (a-d) display contour plots of the misfit reduction, γ, for models that span the blue boxes in Fig. 2a (i.e., perturbed models over a 100 km length scale). Each row displays models at fixed $\Delta\rho$ values of 5.1% (panels a,b) and 8.2% (panels c,d). These $\Delta\rho$ values span the full range of models that satisfied condition (18). Left panels display trade-offs between Δv_s and Δv_p and right panels display trade-offs between Δv_s and Δv_b. Panels (e-h) show γ for test models that meet the criterion described by condition (18) assuming different background models: the standard version of PREM (blue circles, also shown in panels a-d, where dark blue circles distinguish a subset of these models for which $\Delta\rho$ is the minimum possible density jump of 5.1% highlighting the directionality of trade-offs), PREM in which the discontinuity has been artificially shifted upward to 660 km (orange circles, see Fig. 1e), and assuming the background model ak135 (orange shaded range).

While for PREM, the largest differences are seen in $\Delta\rho$ and Δv_p, our results are in much closer agreement with the other two seismic models in $\Delta\rho$ and with ak135 for Δv_p. However, our estimate for Δv_s, while in agreement with PREM and STW105, is substantially different from that of ak135 (which is more than $\sim$2% different). These differences may reflect the fact that that STW105 and PREM were derived from much more similar datasets than ak135, which consisted mainly of short period travel time data sampling the earth's upper mantle beneath continental areas. We note that no single model shows an obvious consistency with our new estimates for all three parameters.

PREM and ak135 provide a good representation of the range of background models available given the differing nature of the datasets used to produce them. However, another notable difference between these reference models is that PREM places the 660 at a depth of 670 km, whereas ak135 places this discontinuity at 660 km. Here we test how robust our results are to the choice of background model, and specifically, the depth of the 660 in the reference model.

Starting from PREM, we artificially adjusted the discontinuity to two depths: 650 and 660 km. On the upper side of the discontinuity, we kept the PREM parameter values down to the new discontinuity depth. On the underside, we imposed the density and wave-speeds that conserve the original percent jump of PREM for each quantity, in order to isolate the effect of the discontinuity depth. We then linearly interpolated to the original PREM values at 670 km depth (Fig. 1(e)). For each case, we produced new composite kernels, $\mathcal{K}_+$, $\mathcal{K}_-$, and $\mathcal{K}_\mathrm{full}$ and repeated the analysis of Sections 2-3.

No model with the discontinuity fixed at 650 km resulted satisfied the condition (18), whereas where the depth of discontinuity was 660 km, 9460 models satisfied this constraint. We show the misfit reduction for the latter distribution of models in Figs 3(e-g) (dark orange circles). For comparison we display the results from the standard analysis (blue circles). By elevating the discontinuity to a shallower depth of 660 km, the trade-offs between Δv_p and Δv_s switch direction, requiring an increase of $\Delta\rho$ and Δv_s and a reduction of Δv_p. Furthermore, it seems that a global average depth of 650 km for the 660 is too shallow to satisfy normal mode data.

These trade-offs between a shallower discontinuity and increase in both $\Delta\rho$ and Δv_s and a reduction of Δv_p seem to be consistent if we repeat the entire exercise with ak135 (Kennett et al., 1995). We find values corresponding to those of the gray bars in Fig. 2a of (5.1–12.8)%, (1.4–4.3)%, and (6.5–9.5)% for $\Delta\rho$, Δv_p and Δv_s, respectively. The same culling exercise that reduced the gray bars to the dark blue boxes in Fig. 2a (for a perturbation length scale of 100 km) did not result in significant changes to these ranges.

For $\Delta\rho$, the ak135 estimate is similarly poorly constrained relative to the gray bars, but shifted to higher density jumps. For Δv_p and Δv_s the span of the ak135 estimates are roughly two-thirds of the gray bars (compare Fig. 3(e-h) with the black bars in Fig. 2a). In Figs 3(e-h) (orange circles), the misfit reductions of the modified PREM model (with a depth of discontinuity of 660 km, as in ak135) reduce towards the ranges spanned by the ak135 result for all parameters except for Δv_b.

These differences between the Backus-Gilbert solutions based on PREM and ak135 are not trivial, and illustrate the strong non-linearity of the problem, combining the effects of the depth of the discontinuity and of the jumps in the three parameters considered. Since ak135 was constrained by a very different type of data, these results likely represent an unrealistic "worst case scenario" when applied to normal mode center frequency data.

4.2 Physical Implications

The characteristics of the 660 phase boundary have had much influence on the conceptual picture of mantle convection. Given that the negative Clapeyron slope implies that the transition shifts to higher pressures at colder temperatures, the idea that this transition is a barrier to general mantle circulation has been proposed extensively to satisfy geochemical constraints on mantle heterogeneity (e.g., Allègre, 1997) and explored dynamically with consideration to how parameters such as mantle viscosity might be affected (e.g., van Keken & Ballentine, 1998). While the picture of mantle convection continues to evolve with better seismic imaging techniques and more sophisticated modeling capabilities, many of these approaches require a background seismic model, and, in many cases, PREM is the model of choice.

As such, revising the globally averaged characteristics above and beneath the 660 in the light of recent data is important for both the geodynamical and seismological communities. Paired with increasingly accurate measurements from the mineral physics literature, a more complete picture of the physical characteristics of this region will be reached.

The mineral physics estimates of Matsui (2000) for $\Delta\rho$ and Δv_{b} of the model mantle compositions of pyrolite and piclogite (Ringwood, 1962; Bass & Anderson, 1984) are included in Figs 2a and 3(a-d) (green crosses and lines, respectively). Our estimates of $\Delta\rho$ are more in line with these mineral physics estimates relative to PREM, being closer to a pyrolitic composition. This is further visualized in Fig. 3(e) where misfit reductions point towards the estimate for pyrolite. However, we do not improve the Δv_{b} fit to either of these possible compositions (Fig. 3h). It seems from our results in Section 4.1 that these same conclusions stand whether we consider a discontinuity at 670 km as in PREM, or 660 km as in ak135.

Non-olivine components of the upper mantle, and in particular, the presence of ilmenite, may affect such jumps across 660 (e.g., Vacher et al., 1998). More recently, Ishii et al. (2018) explored the transition of ringwoodite to garnet and magnesiowüstite, as did Wang et al. (2006). The latter suggested that larger seismic jumps may be attributed to a larger degree of garnet in transforming perovsite. Our results support this inference. Indeed, while our new estimates bring PREM closer to mineral physics estimates for model mantle minerals, these still lie on the higher end of estimates for adiabatic pyrolite (Cammarano et al., 2005).

5 Conclusion

We have used the Backus-Gilbert method to find a combination of normal mode center frequency data that enhances sensitivity to just above and below the 660 discontinuity, for density, P wave-speeds, S wave-speeds. We have determined the best-fitting ranges of jumps in these parameters when assuming PREM as a background model (Fig. 2b). There is significant covariance between these parameters. The corresponding PREM value for our $\Delta\rho$ lies above this range, the Δv_{p} lies below this range, and Δv_{s} lies within this range. When shifting the depth of 660 to 660 km brings out additional trade-offs resulting in a range of acceptable models that span larger values of $\Delta\rho$ and Δv_{s}, and smaller values of Δv_{p}. In these calculations, we also found that the normal mode data do not support a globally averaged phase transition depth as shallow as 650 km depth.

Our results produce a range of values for $\Delta\rho$ and Δv_{b} that are generally higher than those estimated by mineral physicists for the pyrolite model, and in particular even higher than PREM for Δv_{b}, supporting the possibility of a larger proportion of garnet in the transformation to perovskite. Still, the density jump of PREM is at the high end of the acceptable models resulting from our study, which may be important for geodynamicists modeling global convection. Finally, the inability to obtain a consistent result when using ak135 as a reference model may reflect frequency dependence of structure and/or the presence of significant lateral variations around the 660.

Acknowledgments

We thank UC Berkeley for the support of this work and Dr Alexander Robson for sharing the compiled normal mode center frequencies (where the data are available through Deuss et al. (2013), Roult et al. (2010), and Laske and Masters (n.d.)). HL thanks David Al-Attar (University of Cambridge) for some fruitful discussions related to Backus-Gilbert methods.

References

Allègre, C. J. (1997). Limitation on the mass exchange between the upper and lower mantle: the evolving convection regime of the earth. *Earth and Planetary Science Letters*, *150*(1), 1 - 6. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0012821X97000721` doi: https://doi.org/10.1016/S0012-821X(97)00072-1

Andrews, J., & Deuss, A. (2008, June). Detailed nature of the 660 km region of the mantle from global receiver function data. *Journal of Geophysical Research (Solid Earth)*, *113*(B6), B06304. doi: 10.1029/2007JB005111

Backus, G., & Gilbert, F. (1970, March). Uniqueness in the Inversion of Inaccurate Gross Earth Data. *Philosophical Transactions of the Royal Society of London Series A*, *266*(1173), 123-192. doi: 10.1098/rsta.1970.0005

Bass, J. D., & Anderson, D. L. (1984, March). Composition of the upper mantle: Geophysical tests of two petrological models. *Geophysical Research Letters*, *11*(3), 229-232. doi: 10.1029/GL011i003p00229

Birch, F. (1952). Elasticity and constitution of the earth's interior. *Journal of Geophysical Research (1896-1977)*, *57*(2), 227-286. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JZ057i002p00227` doi: 10.1029/JZ057i002p00227

Cammarano, F., Deuss, A., Goes, S., & Giardini, D. (2005, January). One-dimensional physical reference models for the upper mantle and transition zone: Combining seismic and mineral physics constraints. *Journal of Geophysical Research (Solid Earth)*, *110*(B1), B01306. doi: 10.1029/2004JB003272

Castle, J. C., & Creager, K. C. (2000, March). Local sharpness and shear wave speed jump across the 660-km discontinuity. *Journal of Geophysical Research*, *105*(B3), 6191-6200. doi: 10.1029/1999JB900424

Dahlen, F. A., & Tromp, J. (1998). *Theoretical Global Seismology.* Princeton University Press.

Deuss, A., Ritsema, J., & van Heijst, H. (2013, May). A new catalogue of normal-mode splitting function measurements up to 10 mHz. *Geophysical Journal International*, *193*(2), 920-937. doi: 10.1093/gji/ggt010

Dziewonski, A. M., & Anderson, D. L. (1981, June). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, *25*(4), 297-356. doi: 10.1016/0031-9201(81)90046-7

Estabrook, C. H., & Kind, R. (1996, November). The Nature of the 660-Kilometer Upper-Mantle Seismic Discontinuity from Precursors to the PP Phase. *Science*, *274*(5290), 1179-1182. doi: 10.1126/science.274.5290.1179

Frost, D. J. (2008, 06). The Upper Mantle and Transition Zone. *Elements*, *4*(3), 171-176. Retrieved from `https://doi.org/10.2113/GSELEMENTS.4.3.171` doi: 10.2113/GSELEMENTS.4.3.171

Ishii, T., Kojitani, H., & Akaogi, M. (2018, January). Phase relations and mineral chemistry in pyrolitic mantle at 1600-2200 °C under pressures up to the uppermost lower mantle: Phase transitions around the 660-km discontinuity and dynamics of upwelling hot plumes. *Physics of the Earth and Planetary Interiors*, *274*, 127-137. doi: 10.1016/j.pepi.2017.10.005

Kennett, B. L. N., & Engdahl, E. R. (1991, May). Traveltimes for Global Earthquake Location and Phase Identification. *Geophysical Journal International*, *105*, 429-465. doi: 10.1111/j.1365-246X.1991.tb06724.x

Kennett, B. L. N., Engdahl, E. R., & Buland, R. (1995, July). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, *122*(1), 108-124. doi: 10.1111/j.1365-246X.1995.tb03540.x

Kustowski, B., EkströM, G., & DziewońSki, A. M. (2008, June). Anisotropic shear-wave velocity structure of the Earth's mantle: A global model. *Journal of Geophysical Research (Solid Earth)*, *113*(B6), B06306. doi: 10.1029/2007JB005169

Laske, G., & Masters, G. (n.d.). *Reference earth model.* `https://igppweb.ucsd.edu/~gabi/rem.html`.

Masters, G., & Gubbins, D. (2003, November). On the resolution of density within the Earth. *Physics of the Earth and Planetary Interiors*, *140*(1-3), 159-167. doi: 10.1016/j.pepi.2003.07.008

Matsui, M. (2000, September). Molecular dynamics simulation of MgSiO $_3$ perovskite and the 660-km seismic discontinuity. *Physics of the Earth and Planetary Interiors*, *121*(1-2), 77-84. doi: 10.1016/S0031-9201(00)00161-8

Montagner, J.-P., & Anderson, D. L. (1989, December). Constrained reference mantle model. *Physics of the Earth and Planetary Interiors*, *58*(2-3), 205-227. doi: 10.1016/0031-9201(89)90055-1

Morelli, A., & Dziewonski, A. M. (1993, February). Body Wave Traveltimes and A Spherically Symmetric P- and S-Wave Velocity Model. *Geophysical Journal International*, *112*(2), 178-194. doi: 10.1111/j.1365-246X.1993.tb01448.x

Niu, F., & Kawakatsu, H. (1995). Direct evidence for the undulation of the 660-km discontinuity beneath tonga: Comparison of japan and california array data. *Geophysical Research Letters*, *22*(5), 531-534. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/94GL03332` doi: https://doi.org/10.1029/94GL03332

Pijpers, F. P., & Thompson, M. J. (1992, September). Faster formulations of the optimally localized averages method for helioseismic inversions. *Astronomy and Astrophysics*, *262*(2), L33-L36.

Ringwood, A. E. (1962, February). A Model for the Upper Mantle. *Journal of Geophysical Research*, *67*(2), 857-867. doi: 10.1029/JZ067i002p00857

Ringwood, A. E. (1991, August). Phase transformations and their bearing on the constitution and dynamics of the mantle. *Geochimica et Cosmochimica Acta*, *55*(8), 2083-2110. doi: 10.1016/0016-7037(91)90090-R

Robson, A. J. S., & Romanowicz, B. (2019, October). New normal mode constraints on bulk inner core velocities and density. *Physics of the Earth and Planetary Interiors*, *295*, 106310. doi: 10.1016/j.pepi.2019.106310

Roult, G., Roch, J., & Clévédé, E. (2010, March). Observation of split modes from the 26th December 2004 Sumatra-Andaman mega-event. *Physics of the Earth and Planetary Interiors*, *179*(1-2), 45-59. doi: 10.1016/j.pepi.2010.01.001

Shearer, P. M. (2000, January). Upper mantle seismic discontinuities. *Washington DC American Geophysical Union Geophysical Monograph Series*, *117*, 115-131. doi: 10.1029/GM117p0115

Shearer, P. M., & Flanagan, M. P. (1999). Seismic velocity and density jumps across the 410- and 660-kilometer discontinuities. *Science*, *285*(5433), 1545–1548. Retrieved from `https://science.sciencemag.org/content/285/5433/1545` doi: 10.1126/science.285.5433.1545

Shim, S.-H., Duffy, T. S., & Shen, G. (2001, May). The post-spinel transformation in Mg_2SiO_4 and its relation to the 660-km seismic discontinuity. *Nature*, *411*(6837), 571-574.

Stixrude, L., & Lithgow-Bertelloni, C. (2005, August). Thermodynamics of mantle minerals - I. Physical properties. *Geophysical Journal International*, *162*(2), 610-632. doi: 10.1111/j.1365-246X.2005.02642.x

Vacher, P., Mocquet, A., & Sotin, C. (1998, April). Computation of seismic profiles from mineral physics: the importance of the non-olivine components for explaining the 660 km depth discontinuity. *Physics of the Earth and Planetary Interiors*, *106*(3-4), 275-298. doi: 10.1016/S0031-9201(98)00076-4

van Keken, P. E., & Ballentine, C. J. (1998). Whole-mantle versus layered mantle convection and the role of a high-viscosity lower mantle in terrestrial volatile evolution. *Earth and Planetary Science Letters*, *156*(1), 19 - 32. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0012821X98000235` doi: https://doi.org/10.1016/S0012-821X(98)00023-5

Wang, Y., Wen, L., Weidner, D., & He, Y. (2006, July). SH velocity and compositional models near the 660-km discontinuity beneath South America and northeast Asia. *Journal of Geophysical Research (Solid Earth)*, *111*(B7), B07305. doi: 10.1029/2005JB003849

Woodhouse, J. H., & Dahlen, F. A. (1978, 05). The Effect of A General Aspherical Perturbation on the Free Oscillations of the Earth. *Geophysical Journal International*, *53*(2), 335-354. Retrieved from `https://doi.org/10.1111/j.1365-246X.1978.tb03746.x` doi: 10.1111/j.1365-246X.1978.tb03746.x

Xu, F., Vidale, J. E., & Earle, P. S. (2003, January). Survey of precursors to P'P': Fine structure of mantle discontinuities. *Journal of Geophysical Research (Solid Earth)*, *108*(B1), 2024. doi: 10.1029/2001JB000817