AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE
10.1002/2014JB011067

Key Points:
- Complex ULVZ
- Probabilistic study of the ULVZ
- Uncertainty estimation

Supporting Information:
- Readme
- Text S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Figure S5

Correspondence to:
S. Pachhai,
surya.pachhai@anu.edu.au

Citation:
Pachhai, S., H. Tkalčić, and J. Dettmer (2014), Bayesian inference for ultralow velocity zones in the Earth's lowermost mantle: Complex ULVZ beneath the east of the Philippines, *J. Geophys. Res. Solid Earth*, *119*, 8346–8365, doi:10.1002/2014JB011067.

Received 23 FEB 2014
Accepted 6 OCT 2014
Accepted article online 22 OCT 2014
Published online 21 NOV 2014

Bayesian inference for ultralow velocity zones in the Earth's lowermost mantle: Complex ULVZ beneath the east of the Philippines

S. Pachhai[1], H. Tkalčić[1], and J. Dettmer[1,2]

[1]Research School of Earth Sciences, Australian National University, Canberra, ACT, Australia, [2]School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada

Abstract Ultralow velocity zones (ULVZs) are small-scale structures with a sharp decrease in *S* and *P* wave velocity, and an increase in the density on the top of the Earth's core-mantle boundary. The ratio of *S* and *P* wave velocity reduction and density anomaly are important to understanding whether ULVZs consist of partial melt or chemically distinct material. However, existing methods such as forward waveform modeling that utilize 1-D and 2-D Earth-structure models face challenges when trying to uniquely quantify ULVZ properties because of inherent nonuniqueness and nonlinearity. This paper develops a Bayesian inversion for ULVZ parameters and uncertainties with rigorous noise treatment to address these challenges. The posterior probability density of the ULVZ parameters (the solution to the inverse problem) is sampled by the Metropolis-Hastings algorithm. To improve sampling efficiency, parallel tempering is applied by simulating a sequence of tempered Markov chains in parallel and allowing information exchange between chains. First, the Bayesian inversion is applied to simulated noisy data for a realistic ULVZ model. Then, measured data sampling the lowermost mantle under the Philippine Sea are considered. Cluster analysis and visual waveform inspection suggest that two distinct classes of *ScP* (*S* waves converted to, and reflected as, *P* waves) waves exist in this region. The distinct waves likely correspond to lateral variability in the lowermost mantle properties in a NE-SW direction. For the NE area, Bayesian model selection identifies a two-layer model with a gradual density increase as a function of depth as optimal. This complex ULVZ structure can be due to the percolation of iron-enriched, molten material in the lowermost mantle. The results for the SW area are more difficult to interpret, which may be due to the limited number of data available (too few waveforms to appropriately reduce noise) and/or complex 2-D and 3-D structures that cannot be explained properly by the 1-D models required by our inversion approach. In particular, the complex waveforms require highly layered 1-D models to fit the data. These models appear physically unreasonable and suggest that the SW region cannot be explained by 1-D structure.

1. Introduction

1.1. Lowermost Mantle and ULVZs

Seismological studies suggest that the lowermost mantle is at least as complex as the Earth's lithosphere. The known complexities in the lowermost mantle include heterogeneity at a variety of spatial scales, i.e., at short scales of 10 to 100 km [e.g., *Doornbos*, 1974; *Bataille and Lund*, 1996; *Bréger et al.,* 2000; *Rost and Earle*, 2010], intermediate scales of 500 to 1000 km [e.g., *Wysession et al.,* 1999; *Bréger and Romanowicz*, 1998], and long scales of greater than 1000 km [e.g., *Su et al.,* 1994; *Masters et al.,* 1996; *Wysession*, 1996; *Mégnin and Romanowicz*, 2000; *Ritsema et al.,* 2011]. *P* wave tomographic models of the lowermost mantle using correlated waveforms sensitive to the lowermost mantle [*Tkalčić et al.*, 2002; *Young et al.,* 2013] or derived as a part of whole mantle tomography [e.g., *Masters et al.,* 1996; *Mégnin and Romanowicz*, 2000; *Kárason and van der Hilst*, 2001; *Fukao et al.,* 2003; *Houser et al.,* 2008; *Soldati et al.,* 2012; *Obayashi et al.,* 2013] are rare and do not indisputably agree with *S* wave velocity models. Most recently, a multiscale heterogeneous power spectrum of the lowermost mantle has been inferred by Bayesian inversion [*Young et al.,* 2013], bridging between short-scale heterogeneity inferred from observations of scattering [e.g., *Haddon and Cleary*, 1974] and a dominant long-scale degree 2 pattern typically obtained from *S* wave tomographic models [e.g., *Fukao*, 1992; *Su et al.,* 1994; *Li and Romanowicz*, 1996; *Masters et al.,* 1996; *Kennett et al.,* 1998; *Ritsema et al.,* 2011]. In addition, a number of studies supported the existence of compositionally distinct heterogeneities in the lowermost mantle [e.g., *Karato and Karki*, 2001; *Tackley*, 2002; *Tkalčić and Romanowicz*, 2002; *McNamara and Zhong*, 2004].

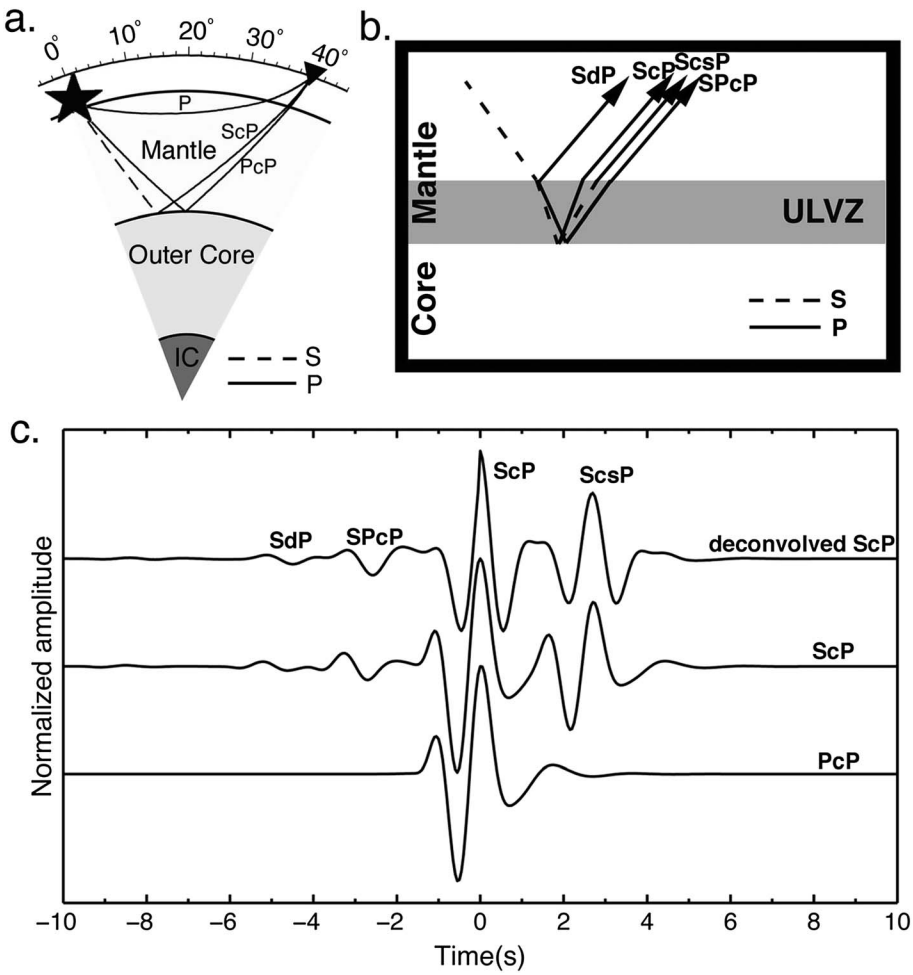

Figure 1. (a) Geometric ray paths for PcP and ScP waves at an epicentral distance of 40°, earthquake source location (star), and receiver location (triangle). (b) Schematic diagram of a 1-D, one-layer ULVZ on top of the CMB and its effect on ScP waveforms. (c) Waveforms for a 20 km thick layer with anomalous elastic properties (*P* and *S* wave velocity decrease of 10% and 30%, respectively, density increase of 10% with respect to ak135), PcP, ScP, and PcP deconvolved from ScP phases. Phases sensitive to the ULVZ are marked as SdP, SPcP, and ScsP. All waveforms are normalized and centered on the maximum amplitude. This example assumes that the PcP phase does not sample the ULVZ.

Apart from these heterogeneities, a strong decrease in *S* and *P* wave velocity with an increase in density has been observed in localized regions of the core-mantle boundary (CMB), which are known as ultralow velocity zones (ULVZs) [e.g., *Garnero et al.,* 1993; *Garnero and Helmberger*, 1995]. The ULVZs were first identified using core-side-diffracted *S* waves (SPdKS) as layers 5–40 km thick, with a 10% decrease in *P* wave velocity [e.g., *Garnero and Helmberger*, 1995; *Helmberger et al.,* 1996; *Williams and Garnero*, 1996]. More recent studies suggest the presence of much thinner layers with decreases in *S* and *P* wave velocity and increases in density [e.g., *Garnero and Helmberger*, 1998; *Rost and Revenaugh*, 2003; *Rost et al.,* 2006; *Idehara et al.,* 2007; *Idehara*, 2011]. Most ULVZs are located at the margins of large low-shear-velocity provinces (LLSVPs) [e.g., *McNamara et al.,* 2010] and play an important role in the generation of mantle upwellings and hotspot volcanism in association with thermal instabilities in overlying LLSVPs [e.g., *Rost et al.,* 2005].

1.2. Common Modeling Practices for ULVZs

Anomalous structures such as ULVZs produce complex signals (i.e., before (precursors) and after (postcursors) the main arrival) in seismic waves reflected from the CMB [e.g., *S* waves converted to, and reflected as, *P* waves (ScP)] (Figure 1). The amplitude and position of the precursor and postcursor depend on the ULVZ properties. For example, the amplitude of the precursor depends on the *P* wave velocity impedance contrast, while the amplitude of the postcursor depends on both *P* and *S* wave velocity impedance contrasts.

Similarly, the positions of precursor and postcursor depend on the layer thickness and *S* wave velocity perturbation. Thus, precursor and postcursor carry information necessary to constrain the thickness, *P* and *S* wave velocity, and density of ULVZs. It is possible to model the impedance contrast at the CMB (by means of PKP, PcP, and ScP waves [e.g., *Vidale and Heldin*, 1998; *Wen and Helmberger*, 1998a; *Rost et al.,* 2005; *Idehara*, 2011]) or the lateral variation of the structure (by means of SPdKS waves; [e.g., *Garnero and Helmberger,* 1995, 1998; *Thorne and Garnero*, 2004]). Finally, the existence of core rigidity zones (CRZs) below the CMB was proposed by studies using ScP [*Rost and Revenaugh*, 2001, 2003] and SPdKS waves [*Garnero and Jeanloz*, 2000].

Forward waveform modeling is widely used to constrain ULVZ properties: density, velocity, and thickness of the ULVZ are adjusted until synthetic seismograms match the main features in the observed waveforms using trial and error or grid search algorithms. It is typically possible to fit observed waveforms with ULVZ models of various elastic properties and layer geometries due to the highly nonunique and nonlinear nature of the inverse problem, which is characterized by strong parameter trade-offs (correlations). Some studies utilize 2-D and 2.5-D (i.e., a 2-D model is used to explain the 3-D ray geometry) modeling of SPdKS waves to constrain the vertical and lateral extent of ULVZ structures [*Helmberger et al.,* 2000; *Rondenay et al.,* 2010; *To et al.,* 2011; *Thorne et al.,* 2013; *Jensen et al.,* 2013]. However, SPdKS waves passing through a strong and narrow ULVZ cannot be distinguished from those passing through a weak and wide ULVZ. This introduces another trade-off between lateral extent and other ULVZ parameters [*Rondenay et al.,* 2010]. In summary, while the shape and extent of ULVZs has been extensively studied, details are still poorly understood because of difficulties stemming from nonuniqueness and nonlinearity of the inverse problem.

1.3. Motivation for This Study

Seismic wave velocity can be decreased either by decreasing the modulus of elasticity or by increasing the density. At CMB temperature and pressure conditions, the 1:3 ratio of *P* to *S* wave velocity reduction favors a partial melting origin of the ULVZ [*Williams and Garnero*, 1996]. On the other hand, the density of the lowermost mantle material can be increased in the presence of heavy material such as iron. The iron can either originate from the core [e.g., *Karato and Karki*, 2001] or from the formation of a banded form of iron due to oxidation at the seafloor and subduction to the CMB [*Dobson and Brodholt*, 2005] resulting in a negative ratio of *S* wave velocity to density contrasts. A negative ratio can be interpreted as a chemical origin of ULVZs. However, it is difficult to uniquely quantify ULVZ parameters because of strong trade-offs (correlations) between ULVZ parameters [e.g., *Garnero and Helmberger*, 1998; *Garnero and Jeanloz*, 2000] and additionally between the ULVZs, CRZs, and core-mantle transition zones [*Garnero and Jeanloz*, 2000; *Rost and Revenaugh*, 2003; *Rondenay et al.,* 2010]. Therefore, uncertainty quantification of the ratio of *S* and *P* wave velocity reductions, as well as the density within a ULVZ is key to inferring whether ULVZs consist of partially molten material, chemically distinct material, or a combination of the two. In addition, the presence of a dense and melted component in ULVZs has important dynamical implications such as percolation of the melt phase of the dense material [*Hernlund and Tackley*, 2007] and stirring of the ULVZs [*Hernlund and Jellinek*, 2010].

The parameter sensitivities of *P* wave velocity and density to ULVZ structure are typically weak, while thickness and *S* wave velocity are more sensitive. In addition, the precursor and postcursor of the main ScP wave can be difficult to identify at common noise levels. Thus, waves with ambiguous precursors can yield both anomalous (i.e., complex structure) and normal (i.e., similar to the ak135 model [*Kennett et al.,* 1995]) structures at the same CMB location [e.g., *Garnero and Vidale*, 1999; *Rost and Revenaugh*, 2001, 2003; *Idehara*, 2011].

Grid search algorithms can find a single solution that minimizes the data misfit but are limited to models with few parameters and do not consider other solutions that may fit the data reasonably well. Therefore, relying on single solutions to represent all data information contained in the observed data is problematic. More sophisticated algorithms such as the neighborhood algorithm [e.g., *Sambridge*, 1999] can be used to eliminate the dependency on a single solution and allow examination of more complex models. Although this algorithm allows computation of ensemble solutions that satisfy predefined criteria, ensemble solutions obtained by this algorithm are based on an arbitrary misfit level (e.g., the best 1000 models), and such ensembles cannot be used for rigorous quantification.

Therefore, a fundamental challenge in ULVZ parameter estimation is to quantify parameter uncertainties and distinguish uncertainty from parameter variability as a function of ULVZ height above the CMB (variability is a measure of structure change while the uncertainty is a measure of our knowledge of the structure).

Uncertainties stem from both measurement (i.e., errors in seismic recordings) and theory errors (i.e., errors resulting from approximations in the model including seismic wave propagation). Uncertainties in ULVZ model parameters were obtained in previous studies either by using goodness of fit [*Rost et al.,* 2005] or by bootstrapping the standard deviation of waveform correlations [*Idehara*, 2011].

This paper develops a rigorous Bayesian approach [e.g., *Mosegaard and Tarantola*, 1995; *Mackay*, 2003; *Tarantola*, 2005] to study the ULVZ. In Bayesian inversion, the solution to an inverse problem is given probabilistically by the posterior probability density (PPD) of model parameters. Bayesian inversion considers prior information (independent information about the parameters which can be expressed as a probability) and updates it with data information that is expressed in terms of the likelihood function. This formulation is particularly useful for nonlinear, nonunique inverse problems. Here uniform prior densities over wide bounds are considered for all parameters such that data information predominantly produces the results. Parameter estimates and uncertainties are inferred from the PPD by marginalization. Furthermore, the PPD ratio of the change in *P* to *S* wave velocity and the change in density to *S* wave velocity is computed to better understand the physical origin of the ULVZ.

The ULVZ model is parameterized by flat-lying, homogeneous layers given by layer thickness, *P* and *S* wave velocity, and density. For observed data, the parameterization is typically not known independently and must be estimated from the data. This process is referred to as model selection. However, the fit to the data improves with the number of parameters so that fit alone cannot be considered a sufficient criterion for model selection. This challenge is greatly exacerbated for Bayesian inference where parameter uncertainty estimation is a main goal: Too complex parameterizations cause overfitting of data, which results in overestimated uncertainties, while too simple parameterizations cause underfitting of data, which results in underestimated uncertainties. Hence, a criterion for an optimal parameterization, such as the Bayesian information criterion (BIC) [*Schwarz*, 1978], is required for model selection [*MacKay*, 2003; *Dettmer et al.,* 2009]. More sophisticated algorithms such as transdimensional inversion [e.g., *Malinverno*, 2002; *Sambridge et al.,* 2006; *Bodin and Sambridge*, 2009; *Dettmer et al.,* 2010a; *Agostinetti and Malinverno*, 2010; *Tkalčić et al.,* 2013a, 2013b] or Bayesian evidence computation [*Dettmer et al.,* 2010b] can be applied at substantially higher computational cost.

Bayesian inference requires specification of a likelihood function which is based on an assumption about the distribution form of the data errors (e.g., Gaussian). Data errors include both measurement and theory errors. Measurement errors are due to the data recording process, while theory errors include model approximations such as flat-lying layers, source complexity, and path effects. Data errors are typically not known independently and must be approximated by residual errors (the difference between observed and predicted data). Data errors are often assumed to be Gaussian-distributed and uncorrelated. However, assuming uncorrelated errors can significantly underestimate parameter uncertainty if residuals are correlated. Here correlated errors are accounted for by a first-order autoregressive (AR1) error model. The AR1 model captures the covariance structure with two parameters (standard deviation and AR1 coefficient).

For nonlinear inverse problems, no analytic solution exists for the PPD and Markov chain Monte Carlo (MCMC) sampling techniques are typically applied to estimate the PPD. Metropolis-Hastings (MH) sampling is the most commonly applied MCMC method. However, the strong nonlinearity and nonuniqueness of some inverse problems can prove challenging for basic MH sampling. In particular, strong parameter correlations and multiple, high-probability modes separated by extensive low-probability regions can cause substantial challenges [*MacKay*, 2003]. Here interacting Markov chains (parallel tempering) are applied to address these challenges [*Geyer*, 1991; *Dettmer et al.,* 2012]: The MH sampling algorithm is augmented by a sequence of increasingly tempered Markov chains to aid sampling efficiency. Heavily tempered chains deemphasize the effect of the likelihood function and sample widely, while less tempered chains are constrained to more local exploration. The chains are simulated in parallel, and information exchange between them substantially improves sampling efficiency.

We apply this hierarchical Bayesian approach to waveform inversion of simulated and observed core-reflected ScP waves to study ULVZ structure. First, the ability to resolve ULVZ parameters in the lowermost mantle is demonstrated with simulated data. Then the inversion is applied to data sensitive to the CMB under the Philippine Sea. The same data have been previously considered and a two-layer ULVZ was inferred using forward modeling of a stack of more than 100 PcP-deconvolved ScP waves within the ScP Fresnel zone at 1 Hz

(spans ~100 km at the CMB) [*Idehara*, 2011]. However, some other studies [e.g., *Rost and Revenaugh*, 2001, 2003; *Rost et al.,* 2006] suggest that short-scale (~10s km), sub-Fresnel zone lateral variation in ULVZs exists. Moreover, significant variation of the ScP and SPdKS waves has been observed for slightly different CMB sampling locations [e.g., *Mori and Helmberger*, 1995; *Rost et al.,* 2006; *Jensen et al.,* 2013]. Therefore, *k*-means cluster analysis is applied here to identify clusters of similar waveforms. The cluster analysis suggests that two different types of waveforms exist and likely correspond to lateral variability in the lowermost mantle properties in a SW-NE direction. For the NE area, the BIC indicates a small density increase between ~4.5 and ~8.0 km above the CMB followed by a strong density anomaly between 0 and ~4.5 km above the CMB. In addition, a well-constrained, strong *S* wave velocity contrast exists between 0 and ~8.0 km but exhibits little change between two layers. From the partial melting and chemical origin perspective, these results suggest that the NE ULVZ could be due to both thermal (partial melting) and chemical origins (iron enrichment). The strong density anomaly in the lower ULVZ layer could be caused by the percolation of iron-enriched, molten material into the lowermost mantle. The results for the SW area are more difficult to interpret, which may be due to the limited number of data available for stacking in that bin and/or complex 2-D and 3-D structure that cannot be explained by the 1-D models required by our inversion approach. Therefore, this paper mainly focuses on the results for the NE region; detailed results for the SW region are given in the supporting information.

2. Inversion Method

2.1. Bayesian Inference

This section considers a brief introduction to Bayesian inversion, extensive treatment can be found elsewhere [*MacKay*, 2003; *Tarantola*, 2005; *Brooks et al.,* 2011]. In Bayesian inversion, model parameters are considered random variables constrained by data and prior information. For nonlinear problems, the solution to the inverse problem is given by approximations of the posterior probability density (PPD), which provides parameter estimates and uncertainties by marginalization. Conditional probabilities are indicated by |, for example, the conditional probability of **x** given **y** as $p(\mathbf{x}|\mathbf{y})$. For a vector of N measured **d** data and a vector of M model parameter **m**, Bayes' theorem is given by

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}, \tag{1}$$

where $p(\mathbf{m}|\mathbf{d})$ is the PPD, and $p(\mathbf{m})$ is the prior probability density. For measured data $\mathbf{d}^{\mathrm{obs}}$, $p(\mathbf{d}^{\mathrm{obs}}|\mathbf{m})$ can be interpreted as the likelihood function $L(\mathbf{m})$ (a function of only **m**). The denominator is a normalizing constant and referred to as Bayesian evidence (discussed later). The high-dimensional PPD combines prior and data information and is the solution to the inverse problem. However, for nonlinear problems, no analytic solutions exist and the PPD must be approximated by sampling methods such as Markov chain Monte Carlo (MCMC) (see section 2.2).

The prior constitutes independent information (derived from sources other than the observed data) about the model parameters and can be obtained from various sources, e.g., from previously published results. We chose uniform prior densities and wide bounds guided by previous studies so that the data predominantly estimate the solution.

The likelihood function of the model parameters introduces data information to the inversion and quantifies the likelihood that a particular parameter vector gave rise to the observed data. The likelihood function must be derived from an assumption about the statistical distribution form of the data errors. Note that data errors are generally not directly accessible in inverse problems and must be approximated by residual errors which combine both measurement (errors as a result of the measurement) and theory errors (errors resulting from approximations in the model including seismic wave propagation).

Here we assume Gaussian data errors and quantify error correlations by an autoregressive process of order 1 (AR1), for details, see *Dettmer et al.* [2012] and *Steininger et al.* [2013]. In this case, the likelihood function can be written as

$$L(\mathbf{m}) = \frac{1}{\sqrt{(2\pi)^{\mathrm{N}}\sigma^{\mathrm{N}}}} \exp\left[-\frac{1}{2\sigma^2}\left|\left(\mathbf{d}^{\mathrm{obs}} - \mathbf{d}(\mathbf{m}) - \mathbf{d}(\mathbf{a})\right)\right|^2\right], \tag{2}$$

where **d**(**a**) are AR (1) predictions that account for correlated errors, and **a** is the model vector containing the autoregressive parameters. In addition, the standard deviation of residual errors (σ) is treated as an unknown. For the case where the noise is treated as uncorrelated, **a** is zero and only σ is an unknown in the inversion.

2.2. MCMC and Interacting Markov Chains

For nonlinear inverse problems, no analytical solution exists for the PPD and numerical sampling is applied here to estimate the high-dimensional PPD. We apply the Metropolis-Hastings (MH) algorithm [*Metropolis et al.,* 1953; *Hastings*, 1970] to draw samples from the PPD by Markov chain Monte Carlo (MCMC) simulation. The MH algorithm simulates a Markov chain by proposing new states **m**′ (realizations of the model vector) based on the current state **m** and a random perturbation drawn from a proposal distribution. The proposed state is then accepted or rejected based on the MH criterion

$$\alpha = \min\left(1, \frac{p(\mathbf{m}')}{p(\mathbf{m})}\left(\frac{L(\mathbf{m}')}{L(\mathbf{m})}\right)^{\beta}\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}\right), \tag{3}$$

where β is a tempering parameter for sampling efficiency (discussed later) and q is the arbitrary but fixed proposal distribution to propose new states based on the current state. If the proposed model is accepted, the current state is updated. If the proposed model is rejected, the current model is retained and the process is repeated. The proposal distribution used here is a Cauchy distribution which has significantly heavier tails than a Gaussian distribution and can result in substantially improved sampling for difficult (e.g., nonlinear and multiple modes) problems due to higher probability of proposing large steps [*Dosso and Wilmut*, 2008].

For highly nonlinear inverse problems, such as waveform inversion, MH sampling can be extremely inefficient. Specifically, highly nonlinear problems can exhibit PPDs with multiple high-probability regions (posterior modes) that are separated by extensive low-probability regions. For these problems, most MCMC proposals perform poorly because large steps are required to transition between modes but such steps are infrequently proposed and rarely accepted. To address this problem, we apply parallel tempering (also referred to as interacting Markov chains). In parallel tempering, the PPD is augmented by a sequence of additional, tempered (raised to a power of <1) probability densities to aid in sampling from the PPD [*Geyer*, 1991; *Dettmer and Dosso*, 2012]. Each tempered density is targeted by a Markov chain, and the group of all chains is run in parallel on a computer cluster. Therefore, this approach requires higher computational effort but convergence is much more rapid for highly nonlinear problems. Here the sequence is chosen such that densities are increasingly tempered by defining the tempering parameter β [see equation (3)] between 0 and 1.

In this scheme, low β values increasingly deemphasize the likelihood function in equation (2), which results in more frequent acceptance of large steps (similar to the high-temperature phases in simulated annealing). Therefore, low chains explore the search space widely. Conversely, high chains (close to 1) preferentially sample high posterior probability regions. Chains provide unbiased samples from the PPD and are used for inferences on parameters of interest. To take advantage of the wide exploration of some chains, exchange updates are introduced, which allow the various Markov chains to exchange information without biasing PPD sampling. To perform an exchange update, random chain pairs are selected and parameter-vector swaps are proposed. Such updates are accepted/rejected based on the MH criterion for exchange updates [*Dettmer and Dosso*, 2012].

2.3. Model Selection

In geophysical inverse problems, the model parameterization is typically not known independently and must be estimated from the data. This process is referred to as model selection, where the goal is to determine the parsimonious (simplest) model that sufficiently explains the observed data. While model selection is of fundamental importance, the process can be extremely challenging and costly for highly nonlinear inverse problems. Model parameter uncertainties obtained from Bayesian inversions are intrinsically tied to the model itself. For example, under-parameterized models lead to simple structure but also produce unrealistically small parameter uncertainties due to theory error. A simple model can only access a limited part of the data space and therefore indicates small uncertainties for the predicted parameters. Furthermore, under-parameterization can lead to biased parameter estimates. On the other hand, overparameterizing the model can lead to spurious/unconstrained complex structure and unrealistically large uncertainties since the parameters are under-constrained by the data. Since the fit to data generally improves (increasing likelihood values) with

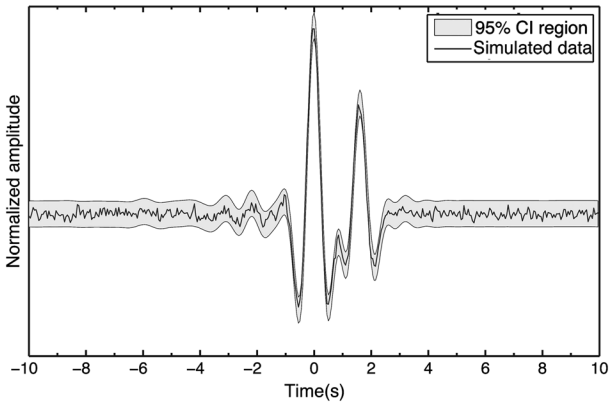

Figure 2. Simulated waveform and 95% confidence intervals computed from the PPD. Uncorrelated Gaussian noise of standard deviation $\sigma = 0.02$ is added to the deconvolved ScP waveform. Waveforms are normalized and centered on the maximum amplitude of the signal.

increase in model complexity, likelihood values alone are insufficient to estimate optimal model complexity.

Bayes' theorem provides a data-based objective criterion for model selection known as evidence, $p(\mathbf{d})$(see equation (1)). When considering multiple models, evidence represents the likelihood of a particular model given observed data **d**. Computation of evidence is by multidimensional integration over likelihood and prior and extremely costly [*Dettmer et al.,* 2009]. However, approximations based on point estimates can be applied in some cases. The Bayesian information criterion (BIC) [*Schwarz*, 1978] is one such approximation for linear problems, but has been successfully applied in some nonlinear cases. The BIC is given by

$$BIC = -2\log\left\{L\left(\mathbf{m}^{\mathrm{ML}}\right)\right\} + M \ \log \ N, \tag{4}$$

where M is the number of model parameters, N the number of data points, $\mathbf{m}^{\mathrm{ML}}$ the maximum likelihood model vector, and L is the likelihood. The first term in this expression quantifies how well the model fits the data while the second term constitutes a penalty term for model complexity. The model with the minimum BIC value is considered the optimal model in the group.

3. Forward Model and Parameterization

Bayesian inversion requires a large number of forward model evaluations (~10^5–10^6), and in typical geophysical inverse problems, these forward computations constitute the main computational cost. That cost limits the algorithm choice for computing synthetic seismograms, and we apply the computationally efficient Wentzel, Kramers, Brillouin, and Jeffreys (WKBJ) method to compute synthetic seismograms [*Chapman and Orcutt*, 1985]. While WKBJ does not provide a complete solution, it allows a straightforward way to isolate a specific seismic phase using input information that describes the ray interactions at discontinuities. The ak135 model [*Kennett et al.,* 1995] is used as a background model and modified by introducing layers on top of the CMB. We prefer to use the ak135 model because it provides a better fit to the PcP and ScP travel times than other available reference models.

In the WKBJ method, a high-frequency signal is synthesized for the medium consisting of radially varying homogeneous layers. In practice, the slowness range is limited either by numerical truncation or by complex slowness values. This discontinuous behavior of the amplitude terms causes endpoint errors on the WKBJ seismogram. However, PcP and ScP waves can be isolated from endpoint errors. In addition, second-order multiples of the ScP waves are included. Green's functions obtained from WKBJ are then convolved with the source time function to compute the final seismogram. The source time function is computed using an expression for the moment rate function presented by *Brüstle and Müller* [1983]. The final waveform for ULVZ structure is obtained by deconvolving the synthetic PcP wave from the ScP wave as shown in Figure 1c. The deconvolution of the PcP wave allows removing the source and raypath effects from the ScP wave. We assume that the PcP wave does not sample the ULVZ (Figure 1c) as the complex PcP waveforms that are suspected of sampling the ULVZ are not used in deconvolution.

Here ULVZs are parameterized by adding homogeneous, flat-lying layer(s) centered on the ScP-wave reflection point of the CMB. Such additional layers are introduced as perturbations from the ak135 model background to a maximum height of 35 km above the CMB. Each layer is parameterized in terms of interface position, *P* wave and *S* wave velocity perturbations, and density perturbation (with respect to the ak135 model). For example, a one-layer model consists of an interface position at some

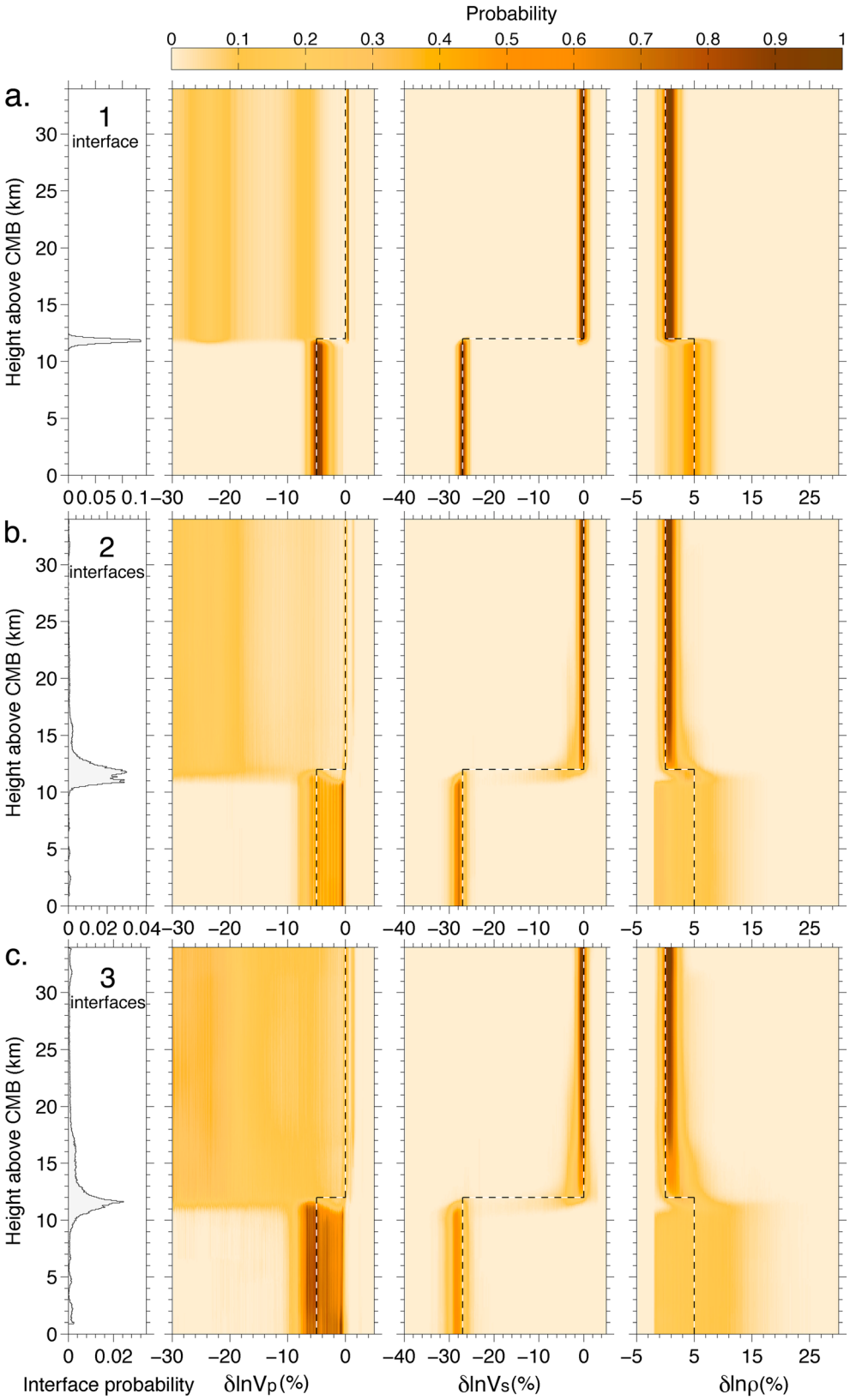

Figure 3. Inversion results for the simulation in terms of profile marginal probability densities for interface probability and perturbations of *P* wave velocity, *S* wave velocity, and density. Models parameterized with (a) one, (b) two, and (c) three interfaces are considered. True parameter values (dashed) are also shown.

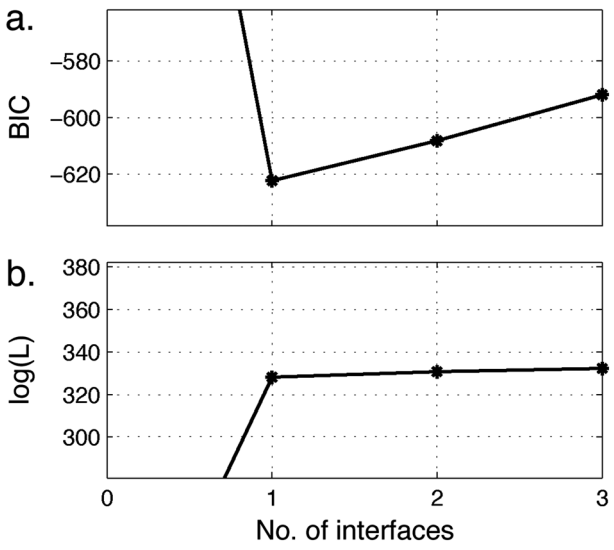

Figure 4. Model selection results in terms of (a) BIC values and (b) logarithmic likelihood values for the simulated data. Four models with zero to three interfaces are considered, and the BIC estimates the one-layer models as optimal. Axes are cropped for better visibility and do not show the values for the zero-layer model.

height above the CMB but below 35 km height and unknown elastic properties above and below that interface. Therefore, that model has seven unknown parameters.

4. Results and Discussion

4.1. Inversion Results for Simulated Data

We first present the results of simulated experiments in which we demonstrate the feasibility of a hierarchical Bayesian inversion and model selection for ULVZ studies. A synthetic, PcP-deconvolved ScP waveform (shown in Figure 2) is prepared for a single-layer ULVZ with a thickness of 11.5 km sitting on the top of the CMB. This layer has two important features suggested in past ULVZ studies beneath the Philippine Islands: a sharp decrease in *P* and *S* wave velocity by 5% and 27%, respectively, and an increase in density by 5% with respect to the background ak135 model [*Idehara*, 2011]. The decrease in velocities is attributed to the presence of a partial melt with a slightly different ratio of *P* and *S* wave velocity perturbation than the typical 1:3 ratio [*Williams and Garnero*, 1996].

Uncorrelated Gaussian random noise with a standard deviation of $\sigma = 0.02$ m/s (i.e., 2% of the maximum amplitude of the waveform) to the synthetic waveform is added. While the measured data show serial correlation in the residual errors, we focus on uncorrelated Gaussian random errors in the simulation to examine the parameter uncertainties for simple noise characteristics (an ideal case). Note that the inversion for observed data includes a hierarchical error model that also addresses serial correlation in the residual errors.

The priors for the inversion are chosen such that the vertical extent of the layered ULVZ can be between 1 and 34 km above the CMB, the change in *P* wave velocity between −30 to 4%, the change in *S* wave velocity from −40 to 4%, and the change in density from −3% to 30%. Ten Markov chains are generated at different temperatures distributed according to a geometric series between 1 and 0. Note that the inversion was also carried out without parallel tempering, but results did not converge to a meaningful solution within reasonable time. Rather, the sampling algorithm was trapped in modes for long periods appearing to have converged to solutions substantially different from the true model. By applying parallel tempering, convergence to the true model was found to be much more rapid. In order to simulate ULVZs with a larger number of layers, the inversion is repeated varying the number of interfaces from 0 to 3 for the same synthetic data. The first third of the samples were discarded as burn-in (ensuring the sample is not biased by the starting point). Samples after the burn-in period are considered for parameter inference. Convergence of the algorithm is monitored by examining MCMC sampling history for various parameters.

Figure 3 shows inversion results in terms of marginal profile densities for the various ULVZ parameters in terms of the height above the CMB. Profile marginal densities are obtained by considering only the parameter of interest as a function of height while integrating out all other parameters in the model. In this process, the height above CMB is required to be discretized and we choose to normalize each depth interval individually to unit area. Figure 3 considers three inversion results for models ranging from 1 to 3 interfaces within the ULVZ. Note that uncertainties generally increase with increasing model complexity. In particular, the model with three interfaces exhibits large uncertainties at 10 km height and significant spurious structure, which is consistent with its overparameterized nature. In addition, the *P* wave velocity uncertainty above, the ULVZ is always large, indicating insensitivity of the data to *P* wave velocity changes.

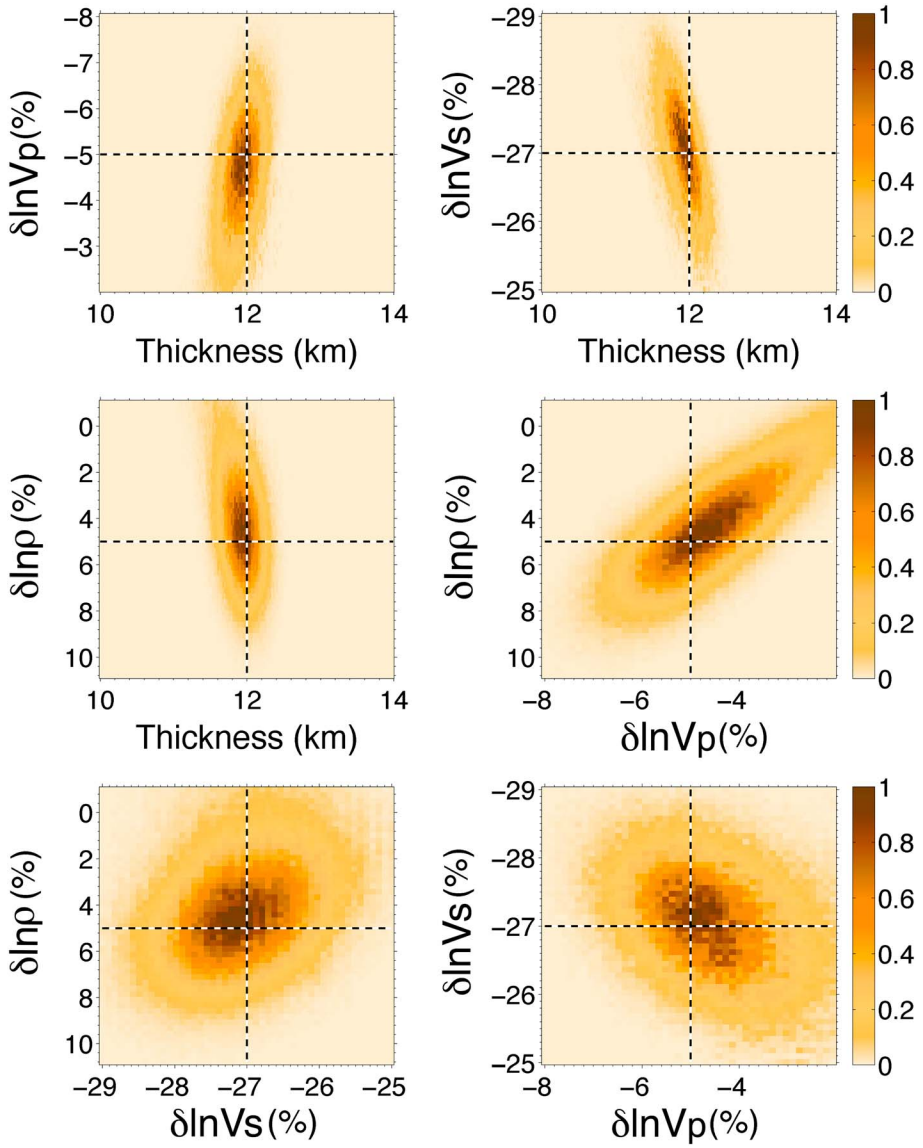

Figure 5. Joint marginal densities of selected parameter pairs for the one-layer model. Dashed lines indicate true model parameters.

The BIC is applied to carry out model selection (Figure 4), and indicates that the single-interface model is optimal. Note that the BIC result coincides with the true model parameterization while likelihood values consistently increase with the number of interfaces (i. e., are insufficient for model selection).

Interparameter relationship for selected parameter pairs of the optimal model (one layer) are presented in Figure 5 by joint marginal densities. The joint marginals are unimodal, and the parameters are recovered well. However, a strong trade-off between parameters exists. In particular, the thickness is strongly correlated with the change in *S* wave velocity. The *P* wave velocity and density are also correlated with each other, while the rest of the parameters are weakly correlated. For example, the increase in thickness is positively correlated with the decrease in *S* wave velocity perturbation. This means that an increase in thickness and a decrease in *S* wave velocity perturbation have similar effects on the position of precursor and postcursor. Similar modeling trade-offs have been suggested for SPdKS waveform in previous ULVZ studies [*Garnero and Helmberger*, 1998; *Garnero and Jeanloz*, 2000]. As a further example, a decrease in *P* wave velocity perturbation is positively correlated with the density perturbation because the *P* wave velocity impedance contrast changes the amplitudes of both precursor and postcursor.

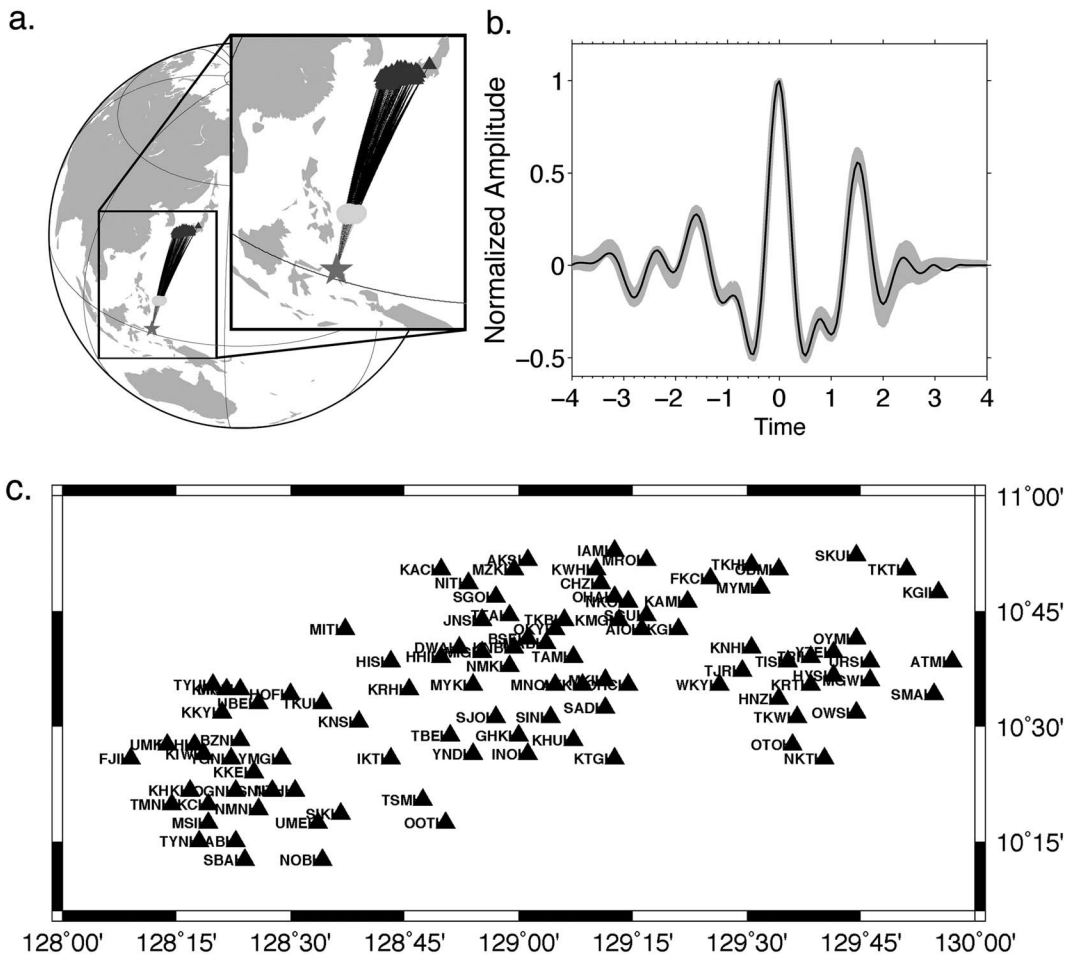

Figure 6. The lowermost mantle region studied here. (a) The Indonesia earthquake location (event id 2003050515, star), surface projections of ScP ray paths (black lines), and Japanese Hi-Net Array stations (triangles) are shown. Reflection points at the CMB (gray circles) are also shown. (b) Observed, stacked (for all stations) and PcP deconvolved ScP waves (black), and range of data predictions for the PPD (gray-shaded area). (c) Projections of station locations at the CMB (triangles) investigated in this study.

4.2. Inversion Results for Observed Data

This section applies the inversion to observed seismic waveforms sensitive to the CMB beneath the east of the Philippine Sea. Figure 6 shows source location, station geometry, and the stacked waveform taken from *Idehara* [2011]. Core-reflected waves (PcP and ScP) are collected from the Hi-net array, which consists of more than 600 borehole short-period stations [*Okada et al.,* 2004].

In our data processing, we test Butterworth filters with different width and corner frequencies and find that a causal band-pass filter with corner frequencies of 0.5 and 1.5 Hz is the optimal filter. We note that we apply filter to relatively long time window (±30 s from the main arrival) so that the central part of the signal remains unaffected. Waveforms are then aligned with peak amplitudes of ScP and PcP waves separately. PcP waves are stacked and the stack is deconvolved from individual ScP waves to remove source and path effects. Note that the deconvolution of *P* from PcP waves would be ideal to eliminate near-source effects as both waves have the same source-radiation pattern. However, the *P* and PcP waves traverse significantly different mantle paths for the source-receiver pair used in this study (Figure 1a). The deconvolved ScP waves are then stacked to enhance the lowermost mantle features. The raypath geometry and CMB reflection points are shown in Figure 6a. The preprocessed and deconvolved ScP waveform (Figure 6b) was previously studied and a two-layer ULVZ model was inferred by forward waveform modeling [*Idehara*, 2011]. *Idehara* [2011] stacked ScP waveforms for a region of piercing points spanning ~2° at the CMB, which is

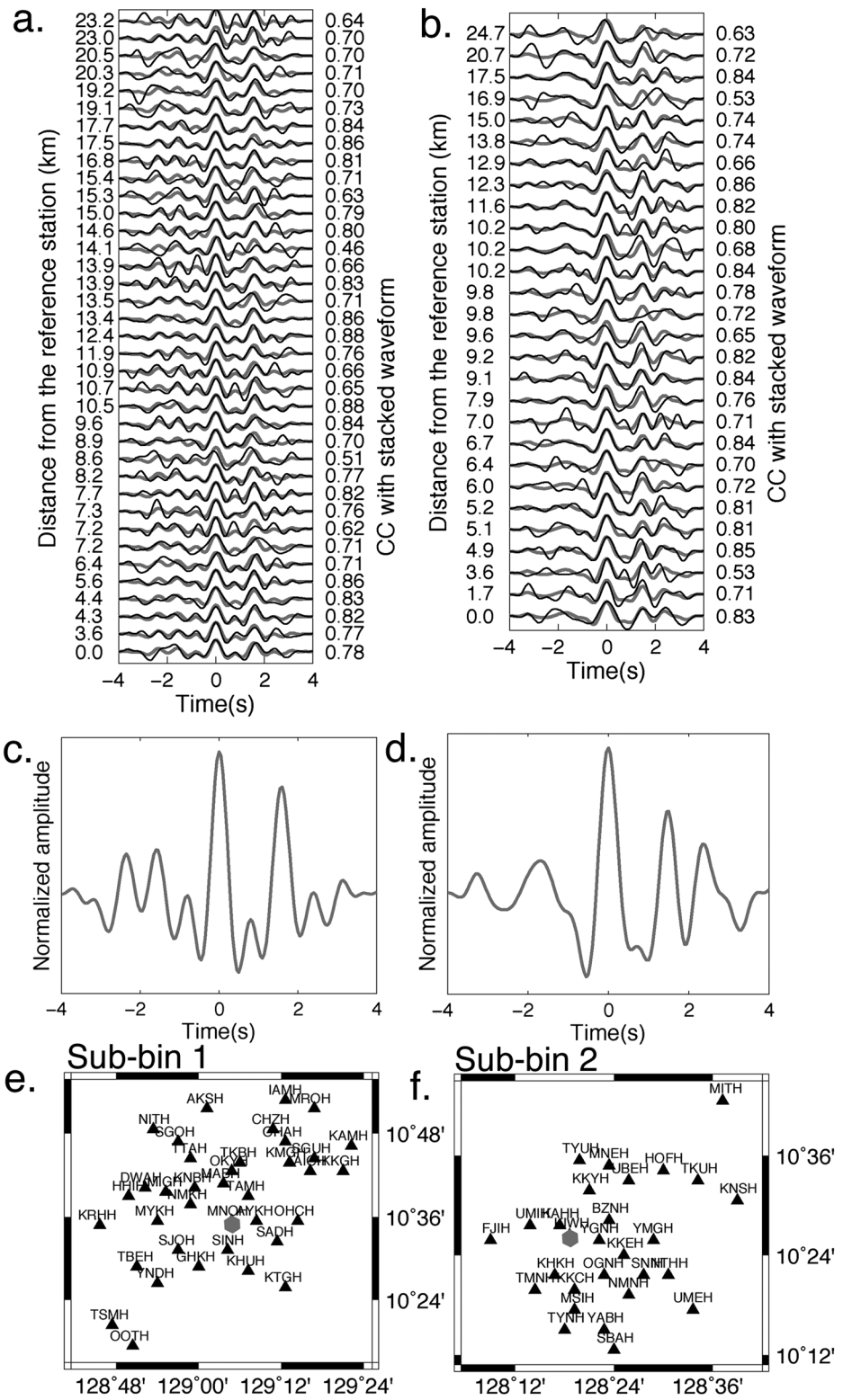

Figure 7. Cluster-analysis results based on cross-correlation coefficients. (a) Observed PcP-deconvolved ScP waves of the NE region (sub-bin 1) and (b) SW region (sub-bin 2) of Figure 6c. The distance of individual CMB piercing points from the cluster centroid (represented by hexagon shape in Figures 7e and 7f) (left axis) and cross-correlation coefficients between stacked and individual waves (right axis) are shown. Gray lines represent waveform stacks of the NE and SW regions. (c, d) Detailed views of stacks. Station projections at the CMB bounce points used to stack waves in the (e) NE and (f) SW regions are also shown.

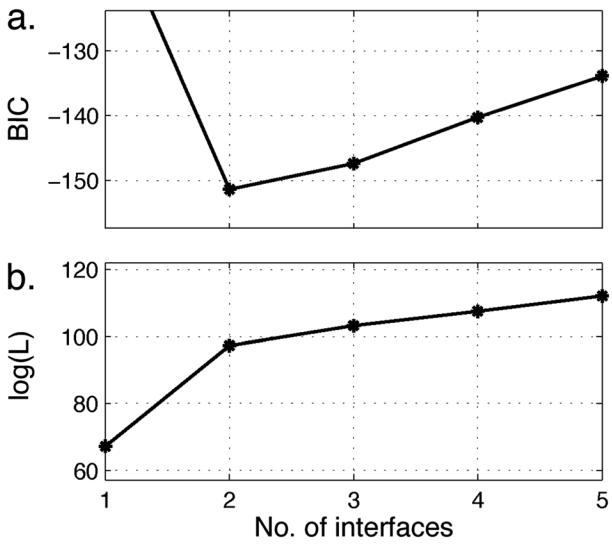

Figure 8. Model selection results for observed data (east of the Philippines) in terms of (a) BIC values and (b) logarithmic likelihood values for the observed data in the NE region (sub-bin 1). Five models with one to five interfaces are considered, and the BIC estimates the two-layer model as optimal. Axes are cropped for better visibility and do not show the BIC value for the one-layer model.

approximately equal to the Fresnel zone at 1 Hz. If a Fresnel zone is considered a zone without discernible lateral changes, a single-stacked waveform can be inverted with an infinite frequency approximation. However, some studies [e.g., *Rost and Revenaugh*, 2001, 2003; *Rost et al.,* 2006; *Idehara et al.,* 2007] suggest that short-scale (~10s km), sub-Fresnel zone lateral variation of ULVZs exists. Furthermore, significant variation of the ScP and SPdKS waves has been observed for slightly different CMB sampling locations [e.g., *Mori and Helmberger*, 1995; *Rost et al.,* 2006; *Jensen et al.,* 2013]. Therefore, *k*-means cluster analysis [*MacQueen*, 1967; *Hartigan*, 1975; *Hartigan and Wong*, 1979] is applied here to identify clusters of similar waveforms. As a distance metric, cross-correlation coefficients are applied to quantify waveform similarity. The *k*-means clustering algorithm minimizes the sum of squared distances between cluster centroid and other waveforms for *k* clusters. Hence, waves of similar cross-correlation coefficients are grouped in the same cluster, while the significantly different waveforms are grouped in separate clusters. A detailed description for the clustering procedure applied to the Philippine Sea data is presented in the supporting information. We find that the data can be divided into two clusters with centroid locations in two separate geographic locations (Figure 7). Note that there are few waveforms with piercing points outside the main clusters that have high cross-correlation coefficients with waveforms within the main clusters, and those waveforms are not considered in the stack. This analysis leaves 37 waveforms for the NE cluster (subbin 1) and 28 for the SW cluster (subbin 2). The individual PcP-deconvolved ScP waves of the two clusters are compared to the stacked waveforms in Figures 7a and 7b, detailed views of the stacked waveforms are also shown in Figures 7c and 7d. The individual waveforms are sorted according to the distance from the reference station at the cluster centroid (see Figure 7). The waveforms in the two clusters show markedly different precursors, postcursors, and frequency content. In subbin 1, higher-frequency content is evident compared to subbin 2. In addition, the precursors of subbin 1 are more consistently present in the waveforms. Postcursors are consistently present in both clusters.

The Bayesian inversion was applied to the stacked waveforms of clusters one and two. While results for subbin 1 are straightforward to interpret, results for subbin 2 are more challenging. This section shows detailed results for subbin 1. The challenging inversion results for subbin 2 preclude interpretation with a 1-D model and are discussed in detail in section 4.3. The complete inversion results for subbin 2 are presented in the supporting information.

Section 4.1 showed that the inversion is largely insensitive to *P* wave velocity perturbations in the top layer (Figure 3). Therefore, priors for the topmost part are set to be narrow and centered around the ak135 background (±2% in *P* and *S* wave velocity perturbations, and ±3% in density perturbations). Other priors are identical with axis bounds of figures for posterior marginal distributions. The model parameters are first optimized by running the inversion with 10 parallel MCMC chains on each of 24 computer cores. For each forward computation, the waveform around the ScP wave is computed including precursors, postcursors, and multiples up to the second order, which is then filtered in the same way as the observed data. This waveform is then deconvolved with the precomputed PcP wave for the same source and receiver. A spectral division deconvolution is stabilized by adding a term in the denominator. This term is also treated as unknown in the inversion and estimated by the data. After the deconvolution, we do not apply filter, which is

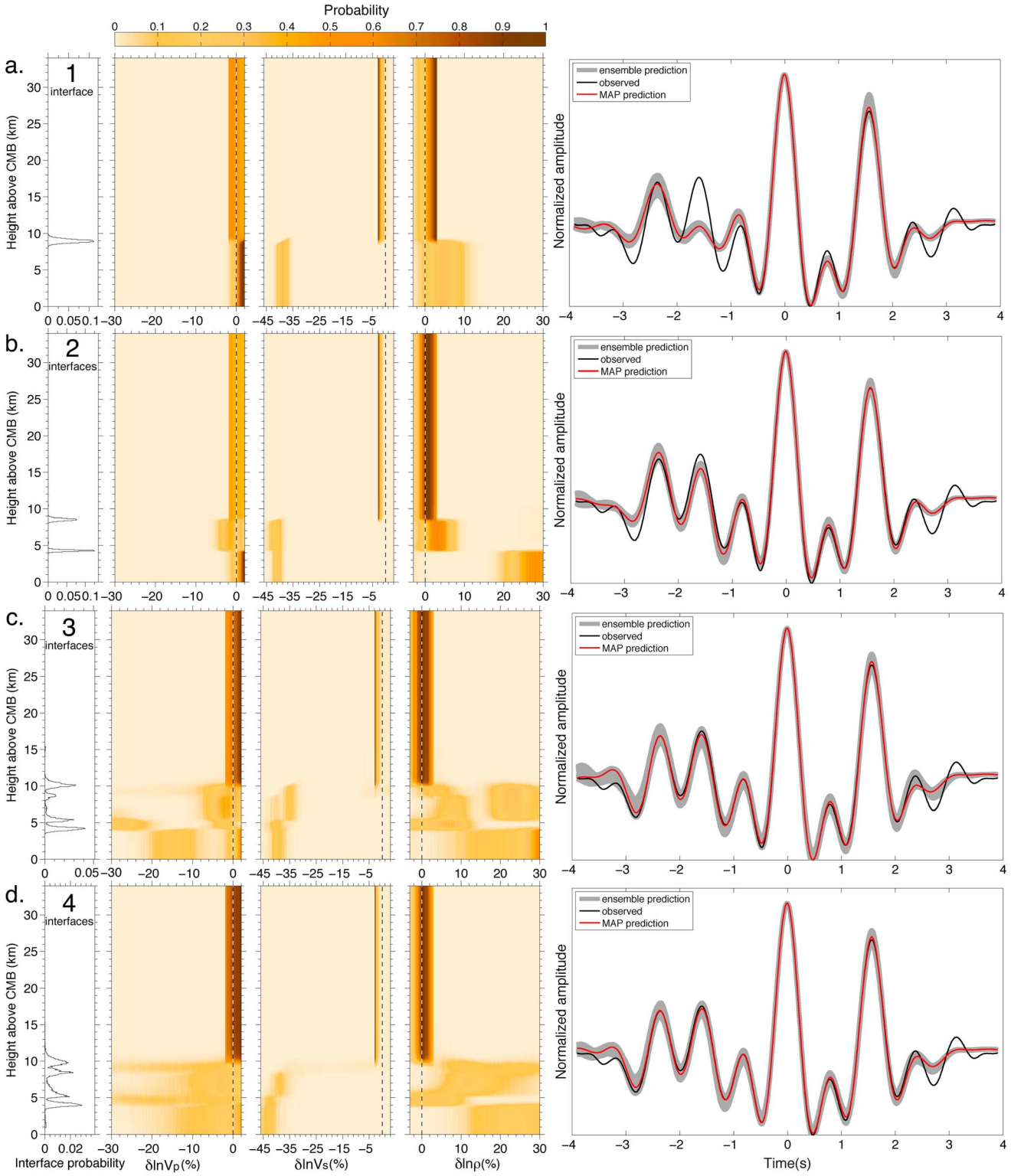

Figure 9. Sub-bin 1 (NE region) inversion results in terms of profile marginal probability densities for interface probability, and perturbations of *P* wave velocity, *S* wave velocity, and density. Models parameterized with (a) one, (b) two, (c) three, and (d) four interfaces are shown. Background ak135 values (dashed) are also shown. Probability densities are normalized to unit area at each depth. Panels on the right show the waveforms used in the inversion (black), maximum a posteriori (MAP) prediction (red), and ensemble of model prediction (gray shading) in the background. Note: Inversions were performed for up to five interfaces, five-layer results are not shown here because the fit to the data did not significantly improve over the four-layer inversion. Priors are uniform over the plot bounds.

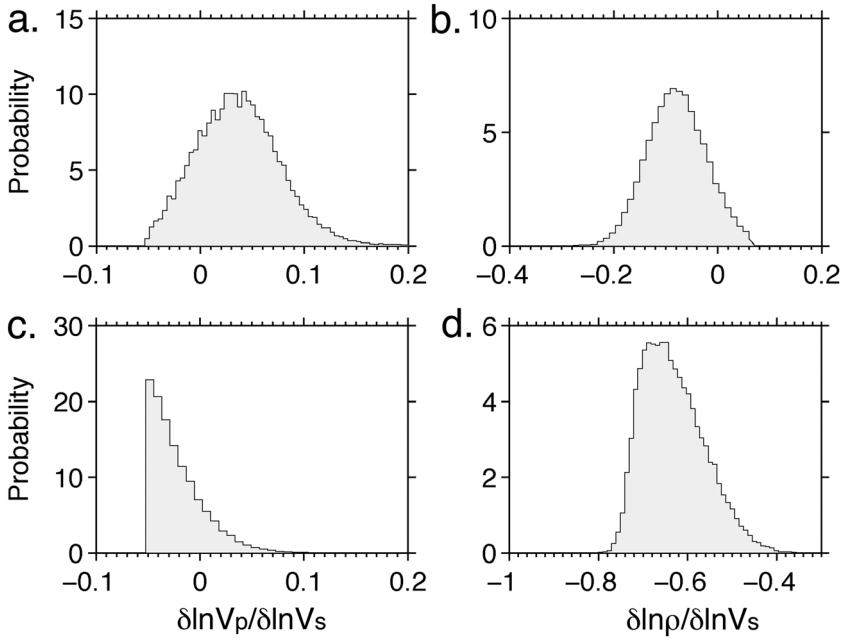

Figure 10. Marginal densities for the ratio of *P* to *S* wave velocity perturbations and density to *S* wave velocity perturbations for sub-bin 1 (Figure 7c). Ratios for the (a, b) top layer and (c, d) bottom layer. Note that these ratios are not directly sampled in the inversion but inferred from the PPD.

consistent with the processing of the observed waveforms. The predicted and observed waveforms are crosscorelated and aligned. The likelihood is computed in a ±4 s time window around the central peak of the deconvolved ScP waveform. Inversions are carried out for five parameterizations from one to five interfaces, and model selection via BIC is applied within that group of models.

Figure 8 shows the BIC results and indicates the two-layer model as optimal for subbin 1. Figure 9 shows inversion results in terms of marginal profile densities for the ULVZ parameters for the one- to four-layer models considered in the BIC computation. Note that the inversion already fits large features in the data (peaks with significant amplitude) with the simple parameterizations. However, with increasing model complexity, fit to the data substantially improves and multiple precursor and postcursor are increasingly fit. This is particularly visible for the two-layer model when an additional interface is introduced at ~4.5 km height above the CMB. For the three- and four-layer models, the misfit decreases less by fitting less significant

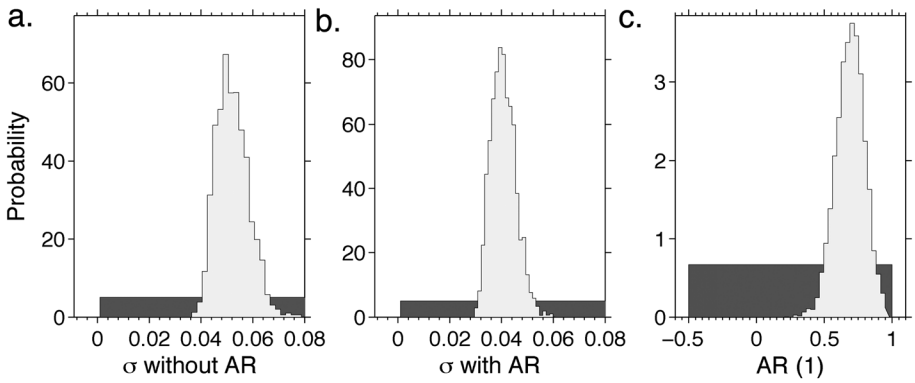

Figure 11. Marginal densities of error parameters for an inversion ignoring correlated errors and for an inversion accounting for correlated errors by including an AR (1) process in the case of sub-bin 1 (Figure 7c). (a) Standard deviation when ignoring correlations, (b) standard deviation when accounting for correlated errors, and (c) AR (1) parameter when accounting for correlated errors.

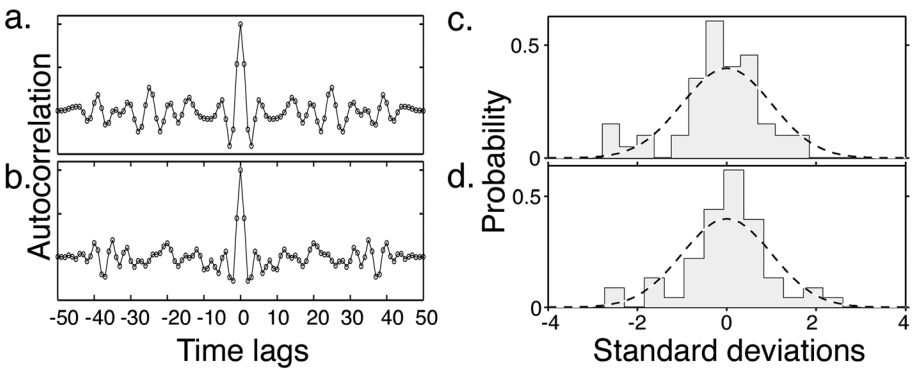

Figure 12. Autocorrelation functions of residuals assuming (a) uncorrelated noise, and (b) correlated noise with AR (1) model in the likelihood (sub-bin 1, Figure 7c). Histograms of the data residuals as a qualitative measure of the Gaussianity of data errors assuming (c) uncorrelated data errors and (d) correlated data errors AR (1). The black dash line on top of the histogram is the theoretical Gaussian distribution.

amplitude features. While the three-layer inversion improves fit to the data, the improvement is not significant enough to justify the additional layer (based on the BIC) and leads to overparameterization.

The ratio of the *P* to *S* wave velocity anomalies ($R_{P/S}$) and density to *S* wave velocity anomalies ($R_{\rho/S}$) are often considered as key factors to quantify the relative contributions of thermal (partial melt) and chemical (iron enrichment) heterogeneity to the observed seismic velocities in the lowermost mantle. In the case of thermal contribution, ($R_{P/S}$) can be ~0.3 and ($R_{\rho/S}$) is close to zero. In contrast, the value of ($R_{P/S}$) is smaller than ~0.3 and the negative value of ($R_{\rho/S}$) requires an additional contribution possibly from the chemical heterogeneity. Figure 10 shows the marginal densities of the ($R_{P/S}$) and ($R_{\rho/S}$) for both layers. The negative value of ($R_{\rho/S}$) (i.e., from 0 to ~ −1.0 due to increased density perturbation of up to 30% and decreased *S* wave velocity perturbation of up to 40%) and smaller value of ($R_{P/S}$) (i.e., less than ~0.2 due to the insignificant change in *P* wave velocity) suggests that both partial melting and iron enrichment are required mechanisms for the physical origin of the ULVZ beneath the east of the Philippines. *Idehara et al.* [2007] has previously observed one-layer ULVZ with an insignificant change in *P* wave velocity (~0%) and strong drop in *S* wave velocity (~40%) in the southwestern part beneath the Philippine-Kalimantan region.

Figure 11 shows marginal densities for the error standard deviation (obtained from inversions with and without AR (1)) and the AR (1) parameter. Note that the standard deviation is much higher when ignoring correlations. This is likely due to some aspects of correlated errors being accounted for by a higher standard deviation. The high values of the AR (1) parameter (up to 0.99) indicate strongly correlated residuals, which are likely due to the data processing (filtering and deconvolution).

To examine the effect of including the AR (1) process in the inversion, residual error analysis is briefly considered here. The residuals are obtained for the maximum a posteriori (MAP) parameter vector of the optimal model (two layers). Inversion results that include the AR (1) process are compared to results ignoring correlated errors (Figure S3). The inversion result significantly changes when we include the AR (1) process. In particular, the density perturbation becomes more physically meaningful and increases toward the bottom of the ULVZ (near the CMB). Randomness is examined by considering the residual autocorrelation function, and Gaussianity is examined by comparing residual histograms to Gaussian probability densities. Figure 12 shows the residual autocorrelation function. When correlated errors are ignored, the autocorrelation shows a wide central peak, indicating dependent errors. When accounting for correlations with the AR (1) model, peak width decreases significantly, indicating that correlations are reasonably accounted for in the inversion. Figure 12 also shows histograms of data residuals for inversions with and without AR (1), and theoretical Gaussian densities are superimposed. Both histograms appear to be reasonably Gaussian distributed. In summary, the residuals indicate that assumptions were reasonably met by the inversion including an AR (1) process, which increases our confidence in the inversion results.

4.3. Discussion of Inversion Results

The physical origin of the velocity reduction in ULVZ regions can be explained by two fundamental causes: an increase in density (e.g., due to a change in the chemistry) and a decrease in elasticity (e.g., due to a

change in the rigidity of the material). *Christensen and Hofmann* [1994] suggest that the enrichment of mid-oceanic ridge basaltic material subducted through the slab slows down the velocity of seismic waves in the lowermost mantle. Similarly, the banded form of iron due to the reaction with atmospheric oxygen, subducted through the slab, can cause the existence of a thin chemically distinct layer on the top of the CMB [*Dobson and Brodholt*, 2005]. A challenging question arises as to how such a thin (~1 km) laminated layer of iron would separate itself from the mantle. More recently *Mao et al.* [2006] suggest that iron-rich postperovskite could account for the anomalous properties of ULVZs. Similarly, a thin layer of magnesium postperovskite (enriched with iron) [e.g., *Wicks et al.,* 2010] can reduce the *S* wave velocity up to 55% at lowermost mantle pressure conditions (~121 GPa). However, the origin of iron in perovskite mineral and its stability at the base of the mantle is highly debated [*Hernlund et al.,* 2005; *Sakai et al.,* 2006; *Otsuka and Karato*, 2012].

On the other hand, the reduction of seismic velocity in the ULVZ is argued to stem from partial melting. The existence of temperatures at the CMB higher than the solidus temperature makes this argument more plausible than the aforementioned mechanisms. While partial melting supports thermal and density anomalies, an issue remains about the complex geometry of the melting [*Hernlund and Tackley*, 2007] and the uncertainty of the temperature at the CMB. The decrease in *P* wave velocity of up to 10% is explained by a fraction of 5–30% of melt depending on the melt geometry [e.g., *Williams and Garnero*, 1996; *Labrosse et al.,* 2007; *Mosenfelder et al.,* 2009].

The optimal (two-layer) model for subbin 1 has density structure increasing with depth (up to ~30%) and a decrease in *S* wave velocity of up to ~40%. The inferred two-layer ULVZ can be explained by a combination of partial melting and iron enrichment in the lowermost mantle. Several mechanisms have been suggested for iron enrichment in ULVZs: Subduction of mid-oceanic ridge basalt [*Christensen and Hofmann*, 1994], subduction of banded form of iron [*Dobson and Brodholt*, 2005], and postperovskite transition [e.g., *Mao et al.,* 2006; *Wicks et al.,* 2010]. However, the morphological instability suggested by *Otsuka and Karato* [2012] appears to be the most feasible mechanism as discussed in the beginning of section 4.3. According to this mechanism, the FeO depleted from the D″ region thickens the FeO enriched material at the top of the core when the chemical equilibrium is achieved between the initial FeO in the mantle and core. The inversion results suggest an increase in density of up to ~30% in the bottom layer. This increase can only explain an *S* wave velocity decrease of up to ~15% assuming that the increase in density does not alter the shear modulus at the lowermost mantle conditions. Thus, the presence of iron-enriched material alone cannot explain the strong decrease in *S* wave velocity of up to 40% inferred in our inversion results. Therefore, melting of iron-enriched material at the CMB temperature condition is required and our results suggest that the dense- and melt-iron-enriched material drains down to the bottom of the ULVZ due to a percolation effect [*Hernlund and Tackley*, 2007].

The results for the SW area (subbin 2) are more difficult to interpret. Inversions are carried out for five parameterizations from one to five interfaces, and model selection via BIC is applied within that group of models. The model selection study (Figure S1) shows significantly decreasing misfit with increasing model complexity for up to five layers. Therefore, the BIC study selects the most complex model considered in the study (five layers) as optimal. However, while the misfit for this model is much better than for the simpler models, the marginal profile densities (Figure S2) appear physically unreasonable with spurious structure and complicated layering. The increasing model complexity causes substantially improved fit of multiple precursor and postcursor, but these improvements are achieved with unphysical structure. Since the waveform selection for this study includes the rejection of any complex PcP waves [*Idehara*, 2011], source and receiver/site effects can be precluded. Therefore, the strong secondary peaks visible in these data unlikely represent source or receiver effects, or contamination by other phases. We conclude that these peaks can result from more complex (i.e., 2-D or 3-D) nature of the ULVZ [*Wen and Helmberger*, 1998a, 1998b; *To et al.,* 2011; *Thorne et al.,* 2013; *Jensen et al.,* 2013]. While extension to 2-D and/or 3-D models is desirable, computational limitations in the Bayesian sampling approach currently prevent such generalizations and are not considered in this work.

5. Conclusion

This paper developed a hierarchical Bayesian inversion including parallel tempering for ScP phases sensitive to the CMB to rigorously quantify ULVZ parameters and uncertainties. Uncertainty quantification is particularly important for ULVZ inversion due to strong nonlinearity and nonuniqueness. Since the

optimal parameterization for the ULVZ is not known independently, the Bayesian information criterion was applied to estimate the optimal model from the data. The approach was first applied to simulated data to study the data information content of ScP phases. Then, the inversion was applied to data recorded by the Hi-Net array in Japan that sample the CMB beneath the east of the Philippines. To account for strongly correlated errors, a first-order autoregressive error model was applied and shown to successfully account for the strong error correlations, which resulted in physically more meaningful results. The BIC identifies a two-layer model as optimal for the NE region. This model exhibits a strong positive density anomaly and a strong negative *S* wave velocity anomaly. The lowermost layer is dense, gravitationally stable, and can be explained by drained, dense, and melt-rich iron material near the CMB. The ratios of the *P* wave velocity to *S* wave velocity perturbations and *S* wave velocity to density perturbations further suggest that both iron enrichment and partial melt are required to explain the observed waveform that samples the NE region.

Acknowledgments
We would like to thank Koki Idehara for providing the data from Hi-Net array and two anonymous reviewers for constructive comments that improved the manuscript substantially. Computations were performed on the Terrawulf II and III cluster, a computational facility supported through the AuScope Australian Geophysical Observing System (AGOS). AusScope is funded under the National Collaborative Research Infrastructure Strategy (NCRIS) and the Education Investment Fund (EIF3), both Australian Commonwealth Government Programs.

References

Agostinetti, P. N., and A. Malinverno (2010), Receiver function inversion by trans-dimensional Monte Carlo sampling, *Geophys. J. Int.*, *181*, 858–872, doi:10.1111/j.1365-246X.2010.04530.x.

Bataille, K., and F. Lund (1996), Strong scattering of short-period seismic waves by the core-mantle boundary and the *P*-diffracted wave, *Geophys. Res. Lett.*, *23*, 2413–2416, doi:10.1029/96GL02225.

Bodin, T., and M. Sambridge (2009), Seismic tomography with the reversible jump algorithm, *Geophys. J. Int.*, *178*, 1411–1436, doi:10.1111/j.1365-246X.2009.04226.x.

Bréger, L., and B. Romanowicz (1998), Three-dimensional structure at the base of the mantle beneath the central Pacific, *Science*, *282*(5389), 718–720, doi:10.1126/science.282.5389.718.

Bréger, L., H. Tkalčić, and B. Romanowicz (2000), The effect of D″ on PKP (AB-DF) travel time residuals and possible implications for inner core structure, *Earth Planet. Sci. Lett.*, *175*, 133–143.

Brooks, S., A. Gelman, G. L. Jones, and X.-L. Meng (Eds.) (2011), *Handbook of Markov Chain Monte Carlo*, Chapman and Hall/CRC Press, Boca Raton, Fla.

Brüstle, W., and G. Müller (1983), Moment and duration of shallow earthquakes from Love-wave modeling for regional distances, *Phys. Earth Planet. Inter.*, *32*, 312–324.

Chapman, C. H., and J. A. Orcutt (1985), The computation of body wave seismograms in laterally homogeneous media, *Rev. Geophys.*, *23*, 105–163, doi:10.1029/RG023i002p00105.

Christensen, U. R., and A. W. Hofmann (1994), Segregation of subducted oceanic crust in the convecting mantle, *J. Geophys. Res.*, *99*, 19,867–19,884, doi:10.1029/93JB03403.

Dettmer, J., and S. E. Dosso (2012), Trans-dimensional matched-field geoacoustic inversion with hierarchical error models and interacting Markov chains, *J. Acoust. Soc. Am.*, *132*(4), 2239–2250.

Dettmer, J., S. E. Dosso, and C. W. Holland (2009), Model selection and Bayesian inference for high-resolution seabed reflection inversion, *J. Acoust. Soc. Am.*, *125*(2), 706–715, doi:10.1121/1.3056553.

Dettmer, J., S. E. Dosso, and C. W. Holland (2010a), Trans-dimensional geoacoustic inversion, *J. Acoust. Soc. Am.*, *128*, 3393–3405.

Dettmer, J., S. E. Dosso, and J. C. Osler (2010b), Bayesian evidence computation for model selection in geoacoustic inversion, *J. Acoust. Soc. Am.*, *128*(6), 3406–3415, doi:10.1121/1.3506345.

Dettmer, J., S. Molnar, G. Steininger, S. E. Dosso, and J. F. Cassidy (2012), Trans-dimensional inversion of micrometer array dispersion data with hierarchical autoregressive error models, *Geophys. J. Int.*, *188*, 719–734, doi:10.1111/j.1365-246X.2011.05302.x.

Dobson, D. P., and J. P. Brodholt (2005), Subducted banded iron formations as a source of ultralow-velocity zones at the core-mantle boundary, *Nature*, *434*, 371–374.

Doornbos, D. J. (1974), Seismic wave scattering near caustics: Observation of PKKP precursors, *Nature*, *247*, 352–353.

Dosso, S. E., and M. J. Wilmut (2008), Uncertainty estimation in simultaneous Bayesian tracking and environmental inversion, *J. Acoust. Soc. Am.*, *124*(82), 82–97, doi:10.1121/1.2918244.

Fukao, Y. (1992), Seismic tomogram of the Earth's mantle: Geodynamic implications, *Science*, *258*, 625–630.

Fukao, Y., A. To, and M. Obayashi (2003), Whole mantle *P* wave tomography using *P* and PP-P data, *J. Geophys. Res.*, *108*(B1), 2021, doi:10.1029/2001JB000989.

Garnero, E. J., and D. V. Helmberger (1995), A very slow basal layer underlying large-scale low velocity anomalies in the lower mantle beneath the Pacific: Evidence from core phases, *Phys. Earth Planet. Inter.*, *91*, 161–176.

Garnero, E. J., and D. V. Helmberger (1998), Further structural constraints and uncertainties of a thin laterally varying ultralow-velocity layer at the base of the mantle, *J. Geophys. Res.*, *103*, 12,495–12,509, doi:10.1029/98JB00700.

Garnero, E. J., and R. Jeanloz (2000), Fuzzy patches on the Earth's core-mantle boundary?, *Geophys. Res. Lett.*, *27*, 2777–2780, doi:10.1029/2000GL008498.

Garnero, E. J., and J. E. Vidale (1999), ScP; A probe of ultralow velocity zones at the base of the mantle, *J. Geophys. Res.*, *26*, 377–380.

Garnero, E. J., D. V. Helmberger, and S. P. Grand (1993), Constraining outermost core velocity with SmKS waves, *Geophys. Res. Lett.*, *20*, 2463–2466, doi:10.1029/93GL02823.

Geyer, C. J. (1991), Markov chain Monte Carlo maximum likelihood, in *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, pp. 156–163, Interface Foundation, Fairfax Station, Va.

Haddon, R. A. W., and J. R. Cleary (1974), Evidence for scattering of seismic PKP waves near the mantle core boundary, *Phys. Earth Planet. Inter.*, *8*(3), 211–234.

Hartigan, J. A. (1975), *Clustering Algorithms (Probability & Mathematical Statistics)*, John Wiley, New York.

Hartigan, J. A., and M. A. Wong (1979), A *k*-means clustering algorithm, *J. R. Stat. Soc., Ser. C*, *28*(1), 100–108.

Hastings, W. K. (1970), Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, *57*(1), 97–109.

Helmberger, D., S. Ni, L. Wen, and J. Ritsema (2000), Seismic evidence for ultralow-velocity zones beneath Africa and eastern Atlantic, *J. Geophys. Res.*, *105*, 23,865–23,878, doi:10.1029/2000JB900143.

Helmberger, D. V., E. J. Garnero, and X. Ding (1996), Modelling two-dimensional structure at the core-mantle boundary, *J. Geophys. Res.*, *101*, 13,963–13,972, doi:10.1029/96JB00534.
Hernlund, J. W., and A. M. Jellinek (2010), Dynamics and structure of a stirred partially molten ultralow-velocity zone, *Earth Planet. Sci. Lett.*, *296*, 1–8.
Hernlund, J. W., and P. J. Tackley (2007), Some dynamical consequences of partial melting in Earth's deep mantle, *Phys. Earth Planet. Inter.*, *162*, 149–163.
Hernlund, J. W., C. Thomas, and J. Tackley (2005), A doubling of the post-perovskite phase boundary and structure of the Earth's lowermost mantle, *Nature*, *434*, 882–886.
Houser, C., G. Masters, P. Shearer, and G. Laske (2008), Shear and compressional velocity models of the mantle from cluster analysis of long-period waveforms, *Geophys. J. Int.*, *174*(1), 195–212, doi:10.1111/j.1365-246X.2008.03763.x.
Idehara, K. (2011), Structural heterogeneity of an ultra-low velocity zone beneath the Philippine Islands: Implications for core-mantle chemical interactions induced by massive partial melting at the bottom of the mantle, *Phys. Earth Planet. Inter.*, *184*, 80–90.
Idehara, K., A. Yamada, and D. Zhao (2007), Seismological constraints on the ultralow velocity zones in the lowermost mantle from core-reflected waves, *Phys. Earth Planet. Inter.*, *165*, 25–46.
Jensen, K. J., M. S. Thorne, and S. Rost (2013), SPdKS analysis of ultralow-velocity zones beneath the western Pacific, *Geophys. Res. Lett.*, *40*, 4574–4578, doi:10.1002/grl.50877.
Kárason, H., and R. D. van der Hilst (2001), Tomographic imaging of the lowermost mantle with differential times of refracted and diffracted core phases (PKP, Pdiff), *J. Geophys. Res.*, *106*, 6569–6587, doi:10.1029/2000JB900380.
Karato, S., and B. Karki (2001), Origin of lateral variation of seismic wave velocities and density in the deep mantle, *J. Geophys. Res.*, *106*, 21,771–21,783, doi:10.1029/2001JB000214.
Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, *122*(1), 108–124.
Kennett, B. L. N., S. Widiyantoro, and R. D. van der Hilst (1998), Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle, *J. Geophys. Res.*, *103*, 12,469–12,493, doi:10.1029/98JB00150.
Labrosse, S., J. Hernlund, and N. Coltice (2007), A crystallizing dense magma ocean at the base of the Earth's mantle, *Nature*, *450*(7171), 866–869, doi:10.1038/nature06355.
Li, X.-D., and B. Romanowicz (1996), Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. Geophys. Res.*, *101*, 22,245–22,272, doi:10.1029/96JB01306.
MacKay, D. J. C. (2003), *Information Theory, Inference, and Learning Algorithms*, Cambridge Univ. Press, Cambridge, U. K.
MacQueen, J. B. (1967), Some methods for classification and analysis of multivariate observations, in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 281–297, Univ. of Calif. Press, Berkeley.
Malinverno, A. (2002), Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem, *Geophys. J. Int.*, *151*(3), 675–688.
Mao, W. L., H.-K. Mao, W. Sturhahn, J. Zhao, V. B. Prakapenka, Y. Meng, J. Shu, Y. Fei, and R. J. Hemley (2006), Iron-rich post-perovskite and the origin of ultralow-velocity zones, *Science*, *312*(5773), 564–565, doi:10.1126/science.1123442.
Masters, G., S. Johnson, G. Laske, and B. Bolton (1996), A shear-velocity model of the mantle, *Philos. Trans. R. Soc. London A*, *354*(1711), 1385–1411.
McNamara, A. K., and S. Zhong (2004), Thermochemical structures within a spherical mantle: Superplumes or piles?, *J. Geophys. Res.*, *109*, B07402, doi:10.1029/2003JB002847.
McNamara, A. K., E. J. Garnero, and S. Rost (2010), Tracking deep mantle reservoirs with ultra-low velocity zones, *Earth Planet. Sci. Lett.*, *299*, 1–9.
Mégnin, C., and B. Romanowicz (2000), The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms, *Geophys. J. Int.*, *143*(3), 709–728.
Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953), Equation of state calculations of fast computing machines, *J. Chem. Phys.*, *21*, 1087–1093.
Mori, J., and D. V. Helmberger (1995), Localized boundary layer below the mid-Pacific velocity anomaly identified from PcP precursor, *J. Geophys. Res.*, *100*, 20,359–20,365, doi:10.1029/95JB02243.
Mosegaard, K., and A. Tarantola (1995), Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res.*, *100*, 12,431–12,447, doi:10.1029/94JB03097.
Mosenfelder, J. L., P. D. Asimow, D. J. Frost, D. C. Rubie, and T. J. Ahrens (2009), The MgSiO3 system at high pressure: Thermodynamic properties of perovskite, postperovskite, and melt from global inversion of shock and static compression data, *J. Geophys. Res.*, *114*, B01203, doi:10.1029/2008JB005900.
Obayashi, M., J. Yoshimitsu, G. Nolet, Y. Fukao, H. Shiobara, H. Sugioka, H. Miyamachi, and Y. Gao (2013), Finite frequency whole mantle *P* wave tomography: Improvement of subducted slab images, *Geophys. Res. Lett.*, *40*, 5652–5657, doi:10.1002/2013GL057401.
Okada, Y., K. Kasahara, S. Hori, K. Obara, S. Sekiguchi, H. Fujiwara, and A. Yamamoto (2004), Recent progress of seismic observation network in Japan—Hi-net, F-net, K-net and KiK-net, *Earth Planets Space*, *56*, xv–xxviii.
Otsuka, K., and S. Karato (2012), Deep penetration of molten iron into the mantle caused by morphological instability, *Nature*, *492*, 243–247.
Ritsema, J., A. Deuss, H. J. van Heijst, and J. H. Woodhouse (2011), S40RTS: A degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements, *Geophys. J. Int.*, *184*(3), 1,223–1,236, doi:10.1111/j.1365-246X.2010.04884.x.
Rondenay, S., V. F. Cormier, and E. M. Van Ark (2010), SKS and SPdKS sensitivity to two-dimensional ultralow-velocity zones, *J. Geophys. Res.*, *115*, B04311, doi:10.1029/2009JB006733.
Rost, S., and P. S. Earle (2010), Identifying regions of strong scattering at the core-mantle boundary from analysis of PKKP precursor energy, *Earth Planet. Sci. Lett.*, *297*, 616–626.
Rost, S., and J. Revenaugh (2001), Seismic detection of rigid zones at the top of the core, *Science*, *294*, 1911–1914, doi:10.1126/science.1065617.
Rost, S., and J. Revenaugh (2003), Small-scale ultralow-velocity zone structure imaged by ScP, *J. Geophys. Res.*, *108*(B1), 2056, doi:10.1029/2001JB001627.
Rost, S., E. J. Garnero, Q. Williams, and M. Manga (2005), Seismological constraints on a possible plume root at the core-mantle boundary, *Nature*, *435*, 666–669, doi:10.1038/nature03620.
Rost, S., E. J. Garnero, and Q. Williams (2006), Fine-scale ultralow-velocity zone structure from high-frequency seismic array data, *J. Geophys. Res.*, *111*, B09310, doi:10.1029/2005JB004088.

Sakai, T., T. Kondo, E. Ohtani, H. Terasaki, N. Endo, T. Kuba, T. Suzuki, and T. Kikegawa (2006), Interaction between iron and post-perovskite at core-mantle boundary and core signature in plume source region, *Geophys. Res. Lett.*, *33*, L15317, doi:10.1029/2006GL026868.

Sambridge, M. (1999), Geophysical Inversion with a neighbourhood algorithm -I. Searching a parameter space, *Geophys. J. Int.*, *138*(2), 479–494.

Sambridge, M., K. Gallagher, A. Jackson, and P. Rickwood (2006), Trans-dimensional inverse problems, model comparison and the evidence, *Geophys. J. Int.*, *167*, 528–542, doi:10.1111/j.1365-246X.2006.03155.x.

Schwarz, G. (1978), Estimating the dimension of a model, *Ann. Stat.*, *6*(2), 461–464.

Soldati, G., L. Boschi, and A. M. Forte (2012), Tomography of core-mantle boundary and lowermost mantle coupled by geodynamics, *Geophys. J. Int.*, *189*, 730–746, doi:10.1111/j.1365-246X.2012.05413.x.

Steininger, G., C. W. Holland, S. E. Dosso, and J. Dettmer (2013), Seabed roughness parameters from joint backscatter and reflection inversion at the Malta Plateau, *J. Acoust. Soc. Am.*, *134*(3), 1833–1842.

Su, W., R. L. Woodward, and A. M. Dziewonski (1994), Degree 12 model of shear velocity heterogeneity in the mantle, *J. Geophys. Res.*, *99*, 6945–6980, doi:10.1029/93JB03408.

Tackley, P. J. (2002), Strong heterogeneity caused by deep mantle layering, *Geochem. Geophys. Geosyst.*, *3*(4), 1024, doi:10.1029/2001GC000167.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia, Pa.

Thorne, M. S., and E. J. Garnero (2004), Inferences on ultralow-velocity zone structure form a global analysis of SPdKS waves, *J. Geophys. Res.*, *109*, B08301, doi:10.1029/2004JB003010.

Thorne, M. S., E. J. Garnero, G. Jahnke, H. Igel, and A. K. McNamara (2013), Mega ultra low velocity zones and mantle flow, *Earth Planet. Sci. Lett.*, *364*, 59–67.

Tkalčić, H., and B. Romanowicz (2002), Short scale heterogeneity in the lowermost mantle: Insights from PcP-P and ScS-S data, *Earth Planet. Sci. Lett.*, *201*(1), 57–68.

Tkalčić, H., B. Romanowicz, and N. Houy (2002), Constraints on D″ structure using PKP(AB-DF), PKP(BC-DF) and PcP-P traveltime data from broad-band records, *Geophys. J. Int.*, *149*, 599–616, doi:10.1046/j.1365-246X.2002.01603.x.

Tkalčić, H., M. K. Young, T. Bodin, S. Nigo, and M. Sambridge (2013a), The shuffling rotation of the Earth's inner core, *Nat. Geosci.*, *6*, 497–502, doi:10.1038/ngeo1813.

Tkalčić, H., T. Bodin, M. K. Young, and M. Sambridge (2013b), On the nature of the *P*-wave velocity gradient in the inner core beneath Central America, *J. Earth Sci.*, *24*(5), 699–705, doi:10.1007/s12583-013-0365-7.

To, A., Y. Fukao, and S. Tsuboi (2011), Evidence for a thick and localized ultra low shear velocity zone at the base of the mantle beneath the central Pacific, *Phys. Earth Planet. Inter.*, *184*, 119–133.

Vidale, J. E., and M. A. H. Heldin (1998), Evidence for partial melt at the core-mantle boundary north of Tonga from the strong scattering of seismic waves, *Nature*, *391*, 682–685.

Wen, L., and D. V. Helmberger (1998a), Ultra-low velocity zones near the core-mantle boundary from broadband PKP precursors, *Science*, *279*, 1701–1703, doi:10.1126/science.279.5357.1701.

Wen, L., and D. V. Helmberger (1998b), A two-dimensional P-SV hybrid method and its application to modeling localized structures near the core-mantle boundary, *J. Geophys. Res.*, *103*, 17,901–17,918, doi:10.1029/98JB01276.

Wicks, J. K., J. M. Jackson, and W. Sturhahn (2010), Very low sound velocities in iron-rich (Mg,Fe)O: Implications for the core-mantle boundary region, *Geophys. Res. Lett.*, *37*, L15304, doi:10.1029/2010GL043689.

Williams, Q., and E. J. Garnero (1996), Seismic evidence for partial melt at the base of Earth's mantle, *Science*, *273*, 1528–1530.

Wysession, M. E. (1996), Large-scale structure at the core-mantle boundary from diffracted waves, *Nature*, *382*, 244–248.

Wysession, M. E., A. Langenhorst, M. J. Fouch, K. M. Fischer, G. I. Al-Eqabi, P. J. Shore, and T. J. Clarke (1999), Lateral variations in compressional/shear velocities at the base of the mantle, *Science*, *284*(5411), 120–125, doi:10.1126/science.284.5411.120.

Young, M. K., H. Tkalčić, T. Bodin, and M. Sambridge (2013), Global *P*-wave tomography of Earth's lowermost mantle from partition modeling, *J. Geophys. Res. Solid Earth*, *118*, 5467–5486, doi:10.1002/jgrb.50391.