



UNIVERSITY OF LEEDS

This is a repository copy of *The most parsimonious ultralow-velocity zone distribution from highly anomalous SPdKS waveforms*.

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/170044/>

Version: Published Version

---

**Article:**

Thorne, M, Leng, K, Pachhai, S et al. (3 more authors) (2021) The most parsimonious ultralow-velocity zone distribution from highly anomalous SPdKS waveforms. *Geochemistry, Geophysics, Geosystems*, 22 (1). e2020GC009467. ISSN 1525-2027

<https://doi.org/10.1029/2020gc009467>

---

© 2020. American Geophysical Union. All Rights Reserved. This is the peer reviewed version of the following article: Thorne, M, Leng, K, Pachhai, S et al. (3 more authors) (2021) The most parsimonious ultralow-velocity zone distribution from highly anomalous SPdKS waveforms. *Geochemistry, Geophysics, Geosystems*, 22 (1). e2020GC009467. ISSN 1525-2027, which has been published in final form at <https://doi.org/10.1029/2020GC009467>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

## RESEARCH ARTICLE

10.1029/2020GC009467

### Key Points:

- We inverted for ULVZ location from highly anomalous SPdKS data using the principle of parsimony
- As much as 19.7% of the CMB by surface area may contain ULVZ-like heterogeneity
- Several mega-ULVZs may exist in the deep Earth

### Supporting Information:

- Supporting Information S1
- Movie S1
- Data Set S1
- Data Set S2
- Data Set S3
- Data Set S4
- Data Set S5
- Data Set S6

### Correspondence to:





M. S. Thorne,  
[michael.thorne@utah.edu](mailto:michael.thorne@utah.edu)

### Citation:

Thorne, M. S., Leng, K., Pachhai, S., Rost, S., Wicks, J., & Nissen-Meyer, T. (2021). The most parsimonious ultralow-velocity zone distribution from highly anomalous SPdKS waveforms. *Geochemistry, Geophysics, Geosystems*, 22, e2020GC009467. <https://doi.org/10.1029/2020GC009467>

Received 6 OCT 2020  
Accepted 8 DEC 2020

## The Most Parsimonious Ultralow-Velocity Zone Distribution From Highly Anomalous SPdKS Waveforms

Michael S. Thorne<sup>1</sup> , Kuangdai Leng<sup>2</sup>, Surya Pachhai<sup>1</sup> , Sebastian Rost<sup>3</sup> , June Wicks<sup>4</sup> , and Tarje Nissen-Meyer<sup>2</sup>

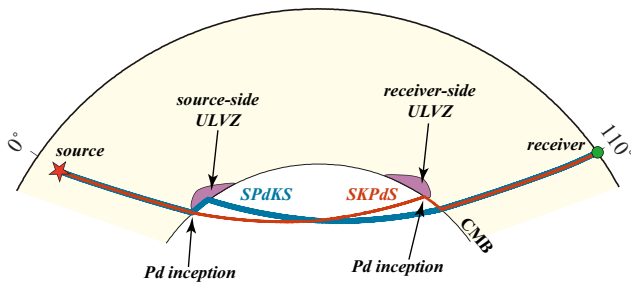
<sup>1</sup>Department of Geology & Geophysics, University of Utah, Salt Lake City, UT, USA, <sup>2</sup>Department of Earth Sciences, University of Oxford, Oxford, UK, <sup>3</sup>School of Earth and Environment, University of Leeds, Leeds, UK, <sup>4</sup>Department of Earth & Planetary Sciences, Johns Hopkins University, Baltimore, MD, USA

**Abstract** The locations of ultralow-velocity zones (ULVZs) at the core-mantle boundary (CMB) have been linked to a variety of features including hot spot volcanoes and large low-velocity province (LLVP) boundaries, yet only a small portion of the CMB region has been probed for ULVZ existence. Here we present a new map of lower mantle heterogeneity locations using a global collection of highly anomalous SPdKS recordings based on a dataset of more than 58,000 radial component seismograms, which sample 56.9% of the CMB by surface area. The inference of heterogeneity location using the SPdKS seismic phase is challenging due to source-versus receiver-side ambiguity. Due to this ambiguity, we conducted an inversion using the principle of parsimony. The inversion is conducted using a genetic algorithm which is repeated several thousand times in order to construct heterogeneity probability maps. This analysis reveals that at probabilities  $\geq 0.5$ , 0.25, and 0.125 up to 1.3%, 8.2%, or 19.7% of the CMB may contain ULVZ-like heterogeneities. These heterogeneities exist in all lower mantle settings, including both high- and low-velocity regions. Additionally, we present evidence that the Samoan ULVZ may be twice as large as previously estimated, and also present evidence for the existence of additional mega-sized ULVZs, such as a newly discovered ULVZ located to the east of the Philippines. We provide new evidence for the ULVZ east of the Philippines through an analysis of ScP records.

**Plain Language Summary** Past deep-Earth studies have identified thin zones of reduced seismic velocities located atop the core-mantle boundary. These zones, referred to as ultralow-velocity zones or ULVZs, have been related to several important features of the Earth's interior such as mantle plumes, continent-sized low-velocity provinces, and large igneous provinces. Nonetheless, we do not know what ULVZs are made of, whether they only exist in specific regions of the core-mantle boundary (CMB), how they link to other mantle properties, or whether all ULVZs are the same. We have assembled a large dataset of seismic waveforms that are sensitive to deep mantle heterogeneity and have identified the waveforms with features indicative of heterogeneous mantle properties, such as ULVZ presence, and search for the simplest distribution of lower mantle heterogeneities. We find that as much as 19.7% of the CMB area must have anomalous seismic properties to explain our dataset. We also use a type of seismic wave reflected off the CMB to verify our initial results in a newly discovered heterogeneous region that lies on the CMB to the east of the Philippines. Our new heterogeneity map provides greater coverage than previous maps and suggests that several mega-sized ULVZs may exist.

## 1. Introduction

For nearly 3 decades, features sitting on top of the core-mantle boundary (CMB) and referred to as ultralow-velocity zones (ULVZs) have been invoked to explain anomalous travel-time delays and the existence of a variety of pre- and post-cursor arrivals observed in seismic data. A clear definition of what constitutes a ULVZ is currently lacking; nevertheless, most papers invoking ULVZ occurrence state that they are regions of low S- and/or P-wave velocities (of the order of 10% or greater decreases) in thin regions (typically less than 40 km is stated) sitting on top of the CMB (Garnero & McNamara, 2008; Ni & Helmberger, 2001; Simmons & Grand, 2002; Yu & Garnero, 2018). However, the range of parameters described as related to ULVZs does not always fit this definition (Bower et al., 2011). Although such ambiguity is disconcerting, strong modeling tradeoffs and nonlinearity make precise determinations of the elastic parameters and morphology of these features challenging. Nonetheless, from a range of independent seismic analyses, it is clear



**Figure 1.** Ray paths for SPdKS (blue) and SKPdS (dark red) for a 500-km deep source (red star) recorded at a sensor 110° in epicentral distance (green circle). The Pd inception point on the source- and receiver-sides is indicated. Also, possible source-side or receiver-side ULVZs are indicated.

that localized regions of low seismic velocities are necessary to explain a host of seismic observations (see Yu & Garnero, 2018 for a recent review of ULVZ-related studies).

Ever since the discovery of ULVZs, their geographic distribution has generated great interest. The first global map of ULVZ location, compiled from 13 different studies, was published by Garnero et al. (1998). This initial map formed the basis for correlating hot spot volcanism with ULVZ locations (Williams et al., 1998). Another global compilation of ULVZ locations was published a little over a decade later (McNamara et al., 2010). In this paper, the authors argued that ULVZs preferentially lie along the edges of large low-velocity provinces (LLVPs). The LLVP edges in turn have been linked to the formation of whole mantle plumes that may give rise to hot spot volcanism (Boschi et al., 2007; McNamara & Zhong, 2005; Thorne et al., 2004; Torsvik et al., 2006; Torsvik et al., 2008). Another comprehensive survey, based on 54 ULVZ studies, was recently conducted

and it argued that ULVZs tend to lie near LLVP edges, but that their link to surface hot spots is less clear (Yu & Garnero, 2018).

Although the correlations presented in the studies outlined above provide some evidence for an overlap between ULVZs and hot spot volcanism or ULVZs and LLVP edges, we note that only a small amount (<20% according to Yu & Garnero, 2018) of the CMB has been probed for ULVZ existence and not all ULVZs are necessarily associated with either phenomena. Since the publication of these studies, additional evidence for ULVZ presence beneath hot spots has been presented. For example, ULVZs may be present beneath the Galapagos, San Felix, Marquesas, Caroline, and Hoggar hot spots (Cottaar & Li, 2019; Kim et al., 2020; Thorne et al., 2020). But, additional new ULVZs beneath Mexico, East Asia, and South America appear to be related to the edges of subducted slab material (Thorne et al., 2019, 2020). Hence, ULVZs may be located beneath major hot spots, but they may also exist near LLVP boundaries that are not co-located with known hot spot volcanism and near the boundaries of past subduction. That is, current evidence does not suggest that ULVZs are principally related to a single type of lower mantle province.

Some of the global surveys discussed above were based on CMB-reflected seismic phases such as ScP, PcP, and ScS. With these phases, there is little ambiguity about the ULVZ location. Nevertheless, it is important to note that even with these phases ULVZs may be mislocated, as a ULVZ location off the great-circle plane or one at a greater epicentral distance from the bounce point location may also give rise to ULVZ-like waveforms (Brown et al., 2015). On the contrary, phases used in detecting ULVZs such as SPdKS and PKP precursors pass through the D'' region on both the source- and receiver-sides of the ray path. Whether the scatterers giving rise to PKP precursors lie on the source- or receiver-side should be distinguishable from the slowness of the precursor (Husebye et al., 1976; King et al., 1974). For other seismic phases such as SPdKS, it is less clear whether the source- or receiver-side structure is responsible for anomalous waveform recordings. For this reason, Yu and Garnero (2018) restricted their global correlations to just seismic phases with unambiguous CMB bounce point locations such as ScP. However, limiting the ULVZ locations to just CMB bounce point phases severely limits the geographic coverage that can be obtained.

ULVZs were first inferred using the SPdKS seismic phase (Garnero et al., 1993). SPdKS has subsequently been used to infer the ULVZ presence in multiple studies (Jensen et al., 2013; Ni & Helmberger, 2001; Ron-denay & Fischer, 2003; Thorne & Garnero, 2004; Vanacore et al., 2016; Wen & Helmberger, 1998). What is commonly referred to as SPdKS is a combination of complementary arrivals called SPdKS and SKPdS, as shown in Figure 1. Some of the down-going S-wave energy from the source converts to a P-wave at the CMB making ScP. At the critical angle for P-wave diffraction (Pd), indicated by the Pd inception point in Figure 1, the conversion becomes SPd. Some of this Pd energy enters the outer core, bottoms, and then on its return to the mantle converts back into an S-wave to generate SPdKS. A similar phenomenon happens on the receiver side of the path to create SKPdS. For a 1D model such as PREM (Preliminary Reference Earth Model [Dziewonski & Anderson, 1981]), SPdKS and SKPdS arrive at the receiver at the same time. However, the existence of a lower mantle structure such as a ULVZ, on either the source- or receiver-side can cause perturbations in timing for either SPdKS or SKPdS. But, in most cases it may be impossible to see

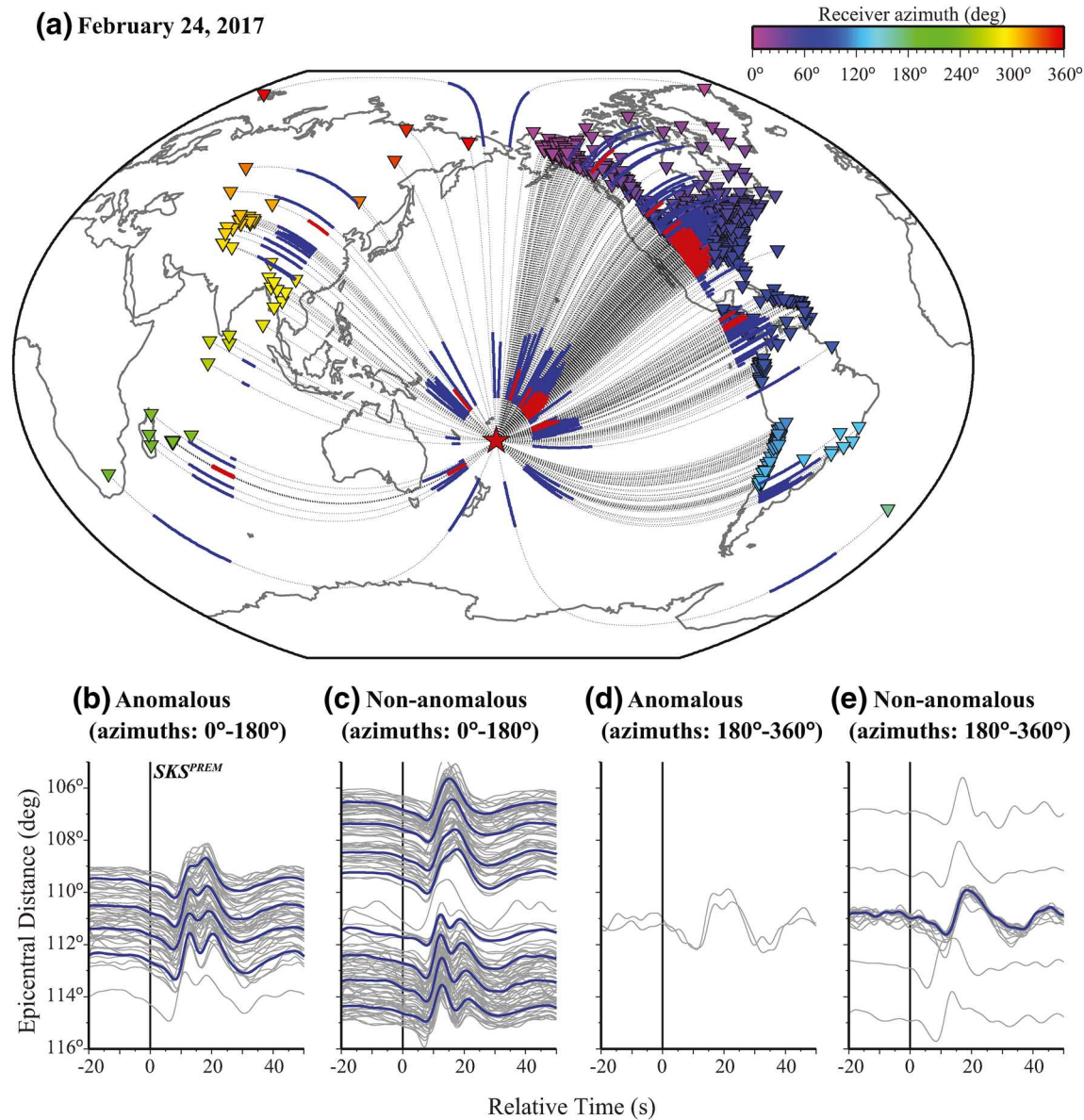
separate SPdKS and SKPdS arrivals (Rondenay et al., 2010; Vanacore et al., 2016). Even if separate SPdKS and SKPdS arrivals are distinguishable, it is not readily apparent whether the structure giving rise to the anomalies occurs as (1) a source-side ULVZ, (2) a receiver-side ULVZ, or (3) a combination of both source- and receiver-side ULVZs.

ULVZ studies using SPdKS have relied on additional arguments to resolve the source- versus receiver-side ambiguity. In particular, these studies have used crossing coverage of Pd paths on the CMB (Ni & Helmberger, 2001; Rondenay & Fischer, 2003) or purely geometric arguments that a ULVZ is more likely to occur on one side of the path because of a higher concentration of anomalous Pd rays (Helmberger et al., 2000; Thorne et al., 2019). Although both arguments appear reasonable, their validity has yet to be confirmed through systematic, synthetic testing. Such testing has not been performed in past studies due to the inability to compute the high-frequency 3-D synthetic seismograms required.

In order to illustrate some of the challenges associated with inferring the ULVZ location from SPdKS data, consider the event shown in Figure 2. For this event, several highly anomalous arrivals were identified in Thorne et al. (2020) that are indicative of ULVZ presence. In Figure 2a, the Pd paths of SPdKS and SKPdS for the highly anomalous arrivals identified are drawn in red, whereas the non-anomalous Pd paths are drawn in blue. Waveforms for anomalous and non-anomalous arrivals are shown in Figures 2b–2e. Waveforms identified as anomalous all show two arrivals, with the second arrival amplitude approximately equal to or greater than the first arrival. Non-anomalous waveforms in the azimuthal range from 0° to 180° are similar in character to the waveforms labeled as *highly anomalous*, but do not get labeled as such due to the strict rules imposed in Thorne et al. (2020). For example, records at distances approximately less than 110° (Figure 2c) do not have a distinct second arrival, although a second arrival is clearly emerging. The deterministic rules prescribed in that paper result in larger distance arrivals not being labeled as highly anomalous as the rules placed a cutoff in the second arrivals amplitude of greater than or equal to 0.8 times the amplitude of the first arrival. On the contrary, non-anomalous arrivals at azimuths greater than 180° do not generally have two arrivals, with the exception of one trace (Figure 2e). Because of this strict rule-based system in identifying anomalies, a strict cutoff was defined to provide a binary label on waveforms. And as such, we then see a hard line between Pd paths (Figure 2a) that are anomalous versus non-anomalous, whereas the reality is likely gradational as is observed in the waveforms. An additional difficulty arises in that the occurrence of anomalous SPdKS waveforms is highly sensitive to the ULVZ position with respect to the point where Pd initiates on the CMB. Nonetheless, this approach highlights those regions in which the waveforms are exceptional in their abnormality, and are thus *highly anomalous*.

Assuming that the ULVZ structure is responsible for the generation of these highly anomalous waveforms, we next desire to know where they are located. If we just decided that each collection of anomalous Pd paths represents a ULVZ location, this would imply that there are 10 ULVZs here for this data example: 4 ULVZs, one each for the SPdKS and SKPdS rays, for azimuths greater than 180° and 6 ULVZs for azimuths less than 180°. However, anomalous data may be explained in a simpler manner. For example, a single anomalous waveform is recorded in China, with a source-side Pd arc beneath the SW Pacific and a receiver-side Pd arc beneath China. A simpler, or more parsimonious, solution would suggest that a single ULVZ either lies beneath the SW Pacific or beneath China, but not beneath both. Unfortunately, from a single waveform it is challenging to determine where this ULVZ lies. Now consider the cluster of anomalous recordings in the azimuthal range from ~70° to 90°. The options are either a large ULVZ beneath North America, or a somewhat smaller ULVZ beneath the Pacific. Typically, the argument is made that the smaller ULVZ is the simpler, most parsimonious, explanation. If additional events exist, then these can also be used, as crossing coverage of anomalous ray paths (on either source- or receiver-side) provide additional evidence for ULVZ location in a specific place.

In this paper, we produce a map of anomalous CMB structure based on SPdKS observations in which the source-/receiver-side ambiguity of the anomaly location inherent in SPdKS observations is taken into account. We do this by first taking the observations of highly anomalous SPdKS waveforms published by Thorne et al. (2020) and calculating the most parsimonious distribution of CMB heterogeneities that could explain these waveforms. This distribution is determined by finding the smallest number of heterogeneities that can be used to explain the observed anomalous waveforms within a predetermined level of misfit. Because the solution is non-unique, we compute the solution for a large number of cases and calculate



**Figure 2.** (a) Great circle arc paths between source (red star) and receivers (inverted triangles) for a 414-km deep event occurring on February 24, 2017 are drawn with dashed black lines. Normal Pd segments of SPdKS and SKPdS are drawn with heavy blue lines. Highly anomalous Pd segments are drawn with heavy red lines. (b) Seismic waveforms identified as anomalous in the azimuthal range from  $0^{\circ}$  to  $180^{\circ}$ . (c) Seismic waveforms identified as non-anomalous in the azimuthal range from  $0^{\circ}$  to  $180^{\circ}$ . (d) Seismic waveforms identified as anomalous in the azimuthal range from  $180^{\circ}$  to  $360^{\circ}$ . (e) Seismic waveforms identified as non-anomalous in the azimuthal range from  $180^{\circ}$  to  $360^{\circ}$ . All seismic traces shown are radial component displacement seismograms. Individual seismic traces are drawn in gray with stacks in  $1^{\circ}$  epicentral distance bins shown in blue.

the probability of finding a heterogeneity in each location where there is data coverage. Here, we compute probability maps of where heterogeneities may exist based solely on SPdKS. These maps may provide focus areas for studies using other CMB-sensitive phases such as ScP, PcP, PKP, ScS, or Sdiff. In addition, we demonstrate that the assumptions inherent in this method are valid through 3-D synthetic testing using the AxisEM3D technique (Leng et al., 2019, 2020) and demonstrate that highly anomalous SPdKS recordings are more likely to be observed when the heterogeneity is located on the source-side of the path. In this approach, we are able to constrain the location of lower mantle heterogeneities, yet we do not constrain the elastic parameters of these regions. As discussed in the next section, past studies have associated highly anomalous SPdKS waveforms with ULVZs. But we cannot state that all anomalous waveforms used in this

study are generated by ULVZs. Hence, these regions may consist of ULVZs but may also correspond to regions with less extreme velocity decrements. Lastly, we show that one of the newly discovered heterogeneous locations, east of the Philippines and north of New Guinea, is consistent with ULVZ presence inferred from the ScP seismic phase, demonstrating the validity of our approach and a link with ULVZs.

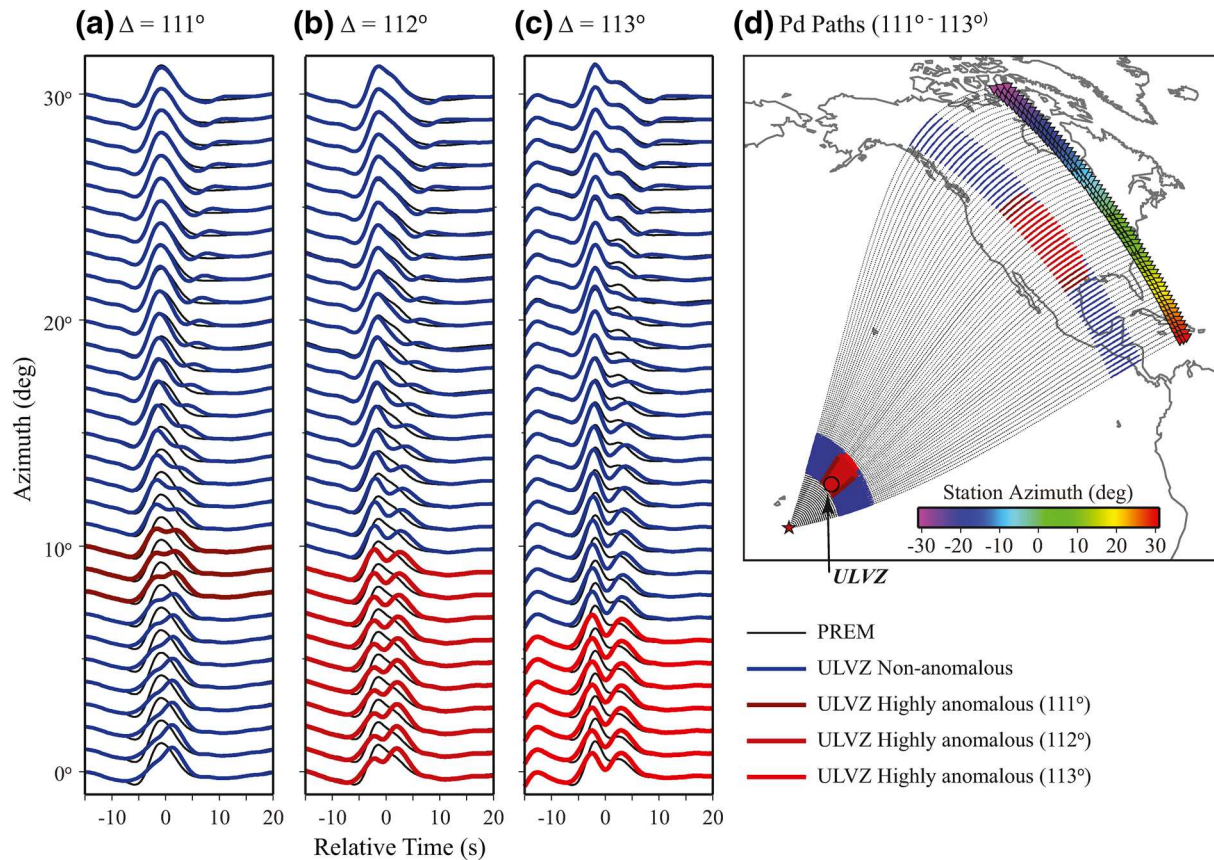
## 2. AxiSEM3D Synthetic Tests

In the PREM model, SPdKS first noticeably bifurcates from SKS at an epicentral distance near  $112^\circ$ . As SPdKS is initiated, it emerges from the shoulder of SKS with a lower amplitude. However, observations of an apparent SPdKS arrival at distances less than  $112^\circ$  and/or SPdKS emergence from SKS with a comparable or larger amplitude than SKS have been observed in association with ULVZs in several past studies (Garnero & Helmberger, 1998; Helmberger et al., 2000; Jensen et al., 2013; Thorne et al., 2013, 2019; Wen & Helmberger, 1998). Thorne et al. (2020) focused on SPdKS recordings in the distance range from  $106^\circ$  to  $115^\circ$ , where such recordings are best observed, and referred to these observations as highly anomalous recordings. Collecting data from events with depths greater than 75 km and occurring between 1990 and 2017, Thorne et al. (2020) retained 58,155 high-quality radial component seismograms in this narrow distance range where the highly anomalous recordings are typically observed. Using an automatic detection algorithm, they found 2,222 anomalous recordings in data where the signal-to-noise (SNR) ratio was greater than or equal to 4. The anomaly detection technique made a binary classification of records based on the records either being clearly anomalous or not. In the case of a clearly anomalous record, the waveform had an additional arrival with respect to the PREM prediction, which stood above the noise level of the seismogram.

These highly anomalous waveforms have been linked to ULVZs on the CMB because of the following considerations:

1. It is unlikely that the waveform distortions are caused by the source-time function. This is because anomalous waveform distortions caused by the source do not also appear in the SKKS arrival, which has a similar takeoff angle (Thorne et al., 2013; Thorne et al., 2019). In addition, normal PREM-like SPdKS waveforms are observed at different azimuths than anomalous waveforms, which is not consistent with a source-time function explanation to the anomalies (see Figure 2).
2. Highly anomalous waveforms are consistently observed for multiple events with crossing Pdiff paths on the CMB (Thorne et al., 2013). This does not necessarily imply a CMB origin to the anomalies, but it is unlikely that all of the events with crossing coverage on the CMB have the same source-time function.
3. It is unlikely that the structure is related to an upper mantle near-source structure. This is because, if the anomalous arrivals were related to a near-source structure, then the anomalous arrivals would move out for different depths. However, in some places we observe events with anomalous arrivals, but that are separated by 100 km or more in depth.
4. It is unlikely that anomalous arrivals are due to near-receiver-structure, as SKS and SKKS have similar paths in the upper mantle and would likely display the same additional arrivals. Given these considerations, it is argued that the origin of the anomalies lies on the CMB.

The binary classification scheme between non-anomalous and highly anomalous waveforms does not identify all the records that may interact with a ULVZ. Yet it may directly highlight where some of the largest ULVZs exist, either in terms of the strength of the velocity anomalies associated with the ULVZ or its size, which tradeoff which each other. This is demonstrated in Figures 3 and 4, which show 3-D synthetic seismograms computed with the AxiSEM3D method (Leng et al., 2019). The 3-D simulations are computed using AxiSEM3D's intrinsic hybrid modeling option (Leng et al., 2020) at a 2-s dominant period, containing energy of up to 1 Hz. Hybrid modeling, also commonly known as wavefield injection and extrapolation (Masson et al., 2013; Robertsson & Chapman, 2000), is a cross-scale method that can enormously speed up wave propagation in a large domain, such as the global Earth (Pienkowska et al., 2020), containing a strong local heterogeneity such as a ULVZ. With AxiSEM3D, the power of hybrid modeling can be maximized by exploiting the different symmetries of the incident wave and scattered wave. The wavefield given by hybrid modeling using AxiSEM3D is the exact 3-D wavefield without any approximations to the model parameters or the wave physics.



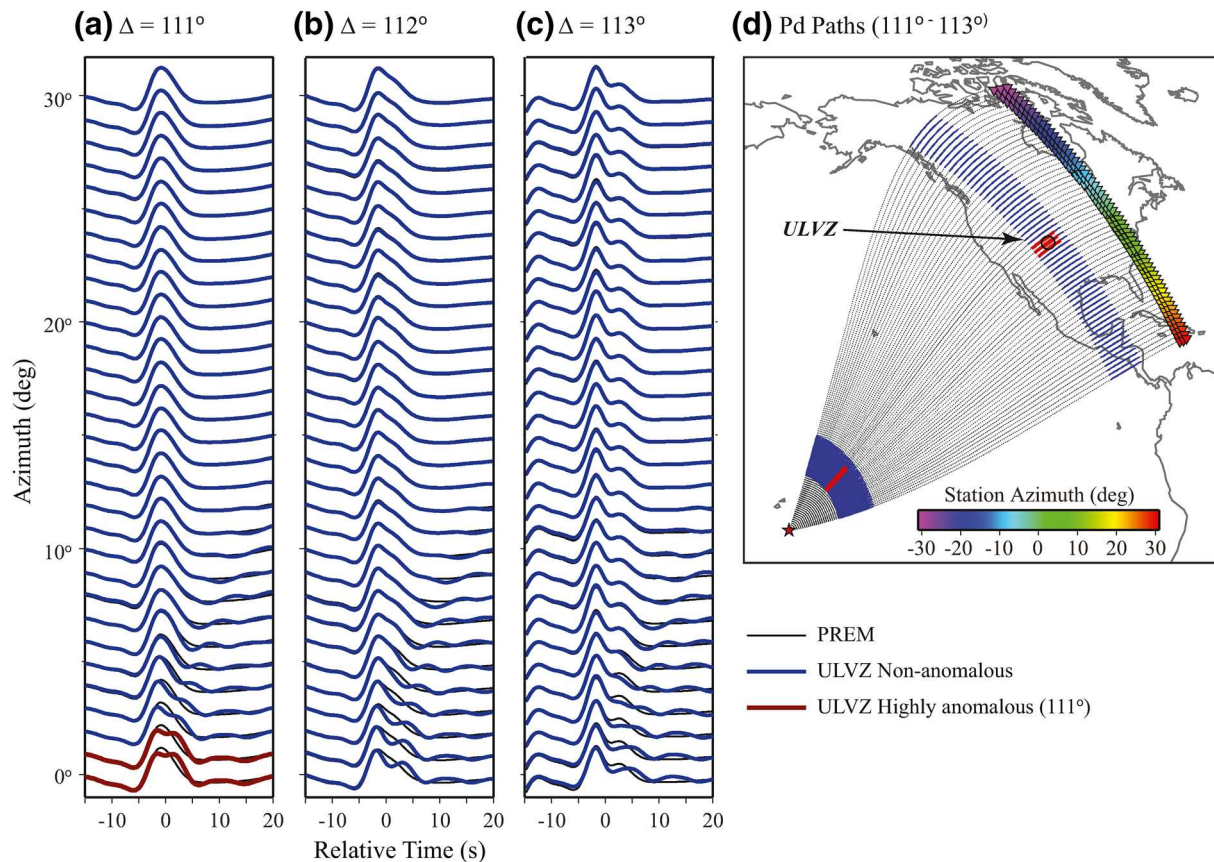
**Figure 3.** 3-D synthetic seismograms for a 3° diameter ULVZ located at latitude =  $-13.56^\circ$  and longitude =  $-167.42^\circ$ . The ULVZ is 20-km thick, with  $\delta V_S = -45\%$  and  $\delta V_P = -15\%$ . Radial component, displacement, synthetic seismograms are shown at azimuths from  $0^\circ$  to  $30^\circ$  with respect to the great circle path between the source and the center of the ULVZ. Azimuths from  $-1^\circ$  to  $-30^\circ$  are not shown as they are symmetric about  $0^\circ$  azimuth. Synthetics are shown at epicentral distances of (a)  $111^\circ$ , (b)  $112^\circ$ , and (c)  $113^\circ$ . In each panel the 1-D PREM synthetic is shown in black. Highly anomalous waveforms for the ULVZ model are drawn in red, and non-anomalous waveforms for the ULVZ model are drawn in blue. Synthetics are filtered with corners between 6 and 40 s to be consistent with data processing from Thorne et al. (2020). (d) Location map. Synthetics are computed for a source (red star) at latitude  $-23.259^\circ$  and longitude  $-178.8^\circ$ . Receivers (inverted triangles) are shown for distances  $111^\circ$ ,  $112^\circ$ , and  $113^\circ$  and are color coded by the receiver azimuth with respect to the great circle arc between source and ULVZ. Pd paths on the CMB are drawn as either non-anomalous (thick blue lines) or highly anomalous (thick red lines, shaded based on epicentral distance, see legend).

Unlike conventional 3-D modeling, the computational cost of AxiSEM3D hybrid modeling scales with the lateral size of the heterogeneity, which in this case is the diameter of the ULVZ. For the 3-D problems discussed in this section, we used a ULVZ model with a diameter of  $3^\circ$  or  $11^\circ$ . At a 2-s dominant period with a 2,400-s record length, these simulations required 21,504 and 78,848 CPU hrs, respectively, to complete.

These synthetics are computed for a ULVZ with identical properties located on either the source-side (Figure 3) or receiver-side (Figure 4) of the path. Using the classification scheme as defined in Thorne et al. (2020), highly anomalous recordings are only found for this ULVZ model in the epicentral distance range from  $111^\circ$  to  $113^\circ$ .

When the ULVZ is located on the source-side of the path (Figure 3), we observe the largest number of highly anomalous waveforms. Regardless of the distance, all highly anomalous waveforms pass through the ULVZ (Figure 3d) or right along its borders. Here we can see that a ULVZ on the source-side of the path shows a tight cluster of highly anomalous Pd arcs (heavy red lines in Figure 3d) on the source-side that is spread out over a much larger geographic area on the receiver-side.

When the ULVZ is located on the receiver-side of the path, for this ULVZ model we only observe highly anomalous recordings at an epicentral distance of  $111^\circ$  (Figure 4a). Here the Pd arcs associated with the



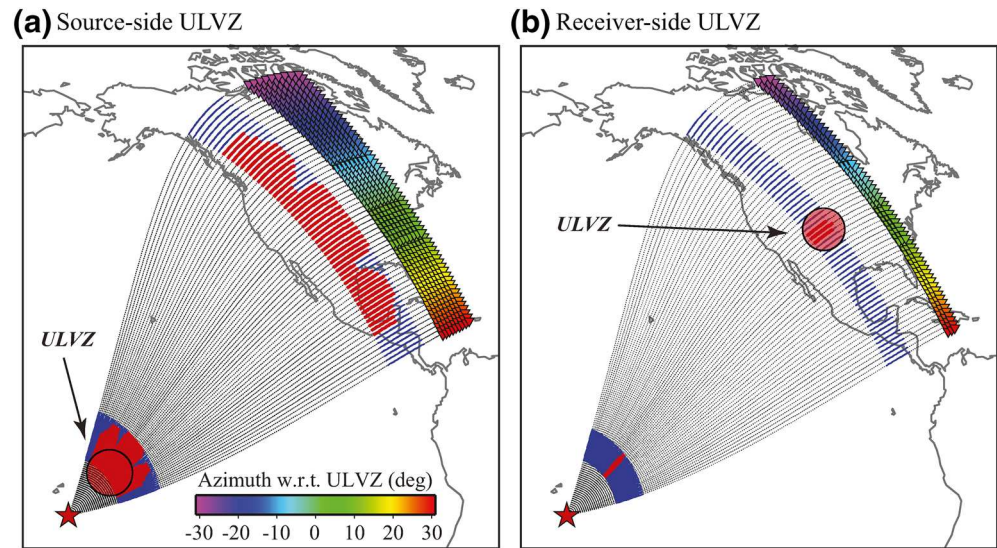
**Figure 4.** 3-D synthetic seismograms for a 3° diameter ULVZ located at latitude = 38.2949° and longitude = -102.1311°. The ULVZ is 20-km thick, with  $\delta V_S = -45\%$  and  $\delta V_P = -15\%$ . Radial component, displacement, synthetic seismograms are shown at azimuths from 0° to 30° with respect to the great circle path between the source and the center of the ULVZ. Azimuths from -1° to -30° are not shown as they are symmetric about 0° azimuth. Synthetics are shown at epicentral distances of (a) 111°, (b) 112°, and (c) 113°. In each panel, the 1-D PREM synthetic is shown in black. Highly anomalous waveforms for the ULVZ model are drawn in red, and non-anomalous waveforms for the ULVZ model are drawn in blue. Synthetics are filtered with corners between 6 and 40 s to be consistent with data processing from Thorne et al. (2020). (d) Location map. Synthetics are computed for a source (red star) at latitude -23.259° and longitude -178.8°. Receivers (inverted triangles) are shown for distances 111°, 112°, and 113° and are color coded by receiver azimuth with respect to great circle arc between source and ULVZ. Pd paths on the CMB are drawn as either non-anomalous (thick blue lines) or highly anomalous (thick red lines, shaded based on epicentral distance, see legend).

anomalous waveforms are still more widely separated on the receiver-side, but are only observed at azimuths that directly pass through the ULVZ (Figure 4d).

In comparing waveforms between the source- and receiver-side ULVZs, we note that far more anomalous waveforms may be detected for a source-side ULVZ. For the synthetic experiments shown here, with 183 receivers (3 epicentral distances  $\times$  61 azimuths), we count a total of 40 versus 3 highly anomalous records when the ULVZ is on the source-side versus the receiver-side of the path. When the ULVZ is on the source-side, we also have a much greater chance of recording and recognizing it as such, as it has the potential to be recorded on a much wider azimuthal range of receivers. As shown in Figure 4, the receiver-side computations were made for a ULVZ centered 94.5° away from the source. We also tested ULVZ models with centers from 90.5° to 106.5° away from the source. In all cases, the maximum number of anomalous waveforms detected was 3. Thus, it is more likely to observe waveform anomalies from a single event for source-side ULVZs, but receiver-side ULVZs may still be manifested by such anomalies, albeit in a narrower azimuthal range.

As shown in Figures 3 and 4, the highly anomalous Pd arcs all pass through the ULVZ, but the actual Pd arc length in these cases is much longer than the ULVZ. In the source-side ULVZ case (Figure 3) the ULVZ is located right at the point of Pd-inception. However, this is not always the case in generating anomalous





**Figure 5.** Locations of anomalous Pd paths for a  $11^\circ$  diameter ULVZ located on (a) the source-side of the path at latitude =  $-13.9038^\circ$  and longitude =  $-167.7937^\circ$ , and on (b) the receiver-side of the path at latitude =  $38.2949^\circ$  and longitude =  $-102.1311^\circ$ . The ULVZ is 25 km thick, with  $\delta V_S = -20\%$  and  $\delta V_P = -15\%$ . Synthetics are computed for a source (red star) at latitude  $-23.259^\circ$  and longitude  $-178.8^\circ$ . Receivers (inverted triangles) are shown for distances where anomalous waveforms are observed. In (a) the distances are from  $107^\circ$  to  $115^\circ$  and in (b) the distances are from  $109^\circ$  to  $111^\circ$ . Receiver locations are color coded by receiver azimuth with respect to the great circle arc between source and ULVZ. Pd paths on the CMB are drawn as either non-anomalous (thick blue lines) or highly anomalous (thick red lines).

waveforms. As noted previously, there are tradeoffs between ULVZ size and velocity contrasts. For example, anomalous Pd paths for a larger,  $11^\circ$  diameter ULVZ are shown in Figure 5. The highly anomalous waveforms for this model are similar to those for the previous model as shown in Figures 3 and 4. But, in order to generate similar seismograms, the position of the ULVZ had to be adjusted and the magnitude of the S-wave velocity reduction had to be reduced. As in the previous case, more anomalous waveforms are observed when the ULVZ is located on the source-side of the path (in this example, we see anomalous waveforms at a wide variety of azimuths from  $\pm 20^\circ$  and from distances between  $107^\circ$  and  $115^\circ$ ). But, note that in this case, the Pd inception starts toward the center of the ULVZ, whereas in the previous case Pd inception was coincident with the ULVZ edge. In this case, only a handful of anomalous recordings are observed on the receiver-side at a much more limited azimuthal ( $\pm 2^\circ$ ) and epicentral distance range (from  $109^\circ$  to  $111^\circ$ ).

In each case shown (Figures 3–5), the Pd paths on the CMB are longer than the physical dimensions of the ULVZs, and where the Pd path occurs with respect to the ULVZ depends on the size of the ULVZ and its elastic parameters. However, the highly anomalous waveforms do constrain the ULVZ azimuthal boundaries.

What we have learned from these synthetic experiments can be summarized in the following points:

1. We are more likely to observe a source-side ULVZ than a receiver-side ULVZ due to receivers at a larger range of azimuths that could show the anomalies.
2. Multiple events interacting with a source-side ULVZ with crossing coverage near the source-side would show anomalous recordings over a potentially larger azimuthal range.
3. Fewer anomalous waveforms are observed when the ULVZ is located on the receiver-side, which could mean that the mechanism by which highly anomalous recordings are generated may differ from source-to-receiver-side. This should be explored further theoretically as to whether one is more feasibly generated than the other.
4. Pd paths of highly anomalous recordings pass through the ULVZ in these experiments, and thus azimuthal bounds can be estimated.

5. Tradeoffs exist between ULVZ size and velocity decrement, so that the ULVZ location along the Pd arc cannot be constrained solely on the presence of highly anomalous waveforms.

### 3. Parsimonious Distributions of Heterogeneities

In this study, we seek to find the most parsimonious distribution of heterogeneities that may explain a given set of highly anomalous SPdKS observations. The principle of parsimony implies that we are looking for the solution that explains the anomalous observations with the least number of heterogeneities. This idea is often utilized in biological studies, in particular with respect to phylogenetic relationships (Fitch, 1971).

To determine the most parsimonious distribution of heterogeneities the approach we use is to (1) define a grid over which we will search for possible heterogeneity locations, (2) select a subset of these grid cells that possibly contain heterogeneities, (3) calculate a misfit between observations and predictions based on the grid cells we are testing to have heterogeneities, and (4) find the distribution of grid cells that contain heterogeneities that fit the observations with the minimum number of heterogeneities, noting that the solution may be non-unique. We provide further details on the procedure in the following paragraphs.

#### 3.1. Misfit Calculation

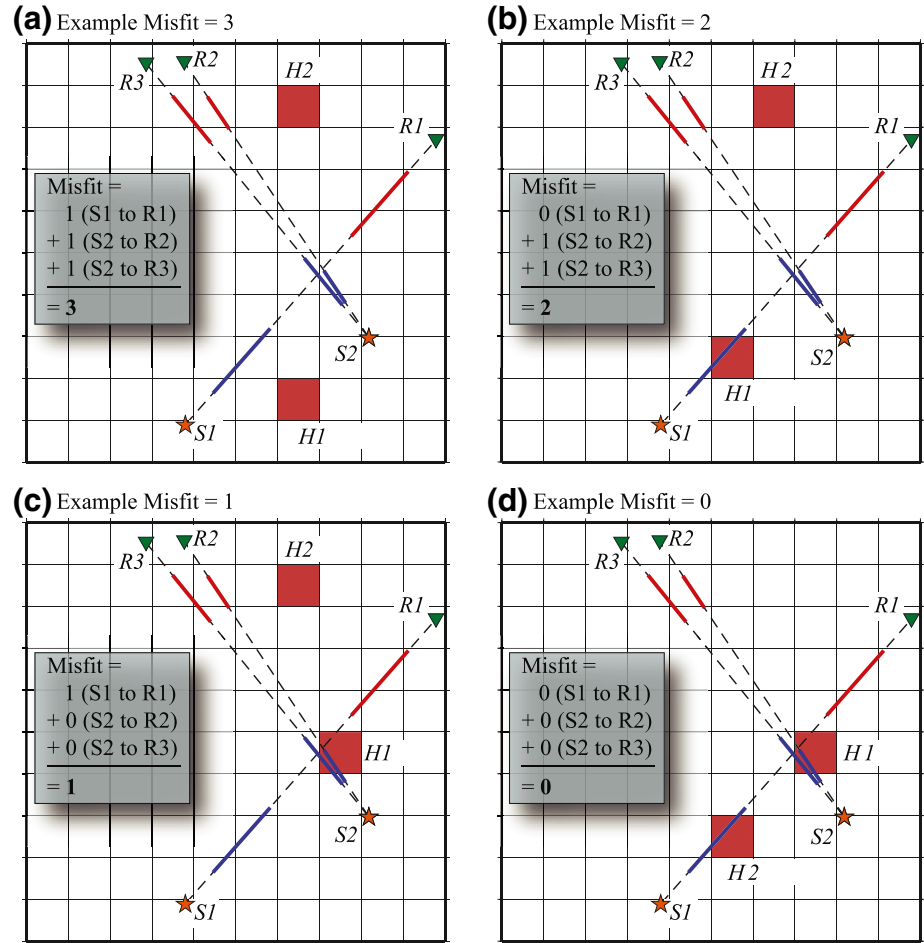
First, we choose a suitable grid size. For the problem of SPdKS, this is discussed further in Section 5. Here, we want to provide a balance between fine enough resolution and the physical sizes of heterogeneities that may effectively generate anomalous SPdKS waveforms. Once a suitable grid has been selected, our procedure is to first number each grid cell. Next, for each highly anomalous observation, we tabulate which grid cells the Pd paths cross through on both the source- (SPdKS) and receiver- (SKPdS) sides. This step makes it efficient to calculate the misfit as described below.

Next, we select a number of heterogeneities to test. An example is shown in Figure 6. Here we have specified a  $10 \times 10$  grid (100 grid cells) and show three example observations from Source 1 (S1) to Receiver 1 (R1) and Source 2 (S2) to Receivers 2 and 3 (R2,R3). In this example, we have decided to test a distribution with two heterogeneities. In Figures 6a–6d, possible heterogeneity locations are shown as solid red squares.

To assess how well a given distribution of heterogeneities fits the observations, we calculate the misfit. We define the misfit as the number of highly anomalous observations that cannot be explained by the given distribution of heterogeneities. Thus, a lower misfit provides a better explanation for the observations. For example, in Figure 6a, we have three highly anomalous observations, and two heterogeneities. But, none of the Pd paths on either the source- or receiver-sides pass through either of these heterogeneities. Thus, the misfit is three which is the largest possible misfit in this example. In example 2 (Figure 6b), the source-side Pd path of the S1-R1 observation crosses through H1. But, neither S2-R2 or S2-R3 are accounted for, so the misfit = 2. In example 3 (Figure 6c), both observations S2-R2 and S2-R3 are accounted for by H1, but S1-R1 is not accounted for and the misfit = 1. The misfit in example 4 (Figure 6d) is 0, as the given heterogeneity distribution accounts for all of the observations. However, non-uniqueness exists in that other heterogeneity distributions could also provide a 0 misfit (e.g., by moving H2 to a grid cell crossing the receiver-side of the path).

In the example given in Figure 6, the most parsimonious distribution of heterogeneities includes two heterogeneities. That is, we cannot explain all of the observations (i.e., we can at best get a misfit = 1 in this case) with only one heterogeneity but we can explain all of the observations (misfit = 0) with more than two heterogeneities. But, by parsimony, the simplest solution is a solution with the least number of heterogeneities, which in this case is 2. Other measures of model fitness have also been introduced in the literature such as the Bayes Information Criteria (BIC) (Pachhai et al., 2014). But, given the same level of misfit, such measures will favor the model with the least number of parameters, or the least number of heterogeneities, and thus the BIC would also favor the most parsimonious solution.

Because non-uniqueness of solutions exists, we seek to quantify which heterogeneity locations are most likely. Ideally, we would calculate the probability that a heterogeneity exists at a particular grid cell by performing a global search through all the possible combinations of heterogeneity locations. The total number of combinations without repetition for a grid with  $n$  elements and  $k$  heterogeneities is given as follows:



**Figure 6.** Example of misfit calculations for possible heterogeneity locations on a  $10 \times 10$  grid for a distribution of two sources (Source 1 and Source 2; red stars) and three receivers (R1, R2, and R3; green triangles). For each observation the great circle path between the source and receiver is drawn with a dashed line. The source-side Pd path is a thick blue line and the receiver-side Pd path is a thick red line. Possible heterogeneity locations (H1 and H2) are indicated by the solid red boxes.

$$C_{n,k} = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

For example, on a  $10 \times 10$  grid with 2 heterogeneities the total number of combinations is as follows:

$$C_{100,2} = \binom{100}{2} = \frac{100!}{2!(100-2)!} = 4,950$$

However, the number of combinations increases rapidly. For 10 possible heterogeneities on the same grid, the number of possible combinations is as follows:

$$C_{100,10} = \binom{100}{10} = \frac{100!}{10!(100-10)!} = 1.731 \times 10^{13}$$

For larger problems, performing a global search may not be computationally feasible. However, it is possible to do a global search for smaller problems and provide a means by which we can test alternate methods.

Figure 7 shows three example problems on a  $10 \times 10$  grid for realistic geometries that demonstrate some of the advantages and challenges with finding heterogeneity locations. The setup for the first problem is given in Figure 7a. Here, we have two sources (labeled S1 and S2) recorded at three stations (labeled R1, R2, and R3). The source-side Pd paths are drawn with heavy blue lines and the receiver-side paths are drawn with heavy red lines. Grid boxes are numbered sequentially starting from the lower left-hand side of the plot (select grid numbers are shown in italics). For this problem, two heterogeneities are required to obtain a 0 misfit. The calculated probabilities are shown in Figure 7d. This problem suffers from non-uniqueness. Only two heterogeneities may be required, but there are several locations where they can exist. One of the heterogeneities must be in one of the four boxes 38, 48, 75, or 85 and thus the probability of a box lying in one of those boxes is  $1/4$ . A heterogeneity in any one of those locations accounts for two of our observations (anomaly from S2-R2 and S2-R3). The other heterogeneity must account for the observation (S1-R1). But with no other records, we cannot say where it lies other than it must lie in one of the eight boxes: 15, 25, 26, 36, 58, 59, 69, or 70. Hence, the probability of a heterogeneity lying in one of those boxes is  $1/8 = 0.125$ . In this example, we can see that there is a greater probability of a heterogeneity existing where more than one observation crosses a grid cell, but with similar paths on source- and receiver-side one cannot distinguish which side of the path the heterogeneity is on.

Example 2 (Figures 7b and 7e) is similar to the first example and two heterogeneities are required to provide a 0 misfit. However, in this case the receiver-side Pd paths are separated further apart. Thus, the most parsimonious solution requires a heterogeneity to be either in box 65 or 74 and on the source-side of the path. This argument has been invoked in several previous heterogeneity studies using SPdKS.

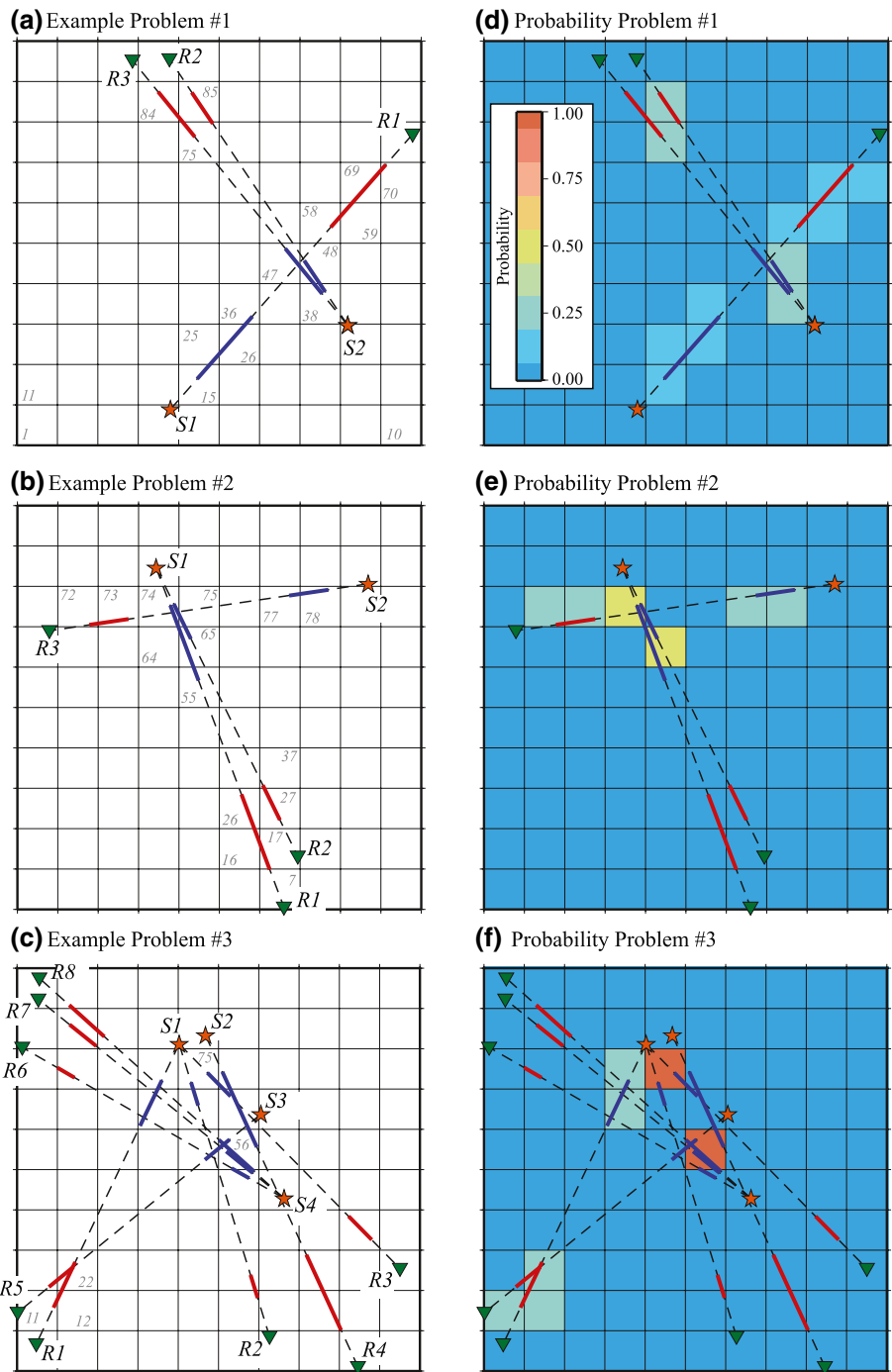
A more complicated example is shown in Figures 7c and 7f. In this case three heterogeneities are required to make the most parsimonious solution. In this case, two heterogeneities must exist (probability = 1) in boxes 56 and 75. A third heterogeneity must exist to explain the observation from S1 to R1. Note that the probability in box 22 does not increase because of the crossing coverage from the S3-R5 observation. This is because this observation is already accounted for with the necessity of a heterogeneity in box 56.

This method provides useful information in that (1) the probability is reduced if one cannot distinguish between source- and receiver-side paths, (2) can distinguish between source- and receiver-side paths if there is a tight clustering of observations on one side of the path, and (3) the method does not give extra bias to grid cells where the observation is already accounted for. However, challenges exist in that (1) the method does not care about the heterogeneity size and thus assumes a heterogeneity of the same size as the grid and (2) the method can split the probability across multiple grid cells for longer Pd path lengths, or Pd paths that lie near the grid cell boundaries. These challenges can be mitigated by performing additional analyses where the grid centers are shifted and by also performing the calculations on different-sized grid cells.

For large-scale problems such as the number of possible heterogeneities on the CMB, a global search is not feasible. For example, searching for 100 possible heterogeneities on a  $4 \times 4^\circ$  grid (4,050 grid cells) we cannot determine the number of possible combinations by direct application of Equation 1. An alternative approach to using a global search is to invert for heterogeneity locations, given the misfit as defined above. In this paper, we invert for heterogeneity locations using a genetic algorithm as described below.

### 3.2. Genetic Algorithm Approach

Genetic algorithms were first developed over the late 1960s and early 1970s as a naïve mimicry of the evolutionary process in solving optimization problems with complex cost functions (Haupt & Haupt, 2003). Genetic algorithms have found use in a wide variety of geophysical applications for decades (Ramillien, 2001; Sambridge & Drijkoningen, 1992; Sen & Mallick, 2018). We chose to use a genetic algorithm approach here as this type of technique is well-suited to the way in which we have set up our problem, because it has the potential to converge to optimal solutions rapidly, and because we can readily find multiple unique solutions.



**Figure 7.** Example problems to demonstrate the probability calculations. Each problem is set up on a  $10 \times 10$  grid. Each grid cell is labeled from 1 to 100 starting from the lower left counter. (a)–(c) Example problem setups. In each example sources (red stars) are labeled S1, S2, S3, or S4. Receivers are indicated with green triangles and labeled R1, R2, ..., R8. Paths from source to receivers are indicated with dashed lines. Source- and receiver-side Pd arcs are drawn with heavy blue or red lines, respectively. In panels (a) and (b), select grid cell numbers are indicated. (d)–(f) The probability of the most parsimonious ULVZ distribution for each grid cell is indicated. In panels (d) and (e), the probability is given for a distribution of 2 ULVZs, which provides a misfit of 0. In panel (f) the probability is given for a distribution of 3 ULVZs, which is required for a misfit of 0.

In adopting a genetic algorithm for this problem, we start as follows. First, we assume that we have  $k$  heterogeneity locations which are described by the grid cell number in which they exist. Each heterogeneity location represents a gene. A chromosome is generated by making an array of such genes. For example, if we had five heterogeneities and 4,050 grid cells a chromosome could look as follows:

$$chromosome = [3042 \ 403 \ 1044 \ 2472 \ 3357]$$

where each number in the chromosome (i.e., each gene) represents a grid cell with a heterogeneity. Our algorithm proceeds as follows:

1. We generate an initial random population of chromosomes referred to as the parent population. For example, we compute 100 parent chromosomes with randomly selected heterogeneity locations
2. We evaluate the fitness of each model (parent) as described previously by our misfit calculation
3. We store the best-fit model, adding it to the mating pool in step 4, but not allowing mutation operations on it as in step 6, which is referred to as elitism
4. We next create a mating pool using rank-biased selection
5. From the mating pool we create children by a 2-point crossover operation of the chromosomes. We used a simple binary representation of the chromosomes, so that after crossover new potential heterogeneity locations are introduced
6. We next apply random mutation to the genes
7. Finally, we evaluate the fitness of the children. If the optimal solution is found (e.g., misfit = 0), then we stop. Otherwise, we replace individuals in the initial parent population with some of the children and return to step 3

We tested the genetic algorithm approach for the  $10 \times 10$  grid problems as shown in Figure 7 and found that we are able to converge to approximately the same probability as the global search within a few hundred iterations of running the genetic algorithm. We will discuss parameterization and convergence in more detail in the next section when we discuss application to the real data.

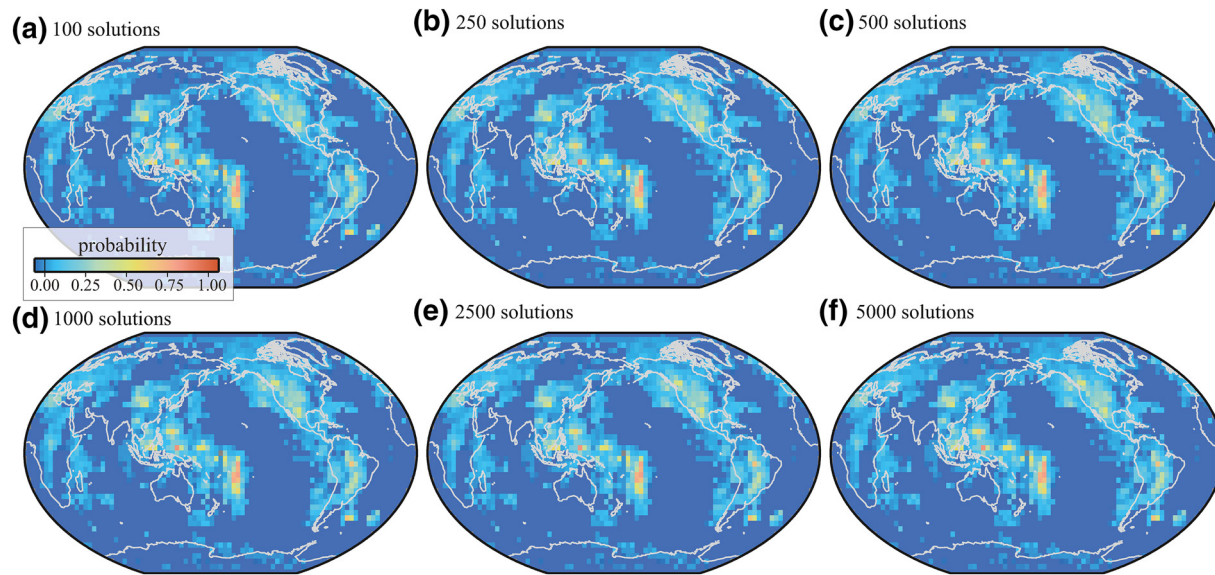
In sum, our technique uses the following procedure:

1. We assume a grid cell size (e.g.,  $4 \times 4^\circ$ ) and number the grid cells sequentially (e.g., from 1 to 4,050 on a  $4 \times 4^\circ$  grid). Grid cells are defined by latitude and longitude
2. For each observation, we tabulate which grid cells have crossing Pd paths on the source- and receiver-sides of the paths
3. We assume a number ( $k$ ) of heterogeneities to test (e.g., we may test for  $k = 100, 120, 140$ , etc., heterogeneities)
4. We use the genetic algorithm to search for the best heterogeneity distribution (minimum misfit) for the given number of heterogeneities
5. We keep searching, using different values of  $k$  until we find the minimum number of heterogeneities ( $k_{\min}$ ) for which we can find a distribution that satisfies our misfit criterion
6. Now that we have determined  $k_{\min}$ , we run the genetic algorithm for multiple iterations using a different starting random seed (parent population) for each iteration, until the probability of a heterogeneity existing at each grid cell has converged

In the next section, we discuss our results of applying this methodology to the highly anomalous SPdKS dataset.

## 4. Results

To determine the most parsimonious distribution of heterogeneities, we select a grid size of either  $4^\circ$  or  $8^\circ$ . The smallest grid size of  $4^\circ$  is selected as a grid size in which we can solve the inverse problem for multiple solutions in a reasonable amount of time, and is also of the order of the smallest heterogeneities that previous studies have demonstrated could produce the anomalous SPdKS waveforms (Jensen et al., 2013; Thorne et al., 2013, 2019; Vanacore et al., 2016).



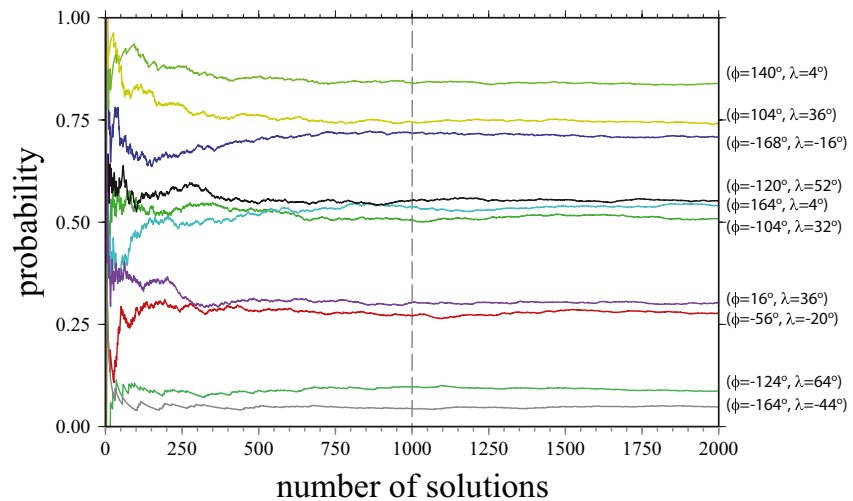
**Figure 8.** Heterogeneity probability for different numbers of solutions. Heterogeneity probability is shown for a search conducted with 116 heterogeneous grid boxes on a  $4 \times 4^\circ$  grid with a 5% misfit. The probability is shown for (a) 100, (b) 250, (c) 500, (d) 1,000, (e) 2,500, and (f) 5,000 unique solutions. The scale bar for probability extends below 0 to include areas with no data coverage in the plot.

For each grid size, we determine the most parsimonious solution as follows. First, we select a level of misfit (either 5%, 2.5%, or 1%). For example, we have 2,222 anomalous observations, and thus a level of misfit of 1% implies that we accept the solution if 22 of these observations are not explained by our solution. For our given misfit level, we allow the genetic algorithm to search for  $k_{\min}$  number of heterogeneities. First, we allow the algorithm to search in multiples of 10 (e.g., 100, 110, 120, etc. heterogeneities). For example, at a 1% misfit level, on a  $4 \times 4^\circ$  grid we found that 156 grid squares must have a heterogeneity in order to explain 2,200 of the 2,222 observations. So, our search first stopped in that we found no solutions when looking for 150 heterogeneities, but did find solutions when looking for 160 heterogeneities. We next refine the search from 150 to 159 possible heterogeneity locations, until we determined that we could not find a solution for 155 heterogeneities but we were able to for 156 heterogeneities. These locations may be contiguous, so this does not mean that we infer 156 unique heterogeneities.

Once we have determined the number of locations we are searching for, we let the genetic algorithm search for solutions. The solutions are non-unique, because (1) we allow for some level of misfit and (2) source-receiver ambiguity exists for anomalous waveforms with low density of anomalies and/or no crossing coverage. We search for many solutions in order to build up our probability map. For this study, we determined that 1,000 unique solutions were sufficient. This was determined by calculating 5,000 unique solutions for a test setup and examining the probability as a function of number of solutions. Probability maps based on different numbers of unique solutions are shown in Figure 8.

From the tests shown in Figure 8 we determined that the probability converges by approximately 1,000 solutions. For example, we cannot distinguish between probability maps for 1,000 solutions (Figure 8d) and 5,000 solutions (Figure 8f). This is further demonstrated in Figure 9, which shows the probability at 10 grid cells chosen to have different final probabilities. The probability shows some variability over the first 300–400 solutions which is referred to as the burn-in period, but is essentially flat for greater than 1,000 solutions.

We determined the most parsimonious heterogeneity distribution for three levels of misfit (5%, 2.5%, and 1%) using a  $4 \times 4^\circ$  grid and a shifted version of the grid in which the grid was shifted by  $2^\circ$  in latitude and longitude. In addition, we repeated the experiment using an  $8 \times 8^\circ$  grid. The results for the  $4 \times 4^\circ$  grid are shown in Figure 10, where we averaged the solutions between the original and shifted versions of the grid using the nearest neighborhood algorithm.



**Figure 9.** Probability at 10 grid cells as a function of the number of solutions for the probability maps shown in Figure 8. The longitude ( $\phi$ ) and latitude ( $\lambda$ ) of the grid cell are displayed on the right of the plot.

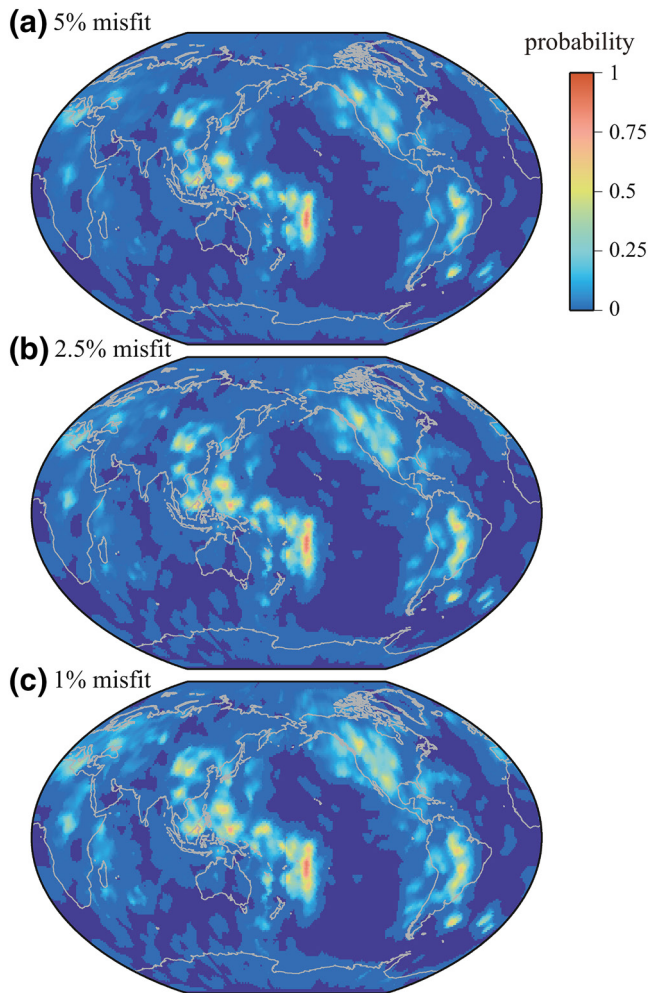
When using a larger misfit, this implies that we are assuming that more of the highly anomalous observations may be in error. As a result, the probability map favors the grid spaces with the largest number of records. Decreasing the misfit forces the calculation to consider more grid spaces as having heterogeneities, and thus we see an increase in possible heterogeneities for the 1% misfit map (Figure 10c) relative to the 5% misfit map (Figure 10a). Thus, the probabilities shown in the 5% misfit map show the most robust solutions with the most numbers of anomalous Pd rays, but the 1% misfit map takes more of the anomalous records into consideration. A contour map of the heterogeneity locations using 1% misfit is shown in Figure 11. An outline of the possible heterogeneity locations is given in Table 1. In this table, we also identify whether or not the location has been identified in a previous study as having ULVZs, excluding Thorne et al. (2020) from previous discovery as the dataset used in that paper is identical to this study.

Some areas stand out as having a probability of  $>0.5$  of heterogeneity existence (Figure 11). Specifically, these areas are (1) beneath East Asia, (2) beneath the South China Sea, (3) two spots East of the Philippines and north of Papua New Guinea, (4) beneath the Caroline hot spot, (5) beneath the Samoa hot spot, (6) east of Vancouver, and (7) three locations possibly forming a line beneath S. America. The area beneath the Samoa hot spot is especially pronounced with a region of probability  $>0.5$  that extends from roughly 350 to 500 km  $\times$  1,600 km along the CMB. However, we see that possible heterogeneities beneath S. America and N. America, if contiguous, may rival the Samoa heterogeneity in size.

Because the probability may get smeared out over several grid points, one should also consider other areas in these maps with probabilities  $<0.5$  as potential sites of heterogeneity. For example, if we consider probabilities  $>0.125$ , then additional heterogeneities are inferred. It is worthwhile to consider all of these spots, as some of the locations have been implicated for ULVZ existence in previous studies. For example, we have a heterogeneity probability  $>0.25$  for areas beneath northern Mexico and Florida (Havens & Revenaugh, 2001; Thorne et al., 2019), New Caledonia (Rost & Revenaugh, 2003; Rost et al., 2005, 2006), and central Africa (Helmlberger et al., 2000). All of these areas have been shown to harbor ULVZs in past studies. Additionally, some areas with probabilities  $>0.125$  may contain ULVZs such as beneath Spain for which ULVZ evidence also exists (Rost & Garnero, 2006). NetCDF grid files of the solutions shown in Figures 10 and 11 are provided in the supplementary online materials.

We also searched for the most parsimonious heterogeneity locations using an  $8 \times 8^\circ$  grid and 5% misfit. As the solution just looks like a lower resolution version of those presented above, we do not show it here. In addition, we also tested a handful of cases with larger numbers of ULVZs than are required. Adding additional ULVZs also does not change the solution much, but increases the non-uniqueness of the solutions and thus the probabilities get further reduced. A further test was conducted in which we relaxed the





**Figure 10.** The probability of heterogeneity existence is shown based on calculations assuming (a) 5% misfit, (b) 2.5% misfit, or (c) 1% misfit. Dark blue regions are areas with no data coverage.

conditions required to define a highly anomalous SPdKS recording. In this test, we allowed all arrivals with amplitudes greater than 0.75 to be considered for classification as highly anomalous. Thorne et al. (2020) had originally set this number to 0.8. This search yielded 2,425 highly anomalous records (relative to 2,222 previously found). However, the resulting distribution of parsimonious heterogeneities was not significantly altered. All tests produced maps similar to those shown above.

## 5. ScP Records Sampling the ULVZ East of the Philippines

The probability maps shown in Figures 10 and 11 provide our best estimate of where heterogeneities are located that explain the most highly anomalous SPdKS waveforms. However, some uncertainty still exists in source- versus receiver-side locations because not all locations have either dense ray coverage or crossing coverage. In some cases, this may result in heterogeneity probability getting distributed across both sides (see examples in Figure 7). Hence, some uncertainty still remains in areas of these maps that have a lower probability. Also, areas identified with high probability does not mean it is heterogeneous; rather that placing a heterogeneity in that location makes for the simplest solution. Hence, these probability maps serve to provide the best estimate of where we expect anomalous seismic heterogeneity using SPdKS, but may also act as a guide of locations to search for heterogeneities with other seismic phases. Once a heterogeneity has been confirmed in a given location, then SPdKS data in conjunction with that phase would ideally be used to model elastic properties.

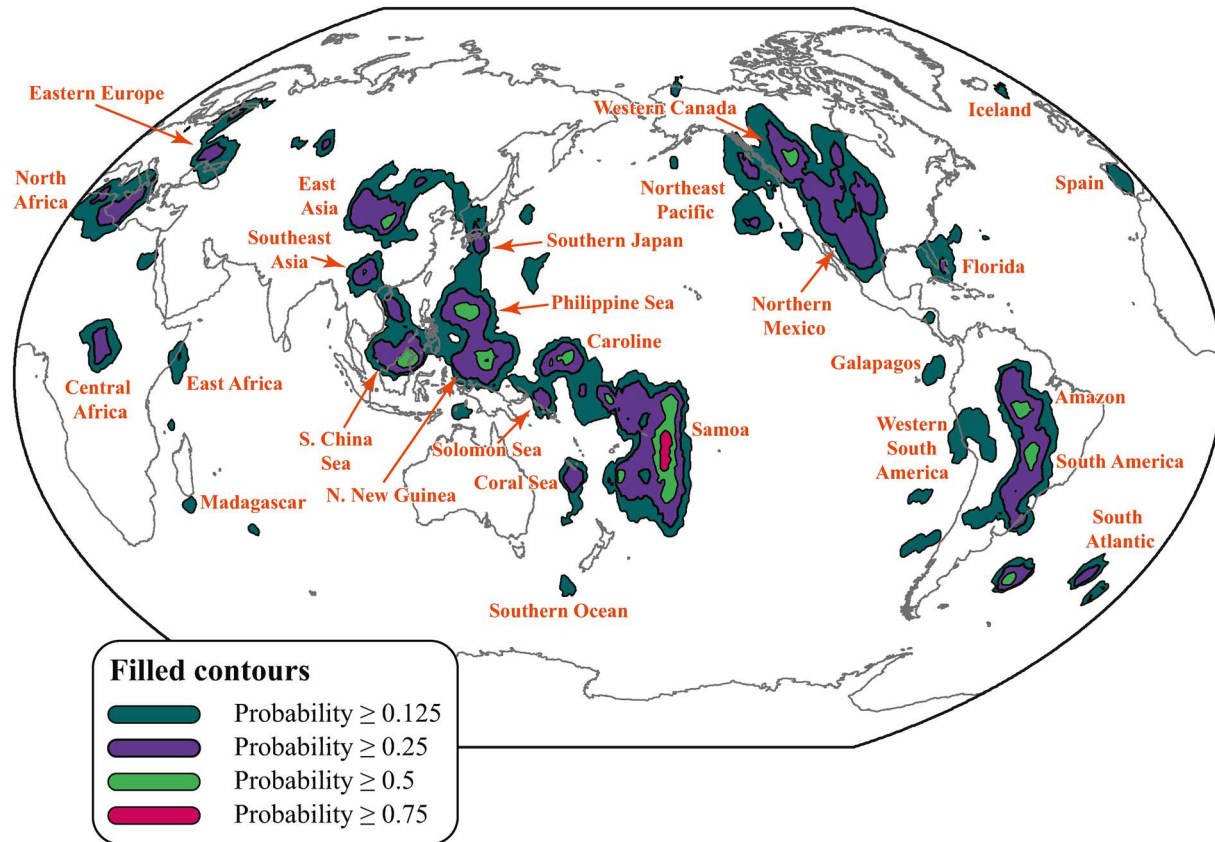
Ideal phases used for confirmation are phases that do not suffer the source-/receiver-side ambiguity that is inherent in SPdKS data. In particular, the seismic phases ScP (Garnero & Vidale, 1999; Hansen et al., 2020; Pachhai et al., 2014; Rost & Revenaugh, 2003), PcP (Havens & Revenaugh, 2001; Revenaugh & Meyer, 1997), ScS (Zhao et al., 2017), and Sdiff (Cottaar & Romanowicz, 2012; Kim et al., 2020) have been successfully used. As searching for heterogeneities with high-frequency arrivals such as ScP and PcP without a priori information can only lead to serendipitous detections due to the small sampling footprint, it is useful

to have targeted regions to search for such as presented in this paper.

We conducted a search for ScP bounce points that coincide with areas of high heterogeneity probability, where previous studies have not indicated ULVZ presence. We found that events occurring along the Mariana Trench and recorded at seismic arrays in Australia are at an ideal location to sample the potential heterogeneities we infer just east of the Philippines and north of New Guinea. We searched for events deeper than 100 km with magnitudes larger than 5.5, recorded between 2010 and 2020 available through the Incorporated Research Institutions for Seismology (IRIS) data management center (DMC).

We searched for promising events based on high SNR and simple P-waveform shape. We found 39, 17, and 7 events meeting these criteria recorded at the Warramunga Array (WRA), Alice Springs Array (ASAR), and Pilbara Array (PBAR), respectively. However, the majority of these events did not have sufficiently simple enough P-waveforms, which are used as a proxy for the source wavelet, to be able to model them adequately. Thus, we narrowed this list down to eight events with the highest SNR and the simplest P-wave. Locations for the eight events we selected are provided in Table 2 and shown in Figure 12. Three ScP bounce points for these events sample the eastern portion of the anomalous region we infer east of the Philippines, one samples the region north of Papua New Guinea, and four sample the region between the two.

### Most Parsimonious ULVZ locations



**Figure 11.** The most parsimonious heterogeneity locations are shown as filled contours based on the 1% misfit map (Figure 10c). Filled contours show the probability of heterogeneity existence at probabilities greater than 0.125 (dark green), 0.25 (purple), 0.5 (light green), and 0.75 (red).

We quantify possible ULVZ parameters (thickness,  $\delta V_p$ ,  $\delta V_s$ , and density) for these data using the Bayesian waveform inversion of ScP waves (Pachhai et al., 2014, 2015). In this approach, ULVZ parameters are assigned prior information, which is combined with the data information to provide the posterior probability densities (PPDs) of ULVZ parameters. The priors are set to be uniform over the parameter bounds so that the PPDs are not biased by our assumption about the ULVZ parameters. The data information is introduced in the inversion through the likelihood function, which is derived assuming a Gaussian distribution of data errors. Here, data errors are approximated by the difference between observed and synthetic ScP waveforms. Before we compute the likelihood function, synthetic and observed waveforms are aligned. Previous studies have aligned waveforms using cross-correlation; however, this may lead to cycle skipping when both P and ScP waveforms have multiple arrivals as observed in these data (e.g., see event #6, Figure 13f). Therefore, we also use a maximum amplitude-based alignment to compute the likelihood. For more details on the derivation of likelihood function, we refer to Pachhai et al. (2014, 2015).

The inversion of ULVZ parameters from the ScP waveforms is a nonlinear problem and no analytical solution exists. Therefore, we apply a parameter sampling approach to compute the PPD of ULVZ parameters. In this approach, we randomly draw new ULVZ parameters from the PPD by perturbing the current ULVZ parameters. The randomly drawn parameters are then accepted or rejected based on the likelihood ratio of the perturbed ULVZ model with respect to the current ULVZ model. In particular, a random number between 0 and 1 is selected. If the likelihood ratio is greater than the random number, the perturbed model is accepted and the model is updated. If the likelihood ratio is lower than the random number, then the perturbed model is rejected and the current ULVZ parameters are retained for the next step and the process is repeated. This procedure is repeated for thousands of iterations until convergence is achieved. Finally,

**Table 1**  
*Potential ULVZs Identified in This Study*

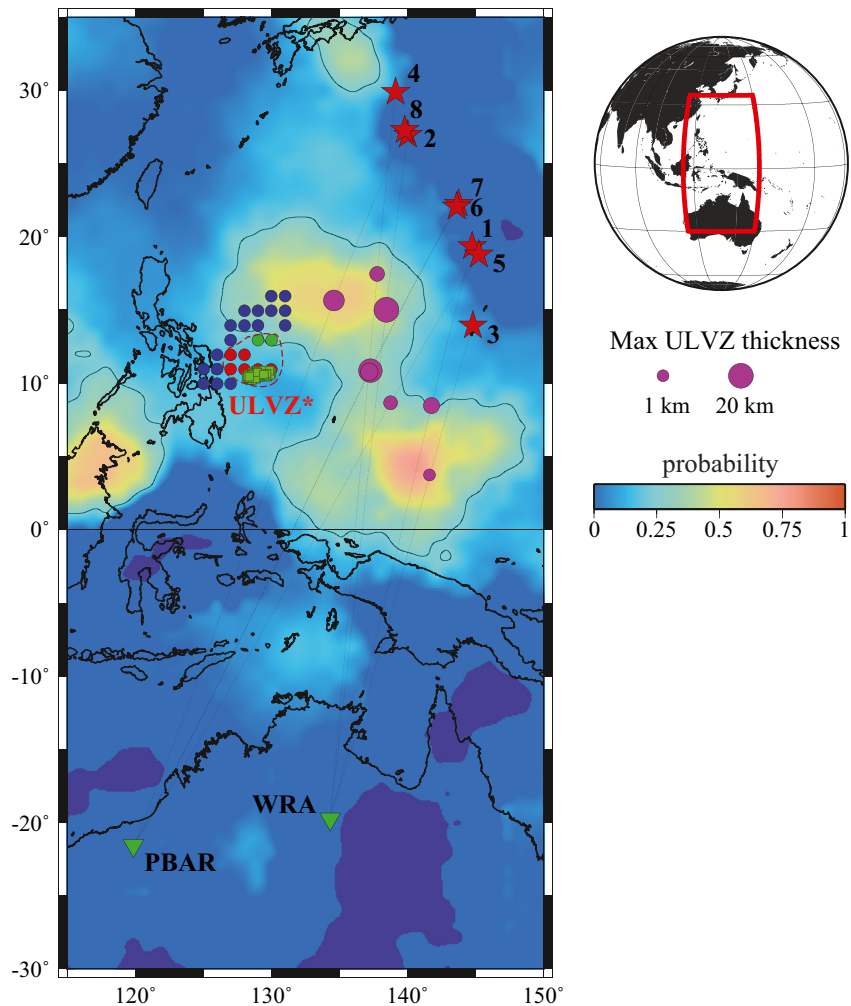
Possible ULVZ location	Previous detections	Probability	References of positive detections
Amazon	Yes	0.5	Zou et al., 2007
Caroline	No	0.5	NA
Central Africa	Yes	0.25	Helmberger et al., 2000; Ni & Helmberger, 2001
Coral Sea	Yes	0.25	Jensen et al. 2013; Idehara et al. 2007; Thomas et al. 2009; Brown et al. 2015; Rost et al. 2006; Rost & Revenaugh 2003; Rost et al. 2005
East Asia	No	0.5	NA
East Africa	No	0.25	NA
Eastern Europe	Yes	0.25	Thomas et al. 1999
Florida	Yes	0.25	Thorne et al. 2019
Galapagos	Yes	0.125	Vanacore and Niu 2011; Cottaar and Li 2019
Iceland	Yes	0.125	Yuan and Romanowicz 2017; Helmberger et al. 1998
North Africa	No	0.25	NA
North of Papua New Guinea	No	0.75	NA
Northeast Pacific	Yes/no	0.25	Revenaugh and Meyer 1997
Northern Mexico	Yes	0.25	Thorne et al. 2019; Havens and Revenaugh 2001
Philippine Sea	No	0.5	NA
Samoa	Yes	0.75	Wen and Helmberger 1998; Thorne et al. 2013; Reasoner and Revenaugh 2000; Zhang et al. 2009
Solomon Sea	No	0.25	NA
South America	Yes	0.5	Vanacore et al. 2016
South China Sea	Yes	0.5	Idehara et al. 2007; Jensen et al. 2013; Yao and Wen 2014
South Atlantic	Yes	0.5	Vanacore et al. 2016; Simmons and Grand 2002; Ni and Helmberger 2001
Southeast Asia	No	0.5	NA
Southern Japan	No	0.25	NA
Southern Ocean	No	0.125	NA
Spain	Yes	0.125	Rost and Garnero 2006
Western Canada	Yes	0.5	Rondenay and Fischer 2003
Western South America	No	0.125	NA

**Table 2**  
*ScP Data Used in This Study*

Event #	Array	Date	Latitude°	Longitude°	Depth (km)
1	WRA	8 Mar 2010	19.35	144.76	447
2	WRA	12 Jan 2011	26.97	140.02	524
3	WRA	18 May 2011	13.93	144.80	154
4	WRA	22 Aug 2016	29.89	139.13	404
5	PBAR	24 Nov 2015	18.77	145.26	587
6	PBAR	5 Jan 2016	22.06	143.61	158
7	PBAR	20 Oct 2016	22.24	143.75	104
8	PBAR	7 Sep 2017	27.28	139.80	451

we extract the maximum likelihood (minimum misfit) model for physical interpretation.

Here, we parameterize the ULVZ as a single layer in 1-D and use the WKBJ method (Chapman & Orcutt, 1985) to compute the Green's functions for the ScP waves. We use the direct P-wave as an approximation of the source-time function, which we attenuate with a  $t^*$  operator similar to Hansen et al. (2020). Then the attenuated P-wave is convolved with the Green's functions and filtered with the same bandpass filter used for the observed ScP waveforms. The velocity and density outside of the ULVZ are allowed to perturb within a narrow band ( $\pm 0.5\%$ ) from the ak135 reference model (Kennett et al., 1995), while velocity and density of the ULVZ layer are allowed to perturb in wider bands based on the previous studies. In particular, the height of the ULVZ is allowed to change from

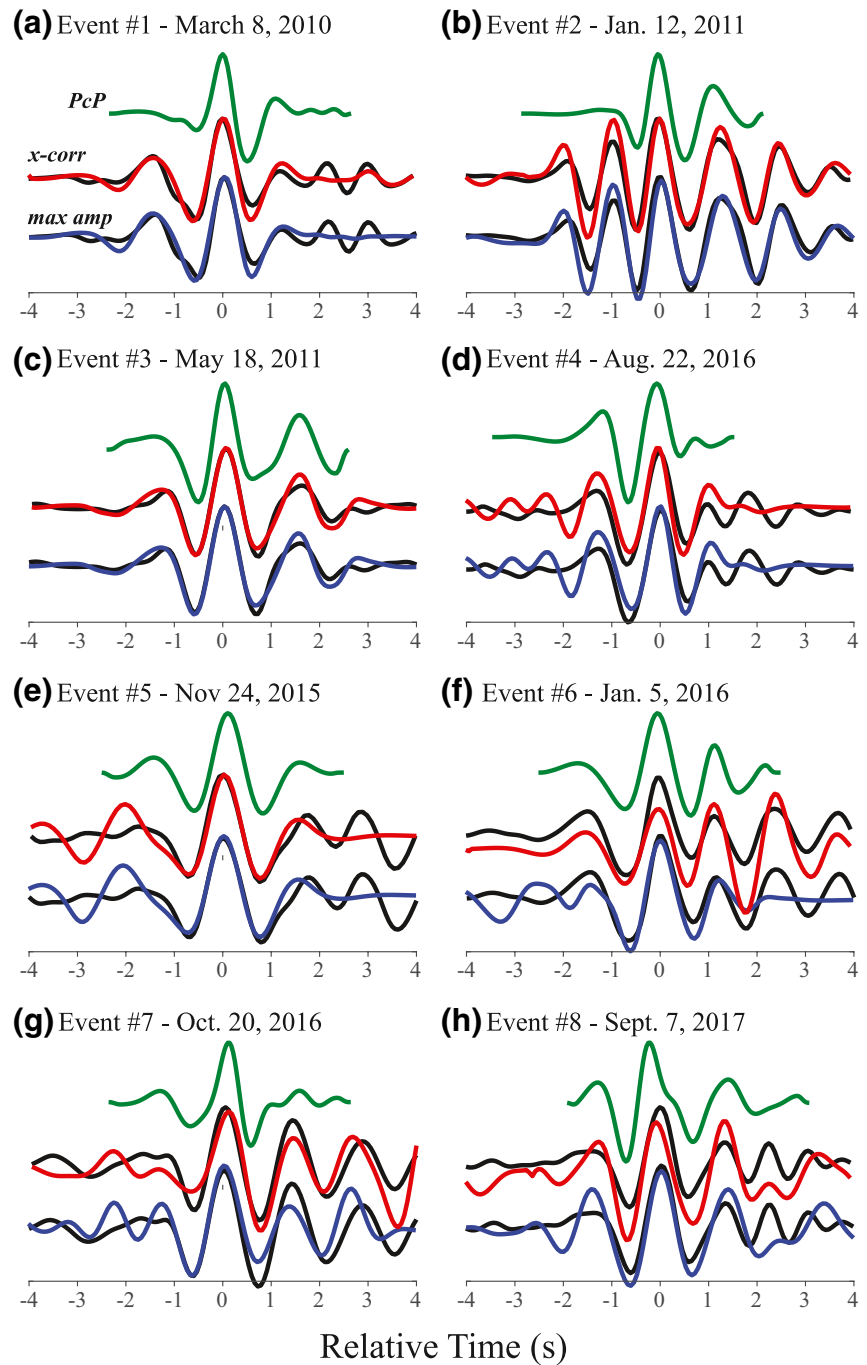


**Figure 12.** Locations of events for ScP data examined in this study are drawn with red stars and numbered according to Table 2. Seismic arrays used are the Warramunga, Australia (WRA), and Pilbara, Australia (PBAR) arrays. Array centers are drawn with inverted triangles. ScP bounce points on the CMB are shown with magenta circles and scaled in size to a maximum modeled ULVZ thickness (see Table 3). The background is the ULVZ probability from this study with a contour drawn around the 0.25 probability. A red dashed line is drawn and labeled ULVZ\*, which shows the approximate location of a previously identified ULVZ (Idehara et al., 2007; Pachhai et al., 2014). Green squares show ScP bounce point locations of Pachhai et al. (2014). ScP bounce points from Idehara et al. (2007) are shown as blue circles (no ULVZ identified), red circles (ULVZ identified), and green circles (maybe ULVZ identified).

0 to 35 km,  $\delta V_p$  is perturbed between  $-30\%$  and  $0\%$ ,  $\delta V_s$  is perturbed between  $-50\%$  and  $0\%$ , and  $\delta\rho$  is perturbed between  $0\%$  and  $30\%$ .

The results of this analysis, based on both alignments, are summarized in Table 3. ScP waveforms and best-fitting ULVZ models are shown in Figure 13. Of these events, events #2, #4, and #7 show the strongest evidence for ULVZ presence with complex ScP waveforms that are not explained by a simple PREM model. Event #2 displays the highest degree of ScP complexity and provides the strongest evidence for ULVZ presence. The best-fit model for this event is for a ULVZ with a thickness of 20 km and  $\delta V_s$  of  $-21\%$  to  $-23\%$ , with little to no  $\delta V_p$  change. Other events examined in this study appear more similar to PREM, yet are challenging to describe solely with the PREM model. These events are better modeled with a thin ULVZ structure.

Thorne et al. (2020) identified 33 events that cross the Philippine region and 32 events that cross the region north of Papua New Guinea that show highly anomalous SPdKS waveforms. In Figures 14 and 15, we show



**Figure 13.** Waveforms for ScP events are analyzed. Vertical component velocity seismograms are shown for each event listed in Tables 1 and 2. In each panel, the top trace (green) is the P-wavelet used that is stretched using the best-fit  $t^*$  operator. The ScP observation is shown in black and the best-fit model where cross-correlation is used for alignment is shown in red, whereas the best-fit model where the maximum amplitude is used for alignment is shown in blue.

data from one of the highest quality of these events (October 7, 2009). Figure 14a shows both source- and receiver-side Pd arcs for the SPdKS paths, which show a tight cluster of anomalous recordings (red Pd arcs) on the source-side that are spread out over a large lateral area on the receiver-side giving indication that source-side velocity heterogeneity is likely. Crossing coverage from additional events also provides evidence for source-side heterogeneity. A detailed plot of the Pd arcs on the source-side is shown in Figure 14b. The

**Table 3**  
ScP Results

Event #	Alignment	h (km)	$\delta V_P$ (%)	$\delta V_S$ (%)	$\delta \rho$ (%)
1	x-corr	6.15	-29.5	-19.3	0.0
1	max amp	6.28	-28.2	-21.4	0.0
2	x-corr	20.95	-4.8	-21.5	27.4
2	max amp	19.67	0.0	-23.2	8.2
3	x-corr	1.95	0.0	-42.2	3.9
3	max amp	1.7	-5.3	-49.6	12.2
4	x-corr	4.10	-9.4	-49.9	29.9
4	max amp	4.07	-7.3	-49.9	29.6
5	x-corr	2.37	0.0	-45.2	5.5
5	max amp	2.46	0.0	-45.4	5.8
6	x-corr	1.85	-3.8	-46.6	28.5
6	max amp	2.46	0.0	-44.9	10.84
7	x-corr	8.7	-1.6	-43.9	0.0
7	max amp	18.6	-28.4	-40.4	0.0
8	x-corr	14.0	-29.6	-45.7	0.0
8	max amp	2.35	-4.1	-49.6	0.0

anomalous Pd arcs span the region of high probability to the east of the Philippines, but miss the ULVZ identified using ScP in previous studies (Idehara et al., 2007; Pachhai et al., 2014). Highly anomalous waveforms that have been identified for this event appear to cluster in two azimuths from 31° to 41° and 41° to 51°. The second azimuth bin (41°–51°) overlaps ScP data discussed above that shows the best evidence for ULVZ existence from the ScP data examined in this study.

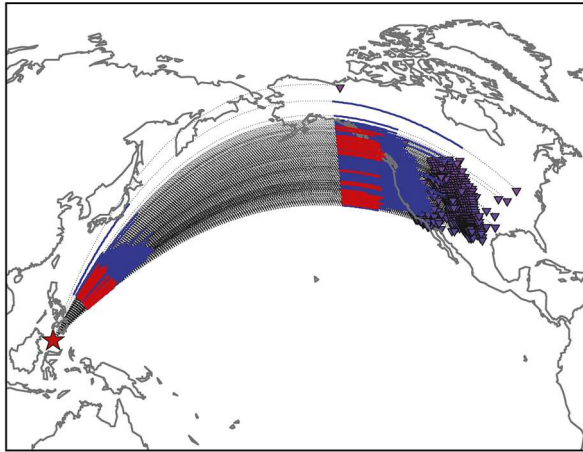
SPdKS waveforms for both azimuth bins are shown in Figure 15. Both azimuthal bins show remarkably similar waveforms, however, those for the larger azimuths show slightly larger delays in SKS and thus we stack waveforms in the separate azimuthal bins in order to preserve the waveform character. These waveforms display complexity that can be modeled with a ULVZ. Here we compare waveforms with 2-D synthetic seismograms computed using the PSVaxi method (Jahnke, 2009). We use a model bank of 975 boxcar-shaped ULVZ models from two sources. First, we use the 675 ULVZ models created for the study of Thorne et al. (2013), augmented by another 306 ULVZ models computed to coarsely span the full range of possible ULVZ parameters. Parameters in the augmented model bank are (1)  $\delta V_S$  varied from -40% to -10% in 10% increments, (2)  $\delta V_P$  varied from -40% to -10% in 10% increments, (3) thickness was allowed to be 10, 20, or 40 km, (4) ULVZ length in the great circle arc direction was allowed to be 3°, 6°, or 12°, and (5) the ULVZ angular position of the left edge was allowed to be 8°, 11°, 14°, 17°, or 20°. The models used in Thorne et al. (2013) were irregularly spaced but were computed for a range of thicknesses from 5 to 30 km, lengths of 1.5°, 3°, 6°, or 12°

and with a range in edge positions from 1° to 19° depending on the ULVZ length. For example, small-scale ULVZ models (e.g., 1.5°) were not computed for small edge positions where SPdKS would not interact with the ULVZ. For that study only 3:1  $\delta V_S$ :  $\delta V_P$  velocity ratios were examined, which is not the case for the augmented models described above, which included both 1:1 and 2:1 velocity ratios.

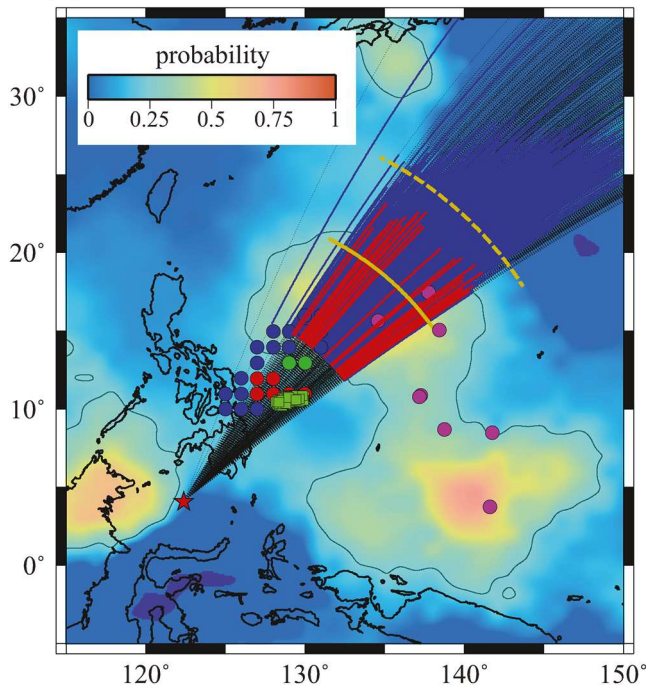
Of all the ULVZ models we compared, several models provided an adequate fit to the data. However, all of the best-fitting ULVZ models had an angular position of 19° or 20°. The length in the great circle arc direction was less well-defined with model lengths of 6° and 12° fitting equally as well. The inferred position of the ULVZ is indicated by the yellow lines in Figure 14 (edge 19°), noting that the far edge of the ULVZ is not well constrained by this event and hence drawn in with a dashed line. Models with thicknesses less than 20 km and  $\delta V_S$  and  $\delta V_P$  combinations with reductions less than 20% fit these data equally well. No models with thickness of over 20 km or velocity reductions greater than 20% adequately fit these data. Comparison between data and synthetic are shown in Figures 15b and 15d for one of the acceptable ULVZ models. The high-amplitude SPdKS arrivals at distances from 110° to 115° are readily reproduced by this ULVZ model; however, this model is not unique and an SPdKS post-cursor is apparent in the synthetics which is not apparent in these data. This post-cursor appears to arise from the far edge of the ULVZ, which could indicate that the length of the ULVZ model in the great circle arc direction is incorrect, the 2-D boxcar-shaped ULVZ model is not correct, or the 3-D ULVZ structure is important.

By themselves, ScP data examined in this study may not be strongly indicative of ULVZ presence, except perhaps for event #2, which has a remarkably complex waveform. However, in combination with SPdKS observations presented here, it is likely that a previously unknown ULVZ exists to the east of the Philippines. This appears to be distinct from the ULVZ described by Idehara et al. (2007) and Pachhai et al. (2014), because (1) Pd paths examined here do not interact with that ULVZ and (2) our waveform modeling of SPdKS indicates a ULVZ farther to the northeast. The ULVZ modeling results presented in this paper must be considered as preliminary as Thorne et al. (2020) found more than 60 events with highly anomalous SPdKS records for the Philippine and Papua New Guinea region, and a thorough search for ScP arrivals in the region has not yet been conducted. Nonetheless, ScP and SPdKS data presented here provide evidence that the inversion results for seismic heterogeneities presented in this paper are meaningful and may be related

(a) Source- and receiver-side paths



(b) Detail of source-side



**Figure 14.** SPdKS paths for an event occurring on October 7, 2009. (a) Source- and receiver-side CMB Pd segments are shown as either highly anomalous (red) or not highly anomalous (blue) as identified in Thorne et al. (2020). The dashed line shows the great circle path between the event (red star) and receivers (inverted triangles). (b) Detail on the source-side region. The background is ULVZ probability from this paper with a contour drawn at the 0.25 probability level. Magenta circles are ScP bounce point locations from this study. Green squares show ScP bounce point locations of Pachhai et al. (2014) in which a ULVZ was located. ScP bounce points from Idehara et al. (2007) are shown as blue circles (no ULVZ identified), red circles (ULVZ identified), and green circles (maybe ULVZ identified). The solid yellow line indicates the location of the near edge of the ULVZ as inferred from 2-D modeling of SPdKS data. The dashed yellow line is drawn 6° away from the solid line.

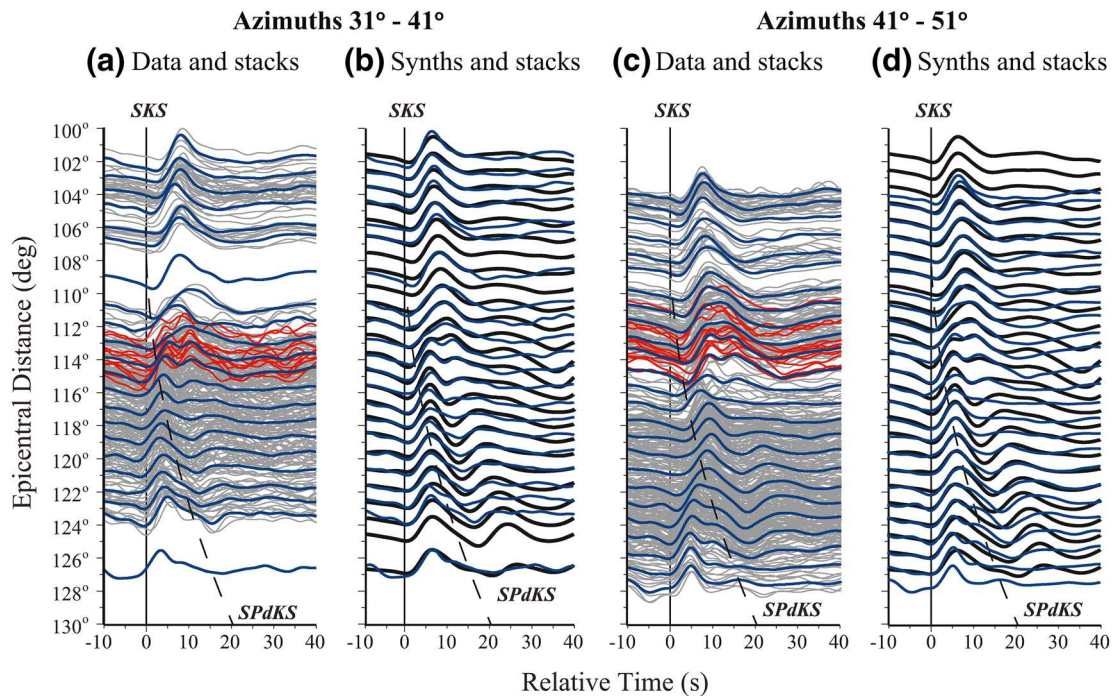
to the ULVZ structure. This region should receive further attention in future studies.

## 6. Discussion and Conclusion

### 6.1. The Nature of Inferred Heterogeneities

The heterogeneities inferred in this paper may be related to ULVZs, although we are hesitant to label all of these features as being ULVZs as we do not attempt to model elastic parameters for any of the features minus one region to the east of the Philippines. Hence, it may be premature to state that all of the features inferred here are indeed ULVZs. A whole range of lower mantle features that could potentially generate anomalous SPdKS waveforms have been observed, including (1) core rigidity zones (Rost & Revenaugh, 2001), (2) lower mantle scatterers (Ma & Thomas, 2020), (3) sharp LLVP edges (Ward et al., 2020), and (4) high-velocity anomalies related to past subduction and/or D'' layering (Whittaker et al., 2016; Zhan et al., 2014). None of these features have previously been shown to generate anomalous SPdKS waveforms; however, it is unlikely that their model spaces have been exhaustively searched and thus they cannot be ruled out here as potentially contributing to some of the inferred heterogeneities. Indeed, ULVZ models that generate the highly anomalous waveforms we observe only inhabit a small subspace of the possible ULVZ models. Nonetheless, the waveform features used to identify these heterogeneities have been modeled as ULVZs in several previous papers (Helmberger et al., 2000; Jensen et al., 2013; Thorne & Garnero, 2004; Thorne et al., 2013, 2019; Wen & Helmberger, 1998) and are also shown in this paper to produce these types of anomalous waveforms (Figures 3 and 4).

The distribution of heterogeneities shown in this paper (e.g., Figure 11) does not differ dramatically from that shown in Thorne et al. (2020), but several distinct differences exist. First, in Thorne et al. (2020), the ULVZ probabilities were principally based on the density of anomalous ray paths. Although the density of rays matters in the present study, it matters less so for the lower misfit calculations. Thus, our 5% misfit map (Figure 10a) looks most similar to the map of Thorne et al. (2020); however, our 1% misfit map (Figure 10c) contains many likely heterogeneities that are not highlighted in Thorne et al. (2020). Second, in the present study, we do not count SPdKS and SKPdS separately. If a heterogeneity is inferred beneath the SPdKS segment, we do not give additional credit to the SKPdS segment. This results in some areas of high probability in Thorne et al. (2020) not showing as high a probability as in Figure 10. Most notably, regions of high probability in Thorne et al. (2020) that do not show as high a probability as in this paper are located beneath the (1) Iceland, (2) Galapagos, (3) western S. America, and (4) the Bowie and Cobb hot spots. For example, with respect to heterogeneities beneath western S. America, Thorne et al. (2020) had a higher concentration of Pd arcs because it counted both source-side Pd arcs for paths going up to Alaska, and source-side Pd arcs for paths going to the western Pacific. However, the paths going to the western Pacific have receiver-side Pd arcs that cross the Samoa ULVZ. Hence, those records are already accounted for by the Samoa ULVZ and thus are not counted again. This results in a lower probability of finding a heterogeneity to the west of S. America.



**Figure 15.** Seismic waveforms and synthetics for the October 7, 2009 event. Seismic waveforms and synthetic seismograms are shown for azimuths from (a) and (b) 31° to 41° or from (c) and (d) 41° to 51°. In each panel radial component displacement seismograms are shown. Seismic data and stacks are shown in (a) and (c) where original seismic traces that are not identified as anomalous in Thorne et al. (2020) are drawn in gray, and seismic traces identified as anomalous are drawn in red. Stacks in 1° epicentral distance bins are drawn in blue. All arrivals are aligned to 0 on the PREM predicted SKS arrival time. In (b) and (d) the data stacks are drawn in blue and synthetic seismograms for a ULVZ model are drawn in black. Synthetic and data stacks are aligned on first break time. The ULVZ model has parameters  $\delta V_s = -15\%$ ,  $\delta V_p = -5\%$ ,  $\delta \rho = +10\%$ , thickness = 15 km, edge position = 19°, and length = 12°. In all panels, radial component displacement seismograms are shown and the PREM-predicted SPdKS arrival time is also indicated.

Heterogeneities determined in this study are dominated by source-side SPdKS recordings. As shown in Section 3, this may not be unexpected, as we are more likely to find a ULVZ on the source-side solely from geometric considerations. However, the nature of anomalous recordings on the source-side may be different than those for receiver-side recordings. The SPdKS wavefield is complex in the presence of a ULVZ. The generation of highly anomalous waveforms appears to be due to (1) constructive interference of scattered arrivals from ULVZ edges (see, e.g., Thorne et al., 2019; Vanacore et al., 2016) and (2) the generation of additional Pd energy from converted phases such as  $SP^{ULVZ}PdKS$  (see Thorne et al., 2020). It is unclear why more anomalous SPdKS-like energy is observed for ULVZs located on the source-sides of the path in the synthetic computations we have done thus far. Additional theoretical work, especially utilizing fully 3-D methods (Leng et al., 2019, 2020), should explore this further. Nonetheless, the identification of heterogeneous structures using SPdKS appears to be biased toward heterogeneities that exist on the source-side of the path. Thus, we have explored nearly 57% by surface area of the CMB for heterogeneities, but we must emphasize that it is more challenging to ascertain the heterogeneity existence in areas only covered by receiver-side SKPdS arrivals. In this study, only a couple of regions are identified as having high probability solely utilizing SKPdS arrivals. One region appears to the east of British Columbia, which has been suggested to contain a ULVZ in one past study (Rondenay & Fischer, 2003). Another location emerges in central Africa, which has also previously been suggested to contain a ULVZ (Helmberger et al., 2000).

In addition to inherent SKPdS waveform effects reducing the visibility of receiver-side ULVZs, our inversion method may also decrease their visibility for the highest misfit inversions. For example, more seismic traces may be interpreted as errors with a 5% data misfit than with a 1% misfit. As a result, the most parsimonious solution for 5% misfit has fewer ULVZs. This is because we can obtain 5% data misfit by excluding potential ULVZs only accounted for by a couple of data traces. Because receiver-side ULVZs will inherently manifest themselves in fewer anomalous seismograms, they will likely be suppressed in the inversions with the



largest data misfit. As a result, ULVZs that may be observed only with SKPdS such as beneath Galapagos or Iceland do not appear as likely in those inversions with a greater misfit.

In this study, our data coverage is 56.9% by CMB surface area. We can estimate the portion of the CMB that contains heterogeneities. Relative to the total surface area of the Earth we find that 0.7%, 4.6%, and 11.2% of the Earth is covered by heterogeneities at probabilities of 0.5, 0.25, and 0.125, respectively. Or relative to the area covered in this study we find 1.3%, 8.2%, and 19.7% of the Earth are covered by heterogeneities at a probability of 1/2, 1/4, and 1/8, respectively. However, there is uncertainty related to the size of the heterogeneities uncovered in this study. In particular, we emphasize that we assumed that heterogeneities could lie anywhere along the Pd arc. But, as demonstrated by synthetic experiments in Figures 3–5 and by the data shown in Figure 14, the heterogeneity may extend outside of the Pd rays. Examination of the synthetic database of ULVZ models used in modeling the Philippine data suggest that there is a fundamental tradeoff between ULVZ size and velocity decrease. This uncertainty will also contribute to uncertainty in how much of the CMB is covered by ULVZs.

## 6.2. Mega-Scale ULVZs

Some of the heterogeneities uncovered in this study may be as large as any previously discovered. Some of the largest ULVZs discovered thus far are located beneath Iceland, Hawaii, and Samoa. The Icelandic ULVZ was modeled as a cylindrical feature on top of the CMB with a diameter of  $800 \pm 50$  km (Yuan & Romanowicz, 2017), whereas the Hawaiian ULVZ was modeled as a cylindrical feature with a diameter of roughly 910 km (Cottaar & Romanowicz, 2012). The Samoan ULVZ has been modeled as being a linear feature on the CMB approximately  $250 \times 800$  km on each side (Thorne et al., 2013). In this study, we see that the heterogeneity associated with the Samoan ULVZ may be roughly twice as large extending to approximately  $480 \times 1,600$  km on each side, using the region of probability  $\geq 0.5$  to estimate its boundaries. Here, we note that this size is not specifically determined through waveform modeling but is inferred as the size of the heterogeneities we found that explain these data on a  $4 \times 4^\circ$  grid. The size of the Samoan ULVZ does appear to be at least 1,600 km in length from north to south. As observed in the synthetic examples, the ULVZ is likely well-constrained azimuthally by the highly anomalous recordings. In this case, we observe highly anomalous recordings, such as those shown in Figure 2, from events extending up and down the Kermadec trench. The dataset used in this paper is considerably larger than that used in Thorne et al. (2013), and as a result a much larger heterogeneity emerges in this present work.

The heterogeneity inferred to lie to the east of the Philippines and north of Papua New Guinea may also represent a single large-scale ULVZ. The regions of probability  $>0.5$  (Figure 11) suggest that two separate ULVZs may exist here. However, highly anomalous SPdKS recordings exist spanning the whole length between these two regions, and thus this may be just a single ULVZ. If so, it would be of the order of  $600 \times 1,200$  km, making it comparable in size to the Samoan ULVZ. Similarly, if contiguous ULVZs exist beneath N. or S. America (e.g., everything with probability  $>0.25$ ) and if these regions prove to have ULVZ-like elastic parameters, then these ULVZs could be larger than the Samoan ULVZ.

Seismic waveform simulations in 2-D suggest that there are tradeoffs between seismic velocity and ULVZ size. That is, in order to replicate the highly anomalous waveforms we observe for ULVZs with these large physical sizes, thinner ULVZs with more modest velocity reductions may be required. That is, if  $\delta V_S$  and  $\delta V_P$  reductions are greater than 20% and thicknesses are greater than 20 km, the predicted SPdKS waveforms may no longer resemble what is observed in real data. This could imply that some previous studies have overestimated the magnitude of the P- and/or S-wave velocity decreases associated with ULVZs. For example, if the Samoa ULVZ is as large as predicted in this study, then an S-wave decrease of the order of 40% as indicated by Thorne et al. (2013) is likely too large of a velocity decrease. We are currently re-examining the elastic parameters associated with the Samoan ULVZ.

The implications of such large-scale ULVZs is not clear. A number of possible ULVZ compositional anomalies have been suggested, including Fe-rich post-perovskite (Mao et al., 2006), Fe-rich ferropericlaste (fp) (Wicks et al., 2010, 2017), and ferric-Fe enriched post-perovskite (Stackhouse & Brodholt, 2008). Of all of these scenarios, some form of Fe-enrichment appears necessary. However, partially molten ULVZs, either

through the melting of Mid-Ocean Ridge Basalt (MORB) (Andraut et al., 2014) or through melting of accumulations of material with lower melting temperature such as fp (Berryman, 2000) may also lead to regions with excess Fe (Pradhan et al., 2015). The effects of Fe-rich materials in the deep mantle have been discussed in past studies, although the effects of large-scale Fe-rich materials may not be fully understood. Because an Fe-rich material such as a Fe-rich fp metalizes at CMB pressures and temperatures, it is expected that its thermal conductivity may increase (Manga & Jeanloz, 1996). Although a recent study has suggested that the thermal conductivity of Fe-rich fp may actually be lower (Hsieh et al., 2018), the effects of large-scale ULVZs with thermal conductivity variations from ambient mantle have not been explored geodynamically. The electrical conductivity of Fe-rich ULVZs is also expected to be larger than ambient mantle, which could affect the geodynamo and alter pole reversals (Glatzmaier et al., 1999) and contribute to core nutation (Buffett et al., 2000). It has been argued in the past that ULVZs may indeed have higher electrical conductivity, because virtual geomagnetic pole (VGP) positions appear to pass through the Americas and Asia due to highly conductive ULVZs being located beneath the central Pacific and Africa (Costin & Buffett, 2004). This may have to be re-examined if the large-scale heterogeneities we infer in this paper beneath Asia and the Americas have ULVZ-like properties, noting that previous studies have already provided evidence for ULVZs in these locations (see Table 1).

However, some caution must be maintained as the initial efforts in modeling SPdKS waveforms in 3-D that were conducted for this study suggest that the ULVZ shape in 3-D is an important factor in controlling the SPdKS waveform shape. In the handful of synthetic experiments conducted in this paper, this was most evident for the large-scale ULVZ (Figure 5a). In this case, the 3-D synthetics were computed for a cylindrical-shaped ULVZ in 3-D. This ULVZ shape focused on the energy passing through it and the waveforms appeared much more anomalous than those computed in the 2-D case. Thus, we can infer that there are also tradeoffs with the ULVZ shape. However, the importance of the ULVZ shape is virtually unknown at this point due to the large computational cost of computing 3-D synthetics and only a handful of computations have been performed thus far. We only computed 3-D synthetics for a perfect cylindrical shape, which may be unrealistic in the real Earth and the focusing effects observed therein may not persist for more realistic ULVZ boundaries. An animation showing differences between 2-D and 3-D wave propagation effects is included in the supplemental materials.

Not all heterogeneities uncovered in this study are necessarily mega-sized ULVZs however. Several smaller heterogeneities (e.g., with diameters of the order of 250 km) are also identified in this study. However, as demonstrated in Figure 4, it is difficult to find small receiver-side ULVZs with the SKPdS phase. Thus, it is likely that this analysis has not uncovered all ULVZs on the receiver-side.

### 6.3. Correlation with Other Features

As discussed in Section 1, previous studies have given much attention to correlations between ULVZs and other phenomena such as hot spot volcanism and LLVP boundaries. In this study, we have greater coverage of the CMB area than previous studies, and thus it is also tempting to examine correlations between our heterogeneity distribution and lower mantle features. However, in this study we only utilize a subset of the data we have collected (constrained to epicentral distances between 106° and 115°) that are highly sensitive to the ULVZ structure, as a first attempt at locating heterogeneities. The entire SPdKS dataset (see Thorne et al., 2020) has approximately 90% coverage of the CMB area, although that is including data at epicentral distances of up to 130°, which are less sensitive to ULVZ structures (Thorne & Garnero, 2004). Correlations with ULVZs and other features are likely most meaningful when the majority of the CMB has been probed and when we have completed the analysis with the full dataset. Nonetheless, we have compared our current heterogeneity distribution with seismic velocities from three tomographic models as described here: (1) S-wave model SEMUCB (French & Romanowicz, 2015) was chosen as it shows broad plumes extending from hot spot volcanoes to the base of the CMB rooted in ULVZs; (2) S-wave model GyP-Sum (Simmons et al., 2010), which was a joint P- and S- inversion and among the first tomographic models to show agreement between LLVPs in P- and S-wave tomography; and (3) P-wave model DETOX-P2 (Hosseini et al., 2019), which is one of the most recent P-wave models based on over 1 million Pdiff waveforms.

Visually, model DETOX-P2 appeared most similar to our ULVZ distribution and as it was primarily derived from Pdiff waveforms one might expect that it is related to low velocities observed in SPdiffKS.

We computed the correlation as outlined in Ray and Anderson (1994) with all heterogeneity regions with probabilities  $\geq 0.125$ , 0.25, and 0.5 (for the 1% misfit distribution). However, none of these distributions were significantly correlated with low or high seismic velocities in any of the models. This is not surprising as we see possible large heterogeneities near LLVPs, such as the Samoa mega-ULVZ. However, other large heterogeneities also exist near regions of past down-wellings such as beneath Mexico, East Asia, or S. America. Due to the proximity of many of these large heterogeneities to past subduction in our distribution, it would similarly be unlikely that they correlate significantly with LLVP boundaries or hot spots, although we did not explore this quantitatively. It is likely that ULVZs are not so simply related to lower mantle phenomena as either being related to hot spot volcanism or dense dregs that have been swept toward LLVP boundaries (Li et al., 2017). Rather, the diversity of locations in which ULVZs are observed argues that either (1) everything we refer to as ULVZ is not all the same thing, (2) we are seeing a feature that is in different stages of its life-cycle, or (3) we are seeing a feature that may arise through multiple different processes in the Earth.

#### 6.4. Conclusion

In this paper, we presented a new probability map of a distribution of heterogeneities based on SPdKS observations. Although we cannot state that these heterogeneities are specifically ULVZs as we do not model the elastic parameters of each heterogeneity; modeling done in this paper and several previous papers have linked the types of SPdKS waveforms used for ULVZs. This distribution of heterogeneities is based on a dataset that covers 56.9% of the CMB by surface area, and suggests that as much as 8%–20% of the CMB may be covered by ULVZ-like heterogeneity. The inferred distribution reveals that all lower mantle provinces may contain heterogeneities including areas associated with plume roots of hot spot volcanoes, LLVPs, and LLVP boundaries, and high seismic velocities areas that may be related to deep subduction. The distribution also reveals that some ULVZs may be larger than previously thought and that several large-scale mega-ULVZs may exist. The consequences of such large ULVZs are unknown but may play an important role in both mantle and core dynamics. This distribution may provide some focus on areas within which to make more targeted searches for ULVZs using other phases.

#### Data Availability Statement

Data processing and analysis used SACTOOLS (Thorne, 2018). The facilities of IRIS Data Services, and specifically the IRIS Data Management Center, were used for access to ScP waveforms, related metadata, and/or derived products used in this study. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience (SAGE) Award of the National Science Foundation under Cooperative Support Agreement EAR-1851048.

#### References

- Andraut, D., Pesce, G., Bouhifd, M. A., Bolfan-Casanova, N., Henot, J.-M., & Mezouar, M. (2014). Melting of subducted basalt at the core-mantle boundary. *Science*, *344*(6186), 892–895. <https://doi.org/10.1126/science.1250466>
- Berryman, J. G. (2000). Seismic velocity decrement ratios for regions of partial melt in the lower mantle. *Geophysical Research Letters*, *27*(3), 421–424. <https://doi.org/10.1029/1999gl008402>
- Boschi, L., Becker, T. W., & Steinberger, B. (2007). Mantle plumes: Dynamic models and seismic images. *Geochemistry, Geophysics, Geosystems*, *8*(10). <https://doi.org/10.1029/2007gc001733>
- Bower, D. J., Wicks, J. K., Gurnis, M., & Jackson, J. M. (2011). A geodynamic and mineral physics model of a solid-state ultralow-velocity zone. *Earth and Planetary Science Letters*, *303*(3–4), 193–202. <https://doi.org/10.1016/j.epsl.2010.12.035>
- Brown, S. P., Thorne, M. S., Miyagi, L., & Rost, S. (2015). A compositional origin to ultralow-velocity zones. *Geophysical Research Letters*, *42*(4), 1039–1045. <https://doi.org/10.1002/2014gl062097>
- Buffett, B. A., Garnero, E. J., & Jeanloz, R. (2000). Sediments at the top of Earth's core. *Science*, *290*(5495), 1338–1342. <https://doi.org/10.1126/science.290.5495.1338>
- Chapman, C. H., & Orcutt, J. A. (1985). The computation of body wave synthetic seismograms in laterally homogeneous media. *Reviews of Geophysics*, *23*(2), 105. <https://doi.org/10.1029/rg023i002p0105>

#### Acknowledgments

The authors thank Editor Claudio Faccenna and two anonymous reviewers for constructive reviews which improved this manuscript. The authors acknowledge the University of Utah Center for High Performance Computing (CHPC) for computer resources and support, as well as the UK's supercomputer facility ARCHER, where some of the modeling was undertaken. Surya Pachhai acknowledges the Argo cluster at the International Center for Theoretical Physics, Trieste, Italy. Michael S. Thorne and Surya Pachhai were partially supported by NSF grant EAR-1723081. Kuangdai Leng was fully supported by NERC grant NE/R012199/1. Tarje Nissen-Meyer and Sebastian Rost were partially supported by NERC grant NE/R012199/1. Figures were drawn using the Generic Mapping Tools (Wessel & Smith, 1998).

- Costin, S. O., & Buffett, B. A. (2004). Preferred reversal paths caused by a heterogeneous conducting layer at the base of the mantle. *Journal of Geophysical Research*, *109*(B6). <https://doi.org/10.1029/2003jb002853>
- Cottaar, S., & Li, Z. (2019). An ultra-low velocity zone beneath the Galapagos hotspot. American Geophysical Union Fall Meeting. Abstract D153A-0043.
- Cottaar, S., & Romanowicz, B. (2012). An unusually large ULVZ at the base of the mantle near Hawaii. *Earth and Planetary Science Letters*, *355*–356, 213–222. <https://doi.org/10.1016/j.epsl.2012.09.005>
- Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, *25*(4), 297–356. [https://doi.org/10.1016/0031-9201\(81\)90046-7](https://doi.org/10.1016/0031-9201(81)90046-7)
- Fitch, W. M. (1971). Toward defining the course of evolution: Minimum change for a specific tree topology. *Systematic Zoology*, *20*(4), 406–416. <https://doi.org/10.2307/sysbio/27.4.401>
- French, S. W., & Romanowicz, B. (2015). Broad plumes rooted at the base of the Earth's mantle beneath major hotspots. *Nature*, *525*(7567), 95–99. <https://doi.org/10.1038/nature14876>
- Garnero, E. J., Grand, S. P., & Helmberger, D. V. (1993). Low P-wave velocity at the base of the mantle. *Geophysical Research Letters*, *20*(17), 1843–1846. <https://doi.org/10.1029/93gl02009>
- Garnero, E. J., & Helmberger, D. V. (1998). Further structural constraints and uncertainties of a thin laterally varying ultralow-velocity layer at the base of the mantle. *Journal of Geophysical Research*, *103*(B6), 12495–12509. <https://doi.org/10.1029/98jb00700>
- Garnero, E. J., & McNamara, A. K. (2008). Structure and dynamics of Earth's lower mantle. *Science*, *320*(5876), 626–628. <https://doi.org/10.1126/science.1148028>
- Garnero, E. J., Revenaugh, J., Williams, Q., Lay, T., & Kellogg, L. H. (1998). Ultralow velocity zone at the core-mantle boundary. *The Core-Mantle Boundary Region*, 319–334.
- Garnero, E. J., & Vidale, J. E. (1999). ScPS<sub>1</sub> mathsemicolon\$ A probe of ultralow velocity zones at the base of the mantle. *Geophysical Research Letters*, *26*(3), 377–380. <https://doi.org/10.1029/1998gl900319>
- Glatzmaier, G. A., Coe, R. S., Hongre, L., & Roberts, P. H. (1999). The role of the Earth's mantle in controlling the frequency of geomagnetic reversals. *Nature*, *401*(6756), 885–890. <https://doi.org/10.1038/44776>
- Hansen, S. E., Carson, S. E., Garnero, E. J., Rost, S., & Yu, S. (2020). Investigating ultra-low velocity zones in the southern hemisphere using an Antarctic dataset. *Earth and Planetary Science Letters*, *536*, 116142. <https://doi.org/10.1016/j.epsl.2020.116142>
- Haupt, R. L., & Haupt, S. E. (2003). *Practical genetic algorithms*. John Wiley & Sons Inc. <https://doi.org/10.1002/0471671746>
- Havens, E., & Revenaugh, J. (2001). A broadband seismic study of the lowermost mantle beneath Mexico: Constraints on ultralow velocity zone elasticity and density. *Journal of Geophysical Research*, *106*(B12), 30809–30820. <https://doi.org/10.1029/2000jb000072>
- Helmberger, D., Ni, S., Wen, L., & Ritsema, J. (2000). Seismic evidence for ultralow-velocity zones beneath Africa and eastern Atlantic. *Journal of Geophysical Research*, *105*(B10), 23865–23878. <https://doi.org/10.1029/2000jb900143>
- Helmberger, D. V., Wen, L., & Ding, X. (1998). Seismic evidence that the source of the Iceland hotspot lies at the coremantle boundary. *Nature*, *396*(6708), 251–255. <https://doi.org/10.1038/24357>
- Hosseini, K., Sigloch, K., Tsekhmistrenko, M., Zaheri, A., Nissen-Meyer, T., & Igel, H. (2019). Global mantle structure from multifrequency tomography using P PP and P-diffracted waves. *Geophysical Journal International*, *220*(1), 96–141. <https://doi.org/10.1093/gji/ggz394>
- Hsieh, W.-P., Deschamps, F., Okuchi, T., & Lin, J.-F. (2018). Effects of iron on the lattice thermal conductivity of Earth's deep mantle and implications for mantle dynamics. *Proceedings of the National Academy of Sciences*, *115*(16), 4099–4104. <https://doi.org/10.1073/pnas.1718557115>
- Husebye, E. S., King, D. W., & Haddon, R. A. W. (1976). Precursors to PKIKP and seismic wave scattering near the mantle-core boundary. *Journal of Geophysical Research*, *81*(11), 1870–1882. <https://doi.org/10.1029/jb081i011p01870>
- Idehara, K., Yamada, A., & Zhao, D. (2007). Seismological constraints on the ultralow velocity zones in the lowermost mantle from core-reflected waves. *Physics of the Earth and Planetary Interiors*, *165*(1–2), 25–46. <https://doi.org/10.1016/j.pepi.2007.07.005>
- Jahnke, G. (2009). *Methods for seismic wave Propagation on Local and global Scales with finite differences* (PhD thesis). Munich: Ludwig Maximilians University.
- Jensen, K. J., Thorne, M. S., & Rost, S. (2013). SPdKS analysis of ultralow-velocity zones beneath the western Pacific. *Geophysical Research Letters*, *40*(17), 4574–4578. <https://doi.org/10.1002/grl.50877>
- Kennett, B. L. N., Engdahl, E. R., & Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, *122*(1), 108–124. <https://doi.org/10.1111/j.1365-246x.1995.tb03540.x>
- Kim, D., Lekić, V., Ménard, B., Baron, D., & Taghizadeh-Popp, M. (2020). Sequencing seismograms: A panoptic view of scattering in the core-mantle boundary region. *Science*, *368*(6496), 1223–1228. <https://doi.org/10.1126/science.aba8972>
- King, D. W., Haddon, R. A. W., & Cleary, J. R. (1974). Array analysis of precursors to PKIKP in the distance range 128 to 142. *Geophysical Journal International*, *37*(1), 157–173. <https://doi.org/10.1111/j.1365-246x.1974.tb02450.x>
- Leng, K., Korenaga, J., & Nissen-Meyer, T. (2020). 3-D scattering of elastic waves by small-scale heterogeneities in the Earth's mantle. *Geophysical Journal International*, *223*(1), 502–525. <https://doi.org/10.1093/gji/ggaa331>
- Leng, K., Nissen-Meyer, T., van Driel, M., Hosseini, K., & Al-Attar, D. (2019). AxiSEM3D: broad-band seismic wavefields in 3-D global earth models with undulating discontinuities. *Geophysical Journal International*, *217*(3), 2125–2146. <https://doi.org/10.1093/gji/ggz092>
- Li, M., McNamara, A. K., Garnero, E. J., & Yu, S. (2017). Compositionally-distinct ultra-low velocity zones on Earth's core-mantle boundary. *Nature Communications*, *8*(1). <https://doi.org/10.1038/s41467-017-00219-x>
- Manga, M., & Jeanloz, R. (1996). Implications of a metal-bearing chemical boundary layer in D for mantle dynamics. *Geophysical Research Letters*, *23*(22), 3091–3094. <https://doi.org/10.1029/96gl03021>
- Mao, W. L., Mao, H.-K., Sturhahn, W., Zhao, J., Prakapenka, V. B., Meng, Y., et al. (2006). Iron-rich post-perovskite and the origin of ultralow-velocity zones. *Science*, *312*(5773), 564–565. <https://doi.org/10.1126/science.1123442>
- Masson, Y., Cupillard, P., Capdeville, Y., & Romanowicz, B. (2013). On the numerical implementation of time-reversal mirrors for tomographic imaging. *Geophysical Journal International*, *196*(3), 1580–1599. <https://doi.org/10.1093/gji/ggt459>
- Ma, X., & Thomas, C. (2020). Small-scale scattering heterogeneities in the lowermost mantle from a global analysis of PKP precursors. *Journal of Geophysical Research*, Vol. *125*(3), e2019JB018736. <https://doi.org/10.1029/2019jb018736>
- McNamara, A. K., Garnero, E. J., & Rost, S. (2010). Tracking deep mantle reservoirs with ultra-low velocity zones. *Earth and Planetary Science Letters*, *299*(1–2), 1–9. <https://doi.org/10.1016/j.epsl.2010.07.042>
- McNamara, A. K., & Zhong, S. (2005). Thermochemical structures beneath Africa and the Pacific Ocean. *Nature*, *437*(7062), 1136–1139. <https://doi.org/10.1038/nature04066>
- Ni, S., & Helmberger, D. V. (2001). Probing an ultra-low velocity zone at the core mantle boundary with P and S waves. *Geophysical Research Letters*, *28*(12), 2345–2348. <https://doi.org/10.1029/2000gl012766>

- Pachhai, S., Dettmer, J., & Tkalčić, H. (2015). Ultra-low velocity zones beneath the Philippine and Tasman Seas revealed by a trans-dimensional Bayesian waveform inversion. *Geophysical Journal International*, 203(2), 1302–1318. <https://doi.org/10.1093/gji/ggv368>
- Pachhai, S., Tkalčić, H., & Dettmer, J. (2014). Bayesian inference for ultralow velocity zones in the Earth's lowermost mantle: Complex ULVZ beneath the east of the Philippines. *Journal of Geophysical Research: Solid Earth*, 119(11), 8346–8365. <https://doi.org/10.1002/2014jb011067>
- Pienkowska, M., Monteiller, V., & Nissen-Meyer, T. (2020). High-frequency global wavefields for local 3D structures by wavefield injection and extrapolation. *Geophysical Journal International*. <https://doi.org/10.1093/gji/ggaa563>
- Pradhan, G. K., Fiquet, G., Siebert, J., Auzende, A.-L., Morard, G., Antonangeli, D., et al. (2015). Melting of MORB at core-mantle boundary. *Earth and Planetary Science Letters*, 431, 247–255. <https://doi.org/10.1016/j.epsl.2015.09.034>
- Ramillien, G. (2001). Genetic algorithms for geophysical parameter inversion from altimeter data. *Geophysical Journal International*, 147(2), 393–402. <https://doi.org/10.1046/j.0956-540x.2001.01543.x>
- Ray, T. W., & Anderson, D. L. (1994). Spherical disharmonies in the Earth sciences and the spatial solution: Ridges hotspots, slabs, geochemistry and tomography correlations. *Journal of Geophysical Research*, 99(B5), 9605–9614. <https://doi.org/10.1029/94jb00340>
- Reasoner, C., & Revenaugh, J. (2000). ScP constraints on ultralow-velocity zone density and gradient thickness beneath the Pacific. *Journal of Geophysical Research*, 105(B12), 28173–28182. <https://doi.org/10.1029/2000jb900331>
- Revenaugh, J., & Meyer, R. (1997). Seismic evidence of partial melt within a possibly ubiquitous low-velocity layer at the base of the mantle. *Science*, 277(5326), 670–673. <https://doi.org/10.1126/science.277.5326.670>
- Robertsson, J. O. A., & Chapman, C. H. (2000). An efficient method for calculating finite-difference seismograms after model alterations. *Geophysics*, 65(3), 907–918. <https://doi.org/10.1190/1.1444787>
- Rondenay, S., Cormier, V. F., & Ark, E. M. V. (2010). SKS and SPdKS sensitivity to two-dimensional ultralow-velocity zones. *Journal of Geophysical Research*, 115(B4). <https://doi.org/10.1029/2009jb006733>
- Rondenay, S., & Fischer, K. M. (2003). Constraints on localized core-mantle boundary structure from multichannel broadband SKS coda analysis. *Journal of Geophysical Research*, 108(B11). <https://doi.org/10.1029/2003jb002518>
- Rost, S., & Garnero, E. J. (2006). Detection of an ultralow velocity zone at the core-mantle boundary using diffracted PKKPab waves. *Journal of Geophysical Research*, 111(B7). <https://doi.org/10.1029/2005jb003850>
- Rost, S., Garnero, E. J., & Williams, Q. (2006). Fine-scale ultralow-velocity zone structure from high-frequency seismic array data. *Journal of Geophysical Research*, 111(B9). <https://doi.org/10.1029/2005jb004088>
- Rost, S., Garnero, E. J., Williams, Q., & Manga, M. (2005). Seismological constraints on a possible plume root at the core-mantle boundary. *Nature*, 435(7042), 666–669. <https://doi.org/10.1038/nature03620>
- Rost, S., & Revenaugh, J. (2001). Seismic detection of rigid zones at the top of the core. *Science*, 294(5548), 1911–1914. <https://doi.org/10.1126/science.1065617>
- Rost, S., & Revenaugh, J. (2003). Small-scale ultralow-velocity zone structure imaged by ScP. *Journal of Geophysical Research*, 108(B1). <https://doi.org/10.1029/2001jb001627>
- Sambridge, M., & Drijkoningen, G. (1992). Genetic algorithms in seismic waveform inversion. *Geophysical Journal International*, 109(2), 323–342. <https://doi.org/10.1111/j.1365-246x.1992.tb00100.x>
- Sen, M. K., & Mallick, S. (2018). Genetic Algorithm with Applications in Geophysics. *Application of soft computing and intelligent methods in geophysics* (pp. 487–533). Springer International Publishing. [https://doi.org/10.1007/978-3-319-66532-0\\_7](https://doi.org/10.1007/978-3-319-66532-0_7)
- Simmons, N. A., Forte, A. M., Boschi, L., & Grand, S. P. (2010). GyPSuM: A joint tomographic model of mantle density and seismic wave speeds. *Journal of Geophysical Research*, 115(B12). <https://doi.org/10.1029/2010jb007631>
- Simmons, N. A., & Grand, S. P. (2002). Partial melting in the deepest mantle. *Geophysical Research Letters*, 29(11). <https://doi.org/10.1029/2001gl013716>
- Stackhouse, S., & Brodholt, J. P. (2008). Elastic properties of the post-perovskite phase of Fe<sub>2</sub>O<sub>3</sub> and implications for ultra-low velocity zones. *Physics of the Earth and Planetary Interiors*, 170(3–4), 260–266. <https://doi.org/10.1016/j.pepi.2008.07.010>
- Thomas, C., Kendall, J.-M., & Helffrich, G. (2009). Probing two low-velocity regions with PKPb-caustic amplitudes and scattering. *Geophysical Journal International*, 178(1), 503–512. <https://doi.org/10.1111/j.1365-246x.2009.04189.x>
- Thomas, C., Weber, M., Wicks, C. W., & Scherbaum, F. (1999). Small scatterers in the lower mantle observed at German broadband arrays. *Journal of Geophysical Research*, 104(B7), 15073–15088. <https://doi.org/10.1029/1999jb900128>
- Thorne, M. S. (2018). SACTOOLS v1.0.0. *Zenodo*. <http://doi.org/10.5281/zenodo.1314738>
- Thorne, M. S., & Garnero, E. J. (2004). Inferences on ultralow-velocity zone structure from a global analysis of SPdKS waves. *Journal of Geophysical Research*, 109(B8). <https://doi.org/10.1029/2004jb003010>
- Thorne, M. S., Garnero, E. J., & Grand, S. P. (2004). Geographic correlation between hot spots and deep mantle lateral shear-wave velocity gradients. *Physics of the Earth and Planetary Interiors*, 146(1–2), 47–63. <https://doi.org/10.1016/j.pepi.2003.09.026>
- Thorne, M. S., Garnero, E. J., Jahnke, G., Igel, H., & McNamara, A. K. (2013). Mega ultra low velocity zone and mantle flow. *Earth and Planetary Science Letters*, 364, 59–67. <https://doi.org/10.1016/j.epsl.2012.12.034>
- Thorne, M. S., Pachhai, S., Leng, K., Wicks, J. K., & Nissen-Meyer, T. (2020). New candidate ultralow-velocity zone locations from highly anomalous SPdKS waveforms. *Minerals*, 10(3), 211. <https://doi.org/10.3390/min10030211>
- Thorne, M. S., Takeuchi, N., & Shiomi, K. (2019). Melting at the edge of a slab in the deepest mantle. *Geophysical Research Letters*, 46(14), 8000–8008. <https://doi.org/10.1029/2019gl082493>
- Torsvik, T. H., Smethurst, M. A., Burke, K., & Steinberger, B. (2006). Large igneous provinces generated from the margins of the large low-velocity provinces in the deep mantle. *Geophysical Journal International*, 167(3), 1447–1460. <https://doi.org/10.1111/j.1365-246x.2006.03158.x>
- Torsvik, T., Steinberger, B., Cocks, L., & Burke, K. (2008). Longitude: Linking Earth's ancient surface to its deep interior. *Earth and Planetary Science Letters*, 276(3–4), 273–282. <https://doi.org/10.1016/j.epsl.2008.09.026>
- Vanacore, E., & Niu, F. (2011). Characterization of the D beneath the Galapagos Islands using SKKS and SKS waveforms. *Earthquake Science*, 24, 87–99. <https://doi.org/10.1007/s11589-011-0772-8>
- Vanacore, E. A., Rost, S., & Thorne, M. S. (2016). Ultralow-velocity zone geometries resolved by multidimensional waveform modeling. *Geophysical Journal International*, 206(1), 659–674. <https://doi.org/10.1093/gji/ggw114>
- Ward, J., Nowacki, A., & Rost, S. (2020). Lateral velocity gradients in the African lower mantle inferred from slowness space observations of multipathing. *Geochemistry, Geophysics, Geosystems*, 21(8), e2020GC009025. <https://doi.org/10.1029/2020gc009025>
- Wen, L., & Helmberger, D. V. (1998). A two-dimensional P-SV hybrid method and its application to modeling localized structures near the core-mantle boundary. *Journal of Geophysical Research*, 103(B8), 17901–17918. <https://doi.org/10.1029/98jb01276>
- Wessel, P., & Smith, W. H. F. (1998). New improved version of generic mapping tools released. *Eos, Transactions American Geophysical Union*, 79(47), 579. <https://doi.org/10.1029/98eo00426>

- Whittaker, S., Thorne, M. S., Schmerr, N. C., & Miyagi, L. (2016). Seismic array constraints on the D discontinuity beneath Central America. *Journal of Geophysical Research: Solid Earth*, *121*(1), 152–169. <https://doi.org/10.1002/2015jb012392>
- Wicks, J. K., Jackson, J. M., & Sturhahn, W. (2010). Very low sound velocities in iron-rich (Mg,Fe)O: Implications for the core-mantle boundary region. *Geophysical Research Letters*, *37*(15). <https://doi.org/10.1029/2010gl043689>
- Wicks, J. K., Jackson, J. M., Sturhahn, W., & Zhang, D. (2017). Sound velocity and density of magnesiowüstites: Implications for ultra-low-velocity zone topography. *Geophysical Research Letters*, *44*(5), 2148–2158. <https://doi.org/10.1002/2016gl071225>
- Williams, Q., Revenaugh, J., & Garnero, E. J. (1998). A correlation between ultra-low basal velocities in the mantle and hot spots. *Science*, *281*(5376), 546–549. <https://doi.org/10.1126/science.281.5376.546>
- Yao, J., & Wen, L. (2014). Seismic structure and ultra-low velocity zones at the base of the Earth's mantle beneath Southeast Asia. *Physics of the Earth and Planetary Interiors*, *233*, 103–111. <https://doi.org/10.1016/j.pepi.2014.05.009>
- Yuan, K., & Romanowicz, B. (2017). Seismic evidence for partial melting at the root of major hot spot plumes. *Science*, *357*, 393–397. <https://doi.org/10.1126/science.aan0760>
- Yu, S., & Garnero, E. J. (2018). Ultralow velocity zone locations: A global assessment. *Geochemistry, Geophysics, Geosystems*, *19*(2), 396–414. <https://doi.org/10.1002/2017gc007281>
- Zhang, Y., Ritsema, J., & Thorne, M. S. (2009). Modeling the ratios of SKKS and SKS amplitudes with ultra-low velocity zones at the core-mantle boundary. *Geophysical Research Letters*, *36*(19). <https://doi.org/10.1029/2009gl040030>
- Zhan, Z., Helmberger, D. V., & Li, D. (2014). Imaging subducted slab structure beneath the Sea of Okhotsk with teleseismic waveforms. *Physics of the Earth and Planetary Interiors*, *232*, 30–35. <https://doi.org/10.1016/j.pepi.2014.03.008>
- Zhao, C., Garnero, E. J., Li, M., McNamara, A., & Yu, S. (2017). Intermittent and lateral varying ULVZ structure at the northeastern margin of the Pacific LLSVP. *Journal of Geophysical Research: Solid Earth*, *122*(2), 1198–1220. <https://doi.org/10.1002/2016jb013449>
- Zou, Z., Leyton, F., & Koper, K. D. (2007). Partial melt in the lowermost mantle near the base of a plume. *Geophysical Journal International*, *168*(2), 809–817. <https://doi.org/10.1111/j.1365-246x.2006.03266.x>