Geophysical Journal International

Geophys. J. Int. (2017) **208,** 10–23
Advance Access publication 2016 October 8
GJI Seismology

doi: 10.1093/gji/ggw383

A Markov chain Monte Carlo with Gibbs sampling approach to anisotropic receiver function forward modeling

Erin A. Wirth,[1,*] Maureen D. Long[1] and John C. Moriarty[2]

[1]*Department of Geology and Geophysics, Yale University, New Haven, CT, USA. E-mail: ewirth@uw.edu*
[2]*Department of Astronomy, Yale University, New Haven, CT, USA*

Accepted 2016 October 5. Received 2016 October 3; in original form 2016 February 26; Editorial Decision 2016 October 4

SUMMARY

Teleseismic receiver functions contain information regarding Earth structure beneath a seismic station. *P*-to-*SV* converted phases are often used to characterize crustal and upper-mantle discontinuities and isotropic velocity structures. More recently, *P*-to-*SH* converted energy has been used to interrogate the orientation of anisotropy at depth, as well as the geometry of dipping interfaces. Many studies use a trial-and-error forward modeling approach for the interpretation of receiver functions, generating synthetic receiver functions from a user-defined input model of Earth structure and amending this model until it matches major features in the actual data. While often successful, such an approach makes it impossible to explore model space in a systematic and robust manner, which is especially important given that solutions are likely non-unique. Here, we present a Markov chain Monte Carlo algorithm with Gibbs sampling for the interpretation of anisotropic receiver functions. Synthetic examples are used to test the viability of the algorithm, suggesting that it works well for models with a reasonable number of free parameters ($<\sim$20). Additionally, the synthetic tests illustrate that certain parameters are well constrained by receiver function data, while others are subject to severe trade-offs—an important implication for studies that attempt to interpret Earth structure based on receiver function data. Finally, we apply our algorithm to receiver function data from station WCI in the central United States. We find evidence for a change in anisotropic structure at mid-lithospheric depths, consistent with previous work that used a grid search approach to model receiver function data at this station. Forward modeling of receiver functions using model space search algorithms, such as the one presented here, provide a meaningful framework for interrogating Earth structure from receiver function data.

Key words: Probability distributions; Body waves; Seismic anisotropy.

1 INTRODUCTION

Receiver functions are a popular tool for characterizing impedance contrasts beneath a seismic station. The *P*-to-*S* receiver function method relies on the notion that the coda after a teleseismic *P*-wave arrival represents a convolution of the incident *P* wave with Earth structure. Deconvolving the vertical component (i.e. incident *P*-wave energy) from the horizontal component wave train (i.e. *P*-to-*S* converted energy, referred to as a Ps phase) yields a time-series, or 'receiver function', containing information about sharp gradients in Earth structure beneath the receiver (e.g. Langston 1979).

Receiver functions have the potential to constrain Earth structure through analysis of Ps phase arrival times, amplitudes and polarities. The arrival time of *P*-to-*SV* phases (i.e. polarized in the source–receiver plane), relative to the incident *P*-wave arrival, reflects both the depth to the discontinuity and the isotropic velocity profile. Additionally, the *P*-to-*SV* phase amplitude is related to the strength and sharpness of the impedance contrast. If Earth structure beneath the receiver includes an inclined interface or a sharp gradient in seismic anisotropy, there will be an additional component of *P*-to-*SH* converted energy, with backazimuthal variations in *P*-to-*SH* phase polarity and amplitude yielding constraints on the geometry of the dipping structure or anisotropy (e.g. Levin & Park 1997). A dipping interface or tilted anisotropic symmetry axis will result in a two-lobed pattern of *P*-to-*SH* converted energy as a function of backazimuth (i.e. two polarity reversals within the full backazimuthal range), while anisotropy with a horizontal

*Now at: Department of Earth and Space Sciences, University of Washington, Seattle, WA, USA.

© The Authors 2016. Published by Oxford University Press on behalf of The Royal Astronomical Society.

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

symmetry axis (assuming hexagonal symmetry) will result in a four-lobed pattern (i.e. a $\sin(2\theta)$ amplitude pattern; e.g. Park & Levin 2016). The amplitude of *P*-to-*SH* converted energy can be used to infer the magnitude of anisotropy and/or the dip of a plunging symmetry axis.

Receiver functions contain a wealth of information regarding Earth structure beneath a receiver; however, reliable interpretation of this information is difficult. Receiver function data are subject to noise, scattering effects due to small-scale heterogeneities and near-surface reverberations that may obscure direct *P*-to-*S* conversions. There are also numerous trade-offs that interfere with interpreting receiver function data. For example, trade-offs exist between layer thickness and (isotropic) *P*- and *S*-wave velocities: increasing layer thickness has the same effect on Ps delay time as decreasing V_S or increasing V_P. Similarly, the amplitude of transverse component *P*-to-*SH* conversions will vary either with changes in the strength of anisotropy or by tilting the anisotropic symmetry axis out of the horizontal plane (e.g. Levin & Park 1997). Additionally, the value of the shape parameter, or ellipticity of the anisotropy, can also influence *P*-to-*SH* amplitude.

We present work to develop and implement a model space sampling algorithm for forward modeling of anisotropic receiver functions. We first review previous work on the interpretation of transverse component receiver functions, and then discuss the application of a non-Bayesian Markov chain Monte Carlo (McMC) approach to this problem. We present a series of synthetic tests to demonstrate the utility of our approach and illustrate trade-offs among different parameters, and present an application of the method to real data. Finally, we discuss the strengths and limitations of an McMC approach for forward modeling of receiver functions, as well as future directions that may further enhance the utility of such an algorithm.

2 INFERRING CONSTRAINTS ON EARTH STRUCTURE FROM RECEIVER FUNCTIONS

The abundance of information contained in receiver functions makes it nearly impossible to infer complex, anisotropic Earth structure through simple visual inspection. Forward modeling of receiver functions, in which a user-defined input model that contains a large number of parameters (i.e. V_P, V_S, density, layer thickness, strike, dip, the strength of anisotropy, azimuth and tilt of the anisotropic symmetry axes, for *each* layer in the model) is used to generate a set of synthetic receiver functions for comparison with actual data, is a common strategy for interpretation. Such an approach allows for more detailed constraints on Earth structure beneath a seismic station than is possible with visual inspection alone.

Many previous studies that have attempted to constrain complex anisotropic structure based on receiver function analysis have relied on forward modeling with a trial-and-error approach (e.g. Park *et al*. 2004; Nikulin *et al*. 2009; Wirth & Long 2012; McCormack *et al*. 2013). This usually consists of comparing synthetic receiver functions to receiver functions generated by actual data, in order to determine how well a given model matches the observations. While the trial-and-error approach has had some successes, it makes it impossible to explore model space in a systematic manner. Additionally, such an approach yields no meaningful constraints on the uniqueness of the solution, or how sensitive the model is to changes in certain parameters. Some workers have attempted to place more quantitative constraints on model parameters using a grid search approach (e.g. Wirth & Long 2014), but at a high computational cost. Alternate approaches to detailed forward modeling have also been used (e.g. Shiomi & Park 2008; Schulte-Pelkum & Mahan 2014; Ford *et al*. 2016); these studies have attempted to extract first-order information about the characteristics of dipping or anisotropic interfaces at depth without carrying out detailed forward modeling or synthetic seismogram computation.

Monte Carlo (MC) methods are a class of algorithms that make use of repeated random sampling to generate a numerical result, and have been applied to a wide array of geophysical inverse problems as a way of mapping out acceptable models in parameter space (e.g. Mosegaard & Sambridge 2002; Sambridge & Mosegaard 2002). A subdivision of MC methods that allows for a more efficient sampling of parameter space is McMC algorithms. McMC methods sample a multidimensional parameter space in a series of small steps (the 'Markov chain'), which approximates the probability distribution of the model space. Highly sampled regions of model space are more likely to contain models that are a good fit (based on a chosen misfit function) to Earth structure.

In an effort to obtain more robust information regarding anisotropic Earth structure, other studies have employed MC algorithms to more efficiently assess parameter constraints and trade-offs. (For a detailed overview of receiver function inversion, mostly limited to isotropic studies, we refer the reader to Bodin *et al*. 2012.) The McMC approaches can generally be divided into two categories. The first, model space sampling, explores parameter space based on a chosen misfit function that reflects a fit of the predictions to data. This results in an ensemble of models, which can be visually evaluated to determine areas of parameter space with low misfit values and discern trade-offs among parameters. Alternatively, Bayesian probabilistic sampling uses statistical techniques to sample from an explicitly defined posterior probability distribution, where $posterior \propto likelihood \times prior$. While Bayesian sampling allows for quantitative estimates of parameter constraints (e.g. Sambridge 1999b; Piana Agostinetti & Malinverno 2010; Bodin *et al*. 2012; Bodin *et al*. 2016), it is more burdensome in requiring an accurately and explicitly defined likelihood and prior probability density function.

In this study, we employ the simple (non-Bayesian) model space sampling approach, in which any desired misfit and likelihood function can be chosen. This methodology yields a distribution of model solutions that are biased towards (assumingly better) models with lower misfit values. While not providing truly quantitative constraints like Bayesian sampling, this approach still provides valuable information regarding trade-offs and the relative constraints of model parameters. This type of approach has been attempted previously with a range of algorithms (and often limited to isotropic structure), with varying degrees of success—including genetic algorithms (e.g. Shibutani *et al*. 1996; Levin & Park 1997; Chang *et al*. 2004), simulated annealing (e.g. Vinnik *et al*. 2004) and neighbourhood algorithms (e.g. Sambridge 1999a; Frederiksen *et al*. 2003; Sherrington *et al*. 2004; Porter *et al*. 2011).

We develop an McMC approach for the forward modeling of anisotropic receiver functions as an alternative strategy to those employed in previous studies. This allows for an efficient, systematic assessment of parameter space (i.e. identification of models with low misfit values) and visual assessment of parameter trade-offs, a valuable improvement over trial-and-error or grid search approaches. We then apply the McMC method to three sets of synthetic receiver functions based on idealized models of Earth structure and actual receiver function data from station WCI in the Granite-Rhyolite province of the eastern North American craton. Finally, we discuss the utility of our approach and future directions that may further enhance the viability of such an algorithm.

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

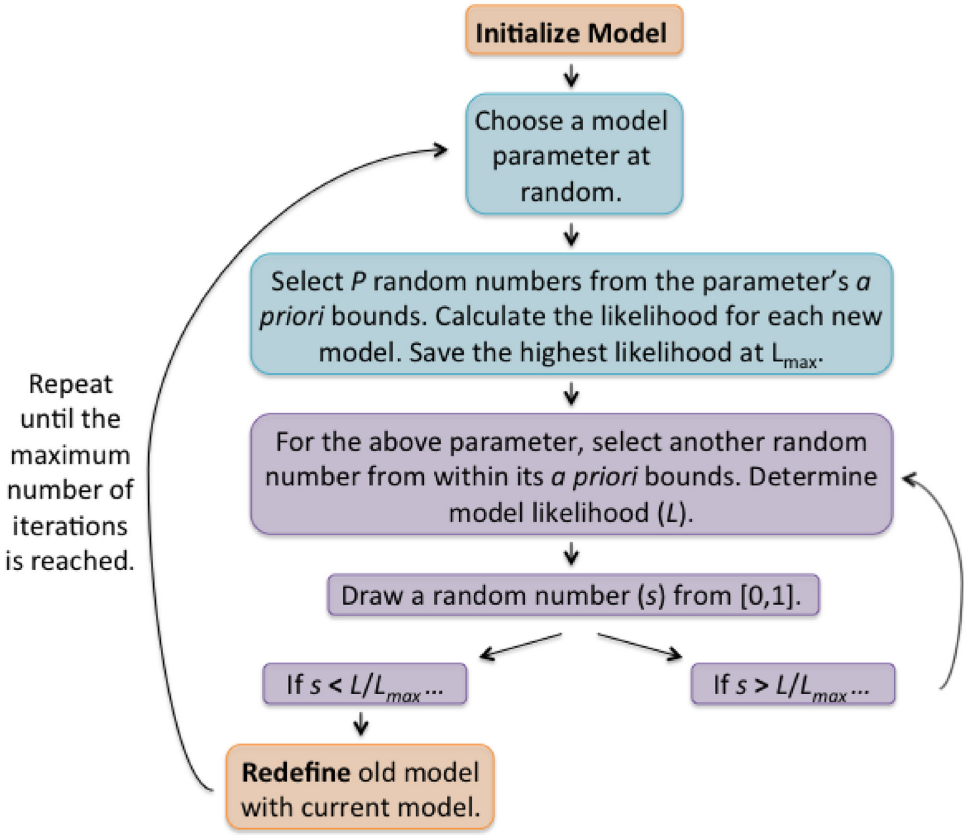

Figure 1. Flow chart of the Markov chain Monte Carlo algorithm with Gibbs sampling used in this study. Blue-shaded boxes correspond to the 'random scan' step in which 1-D conditional, biased probability distributions are created. Purple-shaded boxes represent the 'Gibbs sampling' steps that determine the progression of the Markov chain.

3 METHODOLOGY AND IMPLEMENTATION

When utilizing an McMC approach, some algorithm is necessary to move the Markov chain randomly, but efficiently, through parameter space. Perhaps the most well-known strategy is the Metropolis–Hastings algorithm (Metropolis *et al.* 1953; Hastings 1970). In the Metropolis–Hastings algorithm, a randomly initialized model is chosen and its corresponding likelihood is calculated, $L(\mathbf{m})$. The likelihood of a model, $L(\mathbf{m})$, is equal to the probability of the data assuming that model. The initial model is then perturbed, and the likelihood is calculated for the new perturbed model, $L(\mathbf{m}')$. If $L(\mathbf{m}') > L(\mathbf{m})$, the perturbed model is a better fit, and is accepted as the next model in the chain. However, if $L(\mathbf{m}') < L(\mathbf{m})$, we select a random number (s) from the interval [0,1]. If $s < L(\mathbf{m}')/L(\mathbf{m})$, the perturbed model ($\mathbf{m}'$) is accepted. Conversely, if $s > L(\mathbf{m}')/L(\mathbf{m})$, $\mathbf{m}'$ is rejected and the process continues. This rejection criterion ensures that there is always the possibility to accept a model that is a 'worse fit,' ensuring that the McMC algorithm can escape local minima.

The size of the random perturbation is extremely important in the Metropolis–Hastings algorithm. If the steps are too large, it may result in drastic changes that would cause most new models to be rejected. However, if the perturbation size is too small, the algorithm may remain 'trapped' in a local minimum for an extended period of time. The selection of the perturbation size therefore becomes complicated, particularly with a problem that has a large number of free parameters (as is the case for forward modeling of receiver functions), in which each parameter may have a different optimal step size.

An extension to the Metropolis–Hastings algorithm described above is the Gibbs sampler (e.g. Liu 2008). Gibbs sampling bases its next 'step' on 1-D conditional probability distributions that are global in scale (i.e. it compares the perturbed model, $\mathbf{m}'$, to several other previous models in which the same parameter was perturbed). Therefore, unlike the Metropolis–Hastings algorithm, it does not require any step 'tuning.' Details of the McMC approach with Gibbs sampling, based on work by Korenaga & Karato (2008), are provided below and shown as a flow chart in Fig. 1. For a more comprehensive review of MC methods in general, we refer the reader to Liu (2008).

(1) Model initialization: For each model parameter, draw a random number r_k from the interval [0,1]. Set the value of the parameter to $m_{0,k} = m_k^L + r_k(m_k^U - m_k^L)$, where m^L and m^U are the parameter's *a priori* lower and upper bounds, respectively.

(2) Random scan (Generating a conditional probability distribution for Gibbs sampling): Pick one model parameter at random, m_r. While holding all other model parameters fixed, select P random numbers from the interval $[m_r^L,\ m_r^U]$ and calculate the corresponding likelihood values. Save the highest likelihood as $L_{\max}$.

(3) Gibbs sampling: Pick one random number from the interval $[m_r^L,\ m_r^U]$ and refer to it as m_r. Draw a random number, s, from the interval [0, 1]. If $s < L(m'_r)/L_{\max}$, continue to the next step. Otherwise, repeat this step.

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

(4) Model update: Redefine the old model $\mathbf{m}_0$ with $\mathbf{m}'$. Go back to Step 2, until the maximum number of iterations is reached.

We adopt the likelihood function for Gaussian distributed errors:

$$L(m) \propto \exp\left(-\frac{1}{2}\chi^2(m)\right),$$

where,

$$\chi^2 = \sum \frac{(X_{\text{Data}} - X_{\text{Model}})^2}{\sigma^2}.$$

The variance, σ^2, is not known and so it was adjusted such that the acceptance rate of new models in Step 3 was ∼15–50 per cent; this process is referred to as tempering (e.g. Rothman 1985, 1986; Sambridge 2014). In this case, $L(m)$ is not strictly a likelihood, thus the parameter distributions returned by the McMC algorithm cannot necessarily be interpreted as the underlying probability distribution of the parameters (hereafter referred to as 'biased'). Nevertheless, the algorithm will still explore high likelihood regions of parameter space more than low likelihood regions, thereby achieving the main goal of a more efficient and useful way to extract information from receiver function data. In Section 6.1, we explore alternative approaches and forms of the likelihood function that may offer further improvement.

McMC algorithms are meant to densely sample areas of model space with high fitness values. Therefore, rather than focusing our interpretation solely on the model with the lowest misfit (or highest likelihood value), it is better to get an overall impression of where the Markov chain spent most of its 'time' in model space, or the biased probability distributions of the parameters. To this end, we chose to visualize our searches using a series of 2-D histograms that clearly show what areas of model space were most highly sampled.

We implemented a version of the McMC algorithm described above to explore model space for the forward modeling of anisotropic radial and transverse component Ps receiver functions. We wrote a version of the algorithm that incorporates previously written codes for generating synthetic seismograms in dipping-anisotropic media (Frederiksen & Bostock 2000) and computing receiver functions using the multitaper correlation method (Park & Levin 2000). These codes invoke hexagonal anisotropic symmetry, which is a common and appropriate assumption for upper-mantle anisotropy. The anisotropy parameter (η), which controls the propagation velocity for orientations other than parallel and perpendicular to the anisotropic symmetry axis, is defined such that the 3-D expression of velocity is purely ellipsoidal, consistent with Levin & Park (1997; see also Sherrington *et al*. 2004). As the McMC algorithm with Gibbs sampling easily lends itself to parallel computing, we carried out most runs on Yale University's High Performance Computing cluster using Open MPI running on a total of 32 processors. To evaluate the viability of the McMC approach, we first ran a series of synthetic tests. For a series of user-defined models of Earth structure (including the number of layers, layer thickness, V_P, V_S, density, the strike and dip of each layer, the azimuth and tilt of the anisotropic symmetry axes and the strength of anisotropy), we generated synthetic seismograms and then computed receiver functions that serve as 'data' for our test runs. The McMC search uses the algorithm described above to explore model space, and compares receiver functions generated based on potential models for Earth structure to the receiver functions representing actual (synthetic) 'data.'

4 RESULTS OF SYNTHETIC TESTS

Synthetic tests were performed for three different scenarios with increasing complexity, including a model where parameters in one layer are unknown (i.e. nine free parameters), two layers are unknown (i.e. 18 free parameters), or four layers are unknown (i.e. 36 free parameters). As receiver functions can be visually assessed for the presence of sharp impedance contrasts, yielding an estimate as to how many layers should be modeled, we elect to pre-define the number of layers in each scenario before running the sampling algorithm.

4.1 Case A: 1 'unknown' layer

We applied the McMC algorithm to a hypothetical scenario in which the characteristics of one layer of Earth structure (i.e. nine parameters) are unknown (a sketch of Earth structure for this scenario, along with synthetic receiver function data generated from the model, is shown in Fig. 2). The free parameters, their *a priori* bounds, the actual values of the parameters and the values estimated from the McMC model with the lowest misfit are shown in Table 1. However, we emphasize that it is generally more informative to look at the distribution of parameter values with the lowest model misfit. To this end, we plot a series of distributions illustrating the results of the McMC model space search for this scenario (Fig. 3). (Again, we note that these distributions are *biased*, based on the choice of likelihood function—and therefore not the *true* posterior probability distribution.) The biased distributions for the full parameter sets are represented via a series of 2-D histograms, which demonstrate parameter values that resulted in low model misfit and how different sets of parameters trade-off against each other.

For Case A, the 'unknown' parameters are accurately constrained by the McMC algorithm and our choice of likelihood function. For example, despite the trade-offs that are typically inherent among V_P, V_S and layer thickness (defined as the thickness beneath the station), low model misfits are concentrated around the true parameter values, a result of matching Ps phase timing and amplitude exactly (Figs 3a and b). The parameters that characterize anisotropic structure (i.e. trend and plunge of the anisotropic symmetry axis, strength of the anisotropy) are also highly sampled at the correct values, and the distributions shown in Fig. 3 demonstrate some important points. First, the *trend* of the anisotropic symmetry axis is well constrained (Fig. 3c). The hint of bimodality in the orientation of the anisotropic symmetry axis ('anisotropy trend') is simply due to symmetry considerations. Specifically, since we assume hexagonal symmetry, if the anisotropic symmetry axis is in the horizontal plane and aligned with a backazimuth of θ, the resulting receiver functions are identical to those that would be produced for anisotropy aligned to a backazimuth of $\theta + 180^\circ$.

A second important inference is that the *plunge* of the anisotropic symmetry axis from horizontal (which manifests itself in small changes in Ps phase amplitude) is not as tightly constrained as the azimuthal orientation. This is clearly shown in Fig. 3(c) by the 'strip' of highly sampled regions along all possible *plunges* of the anisotropic symmetry axis, as long as the *trend* is set at ∼120°–150°. This ambiguity may be partially due to the trade-off between the strength of anisotropy and the plunge of the anisotropic symmetry axis, both of which contribute to Ps phase amplitude. Fig. 3(e) provides further evidence of this trade-off, in that models with a plunge of ∼40°–60° exhibit weaker anisotropic strengths (i.e. ∼1–2 per cent), while models with a more modest plunge (∼10°–20°) require stronger anisotropy (i.e. 1–4 per cent). Finally,

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

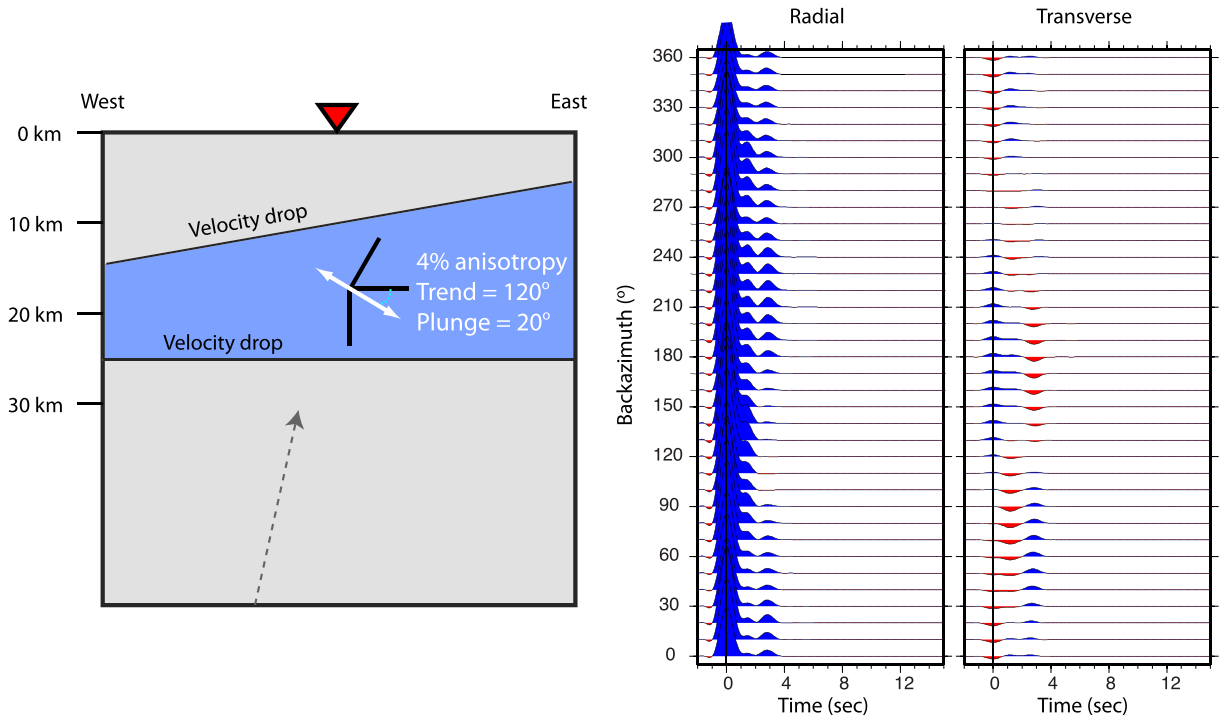

Figure 2. Sketch of model Earth structure for Case A (left), in which the parameters of 1 layer (in colour) are 'unknown.' The corresponding synthetic receiver functions that the McMC code attempts to match are also shown (right). Receiver functions are plotted as a function of time (horizontal axis), where $t = 0$ is the arrival of the *P* wave at the seismic station, and grouped by backazimuth (vertical axis). In general, positive pulses (blue) represent a velocity decrease experienced by the wave as it moves towards the surface (indicated by grey, dashed arrow) and negative pulses (red) represent a velocity increase. Figure is not to scale.

for the Case A synthetic test, we see that the strike and dip of the inclined interface at the top of the anisotropic layer is not particularly well constrained (Fig. 3d). However, for models with steeper dip angles (which make the signature of an inclined interface more apparent in receiver functions), the biased probability distributions suggest (correctly) that the strike of the dipping structure is aligned roughly N-S.

4.2 Case B: 2 'unknown' layers

Next, we evaluated an alternative synthetic scenario in which two layers (i.e. 18 parameters) are unknown using the McMC algorithm (Fig. 4). With a few exceptions, we found that regions of low model misfit (i.e. highly sampled regions) could be used to infer approximate parameter values (Fig. 5). In the upper layer, the estimated magnitude of V_S appears well constrained (Fig. 5b), which is understandable considering that variations in V_S control *P–SV* conversions (i.e. a drop in *S*-wave velocity with no corresponding drop in *P*-wave velocity will still result in a *P*-to-*SV* conversion, while the opposite scenario will not). By interpreting the 2-D histograms of layer thickness versus V_P or V_S simultaneously (i.e. comparing the highly sampled regions in both plots concurrently; Figs 5a and b), layer thickness and V_P can also be reasonably inferred.

In the Case B model, the upper layer was defined as isotropic. This is well captured by the results of the model space search, in that most models sampled had very weak anisotropy in the upper layer (i.e. <3 per cent; Fig. 5e). In models where stronger anisotropy was present, this was compensated for by a shallow plunge of the anisotropic symmetry axis (i.e. resulting in low-amplitude *P*-to-*SH* arrivals; Figs 5c and e). (We note that this result could be modified slightly by a different definition of the shape parameter, η.) Because the anisotropy is so weak in the models favoured by the McMC algorithm, the trend of the anisotropic symmetry axis (Fig. 5c) cannot be interpreted. Finally, the strike and dip of the upper interface appears very well constrained (Fig. 5d).

It is more difficult to infer information about the structure of the lower layer, as the histograms are generally multimodal. This is the case for layer thickness, V_P, and V_S, with a greater layer thickness being compensated for with a faster V_P (Figs 5f and g). However, if the more highly sampled mode is chosen as the more likely solution, V_P and layer thickness are also correctly constrained (Fig. 5f). Using these implied constraints, V_S is correspondingly (and correctly) constrained (Fig. 5g). However, we note that this distinction is subtle, and with real data, users would not likely be truly confident in choosing one mode over the other.

The trend and plunge of anisotropy in the lower layer are also well matched (Fig. 5h), although the strength of anisotropy is significantly underpredicted by the algorithm and chosen likelihood function (Fig. 5j). The strike of the upper boundary is poorly constrained, likely because the dip of the layer is predicted to be extremely low (i.e. essentially a horizontal layer). However, the poor constraints on dipping structure actually result from the fact that the dipping interface does not cause a corresponding prominent feature in the receiver function data; rather, it mainly results in amplitude changes of already existing patterns, and is therefore poorly constrained by the synthetic data.

4.3 Case C: 4 'unknown' layers

Ideally, we would like to be able to apply the McMC algorithm to complicated, real-world data sets that might include a large number of anisotropic and/or dipping layers beneath a station. While it is

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

Table 1. *A priori* parameter bounds and actual 'data' values for each of the synthetic examples. For Case A (1 'unknown' layer), the best-fit values from the McMC algorithm are also given. However, we emphasize that it is more instructive to look at the ensemble of sampled models (e.g. Fig. 3).

	Thickness (km)	V_P (km s^{-1})	V_S (km s^{-1})	Density (kg m^{-3})	Per cent	Trend (°)	Plunge (°)	Strike (°)	Dip (°)		
Case A											
Parameter bounds	10–30	6.0–8.0	3.2–4.6	3.0–3.3	0–10	0–360	0–90	0–360	0–30		
'Data'	15	7.3	4.1	3.1	4	120	20	180	10		
McMC best fit	15	7.5	4.1	3	3	122	33	187	8		
Case B											
	Thickness (km)	V_P (km s^{-1})	V_S (km s^{-1})	Density (kg m^{-3})	Per cent	Trend (°)	Plunge (°)	Strike (°)	Dip (°)		
Parameter bounds Layer 1	5–20	6.0–9.0	3.2–5.0	3.0–3.3	0–20	0–360	0–90	0–360	0–40		
Parameter bounds Layer 2	5–20	6.0–9.0	3.2–5.0	3.0–3.3	0–20	0–360	0–90	0–360	0–40		
'Data'—Layer 1	13	7.5	4.1	3.2	0	0	0	200	15		
'Data'—Layer 2	7	7.0	3.9	3.1	6	280	10	180	20		
Case C											
	Per cent	Trend (°)	Plunge (°)	Strike (°)	Dip (°)		Per cent	Trend (°)	Plunge (°)	Strike (°)	Dip (°)
Parameter bounds Layer 1	0–10	0–360	0–90	0–360	0–30	'Data' Layer 1	3	300	0	180	0
Parameter bounds Layer 2	0–10	0–360	0–90	0–360	0–30	'Data' Layer 2	3	350	15	180	10
Parameter bounds Layer 3	0–10	0–360	0–90	0–360	0–30	'Data' Layer 3	3	70	60	–	–
Parameter bounds Layer 4	0–10	0–360	0–90	0–360	0–30	'Data' Layer 4	3	90	0	–	–
Case D											
	Thickness (km)	V_P (km s^{-1})	V_S (km s^{-1})	Density (kg m^{-3})	Per cent	Trend (°)	Plunge (°)	Strike (°)	Dip (°)		
Parameter bounds Layer 1	10–30	6.0–8.0	3.2–4.6	3.0–3.3	0–10	0–360	0–90	0–360	0–30		
'Data'	15	7.5	4.1	3.2	6	70	40	270	20		

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

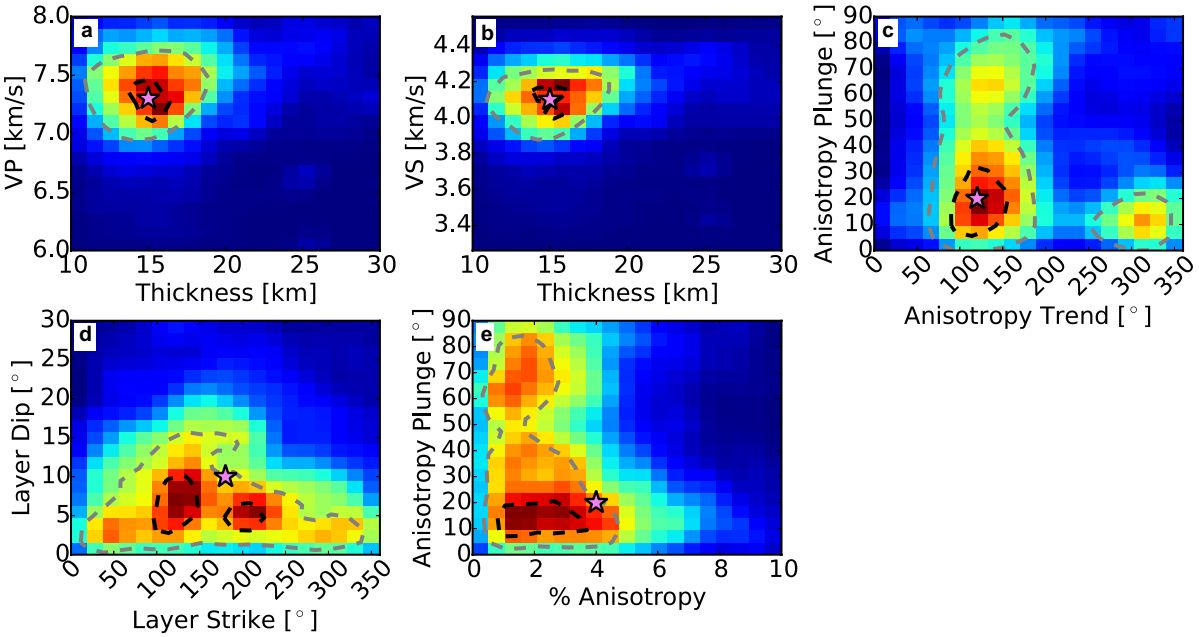

Figure 3. 2-D histograms showing the range of model parameters sampled by the McMC model space search for Case A. Warm colours represent regions that were highly sampled by the Markov chain (i.e. where a solution is likely), and cool colours represent very sparsely sampled regions of parameter space (i.e. where a solution is unlikely). Black and grey dashed lines enclose 10 and 50 per cent of sampled models, respectively. Pink stars denote the true parameter values. In this synthetic example, one layer of the model was solved for (i.e. nine parameters). A 2-D histogram for density is not shown since density has a minimal effect on the properties of the receiver functions. Histograms are smoothed using a Gaussian kernel density estimator.

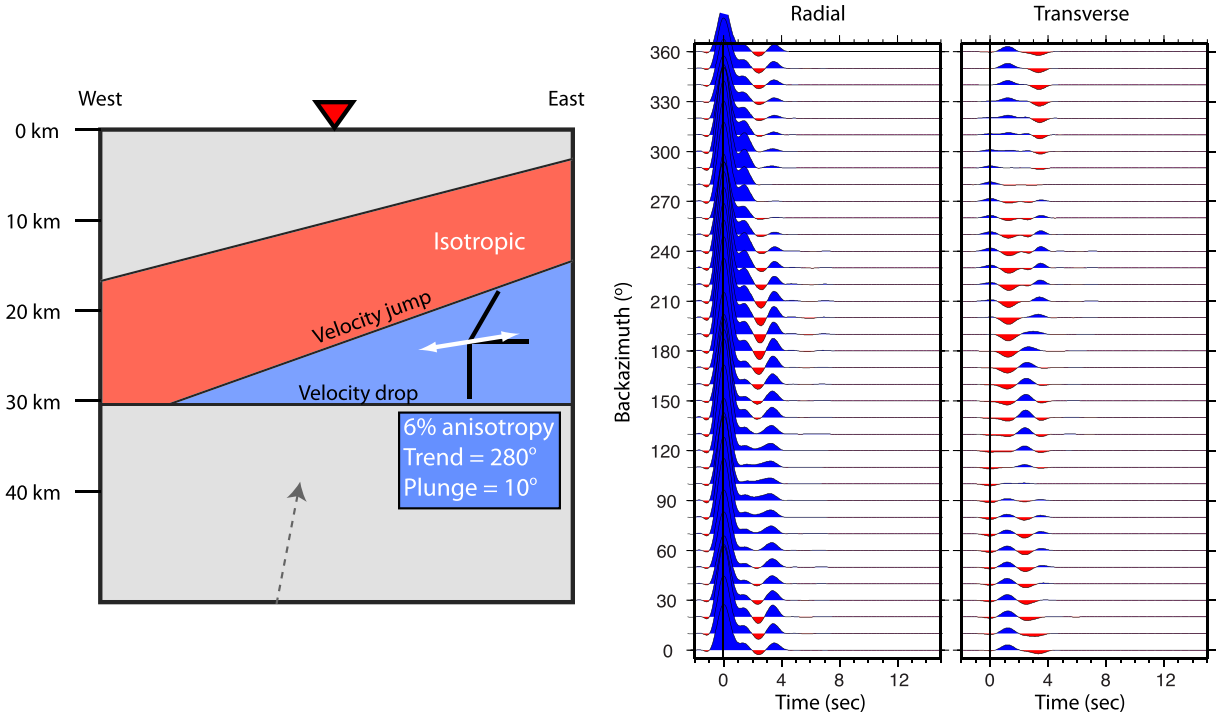

Figure 4. Sketch of model Earth structure for Case B (left), in which the parameters of two layers (in colour) are 'unknown,' for a total of 18 parameters. Corresponding synthetic receiver functions are plotted using the same conventions as in Fig. 2.

already clear from Case B (above) that there is ambiguity in the resolution of several parameters for even a relatively simple two-layer model, we also tested a more complicated four-layer model that is based on structure inferred in a previous study. Our final synthetic model (Case C) is based on the best-fit model for lithospheric structure beneath station WCI in the central United States put forth by Wirth & Long (2014), with a few additional complexities (i.e. an isotropic layer is made anisotropic and dipping structure is added). Model C includes a total of 36 free parameters (Fig. 6) and therefore provides a test of the McMC algorithm's performance for a complex model with a large number of unknown parameters. For our Case C scenario, we find that 2-D histograms exhibit considerable

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

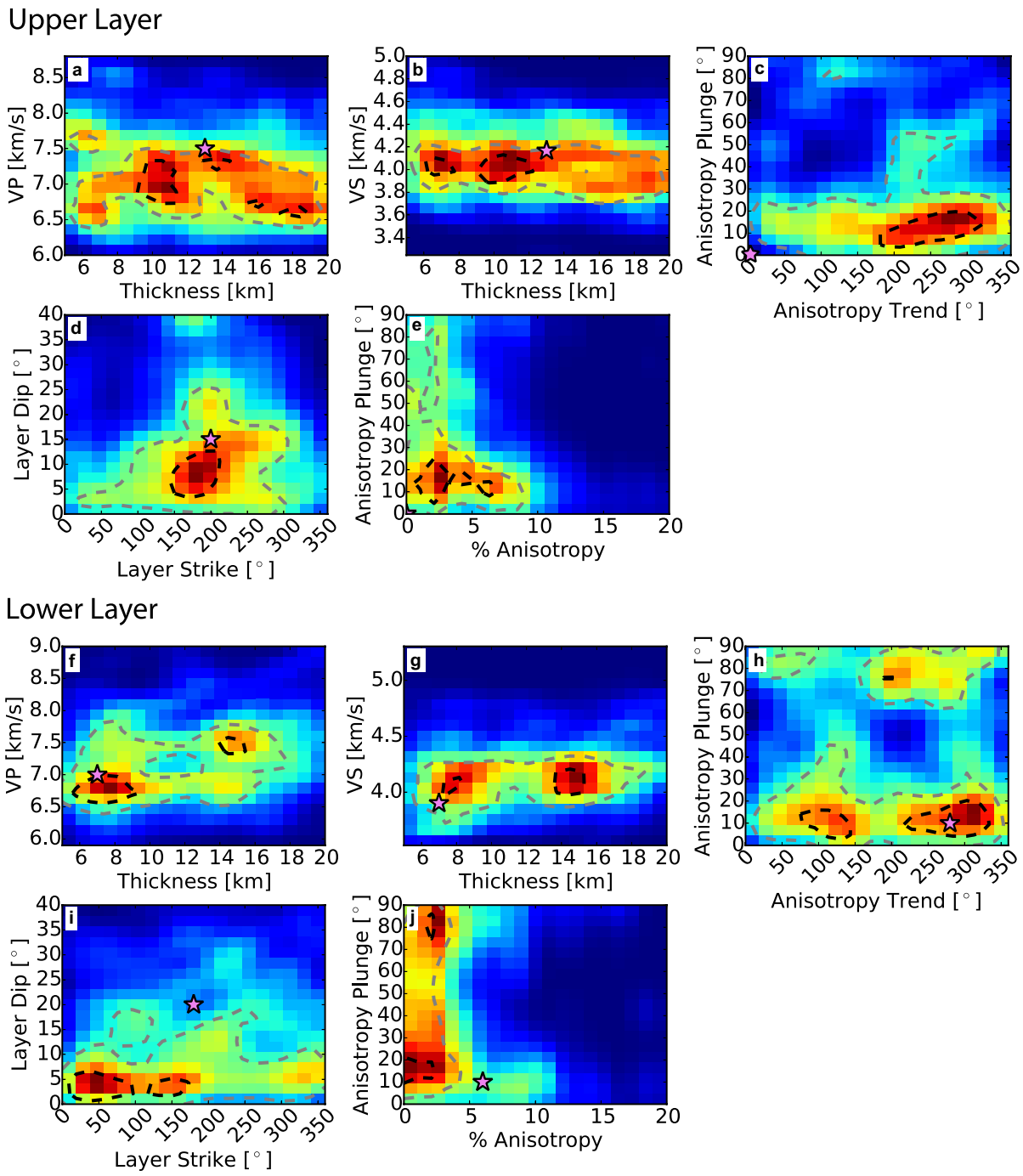

Figure 5. 2-D histograms for the Case B synthetic example. Plotting conventions are the same as in Fig. 3.

scatter (Fig. 7) and are heavily subject to trade-offs among parameters (e.g. layer thickness and isotropic *P*- or *S*-wave velocity). Since transverse component receiver function features are less affected by such serious trade-offs, we focus on those model parameters that describe anisotropy or dipping interfaces, which result in prominent polarity reversals of *P*-to-*SH* converted phases on transverse components and should be relatively robustly constrained by the data. Therefore, Fig. 7 shows only histograms of the trend of the anisotropic symmetry axis and the strike of any dipping interfaces for each layer.

In the uppermost layer ('Layer 1'), the strike of the dipping interface is reasonably well constrained, consistent with the idea that dipping structure results in a prominent polarity reversal in the receiver function data (Fig. 7a). The trend of the anisotropic symmetry axis is less well constrained (perhaps due to the relatively weak anisotropy in the 'actual' data, ∼3 per cent); regardless, the most highly sampled region as illuminated by the 2-D histogram does in fact suggest the correct orientation of anisotropy. Additionally, because the symmetry axis is oriented in the horizontal plane (as it is for all layers in this synthetic model), an orientation that is ±180° should also give the same results (although we note that this does not appear to be a particularly highly sampled region), perhaps explaining some of the apparent ambiguity in Fig. 7(a).

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

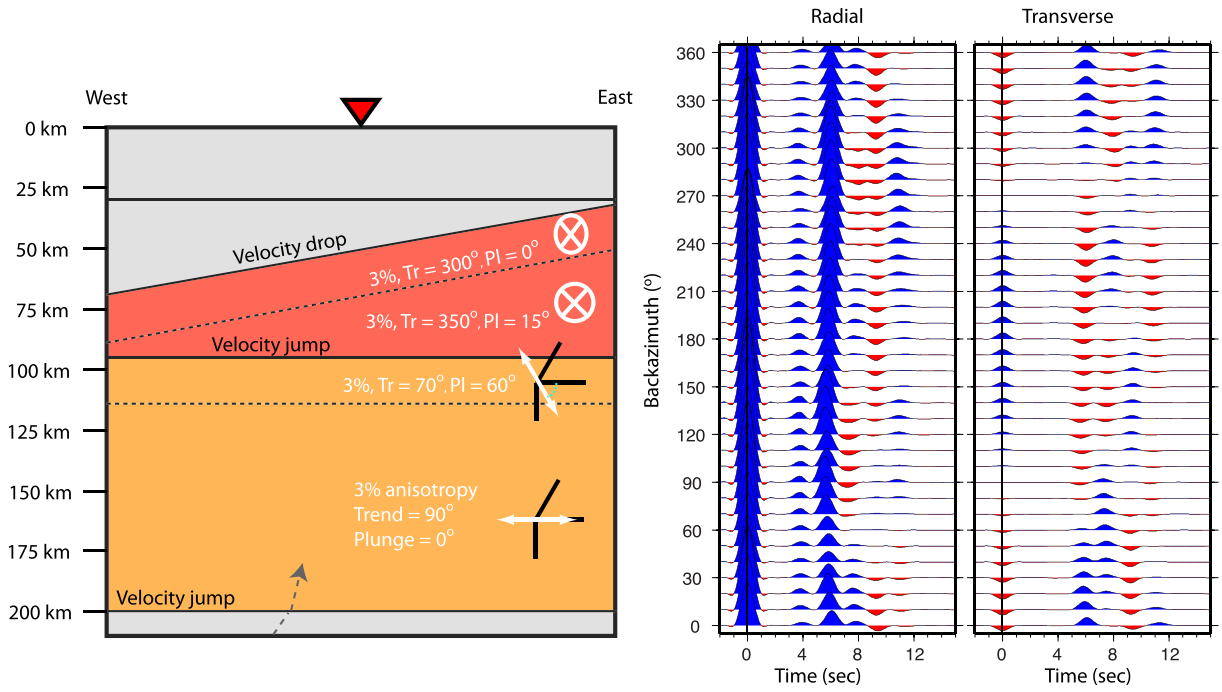

Figure 6. Sketch of Earth structure for Case C (left), in which the parameters of four layers (in colour) are 'unknown'. Dashed interfaces indicate a change in the orientation of anisotropy, without a co-located change in isotropic velocity. Corresponding receiver functions are plotted using the same conventions as in Fig. 2.

In Layer 2, the strike values suggested by the algorithm are insignificant because the dip was predicted to be extremely shallow ($<10^{\circ}$), and the interface is therefore essentially horizontal (Fig. 7b). With this in mind, it is clear that the trend of the anisotropic symmetry axis is somewhat well constrained, if the results for each trend bin are summed across all strikes (although there is a suggestion of multimodality). For Layers 3 and 4, there is no dipping structure, which is apparent in the considerable scatter in McMC models (Figs 7c and d). It is clear that even when the scatter due to the algorithm attempting to assign a strike to a non-dipping layer is taken into account, the trend of anisotropy in Layer 3 is still very poorly constrained by the synthetic data. In Layer 4, the McMC search does show some evidence of higher sampling at the correct orientation of the anisotropic symmetry axis, although the constraints are decidedly weak.

4.4 Summary of inferences from synthetic tests

Synthetic tests demonstrate that if there are some pre-existing constraints on Earth structure and the number of free parameters is small (i.e. akin to Case A with nine free parameters), the McMC algorithm works well for placing relative constraints on anisotropic Earth structure. In more realistic scenarios (e.g. Cases B and C), clear trade-offs begin to emerge between parameters that affect the timing and amplitude of the Ps phase arrivals (e.g. layer thickness, V_P, V_S, per cent anisotropy and plunge of the anisotropic symmetry axis). However, because prominent polarity reversals of *P*-to-*SH* converted energy are not subject to parameter trade-offs, the McMC algorithm presented here can be used to place constrain the orientation of anisotropy at depth, even for models with complex structure.

Our synthetic tests also demonstrate that there are significant benefits to visualizing the distribution of potential models (i.e. by utilizing a suite of 2-D histograms), as opposed to simply examining a singular 'best-fit' model. For instance, inspection of the ensemble of models allows the viewer to directly assess trade-offs between parameters (e.g. layer thickness and isotropic velocity, Fig. 5f; the plunge and strength of anisotropy, Fig. 3e). Furthermore, in certain instances it is useful to view the distribution of several parameters at once. For example, if the strength of anisotropy is predicted to be very weak or have a low plunge angle (e.g. Fig. 5e), both of which would result in low amplitude *P*-to-*SH* conversions for a vertically propagating wave, this suggests that the interpretation of the trend of the anisotropic symmetry axis in this layer (e.g. Fig. 5c) may not be particularly meaningful.

5 APPLICATION TO ACTUAL DATA

We apply the McMC algorithm to actual receiver function data from permanent broad-band station WCI, located at the eastern edge of the North American craton. In previous work, we examined radial and transverse component receiver functions at this station and used a grid search approach to constrain the orientation of anisotropy in the mantle lithosphere (Wirth & Long 2014); our preferred model suggested multiple sharp contrasts in anisotropy at lithospheric depths. This previous result provides a useful comparison to constraints on the orientation of anisotropy provided by the McMC algorithm.

Since actual receiver function data often exhibits considerable noise, and because our synthetic tests suggest that many parameters cannot be constrained for complex models, we chose to only search for the best-fit trends and plunges of anisotropy in four layers of mantle lithosphere (i.e. 50–70, 70–90, 90–110 and 110–200 km depths, based on the results of Wirth & Long (2014), for a total of eight free parameters). The results from the McMC algorithm applied to actual data (Fig. 8) provide reasonable bounds (i.e. distinct regions of parameter space with low model misfit) on the orientation of anisotropy at mid-lithospheric mantle depths, particularly in the depth range from ~70 to 110 km. Since models with steep

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

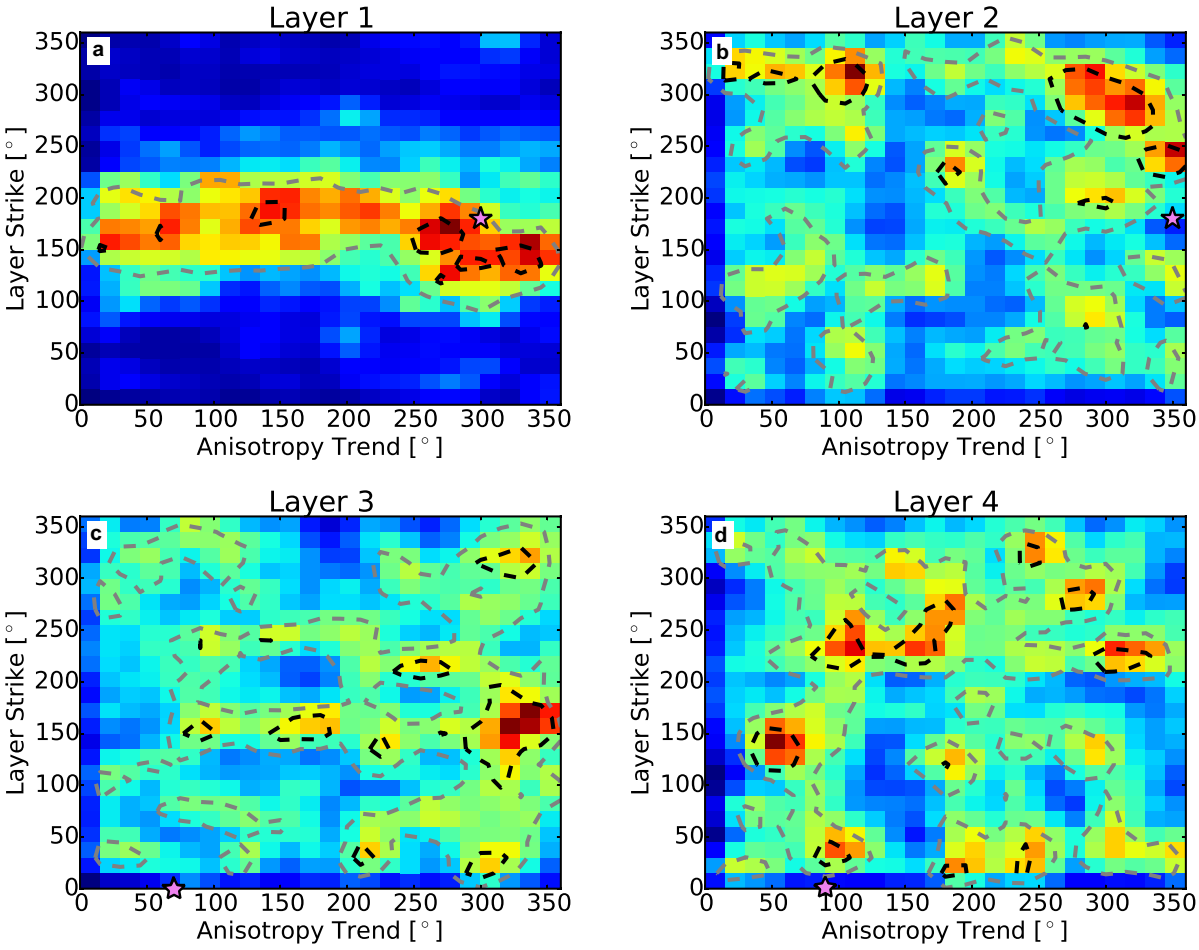

Figure 7. 2-D histograms for a synthetic example in which four layers of the model were solved for (i.e. 36 parameters). Since most parameters were poorly constrained, only the results for *trend* of the anisotropic symmetry axis and *strike* of an inclined interface are shown. Plotting conventions are as described in Fig. 3.

plunges of anisotropy sampled by near-vertically incident rays begin to resemble isotropy and are difficult to distinguish from weak anisotropy, we also plot 1-D histograms of the trend of anisotropy only (e.g. Figs 8c and d). At 70–90 km depth, the algorithm suggests a clear N-S orientation of anisotropy, with a clear 180° periodicity in the probability distribution that indicates a horizontal axis of symmetry (Fig. 8c). At 90–110 km depths, the probability distribution favours an eastward orientation of the anisotropic symmetry axis, but without the 180° periodicity (Fig. 8d). This suggests that the symmetry axis is plunging out of the horizontal plane, but because of the trade-offs inherent to the strength and plunge of anisotropy (Section 4.2), we do not seek to constrain the steepness of this plunge.

A comparison among synthetic transverse component receiver function polarity reversals for the best-fitting model identified by the McMC algorithm, synthetic receiver functions for the best-fitting model from Wirth & Long (2014), and actual data from station WCI (Fig. 9) demonstrates that both the McMC model and the previous grid search approach suggest parameter values that provide a good fit to the data. Additionally, an McMC algorithm is much more computationally efficient that a grid search approach. Overall, this test demonstrates the ability of the McMC algorithm to place constraints on anisotropic Earth structure using actual receiver function data. In addition to robustly constraining certain parameters, such as the orientation of anisotropy, it also provides valuable information on the limitations of the data. That is, the McMC results clearly demonstrate that for certain parameters and depth ranges, receiver function data cannot be used to infer particular aspects of Earth structure.

6 SUGGESTIONS FOR FUTURE IMPLEMENTATIONS

6.1 Alternative misfit evaluations

As previously mentioned, the way in which the likelihood of a modeled receiver function is evaluated is extremely important, as the interpretation of the underlying probability distribution of the parameters is contingent upon the correctness of the likelihood. In the examples presented in this paper, the likelihood was calculated assuming independent Gaussian distributed errors with a variance set to maximize the efficiency of the algorithm. This description of the likelihood is imperfect, as receiver functions are known to exhibit correlated noise (Sambridge 1999b). However, our definition is sufficient for the task of guiding the McMC algorithm to more important regions of parameter space. Future implementations could utilize a properly defined likelihood function that incorporates correlated noise (Sambridge 1999b; Piana Agostinetti &

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

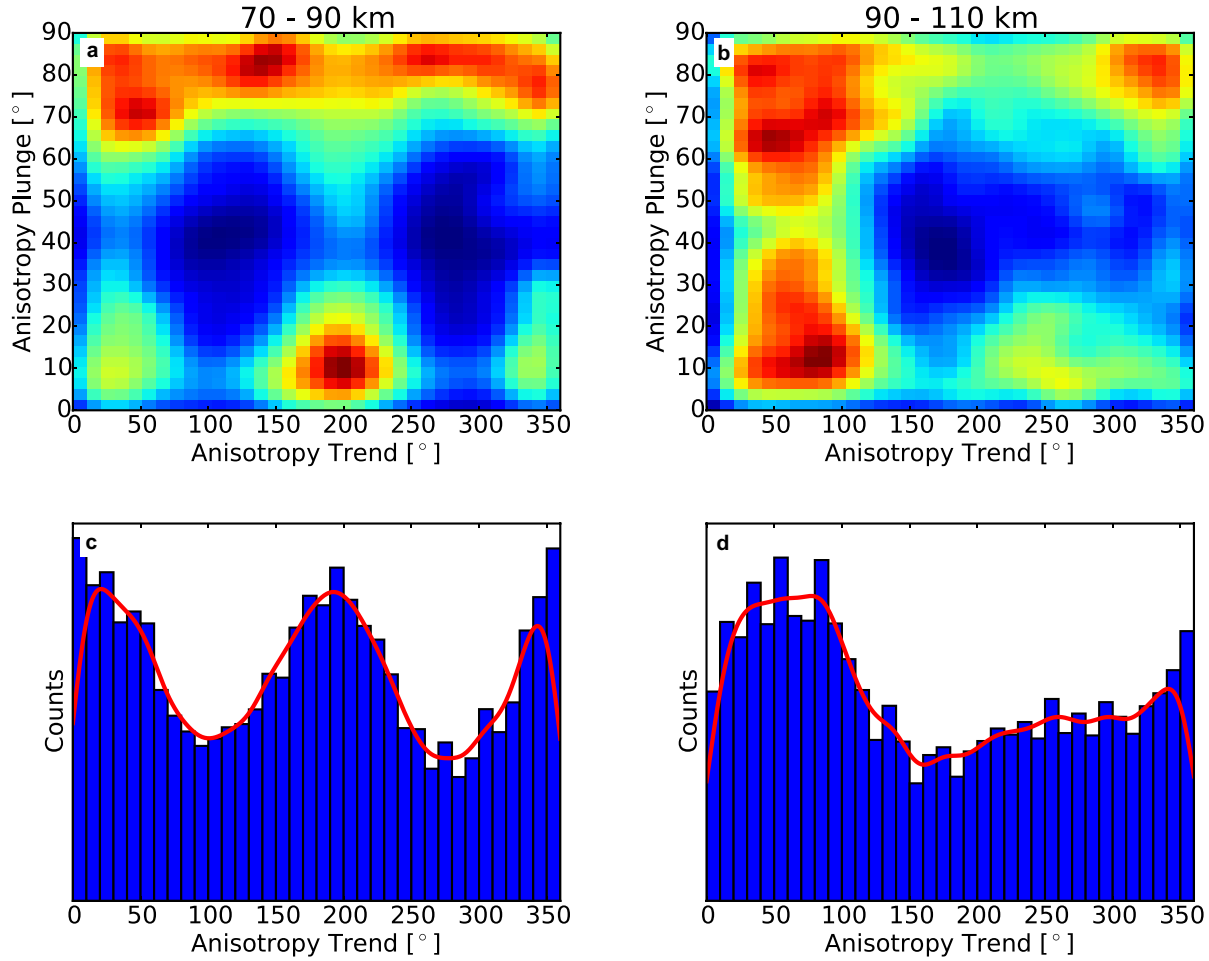

Figure 8. (a) and (b) 2-D histograms for the trend and plunge of anisotropy in two layers of lithospheric mantle, generated using the McMC algorithm and actual receiver function data from station WCI. (c) and (d) 1-D histograms for the trend of anisotropy in the same two layers of lithospheric mantle.

Malinverno 2010; Bodin *et al*. 2012; Dettmer *et al*. 2014), which at the expense of additional complexity at the onset, is rewarded with truly quantitative estimates of parameter constraints and errors.

In keeping with our chosen definition of likelihood, the value of σ could be amended to suit the character of the data at hand. For instance, σ could be correlated with the level of background noise in the actual receiver function data, thereby assigning less weight to fitting receiver functions that correspond to noisy data traces. This would still result in biased probability distributions that are guided by our chosen misfit function, but may provide a more accurate definition of model misfit. Additionally, we note that the ensemble of models produced by the McMC algorithm can be *unbiased* at a later time. For example, Sambridge (1999a) produced a biased model ensemble (i.e. similar to this work) using a neighbourhood algorithm, and later unbiased the ensemble to obtain quantitative parameter constraints, using a properly defined likelihood function (Sambridge 1999b).

A substantially different estimate of model fitness could also be tested. A potential example of this is dynamic time warping (DTW; for a detailed description, see Berndt & Clifford 1994). A pitfall of an L2-norm or χ^2 misfit function is that it aligns the ith point of one time-series (i.e. the actual receiver function data) to the ith point in another time-series (i.e. receiver functions generated by the McMC algorithm). However, in a hypothetical situation in which all the model parameters that describe Earth structure are correct except for one (say, layer thickness), the predicted receiver function *patterns* may provide a good match to the data, but their *timing* would be offset. In this hypothetical case, the 'acceptability' of the model should be high, but a measure of misfit that uses point-by-point Euclidean distances would imply that the model likelihood would be low. In contrast, algorithms such as DTW focus on similarities in the shape of the two time-series by allowing for non-linear alignment (Fig. 10). Using DTW or a similar construction of model fitness could improve the efficiency of the McMC algorithm in finding 'acceptable' fits to the data. In order to provide a proof of concept, we wrote a module for determining receiver function misfit based on DTW and re-ran the McMC algorithm on some simple models for Earth structure (i.e. akin to Cases A and B with 1 or 2 'unknown' layers).

Example ensembles from an McMC run using one such model, Model D, with DTW to evaluate misfit is shown in Fig. 11. For this particular scenario with one 'unknown' layer, DTW does an excellent job constraining the trend and plunge of anisotropy and the strike and dip of the inclined interface (Figs 11c and d). However, it struggles to constrain the layer thickness and isotropic velocities (Figs 11a and b), precisely because it allows for non-linear alignment of waveform patterns. (It is worth noting, however, that the 'best-fit' model has a layer thickness that is off only by ∼5 km, and estimates

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

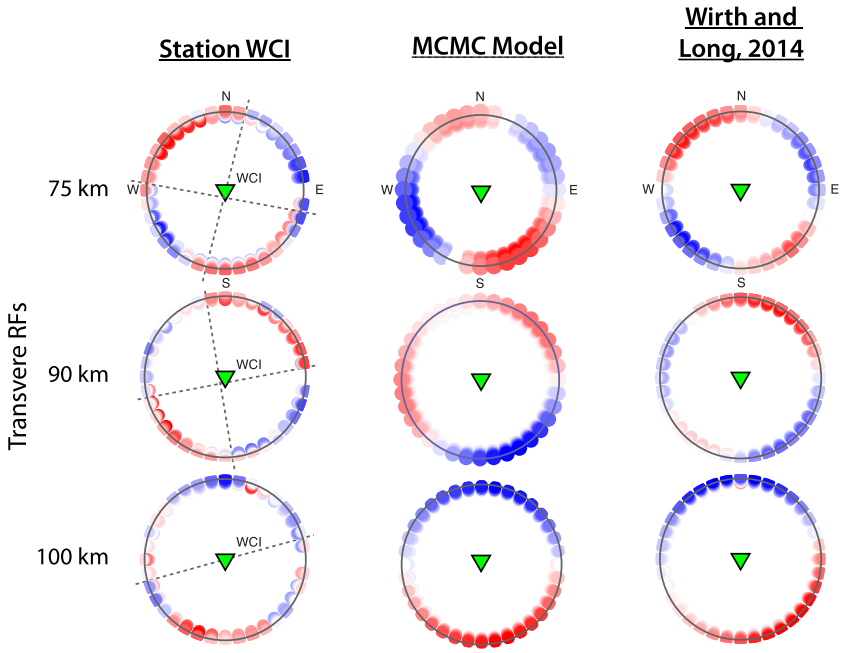

Figure 9. Comparison of *P*-to-*SH* conversions at various depths from station WCI (left), the best-fit model suggested by the McMC algorithm (middle) and the grid search results of Wirth & Long (2014; right). Each circle represents the polarity (red—negative and blue—positive) and amplitude (darker colours indicate greater amplitudes) of the Ps phase at a particular backazimuth. Polarity reversals should occur at backazimuths parallel and perpendicular to the anisotropic symmetry axis.

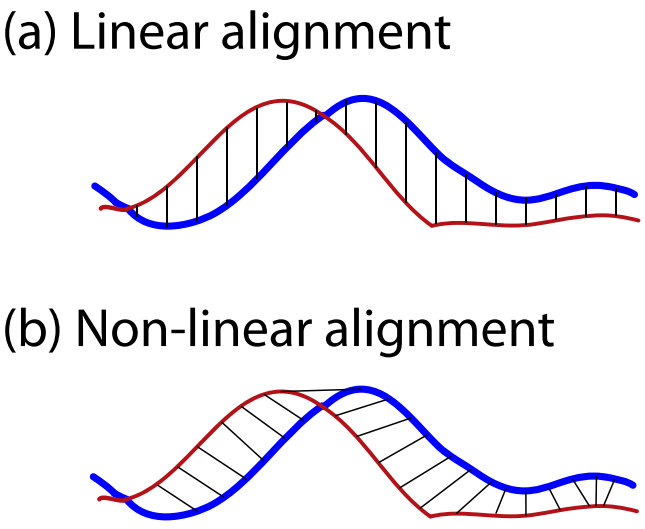

Figure 10. Example of two time-series that are out of phase. Although similar in shape and pattern, an L2-norm or χ^2 misfit would be unrealistically high when comparing these two time-series because of the offset in time (e.g. this offset in time could be due to one inaccurate parameter, for instance an incorrect layer thickness, even if all the other parameters are correct). However, if the misfit between the two time-series is characterized in a non-linear manner (i.e. using dynamic time warping), the misfit between these two time-series would be smaller than the corresponding χ^2 misfit. Figure modified after Rakthanmanon *et al*. (2012).

for layer velocities are off by <0.1 km s^{-1} for both V_P and V_S.) Although DTW did not prove to be a significant improvement over an L2-norm or χ^2 misfit for 1–2 layer models for the tests we ran, it is possible that DTW could potentially provide advantages when applied to complicated Earth structure that results in more complex receiver function patterns. Additionally, it may be a better choice for studies that focus on resolving the orientation of anisotropy or dipping interfaces, but do not attempt to constrain isotropic velocity profiles or precise interface depths.

6.2 Incorporating additional observational constraints

One of the most valuable results from this work is the demonstration that while certain parameters are well constrained by anisotropic receiver function data (e.g. the trend of the anisotropic symmetry axis and the strike of dipping interfaces), even in the case of complex, multilayered Earth structure, other parameters cannot be robustly constrained from receiver function data alone. If precise estimates of parameters such as absolute *P*- and *S*-wave velocities and layer thickness are desired, additional constraints are necessary. To alleviate the trade-offs between seismic velocity and interface depth, some previous studies have jointly inverted receiver function data, which is sensitive to sharp impendence contrasts, and surface wave dispersion data, which is sensitive to vertically averaged seismic velocity (V_S) structure (e.g. Julia *et al*. 2000; Shen *et al*. 2013; Obrebski *et al*. 2015). Future work will seek to incorporate additional constraints from surface wave data into the McMC algorithm presented here to improve the constraints on the aforementioned parameters.

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

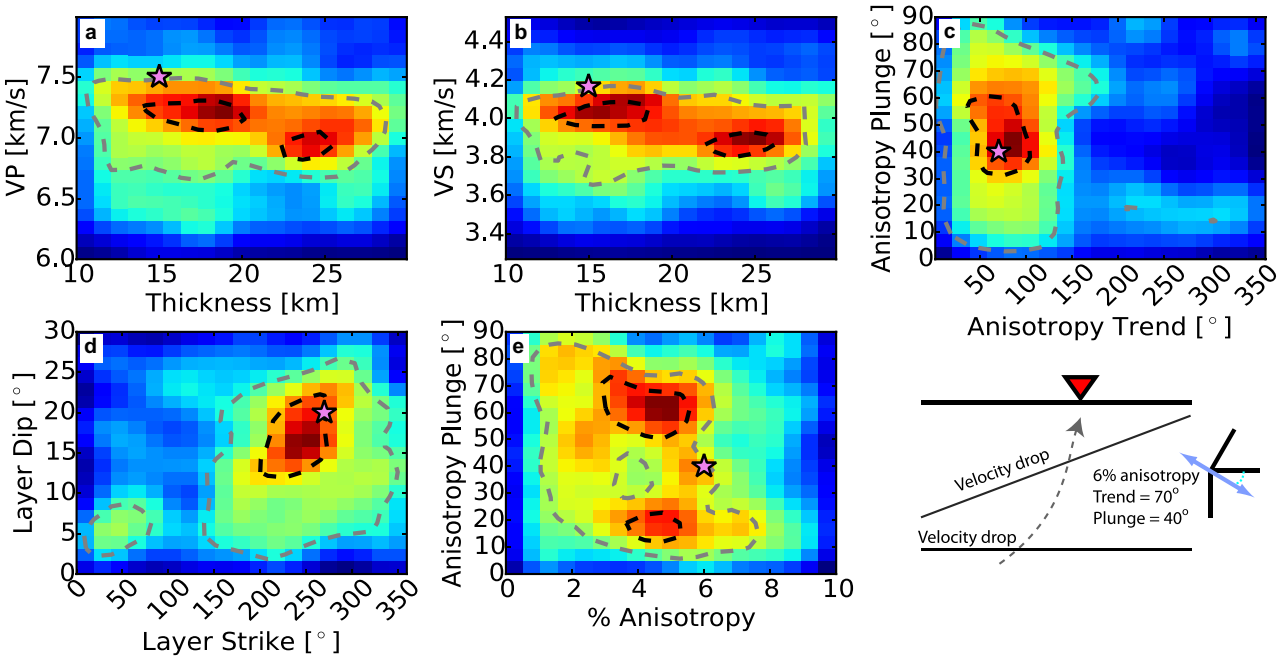

Figure 11. 2-D histograms for a synthetic example (Case D) in which parameters for one layer of the model were unknown (i.e. nine parameters total) and misfit was calculated using dynamic time warping. Plotting conventions are the same as in Fig. 3. Although it does not make a significant difference in terms of the interpretation of Earth structure, we find that parameter trade-offs are most clearly illustrated when we only plot the last third of the Markov chain. Therefore, for instructional purposes, here we only plot the last ∼4000 model iterations. A sketch of Earth structure beneath the receiver is provided in the lower right corner.

Additionally, in some regions, multiple phases (e.g. Moho reverberations) contribute significantly to the receiver function signal. Information from multiple phases could be incorporated into the algorithm presented here by either (1) employing a separate technique for inferring seismic velocity and layer thickness, and using this to guide *a priori* constraints on the parameters, or (2) using the McMC algorithm described above, but explicitly attempting to match multiple phases when generating synthetic receiver functions. In this study, we do not model multiple phases as they were found to not contribute significantly to the receiver functions generated in Wirth & Long (2014), but this could be easily adapted for future implementations.

7 SUMMARY

Model space searches such as the McMC algorithm presented here provide a framework for forward modeling of receiver functions to obtain estimates of parameters that describe Earth structure in a systematic and robust manner. We have developed and implemented an McMC with Gibbs sampling approach for the interpretation of radial and transverse component receiver functions, in an algorithm that is computationally efficient compared grid search approaches and straightforward to run in parallel. Synthetic tests suggest that this algorithm provides useful information regarding a reasonable number of free parameters ($<\sim$20), as long as the parameters being constrained actually contribute to prominent features in the receiver function data. Parameters that result in clear polarity reversals in *P*-to-*SH* converted energy as a function of backazimuth, such as the trend of anisotropy or the strike of a dipping interface, are the most robustly constrained parameters from anisotropic receiver function data. Meanwhile, trade-offs are apparent between parameters such as absolute seismic velocities and interface depth (which contribute to the timing of Ps phase arrivals), and the strength and plunge of anisotropy (which contribute to the amplitude of *P*-to-*SH* conversions). Possible extensions of the McMC algorithm presented here include unbiasing the ensemble of models to obtain quantitative error estimates, incorporating additional constraints from multiple phases or surface wave data to reduce parameter trade-offs, or using alternative misfit functions such as DTW, which may allow this method to be of use in more complex tectonic settings.

ACKNOWLEDGEMENTS

We thank Jun Korenaga for helpful comments and discussion related to this manuscript. Data from the Global Seismographic Network were obtained through the Incorporated Research Institutions for Seismology Data Management Center. Figures were created using the Generic Mapping Tools of Wessel & Smith (1991). This work was partially supported by an National Science Foundation Graduate Research Fellowship (EAW) and National Science Foundation grant EAR-1358325 (MDL). We acknowledge support from the facilities and staff of the Yale University Faculty of Arts and Sciences High Performance Computing Center. We are grateful to Malcom Sambridge and an anonymous reviewer for insightful comments that substantially improved this paper.

REFERENCES

Berndt, D.J. & Clifford, J., 1994. Using dynamic time warping to find patterns in time series, in *Workshop on Knowledge Discovery in Databases*, Vol. 10, pp. 359–370.

Bodin, T., Sambridge, M., Tkalcic, H., Arroucau, P., Gallagher, K. & Rawlinson, K., 2012. Transdimensional inversion of receiver functions and surface wave dispersion, *J. geophys. Res.*, **117,** B02301, doi:10.1029/2011JB008560.

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016

Bodin, T., Leiva, J., Romanowicz, B., Maupin, V. & Yuan, H., 2016. Imaging anisotropic layering with Bayesian inversion of multiple data types, *Geophys. J. Int.,* **206,** 605–629.

Chang, S., Baag, C. & Langston, C., 2004. Joint analysis of teleseismic receiver functions and surface wave dispersion using the genetic algorithm, *Bull. seism. Soc. Am.,* **94,** 691–704.

Dettmer, J., Benavente, R., Cummins, P.R. & Sambridge, M., 2014. Trans-dimensional finite-fault inversion, *Geophys. J. Int.,* **199,** 735–751.

Ford, H.A., Long, M.D. & Wirth, E.A., 2016. Mid-lithospheric discontinuities and complex anisotropic layering in the mantle lithosphere beneath the Wyoming and Superior Provinces, *J. geophys. Res.,* **121**(9), 6675–6697.

Frederiksen, A.W. & Bostock, M.G., 2000. Modelling teleseismic waves in dipping anisotropic structures, *Geophys. J. Int.,* **141,** 401–412.

Frederiksen, A.W., Folsom, H. & Zandt, G., 2003. Neighbourhood inversion of teleseismic Ps conversions for anisotropy and layer dip, *Geophys. J. Int.,* **155,** 200–212.

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

Julia, J., Ammon, C.J., Herrmann, R.B. & Correig, A.M., 2000. Joint inversion of receiver function and surface wave dispersion observations, *Geophys. J. Int.,* **143,** 99–112.

Korenaga, J. & Karato, S., 2008. A new analysis of experimental data on olivine rheology, *J. geophys. Res.,* **113,** B02403, doi:10.1029/2007JB005100.

Langston, C.A., 1979. Structure under Mount Rainier, Washington, inferred from teleseismic body waves, *Bull. seism. Soc. Am.,* **67,** 1029–1050.

Levin, V. & Park, J., 1997. *P-SH* conversions in a flat-layered medium with anisotropy of arbitrary orientation, *Geophys. J. Int.,* **131,** 253–266.

Liu, J.S., 2008. *Monte Carlo Strategies in Scientific Computing,* Springer.

McCormack, K., Wirth, E.A. & Long, M.D., 2013. B-type olivine fabric and mantle wedge serpentinization beneath the Ryukyu arc, *Geophys. Res. Lett.,* **40,** 1697–1702.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E., 1953. Equation of state calculations by fast computing machines, *J. Chem. Phys.,* **21,** 1087–1092.

Mosegaard, K. & Sambridge, M., 2002. Monte Carlo analysis of inverse problems, *Inverse Probl*., **18,** R29–R54.

Nikulin, A., Levin, V. & Park, J., 2009. Receiver function study of the Cascadia megathrust: evidence for localized serpentinization, *Geochem. Geophys. Geosyst.,* **10,** Q07004, doi:10.1029/2009GC002376.

Obrebski, M., Abers, G.A. & Foster, A., 2015. Magmatic arc structure around Mount Rainier, WA, from the joint inversion of receiver functions and surface wave dispersion, *Geochem. Geophys. Geosyst.,* **16,** 178–194.

Park, J. & Levin, V., 2000. Receiver functions from multiple-taper spectral correlation estimates, *Bull. seism. Soc. Am.,* **90,** 1507–1520.

Park, J. & Levin, V., 2016. Anisotropic shear zones revealed by back-azimuthal harmonics of teleseismic receiver functions, *Geophys. J. Int.,* **207**(2), 1216–1243.

Park, J., Yuan, H. & Levin, V., 2004. Subduction-zone anisotropy under Corvallis, Oregon: a serpentinite skidmark of trench-parallel terrane migration?, *J. geophys. Res.,* **109,** B10306, doi:10.1029/2003JB002718.

Piana Agostinetti, N. & Malinverno, A., 2010. Receiver function inversion by trans-dimensional Monte Carlo sampling, *Geophys. J. Int.,* **181,** 858–872.

Porter, R., Zandt, G. & McQuarrie, N., 2011. Pervasive lower-crustal seismic anisotropy in Southern California: evidence for underplated schists and active tectonics, *Lithosphere,* **3,** 201–220.

Rakthanmanon, T., Campana, B., Mueen, A., Batista, G., Westover, B., Zhu, Q., Zakaria, J. & Keogh, E., 2012. Searching and mining trillions of time series subsequences under dynamic time warping, in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 262–270, ACM, New York, NY.

Rothman, D.H., 1985. Nonlinear inversion statistical mechanics, and residual statics correlations, *Geophysics,* **50,** 2784–2796.

Rothman, D.H., 1986. Automatic estimation of large residual statics corrections, *Geophysics,* **51,** 332–346.

Sambridge, M., 1999a. Geophysical inversion with a neighborhood algorithm—I. Searching a parameter space, *Geophys. J. Int.,* **138,** 479–494.

Sambridge, M., 1999b. Geophysical inversion with a neighborhood algorithm—II. Appraising the ensemble, *Geophys. J. Int.,* **138,** 727–746.

Sambridge, M., 2014. A parallel tempering algorithm for probabilistic sampling and multimodal optimization, *Geophys. J. Int.,* **196,** 357–374.

Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.,* **40,** 3–1.

Schulte-Pelkum, V. & Mahan, K.H., 2014. A method for mapping crustal deformation and anisotropy with receiver functions and first results from USArray, *Earth planet. Sci. Lett.,* **402,** 221–233.

Shen, W., Ritzwoller, M.H. & Schulte-Pelkum, V., 2013. A 3-D model of the crust and uppermost mantle beneath the Central and Western US by joint inversion of receiver functions and surface wave dispersion, *J. geophys. Res.,* **118,** 262–276.

Sherrington, H.F., Zandt, G. & Frederiksen, A., 2004. Crustal fabric in the Tibetan Plateau based on waveform inversions for seismic anisotropy parameters, *J. geophys. Res.,* **109**(B2), doi:10.1029/2002JB002345.

Shibutani, T., Sambridge, M. & Kennett, B., 1996. Genetic algorithm inversion for receiver functions with application to the crust and uppermost mantle structure beneath eastern Australia, *Geophys. Res. Lett.,* **23,** 1829–1832.

Shiomi, K. & Park, J., 2008. Structural features of the subducting slab beneath the Kii Peninsula, central Japan: seismic evidence of slab segmentation, dehydration, and anisotropy, *J. geophys. Res.,* **113**(B10), doi:10.1029/2007JB005535.

Vinnik, L., Reigber, C., Aleshin, I., Kosarev, G., Kaban, M., Oreshin, S. & Roecker, S., 2004. Receiver function tomography of the central Tien Shan, *Earth planet. Sci. Lett.,* **225,** 131–146.

Wirth, E.A. & Long, M.D., 2012. Multiple layers of seismic anisotropy and a low-velocity region in the mantle wedge beneath japan: evidence from teleseismic receiver functions, *Geochem. Geophys. Geosyst.,* **13,** Q08005, doi:10.1029/2012GC004180.

Wirth, E.A. & Long, M.D., 2014. A contrast in anisotropy across mid-lithospheric discontinuities beneath the central United States—a relic of craton formation, *Geology,* **42,** 851–854.

Downloaded from http://gji.oxfordjournals.org/ at Yale University on November 30, 2016