

A Bayesian framework for fracture characterization from surface seismic data

S. Ahmad Zamanian, Michael C. Fehler, Xinding Fang and Daniel R. Burns

Earth Resources Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: zamanian@alum.mit.edu

Accepted 2014 August 19. Received 2014 June 28; in original form 2014 February 20

SUMMARY

We describe a methodology for quantitatively characterizing the fractured nature of a hydrocarbon or geothermal reservoir from surface seismic data under a Bayesian inference framework. The method combines different kinds of measurements of fracture properties to find a best-fitting model while providing estimates of the uncertainty of model parameters. Fractures provide pathways for fluid flow in a reservoir, and hence, knowledge about a reservoir's fractured nature can be used to enhance production from the reservoir. The fracture properties of interest in this study (to be inferred) are fracture orientation and excess compliance, where each of these properties are assumed to vary spatially over a 2-D horizontal grid which is assumed to represent the top of a reservoir. The Bayesian framework in which the inference problem is cast has the key benefits of (1) utilization of a prior model that allows geological information to be incorporated, (2) providing a straightforward means of incorporating all measurements (across the 2-D spatial grid) into the estimates at each gridpoint, (3) allowing different types of measurements to be combined under a single inference procedure and (4) providing a measure of uncertainty in the estimates. The observed data are taken from a 2-D array of surface seismic receivers responding to an array of surface sources. Well understood features from the seismic traces are extracted and treated as the observed data, namely the *P*-wave reflection amplitude variation with acquisition azimuth and offset (amplitude versus azimuth data) and fracture transfer function (FTF) data. Amplitude versus azimuth data are known to be more sensitive to fracture properties when the fracture spacing is significantly smaller than the seismic wavelength, whereas FTF data are more sensitive to fracture properties when the fracture spacing is on the order of the seismic wavelength. Combining these two measurements has the benefit of allowing inferences to be made about fracture properties over a larger range of fracture spacing than otherwise attainable. Geophysical forward models for the measurements are used to arrive at likelihood models for the data. The prior distribution for the fracture variables is obtained by defining a Markov random field over the lateral 2-D grid where we wish to obtain fracture properties, where this method for defining the prior has the added benefit of allowing for non-stationarity in the resulting model covariance. The fracture variables are then inferred by application of loopy belief propagation to yield approximations for the posterior marginal distributions of the fracture properties, as well as the maximum *a posteriori* and Bayes least-squares (posterior mean) estimates of these properties. Verification of the inference procedure is performed using a synthetic data set, where the estimates obtained are shown to be at or near ground truth for the full range of fracture spacings for fracture orientation and at low fracture spacings for excess compliance estimates.

Key words: Probability distributions; Fracture and flow; Seismic anisotropy; Computational seismology; Statistical seismology; Wave scattering and diffraction.

1 INTRODUCTION

Fractures are cracks in the Earth's crust through which fluid, such as oil, natural gas or brine, can flow. Knowledge about the presence and properties of fractures in a reservoir can be extremely

valuable, as such information can be used to determine pathways for fluid flow and to optimize production from the reservoir (Sayers 2009; Ali & Jakobsen 2011). Since the presence of fractures in an elastic medium can alter the compliance of the medium and fractures often have a preferred alignment relative to *in situ* stress,

fractures can cause the medium to exhibit anisotropy (Schoenberg & Sayers 1995). The anisotropy has been exploited to give different techniques for determining fracture properties from seismic data, such as reflection amplitude versus offset and azimuth analysis (Sayers & Rickett 1997; Rüger 1998; Lynn *et al.* 2009; Sayers 2009) and shear wave birefringence (Gaiser & Van Dok 2001). These methods, however, are only valid when the fracture spacing is small in comparison to the seismic wavelength, so that seismic waves average over the fractures (Willis *et al.* 2006; Fang *et al.* 2014). This equivalent anisotropic medium assumption, however, breaks down when the spacing between the fractures increases to being on the order of the seismic wavelength. For the case of larger fracture spacings, Willis *et al.* (2006) proposed a technique, referred to as the scattering index method, to estimate the azimuthal orientation (or strike) of a fracture system based on the scattered seismic energy. Fang *et al.* (2014) described a series of modifications to this technique to give a more robust methodology for determining fracture orientation, which is referred to as the fracture transfer function (FTF) method.

In the way of statistical inference methods applied to geophysical problems, Eidsvik *et al.* (2004) gave a Bayesian framework for determining rock facies and saturating fluid by integrating a forward rock physics model with spatial statistics of rock properties. Specifically in the area of fracture characterization, Ali & Jakobsen (2011) used a Bayesian inference framework to infer fracture orientation and density from seismic velocity and attenuation anisotropy data. Sil & Srinivasan (2009) applied a similar Bayesian inference methodology to determine fracture strikes from seismic and well data. All of the aforementioned statistical studies solved the inference problem via Markov chain Monte Carlo (MCMC), a sampling technique which stochastically searches the model space. Furthermore, the data models used in all of these studies follow from the assumption that a medium with closely spaced fractures is an equivalent anisotropic medium.

The aim of our study is to utilize a Bayesian framework to combine different data which, on their own, are informative about certain regimes of fracture spacing, to be able to estimate different fracture properties (particularly, excess fracture compliance and fracture orientation) at a wider range of fracture spacings than otherwise attainable if one data type is used. The Bayesian framework in which the problem is cast furthermore makes it straightforward to encode prior knowledge about either geological features of the reservoir, such as the existence of a discontinuity arising from a geological fault, or known information about the fracture properties or their spatial correlation, where we assume that the fracture properties can vary spatially over a 2-D lateral grid.

We first describe the physical parameters that we use to characterize a fracture system in a reservoir whose properties vary with space. We then present our approach for casting the fracture characterization problem within a Bayesian framework in which we assume that fracture properties within the reservoir are spatially correlated. We capture this spatial correlation by defining a Markov random field (MRF) over the grid of fracture properties, which we use to arrive at the prior distribution for the fracture properties. We then introduce the types of seismic data that we assume are available to characterize fractures, particularly amplitude versus azimuth (Sayers & Rickett 1997; Rüger 1998) and FTF (Fang *et al.* 2014), and describe the methods by which these data are used to place constraints on the properties. The inversion for model parameters is accomplished via loopy belief propagation (LBP; Pearl 1982, 1988; Murphy *et al.* 1999), a numerically scalable approximate inference procedure that yields both the posterior marginal distributions of the properties at each point in space in addition to the Bayes least

squares (BLS) and maximum *a posteriori* (MAP) estimates of the fracture properties. Finally, we demonstrate the applicability of the method for inferring fracture properties using synthetic seismic data.

In this paper, we use **sans-serif** font to denote random variables and **boldface** font to denote vector or matrix quantities.

2 DESCRIPTION OF THE PROBLEM

Consider a set of seismic measurements taken from a 2-D array of surface receivers over a layered medium responding to a set of surface seismic sources. We are interested in inferring from the seismic data whether or not fractures are present in a particular layer of the medium (e.g. the reservoir) and, if so, the properties of the fractures. A simple example of this setup, where the medium consists of flat homogeneous layers, is displayed in Fig. 1. In particular, we would like to infer fracture orientation $\boldsymbol{\varphi} = [\varphi_{ij}]$ and the (base 10) log excess fracture compliance $\mathbf{z} = [z_{ij}]$ spatially over a 2-D m -by- n lateral grid $\mathbf{L} = \{1, \dots, m\} \times \{1, \dots, n\}$ (where i and j index the axes of the grid \mathbf{L}). Each gridpoint corresponds to a square of area ℓ^2 , so that the entire grid corresponds to a region of area $mn\ell^2$. Excess fracture compliance is defined as the overall additional medium compliance (having units of Pa^{-1}) due to the presence of fractures and is the ratio of the compliance of the individual fractures (in m Pa^{-1}) to the fracture spacing (in m) (Schoenberg & Sayers 1995; Dubos-Sallée & Rasolofosaon 2008). We make the simplifying assumptions (1) that the fractures are vertical, so that the fracture orientation φ_{ij} is simply the azimuth (or strike) of the fractures (with respect to north) and (2) that the normal and tangential excess compliances are equal, which may represent gas-filled fractures (Sayers *et al.* 2009), hence we need only infer a single log excess compliance value z_{ij} for each grid node. It is possible that the ratio of normal to tangential fracture compliances may deviate from unity due to mineralization (Sayers *et al.* 2009); while this ratio has been measured using shear wave splitting data (Verdon & Wüstefeld 2013), we are not able to uniquely resolve this ratio from AvAz data. If this ratio is known to differ from unity for a particular fracture system, we may proceed with our analysis inferring for only the normal compliance and setting the tangential compliance according to this ratio. An excess compliance value of 0 at a particular gridpoint is taken to mean there are no fractures at that gridpoint (rendering the value for azimuth arbitrary and meaningless). In order to compare zero and non-zero compliance on a logarithmic scale, we treat an excess compliance

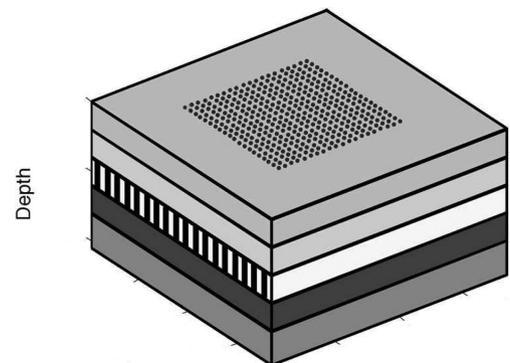


Figure 1. A simple model of the problem setting. The formation consists of five flat homogeneous layers with fractures that may be present in the third layer and measurements obtained from the 2-D array of surface seismic receivers. [Figure modified from Willis *et al.* (2006).]

of zero as 10^{-13}Pa^{-1} , which is geophysically reasonable as this is an insignificantly small value for excess compliance and results in a negligible effect on seismic wave propagation. We assume that the data set is rich enough so that for each gridpoint in \mathbf{L} , there are corresponding source–receiver pairs that sample the point at multiple offsets and acquisition azimuths. We further assume that the background velocity structure of the medium is well understood.

In order to relate the fracture properties $\mathbf{x} = (\mathbf{z}, \boldsymbol{\varphi}) = [\mathbf{x}_{ij}]$ to the seismic trace data set, it is necessary to model seismic data as a function of the fracture properties. Unfortunately, modelling the entire seismic trace data set requires a full elastic 3-D forward simulation of the seismic wavefield, and the computational cost associated with the repeated simulations required to invert for the fracture properties is prohibitively high, so we instead resort to simulating well understood features of the seismic trace data set and treat these features as our observed data \mathbf{y} . In particular, we choose to model P -wave reflection amplitude as a function of acquisition azimuth (at a fixed angle of incidence), also known as amplitude versus azimuth (AvAz) data (Sayers & Rickett 1997; Rüger 1998), and FTF data, as defined by Fang *et al.* (2014). We refer to these observed data with variables \mathbf{y}^{AvAz} and \mathbf{y}^{FTF} , respectively, and let $\mathbf{y} = (\mathbf{y}^{\text{AvAz}}, \mathbf{y}^{\text{FTF}})$. Detailed descriptions of the data and their forward models are detailed in Section 3.2. Both $\mathbf{y}^{\text{AvAz}} = [\mathbf{y}_{ij}^{\text{AvAz}}]$ and $\mathbf{y}^{\text{FTF}} = [\mathbf{y}_{ij}^{\text{FTF}}]$ are defined over the grid \mathbf{L} , in a manner such that to each grid node of fracture properties \mathbf{x}_{ij} there is an associated data vector \mathbf{y}_{ij} .

3 BAYESIAN INFERENCE FRAMEWORK

In order to arrive at an estimation of the fracture properties from the seismic data, we employ a Bayesian inference framework. As mentioned earlier, the Bayesian framework is chosen as it allows us to naturally encode prior information about the fracture properties (and their spatial variation), combine different types of data, and quantify the uncertainty associated with the inferred quantities. The fracture properties and seismic data are treated as random variables, and a stochastic model is used to give the joint distribution of the fracture properties and seismic data $(\mathbf{x}, \mathbf{y}) = ((\mathbf{z}, \boldsymbol{\varphi}), \mathbf{y})$. In particular, we model the fracture properties as discrete random variables where the domain for each of the variables is given by: $10^{\mathbf{z}_{ij}} \in \mathcal{Z} = \{10^{-9.0}, 10^{-9.1}, \dots, 10^{-12.0}, 10^{-13}\}$ (in units of Pa^{-1}) and $\varphi_{ij} \in \mathcal{F} = \{0^\circ, 20^\circ, \dots, 160^\circ\}$, $\forall (i, j) \in \mathbf{L}$. The use of discrete (rather than continuous) random variables makes the inference problem amenable to the general framework of message-passing inference algorithms described in Section 3.3, where the intervals of discretization were picked based on the level of resolution we might reasonably expect to achieve using seismic measurements. The posterior distribution of the fracture properties given the data $p(\mathbf{x}|\mathbf{y})$ is given by Bayes' rule:

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{x})p(\mathbf{y}|\mathbf{x})}{\sum_{\mathbf{x}'} p(\mathbf{x}')p(\mathbf{y}|\mathbf{x}')} \propto p(\mathbf{x})p(\mathbf{y}|\mathbf{x}), \quad (1)$$

where $p(\mathbf{x})$ and $p(\mathbf{y}|\mathbf{x})$ are the prior distribution of the fracture properties and the distribution of the seismic data given the fracture properties, respectively.

While the posterior distribution of eq. (1) is the complete solution to the Bayesian inference problem, exploring this distribution can be intractable due to the high-dimensionality of the fracture properties \mathbf{x} . To glean meaningful inferences from the posterior distribution, one may either choose to obtain point estimators of \mathbf{x} from the posterior or to obtain marginal posterior distributions over some

tractably explorable subsets of the random variables in \mathbf{x} . While point estimators are useful when a single answer to the inference problem is desired, they do not capture the associated estimation uncertainties (which are fully described by the marginal posterior distributions). The most common point estimators are the MAP estimate $\hat{\mathbf{x}}_{\text{MAP}}$ and the BLS (or mean) estimate $\hat{\mathbf{x}}_{\text{BLS}}$. The MAP estimate of the fracture properties minimizes the probability of estimation error and is the overall configuration of the fracture properties that maximizes the posterior distribution, that is

$$\hat{\mathbf{x}}_{\text{MAP}} \in \arg \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}), \quad (2)$$

where set membership is used in place of equality in case the maximizing value is non-unique. In contrast, the BLS estimate of the fracture properties minimizes the expected value of the squared estimation error and is given by the expected value of the fracture properties given the data. Since the BLS estimator takes into account the value of the estimation error it can be more sensitive to outlying values in the posterior distribution than the MAP estimator; this may be an appropriate choice for posterior distributions having significant probabilities over a broad range of values, whereas the MAP estimator may be more appropriate for peaky or multimodal posteriors. The posterior marginal distribution for the fracture properties at a particular node \mathbf{x}_{ij} is given by summation of the posterior distribution over all other variables $\mathbf{x}_{-ij} \triangleq \{x_{kl} : (k, l) \in \mathbf{L} \setminus \{(i, j)\}\}$. So, for example, the posterior marginal for the log excess compliance $p(z_{ij}|\mathbf{y})$ is given by

$$p(z_{ij}|\mathbf{y}) = \sum_{\varphi_{ij}} \sum_{\mathbf{x}_{-ij}} p(z_{ij}, \varphi_{ij}, \mathbf{x}_{-ij}|\mathbf{y}). \quad (3)$$

Computing the posterior marginals has the additional benefit of yielding (with minimal additional computation) the BLS estimates. For example, for the log excess compliance at node (i, j) , we have

$$\hat{z}_{ij, \text{BLS}} = \mathbb{E}[\mathbf{z}_{ij}|\mathbf{y} = \mathbf{y}] = \sum_{z_{ij}} z_{ij} p(z_{ij}|\mathbf{y}). \quad (4)$$

For any reasonably large number of grid nodes mn , the maximization and summation in eqs (2) and (3), respectively, are intractable, hence we must turn to approximate inference algorithms to perform the estimation. We discuss the inference algorithms used to approximate the MAP estimates and posterior marginals in Section 3.3.

3.1 Prior model

Assuming that the fracture properties will not change rapidly with position, it is reasonable to make the properties at one point depend on its nearest neighbours in space. We capture this spatial dependence mathematically by carefully constructing an appropriate prior model for the fracture properties. We arrive at a prior model by defining the set of fracture properties \mathbf{x} as a Markov random field over an undirected graphical model $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ on the 2-D grid \mathbf{L} (Koller & Friedman 2009). Here \mathcal{V} is the set of vertices or nodes of the graphical model, which correspond to partitions of the random variables in \mathbf{x} . We associate with each grid node $(i, j) \in \mathbf{L}$ a vertex in \mathcal{V} corresponding to the pair of random variables $\mathbf{x}_{ij} = (\mathbf{z}_{ij}, \varphi_{ij})$, so that $\mathcal{V} \equiv \mathbf{L}$. The set of edges of the graph is $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, represented as pairs of nodes, which encode dependencies between the random variables. In particular, a set of random variables form a Markov random field over an undirected graph \mathcal{G} if it is globally Markov on \mathcal{G} , meaning that, for any disjoint subsets $A, B, C \subset \mathcal{V}$ such that C separates A from B on \mathcal{G} , we have $\mathbf{x}_A \perp\!\!\!\perp \mathbf{x}_B | \mathbf{x}_C$ (read as \mathbf{x}_A is conditionally independent of \mathbf{x}_B

given \mathbf{x}_C). Here, $\mathbf{x}_A \triangleq \{\mathbf{x}_{ij} : (i, j) \in A\}$ (similarly for \mathbf{x}_B and \mathbf{x}_C), and C separates A from B if every path from any node in A to any node in B passes through C . The Hammersley–Clifford theorem (Hammersley & Clifford 1971; Clifford 1990) states that a distribution is globally Markov on a graphical model \mathcal{G} if it factorizes over the maximal cliques of that graph. Here a clique C of a graph \mathcal{G} is any subset of its vertices ($C \subset \mathcal{V}$) which are fully connected, meaning every vertex in C shares an edge from \mathcal{E} with every other vertex in C . The maximal cliques of a graph \mathcal{G} are its largest possible cliques: C is a maximal clique of \mathcal{G} if it fails to remain a clique when even one more vertex from $\mathcal{V} \setminus C$ is added to C . We say that a distribution factorizes over the maximal cliques of \mathcal{G} if it can be written as a product of functions of the random variables in each maximal clique C of \mathcal{G} .

To proceed with our construction of the prior distribution, we define the edge set \mathcal{E} over the 2-D grid \mathbf{L} so that a particular node shares edges with its four neighbours on the grid \mathbf{L} . The graphical model for \mathbf{x} prior to observing the data is shown in Fig. 2. Intuitively, such a graph structure means that given the fracture properties of the four nearest neighbours of a particular node, knowledge of the fracture properties of the medium elsewhere on the grid will have no impact on our belief about the properties at that node. Since we expect the properties of the medium at a particular point in space to be similar to its surrounding properties, this suggests a prior distribution that penalizes differences between a node and its neighbours. In particular we define the prior distribution for the fracture parameters to be

$$p(\mathbf{z}) \propto \exp \left\{ - \sum_{(ij,kl) \in \mathcal{E}} \beta_{z_{ij,kl}} (z_{ij} - z_{kl})^2 \right\}, \quad (5)$$

and

$$p(\boldsymbol{\varphi}) \propto \exp \left\{ - \sum_{(ij,kl) \in \mathcal{E}} \beta_{\varphi_{ij,kl}} (\varphi_{ij} - \varphi_{kl})^2 \right\}, \quad (6)$$

where $\beta_{z_{ij,kl}}$ and $\beta_{\varphi_{ij,kl}}$ are smoothness parameters. Note that the fracture orientations must be manipulated with a modulo operation to bring the difference within the interval $[-90^\circ, 90^\circ]$, as the azimuth is identical modulo 180° .

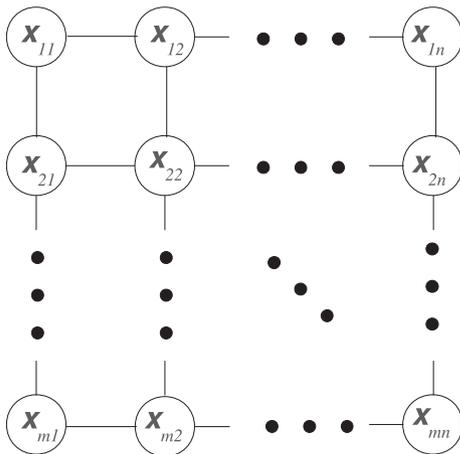


Figure 2. Undirected graphical model \mathcal{G} over which \mathbf{x} is Markov, prior to observing the seismic data. The model is based on the assumption that the values of the fracture properties at one location are dependent on those of its nearest neighbours.

We define the smoothness parameters in terms of an overall (spatially varying) smoothness parameter $\beta_{ij,kl}$ after normalizing by the bin sizes for each variable (so that the degree of smoothness is not dependent on the units of the variables) via $\beta_{z_{ij,kl}} = \beta_{ij,kl}/(0.1)^2$ and $\beta_{\varphi_{ij,kl}} = \beta_{ij,kl}/(20)^2$. Allowing spatial variation in smoothness parameter allows us to encode *a priori* information about discontinuities in the medium, such as those which may arise from a geological fault. In general, we pick a single value $\beta_c > 0$ for the smoothness parameter along edges where there are no known discontinuities and set the smoothness parameter to zero along edges where a fault is known to exist $\beta_f = 0$ (thereby removing those edges from the graph). We experiment with different choices for β_c in Section 4.2. Defining \mathcal{E}_f to be the set of edges where a fault is known to exist, then

$$\beta_{ij,kl} = \begin{cases} \beta_c & \text{if } [(i, j), (k, l)] \in \mathcal{E} \setminus \mathcal{E}_f \\ \beta_f & \text{if } [(i, j), (k, l)] \in \mathcal{E}_f \end{cases}. \quad (7)$$

Treating the two different fracture properties as independent *a priori* gives the overall prior distribution for the fracture properties as

$$p(\mathbf{x}) = p(\mathbf{z}, \boldsymbol{\varphi}) = p(\mathbf{z})p(\boldsymbol{\varphi}). \quad (8)$$

Indeed, we see that $p(\mathbf{x})$ factorizes over the maximal cliques of \mathcal{G} (which are precisely the edges \mathcal{E}) and thus by Hammersley–Clifford, \mathbf{x} is Markov on \mathcal{G} .

This graph theoretic approach to defining the prior distribution via a Markov random field contrasts with the more typical method for defining priors in geophysical inverse problems using stationary covariance functions (typically over an assumed Gaussian random field; Tarantola 2004). While defining a covariance function explicitly encodes the covariances between any two points in space, an MRF instead encodes the conditional independencies and local smoothness structure of the model, thereby implicitly defining a covariance function. Simpson *et al.* (2012) provides a nice connection between a stationary MRF and its resulting stationary covariance function through the use of stochastic partial differential equations. The MRF approach however can be more general as it has the further benefit of allowing non-stationarity in the local smoothness parameters (as in eq. 7) and, hence, in the induced model covariance function and is not restricted to Gaussian fields.

3.2 Likelihood model

As discussed in Section 2, the seismic data used in this study are AvAz and FTF data, which are extracted from the seismic trace data set and denoted by \mathbf{y}^{AvAz} and \mathbf{y}^{FTF} , respectively. We make the assumption that given the fracture parameters \mathbf{x} , the two types of seismic data \mathbf{y}^{AvAz} and \mathbf{y}^{FTF} are conditionally independent, and hence

$$p(\mathbf{y}|\mathbf{x}) = p(\mathbf{y}^{\text{AvAz}}|\mathbf{x})p(\mathbf{y}^{\text{FTF}}|\mathbf{x}). \quad (9)$$

In the remainder of this section, we discuss how we model the data to arrive at the likelihood models $p(\mathbf{y}^{\text{AvAz}}|\mathbf{x})$ and $p(\mathbf{y}^{\text{FTF}}|\mathbf{x})$.

3.2.1 Amplitude versus azimuth data

We suppose we have data for the amplitudes of P – P arrivals reflected from the top of the fractured layer at a full range of acquisition azimuths and source–receiver offsets for each gridpoint $(i, j) \in \mathbf{L}$. We can use ray tracing to determine spatially the gridpoint corresponding to each source–receiver pair as well as to map the offsets to

incidence angles for the wave incident on the top of the fractured layer (where the incidence angle is the angle the incident wave makes with the vertical axis). This gives, for each gridpoint (i, j) a set of P - P reflection amplitudes that vary with incidence angle $\theta \in \Theta$ and acquisition azimuth (relative to north) $\phi^{\text{Acq}} \in \Phi^{\text{Acq}}$, where Θ and Φ^{Acq} are the sets of incidence angles and acquisition azimuths over which the data has been obtained; we denote the reflection amplitudes by $\hat{R}_{ij}^{PP}(\theta, \phi^{\text{Acq}})$. For concreteness, suppose the acquisition azimuths we have are precisely the set $\Phi^{\text{Acq}} = \{0^\circ, 10^\circ, \dots, 170^\circ\}$. In order to be able to compare these amplitudes to the P -wave reflection coefficient, for each incidence angle θ , we normalize the amplitudes by the mean amplitude (taken over acquisition azimuths). This allows us to compare the variation of the reflection coefficient with azimuth (rather than its absolute value):

$$\mathbf{y}_{ij}^{\text{AvAz}} = \mathbf{y}_{ij, \theta, \phi^{\text{Acq}}}^{\text{AvAz}} = \frac{\hat{R}_{ij}^{PP}(\theta, \phi^{\text{Acq}})}{\frac{1}{|\Phi^{\text{Acq}}|} \sum_{\phi \in \Phi^{\text{Acq}}} \hat{R}_{ij}^{PP}(\theta, \phi)}. \quad (10)$$

In order to arrive at a forward model for the P -wave reflection coefficient of the interface above the fractured layer as a function of acquisition azimuth, we make various simplifying assumptions about the formation and the fractured medium. The layers above the fractured layer are assumed to be isotropic and homogeneous and the background medium of the layer in which the fractures exist is assumed to be homogeneous and isotropic with known medium parameters. We assume that the presence of fractures in the fractured layer causes the layer to behave as an equivalent anisotropic medium, which is a geophysically valid assumption when the fracture spacing is small compared to the seismic wavelength (Schoenberg & Douma 1988; Willis *et al.* 2006). In this case, it is reasonable to assume that the presence of a parallel set of vertical fractures causes the medium to exhibit horizontal transverse isotropy (HTI) with a symmetry axis normal to the strike of the fractures (Rüger 1998; Tsvankin 2001). A transverse isotropic medium with a given symmetry axis means seismic wave propagation in all directions that form the same angle with the symmetry axis is equivalent. As such, in an HTI medium resulting from a set of parallel vertical fractures, the plane normal to the symmetry axis (and parallel to the fractures) is referred to as the isotropy plane, as wave propagation is equivalent in all directions in this plane (Tsvankin 2001).

The P -wave reflection coefficient of an interface is defined as the ratio of the reflected P -wave amplitude to the incident P -wave amplitude on the interface. Rüger (1998) derives the P -wave polarization vector and P -wave phase velocities in an HTI medium and uses these to solve a system of perturbation equations for the reflection and transmission coefficients at the interface of two HTI media having the same symmetry axes. The resultant P -wave reflection coefficient is given as a function of the incidence phase angle (θ , the angle the incident P -wave makes with the vertical axis) and the azimuthal phase angle (ϕ , the azimuth of the incident P -wave relative to the symmetry axis), and in terms of the isotropic background and anisotropy parameters, as

$$\begin{aligned} R^{PP}(\theta, \phi) = & \frac{1}{2} \frac{\Delta Z}{\bar{Z}} + \frac{1}{2} \left\{ \frac{\Delta \alpha}{\bar{\alpha}} - \left(\frac{2\bar{\beta}}{\bar{\alpha}} \right)^2 \frac{\Delta G}{\bar{G}} \right. \\ & \left. + \left[\Delta \delta^{(V)} - 2 \left(\frac{2\bar{\beta}}{\bar{\alpha}} \right)^2 \Delta \gamma^{(V)} \right] \cos^2 \phi \right\} \sin^2 \theta \\ & + \frac{1}{2} \left(\frac{\Delta \alpha}{\bar{\alpha}} + \Delta \epsilon^{(V)} \cos^4 \phi + \Delta \delta^{(V)} \sin^2 \phi \cos^2 \phi \right) \sin^2 \theta \tan^2 \theta, \end{aligned} \quad (11)$$

where α is the vertical P -wave velocity, β is the vertical velocity of the S -wave polarized parallel to the isotropy plane, ρ is the medium density, $Z = \rho \alpha$ is the vertical P -wave impedance and $G = \rho \beta^2$ is the vertical shear modulus; these parameters are all from the background isotropic model and are assumed to be known in our analysis. Rüger (2002) and Liu & Martinez (2013) note that this linearized equation for the P -wave reflection coefficient is accurate for small medium contrasts and weak anisotropy at angles of incidence less than 35° . In cases where a more accurate model is required, one may wish to use the approximations of the P -wave reflection coefficient given by Ursin & Haugen (1996) or Pšeničák & Martins (2001). The parameters $\delta^{(V)}$, $\epsilon^{(V)}$ and $\gamma^{(V)}$ are the Thomsen anisotropy parameters defined with respect to the vertical axis (Thomsen 1986); these parameters are identically zero for an isotropic medium, but will depend on the fracture properties for an HTI medium. The parameters in eq. (11) are defined in terms of their relative differences between the upper and lower media $\Delta(\cdot)$ and their average values $(\bar{\cdot})$. So, for example, $\Delta \alpha = \alpha_2 - \alpha_1$ and $\bar{\alpha} = (\alpha_1 + \alpha_2)/2$, where α_1 and α_2 are the vertical P -wave velocities of the upper and lower media, respectively. Since the axis of symmetry is normal to the strike of the fractures (thus having an azimuth relative to north of $\phi_{ij} + 90^\circ$), then with ϕ^{Acq} as the known azimuth of the incident P -wave relative to north, we have

$$\phi = \phi^{\text{Acq}} - \phi_{ij} - 90^\circ. \quad (12)$$

Furthermore, we set the incidence angle θ to the values computed for the AvAz data.

We use the linear slip model of Schoenberg & Sayers (1995) to express the Thomsen anisotropy parameters of the fractured medium in terms of the excess fracture compliance. The details of this derivation are given in Appendix. Combining this with (11) gives the forward model for the P - P reflection coefficient as a function of the fracture parameters at node (i, j) , which we denote by $R_{ij}^{PP}(\theta, \phi^{\text{Acq}}, \varphi_{ij}, z_{ij})$. To make this comparable to the data $\mathbf{y}_{ij}^{\text{AvAz}}$ defined in (10), we process it in the same manner by normalizing by the mean reflection coefficient, for each incidence angle θ , over all acquisition azimuths, giving

$$\bar{R}_{ij}^{PP}(\theta, \phi^{\text{Acq}}, \varphi_{ij}, z_{ij}) = \frac{R_{ij}^{PP}(\theta, \phi^{\text{Acq}}, \varphi_{ij}, z_{ij})}{\frac{1}{|\Phi^{\text{Acq}}|} \sum_{\phi \in \Phi^{\text{Acq}}} R_{ij}^{PP}(\theta, \phi, \varphi_{ij}, z_{ij})} \quad (13)$$

as the deterministic forward model for $\mathbf{y}_{ij, \theta, \phi^{\text{Acq}}}^{\text{AvAz}}$. We note that if the excess fracture compliance is zero, then the P - P reflection coefficient is constant with respect to acquisition azimuth, and hence does not vary with the fracture orientation φ_{ij} . This is consistent with our interpretation of zero compliance to mean the absence of fractures, which indeed renders the value of φ_{ij} arbitrary.

We arrive at a stochastic model for the data by assuming the output of the forward model is perturbed by zero-mean additive independent, identically distributed (i.i.d.) Gaussian noise, so that

$$\mathbf{y}_{ij, \theta, \phi^{\text{Acq}}}^{\text{AvAz}} = \bar{R}_{ij}^{PP}(\theta, \phi^{\text{Acq}}, \varphi_{ij}, z_{ij}) + \mathbf{w}_{ij, \theta, \phi^{\text{Acq}}}, \quad (14)$$

where the $\mathbf{w}_{ij, \theta, \phi^{\text{Acq}}}$ are mutually independent Gaussian random variables that are distributed as $\mathcal{N}(0, \sigma_{ij, \text{AvAz}}^2)$. We estimate the variance $\sigma_{ij, \text{AvAz}}^2$ using synthetic data obtained from a finite difference simulation of the seismic wavefield; the details of the synthetic data are described in Section 4.1. Processing of the synthetic data gives a set of single observations of the data $\mathbf{y}_{ij, \theta, \phi^{\text{Acq}}}^{\text{AvAz}}$ at a single grid node and at a range of incidence angles and acquisition azimuths, where the fracture properties (z_{ij}, φ_{ij}) are known, thus giving independent samples for the noise

$$\mathbf{w}_{ij} = \mathbf{w}_{ij, \theta, \phi^{\text{Acq}}} = \mathbf{y}_{ij, \theta, \phi^{\text{Acq}}}^{\text{AvAz}} - \bar{R}_{ij}^{PP}(\theta, \phi^{\text{Acq}}, \varphi_{ij}, z_{ij}). \quad (15)$$

We estimate $\sigma_{ij,AvAz}^2$ via its maximum likelihood (ML) estimator, which is given by

$$\hat{\sigma}_{ij,AvAz,ML}^2 = \hat{\sigma}_{ij,AvAz,ML}^2(\mathbf{w}_{ij}) = \frac{1}{|\Theta| |\Phi^{Acq}|} \sum_{\theta \in \Theta} \sum_{\phi \in \Phi^{Acq}} w_{ij,\theta,\phi}^2. \quad (16)$$

Note that, in contrast to the ML estimator of the combined variance and mean of a normal random variable, the ML estimator in eq. (16) is unbiased, that is $\mathbb{E}[\hat{\sigma}_{ij,AvAz,ML}^2(\mathbf{w}_{ij})] = \sigma_{ij,AvAz}^2$. Having fully described the stochastic model for the data, we are now in a position to give an expression for the likelihood model for \mathbf{y}^{AvAz} , which is:

$$p(\mathbf{y}^{AvAz} | \mathbf{x}) = \prod_{(i,j) \in \mathcal{L}} \left[\prod_{\theta \in \Theta} \prod_{\phi^{Acq} \in \Phi^{Acq}} \mathcal{N}(y_{ij,\theta,\phi^{Acq}}^{AvAz}; \bar{R}_{ij}^{PP}(\theta, \phi^{Acq}, \varphi_{ij}, z_{ij}), \hat{\sigma}_{ij,AvAz,ML}^2) \right], \quad (17)$$

where $\mathcal{N}(\cdot; \mu, \sigma^2)$ is the Gaussian probability density function (PDF) with mean μ and variance σ^2 .

3.2.2 FTF data

We further suppose that we have what Fang *et al.* (2014) describe as FTF data. We briefly describe the definition of the FTF data and how it is computed, which will result in a natural choice for our data \mathbf{y}^{FTF} and its likelihood model $p(\mathbf{y}^{FTF} | \mathbf{x})$.

Intuitively, the FTF is the transfer function from the seismic wavefield reflected off the top of the fractured layer to the wavefield propagating out of the fractured layer after reflecting off the bottom of this layer. In other words, it quantifies the redistribution of energy of the reflected and scattered seismic wavefield after passing through the fractured layer. A cartoon depicting this is shown in Fig. 3.

FTF is inherently a function of the propagation azimuth of the incident and reflected waves. At fracture spacings on the order of the seismic wavelength, the orientation of the fractures relative to the propagation azimuth has a significant effect on the amplitude of the scattered wavefield reflected off the bottom of the fractured layer. In particular, when fractures are parallel to the propagation azimuth, the fractures tend to act as waveguides, directing more of the scattered energy back to the surface in the direction away from the source. However, when the fractures are normal to the propagation azimuth, the scattered energy is less coherent as the fractures tend to scatter energy in both forward and backward directions. With this in mind, we expect FTF to be maximized at propagation azimuths parallel to the fractures.

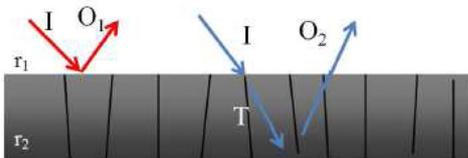


Figure 3. A cartoon depicting the meaning of fracture transfer function for layer r_2 . $I(\omega)$ is the incident wavefield, $T(\omega)$ is the transmitted wavefield into the fractured layer, and $O_1(\omega)$ and $O_2(\omega)$ are the waves reflected by layers above and below the fracture zone, respectively. Theoretically, the fracture transfer function at angular frequency ω is defined as $FTF(\omega) = \frac{O_2(\omega)}{O_1(\omega)}$. [Figure adapted from Fang *et al.* (2014).]

According to the methodology given by Fang *et al.* (2014) and Fang *et al.* (2012), the FTF at a particular spatial gridpoint (i, j) is estimated from surface seismic data by first determining (via ray tracing) all source–receiver pairs corresponding to gridpoint (i, j) . Then, for all source–receiver pairs within the same acquisition azimuth bin ϕ^{Acq} , normal moveout to zero offset is applied to the seismic traces which are then stacked. The result of this procedure gives a single, stacked seismic trace for each acquisition azimuth. The arrivals on the traces corresponding to reflections off the top and bottom of the fractured layer are then located in the stacked trace and windowed, giving windowed arrivals for each acquisition azimuth $o_1^{ij}(t, \phi^{Acq})$ and $o_2^{ij}(t, \phi^{Acq})$, respectively. The Fourier transforms $O_1^{ij}(\omega, \phi^{Acq})$ and $O_2^{ij}(\omega, \phi^{Acq})$ of the windowed arrivals are taken, and we compute the FTF at angular frequency ω and acquisition azimuth ϕ^{Acq} as

$$FTF^{ij}(\omega, \phi^{Acq}) = \frac{O_2^{ij}(\omega, \phi^{Acq})}{O_1^{ij}(\omega, \phi^{Acq})}. \quad (18)$$

This is reduced to a function of only acquisition azimuth by integrating out the angular frequency via a weighted integral. The idea is that frequencies at which there is greater variability in FTF with acquisition azimuth should be given more weight, hence a frequency weighting function $W^{ij}(\omega)$ is defined as the standard deviation of $FTF^{ij}(\omega, \cdot)$ with respect to acquisition azimuth, so that

$$\overline{FTF}^{ij}(\phi^{Acq}) = \int_{\omega} FTF^{ij}(\omega, \phi^{Acq}) W^{ij}(\omega) d\omega. \quad (19)$$

Due to the reasons mentioned above, we expect $\overline{FTF}^{ij}(\phi^{Acq})$ to be maximized at φ_{ij} if fractures are present in the medium. On the other hand, in the absence of fractures, we expect there will not be a unique maximizer for $\overline{FTF}^{ij}(\phi^{Acq})$. Hence, it is natural to define the FTF data used in our analysis as $\mathbf{y}_{ij}^{FTF} = (y_{ij,1}^{FTF}, y_{ij,2}^{FTF})$, where $y_{ij,1}^{FTF} \in \{0, 1\}$ is an indicator variable set to 0 when there is no unique maximizer (within a numerical threshold) for $\overline{FTF}^{ij}(\phi^{Acq})$, and set to 1 otherwise and where $y_{ij,2}^{FTF}$ is set to the acquisition azimuth that maximizes $\overline{FTF}^{ij}(\phi^{Acq})$:

$$y_{ij,2}^{FTF} \triangleq \arg \max_{\phi^{Acq} \in \Phi^{Acq}} \overline{FTF}^{ij}(\phi^{Acq}). \quad (20)$$

If there is no unique maximizing ϕ^{Acq} , we arbitrarily set $y_{ij,2}^{FTF}$ to any one maximizing value.

We define a stochastic forward model for \mathbf{y}_{ij}^{FTF} by first assuming that, given the fracture properties \mathbf{x}_{ij} at node (i, j) , \mathbf{y}_{ij}^{FTF} is conditionally independent of the fracture properties and FTF data at the remaining nodes.

We define ζ_{ij} as the probability that $y_{ij,1}^{FTF}$ correctly predicts whether or not there are fractures present at node (i, j) . Then, given the fracture properties, we model $y_{ij,1}^{FTF}$ as a Bernoulli random variable. Thus, given $z_{ij} = -13$ (i.e. zero excess fracture compliance) then $y_{ij,1}^{FTF} = 0$ with probability ζ_{ij} and $y_{ij,1}^{FTF} = 1$ with probability $1 - \zeta_{ij}$, and given $z_{ij} > -13$ (i.e. non-zero excess fracture compliance) then $y_{ij,1}^{FTF} = 1$ with probability ζ_{ij} and $y_{ij,1}^{FTF} = 0$ with probability $1 - \zeta_{ij}$. That is,

$$p(y_{ij,1}^{FTF} | \mathbf{x}_{ij}; \zeta_{ij}) = \begin{cases} \zeta_{ij} & \text{if } y_{ij,1}^{FTF} = \mathbb{1}_{\{z_{ij} > -13\}} \\ 1 - \zeta_{ij} & \text{if } y_{ij,1}^{FTF} = 1 - \mathbb{1}_{\{z_{ij} > -13\}} \end{cases}, \quad (21)$$

where $\mathbb{1}_{\{\cdot\}}$ is the indicator function defined as

$$\mathbb{1}_{\{A\}} \triangleq \begin{cases} 1 & \text{if } A \\ 0 & \text{otherwise} \end{cases}.$$

Now, given $y_{ij,1}^{FTF}$ and the fracture properties, if either $z_{ij} = -13$, so that there are no fractures present, or if $y_{ij,1}^{FTF} = 0$, so that a unique

preferential scattering direction was not identified, then any value for $y_{ij,2}^{\text{FTF}} \in [0^\circ, 180^\circ]$ is arbitrary, so we model $y_{ij,2}^{\text{FTF}}$ as uniform on the set $[0^\circ, 180^\circ]$. Otherwise, if both $z_{ij} > -13$ and $y_{ij,1}^{\text{FTF}} = 1$, then $y_{ij,2}^{\text{FTF}}$ should be near the true fracture orientation φ_{ij} . As with the AvAz data, we model this by introducing additive zero-mean independent Gaussian noise, so that conditioned on the event $\{z_{ij} \neq 0, y_{ij,1}^{\text{FTF}} = 1\}$, we have:

$$y_{ij,2}^{\text{FTF}} = \varphi_{ij} + v_{ij}, \quad (22)$$

where $v_{ij} \sim \mathcal{N}(0, \sigma_{ij,\text{FTF}}^2)$ and $\{v_{ij} : (i, j) \in L\}$ is a collection of mutually independent random variables. This gives the conditional distribution for $y_{ij,2}^{\text{FTF}}$ as

$$p(y_{ij,2}^{\text{FTF}} | y_{ij,1}^{\text{FTF}}, \mathbf{x}_{ij}; \sigma_{ij,\text{FTF}}^2) = \begin{cases} \frac{1}{180} \mathbb{1}\{y_{ij,2}^{\text{FTF}} \in [0, 180]\} & \text{if } z_{ij} = -13 \text{ or } y_{ij,1}^{\text{FTF}} = 0 \\ \mathcal{N}(y_{ij,2}^{\text{FTF}}; \varphi_{ij}, \sigma_{ij,\text{FTF}}^2) & \text{if } z_{ij} > -13 \text{ and } y_{ij,1}^{\text{FTF}} = 1 \end{cases}. \quad (23)$$

As with the AvAz data, we use synthetic data from finite difference simulations of the seismic wavefield (described in Section 4.1), to obtain a set of K independent samples $\mathbf{D}_{ij} = (\mathbf{x}_{ij}^{(k)}, \mathbf{y}_{ij}^{\text{FTF}(k)})_{k=1, \dots, K}$ of the fracture properties and FTF data at a single gridpoint, arising from the simulation of K different fracture models. The ML estimate for ζ_{ij} is simply the fraction of times $y_{ij,1}^{\text{FTF}(k)}$ correctly takes the value 0 (when $z_{ij}^{(k)} = -13$) or 1 (when $z_{ij}^{(k)} > -13$). For the synthetic data we have obtained, this fraction turns out to be 1. However, in order to preserve stochasticity in detecting the presence of fractures from FTF data, we instead estimate ζ_{ij} under a Bayesian approach by treating it as random variable with a prior distribution that is uniform over $[0, 1]$. Having observed K correct observations (and 0 incorrect observations) of $y_{ij,1}^{\text{FTF}}$, the posterior distribution for ζ_{ij} is a Beta distribution $\zeta_{ij} \sim \text{Beta}(K+1, 1)$. Integrating out ζ_{ij} in the likelihood model for $y_{ij,1}^{\text{FTF}}$, given the data, we have

$$p(y_{ij,1}^{\text{FTF}} | \mathbf{x}_{ij}, \mathbf{D}_{ij}) = \int_0^1 p(\zeta_{ij} | \mathbf{x}_{ij}, \mathbf{D}_{ij}) p(y_{ij,1}^{\text{FTF}} | \mathbf{x}_{ij}, \mathbf{D}_{ij}, \zeta_{ij}) d\zeta_{ij} \quad (24)$$

$$= \int_0^1 p(\zeta_{ij} | \mathbf{D}_{ij}) p(y_{ij,1}^{\text{FTF}} | \mathbf{x}_{ij}, \zeta_{ij}) d\zeta_{ij} \quad (25)$$

$$= \begin{cases} \frac{1}{B(K+1,1)} \int_0^1 \zeta_{ij}^{K+1} (1 - \zeta_{ij})^0 d\zeta_{ij} & \text{if } y_{ij,1}^{\text{FTF}} = \mathbb{1}\{z_{ij} > -13\} \\ \frac{1}{B(K+1,1)} \int_0^1 \zeta_{ij}^K (1 - \zeta_{ij})^1 d\zeta_{ij} & \text{if } y_{ij,1}^{\text{FTF}} = 1 - \mathbb{1}\{z_{ij} > -13\} \end{cases} \quad (26)$$

$$= \begin{cases} \frac{B(K+2,1)}{B(K+1,1)} & \text{if } y_{ij,1}^{\text{FTF}} = \mathbb{1}\{z_{ij} > -13\} \\ \frac{B(K+1,2)}{B(K+1,1)} & \text{if } y_{ij,1}^{\text{FTF}} = 1 - \mathbb{1}\{z_{ij} > -13\} \end{cases} \quad (27)$$

$$= \begin{cases} \frac{K+1}{K+2} & \text{if } y_{ij,1}^{\text{FTF}} = \mathbb{1}\{z_{ij} > -13\} \\ \frac{1}{K+2} & \text{if } y_{ij,1}^{\text{FTF}} = 1 - \mathbb{1}\{z_{ij} > -13\}, \end{cases} \quad (28)$$

where $B(\cdot, \cdot)$ is the beta function. We are left with K' i.i.d. samples where $z_{ij}^{(k)} > -13$ (and $y_{ij,1}^{\text{FTF}} = 1$), giving samples of the additive Gaussian noise

$$\{v_{ij}^{(k')}\}_{k'=1, \dots, K'} = \{\varphi_{ij} - y_{ij,2}^{\text{FTF}(k')} : z_{ij}^{(k')} > -13, y_{ij,1}^{\text{FTF}(k')} = 1\}$$

used to give the ML estimation of the variance

$$\hat{\sigma}_{ij,\text{FTF},ML}^2 = \hat{\sigma}_{ij,\text{FTF},ML}^2 \left(\{v_{ij}^{(k')}\}_{k'=1, \dots, K'} \right) = \frac{1}{K'} \sum_{k'=1}^{K'} (v_{ij}^{(k')})^2. \quad (29)$$

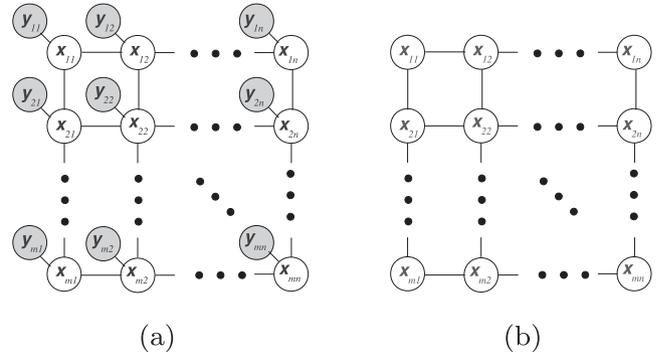


Figure 4. (a) Graphical model showing the Markovianity between the observations \mathbf{y} and the fracture parameters \mathbf{x} . (b) Graphical model for the posterior distribution after removing the observed nodes.

As before, the ML estimator in eq. (29) is unbiased. This gives the likelihood model $p(\mathbf{y}^{\text{FTF}} | \mathbf{x})$ as:

$$p(\mathbf{y}^{\text{FTF}} | \mathbf{x}) = \prod_{(i,j) \in L} p(y_{ij,2}^{\text{FTF}} | \mathbf{x}_{ij}, \mathbf{D}_{ij}) \times p(y_{ij,2}^{\text{FTF}} | y_{ij,1}^{\text{FTF}}, \mathbf{x}_{ij}; \hat{\sigma}_{ij,\text{FTF},ML}^2). \quad (30)$$

3.3 Inference algorithms

We return to the graphical model representation of the distribution, as this will play a key role in the inference algorithms used to obtain the posterior marginals and MAP estimate. Having defined the prior and likelihood models, the posterior distribution is given by eq. (1). We immediately notice that given the fracture properties of a particular grid node (i, j) , the observations \mathbf{y}_{ij} at that grid node are conditionally independent of the remaining fracture properties and observations, that is, with \mathbf{x}_{-ij} and \mathbf{y}_{-ij} defined as in eq. (3), $\mathbf{y}_{ij} \perp \{\mathbf{x}_{-ij}, \mathbf{y}_{-ij}\} | \mathbf{x}_{ij}$; this is depicted in Fig. 4(a). Having observed the data $\mathbf{y} = \mathbf{y}$, the data are no longer random and hence separate nodes for the data are not included in the graphical model for the posterior distribution.

Hence, we can write the posterior distribution in terms of the node and edge potentials of the graph, ψ_{ij} and $\psi_{ij,kl}$, respectively, where the node potentials capture the effect of the data and the edge potentials capture the prior distribution. These potentials are given by:

$$\psi_{ij}(\mathbf{x}_{ij}) = p(\mathbf{y}_{ij}^{\text{AvAz}} | \mathbf{x}_{ij}) p(\mathbf{y}_{ij}^{\text{FTF}} | \mathbf{x}_{ij}) \quad (31)$$

and

$$\psi_{ij,kl}(\mathbf{x}_{ij}, \mathbf{x}_{kl}) = \exp \{-\beta_{z_{ij,kl}}(z_{ij} - z_{kl})^2 - \beta_{\varphi_{ij,kl}}(\varphi_{ij} - \varphi_{kl})^2\}, \quad (32)$$

and the posterior distribution is then given as:

$$p(\mathbf{x} | \mathbf{y}) \propto \prod_{(i,j) \in \mathcal{V}} \psi_{ij}(\mathbf{x}_{ij}) \prod_{(ij,kl) \in \mathcal{E}} \psi_{ij,kl}(\mathbf{x}_{ij}, \mathbf{x}_{kl}). \quad (33)$$

Having fully described the posterior distribution in terms of its graphical model and node and edge potentials, we are able to apply belief propagation algorithms to perform approximate inference of the fracture properties \mathbf{x} .

3.3.1 Loopy belief propagation

Belief propagation (BP) is a technique for performing inference on graphical models which has recently enjoyed much popularity

for use amongst a wide-range of applications (Pearl 1982, 1988; Murphy *et al.* 1999). Originally formulated for tree graphs (i.e. graphs having no cycles), BP refers to message-passing algorithms for computing either marginal distributions (called the sum-product algorithm) or MAP configurations (called the max-product algorithm). In particular, for an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the BP algorithm computes messages [denoted by $m_{ij \rightarrow kl}(x_{kl})$] from each node $(i, j) \in \mathcal{V}$ to every node (k, l) which shares an edge with (i, j) [called a ‘neighbour’ of (i, j)]; these messages capture the beliefs node (i, j) carries about its neighbours. The messages are iteratively propagated from each node to its neighbours, hence the name ‘belief propagation.’

The sum-product variant of the BP algorithm (Koller & Friedman 2009) with node and edge potentials ψ_{ij} and $\psi_{ij,kl}$ is given by the update equations

$$m_{ij \rightarrow kl}^{(0)}(x_{kl}) \propto 1 \quad (34)$$

$$m_{ij \rightarrow kl}^{(t+1)}(x_{kl}) \propto \sum_{x_{ij}} \psi_{ij}(x_{ij}) \psi_{ij,kl}(x_{ij}, x_{kl}) \prod_{uv \in \text{Nb}(ij) \setminus \{kl\}} m_{uv \rightarrow ij}^{(t)}(x_{ij}) \quad (35)$$

$$\hat{p}_{ij}^{(t)}(x_{ij}) \propto \psi_{ij}(x_{ij}) \prod_{uv \in \text{Nb}(ij)} m_{uv \rightarrow ij}^{(t)}(x_{ij}) \quad (36)$$

$\forall (i, j) \in \mathcal{V}, [(i, j), (k, l)] \in \mathcal{E}$, where $\text{Nb}(ij)$ denotes the set of neighbours of node (i, j) in \mathcal{G} and $\hat{p}_{ij}^{(t)}(x_{ij})$ is the estimate of the marginal for node (i, j) at iteration t . One can verify that if the underlying graphical model \mathcal{G} is a tree, then the sum-product algorithm converges to the true marginal distributions in a number of iterations equal to the diameter of the tree (Koller & Friedman 2009). The max-product algorithm is similarly defined by replacing summations with maximizations

$$m_{ij \rightarrow kl}^{(0)}(x_{kl}) \propto 1 \quad (37)$$

$$m_{ij \rightarrow kl}^{(t+1)}(x_{kl}) \propto \max_{x_{ij}} \psi_{ij}(x_{ij}) \psi_{ij,kl}(x_{ij}, x_{kl}) \prod_{uv \in \text{Nb}(ij) \setminus \{kl\}} m_{uv \rightarrow ij}^{(t)}(x_{ij}) \quad (38)$$

$$\hat{p}_{ij}^{(t)}(x_{ij}) \propto \psi_{ij}(x_{ij}) \prod_{uv \in \text{Nb}(ij)} m_{uv \rightarrow ij}^{(t)}(x_{ij}) \quad (39)$$

$\forall (i, j) \in \mathcal{V}, [(i, j), (k, l)] \in \mathcal{E}$, where $\hat{p}_{ij}^{(t)}(x_{ij})$ is the estimate of the node max-marginal $\bar{p}_{ij}(x_{ij})$ for node (i, j) at iteration t . The node max-marginal [at node (i, j)] is defined to be the function of x_{ij} one would obtain by fixing the random variable at node (i, j) to the value x_{ij} and then maximizing the joint distribution $p(\mathbf{x})$ over all other random variables. That is:

$$\bar{p}_{ij}(x_{ij}) \triangleq \max_{\mathbf{x}_{-ij}} p(x_{ij}, \mathbf{x}_{-ij}). \quad (40)$$

It is important to note that the node max-marginals are *not* the marginal distributions, and in fact they are not even probability distributions. However, they can be used to readily obtain the MAP estimate of \mathbf{x} . In particular, if the node max-marginals $\bar{p}_{ij}(x_{ij})$ have unique maximizers x_{ij}^* , then the MAP estimate is simply the vector of these unique maximizing values for each node:

$$\hat{\mathbf{x}}_{\text{MAP}} = [x_{ij}^*]_{ij \in \mathcal{V}} = \left[\arg \max_{x_{ij}} \bar{p}_{ij}(x_{ij}) \right]_{ij \in \mathcal{V}}. \quad (41)$$

Thus, the estimated node max-marginals $\hat{p}_{ij}^{(t)}(x_{ij})$ obtained from the max-product algorithm can be used to approximate the MAP estimate. Again, if \mathcal{G} is a tree, then it can be shown that the max-product algorithm converges to the true node max-marginals and will hence produce the exact MAP estimate.

While BP was originally intended for tree graphs (and indeed converges to the correct result on trees), it can still be applied to graphs which are not trees, such as the grid graph for our posterior distribution. Applying BP to perform inference on a graph with loops is referred to as loopy belief propagation. While LBP is an approximate algorithm, as it does not, in general, converge to the correct answer, it has nonetheless been used extensively in various settings and found to often give very good approximations, particularly on graphs with a relatively sparse edge set (such as our 2-D grid graph) and when the node potentials are strong relative to the edge potentials (Murphy *et al.* 1999). With this in mind, we apply loopy belief propagation on the posterior distribution for the fracture parameters to approximate the MAP configuration and marginal distributions.

4 RESULTS

4.1 Synthetic data

We validate our methodology by performing inference on a synthetic data set. The synthetic data are obtained from a 3-D elastic finite-difference simulation of the seismic wavefield on reservoir models having topology as shown in Fig. 1, and is the same data referenced in the FTF study by Fang *et al.* (2014). Each of the models consists of five flat homogeneous layers, with fractures in the third layer; the remaining layers are isotropic. Following the methodology of Coates & Schoenberg (1995) to simulate discrete fractures, the finite-difference grid cells intersecting individual fractures are modelled as anisotropic, and the individual fractures are spaced uniformly within an isotropic background medium (note the finite-difference grid cells are distinct from and on a much smaller scale than the grid nodes on which the random variables are defined). The isotropic background parameters are given in Table 1. It is important to note that we are *not* modelling the entire fractured layer as anisotropic, but only the individual fractures; hence the validity of this model does not depend on the fracture spacing.

The fractured layer in the models contains a single set of discrete parallel fractures with strike 0° and individual normal and tangential fracture compliances of 10^{-9} m Pa $^{-1}$ (note that the fracture compliance is distinct from the excess fracture compliance, which is the ratio of fracture compliance to fracture spacing). The models differ from one another by fracture spacing, where synthetic data have been obtained from models having fracture spacings of 12, 20, 40, 60, 80 and 100 m, and where the fracture parameters in a particular model are constant over the entire layer. The synthetic seismic trace data set for each model is obtained from a 2-D array of surface

Table 1. Isotropic background parameters for the finite difference synthetic data.

Layer	Thickness (m)	α (m s $^{-1}$)	β (m s $^{-1}$)	ρ (g cm $^{-3}$)
1	200	3000	1765	2.20
2	200	3500	2060	2.25
3	200	4000	2353	2.30
4	200	3500	2060	2.25
5	200	4000	2353	2.30

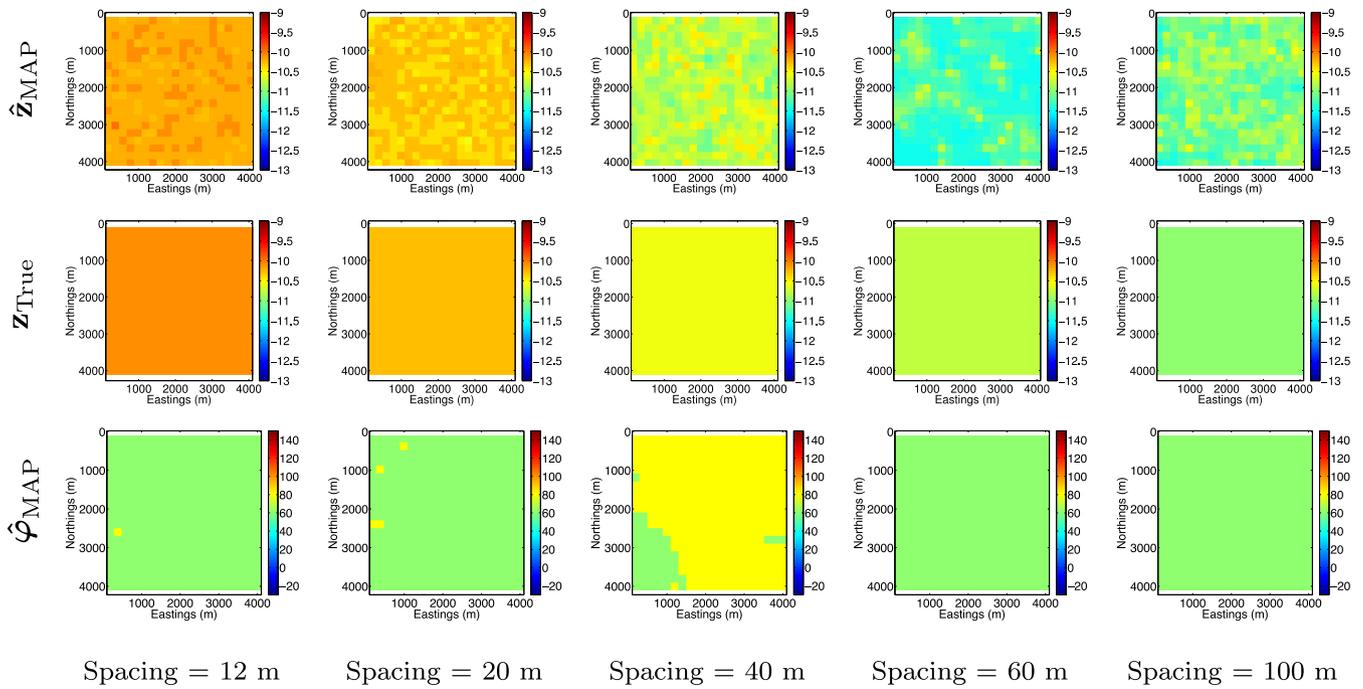


Figure 5. Approximate MAP estimates of the fracture properties computed for models of a single fracture set, with fracture compliance $10^{-9} \text{ m Pa}^{-1}$, fracture strike $\varphi_{ij} = 60^\circ$ and varying fracture spacing, and with smoothness parameter set to $\beta_c = 0.1$.

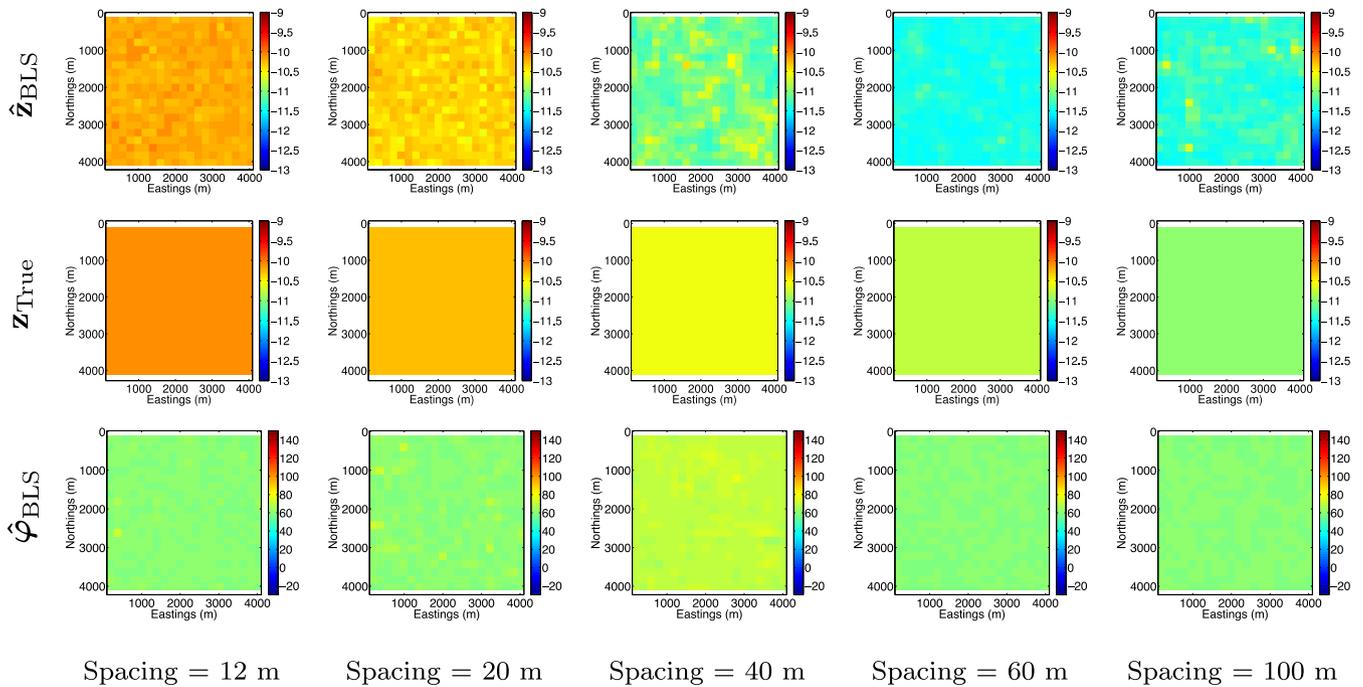


Figure 6. Approximate BLS estimates of the fracture properties computed for models of a single fracture set, with fracture compliance $10^{-9} \text{ m Pa}^{-1}$, fracture strike $\varphi_{ij} = 60^\circ$ and varying fracture spacing, and with smoothness parameter set to $\beta_c = 0.1$.

seismic receivers spaced 4 m apart responding to a single Ricker source wavelet having central frequency of 40 Hz. AvAz data are computed from the seismic trace data set by using ray tracing to compute the arrival time of the P - P arrival reflected from the top of Layer 3 for each receiver and taking the amplitude of this arrival in the seismic trace, where the source-receiver offsets and acquisition azimuths are known. Since the layers are flat and homogeneous, the

reflection from all points on the horizontal grid are equivalent on average. Hence, the data are treated as the generic AvAz data for a single grid node or common depth point (CDP). Similarly, the FTF data are computed according to the procedure in Section 3.2.2 and treated as the generic FTF data corresponding to a single node or CDP. Prior to processing the synthetic data, we perturb the raw seismic traces with zero-mean Gaussian noise, with a standard

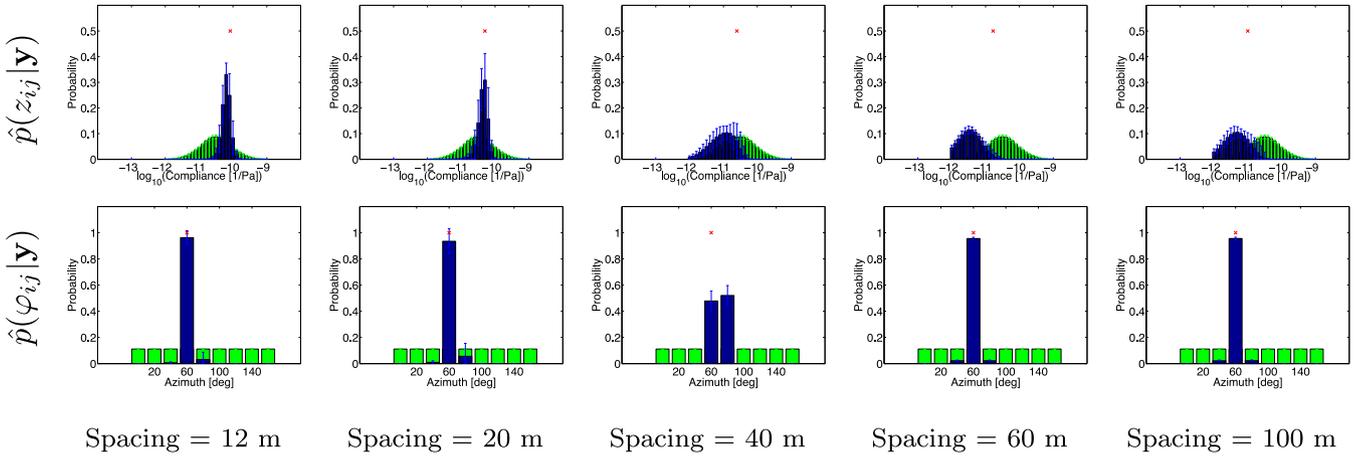


Figure 7. Approximate posterior marginal distributions (blue) of the fracture properties at a single node plotted along with the prior distributions (green). Results are given as mean ± 1 SD over all grid nodes. True value is plotted with a red 'x'. Computed for models of a single fracture set, with fracture compliance 10^{-9} m Pa $^{-1}$, fracture strike $\varphi_{ij} = 60^\circ$ and varying fracture spacing, and with smoothness parameter set to $\beta_c = 0.1$.

Table 2. The rms of residuals between estimates and ground truth (mean taken over all nodes) when estimating a single fracture set with fracture compliance 10^{-9} m Pa $^{-1}$, fracture strike $\varphi_{ij} = 60^\circ$ and varying fracture spacing, and with smoothness parameter set to $\beta_c = 0.1$. For comparison, note that azimuth φ is discretized into 20° bins and log compliance z has been discretized into bins of size $0.1 \log_{10}\text{Pa}^{-1}$.

Fracture spacing	$\epsilon_{\text{rms}, \hat{z}_{\text{BLS}}}$	$\epsilon_{\text{rms}, \hat{\varphi}_{\text{BLS}}}$	$\epsilon_{\text{rms}, \hat{z}_{\text{MAP}}}$	$\epsilon_{\text{rms}, \hat{\varphi}_{\text{MAP}}}$
12 m	$0.132 \log_{10}\text{Pa}^{-1}$	1.27°	$0.123 \log_{10}\text{Pa}^{-1}$	1.00°
20 m	$0.102 \log_{10}\text{Pa}^{-1}$	2.27°	$0.083 \log_{10}\text{Pa}^{-1}$	2.00°
40 m	$0.488 \log_{10}\text{Pa}^{-1}$	10.53°	$0.235 \log_{10}\text{Pa}^{-1}$	18.63°
60 m	$0.670 \log_{10}\text{Pa}^{-1}$	0.05°	$0.530 \log_{10}\text{Pa}^{-1}$	0°
80 m	$0.555 \log_{10}\text{Pa}^{-1}$	7.98°	$0.756 \log_{10}\text{Pa}^{-1}$	0°
100 m	$0.358 \log_{10}\text{Pa}^{-1}$	0.12°	$0.187 \log_{10}\text{Pa}^{-1}$	0°

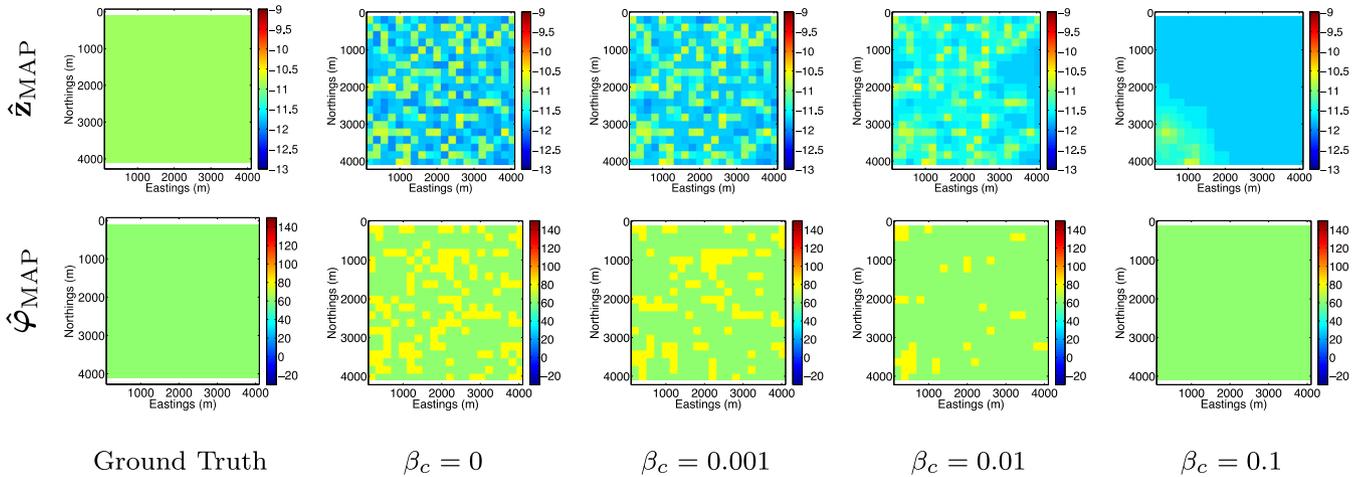


Figure 8. Approximate MAP estimates of the fracture properties computed on a model containing a single set of fractures with fracture compliance 10^{-9} m Pa $^{-1}$, fracture spacing 80 m, and fracture orientation 60° . Ground truth is plotted along with the estimates using various values for the smoothness parameter β_c .

deviation of 5 per cent of the peak amplitude of the data, where a different realization of the noise is used for each CDP gather.

To obtain measurements for the entire grid \mathbf{L} , we predetermine the fracture parameters \mathbf{x} over \mathbf{L} . For each node $(i, j) \in \mathbf{L}$, we are free to choose any fracture strike in the full range of azimuths $[0^\circ, 180^\circ)$, as we can simply rotate the synthetic data from 0° to any desired azimuth φ_{ij} . For the log excess compliance z_{ij} , we are able

to use any value obtained from the models corresponding to excess compliances that can be achieved using fracture compliance of 10^{-9} m Pa $^{-1}$ divided by any of the spacings for which synthetic data have been obtained. Having set the desired fracture properties across the grid, we map the generic synthetic data to specific values across the grid \mathbf{L} . Processing the noisy data according to the procedure described in Section 3.2 results in our vector of noisy measurements

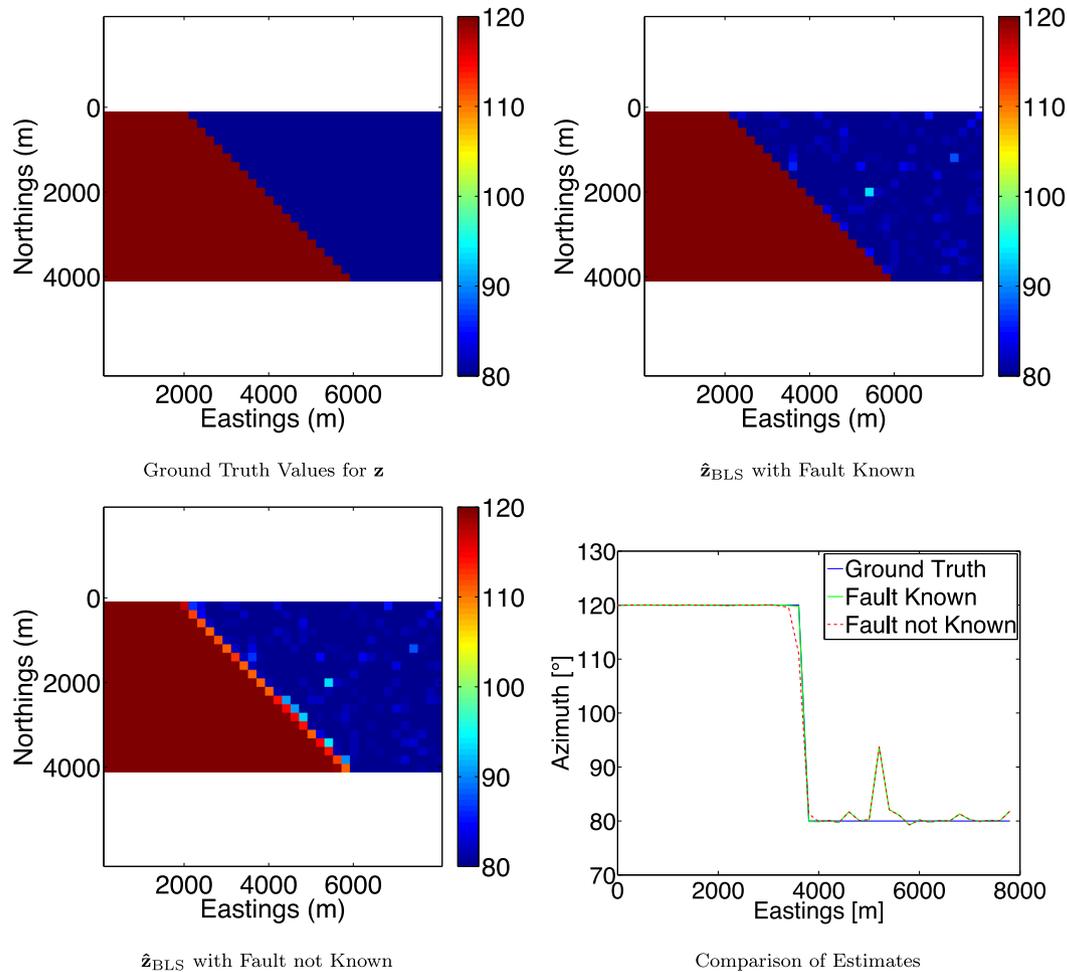


Figure 9. Effect of *a priori* knowledge of the fault on approximate BLS estimates of the fracture azimuth of a model containing two fracture sets. The fractures on the left are at azimuth 120° , spacing 100 m, and fracture compliance 10^{-9} m Pa $^{-1}$. The fractures on the right are at azimuth 80° , spacing 12 m, and fracture compliance 10^{-9} m Pa $^{-1}$. Ground truth is plotted along with the estimates with the smoothness parameter $\beta_c = 0.1$. The fourth pane shows a comparison of the estimates at the horizontal slice north = 2000 m.

\mathbf{y} corresponding to the known fracture properties \mathbf{x} over the entire grid. A grid spacing of $\ell = 200$ m is used for \mathbf{L} .

4.2 Results of inference procedure

We perform the inference on synthetic data arising from two scenarios. The first scenario is given by a 20-by-20 node grid of a single fracture set (so that the fracture properties are constant along the grid), for each of the available fracture spacing models. The second scenario is given by a 20-by-40 node grid of two fracture sets with distinct excess compliances and orientations, separated spatially by a linear discontinuity (such as that which may arise from a vertical, planar fault). The effect of the smoothness parameter β_c on the inference result is investigated by performing the inference for different choices of β_c .

LBP is applied in each case to obtain the approximate MAP configuration and posterior marginal distributions of the fracture properties, the latter of which are used to compute the approximate BLS estimate of the fracture properties. LBP converged for all models in less than 200 iterations, when the smoothness parameter β_c was taken to be less than or equal to 0.1. Choices of the smoothness parameter greater than 0.1 resulted in LBP not converging for some realizations of the noisy data.

We must take care to interpret the results correctly, as we have taken $z_{ij} = -13$ to mean that no fractures are present at node (i, j) , which would render φ_{ij} meaningless and arbitrary. Thus, we compute the posterior marginals for φ_{ij} conditioned on the event $\{z_{ij} > -13\}$, and likewise compute the BLS estimates for these random variables conditioned on the same event. The results of the inference procedure on the single fracture system are plotted in Figs 5–7. The resulting residuals between the estimates and true values are given in Table 2 in terms of the rms residuals over all nodes.

We observe that the inference procedure performs very well at fracture spacings smaller than 40 m. This is to be expected as the forward model for the AvAz data relies on the assumption that fractures cause the medium to behave as an equivalent anisotropic medium, but this assumption breaks down as the fracture spacing becomes comparable to the dominant seismic wavelength (which is 100 m in the fractured layer and 87.5 m in the layer above the fractures). Furthermore, while the scattering assumptions underlying the forward model for the FTF data are valid for fracture spacings on the order of the seismic wavelength, the FTF data in this study only contributes to fracture detection and strike estimation; the actual excess compliance value (given fractures are present) has no bearing on our model for the FTF data. In particular, we

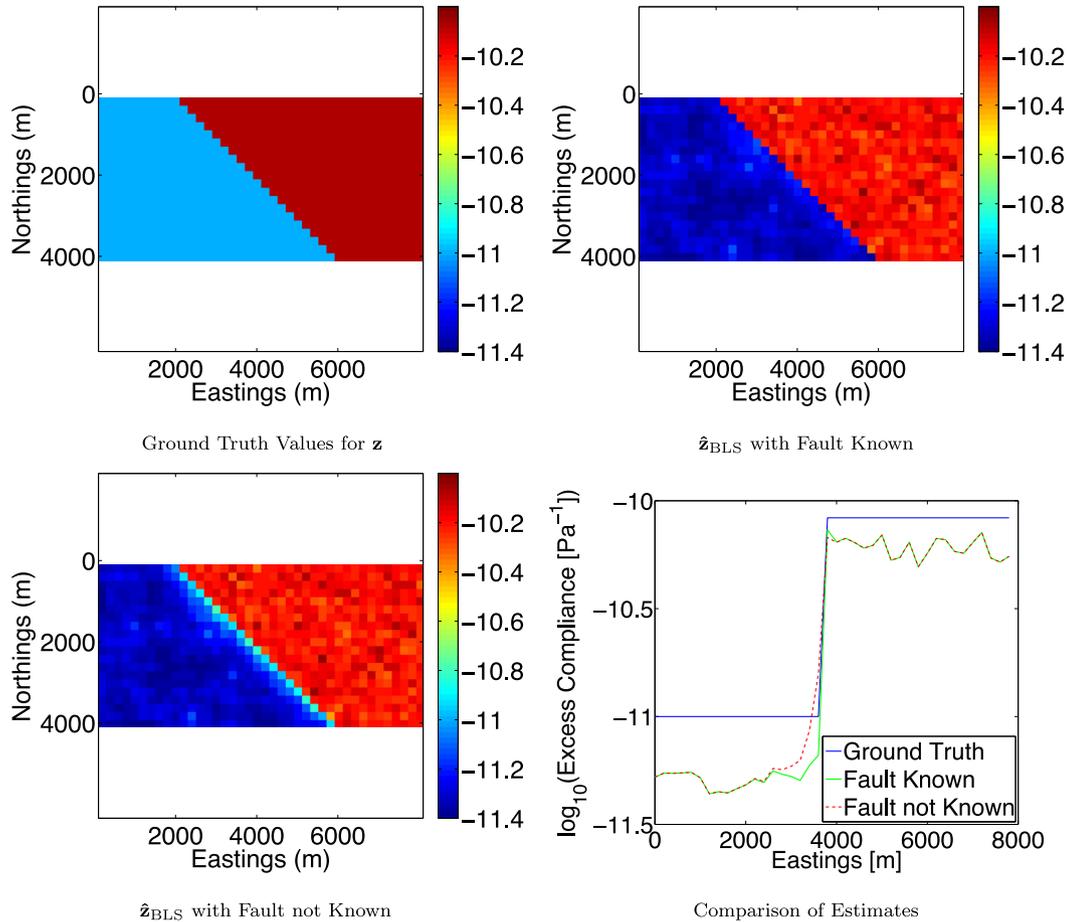


Figure 10. Effect of *a priori* knowledge of the fault on approximate BLS estimates of the fracture excess compliance of a model containing two fracture sets. The fractures on the left are at azimuth 120° , spacing 100 m, and fracture compliance 10^{-9} m Pa $^{-1}$. The fractures on the right are at azimuth 80° , spacing 12 m, and fracture compliance 10^{-9} m Pa $^{-1}$. Ground truth is plotted along with the estimates with the smoothness parameter $\beta_c = 0.1$. The fourth pane shows a comparison of the estimates at the horizontal slice north = 2000 m.

notice from Table 2 that the RMS residuals for estimating log excess compliance grows to multiple bin sizes at fracture spacings of 40 m and larger. The marginal distributions in Fig. 7 convey well what happens at spacings of 40 m and larger. We see that the least best fit to fracture orientation is obtained at 40 m spacing; this is likely because at this mid-range fracture spacing, we see a weaker response from both the AvAz data and the FTF data. However, at larger fracture spacings, the marginal distributions for fracture orientation remain concentrated around the true orientation of 60° , as expected due to both the very simple model for FTF in terms of fracture orientation as well as the assumptions underlying the FTF model remaining strong at larger spacings. We observe that excess compliance tends to be underestimated at fracture spacing values of 40 m and above. This observation is consistent with our intuition for AvAz data; at larger spacings, the AvAz response becomes weaker, and a better fit to the data is found with smaller excess compliances than those resulting from the true fracture compliance and spacing.

In order to investigate the effect of the smoothness parameter β_c on the inference we apply our procedure over a range of choices for β_c on the 80 m spacing model. The effect of the smoothness parameter is most easily seen in the MAP estimates, which are plotted in Fig. 8. Observing the changes in the estimate with increasing smoothness parameter β_c , we see that a higher value of β_c has the effect of denoising the estimates. When β_c is 0, this is identi-

cal to performing the inference on a fully disconnected graph, as the edge potentials will all be identically equal to 1. As such, the estimate at each node will fit only the noisy data corresponding to itself. Increasing β_c strengthens the links between adjacent nodes, and can cause an incorrect fit to noisy data to be less probable; this is particularly true for fracture azimuth, which is estimated correctly at 80 m fracture spacing. Increasing β_c still smooths the estimates for excess compliance, however towards the underestimated value found at 80 m (again due to the weaker AvAz response at high fracture spacings). At $\beta_c = 0.1$, we see a considerable improvement in the estimates for fracture azimuth.

We now turn to the second scenario arising from two sets of fractures with different azimuths and fracture spacings separated by a discontinuity on the grid such as that which may arise from a geological fault. We investigate the effect of *a priori* knowledge of the fault by performing the inference both with and without knowledge of the fault location encoded in the prior with the smoothness parameter set to $\beta_{ij,kl} = 0$ at the fault and $\beta_{ij,kl} = \beta_c = 0.1$ elsewhere. The results are plotted in Figs 9 and 10 for fracture azimuth and log excess compliance, respectively. As evidenced by the figures, when the fault is unknown *a priori*, the estimates for fracture properties are smoothed across the fault. This is particularly undesirable for the estimation of fracture azimuth, where we would otherwise be able to obtain good estimates in both regions. Specifying the location of the fault *a priori* sets the smoothness parameter

at the corresponding edges to 0, and hence we no longer observe this behaviour.

5 CONCLUSIONS AND FUTURE WORK

A methodology for estimation of fracture properties from AvAz and FTF data under a Bayesian inference framework has been presented. The inference is performed by running loopy belief propagation on the 2-D Markov random field of fracture variables. LBP converged relatively quickly on the synthetic data, in under 200 iterations for all models, when using a smoothness parameter less than or equal to 0.1. We have demonstrated that the approximate inference results perform well for both fracture azimuth and excess compliance at low spacings of 12 and 20 m, and continue to give good estimates for fracture azimuth up to 100 m spacing. This is significant, as we are able to estimate the fracture properties in a rigorous manner at a greater range of spacings than would otherwise be attainable.

We further showed that our use of the spatial smoothness prior has the effect of denoising estimates that would otherwise be incorrect. We also demonstrated the capability of this framework to handle prior information about geological features, such as the discontinuity shown in the previous figures. While we presented a very simple case of this, it is not difficult to extend this to more complicated scenarios. Having validated our procedure on synthetic data, the next natural step will be to obtain field data and apply the inference procedure to estimate the desired fracture properties.

One future direction to improve the inference is to relate fracture spacing and compliance to the FTF data. While the data we chose depended only on the presence and orientation of fractures, Fang *et al.* (2014) showed both theoretically and in laboratory experiments, that FTF also contains information about fracture spacing. However, even when using synthetic data, the precise physical relationship between FTF and fracture spacing has been difficult to determine, but a geophysical basis remains for exploring this avenue further. If a reliable forward model can be determined to relate FTF to fracture spacing, then we will be able to move beyond estimating excess fracture compliance to estimation of individual fracture compliances and fracture spacing. A related future direction is to incorporate additional features of the seismic data in the inference procedure. In particular, Zheng *et al.* (2013) describe a theory for using 3-D beam interference to determine fracture properties of a reservoir from reflected seismic *P*-wave data.

ACKNOWLEDGEMENTS

This work was supported in part by Total S.A. The authors would like to thank John Fisher, of MIT CSAIL, for an insightful discussion that provided motivation for the general idea. We also thank Jonathan Kane and an anonymous reviewer for providing helpful guidance in improving the quality of this work.

REFERENCES

Ali, A. & Jakobsen, M., 2011. Seismic characterization of reservoirs with multiple fracture sets using velocity and attenuation anisotropy data, *J. appl. Geophys.*, **75**(3), 590–602.

Clifford, P., 1990. Markov random fields in statistics, in *Disorder in Physical Systems: A Volume in Honour of John M. Hammersley*, pp. 19–32, eds Grimmett, G. & Welsh, D., Oxford Univ. Press.

Coates, R. & Schoenberg, M., 1995. Finitedifference modeling of faults and fractures, *Geophysics*, **60**(5), 1514–1526.

Dubos-Sallée, N. & Rasolofosaon, P., 2008. Evaluation of fracture parameters and fluid content from seismic and well data, in *Proceedings of the SEG Technical Program, Expanded Abstracts*, pp. 1511–1515.

Eidsvik, J., Avseth, P., Omre, H., Mukerji, T. & Mavko, G., 2004. Stochastic reservoir characterization using prestack seismic data, *Geophysics*, **69**(4), 978–993.

Fang, X., Fehler, M., Zhu, Z., Zheng, Y. & Burns, D., 2012. Reservoir fracture characterizations from seismic scattered waves, in *Proceedings of the SEG Technical Program, Expanded Abstracts*, Vol. **31**, pp. 1–6.

Fang, X., Fehler, M.C., Zhu, Z., Zheng, Y. & Burns, D.R., 2014. Reservoir fracture characterization from seismic scattered waves, *Geophys. J. Int.*, **196**(1), 481–492.

Gaiser, J. & Van Dok, R., 2001. Green river basin 3-D/3-C case study for fracture characterization: analysis of PS wave birefringence, in *Proceedings of the SEG Technical Program, Expanded Abstracts*, Vol. **20**, pp. 764–767.

Hammersley, J. & Clifford, P., 1971. Markov fields on finite graphs and lattices (unpublished manuscript).

Koller, D. & Friedman, N., 2009. *Probabilistic Graphical Models: Principles and Techniques*, MIT Press.

Liu, E. & Martinez, A., 2013. *Seismic Fracture Characterization: Concepts and Practical Applications*, EAGE Publications BV.

Lynn, H., Narhari, S.R., Al-Ashwak, S., Kidambi, V., Al-Qadeeri, B. & Al-Khaled, O., 2009. PP azimuthal-amplitudes and -acoustic impedance for fractured carbonate reservoir characterization, in *Proceedings of the SEG Technical Program, Expanded Abstracts*, Vol. **29**, pp. 258–262.

Murphy, K., Weiss, Y. & Jordan, M., 1999. Loopy belief propagation for approximate inference: an empirical study, in *Proceedings of Uncertainty in Artificial Intelligence*, pp. 467–475.

Pearl, J., 1982. Reverend Bayes on inference engines: a distributed hierarchical approach, in *Proceedings of the Second National Conference on Artificial Intelligence*, AAAI, Pittsburgh, PA, pp. 133–136, AAAI Press.

Pearl, J., 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers Inc.

Pšenčík, I. & Martins, J., 2001. Properties of weak contrast pp reflection/transmission coefficients for weakly anisotropic elastic media, *Stud. Geophys. Geod.*, **45**(2), 176–199.

Rüger, A., 1998. Variation of P-wave reflectivity with offset and azimuth in anisotropic media, *Geophysics*, **63**(3), 935–947.

Rüger, A., 2002. *Reflection Coefficients and Azimuthal AVO Analysis in Anisotropic Media*, Society of Exploration Geophysicists.

Sayers, C., 2009. Seismic characterization of reservoirs containing multiple fracture sets, *Geophys. Prospect.*, **57**(2), 187–192.

Sayers, C. & Rickett, J., 1997. Azimuthal variation in AVO response for fractured gas sands, *Geophys. Prospect.*, **45**, 165–182.

Sayers, C.M., Taleghani, A.D. & Adachi, J., 2009. The effect of mineralization on the ratio of normal to tangential compliance of fractures, *Geophys. Prospect.*, **57**(3), 439–446.

Schoenberg, M. & Douma, J., 1988. Elastic wave propagation in media with parallel fractures and aligned cracks, *Geophys. Prospect.*, **36**, 571–590.

Schoenberg, M. & Sayers, C., 1995. Seismic anisotropy of fractured rock, *Geophysics*, **60**(1), 204–211.

Sil, S. & Srinivasan, S., 2009. Stochastic simulation of fracture strikes using seismic anisotropy induced velocity anomalies, *Explor. Geophys.*, **40**, 257–264.

Simpson, D., Lindgren, F. & Rue, H., 2012. In order to make spatial statistics computationally feasible, we need to forget about the covariance function, *Environmetrics*, **23**(1), 65–74.

Tarantola, A., 2004. *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics.

Thomsen, L., 1986. Weak elastic anisotropy, *Geophysics*, **51**(10), 1954–1966.

Tsvankin, I., 2001. *Seismic Signatures and Analysis of Reflection Data in Anisotropic Media*, Vol. 29 of Series Handbook of Geophysical Exploration: Seismic Exploration, Elsevier.

Ursin, B. & Haugen, G.U., 1996. Weak-contrast approximation of the elastic scattering matrix in anisotropic media, *Pure appl. Geophys.*, **148**(3-4), 685–714.

Verdon, J.P. & Wüstefeld, A., 2013. Measurement of the normal/tangential fracture compliance ratio (Z_N/Z_T) during hydraulic fracture stimulation using s-wave splitting data, *Geophys. Prospect.*, **61**, 461–475.

Willis, M., Burns, D., Rao, R., Minsley, B., Toksoz, M. & Vetri, L., 2006. Spatial orientation and distribution of reservoir fractures from scattered seismic energy, *Geophysics*, **71**(5), O43–O51.

Zheng, Y., Fang, X., Fehler, M.C. & Burns, D.R., 2013. Seismic characterization of fractured reservoirs by focusing Gaussian beams, *Geophysics*, **78**(4), A23–A28.

APPENDIX: DERIVATION OF THE THOMSEN ANISOTROPY PARAMETERS FROM EXCESS FRACTURE COMPLIANCE

We use the linear slip model of Schoenberg & Sayers (1995) to express the Thomsen anisotropy parameters of the fractured medium in terms of the excess fracture compliance of the medium. The anisotropy parameters can be expressed in terms of the stiffness tensor of the medium \mathbf{C} as (Rüger 1998):

$$\delta^{(V)} = \frac{(C_{13} + C_{55})^2 - (C_{33} - C_{55})^2}{2C_{33}(C_{33} - C_{55})}, \quad (\text{A1})$$

$$\epsilon^{(V)} = \frac{C_{11} - C_{33}}{2C_{33}}, \quad (\text{A2})$$

$$\gamma^{(V)} = \frac{C_{66} - C_{44}}{2C_{44}}. \quad (\text{A3})$$

We can relate the fracture properties of the medium to the stiffness tensor by computing the excess compliance tensor of the fractures \mathbf{S}_{frac} (which is the contribution of the fractures to the overall medium compliance tensor). Schoenberg & Sayers (1995) show that, under the simplifying assumption that the behaviour of the fracture system is invariant with respect to rotation about the axis normal to the fractures, the excess compliance tensor of the fractures is given by

$$\mathbf{S}_{\text{frac}} = \begin{bmatrix} Z_N & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & Z_T & 0 \\ 0 & 0 & 0 & 0 & 0 & Z_T \end{bmatrix}, \quad (\text{A4})$$

where Z_N and Z_T are the excess normal and tangential compliances of the fracture system, respectively. In our analysis, we assumed that the excess normal and tangential compliances of the fracture system are equal. Thus at grid node (i, j) , we have $Z_N = Z_T = 10^{z_{ij}}$. Keeping with our convention of treating zero excess compliance as 10^{-13}Pa^{-1} , if $z_{ij} = -13$ then we set $Z_N = Z_T = 0$ [corresponding to the case of no fractures at node (i, j)]. Schoenberg and Sayers further show that the overall medium compliance tensor \mathbf{S}_{tot} can be expressed as the sum of the fracture excess compliance tensor \mathbf{S}_{frac} and the background compliance tensor \mathbf{S}_{back} , so that

$$\mathbf{S}_{\text{tot}} = \mathbf{S}_{\text{back}} + \mathbf{S}_{\text{frac}}. \quad (\text{A5})$$

The background compliance tensor is the inverse of background stiffness tensor \mathbf{C}_{back} , which for an isotropic, homogeneous medium is given by (Tsvankin 2001):

$$\mathbf{S}_{\text{back}}^{-1} = \mathbf{C}_{\text{back}} = \begin{bmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix}, \quad (\text{A6})$$

where $\mu = \rho\beta^2$ and $\lambda = \rho\alpha^2 - 2\mu$ are Lamé's parameters. The overall stiffness tensor of the medium \mathbf{C} is then found as the inverse of the overall compliance tensor (Schoenberg & Sayers 1995):

$$\mathbf{C} = \mathbf{S}_{\text{tot}}^{-1} = \begin{bmatrix} M_b(1-d_N) & \lambda(1-d_N) & \lambda(1-d_N) & 0 & 0 & 0 \\ \lambda(1-d_N) & M_b(1-r_b^2 d_N) & \lambda(1-r_b d_N) & 0 & 0 & 0 \\ \lambda(1-d_N) & \lambda(1-r_b d_N) & M_b(1-r_b^2 d_N) & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu(1-d_T) & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu(1-d_T) \end{bmatrix}, \quad (\text{A7})$$

where

$$M_b = \lambda + 2\mu, \quad r_b = \frac{\lambda}{M_b}, \quad 0 \leq d_T = \frac{Z_T \mu}{1 + Z_T \mu} < 1, \quad 0 \leq d_N = \frac{Z_N M_b}{1 + Z_N M_b} < 1.$$

Combining all of the above gives the anisotropy parameters of the fractured medium at node (i, j) , which we denote by $\delta_{z_{ij}}^{(V)}$, $\gamma_{z_{ij}}^{(V)}$, $\epsilon_{z_{ij}}^{(V)}$, in terms of the excess fracture compliance. Using (A7) in (A1), (A2), (A3) and (11) gives the forward model for the P - P reflection coefficient as a function of the fracture parameters at node (i, j) .