Geophysical Journal International Advance Access published May 18, 2016

A Bayesian method for microseismic source inversion

D.J. Pugh[1,2,3*], R.S. White [1] and P.A.F. Christie[2]

[1]Bullard Laboratories, Department of Earth Sciences, University of Cambridge

[2]Schlumberger Gould Research, High Cross, Cambridge

[3]Now at McLaren Applied Technologies, McLaren Technology Centre, Chertsey Road, Woking

*Email: david.j.pugh@cantab.net

In original form 24th September, 2015 , in revised form 18th May, 2016.

SUMMARY

Earthquake source inversion is highly dependent on location determination and velocity models. Uncertainties in both the model parameters and the observations need to be rigorously incorporated into an inversion approach. Here, we show a probabilistic Bayesian method that allows formal inclusion of the uncertainties in the moment tensor inversion. This method allows the combination of different sets of far-field observations, such as P-wave and S-wave polarities

and amplitude ratios, into one inversion. Additional observations can be included by deriving a suitable likelihood function from the uncertainties. This inversion produces samples from the source posterior probability distribution, including a best-fitting solution for the source mechanism and associated probability. The inversion can be constrained to the double-couple space or allowed to explore the gamut of moment tensor solutions, allowing volumetric and other non-double-couple components. The posterior probability of the double-couple and full moment tensor source models can be evaluated from the Bayesian evidence, using samples from the likelihood distributions for the two source models, producing an estimate of whether or not a source is double-couple. Such an approach is ideally suited to microseismic studies where there are many sources of uncertainty and it is often difficult to produce reliability estimates of the source mechanism, although this can be true of many other cases. Using full-waveform synthetic seismograms, we also show the effects of noise, location, network distribution and velocity model uncertainty on the source probability density function. The noise has the largest effect on the results, especially as it can affect other parts of the event processing. This uncertainty can lead to erroneous non-double-couple source probability distributions, even when no other uncertainties exist. Although including amplitude ratios can improve the constraint on the source probability distribution, the measurements are often systematically affected by noise, leading to deviation from their noise-free true values and consequently adversely affecting the source probability distribution, especially for the full moment tensor model. As an example of the application

of this method, four events from the Krafla volcano in Iceland are inverted, which show clear differentiation between non-double-couple and double-couple sources, reflected in the posterior probability distributions for the source models.

Keywords: *Earthquake source observations, Probability distributions, Volcano seismology, Numerical solutions*

1 INTRODUCTION

Earthquakes occur from the creation and propagation of a rupture. Rapid motion produces seismic waves, which are usually observed in the far-field. Source inversion can use such observations to describe the kinematic behaviour of the source by the moment tensor, while accounting for measurement uncertainties. There are other approaches to source inversion, including the use of GPS, InSAR data and near-field strong motion observations (e.g. Lundgren & Salichon 2000; Delouis et al. 2002; Page et al. 2009; Konca et al. 2010; Yokota et al. 2012; O'Toole 2013; Weston et al. 2014; Houlié et al. 2014), although the signals from micro-seismic events may be too small and frequent to be detected, although not always (e.g. O'Toole 2013). Often the seismogram is often the only observations of an earthquake.

For regional and global earthquakes, it is often feasible to pre-calculate databases of Green functions to perform full waveform inversions efficiently (Heimann 2011; Duputel et al. 2012b), although this can also be done on the fly for less complex velocity structures (e.g. Dziewonski et al. 1981; Ekström et al. 2012) and even for microseismic cases (e.g. O'Toole 2013). However microseismic earthquakes are small-magnitude earthquakes often detected on a small local network of receivers. For these cases, it is difficult to compute the Green functions database because the velocity structure is rarely well constrained and the event locations are often distributed throughout the network. Furthermore, velocity

variations in the region may have a large effect on the ray-paths due to the close proximity of the receivers to the sources. Therefore, observations such as P- and S-wave polarities and amplitude ratios are more robust in these regimes, so are commonly used to constrain the source inversion (Reasenberg & Oppenheimer 1985; Hardebeck & Shearer 2002, 2003; Snoke 2003). Such approaches still depend on knowing the velocity model to calculate the azimuths and take-off angles of the rays from the source to the instrument. Many of the current inversions provide a result and an estimate of some misfit or quality parameter.

There are several Bayesian approaches to source inversion. Often these are based on full waveform inversion approaches (e.g. Kennet et al. 2000; Wéber 2006), and can extend the approach to near-field observations and finite-fault models (e.g. Zollo & Bernard 2007; Minson et al. 2013), or other data sources (e.g. O'Toole 2013; Kaeufl et al. 2013). This paper is focussed on the far-field observations such as polarities and amplitude ratios, rather than the full waveform approach, because these measurements are more robust when the velocity model is uncertain. Walsh et al. (2009) explored a method for determining double-couple focal mechanism parameters using P-wave polarities. The approach presented in this paper is consistent with that framework and extends it to different independent observation types and to full moment tensor inversion. This inversion process also enables a comparison of the best fitting double-couple and full moment tensor solutions to be made. This allows a more detailed examination of any non-double-couple components.

The source function of an earthquake is a function of both time and position, although, in the case of microseismic events, the time dependence is modelled as a step function, so it is usually assumed that there is no spatial dependence of the time component. Therefore, the source function can be split into the time-dependent source-time function, $S(t)$, and the spatially dependent moment tensor, $\mathbf{M}$. The moment tensor describes the nine force couples required to define a point source (Fig. 1). While angular momentum is conserved, the moment tensor must be symmetrical and, therefore, there

are six independent components. It may be possible to determine specific cases where angular momentum is not conserved, but these will not be discussed here.

In most cases, especially for teleseismic events, the moment tensor appears to be double-couple or close to double-couple, signifying slip along a fault plane. Some non-double-couple mechanisms have been observed, such as those associated with nuclear explosions (Müller 1973; Ford et al. 2008), and potentially non-double-couple mechanisms in events in volcanic and geothermal regions and other areas associated with induced seismicity, such as hydraulic fracturing (Vasco 1990; Foulger et al. 2004; Templeton & Dreger 2006; Vavryčuk et al. 2008). These non-double-couple mechanisms could arise from processes such as conduit collapse or fracture opening, perhaps associated with fluid movement. However, apparent (but incorrect) non-double-couple characteristics can also be caused by uncertainties in the inversion such as noise in the data or velocity model uncertainties, finite fault effects (Kuge & Lay 1994), as well as the improved fit due to the extra two parameters in the full moment tensor compared to the double-couple source (Panza & Sarao 2000).

The six independent moment tensor components can be written as a six-vector, in a form similar to the Voigt form (Voigt 1910) and that of Chapman & Leaney (2011). Using this notation, it is possible to write the far-field seismic amplitude equation as shown in Eq. 1, where $\mathbf{a}$ is the matrix of station propagation coefficients, $\mathbf{a}_{STA_i}$ (a derivation of the isotropic propagation coefficients can be found in Appendix A), $\mathbf{u}$ is the vector of amplitudes recorded at the receiver stations STA_i and $\tilde{\mathbf{M}}$ is the six-vector form of the

moment tensor **M**.

$$\mathbf{u} = \begin{pmatrix} u_{STA_1} \\ u_{STA_2} \\ u_{STA_3} \\ u_{STA_4} \\ u_{STA_5} \\ \vdots \end{pmatrix} = \begin{pmatrix} \mathbf{a}_{STA_1} \\ \mathbf{a}_{STA_2} \\ \mathbf{a}_{STA_3} \\ \mathbf{a}_{STA_4} \\ \mathbf{a}_{STA_5} \\ \vdots \end{pmatrix} \begin{pmatrix} M_{11} \\ M_{22} \\ M_{33} \\ \sqrt{2}M_{23} \\ \sqrt{2}M_{13} \\ \sqrt{2}M_{12} \end{pmatrix} = \mathbf{a} \cdot \tilde{\mathbf{M}}. \tag{1}$$

The six-vector components are scaled in Eq. 1 so that the six-vector can be normalised to give a normalised moment tensor (Eq. 2), defined following Chapman & Leaney (2011):

$$\sum_{i,j=1}^{3} M_{ij}^2 = 1. \tag{2}$$

Because Eq. 1 is linear, a suitable left pseudo-inverse can be calculated and the moment tensor determined. The seismic amplitude equation (Eq. 1) can also represent the full waveform problem, with **u** signifying the vector of waveforms and the station propagation coefficients (**a**), now time dependent and containing the Green functions and station responses. This is again a linear problem for known Green functions.

When using first-motion polarity information (Y) for source inversion, the signum function prevents a linear inversion approach:

$$Y = \mathrm{sgn}\left(\mathbf{a} \cdot \tilde{\mathbf{M}}\right). \tag{3}$$

For amplitude inversions, the propagation effects are dependent on the velocity structure and attenuation models used, whereas amplitude ratios are less dependent on these, and so are often used in preference to absolute amplitudes (Hardebeck & Shearer 2002, 2003; Snoke 2003). While it is possible to invert directly for the moment tensor using amplitudes or amplitude ratios, multiple different observation types can make it difficult to invert for

the source directly, and impossible when using the easily measured first-motion polarities due to the non-linear signum function (Eq. 3).

2 BAYES' THEOREM AND PROBABILITIES

Inversion methods aim to find the best fitting model parameters for the given data. In the case of source inversion, the probability of obtaining the data is evaluated for the possible sources. The resulting estimates of the probability density function (PDF) can be combined for all the data to approximate the true PDF for the source. This PDF describes the likelihood of observing the data for a particular source given the model, $\mathrm{p}\,(\mathrm{data}\,|\,\mathrm{model})$. However, the value of interest in such an inversion is the probability of the model given the observed data, the posterior PDF. This can be evaluated from the likelihood using Bayes' Theorem (Bayes & Price 1763; Laplace 1812; Sivia 2000):

$$\mathrm{p}\,(\mathrm{model}\,|\,\mathrm{data}) = \frac{\mathrm{p}\,(\mathrm{data}\,|\,\mathrm{model})\ \mathrm{p}\,(\mathrm{model})}{\mathrm{p}\,(\mathrm{data})}. \tag{4}$$

Bayes' theorem links the two PDFs using prior probabilities. This incorporates the prior known information for the data, and can be used to include known constraints such as fault geometries or source type. Quantifying the prior state of knowledge for a system is often not trivial, and Bayes' theorem states that any assumptions made when formulating the prior will be directly reflected in the posterior PDF (Sivia 2000, Chapter 2). Although a commonly used prior is the null prior, which is a uniform distribution and does not change the posterior from the likelihood.

The normalised full moment tensor has five independent components, compared to three in the normalised double-couple tensor. These extra free parameters provide an improved fit, for example, to both data and noise. However, the commonly held view is that most earthquakes are double-couple, so a possible choice of prior could reduce the

tendency for full moment tensor inversions to over-fit double-couple solutions, without taking the step of forcing the solution to be double-couple.

3 UNCERTAINTIES AND MARGINALISATION

Source inversion is particularly sensitive to uncertainties and, because it is usually carried out after several other necessary steps, the effect of the uncertainties may be difficult to understand. The uncertainties can be broken down into several different types: instrument errors, model errors, and background noise. Table 1 shows some of the error types and their origins, although there is an additional source of uncertainty arising from an inability to resolve all the model parameters given the available data.

Although interdependencies between the uncertainties can be explored, no quantitative relationship is known. For such a treatment to be truly rigorous, the variations in these errors must be included throughout the inversion.

The Bayesian formulation allows rigorous inclusion of uncertainties in the problem using marginalisation (Sivia 2000, Section 1.3). Marginalisation removes the dependence of the joint PDF, $\mathrm{p}\,(A, B)$, on one of the variables by integrating over the PDF for that variable:

$$\mathrm{p}\,(A) = \int \mathrm{p}\,(A\,,\,B)\,dB = \int \mathrm{p}\,(A \,|\, B)\,\mathrm{p}\,(B)\,\mathrm{d}B. \tag{5}$$

Consequently, marginalisation can be used to reduce the PDF to a form that is dependent only on the parameters of interest, while taking all the uncertainties into account.

To incorporate the measurement errors in the inversion, for example, it is necessary to produce a PDF for the error (Δ_{mes}) over which to marginalise. The assumption made for the background noise is that the mean and variance are measurable and, therefore, the most ambiguous (maximum entropy) distribution for these measurements is the Gaussian distribution, as can be shown using variational calculus (Appendix B). For de-meaned data, the background noise can be assumed also to have zero mean, and a standard

deviation σ_{mes}, so that the PDF is:

$$\mathrm{p}\left(\Delta_{\mathrm{mes}}\right) = \frac{1}{\sqrt{2\pi\sigma_{\mathrm{mes}}^2}} \mathrm{e}^{-\frac{\Delta_{\mathrm{mes}}^2}{2\sigma_{\mathrm{mes}}^2}}. \tag{6}$$

Location and velocity model uncertainties can have significant effects on the source. A common location method for microseismic events is NonLinLoc (Lomax et al. 2000, 2009), which produces samples from the source location PDF. Velocity model uncertainty can be included in that method by marginalising the location PDF with respect to the velocity models. These approaches can therefore be used to incorporate location and velocity model uncertainty in the source inversion approach as described below.

4 INVERSION APPROACH

The inversion approach described here is designed for microseismic events. In many cases where microseismic events are recorded, it would be difficult and time consuming to calculate the Green functions for all the events due to poorly constrained velocity models and event locations that often vary across the network, increasing the number of calculations. Furthermore, velocity variation has a larger effect on the ray-paths due to the close proximity of the receivers to the source, unlike the teleseismic case. Our approach aims to produce a robust result for the moment tensor, which includes known uncertainties in the measurements. The approach has been tested with polarity and amplitude ratio data, and can easily be extended to other possible data types such as amplitude and the full-waveform fitting, provided a suitable PDF can be described relating the source to the data and the uncertainties.

A necessary caveat when performing any inversion is that any dependence in the data must be accounted for. For example, if combining P-, SH-, and SV-amplitude measurements into amplitude ratios, the ratios used must be independent, so from three measurements it is possible to construct only two independent ratios.

The uncertainties, both in the polarity and amplitude ratio measurements and in the model parameters are included in the source PDF using Bayesian marginalisation (Section 3). This requires an estimate of the PDFs for the different uncertainties, either from estimates on the measurements or from direct sampling of the associated PDF if it is known (e.g. location). Often a Gaussian uncertainty is assumed as the uncertainty PDF, but this may be incorrect, leading to systematic errors in the resultant PDFs.

4.1 Posterior PDF

A probabilistic approach to source inversion requires an approximation to the posterior PDF for the source parameters, $\mathrm{p}\left(\mathbf{M} \mid \mathbf{d}, \mathbf{k}\right)$. The parameters used in this section are summarised in Table 2. The posterior PDF depends on the observations, $\mathbf{d} = (\mathbf{d}', \tau)$, the unknown nuisance parameters that are marginalised over, $\epsilon = (\epsilon', \mathbf{x})$, where $\epsilon' = (\mathbf{A}, \sigma, \varpi)$, and the known parameters, $\mathbf{k} = (\mathbf{s}, \sigma_\tau, \boldsymbol{G})$.

Bayes' theorem (Eq. 4) allows us to rewrite the posterior PDF in terms of the likelihood of the data and the model prior:

$$\mathrm{p}\left(\mathbf{M} \mid \mathbf{d}, \mathbf{k}\right) \propto \mathrm{p}\left(\mathbf{d} \mid \mathbf{M}, \mathbf{k}\right) \mathrm{p}\left(\mathbf{M}\right). \tag{7}$$

Expanding the observations, $\mathbf{d}$, into those used for the source mechanism inversion, $\mathbf{d}'$, and the phase arrival times, τ, gives:

$$\mathrm{p}\left(\mathbf{M} \mid \mathbf{d}', \tau, \mathbf{k}\right) \propto \mathrm{p}\left(\mathbf{d}' \mid \mathbf{M}, \tau, \mathbf{k}\right) \mathrm{p}\left(\mathbf{M}\right), \tag{8}$$

which is dependent on the likelihood of the source type observations, $\mathrm{p}\left(\mathbf{d}' \mid \mathbf{M}, \tau, \mathbf{k}\right)$. This likelihood is a joint distribution over the observed data used for constraining the source, and can be separated into the product over the set of N receivers with observation,

because the measurements at a receiver are independent of the others:

$$\mathrm{p}\left(\mathbf{d}' \mid \mathbf{M}, \tau, \mathbf{k}\right) = \prod_{i=1}^{N} \mathrm{p}\left(\mathbf{d}'_i \mid \mathbf{M}, \tau_i, k_i\right), \tag{9}$$

where the i subscript indicates the coordinates for receiver i. Unfortunately the data likelihood is dependent on the nuisance parameters, so to obtain Eq. 9, the nuisance parameters, ϵ, must be marginalised (Eq. 5) over:

$$\mathrm{p}\left(\mathbf{d}' \mid \mathbf{M}, \tau, \mathbf{k}\right) = \int \prod_{i=1}^{N} \left[\mathrm{p}\left(\mathbf{d}'_i \mid \mathbf{M}, \tau_i, \epsilon_i, k_i\right)\right] \mathrm{p}\left(\epsilon \mid \mathbf{M}, \tau, k\right) \mathrm{d}\epsilon, \tag{10}$$

where the product corresponds to the product of the data PDFs over each receiver. The nuisance parameters can be expanded using the product rule for conditional probabilities:

$$\mathrm{p}\left(a, b, c\right) = \mathrm{p}\left(a|b, c\right) \mathrm{p}\left(b|c\right) \mathrm{p}\left(c\right) \tag{11}$$

The nuisance parameter components are: ϵ', which includes the modelled amplitudes and data uncertainties, and the location $\mathbf{x}$. This gives:

$$\mathrm{p}\left(\mathbf{d}' \mid \mathbf{M}, \tau, \mathbf{k}\right) = \iint \prod_{i=1}^{N} \mathrm{p}\left(\mathbf{d}'_i \mid \mathbf{M}, \tau_i, \epsilon'_i, \mathbf{x}, k_i\right) \mathrm{p}\left(\mathbf{x} \mid \mathbf{M}, \tau, \epsilon', k\right) \mathrm{p}\left(\epsilon' \mid \mathbf{M}, \tau, k\right) \mathrm{d}\mathbf{x}\mathrm{d}\epsilon'. \tag{12}$$

However, the observations, $\mathbf{d}'_i$, depend only on ϵ' and the pick times, while the location, $\mathbf{x}$, depends only on the arrival times, τ, and the known parameters, $\mathbf{k}$. Therefore, the likelihood can be simplified to:

$$\mathrm{p}\left(\mathbf{d}' \mid \mathbf{M}, \tau, \mathbf{k}\right) = \iint \prod_{i=1}^{N} \mathrm{p}\left(\mathbf{d}'_i \mid \tau_i, \epsilon'_i\right) \mathrm{p}\left(\mathbf{x} \mid \tau, k\right) \mathrm{p}\left(\epsilon' \mid \mathbf{M}, \tau, k\right) \mathrm{d}\mathbf{x}\mathrm{d}\epsilon'. \tag{13}$$

Expanding for the nuisance variables in ϵ' using the product rule and splitting these

variables into the receiver dependent and independent parameters gives:

$$\mathrm{p}\left(\mathbf{d}' \mid \mathbf{M}, \tau, \mathbf{k}\right) = \iiint \prod_{i=1}^{N} \left[\int \mathrm{p}\left(\mathbf{d}'_i \mid \mathbf{A}_i, \sigma_i, \varpi_i, \tau_i\right) \mathrm{p}\left(\mathbf{A}_i \mid \mathbf{x}, \mathbf{M}\right) \mathrm{d}\mathbf{A}_i \right] \mathrm{p}\left(\mathbf{x} \mid \tau, k\right) \mathrm{p}\left(\sigma\right) \mathrm{p}\left(\varpi\right) \mathrm{d}\mathbf{x}\mathrm{d}\sigma\mathrm{d}\varpi. \quad (14)$$

The likelihood of observing the data given the source model parameters, the time picks and associated uncertainties, the station locations and the velocity model is marginalised over four nuisance parameters: the theoretical amplitude probability, $\mathrm{p}\left(\mathbf{A}_i \mid \mathbf{x}, \mathbf{M}\right)$, which is dependent on the location and the model parameters; the location PDF, $\mathrm{p}\left(\mathbf{x} \mid \tau, k\right)$; the probability of obtaining the observed measurement errors for the polarities and amplitude ratios, $\mathrm{p}\left(\sigma\right)$; and the probability of an instrument trace reversal, $\mathrm{p}\left(\varpi\right)$.

The theoretical amplitude probability can be written in terms of the dot product between the station propagation coefficients and the moment tensor six-vector, $\mathbf{a}_i \cdot \tilde{\mathbf{M}}$, where the station propagation coefficients depend on the location, $\mathbf{x}$. The PDF for the amplitude for a given source location, $\mathbf{x}_j$, is dependent only on that location, so is given by a delta function, $\delta\left(\mathbf{A}_i - \mathbf{a}_{i;j} \cdot \tilde{\mathbf{M}}\right)$, where $\mathbf{a}_{i;j} = \mathbf{a}_i\left(\mathbf{x}_j\right)$ refers to the station propagation coefficients for the receiver i, associated with the location at $\mathbf{x}_j$:

$$\mathrm{p}\left(\mathbf{d}' \mid \mathbf{M}, \tau, \mathbf{k}\right) = \iiint \prod_{i=1}^{N} \left[\int \mathrm{p}\left(\mathbf{d}'_i \mid \mathbf{A}_i, \sigma_i, \varpi_i, \tau_i\right) \delta\left(\mathbf{A}_i - \mathbf{a}_{i;j} \cdot \tilde{\mathbf{M}}\right) \mathrm{d}\mathbf{A}_i \right] \mathrm{p}\left(\mathbf{x} = \mathbf{x}_j \mid \tau, k\right) \mathrm{p}\left(\sigma\right) \mathrm{p}\left(\varpi\right) \mathrm{d}\mathbf{x}\mathrm{d}\sigma\mathrm{d}\varpi. \quad (15)$$

Integrating over the delta function (Eq. 57) is equivalent to evaluating the theoretical amplitude associated with the location:

$$\mathrm{p}\left(\mathbf{d}' \mid \mathbf{M}, \tau, \mathbf{k}\right) = \iiint \prod_{i=1}^{N} \mathrm{p}\left(\mathbf{d}'_i \mid \mathbf{A}_i = \mathbf{a}_{i;j} \cdot \tilde{\mathbf{M}}, \sigma_i, \varpi_i, \tau_i\right) \mathrm{p}\left(\mathbf{x} = \mathbf{x}_j \mid \tau, k\right) \mathrm{p}\left(\sigma\right) \mathrm{p}\left(\varpi\right) \mathrm{d}\mathbf{x}\mathrm{d}\sigma\mathrm{d}\varpi. \quad (16)$$

The integral over the location uncertainty is not analytic but can be evaluated using a Monte Carlo approach by summing over M hypocentre samples drawn from the location PDF, $\mathrm{p}\left(\mathbf{x} = \mathbf{x}_j \mid \tau, k\right)$:

$$\mathrm{p}\left(\mathbf{d}' \mid \mathbf{M}, \tau, \mathbf{k}\right) = \iint \sum_{j=1}^{M} \prod_{i=1}^{N} \mathrm{p}\left(\mathbf{d}'_i \mid \mathbf{A}_i = \mathbf{a}_{i;j} \cdot \tilde{\mathbf{M}}, \sigma_i, \varpi_i, \tau_i\right) \mathrm{p}\left(\sigma\right) \mathrm{p}\left(\varpi\right) \mathrm{d}\sigma \mathrm{d}\varpi. \quad (17)$$

Eq. 17 includes the location uncertainty in the PDF by converting the integral over $\mathrm{d}\mathbf{x}$ to a sum over the location PDF samples.

Following Walsh et al. (2009), the likelihood for a case with an unknown earth model differs from the known earth model case (Eq. 17). It has an additional Monte Carlo integration over the possible velocity models, with the station propagation coefficients now dependent on the location, $\mathbf{x}$, and the earth model, $\boldsymbol{G}$, where $\mathbf{a}_{i;jk} = \mathbf{a}_i\left(\mathbf{x}_j, \boldsymbol{G}_k\right)$ refers to the station propagation coefficients associated with a location at $\mathbf{x}_j$ and an earth model $\boldsymbol{G}_k$. For Q earth models and M locations, there are $Q \times M$ location and velocity model samples to sum over to carry out the marginalisation:

$$\mathrm{p}\left(\mathbf{d}' \mid \mathbf{M}, \tau, \mathbf{k}\right) = \iint \sum_{k=1}^{Q} \sum_{j=1}^{M} \prod_{i=1}^{N} \mathrm{p}\left(\mathbf{d}'_i \mid \mathbf{A}_i = \mathbf{a}_{i;jk} \cdot \tilde{\mathbf{M}}, \sigma_i, \varpi_i, \tau_i\right) \mathrm{p}\left(\sigma\right) \mathrm{p}\left(\varpi\right) \mathrm{d}\sigma \mathrm{d}\varpi. \quad (18)$$

Assuming the measurements are independent, the data likelihood at receiver i, $\mathrm{p}\left(\mathbf{d}'_i \mid \mathbf{A}_i = \mathbf{a}_{i;jk} \cdot \tilde{\mathbf{M}}, \sigma_i, \varpi_i, \tau_i\right)$, can be expanded as the product of the different PDFs observation types (Sivia 2000, Section 3.5). Sections 4.2 and 4.3 define specific PDFs for the polarity, $\mathrm{p}\left(Y_i \mid \mathbf{A}_i = \mathbf{a}_{i;jk} \cdot \tilde{\mathbf{M}}, \sigma_i, \varpi_i, \tau_i\right)$, and amplitude ratios, $\mathrm{p}\left(\mathbf{R}_i \mid \mathbf{A}_i = \mathbf{a}_{i;jk} \cdot \tilde{\mathbf{M}}, \sigma_i, \tau_i\right)$, giving:

$$\mathrm{p}\left(\mathbf{d}'_i \mid \mathbf{A}_i = \mathbf{a}_{i;jk} \cdot \tilde{\mathbf{M}}, \sigma_i, \varpi_i, \tau_i\right) = \mathrm{p}\left(Y_i \mid \mathbf{A}_i = \mathbf{a}_{i;jk} \cdot \tilde{\mathbf{M}}, \sigma_i, \varpi_i, \tau_i\right) \mathrm{p}\left(\mathbf{R}_i \mid \mathbf{A}_i = \mathbf{a}_{i;jk} \cdot \tilde{\mathbf{M}}, \sigma_i, \tau_i\right). \quad (19)$$

There is no dependence on the trace reversal probability, ϖ_i, in the amplitude ratio likelihood because the observations are unsigned, so are unaffected.

$\mathrm{p}(\sigma)$ and $\mathrm{p}(\varpi)$ are the chosen priors for the measurement uncertainties and the probability of instrument trace reversal respectively. It is possible to remove any dependence on the trace reversal probability, ϖ_i, by using observations from a source with a known mechanism (e.g., a teleseismic earthquake) to calibrate the trace orientation. So the trace reversal probability in Eqs 27 and 59 would be either 0 or 1, or more usefully, any incorrect traces could be inverted so that $\varpi_i = 0$ for all stations.

The likelihoods for a known velocity model (Eq. 17) and unknown velocity models (Eq. 18) have been defined in some detail. The equations include uncertainties in the observations, the location, with all of its associated uncertainties, and the probability of a trace reversal. The source model distribution is fully described by the posterior PDF, $\mathrm{p}(\mathbf{M} \mid \mathbf{d}, \mathbf{k})$, which can be evaluated by multiplying the likelihood (Eqs 17 and 18) with the chosen source prior.

4.2 Polarity PDF

The first-arrival polarity is a commonly used measurement for constraining source inversion (e.g. Reasenberg & Oppenheimer 1985; Hardebeck & Shearer 2002; Snoke 2003), and is measured as the direction of the first peak of the arrival waveform at a receiver (Fig. 2 (a)). It is dependent on correct determination of the arrival onset, as well as a clear impulsive arrival characteristic.

The PDF for polarity, Y, given a polarity observation, y, at an instrument for a given theoretical amplitude, A, is given by a step function such as the Heaviside step function,

$$\mathrm{H}(x) = \int_{-\infty}^{x} \delta(s)\,\mathrm{d}s. \tag{20}$$

This PDF is dependent on the error in the amplitude, Δ_{Y},

$$\mathrm{p}(Y = y \mid A, \Delta_{\mathrm{Y}}) = \mathrm{H}(y\,(A + \Delta_{\mathrm{Y}}))\,. \tag{21}$$

Marginalising Eq. 21 over Δ_{Y}, for a given standard deviation, σ_{Y}, using a Gaussian noise model (Eq. 6) gives:

$$\mathrm{p}\left(Y = y \,|\, A, \sigma_{\mathrm{Y}}\right) = \int_{-\infty}^{\infty} \mathrm{p}\left(Y = y \,|\, A, \Delta_{\mathrm{Y}}\right)\, \mathrm{p}\left(\Delta_{\mathrm{Y}} \,|\, \sigma_{\mathrm{Y}}\right) \mathrm{d}\Delta_{\mathrm{Y}}, \tag{22}$$

$$\mathrm{p}\left(Y = y \,|\, A, \sigma_{\mathrm{Y}}\right) = \int_{-\infty}^{\infty} \mathrm{H}\left(yA + y\Delta_{\mathrm{Y}}\right)\, \frac{1}{\sqrt{2\pi\sigma_{\mathrm{Y}}^2}} \mathrm{e}^{-\frac{\Delta_{\mathrm{Y}}^2}{2\sigma_{\mathrm{Y}}^2}} \mathrm{d}\Delta_{\mathrm{Y}}. \tag{23}$$

The product $y\Delta_{\mathrm{Y}}$ changes the sign of the noise to reflect the polarity, but because the PDF for Δ_{Y} is symmetric about 0, this change in sign has no effect. Eq. 23 can be integrated using this behaviour of the step function:

$$\int_{-\infty}^{\infty} \mathrm{H}\left(x + \Delta\right) f\left(\Delta\right) \mathrm{d}\Delta = \int_{-x}^{\infty} f\left(\Delta\right) d\Delta. \tag{24}$$

So, Eq. 23 can be rewritten as:

$$\mathrm{p}\left(Y = y \,|\, A, \sigma_{\mathrm{Y}}\right) = \int_{-yA}^{\infty} \frac{1}{\sqrt{2\pi\sigma_{\mathrm{Y}}^2}} \mathrm{e}^{-\frac{\Delta_{\mathrm{Y}}^2}{2\sigma_{\mathrm{Y}}^2}} \mathrm{d}\Delta_{\mathrm{Y}}, \tag{25}$$

which due to the symmetry of the normal distribution about the mean, results in:

$$\mathrm{p}\left(Y = y \,|\, A, \sigma_{\mathrm{Y}}\right) = \int_{-\infty}^{yA} \frac{1}{\sqrt{2\pi\sigma_{\mathrm{Y}}^2}} \mathrm{e}^{-\frac{\Delta_{\mathrm{Y}}^2}{2\sigma_{\mathrm{Y}}^2}} \mathrm{d}\Delta_{\mathrm{Y}} = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{yA}{\sqrt{2}\sigma_{\mathrm{Y}}}\right)\right), \tag{26}$$

with the Gauss error function defined as $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2} \mathrm{d}t$. However, there is also a possibility that a manufacturing or user error has occurred and the instrument is incorrectly wired, such that the trace is inverted. Consequently, the PDF for a given polarity is also dependent on the probability of such a trace inversion occurring, ϖ:

$$\mathrm{p}\left(Y = y \,|\, A, \sigma_{\mathrm{Y}}, \varpi\right) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{yA}{\sqrt{2}\sigma_{\mathrm{Y}}}\right)\right)(1 - \varpi) + \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{-yA}{\sqrt{2}\sigma_{\mathrm{Y}}}\right)\right)\varpi \tag{27}$$

This polarity PDF is consistent with that shown by Brillinger et al. (1980) and Walsh et al. (2009).

Because the moment tensor six-vector is normalised, the modelled amplitude can therefore take values between -1 and 1. The polarity standard deviation, σ_{Y}, is related to the amplitude uncertainty compared to the maximum theoretical amplitude at the station (based purely on the event magnitude, source-to-station propagation and receiver coupling). However, this is difficult to calculate, especially for events with poor focal sphere coverage.

Walsh et al. (2009) call σ_{Y} the noise for the arrival, but this does not scale correctly in comparison to the modelled amplitude due to the propagation effects. They treat it as a user-determined value, representing the confidence in the arrival. σ_{Y} could be estimated from the fractional amplitude uncertainty, but this will be greater than or equal to the true value, because the amplitude at a receiver is only ever less than or equal to the maximum theoretical amplitude (accounting for propagation effects). Therefore, this would most likely overestimate the uncertainty. It is clear that the uncertainty value should be station-specific because noise environments at different stations often vary, so the maximum estimate of the event signal-to-noise ratio (SNR) fails to account for the variation across the receivers.

The difficulty in estimating the uncertainty is increased further when polarity picking is done manually, so the uncertainty on the trace is perhaps not known. Due to the difficulty in quantifying the uncertainty, it is best left as a user-defined parameter, as Walsh et al. (2009) do, that reflects the confidence in the arrival polarity pick, which can be mapped to the pick quality. However, Pugh et al. (2016) propose an alternate method for calculating polarity uncertainties that can be included in this framework.

4.3 Amplitude Ratio PDF

The amplitude ratio is more complex than the polarity, and there are different approaches to the measurement (Fig. 2 (b) and (c)), each of which can have different effects (Section 4.3.1). Moreover, as the uncertainty in the amplitude measurement is usually modelled as a Gaussian, the uncertainty in the amplitude ratio cannot also be Gaussian, but is more complex (Section 4.3.2). This results in a non-Gaussian amplitude ratio likelihood (Section 4.3.3).

The modelled amplitude ratio is dependent on the theoretical amplitudes, A_{P}, A_{SH}, and A_{SV}. Each amplitude observation has a measurement uncertainty, and the uncertainty of the ratio is not simply the ratio of the amplitude uncertainties. The PDF for a given ratio observation vector $\mathbf{r} = \left(r^{\mathrm{P}}_{\mathrm{SH}}, r^{\mathrm{P}}_{\mathrm{SV}}\right)$ at an instrument with given theoretical amplitudes A_{P}, A_{SH}, and A_{SV} is two dimensional. However, because the observations are independent, the PDFs are independent and, therefore, the PDFs for each ratio can be multiplied together to give the total PDF for the ratios.

Considering a single observed ratio $r = \frac{x}{y}$ and theoretical amplitudes A_{X} and A_{Y}, the PDF for the ratio is given by the delta function $\delta(x)$, although this is again dependent on the measurement error of the two observations Δ^{X} and Δ^{Y}:

$$\mathrm{p}\left(r = \frac{x}{y} \,|\, A_{\mathrm{X}}, A_{\mathrm{Y}}, \Delta^{\mathrm{X}}, \Delta^{\mathrm{Y}}\right) = \delta\left(\frac{A_{\mathrm{X}} + \Delta^{\mathrm{X}}}{A_{\mathrm{Y}} + \Delta^{\mathrm{Y}}} - \frac{x}{y}\right). \tag{28}$$

4.3.1 Measuring The Amplitude Ratio

There are several approaches for measuring the amplitude of a seismic arrival. The simplest approach is to take the maximum absolute value (A_{max}) from the arrival window. However, the maximum peak-to-peak amplitude (A_{PP}) has a higher SNR for uncorrelated noise (Fig. 2 (b)). A logical extension of the peak-to-peak approach is the root-mean-square (RMS) amplitude of samples within the arrival window (A_{RMS}) (Fig. 2 (c)) because it reduces the dependence of the observation on the (uncorrelated) noise

further by a factor dependent on the window length n,

$$\sigma_{\mathrm{RMS}} \propto \frac{\sigma}{\sqrt{n}}. \tag{29}$$

However, the choice of window length is important because the arrival is finite; therefore, extending the RMS window a long way beyond the noise reduces the effect of the signal by a similar factor of $\sqrt{n}$. As a consequence, it can be difficult to compare RMS observations for different window lengths.

The choice between these different approaches is, therefore, not straightforward. Additionally, the presence of noise on an arrival complicates the matter. The noise leads to an increasing difference from the true value of the amplitude, as shown in Fig. 3. This figure shows the results of a synthetic test of the distribution of amplitude measurements for a synthetic arrival with known amplitude and different levels of Gaussian noise added. The mean of the empirical distribution increases as the noise level increases, moving away from the true value of the arrival amplitude given by the broken vertical lines and the corresponding Gaussian distributions. While a correction could be applied to the mean of the distribution to return it to the true value, the value is dependent on both the noise and the arrival, making it impractical. The width of the distribution also increases as expected, although less than the modelled value from the Gaussian distribution. This deviation from the true value is due to amplitude measurements being determined by the extrema in the arrival. They will therefore tend to increase as the noise increases. If the arrivals of different phases have disparate noise levels, it can cause large deviations from the expected distribution in the ratio (Eq. 59), as shown in Fig. 4. The deviation in the amplitude ratio distribution from the expected distribution is greatest when the difference in the fractional noise (inversely proportional to the true SNR) is largest, especially if the larger value is in the numerator. If the model source parameters reproduce the true amplitude ratio of the arrivals, this systematic deviation in the observed values

due to the noise can reduce the probability of obtaining the correct source parameters, potentially leading to incorrect results (discussed further in Section 6).

Hardebeck & Shearer (2003) found that S-wave/P-wave amplitude ratio uncertainties may be large due to noise and site effects, and are sometimes not effective in improving well-constrained solutions. Fig. 4 shows that the choice of amplitude estimate can have an effect and, although the variation in the distributions of $A_{\max}$ and A_{PP} is small and there is little difference between the two, there can be large differences between $A_{\max}$ and A_{RMS}.

The amplitude observations must be consistent between different receivers, which can be complicated by the different coupling of the receiver components to the ground. This coupling varies between receivers, creating the need foran additional correction.

4.3.2 Amplitude Ratio Noise Model

The observed amplitude ratio does not simply relate to the ratio of the modelled amplitudes, but it is dependent on a propagation correction, Z (c.f. Aki & Richards 2002, Chapter 4 eq. 4.97):

$$R_{\mathrm{SH}}^{\mathrm{P}} = \frac{A_{\mathrm{P}}}{Z\, A_{\mathrm{SH}}}. \tag{30}$$

While this is unity for S amplitude ratios, the coefficient depends on the velocity ratio between the different phases involved.

The simplest form of the propagation correction, assuming constant $V_{\mathrm{P}}/V_{\mathrm{S}}$ ratio along the ray-path, is given by:

$$Z = \left(V_{\mathrm{P}}/V_{\mathrm{S}}\right)^3. \tag{31}$$

Assuming a Gaussian uncertainty for the $V_{\mathrm{P}}/V_{\mathrm{S}}$ ratio, it becomes necessary to determine the distribution of the propagation coefficient (Fig. 5). The mean and standard deviation

are given by:

$$\mu_{\mathrm{Z}} = \mu^3_{V_{\mathrm{P}}/V_{\mathrm{S}}}, \tag{32}$$

$$\sigma_{\mathrm{Z}} = 3\left(V_{\mathrm{P}}/V_{\mathrm{S}}\right)^2 \sqrt{\sigma^2_{V_{\mathrm{P}}/V_{\mathrm{S}}}}, \tag{33}$$

where σ_{Z} can be determined from the Taylor expansion of the correction function:

Fig. 5 shows that a Gaussian distribution is a good approximation to the true distribution of values, provided that the percentage uncertainty on the $V_{\mathrm{P}}/V_{\mathrm{S}}$ ratio is small compared to the ratio $\left(\mu_{V_{\mathrm{P}}/V_{\mathrm{S}}} \gtrsim 9\sigma_{V_{\mathrm{P}}/V_{\mathrm{S}}}\right)$. The $V_{\mathrm{P}}/V_{\mathrm{S}}$ percentage uncertainty is typically small, so the Gaussian noise model is a valid model to use. Fractional uncertainties within 0.2 often encompass all the uncertainty in the Wadati plot of S-to-P delay times versus P-arrival time (Wadati 1933; Wadati & Oki 1933).

However, instead of applying the correction to the modelled S-amplitude, it can be applied to the observed P-amplitude, as in this case. The corrected P-amplitude can be approximated by a Gaussian distribution, with mean and standard deviation given by:

$$\mu_{A_{\mathrm{P}'}} = \mu_{\mathrm{Z}} * \mu_{A_{\mathrm{P}}}, \tag{34}$$

$$\sigma_{A_{\mathrm{P}'}} = \sqrt{A^2_{\mathrm{P}}\sigma^2_{\mathrm{Z}} + Z^2\sigma^2_{A_{\mathrm{P}}}}. \tag{35}$$

Fig. 6 shows the distributions of the corrected P-wave amplitudes, evaluated for a range of $V_{\mathrm{P}}/V_{\mathrm{S}}$ uncertainties and noise levels. For small $V_{\mathrm{P}}/V_{\mathrm{S}}$ uncertainties, the distribution is approximately Gaussian. However, as the uncertainty increases it becomes skewed, although the region where the Gaussian approximation fails is higher than the expected $V_{\mathrm{P}}/V_{\mathrm{S}}$ uncertainties.

It is possible to describe the amplitude ratio in terms of two parameters: the mean

and standard deviation, given by:

$$\mu_{\mathrm{R}} = \frac{\mu_{\mathrm{Z}}\mu_{A_{\mathrm{P}}}}{\mu_{A_{\mathrm{SH}}}} = \frac{\mu_{A_{\mathrm{P}'}}}{\mu_{A_{\mathrm{SH}}}}, \tag{36}$$

$$\sigma_{\mathrm{R}} = R^{\mathrm{P}}_{\mathrm{SH}}\sqrt{\frac{\sigma^2_{\mathrm{P}}}{A^2_{\mathrm{P}}} + \frac{\sigma^2_{\mathrm{SH}}}{A^2_{\mathrm{SH}}} + \frac{\sigma^2_{\mathrm{Z}}}{Z^2}}. \tag{37}$$

In this case, the maximum entropy distribution is the Gaussian distribution (Appendix B).

These two parameters describe a possible Gaussian model for the amplitude ratio uncertainty, with the PDF given by:

$$\mathrm{p}\left(R = |r| \,|\mu_{\mathrm{R}}, \sigma_{\mathrm{R}}\right) = \mathcal{N}\left(r, \mu_{\mathrm{R}}, \sigma_{\mathrm{R}}\right) + \mathcal{N}\left(-r, \mu_{\mathrm{R}}, \sigma_{\mathrm{R}}\right), \tag{38}$$

$$\mathrm{p}\left(R = |r| \,|\mu_{\mathrm{R}}, \sigma_{\mathrm{R}}\right) = \frac{2}{\sqrt{2\pi\sigma^2_{\mathrm{R}}}}\left(\mathrm{e}^{-\frac{(r-\mu_{\mathrm{R}})^2}{2\sigma^2_{\mathrm{R}}}} + \mathrm{e}^{-\frac{(-r-\mu_{\mathrm{R}})^2}{2\sigma^2_{\mathrm{R}}}}\right), \tag{39}$$

defined over $r \geqslant 0$, and the cumulative distribution function (CDF) by:

$$\mathrm{P}\left(R \leqslant |r| \,|\mu_{\mathrm{R}}, \sigma_{\mathrm{R}}\right) = \Phi\left(r, \mu_{\mathrm{R}}, \sigma_{\mathrm{R}}\right) + \Phi\left(-r, \mu_{\mathrm{R}}, \sigma_{\mathrm{R}}\right) - 1, \tag{40}$$

$$\mathrm{P}\left(R \leqslant |r| \,|\mu_{\mathrm{R}}, \sigma_{\mathrm{R}}\right) = \frac{1}{2}\left(\mathrm{erf}\left(\frac{r-\mu_{\mathrm{R}}}{\sqrt{2}\sigma_{\mathrm{R}}}\right) + \mathrm{erf}\left(\frac{-r-\mu_{\mathrm{R}}}{\sqrt{2}\sigma_{\mathrm{R}}}\right)\right), \tag{41}$$

where the Gaussian PDF is given by:

$$\mathcal{N}\left(r, \mu_{\mathrm{R}}, \sigma_{\mathrm{R}}\right) = \frac{1}{\sqrt{2\pi\sigma^2_{\mathrm{R}}}}\mathrm{e}^{-\frac{(r-\mu_{\mathrm{R}})^2}{2\sigma^2_{\mathrm{R}}}}, \tag{42}$$

and the CDF by:

$$\Phi\left(r, \mu_{\mathrm{R}}, \sigma_{\mathrm{R}}\right) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{r-\mu_{\mathrm{R}}}{\sqrt{2}\sigma_{\mathrm{R}}}\right)\right). \tag{43}$$

The inversion approach uses unsigned (absolute) amplitude ratios, so the distribution reflected in $R = 0$ must be included.

Figs 7 and 8 show simulated distributions of the ratio given a Gaussian uncertainty in

the amplitudes and the V_P/V_S ratio, and it is clear that, when the denominator uncertainty is significant, the Gaussian is a poor noise model.

While the Gaussian distribution is the maximum entropy distribution given a mean and a variance, so is consistent with the uncertainty on the amplitude observations, the modelled amplitude ratio is a ratio of two normally distributed parameters. So the distribution of the ratio must be used. Assuming that a Gaussian model is sufficient for the corrected P amplitude, there are, in fact, five parameters: $\mu_{P'}$, μ_{SH}, $\sigma_{P'}$, σ_{SH}, and ρ.

Fieller (1932) derived a distribution for the ratio, $R = \frac{X}{Y}$, of two normally distributed observations with means μ_X and μ_Y, standard deviations σ_X and σ_Y, and correlation ρ. ρ is the correlation between the parameters. The PDF for R is determined from the joint density of X, Y, $g(x, y)$ by:

$$\mathrm{p}(R = r) = \int_{-\infty}^{\infty} |y|\, g(ry, y)\, \mathrm{d}y. \tag{44}$$

Substituting a bivariate normal density for $g(x, y)$ gives the ratio PDF (Hinkley 1969),

$$\mathrm{p}(R = r) = \frac{b(r)\, d(r)}{\sigma_X \sigma_Y a^3(r) \sqrt{2\pi}} \left[\Phi\left(\frac{b(r)}{a(r)\sqrt{1-\rho^2}} \right) - \Phi\left(\frac{-b(r)}{a(r)\sqrt{1-\rho^2}} \right) \right] + \frac{\sqrt{1-\rho^2}}{\pi \sigma_X \sigma_Y a^2(r)} \mathrm{e}^{\left(-\frac{c}{2(1-\rho^2)}\right)}, \tag{45}$$

where $\Phi(x)$is the Gaussian CDF (Eq. 43). The coefficients in Eq. 45,$a(r)$, $b(r)$, c, and

$d(r)$ given by:

$$a(r) = \sqrt{\frac{r^2}{\sigma_{\mathrm{X}}^2} - 2\rho\frac{r}{\sigma_{\mathrm{X}}\sigma_{\mathrm{Y}}} + \frac{1}{\sigma_{\mathrm{Y}}^2}}, \tag{46}$$

$$b(r) = \frac{\mu_{\mathrm{X}} r}{\sigma_{\mathrm{X}}^2} - \rho\frac{\mu_{\mathrm{X}} + \mu_{\mathrm{Y}} r}{\sigma_{\mathrm{X}}\sigma_{\mathrm{Y}}} + \frac{\mu_{\mathrm{Y}}}{\sigma_{\mathrm{Y}}^2}, \tag{47}$$

$$c = \frac{\mu_{\mathrm{X}}^2}{\sigma_{\mathrm{X}}^2} - 2\rho\frac{\mu_{\mathrm{X}}\mu_{\mathrm{Y}}}{\sigma_{\mathrm{X}}\sigma_{\mathrm{Y}}} + \frac{\mu_{\mathrm{Y}}^2}{\sigma_{\mathrm{Y}}^2}, \tag{48}$$

$$d(r) = \mathrm{e}^{\left(\frac{b^2(r) - ca^2(r)}{2(1-\rho^2)a^2(r)}\right)}. \tag{49}$$

This ratio PDF shows a much better fit to the sampled ratio distributions (Fig. 7) than the Gaussian model, and a corresponding better fit to the CDFs (Fig. 8). Unlike the Gaussian PDF, there is a dependence on the numerical values of the errors and Eq. 45 is not symmetrical in the ratio and the means.

For ease of use throughout the rest of this paper, the ratio PDF (Eq. 45) is referred to as $\mathcal{R}_{\mathcal{N}}(r, \mu_{\mathrm{X}}, \mu_{\mathrm{Y}}, \sigma_{\mathrm{X}}, \sigma_{\mathrm{Y}}, \rho)$. The ratio distribution has a CDF given by Eq. 50 (Hinkley, 1969):

$$\mathrm{P}(-\infty < R \leqslant r) = \Phi_{\mathcal{R}_{\mathcal{N}}}(r, \mu_{\mathrm{X}}, \mu_{\mathrm{Y}}, \sigma_{\mathrm{X}}, \sigma_{\mathrm{Y}}, \rho) = L\left(\frac{\mu_{\mathrm{X}} - \mu_{\mathrm{Y}} r}{\sigma_{\mathrm{X}}\sigma_{\mathrm{Y}} a(r)}, -\frac{\mu_{\mathrm{Y}}}{\sigma_{\mathrm{Y}}}; \frac{\sigma_{\mathrm{Y}} r - \rho\sigma_{\mathrm{X}}}{\sigma_{\mathrm{X}}\sigma_{\mathrm{Y}} a(r)}\right) + L\left(\frac{\mu_{\mathrm{Y}} r - \mu_{\mathrm{X}}}{\sigma_{\mathrm{X}}\sigma_{\mathrm{Y}} a(r)}, \frac{\mu_{\mathrm{Y}}}{\sigma_{\mathrm{Y}}}; \frac{\sigma_{\mathrm{Y}} r - \rho\sigma_{\mathrm{X}}}{\sigma_{\mathrm{X}}\sigma_{\mathrm{Y}} a(r)}\right), \tag{50}$$

where $L(u, v; \gamma)$ is the standard bivariate integral:

$$L(u, v; \gamma) = \frac{1}{2\pi\sqrt{1-\gamma^2}} \int_h^\infty \int_k^\infty \mathrm{e}^{-\frac{x^2 - 2\gamma xy + y^2}{2(1-\gamma^2)}} \mathrm{d}x\mathrm{d}y, \tag{51}$$

with zero means and covariance matrix $\mathbf{V}$:

$$\mathbf{V} = \begin{pmatrix} 1 & \gamma \\ \gamma & 1 \end{pmatrix}. \tag{52}$$

To account for the fact that the amplitude ratio used is unsigned (absolute), the PDF used is given by Eq. 53:

$$\mathrm{p}\left(R=|r|\right)=\mathcal{R}_{\mathcal{N}}\left(r,\mu_{\mathrm{X}},\mu_{\mathrm{Y}},\sigma_{\mathrm{X}},\sigma_{\mathrm{Y}},\rho\right)+\mathcal{R}_{\mathcal{N}}\left(-r,\mu_{\mathrm{X}},\mu_{\mathrm{Y}},\sigma_{\mathrm{X}},\sigma_{\mathrm{Y}}\rho\right). \tag{53}$$

Therefore, the likelihood for the absolute ratio is given by:

$$\Phi_{\mathcal{R}_{\mathcal{N}}}\left(|r|,\mu_{\mathrm{X}},\mu_{\mathrm{Y}},\sigma_{\mathrm{X}},\sigma_{\mathrm{Y}},\rho\right)=\Phi_{\mathcal{R}_{\mathcal{N}}}\left(r,\mu_{\mathrm{X}},\mu_{\mathrm{Y}},\sigma_{\mathrm{X}},\sigma_{\mathrm{Y}},\rho\right)+\Phi_{\mathcal{R}_{\mathcal{N}}}\left(-r,\mu_{\mathrm{X}},\mu_{\mathrm{Y}},\sigma_{\mathrm{X}},\sigma_{\mathrm{Y}},\rho\right)-1, \tag{54}$$

where $\mu_{\mathrm{X}},\mu_{\mathrm{Y}}\geqslant 0$. For the amplitude ratio distribution, ρ is assumed to be zero, because the observations are independent.

4.3.3 Amplitude Ratio PDF

The amplitude ratio noise model is given by the ratio PDF (Eq. 53). The ratio PDF is not symmetric in the means and the ratio, but it is straightforward to show that the distribution depends only on the ratio of the means and the percentage errors, and so the uncertainties in Eq. 53 (σ_{X} and σ_{Y}) are determined from the percentage error in the measurement:

$$\sigma=A_{\mathrm{mes}}\frac{\sigma_{\mathrm{mes}}}{A_{\mathrm{model}}}. \tag{55}$$

The likelihood PDF (Eq. 28) can be marginalised using the uncertainty scaling (Eq. 55) and the ratio noise model (Eq. 53) to give:

$$\mathrm{p}\left(R=|r|\,|A_{\mathrm{X}},A_{\mathrm{Y}},\sigma_{\mathrm{X}},\sigma_{\mathrm{Y}}\right)=\int\delta\left(r-\frac{A_{\mathrm{X}}}{A_{\mathrm{Y}}}-\Delta_{\mathrm{R}}\right)\mathrm{p}\left(\frac{A_{\mathrm{X}}}{A_{\mathrm{Y}}}+\Delta_{\mathrm{R}}|A_{\mathrm{X}},A_{\mathrm{Y}},\sigma_{\mathrm{X}},\sigma_{\mathrm{Y}}\right)\mathrm{d}\Delta_{\mathrm{R}}, \tag{56}$$

which can be evaluated using this property of the delta function:

$$\int_{-\infty}^{\infty}\delta\left(s-x\right)f\left(s\right)\mathrm{d}s=f\left(x\right). \tag{57}$$

The distribution for the ratio uncertainty is then given by:

$$\mathrm{p}\left(\frac{A_\mathrm{X}}{A_\mathrm{Y}}+\Delta_\mathrm{R}|A_\mathrm{X},A_\mathrm{Y},\sigma_\mathrm{X},\sigma_\mathrm{Y}\right)=\mathcal{R}_\mathcal{N}\left(\frac{A_\mathrm{X}}{A_\mathrm{Y}}+\Delta_\mathrm{R},A_\mathrm{X},A_\mathrm{Y},\sigma_\mathrm{X},\sigma_\mathrm{Y},0\right)+\mathcal{R}_\mathcal{N}\left(-\frac{A_\mathrm{X}}{A_\mathrm{Y}}-\Delta_\mathrm{R},A_\mathrm{X},A_\mathrm{Y},\sigma_\mathrm{X},\sigma_\mathrm{Y},0\right). \quad (58)$$

Therefore, the marginalised PDF is given by:

$$\mathrm{p}\left(R=|r|\,|A_\mathrm{X},A_\mathrm{Y},\sigma_\mathrm{X},\sigma_\mathrm{Y}\right)=\mathcal{R}_\mathcal{N}\left(r,A_\mathrm{X},A_\mathrm{Y},\sigma_\mathrm{X},\sigma_\mathrm{Y},\rho=0\right)+\mathcal{R}_\mathcal{N}\left(-r,A_\mathrm{X},A_\mathrm{Y},\sigma_\mathrm{X},\sigma_\mathrm{Y},\rho=0\right) \quad (59)$$

4.4 Source PDF Representations

Because the normalised full moment tensor source is five dimensional, it is difficult to represent the source on paper, and there are several different methods of representing the source (Hudson et al. 1989; Riedesel & Jordan 1989; Chapman & Leaney 2011; Tape & Tape 2012a). The same is true of representing the source PDF using the Hudson plot or the fundamental eigenvalue lune (Tape & Tape 2012a).

Care must be taken in interpreting the plots as well as how to plot the source PDF. Both the Hudson plot and the fundamental lune plot only show the source type, rather than any orientation information, so the PDF should first be marginalised with respect to the orientation parameters, when plotting the source PDF using these projections. The most likely source type from the marginalised PDF cannot be linked to any orientation, and it does not necessarily correspond to the maximum probability source from the full (un-marginalised) PDF.

It may be of more use to examine the silhouette of the full PDF projected onto this projection, which shows the maximum probability value from the un-marginalised PDF for each coordinate point on the plot, rather than the integral over the marginalised

parameters as in the marginalised case. This silhouette plot allows deductions to be made about the maximum probability source, as well as its orientation. A similar approach can be applied to the orientation information of a source, considering either the most likely orientation marginalised over the source-type parameters, or the orientation of the most likely source.

A comparison of the silhouette and orientation marginalised Hudson projections is shown in Fig. 9. There is a difference between the marginalised and silhouette plots, with an improved constraint for the marginalised plots. However, while the most likely source type is well constrained, the source type of the most likely moment tensor is not, although in both cases, they do broadly agree.

5 DISTINGUISHING BETWEEN MOMENT TENSOR AND DOUBLE-COUPLE SOLUTIONS

As mentioned above, it is often desirable to understand whether a source is a full moment tensor, or if deviations from the double-couple source are due to over-fitting of the source by the extra free parameters. There are several approaches to distinguishing between the full-moment tensor and double-couple source models, but the method described in this paper allows a clear understanding of the robustness of any non-double-couple component in a source.

One approach to estimating whether the result is due to over-fitting is the Bayesian Information Criterion (BIC), $\mathcal{I}^{\mathrm{B}}$, introduced by Schwarz (1978):

$$\mathcal{I}^{\mathrm{B}} = 2 \ln \mathcal{L}_{max} - k \ln n, \tag{60}$$

where $\mathcal{L}_{max}$ is the maximum likelihood for the model, k is the model dimension, and n is the number of data points used. This provides a method for comparing the double-couple and the full moment tensor models. The absolute values are not useful, but the difference

between values for the models is used as an indicator of evidence for the model with larger BIC. In the literature, differences between 2 and 6 are regarded as evidence for the larger BIC model, between 6 to 10 are considered strong evidence, and differences bigger than 10 are considered very strong evidence (Mukherjee et al. (1998); Jeffreys (1998)).

However, because the approach described in this paper samples the full source PDF using Eqs. 17 or 18, it is possible to use the Bayesian model evidence:

$$\mathrm{p}\left(data|model\right) = \int_x \mathrm{p}\left(data|x\right) \mathrm{p}\left(x|model\right) \mathrm{d}x. \tag{61}$$

This is a more complex approach with a more easily understood result, corresponding to the likelihood of the data given the model type. It requires good sampling of the un-normalised likelihoods to compare between the models, with the higher dimensional models being penalised by the parameter priors. The resultant number corresponds to the data likelihood given the model, with a higher likelihood corresponding to a better model.

From the likelihood (Eqs 17 and 18), the Bayesian model evidence can be evaluated by summing over the probability samples:

$$\mathrm{p}\left(data|model\right) = \sum_x \mathrm{p}\left(data|x\right) \mathrm{p}\left(x|model\right) \Delta x, \tag{62}$$

but care must be taken with the choice of the prior parameterisation. This must correspond to the same parameterisation in which the Monte Carlo samples were drawn, either directly or by correcting both the prior distribution and the Δx values. A Monte Carlo approach can be affected by small sample sizes in the integration, which is sometimes the case when the source PDF is dominated by a few very large probability samples.

The Bayesian evidence for the full moment tensor and double-couple models can be converted into the corresponding posterior values (p_{DC} and p_{MT}) using Bayes' theorem (Eq. 4) for a suitable prior such as the uniform prior $p = 0.5$. The posterior probabilities

can be normalised together because the source is either a full moment tensor or a double-couple source. This is similar to a test of the hypothesis that an event is non-double-couple. Consequently, the approaches for determining statistical significance levels, well known in hypothesis testing can be applied, similar to Horálek et al. (2010) using the F-test. The resultant probabilities help distinguish between whether or not a source is double-couple, providing discrimination between different types of events, and, therefore between different physical processes and source interpretations (Baig & Urbancic 2010).

6 SYNTHETIC EXAMPLES

A set of synthetic seismograms was computed from randomly generated moment tensor sources using a finite difference approach (Bernth & Chapman 2011) for a one dimensional velocity model expanded into three dimensions. Random Gaussian noise was added to the traces in varying signal-to-noise ratios, close to a targeted ratio, although real-world events often have larger variation in signal-to-noise ratio across a network. The synthetic examples are used to explore the dependency of the PDF solutions on different types of uncertainty, for both the fully constrained double-couple space and the full moment tensor space. The chosen prior for each case was the uniform prior, so no preferred source mechanism, or orientation was specified.

6.1 Processing and Inverting the Data

The synthetic data (Fig. 10) were manually picked for both P and S arrival times and P polarities, and then located using NonLinLoc (Lomax et al. 2000, 2009) and a simplified version of the known velocity model. The arrival time picks were used to automatically window and measure P-, SH- and SV-wave amplitudes. A simple Monte Carlo random search algorithm was used to produce samples of the posterior PDFs (Figs 11 - 17), calculated as in Section 4.1, in both double-couple space and the unconstrained full moment tensor space, using sample sizes in the range of $10^5 - 10^8$ for the double-couple

constrained solutions and $10^7 - 10^{10}$ for the full moment tensor solutions, running on a small computing cluster for up to an hour for the inversions with the largest sample sizes. Two sets of inversions were carried out, one using only polarity data and the second including amplitude ratios.

6.2 Network Dependency of the Solutions

The network geometry can lead to uncertainties in the source inversion, because the sampling of the focal sphere by the receivers often does not constrain the source well, leading to an increased range of possible solutions (Vavryčuk 2007; Godano et al. 2009; Kim 2011). Increasing the number of receivers does not always improve the solution, since the receivers need to improve the sampling of the focal sphere to improve the constraint. It is possible to invert for the source from a small number of waveforms, but this is not always stable and can lead to large uncertainties (Kim et al., 2000; Vavryčuk, 2007).

An event generated from a double-couple source was located and inverted for different numbers of stations and geometries (Fig. 11), with no location uncertainty included (Eq. 17 with no additional location samples, $M = 1$). The results show that the range of possible solutions increases as the constraints from the station geometry decrease, either because the number of stations decreases or the distance between them increases. However, in all cases, the addition of amplitude ratios into the solutions reduces the range of possible solutions, sharpening the PDF.

The double-couple solution for the polarity-only inversion is well constrained, with orientation similar to the source and little variation in the range of solutions. But as the number of stations decreases, the range of non-zero solutions increases, with deviation from the source. The full moment tensor solutions show a wider range of solutions as the number of receivers decreases, along with some deviation of the PDF from the double-couple point.

Including amplitude ratios reduces the dependence of the solutions on the network

distribution. The double-couple solutions are well constrained, with orientations similar to the source, and there is little variation in the range of solutions, even as the number of stations reduces. In both cases, the station geometry is more important than the number of stations, as a low station coverage case may still provide well-constrained solutions if there are sufficient constraining station positions (cf. Panza & Sarao 2000; Šílený et al. 1996). For full moment tensor solutions, the nodal regions of the radiation pattern are not necessarily planar, unlike the double-couple solution and, therefore, more stations are required to constrain the solution.

All of the solutions in Fig. 11 using polarities and amplitude ratios had high values for the posterior model probabilities (Section 5) for the full moment tensor model, despite the solutions visually resembling double-couple solutions. The PDFs are dominated by a few high-probability samples in the full moment tensor case, which can lead to uncertainties in the Monte Carlo integration approach used to calculate the Bayesian evidence, as discussed in section 5. The very large likelihoods associated with these samples may be due to mischaracterisation of the amplitude ratios and associated uncertainties as shown in Section 4.3.1. The posterior model probabilities for the inversions using only polarities are significant for the double-couple source model, suggesting that it is the more likely source model.

The results shown here are consistent with those of (Vavryčuk 2007; Godano et al. 2009; Kim 2011) showing that the position of the receivers is important for constraint of the source PDF, not just the number of receivers.

6.3 Noise Dependency of the Solutions

The effects of noise on the inversion are difficult to test, because noise has a large effect on the different steps leading up to the source inversion. Fig. 12 shows the solutions from inverting the synthetic event with 22 stations for varying levels of uncertainty in the data, corresponding to the likelihood in Eq. 17 with no additional location samples ($M = 1$).

The noise was not applied to the full workflow (arrival picking and location) but only to the polarity and amplitude ratio data used in the source inversion.

As the SNR decreases, the source solutions increase in range, showing some rotation away from the true source, and the full moment tensor PDF deviates from the double-couple point. The solutions using amplitude ratios show more uncertainty and skew as the SNR decreases. This is consistent with the results shown in Section 4.3.1, where the noise introduced a systematic deviation into the amplitude measurements. which may lead to the deviation of the full moment tensor solution from the true source. Alternatively, the polarity-only inversions show a wider range of possible solutions, with decreasing maximum probability. However, the orientation remains similar for the double-couple case, and the full moment tensor solutions show a wider range of possible solutions.

The Bayesian evidence for the solutions varied, with very low double-couple posteriors for the polarity and amplitude ratio examples with SNR $\leqslant 10$, with posterior values $p_{\mathrm{DC}} \leqslant 0.02$, despite the solutions appearing to be good double-couple solutions from a visual inspection. Similarly to the solutions in Fig. 11, these full moment tensor solutions are dominated by a few high-probability samples, which can lead to uncertainties in the Monte Carlo integration approach used to calculate the Bayesian evidence. However, the polarity-only cases have much more significant Bayesian evidence values, with the lowest value $p_{\mathrm{DC}} = 0.77$ for the SNR $= \infty$ case, and the values increasing with increasing noise level.

The noise has a larger overall effect on the workflow than that explored above, as it also affects the ability to make the arrival pick and to determine the location precisely. To investigate this, the synthetic event was manually reprocessed at different noise levels, and then inverted for the source using the likelihood in Eq. 17 with no additional location samples ($M = 1$). The noise levels were not uniform across each waveform, but the noise was added so that the average SNR was consistent with the desired SNR values. Adding

noise reduces the number of confident picks, increasing the range of possible solutions, partially due to the network distribution (Section 6.2).

As with the solution with only increased uncertainty in the inversion data (Fig. 12), the maximum probability of the full moment tensor solutions deviates from the double-couple source as the SNR decreases (Fig. 13). The lowest SNR case shows very poor constraint of the double-couple solutions with a much larger range than observed in Fig. 12. This uncertainty is much more pronounced because of the reduced number of stations. The double-couple solutions are consistent with the source, and those of Fig. 12 for the polarity-only inversion. However, there is a much wider range of possible solutions in both the double-couple and the full moment tensor solutions for the results using both polarities and amplitude ratios, as well as large deviations from the true source values. This is probably due to the increased difficulty of measuring the amplitudes as the noise level increases and the deviation of the measurement as the noise level increases, as described in Section 4.3.1.

The distribution from the full moment tensor inversion shows that adding noise increases the range of possible non-double-couple solutions, as well as moving the maximum probability full moment tensor source away from the double-couple point. This may be a possible explanation for some reported non-double-couple type sources. The effect of the noise on the solutions using amplitude ratios is much stronger than that of the station distribution, although increasing the noise level does change the station distribution as well because some arrivals cannot be seen above the background noise, leading to an often strong effect on polarity-only inversions. Consequently, at higher noise levels, the network distribution has an even greater effect, leading to more deviation from the true source as the number of stations is reduced.

A few examples show multi-modal PDFs in the full moment tensor case, especially when including amplitude ratios. Although it can be expected that the source PDF does not have to be mono-modal, some of the multi-modality may arise from the amplitude

ratio distributions, shown in Fig. 14, which demonstrates that the solutions using different amplitude ratios can often be inconsistent. So care must be taken when including them in any inversion, especially at higher noise levels.

The posterior model probabilities (Section 5) for these solutions are more consistent with the expected values. Although, again, the $SNR = \infty$ case using both polarity and amplitude-ratio observations has a very low likelihood for the double-couple model, the PDF is dominated by a few very-high-probability samples, which can lead to uncertainties in the Monte Carlo integration approach used to calculate the Bayesian evidence. However, the remaining solutions have values more consistent with a visual interpretation of the source PDFs, and all suggest that the double-couple model cannot be discarded.

6.4 Location Dependency of the Solutions

To investigate the location dependency of the recovered source PDF (using Eq. 17), the synthetic source from Section 6.2 was used to generate synthetic seismograms for different source depths. These synthetics were then inverted for a randomly generated network geometry of 39 stations at a signal-to-noise ratio of 20, while including the location uncertainty as described in section 4.1. The effect of this uncertainty on the station angles can be seen in Fig. 15. In this case, the location is well constrained for all the possible solutions and there is not much uncertainty in the source-receiver ray take-off angles and azimuths. However, were the noise level to increase, this uncertainty would also increase.

The double-couple constrained solutions mainly show a small range of possible orientations, although there is also an effect on the network distribution, with the ray-paths moving from being equatorial on the focal sphere for shallow depths to more vertical paths as the depth increases. This leads to additional uncertainties arising from clustering of stations on the focal sphere in the 7.0 km depth example, especially for the polarity-only inversion, which is very poorly constrained. The full moment tensor solu-

tions show a similar small range of solutions, and most are close to double-couple. The sources at 0.5 km and 7.0 km depths have the largest range of solutions, linked to the specific network geometries. As in previous cases, including amplitude ratios in the inversion greatly improves the source constraint.

The posterior model probabilities are all consistent with the double-couple source, apart from the solution using both polarities and amplitude ratios at 1.0 km depth: This again is not well constrained due to the domination of the full moment tensor PDF by a few very-high-probability samples.

It is clear that the location uncertainty for well-constrained events with low noise has a small effect, although, as the noise level increases, so will the location uncertainty and its effects, as can be seen in the examples shown in Section 7. Fig. 16 shows that there is little difference between the solutions using location uncertainty and just using the maximum likelihood location, and there is only a slight change to the posterior model probabilities. However, if the location is not well-constrained, the resultant effect on the source PDF can be much larger.

The results shown here are consistent with studies of the effects of location uncertainty on source inversion, especially for full-waveform based inversions, such as Duputel et al. (2012a) and Šílený (2009), who show that in some cases, location uncertainty can cause large variations in the solution.

6.5 Velocity Model Dependency of the Solutions

The synthetic event from Section 6.4 at a source depth of 1.5 km was inverted for a range of velocity models, randomly perturbed from the true model (Fig. 17) at two perturbation levels, while including the location uncertainty using Eq. 18. This location uncertainty was then sampled to create a location distribution for the range of velocity models for each perturbation level. Variation in the model leads to increased uncertainty in the source-receiver ray-path angles (Fig. 17), especially for the larger velocity model perturbation, in

contrast to the example in Fig. 15. This increased uncertainty causes a slightly increased range of possible orientations for the double-couple constrained solution, although the maximum probability solution is still close to the true source. The full moment tensor solution shows a small increase in the range of possible source types compared to just the location uncertainty, although there is little effect due to the low noise level. As a consequence, velocity model uncertainty does affect both the resultant location, and the source PDF, although if the location remains well constrained, the effects are less than those of the network distribution and background noise.

The posterior model probability for the double-couple source model using only polarity data is 0.86 and 0.76 for the 3% and 10% perturbations, respectively. Including amplitude ratios leads to $p_{\mathrm{DC}} = 1$ in both cases.

For these examples, a simple one-dimensional velocity model was used, and it was assumed to be consistent for all stations. It is possible using Eq. 18 to allow the velocity models to vary independently for each station and to combine these, although this would increase the number of location samples required in the inversion. Extending the velocity model uncertainties to full three-dimensional heterogeneous models may make the required number of models impossibly large to include using this Monte Carlo approach.

Velocity model uncertainties have been shown in the literature to lead to non-double-couple components (Šílený 2004; Šílený & Milev 2006; Vavryčuk 2007; Šílený 2009), so including the velocity model uncertainty can be important when trying to estimate the source type. If the model is well-constrained, the uncertainty on the source PDF is minimal. Topography, site effects and near surface structure can have additional affects on both the source inversion (Šílený et al. 2001), as well as the location of the event (Bean et al. 2008; O'Brien et al. 2011).

7 KRAFLA EVENTS

The inversion approach was used on four local seismic earthquakes from the Krafla geothermal region of Iceland. These events were at shallow depths ranging from 0.9 km to 1.6 km and were part of a larger group detected on a temporary network between 2009 and 2010 (Fig. 18). The inversion was carried out using both the constrained double-couple and unconstrained full moment tensor source models, using P polarities only as both the P/SH and P/SV amplitude ratios were difficult to measure, with very large uncertainties. The location uncertainties from NonLinLoc were included in the inversion (Eq. 17).

Fig. 19 shows both the range of double-couple solutions and the unconstrained full moment tensor inversions for these events. The double-couple results for the events B-D have large probability ranges and, even though the full moment tensor solutions do not peak at the double-couple point, they still have a significant value there. These solutions resemble those for the noise uncertainty (Fig. 13), especially the u and v distributions.

Event A is clearly non-double-couple and, although it is possible to fit some double-couple solutions to this source, these have very low probabilities, with very significant misfit from the observed data. The full moment tensor solutions show a small sharp peak near the closing tensile crack point (see Fig. 9(a)), with a large region of very-low-probability solutions reaching the double-couple point. The full moment tensor source PDF for event A is clearly different from the other three solutions, which are described well by a double-couple source type, when the effects shown in Fig. 13 are considered.

Events C and D appear to have a bi-modality in the full moment tensor solutions, but this is partly due to the conversion of the moment tensor into the Hudson u and v coordinates (Tape & Tape (2012b)).

The posterior model probabilities for the double-couple source model are consistent with this interpretation with a very low value (0.002) for Event A and values bigger than

0.7 for the others (Table 3). We therefore conclude that Event A is a non-double-couple closing crack source, but Events B-D can be described best by a double-couple failure.

Table 3 also shows the evaluation of the BIC values (Eq. 60) for the double-couple model, which suggest that there is strong evidence for the double-couple model for events B-D and very strong evidence for a non-double-couple model for event A. However, the values of the BIC, and the associated differences are not as informative as the posterior values from the Bayesian evidence, which provide more of a quantitative estimate of the support for the different model types.

These specific results are discussed in more depth in Mildon et al. (2016), but these results show that non-double-couple (closing crack) events can be detected in close proximity to double-couple events, with the complex mechanism arising in a geothermal area with complex fluid motions (Gudmundsson & Arnórsson 2002) and a deflating magma chamber, and pockets of melt hypothesised to be close to the earthquake locations (Schuler et al. 2015; Mildon et al. 2016).

8 SUMMARY AND DISCUSSION

This paper presents a method for source inversion based on a Bayesian approach. This approach includes the uncertainties in the inversion in a rigorous framework, allowing for expansion of data types used, along with improved uncertainty determination. The method also allows the inclusion of location and model uncertainties. However, because the effects of these uncertainties on the station propagation coefficients are not independent, methods of generating the appropriate PDFs such as that proposed in Section 3 must be explored. The use of prior distributions in the inversion can be extended to determine whether any apparent non-double-couple component is introduced as an artefact due to data uncertainty and the larger number of free parameters available for fitting compared to the double-couple solution.

When tested on synthetic events with a known source, the results are consistent and

provide a useful approximation of the source PDF, with the estimate including the effects of the known uncertainties. The inversions of the synthetic events agree closely with the source mechanisms, with greater uncertainty arising as the noise level is increased and the number of stations reduced. Neither the velocity nor location uncertainty preclude the recovery of the source mechanism, although increasing the noise level is likely to increase the effects of these uncertainties.

Full-moment tensor inversion of the synthetic events showed deviations in the PDF maxima from the true double-couple source, which is a possible explanation for deviatoric non-double-couple source observations. Including amplitude ratios improves the constraint of the source PDF, but the systematic deviation of the amplitude measurements due to noise can lead to deviations in the resultant PDF from the true value, so care must be taken when including them in an inversion. Furthermore the low SNR (high noise) cases of Fig 13 show that source PDFs can be multimodal, and so maximum probability approaches are unlikely to produce a good description of such a solution. Real-world amplitude ratio measurements can have additional complications such as site and near-surface effects, S-wave splitting and anisotropy and non-Gaussian noise, which can make inclusion of the amplitude ratios difficult.

The inversions of four events from the Krafla region of Iceland have double-couple solutions for Events B-D and a non-double-couple solution for Event A, supported by the posterior model probabilities from the Bayesian evidence. These values agree with the corresponding BIC values, although the posterior probabilities provide a better qualitative estimate of the support for the different model types. Both the posterior model probabilities and the BIC values support a double-couple source model for events B-D despite the full moment tensor PDF having a strong peak away from the double-couple point, suggesting that these are good estimators of the model type. The posterior model probabilities for the double-couple model have the lowest value for Event B, which is consistent with visual interpretation of the observed source PDFs, unlike the values for the

BIC, which has a smaller value for Events B and D. The non-double-couple Event A has a very sharp peak in the closing tensile crack source region, with a very low probability of fitting a double-couple solution.

The solutions generated using this approach suggest that a source inversion should examine the full source PDF, which allows a clear interpretation of the possible source types, and accounts for uncertainties as well as any multi-modality. All the explored uncertainty types affect the source PDF, although the dominant effect is the noise. The noise also affects the other types of uncertainties, amplifying the effects.

The posterior model probabilities provide a useful quantitative estimate of whether the source is non-double-couple, although they can be very uncertain if the source PDF is dominated by a few very-high-probability samples. This is probably due to uncertainties from the small sample size in the Monte Carlo integration approach used to calculate the Bayesian evidence. However, for more realistic PDFs with larger uncertainties, the values correspond well to a visual interpretation of the source PDF, suggesting that it is a good approach for distinguishing between double-couple and non-double-couple sources.

ACKNOWLEDGEMENTS

This work was funded under a Natural Environment Research Council (NERC) studentship as a CASE award with Schlumberger. Seismometers were borrowed from the NERC SEIS-UK (loan 842). We thank Mike Williams for his help and advice, as well as Zoe Watson and Jon Tarasewicz for processing and picking the Krafla data. We would also like to thank our reviewers, Andrew Valentine and one anonymous reviewer for helping improve this manuscript. Department of Earth Sciences, Cambridge contribution number ESC3602

References

Aki, K. & Richards, P. G., 2002. *Quantitative Seismology*, University Science Books, 2nd edn.

Baig, A. & Urbancic, T. I., 2010. Microseismic moment tensors: A path to understanding frac growth, *The Leading Edge*, **29**(3), 320–324.

Bayes, T. & Price, R., 1763. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S., *Philosophical Transactions of the Royal Society of London*, **53**, 370–418.

Bean, C. J., Lokmer, I., & O'Brien, G. S., 2008. Influence of near-surface volcanic structure on long-period seismic signals and on moment tensor inversions: Simulated examples from Mount Etna, *Journal of Geophysical Research: Solid Earth*, **113**(B08308).

Bernth, H. & Chapman, C., 2011. A comparison of the dispersion relations for anisotropic elastodynamic finite-difference grids, *Geophysics*, **76**(3), WA43–WA50.

Brillinger, D. R., Udias, A., & Bolt, B. A., 1980. A probability model for regional focal mechanism solutions, *Bulletin of the Seismological Society of America*, **70**(1), 149–170.

Chapman, C. H., 2004. *Fundamentals of Seismic Wave Propagation*, Cambridge University Press.

Chapman, C. H. & Leaney, W. S., 2011. A new moment-tensor decomposition for seismic events in anisotropic media, *Geophysical Journal International*, **188**(1), 343–370.

Cover, T. M. & Thomas, J. A., 2005. *Elements of Information Theory*, Wiley, 2nd edn.

Dahm, T., 1996. Relative moment tensor inversion based on ray theory: theory and synthetic tests, *Geophysical Journal International*, **124**(1), 245–257.

Delouis, B., Giardini, D., Lundgren, P., & Salichon, J., 2002. Joint inversion of In SAR, GPS, teleseismic, and strong-motion data for the spatial and temporal distribution of earthquake slip: application to the 1999 Izmit mainshock, *Bulletin of the Seismological Society of America*, **92**(1), 278–299.

Duputel, Z., Rivera, L., Fukahata, Y., & Kanamori, H., 2012a. Uncertainty estimations for seismic source inversions, *Geophysical Journal International*, **190**(2), 1243–1256.

Duputel, Z., Rivera, L., Kanamori, H., & Hayes, G., 2012b. W phase source inversion for moderate to large earthquakes (1990-2010), *Geophysical Journal International*, **189**(2), 1125–1147.

Dziewonski, A. M., Chou, T. A., & Woodhouse, J. H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *Journal of Geophysical Research*, **86**(B4), 2825–2852.

Einarsson, P. & Sæmundsson, K., 1987. Earthquake epicenters 1982-1985 and volcanic systems in Iceland, in *Í hlutarins edli, Festscript for Th. Sigurgeirsson*, ed. Sigfússon, T. I., Menningarsjódur.

Ekström, G., Nettles, M., & Dziewonski, A. M., 2012. The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes, *Physics of the Earth and Planetary Interiors*, **200-201**.

Fieller, E. C., 1932. The distribution of the index in a normal bivariate population, *Biometrika*, **24**(3-4), 428–440.

Ford, S. R., Dreger, D. S., & Walter, W. R., 2008. Identifying isotropic events using a regional moment tensor inversion, in *Proceedings of the 30th Monitoring Research Review: Ground-Based Nuclear Explosion Monitoring Technologies, 23-25 Sep 2008, Portsmouth, VA sponsored by the National Nuclear Security Administration (NNSA) and the Air Force Research Laboratory (AFRL).*

Foulger, G. R., Julian, B. R., Hill, D. P., Pitt, A. M., Malin, P. E., & Shalev, E., 2004. Non-double-couple microearthquakes at Long Valley caldera, California, provide evidence for hydraulic fracturing, *Journal of Volcanology and Geothermal Research*, **132**(1), 45–71.

Godano, M., Regnier, M., Deschamps, A., Bardainne, T., & Gaucher, E., 2009. Focal mechanisms from sparse observations by nonlinear inversion of amplitudes: method

and tests on synthetic and real data, *Bulletin of the Seismological Society of America*, **99**(4), 2243–2264.

Gudmundsson, B. T. & Arnórsson, S., 2002. Geochemical monitoring of the Krafla and Námafjall geothermal areas, N-Iceland, *Geothermics*, **31**(2), 195–243.

Hardebeck, J. L. & Shearer, P. M., 2002. A new method for determining first-motion focal mechanisms, *Bulletin of the Seismological Society of America*, **92**(6), 2264–2276.

Hardebeck, J. L. & Shearer, P. M., 2003. Using S / P amplitude ratios to constrain the focal mechanisms of small earthquakes, *Bulletin of the Seismological Society of America*, **93**(6), 2434–2444.

Heimann, S., 2011. *A robust method to estimate kinematic earthquake source parameters*, Ph.D. thesis, Hamburg.

Hinkley, D. V., 1969. On the ratio of two correlated normal random variables, *Biometrika*, **56**(3), 635–639.

Horálek, J., Jechumtálová, Z., Dorbath, L., & Šílený, J., 2010. Source mechanisms of micro-earthquakes induced in a fluid injection experiment at the HDR site Soultz-sous-Forêts (Alsace) in 2003 and their temporal and spatial variations, *Geophysical Journal International*, **181**, 1547–1565.

Houlié, N., Dreger, D., & Kim, A., 2014. GPS source solution of the 2004 Parkfield earthquake., *Scientific reports*, **4**(3646).

Hudson, J. A., Pearce, R. G., & Rogers, R. M., 1989. Source type plot for inversion of the moment tensor, *Journal of Geophysical Research*, **94**(B1), 765–774.

Jeffreys, H., 1998. *Theory of probability*, Oxford Univ Press, 3rd edn.

Kaeufl, P., Valentine, A. P., O'Toole, T., & Trampert, J., 2013. A framework for fast probabilistic centroid - moment-tensor determination - Inversion of regional static displacement measurements, *Geophysical Journal International*, **196**(3), 1676–1693.

Kennet, B. L. N., Marson-Pidgeon, K., & Sambridge, M. S., 2000. Seismic Source Characterization using a Neighbourhood Algorithm, *Geophysical Research Letters*, **27**(20),

3401–3404.

Kim, A., 2011. Uncertainties in full waveform moment tensor inversion due to limited microseismic monitoring array geometry, *SEG San Antonio 2011 Annual Meeting*, (5), 1509–1513.

Kim, S. G., Kraeva, N., & Chen, Y. T., 2000. Source parameter determination of regional earthquakes in the Far East using moment tensor inversion of single-station data, *Tectonophysics*, **317**, 125–136.

Knopoff, L. & Randall, M. J., 1970. The compensated linear-vector dipole: a possible mechanism for deep earthquakes, *Journal of Geophysical Research*, **75**(26), 4957–4963.

Konca, A. O., Leprince, S., Avouac, J.-P., & Helmberger, D. V., 2010. Rupture process of the 1999 Mw 7.1 Duzce earthquake from joint analysis of SPOT, GPS, InSAR, strong-motion, and teleseismic data: a supershear rupture with variable rupture velocity, *Bulletin of the Seismological Society of America*, **100**(1), 267–288.

Kuge, K. & Lay, T., 1994. Systematic non-double-couple components of earthquake mechanisms: The role of fault zone irregularity, *Journal of Geophysical Research*, **99**(B8), 15457–15467.

Laplace, P. S., 1812. *Théorie analytique des probabilités*.

Lomax, A., Virieux, J., Volant, P., & Berge, C., 2000. *Probabilistic earthquake location in 3D and layered models: Introduction of a Metropolis-Gibbs method and comparison with linear locations*, pp. 101–134, Advances in Seismic Location, Kluwer.

Lomax, A., Michelini, A., & Curtis, A., 2009. *Earthquake Location, Direct, Global-Search Methods*, pp. 2449–2473, Encyclopedia of Complexity and System Science, Part 5, Springer.

Lundgren, P. & Salichon, J., 2000. Joint inversion of InSAR and teleseismic data for the slip history of the 1999 Izmit (Turkey) earthquake Izmit / Sapanca, *Geophysical Research Letters*, **27**(20), 3389–3392.

Mildon, Z. K., Pugh, D. J., Tarasewicz, J., White, R. S., & Brandsdóttir, B., 2016. Clos-

ing crack earthquakes within the Krafla caldera, North Iceland, *Geophysical Journal International*, (submitted).

Minson, S. E., Simons, M., & Beck, J. L., 2013. Bayesian inversion for finite fault earthquake source models I-theory and algorithm, *Geophysical Journal International*, **194**(3), 1701–1726.

Mukherjee, S., Feigelson, E. D. F., Babu, G. J., Murtagh, F., Fraley, C., & Raftery, A., 1998. Three types of gamma-ray bursts, *The Astrophysical Journal*, **508**, 314–327.

Müller, G., 1973. Seismic moment and long-period radiation of underground nuclear explosions, *Bulletin of the Seismological Society of America*, **63**(3), 847–857.

O'Brien, G. S., Lokmer, I., De Barros, L., Bean, C. J., Saccorotti, G., Metaxian, J. P., & Patane, D., 2011. Time reverse location of seismic long-period events recorded on Mt Etna, *Geophysical Journal International*, **184**(1), 452–462.

O'Toole, T. B., 2013. *Studies of earthquakes and microearthquakes using near-field seismic and geodetic observations*, Ph.D. thesis, University of Oxford.

Page, M. T., Cust, S., Archuleta, R. J., & Carlson, J. M., 2009. Constraining earthquake source inversions with GPS data 1: resolution based removal of artifacts, *Journal of Geophysical Research*, **114**(B01314).

Panza, G. F. & Sarao, A., 2000. Monitoring volcanic and geothermal areas by full seismic moment tensor inversion: are non-double-couple components always artefacts of modelling?, *Geophysical Journal International*, **143**(2), 353–364.

Pugh, D.J., White, R.S. & Christie, P.A.F., 2016. Automatic Bayesian polarity determination, *Geophysical Journal International*, (submitted).

Reasenberg, P. A. & Oppenheimer, D., 1985. FPFIT, FPPLOT and FPPAGE: Fortran computer programs for calculating and displaying earthquake fault-plane solutions - OFR 85-739, Tech. rep., USGS.

Riedesel, M. A. & Jordan, T. H., 1989. Display and assessment of seismic moment tensors, *Bulletin of the Seismological Society of America*, **79**(1), 85–100.

Schuler, J., Greenfield, T., White, R. S., Roecker, S. W., Brandsdóttir, B., Stock, J. M., Tarasewicz, J., Martens, H., & Pugh, D., 2015. Seismic imaging of the shallow crust beneath the Krafla central volcano , NE Iceland, *Journal of Geophysical Research: Solid Earth*.

Schwarz, G., 1978. Estimating the dimension of a model, *The Annals of Statistics*, **6**(2), 461–464.

Sivia, D. S., 2000. *Data Analysis: A Bayesian Tutorial*, Oxford Univ Press.

Snoke, J. A., 2003. FOCMEC: FOCal MEChanism determinations, Tech. rep.

Tape, W. & Tape, C., 2012a. A geometric setting for moment tensors, *Geophysical Journal International*, **190**(1), 476–498.

Tape, W. & Tape, C., 2012b. A geometric comparison of source-type plots for moment tensors, *Geophysical Journal International*, **190**(1), 499–510.

Templeton, D. C. & Dreger, D. S., 2006. Non-double-couple earthquakes in the Long Valley volcanic region, *Bulletin of the Seismological Society of America*, **96**(1), 69–79.

Vasco, D. W., 1990. Moment-tensor invariants: searching for non-double- couple earthquakes, *Bulletin of the Seismological Society of America*, **80**(2), 354–371.

Vavryčuk, V., 2007. On the retrieval of moment tensors from borehole data, *Geophysical Prospecting*, **55**(3), 381–391.

Vavryčuk, V., Bohnhoff, M., Jechumtálová, Z., Kolar, P., & Šílený, J., 2008. Non-double-couple mechanisms of microearthquakes induced during the 2000 injection experiment at the KTB site, Germany: A result of tensile faulting or anisotropy of a rock?, *Tectonophysics*, **456**(1-2), 74–93.

Voigt, W., 1910. *Lehrbuch der Kristallphysik*, B.G. Teubner.

Šílený, J., 2004. Regional moment tensor uncertainty due to mismodeling of the crust, *Tectonophysics*, **383**(3-4), 133–147.

Šílený, J., 2009. Resolution of non-double-couple mechanisms: simulation of hypocenter mislocation and velocity structure mismodeling, *Bulletin of the Seismological Society*

of America, **99**(4), 2265–2272.

Šílený, J. & Milev, A., 2006. Seismic moment tensor resolution on a local scale: simulated rockburst and mine-induced seismic events in the Kopanang gold mine, South Africa, *Pure and Applied Geophysics*, **163**(8), 1495–1513.

Šílený, J., Campus, P., & Panza, G. F., 1996. Seismic moment tensor resolution by waveform inversion of a few local noisy records - I. Synthetic tests, *Geophysical Journal International*, **126**, 605–619.

Šílený, J., Pšenčík, I., & Young, R. P., 2001. Point-source inversion neglecting a nearby free surface: Simulation of the underground Research Laboratory, Canada, *Geophysical Journal International*, **146**(1), 171–180.

Wadati, K., 1933. On the travel time of earthquake waves. (Part II), *The Geophysical Magazine*, **VII**, 101–111.

Wadati, K. & Oki, S., 1933. On the travel time of earthquake waves. (Part III), *The Geophysical Magazine*, **VII**, 113–137.

Walsh, D., Arnold, R., & Townend, J., 2009. A Bayesian approach to determining and parametrizing earthquake focal mechanisms, *Geophysical Journal International*, **176**(1), 235–255.

Wéber, Z., 2006. Probabilistic local waveform inversion for moment tensor and hypocentral location, *Geophysical Journal International*, **165**(2), 607–621.

Weston, J., Ferreira, A. M. G., & Funning, G. J., 2014. Joint earthquake source inversions using seismo-geodesy and 3-D earth models, *Geophysical Journal International*, **198**(2), 671–696.

Yokota, Y., Kawazoe, Y., Yun, S., Oki, S., Aoki, Y., & Koketsu, K., 2012. Joint inversion of teleseismic and InSAR datasets for the rupture process of the 2010 Yushu, China, earthquake, *Earth, Planets and Space*, **64**(11), 1047–1051.

Zollo, A. & Bernard, P., 2007. Fault mechanisms from near-source data: joint inversion of S polarizations and P polarities, *Geophysical Journal International*, **104**(3), 441–451.

A STATION PROPAGATION COEFFICIENTS

In the case of a homogeneous isotropic velocity structure, the Green functions are known for the different components. In other cases, it is necessary to determine the ray angles between the source and the receiver for an appropriate geological structure.

Considering a system of orthogonal basis vectors on the sphere with ϕ the angle from the x axis and θ the angle from the z axis gives Γ the radial unit vector, Θ the unit vector along lines of longitude, and Φ the equatorial unit vector:

$$\mathbf{\Gamma} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}, \tag{A.1}$$

$$\mathbf{\Theta} = \begin{pmatrix} \cos\theta\cos\phi \\ \cos\theta\sin\phi \\ \sin\theta \end{pmatrix}, \tag{A.2}$$

$$\mathbf{\Phi} = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix}. \tag{A.3}$$

For a point with a take-off angle θ and azimuth ϕ, the displacement components are (following Aki & Richards 2002):

$$u_{\mathrm{P}} = \frac{1}{4\pi\rho\alpha^3 r}\left(\mathbf{\Gamma}^{\mathsf{T}}\mathbf{M}\mathbf{\Gamma}\right)\mathbf{\Gamma} = \mathcal{F}_{\mathrm{P}}\left(\mathbf{\Gamma}^{\mathsf{T}}\mathbf{M}\mathbf{\Gamma}\right)\mathbf{\Gamma}, \tag{A.4}$$

$$u_{\mathrm{SV}} = \frac{1}{4\pi\rho\beta^3 r}\left(\mathbf{\Theta}^{\mathsf{T}}\mathbf{M}\mathbf{\Gamma}\right)\mathbf{\Theta} = \mathcal{F}_{\mathrm{SV}}\left(\mathbf{\Theta}^{\mathsf{T}}\mathbf{M}\mathbf{\Gamma}\right)\mathbf{\Theta}, \tag{A.5}$$

$$u_{\mathrm{SH}} = \frac{1}{4\pi\rho\beta^3 r}\left(\mathbf{\Phi}^{\mathsf{T}}\mathbf{M}\mathbf{\Gamma}\right)\mathbf{\Phi} = \mathcal{F}_{\mathrm{SH}}\left(\mathbf{\Phi}^{\mathsf{T}}\mathbf{M}\mathbf{\Gamma}\right)\mathbf{\Phi}, \tag{A.6}$$

where $\mathcal{F}$ is the propagation effect, containing the effects from geometric spreading and the velocity structure.

The radiation components (Chapman 2004, Chapter 4) can be given by:

$$\mathcal{R}_{\mathrm{P}} = \mathbf{\Gamma}^{\mathsf{T}}\,\mathbf{M}\,\mathbf{\Gamma}, \tag{A.7}$$

$$\mathcal{R}_{\mathrm{SV}} = \mathbf{\Theta}^{\mathsf{T}}\,\mathbf{M}\,\mathbf{\Gamma}, \tag{A.8}$$

$$\mathcal{R}_{\mathrm{SH}} = \mathbf{\Phi}^{\mathsf{T}}\,\mathbf{M}\,\mathbf{\Gamma}. \tag{A.9}$$

The radiation components can be used to determine the coefficients of the different moment tensor components. This allows a calculation of the amplitude for the different components at the source, which can be used to help determine the moment tensor. The amplitude of the P component at the source is given by:

$$\mathcal{R}_{\mathrm{P}} = \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{12} & M_{22} & M_{23} \\ M_{13} & M_{23} & M_{33} \end{pmatrix} \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}, \tag{A.10}$$

$$\mathcal{R}_{\mathrm{P}} = M_{11}\sin^2\theta\cos^2\phi + M_{22}\sin^2\theta\sin^2\phi + M_{33}\cos^2\theta + M_{12}\sin^2\theta\sin 2\phi + M_{13}\sin 2\theta\cos\phi + M_{23}\sin 2\theta\sin\phi. \tag{A.11}$$

Similarly for the SH and SV components:

$$\mathcal{R}_{\mathrm{SV}} = \frac{1}{2}M_{11}\sin 2\theta\cos^2\phi + \frac{1}{2}M_{22}\sin 2\theta\sin^2\phi - \frac{1}{2}M_{33}\sin 2\theta + M_{12}\sin 2\theta\cos\phi\sin\phi + M_{13}\cos 2\theta\cos\phi + M_{23}\cos 2\theta\sin\phi, \tag{A.12}$$

$$\mathcal{R}_{\mathrm{SH}} = -\frac{1}{2}M_{11}\sin\theta\sin 2\phi + \frac{1}{2}M_{22}\sin\theta\sin 2\phi + M_{12}\sin\theta\cos 2\phi - M_{13}\cos\theta\sin\phi + M_{23}\cos\theta\cos\phi. \quad \text{(A.13)}$$

These coefficients describe the amplitude of the different components at the source. However, propagation effects need to be accounted for when extending this further to a measurement at a receiver. The coefficients produced here are the same as those reported by Dahm (1996).

These amplitude coefficients can be written in terms of the moment tensor six-vector (Eq. 1):

$$\begin{pmatrix} \mathcal{R}_{\mathrm{P}} \\ \mathcal{R}_{\mathrm{SV}} \\ \mathcal{R}_{\mathrm{SH}} \end{pmatrix} = \begin{pmatrix} \sin^2\theta\cos^2\phi & 1/2\sin 2\theta\cos^2\phi & -\frac{1}{2}\sin\theta\sin 2\phi \\ \sin^2\theta\sin^2\phi & 1/2\sin 2\theta\sin^2\phi & \frac{1}{2}\sin\theta\sin 2\phi \\ \cos^2\theta & -1/2\sin 2\theta & 0 \\ 1/\sqrt{2}\sin 2\theta\sin\phi & 1/\sqrt{2}\cos 2\theta\sin\phi & 1/\sqrt{2}\cos\theta\cos\phi \\ 1/\sqrt{2}\sin 2\theta\cos\phi & 1/\sqrt{2}\cos 2\theta\cos\phi & -1/\sqrt{2}\cos\theta\sin\phi \\ 1/\sqrt{2}\sin^2\theta\sin 2\phi & 1/\sqrt{2}\sin 2\theta\cos\phi\sin\phi & 1/\sqrt{2}\sin\theta\cos 2\phi \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} M_{11} \\ M_{22} \\ M_{33} \\ \sqrt{2}M_{23} \\ \sqrt{2}M_{13} \\ \sqrt{2}M_{12} \end{pmatrix}. \quad \text{(A.14)}$$

Consequently the station propagation coefficients are given by:

$$\mathbf{a}_{\mathrm{P}} = \begin{pmatrix} \sin^2\theta\cos^2\phi \\ \sin^2\theta\sin^2\phi \\ \cos^2\theta \\ 1/\sqrt{2}\sin 2\theta\sin\phi \\ 1/\sqrt{2}\sin 2\theta\cos\phi \\ 1/\sqrt{2}\sin^2\theta\sin 2\phi \end{pmatrix}^{\mathsf{T}}, \quad \text{(A.15)}$$

$$\mathbf{a}_{\mathrm{SV}} = \begin{pmatrix} 1/2 \sin 2\theta \cos^2 \phi \\ 1/2 \sin 2\theta \sin^2 \phi \\ -1/2 \sin 2\theta \\ 1/\sqrt{2} \cos 2\theta \sin \phi \\ 1/\sqrt{2} \cos 2\theta \cos \phi \\ 1/\sqrt{2} \sin 2\theta \cos \phi \sin \phi \end{pmatrix}^{\mathsf{T}}, \tag{A.16}$$

$$\mathbf{a}_{\mathrm{SH}} = \begin{pmatrix} -\frac{1}{2} \sin \theta \sin 2\phi \\ \frac{1}{2} \sin \theta \sin 2\phi \\ 0 \\ 1/\sqrt{2} \cos \theta \cos \phi \\ -1/\sqrt{2} \cos \theta \sin \phi \\ 1/\sqrt{2} \sin \theta \cos 2\phi \end{pmatrix}^{\mathsf{T}}. \tag{A.17}$$

B MAXIMUM ENTROPY DISTRIBUTION FOR MEAN AND VARIANCE

Consider a function $g(x)$ with mean μ and variance σ^2. The entropy is given by (e.g. Cover & Thomas 2005):

$$S = \int_{-\infty}^{\infty} g(x) \ln(g(x)) \, \mathrm{d}x, \tag{B.1}$$

with the maximum entropy at the stationary point of the functional:

$$F\left(x, g(x), g'(x)\right) dx = g(x) \ln(g(x)). \tag{B.2}$$

There are two constraints to apply: $g(x)$ is normalised:

$$\int_{-\infty}^{\infty} G\left(x, g(x), g'(x)\right) \mathrm{d}x = \int_{-\infty}^{\infty} g(x) \, \mathrm{d}x = 1, \tag{B.3}$$

and the variance is given by:

$$\int_{-\infty}^{\infty} H\left(x, g(x), g'(x)\right) \mathrm{d}x = \int_{-\infty}^{\infty} (x-\mu)^2 \, g(x) \, \mathrm{d}x = \sigma^2. \tag{B.4}$$

The constrained stationary value can be found by maximising the Euler-Lagrange functional:

$$L\left(x, g(x), g'(x)\right) = \int_{-\infty}^{\infty} F + \lambda G + \nu H, \tag{B.5}$$

$$\frac{\partial F}{\partial g} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial g'} + \lambda\left(\frac{\partial G}{\partial g} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial G}{\partial g'}\right) + \nu\left(\frac{\partial H}{\partial g} - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial H}{\partial g'}\right) = 0. \tag{B.6}$$

where λ and ν are Lagrange undetermined multipliers. The Euler-Lagrange equation for this functional is given by:

$$\ln\left(g(x)\right) + 1 + \lambda + \nu(x-\mu)^2 = 0, \tag{B.7}$$

because F, G, and H are independent of $g'(x)$. This can be solved to give an expression for $g(x)$:

$$g(x) = \mathrm{e}^{-\nu(x-\mu)^2 - \lambda - 1}. \tag{B.8}$$

Applying the constraints on $g\,(x)$ (Eqs B.3 and B.4) gives a relationship for λ and ν:

$$\mathrm{e}^{-\lambda-1}\int_{-\infty}^{\infty}\mathrm{e}^{-\nu(x-\mu)^2}\mathrm{d}x = 1,$$

$$\therefore \mathrm{e}^{-\lambda-1} = \sqrt{\frac{\nu}{\pi}}, \tag{B.9}$$

$$e\mathrm{e}^{-\lambda-1}\int_{-\infty}^{\infty}(x-\mu)^2\,\mathrm{e}^{-\nu(x-\mu)^2}\mathrm{d}x = \sigma^2,$$

$$\therefore \frac{\mathrm{e}^{-\lambda-1}}{2}\sqrt{\frac{\pi}{\nu^3}} = \sigma^2, \tag{B.10}$$

$$\nu = \frac{1}{2\sigma^2}, \tag{B.11}$$

$$\mathrm{e}^{-\lambda-1} = \frac{1}{\sqrt{2\pi\sigma^2}}. \tag{B.12}$$

Substituting for these into Eq. B.8 gives a Gaussian distribution for the maximum entropy distribution for a mean and variance:

$$g\,(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{\frac{-(x-\mu)^2}{2\sigma^2}}. \tag{B.13}$$

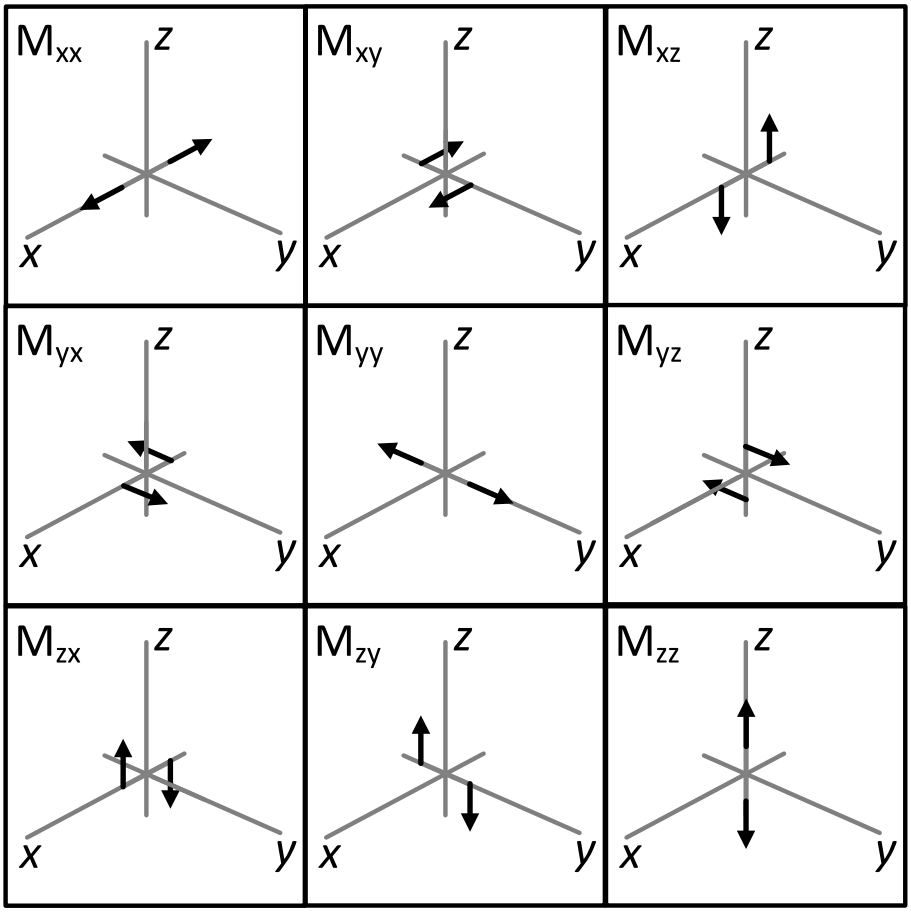

Figure 1. Force couple description of a moment tensor following Aki & Richards (2002, p51). Force couples are used so that linear momentum is conserved. To conserve angular momentum, it is necessary to constrain the tensor to be symmetric.

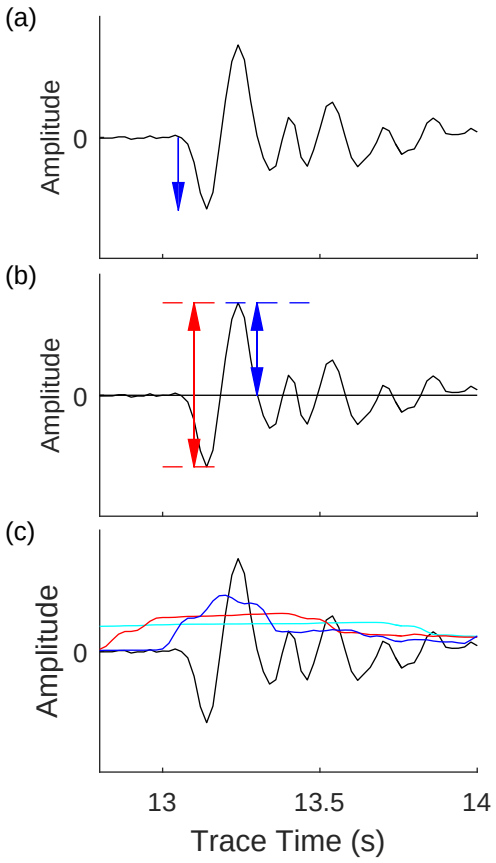

Figure 2. Examples of the different measurements used in section 4. The seismogram (black) examples are a synthetic waveform example. In (a), the blue arrow shows the direction of the first motion polarity of the arrival. (b) shows the peak-to-peak (red arrow) and maximum value (blue arrow) amplitude measurements. (c) shows the windowed root-mean-squared value over 3 different central windows: 0.1 second (blue line), 0.3 seconds (red line) and 0.6 seconds (cyan line), overlain on the seismogram (black line).

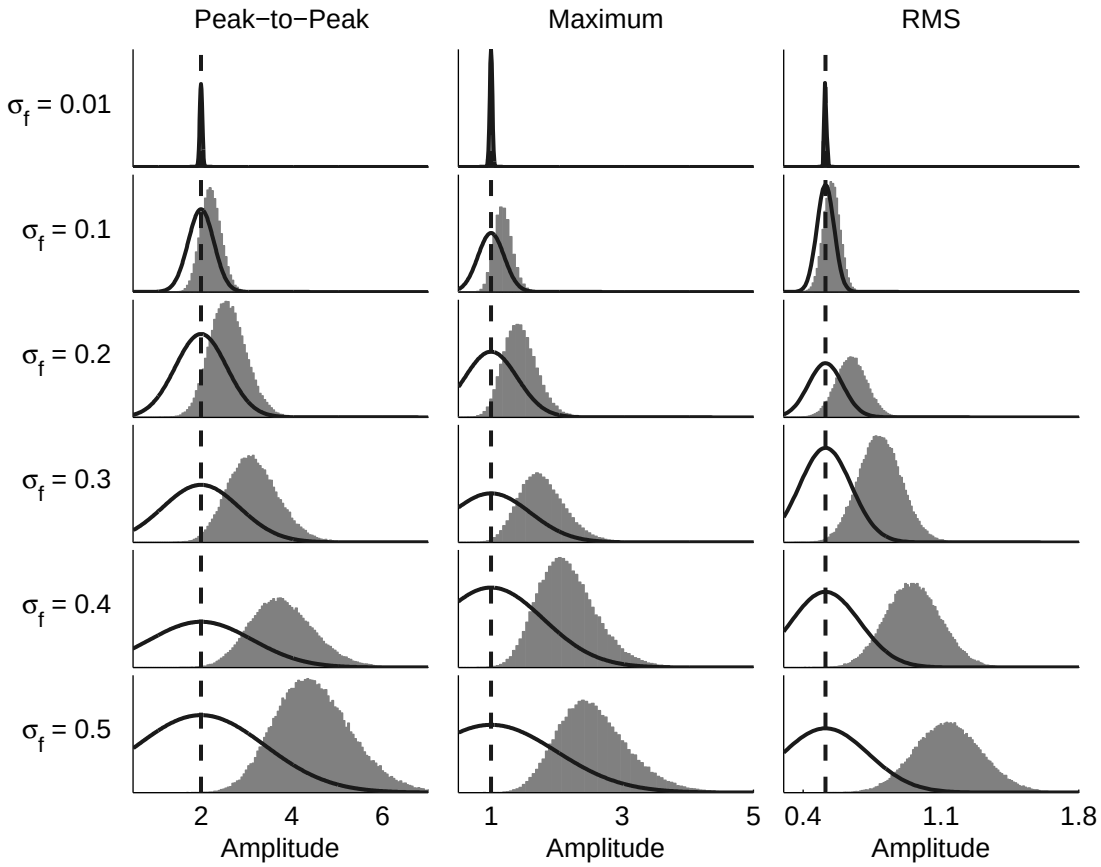

Figure 3. Histogram of amplitude measurement distributions (in grey) for different fractional noise levels, σ_f, using different measurement approaches. The vertical dashed lines are the true amplitude values, and the solid lines are the assumed Gaussian distributions for the given uncertainty. The fractional noise corresponds to the inverse of the true SNR value. The vertical scales are independent between the plots because the amplitude distributions are normalised so that the area integral is 1.

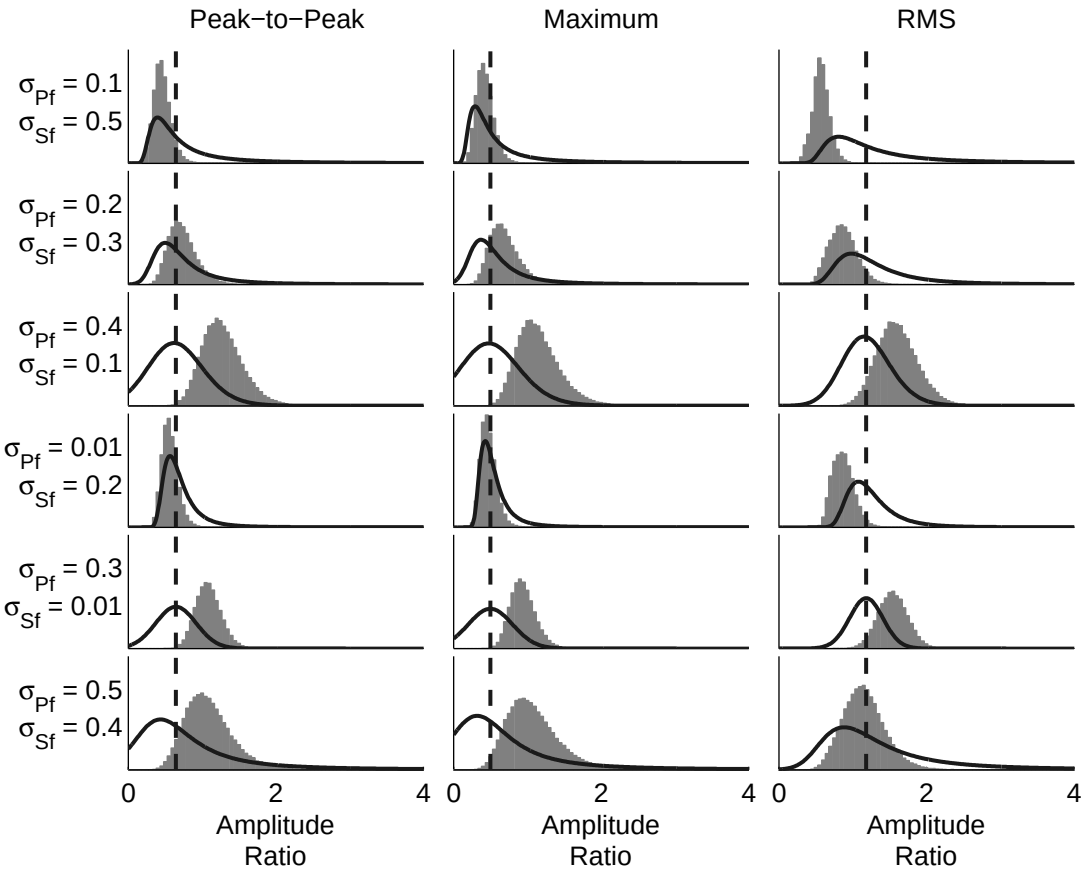

Figure 4. Histogram of amplitude ratio distributions (in grey) with differing fractional noise levels for the P, σ_{Pf}, and S, σ_{Sf}, arrivals using different measurement approaches. The vertical dashed lines correspond to the correct values and the solid lines are the assumed distributions of the observed ratios given the correct amplitudes and uncertainties, corresponding to the ratio PDF as defined in Section 4.3.2. The fractional noise corresponds to the inverse of the true SNR value. The vertical scales are independent between the plots because the amplitude ratio distributions are normalised so that the area integral is 1.

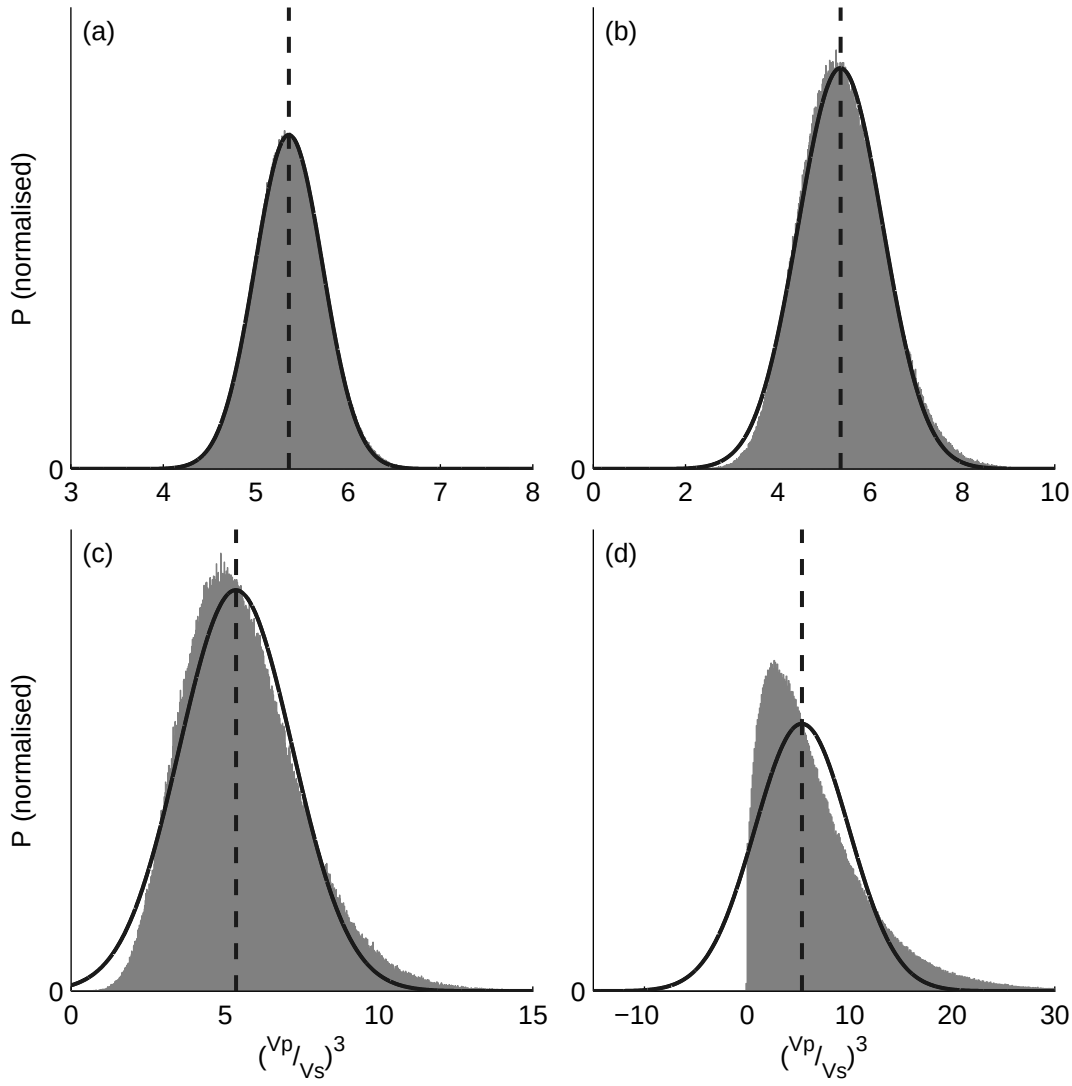

Figure 5. Histograms of the propagation correction (Eq. 31) distributions (in grey) assuming Gaussian error on $^{V_P}/_{V_S}$. The solid lines correspond to $\mathcal{N}(Z, \mu_Z, \sigma_Z)$ calculated according to Eqs 32 and 33. The dashed vertical lines correspond to μ_Z, and the histograms show the actual distributions calculated from 1 000 000 samples from Gaussian distributions for $^{V_P}/_{V_S}$ with $\mu_{V_P/V_S} = 1.75$. The standard deviation was varied, taking values of (a) $\sigma_{V_P/V_S} = 0.04$, (b) $\sigma_{V_P/V_S} = 0.1$, (c) $\sigma_{V_P/V_S} = 0.2$, and (d)$\sigma_{V_P/V_S} = 0.5$. The vertical scales are independent between the plots because the amplitude distributions are normalised so that the area integral is 1.

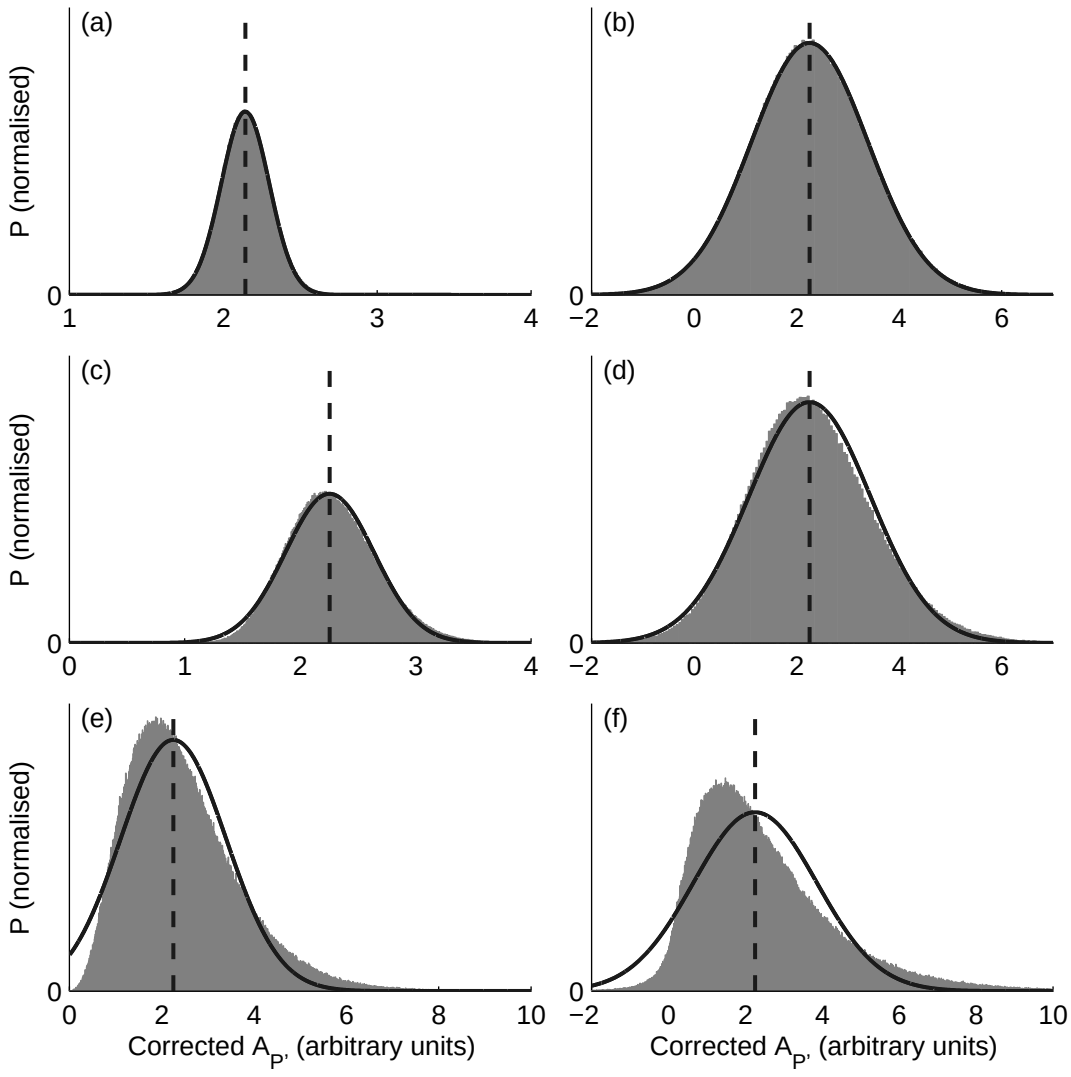

Figure 6. Histograms of the propagation corrected model P amplitude $A_{\mathrm{P}'}$ distributions (in grey). The solid lines correspond to $\mathcal{N}\left(A_{\mathrm{P}'}, \mu_{A_{\mathrm{P}'}}, \sigma_{A_{\mathrm{P}'}}\right)$ calculated according to Eqs 34 and 35. The dashed vertical lines correspond to $\mu_{A_{\mathrm{P}'}}$, and the histograms show the actual distribution calculated from 1 000 000 samples from Gaussian distributions for A_{P} and $V_{\mathrm{P}}/V_{\mathrm{S}}$ with $\mu_{V_{\mathrm{P}}/V_{\mathrm{S}}} = 1.75$. The left column shows the distributions for a P amplitude SNR = 40, the right column SNR = 2. (a) and (b) $\sigma_{V_{\mathrm{P}}/V_{\mathrm{S}}} = 0.04$, (c) and (d) $\sigma_{V_{\mathrm{P}}/V_{\mathrm{S}}} = 0.1$, and (e) and (f)$\sigma_{V_{\mathrm{P}}/V_{\mathrm{S}}} = 0.3$. The vertical scales are independent between the plots as the amplitude distributions are normalised so that the area integral is 1.

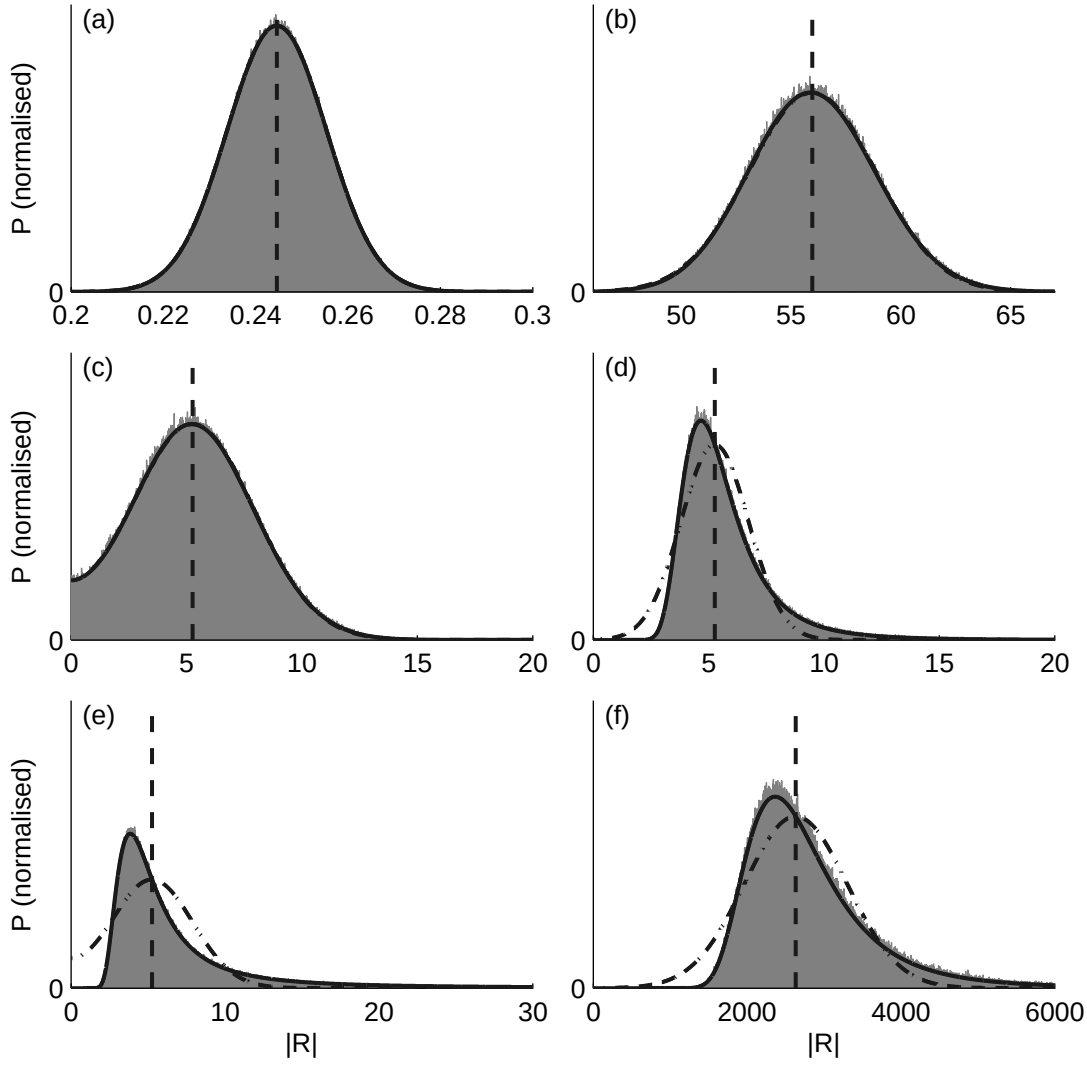

Figure 7. Histograms of the model ratio R distributions (in grey). The dash-dot lines correspond to the Gaussian noise model (Eq. 38). The solid lines correspond to the ratio distribution (Eq. 53), again adding on the part corresponding to the distribution reflected in $R = 0$ to account for the fact that the unsigned (absolute) ratio is used. The vertical dashed lines show the true value, μ_R, and the histograms show the actual distribution calculated from 1 000 000 samples from Gaussian distributions for A_P, A_S, and $^{V_P}/_{V_S}$, with $\mu_{v_{P/V_S}} = 1.74$ and $\sigma_{v_{P/V_S}} = 0.1$. The P and S uncertainties were varied: (a) and (b) have low P and S amplitude uncertainties (SNR > 10), (c) $\mathrm{SNR_P} = 2$, $\mathrm{SNR_S} = 20$, (d) $\mathrm{SNR_P} = 15$, $\mathrm{SNR_S} = 3.8$, (e) $\mathrm{SNR_P} = 200$, $\mathrm{SNR_S} = 2$, and (f) $\mathrm{SNR_P} = 1000$, $\mathrm{SNR_S} = 4$. The vertical scales are independent between the plots as the amplitude distributions are normalised so that the area integral is 1.

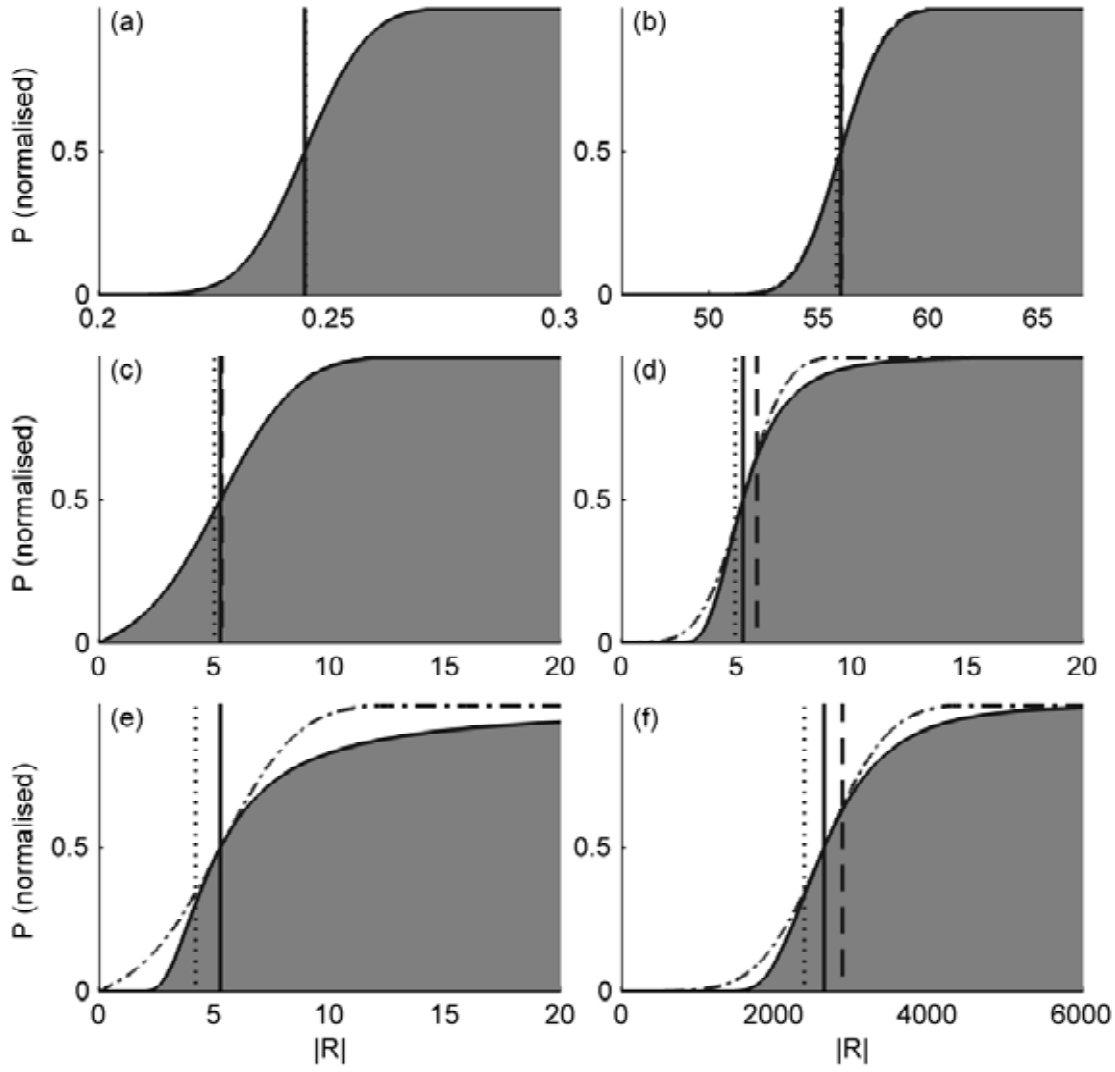

Figure 8. Histograms of the cumulative distribution of the model ratio R (in grey). The dashed-dot lines correspond to the Gaussian noise model CDF (Eq. 40). The solid lines show the ratio distribution CDF (Eq. 54). The solid vertical lines correspond to μ_R and the distribution medians, the dotted vertical lines are an estimate of the distribution modes and the dashed vertical lines are the distribution means. The histograms show the actual distribution calculated from 1 000 000 samples from Gaussian distributions for A_P, A_S, and v_P/v_S, with $\mu_{v_P/v_S} = 1.74$ and $\sigma_{v_P/v_S} = 0.1$. The P and S uncertainties were varied: (a) and (b) have low P and S amplitude uncertainties (SNR > 10), (c) $SNR_P = 2$, $SNR_S = 20$, (d) $SNR_P = 15$, $SNR_S = 3.8$, (e) $SNR_P = 200$, $SNR_S = 2$, and (f) $SNR_P = 1000$, $SNR_S = 4$.

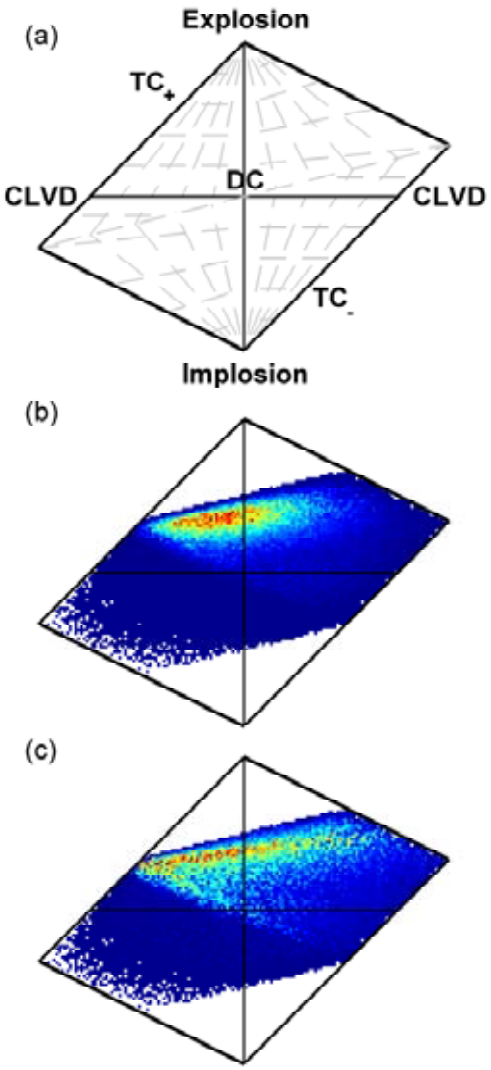

Figure 9. Comparison of marginalised and silhouette plots of the source PDF types for an example real event. (a) shows the position of different source types on the Hudson plot, with the double-couple point located where the axes cross. Several source types are shown on the plot: the compensated linear vector dipole (CLVD) end members (Knopoff & Randall 1970), the implosive and explosive sources, the positive and negative tensile crack sources (TC_+ and TC_-) for a Poisson's ratio of 0.25, and the double-couple source (DC). (b) shows the marginalised source PDF, and (c) the silhouette plot of the source PDF for the equal area Hudson plot. Red shows high probability, and blue low (non-zero) probability.

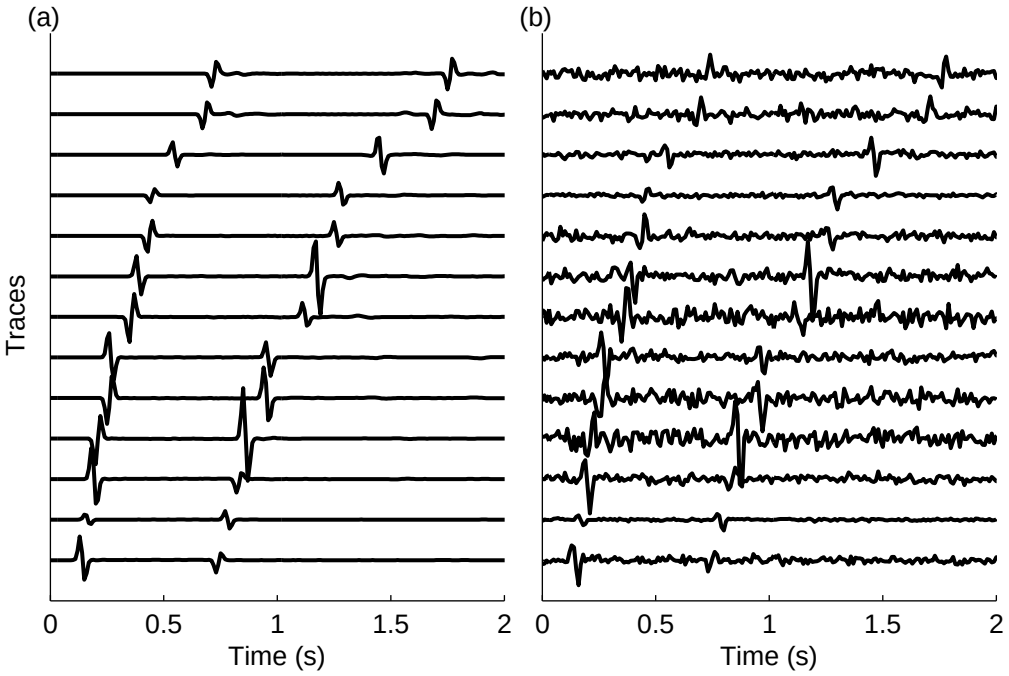

Figure 10. Example vertical component synthetic data for a double-couple source with added Gaussian noise. (a) shows the synthetic data, and (b) the same data with the added noise with a SNR of 3. The frequency content of the signal is band-pass filtered between 1-40 Hz.

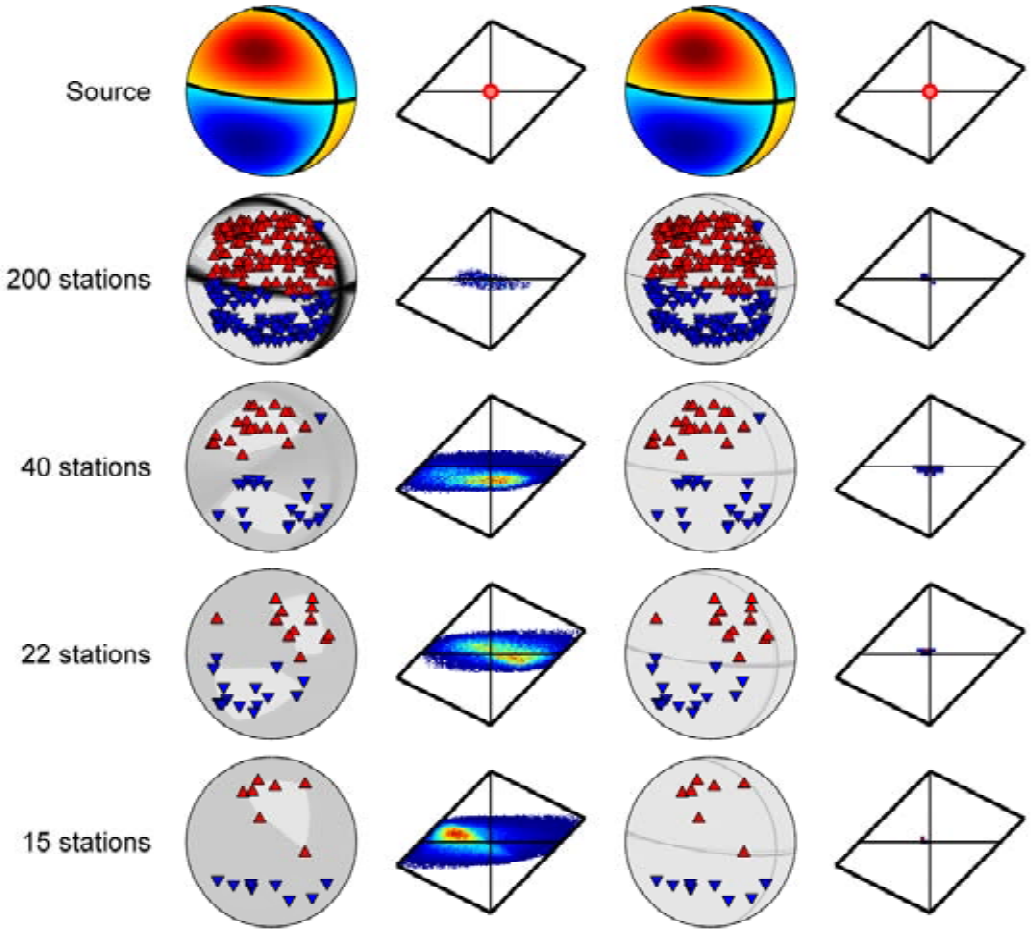

Figure 11. Lower hemisphere equal area projections and Hudson plots of the source PDF for a synthetic double-couple source with a range of station numbers and geometries. The first and third columns show the source PDF for the solution constrained to be double-couple only. The second and fourth columns show the source PDF for the full moment tensor solution. The first two columns show the solutions for inversions using only polarity data, and the second two columns show the solutions using polarity and amplitude ratio data. Manually picked station first motions are given by upward red or downward blue triangles. For the focal sphere plots, possible fault planes are given by dark lines. The most likely fault planes are given by the darkest lines. For the Hudson plots, high probability is red and low probability is in blue. The positions of the different source types on the Hudson plot are shown in Fig. 9. The posterior model probabilities for the double-couple source model determined from the Bayesian evidence are, from top to bottom, 0.97, 0.40, 0.77, and 0.63, for polarity only inversions and 0 for the inversions including amplitude ratios, except for the 15 receiver examples which has $p_{\mathrm{DC}} = 1$.

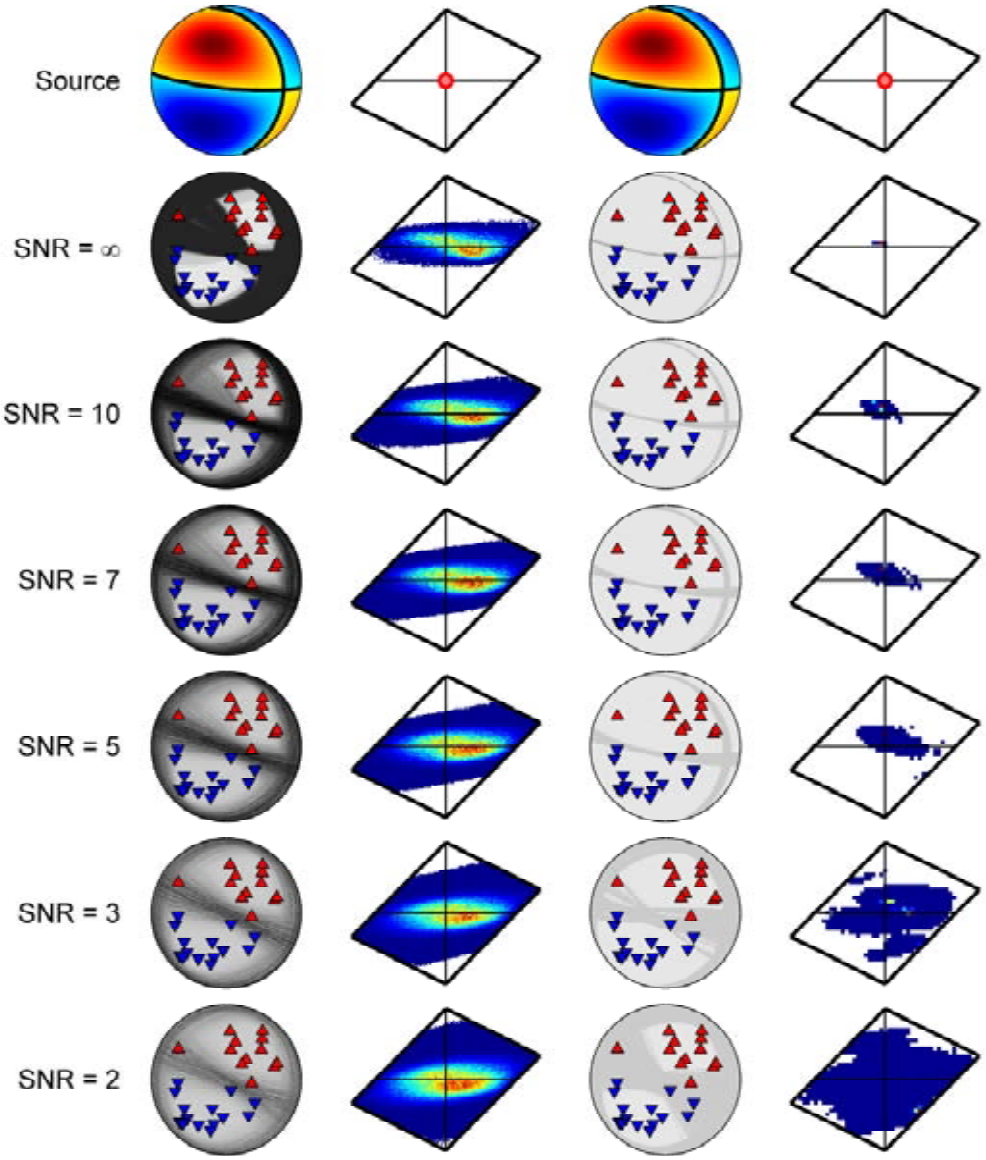

Figure 12. Lower hemisphere equal area projections and Hudson plots of the source PDF for a synthetic double-couple source for a range of data uncertainties, corresponding to SNR = ∞, SNR = 10, SNR = 7, SNR = 5, SNR = 3, and SNR = 2. The first and third columns show the source PDF for the solution constrained to be double-couple only. The second and fourth columns show the source PDF for the full moment tensor solution. The first two columns show the solutions for inversions using only polarity data, and the second two columns show the solutions using polarity and amplitude ratio data. Manually picked station first motions are given by upward red or downward blue triangles. For the focal sphere plots, possible fault planes are given by dark lines. The most likely fault planes are given by the darkest lines. For the Hudson plots, high probability is red and low probability is in blue. The positions of the different source types on the Hudson plot are shown in Fig. 9. The posterior model probabilities for the double-couple source model determined from the Bayesian evidence are, from top to bottom, 0.77, 0.78, 0.80, 0.84, 0.91, and 0.94, for the polarity only inversion, and 1.0, 0.01, 0, 0, 0.02, and 0.01, for the polarity and amplitude ratio inversions.

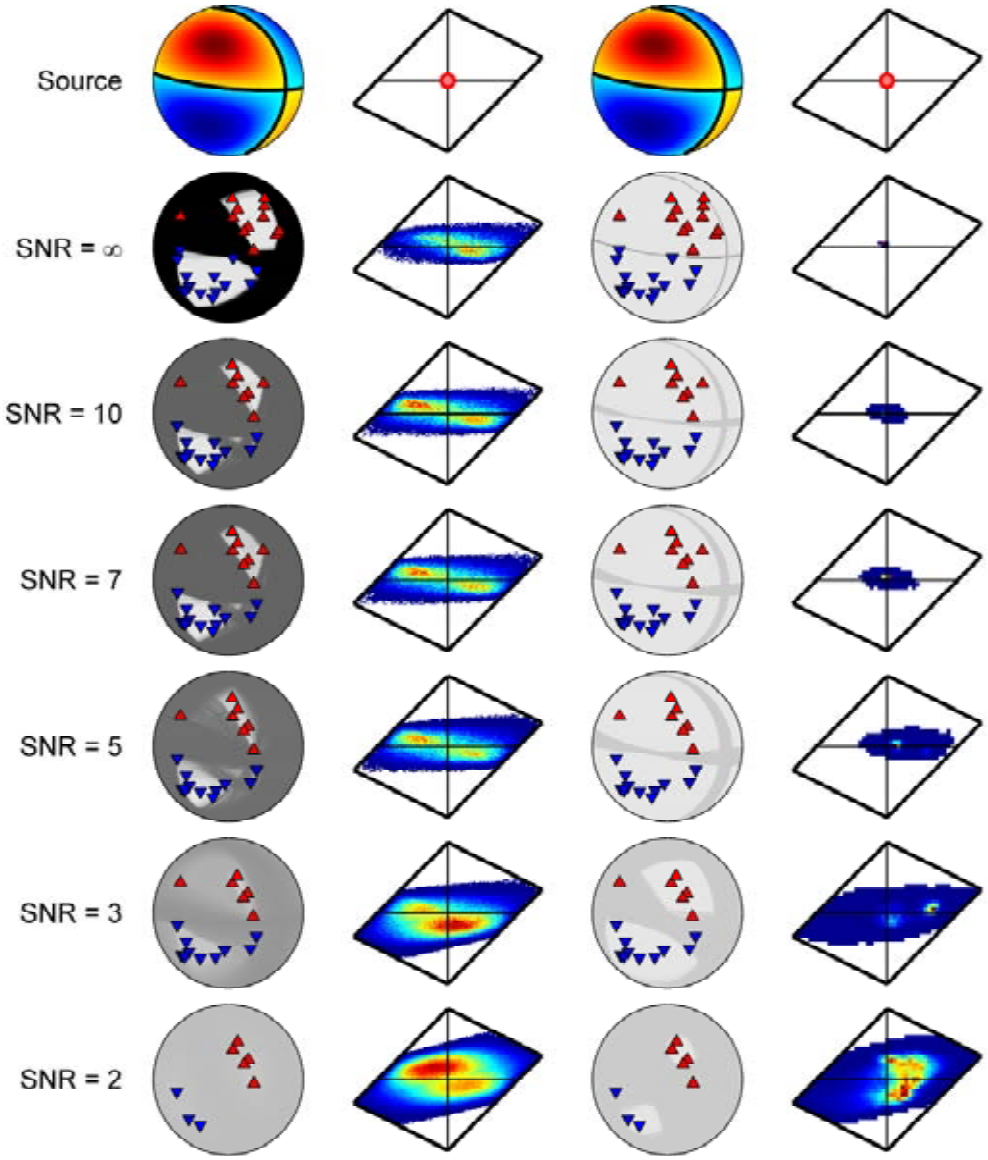

Figure 13. Lower hemisphere equal area projections and Hudson plots of the source PDF for a synthetic double-couple source with a range of noise levels, SNR = ∞, SNR = 10, SNR = 7, SNR = 5, SNR = 3, and SNR = 2. The first and third columns show the source PDF for the solution constrained to be double-couple only. The second and fourth columns show the source PDF for the full moment tensor solution. The first two columns show the solutions for inversions using only polarity data, and the second two columns show the solutions using polarity and amplitude ratio data. Manually picked station first motions are given by upward red or downward blue triangles. For the focal sphere plots, possible fault planes are given by dark lines. The most likely fault planes are given by the darkest lines. For the Hudson plots, high probability is red and low probability is in blue. The positions of the different source types on the Hudson plot are shown in Fig. 9. The posterior model probabilities for the double-couple source model determined from the Bayesian evidence are, from top to bottom, 0.77, 0.71, 0.71, 0.77, 0.64, and 0.62, for the polarity only inversion, and 0, 0.99, 0.70, 0.84, 0.82, and 0.58, for the polarity and amplitude ratio inversions.

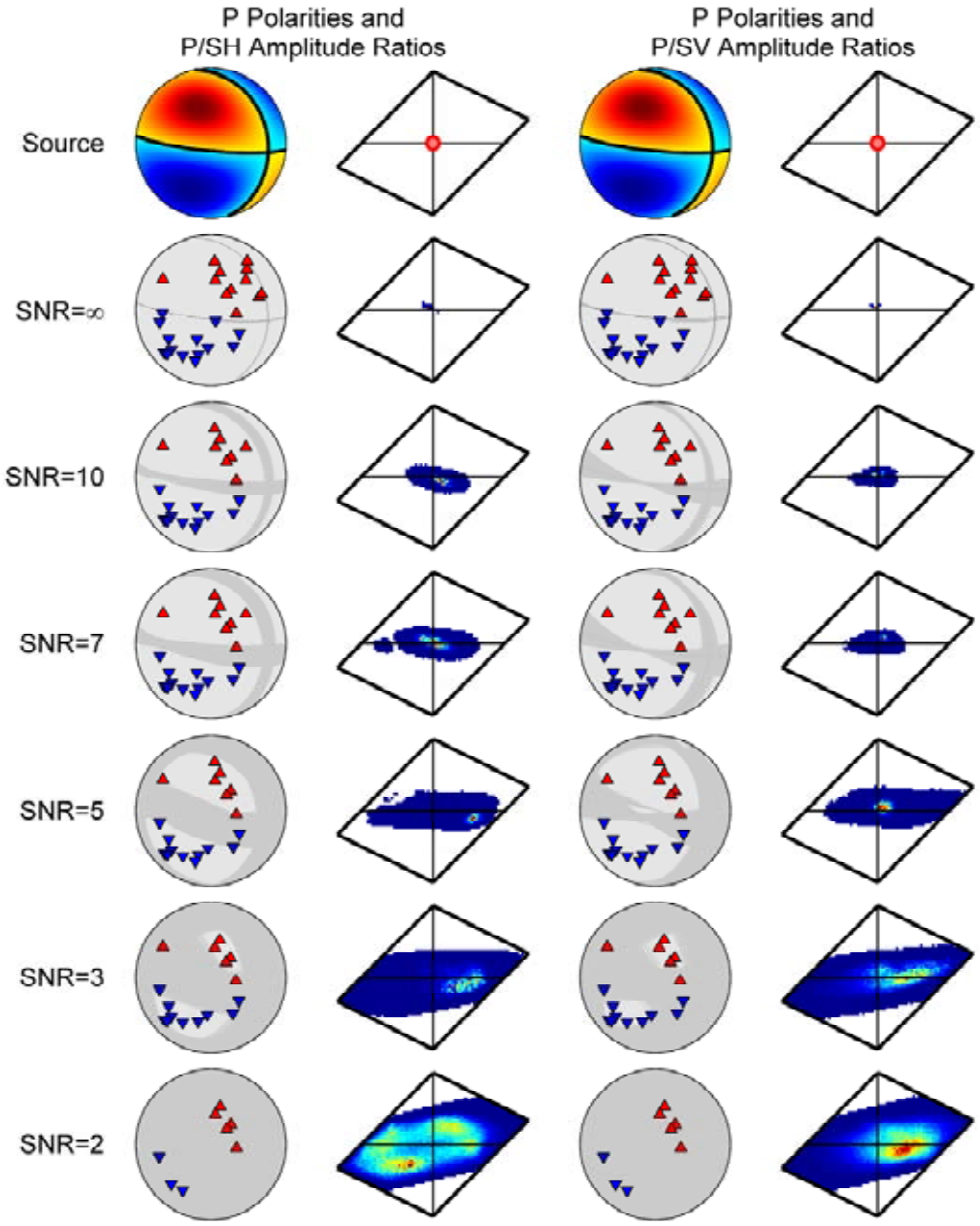

Figure 14. Lower hemisphere equal area projections and Hudson plots of the source PDF for a synthetic double-couple source using P polarity and either P/SH- or P/SV-amplitude ratios. The first and third columns show the source PDF for the solution constrained to be double-couple only. The second and fourth columns show the source PDF for the full moment tensor solution. Manually picked station first motions are given by red (up) or blue (down) points, with darker points corresponding to more likely locations. For the focal sphere plots, possible fault planes are given by dark lines. The most likely fault planes are given by the darkest lines. For the Hudson plots, high probability is red and low probability is in blue. The positions of the different source types on the Hudson plot are shown in Fig. 9.

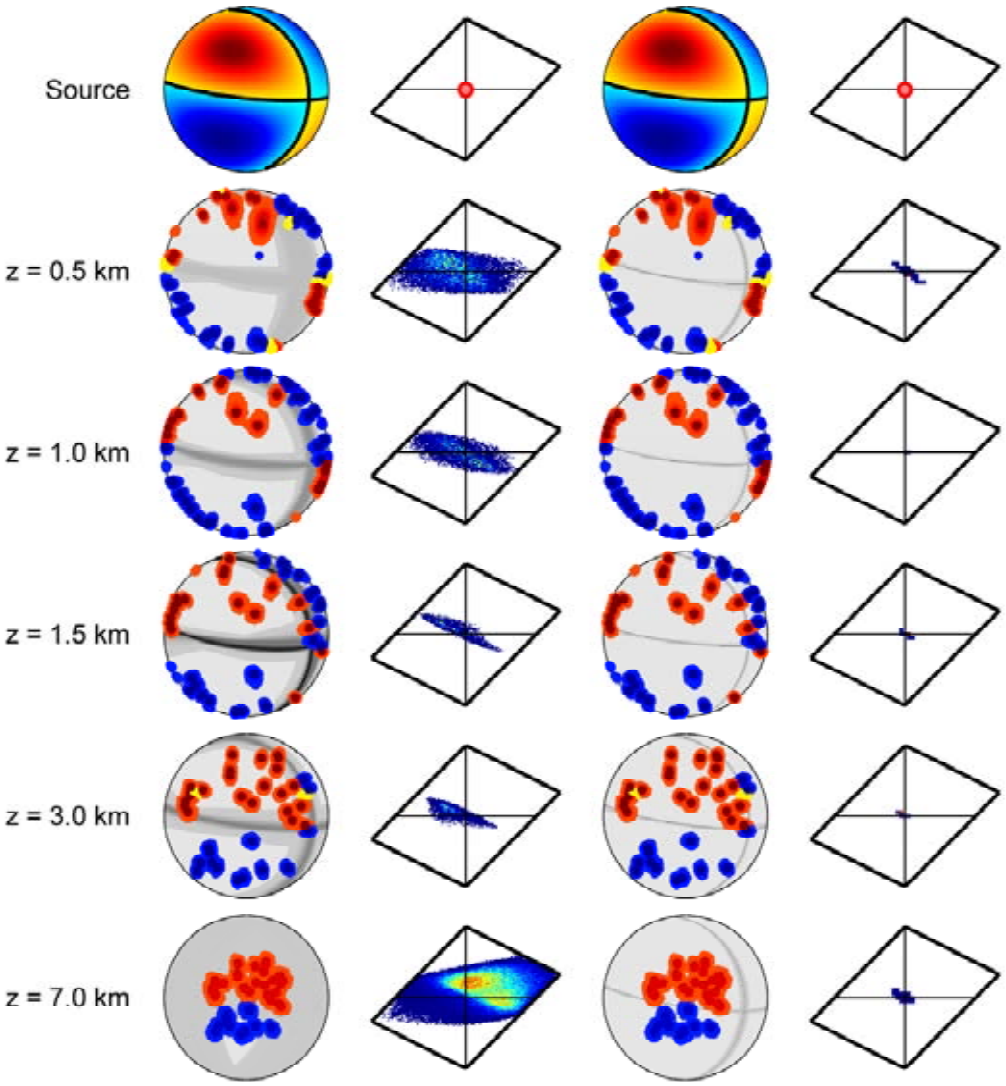

Figure 15. Lower hemisphere equal area projections and Hudson plots of the source PDF, marginalised over the location uncertainty, for a synthetic double-couple source at a range of source depths in the centre of an array, z = 0.5 km, z = 1 km, z = 1.5 km, z = 3 km, and z = 7 km. The first and third columns show the scatter of station ray positions on the focal sphere for the NonLinLoc location PDF and the source PDF for the solution constrained to be double-couple only. The second and fourth columns show the source PDF for the full moment tensor solution. The first two columns show the solutions for inversions using only polarity data, and the second two columns show the solutions using polarity and amplitude ratio data. Manually picked station first motions are given by red (up) or blue (down) points, with darker points corresponding to more likely locations. For the focal sphere plots, possible fault planes are given by dark lines. The most likely fault planes are given by the darkest lines. For the Hudson plots, high probability is red and low probability is in blue. The positions of the different source types on the Hudson plot are shown in Fig. 9. The posterior model probabilities for the double couple source model determined from the Bayesian evidence are, from top to bottom, 0.80, 0.91, 0.74, 0.95, and 0.38, for the polarity only inversion, and 0.99, 0, 1, 1, and 1, for the polarity and amplitude ratio inversions.

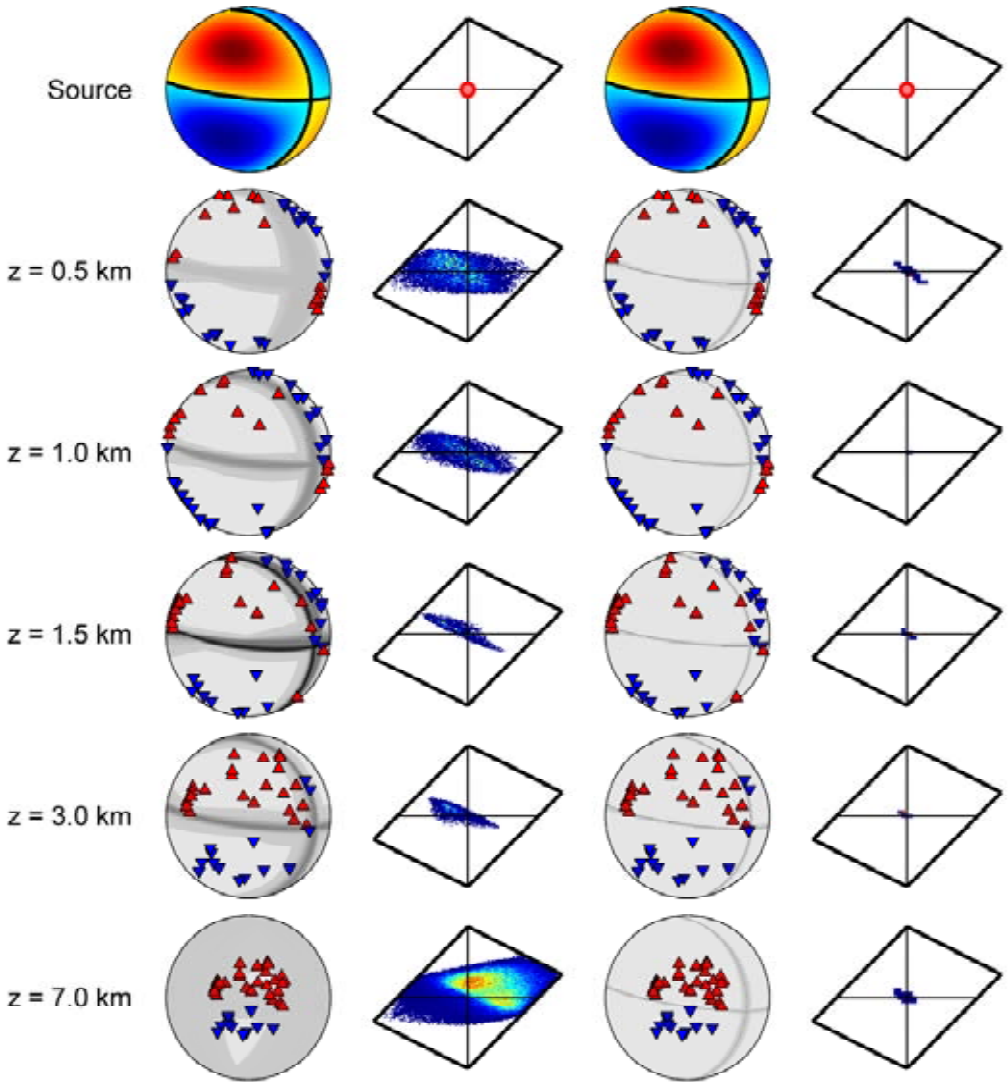

Figure 16. Lower hemisphere equal area projections and Hudson plots of the source PDF, without location uncertainty, for a synthetic double-couple source at a range of source depths in the centre of an array, z = 0.5 km, z = 1 km, z = 1.5 km, z = 3 km, and z = 7 km. The first and third columns show the source PDF for the solution constrained to be double-couple only. The second and fourth columns show the source PDF for the full moment tensor solution. The first two columns show the solutions for inversions using only polarity data, and the second two columns show the solutions using polarity and amplitude ratio data. Manually picked station first motions are given by red (up) or blue (down) markers. For the focal sphere plots, possible fault planes are given by dark lines. The most likely fault planes are given by the darkest lines. For the Hudson plots, high probability is red and low probability is in blue. The positions of the different source types on the Hudson plot are shown in Fig. 9. The posterior model probabilities for the double couple source model determined from the Bayesian evidence are, from top to bottom, 0.80, 0.89, 0.63, 0.95, and 0.38, for the polarity only inversion, and 0.99, 0, 1, 1, and 1, for the polarity and amplitude ratio inversions.

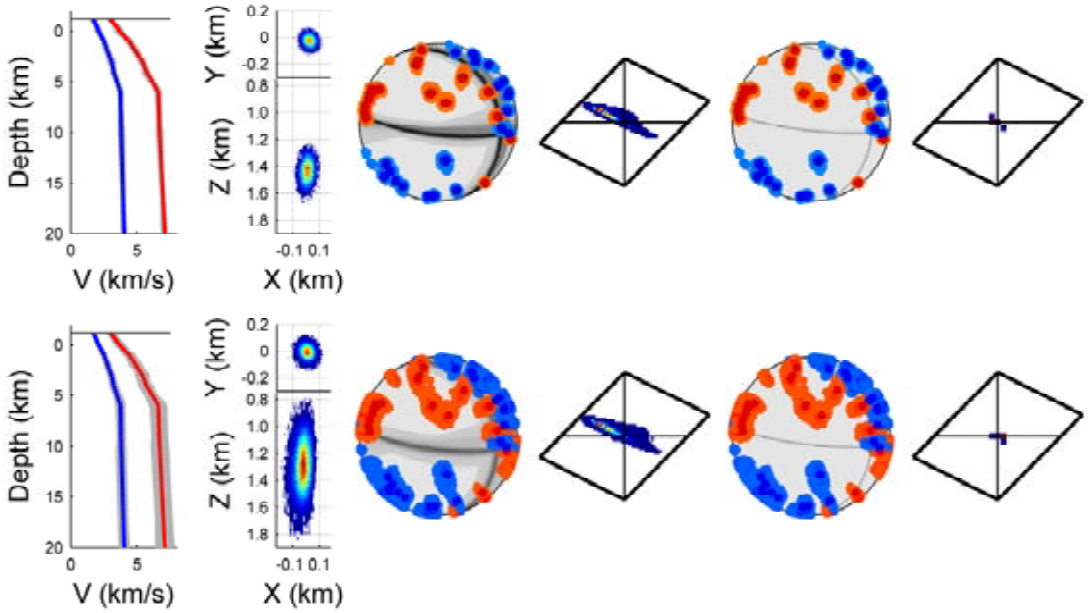

Figure 17. Lower hemisphere equal area projections and Hudson plots of a synthetic double-couple source at a depth of 1.5 km located using two sets of velocity models perturbed from the true model. The first column shows the range of velocity models, with a 3% perturbation for the top row and a 10% perturbation for the bottom row. The second column shows the location distribution (blue indicates low probability and red high probability) for the NonLinLoc location PDF, in an XY plan view and XZ depth section. The third and fifth columns show the scatter of station ray positions on the focal sphere for the NonLinLoc location PDF and the source PDF for the solution constrained to be double-couple only. The fourth and sixth columns show the source PDF for the full moment tensor solution. The first two columns of source PDFs show the solutions for inversions using only polarity data, and the second two show the solutions using polarity and amplitude ratio data. Manually picked station first motions are given by red (up) or blue (down) points, with darker points corresponding to more likely locations. For the focal sphere plots, possible fault planes are given by dark lines. The most likely fault planes are given by the darkest lines. For the Hudson plots, high probability is red and low probability is in blue. The positions of the different source types on the Hudson plot are shown in Fig. 9. The posterior model probabilities for the double couple source model determined from the Bayesian evidence are, for the 0.86 for the 3% perturbation and 0.76 for the 10% perturbation, for the polarity only inversion, and 1.0 for both perturbations for the polarity and amplitude ratio inversions. 1

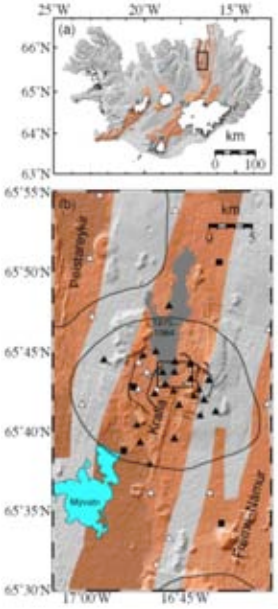

Figure 18. Map of the receiver locations for the Krafl a temporary network for the events in Section 7. (a) Shows a tectonic map of Iceland (Einarsson & Sæmundsson 1987) with brown shading indicating the location of fi ssure swarms associated with volcanic systems. Glaciers are shown with white fi ll. The black rectangle shows location of (b) in the Northern Volcanic Zone. (b) Map of the seismic network. Brown shading shows extent of main fi ssure swarms associated with diff erent volcanic systems. The solid black outlines show the extent of central volcanoes, and the inner sub-circular outline at Krafla indicates the extent of the caldera. Black and white symbols show the seismometer locations: black squares show the permanent Icleandic Meteorological Offi ce seismometers; solid black triangles show temporary seismometers operating from July 2009 and white triangles show additional seismometer sites used during July 2010 to Sept 2011. The black rectangle shows the region where the seismicity occurred.

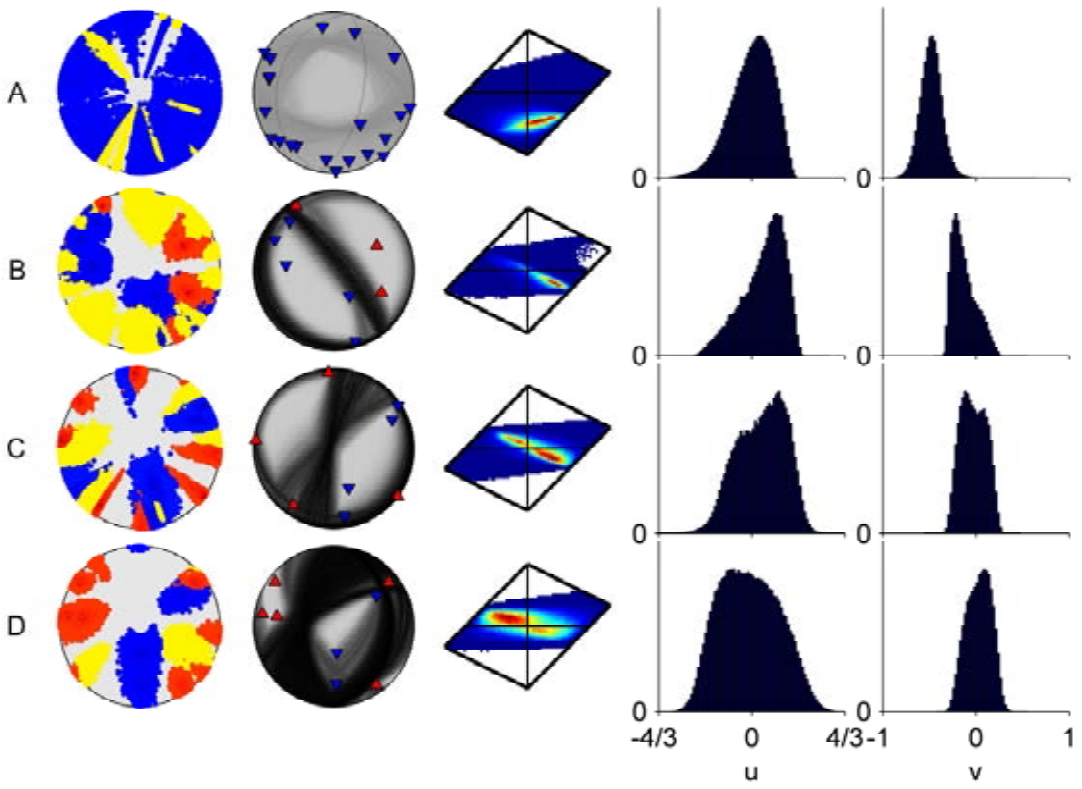

Figure 19. Lower hemisphere equal area projections, Hudson plots and marginalised distributions of the Hudson u and v parameters of four events from the Krafla geothermal region of Iceland. The first column shows the station ray position distribution from the NonLinLoc location PDF. The second column shows the maximum likelihood location station positions and the source PDF for the solution constrained to be double-couple only. The third column shows the source PDF for the full moment tensor solution. The fourth and fifth columns show the marginalised distributions of the Hudson u and v parameters, respectively. Manually picked station first motions are given by red (up) or blue (down) markers, with darker points corresponding to more likely locations in the ray position distribution. For the focal sphere plots, possible fault planes are given by dark lines. The most likely fault planes are given by the darkest lines. For the Hudson plots, high probability is red and low probability is in blue. The positions of the different source types on the Hudson plot are shown in Fig. 9. The posterior model probabilities for the double-couple source model determined from the Bayesian evidence and the Bayesian information criterion values are both given in Table 3.

Uncertainty	**Description**
Velocity Model	Used throughout an inversion, poor choice of model can lead to large errors
Pick Time	Depends on instrument timing errors, background noise and human subjectivity
Instrument and Background Noise	Noise affects all the observations
Hypocentre	Dependent on velocity model used and pick times
Polarity	Dependent on pick time, instrument orientation error, and background noise
Instrument Location	Affects positioning on the focal sphere for a given hypocentre and model
Instrument Coupling and Site Effects	Affects measured seismic wave amplitudes
Instrument Orientation	Errors in orientation of instrument components impact many measurements
Source Angle Error	Dependent on hypocentre, instrument location and velocity model

Table 1. Some uncertainties that affect source inversion and their origins

Symbol	Parameter
$\mathbf{Y}$	Observed polarities
$\mathbf{R}$	Observed amplitude ratios
τ	Pick arrival times
$\mathbf{x}$	Event location
$\mathbf{A}$	Theoretical amplitudes
σ	Observed measurement errors
σ_τ	Pick arrival time uncertainties
ϖ	Probability of reversed trace
$\mathbf{a}$	Station angle six-vector
$\mathbf{M}$	Moment tensor
$\boldsymbol{G}$	Earth model
$\mathbf{s}$	Station locations

Table 2. Inversion parameters

Event	p_{DC}	$\mathcal{I}^{\mathrm{B}}_{\mathrm{DC}} - \mathcal{I}^{\mathrm{B}}_{\mathrm{MT}}$
A	0.002	−4.3
B	0.84	4.1
C	0.81	4.9
D	0.73	4.1

Table 3. Posterior model probabilities and Bayesian information criterion values, $\mathcal{I}^{\mathrm{B}}$ (Section 5), for the Krafla events shown in Fig. 19.