Geophysical Journal International

Geophys. J. Int. (2018) 212, 1963-1985 Advance Access publication 2017 December 4 GJI Seismology

doi: 10.1093/gji/ggx517

Waveform-based Bayesian full moment tensor inversion and uncertainty determination for the induced seismicity in an oil/gas field

Chen Gu,¹ Youssef M. Marzouk² and M. Nafi Toksöz¹

¹Earth Resources Laboratory, Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail: guchch@mit.edu

²Aerospace Computational Design Laboratory, Department of Aeronautics and Astronautics Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Accepted 2017 November 30. Received 2017 November 23; in original form 2016 June 25

SUMMARY

Small earthquakes occur due to natural tectonic motions and are induced by oil and gas production processes. In many oil/gas fields and hydrofracking processes, induced earthquakes result from fluid extraction or injection. The locations and source mechanisms of these earthquakes provide valuable information about the reservoirs. Analysis of induced seismic events has mostly assumed a double-couple source mechanism. However, recent studies have shown a non-negligible percentage of non-double-couple components of source moment tensors in hydraulic fracturing events, assuming a full moment tensor source mechanism. Without uncertainty quantification of the moment tensor solution, it is difficult to determine the reliability of these source models. This study develops a Bayesian method to perform waveform-based full moment tensor inversion and uncertainty quantification for induced seismic events, accounting for both location and velocity model uncertainties. We conduct tests with synthetic events to validate the method, and then apply our newly developed Bayesian inversion approach to real induced seismicity in an oil/gas field in the sultanate of Oman-determining the uncertainties in the source mechanism and in the location of that event.

Key words: Joint inversion; Probability distributions; Waveform inversion; Induced seismicity.

1 INTRODUCTION

Induced microearthquakes occur widely in conventional and unconventional oil/gas fields. The study of induced seismicity is of great importance in monitoring and understanding the processes of hydraulic fracturing, fluid injection and oil/gas extraction (Maxwell et al. 2014; Shapiro 2015). Inverting for the source mechanism of a seismic event is an essential effort in this context. Determining the source mechanisms of induced earthquakes can give the stress and fault orientation in the field (Vavryčuk 2014).

The source mechanisms of the great majority of tectonic earthquakes can be described by a double-couple (DC), corresponding to a shear fracture. However, some events in volcanic, geothermal and salt dome areas exhibit more complex source mechanisms with non-DC components, such as volumetric components (ISO) and compensated linear vector dipoles (CLVD; Cespuglio et al. 1996; Panza & Saraò 2000; Templeton & Dreger 2006; Nayak & Dreger 2014). These non-DC components have also been related to mine collapses and nuclear explosions (Ford et al. 2009a,b). Previous research on the analysis of induced seismic events in conventional oil/gas fields assumed a DC source mechanism (Li et al. 2011a,b; Li 2013). Assuming full moment source mechanisms, Horálek et al. (2010) found a DC-dominated source mechanism of the induced microearthquakes in a geothermal area during massive fluid injection. However, recent studies have shown interest in non-DC components of source moment tensors in hydraulic fracturing events (Vavryčuk et al. 2008; Šílený et al. 2009; Warpinski & Du 2010; Song & Toksöz 2011; Song 2013; Vavryčuk 2015).

Many studies have implemented inversion of the full moment tensor by least-squares (LSQ) or regularized LSQ methods (Sipkin 1982; Šílený et al. 1992, 1996). However, LSQ methods have limitations in estimating and interpreting the uncertainty of moment tensor solutions, since they only search for the best moment tensor solution and do not naturally yield probability distributions on the solution. An important question for full moment tensor inversion is whether the non-DC components are real. Some research has applied F-tests to check the significance of the non-DC components (Templeton & Dreger 2006; Šílený et al. 2009; Horálek et al. 2010; Nayak & Dreger 2014).

LSQ methods do also allow for a limited form of uncertainty quantification, based on the Hessian of the misfit function near the LSQ point estimate. But this is only a local and linearized estimate of uncertainty, and can be difficult to interpret in the regularized case. Alternatively, to address uncertainty in the moment tensor due to data noise or imperfect station coverage, many LSQ-based moment tensor inversion studies have applied resampling methods to the data, such as Monte Carlo noise realization methods and jackknife tests (Šílený *et al.* 2009; Stierle *et al.* 2014a,b); here, uncertainties in the moment tensor were estimated by statistically analysing solutions obtained from the resampled data. Du & Warpinski (2011) theoretically derived the uncertainty of the focal plane solutions of seismic sources due to Gaussian noise contamination. Additionally, to characterize uncertainties due to misspecification of the earthquake location and the velocity model perturbation tests and observed the impact of these perturbations on moment tensor solutions (Stierle *et al.* 2014a,b). These tests only show how much the inverted parameters change given a few prescribed perturbations to the earthquake location and velocity model; they do not provide a statistical characterization of uncertainty in the inverted parameters.

Alternatively, Nayak & Dreger (2014) jointly assessed the sensitivities of the moment tensor and location by statistically analysing all the feasible moment tensor and location solutions that yielded variance reduction (i.e. waveform matching) better than a threshold. This method provides somewhat more comprehensive estimates of the uncertainties of the moment tensor and location, but it is difficult to develop a criterion for choosing the threshold.

Compared to LSQ methods, Bayesian inversion methods naturally quantify the uncertainty in model parameters by characterizing a posterior probability distribution over the parameter space (Tarantola 2005; Kaipio & Somersalo 2006; Stuart 2010). Some studies have conducted Bayesian moment tensor inversion on moderate and large earthquakes. Duputel *et al.* (2012) introduced a Bayesian moment tensor inversion method to estimate the uncertainties of source mechanisms for large earthquakes ($M_w \ge 6.0$) using W phase waveforms from a global seismic network. That study did not determine the uncertainty of seismic locations jointly with the moment tensor. Stähler & Sigloch (2014, 2016) presented a detailed study of probabilistic moment tensor and depth inversion from *P* and *S* waveforms, as well as error modelling and station covariances during seismic source inversion. A very recent paper by Mustać & Tkalčić (2016) developed a Bayesian full moment tensor inversion for a moderate-size earthquake with a well-studied source mechanism using a regional seismic network. In that research, the uncertainties of both the seismic location and the moment tensor were studied by implementing an outer Markov chain to sample the location parameters.

In this study, we introduce a waveform-based Bayesian full moment tensor inversion method. Our method analyses the uncertainties of both the seismic moment tensor and location; moreover, we also account for velocity model uncertainties in a rigorous Bayesian fashion. Unlike the Bayesian method implemented by Mustać & Tkalčić (2016), we sample the source location, velocity, and moment tensor parameters using a single Markov chain; this approach reduces computational cost and provides more accurate uncertainty estimates, particularly for the source location. Moreover, we use the conditionally Gaussian structure of the parameter posterior to solve portions of the inverse problem analytically, reducing the dimension of the sampling problem and allowing the impact of location and velocity uncertainties on the moment tensor uncertainty to be explicitly quantified. We first validate the method using synthetic data before applying it to a selected induced event in an oil/gas field in Oman (Fig. 1). We determine the full moment tensor of the induced seismicity from a conventional oil/gas field. The seismicity of this field and source mechanisms of events using DC assumptions have been studied extensively (Sarkar 2008; Li *et al.* 2011a,b). To better quantify the uncertainties, we use the newly developed waveform-based Bayesian method for full moment tensor inversion, source location, 1-D velocity model inference, and uncertainty quantification.

2 METHODOLOGY

2.1 Full moment tensor and waveform modelling

The source mechanisms of a seismic event can be represented by a 3×3 symmetric matrix M,

$$\boldsymbol{M} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}$$
(1)

where each element of the matrix presents a force couple. M is the full moment tensor representation of a seismic source mechanism.

The matrix M can be converted to fault plane solutions, which are very important for analysing and interpreting the induced seismicity. The following equations define the source-related quantities (Vavryčuk 2001, 2011):

$$\epsilon = -\frac{|\lambda_{\min}|}{|\lambda_{\max}|},$$
(2)

$$M_0 = \frac{1}{\sqrt{2}} \left(\sum_{j=1}^{N} \sum_{k=1}^{N} M_{jk}^2 \right) \quad , \tag{3}$$



Figure 1. (a) Map view and side view of the stations and located events for both near-surface network and downhole network (Sarkar 2008). The red dots denote the location of the detected events, and the green triangles show the location of the stations. The black lines are the identified faults. The green triangles (VA11, VA21, VA31, VA41 and VA51) are the five near-surface stations. These stations are located in shallow boreholes 150 m below the surface. (b) The beach ball representation of DC moment tensor inversion results from Li *et al.* (2011a).

$$ISO = \frac{1}{3} \frac{Tr(m_{jk})}{M_0},$$
(4)

$$CLVD = 2\epsilon(1 - |ISO|), \tag{5}$$

$$DC = 1 - |ISO| - |CLVD|, \tag{6}$$

$$u' = \frac{1}{\sqrt{2}}(T+P), v' = \frac{1}{\sqrt{2}}(T-P).$$
(7)

Here λ_{\min} and λ_{\max} are the minimum and maximum eigenvalues of *M*; *T* and *P* correspond to the eigenvectors of the maximum and minimum eigenvalues, which indicate the maximum tensile and compressional stress directions; ISO, CLVD and DC denote the percentages of source components; *u'* is the fracture plane normal vector; and *v'* is the slip vector.

Before performing Bayesian inversion, we must create a forward operator that relates the moment tensor, the location of the seismic source, and the velocity model to the observed seismograms. To do so, we first construct a Green's function library, and calculate the synthetic seismograms for a point moment tensor source using the discrete wavenumber method (Bouchon 1981, 2003). The synthetic seismogram v_i^n of the *i*th component at the *n*th geophone at location x_r^n is modelled by

$$v_i^n(\mathbf{x}_r^n, \mathbf{x}_s, V, t) = \sum_{j=1}^3 \sum_{k=1}^3 M_{jk} G_{ij,k}^n(\mathbf{x}_r^n, \mathbf{x}_s, V, t) * s(t) + e_i^n(t),$$
(8)

where $G_{ij,k}(\mathbf{x}_r^n, \mathbf{x}_s, \mathbf{V}, t)$ is the spatial derivative of the Green's function of the *i*th component at the *n*th geophone of the location \mathbf{x}_r^n due to a point moment tensor source at the location \mathbf{x}_s using the 1-D velocity model \mathbf{V} , s(t) is the source time function, and $e_i^n(t)$ is the noise perturbation of the *i*th component at the *n*th geophone. Since \mathbf{M} is a symmetric matrix, we can represent it with a vector \mathbf{m} of six elementary moment tensor parameters, that is, $m_1 = M_{11}$, $m_2 = M_{22}$, $m_3 = M_{33}$, $m_4 = M_{12}$, $m_5 = M_{13}$, and $m_6 = M_{23}$. Concatenating all the seismograms and corresponding Green's functions (here the Green's function has been convolved with the source time function) and noise perturbations,

1966 C. Gu, Y.M. Marzouk and M.N. Toksöz

we rewrite eq. (8) as

Eq. (9) is then represented more compactly as

$$d = G(x, V)m + e, x \in \mathbb{R}^3, m \in \mathbb{R}^6$$

where x denotes the source location x_s in eq. (8), d is the concatenation of all the waveform vectors v_i^n , G(x, V) is the concatenation of all the synthetic seismograms of the six elementary moment tensor parameters, and e is the concatenation of all the noise vectors. The objective of Bayesian inversion is to infer the parameters m, x, and V in eq. (10) and to quantify their uncertainties, based on the waveform data d.

(10)

2.2 Bayesian formulation: prior, likelihood and posterior

This section details the Bayesian formulation of the inversion problem. Computational strategies for solving this problem are described in Section 2.3.

We begin by applying Bayes' rule to the parameters m, x, and V given data d:

$$P(m, x, V|d) = \frac{P(d|m, x, V)\pi_0(m)\pi_0(x)\pi_0(V)}{P(d)}.$$
(11)

Here $\pi_0(\mathbf{m})$, $\pi_0(\mathbf{x})$, and $\pi_0(\mathbf{V})$ are the prior probability densities of the moment tensor \mathbf{m} , location \mathbf{x} , and velocity model \mathbf{V} , respectively; $P(\mathbf{d}|\mathbf{m}, \mathbf{x}, \mathbf{V})$ is the likelihood function; $P(\mathbf{d})$ is the evidence (also called the marginal likelihood); and $P(\mathbf{m}, \mathbf{x}, \mathbf{V}|\mathbf{d})$ is the posterior probability density. As described in the introduction, given the available data, a model for the noise or errors in that data and any prior information about the parameters, the Bayesian posterior distribution quantifies uncertainty in \mathbf{m} , \mathbf{x} , and \mathbf{V} . Next we detail the key terms on the right-hand side of Bayes' rule.

The priors $\pi_0(\mathbf{m})$, $\pi_0(\mathbf{x})$ and $\pi_0(\mathbf{V})$ represent any information available about \mathbf{m} , \mathbf{x} and \mathbf{V} before performing the inversion. Here we use improper uniform probability distributions for \mathbf{m} and \mathbf{x} , that is, $\pi_0(\mathbf{m}) = \text{const}$, $\pi_0(\mathbf{x}) = \text{const}$. We use a proper uniform prior for \mathbf{V} , that is, $\pi_0(\mathbf{V}) = \text{constant}$ within a prescribed range and zero outside of that range. Improper priors are not integrable and are considered particularly 'non-informative' (Sivia & Skilling 2006); they can successfully be used in Bayesian inference as long as the posterior distribution is proper, which is guaranteed by the expressions below. The range for the proper prior on \mathbf{V} is justified in the context of specific examples below.

The form of the likelihood function P(d|m, x, V) follows directly from eq. (10), and depends on the probability distribution P_e of the additive error e:

$$P(d|\boldsymbol{m},\boldsymbol{x},\boldsymbol{V}) = P_e(\boldsymbol{d} - \boldsymbol{G}(\boldsymbol{x},\boldsymbol{V})\boldsymbol{m}).$$
⁽¹²⁾

In this paper we assume that the error is Gaussian, with zero mean and covariance matrix Σ_e :

$$\boldsymbol{e} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{e}}). \tag{13}$$

The covariance matrix Σ_e is allowed to be full, that is, not necessarily diagonal. The error *e* models any residual variation, including both the observation error and possible model inadequacy, for example, imperfections of the Green's function model resulting from any neglected physics (Kennedy & O'Hagan 2001; Kaipio & Somersalo 2007).

With the Gaussian probability distribution of *e*, the likelihood function can be written as

$$P(\boldsymbol{d}|\boldsymbol{m},\boldsymbol{x},\boldsymbol{V}) = \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{\Sigma}_e}} \exp\left[-\frac{1}{2}\left(\boldsymbol{d} - \boldsymbol{G}(\boldsymbol{x},\boldsymbol{V})\boldsymbol{m}\right)^T \boldsymbol{\Sigma}_e^{-1}\left(\boldsymbol{d} - \boldsymbol{G}(\boldsymbol{x},\boldsymbol{V})\right)\boldsymbol{m}\right],\tag{14}$$

where N is the total number of data samples used for inversion.

Given the above prior and likelihood function, the joint posterior probability density of m, x and V is

$$P(\boldsymbol{m}, \boldsymbol{x}, \boldsymbol{V}|\boldsymbol{d}) = cP(\boldsymbol{d}|\boldsymbol{m}, \boldsymbol{x}, \boldsymbol{V}),$$
(15)

where *c* is a normalization constant.

2.3 Algorithms for posterior sampling

-

A general approach for characterizing the posterior distribution of m, x, and V is Markov chain Monte Carlo (MCMC) sampling. The Metropolis-Hastings algorithm (Metropolis et al. 1953; Hastings 1970), a particular kind of MCMC scheme, constructs a Markov chain that asymptotically samples from the posterior. Metropolis-Hastings requires the ability to evaluate the joint posterior density of m, x and V up to a normalizing constant, exactly as specified in eq. (15). Although this is in principle sufficient to run a Metropolis-Hastings sampler, the linear dependence on m and nonlinear dependence on x and V of the Green's function G(x) result in a complex structure of the joint posterior distribution of *m*, *x* and *V*. Thus a generic MCMC approach will encounter difficulty: slow mixing, and an inaccurate characterization of the posterior. We instead design an MCMC approach that makes more careful use of the problem structure.

2.3.1 Direct sampling for m at fixed x and V

The first element of our approach is the ability to sample *m* directly from its full conditional distribution $P(m|d, x^*, V^*)$, without resorting to MCMC. Since, for a given x^* and V^* , the modelling waveform $G(x^*, V^*)m$ depends linearly on m, $P(m|d, x^*, V^*)$ can be determined analytically:

$$m|d, x^*, V^* \sim \mathcal{N}(\mu_m(d, x^*, V^*), \Sigma_m(d, x^*, V^*))$$
(16)

$$\mu_m(d, x^*, V^*) = \left[G^T(x^*, V^*) \Sigma_e^{-1} G(x^*, V^*) \right]^{-1} G^T(x^*, V^*) \Sigma_e^{-1} d$$
(17)

$$\Sigma_m(d, x^*, V^*) = \left[G^T(x^*, V^*) \Sigma_e^{-1} G(x^*, V^*) \right]^{-1}.$$
(18)

If one wishes to ignore uncertainty in the source location and velocity model, that is, to assume a fixed x^* and V^* , then this approach characterizes uncertainty in the moment tensor m under such an assumption. But it can also be used as a building block for an approach that jointly characterizes the uncertainty in m, x and V. We describe this approach next.

2.3.2 Marginal-then-conditional sampling of the joint distribution

To incorporate variation in x and V into the inverse problem, we first obtain the marginal posterior probability distribution $P(x^*, V^*|d)$ for any given x^* and V^* :

$$P(\mathbf{x}^*, \mathbf{V}^*|\mathbf{d}) \propto P(\mathbf{d}|\mathbf{x}^*, \mathbf{V}^*) = \int_{-\infty}^{\infty} P(\mathbf{d}|\mathbf{m}, \mathbf{x}^*, \mathbf{V}^*) \pi_0(\mathbf{m}) \mathrm{d}\mathbf{m} \,.$$
(19)

We substitute above for $P(d|m, x^*, V^*)$ using eq. (14) and obtain

$$P(d|\mathbf{x}^{*}, V^{*}) = \frac{1}{\sqrt{(2\pi)^{N} \det \Sigma_{e}}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2} (d - G(\mathbf{x}^{*}, V^{*})\mathbf{m})^{T} \Sigma_{e}^{-1} (d - G(\mathbf{x}^{*}, V^{*})\mathbf{m})\right] d^{6}\mathbf{m}$$

$$= \frac{\exp\left[-\frac{1}{2} \sum_{i=1}^{N} (\Sigma_{e}^{-1})_{ii} d_{i}^{2}\right]}{\sqrt{(2\pi)^{N} \det \Sigma_{e}}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2} \sum_{j,k=1}^{6} A_{j,k}m_{j}m_{k} + \sum_{j=1}^{6} B_{j}m_{j}\right] d^{6}\mathbf{m}$$

$$= \frac{\exp\left[-\frac{1}{2} \sum_{i=1}^{N} (\Sigma_{e}^{-1})_{ii} d_{i}^{2}\right]}{\sqrt{(2\pi)^{N} \det \Sigma_{e}}} \sqrt{\frac{(2\pi)^{6}}{\det A}} e^{\frac{1}{2}B^{T}A^{-1}B}$$

$$= \frac{\exp\left[-\frac{1}{2} \sum_{i=1}^{N} (\Sigma_{e}^{-1})_{ii} d_{i}^{2}\right]}{\sqrt{(2\pi)^{N-6} \det \Sigma_{e} \det A}} e^{\frac{1}{2}B^{T}A^{-1}B},$$

(20)

x = 6.4 km, y = 5.4 km, z = 1.0 km

Strike = 40°, Dip = 50°, Rake = 280° DC%: 61, CLVD%: 17, ISO%: 21, α = 10°



Figure 2. Synthetic data example of Section 3. (a) The configuration of source and stations distribution in the top and side views. The red star denotes the source, the green triangles denote the stations. (b) The beach ball shows the mechanism of the synthetic source. The green triangles denote the projection of the five stations on the source focal plane. (c) The detailed 1-D velocity profile from the ultrasonic logging is shown by solid red (*P*-velocity) and blue (*S*-velocity) lines. The light red and blue regions show ± 5 per cent perturbations of this velocity profile.

where

$$A = \sum_{i=1}^{N} \left(\Sigma_{e}^{-1} \right)_{ii} \begin{bmatrix} G_{i1}^{2} & G_{i1}G_{i2} & G_{i1}G_{i3} & G_{i1}G_{i4} & G_{i1}G_{i5} & G_{i1}G_{i6} \\ G_{i2}G_{i1} & G_{i2}^{2} & G_{i2}G_{i3} & G_{i2}G_{i4} & G_{i2}G_{i5} & G_{i2}G_{i6} \\ G_{i3}G_{i1} & G_{i3}G_{i2} & G_{i3}^{2} & G_{i1}G_{i4} & G_{i1}G_{i5} & G_{i3}G_{i6} \\ G_{i4}G_{i1} & G_{i4}G_{i2} & G_{i4}G_{i3} & G_{i4}^{2} & G_{i1}G_{i5} & G_{i4}G_{i6} \\ G_{i5}G_{i1} & G_{i5}G_{i2} & G_{i5}G_{i3} & G_{i5}G_{i4} & G_{i2}^{2} & G_{i5}G_{i6} \\ G_{i6}G_{i1} & G_{i6}G_{i2} & G_{i6}G_{i3} & G_{i6}G_{i4} & G_{i6}G_{i5} & G_{i6}^{2} \end{bmatrix},$$

$$(21)$$

and

$$\boldsymbol{B} = \sum_{i=1}^{N} \left(\boldsymbol{\Sigma}_{e}^{-1} \right)_{ii} d_{i} \begin{bmatrix} G_{i1} \\ G_{i2} \\ G_{i3} \\ G_{i4} \\ G_{i5} \\ G_{i6} \end{bmatrix}.$$
(22)

Since the waveform data d typically contain a very large number of samples compared to first-*P* amplitude data or *P* (or *S*) arrival time data, small perturbations in x^* and V^* can lead to enormous variations in the value of the marginal likelihood $P(x^*, V^*|d)$. The posterior thus tends to concentrate very tightly, and this situation is particularly vulnerable to model misspecification (i.e. errors in the assumed structure of the likelihood function or forward model), as described in Miller & Dunson (2015). When the posterior concentrates around one or a few points in the parameter space, any misspecification of the likelihood causes the posterior to underestimate uncertainty. For instance, in Mustać & Tkalčić (2016), waveform-based posterior location sampling concentrates at only a few locations for both synthetic and real data. To



Figure 3. Synthetic data example of Section 3: log marginal likelihood log $P_{\gamma}(d|x, V)$ as a function of the source location x, with V fixed to the true velocity model.

achieve more robust Bayesian inference in the current large-data setting, we introduce a *coarsened* marginal posterior probability distribution following the formulation of (Miller & Dunson 2015). This coarsened posterior simply tempers the marginal likelihood function:

$$P_{\gamma}(\mathbf{x}^{*}, \mathbf{V}^{*}|\mathbf{d}) \propto (P(\mathbf{d}|\mathbf{x}^{*}, \mathbf{V}^{*}))^{1/\gamma} \pi_{0}(\mathbf{x}^{*}) \pi_{0}(\mathbf{V}^{*}), \quad \gamma > 1.$$
(23)

The analytical expression in eq. (20) lets us evaluate the coarsened joint marginal posterior probability density P_{γ} at any x^* and V^* in the region of interest. We can then use MCMC to generate samples of x^* and V^* from $P_{\gamma}(x^*, V^*|d)$. For simplicity we will also sometimes refer to the coarsened likelihood $P_{\gamma}(d|x, V) := (P(d|x, V))^{1/\gamma}$.

Our MCMC scheme for these two variables can be viewed as a Metropolis-within-Gibbs algorithm that sequentially updates x^* and V^* within each iteration, holding one variable fixed at its current value while applying a Metropolis update to the other. To update x^* , we use the adaptive Metropolis (AM) method of Haario *et al.* (2001). The AM scheme adjusts the covariance matrix of the Metropolis proposal distribution after each step of the MCMC chain, based on all previous samples of x:

$$C_{n_0}^* = s_d \operatorname{Cov}(\mathbf{x}_0, \dots, \mathbf{x}_{n_0}) + s_d \epsilon_0 I_d .$$
⁽²⁴⁾

Here $C_{n_0}^*$ is the updated proposal covariance matrix at step n_0 , I_d is the *d*-dimensional identity matrix, $\epsilon_0 > 0$ is a small constant to make $C_{n_0}^*$ positive definite, d = 3 is the dimension of \mathbf{x} , and $s_d = 2.4^2/d$. The value of the scaling parameter s_d is a standard choice from Gelman *et al.* (1996), where it is shown to optimize the mixing properties of the Metropolis search. It is important to note that this value might affect the efficiency of MCMC, but not the posterior distribution itself. In practice, rather than letting \mathbf{x}^* vary continuously in \mathbb{R}^3 , we select a discrete tensor grid of possible \mathbf{x}^* values, uniformly distributed over a cube that contains the bulk of the posterior. Otherwise, if \mathbf{x}^* were left continuous, we would have to evaluate a new Green's function for each proposed value of the location, 'on the fly' during MCMC. Doing so would be computationally prohibitive.

Similar computational constraints apply to the velocity model V^* , but they are even more severe, as the velocity model is of much higher dimension (e.g. two velocities and 17 layers in the examples below). Since we cannot afford a regular grid of V^* values covering this space, we instead consider a Monte Carlo ensemble of discrete 1-D velocity models, generated from the uniform prior on V. In particular, our prior encodes perturbations of $\pm s$ per cent in the P and S velocities in each layer, where the perturbations are independent across layers. Thus, for



Figure 4. Synthetic data example of Section 3. (a) Log marginal likelihood log $P_{\gamma}(\boldsymbol{d}|\boldsymbol{x}^{p}, V^{q})$, for a discrete set of source locations and velocity models, indexed by p and q. (b) Log marginal likelihood log $P_{\gamma}(\boldsymbol{d}|\boldsymbol{x}, V^{q})$ as function of the source location \boldsymbol{x} for a particular realization of the velocity model, $V^{q=500}$. (c) Log marginal likelihood log $P_{\gamma}(\boldsymbol{d}|\boldsymbol{x}^{p}, V)$ for each velocity model in the prior ensemble, with the source fixed at $\boldsymbol{x}^{p=63} = [6.4, 5.4, 1.0]$ km.

the qth velocity model V^q in the ensemble, the P and S velocities at the *l*th layer are sampled from

$$V_{l,P}^{q} \sim \mathcal{U}\left(V_{l,P}^{0}(1-s \text{ per cent}), V_{l,P}^{0}(1+s \text{ per cent})\right)$$

$$V_{l,S}^{q} \sim \mathcal{U}\left(V_{l,S}^{0}(1-s \text{ per cent}), V_{l,S}^{0}(1+s \text{ per cent})\right), \qquad (25)$$

where $V_{l,P}^0$ and $V_{l,S}^0$ are the initial unperturbed *P* and *S* velocities. Setting up this ensemble can be thought of as replacing the continuous prior distribution with its discrete or empirical approximation. For posterior sampling, we then sample V^* values from this ensemble using Metropolis independence updates: choose an index randomly (proposal step) and accept or reject this velocity model using a Metropolis step.

For each sampled x^* and V^* , we then directly sample m from its Gaussian full conditional distribution (eq. 18). This overall algorithm is called marginal-then-conditional sampling (Fox & Norton 2015), and effectively yields a single Markov chain that explores P(m, x, V|d). The posterior distribution of m given d can be extracted by marginalization

$$P(\boldsymbol{m}|\boldsymbol{d}) = \int P(\boldsymbol{m}|\boldsymbol{x}^*, \boldsymbol{V}^*, \boldsymbol{d}) P_{\gamma}(\boldsymbol{x}^*, \boldsymbol{V}^*|\boldsymbol{d}) \mathrm{d}\boldsymbol{x}^* \mathrm{d}\boldsymbol{V}^*,$$
(26)

which is automatic given samples of m from the chain.

3 SYNTHETIC TEST

To validate our methods, we first applied our Bayesian inversion method to synthetic data. The configuration of the seismic source and stations is shown in Fig. 2. The synthetic source is located at x = [6.405, 5.405, 1.005] km. The source mechanism is set to be: Strike = 50°, Dip = 40°, Rake = 280°, DC per cent = 61 per cent, CLVD per cent = 17 per cent, ISO = 21 per cent, $\alpha = 10^\circ$. Data are generated assuming a layered velocity model obtained from well-logging, shown in Fig. 2(c). Gaussian noise perturbations with a relative standard deviation of 10 per cent are added to the synthetic data.

We apply three inversion procedures to the synthetic data: (1) direct sampling for m, with x fixed at the true location and V fixed at the true velocity model; (2) MCMC sampling for x coupled with conditional sampling for m, with V fixed at the true velocity model; (3) MCMC sampling for x and V, coupled with conditional sampling for m. These procedures allow for successively more comprehensive quantification of uncertainty, by lifting the 'truth' assumptions on x and V.



Figure 5. Synthetic data example of Section 3: MCMC chains and posterior distribution of x and V for inversion procedures #2 and #3. (a) MCMC chain (left) and histogram (right) of x sampled from the coarsened posterior distribution $P_{\gamma}(x|d)$, with velocity model fixed to the truth. The red lines on the right show a Gaussian fit to the histogram. We mark the starting location of each chain with a black dot. However, this point does not affect the observed distribution of the chain, since these chains mix well and are sufficiently long. (b) MCMC chain (left) and histogram (right) of x sampled from the coarsened posterior distribution $P_{\gamma}(x, V|d)$, where the velocity V is now uncertain. The red lines on the right show a Gaussian fit to the histogram. (c) MCMC chain (left) and histogram (right) of velocity model indices q, for V^q sampled from the posterior distribution $P_{\gamma}(x, V^q|d)$. (d) Posterior samples of V^q ; the blue and red regions show the sampled P- and S- velocity models and the black lines show the true P- and S- velocity model.

For procedure #2, we form a hypercube of possible locations centred on $x_0 = [6.4, 5.4, 1.0]$ km and fill it with a 19 × 19 × 19 grid, corresponding to a grid spacing of 10 m. We then pre-calculate the Green's functions for each possible location in this grid. Note that we do not include the true location in this set of locations, in order to avoid an inverse crime (Kaipio & Somersalo 2007). The location x is sampled according to the coarsened posterior (with likelihood $P_{\gamma}(d|x, V^*)$) with V^* fixed at the true velocity model. We use a value of $\gamma = 900$ for coarsening; the procedure for choosing γ is described in Appendix A. Fig. 3 shows the log marginal likelihood log $P_{\gamma}(d|x, V^*)$ as a function of the location x, for the fixed true velocity model. We see that it is largest in the vicinity of the true location.

For procedure #3, we generate 10^3 prior velocity models by randomly perturbing the true velocity model by s = 5 per cent, according to eq. (25); the range of these perturbations is illustrated by light red (P) and light blue (S) regions in Fig. 2(c). As above, we discretize the locations around $x_0 = [6.4, 5.4, 1.0]$ km with a $5 \times 5 \times 5$ grid, now with a grid spacing of 30 m, and pre-calculate the Green's functions for each possible location grid and velocity model. Again, to avoid an inverse crime, we do not include the true location in the location grid, nor do we include the true velocity model in the generated velocity ensemble. Then x and V are sampled via the algorithms described in Section 2.3.2, using the coarsened marginal likelihood $P_{\gamma}(d|x, V)$, again with $\gamma = 900$. Fig. 4(a) shows the resulting log marginal likelihood $\log P_{\gamma}(d|x^p, V^q)$ at the *p*th location x^p (p = 1, 2, ..., 125) with the *q*th velocity model V^q ($q = 1, 2, ..., 10^3$). Two cross-sections—one for fixed velocity with q = 500 and the other for fixed location with p = 63—are shown in Figs 4(b) and (c), respectively.

Figs 5(a) and (b) show the MCMC chains of source location x for procedures #2 and #3, respectively. The MCMC chain and posterior distribution of V constructed by procedure #3 are shown in Figs 5(c) and (d). A key observation is that posterior distribution of x is broader—that is, the MCMC samples a wider location range—when the velocity is varied than when the velocity is fixed at the true model. This means that the inclusion of velocity uncertainties in procedure #3 increases uncertainty in the location. Note that procedures #2 and #3



Figure 6. Synthetic data example of Section 3: MCMC chains and posterior distributions of m for the three inversion procedures. (a) MCMC chain of m for procedure #1, with source location and velocity model fixed to their true values; (b) MCMC chain of m for procedure #2, with the source location uncertain and the velocity model fixed to its true value; (c) MCMC chain of m for procedure #3, with the source location and velocity model uncertain. (d–f) Scatter plots of the components of m corresponding to the MCMC chain/inversion procedure immediately above.

employ the same level of coarsening, so this location uncertainty increase is entirely due to the inclusion of velocity uncertainties. Though the uncertainties are different, both posteriors are centred roughly on the true value of x.

Fig. 6 focuses on the posterior distribution of moment tensor parameters; it shows trace plots of the MCMC chain of *m* and scatter plots of the posterior distribution P(m|d), obtained via the three different inversion procedures. Direct sampling for *m*, with *x* and *V* fixed at the true location and true velocity model, shows the smallest uncertainty. Including uncertainty in the location *x*, with *V* fixed at the true velocity model, increases the posterior uncertainty of *m*. Allowing both *x* and *V* to be uncertain increases the marginal posterior uncertainty of *m* even further. Note that all three posteriors are centred on more or less the same values of *m*. The true value of *m*, used to generate the data, is $m = [2.08, 2.16, -1.70, -1.64, 0.52, -0.93] \times 10^{11}$ Nm. The posterior mean moment tensor parameters for procedures #1, #2 and #3 are shown in Table 1. The m_1 , m_2 and m_3 , the three diagonal elements of the moment tensor matrix representing three force couples, present larger uncertainty compared to m_4 , m_5 and m_6 , which are off-diagonal elements presenting three DCs. This observation indicates that the uncertainty of non-DC components is larger than DC components.

The computational cost of running 10^5 MCMC iterations is a few minutes for procedure #1 and about half an hour for procedures #2 and #3 on a MacBook laptop; these costs are negligible compared to the forward Green's function calculations for each station, with thousands of velocity models and possible source locations. For example, the computational time of the forward Green's function calculation is 3–4 d for five stations with a 5 × 5 × 5 location grid and 10^3 velocity models on the MIT engaging cluster.

Table 1. The source mechanism results for the synthetic event. Column 1 shows the true source parameters and Columns 2–7 show the posterior mean and standard deviation (std dev) source parameters for procedures #1, #2 and #3.

	True value	Procedure #1		Procedure #2		Procedure #3	
		Posterior mean	Posterior std dev	Posterior mean	Posterior std dev	Posterior mean	Posterior std dev
$m_1 (10^{11} \text{ Nm})$	2.08	2.12	0.04	2.04	0.21	1.69	0.34
$m_2 (10^{11} \mathrm{Nm})$	2.16	2.22	0.04	2.13	0.20	1.83	0.32
$m_3 (10^{11} \text{ Nm})$	-1.70	-1.66	0.04	-1.54	0.26	-1.24	0.34
$m_4 \ (10^{11} \ \text{Nm})$	-1.64	-1.63	0.02	-1.57	0.07	-1.39	0.16
$m_5 (10^{11} \mathrm{Nm})$	0.52	0.52	0.01	0.50	0.04	0.44	0.12
$m_6 (10^{11} \mathrm{Nm})$	-0.93	-0.93	0.01	-0.91	0.06	-0.84	0.11
$M_0 (10^{11} \mathrm{Nm})$	2.94	3.15	0.02	3.02	0.14	2.55	0.29
Strike (°)	50	49.6	0.3	49.8	1.0	49.8	3.4
Dip (°)	40	39.7	0.3	39.9	0.3	40.7	2.9
Rake (°)	280	279.5	0.4	280.1	2.0	281.6	4.9
DC (per cent)	61	61.4	1.0	60.3	4.8	55.4	8.0
CLVD (per cent)	17	16.2	0.7	17.1	2.6	21.4	5.2
ISO (per cent)	21	22.3	0.8	22.5	4.4	23.1	6.7
α (°)	10	9.5	0.4	10.2	1.8	13.2	3.6

MCMC chains (left) and the histogram (right) of x MCMC chains (left) and the histogram (right) of x with the true velocity model for the real data with the 5% perturbed velocity model ensemble for the real data 6.7 $\mu_{\rm F} = 6.518, \sigma_{\rm F} = 0.036$ 6.6 East (km) 6.5 6.4 = 6.600. σ₋ = 0.09 6.3 5.2 = 5.085, σ_N = 0.041 = 5.112, $\sigma_{\rm N}$ = 0.020 μ_N (k) 4 10 5.10 5.05 $\mu_{\rm V} = 0.926, \sigma_{\rm V} = 0.073$ = 0.974, $\sigma_{\rm V}$ = 0.038 1.1 Vertical (km) 0.9 0.8 100 2 0 2 6 8 100 MCMC iteration MCMC iteration ×10⁴ Frequency ×10⁴ (b) ×10⁴ Frequency (a) ×10⁴ The MCMC chain (left) and the posterior distribuition of V (right) The posterior distribuition of V presented in 1-D velocity profiles for the real data for the real data Vs V_P 1000 0 900 1600 Well-logging 500 800 Model 1400 ۳. П 1000 700 Velocity model index 1200 1500 600 Frequency (E 1000 500 <u></u> 2000 800 400 2500 600 300 3000 V 400 200 s 3500 200 100 4000 L 0 10 0 1000 2000 8 2 5 7 3 4 6 (d) (c) MCMC iteration $imes 10^4$ Frequency Velocity (km/s)

Figure 7. Real data example of Section 4: MCMC chains and posterior distribution of x and V for inversion procedures #2 and #3. The first 4×10^4 iterations of the MCMC chains for procedures #2 and #3 are discarded as burn-in. (a) MCMC chain (left) and histogram (right) of x sampled from the coarsened posterior distribution $P_{\gamma}(x|d)$, with velocity model fixed to the well logging model. The red lines on the right show a Gaussian fit to the histogram. We mark the starting location of each chain with a black dot. However, this point does not affect the observed distribution of the chain, since these chains mix well and are sufficiently long. (b) MCMC chain (left) and histogram (right) of x sampled from the coarsened posterior distribution $P_{\gamma}(x, V|d)$, where the velocity V is now uncertain. The red lines on the right show a Gaussian fit to the histogram. (c) MCMC chain (left) and histogram (right) of velocity model indices q, for V^q sampled from the posterior distribution $P_{\gamma}(x, V^q|d)$. (d) Posterior samples of V^q ; the blue and red regions show the sampled P- and S- velocity models and the black lines show the true P- and S- velocity model.



Figure 8. Real data example of Section 4: MCMC chains and posterior distributions of m for the three inversion procedures. The first 4×10^4 iterations of the MCMC chains for procedure #2 and #3 are discarded as burn-in. (a) MCMC chain of m for procedure #1, with source location and velocity model fixed to [6.50, 5.10, 0.97] km and the well logging obtained model; (b) MCMC chain of m for procedure #2, with the source location uncertain and the velocity model fixed to the well logging obtained model; (c) MCMC chain of m for procedure #3, with the source location and velocity model uncertain. (d–f) Scatter plots of the components of m corresponding to the MCMC chain/inversion procedure immediately above.

4 RESULTS FOR REAL DATA

Now we present inversion results with real field data. The data employed here are from an oil and gas field in the Sultanate of Oman, studied extensively by Sarkar (2008) and Li *et al.* (2011a,b). The events were located using the NonLinLoc algorithm, which utilizes a probabilistic, nonlinear, global search earthquake location algorithm (Lomax *et al.* 2000). The seismicity data are from the surface monitoring network shown in Fig. 1. For this network, five surface stations, instrumented with SM-6B geophones (125 Hz sampling rate), have been set up since 1999. The data used in the Oman studies consist of 800 events located at the surface network. This field is dominated by two fault systems: the northeast-southwest trending main faults and northwest–southeast trending auxiliary system. The accurate and detailed 1-D *P*-wave velocity models are from the sonic logging data (Sarkar 2008). The *S*-wave velocity models are from the double-difference tomography results (Zhang *et al.* 2009).

We selected one event from the surface monitoring network. The data for inversion are the vertical component seismograms from five stations, since the horizontal components are not accurate because of the unknown orientation of seismometers. We only fit a portion of the seismograms from each station. The fitted seismogram contains a *P*-segment which involves two periods before the *P*-wave arrival and four to eight periods after *P*-wave arrival, and an *S*-segment which involves 2 to six periods after the *S*-wave arrival. In this paper, we have not constructed an algorithm to select the windows automatically; we set up the window lengths manually in the input. (Automatic window selection would be an interesting topic to address in the future, to make the code production-ready.) An independent time shift for the *P* and *S* segments for each station is applied before Bayesian inversion to mitigate the effects of imperfect Green's functions (Zhu & Helmberger 1996).



Figure 9. Real data example of Section 4: source mechanism and waveform fitting results for the three inversion procedures. The first 4×10^4 iterations of the MCMC chains for procedures #2 and #3 are discarded as burn-in. (a) Left: the comparison of the mean posterior predicted (blue) and real (red) data for the separated *P*- and *S*- wave segments. The purple shading areas show the 10^5 posterior predicted waveforms. The mean and range of the variance reduction (VR) for each station are shown in the figure. The full records are plotted in light grey lines as a reference. The unit of the waveform is nm s⁻¹. All the waveforms are bandpass filtered from 3 to 8 Hz. Middle: the table of the statistics of source parameters. Right top: the focal plane projection of the mean source mechanism from procedure #1 at the fixed x = [6.50, 5.10, 0.97] km with fixed *V* from well logging. The dashed line shows the mean value of the fault plane solutions. The green triangles denote the five stations. Right bottom: the green region shows the uncertainty of the fault plane solutions and the tensile and compressional stress direction. (b and c) Panels for procedures #2 and #3.

For Bayesian inversion, to account for the maximal effects of background observational noise, we use a diagonal covariance $\Sigma_e = \sigma^2 I$ with σ equal to 10 per cent of the maximum amplitude over all the stations. The real background noise is at most a few percentage of the maximum amplitude over all the stations. We use diagonal noise covariance matrices in the main text; a discussion of the effects of non-diagonal noise covariance matrices is given in Appendix B. An additional inversion study employing *downsampled* data, with diagonal noise covariance matrices, is described in Appendix C. Procedures #2 and #3 are both applied with $\gamma = 25$, which is a smaller value than in the synthetic tests. The tuning of γ is described in Appendix A.

1976 C. Gu, Y.M. Marzouk and M.N. Toksöz

The three inversion and uncertainty quantification procedures described Section 3 are applied here to the real data. We show the MCMC chains and posterior distributions of x and V for sampling procedures #2 and #3 in Fig. 7. Note that, in the results of procedure #3, uncertainty in the source location broadens when uncertainty in V is included. The bottom plots in Fig. 7 show the posterior distribution of the velocity model. On the left (Fig. 7c), we see that many velocity models from the prior ensemble are sampled, yet the data favour some over others, as reflected by the sampling frequencies in the histogram. Fig. 7(d) shows the marginal posterior of the velocity at each depth; we see that there remains substantial uncertainty in the velocities, but that the posterior is largely centred on the well-logging model, which is depicted with a dashed line.

The MCMC chains and posterior distributions of m for the three sampling procedures are compared in Fig. 8. These results quantify the increase in the uncertainty of the moment tensor m as errors in the model parameters x and V are considered; these errors directly affect the Green's function calculations. Using the same waveform data, with the inclusion of more realistic uncertainties in the model, we predict more uncertainty for the moment tensor.

Fig. 9 shows a comparison between the real waveforms and the posterior predicted waveforms, as well as the recovered source parameters and their uncertainties; the three procedures correspond to Figs 9(a)-(c), respectively. The waveform matching, a table summarizing the inference of each source parameter, and a beach ball representing the source mechanisms are shown in the left, middle, and right panels of each row. For the left panels, we plot the posterior predicted waveforms as purple shaded areas and the mean predicted waveforms as bold red lines. The posterior predicted waveform region (purple areas) are narrow, and the mean posterior predicted waveforms (blue lines) match well with the real data (red lines) for the three procedures. The purple shaded areas increase in size from procedure #1 to #2 to #3. As shown in the tables, the standard deviations of the source parameters also increase from procedure #1 to #2 to #3. The fault plane geometry and DC components show relatively smaller uncertainties compared to isotropic and CLVD components; this observation applies to all three procedures. For the beach ball representation, the posterior sampled fault plane solution (green areas) has the same increasing trend from #1 to #2 to #3. These trends show that a wider range of moment tensor parameters can fit the data well, with the inclusion of the uncertainties of the location and velocity model. Note that the mean source mechanisms (the red beach balls in Figs 9a and b) for procedure #1 and #2 are close; however, the mean source mechanism (the red beach balls in Fig. 9c) for procedure #3 shows a significant difference. This observation shows that the uncertainty of velocity models can be a significant source of uncertainty in source mechanisms. Including more waveform data would mitigate the effects of location and velocity model uncertainties, reduce the uncertainties of the moment tensor parameters, and thus provide a more reliable moment tensor solution. In this study, however, we only have a very sparse surface network; with the inclusion of only 5 per cent uncertainties in the velocity model, we cannot tightly constrain the moment tensor parameters.

5 CONCLUSION

The Bayesian moment tensor inversion method works well for recovering the source mechanisms and location from the seismograms. This is validated by our synthetic study. Running a single Markov chain with our marginal-then-conditional sampling approach significantly reduces the dimension of the sampling problem and its computational costs. In addition, the Bayesian method naturally lets us extract uncertainties of the source parameters and source location from the posterior distributions of these parameters. In general, the fault geometry and DC component of the moment tensor are determined most accurately. The uncertainties of the isotropic and CLVD components are relatively larger than those of the DC components.

Based on the synthetic study and the study of an induced seismic event from an oil/gas field, we conclude that Bayesian uncertainty quantification of full moment tensor solutions is a useful tool for estimating how reliable a source mechanism model is. This paper also shows how to include velocity model uncertainties into a fully Bayesian inversion framework. We demonstrate that the inclusion of velocity uncertainties broadens the uncertainties of both the moment tensor and location, and provides more realistic uncertainty bounds.

ACKNOWLEDGEMENTS

We thank Petroleum Development Oman (PDO) for providing field data. We acknowledge the Kuwait Foundation for the Advancement of Sciences and the Kuwait-MIT Center for Natural Resources and the Environment for their support during this work.

REFERENCES

- Bouchon, M., 1981. A simple method to calculate Green's functions for elastic layered media, *Bull. seism. Soc. Am.*, **71**(4), 959–971.
- Bouchon, M., 2003. A review of the discrete wavenumber method, *Pure appl. Geophys.*, **160**(3–4), 445–465.
- Cespuglio, G., Campus, P. & Šílený, J., 1996. Seismic moment tensor resolution by waveform inversion of a few local noisy record—II. Application to the Phlegraean fields (southern Italy) volcanic tremors, *Geophys. J. Int.*, 126(3), 620–634.
- Du, J. & Warpinski, N.R., 2011. Uncertainty in FPSs from moment-tensor inversion, *Geophysics*, 76(6), WC65–WC75.
- Duputel, Z., Rivera, L., Fukahata, Y. & Kanamori, H., 2012. Uncertainty estimations for seismic source inversions, *Geophys. J. Int.*, **190**(2), 1243– 1256.
- Ford, S.R., Dreger, D.S. & Walter, W.R., 2009a. Identifying isotropic events using a regional moment tensor inversion, *J. Geophys. Res.*, 114(B1), 01306.
- Ford, S.R., Dreger, D.S. & Walter, W.R., 2009b. Source analysis of the memorial day explosion, Kimchaek, North Korea, *Geophys. Res. Lett.*, 36(21), L21304.
- Fox, C. & Norton, R.A., 2015. Fast sampling in a linear-Gaussian inverse problem, SIAM/ASA J. Uncertai. Quantification, 4, 1192– 1218.

- Gelman, A., Roberts, G.O. & Gilks, W.R., 1996. Efficient metropolis jumping rules, in *Bayesian Statistics 5*, pp. 599–608, eds Bernardo, J. M., Berger, J.O., Dawid, A. P. & Smith, A.F.M., Clarendon Press.
- Haario, H., Saksman, E. & Tamminen, J., 2001. An adaptive metropolis algorithm, *Bernoulli*, 7, 223–242.
- Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, 57(1), 97–109.
- 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-Sorêts (Alsace) in 2003 and their temporal and spatial variations, *Geophys. J. Int.*, **181**(3), 1547–1565.
- Kaipio, J. & Somersalo, E., 2006. Statistical and Computational Inverse Problems, vol. 160, Springer Science & Business Media.
- Kaipio, J. & Somersalo, E., 2007. Statistical inverse problems: discretization, model reduction and inverse crimes, *J. Comput. Appl. Math.*, **198**(2), 493–504.
- Kennedy, M.C. & O'Hagan, A., 2001. Bayesian calibration of computer models, J. R. Stat. Soc. B, 63(3), 425–464.
- Kolb, J.M. & Lekić, V., 2014. Receiver function deconvolution using transdimensional hierarchical Bayesian inference. *Geophys. J. Int.*, **197**(3), 1719–1735.
- Li, J., 2013. Study of Induced Seismicity for Reservoir Characterization, *PhD thesis*, Massachusetts Institute of Technology.
- Li, J., Sadi Kuleli, H., Zhang, H. & Nafi Toksöz, M., 2011a. Focal mechanism determination of induced microearthquakes in an oil field using full waveforms from shallow and deep seismic networks, *Geophysics*, 76(6), WC87–WC101.
- Li, J., Zhang, H., Sadi Kuleli, H. & Nafi Toksoz, M., 2011b. Focal mechanism determination using high-frequency waveform matching and its application to small magnitude induced earthquakes, *Geophys. J. Int.*, 184(3), 1261–1274.
- Lomax, A., Virieux, J., Volant, P. & Berge-Thierry, C., 2000. Probabilistic earthquake location in 3D and layered models, in *Advances in Seismic Event Location*, pp. 101–134, eds Thurber, C.H. & Rabinowitz, N., Springer.
- Maxwell, S. et al., 2014. Microseismic Imaging of Hydraulic Fracturing: Improved Engineering of Unconventional Shale Reservoirs, no. 17, SEG Books.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E., 1953. Equation of state calculations by fast computing machines, J. Chem. Phys., 21(6), 1087–1092.
- Miller, J.W. & Dunson, D.B., 2015. Robust Bayesian inference via coarsening, arXiv:1506.06101.
- Mustać, M. & Tkalčić, H., 2016. Point source moment tensor inversion through a Bayesian hierarchical model, *Geophys. J. Int.*, 204(1), 311– 323.
- Nayak, A. & Dreger, D.S., 2014. Moment tensor inversion of seismic events associated with the sinkhole at Napoleonville salt dome, Louisiana, *Bull. seism. Soc. Am.*, **104**(4), 1763–1776.
- Panza, G.F. & Saraò, A., 2000. Monitoring volcanic and geothermal areas by full seismic moment tensor inversion: are non-double-couple components always artefacts of modelling?, *Geophys. J. Int.*, **143**(2), 353–364.
- Sambridge, M., 1999. Geophysical inversion with a neighbourhood algorithm–II. Appraising the ensemble[J], *Geophys. J. Int.*, **138**(3), 727–746.
- Sarkar, S., 2008. Reservoir monitoring using induced seismicity at a petroleum field in Oman, *PhD thesis*, Massachusetts Institute of Technology.
- Shapiro, S.A., 2015. Fluid-Induced Seismicity, Cambridge Univ. Press.
- Šílený, J., Panza, G. & Campus, P., 1992. Waveform inversion for point source moment tensor retrieval with variable hypocentral depth and structural model, *Geophys. J. Int.*, **109**(2), 259–274.

- Šílený, J., Campus, P. & Panza, G., 1996. Seismic moment tensor resolution by waveform inversion of a few local noisy records—I. Synthetic tests, *Geophys. J. Int.*, **126**(3), 605–619.
- Šílený, J., Hill, D.P., Eisner, L. & Cornet, F.H., 2009. Non–double-couple mechanisms of microearthquakes induced by hydraulic fracturing, *J. geo*phys. Res., **114**(B8), B08307.
- Sipkin, S.A., 1982. Estimation of earthquake source parameters by the inversion of waveform data: synthetic waveforms, *Phys. Earth planet. Inter.*, 30(2–3), 242–259.
- Sivia, D. & Skilling, J., 2006. Data Analysis: A Bayesian Tutorial, OUP Oxford.
- Song, F., 2013. Microseismic mapping and source characterization for hydrofracture monitoring: a full-waveform approach, *PhD thesis*, Massachusetts Institute of Technology.
- Song, F. & Toksöz, M.N., 2011. Full-waveform based complete moment tensor inversion and source parameter estimation from downhole microseismic data for hydrofracture monitoring, *Geophysics*, 76(6), WC103– WC116.
- Stähler, S. & Sigloch, K., 2014. Fully probabilistic seismic source inversion-part 1: efficient parameterisation, *Solid Earth*, 5(2), 1055– 1069.
- Stähler, S.C. & Sigloch, K., 2016. Fully probabilistic seismic source inversion-part 2: modelling errors and station covariances, *Solid Earth*, 7(6), 1521–1536.
- Stierle, E., Bohnhoff, M. & Vavryčuk, V., 2014a. Resolution of nondouble-couple components in the seismic moment tensor using regional networks—II: application to aftershocks of the 1999 Mw 7.4 Izmit earthquake, *Geophys. J. Int.*, **196**(3), 1878–1888.
- Stierle, E., Vavryčuk, V., Šílený, J. & Bohnhoff, M., 2014b. Resolution of non-double-couple components in the seismic moment tensor using regional networks—I: a synthetic case study, *Geophys. J. Int.*, **196**(3), 1869–1877.
- Stuart, A.M., 2010. Inverse problems: a Bayesian perspective, *Acta Numer.*, 19, 451–559.
- Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation, SIAM.
- Templeton, D.C. & Dreger, D.S., 2006. Non-double-couple earthquakes in the long valley volcanic region, *Bull. seism. Soc. Am.*, 96(1), 69–79.
- Vavryčuk, V., 2001. Inversion for parameters of tensile earthquakes, J. geophys. Res., 106(B8), 16 339–16 355.
- Vavryčuk, V., 2011. Tensile earthquakes: theory, modeling, and inversion, J. geophys. Res., 116(B12), B12320.
- Vavryčuk, V., 2014. Iterative joint inversion for stress and fault orientations from focal mechanisms, *Geophys. J. Int.*, **199**(1), 69–77.
- Vavryčuk, V., 2015. Moment tensor decompositions revisited, J. Seismol., 19(1), 231–252.
- Vavryčuk, V., Bohnhoff, M., Jechumtálová, Z., Kolář, 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), 74–93.
- Warpinski, N.R. & Du, J., 2010. Source-mechanism studies on microseismicity induced by hydraulic fracturing, in SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers.
- Wolff, U., Collaboration, A. et al., 2004. Monte carlo errors with less errors, Comput. Phys. Commun., 156(2), 143–153.
- Zhang, H., Sarkar, S., Toksöz, M.N., Kuleli, H.S. & Al-Kindy, F. 2009. Passive seismic tomography using induced seismicity at a petroleum field in Oman, *Geophysics*, 74(6), WCB57–WCB69.
- Zhu, L. & Helmberger, D.V., 1996. Advancement in source estimation techniques using broadband regional seismograms, *Bull. seism. Soc. Am.*, 86(5), 1634–1641.

APPENDIX A: THE TUNING OF γ

As described in Section 2.3.2, standard Bayesian inference procedures can be sensitive to errors in the model of the data-generating process. These errors are called 'model misspecification' in the statistics literature. This sensitivity is often safely ignored when the data are few



Figure A1. The posterior ranges of P and S velocities in the real data example, defined by the velocity perturbation percentages s_P and s_S , as a function of the coarsening parameter γ . The full velocity range is preserved when $s_P = 5$ and $s_S = 5$.

in number, but large data sets—like the full waveform data used here—exacerbate this sensitivity and can make the Bayesian posterior particularly non-robust in the presence of model misspecification. In particular, posteriors in this setting tend to underestimate uncertainty, that is, to concentrate more than is warranted. To mitigate this problem, we employ the posterior coarsening method suggested by Miller & Dunson (2015) for robust Bayesian inference.

In eq. (23), we introduce a parameter γ to flatten the standard marginal likelihood P(d|x, V). The coarsened posterior associated with the likelihood $P_{\gamma}(d|x, V)$ yields samples of x and V that cover a geophysically realistic range, that is, tens of metres for the source location and perturbations of a few per cent in the 1-D velocity model. As a side benefit, coarsening can also improve the mixing of the MCMC chains for x and V.

The choice of γ is of course quite important in this context. If γ is too small, $P_{\gamma}(d|x, V)$ will still be very restrictive and cause the posterior on x and V to concentrate only on a few values. If γ is too large, $P_{\gamma}(d|x, V)$ will be too flat and the data will lose the ability to constrain x and V, which means that the marginal posterior distribution of x and V would revert to the prior. From a theoretical perspective, the exponent γ relates to the allowed discrepancy between the modelled and true data distributions (Miller & Dunson 2015). But in practice, the relationship between γ and the allowed discrepancy involves constants that are difficult to set *a priori*. We instead take a practical approach and set γ to the smallest value that does not shrink the range of velocities in the joint posterior. The posterior ranges of the P and S velocities are defined by the posterior velocity perturbation percentages s_P and s_S , as follows:

$$s_{P} = 100 * \min_{l} \left\{ \frac{\max_{i=n_{b}+1,\dots,N} |(V_{l,P}^{i} - V_{l,P}^{0})|}{V_{l,P}^{0}} \right\}$$

$$s_{S} = 100 * \min_{l} \left\{ \frac{\max_{i=n_{b}+1,\dots,N} |(V_{l,S}^{i} - V_{l,S}^{0})|}{V_{l,S}^{0}} \right\},$$
(A1)

where $V_{l,P}^i$ and $V_{l,S}^i$ are, respectively, the *P* and *S* velocities at the *l*th layer sampled at the *i*th MCMC step; n_b is the number of MCMC steps discarded as burn-in; *N* is the total number of MCMC steps; and $V_{l,P}^0$ and $V_{l,S}^0$ are the *P* and *S* velocities obtained from well logging.

We consider the posterior velocity range to be full when $s_P = s_S = 5$, corresponding to the 5 per cent velocity perturbations prescribed in both the synthetic and real-data examples (see eq. 25). Fig. A1 illustrates our γ tuning approach in the real field data case. We vary γ from 1 to 100 and calculate the posterior velocity perturbation percentages s_P and s_S . For $\gamma \ge 25$, the sampled range does not change significantly; for $\gamma = 20$ (and below), this range begins to shrink. We thus choose $\gamma = 25$, to allow the data to have maximum impact on the posterior while preserving the full range of velocities. Note that since this approach depends on having an uncertain velocity, we tune γ in the context of procedure #3 and then simply use the same value of γ in procedure #2.

APPENDIX B: NON-DIAGONAL NOISE COVARIANCE MATRIX

In this appendix, we test the impact of non-diagonal noise covariance matrices. Non-diagonal noise covariance matrices are obtained by fitting the average pre-event background noise (Sambridge 1999; Kolb & Lekić 2014; Mustać & Tkalčić 2016). Fig. B1 shows two different fittings of the autocorrelation of the average background noise, with the exponential function (Σ_e^1) and exponential cosine function (Σ_e^2) ,

$$(\Sigma_e)_{ij}^1 = \sigma^2 \exp\left(-\frac{|i-j|}{r_1}\right)$$
(B1)



Figure B1. (a) Autocorrelations of five pre-event noise series from five stations. The back dot line shows the average autocorrelation. (b) The exponential and exponential cosine fitting of the average autocorrelation. (c) The realistic covariance matrix generated from the best-fit exponential cosine function. Note that the sampling rate is 125 Hz.

$$\left(\Sigma_{e}\right)_{ij}^{2} = \sigma^{2} \exp\left(-\frac{|i-j|}{r_{2}}\right) \cos\left(-2\pi \frac{(i-j)}{L_{2}}\right),\tag{B2}$$

where (i - j) is the time lag between sample *i* and sample *j*, r_1 is the decay factor of the exponential function, r_2 is the decay factor of the exponential cosine function, and L_2 is the period of the exponential cosine function. We use the best-fit exponential cosine function to generate the non-diagonal error covariance matrix, as illustrated in Fig. B1.

Using the non-diagonal error covariance matrix of Fig. B1, we perform the inversion procedures described in the main text. We plot the source mechanism results and their uncertainties in Figs B2–B4. With the same noise level as the diagonal error covariance matrix (such that σ in eq. B2 equals 10 per cent of the maximum amplitude over all the stations), introducing the non-diagonal error covariance matrix results in the same increasing trend in the uncertainty of *m* from inversion procedure #1 to #2 to #3. The mean posterior source mechanisms from inversion procedures #1 and #2 are similar.

One interesting observation is that although the predicted source mechanisms differ somewhat between the diagonal and non-diagonal noise covariance matrices for inversion procedures #1 and #2, the predicted source mechanisms for these two noise covariance matrices are very similar under inversion procedure #3. The non-diagonal matrix weighs adjacent waveform matching errors based on the autocorrelation of the background noise. The discrepancy between the posterior predicted source mechanisms with the diagonal and non-diagonal noise covariance matrices shows that this weighing can be important. But the similarity of the predicted source mechanisms for the two covariance matrices using inversion procedure #3 indicates that the impact of this weighing on the uncertainties of the source mechanisms is small compared to the impact of velocity model uncertainty, which has been discussed in the main text.



MCMC chains (left) and the histogram (right) of ${\bf x}$ with the 5% perturbed velocity model ensemble for the real data and noise



Figure B2. Real data example with non-diagonal noise covariance matrix: MCMC chains and posterior distribution of x and V for inversion procedures #2 and #3. (a) MCMC chain (left) and histogram (right) of x sampled from the coarsened posterior distribution $P_{\gamma}(x|d)$, with velocity model fixed to the well logging model. The red lines on the right show a Gaussian fit to the histogram. We mark the starting location of each chain with a black dot. However, this point does not affect the observed distribution of the chain, since these chains mix well and are sufficiently long. (b) MCMC chain (left) and histogram (right) of x sampled from the coarsened posterior distribution $P_{\gamma}(x, V|d)$, where the velocity V is now uncertain. The red lines on the right show a Gaussian fit to the histogram. (c) MCMC chain (left) and histogram (right) of velocity model indices q, for V^q sampled from the posterior distribution $P_{\gamma}(x, V|d)$. (d) Posterior samples of V^q ; the blue and red regions show the sampled P- and S-velocity models and the black lines show the true P- and S-velocity model.



Figure B3. Real data example with non-diagonal noise covariance matrix: MCMC chains and posterior distributions of m for the three inversion procedures. (a) MCMC chain of m for procedure #1, with source location and velocity model fixed to their true values; (b) MCMC chain of m for procedure #2, with the source location uncertain and the velocity model fixed to its true value; (c) MCMC chain of m for procedure #3, with the source location and velocity model uncertain. (d–f) Scatter plots of the components of m corresponding to the MCMC chain/inversion procedure immediately above.



Figure B4. Real data example with non-diagonal noise covariance matrix: source mechanism and waveform fitting results for the three inversion procedures. (a) Left: the comparison of the mean posterior predicted (blue) and real (red) data for the separated *P*- and *S*-wave segments. The purple shading areas show the 10⁵ posterior predicted waveforms. The mean and range of the variance reduction (VR) for each station are shown in the figure. The full records are plotted in light grey lines as a reference. The unit of the waveform is nm s⁻¹. All the waveforms are bandpass filtered from 3 to 8 Hz. Middle: the table of the statistics of source parameters. Right top: the focal plane projection of the mean source mechanism from procedure #1 at the fixed x = [6.50, 5.10, 0.97] km with fixed *V* from well logging. The dashed line shows the mean value of the fault plane solutions. The green triangles denote the five stations. Right bottom: the green region shows the uncertainty of the fault plane solutions and the tensile and compressional stress direction. (b and c) Panels for procedure #2 and #3.

APPENDIX C: DOWNSAMPLED DATA

In this appendix, we investigate the impact of data downsampling on the inversion results. We downsample the input data from 125 to 25 Hz and apply the three inversion procedures to the downsampled data, using the diagonal error covariance matrix defined in Section 4. Via the tuning method described in Appendix A, we find that a value of $\gamma = 7$ preserves the full velocity range. Downsampling the data thus reduces



Figure C1. Real data example with downsampling to 25 Hz: MCMC chains and posterior distribution of x and V for inversion procedures #2 and #3. The first 2.5 × 10⁴ iterations of the MCMC chains for procedures #3 are discarded as burn-in. (a) MCMC chain (left) and histogram (right) of x sampled from the coarsened posterior distribution $P_{\gamma}(x|d)$, with velocity model fixed to the well logging model. The red lines on the right show a Gaussian fit to the histogram. We mark the starting location of each chain with a black dot. However, this point does not affect the observed distribution of the chain, since these chains mix well and are sufficiently long. (b) MCMC chain (left) and histogram (right) of x sampled from the coarsened posterior distribution $P_{\gamma}(x, V|d)$, where the velocity V is now uncertain. The red lines on the right show a Gaussian fit to the histogram. (c) MCMC chain (left) and histogram (right) of velocity model indices q, for V^q sampled from the posterior distribution $P_{\gamma}(x, V^q|d)$. (d) Posterior samples of V^q ; the blue and red regions show the sampled P- and S-velocity models and the black lines show the true P- and S-velocity model.

the value of the coarsening parameter γ ; this is expected because downsampling reduces the size of the data set and thus the impact of model misspecification. The source mechanism results and uncertainties are shown in Figs C1–C3.

Similar to the results obtained with the original (non-downsampled) data, results obtained with downsampled data and inversion procedures #1, #2 and #3 show that the inclusion of location uncertainties increases the uncertainty of the moment tensor, and that inclusion of velocity uncertainties increases the uncertainties of both the moment tensor and the location. For inversion procedures #1 and #2, the source mechanisms obtained with downsampled data and a diagonal error covariance matrix differ from those obtained with the original non-downsampled data and either a diagonal or non-diagonal covariance matrix. Similarity of the predicted source mechanisms for all three data/covariance scenarios under inversion procedure #3, however, indicates that velocity model uncertainty has a larger impact than data sampling and weighing.



Figure C2. Real data example with downsampling to 25 Hz: MCMC chains and posterior distributions of m for the three inversion procedures. The first 2.5×10^4 iterations of the MCMC chains for procedure #3 are discarded as burn-in. (a) MCMC chain of m for procedure #1, with source location and velocity model fixed to its true values; (b) MCMC chain of m for procedure #2, with the source location uncertain and the velocity model fixed to its true value; (c) MCMC chain of m for procedure #3, with the source location and velocity model uncertain. (d–f) Scatter plots of the components of m corresponding to the MCMC chain/inversion procedure immediately above.



Figure C3. Real data example with downsampling to 25 Hz: source mechanism and waveform fitting results for the three inversion procedures. The first 2.5×10^4 iterations of the MCMC chains for procedure #3 are discarded as burn-in. (a) Left: the comparison of the mean posterior predicted (blue) and real (red) data for the separated *P*- and *S*-wave segments. The purple shading areas show the 10^5 posterior predicted waveforms. The mean and range of the variance reduction (VR) for each station are shown in the figure. The full records are plotted in light grey lines as a reference. The unit of the waveform is nm s⁻¹. All the waveforms are bandpass filtered from 3 to 8 Hz. Middle: the table of the statistics of source parameters. Right top: the focal plane projection of the mean source mechanism from procedure #1 at the fixed x = [6.50, 5.10, 0.97] km with fixed *V* from well logging. The dashed line shows the mean value of the fault plane solutions. The green triangles denote the five stations. Right bottom: the green region shows the uncertainty of the fault plane solutions and the tensile and compressional stress direction. (b and c) Panels for procedures #2 and #3.