



1

# 2 An MCMC Bayesian full moment tensor inversion constrained

## 3 by first-motion polarities and double couple percent

- 4
- 5 Mehrdad Pakzad<sup>1</sup>, Mahnaz Khalili<sup>2</sup>, Shaghayegh Vahidravesh<sup>2</sup>
- 6 <sup>1</sup>Institute of Geophysics, University of Tehran, Tehran, 1435944411, Iran
- 7 <sup>2</sup>Graduate Student of Geophysics, Institute of Geophysics, University of Tehran, Tehran, 1435944411, Iran
- 8 Correspondence to: Mehrdad Pakzad (pakzad@ut.ac.ir)
- 9

10 Abstract. Monte Carlo Markov chain (MCMC) samplings can obtain a set of samples by directed random walk, 11 mapping the posterior probability density of the model parameters in Bayesian framework. We perform earthquake 12 waveform inversion to retrieve focal angles or the elements of moment tensor and source location using a Bayesian 13 MCMC method with the constraints of first-motion polarities and double couple percentage using full Green functions 14 and data covariance matrix. The algorithm tests the compatibility with polarities and also checks the double couple 15 percentage of every site before the time-consuming synthetic seismogram computation for every sample of moment 16 tensor of every trial source position. Other than large earthquakes, the method is especially suitable for weak events 17 (M < 4) that their focal mechanisms cannot be well-constrained by polarities or seismograms alone, unless a dense 18 local network is available; something that is generally occasional. Two- and one-station solutions show more 19 agreement with all-station solution if polarity and DC% constraints are employed. In order to examine the validity of 20 the method, two events with the independent focal mechanism solutions are utilized. Furthermore, we also calculate 21 data covariance matrix from pre-event noise and Green function uncertainty to obtain the errors of focal mechanisms. 22

### 23 1 Introduction

It should be taken into consideration that most of the time due to lack of recording or noisy content of the records in long epicentral distances, determination of the focal mechanisms of weak events are difficult. Moreover, the signal to noise ratio (SNR) for microseismic events at long periods is low, therefore, small events have to be investigated at high frequencies. Accordingly, more high frequency velocity models are required. A suite of approaches has been introduced to tackle the issues, some of which utilize a Markov chain Monte Carlo (MCMC) method in Bayesian framework.

30 Among other methods, two-step method of Šílený et al. (1992) consists of an iteration of a linear inversion step with

31 fixed depth and velocity model and successive perturbation of both inside a set, bounded between two depths and two

32 structural models. Mao et al. (1994) used their method for high frequency data up to 10 Hz. They achieved this by

33 Green's functions calculation for an inhomogeneous medium with detailed structure. Wéber studied low-magnitude

34 earthquakes through a series of papers. His probabilistic procedure solves the nonlinearity problem of using

- 35 hypocentral location as model parameter. Routinely determined locations are usually not accurate enough in short
- 36 epicentral distances where the weak events are recorded. A priori hypocenter distribution is given by observed arrival





37	times, and it is employed in a Bayesian formulation with likelihood function constructed by observed waveforms, then
38	the posterior hypocenter distribution is mapped by octree importance sampling (Lomax and Curtis, 2001). The
39	posterior probability density function (hereafter PPD) of moment tensor rate functions are sampled by a large number
40	of bootstrapped data sets with the rate functions linearly inverted using hypocenters randomly chosen from the
41	posterior hypocentral probability density function (Wéber, 2006). Stähler and Sigloch (2014) proposed a probabilistic
42	framework that samples earthquake depth, moment tensor (MT), and source time function with the neighborhood
43	algorithm. Mustać and Tkalčić (2016) used two chains approach for sampling location and MT parameters. They also
44	treated noise as a free parameter in the inversion. Ito et al. (2016) estimated the probability density functions of fault
45	parameters using MCMC method for the 2004 Sumatra-Andaman earthquake. Gu et al. (2018) applied one Markov
46	chain technique for their waveform-based Bayesian full moment tensor inversion for small earthquakes. They
47	performed source relocation, full moment tensor inversion and uncertainty analysis. In their study, Marginal-then-specific sampling strategy
48	conditional sampling of the joint distribution was first obtained for any given location and velocity model, then forbe stated like that
49	each sampled location and velocity model they directly sampled MT from its conditional distribution. Wéber (2018)
50	introduced a method called JOWAPO (joint waveform and polarity) inversion. The method constructs a posterior
51	probability density of strike, dip and rake and maps it by octree importance sampling. The PPD consists of a null a
52	priori information and two likelihood functions for polarities and waveforms. For the details about the polarity would be introduction
53	likelihood refer to Brillinger, 1980; Walsh et al. 2009 and Wéber 2018. Comparing to waveform data, the information but rather in a method section. Or its structure
54	content of first-motion polarities of body waves is low, that is why a dense coverage of focal sphere is required for a could be changed to keep in introduction.
55	reliable result. On the other hand, for high frequency weak events, available velocity distributions are usually not What is that?
56	detailed enough to model their waveforms and retrieve the focal mechanisms, that is, waveforms can be modelled velocity model?
57	convincingly just for relatively close stations to receive a quite dependable focal mechanisms solution for near station
58	earthquakes. However, seismic networks are not usually dense enough to make sufficient data available for inversion. Language!
59	Therefore, combining polarity data with near-station records can be helpful. In the case of a small event occurrence
60	and with low number of stations, the objective cannot be more than to retrieve its DC focal mechanism with the
61	uncertainty. Earthquakes source inversion is relevant to the location determination and also velocity models.
62	Uncertainty in both the model parameters (here DC mechanisms), first motion observations and seismic waveform
63	should be merged by an inversion technique. In this regard, the most suitable inversion method is Bayesian sampling
64	producing an ensemble of DC focal mechanisms based on the posterior probability distribution. (Wéber, 2018).
65	As to the constraints, various methods adopted them for retrieving focal mechanisms of weak events in sparse
66	networks. For example, the phase and waveform amplitude can be combined with the first-motion P polarities and
67 if the 68	average S/P amplitude ratios (Li et al., 2011). The focal mechanisms <u>obtained</u> by a broad set of the first-motion focal mechanisms have been obtained they have been constrained up to some degree already polarities can be constrained by a single-station waveform inversion (Fojtiková and Zahradník, 2014). In this study
69	we perform waveform inversion but constrain it by first motion polarities and DC% for tectonic earthquakes. The
70	method can work with strike, dip and rake and also for the elements of MT as the model parameters; therefore for non-
71	tectonic earthquakes, DC% constrain can be eliminated. Here we describe full moment tensor and location inversion.
72	In the following sections, after a brief introductory overview of the intended methods used in this study, the performed
73	synthetic tests are described. It is preceded with the testing the method on two earthquakes in Switzerland and Iran.





74		
75	2 Method	
76		
77	The PPD is computed using the Bayesian rule to the parameters $\boldsymbol{m}$ , that can be strike, dip and rake or elements of MT	
78	and $\boldsymbol{x}$ , the location; given polarities, $\boldsymbol{P}$ and waveforms, $\boldsymbol{d}$	
79		
80	$\sigma(\boldsymbol{m}, \boldsymbol{x}   \boldsymbol{P}, \boldsymbol{d}) \propto \rho(\boldsymbol{x}) \rho(\boldsymbol{m}) L_{\boldsymbol{P}}(\boldsymbol{P}   \boldsymbol{m}, \boldsymbol{x}) L_{\boldsymbol{d}}(\boldsymbol{d}   \boldsymbol{m}, \boldsymbol{x}), $ (1)	
81		
82	where, $\rho(\mathbf{x})$ and $\rho(\mathbf{m})$ are the prior information about $\mathbf{x}$ and $\mathbf{m}$ ; $L_p(\mathbf{P} \mathbf{m}, \mathbf{x})$ and $L_d(\mathbf{d} \mathbf{m}, \mathbf{x})$ are the likelihood	
83	functions for polarities and waveforms. The uniform distribution assumptions are considered for both of the prior	
84	probability densities, that is, all trial locations have equal chance before considering data and the boundary values of	
85	the coefficients of elementary seismograms are set to -1.5 and 1.5. Unlike Wéber (2018), we only benefit from the	how are these bounds determined? need to intro-
86	this is a strong assumption, but unclear from the text and needs better description reliable polarities as constraints for the inversion, therefore we consider $L_P(P m, x)$ to be equal to one. The Gaussian	duce the source parameter- isation first. A common parameterisation is +-sqrt(2) in the trace and
87	model waveform likelihood is given by	
88		+-1 in the off-diagonals (e.g. Vavrycuk 2015
89	$L_d(\boldsymbol{d} \boldsymbol{m},\boldsymbol{x}) \propto \exp[-\frac{1}{2} \left(\mathbf{G}(\boldsymbol{x})\boldsymbol{m} - \boldsymbol{d}\right)^T \mathbf{C}_{\mathrm{D}}^{-1} \left(\mathbf{G}(\boldsymbol{x})\boldsymbol{m} - \boldsymbol{d}\right)], \tag{2}$	decompositions revisited, or Staehler et al 2014
90	how is G calculated? why spatial derivative?	Fully probabilistic seismic source inversion
91	where $\mathbf{G}(\mathbf{x})$ is the spatial derivative of the Green's function at the source location $\mathbf{x}$ and $\mathbf{C}_{\mathrm{D}} = \mathbf{C}_{\mathrm{d}} + \mathbf{C}_{\mathrm{T}}$ , that is,	<ul> <li>Part 1 :</li> <li>Efficient parameterisation</li> </ul>
92	waveform uncertainties and theoretical uncertainties combined by adding the respective covariance operators to obtain	
93	need to be more precise, it is data error and theory error due to mathematical formulation of the problem the total covariance matrix (Tarantola, 1987).	
94	The inverse problem is linear in $m$ and nonlinear in $x$ , which results in complex structure of the joint posterior	
95	distribution of the model parameters. In their waveform-based Bayesian full moment tensor inversion, Gu et al. (2018)	
96	designed an MCMC approach to incorporate variation in $\boldsymbol{x}$ into the problem. They first obtain the marginal posterior	
This whole para- 97 graph is unclear	probability distribution $\sigma(x^* d)$ for any given x and use it to calculate the Metropolis acceptance ratio. The adaptive	
with misfortunate 98 use of terminology	Metropolis method of Haario et al. (2001) is used to draw a new proposal model $x$ . Then for each sampled $x$ , they	
and mixing of 99 concepts	directly sample $m$ from its analytical covariance matrix. The algorithm is called marginal-then-conditional sampling	repetition, of the work of Gu et al. 2018
100	(Fox & Norton, 2015) that only needs one Markov chain to explore the posterior probability density. Employing	rewriting
101	polarities in the inversion also makes finding $m$ nonlinear. We implement two chains for sampling location and MT	
102	parameters. The second chain to sample MT is inside the first one which samples location. The procedure to sample	
103	$\boldsymbol{x}$ in the first chain is the same as used in Gu et al. (2018), that is a metropolis test which is used to determine whether	
104	to accept or reject a trial $x$ according to marginal posterior distribution for any sampled $x$ without reference to the	
105	values of MT; but for the inner sampling of $m$ , we explore $L_d(d m, x^*)$ in Eq. (2) by Metropolis-Gibbs sampler	
106	described by Lomax et al. (2000). They employ Metropolis-Gibbs Sampling algorithm for probabilistic earthquake	
107	location in 3D space (NonLinLoc program), here we use it in 6D space to retrieve MT. The procedure explores the	
108	PPD by directed walk towards high likelihood regions. The new walk site $m_{new}$ is obtained from the current site	
109	$m_{curr}$ , by adding a vector of arbitrary direction $dm$ , with length $l$ . The new site is accepted, if $\sigma(m_{new}, x^* d) \ge 1$	
110	$\sigma(\mathbf{m}_{curr}, \mathbf{x}^*   \mathbf{d})$ , otherwise the new site is accepted with probability $\sigma(\mathbf{m}_{new}, \mathbf{x}^*   \mathbf{d}) / \sigma(\mathbf{m}_{curr}, \mathbf{x}^*   \mathbf{d})$ . In order to	



The polarity

appendix or supplement.

The polarity

constraint is

to the other

publications.

constraint needs to be



111 achieve a good coverage of PPD, determination of the step size l is essential. The algorithm does this adaptively in 112 stages. In the first stage called the learning stage, the step size is constant and relatively large enough to explore all 113 the solution space and to wander towards high likelihood regions. In the equilibration stage, the searching of high likelihood regions can continue or these regions may begin to be searched for the optimum point. To achieve that, l is 114 this section is again a mixture of previous work and work related to this set equal to  $f_s(S_{m1}S_{m2}S_{m3}S_{m4}S_{m5}S_{m6}/N_s)^{1/6}$ , where  $f_s = 16$  is the scaling factor and S stands for standard 115 article. It needs better 116 deviation.  $N_{\rm s}$  is the number of samples to be accepted during the saving stage. In the final saving stage the step size is structure to see whats new 117 fixed at its final value from the previous stage and the walk can continue to explore high likelihood regions (Lomax 118 et al., 2000). The polarity constraint and arbitrary DC% condition for tectonic earthquakes are applied inside the 119 second chain, that is, after sampling MT, the polarity and DC% tests are performed for compliance and if the conditions stated with an equation. If it are fulfilled, then the metropolis test is performed, otherwise the sample is rejected. has been done 120 before (stated in introduction) 121 We employ Vackář et al. (2017) method for calculating  $C_d$ . Data covariance matrix is constructed from pre-event it needs a shor 122 noise which allows an automated weighting of the records according to their SNR. In other words, it plays the role of 123 automated frequency filter containing noisy frequency ranges in the frequency domain. The noise generation is claimed to be 124 supposed to be a random Gaussian stationary process. Therefore, with additional ergodicity assumption taken into an important part of the worl 125 account, the covariance function is estimated from a time series autocorrelation. This matrix can be assigned to one and the author cannot rely on the 6 station by calculating the covariance function from the cross correlations of three components. This way, each station have nine matrix blocks. It can also be assumed that noises at distant stations for high frequencies are not correlated, 127 128 so that the off-diagonal blocks in the main covariance matrix have zero values. The other source of error is theoretical. 129 Green function uncertainty is mostly related to the random time shifts of the data, accordingly this feature can be 130 employed to obtain approximate covariance matrix for  $C_T$  (Hallo and Gallovič, 2016; Hallo et al., 2017).  $C_d$  takes precedence for weaker earthquakes (with significant uncorrelated noise) that is while  $C_T$  is dominant for stronger 131 That is again a strong (uncorrelated noise free) earthquakes. In the following we apply the method to two earthquakes with M<sub>W</sub> 3.7 and 3.8; 132 statement to make, but there is no reference given. Whether the 133 for both,  $C_d$  is dominant. theory errors are dominant depends mostly on how wel 134 The computational cost of running the code on a 2.60 GHz Dual-Core CPU, 4G memory PC, for 1000 iterations in the Earth structure i.e. the velocity model is known ( 135 location chain and  $10^5$  iterations in MT chain, is less than a minute to a few minutes for each  $\gamma$  explained in the for seismic waveform data). It is thus often very 136 synthetic test section below. The speed conversely depends on the number of station/components and is proportional complicated to even seperate these two 137 to the number of restrictive polarities. For example, in one of the applications below with 24 seismogram components, components. 14 polarities, 9261 potential sources, and the starting point at the farthest corner of the location grid, the time is about 138 139 3.5 minutes. Without the constraints the time may increase even to hours.

which constraints?

141 **3** Synthetic test

142

- 143 We perform several synthetic tests to confirm the validity of the method. The configuration of the stations in the
- 144 synthetic tests is identical to the recording stations of Sargans earthquakes used also as an application (Fig. 1).
- 145 The elements of MT in NED coordinate system used are as follows:  $m_{xx} = -2.7645e+16$ ,  $m_{yy} = 3.2959e+15$ ,  $m_{zz} = -2.7645e+16$ ,  $m_{yy} = 3.2959e+15$ ,  $m_{zz} = -2.7645e+16$
- 146 2.4349e+16,  $m_{xy} = 1.1381e+18$ ,  $m_{xz} = 1.8408e+17$ ,  $m_{yz} = 3.6964e+17$ , with about 86%, 14% and 0%, DC, CLVD
- 147 and isotropic components. The strike, dip and rake are equal to 89.05°, 72.74° and 171.82° successively. The synthetic



177

178



(3)

148 location is located at x = (1 km N, -1 km E, 6 km down) with respect to Sargans earthquake epicenter. We used 9261 149 trial sources with equal step of 1 km from -10 to 10 km for horizontal coordinates and between 1 and 21 km for depth. 150 The results with different Signal-to-noise ratio (SNR) are shown in Table 1-3. SNR is here defined as the power of 151 signal divided by the power of white noise. A Butterworth filter with the frequency range 0.02 - 0.15 Hz is applied to 152 both the noise and synthetic data. The inversion is performed with the same velocity model as used to produce the 153 synthetic data. In the tables, beachballs of the solutions (in red) are illustrated with the true mechanism used for 154 creating synthetic seismograms in green color. Kagan angles (Kagan, 1991) are the angles of rotation between two 155 nodal planes of the solutions and the true mechanism. In Table 1 we utilize all stations in the inversion while in Tables 156 2 and 3 we just used two nearby stations, LIENZ and SGT04. Table 2 and 3 differ in using the constraints of polarity 157 and DC% > 70 in Table 2. The results presented in Table 1 and Table 2 are more close to each other; although in the 158 latter, we just used two stations. On the other hand, the results in Table 3 shows that the solutions deteriorate more, in 159 terms of Kagan angles and deviatoric part due to the lack of polarity and DC% constraints. For example, for SNR 160 equal to 0.5, Kagan angle is 8°, in case of applying the constraints, while it increases to 30° otherwise. For the case of 161 using all stations, we calculate the location as model parameter (Table 1) while for two-station cases we fix the location 162 to the one obtained for all-station computation (Table 2 and 3). 163 The outer chain consists of drawing samples by the adaptive Metropolis method and calculating the marginal posterior 164 probability for any given location and performing the acceptance test which is a Metropolis test. The iteration is repeated for 1000 times, however after few hundred steps, the optimum location is found. In the synthetic test without 165 166 noise, 38 iterations were enough for location parameters to converge. Similar to Mustać and Tkalčić (2016) the visited 167 potential locations, as well as the accepted location solution with the increasing likelihood and the optimum one are 168 shown in Fig. 2. Both the 3D and 2D views are illustrated. The starting search point is x = (0, 0, 10) representing the 169 beginning of the lines connecting the accepted solutions; finally ending with the maximum a posteriori solution 170 encircled by green squares in 2D views. In 3D view, the accepted Metropolis locations are drawn by green cubes. As 171 is presented in the figure, the concentration of high probability sites (larger cubes and squares) are around the optimum 172 solution, and the accepted solutions find their ways around it.

173 We applied the posterior coarsening method introduced by Miller and Dunson (2015) to reduce the sensitivity of x to 174 noise. If the dataset is large, the marginal likelihood value changes substantially for small variations of  $\boldsymbol{x}$ . A coarsened 175 marginal posterior probability distribution can remedy the problem, which is raising the marginal likelihood to the 176 power of  $1/\gamma$  with  $\gamma > 1$  (Eq. (3)).

#### thats tempering the likelihood similar to Parallel Tempering (Geyer 1992)

 $\sigma_{\gamma}(d|x) \coloneqq (\sigma(d|x))^{1/\gamma}$ here comes another concept in the synthetic test setup section that needs transfer to the method section. There the authors describe the sampling algorithm to be Metropolis-Gibbs. Parallel Tempering can be done with

179 180  $\sigma_{\gamma}(d|x)$  is more flat for larger  $\gamma$  and data cannot constrain source location, on the other hand, for small  $\gamma$ ,  $\sigma_{\gamma}(d|x)$ 181 causes the posterior on x to be limited to a few values. The former causes the marginal posterior distribution of x to 182 degenerate to the prior, and the latter situation is susceptible to noise (Gu et al., 2018). That is why the adjustment of 183  $\gamma$  is necessary, especially for obtaining optimum depth; the horizontal source coordinates are less sensitive to the 184 noise. For investigating source location variation, we plot the mean of MCMC trace versus  $\gamma$  (Fig. 3). The calculations





185 are performed for two cases, one in which starting point is near to the synthetic data source location (Fig.3, left panel) and the other with starting point in the place of the farthest location node of the  $20 \times 20 \times 20$  km grid (Fig. 3, right 186 panel). For both of the cases, SNR is 2. In the former situation the mean equals to the input location for all values of 187 188 low  $\gamma$ . Epecially for horizontal coordinates of the location; the means do not change for  $\gamma$  up to 300, while the depth 189 is more sensitive to the value. In the latter condition we have much longer burn-in period that are discarded before 190 plotting. The source location lastly reaches the correct input location, but there is a value of  $\gamma$  below which the range begin to shrink and the curves of source location range versus  $\gamma$  show trends. This value can be chosen as the optimum 191 192 values of  $\gamma$  shown by the black circles. For this case this optimum value is 50. The figures also illustrate the standard 193 deviations by gray shaded error bars (Campbell, 2009) which show the increment for larger  $\gamma$ s. That is due to more 194 flat  $\sigma_{y}(d|x)$  and failure of data to constrain x that is visible in the plot of vertical coordinate of the location, but 195 happens also for horizontal coordinates for higher ys than 300 (not shown in the figures). 196 Figure 4 shows the random walk in the focal angles' solution space utilizing all stations with no usage of noise (first

197 row in Table 1). The strikes, dips and rakes are calculated from the actual random walk in MT space. For simplicity 198 we only show the search in focal angles' space. The unvisited sites are shown by gray color and low and high 199 probability areas are depicted by a range of hot pallet colors from white to black. The start and end of the overall 200 search are illustrated by the green arrow and circle, respectively. The green lines show the path of all accepted focal It makes figure 4 just busy. 201 angles for all accepted locations. The total number of tested sites are  $10^5$ , however in the figure we only demonstrate the proposed and accepted sites which pass through the test of polarity and DC% in terms of the value of PPD. The 202 203 accepted focal angles and the relevant path for all accepted trial locations are shown by green circles and lines. There 204 are six accepted locations with the increasing likelihood out of 1000 tested locations.

205

#### 206 **4** Application

207 We present the results of applying the method on two small (M<sub>w</sub> 3.6 and 3.8) events with available independent focal 208 mechanism solutions. The first earthquake, which was also used in the synthetic tests above, is a Switzerland event 209 near Lichtenstein border. The second one is an Iranian event happened near the capital, Tehran, called Malard 210 earthquake.

211

#### 212 4.1 Sargans Earthquake

213 The first earthquake is an M<sub>W</sub> 3.6 earthquake at Sargans, Switzerland which happened on December 27, 2013 at 214 07:08:28 UTC. Figure 1 shows the reference DC solution retrieved by Bayesian ISOLA (Vackář et al., 2017) with the 215 mechanism, strike, dip and rake equal to 91/183, 78/79 and 169/12. We use 14 first-motion polarities to constrain the 216 solution resulted from broadband station inversion including the polarities of four other stations: GEA0, INS7, TMO20 217

and TMO22, not shown on Fig. 1 due to their larger epicentral distances comparing to the other illustrated stations. 218 Firstly, we test the method using all stations and all polarities. We filtered the records in frequency range 0.02 to 0.15

- Hz by Butterworth filter and inverted in the displacement domain. The results are presented in Fig. 5 to 10 and Table 219
- 220 4. Figure 5 shows the tested locations for 1000 iterations.

Sure the model parameters are dependent on Gamma, BUT valid samples from the PPD can only obtained when Gamma==1 !!! Also see Sequential Monte Carlo sampling methods! e.g.: DelMoral et al. 2006 Sequential Monte Carlo samplers

I suggest the authors use 2d marginals to show only the valid PPD with Gamma = 1 What is the path of exploration of the sampler adding in process understanding?

I miss synthetic waveform misfit plots of the input data and the obtained solution



221



222 solution beginning from  $\gamma = 1$ . Sargans earthquake does not show the shrinkage part even with the starting point in 223 the left top most corner of the location grid, that is, away from the optimum source location. Again the calculations is 224 performed after discarding the burn-in samples and the full source location range is the box with vertices  $\boldsymbol{x} = (-10, -10)$ 225 10, 1) and x = (10, 10, 21), with 9261 trial source positions. 226 Figure 7 is an illustration of the inner chain searching for optimum MT for any accepted solution. From among 10<sup>5</sup> 227 tested moment tensors for each given location only 638 sites go through the CPU intensive synthetic seismograms 228 calculations due to polarity and DC% test. For example, for the last and optimum source location, there are 149 MT 229 sites in this event. 230 As an example, all visited focal angles and accepted solutions with higher likelihood for LIENZ and SGT04 stations

The selected  $\gamma$  for Sargans earthquake is 35 (Fig. 6). Actually, there is a range of values that gives identical location

inversion are shown in Fig. 8. In two-station calculations, the location is fixed to the estimated value of all-stationresult, therefore Fig. 8 contains less visited sites.

233 The DC solution of Sargans earthquake is a strike-slip mechanism. It is obtained for full  $C_{\rm D}$ , that is, considering both 234 data and theoretical uncertainties and in the displacement domain (Fig. 9). For this event data uncertainty is dominant 235 over the Green function uncertainty. The waveform comparisons are illustrated for standardized data, that is, original 236 waveforms multiplied by Cholesky decomposition of the  $C_D$  (Fig. 10). Covariance matrix plays the role of automatic 237 frequency filter reducing the effect of noisy part of the spectrum, thus improving the result (Vackář, et al., 2017). 238 Variance reductions is 0.82 and strike, dip and rake are, 88/180, 80/80 and 170/10 with the magnitude of M<sub>w</sub> 3.6. That 239 is in comparison with inverting without covariance matrix or with the diagonal one whose elements are chosen to be 240 the mean squared value of the waveforms with calculated variance reduction of 0.57. The event is a shallow earthquake 241 with estimated 6 km hypocentral depth and horizontal shift of 0.5 and 1 km to the north and west of the epicenter.

242 Table 4 contains the result of the inversion for two- and one-station. Only two nearby stations are used and both of 243 the solutions with and without the constraints of polarity and DC% are determined. The first row of the table contains 244 the result of the inversion using all stations (red nodal lines) with the solution of Bayesian ISOLA also depicted in 245 green. Kagan angle in the first row is the comparison made with Bayesian ISOLA solution, but other angles are 246 determined in comparison with our own all-station solution. Although the two-station no-constraints DC solutions are 247 better in terms of Kagan angle but deviatoric solutions deteriorate. One-station results become worse both in regard 248 to Kagan angle and deviatoric part of the MT. Overall, as is the case with synthetic tests, polarity and DC% constraint can help to obtain better results when using lower number of stations. 249

250

#### 251 4.2 Malard earthquake

Here we apply the method on the second event happened around the town of Malard near Tehran, Iran, with M<sub>W</sub> 3.8,
on December 26, 2017 at 21:24:34 UTC (Fig 11). The reference solution of this event is our solution, that is, the

result of inversion by ISOLA (Zahradník, and Sokos, 2019) utilizing all shown stations that resulted in strike, dip and

255 rake equal to 24/118, 56/83 and -7/-145.

Figure 12 shows the plot of source location versus  $\gamma$  for Malard earthquake. The chosen  $\gamma$  is 10 and the source location

found is x = (-1, 4, 12), which is near to the location found by ISOLA using all stations, that is x = (-3, 3, 11.8). The





- north-east horizontal location have a small shrinkage part, while it does not exist for east-west location. The shrinking
  stage is longer for the vertical component of the location. The lack of shrinking stage for Sargans earthquake and it
  existence for Malard event could be due to higher level of noise in case of Malard event and the low number of station-
- components used for its calculation.
- 262 In order to apply the method on this earthquake, we utilize 21 first-motion polarities from broadband and short-period
- 263 records. The observed seismograms of HSB, VRN, JIR1, FIR and QSDN stations are filtered to frequency ranges
- 264 0.04-0.17, 0.04-0.08, 0.055-0.085, 0.055-0.085 and 0.055-0.08 to gain better waveform fit (Variance reduction =
- 265 0.79). The resulted strike, dip and rake are 26/119, 58/84 and -7/-148 (Fig 13).
- 266 We also determine two- and one-station solutions for this event. The results for this event show the advantage of the
- 267 constraints of polarity and DC% again. Of course, for all the cases, only one polarity is enough to constrain the solution
- to the optimum solution. That is except in the case of using the single station of VRN, in which more polarity
- constraints are needed for better compatibility with all station solution.
- 270

### 271 5 Conclusion

- We employed Bayesian framework using an MCMC algorithm to retrieve full moment tensor and source location of earthquakes by applying the constraints of polarity and DC%. The results show that the constraints can help to obtain better results in case of restricting the number of broadband stations to two or one. This is helpful, for example, when many short-period stations and therefore many polarities are available but the broadband network is sparse. The obtained results indicate that despite the low magnitude of the selected earthquakes, the employed approach could be reliable for retrieving location and moment tensors. The study added some methodical insights to the broad suite of similar methods including the two chain approach used comprising Metropolis-Gibbs Sampling algorithm and the
- coarsened likelihood for the parameter of source location.
- 280

#### 281 Data availability

- The data used in this study are freely available from Switzerland (Swiss Seismological Service (SED) at ETH Zürich
  1983), ZAMG (Vienna), International Institute of Earthquake Engineering and Seismology (IIEES) and Iranian
  Seismological Center (IRSC) of Institute of Geophysics, University of Tehran.
- 285

#### 286 Author contribution

287 M. Pakzad developed the code, did the synthetic tests and prepared the manuscript. M. Khalili performed the calculations of Sargans earthquake and helped making the figures. Sh. Vahidravesh performed the calculations of the
289 Malard earthquake and helped making the figures.

290

## 291 Acknowledgments

- 292 Data used in the study is from Switzerland (Swiss Seismological Service (SED) at ETH Zürich 1983), ZAMG
- 293 (Vienna), International Institute of Earthquake Engineering and Seismology (IIEES) and Iranian Seismological Center
- 294 (IRSC) of Institute of Geophysics, University of Tehran.





#### 295

#### 296 References

- Brillinger, D.R., Udias, A. and Bolt, B.A.: A probability model for regional focal mechanism solutions, B SEISMOL
  SOC AM. 70 (1): 149–170, 1980.
- 299 Fojtíková, L. and Zahradník, J.: A new strategy for weak events in sparse networks: The first-motion polarity
- solutions constrained by single-station waveform inversion, SEISMOL RES LETT: 85(6), 1265-1274.
   https://doi.org/10.1785/0220140072, 2014.
- **302** Fox, C. & Norton, R.A.: Fast sampling in a linear-Gaussian inverse problem, SIAM/ASA J. Uncertainty.
- 303 Quantification, 4, 1192–1218. <u>https://doi.org/10.1137/15M1029527</u>, 2015.
- Gu, C., Marzouk, Y.M. and Toksöz, M.N.: Waveform-based Bayesian full moment tensor inversion and uncertainty
   determination for the induced seismicity in an oil/gas field, GEOPHYS J INT, 212(3), 1963-1985.
   https://doi.org/10.1093/gji/ggx517, 2018.
- Hallo, M. and Gallovič, F.: Fast and cheap approximation of Green function uncertainty for waveform-based
  earthquake source inversions, GEOPHYS J INT, 207(2), 1012-1029. <u>https://doi-org.eres.qnl.qa/10.1093/gji/ggw320</u>,
  2016.
- Hallo, M., Asano, K. and Gallovič, F.: Bayesian inference and interpretation of centroid moment tensors of the 2016
  Kumamoto earthquake sequence, Kyushu, Japan, EARTH PLANETS SPACE, 69(1), p.134.
  https://doi.org/10.1186/s40623-017-0721-4, 2017.
- Haario, H., Saksman, E. & Tamminen, J.: An adaptive metropolis algorithm, BERNOULLI, 7, 223–242.
   <u>https://doi.org/10.2307/3318737</u>, 2001.
- 315 Ito, T., Gunawan, E., Kimata, F., Tabei, T., Meilano, I., Agustan, Ohta, Y., Ismail, N., Nurdin, I., Sugiyanto, D.: Co-
- seismic offsets due to two earthquakes (Mw 6.1) along the Sumatran fault system derived from GNSS measurements,
   EARTH PLANETS SPACE 68, 57, <u>http://dx.doi.org/10.1186/s40623-016-0427-z</u>, 2016.
- 318 Kagan, Y.Y.: 3-D rotation of double-couple earthquake sources, GEOPHYS J INT, 106(3), 709-716,
   319 <u>https://doi.org/10.1111/j.1365-246X.1991.tb06343.x</u>, 1991.
- Campbell, R., 2009: Shaded error bars (<u>https://github.com/raacampbell/shadedErrorBar</u>), MATLAB Central File
   Exchange, last access: 1 October 1, 2019.
- 322
- Li, J., Zhang, H., Sadi Kuleli, H. and Nafi Toksoz, M.: Focal mechanism determination using high-frequency
   waveform matching and its application to small magnitude induced earthquakes, GEOPHYS J INT, 184(3), 1261 1274, https://doi.org/10.1111/j.1365-246X.2010.04903.x, 2011.
- Lomax, A. and Curtis, A.: Fast, probabilistic earthquake location in 3D models using oct-tree importance sampling.
   In Geophys. Res. Abstr, Vol. 3, 955, 2001.
- 328 Lomax, A., Virieux, J., Volant, P. and Berge-Thierry, C.: Probabilistic earthquake location in 3D and layered models.
- In Advances in seismic event location, 101-134, SPRINGER, Dordrecht, <u>https://doi.org/10.1007/978-94-015-9536-</u>
   0 5, 2000.
- 331 Mao, W.J., Panza, G.F. & Suhadolc, P.: Linearized waveform inversion of local and near-regional events for source
- 332 mechanism and rupturing processes, GEOPHYS J INT, 116, 784–798, <u>https://doi.org/10.1111/j.1365-</u>
   333 246X.1994.tb03296.x, 1994.
- 334 Miller, J.W. and Dunson, D.B.: Robust Bayesian inference via coarsening, J AM STAT ASSOC, 1-13,
- 335 https://doi.org/10.1080/01621459.2018.1469995, 2018.





- 336 Mustać, M. & Tkalčić, H.: Point source moment tensor inversion through a Bayesian hierarchical model, GEOPHYS
   337 J INT, 204(1), 311–323, https://doi.org/10.1093/gji/ggv458, 2016.
- 338 Šílený, J., Panza, G.F. & Campus, P.: Waveform inversion for point source moment tensor retrieval with variable
- hypocentral depth and structural model, GEOPHYS J INT, 109, 259–274, <a href="https://doi.org/10.1111/j.1365-246X.1992.tb00097.x">https://doi.org/10.1111/j.1365-246X.1992.tb00097.x</a>, 1992.
- Stähler, S. and Sigloch, K.: Fully probabilistic seismic source inversion-Part 1: Efficient parameterization, SOLID
   EARTH, (2), 1055-1069, <u>https://doi.org/10.5194/se-5-1055-2014, 2014, 2014.</u>
- 343 Swiss Seismological Service (SED) at ETH Zürich: National Seismic Networks of Switzerland; ETH Zürich,
- 344 https://doi.org/10.12686/sed/networks/ch, 1983.
- Vackář, J., Gallovič, F., Burjánek, J., Zahradník, J., & Clinton, J.: Bayesian ISOLA: New tool for automated centroid
   moment tensor inversion, GEOPHYS J INT, 210(2), 693–705, <a href="https://doi.org/10.1093/gji/ggx158">https://doi.org/10.1093/gji/ggx158</a>, 2017.
- Walsh, D., Arnold, R. and Townend, J.: A Bayesian approach to determining and parametrizing earthquake focal
  mechanisms, GEOPHYS J INT, 176(1), 235-255, <u>https://doi.org/10.1111/j.1365-246X.2008.03979.x</u>, 2009.
- Wéber, Z.: Probabilistic local waveform inversion for moment tensor and hypocentral location, GEOPHYS J INT,
   165, 607–621, <u>https://doi.org/10.1111/j.1365-246X.2006.02934.x</u>, 2006.
- Wéber, Z.: Estimating source time function and moment tensor from moment tensor rate functions by constrained L1
   norm minimization, GEOPHYS J INT, 178, 889–900, <u>https://doi.org/10.1111/j.1365-246X.2009.04202.x</u>, 2009.
- Wéber, Z.: Probabilistic joint inversion of waveforms and polarity data for double-couple focal mechanisms of local
   earthquakes, GEOPHYS J INT, 213(3), 1586-1598, <u>https://doi.org/10.1093/gij/ggy096</u>, 2018.
- 355 Tarantola, A.: Inverse Problem Theory, ELS. ISBN 9780444599674, 1987.
- Zahradník, J., Sokos, E.N.: ISOLA code for multiple-point source modeling review, in: Moment Tensor Solutions A Useful Tool for Seismotectonics, edited by: D'Amico S., Springer, Berlin, Heidelberg, Germany, https://doi.org/10.1007/978-3-319-77359-9, 2019.
- 359
- 360
- 361
- 362
- 363
- 364
- 365
- 366
- 367
- 368
- 369
- 370
- 371
- 372
- 373





Table 1: Results of the synthetic tests with different SNR using all stations. The plots are equal-area Lambert-Schmidt

377 378 projections, lower hemisphere with compressional and dilatational polarities, in black and white respectively. The Compressional quadrants are shaded. The input focal mechanism nodal lines are in green and the solutions' nodal lines are

in red. VR stands for variance reduction

	Data	SNR	Strike (°)	Dip (°)	Rake (°)	DC%	CLVD%	VR	Kagan angle (°)	DC plot	Deviatoric plot
	0C%>70	No noise	87/179	75/83	173/15	76	19	0.99	4	× •	
	rities and <b>D</b>	1.0	84/177	67/82	171/23	85	10	0.60	8		
	s With pola	0.5	82/177	65/81	170/26	79	8	0.25	10		
	All station	0.1	332/112	60/37	-67/-124	87	12	0.005	63		





397

 

 Table 2: Mechanisms obtained using different SNR applied to the synthetic data. Only two stations (LIENZ and SGT04) are used in the inversion. The source is fixed to the one obtained in all stations computation.

 398

399

	Data	SNR	Strike (°)	Dip (°)	Rake (°)	DC%	CLVD%	VR	Kagan angle (°)	DC plot	Deviatoric plot
	nd DC% >	No noise	86/179	69/84	173/21	73	3	0.99	5	× • • •	
CT01 With a claimin	polarities a	1.0	81/176	71/77	167/19	86	14	0.58	8		
	GT04 With 7(	0.5	81/176	71/77	167/19	86	14	0.24	8		
	LIENZ + S	0.1	322/151	57/33	-95/-83	73	12	0.008	85		

400

401

402 Table 3: Results of the same inversion as in Table 2, but no polarity or DC% constraints are employed. The source is fixed 403 to the one obtained in all stations computation.

Data	SNR	Strike (°)	Dip (°)	Rake (°)	DC%	CLVD%	VR	Kagan angle (°)	DC plot	Deviatoric plot
rities and	No noise	87/180	72/80	170/18	89	4	0.99	2		
ithout pola DC%	1.0	79/180	63/70	157/29	72	6	0.59	17		
SGT04 W	0.5	66/175	55/65	149/40	49	9	0.27	30		
LIENZ +	0.1	39/152	56/59	142/40	48	1	0.03	48		





405Table 4: Moment tensor solutions using different data sets employing full CD for Sargans earthquake. The first row is the406reference solution of Fig. 9 resulting from the inversion of all stations with polarity and DC% constraints. Two-station407solutions are close to the reference one in terms of deviatoric mechanism. Two datasets of LIENZ + SGT04 and SGT04 +408PANIX show identical DC solution to the reference result. That is while the one-station dataset mechanisms are often badly

estim	estimated.											
	et e for all ti range r.	esse tables it ather than jus	would be ( ()	(°) Rake (°) tather necessa mum aposteri	DC% DC or valu	CL VD% osle wou (s)	M M an unce	N Prtainty	Kagan angle (°)	DC plot	Deviatoric plot	
	Reference (Fig. 9)	88/180	80/80	170/10	72	13	3.7	0.82	4			
DC% > 70	LIENZ + SGT04	80/178	63/74	163/28	71	25	3.6	0.80	19			
rities and I	SGT04 + PANIX	272/180	83/71	-161/-8	77	21	3.6	0.83	20			
With pola	LIENZ	272/181	87/80	-170/-3	78	-21	3.6	0.82	14			
	SGT04	80/178	63/74	163/28	71	25	3.6	0.79	19			
DC%	LIENZ + SGT04	92/184	81/77	166/10	21	39	3.8	0.83	6			
es and free	SGT04 + PANIX	84/182	70/70	159/21	41	26	3.7	0.86	15			
out polaritie	LIENZ	12/104	83/76	166/7	42	-16	3.7	0.84	75			
Withc	SGT04	82/184	66/64	151/27	28	39	3.6	0.84	23			





416Table 5: Moment tensor solutions using different datasets employing full CD for Malard earthquake. The first row is the417solution considered as reference (red) shown also in Fig. 13 resulted from the inversion of five stations with polarity and418DC% constraints. The green nodal lines in the rows other than the first row are the reference solution.

_	/		reen nouur r	mes in ent	roms other					01100 001	utioni	
		Data	Strike (°)	Dip (°)	Rake (°)	DC%	CLVD%	Mw	VR	Kagan angle (°)	DC plot	Deviatoric plot
		Reference Fig. 13	26/119	58/84	-7/-148	87	9	3.6	0.79	3		$\bigcirc$
	DC%>70	HSB + VRN	26/122	65/78	-14/-154	86	3	3.7	0.79	10		
	rities and I	VRN + JIR1	20/112	60/87	-3/-150	79	3	3.7	0.46	8		
	With pola	HSB	26/122	65/78	-14/-154	86	3	3.7	0.79	10		
		VRN	22/289	79/79	12/168	76	14	3.6	0.26	29		
	DC%	HSB + VRN	28/123	74/73	162/17	79	16	3.7	0.80	90		
	es and free	VRN + JIR1	25/293	46/88	-177/-44	70	11	3.7	0.47	91		
1 1	out polaritie	HSB	28/122	74/73	162/17	79	16	3.7	0.80	90		
	Withc	VRN	141/324	61/29	89/93	10	88	3.7	0.46	95		

419







 421
 9'00'E
 9'30'E
 10'00'E

 422
 Figure 1: Map related to Mw 3.6 Sargans, Switzerland earthquake, near Liechtenstein border, applied in the synthetic tests

422 Figure 1: Map related to Mw 3.6 Sargans, Switzerland earthquake, near Liechtenstein border, applied in the synthetic tests 423 and as the method application in the following sections. The independent beachball solution (retrieved using all stations by 424 Bayesian ISOLA (Vackář et al., 2017)) are inserted at the epicenter and the triangles indicate the station locations. Black 425 lines show countries' borders and lake shores.

426

427







429

Figure 2: 1000 random walks to reach the maximum a posteriori location in three and two dimensional views for the synthetic test using all stations without noise. After 38 iterations, the walker reaches the optimum point. Cubes and squares show proposed locations colored according to their iteration number and sized in keeping with the likelihood value, that is, largest cubes and squares indicate greater than 1% maximum a posteriori location, etc. The green cubes in 3D view show the accepted movements of the walker in location space with the increasing likelihoods, with their last optimum one encircled by green squares in 2D views. There are seven accepted solutions with the increasing likelihood that their paths are shown by green lines, reaching to the input location.

- 437
- 438 439
- 440
- -
- 441
- 442
- 443
- 444
- 445







446

Figure 3: Source location and shaded error bars versus  $\gamma$  for SNR equal to 2.0 in the frequency range 0.02 - 0.15 Hz for the synthetic test. The source locations' ranges are the mean of MCMC traces. The standard deviations are in gray, the red line illustrate the correct input location. The source location in the left panel (a to c) belong to the calculations with the starting point x = (0 km N, 0 km E, 10 km down) near to the input location in the synthetic test while the right panel (d to f) shows the source locations for the farthest starting point, that is x = (-10, -10, 1). The circle show the selected  $\gamma = 50$  for this test. What does the solid black line show?

What does the solid black line show? With gamma=50 the sampled ensemble is not a proper estimate

of the PPD, see my comment above. The ensemble at gamma = 1 is !







453

Figure 4: 2174 polarity and DC% tested focal angles out of 10<sup>5</sup> ones in the inner Markov chain shown by squares, colored according to the values of posterior probability density obtained applying all stations polarities and waveforms observed for the synthetic test with no noise. The steps belong to all six accepted source locations with the increasing likelihood in the outer location chain. The green lines demonstrate the accepted random walks. The green arrows represent the first point passing through the condition of larger likelihood; small green circles are subsequent points and finally the large green circles show the location of the optimum (maximum likelihood) focal mechanism (in total 87 sites for all accepted sources).







Figure 5: 409 accepted random walks to reach the maximum a posteriori location in three and two dimensional views for the inversion of Sargans earthquake data using all stations with full C<sub>D</sub> covariance matrix. 37 out of 2000 iterations are enough to find the maximum a posteriori location. The center of the Cartesian coordinate is 47.057°N and 9.486°E. There are five accepted solutions with the increasing likelihood that their paths are shown by green lines reaching to the optimum point. For the explanations of the symbols refer to Fig. 2.







478

479 Figure 6: Source location ranges and standard deviations versus  $\gamma$  for Sargans earthquake. The source locations are the 480 mean of MCMC trace shown in black line for each  $\gamma$  and the shaded error bars are in gray. The circle shows the selected 481  $\gamma$ .

482

483

484







Figure 7: 638 visited sites on 2D square lattices of focal angles' space, searching for maximum likelihood point of the PPD which obtained using all stations polarities and waveforms observed and for a Sargans earthquake. The steps belong to all five accepted source locations in the outer chain. The data and theoretical errors are used in the inversion in the form of full C<sub>D</sub>. There are in total 34 sites for all accepted sources. 1000 location sites and 10<sup>5</sup> MT values are polarity and DC% tested before CPU intensive synthetic seismograms calculations, leading to only 638 visited sites. For the symbols see Fig. 4.







493

Figure 8: Random walk of 122 trial steps on 2D square lattices of focal angles' space, searching for maximum likelihood point of the posterior probability density (PPD) which obtained using all stations polarities and waveforms observed at two stations, LIENZ and SGT04, for a Sargans earthquake. The location is fixed to the one estimated from all-station calculation. The data and theoretical errors are used through the inversion in the form of full C<sub>D</sub>. There are in total five sites with the increasing likelihood values. The location site is fixed to x = (0.5 km N, -1 km E, 6 km down) and 10<sup>5</sup> MT sites are polarity and DC% tested before CPU intensive synthetic seismograms calculations, leading to only 122 visited sites. For the symbols see Fig. 4.

501

502

- 503
- 504
- 505
- 506

- 508
- 509







- Figure 9: Moment tensor solution in case of using all stations and polarities in the inversion. The inversion is performed in
- 511 512 513 514 the displacement domain applying full covariance matrix of both data and Green functions uncertainties. The solution is shown by red nodal line; the errors are in black and the independent solution of Bayesian ISOLA (Vackář, et al., 2017) is
- in green. (a) DC focal mechanism solution. (b) Deviatoric part.
- 515
- 516











Figure 10: Comparison of standardized observed (blue) and synthetic (red) displacement seismograms using all stations 520 521 522 and polarities in the inversion involving data and Green functions uncertainties. The waveforms are normalized by means of the largest component of each station; that is, the largest component of each station is 1 and the numbers on the right are the maximum amplitudes in m. The station codes, epicentral distances and azimuths are shown on the left. The variance

- reduction using all seismograms is 0.82.







locations. Black lines show countries' borders and lake shores.

552

553

554

555 556

557



Figure 11: Map related to Mw 3.8 Malard earthquake near Tehran, Iran, used as the method application. The independent

beachball solution (retrieved using all stations by ISOLA) are inserted at the epicenter and the triangles show the station

Figure 12: Source location range for different γs calculated by Malard earthquake data. The filled circle illustrates the
 selected γ value for computing optimum source location.







561

Figure 13: Focal mechanism solution of Malard earthquake obtained by five-station broadband waveform inversion with
 full C<sub>b</sub>, constrained by first-motion polarities and DC% > 70 (red). The independent ISOLA solution obtained by 13 stations
 are illustrated in green and the errors are in black. a) DC part with polarities. b) Deviatoric part.

566



