On the Use of Data Noise as Site-Specific Weight Parameter in a Hierarchical Bayesian Moment Tensor Inversion: The Case Study of The Geysers and Long Valley Caldera Earthquakes

Mustać, Marija; Tkalčić, Hrvoje

Source / Izvornik: **Bulletin of the Seismological Society of America, 2017, 107, 1914 - 1922**

Journal article, Accepted version
Rad u časopisu, Završna verzija rukopisa prihvaćena za objavljivanje (postprint)

https://doi.org/10.1785/0120160379

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:487708

On the Use of Data Noise as Site-Specific Weight Parameter in a Hierarchical Bayesian Moment Tensor Inversion: The Case Study of The Geysers and Long Valley Caldera Earthquakes

Marija Mustać[1], Hrvoje Tkalčić[1]

[1] Research School of Earth Sciences,

The Australian National University,

Canberra ACT *2601*, Australia

This manuscript includes an electronic supplement containing details about the forward modeling and Bayesian information criterion. Additionally, it contains figures showing parameterization of the noise covariance matrix, algorithm flowchart, sampled centroid locations in non-linear inversions of The Geysers earthquake and posterior probability distributions of noise hyperparameters for the synthetic test and Long Valley caldera earthquake.

This is a post-print version accepted by the journal, but does not contain journal layout.

Mustać, Marija; Tkalčić, Hrvoje, On the use of data noise as site-specific weight parameter in a hierarchical Bayesian moment tensor inversion: The case study of The Geysers and Long Valley caldera earthquakes, Bulletin of the Seismological Society of America, 107(4), 1914–1922, 2017, DOI: 10.1785/0120160379,

Abstract

We expand a method for seismic moment tensor inversion using probabilistic Bayesian inference, which yields parameter uncertainties and includes a thorough treatment of noise in the data, to include additional noise parameters that weight the contributions of particular stations. In a synthetic test, we show that having individual noise parameters for each station gives an optimal fit to the data. The noise determines the level of data fit at each station, and in turn weights their contribution in the final solution. Apart from the noise level, an empirically determined data covariance matrix accounts for noise correlations present in waveform data. This improves the estimate of the centroid location and the non-double-couple components. We apply the method to two earthquakes, one from a volcanic (Long Valley caldera) and another from a geothermal (The Geysers) environment in California, which are likely to have non-double-couple components in the source mechanism. We confirm a significant isotropic component for the Long Valley caldera earthquake. Implementing a cosine data covariance matrix reduces the trade-off between the isotropic and compensated linear vector dipole components for The Geysers earthquake, and yields considerably higher non-double-couple components. This shows the importance of adequate noise treatment for earthquakes in complex tectonic environments.

Introduction

Earthquakes of small and moderate magnitudes can be effectively explained using the point source approximation (Aki and Richards, 2002). Even small earthquakes in environments such as volcanic or geothermal areas can have complex mechanisms including tensile faulting and fluid flow (Miller et al., 1998). Mechanisms other than simple shear dislocation are included in the second-rank seismic moment tensor (MT), usually decomposed into a double-couple (DC), compensated linear vector dipole (CLVD) and isotropic (ISO) components (e.g. Jost and Herrmann, 1989). In recent years, an increasing number of studies focused on non-double-couple components of the seismic MT (e.g. Miller et al., 1998; Panza and Saraó, 2000; Vavryčuk, 2004; Tkalčić et al., 2009).

Uncertainty estimates are important when examining earthquakes with large non-DC components, as the percentages of CLVD and ISO components are known to vary significantly with small perturbations of source position and time (Zahradník et al., 2008). Uncertainties have been estimated in linearized inversions (not including the source location) (e.g. Riedesel and Jordan, 1989; Vasco, 1990; Zahradník and Custodio, 2012), but

analytical expressions cannot be obtained when the location is also inverted for. Thus, they are being computed using probabilistic methods (e.g. Šílený, 1998; Ford et al., 2009; Fichtner and Tkalčić, 2010; Křížová et al., 2013), including the Bayesian inversion. In a Bayesian framework, the model parameters are treated as random variables and the solution is given as a posterior probability density of model parameters, taking into account prior independent knowledge of the model. The prior distribution is updated with a likelihood function that incorporates information from the data. Walsh et al. (2009) used generalized Matrix Fisher distributions to parameterize focal mechanism uncertainties using wave polarity data. Data functionals of seismic waveforms have been analyzed by Lee et al. (2011), while Duputel et al. (2012) and Stähler and Sigloch (2014) used waveform data.

Apart from uncertainty assessment, another advantage of the Bayesian inversion is its ability to account for noise in the data making it a free parameter in the inversion. Adequate evaluation of the amount and distribution of the noise enables an appropriate fit to the data, in contrast to ordinary least squares method that maximizes the fit to observations, discarding the noise. In turn, noise assessment regulates the model complexity. As pointed out by Scales and Snieder (1998), noise is everything the model cannot fit, i.e. it includes both measurement and theory errors. Duputel et al. (2012) used pre-event noise levels as noise variances for each trace, and emphasized the importance of non-diagonal elements of the covariance matrix to account for noise correlations, i.e. the interdependence of data errors. Including noise variances is similar to weighting station contributions based on the pre-event noise (e.g. Zahradník and Custodio, 2012) or epicentral distance (e.g. Kubo et al., 2002; Scognamiglio et al., 2009) in linearized inversions. Our assumptions on the noise define the level of data fit expected from the model and can be crucial in estimating correct parameter uncertainties.

We expand the method of Mustać and Tkalčić (2016), which uses a hierarchical Bayesian inversion to compute the seismic MT, its location and related uncertainties, by including additional hyperparameters to weight the contributions of noise in the data specific to each seismic station. Seismograms from different stations included in the inversion can have different levels of noise because of their epicentral distances, proximity to the sea, different geological structures beneath stations, human activities, as well as adequacy of the structure model for that particular path. Our algorithm determines the level of data fit for each station (or a group of stations), which influences the complexity of the solution. The optimal number of noise parameters is determined using the Bayesian Information Criterion (BIC) (Schwarz, 1978, more details are available in the electronic supplement to this article). Sampling of the parameter space is performed using two Markov chains:

one for the location and another for the MT parameters and the noise because the location parameters make the problem non-linear. We run a separate, inner Markov chain for every location in the outer one. This results in advantageous sampling the locations with highest likelihoods multiple times with different initial values of the MT parameters and the noise, similar to sampling with parallel chains. The inversion gives an ensemble of solutions that can be utilized to estimate model uncertainties.

The algorithm is designed to invert waveforms of moderate-size earthquakes and explosions at regional distances. We use long-period (20-50 s) data to reduce the effects of Earth structure and perform a thorough treatment of the noise, including its correlation. We test the importance of including multiple noise parameters in a synthetic experiment with real noise added to the waveforms. Subsequently, we apply the method to earthquakes from a volcanic (Long Valley caldera) and a geothermal (The Geysers) area in California. Before inverting seismograms of real events, we perform a linear inversion on a number of depths and search for a time shift between the data and synthetics based on 1D model Green's functions. These time shifts define the centroid origin time and accommodate for an imperfect structure model.

Method

Bayesian inversion

In a Bayesian approach, the inversion yields the posterior distribution of model parameters $p(\mathbf{m}|\mathbf{d})$, based on a prior distribution $p(\mathbf{m})$ and the likelihood function $p(\mathbf{d}|\mathbf{m})$, where **m** are the model parameters, **d** is the data and $x|y$ denotes the probability of x given a value of y. The posterior is defined by the Bayes' theorem (Bayes and Price, 1763)

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}, \tag{1}$$

where $p(\mathbf{d})$ is a constant, known as the Bayesian evidence. It normalizes the posterior distribution so its integral over the model space equals to unity.

The model consist of three parameters for the location (longitude, latitude and depth) and six parameters for the MT, based on the approach of Kikuchi and Kanamori (1991). Slight modifications were implemented in the frequency-wavenumber code *AXITRA* to create the Green's functions (Bouchon, 1981; Cotton and Coutant, 1997) and convolve them with six elementary tensors to yield six elementary seismograms $\mathbf{E}^n$. Thus, the forward modeling involves computing synthetic seismograms $u_i(t)$ as a linear combination

of the elementary seismograms

$$u_i(t) = \sum_{n=1}^{6} a_n E_i^n. \tag{2}$$

Coefficients a_n define the MT, and one of them isolates the isotropic component. More details can be found in the electronic supplement to this article and in Mustać and Tkalčić (2016). The noise is defined by additional hyperparameters determining the value of data errors. The number of noise hyperparameters is determined prior to the inversion; it can be between one and the number of stations (see the following section). The noise values could be approximated using the signal-to-noise ratio (SNR) (e.g. Duputel et al., 2012; Zahradník and Custodio, 2012). Still, the noise hyperparameters also include the theory error so we do not use the SNR to define the prior distribution of the noise. The algorithm can always converge to the SNR value, and we prefer this over restricting our prior. Furthermore, the choice of the prior distribution should not affect the inversion outcome (e.g. Sivia and Skilling, 2006).

Thus, a uniform prior distribution with a broad range of values is used for all model parameters. To speed up the inversion, particularly when analyzing multiple events in a certain area, we pre-compute the elementary seismograms for a discrete set of points around a reported location. All the locations have equal prior probability. For the MT parameters, we employ a uniform distribution between $-1.5 * M_0$ to $1.5 * M_0$, where M_0 is a previously determined value for the scalar moment.

Hierarchical aspect

The likelihood function $p(\mathbf{d}|\mathbf{m})$ quantifies how well does a given model $\mathbf{m}$ reproduce the observed data $\mathbf{d}$. We use a Gaussian distribution for the likelihood

$$p(\mathbf{d}|\mathbf{m}) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{C_D}|}} \exp\left[-\frac{1}{2}(\mathbf{G}(\mathbf{m}) - \mathbf{d})^T \, \mathbf{C_D}^{-1} (\mathbf{G}(\mathbf{m}) - \mathbf{d})\right], \tag{3}$$

where **G**(**m**) are the modeled seismic waveforms, N is the number of data points, $\mathbf{C_D}$ is the data covariance matrix and $|\mathbf{C_D}|$ its determinant. The matrix $\mathbf{C_D}$ quantifies the total data uncertainty and the covariability between data errors. Since we are inverting for the location parameters, the principal source of theory error is the lack of knowledge on the Earth structure (predominantly the velocity profile). We reduce this effect using long-period waveforms (20-50 s) that are less sensitive to small-scale heterogeneities. A part of the theory error might still be mapped into the free parameters of the covariance

matrix. The measurement error stems from ambient noise, instrument imperfections (e.g. timing errors, misalignment of components, changes in the instrument response) and data processing. Instrument-related errors should be detected and accounted for separately, but we can estimate the covariance matrix using pre-event noise traces, processed the same way as the data.

The cross-diagonal elements of the $\mathbf{C_D}$ matrix can be related to noise auto-correlations, which show a distinct character with prominent side lobes that depend on the frequency content of the noise. Also, there is a difference in auto-correlations of the horizontal and vertical components (Fig. S1, available in the electronic supplement to this article). We construct the $\mathbf{C_D}$ matrix as block diagonal, where each block $\mathbf{C_n}$ corresponds to one seismogram. To ensure the matrix is invertible, it is parameterized using two attenuated cosine functions

$$(\mathbf{C_n})_{ij} = \sigma_n^2 \left[b \cdot \exp\left(-\frac{|i-j|}{r_{e1}}\right) \cos\left(\frac{2\pi(i-j)}{L_1}\right) + (1-b) \cdot \exp\left(-\frac{|i-j|}{r_{e2}}\right) \cos\left(\frac{2\pi(i-j)}{L_2}\right) \right], \tag{4}$$

where $|i-j|$ is the time difference between samples *i* and *j*, *b* determines the amplitude of the cosine functions, L_1 and L_2 their periods, and r_{e1} and r_{e2} their exponential decay. Due to the similarities of the auto-correlations on different stations, it is sufficient to define the "shape" (the part within the square bracket) of only two $\mathbf{C_n}$ matrices (one for the horizontal and one for the vertical components). Including the coefficients b, r_{e1}, L_1, r_{e2} and L_1 as parameters in the inversion would require computing $|\mathbf{C_D}|$ and $\mathbf{C_D}^{-1}$ with every perturbation of these parameters, which would make the inversion computationally expensive. They are computed prior to the inversion as an average over horizontal and vertical noise auto-correlations on the stations from the examined region using the *Hyper-sweep* code (see the Data and Resources section) that performs a grid search in a multidimensional space. Each block $\mathbf{C_n}$ is multiplied by a variance σ_n^2, which quantifies the contribution of a particular seismogram in the inversion (the value of the cosine functions on the diagonal is one). The variances are defined as a percent of data root mean square (rms) on a particular station and can have values between zero and 500 per cent rms. Their value is a hyperparameter in the inversion. We compare the performance of a cosine covariance matrix with a diagonal one.

The algorithm we utilize here is summarized in Fig. S2, available in the electronic supplement to this article. Before performing the inversion, the shape of the data covariance matrix $\mathbf{C_D}$ and the elementary seismograms for all locations need to be computed.

We also determine the number of iterations beforehand, based on previous tests. For each location in the outer Markov chain, the elementary seismograms are read and an inner chain samples the MT parameters and the noise. Synthetic waveforms obtained by forward modeling using random perturbations of an initial model are compared to the waveform data, yielding the posterior probability distribution of model parameters $p(\mathbf{m}|\mathbf{d})$. This is performed for different numbers of noise parameters and the BIC used to determine their optimal number. The lowest BIC value indicates the optimal number of noise parameters, as explained in the electronic supplement to this article.

Synthetic tests with real noise

To show the performance of the algorithm in a realistic setting and explore the model selection using the BIC, we construct synthetic data using real station locations and add different amounts of noise on different stations. We take the location of an earthquake in Long Valley caldera, California (Dreger et al., 2000; Minson and Dreger, 2008) and five stations from the Berkeley Digital Seismic Network (BDSN) (shown in black in Fig. 1). The Green's functions are created using the SoCal model, an average structure model of southern California commonly used in MT inversions (Dreger and Helmberger, 1990), with 0.025° ($\sim$ 2.5 km) spacing for the epicentral coordinates and 1 km spacing in depth. The waveforms have 1 Hz sampling and 200 s duration. A complex mechanism (55% DC, 34% CLVD, and 11% ISO components) is used to create synthetic seismograms. Additionally, we add real noise from a quiet period (without earthquakes with magnitudes above 2.0), processed in the same way as the data (band-pass filtered between 20 and 50 s). We add noise from the north-south and vertical component seismograms to the horizontal and vertical components of the synthetic data, respectively. The amount of noise is defined as a per cent of the data rms; we add from 10 to 50%.

We perform inversions with both diagonal and cosine covariance matrices and one to five noise parameters, then use the BIC to determine the optimal number of noise parameters. When incorporating two, three and four parameters, we perform inversions with all possible station combinations (15, 25 and 10 combinations, respectively) and average the BIC value. The BIC decreases as the number of noise parameters increases, indicating that the number of noise parameters equal to the number of stations is optimal (Fig. 2 a). The maximum *a posteriori* probability (MAP) location varies in inversions with different number of free parameters when using a diagonal $\mathbf{C_D}$, but implementing a cosine $\mathbf{C_D}$ reduces the ambiguity and all inversions converge to the input location.

When we examine the MT solutions from inversions with one and five noise param-

eters, the DC parts are fairly similar (Fig. 2 b). To examine this further, we compare the variance reduction ($\mathrm{VR} = 1 - \frac{\int (\mathbf{d}-\mathbf{G}(\mathbf{m}))^2}{\int \mathbf{d^2}}$) of the MAP solutions with the same expression VR_0, computed using synthetic data without noise as $\mathbf{G}(\mathbf{m})$. We get a higher value ($\mathrm{VR} > \mathrm{VR}_0$) for most inversions, i.e. they are fitting the noise and explaining more features of the waveform than necessary. This overfitting is reduced when multiple noise parameters are used, and diminishes when a cosine $\mathbf{C_D}$ is implemented. The amount of input noise was determined taking into account its correlated nature so the noise values retrieved with the diagonal $\mathbf{C_D}$ are somewhat larger (Fig. S7, available in the electronic supplement to this article). Their relative values are similar, except that the noise on BKS is higher than the noise on ORV. The inversion with a cosine $\mathbf{C_D}$ retrieved the noise values reasonably well, except for stations CMB and ORV, where it was somewhat overestimated and underestimated, respectively (Fig. S8, available in the electronic supplement to this article). This might be due to extreme values of the input noise on these stations (10 and 50%). The inversion with a single noise parameter had the MAP value of 30%, an average of the input noises.

The main difference in the MT solutions is in the non-DC components of the moment tensor (visualized on the lune plots of Tape and Tape (2013) in Fig. 2 d). Both inversions with the cosine $\mathbf{C_D}$ retrieve the input non-DC components within uncertainty. Increasing the number of noise parameters reduced the discrepancy in non-DC components when the diagonal $\mathbf{C_D}$ was used.

Application to non-DC earthquakes

We apply the method to two earthquakes from complex tectonic environments. Initially, we extend the analysis by Mustać and Tkalčić (2016) of a 30 November 1997 M_W 4.9 earthquake in the Long Valley caldera (LVC) that occurred during an inflation episode (Dreger et al., 2000). Subsequently, we examine a 20 October 2006 M_W 4.6 earthquake in The Geysers geothermal field, where the exploitation for electrical power was accompanied by an increase in the number of earthquakes, many of which have substantial non-DC components (Johnson, 2014).

Long Valley caldera

Recent unrest in the LVC began in 1978 and included continuous seismic activity, uplift of the resurgent dome, emissions of CO_2 and other evidence of volcanic unrest (Foulger et al., 2004). Seismic activity intensified in 1997, when a number of earthquakes with

moment magnitudes above 4.5 occurred. Four of them had anomalous non-DC radiation, related to hydrothermal or magmatic processes (Dreger et al., 2000). In Mustać and Tkalčić (2016), we have examined one of them using a single noise parameter for all stations and confirmed a high ISO component. Now, we examine solutions obtained with multiple noise parameters.

The BIC shows that using an individual noise parameter for each station gives the optimal fit to the data (Fig. 3 a). The MAP location in the inversions with five noise parameters is at a depth of 7 km, 1 km deeper than with a single noise parameter, and closer to the Council of the National Seismic System (CNSS) hypocentral depth of 7.1 km. The epicenter in all cases is 3.5 km SE of the reported location. The relative noise values in the inversions with different assumptions about the noise are not the same. The inversion with a diagonal $\mathbf{C_D}$ yields the highest noise for KCC and PKD, while taking the noise correlations into account in the inversion with a cosine $\mathbf{C_D}$ results in the highest noise for ORV and BKS (Fig. S9, available in the electronic supplement to this article). The noise values from all inversions are significantly higher than values obtained by dividing the pre-event noise rms with data rms. The latter were 3, 2, 17, 4 and 16% for BKS, CMB, KCC, ORV and PKD, respectively. Such a difference most likely arises because of the theory error. The MAP moment tensors are quite similar, with the biggest difference occurring for the non-DC components (Fig. 3 b and d). However, they all confirm a high ISO component from the previous studies, probably related to the presence of high-pressure fluids from the magmatic body (Dreger et al., 2000).

The Geysers

Located in the Coastal Ranges of Northern California, The Geysers geothermal field lies in the area of regional shear stress field. However, most of the seismicity, dominated by microearthquakes, is related to steam production and water injection (Majer et al., 2007). A number of studies found earthquakes with large ISO components, both positive and negative, as well as large CLVD components (e.g. Ross et al., 1999; Johnson, 2014; Guilhem et al., 2014). For the first time, we apply a Bayesian method to one of the earthquakes in this region.

When performing inversions of The Geysers earthquake, elementary seismograms were computed using the GIL7 structure model (e.g. Pasyanos et al., 1996) because it yielded a better fit than the SoCal model in linearized inversions. We invert data from ten BDSN stations (white stations in Fig. 1). Due to the shallow depth of this event (3.46 km), the grid for elementary seismograms is refined to have 500 m spacing in depth

(0.025° spacing for the epicentral coordinates remains as before). We take seismograms of longer duration (220 s) because the epicentral distance to some stations has increased. Since we are using more stations for the inversion, the number of combinations of noise parameters on these ten stations (i.e. the number of ways to group the stations) drastically increases so we perform inversions using only one and ten noise parameters.

The algorithm was unsuccessful in retrieving a plausible centroid location for this earthquake (Figs. S3-S6, available in the electronic supplement to this article). Inversions with the cosine $\mathbf{C_D}$ converged at a boundary of pre-computed Green's functions, at a depth of 10 km. This most likely occurred because of the smaller size (the scalar moment is three orders of magnitude smaller than for the LVC earthquake) and shallow depth of this event, as determined by CNSS. A depth of 10 km is greater than depths of other earthquakes observed in The Geysers field. Furthermore, a cooling magma body is suggested to lie at 7-10 km depth (e.g. Truesdale et al., 1993) so it is less likely to expect earthquakes at those depths. Thus, we also perform the inversion with the location fixed at the CNSS hypocenter.

The inversions with different assumptions for the noise agree on high noise values on HAST, HATC and SUTB, and low noise on BKS and MHC, but the relative values are again quite different (Fig. 4 a). We get much larger differences between solutions obtained with a diagonal and a cosine $\mathbf{C_D}$ than for the LVC earthquake (Fig. 4 b). DC components of the MAP solutions are similar for inversions with one and ten noise parameters and a particular parameterization of $\mathbf{C_D}$, but inversions with a cosine $\mathbf{C_D}$ yield about 10° smaller dip angles and 10 and 17° difference in rake angles (for the N-S and E-W striking nodal planes, respectively) than inversions with a diagonal $\mathbf{C_D}$. Once more, the largest difference in the solutions is in the non-DC components. Accounting for noise correlations with the cosine $\mathbf{C_D}$ resulted in significantly higher CLVD component and slightly lower ISO component (Fig. 4 c). This led to poorer waveform fit (Fig. 4 a), but, as we showed in the synthetic experiment, inversions with a diagonal $\mathbf{C_D}$ are prone to overfitting the data. Furthermore, implementing a cosine $\mathbf{C_D}$ reduced the trade-off between the ISO and CLVD components that is observed in solutions with a diagonal $\mathbf{C_D}$.

The positive CLVD obtained in these inversions is consistent with crack openings. However, sources that involve only shear slip and an opening crack lie on the line connecting the DC and +Crack points on the lune plot (Tape and Tape, 2013). Thus, inversions with a cosine $\mathbf{C_D}$ suggest an additional mechanism that reduces the volumetric component, such as fluid extraction. Microearthquakes with such mechanism have previously been observed in The Geysers region (e.g. Ross et al., 1996, 1999) and the Hengill-Grensdalur volcanic complex in Iceland (e.g. Julian et al., 1997). Uncertainties for The

Geysers earthquake are larger than for the LVC earthquake, most likely due to the shallow depth (determined by CNSS) and considerably smaller magnitude of this event. This intensifies the need for a thorough noise treatment in the inversion.

Conclusions

We augmented a hierarchical Bayesian approach to the point source MT regional inversion. In addition to implementing an empirically estimated cosine data covariance matrix, whose non-diagonal elements account for interdependence of errors present in the waveform data, we include individual variances for each station as free parameters in the inversion. This allows the data themselves to modulate the contribution of each station. Implementing a cosine covariance matrix increased the ability to retrieve the correct centroid location and non-double-couple components in the synthetic experiment. This was not the case for The Geysers earthquake, where the location had to be fixed to the hypocenter. The Bayesian Information Criterion shows the advantage of using individual variances for each station, as opposed to fewer noise parameters.

The solutions for the LVC earthquake did not differ significantly when individual station variances were included, but there is a difference in the non-DC components. Furthermore, the BIC again favors inversions with individual noise parameter for each station. A smaller earthquake in The Geysers region required fixing the centroid location. A trade-off between the ISO and CLVD components can be seen when using a diagonal $\mathbf{C_D}$, but it is resolved when using the cosine one. Our preferred solution, obtained with a cosine $\mathbf{C_D}$ and individual noise parameters, has large ISO and CLVD components that indicate tensile failure due to fluid pressure, and suggest fluid extraction.

Data and Resources

Data for this study come from the Berkeley Digital Seismic Network (BDSN), doi:10.7932/BDSN, operated by the UC Berkeley Seismological Laboratory, which is archived at the Northern California Earthquake Data Center (NCEDC), doi: 10.7932/NCEDC (last accessed July 2015). Figures were made with the General Mapping Tools (www.soest.hawaii.edu/gmt [last accessed February 2014]; Wessel and Smith. (1995)). This study makes use of the computer package *Hyper-sweep* (http://www.iearth.org.au/codes/Hyper-sweep/) which was made available with support from the Inversion Laboratory (ilab). Ilab is a program for construction and distribution of data inference software in the geosciences supported by

AuScope Ltd, a non-profit organization for Earth Science infrastructure funded by the Australian Federal Government.

Acknowledgments

M. Mustać was supported by an Australian National University Research Scholarship and AE Ringwood Supplementary Scholarship. The research was also supported by grant no. FA9453-13-C-0268. We thank D. S. Dreger and A. Chiang from the University of California, Berkeley for discussions and assistance. We are grateful to O. Coutant for assistance with *AXITRA* code, B.L.N. Kennett for valuable discussions and the anonymous reviewers for comments that helped improve this manuscript.

References

Aki, K. and Richards, P. (2002). *Quantitative Seismology*, volume 2. University science books, Sausalito, California.

Bayes, M. and Price, M. (1763). An essay towards solving a problem in the doctrine of chances. *Philos. Trans. R Soc. London*, 53:370–418.

Bouchon, M. (1981). A simple method to calculate Green's functions for elastic layered media. *Bull. Seismol. Soc. Am.*, 71(4):959–971.

Cotton, F. and Coutant, O. (1997). Dynamic stress variations due to shear faults in a plane–layered medium. *Geophys. J. Int.*, 128(3):676–688.

Dreger, D. and Helmberger, D. (1990). Broadband modeling of local earthquakes. *Bull. Seismol. Soc. Am.*, 80(5):1162–1179.

Dreger, D., Tkalčić, H., and Johnston, M. (2000). Dilational processes accompanying earthquakes in the Long Valley caldera. *Science*, 288(5463):122–125.

Duputel, Z., Rivera, L., Fukahata, Y., and Kanamori, H. (2012). Uncertainty estimations for seismic source inversions. *Geophys. J. Int.*, 190(2):1243–1256.

Fichtner, A. and Tkalčić, H. (2010). Insights into the kinematics of a volcanic caldera drop: Probabilistic finite–source inversion of the 1996 Bárdarbunga, Iceland, earthquake. *Earth Planet. Sci. Lett.*, 297(3):607–615.

Ford, S., Dreger, D., and Walter, W. (2009). Identifying isotropic events using a regional moment tensor inversion. *J. Geophys. Res.*, 114:B01306.

Foulger, G. R., Julian, B. R., Hill, D. P., Pitt, A. M., Malin, P. E., and Shalev, E. (2004). Non–double–couple microearthquakes at Long Valley caldera, California, provide evidence for hydraulic fracturing. *J. Volcanol. Geotherm. Res.*, 132:45–71.

Guilhem, A., Hutchings, L., Dreger, D. S., and Johnson, L. R. (2014). Moment tensor inversions of M$\sim$3 earthquakes in the Geysers geothermal fields, California. *J. Geophys. Res. Solid Earth*, 119(3):2121–2137.

Johnson, L. (2014). Source mechanisms of induced earthquakes at The Geysers geothermal reservoir. *Pure Appl. Geophys.*, 171(8):1641–1668.

Jost, M. L. and Herrmann, R. B. (1989). A student's guide to and review of moment tensors. *Seismol. Res. Lett.*, 60(2):37–57.

Julian, B. R., Miller, A. D., and Foulger, G. R. (1997). Non–double–couple earthquake mechanisms at the Hengill–Grensdalur volcanic complex, southwest Iceland. *Geophys. Res. Lett.*, 24(7):743–746.

Kikuchi, M. and Kanamori, H. (1991). Inversion of complex body waves — III. *Bull. Seismol. Soc. Am.*, 81(6):2335–2350.

Křížová, D., Zahradník, J., and Kiratzi, A. (2013). Resolvability of isotropic component in regional seismic moment tensor inversion. *Bull. Seismol. Soc. Am.*, 103(4):2460–2473.

Kubo, A., Fukuyama, E., Kawai, H., and Nonomura, K. (2002). NIED seismic moment tensor catalogue for regional earthquakes around Japan: quality test and application. *Tectonophysics*, 356(1):23–48.

Lee, E.-J., Chen, P., Jordan, T., and Wang, L. (2011). Rapid full–wave centroid moment tensor (CMT) inversion in a three–dimensional earth structure model for earthquakes in Southern California. *Geophys. J. Int.*, 186(1):311–330.

Majer, E. L., Baria, R., Stark, M., Oates, S., Bommer, J., Smith, B., and Asanuma, H. (2007). Induced seismicity associated with enhanced geothermal systems. *Geothermics*, 36(3):185–222.

Miller, A. D., Julian, B. R., and Foulger, G. R. (1998). Non–double–couple earthquake 2. Observations. *Rev. Geophys.*, 36(4):551–568.

Minson, S. and Dreger, D. (2008). Stable inversion for complete moment tensors. *Geophys. J. Int.*, 174(2):585–592.

Mustać, M. and Tkalčić, H. (2016). Point source moment tensor inversion through a Bayesian hierarchical model. *Geophys. J. Int.*, 204(1):311–323.

Panza, G. F. and Saraó, 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.*, 143:2265–2272.

Pasyanos, M., Dreger, D. S., and Romanowicz, B. (1996). Toward real–time estimation of regional moment tensors. *Bull. Seismol. Soc. Am.*, 88(5):1255–1269.

Riedesel, M. and Jordan, T. (1989). Display and assessment of seismic moment tensors. *Bull. Seismol. Soc. Am.*, 79(1):85–100.

Ross, A., Foulger, G. R., and Julian, B. R. (1996). Non–double–couple earthquake mechanisms at the Geysers geothermal area, California. *Geophys. Res. Lett.*, 23(8):877–880.

Ross, A., Foulger, G. R., and Julian, B. R. (1999). Source processes of industrially–induced earthquakes at The Geysers geothermal area, California. *Geophysics*, 64(6):1877–1889.

Scales, J. and Snieder, R. (1998). What is noise? *Geophysics*, 63(4):1122–1124.

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.*, 6(2):461–464.

Scognamiglio, L., Tinti, E., and Michelini, A. (2009). Real–time determination of seismic moment tensor for the Italian region. *Bull. Seismol. Soc. Am.*, 99(4):2223–2242.

Šílený, J. (1998). Earthquake source parameters and their confidence regions by a genetic algorithm with a 'memory'. *Geophys. J. Int.*, 134(1):228–242.

Sivia, D. and Skilling, J. (2006). *Data analysis: A Bayesian tutorial*. Oxford University Press, New York.

Stähler, S. C. and Sigloch, K. (2014). Fully probabilistic seismic source inversion–Part 1: Efficient parameterisation. *Solid Earth*, 5(2):1055–1069.

Tape, W. and Tape, C. (2013). The classical model for moment tensors. *Geophys. J. Int.*, 195:1701–1720.

Tkalčić, H., Dreger, D., Foulger, G., and Julian, B. (2009). The puzzle of the 1996 Bárdarbunga, Iceland, earthquake: No volumetric component in the source mechanism. *Bull. Seismol. Soc. Am.*, 99(5):3077–3085.

Truesdale, A. H., Walters, M., Kennedy, M., and Lippmann, M. (1993). An integrated model for the origin of The Geysers geothermal field. *Geotherm. Resour. Counc. Trans.*, 17:273–280.

Vasco, D. W. (1990). Moment–tensor invariants: Searching for non–double–couple earthquakes. *Bull. Seismol. Soc. Am.*, 80(2):354–371.

Vavryčuk, V. (2004). Inversion for anisotropy from non–double–couple components of moment tensors. *J. Geophys. Res. Solid Earth*, 109(B07306).

Walsh, D., Arnold, R., and Townend, J. (2009). A Bayesian approach to determining and parametrizing earthquake focal mechanisms. *Geophys. J. Int.*, 176(1):235–255.

Wessel, P. and Smith., W. (1995). New version of the generic mapping tools released. *Eos Trans. Am. Geophys. Union*, 76:329.

Zahradník, J. and Custodio, S. (2012). Moment tensor resolvability: Application to southwest Iberia. *Bull. Seismol. Soc. Am.*, 102(3):1235–1254.

Zahradník, J., Sokos, E., Tselentis, G.-A., and Martakis, N. (2008). Non–double–couple mechanism of moderate earthquakes near Zakynthos, Greece, April 2006; explanation in terms of complexity. *Geophys. Prospect.*, 56(3):341–356.

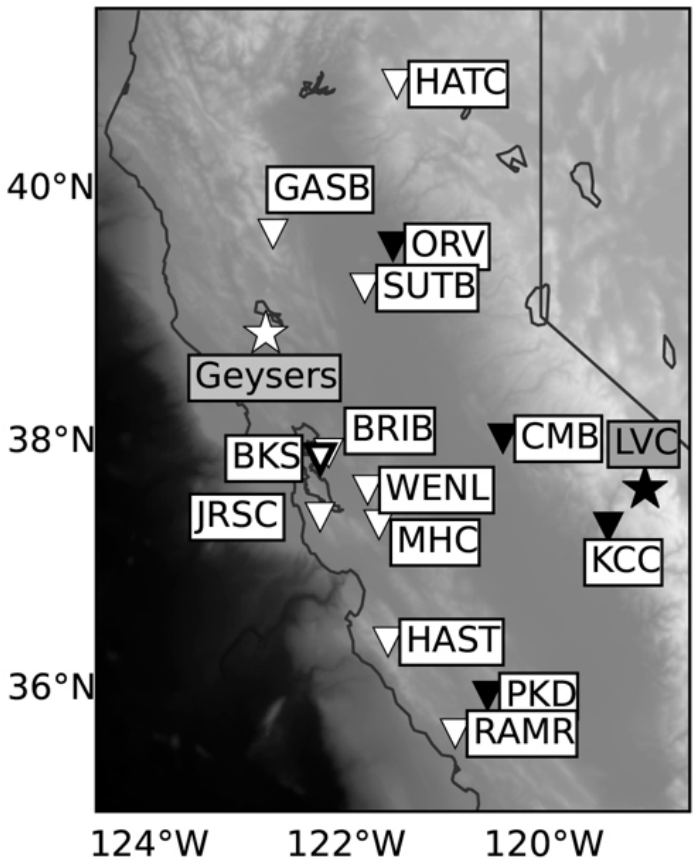

Figure 1: Map of the studied region with the LVC (black star) and The Geysers earthquake locations (white star), as well as stations used in the inversions of the LVC earthquake (black inverted triangles) and The Geysers earthquake (white inverted triangles). Station BKS was used in the analysis of both earthquakes.

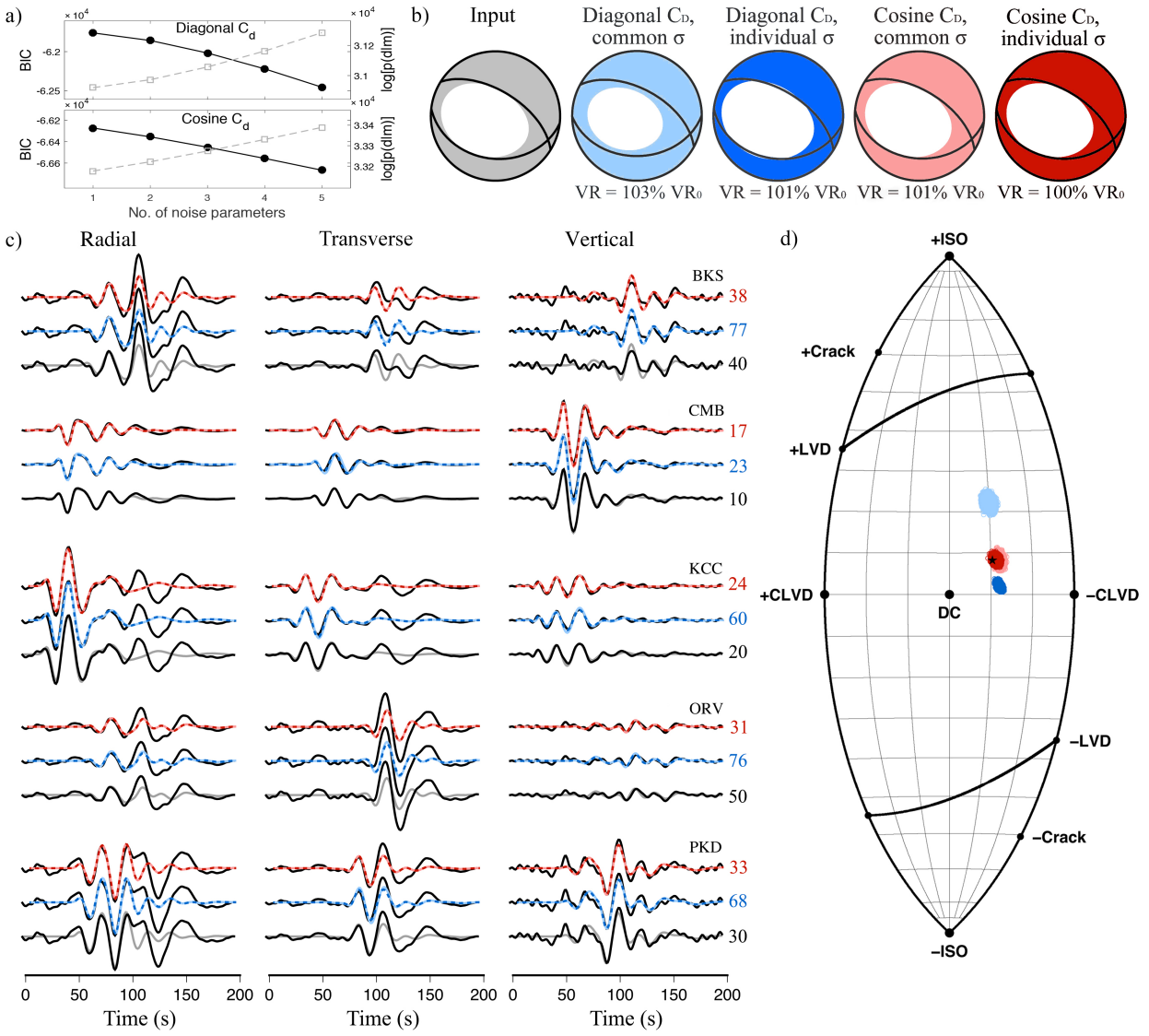

Figure 2: Comparison of solutions with different assumptions for $\mathbf{C_D}$. (a) (Gray) The logarithmic likelihood and (black) the BIC values for inversions with (top) a diagonal, and (bottom) an attenuated cosine covariance matrix for different numbers of noise hyperparameters in the inversion. BIC and likelihood values for two, three, and four noise parameters are averages from all inversions. (b) Input mechanism and MAP solutions from inversions with a diagonal and cosine $\mathbf{C_D}$ and one (common σ) and five (individual σ for each station) noise parameters. The numbers below beachballs show that most inversions have higher variance reduction ($\mathrm{VR} = 1 - \frac{\int (\mathbf{d}-\mathbf{G}(\mathbf{m}))^2}{\int \mathbf{d^2}}$) than VR_0, computed between the data with and without noise. (c) (Gray) data without noise, (black) data with noise and synthetic seismograms from MAP solutions, colored as in (b). Seismograms from inversions with multiple noise parameters are plotted with a dashed line. Numbers on the right hand side of the seismograms show the input level of noise (black) and the MAP values obtained in inversions with a diagonal (blue) and a cosine (red) $\mathbf{C_D}$ matrix for each station. (d) Lune source-type diagram showing the input mechanism (black star) and ensembles of solutions from the same four inversions, colored as in (b).

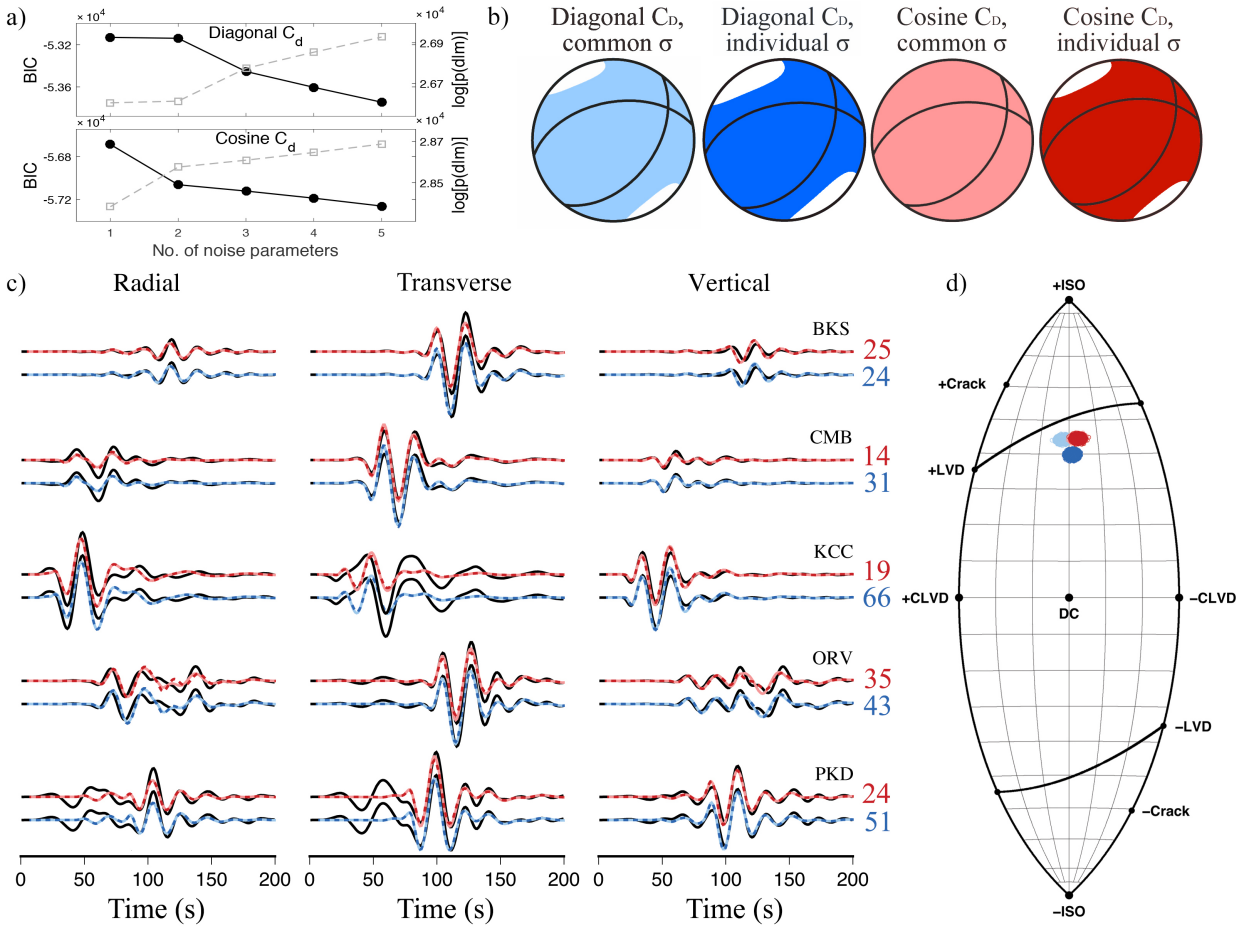

Figure 3: (a) (Gray) The logarithmic likelihood and (black) the BIC, (b) MAP solutions, (c) observed seismograms (black) and synthetic seismograms with added MAP noise values (blue for a diagonal $\mathbf{C_D}$ matrix and red for a cosine $\mathbf{C_D}$ matrix), and (d) Lune source-type diagram for the LVC earthquake ensemble solutions. The color scheme for the ensemble solutions is given in (b).

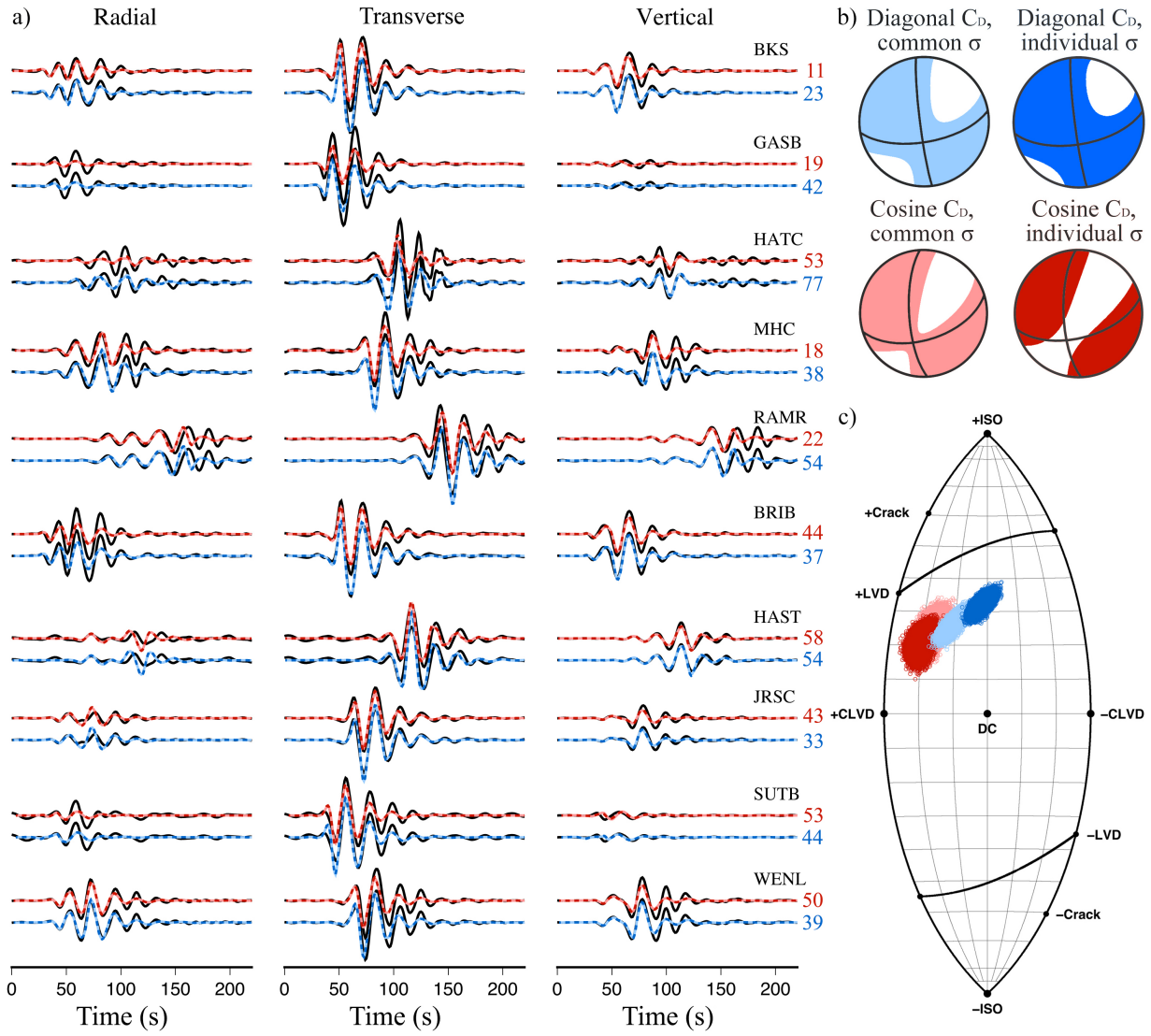

Figure 4: (a) Observed seismograms (black) and synthetic seismograms with added MAP noise values (blue for a diagonal $\mathbf{C_D}$ matrix and red for a cosine $\mathbf{C_D}$ matrix), (b) MAP solutions, and (c) Lune source-type diagram for The Geysers earthquake ensemble solutions. The color scheme for the ensemble solutions is given in (b).