AZIMUTHAL ELASTIC INVERSION FOR FRACTURE CHARACTERIZATION

By

Salam AlNabulsi, Ehsan Kamalinejad, Justin Meskas, Jing (Catherine) Wang, Ke Yin,
Mentor: Jon Downton

IMA Preprint Series #2399

(June 2012)

INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS
UNIVERSITY OF MINNESOTA
400 Lind Hall
207 Church Street S.E.
Minneapolis, Minnesota 55455-0436
Phone: 612-624-6066 Fax: 612-626-7370
URL: http://www.ima.umn.edu

Azimuthal Elastic Inversion for Fracture Characterization*

Salam AlNabulsi†, Ehsan Kamalinejad‡, Justin Meskas§,
Jing (Catherine) Wang¶, Ke Yin‖, Mentor: Jon Downton**

Abstract

In this paper, we investigated azimuthal elastic inversion problem using Monte Carlo type methods: Simulated Annealing (SA) and Markov Chain Monte Carlo (MCMC). It is demonstrated in numerical simulation that these two approaches are capable of resolving 5 and even 6 parameters in the forward model, which is considered complicated for its non-linearity ([2]).

1 Introduction

Azimuthal amplitude versus angle (AVAZ) has proved to be an important tool in characterizing fracture distributions and direction for hydrocarbon reservoirs (e.g. Al-Marzoug et al., 2004 [1]; Gray and Todorovic-Marinic, 2004 [6]). Rüger (2002 [9]) shows that the amplitude R versus azimuth ϕ for narrow angles of incidence θ for an isotropic half-space over an HTI anisotropic half-space is

$$R(\theta,\phi) = A + (B_{iso} + B_{ani}\sin^2(\phi - \phi_{iso}))\sin^2\theta \tag{1}$$

where A is P-wave impedance reflectivity, B_{iso} the isotropic gradient, B_{ani} the anisotropic gradient and $\phi_{sym} = \phi_{iso} + 90°$ is the symmetry axis of the HTI anisotropic media. The anisotropic gradient is related to the crack density. Note that equation (1) is nonlinear, thus complicating the inference of the desired parameters. Some linearization techniques have been proposed by Downton (2001 [3]) and Jenner (2002 [7]), and the uncertainty estimation has been studied by Downton (2006 [2]). In this paper, we apply Monte Carlo type of methods like Simulated Annealing (SA) and Markov Chain Monte Carlo (MCMC) to invert the nonlinear forward model (1) with 4 parameters. A statistical analysis of the marginal distributions and covariance of pairs of parameters is performed to show the resolvability

*This work is done in IMA Math Modeling in industry Workshop in University of Calgary, Canada, June 17-27, 2012

†University of Calgary, Email: ssalnabu@ucalgary.ca

‡University of Toronto, Email: ehsan.kamalinejad@gmail.com

§Simon Fraser University, Email: jmeskas@sfu.ca

¶University of Alberta, Email: jw14@ualberta.ca

‖Georgia Institute of Technology, Email: kyin@math.gatech.edu

**Hampson−Russell Software & Services, A CGGVeritas Company, Email: Jon.Downton@cggveritas.com

of parameters. It is also demonstrated with simulated data that our approaches can be applied to more complete models (see [4])

$$R(\theta,\phi) = A_0 + B_0 \sin^2\theta + C_0 \sin^2\theta \tan\theta + B_2 \cos(2(\phi - \phi_{iso})) \sin^2\theta + C_2 \cos(2(\phi - \phi_{iso})) \sin^2\theta \tan^2\theta + C_4 \cos(4(\phi - \phi_{iso})) \sin^2\theta \tan^2\theta \quad (2)$$

where $(A_0, B_0, C_0, B_2, C_2, C_4) = f(R_p, R_s, \delta\Delta_N, \delta\Delta_V, \delta\Delta_H)$, and f is a linear transformation.

2 Metropolis Simulated Annealing

Simulated Annealing (See [5]) is required when a global extremum of an objective function is sought after. It is an optimization technique to avoid local extrema and focus on global extrema. We use Metropolis Simulated Annealing to find the global minimum of error function $|d - G(m)|^2$. The Metropolis algorithm is as follows:

- Pick an initial approximated model m_0
- Recursively perturb m_j, where j = 0,1,2...N, with a random normal perturbation to get m_{j+1}
- Compare the L^2 norms of the error between the recorded seismic data and the data produced by the current and perturbed model.
- Calculate $E_j = |d - G(m_j)|^2$ and $E_{j+1} = |d - G(m_{j+1})|^2$
- If $E_{j+1} < E_j$, accept the new model.
- If $E_{j+1} > E_j$, only accept the new model with an acceptance ratio of $P = exp((E_j - E_{j+1})/T)$, where T is the temperature.
- Perturb m_{j+1} and continue the recursion process until either a tolerance is obtained or enough iterations have be achieved

This method leaves ambiguity for the size of the perturbation and for what temperature function to use. Both of these must be tweaked by the user to each specific problem. The Fast Metropolis algorithm is a altered version which uses Cauchy distributions to perturb the model, and this is implemented in Section 2.2. The user can choose the temperature profile, and Figure 1 shows some common choices for the temperature function. The higher the temperature the more likely a model m_{j+1} is accepted even if the error is larger. This creates volatility in the updating process which allows the model to escape from the local extremum and perhaps move towards the global extremum. In Section 2.1 the $\frac{1}{x}$ temperature function is used, while in Section 2.2 the $\frac{1}{log(x)}$ temperature function is chosen. The implications of these choices are discussed further in the respective sections. One thing to note is that as the number of iterations of Metropolis Simulated Annealing goes up, with the choice of the $\frac{1}{log(x)}$ temperature function, the probability that the global extrema is reached approached 1.

Figure 2 shows a fictitious 2D objective function. It has a local minimum and a global minimum. The lowest point at the global minimum is sought after. To give an example of the implimentation of Metropolis Simulated Annealing we could pick a point for m_j, say the top left cross marked in the plots of Figure 2. If the perturbation m_{j+1} is in the direction of either minima the model will automatically update to this new point. If the point that is randomly selected is up the slope

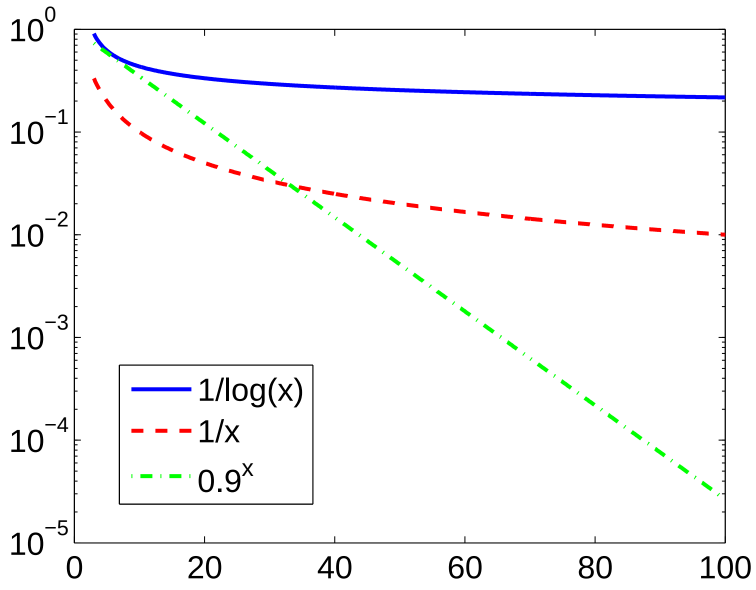

Figure 1: Three common temperature functions used in Metropolis Simulated Annealing

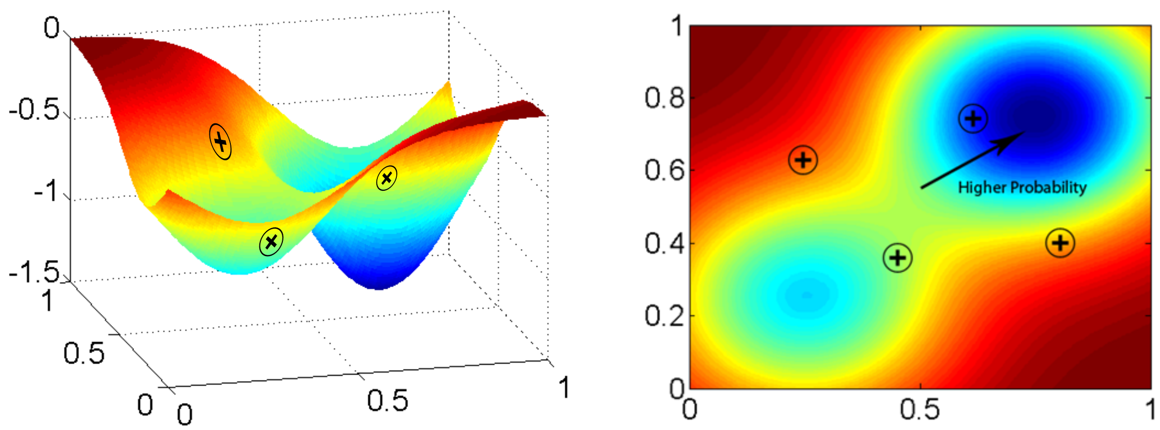

Figure 2: Surface and contour plots of a 2D objective function

in the red region, the model will update with an acceptance ratio that is large for large temperature. A good temperature function must be chosen to allow for the perturbations to take the model out of a local minimum and have a chance to go into the global minimum.

2.1 Four Parameter Study

As discussed previously in introduction, the geophysical problem gets reduced down to solving the inverse problem by solving for the model $m_j = [A, B_{iso}, B_{ani}, \phi_{iso}]$ from the seismic data, d. This is obtained by implementing Metropolis Simulating Annealing from Section 2. The temperature function chosen for this problem is $\frac{1}{x}$. This was optimal for this problem because the cooling rate is fast and a global minimum is obtained quickly and accurately. Luckily this problem does not appear to have any significant local minima. On the other hand, the problem has an unfortunate side effect of having two global minima. Therefore the simulations would converge on two different equally valid models. Both models are accepted because of a geophysical reason which has the orientation of the fracture, ϕ_{iso}, off by $90°$.

Each point in the scatter plots of Figure 3 depicts the point that had the lowest error during a simulation. This is repeated many times to give distributions of the parameter values for the model. Figure 3 shows the scatter plots of every combination of two parameters for both solutions. We can use this to show the correlation between parameters which can give insight into the geophysical problem. Another statistical technique is to look at the histograms of each parameter as seen in Figure 4. We can see from the figure that the parameters B_{iso} and B_{ani} have a normal distribution which validates the reliability of the model. The A and ϕ_{iso} parameters show a skewed distribution which shows unreliability of the model. This skewedness could of been caused by the sample size being too small among other possibly reasons.

The seismic data, d, must be created artificially. This is done by fixing the four parameters and creating the seismic data from Equation (1). Some noise is added to this data that accounts for the noise that is accompanied by real data acquisition. These four fixed parameters double as an answer to compare the model we obtain by Metropolis Simulated Annealing. These four parameters are set to be: $A = 0.1$, $B_{iso} = -0.1$, $B_{ani} = -0.2$, and $\phi_{iso} = 25°$. As seen by the axes in Figures 3 and 4, we have parameters that are centred on slightly different values than the true answers. These deviations are caused by the noise added to the system. Solution 2 has ϕ_{iso} with a $90°$ rotation which causes B_{ani} to be negative of the true value.

2.2 Six Parameter Study

The implementation of simulated annealing for the four parameter model can be generalized to any number of parameters. This particular section discusses a six parameter model. The six parameters we search for in the inverse problem are $m_j = [R_p, R_s, \Delta\delta_N, \Delta\delta_V, \Delta\delta_H, \phi_{iso}]$. Looking at equation (2) we must find the parameters $A_0, B_0, C_0, B_2, C_2, C_4$. In the following study we only look at the geophysical parameters for m_j and a direct calculation can convert one set of parameters to the other.

In Figure 5 we can see how these parameters relate to one another. Similarly to the four parameter study we are interested in the correlation among parameters. One

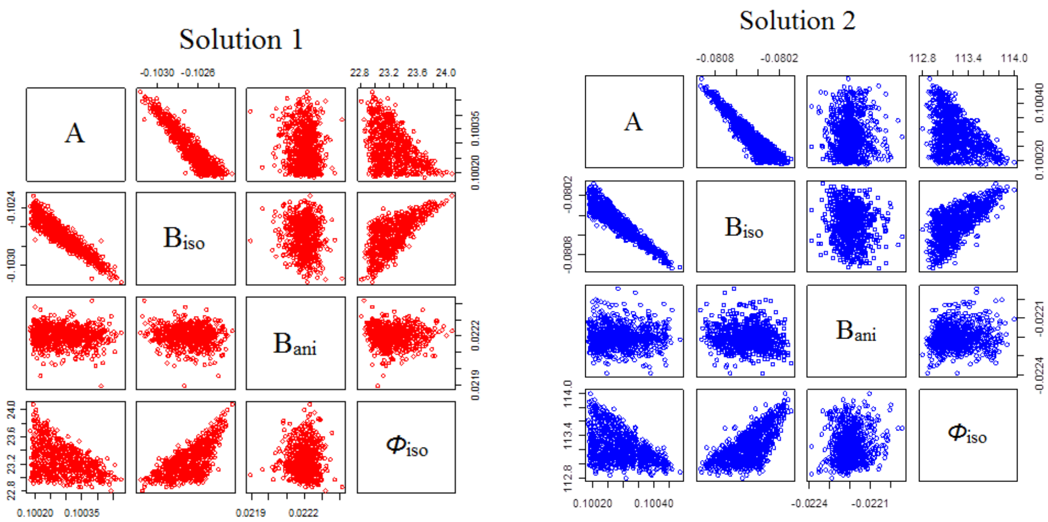

Figure 3: Scatter plots of the two solutions for the 4 parameter problem. Each particular scatter plot involves the parameters that are in the same column and row. For example the bottom left is the A vs. ϕ_{iso} scatter plot.

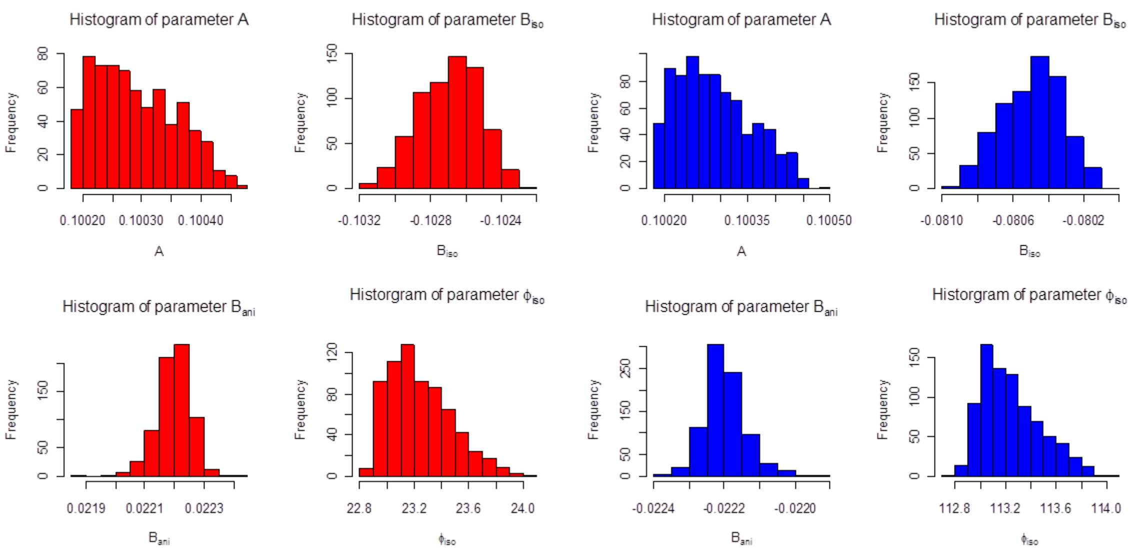

Figure 4: Histograms of each parameter for both solutions.

negative about this problem is the variability of the R_s, R_p, $\Delta\delta_N$,$\Delta\delta_V$ and $\Delta\delta_H$. The simulation which produced this variability had noise added to the simulated data which was a factor of $\frac{1}{8}$ the magnitude of the data. This shows that either the Metropolis Algorithm has complications that went unaware to the authors, or that the problem is ill-posed when it gets extended to 6 parameters. Since seismic data that is collected in practice could have noise up to $\frac{1}{2}$ the magnitude of the seismic data, a model that can resolve this is wanted. The current geophysical readings using seismic may be under resolved and more data may need to be collected to resolve additional parameters. When the noise is too large, the local and global minimum can become indistinguishable. For the simulation shown in Figures 5 and 6 are only half of the models discovered because the local minimum was captured as many times as the global minimum. This could be a large problem when trying to resolve the parameters.

During this study, the Metropolis algorithm was upgraded to contain two improvements. The first modification, from an idea from [10], changes the normal distribution of the perturbation of the model to a Cauchy distribution. The second modification was to use the temperature function of $\frac{1}{log(x)}$, which is a common temperature function choice. The Cauchy distribution has a significant increase in size of its tails compared to the normal distribution which allows for large movements in the parameter space. This gives a higher probability to escape the local minimum and end up in the global minimum. The choice of the temperature function being $\frac{1}{log(x)}$ allows for more volatility with the model perturbations since the temperature takes a long time to decrease. The combination of these two allows the model perturbations to move freely in and out of both the local and global minimum. During one simulation a global minimum can be captured many times. The down side to these modifications is once it is in the global minimum it does not stay there for long and potentially does not get a good approximation to the exact global minimum. One option is to take the best model after using Metropolis algorithm and to quickly converge it to the global minimum with something like conjugate gradient, this was not done. This may give a better approximation to the $\Delta\delta$ parameters as discussed above.

3 Markov Chain Monte Carlo

3.1 Description of the algorithm and its application

Given the forward model

$$\mathbf{d} = \mathbf{G}(\mathbf{m}) + \xi$$

where

$$\mathbf{m} = (m_1, \ldots, m_n)$$

are the set of unknown parameters that we want to estimate form the seismic data $\mathbf{d}$, and ξ is the noise.

In the previous section we try to find parameter $\mathbf{m}$ that minimize the misfit

$$e(\mathbf{m}) = |\mathbf{d} - \mathbf{G}(\mathbf{m})|^2 \tag{3}$$

using Simulated Annealing (SA). In existence of noise, $\mathbf{d}$ is a random variable, and from statistical point of view, we may consider parameters as random variables as

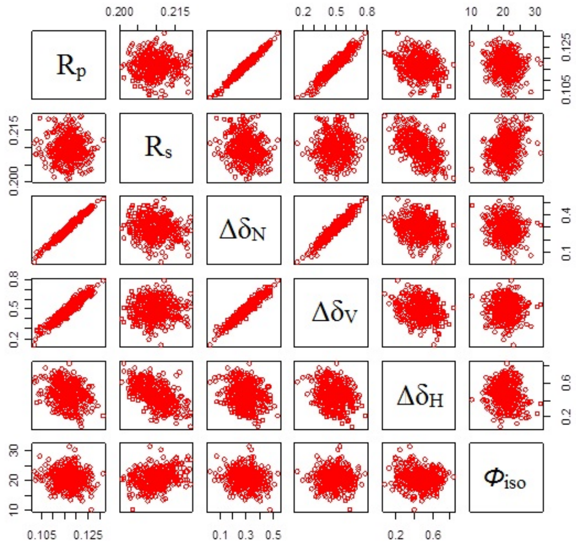

Figure 5: Scatter plots of 6 parameter study. Each particular scatter plot invloves the parameters that are in the same column and row. For example the bottom left is the R_p vs. ϕ_{iso} scatter plot.

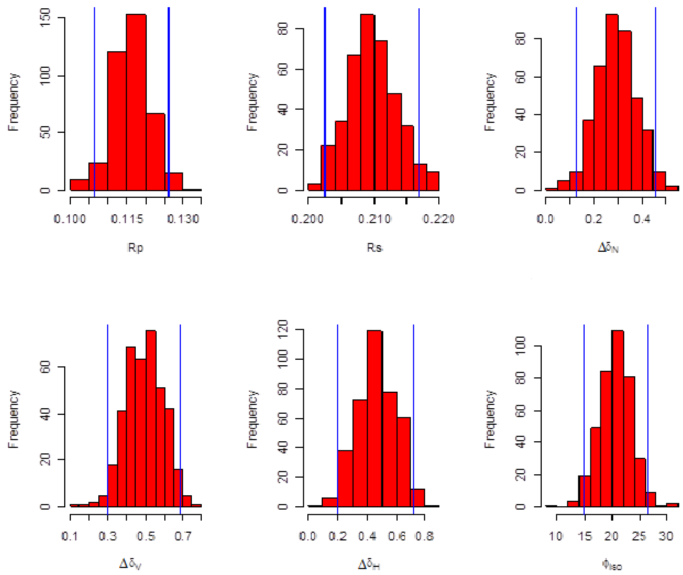

Figure 6: Histograms of each parameter for the six parameter study. 95% confidence intervals are marked between two blue lines.

well. The distribution of the parameters given an observation of $\mathbf{d}$ is actually the posterior probability distribution in Bayesian statistics, which is proportional to the product of likelihood function and the prior, i.e.,

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m})p(\mathbf{m}) \tag{4}$$

where the likelihood is determined by the distribution of noise(see [10] and reference therein). Here we assume that the noise ξ follows normal distribution $\mathcal{N}(0, \sigma^2\mathbf{I})$, then

$$p(\mathbf{d}|\mathbf{m}) \propto \exp\{-\frac{e(\mathbf{m})}{2\sigma^2}\} \tag{5}$$

Due to the prior $p(\mathbf{m})$, the posterior pdf $p(\mathbf{m}|\mathbf{d})$ is generally non-Gaussian. However, by Markov Chain Monte Carlo (MCMC) method(see [8]), we can construct a Markov chain $\{\mathbf{m}^{(k)}\}_{k=1}^{\infty}$ that has the distribution $p(\mathbf{m}|\mathbf{d})$ as its equilibrium distribution, bypassing the difficulty of knowing exactly what the distribution is. From these samples we can analyze the posterior distribution by statistical methods.

In general, the MCMC algorithm for this model is as follows:

1. Initialize samples $\mathbf{m}^{(0)} = (m_1^{(0)}, \ldots, m_n^{(0)})$ and $\sigma^{(0)}$.
2. In each iteration k, a sample of $m^{(k)}$, namely $m_i^{(k)}$ is generated based on

$$m_1^{(k)}, \ldots, m_{i-1}^{(k)}, m_{i+1}^{(k-1)}, \ldots, m_n^{(k-1)}$$

   and the nuisance parameter $\sigma^{(k)}$ is sampled based on $m_1^{(k)}, \ldots, m_n^{(k)}$.
3. Samples from the beginning of the chain (the burn-in period) are thrown away. In practice, the first few hundreds or thousands samples are considered burn-in. The remaining samples $\{\mathbf{m}^{(k_0)}, \mathbf{m}^{(k_0+1)} \ldots, \}$ are used to study the probability distribution of $\mathbf{m}$.

There are two implementations of step 2 that are used here. One of them is the Gibbs sampling, if the conditional distribution of the parameter we update is known; otherwise the Metropolis-Hastings method is applied.

$\mathbf{G}(\mathbf{m})$ is nonlinear in general, but it may be linear with respect to a subset of parameters, which is denoted by $\mathbf{m_l}$, and $\mathbf{m}$ is decomposed as

$$\mathbf{m} = \mathbf{m_l} \bigcup \mathbf{m_o}$$

Assuming non-informative prior $p(\mathbf{m_l})$, it can be shown that

$$p(\mathbf{m_l}|\mathbf{m_o}, \sigma^2) \propto \exp\{-\frac{e(\mathbf{m})}{2\sigma^2}\}$$

which is a multivariate normal distribution $\mathcal{N}(\mu, \Sigma)$ where

$$\mu = (C^tC)^{-1}C^t\mathbf{d}$$

and

$$\Sigma = \sigma^2(C^tC)^{-1}$$

here C is the Jacobian matrix of $\mathbf{G}(\mathbf{m_l}, \mathbf{m_o})$ with respect to $\mathbf{m_l}$.

Then with non-informative prior $p(\sigma)$, $P(\sigma^2|\mathbf{m})$ is the inverse-gamma distribution with shape parameter

$$\alpha = \frac{N-1}{2}$$

and scale parameter

$$\beta = \frac{2}{e(\mathbf{m})}$$

Therefore, $\mathbf{m_l}$ and σ^2 can be sampled from their respective conditional distribution, which is exactly Gibbs sampling.

The observation $\mathbf{d}$ is nonlinear with respect to $\mathbf{m_o}$, hence the conditional probability of $\mathbf{m_o}$ is difficult to obtain. However, the ratio of conditional probabilities of two samples of $\mathbf{m_o}$ has exact formula

$$\gamma = \frac{P(\mathbf{m_o}'|(\mathbf{m_l}, \sigma^2))}{P(\mathbf{m_o}|(\mathbf{m_l}, \sigma^2))} \tag{6}$$

$$= \exp\{-\frac{e(\mathbf{m_l}, \mathbf{m_o}') - e(\mathbf{m_l}, \mathbf{m_o})}{2\sigma^2}\} \tag{7}$$

Hence, $\mathbf{m_o}$ can be sampled by Metropolis-Hastings method in this way:

1. Generate a new sample $\mathbf{m_o}' = \mathbf{m_o} + stepsize \cdot randn$ where $randn \sim \mathcal{N}(0, \mathbf{I})$;
2. Use γ in (6) as the acceptance ratio to accept $\mathbf{m_o}'$ or reject it. When the new sample is rejected, the previous sample is kept as the current state in the Markov chain.

3.2 Numerical examples of MCMC

Here we apply the MCMC method to the five-parameter model (2), where the unknown parameter set is $\mathbf{m} = (\delta\Delta_T, \delta\Delta_N, R_p, R_s, \phi_{iso})$ and

$$\mathbf{m_l} = (\delta\Delta_T, \delta\Delta_N, R_p, R_s), \quad \mathbf{m_0} = \phi_{iso}$$

The model parameters are $g = 0.25, \chi = 1 - 2g, e_0 = 0.25, R_d = e_0/(1 + e_0)R_p\delta\Delta_V = \delta\Delta_H = \delta\Delta_T$, and the ground truth of unknown parameters are $\delta\Delta_T = 0.2, \delta\Delta_N = 0.1, R_p = 0.10, R_s = 0.20, \phi_{iso} = 27°$. The seismic data are simulated with 8326 input values of (θ, ϕ) with noise level SNR = 50.

Initially, $\delta\Delta_T, \delta\Delta_N, R_p, R_s$ are generated from $\mathcal{U}(0, 1)$, and ϕ_{iso} is drawn from $\mathcal{U}(0, \pi)$. Then a Markov Chain with totally 100,000 states is generated, with first 1000 the burn-in period. The remaining 99,000 states are considered as samples from posterior probability distribution $P(\mathbf{m}|\mathbf{d})$. The histogram of each unknown parameter and scatter plot of each pair of these parameters are shown in Figure 7 and Figure 8 respectively.

4 Conclusion

In geophysics there is a need to uncover knowledge of the material below the Earth's surface. Seismic data can be used mathematically to discover these properties. We implemented two well-known mathematical optimization algorithms to solve a problem in geophysics. Both Metropolis Simulated Annealing and Markov Chain Monte

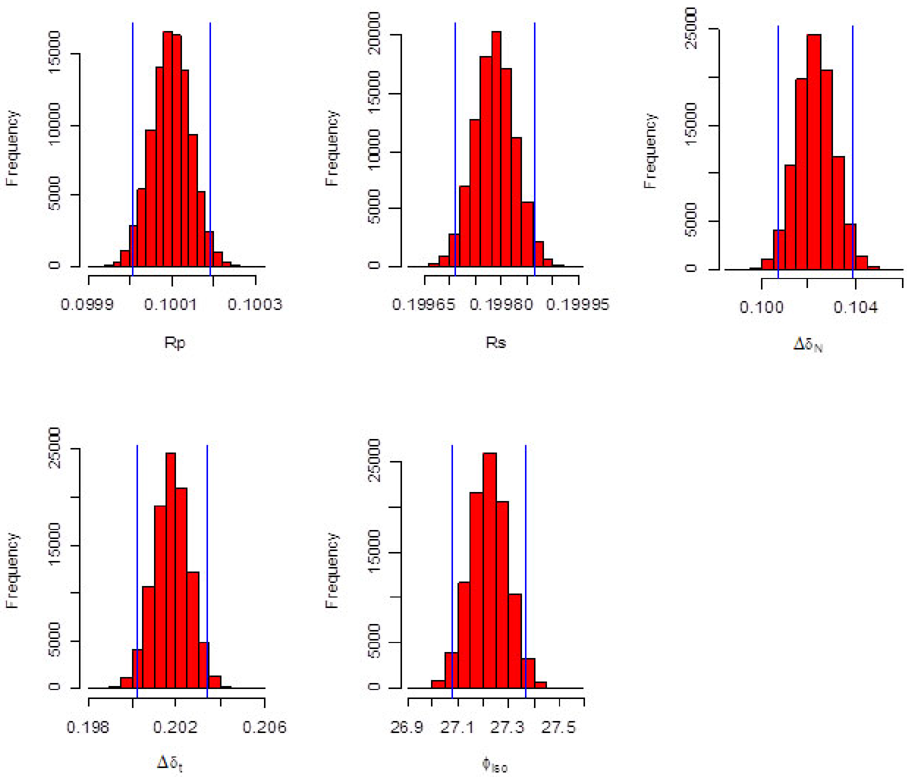

Figure 7: The histogram of unknown parameters $\delta\Delta_T, \delta\Delta_N, R_p, R_s, \phi_{iso}$. They are asymptotically normal distributed, and the 95% confidence interval is marked between two blue lines. You can see that the estimated mean value for each parameter deviates the ground truth by an amount of 2%.

Carlo proved adequate in solving for parameters given seismic data. Both are easily parallelizable and can be very efficient computationally given enough cpu cores. Markov Chain Monte Carlo does have an upper hand over Metropolis Simulated Annealing when many more parameters are introduced into the model because Metropolis Simulated Annealing will reject more perturbations when the dimensions increase. There are many mathematical techniques that are and still need to be implemented on geophysical problems.

Acknowledgments

We would like to thank PIMS, IMA, University of Calgary, NSERC, mprime and Government of Alberta for organizing this wonderful workshop. The first five authors are supported by travel grant from IMA and PIMS during this workshop.

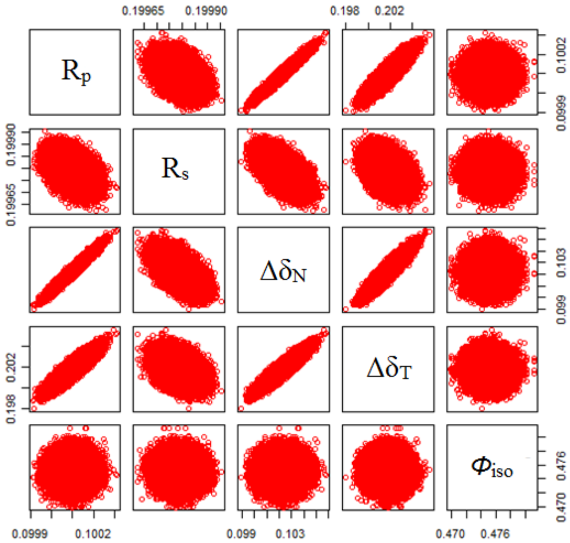

Figure 8: The scatter plot of pairs of parameters in $\{\delta\Delta_T, \delta\Delta_N, R_p, R_s, \phi_{iso}\}$. From this plot you can see that R_p and $\delta\Delta_N$, R_p and $\delta\Delta_T$, $\delta\Delta_T$ and $\delta\Delta_N$ are strongly positively correlated, but ϕ_{iso} is not correlated with any other parameters.

References

[1] A.M. Al-Marzoug, F.A. Neves, J.J. Kim, and E.L. Nebrija. P-wave anisotropy from azimuthal velocity and avo using wide-azimuth 3d seismic data. In *2004 SEG Annual Meeting*, 2004.

[2] J. Downton and D. Gray. Avaz parameter uncertainty estimation. In *2006 SEG Annual Meeting*, 2006.

[3] J. Downton and L. Lines. Avo feasibility and reliability analysis in the presence of random noise. 2001.

[4] J. Downton, B. Roure, L. Hunt, S. Reynolds, and S. Hadley. Predicting natural fractures in the tight nordegg gas sandstone of west central alberta using azimuthal fourier coefficients. 2011.

[5] V. Granville, M. Krivánek, and J.P. Rasson. Simulated annealing: A proof of convergence. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(6):652–656, 1994.

[6] D. Gray and D. Todorovic-Marinic. Fracture detection using 3d azimuthal avo. *CSEG Recorder*, 29:5–8, 2004.

[7] E. Jenner. Azimuthal avo. *The Leading Edge*, 21(8):782–786, 2002.

[8] J.S. Liu. *Monte Carlo strategies in scientific computing*. Springer Verlag, 2008.

[9] A. Rüger. *Reflection coefficients and azimuthal AVO analysis in anisotropic media*, volume 10. Soc of Exploration Geophysicists, 2002.

[10] M.K. Sen and P.L. Stoffa. Bayesian inference, gibbs' sampler and uncertainty estimation in geophysical inversion1. *Geophysical Prospecting*, 44(2):313–350, 1996.