

# A Bayesian Approach for Seismic Inversion at Roncott Research Area

Yaoting Lin<sup>1,\*</sup>, Wei Zhou<sup>2,3</sup> and Wenyuan Liao<sup>1</sup>

<sup>1</sup>Department of Mathematics and Statistics, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4

<sup>2</sup>Ministry of Education Key Laboratory of Oil and Gas Resources and Exploration Technology, (Yangtze University), Wuhan, Hubei, 430100

<sup>3</sup>School of geophysics and petroleum resources, Yangtze University, Wuhan, Hubei, 430100

\*Corresponding author

**Abstract**—In this paper, we embedded the genetic algorithm (GA) into the inner loop of the simulated annealing (SA) with a special design. The new method will boost the tunability of the searching process by providing two scales of regulation in seismic inversion problem. Moreover, a quantified uncertainty of the inversion result can be obtained when we put this strategy under Bayesian framework. Real data tests are conducted to support the theoretical calculation. Based on the conventional sparse spike inversion results, as a part of the prior information, our proposed method presents a superior quality and convincible uncertainty description.

**Keywords**—genetic algorithm; simulated annealing; optimal regulation; bayes theory; seismic inversion

## I. INTRODUCTION

A fundamental role of seismic inversion is to give a quantitative inference of the earth's interior physical properties. Due to the inherently complex and heterogeneous media and tectonic structures, it is rarely possible to obtain a deterministic solution of the inverse problem based on the indirectly related surface data and measurements only. From the perspective of statistic, we are unable to or should not apply a deductive reasoning to prove or disprove a certain model or some values of those model parameters<sup>[1]</sup>. So, it is more appropriate to establish a statistic framework in which we can cast an inverse problem and solve it by constructing the posterior probability density (PPD). In a Bayesian inference framework, the PPD gives all the consistent models a description with quantitative uncertainties, which allows us to directly calculate the probability of the target model or value of those parameters<sup>[1]</sup>.

The Bayes analysis was firstly proposed by Bayes in 1763, and was carefully discussed by Tarantola in 1987 since it began to be tried for geophysical inverse problems. In the works of Mosegaard, Tarantola, Gouveia and Moraes from 1995 to 1996, the Bayesian analysis was tested as an alternative to the traditional deterministic process. In 1996, Gouveia incorporated various experimental noises into the model and data covariance with subjectivity. Until 1998 and 1999, in some high-dimensional problems, Markov chain Monte Carlo methods have been widely used for geophysical inversion problems as well. The searching process based on Genetic Algorithms (GA) has been proved as a convergence procedure to a certain stationary distribution<sup>[2]</sup> and can be adapted as sampling tools<sup>[3]</sup>. In addition, GA and Simulated Annealing (SA) were

incorporated in the framework of MCMC. It improved the optimization and PPD simulation results<sup>[4-6]</sup>. The thought of Simulated Annealing (SA) method is introduced here as the framework of the processing. SA method was proposed in 1983<sup>[7]</sup>. It was pointed out that the annealing in solids provides a framework for optimization of the properties of very large and complex systems. Černý discussed about the analogy with thermodynamics that can offer an insight into the optimization problems and can suggest efficient algorithms for solving the Traveling Salesman Problem (TSP) based on importance sampling<sup>[8]</sup>. Granville, Krivanek and Rasson proved the convergence of SA procedure when the decision to change the current configuration is blind of the cost of the new configuration in 1994. Velis fully discussed the application of simulated annealing in seismic problems<sup>[9]</sup>.

On the other hand, many commonly used inversion methods are gradient-based. The discretization of those continuous parameter fields makes the optimization method hard to locate the true solution in preferred precision and with corresponding quantified uncertainty. To overcome this issue, a great deal of works has been done in introducing the stochastic method into seismic inversion problem. More hybrid methods, such as GA<sup>[10,11]</sup>、SA、PSO (Particle Swarm Optimization), are proposed to deal with such problems. However, they rarely consider the regulation in the optimal process, which can enable us to adjust the process in a balance between efficiency and accuracy.

The new method proposed in this paper aims to resolve this issue through providing a solution procedure that is easy to operate. It has been shown that this new combination strategy can boost the tunability of the whole searching process. As we embedded the GA process into the inner loop of SA with a special designed junction, the powerful global searching capacity of GA and the outstanding local searching ability of SA cooperated efficiently. As a result, it provides a two-scale regulation and an interface for external interference. Numerical results and corresponding confidences are performed in both model test and real data test.

## II. MODIFIED METHOD UNDER BAYESIAN FRAMEWORK

### A. The Basic Principles of this Theory Include:

Firstly, we establish the prior distribution of the unknown parameters, set up the forward model;

Secondly, we generate samples obeying the posterior distribution of the unknown parameters through the Bayes Formula;

Thirdly, statistical decision is made based on the analysis of the posterior distribution.

#### B. Traditional Methods Based on Likelihood Function

There are two problems in the traditional methods based on likelihood function: (1) generally it leads to the local optimal result, and will be trapped there since the results around are worse than this local one; (2) can the likelihood function value fairly judge a result? Especially in the problems of multiple solutions and random errors.

Here is how does the Bayes theory take these two problems into consideration: the evaluation of a result should be based on both the prior information and the likelihood function value. That means, a better result should have a better likelihood function value and fit better with the prior information. These two indexes can control the searching process and avoid it to be trapped into some local solutions. So, the likelihood function value is no longer a dictatorial parameter in optimization of the inversion problems.

In probability and statistics, the Bayes theorem provides a connection between the conditional probability and its inverse. For event A and B, if  $P(B) \neq 0$ , then:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \quad (1)$$

Here the  $P(A)$  indicates the prior probability of event A,  $P(A|B)$  indicates the posterior probability of event A when the event B has already happened. The  $P(B|A)$  can be interpreted as the support degree of the event A for event B, or the probability of the appearance of the event B when we already knew that the event A has happened. Or more directly, the  $P(B|A)$  functions as the likelihood value. At last,  $P(B)$  has little relation with the posterior probability and generally works as a constant value.

In seismic inversion problem, we can make the formula (1) more specific as:

$$P(m|d, I) = \frac{P(d|m, I)*P(m|I)}{P(d|I)}. \quad (2)$$

Here  $d$  is the observed dataset,  $I$  is the prior information and  $m$  is the model. Moreover, we can further define it as:

$$\begin{aligned} P(m|I) &= P(m_l = m_{l-1} + \Delta m_l | I) \\ &= P(\Delta m_l = m_l - m_{l-1} | I) = P(\Delta m_l | I). \end{aligned}$$

Here we use the disturbance to represent the prior probability in model iteration process. At the same time, we can use second and higher order terms to set up the likelihood function (that is the  $P(d|m, I)$  in formula (2)) under the assumption that the higher order terms and disturbance satisfy Gaussian distribution. The Taylor series expansion of  $d$  at  $m_k$ :

$$D = S + \sum_{k=0}^{n-1} \Delta m_k \frac{\partial S}{\partial m_k} + \frac{1}{2} \Delta m_k^2 \left( \sum_{k=0}^{n-1} \frac{\partial^2 S}{\partial m_k \partial m_k} \right) + \dots \quad (3)$$

We name  $\Delta d = D - S$ ,  $G(m_k) = \frac{\partial S}{\partial m_k}$ , then the likelihood function can be set as:

$$P(d|m, I) = \frac{1}{(\sigma\sqrt{2\pi})^N} * \exp \left[ -\frac{\sum_{n=1}^N (\Delta d - G(m_k) * \Delta m_k)^2}{2\sigma^2} \right] \quad (4)$$

Finally, we can rewrite the prior probability  $P(m|I)$  and likelihood function and posterior probability as follow:

$$\begin{aligned} P(m|d, I) &= \frac{1}{(\sigma\sqrt{2\pi})^N} * \exp \left[ -\frac{\sum_{n=1}^N (\Delta d - G(m_k) * \Delta m_k)^2}{2\sigma^2} \right] * \\ &\quad \frac{1}{\sqrt{|C_{\Delta m_k}|^3 * 2\pi^{\frac{3}{2}}}} \exp \left[ -\frac{(\Delta m_k)^2}{2C_{\Delta m_k}} \right] \end{aligned} \quad (5)$$

Here  $N$  is the length of observed data,  $C_{\Delta m_k}$  is the variance of model disturbance.

It should be clear that the posterior distribution of the parameters can only be estimated, instead of deterministically and analytically obtained. An important and widely used method is MCMC (Markov chain Monte Carlo) which can lead to the approximate posterior distribution from the prior distribution based on appropriate sampling method [12-14].

Here in this paper, we chose the genetic annealing evolutionary method with some modifications. The power of this combination lies in its strong global searching ability from the genetic algorithm and the precise local searching ability from the simulated annealing algorithm. Here, GA is embedded into SA algorithm to replace the original inner loop in order to take the advantages of both methods.

In standard SA, the searching process under the same temperature is a time homogeneous Markov chain. With GA method providing a series of parallel searching processes in the inner loop, the individuals of the Markov chain could be replaced by the population. So, it turns into a Markov chain of a whole population instead of individuals. The development of the generation in current temperature will achieve a steady state, in which the individuals can describe the solution in the form of a probability distribution.

In the outer loop of SA, according to the Metropolis-Hastings algorithm, an evaluation value is needed to determine the proposed population to be accepted or rejected. So here in this paper, we modified the selection and mutation operation methods in GA part. For each individual in the current proposed population, we partitioned its gene matrix into fragments. For each location of the gene matrix, we selected the best fragment through the whole population. As shown in Fig.1, the gray value presents the evaluation of a fragment. We picked out those fragments with high value and put them together to generate a new individual with all the best performed gene fragments. Actually, we selected the best gene fragments and assembled them. So, the artificial new individual, called as ‘King’ in this paper, accumulated nearly most of the preponderant schemas. It indicates the optimum individual the current population can

provide. As the population development controlled by the genetic simulated annealing algorithm, the Markov chain composed by the samples of the process will present the status of the optimization results (shown in Fig. 2).

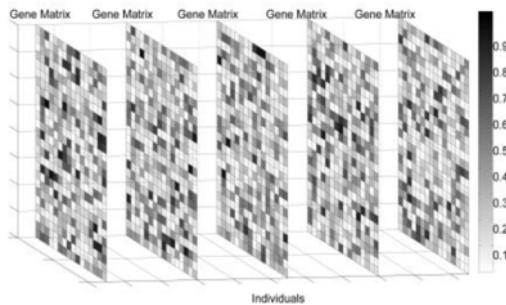


FIGURE I. GENE MATRIX

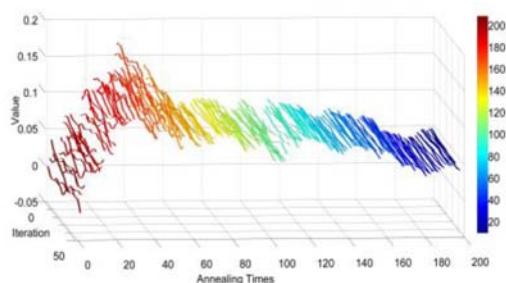


FIGURE II. ANNEALING PROCESS

As we can see from Fig. 2, in the state of high temperature, the result's evaluation value shows strong stochastic volatility. It means the global searching ability has the predominance. As the temperature decreases, the result moves to the real solution asymptotically. A suitable cooling strategy will make this process smooth and rapid. In the state of low temperature, the value shows weak stochastic volatility. It means the local searching ability rises to the predominance, which leads to a more precise result with relatively low deviation.

### III. APPLICATION

The Roncott research oil field is the main development zone of Saskatchewan, which is Canada's second-largest oil and gas producing province accounting for 20% of Canada total production. Roncott oilfield area is located within the province with reservoir distribution of Paleozoic carbonate strata. Under the influence of traps and trap ranges in size, final moisture content in oil production and changes in the reservoir is very large, but thin and dispersed oil concentration, facies changes and oil complex, strong segmentation, resulting in poor comparability between wells and wells. Hence, the conventional methods of seismic reservoir prediction have been proved difficult to meet the development needs of the block.

Numerous oil-bearing layers and great lateral variation of sand body led to the unclear reservoir distribution, complex oil-water relation and high dispersion of remaining oil. High quality inversion results can provide a solid foundation for the research of reservoir and fluid spatial distribution and other further development. In this work we take the 3D seismic data of this block shown in Fig.3, in which the color lines are the target

horizons with seismic data frequency around 50Hz.

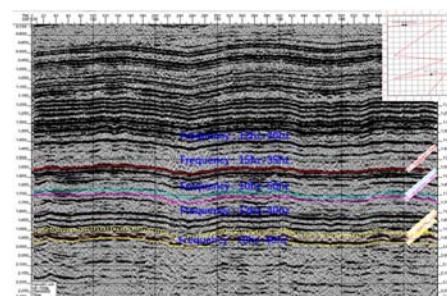


FIGURE III. TARGET SEISMIC SECTION

We selected part of the seismic data (about 1.25s to 2.00s) as the target strata for inversion. Figure 4 is the seismic impedance inversion section. This is a general seismic impedance inversion section, which is one of the most important means of reservoir prediction in oil and gas exploration and development. Here the warm tone (red) indicates the higher impedance while the cold tone (blue) indicate the lower impedance.

By study the Bayesian approach to seismic inversion based on GA-SA algorithm, we can see the result of the inversion section is structured and show more details. Especially on the bottom of the seismic-based, we can clearly find there are two favorable reservoirs. Low impedance of pore development indicates higher porosity.

Figure 4 is the normal time-domain seismic inversion. Sometimes we need to convert the time domain to depth domain because seismic data is the time-domain, and the well logs data is the depth-domain. We will transform the different domain which make it more convenient for researchers to provide more intuitive research data on the future deployment well locations. Figures 5-7 show seismic velocity inversion, seismic density inversion and seismic anisotropic inversion of strength in depth domain, respectively.

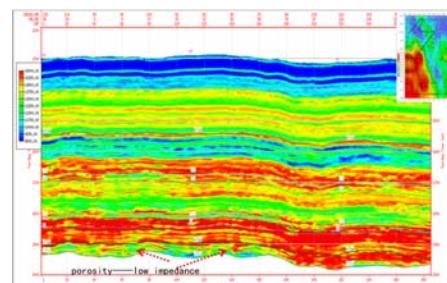


FIGURE IV. SEISMIC IMPEDANCE INVERSION SECTION

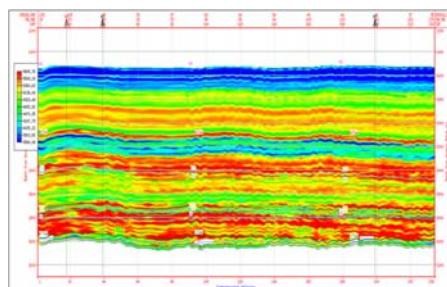


FIGURE V. SEISMIC VELOCITY INVERSION IN DEPTH DOMAIN

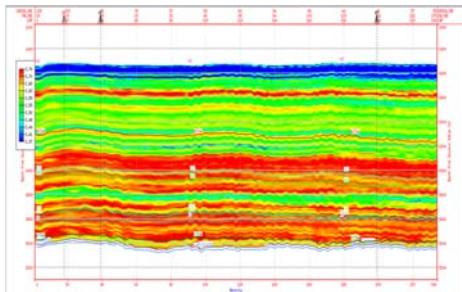


FIGURE VI. SEISMIC DENSITY INVERSION IN DEPTH DOMAIN

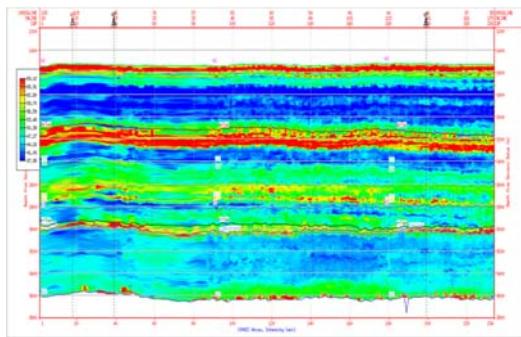


FIGURE VII. SEISMIC ANISOTROPIC INVERSION OF STRENGTH IN DEPTH DOMAIN

From the three depth-domain inversion sections, we can see that, using velocity inversion, density inversion and anisotropic inversion, only a small local structure is presented (black type explosion). The result of such a study confirm that there is a favorable local oil-bearing structure. The result from the Bayesian stochastic inversion method can be promoted to a certain extent. In some extreme cases, this method would present a result at least the same as, if not greater than, the prior model.

#### IV. CONCLUSION AND DISCUSSION

We have presented a modification seismic inversion method based on GA and SA in the framework of Bayes theory. Precision control on two different scales is provided through this design which improves the generalization ability. The method proposed in this paper is stable enough to estimate feasible solutions and provide a corresponding uncertainty analysis. It can promote the present inversion results of other methods.

We should point out that the Bayes stochastic inversion method promotes the quality of the inversion result based on previous works and achievements, including the results from former processing, or results of other methods, or in the worst-case scenario, the interpolation of the well log data. We cannot say it will provide a resolution-improved result without considering the prior information. We hope all the previous information act as a positive influence on searching process instead of the negative one, like the frequency band constrains in sparse spike inversion. The conventional result function as the expectation of the initial model set in our searching process. On the contrary, conventional inversion methods use low-frequency model, which decisively influent the whole searching process in the initial stages. Here in this paper, low-frequency model, or those low quality previous inversion results, function only in prior information and influences the acceptance probability of a

new generated proposal solution. However, study the Bayesian on GA-SA algorithm in the Roncott oilfield area, at least we can be sure that the method is practical in the Roncott oilfield area.

#### ACKNOWLEDGMENT

The work of the first author is jointly supported by Mitacs and Deep Treasure Corp. through Elevate Program. The second author's work is supported by the Key Laboratory of Oil and Gas Resources and Exploration Technology (Yangtze University). The third author's work is supported by Canada NSERC DG grant.

#### REFERENCES

- [1] Tiancong Hong, Mribal K. Sen, "A new MCMC algorithm for seismic waveform inversion and corresponding uncertainty analysis." *Geophysical Journal International*, vol. 177, pp. 14-32, 2009.
- [2] L. Davis, *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, New York, NY, USA, 1991.
- [3] Holmes, C., Mallick, B.K, "Bayesian radial basis functions of variable dimension." *Neural Computation*, vol. 10, pp. 1217-1235, 1998.
- [4] Seok-Hoon Oh, Byung-Doo Kwon, "Geostatistical approach to Bayesian inversion of geophysical data: Markov chain Monte Carlo method," *Earth Planets and Space*, vol. 53, pp. 777-791, 2001.
- [5] Knud Skou Cordua, Thomas Mejer Hansen, Klaus Mosegaard, "Monte Carlo full-waveform inversion of crosshole GPR data using multipole-point geostatistical a priori information," *GEOPHYSICS*, vol. 77(2), pp. 9-31, 2012.
- [6] James Martin, Lucas C. Wilcox, Carsten Burstedde, "A stochastic newton MCMC method for large-scale statistical inverse problems with application to seismic inversion," *SIAM J. SCI. COMPUT*, vol. 34(3), pp. 1460-1487, 2012.
- [7] Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P, "Optimization by Simulated Annealing," *Science*, vol. 220, pp. 671-680, 1983.
- [8] Černý, V, "Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm," *Journal of Optimization Theory and Applications*. vol. 45, pp. 41-51, 1985.
- [9] Pérez, D. O., Velis, D. R. and Sacchi, M. D, "Three-term inversion of prestack seismic data using a weighted l2, 1 mixed norm," *Geophysical Prospecting*, pp. 1-19, 2017.
- [10] Michael J.Brusco, "A comparison of simulated annealing algorithms for variable selection in principal component analysis and discriminant analysis," *Computational Statistics and Data Analysis*, vol. 77, pp. 38-53, 2014.
- [11] Maximilino J.Bezada, Colin A.Zelt, "Gravity inversion using seismically derived crustal density models and genetic algorithms: an application to the Caribbean-South American Plate boundary," *Geophysical Journal International*, vol. 185, pp. 577-592, 2011.
- [12] Thomas Mejer Hansen, Knud Skou Cordua, "Inverse problems with non-trivial priors: efficient solution through sequential Gibbs sampling," *Comput Geosci*, vol. 16, pp. 593-611, 2012.
- [13] Majken C. Looms, Thomas M. Hansen, Knud S. Cordua, "Geostatistical inference using crosshole ground-penetrating radar," *GEOPHYSICS*, vol. 75(6), pp. 29-41, 2010.
- [14] Jorge A. Sanchez-Ranirez, David Wolf, Carlos Torres-Verdin, "Synthetic and field examples of the Bayesian stochastic inversion of gamma-ray, density, and resistivity logs," *Annual Logging Symposium*, vol. 51, 2010.