MCMC for Large-Scale Geosteering Inversion with A Scalable MPI Implementation

Han Lu, Qiuyang Shen, Xuqing Wu, Jiefu Chen, Xin Fu
1 ECE, 2 ILT
University of Houston

***Abstract*—Inverse problems arise in many fields of science focusing on the process that explores the causal factors from which a set of measurements are observed. Compared with deterministic methods, statistical inversion is more capable of finding global optima of nonlinear inverse problems. In this paper, we propose a statistical approach based on the Markov chain Monte Carlo (MCMC) method and its implementation with the scalable dataset under the parallel environment Message Passing Interface (MPI). Traditional MCMC method launches one chain at a time. Our approach can simultaneously launch multiple Markov chains for one inverse problem. By calculating the data misfit and applying proper clustering algorithms, we can get a good estimation of the formation parameters from different inversion results. Numerical experimental evidences show that our parallel approach has better performance than traditional approaches. Meanwhile, it requires much less computational time with appropriate resource allocation.**

I. Introduction

Inverse problem is an important kind of mathematical problems in science and mathematics because they let us know parameters that can't be directly observed. In this study, we focus on the geo-steering inverse problems. Geo-steering is a technique to actively adjust the direction of drilling based on real-time well logging data such as directional electromagnetic (EM) logging measurements [1]. This process enables drillers to efficiently reach the target zone and actively respond to the geological changes so that they can maintain the maximal reservoir contact. Directional resistivity logging-while-drilling (LWD) tools are widely used in geo-steering because of its relatively large depth of detection as well as azimuthal sensitivity [1, 2]. As shown in Figure 1, T1, T2, T3 and T4 are the transmitters whose moments are with the tool axis, while T5 and T6 are transverse antennas that are perpendicular to the tool axis. Similarly, R1 and R2 are the receivers directing along the tool axis. R3 and R4 are the verical receiver antennas with directional sensitivity.

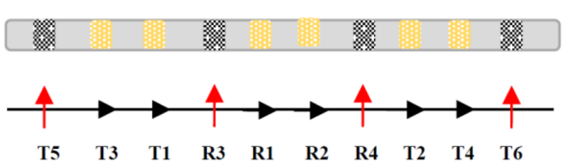

Fig. 1. The structure of an azimuthal resistivity LWD tool.

II. Background

The objective of the inverse process is to estimate model parameters through observed measurement. The geo-steering inversion output contains resistivity (or conductivity), location, and thickness of each underground formation layer. A reconstructed earth model can be drawn base on these parameters. Geo-steering inversion is very important information as it can be used to adjust trajectory of the drilling tool to minimize the gas or oil breakthrough and maximize economic production.

A. The statistical inversion methods

The statistical inversion is a branch of Bayesian statistics. The general starting point is the pair of random variables x and y, where x is the parameters of the object to be inverted and y denotes the measurement data. The relationship between observed output $\tilde{y}$ and corresponding model parameters can be represented as:

$$\tilde{y} = f(x) + \varepsilon \tag{1}$$

where f is a forward map and ε is the additive noise that usually has distribution: $\varepsilon \sim N(0, \sigma^2 I)$. The inverse problem can then be solved by Bayes' Theorem. The likelihood can be deduced as:

$$p(\tilde{y}|x) \sim \mathcal{N}(\tilde{y} - f(x), \sigma^2 I) \tag{2}$$

Suppose the prior distribution of x is governed by a zero-mean isotropic Gaussian distribution such that $p(x) \sim \mathcal{N}(0, \beta^2 I)$, then the posterior distribution of x is given by:

$$p(x|\tilde{y}) \sim \mathcal{N}(\tilde{y} - f(x), \sigma^2 I)\mathcal{N}(0, \beta^2 I) \tag{3}$$

The solution of x can be sampled and estimated according to the distribution function $p(x|\tilde{y})$.

B. The MCMC sampling method

The Metropolis-Hastings (MH) algorithm [3] is a popular MCMC method. A MH step of invariant distribution $p(x)$ and proposal distribution $q(x^*|x)$ involves sampling a candidate value x^* given the current value x according to $q(x^*|x)$. The Markov chain moved towards x^* with an accept rate, otherwise it remains at x. Algorithm 1 presents the MH algorithm for sampling from the posterior distribution:

Algorithm 1: The Metropolis-Hastings algorithm for sampling from $p(x|\tilde{y})$

input: initial value $x^{(0)}$, jumping function $q(x^{(i)}|x^{(j)})$
output: $x^{(k)}$ where $k \leq K$
begin
  initialize with arbitrary value $x^{(0)}$
  while length of MCMC chain < pre-defined length K
do
    Generate $x^{(k)}$ from $q(x^{(k)}|x^{(k-1)})$
    $\mathcal{A}(x^k, x^{k-1}) = \min\{1, \frac{p(x^{(k)}|\tilde{y})}{p(x^{(k-1)}|\tilde{y})}\}$
    Generate $\mathcal{A}_0$ from uniform distribution $\mathcal{U}(0,1)$
    if $\mathcal{A}_0 < \mathcal{A}(x^k, x^{k-1})$ then
      Keep $x^{(k)}$
    else
      $x^{(k)} = x^{(k-1)}$
    end
    save $x^{(k)}$ in the chain
  end
end

C. DBSCAN clustering method

Multiple chain MCMC geo-steering inversion may generate a set of inverted earth models. Here we use Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to cluster and screen the results of multiple chains. DBSCAN is a density-based algorithm: given a set of points in some space, it groups together points that are closely packed together (points with many neighbors). In our application, we simply chose the largest cluster as the screening result.

D. Parallel multiple chain MCMC method

For complicated geo-steering inversion problems, oftentimes a single MCMC chain cannot converge to global optimum with a fixed chain length. Running multiple chains can significantly enhance the possibility of achieving global solution. Figure 2 shows how we get reliable results from multiple chains.

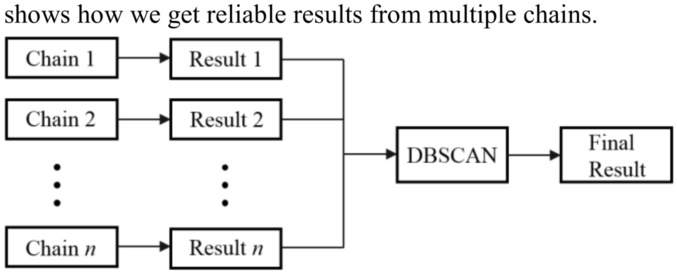

Fig. 2. The procedure of getting result from multiple chains.

III. Examples

QUESO (Quantification of Uncertainty for Estimation, Simulation and Optimization) is a parallel C++ statistical library for Bayesian Inference [5]. We use QUESO to run multiple chains for the geo-steering inversion under MPI. Shen et al [4] proposed a MCMC method on big-data platform and achieved good performance on a 3-layer problem. In this section, I will show the performance of our method on a more complicated 5-layer geo-steering problem with stronger nonlinearity. The earth model in Figure 3(a) is a 5-layer model with the resistivities of 10 Ω·m, 1 Ω·m, 100 Ω·m, 10 Ω·m and 50 Ω·m for the layers from top to bottom. The tool keeps moving in the middle layer, as represented by the black dot line. Both the point-wise inversion and multi-MCMC chains are parallelized and there are 80 points in total. We run 30 MCMC chains for each point and inverse 10 points at each time. In this test, we assume the tool's relative dip angle is fixed as 90 degrees. Figure 3(b) and 3(c) show the results of deterministic inversion and statistical inversion respectively. From these results we can see that parallel multi-MCMC method has better result than deterministic method.

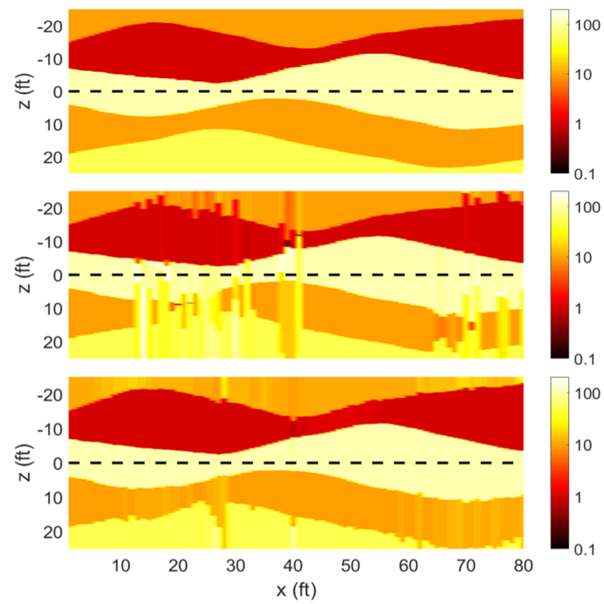

Fig. 3. (a) Real earth model; (b) Reconstructed model by deterministic inversion; (c) Reconstructed model by parallel mutil-MCMC method.

IV. Ackowledgement

This material is based upon work supported by the U.S. Department of Energy, Office of Science, and Advanced Scientific Computing Research, under Award Numbers DE-SC0017033.

References

[1] Q. Li, D. Omeragic, L. Chou, L. Yang and K Duong, "New directional electromagnetic tool for proactive geosteering and accurate formation evaluation while drilling," SPWLA 46th Annual Logging Symposium, Society of Petrophysicists and Well-Log Analysts, 2005.

[2] M. S. Bittar, "Electromagnetic wave resistivity tool having a titled antenna for determining the horizontal and vertical resistivities and relative dip angle in anisotropic earth formations," U.S. Patent No. 6,163,155. 19 Dec. 2000.

[3] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," Biometrika 57.1: 97-109, 1970.

[4] Q. Shen, X. Wu, J. Chen, Z. Han, "Distributed Markov Chain Monte Carlo Method on Bid-data Platform for Large-Scale Geosteering Inversion Using Directional Electromagnetic Well Logging Measurements," Applied Computational Elctromagnetics Society Journal 32.5, 2017.

[5] D. McDougall, N. Malaya, RD. Moser, "The Parallel C++ Statistical Library for Bayesian Inference: QUESO," arXiv preprint arXiv:1507.00398. 2015 Jul 2.