Distributed Markov Chain Monte Carlo Method on Big-Data Platform for Large-Scale Geosteering Inversion Using Directional Electromagnetic Well Logging Measurements

Qiuyang Shen [1], Xuqing Wu [2], Jiefu Chen [1], and Zhu Han [1]

[1] Department of Electrical and Computer Engineering
University of Houston, Houston, Texas 77204, USA
qshen4@uh.edu, zhan2@uh.edu, jchen82@central.uh.edu

[2] Department of Information and Logistics Technology
University of Houston, Houston, Texas 77204, USA
xwu8@central.uh.edu

Abstract **─** Inversion problems arises in many fields of science focusing on the process that explores the causal factors from which a set of measurements are observed. Statistical inversion is an alternative approach compared to deterministic methods with better capability to find optimal inverse values. Due to the increasing volume of data collections in the oil and gas industry, statistical approaches show its advantage on the implementation of large-scale inverse problems. In this paper, we address on the solution of big-data-scale inverse problems. After examining both conventional deterministic and statistical methods, we propose a statistical approach based on the Markov Chain Monte Carlo (MCMC) method and its implementation with the scalable dataset on the big data platform. The feasibility and methods to apply statistical inversion on the big data platform is evaluated by examining the use of parallelization and MapReduce technique. Numerical evidence from the simulation on our synthetic dataset suggests a significant improvement on the performance of inversion work.

Index Terms **─** Big data, geosteering, MapReduce, MCMC, multiple chains, well logging.

I. INTRODUCTION

Obtaining a reliable and detailed information about the earth's subsurface is of great challenging and requirement for the scientists and engineers to figure out the interior structures and then to be served for economic exploitation or geological prediction [1]. Inversion process is an organized set of mathematical techniques for projecting data to obtain knowledge about the physical world on the basis of inference drawn from observations [2]. The observations consist of a set of measurements from the real world. In this article, we are focusing on the background of geosteering inverse problems. Geosteering is a technique to actively adjust the direction of drilling, often in horizontal wells, based on real-time formation evaluation by using directional electromagnetic (EM) logging measurements [3]. This process enables drillers to efficiently reach the target zone and actively respond while drilling to geological changes in the formation so they can maintain the maximal reservoir contact. Among different types of logging techniques, azimuthal resistivity logging-while-drilling (LWD) tools are widely used in geosteering to provide electromagnetic measurements [3, 4]. A schematic model diagram of an azimuthal resistivity LWD tool is shown in Fig. 1.

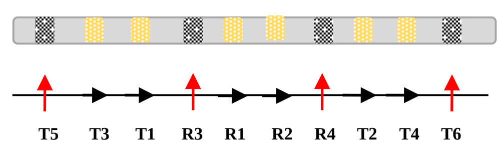

Fig. 1. The structure and schematic of an azimuthal resistivity LWD tool. T1, T2 T3 and T4 are the transmitters whose moments are with the tool axis, while T5 and T6 are transverse antennas that perpendicular to the tool axis. Similarly, R1 and R2 are the receivers directing along the tool axis. R3 and R4 are the vertical receiver antennas with directional sensitivity.

A set of antennas with different polarizations and working frequencies are installed on the azimuthal resistivity tool. They play the roles as transmitters and receivers that always in pairs and symmetric to the center of the tool. Each pair of symmetric antennas can provide a group of electromagnetic signals while the tool is rotating and drilling forwards. The signals which have a sensitivity to layer interfaces and medium resistivity are collected as measurements, and serve for inverse work.

Submitted On: December 24, 2016
Accepted On: January 28, 2017

The formation of an earth model is usually reconstructed through inverted parameters such as depth-to-boundary and layer resistivity. However, a huge uncertainty of parameters is introduced due to constraint propagation of EM signals when the tool is far from detecting boundary [5]. In recent years, a new generation of azimuthal resistivity logging-while-drilling tools, which have longer spacing between antennas and lower working frequency, with a much larger depth of investigation have emerged on the market.

Though more advanced techniques are applied on logging tools, the geosteering inverse problems are always challenging since given measurements are generally finite and imprecise and infinite number of solutions to the inverse problem exists due to the sparsity and uncertainty of the data, and incomplete knowledge of operating circumstance [6]. In the view of Bayes, the optimal solution of inverse parameters can be extracted from the solution set by the statistical characteristics of model structures, whereas the analytical solution for the posterior distribution is not always available anymore [7, 8]. In practice, the Markov Chain Monte Carlo (MCMC) method is one of the most popular sampling approaches to draw the samples from an unknown distribution. The MCMC method guarantees asymptotically exact solution for recovering the posterior distribution [9], though the computational cost is inevitably high while the most MCMC algorithms suffer from low acceptance rate and slow convergence with long burn-in periods. Launching multiple Markov chains is a choice to improve the sampling quality since that the samples from an individual chain have serial independence. A multiple chain MCMC method weakens the correlation of each sample, and thus improves the possibility of convergence [10]. However, the computational cost will not be reduced while traditionally each Markov chain runs in sequence. Fortunately, the strategy of multiple chains is well suited to the platform with support of parallel computing. It comes to be one of our interests to deploy the multiple chains MCMC inversion in an efficient way.

Meanwhile, more accurate and deeper investigation can produce a larger amount of measuring data. Though these data play a critical role in a better reconstruction of earth model, the sheer volume of data and the high dimensional parameter spaces involved in the inversion process means that the statistical methods can scarcely keep up with the demand to deliver in-time information for the decision-making. The current situation conforms to the trend of the age of big data, in which data's volume, variety and velocity are defined as three primary criterions [11]. Hence, the objective of this article is to find out and examine an appropriate strategy to apply the big data techniques on the scalable statistical MCMC method of multiple chains. Moreover, we will show how this scheme can be fit into the MapReduce programming model to take advantage of the potential speed up.

II. INVERSION ALGORITHMS

The adjusting of real time geosteering process is determined by the current working condition of the position and attitude of the tool to minimize the gas or oil breakthrough and maximize the economic production. The operations are relying much on the output from geosteering inversion, from which a group of parameters are generated to draw a reconstructed earth model with the constraint of the measurements. In the real job of geosteering, modeling and inversion are applied to 1D model, in which the layer interfaces are assumed infinitely extended and parallel to each other. A group of results containing distance-to-boundaries and resistivity of each layers are collected to represent the earth model at the current tool position. Along with the trajectory of the tool, inversion is conducted at a fixed interval of distance. Thus the inversed model parameters are grouped together to draw a whole subsurface profile.

A. The deterministic inversion methods

The deterministic inversion method in geosteering applications fit the model function to measured data by minimizing an error term between the forward model responses and observed measuring data. Assume a geosteering tool provides N measurements denoted by $m \in R^N$, while $x \in R^M$ represents M earth model parameters inverted from the measured data. A computation model function or so-called forward function $S: R^M \to R^N$ is designed to synthesize N responses from M model parameters. The forward transformation from model parameters to the responses, which is denoted as $S(x) \in R^N$, is a result of the interaction of a physical system defined by those model parameters. The objective of the inverse problem is to infer the model parameters through observed measurement. A good agreement between the response of the forward model and measured data will be reached if the inversed parameters of the physical model are accurate. The difference between the forward response and measurements is defined as data misfit $F(x)$, which is defined as:

$$F(x) = S(x) - m. \tag{1}$$

Since both forward responses and measurements are vectors, a cost function is defined as the square of L2 norm of the misfit function $F(x)$ as follows:

$$f(x) = \sum_{i=1}^{N} F_i^2(x) = \|F(x)\|_2^2, \tag{2}$$

where $f(x)$ is the cost function representing the magnitude of the data misfit. Hence, the inverse problem is to find the optimal model parameters, x, which minimize the cost function given a forward model function and measurements. Mathematically, this problem is presented as:

$$\underset{x \in R^M}{\text{minimize}} \;\; f(x). \tag{3}$$

This is an unconstraint nonlinear least-square minimization problem. Many iterative numerical algorithms, such as gradient decent method, Gauss-

Newton method, and the Levenberg-Marquardt algorithm (LMA), have been well established to solve this least-square problem [12, 13].

Though most of the aforementioned optimization algorithms are robust and have satisfying performance for convergence, most inverse problems are ill-posed with infinite solutions given finite measurements. In many nonlinear cases, the solution of gradient-based methods is highly dependent on the initial guess or getting trapped in a local minimum. Improvement can be made by using regularizer as an additional constraint. For example, the objective function can be regularized subjecting to the minimization of neighboring inverse points. For example, total variation (TV) regularized inverse problems can be written as:

$$\underset{x \in R^M}{\text{minimize}} \ \|S(x) - m\|_2^2$$
$$\text{s.t} \ \left|m - m_{pre}\right| < \lambda, \tag{4}$$

where the constraint λ regularizes the maximum change between neighboring inversed parameters m_{pre} and current model parameters, m [14]. However, this kind of methods rely much on the assumption that the prior underground information has been acquired fully and the ill inverse only happens in some of inverse points. Hence, the limitation of deterministic inversion methods becomes more prominent especially with the increasing scale of the problem.

B. The statistical inversion methods

Statistical methods arise as an alternative approach to deal with many ill-posed scientific inverse problems. This sort of methods based on the Bayes' theorem has attracted many attentions nowadays. It can be concluded as a method to obtain the posterior distribution from which the solution is deduced after combining the likelihood and the prior. The assumptions made by the forward model $y = f(x)$ (y is data and x denotes the earth model parameter) may not include all factors that affect measurements. Suppose the noise is additive and comes from external sources, the relationship between observed outputs $\tilde{y}$ and corresponding model parameters can be represented as:

$$\tilde{y} = f(x) + \varepsilon, \tag{5}$$

where ε denotes an additive noise. The experiments empirically suggest the additive noise usually follows a zero-mean Gaussian random distribution: $\varepsilon \sim \mathcal{N}(0, \sigma^2 I)$. With the given hypothesis to model parameters x and observed data $\tilde{y}$, the likelihood can be deduced as:

$$p(\tilde{y}|x) \sim \mathcal{N}(\tilde{y} - f(x), \sigma^2 I). \tag{6}$$

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

$$p(x|\tilde{y}) \sim \mathcal{N}(\tilde{y} - f(x), \sigma^2 I)\mathcal{N}(0, \beta^2 I). \tag{7}$$

It suggests that the posterior distribution of model parameters x given observations $\tilde{y}$ can be obtained by calculating the product of two Gaussian distributions. The solution of x can be sampled and estimated according to the probability distribution function $p(x|\tilde{y})$. It is an effective way to overcome the shortcomings of deterministic inversion especially when the problems are underdetermined (ill-posed) because of the large parameter space and the sparsity of the measurements. The earth model parameters are determined by sampling from posterior distribution $p(x|\tilde{y})$ while the measurements $\tilde{y}$ have been acquired.

Drawing samples from posterior distribution $p(x|\tilde{y})$ is challenging when f indicates a non-linear mapping relationship between x and $\tilde{y}$ since the analytical solution for the posterior distribution is not always available. In practice, the MCMC method is one of the most popular sampling approaches to draw the samples from an unknown distribution while the state of the chain after a number of steps can reach its equilibrium. Then the samples are selected to draw a desired distribution.

In the next section, we will first elaborate the implementation of the MCMC sampling method by the Metropolis-Hastings algorithm, and then explore the applicability pf the parallel computing scheme that suits MCMC statistical inversion and is scalable to large inverse problems. Finally, an application of the MapReduce model, a prevailing scheme on the big data platform, and its implementation on the statistical inverse problems, will be introduced.

III. LARGE-SCALE GEOSTEERING INVERSE PROBLEMS

The Metropolis-Hastings (MH) algorithm [15] is a popular MCMC method. In brief description, a MH step of invariant distribution $p(x)$ and proposal distribution or jumping function $q(x^*|x)$ involves sampling a candidate value x^* given the current value x according to $q(x^*|x)$. The Markov chain then moves towards x^* with the following acceptance probability:

$$\mathcal{A}(x, x^*) = \min\left\{1, \frac{p(x^*)q(x|x^*)}{p(x)q(x^*|x)}\right\}, \tag{8}$$

otherwise it remains at x. In the random walk MH algorithm, a zero-mean normal distribution is a popular choice of $q(x^*|x)$ as a symmetric candidate-generating function, which helps reduce the moving probability to $p(x^*)/p(x)$. 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 $q\left(x^{(i)}\middle|x^{(j)}\right)$
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 $q\left(x^{(k)} \middle| x^{(k-1)}\right)$
$\mathcal{A}\left(x^{(k)}, x^{(k-1)}\right) = \min\left\{1, \frac{p\left(x^{(k)} \middle| \tilde{y}\right)}{p\left(x^{(k-1)} \middle| \tilde{y}\right)}\right\}$

Generate $\mathcal{A}_0$ from uniform distribution $\mathcal{U}(0,1)$
if $\mathcal{A}_0 < \mathcal{A}\left(x^{(k)}, x^{(k-1)}\right)$ **then**
keep $x^{(k)}$
else
$x^{(k)} = x^{(k-1)}$
end
save $x^{(k)}$ in the chain
end
end

A. Distributed multiple chains MCMC methods

Although the MCMC method guarantees asymptotically exact recovery of the posterior distribution as the number of posterior samples grows, it may suffer from a large number of “burn-in” steps to reach the equilibrium and slow convergence [10]. Noted that the time cost may be prohibitively high for the inverse problem when the forward model computation is required by every sample drawn from the MH, which may take at least O(N) operations to draw one sample [9]. Meanwhile, with the increasing space of model parameters, burn-in period may reach over thousands of steps. As both data volume and the space of model parameters increase explosively, a dispersion manner of the MCMC algorithms, which distributes computation to multi-processor and multi-machine, is urgently required. It is in our interests to investigate and provide answers to these questions on how to extend the proposed framework to serve for our statistical inverse problems that involve in big volume of data measurements, variety of measurement types and require high speed on data processing.

The essence of the distributed implementation is enlightened by the application of multiple chains (or multiple sequence) of the MCMC method proposed by Gelman and Rubin, 1992. It is based on the idea that the samples from a sequence of chain has a tendency to be unduly influenced by slow-moving realization of the iterative simulation. Whereas a multiple starting points are needed to avoid inferences being influenced [10]. The evidence apparently shows in geosteering statistical inversions and it is not generally possible to reach convergence by running a single chain MH algorithm. The main difficulty is that the random walk can remain for many iterations in a region heavily influenced by the starting distribution, especially in the case of sampling multidimensional random variables with a strong correlation. Hence, a multiple chains MCMC method has been proposed by starting multiple independent chains of the iterative sampling at multiple starting points. The target distribution and the estimation of model parameters can be obtained more quickly by the samples using between-sequence as well as within-sequence information [10]. Once all sequences reach the maximum chain length, an easy estimation of model parameters can be accomplish by sampling results from each chain instead of an inference from the time-series structure of samples from one single chain.

However, efficiency cannot be achieved while multiple chains are running on a single thread since all sequences are queued up and executed one by one. Due to a rapid expansion of computer science, the parallel computing technique is well developed serving for the parallelized tasks [16]. The parallelism can be supported not only within the level of a single machine with multiple-cores or multiple-processors, but also on the scale of clusters, grids and clouds. The multiple chains MCMC method is able to take full advantage of data and task parallelism. A simple strategy of a distributed MCMC method is built on the parallelization of multiple chains, which distributes the data and the task of the MCMC sampling method to multiple processing units.

Nevertheless, hundreds of measured data are collected for one measuring point by the geosteering tool and tons of measurements will be yield while the operating region may extend to thousand feet long. Therefore, the computation structure of the geosteering inverse problem needs to be able to accommodate multiple level of parallelism. The parallelization scheme of geosteering inversion is depicted in Fig. 2. The point-wise inversion tasks with respective measured data are distributed to multiple cluster nodes by the master node. Within each task, multiple chains are launched and run in parallel on the multiple processors or cores. Saturated with huge volumes of collected data, the computation power is well utilized by distributing the task and data to every computing node and thread. A facile solution for solving the large-scale statistical inverse problem in parallel is to take the advantage of the big data platform.

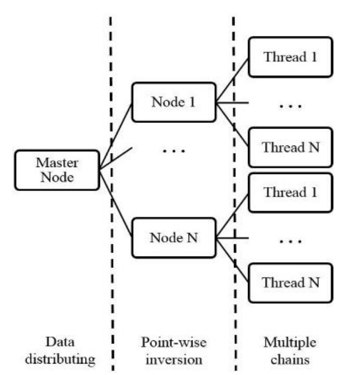

Fig. 2. Parallelization of large-scale statistical geosteering inverse problems on distributed clusters.

B. Large-scale geosteering inversion on MapReduce

To keep up with the requirement for large-scale parallel computing, we need a platform that allows easy access to data and is transparent to its scheduling process. Moreover, the platform should be able to dynamically allocate it computing resources, such as the number of cluster nodes or processing units, and achieve high scalability with low implementation costs. MapReduce is an idea model to meet these requirement.

MapReduce is a programming model and an associated implementation for processing and generating large data sets with a parallel, distributed algorithm on a cluster [17]. It serves for a purpose that allows programmers without any experience with parallel and distributed systems to easily utilize resources in a distributed computing environment. The model is composed of a map function which performs a preprocessing to the data collection, and a reduce function that performs a summary operation. A simple diagram of MapReduce is shown in Fig. 3. A master node is responsible for generating initial starting value set and assigning each starting value to a map function. The mappers filter and sort the allocated data and generate associative value. These intermediate outputs are sent to the reducers which aggregate the data and make calculation for the final outputs.

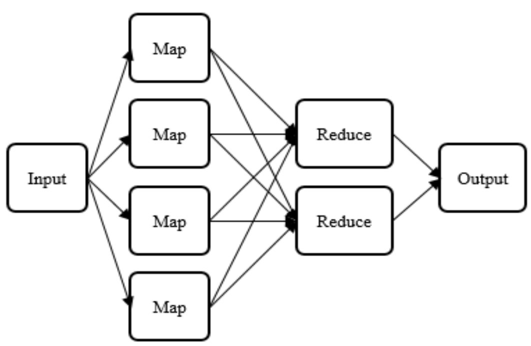

Fig. 3. Diagram of MapReduce programming model.

In the case of geosteering inverse problems, the MapReduce programming model can be suitably applied on the multiple chains MCMC method. Initially, multiple sets of initial values of model parameters are generated by a master node as the starting value for multiple chains. Each set is sent to a mapper where MCMC are running simultaneously for all sets. Samples generated by each chain are cached locally on map nodes for subsequent map-reduce invocations. After intermediate data is collected, the master node coordinates the mappers and reducers and in turn invokes the reducers to process samples from Markov chains. The reducer will calculate the variance within and between chains, evaluate convergence and statistical characteristics of all samples and output them as the final result.

IV. EXAMPLES AND DISCUSSIONS

In this section, we will evaluate the computation and convergence performance of the parallel multi-chain statistical geosteering inversion. We will take a synthetic earth model as an instance to examine the scheme. It is noted that the synthetic data is generated based on analytical EM wave solutions for the 1D multi-layer model with dipole transmitters and receivers. The speed of inversions depends largely on the forward program. In the experiment, the forward model is treated as a black box that returns a group of measurements given earth model parameters. The test ran on a DELL PowerEdge T630, with dual Intel Xeon E5-2667 V3 3.2GHz 20M Cache 8C/16T, and 8x16GB RDIMM 2133MT/s.

To verify the computational efficiency of the parallel computing, we firstly examine the time cost by launching multiple chains at fixed length on different numbers of processing units. The inversion is conducted on a single point of a three-layer model with five parameters: the resistivities of the upper, middle and lower layers are 10 Ω·m, 50 Ω·m and 1 Ω·m, respectively. The distance to the upper and lower boundary is 7 ft and 10 ft. Figure 4 shows the relationship between the time cost and the number of processing units. We archived significant performance gain when increasing processing unit from one to four. And the improvement is observed for all chain lengths. We also noted that the performance gain became less when adding more processing units. We think the execution latency variation is due to other shared resource conflicts when running the test on a shared memory platform.

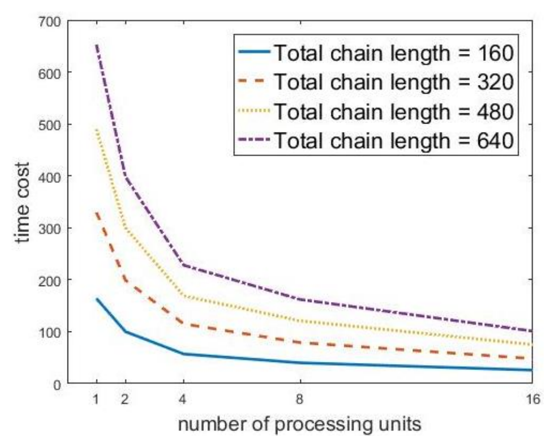

Fig. 4. Comparison of time cost by different total length of chain on different number of processing units.

To examine the convergence of inversion, we run two instances separately. First, we tested a single chain with 640 iterations. Second, we run eight chains simultaneously with the same number of iterations. We compare the accuracy of the inversed results through the unnormalized parameters misfit, a L2 norm of the difference between the inverted model parameters and the true model parameters. In Fig. 5, the result shows that both instances have reached the equilibrium state after a burn-in period. The parameter misfit of multiple chains is smaller and inversed results are more close to the ground truth in comparison with the single chain result. The test demonstrates the advantage of using multiple chains to reduce negative effects of parameter correlation and improve convergence accuracy. It should be noted that the time cost for running eight chains simultaneously is nearly the same as running a single chain on our parallel testing bed.

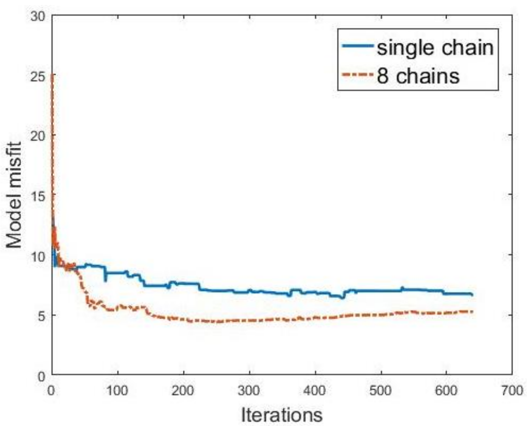

Fig. 5. Comparison of model misfit convergence.

Our next test is to reconstruct a subsurface profile by applying the parallel multi-chain MCMC statistical inversion to a three-layer model. The earth model in Fig. 6 (a) is a three-layer model with the resistivities of 10 Ω·m, 50 Ω·m and 1 Ω·m in the upper, middle and lower layer respectively. The central green line indicates the tool's navigation trajectory. In this case, we assume the tool dip angle is fixed as 90 degrees. The inversion is conducted at each 1ft while the total working region extends to 80 ft. The depth to either the upper or lower boundary is varying on different position. The largest depth to boundary (D2B) is around 18 ft and the smallest is around 2 ft. Both operations, point-wise inversion and multi-chain MCMC sampling, are parallelized. We validate the feasibility of the implementation for large-scale inverse problems on the big data scheme. Figure 6 (b) shows the inverse result which provides a successful and satisfying reconstruction of the earth model compared with the Fig. 6 (a).

Figure 7 also shows the quantified interpretation for the uncertainty of inversed results. The variance at each inversed point denotes the uncertainty of the estimation of the parameter. One can see from Fig. 7 that when the tool is far from the top boundary (the beginning part), the earth model cannot be well reconstructed. In other words, the uncertainty involved in solving the inverse problem is high. The reason is that the sensitivity of the measurements is relatively poor when the tool is far from the layer interface. As the tool moves forward and the top boundary bends downward, the inversion can clearly resolve both the upper and lower boundaries.

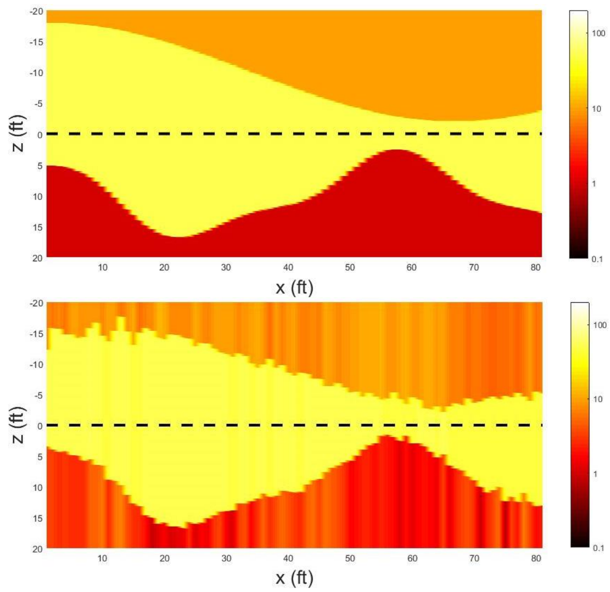

Fig. 6. (a) A synthetic three-layer earth model in 80 ft horizontal region, and (b) the inverse earth model.

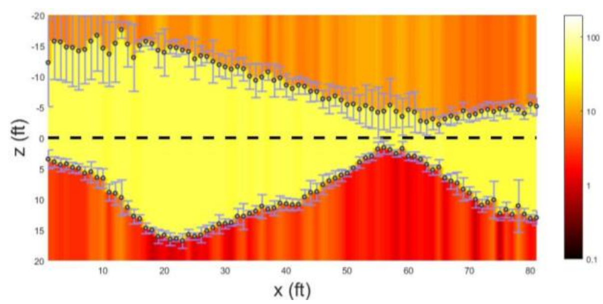

Fig. 7. The inverse earth model with uncertainty.

V. CONCLUSION

Since more advanced measuring technologies have been commercialized and pushed into the market around these years, the increasing volume of measured data undoubtedly will require a more efficient solution for solving large-scale inverse problems. The proposed work in this articles takes full advantage of the latest technology in big data in exploration of computationally efficient methods for solving statistical inverse problems. The multiple chains MCMC method is well suited for using the MapReduce framework in a distributed or cloud environment. Our simulation verifies the feasibility of solving large-scale inverse problem on a big data platform. The techniques developed in this research could have a large positive impact in exploring computationally efficient method to solved real-time statistical inverse problem.

REFERENCES

[1] M. Moorkamp, *Integrated Imaging of the Earth: Theory and Applications*. John Wiley & Sons, 2016.

[2] M. S. Zhdanov, *Inverse Theory and Applications in Geophysics*. Elsevier, 2015.

[3] 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.

[4] M. S. Bittar, "Electromagnetic wave resistivity tool having a tilted 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.

[5] W. G. Lesso Jr. and S. V. Kashikar, "The principles and procedures of geosteering," *SPE/IADC Drilling Conference*, *Society of Petroleum Engineers*, 1996.

[6] K. Key, "1D inversion of multicomponent, multifrequency marine CSEM data: Methodology and synthetic studies for resolving thin resistive layers," *Geophysics*, vol. 74, no. 2, F9-F20, 2009.

[7] A. M. Stuart, "Inverse problems: A Bayesian perspective," *Acta Numerica*, vol. 19, pp. 451-559, 2010.

[8] B. T. Knapik, A. W. van Der Vaart, and J. H. Van Zanten, "Bayesian inverse problems with Gaussian priors," *The Annals of Statistics*, pp. 2626-2657, 2011.

[9] W. Neiswanger, C. Wang, and E. Xing, "Asymptotically exact, embarrassingly parallel MCMC," *arXiv preprint arXiv: 1311.4780*, 2013.

[10] A. Gelman and D. B. Rubin, "Inference from iterative simulation using multiple sequences," *Statistical Science*, pp. 457-472, 1992.

[11] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, "Big data: The next frontier for innovation, competition, and productivity," 2011.

[12] K. Levenberg, "A method for the solution of certain problems in least squares," *Quart. Appl. Math.*, vol. 2, pp. 164-168, 1944.

[13] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM J. Appl. Math.*, vol. 11, pp. 431-441, 1963.

[14] Q. Shen, J. Chen, and Z. Han, "Total variation denoising of azimuthal resistivity inversion in geosteering," *SEG Technical Program Expanded Abstracts 2016*, *Society of Exploration Geophysicists*, pp. 719-723, 2016.

[15] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97-109, 1970.

[16] L. Murray, "Distributed Markov chain Monte Carlo," *Proceedings of Neural Information Processing Systems Workshop on Learning on Cores, Clusters and Clouds*, vol. 11, 2010.

[17] J. Dean and S. Ghemawat, "MapReduce: simplified data processing on large clusters," *Communications of the ACM*, vol. 51, no. 1, pp. 107-113, 2008.

Qiuyang Shen received his B.S. degree in Telecommunication Engineering from Huazhong University of Science and Technology in 2015. Now he is a second year Ph.D. student at the Department of Electrical and Computer Engineering, University of Houston, co-advised by Dr. Jiefu Chen and Dr. Zhu Han. His research interests include geophysical inversions, data optimization and machine learning.

Xuqing Wu received the B.S. degree in Electronic Engineering from the University of Science and Technology Beijing in 1995, the M.S. degree in Computer Science from the Carleton University in 2001, and the Ph.D. degree in Computer Science from University of Houston in 2011.

From 2001 to 2007, he was a Software Engineer for EMC/INTEL/HP. He was a Data Scientist of Schlumberger before joining the University of Houston in 2015. Currently, he is an Assistant Professor in the Information and Logistics Department at the University of Houston, Texas. His research interests include multiphysics interpretation, machine learning, computer vision, high performance, mobile and cloud computing.

Jiefu Chen is an Assistant Professor at the Department of Electrical and Computer Engineering, University of Houston. Previously, he worked for 4.5 years as a Staff Scientist for Weatherford International. His current research interests include electromagnetic modeling and inversion, subsurface wireless communication, and well logging. Chen holds a B.S. degree in Engineering Mechanics from Dalian University of Technology and a Ph.D. degree in Electrical Engineering from Duke University.

Zhu Han received the B.S. degree in Electronic Engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Maryland, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a Research Associate at the University of Maryland. From 2006 to 2008, he was an Assistant Professor at Boise State University, Idaho. Currently, he is a Professor in the Electrical and Computer Engineering Department as well as in the Computer Science Department at the University of Houston, Texas. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. Han received an NSF Career Award in 2010, The Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (Best Paper Award in IEEE JSAC) in 2016, and several Best Paper Awards in IEEE conferences. Currently, Han is an IEEE Communications Society Distinguished Lecturer.