Improving efficiency of traveltime tomography by stochastic optimization
Mengyao Sun*, Jie Zhang, Wei Zhang. University of Science and Technology of China (USTC)

Summary
The first-arrival traveltime tomography is an important approach for solving the near-surface imaging problem, from which results are needed to calculate statics. The method involves inverting massive traveltime picks and therefore may require heavy computation. We apply mathematics methods of Sample Average Approximation (SAA) and Stochastic Approximation (SA) to improve the efficiency of traveltime tomography. SAA and SA are the two approaches to enhance the computational efficiency theoretically, which are realized by selecting a small portion of data randomly to perform the inversion as opposed to all data and that leads to a consequence of saving the computational memory, which can save about 95% cost. However, by applying the SA method, the cost of both computational memory and time can be saved at about 95% and 75%, respectively.

Introduction
Computational cost is an inevitable issue that should be considered for geophysical imaging problem. Especially for the industry today, more computation techniques have been developed and the data that needs to be processed becomes much larger, which requires solutions for reducing the computational cost, and the result accuracy should not be compromised at the same time.

Motivated by the mathematical concept of Sample Average Approximation (SAA) (Kleywegt et al., 2002) and Stochastic Approximation (SA) (Nemirovski et al., 2009), we intend to enhance the efficiency of the first-arrival traveltime tomography (Zhang and Toksöz, 1998) by applying these methods. This research is especially meaningful for dealing with large 3D datasets. Herrmann et al. (2011) reduce the computational cost by introducing SA and compressive sensing for Full Waveform Inversion (FWI) problem (Tarantola, 1984). Herrmann (2011) also applies a similar method to an efficient least-squares migration. We focus on reducing the computational cost of the traveltime tomography by applying the concepts of SAA and SA. Our tests verify that both the SAA and the SA methods can be used to reduce the computational cost of traveltime tomography. The different advantages of these two methods will also be analyzed in this research. We further discuss about how to select data and how to deal with the cases with different acquisition geometry.

Objective function
The first-arrival traveltime tomography is described as follows:

\[ \varphi(m) = \min \left\{ \sum_{i=1}^{K} \| d_i - G(m; q_i) \|_2 + \lambda R(m) \right\} \]  (1)

here, \( m \) is the model slowness; \( d_i \) is the observed traveltime data which is corresponding to the \( i^{th} \) source; while \( G(m; q_i) \) is the calculated traveltimes which are obtained by a wavefront raytracing technique. \( \lambda \) is a constant parameter for balancing the data misfit and the regularization \( R(m) \). \( K \) is the number of source, \( q_i \) represents the \( i^{th} \) source.

Cost reduced first-arrival traveltime tomography:

\[ \hat{\varphi}(m) = \min \left\{ \sum_{i=1}^{K} \| W_i d_i - G(m; W_iq_i) \|_2 + \lambda B(m) \right\} \]  (2)

where \( W_i \) is an operator that should extract data from the observed traveltime data, \( j \) is the sequence number of inversion iteration.

SAA and SA

In this cost reduced first-arrival traveltime tomography, we introduce an operator \( W_j \) (Equation 2) and ensure the inversion a stochastic optimization problem. SAA is an efficient approach for solving the problems with the following form:

\[ \min_{m} \{ \varphi(m) = E_P[\varphi(m; W)] \} \]  (3)

here, \( E \) represents the mathematical expectation; \( W \) is a random vector with probability distribution \( P \). Kleywegt et al. (2002) propose a solution for this type of problem. For this problem, assuming that \( W^1, \ldots, W^N \) are \( N \) random samples of \( W \) with the property of independent and identically distributed (i.i.d). Then the corresponding sample average function is

\[ \hat{\varphi}_N(m) = \frac{1}{N} \sum_{i=1}^{N} \varphi(m; W^i) \]  (4)

Then the sample average approximation problem is:
Improving efficiency of traveltime tomography by stochastic optimization

\[ \min_m \phi_g(m) \quad (5) \]

Therefore, the problem of minimizing the objective function \( \phi(m) \) is converted to minimize \( \phi_g(m) \). The convergence of this sample average approximation problem (Equation 5) is verified by Kleywegt et al. (2002).

In the synthetic tests, we verify that the SAA method can be used to solve the first-arrival traveltime tomography. That motivates us to apply the SA method to realize the further computational cost reduction. The main idea of SA is to update the operator \( W_j \) (Equation 2) at every iteration. That means \( W_j \neq W_{j+1} (j = 1, \ldots, n - 1) \), \( n \) is the max iteration number of the first-arrival traveltime tomography. For the SAA method, \( W_j \) is fixed in every iteration.

**Synthetic test**

Figure 1a shows the true model of the first synthetic test. The model length is 5120 m, the depth is 620 m. There are 120 shots and 250 receivers in total, which are shown in Figure1 by red and yellow dots. The spacing of shots and receivers is 40 m and 20 m, respectively. A common spread of receivers is fixed and covers the entire surface area of the model. Figure 1b shows the inversion result with full data (120 shots), 20 iterations, and we will take the result obtained by full data as the standard result in the following.

First, we test the SAA method on this synthetic model. We perform 20 inversions. Each inversion includes 20 iterations. That is, we perform 400 iterations in total. In each inversion, we extract 4 shots randomly and the extraction pattern follows i.i.d. That means the extraction pattern is different in each inversion. While in certain inversion, the extraction pattern is fixed. Figure 1c is the solution obtained by the SAA method. A good similarity presents in Figure 1b and Figure 1c. We also show the inverted results of the first two inversions using the SAA method in Figure 1e and Figure 1f, their difference indicates that we can use just a part of data to obtain a satisfied approximate solution. This is because of the SAA method but not data redundancy. Then we introduce the SA method. We use the 20 same extraction patterns as the SAA test and perform 20 iterations. Every iteration, we update the extraction pattern. Figure 1d displays the approximated solution obtained by SA. Both SAA and SA give a good approximated solution comparing with the standard result.

![Figure 1](image-url)

Figure 1: (a) The true model of the first synthetic test. (b) The standard result. (c) The SAA approximated result. (d) The SA approximated result. (e) The first inversion result of the SAA test. (f) The second inversion result of the SAA test, the other 18 inversion results are omitted here.
Improving efficiency of traveltime tomography by stochastic optimization

The reason that SA or SAA gives an approximated solution but not the same result as the standard approach is that some strict mathematical conditions of SAA and SA cannot be satisfied in practical implementation. For SAA, practically, one cannot perform infinite number of inversions, which has to be replaced by finite number of inversions. Herrmann et al. (2011) discuss the convergence of the SA method, they point out that one pre-condition in SA cannot be guaranteed strictly in the model updating procedure. However, it still shows better performance on the practical implementation. We deal with the similar problem as Herrmann et al. (2011) do. Therefore, to obtain the same solution as the standard result, one just needs perform few iterations with full data as input after the application of the SA and the SAA methods. As our tests, the iteration number with full data will not be very large, five iterations should be enough. Eventually, we perform five full data iterations using the approximated solutions which are obtained by applying SAA and SA as the initial models. The final results of these three methods are almost the same but the time consuming is 105.012 s, 95.571 s and 26.373 s, which is corresponding to full data, SAA and SA, respectively.

We assume that the total shot number of a case is \( F \), the extracted shot number is \( F_e \). For the SAA method, before the final \( M \) iterations, full data tomography is implemented, \( T \) inversions SAA tomography with \( N \) iterations in each inversion should be performed. However, for the SA method, one just need run \( N \) iterations of SA tomography and \( M \) iterations of full data tomography in one inversion. Therefore, the save of computational cost can be described as following equations:

\[
\text{SAA: } \text{save} = 1 - \frac{T NE + MF}{F(N + M)} \times 100\% \quad (6)
\]

\[
\text{SA: } \text{save} = 1 - \frac{F(N + M)}{F(N + M)} \times 100\% \quad (7)
\]

Obviously, SA is superior to SAA in terms of the computational time cost. Although SA is lack of convergence and instable with respect to noise and it is not ready for extending to more sophisticated problems, it still has been practically verified possessing the ability of obtaining a good approximated solution (Herrmann et al., 2011). Therefore, for the following tests, we compare the performance of the SA method with the standard approach. While, SAA is more stable from a practical point of view, even the speed might be slower than SA or the same as the standard approach, it still owns the advantage of dividing a large case into several small cases and saving the computational memory as the result.

In real cases, the acquisition geometry that we design above is not common. Usually, the receivers will not cover the entire surface area but with a fixed number for every shot. For such cases, there is a stricter extracting rule should be followed if we plan to use SA to reduce the computational cost. In the above synthetic SA test, we extract 4 shots for every iteration and we perform 20 iterations. Even there is no repeated extraction, we just extract 80 shots in total for doing the inversion but not the full data, 120 shots, and it gives a satisfying approximated result. Furthermore, we test that even we just extract 1 shot for every iteration, it still could give a satisfying approximation. However, if the acquisition geometry is changed, one should follow a stricter rule: the total extracted shot number for the entire inversion procedure should not be smaller than the total shot number at least. Since the total iteration number is usually a fixed value, the number of extracted shot can be calculated. For example, Figure 2a shows the second synthetic true model, the length and depth are 33000 m and 790 m, respectively. There are 550 shots with the spacing of 60 m, 200 receivers for every shot and the spacing is 30 m. We totally perform 25 iterations. The first 20 iterations employ the SA method. Every iteration we should extract 28 shots. The final 5 iterations is implemented with the full data as the input. However, the standard result is obtained by using the full data during the entire inversion. Figure 2b and Figure 2c display the standard result and the result obtained by the SA method, respectively. The two different methods show the almost same inversion results. Nevertheless, the computational cost is significantly different. According to the Equation (7), the SA method saves about 76% computational cost, and also 95% memory in the first 20 iterations.

![Figure 2](image.png)

Figure 2: The second synthetic test. (a) The true model. (b) The standard result. (c) The SA result.
Real data application

We also apply the SA method to a set of real data. There are 558 shots and 1124 receivers in this real data. The shot spacing and receiver spacing are 60 m and 30 m, respectively. For every shot, there are about 100-250 receivers around it. Therefore, this case is similar to the second synthetic test. The initial model is built by generalized linear inversion (GLI) method with slight smoothing (Figure 3a). We use a same initial model to perform the two different inversion schemes. We totally perform 25 iterations. For the SA method, in the first 20 iterations, we extract 28 shots for every iteration, finally we perform 5 iterations full data inversion. The standard result is obtained by employing the full data in the entire inversion procedure. Figure 3b and Figure 3c show the inverted results obtained by these two different methods. We perform several tests to illustrate the relationship between the extraction percentage and time consuming (Figure 4). The yellow triangle represents the extraction percentage (5%, 28 shots) in this real data test, the time consuming is 119.409 s. While, for the standard method, the time cost is 471.409 s. In this real data application, if the extraction percentage is lower than 5%, although the computational time cost is lower, one cannot obtain a good approximated solution and therefore cannot ensure the result accuracy after the last five iterations of full data inversion. While, if it is higher than 5%, the result quality can be ensured but the time consuming will be higher.

Figure 3: Real data application. (a) The initial model. (b) The standard result. (c) The SA result.

Figure 4: The relationship between extraction percentage and time consuming. The yellow triangle means the extraction percentage (5%) we choose in this real data application.

Conclusions

We propose a new first-arrival traveltime tomography scheme that intends to reduce the computational cost. Through the first synthetic test, we verify that SAA and SA can be used to solve the traveltime tomography problem even some strict mathematical conditions cannot be satisfied perfectly in practical geophysics problems. We also give an empirical suggestion about how to extract the data for different acquisition geometry. Extracting data too roughly will decrease the convergence speed for the entire inversion procedure. With the import of the SA method, we decrease the computational time cost about 75% for the real case and also 95% computational memory in the first 20 iterations. Furthermore, although the ability of SAA for enhancing the efficiency is lower than SA, SAA is more stable and also it owns the ability of saving computational memory for large datasets. In the future, we will apply this scheme to 3D cases. Since the acquisition geometry is more complex in 3D cases, many conditions should be evaluated.

Acknowledgements

We appreciate the support from GeoTomo allowing us to use the software package and perform this research.
EDITED REFERENCES
Note: This reference list is a copyedited version of the reference list submitted by the author. Reference lists for the 2016 SEG Technical Program Expanded Abstracts have been copyedited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES