Global optimization based on quasi-Monte Carlo methods for seismic inversion

Juarez S. Azevedo (Federal University of Recôncavo, Brazil), Saulo P. Oliveira (Federal University of Paraná, Brazil), and Wilson M. Figueiró (Federal University of Bahia, Brazil)

Abstract

We propose a new methodology based on quasi-Monte Carlo method (QMC) to solve the nonlinear inverse problem of determining the seismic velocity field from traveltime data. Forward modeling is performed with ray tracing. We compare QMC with the standard implementation of the Monte Carlo method (MC) considering synthetic two-dimensional data.

Keywords: Global Optimization, Quasi-Monte Carlo Method, Ray Tracing, Traveltime, Seismic Inversion.

Introduction

The Monte Carlo method is a classical global optimization technique for highly non-linear inverse problems (Sen and Stoffa, 2013; Tarantola, 2005). We randomly generate a large number of realizations of the parameter models and select the realization with lowest misfit. However, this method is computationally very expensive since it requires many realizations in order to have an accurate solution. Another critical issue is that the realizations may not be perfectly equidistributed (not uniformly spread throughout the search space), so that the generation-acceptance/rejection associated with minimization of misfit requires a huge amount of computational work.

In order to overcome the computational cost problem and distribute more evenly the variables in the field, we propose the quasi-Monte Carlo method (Niederreiter, 1992). In this approach, the parameter realizations are constructed from low-discrepancy sequences rather than randomly chosen.

We validate this method with a two-dimensional seismic inversion problem. We seek the velocity field that minimizes the misfit of computed and observed traveltimes data in a source-receiver array over the surface. Forward modeling of traveltimes is performed with ray tracing and the velocity field is parameterized by two-dimensional, cubic polynomials. Herein, we focus on low-discrepancy sequences of Sobol type (Joe and Kou, 2008).

Quasi-Monte Carlo method

The QMC method is based on *low discrepancy sequences*. Discrepancy intuitively represents a measure of deviation from uniformity of a sequence of points on an unit cube $I = [0,1)^k$. As a consequence, realizations generated by this kind of sequence (low-discrepancy) are more uniform than random sequences making them well correlated. Formally, the discrepancy of a sequence can be defined as follows: let $S = \{x_1, x_2, \ldots, x_N\}$ be the first N terms of the sequence, and let J be an arbitrary subset of the unit cube I. If $T(J)$ is the number of terms in S that fall into J, i.e., the number of terms in $S \cap J$, and $V(J)$ is the volume of J, then the discrepancy $D(S)$ is given as

$$D(S) = \sup_{J} \left\| \frac{T(J)}{N} - V(J) \right\|. \qquad (1)$$

This means that the discrepancy is the largest difference between volume of subsets of the unit cube by the fraction of points in the subsets. The error of QMC for N realizations of dimension M decays with order $O(N^{-1}(\log N)^M)$, while in MC the error order is $O(N^{-1/2})$ (Fig. 1), for details see Niederreiter (1978).
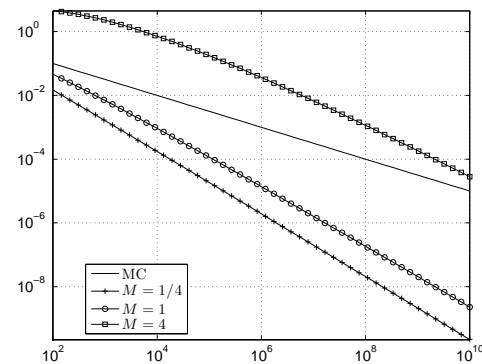

Figure 1: Order of error decay for QMC ($M = 1/4, 1, 4$) and MC methods.

Fig. 2 illustrates the histograms in two-dimensional media involving the Sobol sequence and the MC method for 2000 realizations. Note that the Sobol points are better equidistributed in a given volume than pseudo-random numbers (MC method).

Ray tracing

The theoretical basis of this section is given by Červený (2001) and Santos and Figueiró (2011). We assume that the vector position of a ray trajectory is determined by the parametric equations

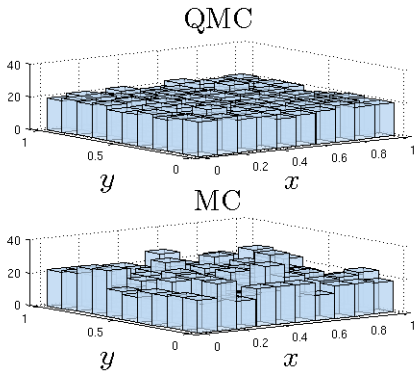$$\mathbf{x} = \mathbf{x}(\tau) = (x(\tau), z(\tau)), \qquad (2)$$

Figure 2: Data point histogram $2D$ with 2000 realizations from QMC (above) and MC (below) methods.

where τ represents a ray parameter. This trajectory is obtained by the following equations:

$$\begin{cases} \dfrac{d\mathbf{x}}{d\tau} &= \mathbf{p}(\tau) \\ \dfrac{d\mathbf{p}}{d\tau} &= \dfrac{1}{2}\vec{\nabla}\left(\dfrac{1}{V^2(x,z)}\right). \end{cases} \quad (3)$$

Here $\mathbf{p}(\tau)$ is the slowness vector that is tangent to the ray trajectory at τ and $V(x,z)$ is the velocity of the seismic wave at each point (x,z) of field. In order to obtain a numerical algorithm allowing the tracing of rays, we employ a first-order Taylor expansion of system (3):

$$\begin{cases} \mathbf{x}(\tau+\Delta\tau) &= \mathbf{x}(\tau)+\mathbf{p}(\tau)\Delta\tau \\ \mathbf{p}(\tau+\Delta\tau) &= \mathbf{p}(\tau)+\dfrac{1}{2}\vec{\nabla}\left(\dfrac{1}{V^2(x(\tau),z(\tau))}\right)\Delta\tau. \end{cases} \quad (4)$$

Moreover, travel time along the ray trajectory can be obtained by expression

$$\frac{dT}{d\tau} = \frac{1}{V^2}, \quad (5)$$

which is calculated numerically by the following equation:

$$T(x_{N+1},z_{N+1}) = \sum_{i=0}^{N}\Delta T_i = \sum_{i=0}^{N}\frac{1}{V_i}\|\mathbf{x}_{i+1}-\mathbf{x}_i\|_2, \quad (6)$$

where V_i represents the wave velocity at $\mathbf{x}_i = (x_i,z_i)$. To solve Eqs. (4)-(6) for the travel time, initial conditions should be established. We consider

$$\mathbf{p}(\mathbf{0}) = \frac{1}{V(0,0)}(cos(\theta),sin(\theta)),$$

for θ evenly distributed in $[0,\pi]$. Once this has been done, at the end of each step in the process of ray trajectory construction, the slowness vector $\mathbf{p}$ must be updated to satisfy the eikonal equation:

$$\|\mathbf{p}\|_2^2 = p_1^2 + p_2^2 = \frac{1}{V^2(x,z)}, \quad (7)$$

where $\mathbf{p} = (p_1,p_2)$ with $p_1 = \|\mathbf{p}\|_2 cos(\theta)$, $p_2 = \|\mathbf{p}\|_2 sin(\theta)$ and θ is the angle between $\mathbf{p}(\mathbf{0})$ and the positive orientation of the x-axis.

Inversion algorithm

In order to introduce the global optimization method based on quasi-Monte Carlo realizations, we firstly consider a model vector $\mathbf{m}$ written as

$$\mathbf{m} = (m_1,m_2,m_3,\ldots,m_L)^T, \quad (8)$$

and a vector $\mathbf{d}^{obs} = (d_1,d_2,d_3,\ldots,d_K)^T$, associated with the following expression

$$e(\mathbf{m}) = \mathbf{d}^{obs} - \mathbf{g}(\mathbf{m}) = \mathbf{d}^{obs} - \mathbf{d}^{cal}.$$

Vectors $\mathbf{d}^{obs}$, $\mathbf{d}^{cal}$ are, respectively, the observed data and the calculated data for the model $\mathbf{m}$, and $\mathbf{g}$ is the forward modeling operator. The global optimization method developed in Sen and Stoffa (2013) is a procedure to get the global minimum of the objective function defined by

$$\phi(\mathbf{m}) = e(\mathbf{m})^T e(\mathbf{m}) = (\mathbf{d}^{obs} - \mathbf{g}(\mathbf{m}))^T(\mathbf{d}^{obs} - g(\mathbf{m})). \quad (9)$$

We present below a brief summary of the algorithm:

1. Choose a number of parameter realizations N and generate a set S with N samples;

2. Search for a new $\mathbf{m}_P \in S$ using the current model;

3. Evaluate the objective function at this new model, and update it;

4. Repeat 2, 3 until it reaches an optimal value for the error or maximum number of realizations.

For further improvement of the parameters, the model provided by the global optimization algorithm serves as the initial guess of a Gauss-Newton local search.

Numerical Experiments

In this section, we consider a two-dimensional velocity field, heterogeneous and isotropic, with horizontal length of 32.0 km and depth of 4.0 km, whose velocities are between 1.0 km/s and 6.0 km/s. The source-receiver array has 5 sources and 33 receivers, both equally distributed on observation surface ($z=0$).

Fig. 3 illustrates this target velocity field (in color scale) and the seismic rays that arrived first at the receivers (x_r) from each source (x_s). The travel times $t^{obs}(m,n)$ of the rays from the k-th source that arrive first at the l-th receiver, which are illustrated in Fig. 4, will form the vector of observed data $\mathbf{d}^{obs}$ for the study of inverse modeling. Note that the total number of observations is the number of sources times the number of receivers, i.e., $K = 5 \times 33 = 165$.

Regarding the QMC/MC inversion algorithm, each component m_i of the model vector $\mathbf{m}$ represents the velocity at the i-th block of an 8×4 grid of the medium (thus the number of parameters is $L = 8 \times 4 = 32$). In order to perform the ray tracing in the forward step, we need the velocity field to be smooth. For this purpose, we find the cubic polynomial

$$\begin{aligned} V(x,z) = &C_{00}+C_{10}x+C_{01}z+C_{20}x^2+C_{11}xz \\ &+C_{02}z^2+C_{30}x^3+C_{21}x^2z+C_{12}xz^2+C_{03}z^3, \end{aligned} \quad (10)$$
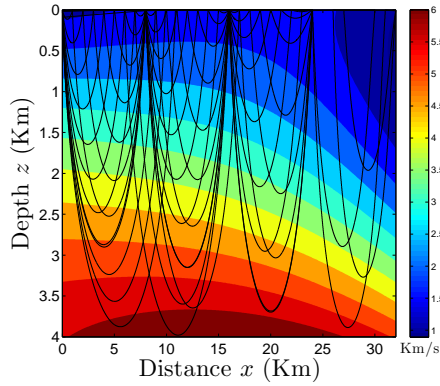
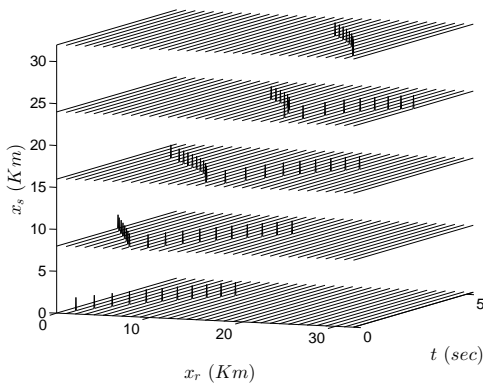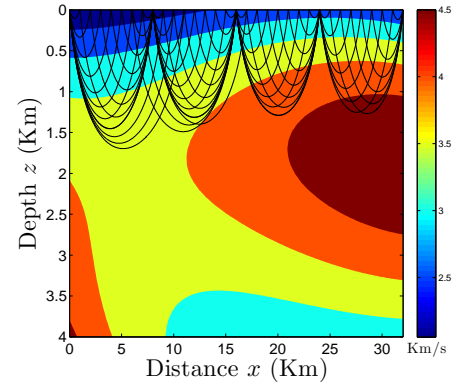Figure 3: Ray tracing on the target model.

(a) MC

Figure 4: Records of the traveltimes of first arrivals for the target model. The receivers and arrival times are indicated by the x- and y-axis, respectively. Each plane in the z-axis represents a single source.

(b) QMC

Figure 5: Ray tracing on inverted model, obtained by (a) MC and (b) QMC methods considering $N = 5000$ realizations.

that best fits the block model in the least-squares sense. This step provides the travel times $t^{cal}(m,n)$ of the rays from the m-th source that arrive first at the n-th receiver, which are the components of the observed data vector $\mathbf{d}^{obs}$.
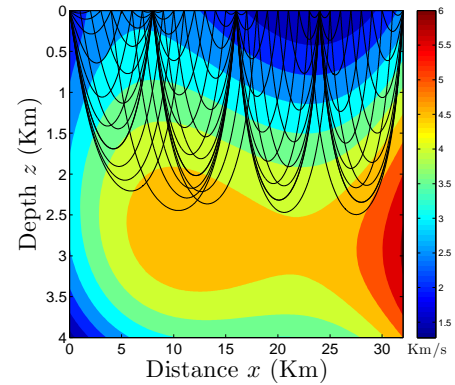
Figs. 5(a)-5(b) show the inversion results obtained from MC and QMC after 5000 realizations. Observe that there is a discrepancy in the model of velocity field on MC, making difficult their geological interpretation. However, such a discrepancy is milder in the model generated by QMC, so that the geological characterization of the velocity field becomes more easily understood. Besides, the seismic rays in the QMC model reach deeper region of the model than in the MC method. Nevertheless, we note bad estimation of the velocity for deeper regions in the QMC method, that can be explained by weak coverage of the rays of such regions in the foward modeling.

Fig. 6 shows the misfit decay of MC and QMC methods. Here, the misfit is defined by root mean square (RMS):

$$RMS = \sqrt{\frac{\sum\limits_{m,n}|t^{obs}(m,n) - t^{cal}(m,n)|^2}{\sum\limits_{m,n}|t^{obs}(m,n)|^2}} \times 100\%. \qquad (11)$$

The convergence patterns of QMC and MC were very similar, but we point out the rapid misfit decrease of QMC in the first 100 iterations.

Figs. 7(a)-7(b) show the arrival times obtained from the inverted models. The best results were obtained at the first half of the receiver array. Note that the lateral variation of the target model (and also the inverted models) is lower below these receivers.

Conclusions

In this paper, we use ray tracing in inverse problems associated with parameterization of the velocity field admitting synthetic data sets. We have applied the global optimization method in light of the quasi-Monte Carlo methods to solve this problem, using Sobol sequences to perform the inversion of traveltimes in synthetic data sets and compare their results with MC.

Although we did not obtain inverted models identical to their target models, the methodology employed on QMC demonstrated excellent ability to highlight the main features of the velocity field, representing the geophysical model well, although that are artifacts that may be linked to the generation of rays that are not sufficiently deep.

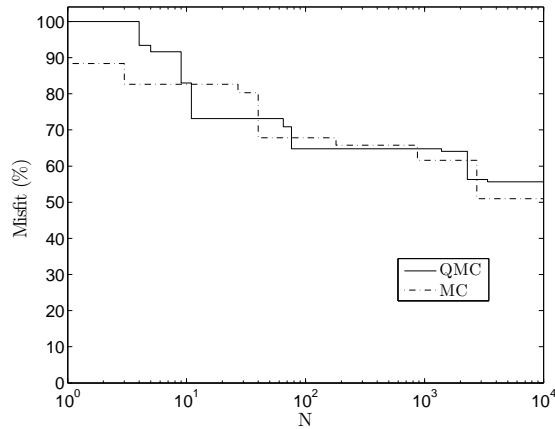It is also important to note that CPU time required to

Figure 6: Misfit between synthetic and computed data for MC and QMC methods.

perform of the computational process in both methods had a similar performance so we have not done a detailed study on this subject. Regarding misfit, the quasi-Monte Carlo method presented a fast decay in the first 100 iterations and eventually delivers a lower misfit Monte Carlo between 100 and 1000 realizations. From this point on, the convergence becomes extremely slow as envisaged in the literature (see Azevedo and Oliveira (2012); Graham et al. (2011)).

For future work, we intend to build quasi-random sequences from weighted averages of low-discrepancy sequences and pseudo-random numbers, with a weighting factor $0 \leq q \leq 1$. In particular, $q = 0$ leads to a completely free MC method, whereas $q = 1$ yields a QMC method.

References

Azevedo, J. S., and S. P. Oliveira, 2012, A numerical comparison between quasi-Monte Carlo and sparse grid stochastic collocation methods: Commun. Comput. Phys., **12**, 1051–1069.

Červený, V., 2001, Seismic Ray Theory: Cambridge University Press, Cambridge, UK.

Graham, I. G., F. Y. Kuo, D. Nuyens, R. Scheichl, and I. H. Sloan, 2011, Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications: J. Comput. Phys., **230**, 3668–3694.

Joe, S., and F. Y. Kou, 2008, Construction of Sobol' sequence with better two-dimensional projections: SIAM J. Sci. Comput., **30**, 2635–2654.

Niederreiter, H., 1978, Quasi-Monte Carlo methods and pseudo-random numbers: Bull. Amer. Math. Soc., **84**, 957–1041.

——, 1992, Random number generation and quasi-Monte Carlo methods: SIAM.

Santos, V. G. B., and W. M. Figueiró, 2011, Seismic ray tomography using L1 integral norm: Revista Brasileira de Geofísica, **29**, 347–358.

Sen, M. K., and P. L. Stoffa, 2013, Global Optimization Methods in Geophysical Inversion: Cambridge University Press.

Tarantola, A., 2005, Inverse Problem Theory and Methods for Model Parameter Estimation: SIAM.
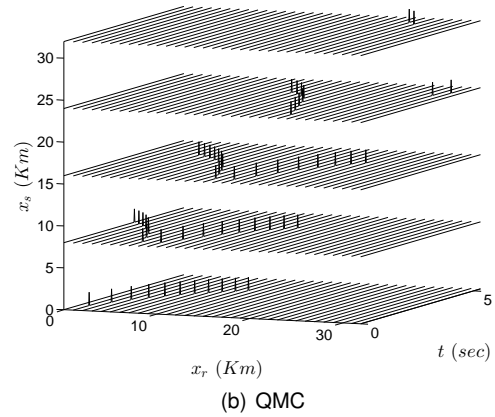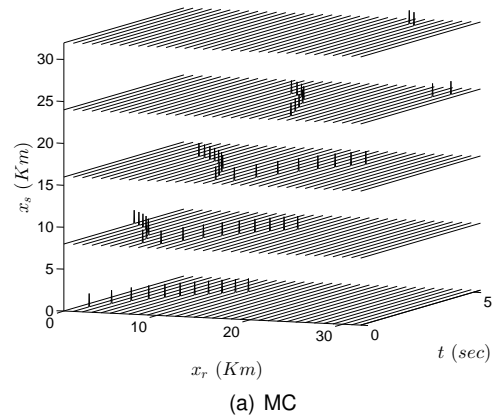
(a) MC

(b) QMC

Figure 7: Travel times from source to receiver obtained by QMC and MC methods considering $N = 5000$ realizations. The receivers and arrival times are indicated by the *x*- and *y*-axis, respectively. Each plane in the *z*-axis represents a single source.

Acknowledgments