Efficient monochromatic traveltome sensitivity kernel calculation using random-boundary condition

Introduction

For wave-equation based inversion methods, such as full waveform inversion (FWI, Tarantola, 1984) and wave-equation traveltome tomography (Luo and Schuster, 1991), the functional gradient, which is commonly derived using the adjoint state method, can be efficiently calculated by the integration of an incident wavefield and an adjoint wavefield. However, the incident and adjoint wavefield must be accessed simultaneously to calculate the gradient. In reverse time migration (RTM), serval solutions have been proposed, such as the optimal checkpointing method (Symes, 2007), the random boundary condition (Clapp, 2009), the source wavefield reconstruction (Dussaud et al., 2008; Feng et al., 2011).

In this paper, we are focused on the efficient calculation of traveltome sensitivity kernel (TSK, Feng et al., 2019). Computationally, we should balance the computational cost and the memory size for 3-D problems. Unlike FWI or RTM, the goal of traveltome inversion is to update the large-scale velocity perturbation using traveltome information, which means that only the low wavenumber components of the traveltome functional gradient are necessary. Therefore, we can combine the advantages of the random boundary condition and the monochromatic sensitivity kernel, the develop an efficient gradient calculation strategy for industrial-scale 3-D traveltome inversion. Next, we will demonstrate in detail how to accelerate the gradient calculation and reduce memory occupancy.

Method

The traveltome sensitivity kernel or traveltome derivative (Luo and Schuster, 1991; Feng et al., 2019) can be calculated by the integration of an incident wavefield and a traveltome adjoint wavefield:

\[ K(x; x_i, x_s) = M(x) \int_0^{\tau_{max}} u(x_i, t; x_s) \lambda(x_i, t; x_s) dt \]

where \( u(x_i, t; x_s) \) and \( \lambda(x_i, t; x_s) \) is respectively the incident wavefield and the traveltome adjoint wavefield, \( M(x) \) is a space-domain function depending on the model parameterization.

Apparently, the expression of the sensitivity kernel looks like the RTM image or FWI gradient. However, the most distinguished difference between traveltome sensitivity kernel and FWI gradient is the wavenumber components: the FWI gradient updates the high wavenumber model perturbation, while the traveltome gradient updates the low wavenumber model along the wave-path.

Typically, the source wavefield reconstruction strategy (Table 1) can be used to calculate of the sensitivity kernel, which includes one forward wave propagation and two backward wave propagation. For large-scale 3-D problems, the size of wavefield at boundaries would be around 10GB ~ 100GB, which is still quite challenging for modern HPCs. Using the Parseval theorem, the sensitivity kernel can be transformed into frequency domain integration, in the form

\[ K(x; x_i, x_s) = 2M(x) \text{Re}\left\{ \int_0^{f_{max}} (u(x, f; x_s))^* \lambda(x, f; x_s) df \right\} \]

where the upper script * stands for taking the conjugate, \( u(x, f; x_s) \) and \( \lambda(x, f; x_s) \) are the frequency-domain incident- and adjoint wavefields.

Mathematically, equation (1) and (2) is equivalent. For traveltome inversion, traveltome misfits are back- propagated along the wave path defined by the band-limited wave propagator. In FWI, we can use the frequency-domain multi-scale inversion strategy starting from low frequency. Similarly, we can also back-propagate the traveltome misfits along with the monochromatic traveltome sensitivity kernel (MTSK) instead of the band-limited kernel (also referred to as the finite-frequency sensitivity kernel). Based on (2), the MTSK is defined as:

\[ K(x; x_i, x_s, f) = 2M(x) \text{Re}\left\{ (u(x, f; x_s))^* \lambda(x, f; x_s) \right\} \]
Recall that the fundamental idea of random boundary condition is to generate incoherently scattered wavefield from the random boundary region. We can apply random boundary condition to the propagation of the incident- and adjoint wavefield and calculate the frequency-domain wavefield on the fly. Therefore, source wavefield reconstruction is eliminated because we only need the two monochromatic wavefields. The new algorithm for MTSK calculation is shown in Table 2.

Compared with source wavefield reconstruction (Table 1), the proposed method (Table 2) has the following advantages:

1. Only two wavefield propagations are needed, one for frequency-domain incident wavefield calculation and the other for adjoint wavefield;
2. No absorbing condition is necessary, which further decreases the computational complexity and cost.

### Table 1 Source wavefield reconstruction strategy for TSK / FWI gradient calculation.

```plaintext
for t=0 to t=tmax do
    forward propagate source wavefield
    save source wavefield at boundaries
end for
```

```plaintext
for t=tmax to t=0 do
    backward propagate source wavefield
    backward propagate adjoint wavefield
    calculate the gradient
end for
```

### Table 2 Random boundary condition for MTSK calculation.

```plaintext
for t=0 to t=tmax do
    forward propagate source wavefield
    calculate the discrete Fourier integration of source wavefield for a given frequency f:
    \[ u(x, f; x) = \sum_{k=1}^{Nt} u(x, k\Delta t; x) \exp(-i2\pi f k t) \]
end for
```

```plaintext
for t=tmax to t=0 do
    backward propagate adjoint wavefield
    calculate the discrete Fourier integration of adjoint wavefield for a given frequency f:
    \[ \lambda(x, f; x) = \sum_{k=1}^{Nt} \lambda(x, k\Delta t; x) \exp(-i2\pi f k t) \]
end for
```

Calculate the final MTSK: \( K(x; x, f) = 2M(x) \text{Re}\left( \left( u(x, f; x) \right)^* \lambda(x, f; x) \right) \)

### Examples

To demonstrate the computational advantages of the proposed method, we calculate the TSK and MTSK in a homogeneous model (Figure 1a). The size of the velocity model is 401 by 401, and the grid interval is 10 m. The source and receiver coordinate is (sx, sy, sz) = (1000 m, 2000 m, 10 m) and (gx, gy, gz) = (3000 m, 2000 m, 10 m). We use a rigid boundary condition (i.e. \( u(t, z=0) = 0 \)) on the top boundary and the random boundary condition for the rest five boundaries. The source function is a Ricker wavelet with 10 hz dominant frequency. The forward stepping is 1 ms, and the record length is 1 s.

We test our method on a workstation with 2 Intel Xeon E5-2683 V3 CPU (14 cores) and 256 GB DDR4 memory. Figure 1 b-d show the band-limited TSK calculated by equation (1), (2) and the monochromatic TSK (equation (3)) using random boundary condition. Obviously, the introduction of random boundary condition will not cause coherent noise that contaminates the wave-path. As theoretically predicted, the shape of the monochromatic traveltime sensitivity kernel (Figure 1d) is spatially oscillating outside the first-Fresnel zone, while only the first-Fresnel zone is visible in the band-limited TSK (Figure 1b and 1c).

Although the single MTSK is highly oscillating, the traveltime functional gradient will still be dominated by the low-wavenumber components because the spatial oscillations will be cancelled due to the multi-shot, multi-receiver seismic acquisition system. To validate our judgment, we also calculate the traveltime functional gradient in a transmission experiment (100 shots and 100 receivers evenly distributed along the four edges of the 2-D velocity model). The gradient calculated by the conventional method (finite-frequency gradient, equation (1)) and the monochromatic kernel (equation (3)) is shown in Figure 2a and 2b respectively. We can see that although the magnitude is different, the shape has no vision disparity. Figure 2c and 2d show the single-shot gradient of Figure 2a and 2b. From Figure 2d
we notice that the low wavenumber component looks like the finite-frequency gradient (Figure 2c), making it practical to use the monochromatic gradient for travelt ime inversion.

To quantitatively compare the memory saving of the proposed method over the source wavefield re-computation strategy, we list the memory requirements of different methods in Table 3, demonstrating the significant memory-saving advantage of the proposed method.

Table 3 Comparison of memory requirement of different implementations (N is the model grid number (assume the model is cubic), Ns is record length, l is spatial-order for the finite-difference algorithm, Npml is additional memory cost introduced by the PML absorbing boundary).

<table>
<thead>
<tr>
<th>Method</th>
<th>Memory size ($N=512$, $Ns=4000$, $l=5$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equation (1)</td>
<td>$(N^2<em>6</em>Ns<em>l + N^2</em>8+N_{pml})*4$Bytes $&gt;$ 121.2GB</td>
</tr>
<tr>
<td>Equation (3)</td>
<td>$(N^3*8)<em>4$Bytes + $(2</em> N^3)*8$Bytes = 6.0GB</td>
</tr>
</tbody>
</table>

Figure 1 (a): the velocity model with random boundaries; (b) and (c) are the band-limited TSK calculated by equation (1) and (2), respectively; (d) is the MTSK calculated by the proposed method.

Discussion and Conclusion

We propose an efficient strategy for calculating the monochromatic traveltime sensitivity kernel using time-domain propagator and discrete Fourier integration, where the random boundary condition is implemented for wave-equation simulation.

Compared with the source wavefield re-computation strategy, the computational cost of the proposed method is less than 2/3 of the re-computation strategy because only two wave-equation propagation steps are included. Moreover, the computational complexity and cost of wave-equation propagation with random boundary condition is less than propagation with an absorbing boundary condition (e.g., the PML boundary condition).

Besides the savings of computational cost, the reduced memory requirement is much more attractive. We have demonstrated that for industry-scale problems, the major memory occupancy is the storage of the wavefield at the six boundaries, whose size can be more than 100GB. On the other hand, if we calculate the monochromatic gradient, only two complex-valued 3-D arrays storing the frequency-
domain wavefields are needed, which eliminates the need to store the boundary wavefields. Therefore, the proposed method is very promising for 3-D wave-equation traveltime inversion. The application of traveltime inversion using a monochromatic gradient will be submitted to a companion abstract.

![Figure 2 Comparison of the traveltime functional gradient calculated by band-limited TSK (a) and monochromatic TSK (b, f=15Hz); (c) and (d) are corresponding single-shot gradient.](image)

**Figure 2** Comparison of the traveltime functional gradient calculated by band-limited TSK (a) and monochromatic TSK (b, f=15Hz); (c) and (d) are corresponding single-shot gradient.

**Acknowledgments**

The authors appreciate the helpful discussion with Dr. Liu Yuzhu for monochromatic traveltime inversion, and we are grateful to our sponsors for their financial support. WPI’s research is also financially supported by the National Natural Science Foundation of China (42074143, 41774126), the National Key R&D Program of China (2018YFA0702503, 2019YFC0312004), the National Science and Technology Major Project (2016ZX05024-001 and 2016ZX05006-002), Southern Marine Science and Engineering Guangdong Laboratory (Zhanjiang)(ZJW-2019-04), and the SINOPEC Key Laboratory of Geophysics (33550006-19-FW0399-0041).

**References**


