Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.OA-2022-0111

Vol. **16**, No. 2, pp. 277-297
May 2023

On Quadratic Wasserstein Metric with Squaring Scaling for Seismic Velocity Inversion

Zhengyang Li[1], Yijia Tang[2], Jing Chen[3,*] and Hao Wu[1]

[1] *Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China*
[2] *School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai 200240, China*
[3] *Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 639798*

Received 14 July 2022; Accepted (in revised version) 30 January 2023

Abstract. The quadratic Wasserstein metric has shown its power in comparing probability densities. It is successfully applied in waveform inversion by generating objective functions robust to cycle skipping and insensitive to data noise. As an alternative approach that converts seismic signals to probability densities, the squaring scaling method has good convexity and thus is worth exploring. In this work, we apply the quadratic Wasserstein metric with squaring scaling to regional seismic tomography. However, there may be interference between different seismic phases in a broad time window. The squaring scaling distorts the signal by magnifying the unbalance of the mass of different seismic phases and also breaks the linear superposition property. As a result, illegal mass transportation between different seismic phases will occur when comparing signals using the quadratic Wasserstein metric. Furthermore, it gives inaccurate Fréchet derivative, which in turn affects the inversion results. By combining the prior seismic knowledge of clear seismic phase separation and carefully designing the normalization method, we overcome the above problems. Therefore, we develop a robust and efficient inversion method based on optimal transport theory to reveal subsurface velocity structures. Several numerical experiments are conducted to verify our method.

AMS subject classifications: 49N45, 65K10, 86-08, 86A15

Key words: Optimal transport, Wasserstein metric, waveform inversion, seismic velocity inversion, squaring scaling.

---

*Corresponding author. *Email addresses:* `hwu@tsinghua.edu.cn` (H. Wu), `yijia_tang@sjtu.edu.cn` (Y.J. Tang), `jing.chen@ntu.edu.sg` (J. Chen), `lizhengy18@mails.tsinghua.edu.cn` (Z.Y. Li)

1. Introduction

Seismic waveform inversion has been receiving wide attention in past decades [4, 14, 29, 32, 38, 42, 46, 47] due to its high-resolution imaging capability. The major goal is to find optimal model parameters that minimize the discrepancy between synthetic and observed seismic signals. In mathematics, it can be formulated as a partial differential equation (PDE) constrained optimization problem, which consists of two key ingredients [41]: the forward modeling of wave propagation and updating model parameters. In previous decades, limited by computing power, most tomography methods simulated wave propagation based on the ray theory. The high frequency assumption ignores finite frequency phenomena such as wave-front healing and scattering [17], and thus, leads to low-resolution inversion results [32]. With the rapid development of computing power and the forward modeling method, more accurate synthetic signals could be computed by numerically solving wave equations. It enables us to obtain high-resolution subsurface velocity structures from the inversion of waveform data, which could provide guidance information for seismic hazard assessment [37] and exploration geophysics [41].

The discrepancy between synthetic and observed seismic waveforms is usually measured using the L^2 metric [32, 37, 38, 41]. However, it suffers from the well-known cycle skipping problem [41] so that the solution may be trapped in local minima during the iteration, leading to incorrect inversion results. To overcome the problem, many methods have been proposed to modify the objective function, e.g., the envelope objective function [5], the cross-correlation-based objective function [23], and the deconvolution-based objective function [22]. In addition, the quadratic Wasserstein (W_2) metric from the Optimal Transport (OT) theory [36, 39, 40] has received wide attention in recent years due to its nice properties and has been applied to many seismic inverse problems such as earthquake location and seismic tomography [7, 10–13, 46, 47, 49]. This metric measures the difference between two probability distributions by minimizing the transport cost from one distribution to the other, which is insensitive to data noise and preserves the convexity regarding data shift, dilation, and partial amplitude change [10, 11, 14]. Thanks to these advantages, the reconstruction of velocity models can succeed even if the initial model is far from the real model [14, 27, 28].

However, it is not straightforward to apply the quadratic Wasserstein metric in the seismic waveform inversion [13]. The main reason is that the seismic signals are signed. Thus, a key ingredient in the application of the quadratic Wasserstein metric to seismic waveform inversion is converting seismic waveforms to probability distributions. Various scaling techniques are developed to deal with this problem, e.g., linear scaling [47], squaring scaling [7], exponential scaling [33], and graph-space transform [26]. Moreover, there are also some other metrics based on the OT theory that have been applied to seismic inverse problems, e.g., the Wasserstein-Fisher-Rao metric and the Kantorovich-Rubinstein norm, which relax the mass conservation constraint [27, 28, 49]. Though the quadratic Wasserstein metric has been applied widely