Full waveform inversion with unbalanced optimal transport distance

Da Li, Michael P. Lamoureux and Wenyuan Liao

da.li1@ucalgary.ca

Introduction

In this work we introduce the unbalanced optimal transport (UOT) distance with Kullback–Leibler divergence to full waveform inversion problem to mitigate the cycle-skipping issue and reduce the nonlinearity during the optimization. An entropy regularization and a scaling algorithm have been used to compute the distance and its gradient efficiently. Two normalization methods which transform the seismic signals into non-negative functions have been compared and the gradient of objective function has been derived through adjoint state method. Three numerical examples in one and two dimension are provided.

Optimal transport problem

Let $X = Y = \{x_1, x_2, \cdots, x_N\} \subset \mathbb{R}^d$, $\mu = \sum_i f_i \delta_{x_i}$, $\nu = \sum_i g_i \delta_{x_i}$. Define cost matrix C as $C_{i,j} = |x_i - x_j|^2$. The optimal transport problem between μ and ν in discrete form is:

$$\min_{T \in \mathbb{R}^{N\times N}} \langle T, C \rangle = \sum_{i,j=1}^{N} T_{i,j} C_{i,j}, \text{ s.t. } T\mathbf{1}_N = f,\ T^T \mathbf{1}_N = g. \tag{1}$$

Regularized unbalanced optimal transport

Given cost matrix C, regularization coefficients ε and ε_m, the regularized unbalanced optimal transport distance between $f, g \in \mathbb{R}_+^N$ is:

$$\begin{aligned} W^2_{2,\varepsilon,\varepsilon_m}(f,g) = \min_{T\in\mathbb{R}^{N\times N}} & \langle T, C\rangle - \varepsilon E(T) \\ & + \varepsilon_m KL(T\mathbf{1}_N | f) + \varepsilon_m KL(T^T\mathbf{1}_N | g) \end{aligned} \tag{2}$$

- Entropy regularization term $E(T)$ is to increase the computational efficiency.
- The $KL(\cdot|\cdot)$ is Kullback-Leibler divergence as a mass balancing term.

Iterative scaling algorithm

In order to solve UOT distance and its gradient, a scaling algorithm is used here. For problem in equation (2), given matrix K with $K_{i,j} = e^{-C_{i,j}/\varepsilon}$. Starting with an initial value $v^{(0)} = \mathbf{1}_N$, compute iteratively with:

$$u_i^{(n+1)} = (f_i / \sum_j K_{i,j} v_j^{(n)})^{\varepsilon_m/(\varepsilon_m+\varepsilon)},$$

$$v_j^{(n+1)} = (g_j / \sum_i K_{i,j} u_i^{(n+1)})^{\varepsilon_m/(\varepsilon_m+\varepsilon)}.$$

Then, $T^*_{i,j} = u_i^* K_{i,j} v_j^*$. The gradient of (2) is:

$$\nabla_{f_i} W^2_{2,\varepsilon,\varepsilon_m}(f,g) = -\varepsilon_m \left(e^{-\phi_i^*/\varepsilon_m} - 1 \right),$$

where $\phi_i = \varepsilon \log u_i$.

Normalizations with linear and exponential transform are given by:

$$h_{\text{linear},k}(f) = f + k,$$

$$h_{\exp,k}(f) = e^{kf}.$$

Shifted Ricker example

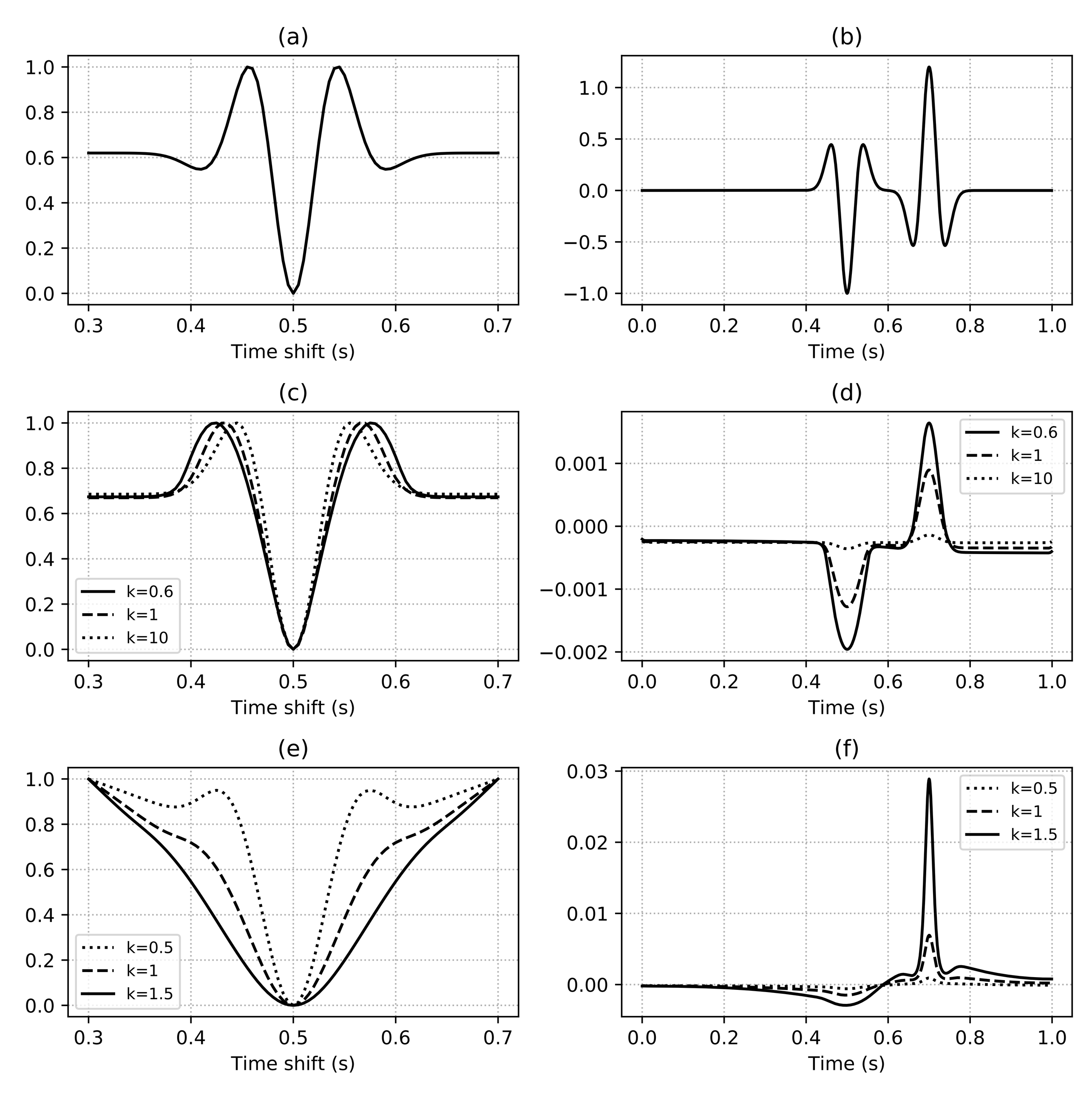

Figure: (a), (b): misfit function and adjoint source using L_2 distance. (c), (d): misfit functions and adjoint sources using UOT distance with linear normalization. (e), (f): misfit functions and adjoint sources using UOT distance with exponential normalization.

Single layer model

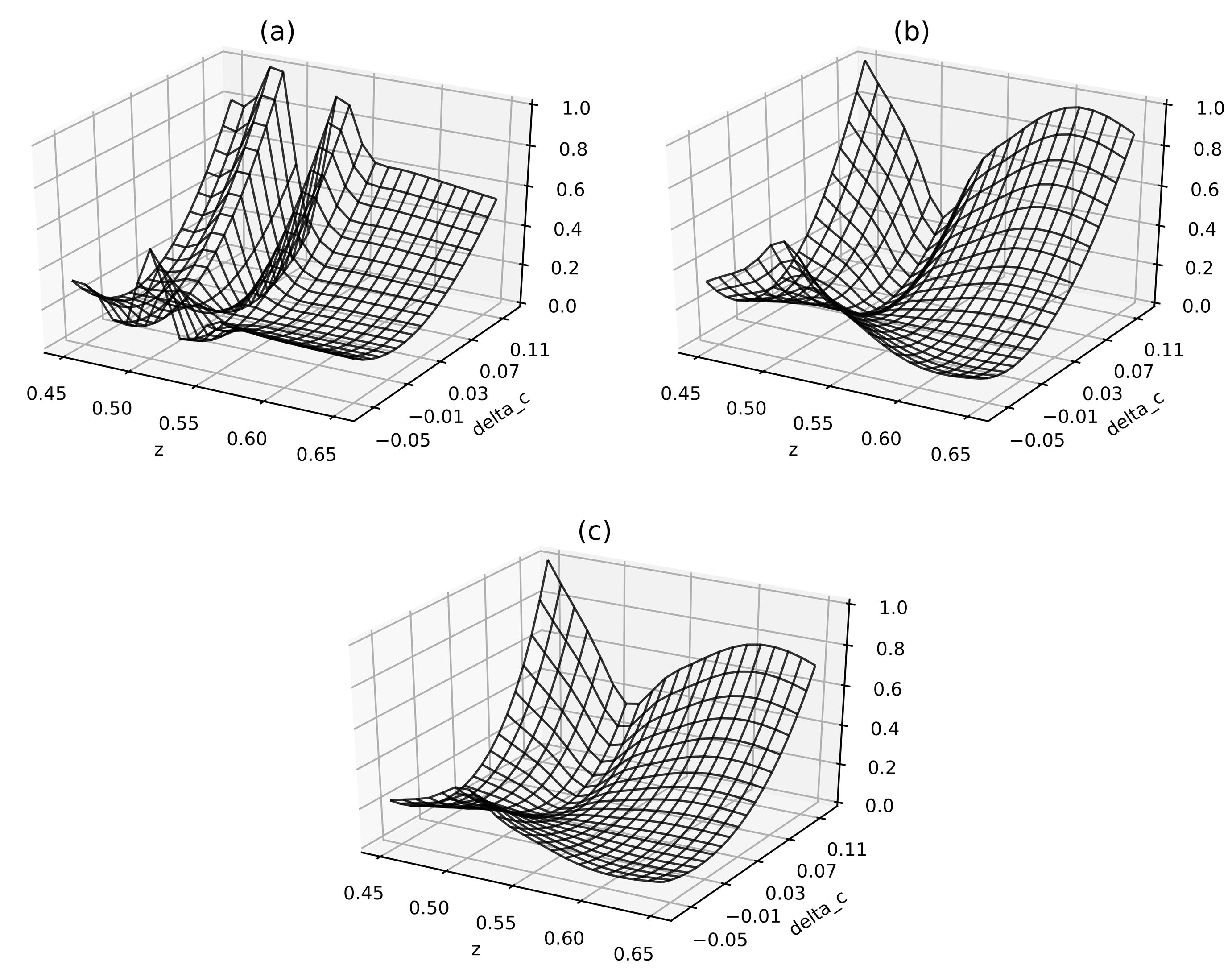

Figure: Set the velocity model $c(\delta c, z) = c_0(x, z) + \delta c H(z)$ in a region with 1km wide and 1km deep. We use 51 receivers at the top of the region, and a source with 10 Hz Ricker wavelet located at $x = 0.5$km, $z = 0.05$km. (a), (b), (c): misfit function by using L_2 distance, UOT distance with linear and exponential normalization respectively.

2D crosshole model

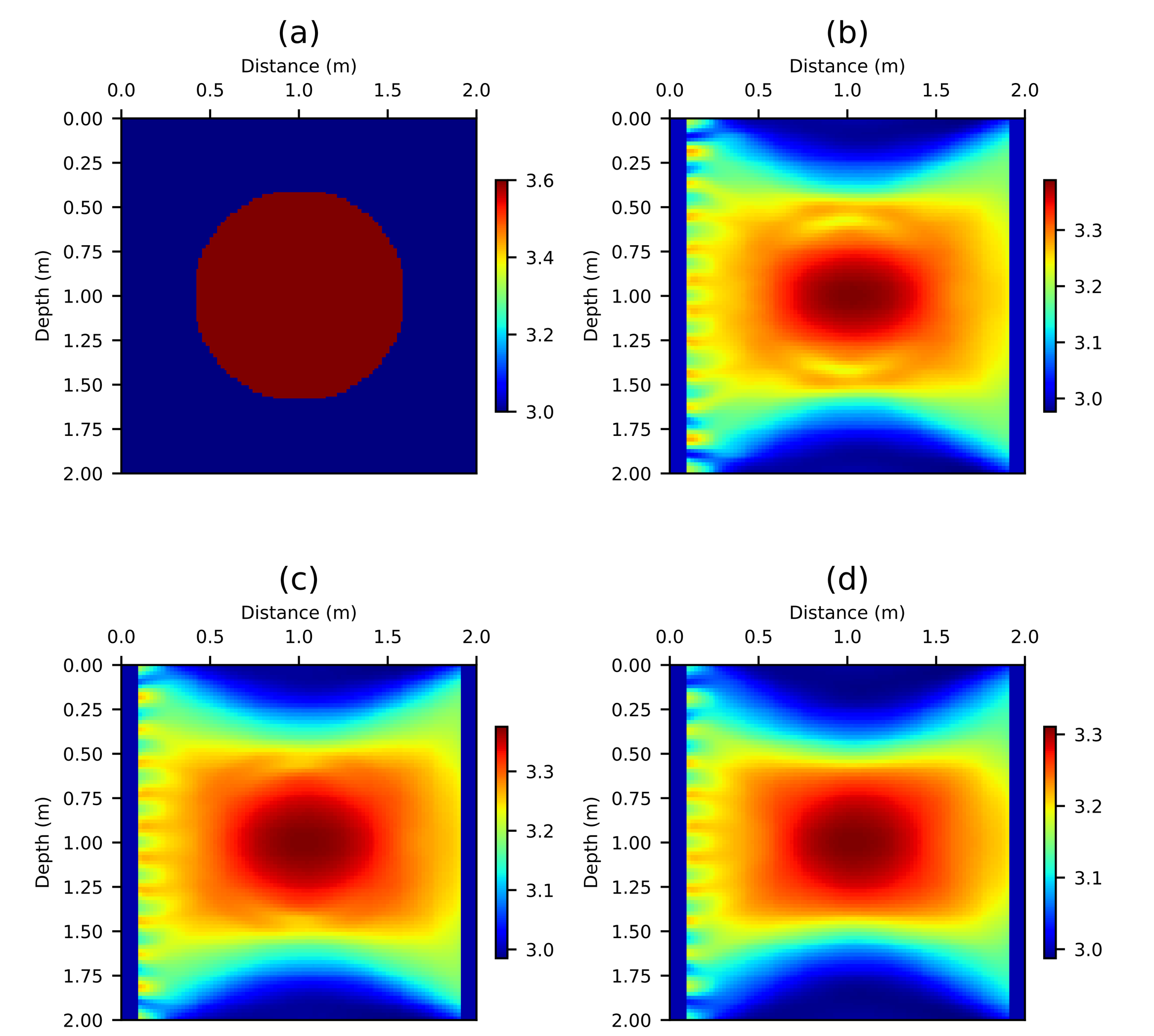

Figure: (a): the true velocity model. There are 11 sources with 10 Hz Ricker wavelet which are equally spaced on the left side and 101 receivers on the right side. (b), (c), (d): inverse results of gradient descent after 5 iterations with L_2 distance, UOT distance with linear and exponential normalization respectively.

Main reference

Chizat, Lenaic and Peyré, Gabriel and Schmitzer, Bernhard and Vialard, François-Xavier, (2018) *Scaling algorithms for unbalanced optimal transport problems*. Mathematics of Computation, 2018, 87(314): 2563-2609.

NSERC CRSNG

UNIVERSITY OF CALGARY
FACULTY OF SCIENCE
Department of Geoscience