A graph-space approach to optimal transport for full waveform inversion

L. Métivier[1,2], A. Allain[1], R. Brossier[2], Q. Mérigot[3], E. Oudet[1], J. Virieux[2]
[1]*CNRS, Univ. Grenoble Alpes, LJK, F-38058 Grenoble, France*
[2]*Univ. Grenoble Alpes, ISTerre, F-38058 Grenoble, France*
[3]*Univ. Paris Sud, LMO, F-91405 Orsay, France*

SUMMARY

Optimal transport distances have been recently proposed to design more convex misfit functions in the frame of full waveform inversion. The main difficulty for this strategy is the non-positivity of seismic data: the optimal transport is developed for the comparison of positive quantities. We review the proposed strategies to overcome this difficulty and put in light their limitations: in particular, those applicable to field data lose the convexity property. On this basis, we propose an alternative approach, consisting in measuring the optimal transport distance between the discrete graph of seismic traces. This makes possible to apply optimal transport to real seismic data, while preserving the convexity. Two synthetic case studies illustrate the interest of the method. The main drawback is related to the computational cost of the current implementation: more powerful techniques based on this idea will soon be investigated.

INTRODUCTION

Optimal transport (OT) based misfit functions have recently been proposed for full waveform inversion (FWI), a high resolution seismic imaging method based on the minimization of the discrepancy between synthetic and observed data. FWI is now recognized as an important tool for exploration, for its ability to provide high resolution velocity models (see Plessix and Perkins, 2010; Warner et al., 2013; Vigh et al., 2014; Operto et al., 2015, for instance). However, its strong dependency to the accuracy of the initial velocity model often requires careful data processing for successful applications to field data. This dependency is related to the local optimization process on which it relies together with the non-convexity of the conventional least-squares (L^2) misfit function used to measure the distance between observed and synthetic data.

When used to compare positive functions, the OT distance has the interesting property to be convex with respect to shifted patterns: the OT distance between two shifted Gaussian functions increases as the shift between the two Gaussian increases. This convexity with respect to shifted patterns could be used, in the frame of seismic data, as a proxy for the convexity with respect to velocity parameters, as intermediate to large scale perturbations of these parameters mainly influence the kinematic of the wave propagation (Jannane et al., 1989), resulting in shifting in time the seismic signal. This was the main motivation for introducing OT in the frame of FWI (Engquist and Froese, 2014).

However the application of OT to seismic data is not straightforward. The OT theory has been developed for the comparison of positive functions. The seismic data being oscillatory with positive and negative amplitudes, the positivity assumption is breaking down. The propositions made so far to apply OT to seismic data can be divided into two categories. The first considers prior processing of the data to bring back to the comparison of positive quantities (Engquist and Froese, 2014; Engquist et al., 2016; Qiu et al., 2017). The second is based on a specific OT distance, namely the 1-Wasserstein distance, which naturally extends to the comparison of non-positive functions (Métivier et al., 2016a,b,c).

In this study, we provide an overview and a comparison of these strategies. The first conclusion is that only for the 1-Wasserstein distance, a fast algorithm allowing to compare the shot-gathers as 2D images has been proposed instead of a trace-by-trace process, giving the possibility to exploit the coherency of the data not only along the time axis but also along the receiver axis. The second conclusion is that none of the proposed approaches can practically ensure positivity and mass conservation, while keeping the convexity property with respect to shifted patterns which motivated the introduction of the OT at first.

On this basis, we propose a novel strategy trying to reconcile all these requirements, where the data is considered, after discretization, as point clouds in a higher dimension space: the graph space. We show that this approach naturally yields positivity, mass conservation, and preserves the convexity with time shifts. Two preliminary tests in a 2D acoustic time-domain settings are provided to assess the interest of this strategy. The main limitation of the method is, for now, its computation cost. However, more dedicated numerical strategies could be employed to make the method feasible for realistic scale applications.

OT FOR SEISMIC DATA: A SHORT REVIEW

For two positive functions $g(x), h(x)$ defined on a space $X \subset \mathbb{R}^d$, satisfying the mass conservation assumption

$$\int_x g(x)dx = \int_x h(x)dx, \tag{1}$$

the optimal transport problem, under the formulation of Kantorovich (1942), is defined as

$$\min_{\gamma(x,x')\in\Pi_{g,h}} \int_x \int_{x'} \gamma(x,x')\|x-x'\|dxdx', \tag{2}$$

where $\|.\|$ is the Euclidean norm of $\mathbb{R}^d$ and

$$\Pi_{g,h} = \left\{ \gamma(x,x'),\ \int_x \gamma(x,x')dx = h(x'),\ \int_{x'} \gamma(x,x')dx' = g(x) \right\} \tag{3}$$

is the space of couplings $\gamma(x,x')$ between an initial and a final mass distribution $g(x)$ and $h(x)$. Intuitively, OT seeks the coupling between g and h which minimizes the effort to transport the mass from the configuration g to the configuration h. The effort is measured as the sum of each mass quantity $\gamma(x,x')$ which is moved between x and x', multiplied by the "ground"

10.1190/segam2018-2997001.1

distance $\|x-x'\|$ between these points. The OT problem defines the family of p-Wasserstein distances (analogously to the L^p distances)

$$W_p(g,h) = \left(\min_{\gamma(x,x')\in\Pi_{g,h}} \int_x \int_x' \gamma(x,x') \|x-x'\|^p dx dx' \right)^{1/p}. \tag{4}$$

Consider two seismic traces $d_{cal}(t)$ and $d_{obs}(t)$. These time functions are oscillatory: the positivity assumption required for OT is not satisfied. However, they satisfy the mass conservation assumption. The total mass in OT sense is their time-integral, which corresponds to their zero frequency. We know that the zero frequency has no energy in seismic data:

$$\int_t d_{cal}(t)dt = \int_t d_{obs}(t)dt = 0. \tag{5}$$

Three propositions have been made in the literature to overcome the non-positivity issue, focused on the 2-Wasserstein distance. The first consists in considering the positive and negative part of the traces separately, defining the misfit function

$$f_{+/-}(d_{cal},d_{obs}) = W_2^2(d_{cal}^+,d_{obs}^+) + W_p^p(d_{cal}^-,d_{obs}^-), \tag{6}$$

where $.^+$ and $.^-$ are the positive and negative part operators respectively (Engquist and Froese, 2014). The second consists in an affine scaling of the data using a positive constant c large enough to make both $d_{cal}(t)$ and $d_{obs}(t)$ positive (Engquist et al., 2016). The corresponding misfit function is

$$f_{scal}(d_{cal},d_{obs}) = W_2^2(d_{cal}+c,d_{obs}+c). \tag{7}$$

The third consists in comparing the exponential of the data (Qiu et al., 2017), defining the misfit function

$$f_{exp}(d_{cal},d_{obs}) = W_2^2(e^{d_{cal}},e^{d_{obs}}). \tag{8}$$

A fourth proposition is based on the dual formulation of the 1-Wasserstein distance (Métivier et al., 2016b,a,c)

$$W_1(g,h) = \max_{\varphi\in \mathrm{Lip}_1(X)} \int_x \varphi(x)(g(x)-h(x))dx, \tag{9}$$

where $\mathrm{Lip}_1(X)$ is the space of 1 Lipschitz function

$$\mathrm{Lip}_1(X) = \left\{ \varphi(x),\ \forall (x,x') \in X\times X,\ |\varphi(x)-\varphi(x')| \le \|x-x'\| \right\}. \tag{10}$$

For g and h satisfying the mass conservation assumption, even if they are not positive, the dual problem (9) has a solution. Therefore, we can define the misfit function

$$f_{dual}(d_{cal},d_{obs}) = W_1(d_{cal},d_{obs}). \tag{11}$$

We select four criterion to compare these misfit functions: (I) convexity with respect to time shifts, (II) applicability to field data, (III) differentiability, (IV) numerical strategy to solve the OT problem.

(I) To assess the convexity of the misfit functions, we perform a simple test where $d_{cal}(t)$ and $d_{obs}(t)$ are two shifted in time Ricker functions. The value of the misfit functions, depending on this time shift, are presented in Figure 1. The convexity (or at least a sensitivity to large time shifts) is ensured only using $f_{+/-}$ and f_{exp}: f_{scal} and f_{dual} are not convex. For f_{scal}, the reason is that adding a constant c creates artificial mass everywhere, making possible local mass transfers to map d_{cal} onto d_{obs} through OT, instead of transferring mass along the time axis. For f_{dual}, the same explanation is valid: from equation (9) we see that f_{dual} is insensitive to the addition of a constant to the data: $f_{dual}(d_{cal}+c,d_{obs}+c) = f_{dual}(d_{cal},d_{obs})$.

(II) The mass conservation is satisfied by $d_{cal}(t)$ and $d_{obs}(t)$. However, there is no reason this should be the case for their positive and negative parts, or for the exponential of the data. Therefore $f_{+/-}$ and f_{exp} can not be applied to field data, unless specific unbalanced OT solvers are used (see Chizat et al., 2018, for instance). But using such solvers might in turn degrades the convexity property.

(III) The differentiability is an issue for $f_{+/-}$. The negative and positive part operators are not differentiable, therefore $f_{+/-}$ is not differentiable with respect to d_{cal}. This is an issue for using it in the frame of FWI, where its gradient should be used to update the subsurface model.

(IV) The three strategies $f_{+/-}, f_{scal}, f_{exp}$ rely on the 2-Wasserstein distance, and make use of the analytic formula available for the solution of 1D OT problems (based on the inverse of the cumulative functions associated with $d_{cal}(t)$ and $d_{obs}(t)$). For this reason, they are very efficient but restricted to trace by trace comparison. On the other hand, an efficient numerical solver, based on a proximal splitting algorithm, has been developed for f_{dual}, which makes possible to consider whole 2D/3D gathers and to account for the lateral coherency of the data along receiver/source axis in the seismic data cube (Métivier et al., 2016c).

While only $f_{+/-}$ and f_{exp} maintain the convexity property, they seem not adapted to the application to field data because of the mass conservation issue. Also $f_{+/-}$ is not differentiable. On the other hand, f_{scal} and f_{dual} lose the convexity but can be applied to seismic data, with the possibility for f_{dual} to consider 2D or 3D shot gather simultaneously. Satisfactory results have been obtained by applying these techniques to realistic synthetic data, mitigating the non-convexity of the standard L^2 approach. Yet, one could wonder if there is a possibility to apply OT to seismic data while still benefiting from its convexity.

A GRAPH SPACE APPROACH

Formulation

In order to overcome the above-mentioned limitations, we propose the following graph space approach. We consider a seismic trace $d(t)$ as a set of discrete points $\{(t_i,d_i),\ i=1,\dots N\}$, where t_i is the time discretization and $d_i = d(t_i)$. This ensemble of points constitutes the discrete graph of the function $d(t)$. A graph space transform transfer the 1D time trace into a sum of Dirac delta points in the 2D graph space, one axis representing the time, the other the amplitude. To ensure the differentiability of the misfit function, we introduce a smooth graph space transform where the Dirac delta functions are approximated through 2D Gaussian functions, defined by

$$\begin{array}{rccc} \mathscr{G}_\sigma : & \mathbf{d} & \longrightarrow & \mathscr{G}_\sigma(d) = d^{\mathscr{G}_\sigma}(x,t) \\ & \mathbb{R}^N & \longrightarrow & \mathbb{C}^\infty(\mathbb{R},\mathbb{R}_*^+), \end{array} \tag{12}$$

with

$$d^{\mathscr{G}_\sigma}(x,t) = \frac{1}{2\pi\sigma_x\sigma_t N} \sum_{i=1}^N \exp\left(-\frac{(t-t_i)^2}{2\sigma_t^2}\right) \exp\left(-\frac{(x-d_i)^2}{2\sigma_x^2}\right), \tag{13}$$

where $\sigma = (\sigma_t,\sigma_x)$, the quantities (σ_t,σ_x) being user-defined

10.1190/segam2018-2997001.1

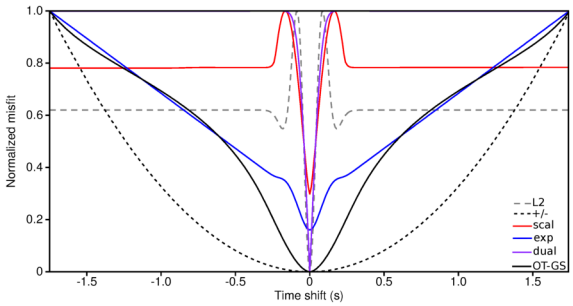

Figure 1: Comparison of misfit functions depending on time shifts for two shifted in time Ricker signals. Gray dotted line: L^2 misfit function. Black dotted line: $f_{+/-}$. Red solid line: f_{scal}. Blue solid line: f_{exp}. Purple solid line: f_{dual}. Black solid line: f_{GS} with $\tau = 4$.

scaling parameters, and $\mathbb{C}^\infty(\mathbb{R},\mathbb{R}_*^+)$ being the set of positive infinitely differentiable functions of $(\mathbb{R})$. We thus introduce the misfit function

$$f_{GS}(d_{cal},d_{obs}) = W_1\left(\mathscr{G}_\sigma(d_{cal}),\mathscr{G}_\sigma(d_{obs})\right). \qquad (14)$$

We focus on the W_1 distance to make possible the use of the 2D solver we have developed in (Métivier et al., 2016c). An efficient 2D OT solver is indeed mandatory in this approach as OT is used to compare the 2D functions resulting from the graph transform.

The interest of this graph transform approach is to guarantee the comparison of positive quantities, with mass conservation (the total mass of both functions is equal to 1 with proper normalizations of the Gaussian functions), while maintaining the convexity with respect to deformation both in the time and amplitude axis. To this respect, a scaling parameter τ is introduced to control the transport cost along these two dimensions. Increasing τ results in penalizing more the transport in the amplitude direction, increasing the sensitivity to time shifts.

The main drawback is related to the potentially high computational cost, as a 2D transport problem needs to be solved for each seismic trace. However this computational cost can be mitigated by decimating in time the 1D trace prior to the transformation in the graph space, and by controlling the number of discrete points introduced to represent the amplitude axis. Also limiting the number of iterations of the proximal splitting technique used to compute W_1 can help reducing the computational cost.

In Figure 2, we present the misfit function f_{GS} for the two shifted in time Ricker functions of the previous example. Different values of τ are used: as can be seen, we recover convex misfit functions. Also the sensitivity to large time shifts increases as τ increases.

Case study 1: cross-hole experiment

To assess the properties of the GS strategy, we first investigate a simple 2D acoustic cross-hole experiment. We use 48 sources located in a 100 m deep borehole, and 48 receivers located in a second borehole, 50 m away from the first borehole. The exact velocity model is homogeneous at 2000 m.s^{-1}. The source wavelet is a Ricker centered at 250 Hz. We consider 3 different homogeneous initial models: (1) $v_P = 1900$ m.s^{-1},

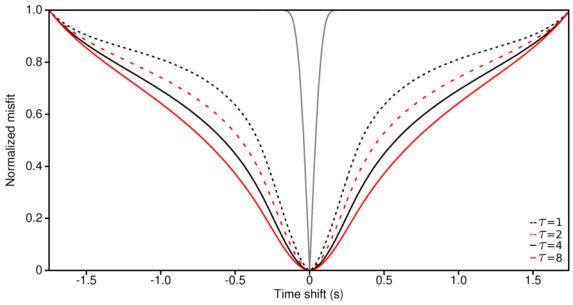

Figure 2: GS misfit function for increasing values of τ. Black dotted line: $\tau = 1$. Red dotted line $\tau = 2$. Black solid line $\tau = 4$. Red solid line $\tau = 8$.

(2) $v_P = 1800$ m.s^{-1} (3) $v_P = 2200$ m.s^{-1}. A 2D Gaussian smoothing is applied to the gradient. The correlation length in each direction is equal to a fraction of the local wavelength (approximately 8 m). Because of this relatively strong smoothing, we use a nonlinear conjugate gradient optimizer (Métivier and Brossier, 2016). For the GS approach, we use 2000 and 100 points for the time and amplitude discretization respectively, with $(\sigma_t,\sigma_x) = (0.001, 0.005)$. We also limit the proximal splitting algorithm to compute the W_1 distance in the graph space to 25 iterations to save computation time. The control parameter τ is set to 1.

We compare the results obtained using a L^2 misfit function, f_{dual} and f_{GS} in Figure 3. Starting from the initial model (1) (Fig. 3b), the three misfit function yield a reliable estimation of the P-wave velocity (Fig. 3c-e). Starting from the initial models (2) and (3), only the GS approach provides a satisfactory estimation (Fig. 3f).

In Figure 4 a single trace is extracted, corresponding to the source and receiver at 49 m depth. For each plot, the traces computed in the exact, initial and reconstructed model using one of the three misfit functions are superimposed. Starting from the initial models (2) and (3), the L^2 and f_{dual} misfit functions are not able to shift the main arrival towards its correct time. Conversely, using f_{GS}, the calculated trace nicely fit the observed data (Fig. 4c,f,i). This simple experiment confirms the promising potentiality of the graph space approach to mitigate the non-convexity of the misfit function.

Case study 2: Marmousi experiment

We consider the Marmousi 2 synthetic case study, in acoustic settings (Martin et al., 2006). We use a fixed-spread surface acquisition at 50 m depth with 128 sources located each 130 m from $x = 0.05$ km to $x = 16.7$ km, and 168 receivers, located each 100 m from $x = 0.05$ km to $x = 16.8$ km. The observed data is computed on a 6 s time window, with a free surface boundary condition at the water/air interface. The source wavelet is a Ricker centered on 6 Hz and high-pass filtered to remove the energy below 3 Hz. We use the same gradient smoothing strategy as previously : the correlation length goes from 80 m to 320 m in this example. The l-BFGS from the SEISCOPE optimization toolbox is used (Métivier and Brossier, 2016). The initial model is 1D and it linearly increases in depth from 1500 m.s^{-1} to 3200 m.s^{-1} (Fig. 5b), resulting in a severe underestimation of the velocity at depth.

10.1190/segam2018-2997001.1

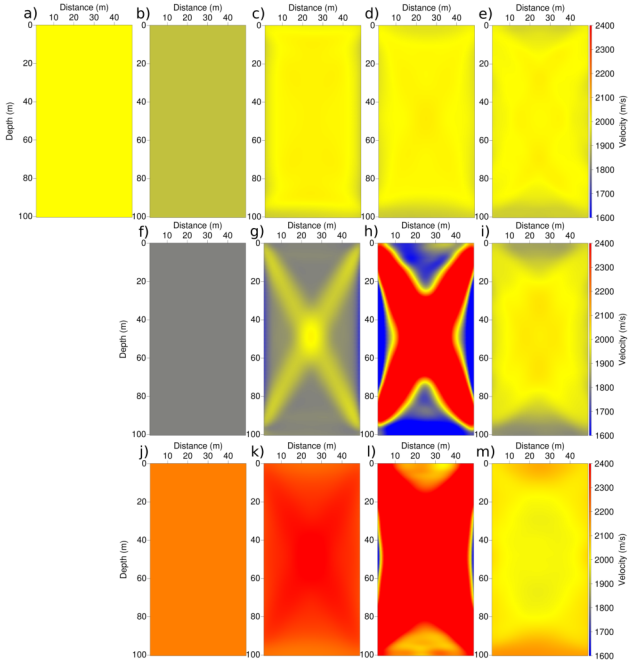

Figure 3: Cross-hole experiment: (a) exact model at 2000 m.s^{-1}. (b) Initial model at 1900 m.s^{-1}, and corresponding reconstructed models using (c) the L^2 misfit function, (d) f_{dual} and (e) f_{GS}. (f) Initial model at 1800 m.s^{-1}, and corresponding reconstructed models using (g) the L^2 misfit function, (h) f_{dual} and (i) f_{GS}. (j) Initial model at 2200 m.s^{-1}, and corresponding reconstructed models using (k) the L^2 misfit function, (l) f_{dual} and (m) f_{GS}.

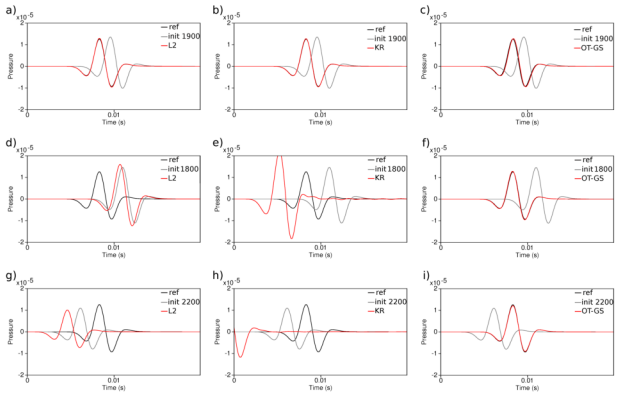

Figure 4: Comparison of traces from the shot and the receiver at $z = 49$ m depth. The seismic traces are computed in the exact, initial and estimated models. For the initial model at $v_P = 1900$ m.s^{-1}, (a) L^2 misfit function, (b) f_{dual}, (c) f_{GS}. For the initial model at $v_P = 1800$ m.s^{-1}, (d) L^2 misfit function, (e) f_{dual}, (f) f_{GS}. For the initial model at $v_P = 2200$ m.s^{-1}, (g) L^2 misfit function, (h) f_{dual}, (i) f_{GS}.

We compare the results obtained using a L^2 misfit function, f_{dual} and f_{GS} misfit functions in Figure 5. For the GS approach, we use 650 and 200 points for the time and amplitude

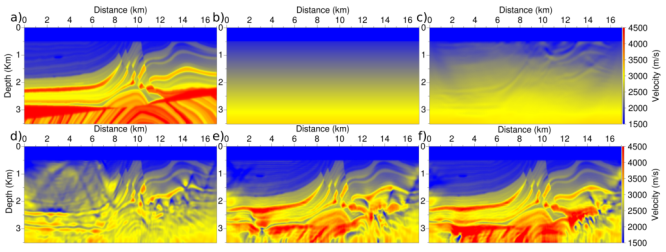

Figure 5: (a) Exact Marmousi model, (b) 1D initial model, (c) L^2 FWI result, (d) f_{dual} FWI result, (e) f_{GS} FWI result with $\tau = 1.5$, (f) f_{GS} final result with $\tau = 0.0025$.

discretization respectively, with $(\sigma_t, \sigma_x) = (0.001, 0.0025)$. We perform 40 iterations of the proximal splitting algorithm to compute the W_1 distance. The control parameter τ is first set to 1.5. As can be seen, both the L^2 and f_{dual} approaches fail to provide meaningful reconstruction of the velocity model, starting from this crude initial guess. On the contrary, the f_{GS} approach yields a much better constrained model, mainly in the central part where the illumination of the medium is maximum.

The Figure 2 for the Ricker experiment suggests that a hierarchical approach with decreasing values of τ could be used. We thus re-start from the final estimation obtained with f_{GS} and run several inversions with τ decreasing from 0.5, 0.25, 0.1, 0.05 to 0.025, meaning that the amplitude is more and more considered. The final estimation is presented in Figure 5f. The velocity estimation is well refined, especially in the less illuminated zones. Note that we tried to use the L^2 and f_{dual} misfit function starting from the model in Figure 5e with no notable improvements.

CONCLUSION AND PERSPECTIVES

To overcome limitations of previously proposed approaches to apply OT to FWI, we apply OT to the graph of the seismic traces using a smooth graph-transform operator. This approach yields promising results for two synthetic case studies: a cross-hole experiment and the Marmousi case study. Its main drawback is its computational cost, as for each seismic trace, a 2D OT problem has to be solved. On the Marmousi example, computing one gradient is approximately 6 times more expensive than computing the L^2 gradient. Further work will be dedicated to reducing this cost, possibly through alternative OT approaches, using entropic regularization (Benamou et al., 2015) or the auction algorithm (Bertsekas and Castanon, 1989).

ACKNOWLEDGMENTS

This study was partially funded by the SEISCOPE consortium (*http://seiscope2.osug.fr*), sponsored by AKERBP, CGG, CHEVRON, EXXON-MOBIL, JGI, PETROBRAS, SCHLUMBERGER, SHELL, SINOPEC, STATOIL and TOTAL. This study was granted access to the HPC resources of CIMENT infrastructure (https://ciment.ujf-grenoble.fr) and CINES/IDRIS/TGCC under the allocation 046091 made by GENCI

10.1190/segam2018-2997001.1

REFERENCES

Benamou, J.-D., G. Carlier, M. Cuturi, L. Nenna, and G. Peyré, 2015, Iterative Bregman Projections for regularized transportation problems: SIAM Journal on Scientific Computing, **37**, A1111–A1138, https://doi.org/10.1137/141000439.

Bertsekas, D. P., and D. Castanon, 1989, The auction algorithm for the transportation problem: Annals of Operations Research, **20**, 67–96, https://doi.org/10.1007/BF02216923.

Chizat, L., G. Peyré, B. Schmitzer, and F. X. Vialard, 2018, An interpolating distance between optimal transport and Fisher–Rao metrics: Foundations of Computational Mathematics, **18**, 1–44, https://doi.org/10.1007/s10208-016-9331-y.

Engquist, B., and B. D. Froese, 2014, Application of the Wasserstein metric to seismic signals: Communications in Mathematical Science, **12**, 979–988, https://doi.org/10.4310/CMS.2014.v12.n5.a7.

Engquist, B., B. D. Froese, and Y. Yang, 2016, Optimal transport for seismic full waveform inversion: Communications in Mathematical Sciences, **14**, 2309–2330, https://doi.org/10.4310/CMS.2016.v14.n8.a9.

Jannane, M., W. Beydoun, E. Crase, D. Cao, Z. Koren, E. Landa, M. Mendes, A. Pica, M. Noble, G. Roeth, S. Singh, R. Snieder, A. Tarantola, and D. Trezeguet, 1989, Wavelengths of Earth structures that can be resolved from seismic reflection data: Geophysics, **54**, 906–910, https://doi.org/10.1190/1.1442719.

Kantorovich, L., 1942, On the transfer of masses: Doklady Academy Nauk USSR, **37**, 7–8.

Martin, G. S., R. Wiley, and K. J. Marfurt, 2006, Marmousi2: An elastic upgrade for Marmousi: The Leading Edge, **25**, 156–166, https://doi.org/10.1190/1.2172306.

Métivier, L., and R. Brossier, 2016, The SEISCOPE optimization toolbox: A large-scale nonlinear optimization library based on reverse communication: Geophysics, **81**, no. 2, F11–F25.

Métivier, L., R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, 2016a, Increasing the robustness and applicability of full waveform inversion: an optimal transport distance strategy: The Leading Edge, **35**, 1060–1067, https://doi.org/10.1190/tle35121060.1.

Métivier, L., R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, 2016b, Measuring the misfit between seismograms using an optimal transport distance: Application to full waveform inversion: Geophysical Journal International, **205**, 345–377, https://doi.org/10.1093/gji/ggw014.

Métivier, L., R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, 2016c, An optimal transport approach for seismic tomography: Application to 3D full waveform inversion: Inverse Problems, **32**, 115008, https://doi.org/10.1088/0266-5611/32/11/115008.

Operto, S., A. Miniussi, R. Brossier, L. Combe, L. Métivier, V. Monteiller, A. Ribodetti, and J. Virieux, 2015, Efficient 3- D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: Application to Valhall in the visco-acoustic vertical transverse isotropic approximation: Geophysical Journal International, **202**, 1362–1391, https://doi.org/10.1093/gji/ggv226.

Plessix, R. E., and C. Perkins, 2010, Full waveform inversion of a deep water ocean bottom seismometer dataset: First Break, **28**, 71–78, https://doi.org/10.3997/1365-2397.2010013.

Qiu, L., J. Ramos-Martinez, A. Valenciano, Y. Yang, and B. Engquist, 2017, Full-waveform inversion with an exponentially encoded optimal-transport norm: 87th Annual International Meeting, SEG, Expanded Abstracts, 1286–1290, https://doi.org/10.1190/segam2017-17681930.1.

Vigh, D., K. Jiao, D. Watts, and D. Sun, 2014, Elastic full-waveform inversion application using multicomponent measurements of seismic data collection: Geophysics, **79**, no. 2, R63–R77, https://doi.org/10.1190/geo2013-0055.1.

Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. Stekl, L. Guasch, C. Win, G. C-. Roy, A. Bertrand, 2013, Anisotropic 3D full-waveform inversion: Geophysics, **78**, no. 2, R59–R80, https://doi.org/10.1190/geo2012-0338.1.

10.1190/segam2018-2997001.1