Snapshots of modern mathematics
from Oberwolfach

№ 4/2022

Seeing through rock with help from optimal transport

Christina Frederick[1] • Yunan Yang[2]

Geophysicists and mathematicians work together to detect geological structures located deep within the earth by measuring and interpreting echoes from manmade earthquakes. This inverse problem naturally involves the mathematics of wave propagation, but we will see that a different mathematical theory – optimal transport – also turns out to be very useful.

1 Introduction

> *"Most people, if you describe a train of events to them will tell you what the result will be. There are few people, however that if you told them a result, would be able to evolve from their own inner consciousness what the steps were that led to that result. This power is what I mean when I talk of reasoning backward."*
>
> – Sherlock Holmes in *A Study in Scarlet*
> Sir Arthur Conan Doyle (1887)

Everyday life is full of *inverse problems*: problems or situations that can be addressed by collecting clues and using backward reasoning to narrow down a specific sequence of events that can solve a mystery. For example, if you lost your

[1] The first author was partially supported by NSF DMS-1720306.
[2] The second author was partially supported by NSF DMS-1913129.

keys somewhere in your house, you would start the detective work by forming a mental image that recreates what you were doing the last time you saw them. Then, starting with the last known location, you would systematically visit rooms in the house, performing a thorough check before completely eliminating one room and moving on to another, until you find the location where they were last stored. But what if the mystery is more complicated? What if the goal of your search isn't visible?

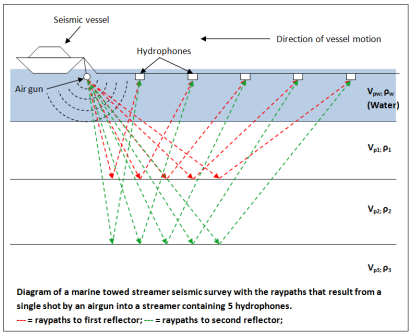
People often visit a doctor when something is wrong; they may be sick, in pain, or perhaps notice unusual patterns in their health. In these cases, the doctor must play the role of Sherlock Holmes and solve mysteries that are much more complex than the case of missing keys. A brain tumor, for example, can't always be easily spotted with the human eye. Instead of dusting for fingerprints at the scene, a doctor collects evidence by measuring the patient's vitals and documenting any symptoms, lifestyle changes, and family history. Then, the information is used to piece together possible causes and form an initial treatment plan.

Unfortunately, detectives may follow red herrings that lead them on the wrong path, and this is also true for medical doctors. There can be many different conditions that produce the same symptoms, and this is why follow-up appointments are needed to confirm whether the initial diagnosis was correct and the treatment was successful. If not, the plan can be updated, and the process repeated until the patient is better. In both of these real-world detective cases, a combination of collected evidence, backward reasoning through trial and error, and a bit of luck can illuminate the investigation.

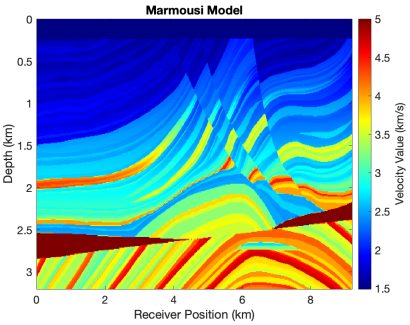
Like Sherlock Holmes, researchers in almost every field in science and engineering use indirect measurements or observational data from complex physical or biological processes to gain an understanding of a quantitative problem. Applications of inverse problems range from finding the energy resources of the Earth by recording shock waves from manmade earthquakes (see Figure 1) to checking the development of a fetus using ultrasound (see Figure 2), and even mapping the surfaces of planets in the solar system with radar imaging.

To formulate these inverse problems mathematically, we use m to denote a parameter of interest that, if known, solves the mystery. Finding the unknown m involves an investigation process that generates observable data d containing clues that ideally lead to a unique choice of m. The *forward model* is characterized by the map F that describes the process that transforms the input parameter m to the observable d:

$$d = F(m). \tag{1}$$



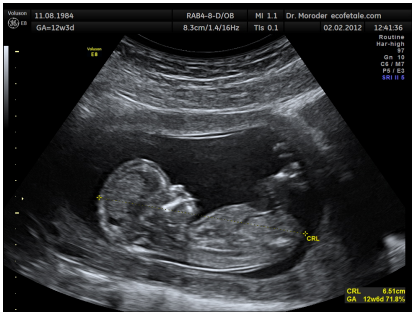
(a)



(b)

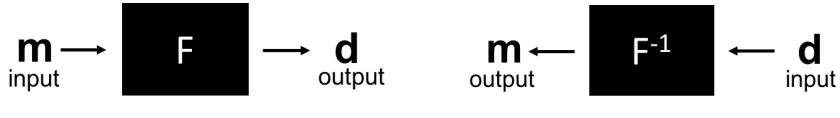
Figure 1: (a) A diagam of the seismic survey for subsurface exploration that uses sound waves to generate seismic data collected by receivers. (b) Geophysical properties of the Marmousi profile [13]. This graph gathers values of parameters that we aim to reconstruct by solving inverse problems.

(a)



(b)

Figure 2: (a) A doctor performs an ultrasound examination on a patient.
(b) Ultrasound image of a fetus at 12 weeks of pregnancy in a sagittal scan.



(a) (b)

Figure 3: (a) The forward map from input parameters to observables. (b) The inverse problem is to use collected data to find unknown model parameters. In both seismic imaging (Figure 1) and ultrasound imaging (Figure 2), F describes physical wave propagation processes.

The inverse problem is to "undo" the forward process, that is, find a Sherlock Holmes map F^{-1} that traces the observable data d back to the parameter m:

$$m = F^{-1}(d). \tag{2}$$

Such a map is more rigorously called "inverse map". The *inverse map* of F is the map F^{-1} that fulfills $F^{-1} \circ F = F \circ F^{-1} = \mathrm{Id}$, where Id is the identity map that maps an object to itself. Thus, $F^{-1}(d) = F^{-1}(F(m)) = m$. Figure 3 illustrates the relationship between the forward and the inverse maps.

In the case of ultrasound imaging, the unknown m can represent the properties of fetal muscles, tendons, and organs of the fetus. Close monitoring of m can give doctors a good sense of any developmental abnormalities that need to be addressed in the early stages. The ultrasound image plays the role of the observational data d generated by the forward map $F(m)$, where F is based on a known mathematical model of ultrasound wave propagation.

For more general inverse problems, desired parameters m could be physical properties of a medium (such as density, electrical conductivity, heat conductivity), geometrical information (such as the location, shape, and structure of intrusions, defects), or sources (of heat, waves, potential difference, pollution). Examples of the forward processes include optical or acoustic imaging, lab experiments, and electrical probing. In these cases, F is based on a known mathematical forward model that describes the governing equations of various physical phenomena, such as the propagation of sound, heat, seismic waves, or electromagnetic waves. Depending on the application, forward models can also include empirical models, statistical models, or machine learning models.

Solving the inverse problem is hard. First, the data-collection process can be flawed: images can be blurry, measurements can be imprecise, and samples can become contaminated. Inadequate data can produce multiple or even infinitely many solutions to the inverse problem. Secondly, inverse problems can also be mathematically unstable, meaning that small changes in the measurements could be produced by parameters that differ dramatically. In other words, even

if the data is collected with a high-precision instrument, there can still be wide margins of error in the parameter recovery. To make things even worse, it is rarely possible or practical to find the exact mathematical formulas for the inverse map F^{-1}, so inverse problems are almost always solved indirectly.

Why do so many mathematicians and researchers work on inverse problems despite all of these difficulties? Advances in the field of inverse problems make a tremendous impact on our daily lives, improving our ability to monitor the health of a fetus, locate energy sources within the earth, and optimally design airplanes. In addition, inverse problems also appeal to mathematicans interested in the development of elegant theory and analysis. Furthermore, there are a seemingly endless number of challenging open problems to be explored that arise from both the existing theory and also from new applications. The rest of this snapshot focuses on the investigation of an inverse problem in seismic exploration that led to new scientific and mathematical discoveries.

2 Seismic inversion

Inverse problems play a central role in reflection seismology, where the aim is to investigate the Earth's deep subsurface layers using seismic data. The most common form of seismic data is a *seismogram*, which is a recording of the phase and amplitude of waves caused by the ground shaking following an earthquake or explosion. Seismic data has successfully been used to uncover intrinsic features of the Earth, including the crust-mantle-core structure. A current active area of research uses seismic data to solve challenging problems in *exploration seismology*, or the quest to find new energy sources such as oil, gas, and coal reserves.

Detectives in exploration seismology rely on the knowledge that sound waves travel in different materials with dramatically different speeds, called *wave propagation speeds*, given in meters per second (m/s). Sound waves travel at a speed of about 1500 m/s in the sand, 2100 m/s in coal, and 4800 m/s in shale. Therefore, once the wave speeds of all the subsurface layers are known, the material types are also revealed.

To determine the sound speed of the subsurface layers, seismologists conduct an experiment starting with a loud explosion occurring at time $t = 0$ seconds. The traveling waves propagate through the Earth and are partially reflected back to recording instruments, called receivers, located at multiple seismic stations. The receivers then track the time-history measurements of both the phase and amplitude.

Figure 1a is a diagram of a seismic marine exploration survey. Such exploration surveys are carried out by seismic vessels, which use air guns to generate seismic waves that propagate through water and then earth. Hydrophones are

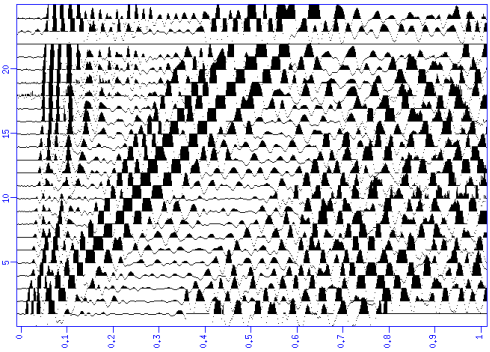


Figure 4: Example of seismic data taken at 25 receiver locations l_i (vertical axis), over the times t with $0 \leq t \leq 1$ (horizontal axis).

placed on the surface of the water around the vessel in order to measure the seismic waves reflected back from the earth at different depths. The goal of this type of seismic survey is to reconstruct geophysical properties of the underground layers. The results can then be compiled into a velocity image, such as the "Marmousi model", which is shown in Figure 1b. This is the application we will focus on in the paper.

Now, we can formulate the forward problem mathematically as $d = F(m)$. We represent the Earth's subsurface by a coordinate system (x, y, z), where for any location l the coordinates x and y represent the horizontal and vertical distance from a fixed origin, and z represents the depth in kilometers (km). The unknown wave propagation speed is denoted by a function $m(x, y, z)$ that varies based on the location inside the Earth. Then, the physical laws governing wave propagation in different media are captured in the forward operator $F(m) = d$, where d is the observable seismic data taken at N receiver locations $l_1, \ldots, l_N$ where each location l_i is described by coordinates $l_i = (x_i, y_i, 0)$ over the times t with $0 \leq t \leq T$. That is, $d_i(t) = u(l_i, t)$ where $u(l, t)$ solves the acoustic wave equation

$$m(l)\frac{\partial^2 u(l,t)}{\partial t^2} - \triangle u(l,t) = s(l,t) \tag{3}$$

for a given source $s(l, t)$ and appropriate boundary conditions (this equation is used to model the empirical data shown in Figure 4).

The seismic inversion process $m = F^{-1}(d)$ can be a cumbersome undertaking for both geophysicists and applied mathematicians. Since the waves with higher propagation speeds will reach the receivers earlier, an excellent way to start decoding the recorded seismic data is to analyze patterns in the arrival times of the reflected waves. However, the complicated, layered structure of the Earth's

subsurface can add a great deal of difficulty in the interpretation, and more advanced mathematical tools are needed in order to accurately reconstruct the wave propagation speed distribution in each subsurface layer.

3 Solving the inverse problem via optimization

Instead of the (often impossible) task of directly finding m by applying F^{-1}, inverse problems can be solved indirectly. A common way of solving an inverse problem is to approximate the true model parameter m^* by solving an *optimization problem*. This method essentially boils down to trial and error. For each guess m, a prediction $F(m)$ is generated. A *misfit function* $J(m)$ is used to measure how well this prediction matches the actual observed data d. Misfit functions typically measure the difference between $F(m)$ and d, and minimizing J finds the best possible choice of m:

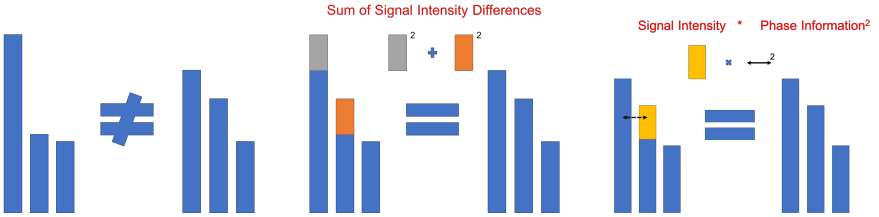
$$m^* = \underset{m}{\operatorname{argmin}}\, J(F(m), d), \tag{4}$$

where the function argmin returns the value of m that produces the smallest J. A significant advantage of (4) over (2) is that the process of trial and error can be performed efficiently and intelligently using advanced optimization techniques.

The most common misfit function used in inverse problems is the *least-squares* or L^2 *norm* which computes the sum of squared difference between observed data d and the simulated prediction $F(m)$ generated by a guess of the model parameter m:

$$J_{L^2}(F(m), d) = \sum_{i=1}^{N} |F_i(m) - d_i|^2, \tag{5}$$

where $F(m)$ and d are vectors with N elements $\{F_i(m)\}_{i=1}^N$ and $\{d_i\}_{i=1}^N$. The L^2 norm defined above is a point-by-point comparison, meaning it accumulates differences in the data and prediction for each coordinate. This way of measuring differences can cause problems for inversion because it is sensitive to noisy data and can lead to a modeling error called *overfitting*, in which a model is too closely fit to data. Another drawback of the L^2 norm is that if the recorded data contains oscillatory patterns, like in the case of seismic signals, the misfit function can contain multiple local minima that can cause the optimization solver described in (4) to stop prematurely and return an incorrect parameter approximation [4].



(a) Two signals **(b)** L^2 norm calculation **(c)** W_2 metric calculation

Figure 5: (a) Two signals f and g; (b) After subtraction (gray) and addition (orange) in signal intensity, two signals are equal. The L^2 norm is the sum of the squared differences; (c) After mass transportation, the two signals are equal. The W_2 metric is the multiplication of the displaced signal intensity and transport distance squared.

4 Optimal transport and the Wasserstein metric

To overcome the challenges caused by using the conventional L^2 norm for measuring the difference between predicted data $F(m)$ and observed data d, we can replace it with a new metric, that is a new notion of distance, with more desirable mathematical properties such as "convexity" and "stability". The graph of a *convex* function is curved outward, like the parabola $y = x^2$. A nonconvex function behaves like $y = sin(x)$, where there are multiple "ups" and "downs". Convex functions have only one global minimum, making convexity a desirable property of a misfit function. *Stability* is another desirable property - a distance function is stable if small changes in the input result in small changes in the measured distance. This means that small errors in the data won't change the outcome dramatically.

Engquist and Froese [8] first proposed to use an alternative notion of distance, called the "Wasserstein metric", derived from optimal transport theory. Their research, focused on seismic inversion, found that the use of the Wasserstein metric mitigated some common difficulties of the least-squares formulation, including sensitivity to noise and non-convexity of the misfit function from the last section. Since the release of the work [8], there has been fruitful activity in further incorporating metrics derived from optimal transport in seismic inversion. In the past four years, these ideas have been studied and implemented in both academia [2, 6, 9, 17, 18, 20, 23, 32] and industry [22, 24, 26] with field data applications, for example, under the North Sea and in Brazil.

The Wasserstein metric is based on the theory of optimal transportation. The optimal transport problem can be viewed as a quest for the most efficient way of transforming one distribution of mass to another. This problem was posed

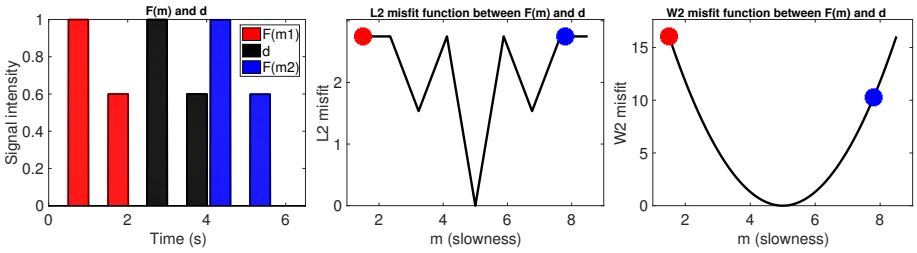
by the French mathematician Gaspard Monge (1746 – 1818) in 1781 [19] and later expanded by the Soviet mathematician and economist Leonid Kantorovich (1912 – 1986) during World War II [14]. Optimal transport-related techniques are both nonlinear and algorithmic, and involve exploring a model in terms of both signal intensities (how large the peaks in the signal are) and locations (the times when the signal has a peak). The significant contributions of the mathematical analysis of optimal transport problems since the 90's [5, 10, 12, 27, 30, 31] together with recent advancements in numerical methods [3, 7, 16, 21, 25] have driven the development of numerous modern data analysis techniques for estimation, detection, and imaging problems [15]. The optimal transport cost function, called the Wasserstein distance belongs to a class of mathematically well-defined distances [30].

The cartoons in Figure 5 illustrate the differences between the quadratic Wasserstein metric W_2 and the L^2 norm. Figure 5a shows two different signals. One can interpret the L^2 distance in the following way. First, we transform the left signal to match the one on the right by subtracting the extra intensity in the first bar (gray) and adding more intensity in the second bar (orange); see Figure 5b. By definition (5), the L^2 norm is the sum of the squared signal intensity differences. For the W_2 metric, we match the two signals by an entirely different transformation. Instead of creating or destroying the signal intensities, we conserve the total intensity and transport the extra intensity in the first bar to the second bar, as indicated by the yellow block in Figure 5c. By definition, the W_2 metric is the *total transport cost*, which is calculated as the sum of the amount moved at every bar multiplied by the squared distance of the location change. If there are multiple transport plans, we pick the one that gives the minimum transport cost. The optimal transport cost is the Wasserstein distance, which comes naturally from the optimal transport problem.

5 Optimal transport for seismic inversion

The calculation of the W_2 metric represents a central "non-local" idea in data comparison and analysis. Optimal transport-based methods compare the observed and simulated data globally and thus include important "phase information" in addition to the intensity differences. This makes optimal transport a powerful tool for comparing waveforms, in which significant differences originate from phase mismatches caused by different wave propagation speeds. We will show the clear advantage of W_2 over L^2 in the next example.

Figure 6a shows measurements of the same source propagated through materials with different wave speeds. For better illustration, we choose the model parameter m to be squared *slowness*, a physical quantity directly related



(a) **(b)** **(c)**

Figure 6: (a) Black: signal d; red and blue: signals generated by model parameter m_1 and m_2 respectively; (b) $J_{L^2}(F(m), d)$ for different m (slowness); (c) $J_{W_2}(F(m), d)$ for different m (slowness); the red and blue marks in (b) and (c) represent the misfits between $F(m_1)$ and d and $F(m_2)$ and d under the norms.

to wave speed through the relationship

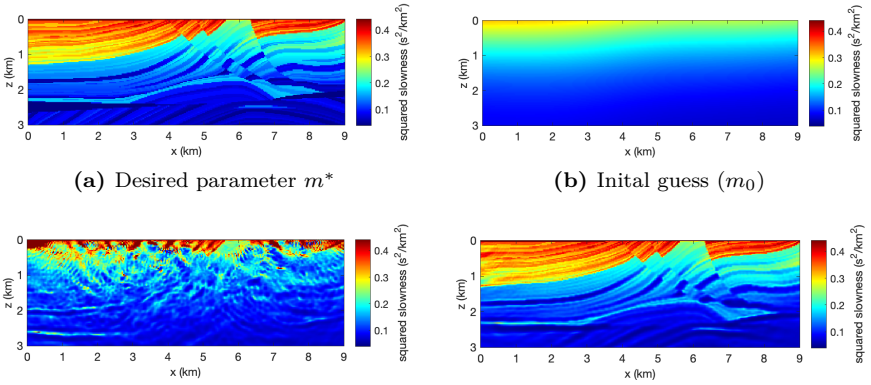
$$\text{squared slowness} = \frac{1}{(\text{wave speed})^2}. \tag{6}$$

Without loss of generality, we assume the black signal in Figure 6a represents the observed waveform d, generated by a true parameter m^*. Our ultimate goal is to reconstruct $m^* = 5$ based on d. The red and blue signals represent predictions (simulated waveforms) computed for two different values of slowness, m_1 and m_2 respectively. Notice in Figure 6a that the red signal indicates an earlier arrival time than the blue one. The different signals here have the same intensities, the discrepancy measured by the L^2 and W_2 distances is therefore solely caused by the different arrival times.

This means that the corresponding signal propagated through a material of faster wave speed and therefore smaller slowness: $m_1 < m_2$. Similarly, we observe that $|m_1 - m^*| > |m_2 - m^*|$ because the gap in the arrival time between $F(m_1)$ and d is larger than the gap in the arrival time between $F(m_2)$ and d.

To solve the inverse problem by optimization (4), one needs to compute the misfit between $F(m)$ and d first. A good misfit $J(F(m), d)$ can be considered as a "ruler" that measures the data misfit based on the choice of model parameters m, emphasizing the differences in the misfit when m changes. Figure 6b shows the shape of the "ruler" if we use the L^2 norm (5) as the misfit. However, the "measurement" for the "ruler" (that is, $J_{L^2}(F(m), F(m^*))$) doesn't always increase as $|m - m^*|$ increases. In particular, we read

$$J_{L^2}(F(m_1), F(m^*)) = J_{L^2}(F(m_2), F(m^*)) \tag{7}$$



(a) Desired parameter m^*

(b) Inital guess (m_0)

(c) Recovered result using L^2

(d) Recovered result using W_2

Figure 7: Inversion of the benchmark Marmousi profile [29].

while $|m_1 - m^*| > |m_2 - m^*|$. Let us illustrate that: For example, the value $m = 4$ is closer to the actual value ($m^* = 5$) than $m = 3$. Yet, using the L^2 norm, $m = 4$ has a larger misfit value than $m = 3$. We have a local minimum of the L^2 misfit function at $m = 3$ that can trick the optimization into returning an erroneous result. Clearly, the L^2 norm is not as reliable at measuring the data misfit for waveform inversion under the optimization formulation (4).

On the contrary, if we use the W_2 metric as the ruler, the measurement shown in Figure 6c increases monotonically as the value $|m - m^*|$ increases. Also, $J_{W_2}(F(m_1), F(m^*)) > J_{W_2}(F(m_2), F(m^*))$, which represents the fact that $|m_1 - m^*| > |m_2 - m^*|$. Compared with the L^2 norm, the W_2 metric has a better characterization of the differences between different wave speed parameters. This simple example motivates us to apply the W_2 metric to real world seismic inversion problems.

6 Numerical results

The discussion and comparison in the previous section illustrate the promising features of W_2 that overcome the issues of the L^2 norm in seismic inversion. The example in Figure 6 only has one single unknown parameter m. In reality, m can represent a vector of thousands of independent variables. In industry, the W_2 metric and other optimal transport-based approaches have been applied to handle more complicated large-scale problems efficiently. In this section, we will show inversion results using both the L^2 norm and the W_2 metric for a subsurface geometry and parameter profile based on the *Marmousi profile*.

The Marmousi profile is a synthetic geological profile designed in 1988 to help researchers better understand practical aspects of seismic inversion. Based on the geological properties in a region of northwestern Angola on the Atlantic Coast of West Africa of size 3 km in depth and 9 km in width, the Marmousi profile contains many subsurface layers, steep dips, and strong velocity gradients, and is used to produce complex seismic data that require advanced processing techniques to solve the inversion problem correctly. Since its creation, the Marmousi model has been used to generate standard benchmark datasets for methods and algorithms [13].

Figure 7a contains plots of the original P-wave[3] squared slowness of the Marmousi profile. We take this to be the "true" model parameter m^* that generates the observable data d. To compute predictions $F(m) = u(l_i, t)$, corresponding to receiver locations $\{l_i\}_{i=1}^N$, and times $0 \leq t \leq 1$ where $u(l, t)$ satisfies (3), we use a numerical forward solver for the acoustic wave equation (3). The optimization procedure for both L^2 and W_2 is initialized with a guess of the parameter m_0 (Figure 7b) and is terminated after 300 iterations. Figure 7c shows the inversion result of using the traditional least-squares L^2 method and Figure 7d shows the final result of using the W_2 misfit function. Again, the result using the L^2 norm has spurious high-frequency artifacts while the inversion using the W_2 metric correctly inverts most details in the true model. Mathematically speaking, this comes from the convexity of the W_2 metric and the nonconvexity of the L^2 norm.

7 Summary

Surprisingly often, mathematical ideas that, on the surface, may seem like they have nothing to do with a particular application, turn out to be strikingly useful. This snapshot highlights how new mathematical insights inform science by illustrating how the idea of applying optimal transport theory and its related Wasserstein metric to inverse problems arose from interdisciplinary research between numerical analysts and geophysicists.

One of the main inverse problems in reflection seismology aims to reveal geological properties of rock located deep within the earth from measured data taken after manmade earthquakes. The standard procedure for solving the inverse problem is: 1) simulate seismic data from predicted geological properties, 2) compare the simulated data with the collected data, 3) improve the prediction,

[3] The P-wave (primary wave or pressure wave) is one of the two main types of elastic body waves, called seismic waves in seismology. P-waves travel faster than other seismic waves and hence are the first signal from an earthquake to arrive at any affected location or at a seismograph. When it comes to reconstruction, it is important to reconstruct the squared slowness that P-waves have as they travel.

and repeat. There are many aspects of this procedure that can be analyzed mathematically, and we focused on new developments that impact step 2). The most common way of evaluating predictions is to measure the least-squares distance between the predicted and collected data. Mathematicians are well aware that several notions of distance can exist, and the use of different metrics can lead to desirable properties of the optimization problem. Noticing that convexity with respect to translation was important, the least-squares norm was replaced by the Wasserstein metric, in which the difference between signals is calculated by how much effort it takes to transport one signal to another.

As it turned out, the idea to replace the least-squares norm with the Wasserstein metric became a turning point in the plot. The choice of the Wasserstein metric and its ability to account for geometric information in the signal (such as a translation) led to a new way of solving the inverse problem in notoriously difficult settings (see Figure 7). Although each inverse problem comes with a unique set of assumptions, challenges, and applications, many inverse problems are tackled by minimizing the difference between predictions and measured data. This means that the ideas developed for inverse problems in reflection seismology can extend well beyond the field. There are many more related problems and mathematical mysteries to be addressed by exploring other disciplines and drawing connections between them. Applications of optimal transport theory and the Wasserstein metric are currently being explored in optics [11], inverse problems in medical imaging [28], and even machine learning [1].

Image credits

Figure 1a "Diagram of a marine seismic survey.png". Author: Nwhit. Licensed under Creative Commons Attribution-Share Alike 4.0 via Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Diagram_of_a_marine_seismic_survey.png, visited on September 16, 2019.

Figure 2a "Aparelhodeultrassom.jpg". Author: Leval. Licensed under Creative Commons Attribution-Share Alike 4.0 via Wikimedia Commons. https://commons.wikimedia.org/wiki/File:Aparelhodeultrassom.jpg, visited on September 16, 2019.

Figure 2b "CRL Crown rump length 12 weeks ecografia Dr. Wolfgang Moroder.jpg". Author: Wolfgang Moroder. Licensed under Creative Commons Attribution-Share Alike 4.0 via Wikimedia Commons. https://commons.wikimedia.org/wiki/File:CRL_Crown_rump_length_12_weeks_ecografia_Dr._Wolfgang_Moroder.jpg, visited on September 16, 2019.

Figure 4 "KR-example.png". Author: DavidMP1. Licensed under Creative Commons Attribution-Share Alike 4.0 via Wikimedia Commons. https://commons.wikimedia.org/wiki/File:KR-example.png, visited on October 27, 2019.

References

[1] M. Arjovsky, S. Chintala, and L. Bottou, *Wasserstein generative adversarial networks*, Proceedings of the 34th International Conference on Machine Learning, Proceedings of Machine Learning Research, vol. 70, PMLR, 06–11 Aug 2017, pp. 214–223.

[2] M. Ballesio, J. Beck, A. Pandey, L. Parisi, E. von Schwerin, and R. Tempone, *Multilevel Monte Carlo acceleration of seismic wave propagation under uncertainty*, International Journal on Geomathematics **10** (2019), no. 22.

[3] J.-D. Benamou, B. D. Froese, and A. M. Oberman, *Numerical solution of the optimal transportation problem using the Monge-Ampère equation*, Journal of Computational Physics **260** (2014), 107–126.

[4] W. B. Beydoun and A. Tarantola, *First Born and Rytov approximations: Modeling and inversion conditions in a canonical example*, The Journal of the Acoustical Society of America **83** (1988), no. 3, 1045–1055.

[5] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Communications on Pure and Applied Mathematics **44** (1991), no. 4, 375–417.

[6] J. Chen, Y. Chen, H. Wu, and D. Yang, *The quadratic Wasserstein metric for earthquake location*, Journal of Computational Physics **373** (2018), 188 – 209.

[7] M. Cuturi, *Sinkhorn distances: Lightspeed computation of optimal transport*, Proceedings of the 26th International Conference on Neural Information Processing Systems, NIPS'13, vol. 2, 2013, pp. 2292–2300.

[8] B. Engquist and B. D. Froese, *Application of the Wasserstein metric to seismic signals*, Communications in Mathematical Sciences **12** (2014), no. 5, 979–988.

[9] B. Engquist, B. D. Froese, and Y. Yang, *Optimal transport for seismic full waveform inversion*, Communications in Mathematical Sciences **14** (2016), no. 8, 2309–2330.

[10] L. C. Evans and W. Gangbo, *Differential equations methods for the Monge-Kantorovich mass transfer problem*, vol. 653, Memoirs of the American Mathematical Society, 1999.

[11] Z. Feng, B. D. Froese, and R. Liang, *Freeform illumination optics construction following an optimal transport map*, Applied Optics **55** (2016), no. 16, 4301–4306.

[12] W. Gangbo and R. J. McCann, *The geometry of optimal transportation*, Acta Mathematica **177** (1996), no. 2, 113–161.

[13] T. Irons, *Marmousi model*, https://reproducibility.org/RSF/book/data/marmousi/paper.pdf, visited on March 1, 2022.

[14] L. V. Kantorovich, *On the Translocation of Masses*, Journal of Mathematical Sciences **133** (2006), 1381–1382.

[15] S. Kolouri, S. Park, M. Thorpe, D. Slepčev, and G. K. Rohde, *Transport-based analysis, modeling, and learning from signal and data distributions*, arXiv preprint arXiv:1609.04767 (2016).

[16] W. Li, S. Osher, and W. Gangbo, *A fast algorithm for Earth Mover's Distance based on optimal transport and L_1 type Regularization*, arXiv preprint arXiv:1609.07092 (2016).

[17] L. Métivier, R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, *Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion*, Geophysical Journal International **205** (2016), no. 1, 345–377.

[18] ———, *An optimal transport approach for seismic tomography: application to 3D full waveform inversion*, Inverse Problems **32** (2016), no. 11, 115008.

[19] G. Monge, *Mémoire sur la théorie des déblais et des remblais*, Imprimerie Royale, 1781.

[20] M. Motamed and D. Appelö, *Wasserstein metric-driven Bayesian inversion with application to signal processing*, International Journal for Uncertainty Quantification **9** (2019), no. 4, 395–414.

[21] A. M. Oberman and Y. Ruan, *Solution of optimal transportation problems using a multigrid linear programming approach*, Journal of Computational Mathematics **38** (2020), no. 6, 933–951.

[22] R. Poncet, J. Messud, M. Bader, G. Lambaré, G. Viguier, and C. Hidalgo, *FWI with optimal transport: a 3D implementation and an application on a field dataset*, Conference Proceedings, 80th EAGE Conference and Exhibition 2018, no. 1, European Association of Geoscientists & Engineers, 2018, pp. 1–5.

[23] M. A. Puthawala, C. D. Hauck, and S. J. Osher, *Diagnosing forward operator error using optimal transport*, Journal of Scientific Computing **80** (2019), no. 3, 1549–1576.

[24] L. Qiu, J. Ramos-Martínez, A. Valenciano, Y. Yang, and B. Engquist, *Full-waveform inversion with an exponentially encoded optimal-transport norm*, SEG Technical Program Expanded Abstracts 2017, Society of Exploration Geophysicists, 2017, pp. 1286–1290.

[25] J. Rabin, G. Peyré, J. Delon, and M. Bernot, *Wasserstein barycenter and its application to texture mixing*, International Conference on Scale Space and Variational Methods in Computer Vision, Springer, 2012, pp. 435–446.

[26] J. Ramos-Martínez, L. Qiu, J. Kirkebø, A. A. Valenciano, and Y. Yang, *Long-wavelength FWI updates beyond cycle skipping*, SEG Technical Program Expanded Abstracts 2018, Society of Exploration Geophysicists, 2018, pp. 1168–1172.

[27] F. Santambrogio, *Optimal transport for applied mathematicians*, Progress in Nonlinear Differential Equations and Their Applications, vol. 87, Birkhäuser, 2015.

[28] Z. Su, W. Zeng, Y. Wang, Z.-L. Lu, and X. Gu, *Shape classification using Wasserstein distance for brain morphometry analysis*, International Conference on Information Processing in Medical Imaging, Springer, 2015, pp. 411–423.

[29] R. Versteeg, *The Marmousi experience: Velocity model determination on a synthetic complex data set*, The Leading Edge **13** (1994), no. 9, 927–936.

[30] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, 2003.

[31] ———, *Optimal transport: old and new*, vol. 338, Springer, 2009.

[32] Y. Yang, B. Engquist, J. Sun, and B. F. Hamfeldt, *Application of optimal transport and the quadratic Wasserstein metric to full-waveform inversion*, Geophysics **83** (2018), no. 1, R43–R62.

Christina Frederick *is an assistant professor in the Department of Mathematical Sciences at New Jersey Institute of Technology.*

Yunan Yang *is an assistant professor in the Department of Mathematics at Cornell University.*

Mathematical subjects
Numerics and Scientific Computing

Connections to other fields
Chemistry and Earth Science, Engineering and Technology, Physics

DOI
10.14760/SNAP-2022-004-EN

Snapshots of modern mathematics from Oberwolfach provide exciting insights into current mathematical research. They are written by participants in the scientific program of the Mathematisches Forschungsinstitut Oberwolfach (MFO). The snapshot project is designed to promote the understanding and appreciation of modern mathematics and mathematical research in the interested public worldwide. All snapshots are published in cooperation with the IMAGINARY platform and can be found on www.imaginary.org/snapshots and on www.mfo.de/snapshots.

ISSN 2626-1995

Mathematisches Forschungsinstitut
Oberwolfach gGmbH
Schwarzwaldstr. 9–11
77709 Oberwolfach
Germany

Director
Gerhard Huisken

Mathematisches
Forschungsinstitut
Oberwolfach

IMAGINARY
open mathematics