Optimal Transportation for Data Assimilation
Nelson Feyeux, Maëlle Nodet, Arthur Vidard

To cite this version:
Nelson Feyeux, Maëlle Nodet, Arthur Vidard. Optimal Transportation for Data Assimilation. 5th International Symposium for Data Assimilation (ISDA 2016), Jul 2016, Reading, United Kingdom. 2016. hal-01349637

HAL Id: hal-01349637
https://hal.archives-ouvertes.fr/hal-01349637
Submitted on 28 Jul 2016
**DATA ASSIMILATION**

Given

- a physical system and its state \(x(t, x, y)\);
- partial observations of the system \((\mathcal{Y})\);
- a (numerical) model \(\mathcal{M}\) simulating the evolution of \(x\);

Can we estimate the initial condition \(x_0\) of the system?

\[
\text{Variational data assimilation consists in retrieving } \theta_0 \text{ by minimizing }
\]

\[
\mathcal{J}(\theta_0) := \sum_i d\left( \frac{\partial \mathcal{M}(\theta_0)}{\partial \theta} \right) y_i^n + \omega d(x_0, x_i^n)^2. 
\]

(1)

It is common for the distance \(d\) to be a weighted \(L^2\) distance. Our main goal is to use the Wasserstein distance \(W_2\) instead, which seems very interesting when dealing with dense data (see right panel). The Wasserstein cost function writes

\[
\mathcal{J}_W(\theta_0) := \sum_i W_2(\mathcal{H}(\theta_0), y_i^n) + \omega W_2(x_0, x_i^n)^2. 
\]

(2)

**RESULTS ON A SHALLOW-WATER EQUATION**

Let the model \(\mathcal{M}\) be a Shallow-Water equation, with initial condition \((h_0, u_0)\),

\[
\mathcal{M} : \begin{cases} 
\frac{\partial h}{\partial t} + \text{div}(hu) = 0 \\
\frac{\partial u}{\partial t} + u \cdot \nabla u = -g \nabla h.
\end{cases}
\]

We control the initial condition \(h_0\) only, thanks to the Wasserstein cost function \(\mathcal{J}_W\). We set \(u_0 = 0\).

The observations of \(\mathcal{H}^{2\times 2}\) at times \(t_i\)

**SPECIFICITIES ON USING THE WASSERSTEIN DISTANCE**

- The Wasserstein distance is only defined for probability measures, i.e. \(\rho \geq 0\) s.t.
  \[
  \int \rho = 1
  \]

- Relaxations of the latter constraint are possible, however complex:
  - the \(W_2\) interpolation works well if \(\rho_0\) and \(\rho_1\) are of distinct support;
  - when \(\mathcal{J}(\rho) \to \min_{\rho_0} \mathcal{J}(\rho_0)\), then there is only weak convergence of \(\rho\) to \(\rho_0\): oscillations or diracs can occur!
  - Computing the Wasserstein distance is expensive [Peyré, Papadakis, Oudet, 2013].

**OPTIMAL TRANSPORTATION AND THE WASSERSTEIN DISTANCE**

For two functions \(\rho_0(x)\) and \(\rho_1(x)\), the square of the Wasserstein distance \(W_2(\rho_0, \rho_1)^2\) is defined as the minimal kinetic energy necessary to transport \(\rho_0\) to \(\rho_1\),

\[
W_2(\rho_0, \rho_1)^2 := \inf \int_{\Omega} \rho \|\nabla \Phi\|^2 dx.
\]

The minimization of \(\mathcal{J}_W\) is performed through a gradient descent, using the Wasserstein gradient, arising from the use of the following Wasserstein scalar product depending on \(\rho_0\),

\[
\langle \eta, \eta' \rangle_{\mathcal{H}} := \int_{\Omega} \rho_0 \nabla \Phi \cdot \nabla \Phi' dx.
\]

**Results:**

- Analysis of \(h_0\) of the assimilation when using the Euclidean \(L^2\) or the Wasserstein \(W_2\) distance.

**Values of \(h\) and \(u\) for the background and true states, as well as analysis for Euclidean and Wasserstein distances, at time \(t\) = \(t_0\).**

This work has been supported by the region Rhône-Alpes.