Multiphysics optimal transportation and image analysis

Afaf Bouharguane, Emmanuel Maitre, Edouard Oudet, Nicolas Papadakis

Mathematical Modelling and Numerical Analysis Will be set by the publisher
Modélisation Mathématique et Analyse Numérique

MULTIPHYSICS OPTIMAL TRANSPORTATION AND IMAGE ANALYSIS

Afaf Bouharguane, Emmanuel Maitre, Édouard Oudet and Nicolas Papadakis[1]

Abstract. Benamou and Brenier formulation of Monge transportation problem [3] has proven to be of great interest in image processing to compute warpings and distances between pair of images [1]. In some applications, however, the built-in minimization of kinetic energy does not give satisfactory results. In particular cases where some specific regions represent physical objects, it does not make sense, as produces genuine optimal transport, to split, merge or arbitrarily deform these regions along the optimal path. The aim of this work is to introduce several extended energies to take care of physical properties of the image in the interpolation process. We present algorithms to compute approximations of the corresponding generalized optimal transportation plans.

1991 Mathematics Subject Classification. ???, ???

Received: date / Accepted: date.

1. Introduction

1.1. Context in image processing

Optimal transportation has found a wide field of application in image interpolation and registration, since pioneering works of Benamou and Brenier [3] who introduced an algorithm based on the minimization of the kinetic energy by flows which preserve the mass. By structure, optimal transportation does not preserve image regions along the optimal interpolation path. Consequently it is actually not difficult to exhibit test cases where the algorithm produces a path of images where high density regions split at the beginning before merging back at its end (see [3] for an example). However, in some applications to image interpolation this behaviour is not desirable. Suppose for instance that the images represent the evolution in time of a droplet of oil in water. The droplet can obviously split during time; however, this phenomenon has a physical cost which is not taken into account in the interpolation given by the genuine optimal transportation model. For an elastic membrane, we would expect no splitting at all. This article aims at studying how some physics can be added to the optimal transportation theory, how to construct algorithms to compute solutions to the corresponding optimization problems and how to apply the proposed methods to image interpolation. We introduce several extended models of optimal transportation, which will amount to add some energy terms to the classical kinetic energy,

Keywords and phrases: ...

[1] Laboratoire Jean Kuntzmann,
Grenoble University and CNRS
e-mail: afaf.bouharguane@math.u-bordeaux1.fr, emmanuel.maitre@imag.fr, edouard.oudet@imag.fr, nicolas.papadakis@imag.fr
(corresponding author) *Present address of A. Bourhaguane and N. Papadakis:* Institut de Mathématiques de Bordeaux, Université Bordeaux 1.

or add constraints on the admissible space for minimizers. Note that due to this generalized form, the Euler-Lagrange equations are not of Monge-Ampere type as genuine optimal transportation, and algorithms based on this formulation [8], and recently [6, 7], can not be used in our context. This work focus on algorithms, with their numerical validation. Study of the mathematical structure of spaces to which the generalized transport plans belong is an undergoing work.

1.2. **Quick introduction to optimal transportation**

Let Ω be an open bounded domain and let us consider the Monge problem of pushing one probability measure μ to another probability measure ν, through a transportation map which minimizes some cost. We assume the measures μ and ν to be absolutely continuous with respect to the Lebesgue measure, of densities ρ_0 and ρ_1, nonnegative on Ω. As probability measures we have

$$\int_{R^n} \rho_0(x)dx = \int_{R^n} \rho_1(x)dx = 1.$$

In application to image processing, these densities will correspond to gray levels, and in general this condition would not be satisfied. However means to cope with the general case exist [2]. A map $T : \mathbb{R}^n \to \mathbb{R}^n$ is a transfer map from ρ_0 to ρ_1 if for all bound $A \subset \mathbb{R}^n$,

$$\int_A \rho_1(x)dx = \int_{\{T(x)\in A\}} \rho_0(x)dx. \tag{1}$$

If T is a $\mathcal{C}^1$ mapping, then by a change of variables this is equivalent to

$$\det(\nabla T(x))\rho_1(T(x)) = \rho_0(x),$$

which is under-determined. Let $\Gamma(\rho_0, \rho_1)$ be the set of mappings T transfering ρ_0 on ρ_1. The L^p Kantorovich-Wassertein distance between ρ_0 and ρ_1 is then defined by

$$d_p(\rho_0, \rho_1)^p = \inf_{T\in\Gamma(\rho_0,\rho_1)} \int |T(x) - x|^p \rho_0(x)dx.$$

The L^p Monge-Kantorovitch problem (MKP) corresponds to find a mapping T such that this infimum is achieved.

In the case $p = 2$, the problem admits an unique solution (see e.g. Villani [12] page 66), which is the gradient of a convex fonctional from Ω to $\mathbb{R}$:

$$T(x) = \nabla\Psi(x).$$

The convex function Ψ is solution of Monge-Ampère equation:

$$\det(D^2\Psi)\rho_1(\nabla\Psi(x)) = \rho_0(x).$$

This equation being highly nonlinear, numerical methods to solve the MKP problem based on discretization of the Monge-Ampère equation have already been investigated [6–8, 10]. In application to image morphing problem, it is relevant to seek a time-dependent family of mappings $T(\cdot, t)$ transfering continuously ρ_0 to ρ_1. In [3] the authors introduced a fluid mechanics formulation of MKP, by adding a new dimension to the original problem (the time). The idea is to consider an arbitrary time interval $[0, t_m]$ and all functions $\rho(x, t) \geq 0$ and vector fields $v(x, t) \in \mathbb{R}^n$ solution of the continuity conditions with prescribed initial and final densities:

$$\partial_t\rho + \operatorname{div}(\rho v) = 0, \qquad \rho(x, 0) = \rho_0(x), \quad \rho(x, t_m) = \rho_1(x). \tag{2}$$

Then we have :

Theorem 1 (Benamou-Brenier)**.** *In the case $p = 2$ the KW distance between ρ_0 and ρ_1 is such that:*

$$d_2(\rho_0, \rho_1)^2 = \inf t_m \int_{\mathbb{R}^N} \int_0^{t_m} \rho(x,t)|v(x,t)|^2 dxdt$$

the infimum being taken on ρ, v verifying (2).

This approach is numerically solved using a saddle-point problem based on a augmented Lagrangian method. In the last few years, others applications of optimal transportation methods to image analysis have been proposed. For instance, instead of solving the saddle-point problem directly, Angenent et al. derived a novel gradient flow for the computation of the optimal transport map [1]. Unfortunately, all these methods do not take into account the physics of the represented objects in a image. Indeed, most of them compute an optimal way to deform one of the images to the other, by minimizing the mean distance between each displaced pixel. Therefore no energy is attached to image regions which could represent physical objects, and some regions can split into pieces along the optimal transportation way, as illustrated in the figure 1 Therefore, in this paper, we propose and study generalized optimal transport models which will attach a multiphysics model to the images to be interpolated or registrated.

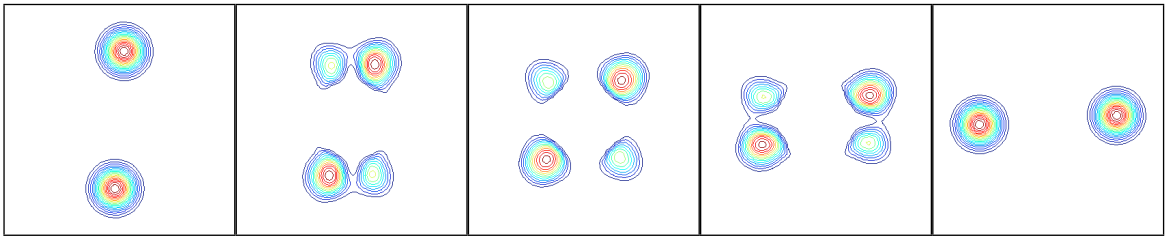


FIGURE 1. Plot of the isolevels of the density $\rho(t)$ along the optimal path computed with the approach of [3]. The physics included in the image is not conserved along this path and the mass can split.

The starting point of this article is the Benamou-Brenier approach which is a fluid mechanics formulation of the L^2-MKP [3]. We then consider $\Omega = (0,1)^2$ with periodic boundary conditions (i.e. the torus $\mathbb{R}^2/\mathbb{Z}^2$). Still considering a time interval $[0, t_m]$, we set $Q_m = \Omega \times (0, t_m)$. In order to minimize the energy under the constraint (2), we first introduce the new variables ρ and $m = \rho v$ (into which the constraint expresses linearly) and we consider the problem:

$$\inf_{\rho,m} t_m \int_{\mathbb{R}^N} \int_0^{t_m} \frac{|m|^2}{2\rho} dxdt$$

under the constraint

$$\partial_t \rho + \mathrm{div}(m) = 0, \qquad \rho(x,0) = \rho_0(x), \quad \rho(x,t_m) = \rho_1(x). \tag{3}$$

The remaining of this paper is organized as follows: in the next section, we consider a new optimal transport taking into account the surface tension. We will see that this additional term still allows some regions of image to split during the optimal transportation path. To tackle this issue, we then consider in section 3 a generalized optimal transport with incompressibility constraint. This first approach is nevertheless time consuming so that we also propose in section 4 a penalization method to impose incompressibility. We finally extend the penalization approach and propose energies modeling general physical constraints, by taking into account velocity priors such as translation or rigidity. The paper ends with a conclusion and some perspectives in section 5.

2. Generalized energy functionals: the surface tension case

Assume that the images to interpolate represent drop of a liquid into another. It is natural to associate to such system a surface tension energy, that we could try to minimize during an optimal transport path. Using a Korteweg-type energy, his amounts to consider

$$\inf_{\rho,m} t_m \int_{\mathbb{R}^N} \int_0^{t_m} \frac{|m|^2}{2\rho} + \chi|\nabla\rho| dxdt,$$

for a given surface tension coefficient $\chi > 0$, still under the constraint (3). The associated Lagrangian is given by

$$L(\phi,\rho,m) = \int_0^{t_m} \int_\Omega \frac{|m|^2}{2\rho} + \chi|\nabla\rho| - \partial_t\phi\rho - \nabla\phi\cdot m dxdt - \int_\Omega \phi(0,x)\rho_0(x) - \phi(t_m,x)\rho_1(x)dx. \tag{4}$$

Given two densities ρ_0 et ρ_1, the minimization problem is equivalent to the saddle-point problem:

$$\inf_{\rho,m}\sup_{\phi} L(\phi,\rho,m).$$

Arguing as in [3] we introduce dual variables $(a,b) \in \mathbb{R}\times\mathbb{R}^d$ such that

$$\frac{|m|^2}{2\rho} = \sup_{(a,b)\in K} a(t,x)\rho(t,x) + b(t,x)\cdot m(t,x),$$

with

$$K = \left\{(a,b) : \mathbb{R}\times\mathbb{R}^2 \to \mathbb{R}\times\mathbb{R}^2, \quad a + \frac{1}{2}|b|^2 \le 0 \text{ on } \mathbb{R}\times\mathbb{R}^2\right\}.$$

As well we introduce a dual variable $c \in \mathbb{R}^d$, associated to the surface tension term, using

$$|\nabla\rho| = \sup_{|c|\le 1} c\cdot\nabla\rho.$$

For sake of clarity, we set $\mu = (\rho, m)$ and $q = (a,b)$, and introduce the support function of K, F such that $F(q) = 0$ for $q \in K$ and $F(q) = +\infty$ otherwise. Therefore we have $\frac{|m|^2}{2\rho} = \sup_{q\in K} \mu\cdot q = \sup_q -F(q) + \mu\cdot q$. Likewise we introduce the indicator function of the closed unit ball of $\mathbb{R}^2$ by setting $H(c) = 0$ if $|c| \le 1$ and $+\infty$ otherwise. At last we set

$$G(\phi) = \int_\Omega \phi(0,x)\rho_0(x) - \phi(t_m,x)\rho_1(x)dx.$$

Still following [3], we show that our saddle point problem can be written as

$$\sup_{\mu}\inf_{\phi,q,c} F(q) + G(\phi) + H(c) + \langle\mu, \nabla_{t,x}\phi - q\rangle - \chi\langle c, \nabla\rho\rangle, \tag{5}$$

where the brackets stand for the $L^2(Q_m)$ scalar product, the variables μ, q are taken in $L^2(Q_m)^{d+1}$, c in $L^2(Q_m)^d$ and ϕ in $H^1(Q_m)$. Indeed we first have

$$-\inf_{\rho,m}\sup_{\phi} L(\phi,\rho,m) = \sup_{\rho,m}\inf_{\phi} G(\phi) + \int_0^{t_m}\int_\Omega \mu\cdot\nabla_{t,x}\phi - \sup_{q\in K}\mu\cdot q - \chi\sup_{|c|\le 1} c\cdot\nabla\rho\, dxdt.$$

Then we note that

$$\int_0^{t_m}\int_\Omega \sup_{q\in K}\mu\cdot q dxdt = \sup_{q\in L^2(Q_m)^{d+1}} -F(q) + \int_0^{t_m}\int_\Omega \mu\cdot q dxdt = \sup_{q\in L^2(Q_m)^{d+1}} -F(q) + \langle\mu, q\rangle,$$

and likewise that

$$\int_0^{t_m}\int_\Omega \chi \sup_{|c|\le 1} c\cdot\nabla\rho\, dxdt = \sup_{c\in L^2(Q_m)^d} -H(c)+\chi\int_0^{t_m}\int_\Omega c\cdot\nabla\rho dxdt = \sup_{c\in L^2(Q_m)^d} -H(c)+\chi\left\langle c,\nabla\rho\right\rangle,$$

which gives (5). Introducing $D(c)=\chi(\mathrm{div}(c),0)$, we can rewrite expression (5) as:

$$\sup_{\mu}\inf_{\phi,q,c} F(q)+G(\phi)+H(c)+\left\langle \mu,\nabla_{t,x}\phi-q+D(c)\right\rangle, \tag{6}$$

where

$$\int_0^{t_n}\int_{\partial\Omega}\rho(t,x)c(t,x)dt$$

vanishes since we consider periodic boundary conditions in space. We now aim at finding a saddle-point of this problem which corresponds to a standard form of [9] in order to apply augmented Lagrangian techniques. The formal optimal condition for this problem are:

$$\begin{cases} \partial_t\phi+\frac{|m|^2}{2\rho^2}+\chi\,\mathrm{div}(c)=0 & \text{in } [0,t_m]\times\Omega\\ \partial_t\rho+div(m)=0 & \text{in } [0,t_m]\times\Omega\\ \frac{m}{\rho}=\nabla\phi & \text{in } [0,t_m]\times\Omega\\ \rho(0,.)=\rho_0 & \text{in } \Omega\\ \rho(t_m,.)=\rho_1 & \text{in } \Omega \end{cases}$$

Observing that the variable m can be eliminated, the optimality conditions can be rewritten in term of ρ, ϕ and c as:

$$\begin{cases} \partial_t\phi+\frac{|\nabla\phi|^2}{2}+\chi\,\mathrm{div}(c)=0 & \text{in } [0,t_m]\times\Omega\\ \partial_t\rho+div(m)=0 & \text{in } [0,t_m]\times\Omega\\ \rho(0,.)=\rho_0 & \text{in } \Omega\\ \rho(t_m,.)=\rho_1 & \text{in } \Omega \end{cases} \tag{7}$$

We then define the augmented Lagrangian by introducing $r>0$:

$$L_r(\phi,q,c,\mu)=F(q)+G(\phi)+H(c)+\left\langle\mu,\nabla_{t,x}\phi-q+D(c)\right\rangle+\frac{r}{2}\left\langle\nabla_{t,x}\phi-q+D(c),\nabla_{t,x}\phi-q+D(c)\right\rangle. \tag{8}$$

2.1. **The algorithm**

We consider the following iterative algorithm to compute this saddle point numerically: This algorithm builds from $(\phi^{n-1},q^{n-1},\mu^n,c^{n-1})$ the next iterate.

Step 1: $\phi^n=\arg\min L_r(\,.\,,q^{n-1},c^{n-1},\mu^n)$
Step 2: $c^n=\arg\min L_r(\phi^n,q^{n-1},\,.\,,\mu^n)$
Step 3: $q^n=\arg\min L_r(\phi^n,\,.\,,c^n,\mu^n)$
Step 4: $\mu^{n+1}=\arg\max L_r(\phi^n,q^n,c^n,\,.\,)$

Let us describe which Euler-Lagrange equations are associated to these optimization problems:

Step 1: Differentiation of the Lagrangian with respect to ϕ reads formally

$$dL_r(\phi,q,c,\mu)(\psi)=G(\psi)+\left\langle\mu,\nabla_{t,x}\psi\right\rangle+r\left\langle\nabla_{t,x}\phi-q+D(c),\nabla_{t,x}\psi\right\rangle$$

Taking first ψ smooth and with compact support in Q_m, so that $G(\psi)=0$, we get the PDE to be verified by the minimizer ϕ^n in Q_m:

$$-r\Delta_{t,x}\phi^n=\mathrm{div}_{t,x}(\mu^n-rq^{n-1}+rD(c^{n-1})).$$

Then we derive the boundary conditions: for all ψ,

$$G(\psi) + \int_{\partial Q_m} N_{t,x} \cdot \left(\mu^n + r(\nabla_{t,x}\phi^n - q^{n-1} + D(c^{n-1}))\right) \psi d\Sigma(x,t) = 0,$$

which means

$$\int_\Omega \psi(0,x)\rho_0(x) - \psi(t_m,x)\rho_1(x) + \left[\rho^n(t_m,x) + r\left(\partial_t\phi(t_m,x) - a^{n-1}(t_m,x) + \chi \operatorname{div}(c^{n-1}(t_m,x))\right)\right]\psi(t_m,x)$$
$$-\left[\rho^n(0,x) + r\left(\partial_t\phi(0,x) - a^{n-1}(0,x) + \chi \operatorname{div}(c^{n-1}(0,x))\right)\right]\psi(0,x)dx + \int_0^{t_m}\int_{\partial\Omega} n\cdot(m^n + r(\nabla\phi - b^{n-1}))\psi(x,t)d\sigma dt = 0$$

Since we assumed periodic boundary conditions in space, the last integral vanishes (we could also have expressed there a Neumann condition), and we get as initial and final conditions on ϕ^n:

$$r\partial_t\phi^n(0,x) = r(a^{n-1}(0,x) - \chi \operatorname{div}(c^{n-1}(0,x))) + \rho_0(x) - \rho^n(0,x)$$
$$r\partial_t\phi^n(t_m,x) = r(a^{n-1}(t_m,x) - \chi \operatorname{div}(c^{n-1}(t_m,x))) + \rho_1(x) - \rho^n(t_m,x).$$

Step 2: Similarly, the differentiation of the Lagrangian with respect to c reads formally

$$dL_r(\phi,q,c,\mu)(\psi') = H(\psi') + \langle \rho, \chi \operatorname{div}(\psi')\rangle + r\left\langle \partial_t\phi - a + \chi \operatorname{div}(c), \chi \operatorname{div}(\psi')\right\rangle.$$

Taking first ψ' with compact support in Q_m and values in the unit ball B, so that $H(\psi') = 0$, we get the PDE to be verified by the minimizer $\tilde{c}$ in Q_m:

$$r\chi\nabla \operatorname{div}(c) = -\nabla(\rho^{n-1} + r\partial_t\phi^n - ra^{n-1}).$$

The operator ∇ div being difficult to deal with, we rather consider Uzawa's algorithm which gives, for an algorithmic timestep $\tau > 0$:

$$c^n = P_B(c^{n-1} - \tau\nabla_c L_r(\phi^n, q^{n-1}, c^{n-1}, \mu^n)),$$

where P_B is the projection on the unit ball B. Hence the computation of c^n is done in two steps. The first update is given with the following explicit scheme:

$$\tilde{c} = c^{n-1} + \tau\nabla(\rho^n + r(\partial_t\phi^n - ra^{n-1} + \chi \operatorname{div}(c^{n-1}))),$$

and c^n is finally obtained as the projection of $\tilde{c}$ onto the unit ball so that $c^n = \tilde{c}$ if $|\tilde{c}| \le 1$ and $c^n = \frac{\tilde{c}}{|\tilde{c}|}$ otherwise.

Step 3: Is identical close to step 2 of [3]. The coefficients a and b being first updated as:

$$\tilde{a} = \partial_t\phi^n + \chi \operatorname{div}(c^n) + \frac{\rho^n}{r}, \quad \tilde{b} = \nabla\phi^n + \frac{m^n}{r}.$$

The vector $\tilde{q} = (\tilde{a}, \tilde{b})$ being projected onto K to obtain a^n and b^n.

Step 4: Corresponds to the update of μ through Uzawa algorithm involving the algorithmic timestep τ (taken as $\tau = r$ in [3]):

$$\rho^{n+1} = \rho^n + r(\partial_t\phi^n - a^n + \chi \operatorname{div}(c^n)), \quad m^{n+1} = m^n + r(\nabla\phi^n - b^n).$$

2.2. **Numerical experiments**

To be able to produce numerical estimates and also for the practical purpose of stopping the computation we need to define a convergence criterium. The optimality conditions (7) are useful for that purpose. We can indeed use the residual of the Hamilton-Jacobi equation, namely

$$res^n = \partial_t \phi^n + \frac{|\nabla \phi^n|^2}{2} + \chi \operatorname{div}(c^n)$$

which is a by-product of the algorithm. This quantity converges to 0 as we approach the solution of problem. The normalized convergence criterium used is

$$crit^n = \sqrt{\frac{\int_0^{t_m} \int_\Omega \rho^n |res^n|}{\int_0^{t_m} \int_\Omega \rho^n |\nabla \phi^n|^2}}.$$

Concerning the algorithmic timestep τ used in the updates of c and μ, we have to choose it according to the value of χ. We take in practice $\tau = r = 0.1$ if $\chi < 1$ and $\tau = r = \frac{1}{10\chi}$ otherwise.
We present in Figures 2 and 3 numerical tests performed on the unit square with periodic boundary conditions in space. The space-time domain is discretized using a regular $64 \times 64 \times 64$ grid. In order to get a good convergence of the algorithm, we impose a number of iterations about 3000.
Figure 2 displays the optimal transportation to the gaussian for the classical Benamou-Brenier algorithm i.e for $\chi = 0$. Because of the periodic boundary conditions, the optimal transfer is not a simple translation (see [3]) but a splitting of the gaussian function into two pieces. Figure 3 shows the optimal way for the same gaussian when we associate the surface tension to the kinetic energy. We remark that this generalization still allows to the gaussian to split into two. This stems from the fact that $|\nabla \rho|$ measures only the length of level sets of the density and so minimizing such quantity leads to a mass flattening.

Remark 2.1. Let us explain more precisely what is occuring on a simple but illustrative example. Assume for instance that the density ρ is a smooth variation from a constant density ρ_i inside a closed curve of length ℓ to a constant density ρ_o outside that curve. Let ϕ be a signed distance function to this curve, negative inside. One way to express this variation is using a smooth increasing function H such that $H = 0$ on $(-\infty, -1)$, $H = 1$ on $(1, +\infty)$. Let ζ be its derivative (which is a cut-off function). Set $\rho = \rho_o H(\frac{\phi}{\varepsilon}) + \rho_i(1 - H(\frac{\phi}{\varepsilon}))$ with $\varepsilon > 0$ which gives a transition of length 2ε between the two constants values. The energy considered so far was $\int_\Omega |\nabla \rho| dx = (\rho_o - \rho_i) \int_\Omega |\nabla \phi| \frac{1}{\varepsilon} \zeta(\frac{\phi}{\varepsilon}) dx$. A quick analysis which is standard in the level-set setting shows that the last integral is equivalent to ℓ as $\varepsilon \to 0$. Therefore the energy considered is essentially $\ell \times (\rho_o - \rho_i)$. Creating two copies of the same curve with half jump would therefore give the same energy. This is what we observe on our tests. However, in contrast, in the incompressible case, the conservation equation being a transport equation, area of level sets of ρ should be conserved. This would prevent this decomposition to occur, and the penalization of $|\nabla \rho|$ therefore would, as expected, prevent the density from splitting.

Therefore, in order to preserve the specificity of initial data, we consider in the next sections others energies which will allow a better representation to physical objects of images.

In Figure 4 we check the convergence of iteration process for this algorithm for $\chi = 0$ and $\chi = 0.1$. We consider runs over large number of iterations and we analyze the Wassertein distance and the L^2-distance for the surface tension:

$$d_{wass} = \int_\Omega \int_0^T \frac{|m|^2}{2\rho} \, dxdt, \qquad d_2 = \int_\Omega \int_0^T |\nabla \rho| \, dxdt. \tag{9}$$

Numerical tests indicate a convergence of our new functional. Note that the presence of the surface tension leads to a deterioration of the convergence rate in comparison to Benamou-Brenier algorithm.

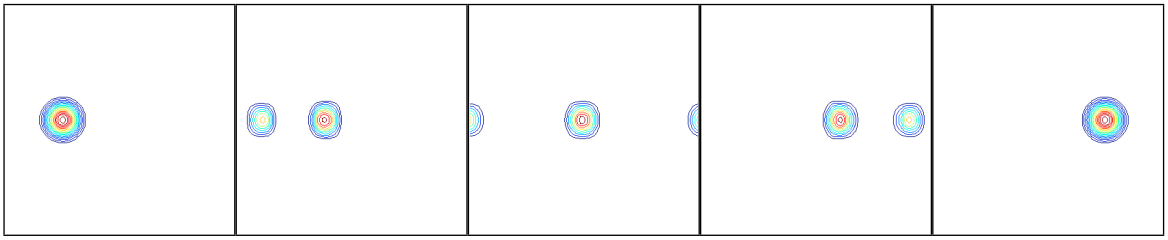


FIGURE 2. Plot of the isolevels of the density $\rho(t)$ along the optimal path computed with the approach of [3].

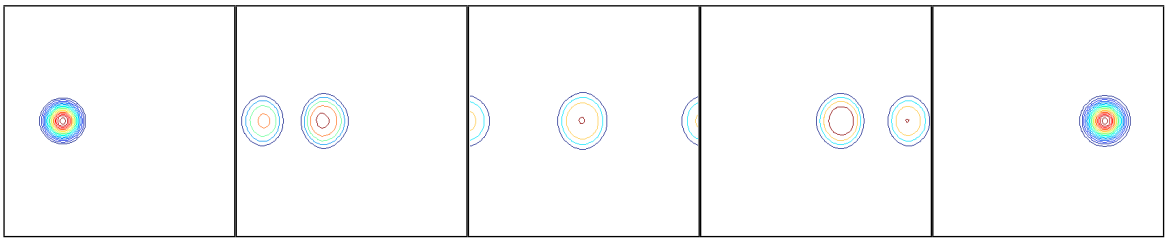


FIGURE 3. Plot of the isolevels of the density $\rho(t)$ along the optimal path computed with the surface tension penalization and $\chi = 0.05$. The new term does not prevent from mass splitting and also flatten the mass.

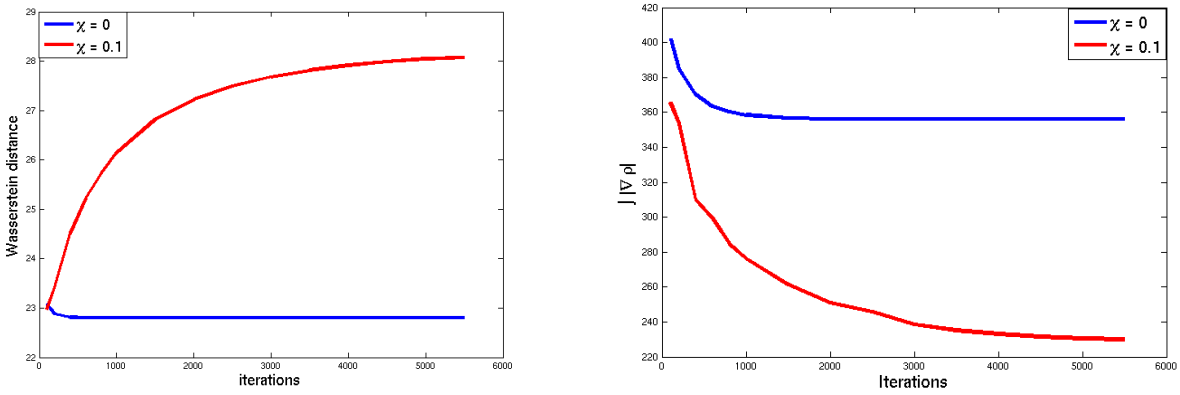


FIGURE 4. Convergence of the iterative process with (in red) and without (in blue) the surface tension term.

3. Optimal transportation with incompressibility constraint

The aim of this section is to generalize the optimal transportation by taking into account incompressibility condition. Hence, the functional to minimize is still given by:

$$\inf_{\rho,m,v} t_m \int_{\mathbb{R}^N} \int_0^{t_m} \frac{|m|^2}{2\rho}\, dxdt, \tag{10}$$

under the well-known constraint (3). The difference is that new constraints are added, namely:

$$\begin{cases} \rho v & = m, \\ \operatorname{div} v & = 0, \end{cases} \tag{11}$$

that link the velocity to the previous state variables and impose the incompressibility of the flow. In the case where $\operatorname{div} v = 0$, the constraint (3) becomes

$$\partial_t \rho + v \cdot \nabla \rho = 0, \tag{12}$$

where the solution of this continuity equation is given by

$$\rho(t,x) = \rho_0(Y(t,x)), \tag{13}$$

where Y designs a characteristic curve. So, in particular the measure of level sets of ρ should be conserved. This will be validated in our numerical tests (see figure 7).

Let us remark that the incompressibility constraint involves reintroducing the velocity v and linking it with the variables (ρ, m). As a consequence, the functional will not be convex anymore with respect to (ρ, m, v) so that we will only be able to compute a local minima of the problem. One could for instance extend the formulation with the surface tension and minimize $|\nabla m|$, without reintroducing the velocity v in order to stay convex in (ρ, m). However, even considering piecewise constant functions ρ, the obtained solution is not satisfying in practice with such constraint on m. As above, we consider the augmented lagrangian method. Formulation of the saddle point problem being similar, we just give the main steps of the algorithm.

First of all, we introduce two new dual lagrangian variables to deal with the two new constraints defined in (11) Introducing $s = (s_1, s_2)$, the constraint $m = \rho v$ will be imposed by maximizing over s:

$$\sup_s (\rho v - m)s = \sup_s \langle \mu, E(v,s) \rangle\,, \tag{14}$$

where

$$E(v,s) = \begin{bmatrix} < v, s > \\ -s \end{bmatrix}.$$

Similarly, the incompressibility constraint $\operatorname{div} v = 0$ will be treated with a scalar variable p. It can be noted that:

$$\int_\Omega p \operatorname{div} v dx = -\int_\Omega v \cdot \nabla p dx,$$

the boundary conditions vanishing, as we consider either a periodic domain or dirichlet conditions (with a slipping flow on the frontier) for the velocity v. Hence, the associated Lagrangian reads:

$$L(\phi,\rho,m,v,s,p) = \int_0^{t_m} \int_\Omega \frac{|m|^2}{2\rho} + p \operatorname{div} v + \langle \mu, E(v,s) \rangle - \partial_t \phi \rho - \nabla \phi \cdot m dx dt - \int_\Omega \phi(0,x)\rho_0(x) - \phi(t_m,x)\rho_1(x) dx. \tag{15}$$

As above, the minimization problem consists in solving the following saddle-point problem

$$\inf_{\rho,m,v} \sup_{\phi,s,p} L(\phi,\rho,m,v,s,p), \tag{16}$$

for two densities given ρ_0 and ρ_1.
Arguing as previously and with the same notations, this saddle point problem can be written as

$$\sup_{\mu,v}\inf_{\phi,q,s,p} F(q)+G(\phi)+H(s)+\langle\mu,\nabla_{t,x}\phi-q-E(v,s)\rangle-p\,\mathrm{div}\,v. \tag{17}$$

We then consider, for $r>0$, the following augmented Lagrangian:

$$\begin{aligned} L_r(\phi,q,\mu,v,s,p) &= F(q)+G(\phi)+H(s)+\langle\mu,\nabla_{t,x}\phi-q-E(v,s)\rangle-p\,\mathrm{div}\,v \\ &+ \frac{r}{2}\langle\nabla_{t,x}\phi-q-E(v,s),\nabla_{t,x}\phi-q-E(v,s)\rangle+\frac{r}{2}\langle\nabla p,\nabla p\rangle, \end{aligned}$$

where another augmented term as been added to deal with the incompressibility constraint.

3.1. **Algorithm**

Therefore, the algorithm can be summarized as follow: from $(\phi^n,q^{n-1},\mu^n,v^n,s^{n-1},p^{n-1})$, the next iteration is given by

Step 1 $s^n=\arg\min L_r(\phi^{n-1},q^{n-1},\mu^n,v^n,\,.\,,p^{n-1})$
Step 2 $\phi^n=\arg\min L_r(\,.\,,q^{n-1},\mu^n,v^n,s^n,p^{n-1})$
Step 3 $q^n=\arg\min L_r(\phi^n,\,.\,,\mu^n,v^n,s^n,p^{n-1})$
Step 4 $p^n=\arg\min L_r(\phi^n,q^n,\mu^n,v^n,s^n,\,.\,)$
Step 5 $v^{n+1}=\arg\max L_r(\phi^n,q^n,\mu^n,\,.\,,s^{n+1},p^n)$
Step 6 $s^n=\arg\min L_r(\phi^n,q^n,c^n,\mu^n,v^{n+1},\,.\,,p^n)$
Step 7 $\mu^{n+1}=\arg\max L_r(\phi^n,q^n,\,.\,,v^{n+1},s^n,p^n)$

Let us give an idea of resolutions to these optimization problems.
Step 1 The differentiation of the augmented Lagrangian with respect to s gives

$$dL_r(\phi,q,\mu,v,s,p)(s')=H(s')+\left\langle\mu+r\left(\nabla_{t,x}\phi-q-E(v,s)\right),-\begin{bmatrix}<v,s'>\\ -s'\end{bmatrix}\right\rangle. \tag{18}$$

Taking s' such that $H(s')=0$, the minimizer s^n verifies:

$$\left(\rho+r(\partial_t\phi-a-<v,s^n>)\right)v=\left(m+r(\nabla_x\phi-b+s^n)\right), \tag{19}$$

which gives:

$$r\lambda(Id+v\otimes v)s^n=\left(\rho+r(\partial_t\phi-a)\right)v-\left(m+r(\nabla_x\phi-b)\right), \tag{20}$$

so the update is:

$$s^n=(Id+v^n\otimes v^n)^{-1}\Big(\frac{\rho^nv^n-m^n)}{r}+\left(\nabla_t\phi^n-a^{n-1}\right)v^n-\nabla_x\phi^n+b^{n-1}\Big), \tag{21}$$

Step 2 Differentiating the augmented lagrangian with respect to ϕ, the minimizer must be verify the following Poisson equation:

$$-r\Delta_{t,x}\phi^n=\mathrm{div}_{t,x}\left(\mu^n-rq^{n-1}-rE(v^n,s^n)\right), \tag{22}$$

with the following boundary conditions:

$$\begin{aligned} r\partial_t\phi^n(0,x)&=r(a^{n-1}(0,x)+v^n(0,x)\cdot s^n(0,x))+\rho_0(x)-\rho^n(0,x)\\ r\partial_t\phi^n(t_m,x)&=r(a^{n-1}(t_m,x)+v^n(t_m,x)\cdot s^n(t_m,x))+\rho_1(x)-\rho^n(t_m,x). \end{aligned}$$

Step 3 To update q, we first compute

$$\tilde{a} = \partial_t \phi^n - v^n \cdot s^n + \frac{\rho^n}{r}, \qquad \tilde{b} = \nabla \phi^n + s^n + \frac{m^n}{r}, \tag{23}$$

and we next project the vector $(\tilde{a}, \tilde{b})$ onto K to get q^n.

Step 4 The differentiation of the augmented Lagrangian with respect to p gives simply $\Delta_x p = -\frac{1}{r} \operatorname{div} v$. We therefore have to solve a $2D$ Poisson equation for each time step t to obtain p^n.

Step 5 The differentiation of the augmented Lagrangian with respect to v gives

$$dL_r(\phi, q, \mu, v, s, p)(v') =< -\rho s - \nabla p + r(\partial_t \phi - q - \lambda v \cdot s), v' > . \tag{24}$$

We then perform a gradient descent over v using this differentiation:

$$v^{n+1} = v^n + \tau \left(-\rho^n s + \nabla p^n + r(\partial_t \phi^n - q^n - v^n \cdot s^n) s^n \right), \tag{25}$$

where $\tau > 0$ is a time step.
Step 6 Identical to step 1. Second update of s to impose the constraint $m = \rho v$ after updating v.

Step 7 We finish by updating μ as follow:

$$\rho^{n+1} = \rho^n + \tau(\partial_t \phi^n - a^n - < v^n, s^n >), \qquad m^{n+1} = m^n + \tau(\nabla \phi^n - b^n + s^n). \tag{26}$$

3.2. **Numerical results**

Note that from now on, we will consider Dirichlet conditions in the experiments, so that periodicity will not be involved anymore in the eventual mass splitting. In Figure 5, we show that such conditions prevents the method of [3] from mass splitting in the case of a single Gaussian to transport.

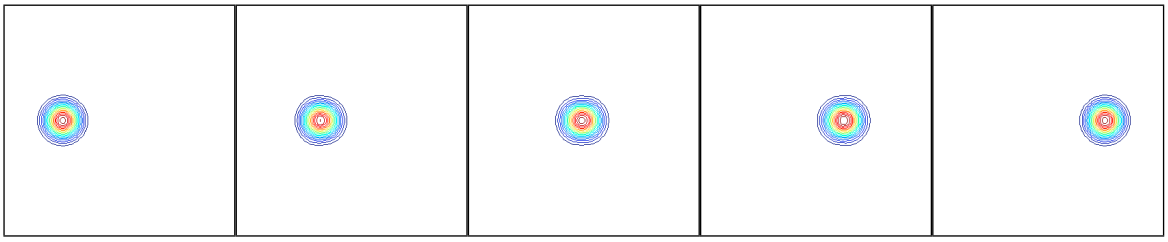
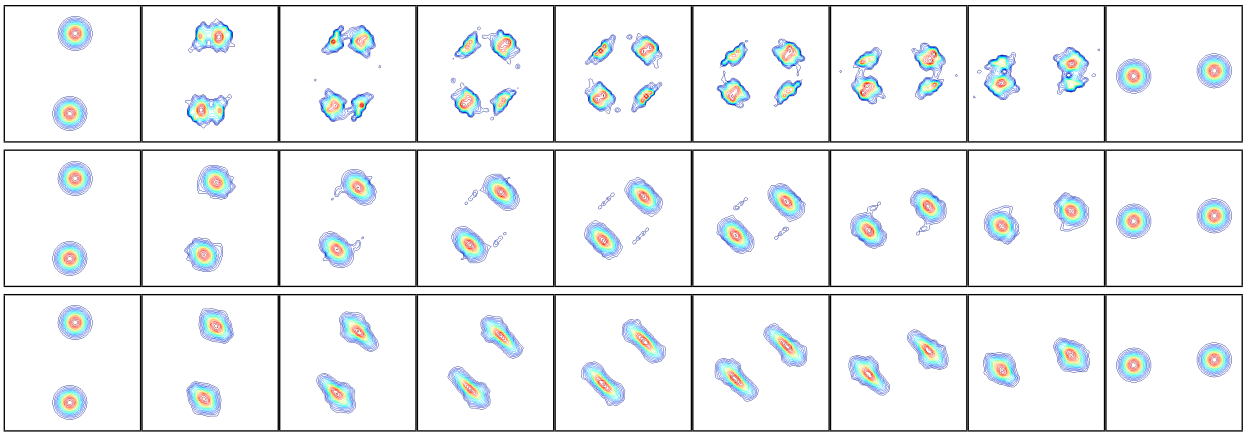


FIGURE 5. Plot of the isolevels of the density $\rho(t)$ along the optimal path computed with the method of [3] and Dirichlet conditions, which prevents from mass splitting.

We therefore consider a more complex example including two Gaussian in each density, as illustrated in the introduction of the paper in Figure 1. The parameters of the algorithm are here set to $\tau = r = 1$. It is important to recall that the obtained results depend on the initialization, as the model is no more convex. As we do not want to use complex initialization models, we observed that a slight modification of the energy allows converging to more suitable local minima. In practice, the Wasserstein distance is minimized in parallel with the verification of the constraints. By multiplying the Wasserstein distance defined in (10) with a parameter $0 < \beta < 1$, we can allow the constraints to be fulfilled in priority, the Wasserstein distance being minimized

latter in the iterations. The influence of the parameter β is illustrated in figure 6. It shows that even if the algorithm is parameter free, the non convexity implies an appropriate tuning of β to obtain a completion of the incompressibility constraint.

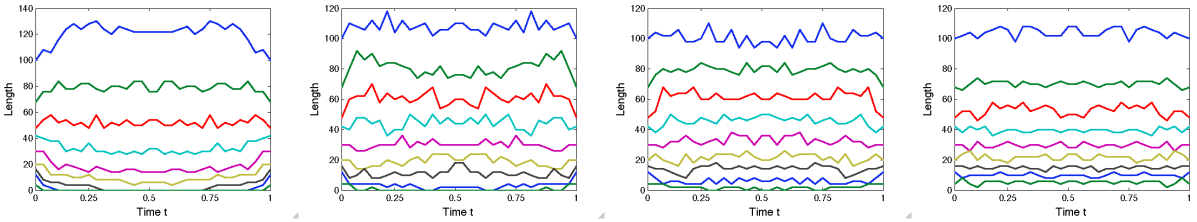


$\rho(0) = \rho_0$ $\rho(1/8)$ $\rho(1/4)$ $\rho(3/8)$ $\rho(1/2)$ $\rho(5/8)$ $\rho(3/4)$ $\rho(7/8)$ $\rho(1) = \rho_1$

FIGURE 6. Two gaussian experiments with incompressibility constraint. Plot of the isolevels of the density $\rho(t)$ along the optimal path computed with the incompressibility constraint. The three lines correspond to experiments realized with $\beta = 1$, $\beta = 10^{-2}$ and $\beta = 10^{-3}$.

An important property of incompressible flow is that the level lines of the transported density must be preserved along time. Hence, we give in Figure 7 a sketch of the evolution of the length of the upper level lines of the density ρ obtained for different values of β. It shows that decreasing β allows reaching constant level lines. In addition with the norm of the divergence of the velocity, given in Figure 8, one can see that the process almost allows approximate computation of divergence free transports.

Finally note that the convergence criteria obtained from the Hamilton-Jacobi residual is here no more pertinent to check the convergence of the process. As consequence, we just use a threshold on the norm between successive values of the variables (ρ, m, v).



Method of [3] $\beta = 1$ $\beta = 10^{-2}$ $\beta = 10^{-3}$

FIGURE 7. Two gaussian experiments with incompressibility constraint. Evolution of the length of the upper level lines of the estimated density along time t: $|\rho(x,t) > i/10|$, for $i = 1 \cdots 9$. The left plot is the optimal transport of [3] shown in Figure 1, and the other ones correspond to the proposed approach with an incompressibility constraints and decreasing values of β. Decreasing β allows better completion of the constraint since the level lines are more and more preserved along the computed path.

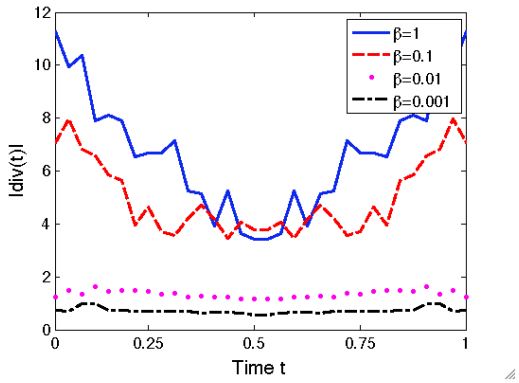


FIGURE 8. Two gaussian experiments with incompressibility constraint. The evolution of the norm of the divergence of the estimated velocity v is given for decreasing values of β.

Remark

Note that the constraint $m = \rho v$ can be relaxed by adding the term $\lambda|m - \rho v|$ to the energy (28) to minimize, with $\lambda > 0$. This can be done using the dual formulation:

$$\lambda \max_{s \in Z} (m - \rho v)s,$$

which involve a projection P_Z of the previously introduced variable s in the unit ball at each update:

$$P_Z(s) = \begin{cases} s & \text{if } ||s|| < 1 \\ s/||s|| & \text{otherwise} \end{cases} \tag{27}$$

As the problem is separately convex in (ρ, m) and in v, such penalization allows to ensure the convergence of the minimization process [11], whereas the convergence is not guaranteed with the strict constraint. However, we observed in all our experiments that the dual formulation of the constraint does not avoid the algorithm from converging (see Figure 9) and it also prevents from fixing an additional parameter λ.

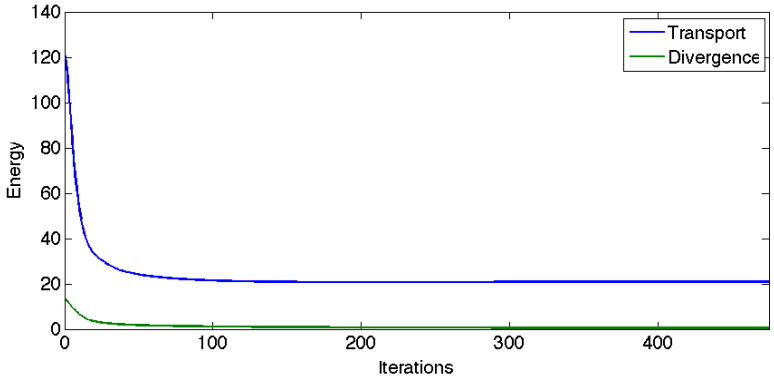


FIGURE 9. Plot of the transport energy (in blue) and the norm of the divergence of v (in green) along the iterations. The process numerically converges.

4. Regularizing the transport map through penalization

In this section, we generalize optimal transportation by taking into account others physicals criterion. Another way to treat the physical conditions consists in penalizing a specific norm derived from the velocity v. This is a relaxed approach with respect to the previous lagrangian formulation of the constraints. Denoting as $v = (v_1, v_2)$, we can mention the following penalized models $M(v)$:

- Incompressibility with the norm of the divergence of v, $M(v) = ||\operatorname{div} v||^2$
- Rigidity with the norm of the deformation tensor of v, $M(v) = \left|\left|\frac{\nabla v + \nabla v^t}{2}\right|\right|^2$
- Translation with the norm of the gradients of v: $M(v) = \sum_i ||\nabla v_i||^2$, that allows obtaining continuous $\mathcal{C}^1$ velocity field v
- Piecewise translations with the total variation of v, $M(v) = \sum_i |\nabla v_i|$

Such approach will lead to minimize

$$\inf_{\rho,m,v} t_m \int_{\mathbb{R}^N} \int_0^{t_m} \left(\frac{|m|^2}{2\rho} + \alpha M(v) \right) dxdt, \tag{28}$$

under the constraints (3) and $m = \rho v$ and where the parameter $\alpha > 0$ monitors the penalization weight. Hence, depending on what should be preserved during optimal transportation on specific data, one can choose among these formulations. Of course, it is also possible to associate others terms to the kinetic energy but globally the methods of numerical resolutions remain the same.

4.1. The algorithm

Let us detail the modifications involved with such penalization model $M(v)$. The associated lagrangian is

$$L(\phi, \rho, m, v, s) = \int_0^{t_m} \int_\Omega \frac{|m|^2}{2\rho} + \alpha M(v) + \langle \mu, E(v,s) \rangle - \partial_t \phi \rho - \nabla \phi \cdot m dx dt - \int_\Omega \phi(0,x)\rho_0(x) - \phi(t_m, x)\rho_1(x) dx. \tag{29}$$

As in the previous section, we rewrite the saddle-point problem and next consider the associated augmented lagrangian. Globally, the algorithm to solve the saddle problem is similar to the previous one. The step 4 is removed as there is no longer a dual variable p for the divergence constraint, and the update of v in step 5 becomes:

$$v^{n+1} = v^n + \tau \left(-\rho^n s - \alpha \partial M^* \partial M v^n \right) + r(\partial_t \phi^n - q^n - \lambda v^n \cdot s^n) s^n \big), \tag{30}$$

where ∂M^* is the adjoint operator of ∂M, the linear tangent operator of M. We can now give example of $\partial M^* \partial M$ for the different models previously introduced.

- Incompressibility: $M(v) = ||\operatorname{div} v||^2$, so that $\partial M^* \partial M v = -\nabla(\operatorname{div} v)$. With respect to the previous strict constraint on incompressibility, we no more have to solve the Poisson equation $-\Delta p = \operatorname{div} v$ for each time step t, which makes the approach faster. Nevertheless, the strong incompressibility constraint without penalization allows a better conservation to level sets as can be seen in Figure 11.
- Rigidity: $M(v) = \left|\left|\frac{\nabla v + \nabla v^t}{2}\right|\right|^2$, $\partial M^* \partial M v = -[v_{1xx}; v_{2yy}] - [v_2; v_1]_{xy} - \frac{[v_2; v_1]_{xy}}{2}$
- Translation: $M(v) = \sum_i ||\nabla v_i||^2$, and $\partial M^* \partial M v = -\Delta_x v$
- Piecewise translations: $M(v) = \sum_i |\nabla v_i|$. The dual formulation of the total variation is here needed to minimize such term. We recall that

$$|\nabla v_i| = \sup_{\xi_i \in Z} \nabla v_i \xi_i = \sup_{\xi_i \in Z} -v_i \operatorname{div}(\xi_i) \tag{31}$$

where $\xi_i = (\xi_i^1, \xi_i^2) \in Z = \{z = (z^1, z^2), |z| \leq 1\}$. We then have an additional step to update this new dual variable ξ_i with $\xi_i^{n+1} = \xi_i^n + \tau \nabla v_i$ and finally obtain $\partial M^* \partial M v_i = -\operatorname{div}(\xi_i)$.

4.2. **Numerical results**

For the penalization approach, we chose large values of α (typically 10^3 or more) to see the desired effect on the estimations. The parameters are then chosen as $\tau = r = \frac{1}{10\alpha}$ to have a stable process. As in the previous section, the residual of the Hamilton-Jacobi equations is not pertinent to check the convergence of the process. First of all, we compare in Figure 10 the results obtained with the incompressible, translation and rigid penalizations.

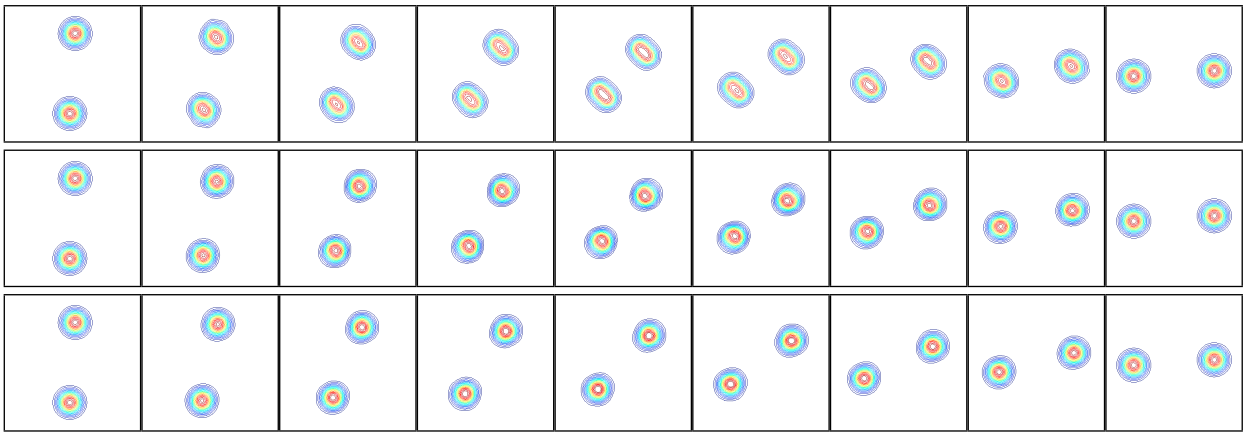
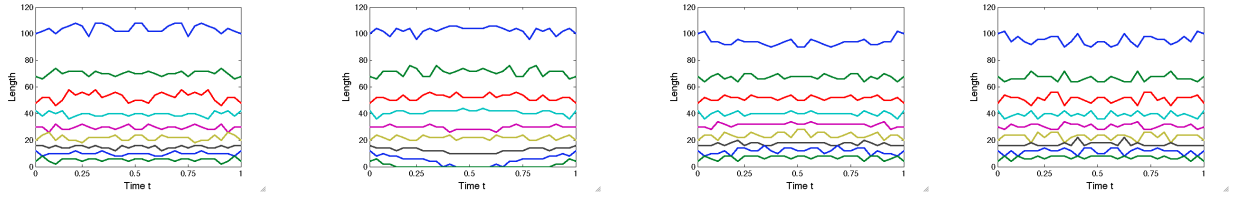


FIGURE 10. Two gaussian experiments with penalization. Plot of the isolevels of the density $\rho(t)$ along the optimal path computed with the different penalization models. The top line is realized with incompressibility penalization and the second with a translation penalization (through the minimization of the L^2 norm of each component of the velocity field) and the last one with a rigid penalization.

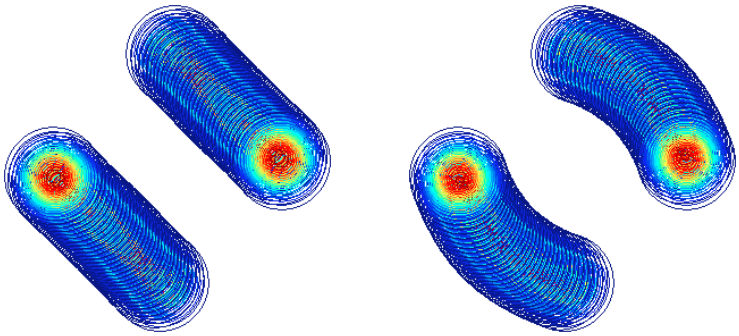
With respect to the previous estimation with the incompressible constraints, the results obtained with the incompressible penalization are visually more satisfying as the instabilities that were appearing (Figure 6) with large values of β are not present in this relaxed approach. The previous constraint method also presents the drawback to be more time consuming, as 2D Poisson equations have to be solved at each iteration. Nevertheless, it can be observed in Figure 11 that the constrained approach allows a better conservation of the level set measures than the incompressible penalization.



Incompressible constraint | Incompressible penalization | Translation penalization | Rigid penalization

FIGURE 11. Two gaussian experiment. Evolution of the length of the upper level lines of the estimated density along time t: $|\rho(x,t) > i/10|$, for $i = 1 \cdots 9$. The left plot is obtained with the incompressible constraint (already seen in Figure 7), the middle one is with an incompressible penalization and the right one corresponds to a rigid penalization. Even if the penalization methods gives better visual results, the level lines are better preserved with the strong incompressible constraint.

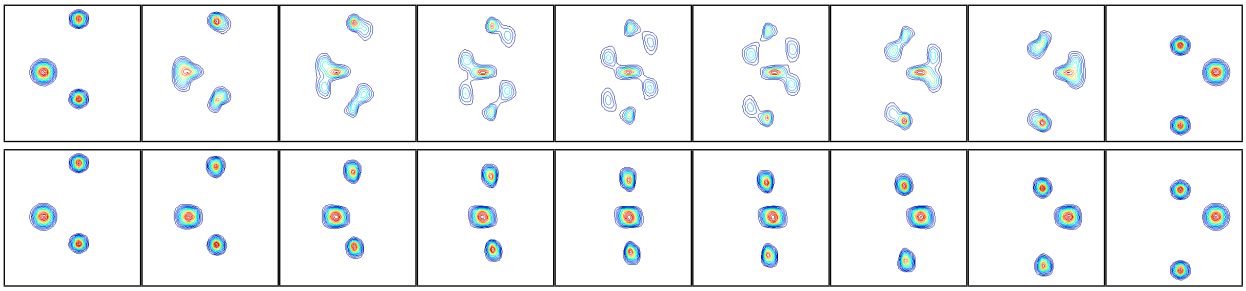
It is also important to underline that both translation and rigid penalizations allow conserving the exact shapes of the two Gaussians along time. Furthermore, one can see in Figure 12 representing both computed paths, that the rigid penalization really performs a rotation and not a translation, so that the optimal path is no more composed of straight lines.



Translation penalization Rigid penalization

FIGURE 12. Two gaussian experiment. Plot of the whole trajectory computed with the translation (on the left) and the rigid (on the right) penalization models. The rigidity here involves a real rotation.

Next, we show in Figure 13 a comparison between the approach of [3] and a piecewise translation penalization for an example showing three Gaussians. This shows that the proposed approach is able to connect properly the different objects composing the scene.



$\rho(0) = \rho_0$ $\rho(1/8)$ $\rho(1/4)$ $\rho(3/8)$ $\rho(1/2)$ $\rho(5/8)$ $\rho(3/4)$ $\rho(7/8)$ $\rho(1) = \rho_1$

FIGURE 13. Three gaussian experiment. Plot of the isolevels of the density $\rho(t)$ along the optimal path computed with different approaches. The first line is the optimal transport of [3], the second is the proposed approach with a piecewise translation penalization that allows conserving the nature of the three transported objects.

In the example of Figure 14 that presents a rotating bar, the rigid penalization allows to recover a quasi-rotation, which better preserves the prior physics with respect to pure optimal transport. It can also be observed in Figure 15 that the level lines of the estimated density are preserved with the penalization approaches. The same observations can be made in the experiments of Figures 16 and the corresponding level line length of Figures 17.

However, if we slightly modify these data, the penalization process is never able to preserve the shape contained in the data (see Figure 18). The β trick used in the previous section is here not able to provide satisfying solutions. This is due to the fact that the energy to minimize is non convex and the path we would like to estimate is too far from the optimal transport.

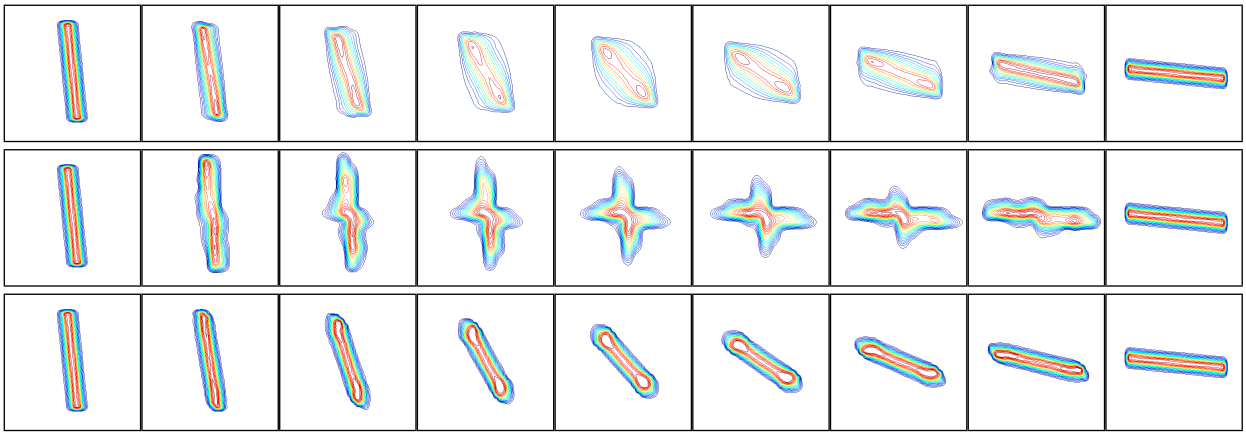
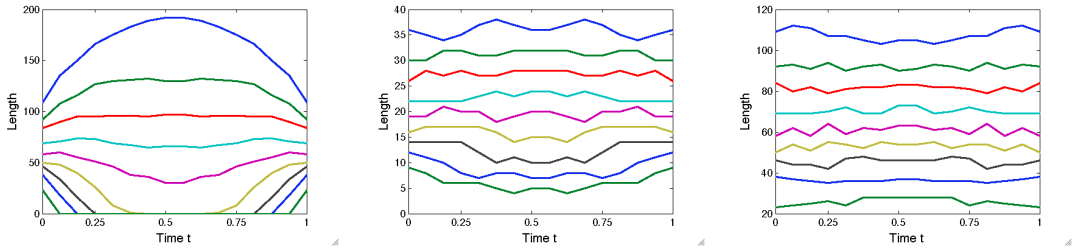


FIGURE 14. Bar experiment. Plot of the isolevels of the density $\rho(t)$ along the optimal path computed with different approaches. The first line is the optimal transport of [3], the second is the proposed approach with an incompressible penalization and the last one with a rigid penalization that allows conserving the nature of the object to transport.



Method of [3] Incompressible penalization Rigid penalization

FIGURE 15. Bar experiment. Evolution of the length of the upper level lines of the estimated density along time t: $|\rho(x,t) > i/10|$, for $i = 1 \cdots 9$. The left plot is the optimal transport of [3], the middle one is the proposed approach with an incompressible penalization and the right one corresponds to the proposed approach with a rigid penalization. Penalization allows the level lines to be preserved along the computed path.

5. Conclusion and perspectives

In this paper, we study generalized optimal transport models which attach a multiphysics model to the images to be interpolated or registrated. This investigation has been necessary as, in lot of applications to image interpolation, the results obtained using a simple minimization of a kinetic energy under some constraints do not preserve image characteristics along the optimal interpolation path, which is not physic. We then propose numerical methods to solve these new optimization problems, where the augmented Lagrangian numerical

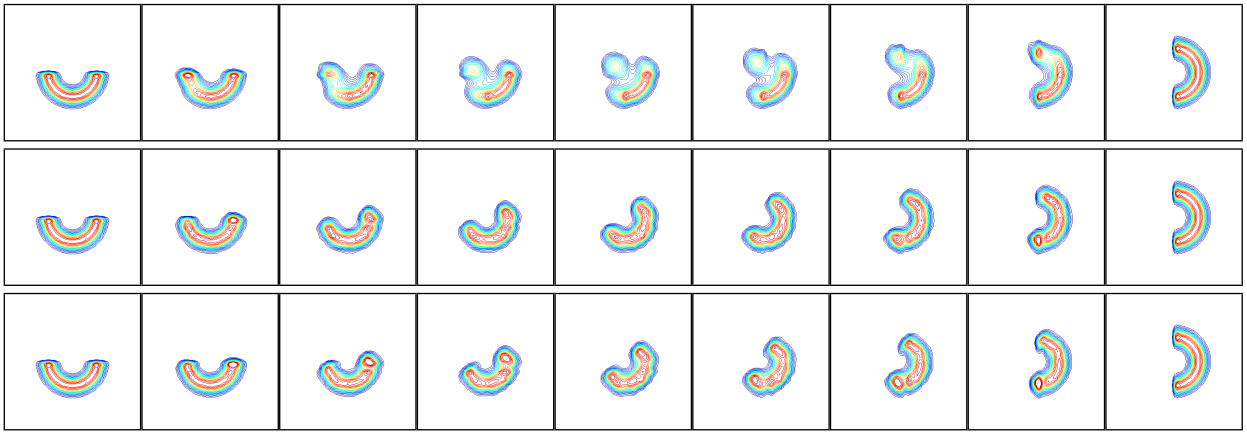
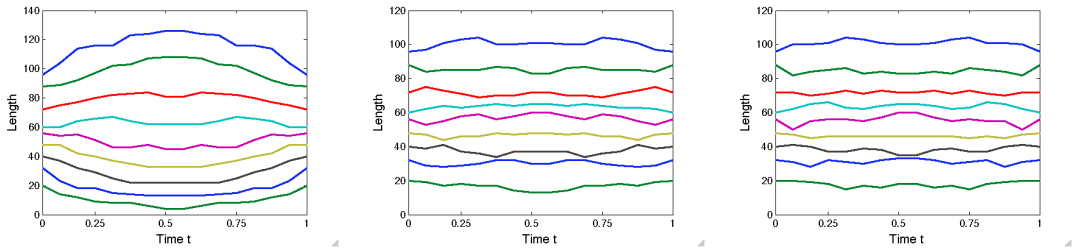


FIGURE 16. Curve experiment. Plot of the isolevels of the density $\rho(t)$ along the optimal path computed with different approaches. The first line is the optimal transport of [3], the second is the proposed approach with an incompressible penalization and the last one with a rigid penalization that allows conserving the nature of the object to transport.



Method of [3] Incompressible penalization Rigid penalization

FIGURE 17. Curve experiment. Evolution of the length of the upper level lines of the estimated density along time t: $|\rho(x,t) > i/10|$, for $i = 1 \cdots 9$. The left plot is the optimal transport of [3], the middle one is the proposed approach with an incompressible penalization and the right one corresponds to the proposed approach with a rigid penalization. Penalization allows the level lines to be preserved along the computed path.

method introduced by Benamou-Brenier has been adapted. These generalizations are not limited by the particular expressions we consider here. Indeed, others physical terms can also be considered taking into account more complex physics.

This work is a part of the ANR project TOMMI (Transport Optimal et Modèles Multiphysiques de l'Image). Another goal of this project consists of developing efficient algorithms and numerical schemes for these models. In particular, we here used the framework of Benamou-Brenier, but other minimization schemes are currently studied in order to better control the convergence of the process and have a better control on he obtained local minimas.

Acknowledgment: This work is supported by French Agence Nationale de la Recherche, ANR Project TOMMI (ANR 2011 BS01 014 01) and by Joseph Fourier University, through MSTIC grant MENTOL.

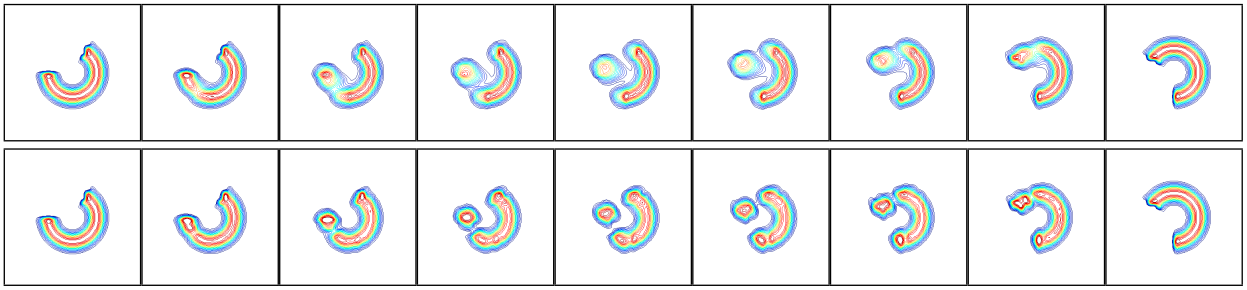


FIGURE 18. Curve experiment 2. Plot of the isolevels of the density $\rho(t)$ along the optimal path computed with two approaches. The first line is the optimal transport of [3], the second is the proposed method with a rigid penalization that is here not able to rotate the shape.

REFERENCES

[1] S. Agenent, S. Haker and A. Tannenbaum, *Minimizing flows for the Monge-Kantorovich problem*, SIAM J. Math. Analysis, 35:61–97 (2003)

[2] J.-D. Benamou, *Numerical resolution of an "unbalanced" mass transfer problem*, ESAIM: Mathematical Modelling and Numerical Analysis Vol. 37, No 5, 851–868 (2003)

[3] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution of the Monge-Katorovich mass transfer problem*, Numer. Math. 84:375–393 (2000)

[4] J.-D. Benamou and Y. Brenier, *Mixed L^2-Wasserstein optimal mapping between prescribed density functions*, J. Opt. Theory and Applications, 111, (2):255–271 (2001)

[5] J.-D. Benamou, Y. Brenier and K. Guittet *Numerical Analysis of a Multi-Phasic Mass Transport Problem*, Int. J. Numer. Meth. Fluids 40:21–30 (2002)

[6] JD Benamou, B Froese, A Oberman, *Numerical solution of the second boundary value problem for the Elliptic Monge-Ampère equation*, HAL preprint on hal.inria.fr (2012)

[7] J.-D. Benamou, B. Froese, A. Oberman, *Numerical Solution of the Optimal Transportation Problem via viscosity solutions for the Monge-Ampère equation*, arXiv preprint arXiv:1208.4873 (2012)

[8] E.J. Dean and R. Glowinski, *Numerical methods for fully nonlinear elliptic equations of the Monge-Ampère type*, Computer methods in applied mechanics and engineering, 195 (13-16):1344–1386 (2006)

[9] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Applications to the Solution of Boundary Value Problems*, Studies in Mathematics and its Applications 15, North-Holland, Amsterdam (1983)

[10] G. Loeper and F. Rapetti, *Numerical solution of the Monge-Ampère equation by a Newton's algorithm*, C. R. Math. Acad. Sci. Paris, 340:319–324 (2005)

[11] P. Tseng, *Convergence of a block coordinate descent method for nondifferentiable minimization*, Journal of Optimization Theory and Applications, 109(3):475–494 (2001)

[12] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, Vol. 50, AMS (2003)