ESAIM: M2AN 49 (2015) 1577–1592
DOI: 10.1051/m2an/2015024

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

NUMERICAL SOLUTION OF THE MONGE–KANTOROVICH PROBLEM BY DENSITY LIFT-UP CONTINUATION

Afaf Bouharguane[1], Angelo Iollo[1] and Lisl Weynans[1]

Abstract. We present an numerical method to solve the L^2 Monge–Kantorovich problem. The method is based on a continuation approach where we iteratively solve the linearized mass conservation equation, progressively decreasing a constant lift-up to map compact support densities in the limit. A Lagrangian as well as an Eulerian integration scheme are proposed. Several examples relative to the transport of two-dimensional densities are investigated, showing that the present methods can significantly reduce the computational effort.

Mathematics Subject Classification. 68U01, 65K05.

Received March 13, 2015.
Published online November 5, 2015.

1. Introduction

Optimal transportation is increasingly used to model problems in mechanics, physics, image analysis and other fields, see *e.g.* [12] and references therein. Because of all these applications, this old topic first introduced by Monge in 1781 [8], has attracted considerable attention these last years especially from a numerical point of view [1–3, 5, 6, 9, 10]. Indeed, compared to the theoretical results already obtained, the discrete solution of this problem still poses challenging problems in terms of computational burden and accuracy.

In this work we focus on the numerical solution of the L^2 Monge–Kantorovich problem (MKP) defined as follows. Let $\rho_0(\xi), \rho_1(x)$ be two smooth enough non-negative scalar density functions with compact support Ω_0 and Ω_1, where $\xi, x \in \mathbb{R}^d$ and d is the space dimension. We assume that

$$\int_{\Omega_0} \rho_0(\xi)\mathrm{d}\xi = \int_{\Omega_1} \rho_1(x)\mathrm{d}x.$$

Let $X : \Omega_0 \to \Omega_1$ a smooth one-to-one map such that $X(\xi)$ realizes the transfer of ρ_0 onto ρ_1, *i.e.*, a map that satisfies the following Jacobian equation:

$$\rho_0(\xi) = \det(\nabla X(\xi))\rho_1(X(\xi)).$$

Keywords and phrases. Optimal transport, Monge–Kantorovich problem, numerical solution, Newton method, continuation approach.

[1] Institut de Mathématiques de Bordeaux, UMR 5251 CNRS, Université de Bordeaux and Equipe-projet MEMPHIS, Inria Bordeaux Sud-Ouest, 33405 Talence, France. afaf.bouharguane@math.u-bordeaux1.fr

This equation is underdetermined with respect to $X(\xi)$ and a solution is selected among all possible maps by introducing the following L^2 Kantorovich–Wasserstein distance:

$$\inf_X \int_{\Omega_0} \rho_0(\xi)|X(\xi) - \xi|^2 \mathrm{d}\xi.$$

The L^2 MKP corresponds to finding a map X^* such that this infimum is achieved. It has been proved that this problem admits a unique solution [4, 11, 12], which is the gradient of a.e. convex function $\Psi : \Omega_0 \to \mathbb{R}$:

$$X^*(\xi) = \nabla\Psi(\xi).$$

In the scientific computing literature, there exist two class of methods to approximate this problem. The first one is based on a direct solution of the Monge–Ampere equation (MAE):

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

The difficulty in approaching the problem in this way is that the boundary conditions of this equation are not known *a priori*. Instead, the solution must verify the constraint $X(\Omega_0) = \Omega_1$. For the solution of the MAE with Dirichlet b.c. a recent numerical study in two dimensions is discussed in [5], where the solution is obtained via a least-square formulation. The full MKP solution via a MAE was considered in [3]. The numerical method employed is based on the solution of MAE with boundary conditions that are iteratively updated to converge to the MKP solution.

In a recent work, Saumier *et al.* [10] propose a Newton algorithm to solve the L^2 MKP for smooth periodic densities bounded away from zero. This method is an extension of the scheme proposed by Loeper and Rapetti for the solution of the MAE on a torus [7] for which convergence is ensured. The convergence of this Newton algorithm relies on the assumption that the initial and final densities are bounded away from zero, as the under-relaxation parameter of the Newton update is vanishingly small when the density support becomes compact.

Another class of methods moves on from ideas of continuum mechanics. Benamou and Brenier (BB) numerically solved the MKP by using an augmented Lagrangian method [2]. In their formulation a temporal dimension is introduced so that, given Π: $[0,1] \times \Omega_0 \to \mathbb{R}^d$, with $\Pi(0,\xi) = \xi$, $\Pi(1,\xi) = X(\xi)$, $x = \Pi(t,\xi)$ and $\partial_t\Pi = v(t,x)$, the MKP amounts to the solution of

$$\inf_{\rho,v} \int_{\mathbb{R}^d} \rho(t,x)|v(t,x)|^2 \,\mathrm{d}x,$$

where the infimum is taken among all densities $\rho(t,x) \geq 0$ and velocity fields $v(t,x) \in \mathbb{R}^d$ satisfying the continuity equation

$$\partial_t\rho + \nabla\cdot(\rho v) = 0,$$

and the initial and final conditions:

$$\rho(0,\cdot) = \rho_0,\ \rho_1(1,\cdot) = \rho_1.$$

The BB method results in a robust and viable discrete minimization problem under constraints which admits a unique solution. However, since it is a gradient method in space-time, the computational cost may be relevant. Also, numerical diffusion of the transported densities is observed in the simulation of the transport for intermediate times. A recent improvement of the minimization method at the base of the BB algorithm is proposed in [9]. Using proximal splitting schemes the authors were able to solve difficult transport problems in presence of geometric constraints.

A yet different approach is devised in [1]. The idea is to first consider a mass preserving mapping, not necessary optimal, between the initial and final distributions and then to solve a PDE up to steady state in order to rearrange the non-optimal mapping into an optimal one. It is shown that also this approach leads to a gradient-based minimization problem for which many gradient steps are usually needed to converge.

In this paper we investigate two computationally efficient algorithms to solve the optimal mass transfer problem for compactly supported densities. These schemes rely on a continuation approach that progressively reduces a constant lift-up of the mapped densities. Two schemes are presented: a Lagrangian scheme which is an extension of the Newton method introduced in [7,10] and its Eulerian counterpart, based on a remeshed particle method, in the continuum mechanics setting proposed by BB. In applications where the Newton method alone would not converge, these approaches prove to be effective and preserve the compactness of the supports in the limit.

We validate these numerical schemes on several test cases involving compactly supported densities, including an example from medical imaging.

2. The general setting

2.1. Newton iteration

The main idea of solution relies on the assumption that we dispose of a mapping at iteration n that is a perturbation of the optimal mapping. We derive a linear perturbation equation that is used to iteratively improve the initial guess. Let us assume that the optimal mapping is

$$X_o(\xi) = \nabla_\xi \Psi_o$$

and that the mapping obtained at iteration n is

$$X^n(\xi) = \nabla_\xi \Psi_o + \nabla_\xi \Psi_\epsilon^n.$$

where the error Ψ_ϵ^n satisfies $\|\Psi_\epsilon^n\|_2 \approx \epsilon$. We define $\rho_0^n(\xi)$ as the initial density at iteration n that mapped by $X^n(\xi)$ gives the exact final density $\rho_1(x)$. Then, taking a first-order Taylor expansion, we have

$$\begin{aligned}\rho_0^n(\xi) &:= \rho_1(X^n(\xi))\det\left(\nabla_\xi X^n(\xi)\right)\\ &= \underbrace{\rho_1(X_o)\det\left(\nabla_\xi X_o\right)}_{=\rho_0(\xi)} + \det\left(\nabla_\xi X_o\right)\rho_1(X_o)\mathrm{Tr}\left((\nabla_\xi X_o)^{-1}\nabla_\xi^2\Psi_\epsilon^n\right)\\ &\quad + \det\left(\nabla_\xi X_o\right)\nabla_x\rho_1(X_o)\cdot\nabla_\xi\Psi_\epsilon^n + o(\epsilon),\end{aligned}$$

where Tr denotes the matrix trace operator.

At first order in ϵ, we have

$$\frac{\rho_0^n(\xi)-\rho_0(\xi)}{\det\left(\nabla_\xi X_o\right)} \approx \rho_1(X_o)\mathrm{Tr}\big((\nabla_\xi X_o)^{-1}\nabla_\xi^2\Psi_\epsilon^n\big) + \nabla_x\rho_1(X_o)\cdot\nabla_\xi\Psi_\epsilon^n. \tag{2.1}$$

The mapping update is then found by the iteration

$$X^{n+1} = X^n - \alpha\,\nabla\Psi_\epsilon^n,$$

that converges to X_o as a geometric series, for $\alpha \in [0,1]$, in the linearized approximation [7,10].

Remark 2.1. If the approximated mapping is a perturbation of identity, *i.e.*,

$$X^n(\xi) = \xi + \nabla_\xi\Psi_\epsilon^n$$

then equation (2.1) reduces to

$$\rho_0^n(\xi) - \rho_0(\xi) \approx \nabla_\xi\cdot\left(\rho_1(X^n(\xi))\nabla_\xi\Psi_\epsilon^n\right) \tag{2.2}$$

that is equivalent to a semidiscretization in time of the continuity equation written in Eulerian form.

2.2. The initial mapping

The choice of the mapping used to initialize the Newton method can be important to reduce the number of iterations to converge. This is especially true when the Wassterstein distance is *a priori* large. A first way to initialise this mapping is to consider a modification of the approach proposed in [1]. As a first step, we compute a mapping that transports ρ_0 onto ρ_1. Without loss of generality we assume that $\Omega = [-\Lambda, \Lambda] \times [-\Lambda, \Lambda]$, $\Lambda \in \mathbb{R}$. We assume also that the initial and final densities have compact support included in Ω. Let $a : [-\Lambda, \Lambda] \to \mathbb{R}$ be the solution of

$$a'(\xi_1) \int_{-\Lambda}^{\Lambda} \rho_1(a(\xi_1), x_2)\, dx_2 = \int_{-\Lambda}^{\Lambda} \rho_0(\xi_1, \xi_2)\, \mathrm{d}\xi_2 \tag{2.3}$$

with $x = X(\xi)$, $\xi = (\xi_1, \xi_2)$, $x = (x_1, x_2)$ and $a(-\Lambda) = -\Lambda$. Also let $b : \Omega \to \mathbb{R}$ be solution of the parametric ordinary differential equation

$$a'(\xi_1) \frac{\partial b(\xi_1, \xi_2)}{\partial \xi_2} \rho_1(a(\xi_1), b(\xi_1, \xi_2)) = \rho_0(\xi_1, \xi_2) \tag{2.4}$$

with $b(\xi_1, -\Lambda) = -\Lambda$, $\forall \xi_1 \in [-\Lambda, \Lambda]$. The mapping $X_\omega = (a(\xi_1), b(\xi_1, \xi_2))$ takes ρ_0 onto ρ_1 by construction. Other choices of $a(\xi)$ respecting monotonicity, regularity and appropriate initial conditions are possible.

The mapping $X_\omega = a(\xi_1)\, b(\xi_1, \xi_2)$ is not in general the gradient of a complex potential, while this requirement is a necessary condition for optimality. However, this mapping can always be decomposed as the sum of an irrotational field and a solenoidal field:

$$X_\omega = \nabla_\xi \Psi_i + \nabla \times A. \tag{2.5}$$

Hence, the actual initial mapping $\nabla_\xi \Psi_i$ is found by solving $\nabla \times \nabla \times A = \nabla \times X_\omega$ with homogeneous Neumann boundary conditions and computing $\nabla_\xi \Psi_i = X_\omega - \nabla \times A$. In two space dimensions this amounts to the solution of one Laplace equation.

A different way to initialise the mapping is to rely on the following discretisation of the continuity equation in Eulerian form:

$$\nabla_\xi \cdot \left(\frac{\rho_1 + \rho_0}{2} \nabla_\xi \Psi_i \right) = \rho_0(\xi) - \rho_1(\xi)$$

with homogeneous Dirichlet boundary conditions. This mapping is defined modulo a rigid translation that can eventually be inferred by the average displacement of the first moment of the initial and final distributions. In the following numerical illustrations we have observed that both approaches reduce the number of iteration to convergence.

2.3. Regularization and continuation

Equations (2.1) and (2.2) are well defined only on the support of the initial and final densities. Therefore, in actual computations involving compact support densities, the initial and final densities are regularized by adding a small constant ζ so that the leading order differential operator is well defined on the entire domain:

$$\left(\rho_1 \left(X^n\right) + \zeta\right) \mathrm{Tr}\left(\left(\nabla_\xi X^n\right)^{-1} \nabla_\xi^2 \Psi_\epsilon^n \right)$$

with $\zeta \approx 10^{-3}$ in practice.

When the support of the initial and final density tends to be compact and correspondingly the Wasserstein distance d_w increases, the under-relaxation parameter α of the Newton method in Section 2.1 becomes excessively small to ensure convergence [7, 10]. In practice the residuals are hanging and convergence is never reached. In those cases it is possible to solve the problem by continuation. We start by observing that

$$\lim_{\zeta \to \infty} \frac{\rho_0(\xi) + \zeta}{\rho_1(x) + \zeta} = \det\left(\nabla_\xi X\right) = 1$$

and hence $\lim_{\zeta\to\infty} d_w = 0$. Therefore, for ζ large enough the condition $X_o(\xi) \approx \xi$ is verified.

The idea is then to add to the initial and final densities a constant such that their Wasserstein distance is reduced and hence the Newton algorithm converges. This constant is then iteratively brought to $O(10^{-3})$ for unitary initial and final densities.

Taking $\zeta^k \in \mathbb{R}^+$, $\forall k \in \mathbb{N}$, we define

$$\rho_0^k = \rho_0 + \zeta^k \text{ and } \rho_1^k = \rho_1 + \zeta^k.$$

Once the optimal mapping Ψ_o^k for ζ^k is determined, we pose $\zeta^{k+1} = (1-\beta)\zeta^k$ with $\beta \in [0,1]$ and in practice we choose $\beta = 0.1$. The optimal mapping Ψ_o^k found for ζ^k is then taken as the initial mapping Ψ_i^{k+1} for the next optimal transport problem with ζ^{k+1}. The overall numerical method for the continuation approach is given in Algorithm 1.

Algorithm 1. Continuation algorithm.

1. $k = 0$, ζ^0 given, $\rho_0^0 = \rho_0 + \zeta^0, \rho_1^0 = \rho_1 + \zeta^0$;
2. Solve the L^2 MKP as in Section 2.1 to obtain Ψ_o^k ;
3. $\zeta^{k+1} = (1-\beta)\zeta^k$;
4. $\rho_0^{k+1} = \rho_0 + \zeta^{k+1}$, $\rho_1^{k+1} = \rho_1 + \zeta^{k+1}$;
5. $\Psi_i^{k+1} = \Psi_o^k$;
6. $k = k + 1$;
7. go to 2 while $\zeta^k > 10^{-3}$.

Remark 2.2. This approach assumes that the solution of the MAE can be continued, *i.e.*, there exists a neighboring solution that results from a small perturbation of the initial and final densities. Let us note:

$$F : (\rho_0, \rho_1, \Psi, \zeta) \to (\rho_0 + \zeta) - (\rho_1(\nabla\Psi) + \zeta)\det(\nabla^2\Psi).$$

Thanks to results stated in [10] and Theorem 2.2 of [7] (Caffarelli's theorem), $D_\zeta F$ is invertible and bounded, as it corresponds to the same linearized operator in Section 2.1. Existence of a neighboring solution is then ensured by the implicit function theorem.

In order to give a first estimate of ζ^0, it is natural to assume that it is proportional to the Wasserstein distance d_w between the two densities. In the following test cases we used the estimate

$$\zeta^0 \propto \int \rho_0 |\nabla_\xi \Psi_i - \xi|^2 \, \mathrm{d}x,$$

where $\nabla\Psi_i$ is the initial mapping, as explained in Section 2.2. For example, for test case TC3 (see Sect. 4) we took $\zeta = 1$ as the Wasserstein distance approximation above is of the same order.

Similar results are obtained for an estimation of d_w that can be computed *a priori*. Let

$$I_0 = \left(\int_{\Omega_0} \rho_0 \xi^2 \mathrm{d}\xi\right)^{1/2} \quad I_1 = \left(\int_{\Omega_1} \rho_1 x^2 \mathrm{d}x\right)^{1/2}.$$

Then, applying the Cauchy–Schwartz inequality we have that $d_w \geq |I_1 - I_0|$. Also, by the triangular inequality we have $d_w \leq I_1 + I_0$. These inequalities can give an estimate of d_w. For example, when $I_1 \gg I_0$, I_1 is a good approximation of d_w. This is the case when the origin of the coordinate system is the center of mass of ρ_0 and the supports of the two densities are disjoint.

3. Discrete solution

In this section, we detail two solution approaches to solve the L^2 MKP for a given ζ^k. A Lagrangian approach, that can readily be coded as it implies the solution of a rather simple elliptic PDE and an Eulerian approach, that can more easily be applied in more general geometries.

3.1. Lagrangian approach

According to previous sections, the following iterative method (Algorithm 2) is considered. In [7, 10] it is

Algorithm 2. Lagrangian iterative algorithm.

1. $n = 0$;
2. compute $X^0 = \nabla\Psi_i$ as explained in Section 2.2;
3. $\rho_0^n(\xi) = \rho_1(X^n(\xi)) \det \nabla_\xi X^n(\xi)$;
4. compute Ψ_ϵ^n by solving equation (2.1);
5. $X^{n+1} = X^n - \alpha \nabla\Psi_\epsilon^n$;
6. $n = n + 1$;
7. go to 3 if convergence is not attained;

shown that this algorithm always converges for sufficiently small damping coefficients (*i.e.* for α small), when the distributions ρ_0 and ρ_1 are sufficiently smooth, periodic and bounded away from zero. Algorithm 2 then converges as a geometric series, *i.e.*,

$$\Psi_\epsilon^n \approx \Psi_o \left(1 - (1-\alpha)^n\right). \tag{3.1}$$

Therefore, defining the normalized residual as $r = \|\Psi_\epsilon^n - \Psi_o\|/\|\Psi_o\|$, we have:

$$\log r = n \log(1-\alpha), \tag{3.2}$$

where $\log(1-\alpha)$ is the rate of convergence of the iterative scheme.

To approximate the problem, we discretize equation (2.1) by a standard second-order finite difference scheme on a Cartesian grid. At the discrete level, we impose homogeneous Dirichlet boundary conditions to equation (2.1). In practice, since the initial and final densities of the problem without the lift-up ζ have compact support, the boundary conditions imposed at the borders of the computational domain do not significantly affect the solution as $\zeta \to 0$. Mass transportation is performed by the continuity equation in Lagrangian coordinates.

In the next sub-section we will provide an alternative approach where the continuity equation is integrated in Eulerian coordinates. For numerical examples, see Section 4, we will assume that the convergence is attained when a threshold criterion on $\|\rho_0^n - \rho_0\|_\infty$ is satisfied.

3.2. Eulerian approach

In the Eulerian framework, the solution of the L^2 MKP consists in determining the initial velocity $u(0,.) = u_0$ satisfying:

$$\begin{cases} \partial_t(\rho) + \nabla \cdot (\rho u) = 0 & (3.3) \\ \partial_t(\rho u) + \nabla \cdot (\rho u \otimes u) = 0 & (3.4) \\ \rho(0,\cdot) = \rho_0, \quad \rho(1,\cdot) = \rho_1. & (3.5) \end{cases}$$

The optimal initial velocity u_0 is written under the form $u_0 = \nabla\Psi$, where the potential Ψ is the Lagrange multiplier of constraints (3.3) and (3.5). We propose an alternative solution method for u_0 by adapting Algorithm 2 to the Eulerian framework.

Starting from an initial guess for u_0, we numerically solve equations (3.3)–(3.4) as detailed hereafter, using $\rho(0,.) = \rho_0$ as the initial condition for density. We get a final density $\tilde{\rho}_1$ and a final velocity field that we denote $\tilde{u}_1$. Since in general the initial velocity guess does not correspond to the optimal solution, the density $\tilde{\rho}_1$ is different from ρ_1. The next step is to find a potential velocity field transporting the data $\tilde{\rho}_1$ on ρ_1. This velocity field is considered as a correction to the initial mapping.

To find the additional velocity field, we approach the continuity equation (3.3) by the following elliptic problem:

$$\nabla \cdot \left(\frac{\rho_1 + \tilde{\rho}_1}{2} \nabla \tilde{\Psi} \right) = \tilde{\rho}_1 - \rho_1.$$

This problem is solved using classical second-order centered finite-differences. From the potential $\tilde{\Psi}$ we compute a velocity field $\tilde{u} = \nabla \tilde{\Psi}$, which can be seen as the velocity field necessary to advect $\tilde{\rho}_1$ to ρ_1. As we want to use it as a correction to u_0, we have to advect it backward in time to make it match with the initial density ρ_0. We therefore perform an integration backward in time of the system (3.3)–(3.4) with final conditions $\tilde{\rho}_1$ and $\tilde{u}_1$, obtained by the forward numerical computation. The result of the backward advection of $\tilde{u}$, that we denote $\tilde{u_0^n}$, is added to u_0 as a corrective term. The full method is detailed in Algorithm 3.

The numerical resolution of equations (3.3)–(3.4) is performed using a remeshed particle method, similar to the one developed in [13] for compressible Euler equations. In this class of methods, the fluid is discretized on small masses concentrated on points. These points, which are called particles, are displaced in a Lagrangian way. New particles, uniformly distributed at the nodes of an underlying grid, are created at regular time intervals by a conservative interpolation of the existing particles, what is usually called remeshing the particles.

Because the particles themselves are moved in a Lagrangian way, remeshed particle methods are submitted to less restrictive stability conditions, in the context of advection problems, than more classical grid-based methods such as finite-differences or finite-volume methods. Indeed their stability condition is typically proportional to the inverse of the velocity gradient (meaning that the particles trajectories do not cross), instead of a classical CFL condition. In the context of the method that we present here for optimal mass transport, we need to compare the density computed at the final time to the final exact density ρ_1. Therefore, the use of larger time steps is an advantage because it means that the final state is computed with less numerical dissipation due to the temporal integration than in the case of more classical grid-based methods.

Algorithm 3. Eulerian iterative algorithm.

1. $n = 0$;
2. initialize u_0;
3. for initial data ρ_0 and $u_0 = X^n$, compute ρ_1^n and u_1^n by numerically solving system (3.3)–(3.4);
4. compute $\tilde{\Psi}$ by solving equation:
$$\nabla \cdot \left(\frac{\rho_1 + \tilde{\rho}_1}{2} \nabla \tilde{\Psi} \right) = \tilde{\rho}_1 - \rho_1$$
5. integrate backward in time system (3.3)–(3.4) with final conditions $\tilde{\rho}_1$ and $\tilde{u}_1$ to advect backward the velocity $\tilde{u}$ and get at initial time the velocity correction $\tilde{u}_0^n$ matching with ρ_0;
6. $X^{n+1} = X^n + \alpha\, \tilde{u}_0^n$;
7. $n = n + 1$;
8. go to 3 if convergence is not attained;

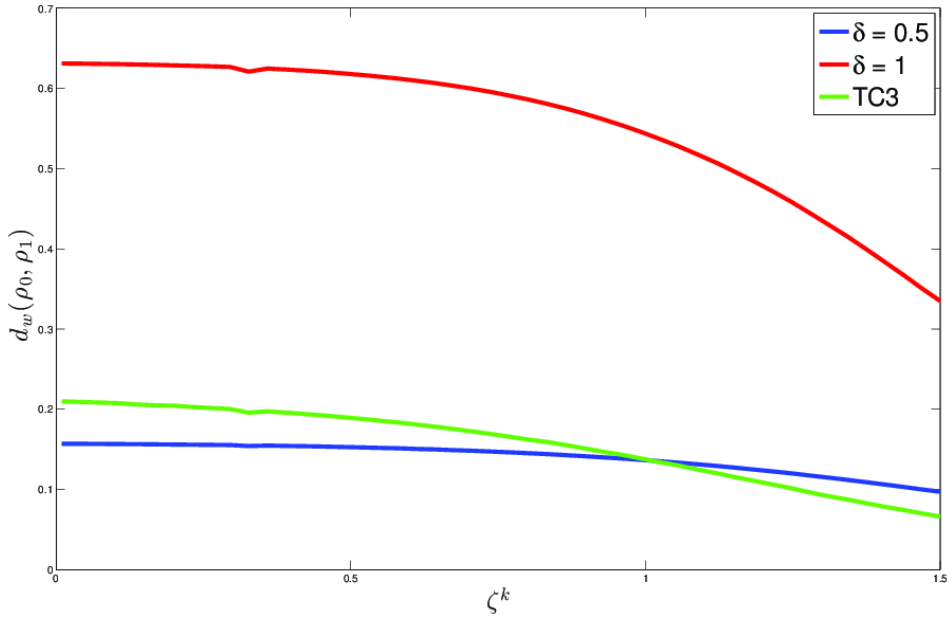

FIGURE 1. Wasserstein distance for ζ^k increassing for the test case TC3 (green line) and for the test case TC0 with $\delta = 0.5, 1$ (blue and red lines). (Color online).

4. Numerical illustrations

4.1. Convergence by continuation

We consider the following test case (TC) with distributions:

$$\text{TC0:} \quad \rho_0 = \exp\left(-|x|^2\right) + \zeta; \quad \rho_1 = \exp\left(-|x+\delta|^2\right) + \zeta;$$

where ζ is the regularization parameter. As shown in [7,10], the Newton method converges for all smooth enough periodic densities bounded away from zero, provided that the under-relaxation parameter α is sufficiently small. For TC0, when $\delta > 0.5$ the under-relaxation parameter α is so small that the Newton method does not converge in practice. Thanks to continuation approach, we can still compute in an efficient way the optimal transportation between these two distributions. We display in Figure 1, the evolution of the d_w as a function of ζ^k for test cases TC0 and TC3 (see below). It is seen that as expected for increasing ζ^k, the Wasserstein distance decreases. The density lift-up reduces d_w and speeds up convergence. Figure 2 shows that there is no impact on the optimal mapping $\nabla\Psi$ if one regularizes the initial and final densities by adding a small constant $\zeta = 10^{-3}$ or $\zeta = 10^{-4}$, confirming that under a certain threshold the regularization has not impact on the optimal mapping.

4.2. Transport tests

In this section, we provide several numerical validations of our methods. We perform the following test cases:

- TC1. We consider two density distributions having a non-negligible support intersection. We take two Gaussians of unit mass, same variance and displaced of 0.2 in the vertical direction. The final density has a cross correlation of 0.5, the initial is isotropic. We have:

$$\rho_0(\xi) = 3.97887\mathrm{e}^{-12.5\left(\xi_1^2+(\xi_2+0.1)^2\right)}$$
$$\rho_1(x) = 4.59441\,\mathrm{e}^{-16.66666\left(x_1^2+x_1(x_2-0.1)+(x_2-0.1)^2\right)}.$$

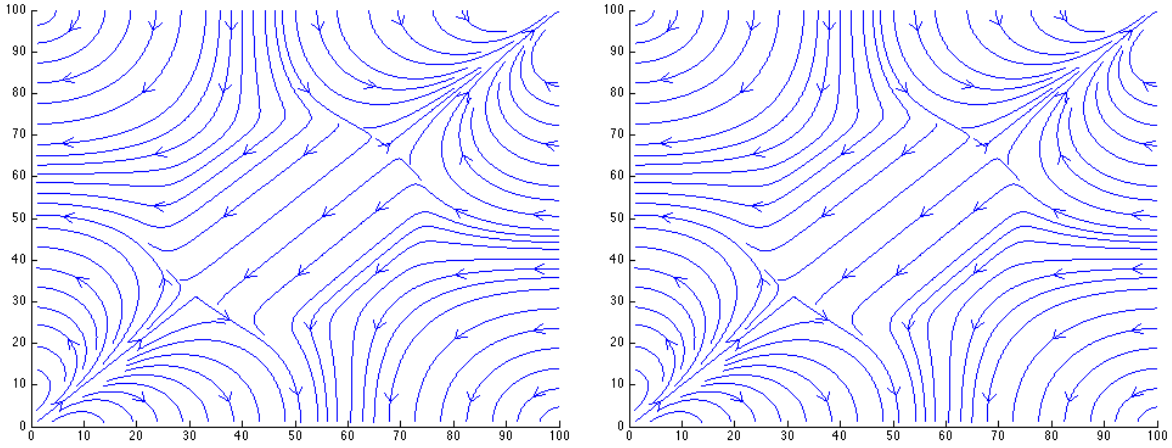

FIGURE 2. Stream curves corresponding to $\nabla_\xi \Psi$ the optimal mapping for $\zeta = 10^{-3}$ (*right*) and $\zeta = 10^{-4}$ (*left*) for the test case TC0 with $\delta = 0.1$.

- TC2. This is similar to TC1, but the centroid of the initial density is further displaced. We take the initial and final density distributions:

$$\rho_0(\xi) = 1.98944 \left(\mathrm{e}^{-12.5\left((\xi_1-0.2)^2+(\xi_2+0.1)^2\right)} + \mathrm{e}^{-12.5\left((\xi_1+0.2)^2+(\xi_2+0.1)^2\right)} \right)$$
$$\rho_1(x) = 3.97887\, \mathrm{e}^{-12.5\left(x^2+y^2\right)}.$$

- TC3. In this case there is mass separation and virtually no intersection between the initial and final density support. The initial and final density are

$$\rho_0(\xi) = 1.98944 \left(\mathrm{e}^{-12.5\left((\xi_1-0.5)^2+(\xi_2+0.3)^2\right)} + \mathrm{e}^{-12.5\left((\xi_1+0.5)^2+(\xi_2+0.3)^2\right)} \right)$$
$$\rho_1(x) = 3.97887\, \mathrm{e}^{-12.5\left(x^2+y^2\right)}.$$

- TC4. A final case showing the robustness of the Newton iteration is presented. The mapping between two scans of 168^2 pixels relative to the abdomen of a breathing patient, see Figure 11, is determined.

Test case TC1, TC2, TC3 are performed using both the Lagrangian and the Eulerian schemes. We consider a square domain with $\Lambda = 2$ discretized using a uniform 200×200 grid. Compared to the Lagangian case, the corresponding Eulerian TCs are rotated of $\pi/4$ in order to show the accuracy of the particle remeshing in a transverse direction with respect to the grid.

For TC1, with $\alpha = 0.2$, the expected rate of convergence is of 0.09691, see (3.2). The fitting of the convergence curve with a straight line gives a convergence rate of 0.0958315 for the Algorithm 2 (Lagrangian), see Figure 3. The relative error in the max norm is of the order of 10^{-3} after 23 iterations. The solution with Algorithm 3 (Eulerian) is given in Figure 4.

In the next test case (TC2), the expected rate of convergence is of.0457575 ($\alpha = 0.1$), the fitting of the convergence curve with a straight line gives a convergence rate of 0.444000 for Algorithm 2 (Lagrangian), see Figure 5. The solution with Algorithm 3 (Eulerian) is given in Figure 6.

Finally, in TC3 the Lagrangian scheme was initialized with a regularization $\zeta = 1$ and in 40 steps of the continuation algorithm it was reduced to 10^{-3}. The initial condition along with the optimal mapping stream curves are shown in Figure 7. The intermediate and final solution and error with Algorithm 2 (Lagrangian) are shown in Figures 8 and 9. The solution with Algorithm 3 (Eulerian) is given in Figure 10.

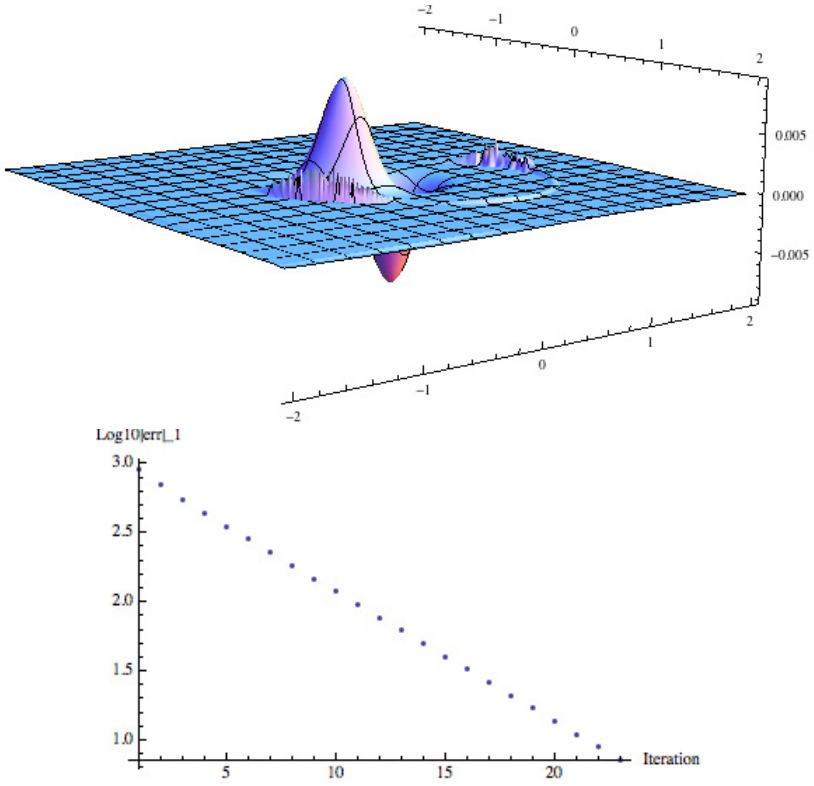

FIGURE 3. TC1. Error after 23 iterations and convergence for Algorithm 2.

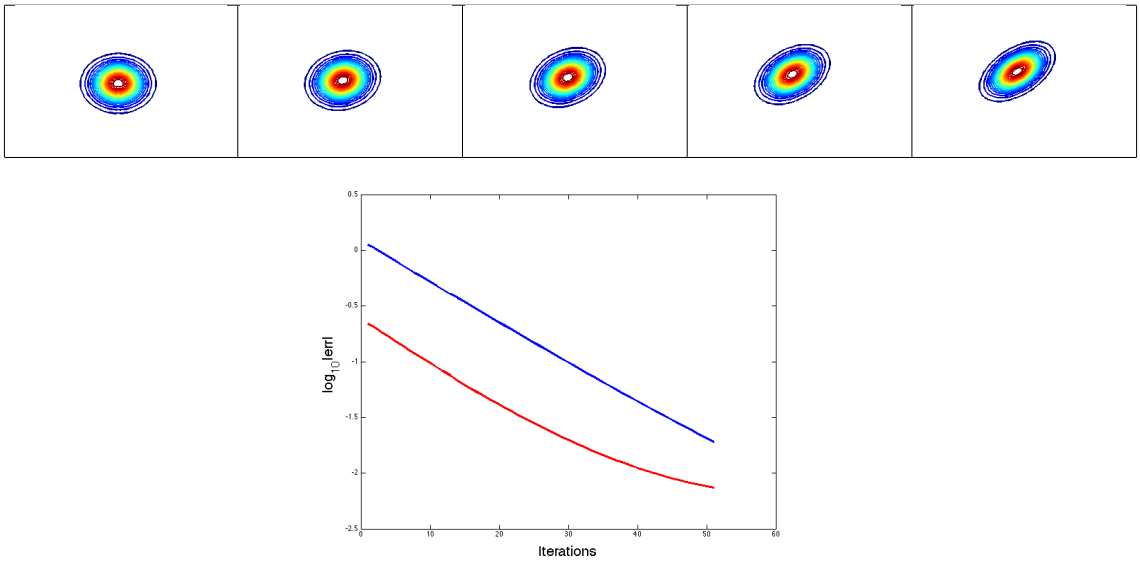

FIGURE 4. TC1. *Top picture row*: Plot of density isolines $\rho(t,x)$ for $t = 0, 0.25, 0.5, 0.75, 1$ along with the optimal path computed by Algorithm 3. *Bottom*: Red (resp. blue) line represents the error in L^1 norm (resp. L^∞ norm). (Color online).

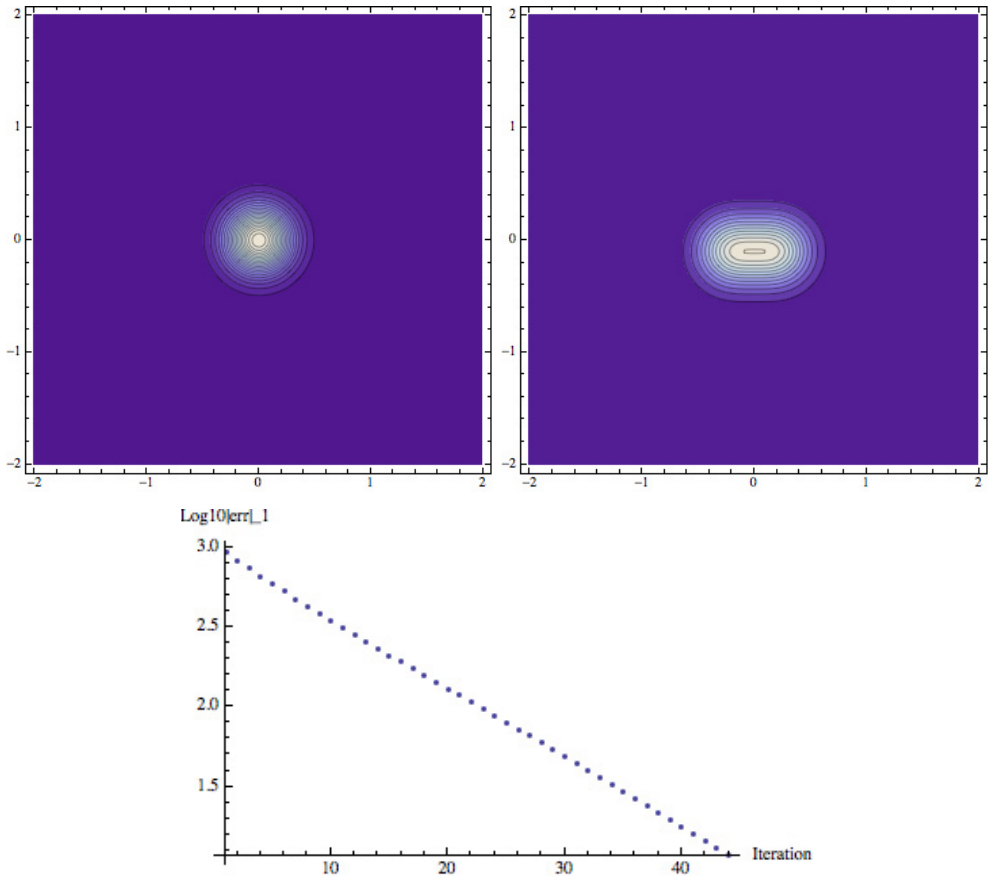

FIGURE 5. TC2: Initial distribution, final distribution and convergence for Algorithm 2. Isolines from 0 to the max spaced of 0.2.

The last case presented, TC4, is relative to front abdomen sections of a breathing patient, Figure 11 (courtesy of B.D. de Senneville). The figure shows the initial velocity field relative to the optimal transport between the two subsequent scans. In certain focalized beam therapies, like focalized ultrasounds, it is crucial to accurately predict the movement of a patient in order to calibrate the displacement of the targeted region. In applications, real-time optical flow techniques based on heuristic arguments are employed. The methods presented here are not real time, although the solution of the MKP took a few seconds in this case. However, the solution of the MKP to determine the displacement between two images offers the advantage over optical flow techniques of being objective in the sense that it is based on a clearly identified model.

5. Discussion and conclusions

In this paper, we proposed two algorithms to compute in an effecient way the numerical solution of the L^2-Monge–Kantorovich problem for smooth enough densities with compact support. We observe that Algorithms 2 and 3 are suitable (without continuation) for transports where the distance travelled by the elementary masses is small compared to the characteristic length of the density distributions.

Compared to Algorithms 2 and 3, the BB approach has the advantage of solving for a space-time saddle point. Hence the trajectories are not identically straight line, as in the case of the present method, and the

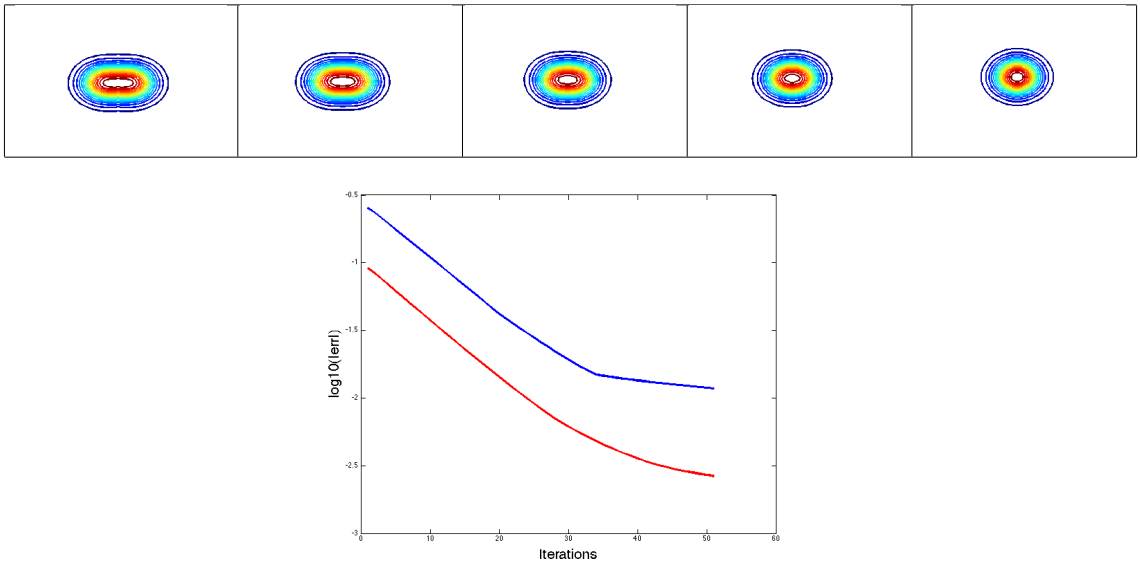

FIGURE 6. TC2. *Top picture row*: Plot of the isolines of the density $\rho(t,x)$ for $t = 0, 0.25, 0.5, 0.75, 1$ with Algorithm 3. *Bottom*: Convergence of Algorithm 3. Red (resp. blue) data represent the $\log_{10}(||\rho_1^n - \rho_1||_1)$ (resp. $\log_{10}(||\rho_1^n - \rho_1||_\infty)$) as a function of the number of iterations. (Color online).

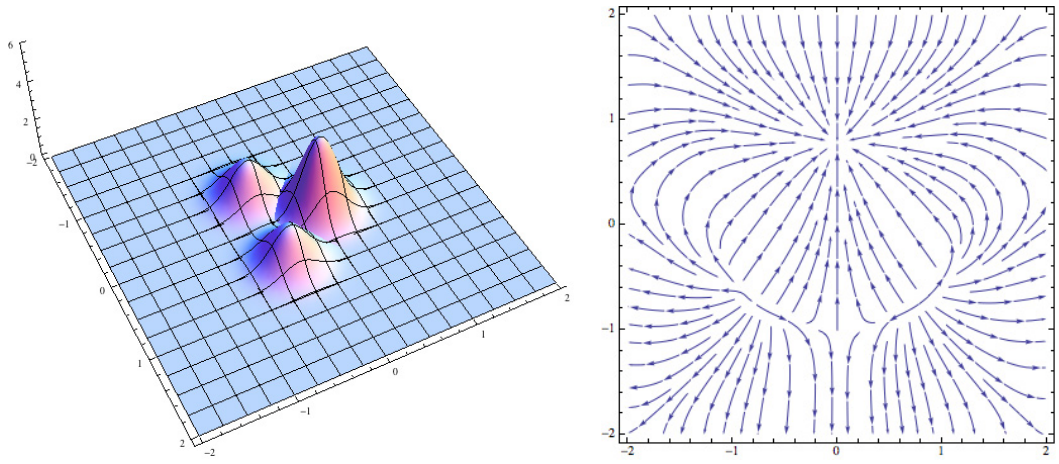

FIGURE 7. TC3. Initial and final densities and stream curves corresponding to $\nabla_\xi \Psi_o$, the optimal mapping obtained with Algorithm 2.

problem remains well conditioned also for large mass displacements as shown in [9]. This is of course at the price of a costly computational problem in $d+1$ dimensions.

In order to show the computational advantage of the Newton iteration for suitable cases, we compare the CPU time to solve the MKP by the BB method, to the CPU time of algorithms presented here. Since the convergence criteria are different for the two methods, we will determine the CPU time so that the initial criteria will be divided by 30, 60 and 100.

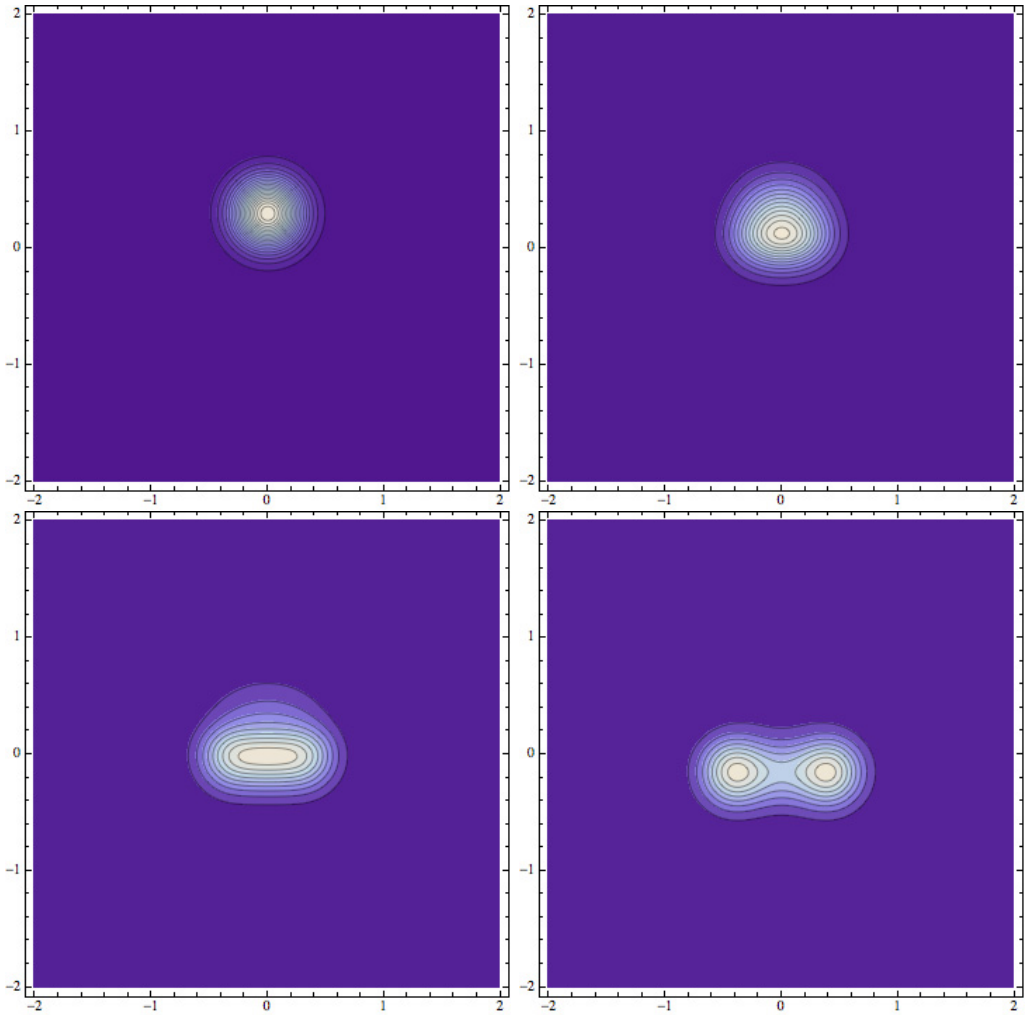

FIGURE 8. TC3. Intermediate densities obtained with Algorithm 2 for $X_\alpha = \xi + \chi(\nabla_\xi \Psi_o - \xi)$ with $\chi = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\}$ from left to right, top to bottom. Isolines from 0 to the max spaced of 0.2.

We recall that the convergence criterion of the BB method is based on the following residual of the Hamilton–Jacobi equation (see [2]):

$$res^n = \partial_t \phi^n + \frac{|\nabla \phi^n|}{2},$$

where ϕ is the Lagrange multiplier of constraints (3.3) and (3.5). The convergence criterion is given by:

$$crit_b^n = \sqrt{\frac{\int_0^1 \int_\Omega \rho^n |res^n|}{\int_0^1 \int_\Omega \rho^n |\nabla \phi^n|^2}}.$$

For the Picard iteration method presented here, we consider the error between the given final distribution ρ_1 and the final distribution ρ_1^n at iteration n:

$$crit^n = ||\rho_1 - \rho_1^n||_\infty$$

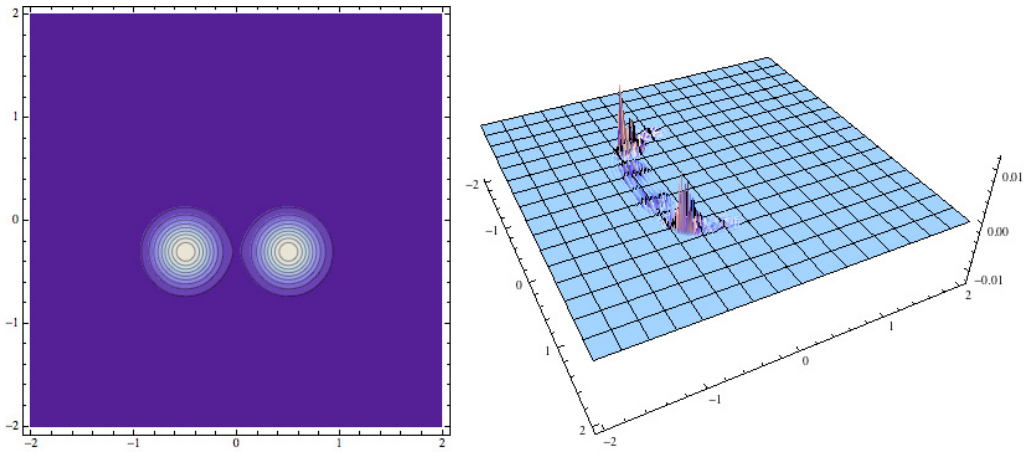

FIGURE 9. TC3. Final density obtained with Algorithm 2 ($\alpha = 1$) and error.

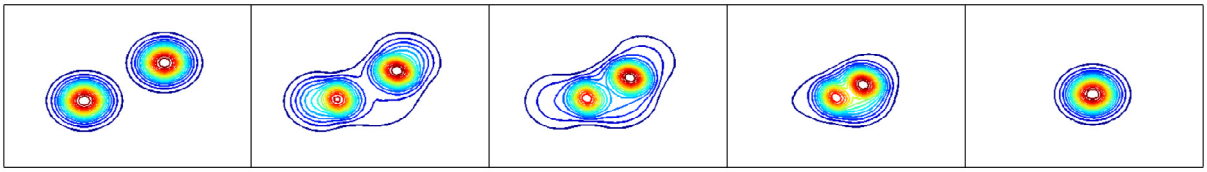

FIGURE 10. TC3. Plot of the isolines of the density $\rho(t, x)$ along the optimal path computed with Algorithm 3.

TABLE 1. CPU time: BB method *vs*. Our methods.

	BB method		Eulerian method		Lagrangian method	
	$crit_b^n$	CPU time	$crit^n$	CPU time	$crit^n$	CPU time
Initial Criteria (IC)	0.707107	0 m0 s	2.063231	0 m0 s	2.063231	0 m0 s
$\approx IC/30$	0.023547	2 m36 s	0.06004	0 m4.13 s	0.060026	0 m2.6 s
$\approx IC/60$	0.011784	5 m12 s	0.030851	0 m5.32 s	0.034312	0 m3.2 s
$IC/60 < IC \leq IC/100$	0.009841	19 m31 s	0.018421	0 m6.22 s	0.027887	0 m3.74 s

in the Eulerian case. For the Lagrangian case we consider the equivalent error based on the initial density. We consider TC1. For the BB algorithm, we discretize the time domain using 32 nodes and the space domain with 200×200 grid points. We take optimal numerical parameters in the Uzawa iteration in order to converge as quickly as possible.

Table 1 shows the evolution of the CPU time for different convergence thresholds. We can see that CPU time is very small for the present methods compared to the BB method. Note also that it is impossible to reduce the initial criteria by 100 in a reasonable time for BB. This is the reason we consider a convergence $IC/60 < IC \leq IC/100$ in the table above. Usually, in order to get to convergence of the BB algorithm ($IC/100$), we have about 3000 Uzawa iterations with a computing time larger than three hours.

Cases where the Wasserstein distance between the densities is small are particularly favorable for the Newton solution because there is no need to regularize the solution by adding an initial constant ζ to both the initial and final distributions. In cases where the Wasserstein distance is larger, like TC3, the computational advantage of the continuation method with respect to the BB scheme is reduced (initially, $\zeta = 1$). For example, in TC3

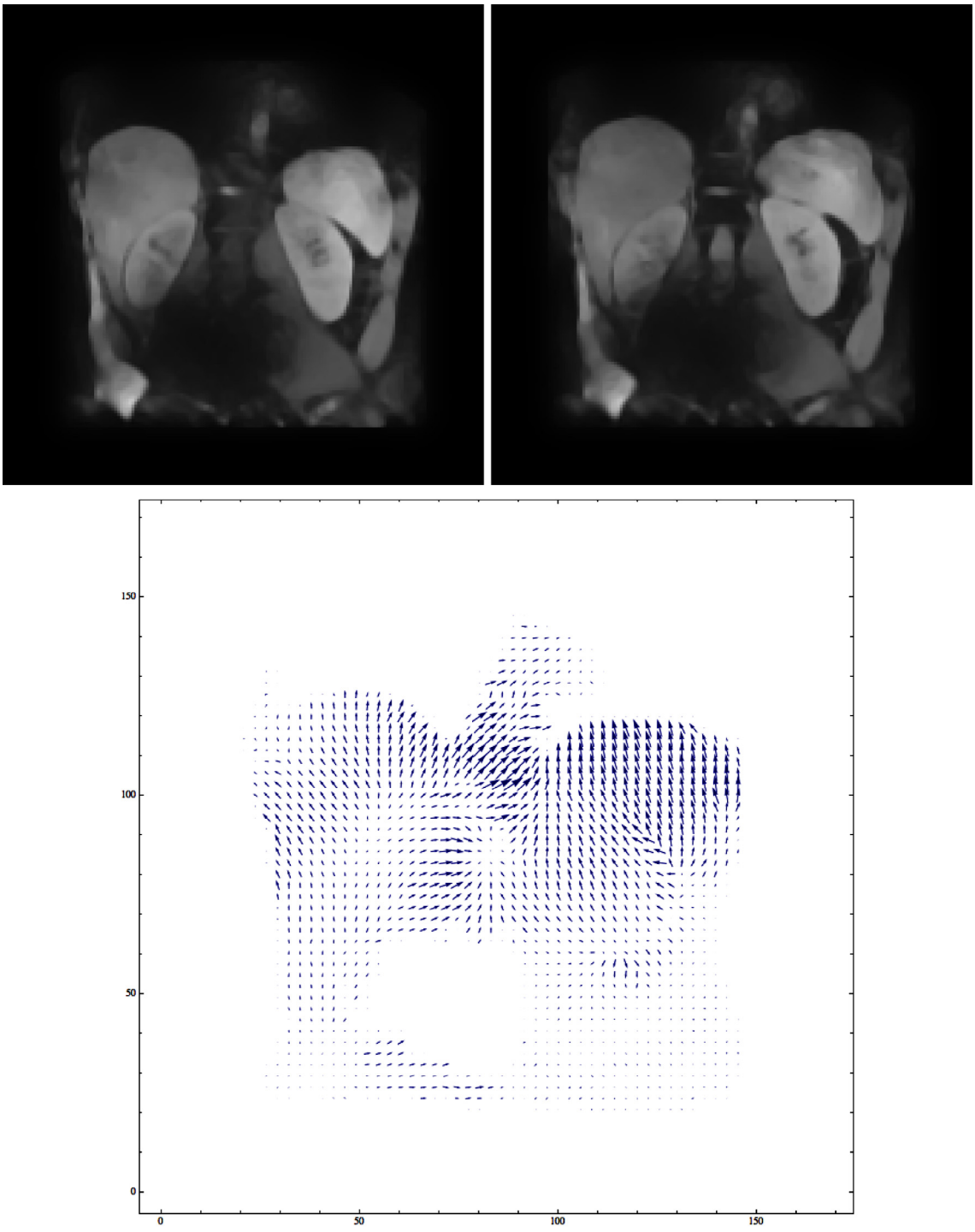

FIGURE 11. TC4. *Top row*: Initial and final grey-scale densities. *Bottom*: Optimal mapping.

the solution is obtained with the Lagrangian scheme in about 5 min CPU time, where the BB scheme takes 20 min to get to $IC/60 < IC \leq IC/100$. Still, the continuation method is more advantageous, whereas the pure Newton iteration does not converge in this case.

In conclusion, the continuation method that we presented can lead to a significant improvements in terms of convergence rate over the BB scheme, in particular when the Wasserstein distance is small. This can be a significant advantages for large three-dimensional problems in imagery and computational mechanics.

References

[1] S. Angenent, S. Haker and A. Tennenbaum, Minimizing flows for the Monge–Kantorovich problem. *SIAM J. Math. Anal.* **35** (2003) 61–97.

[2] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Mat.* **84** (2000) 375–393.

[3] J.-D. Benamou, A. Oberman and B. Froese, *Numerical solution of the second boundary value problem for the Elliptic Monge-Ampère equation*. Rapport de recherche (2012).

[4] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions. *Commun. Pure Appl. Math.* **64** (1991) 375–417.

[5] E.J. Dean and R. Glowinski, Numerical methods for fully nonlinear elliptic equations of the mongeampre type. *Comput. Methods Appl. Mech. Eng.* **195** (2006) 1344–1386.

[6] A. Iollo and D. Lombardi, A Lagrangian scheme for the solution of the optimal mass transfer problem. *J. Comput. Phys.* **230** (2011) 3430–3442.

[7] G. Loeper and Francesca Rapetti, Numerical solution of the Monge-Ampere equation by a newton's algorithm. *C. R. Acad. Sci. Paris, Ser. I* **340** (2005) 319–324.

[8] G. Monge, Memoire sur la théorie des déblais et des remblais. Histoire de l'Académie des Sciences de Paris (1781).

[9] N. Papadakis, G. Peyré and E. Oudet, Optimal transport with proximal splitting. *SIAM J. Imaging Sci.* **7** (2014) 212–238.

[10] L.-P. Saumier, M. Agueh and B. Khouider. An efficient numerical algorithm for the l2 optimal transport problem with periodic densities. *IMA J. Appl. Math.* (2013).

[11] C. Villani, Topics in Optimal Transportation. American Mathematical Society, 1st edition (2003).

[12] C. Villani, Optimal Transport, old and new. Springer-Verlag, 1st edition (2009).

[13] L. Weynans and A. Magni, Consistency, accuracy and entropy behaviour of remeshed particle methods. *ESAIM: M2AN* **47** (2013) 57–81.