Tomographic reconstruction from a few views: a multi-marginal optimal transport approach

I. Abraham [*], R. Abraham, M. Bergounioux [†], G. Carlier, [‡]

September 18, 2014

Abstract

In this article, we focus on tomographic reconstruction. The problem is to determine the shape of the interior interface using a tomographic approach while very few X-ray radiographs are performed. We use a multi-marginal optimal transport approach. Preliminary numerical results are presented.

Keywords: tomographic reconstruction, multi-marginal optimal transport.

1 Introduction

In this article, we focus on a specific application of tomographic reconstruction for a physical experiment whose goal is to study the behavior of a material under a shock. The problem is to determine the density. For this purpose, very few X-ray radiographs are performed, and the density of the object must then be reconstructed using a tomographic approach (see Figure 1.1).

In [3] we mentioned that several techniques exist for tomographic reconstruction, providing an analytic formula for the solution (see for instance [15] or [13]) as soon as a large number of projections of the object, taken from different angles, are available. There is a huge literature about theoretical and practical aspects of the problem of reconstruction from projections, the applications of which concern medicine, optics, material science, astronomy, geophysics, and magnetic resonance imaging (see [6]). When only few projections are known, these methods cannot be used directly, and some alternative methods have been proposed to reconstruct the densities (see for instance [12]).

As in any tomographic reconstruction process, this problem leads to an ill-posed inverse problem (see [20]). As X-rays must cross a very dense object and only a few number of them arrive at the detector, it is therefore necessary to add some amplification devices and very sensitive detectors, which cause a high noise level [25, 24] .

---

[*]CEA Ile de France- BP 12, 91680 Bruyères le Châtel, France `Isabelle.Abraham@cea.fr`.

[†]Université d'Orléans, UFR Sciences, MAPMO, UMR 7349, Route de Chartres, BP 6759, 45067 Orléans cedex 2, France, `romain.abraham@univ-orleans.fr, maitine.bergounioux@univ-orleans.fr`.

[‡]CEREMADE, UMR CNRS 7534, Université Paris IX Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, France `carlier@ceremade.dauphine.fr`.
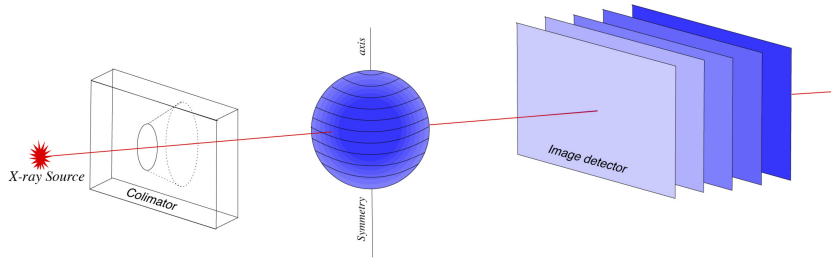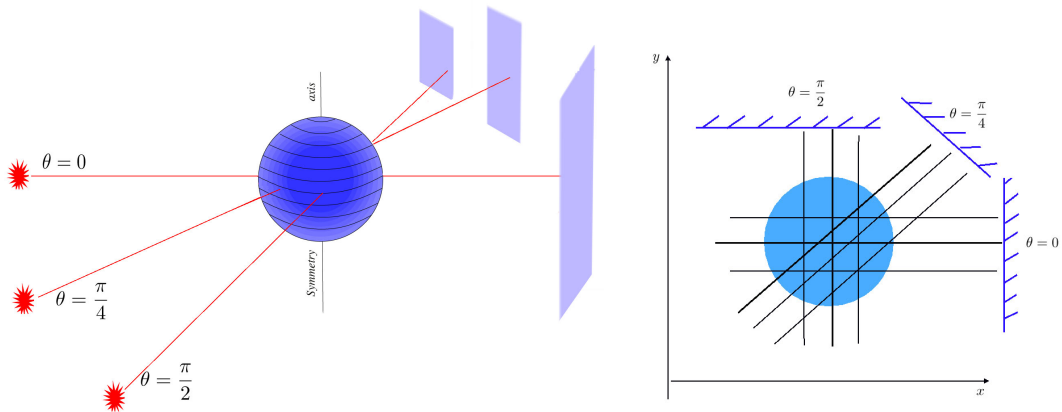
Figure 1.1: Tomography experiment

Figure 1.2: Different projections around the tomography axis

The tomographic reconstruction with few views problem has been widely studied. If a large number of radiographs is available, we can use several efficient methods that lead to exact formulas to compute the solution (see [21, 22] or [15]).

Missing data problems can been studied with such methods as well ([22], chapter 6 or [26]). It is the case, for example, when the object is measured on a subset of its support (so-called inner problem, see for example [9]). These techniques, as, for instance, the back-filtered projection (in the full case) or the back-projection for the projection derivatives (in the missing data case [23]) require a fine sampling of measures (here radiographs) to be performing ([22], chapter 4). Therefore, they are not useful in the case where few projection data are available.

The number of available projections (views) is closely related to the ill-posedness of the reconstruction problem. Indeed, the smaller the number of data is, the larger is the kernel of the related operator. Roughly speaking, there are an infinity of solutions and this infinity is linked to the kernel dimension. Some methods have been proposed that allow a partial reconstruction of the object [12]. In the case where we deal with specific objects there exists methods selecting a solution with respect to some prior : in [17], [16] the authors use a bayesian model while an optimization approach is used in [5],[4] where the problem is modelled as a minimal cost flow problem.

Recently a Mumford-Shah like method has been investigated [19, 27]. In [3] we have assumed that the initial physical setup is axially symmetric so that a single radiograph suffices in theory

to reconstruct the 3D object. The inverse problem remains ill-posed : existence and uniqueness of a solution are ensured but there is a lack of stability. However, interesting results have been obtained with a variational method ([7, 3]).

In the present paper, we do not assume that the object is axially symmetric any longer but we have more than one radiograph. However, due to the experimental setup, we only deal with very few radiographs, taken from three angles that we suppose to be $0, \frac{\pi}{4}$ and $\frac{\pi}{2}$ for sake of simplicity. So the regularization term to choose is not straightforward. The previously quoted methods are efficient as soon as we have much more data sets (projections) than we have. In [1] we used a variational method involving regularization terms that were not necessarily consistent with the physical point of view and numerical results were not convincing.

In the present paper we investigate a different modelling using optimal transport method. To our knowledge, this point of view is new in this context. We consider that the data are twofold : on one hand, one has the finite number of projected images of the objet to recover that lie in a n-1 D space (if the object is a nD one). On the other hand, we assume that we have a prior ρ_0 that can be the result of numerical simulation. We want to transport ρ_0 to some *better* object while fitting the projected data.

The paper is organized as follows. We first make this optimal transport model precise. Section 3 is devoted to the primal and dual formulation of the problem. We give a multi-marginal formulation in section 4. In last section we present numerical hints an some preliminary results.

2   An optimal transport model

In what follows, we assume that the X-sources are far enough from the object so that we may assume that the X-rays are parallel. Therefore we can separate the horizontal planes and reconstruct them independently (see Figure 2.1).
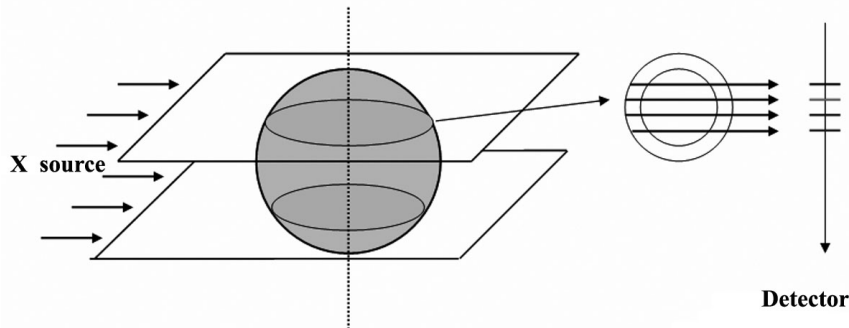

Figure 2.1: Parallel X-rays : the information along a detector segment depends on a planar slice of the object.

As mentioned in [3] that radiography measures the attenuation of X-rays through the object and the reconstruction of the object requires the inversion of the Radon transform restricted to any horizontal slice. Therefore, we focus now on the inversion in the 2D framework.

We do not formulate as a least squares problem any longer. So the Radon operator is not used in the fitting data term as it was in [1, 3, 7]. It plays a hidden role via the projected data and we do not want to inverse it. Our philosophy is different : we want to transport an objet ρ_0 to another that has the desired projected data. In the sequel we use the optimal

transport formalism. Moreover, we denote by bold letters the vectors and by regular ones the real numbers. The data of the tomographic reconstruction problem with a few views are:

- a prior ρ_0 that is a probability measure on $\mathbb{R}^d$ that has a finite second moment,

- a set of (unitary) directions $\mathbf{d}_i \in S^{d-1}$, $i = 1, \ldots, k$ and a corresponding collection of probability measures on the real line, π_i, $i = 1, \ldots, k$ having finite second moments.

We wish to reconstruct a probability measure ρ from the prior ρ_0 and the knowledge of the π_i's that correspond to X-ray measurement on lines whose directions are $\mathbf{d}_i$. Due to noise in the measurement, it is generally not the case that ρ_0 and the π_i's match perfectly, that is we do not assume that π_i coincides with the measure $\Pi_{\mathbf{d}_i \#} \rho_0$ defined by

$$\int_{\mathbb{R}} \varphi(t) \Pi_{\mathbf{d}_i \#} \rho_0(dt) = \int_{\mathbb{R}^d} \varphi(\mathbf{d}_i \cdot \mathbf{x}) \rho_0(d\mathbf{x}), \ \forall \varphi \in C_c(\mathbb{R}), \tag{2.1}$$

where $C_c(\mathbb{R})$, is the space of continuous functions with compact support.

Recall that if T a map from $\mathbb{R}^n$ to $\mathbb{R}^p$ then $T_\# \mu$ denotes the push-forward of the measure μ through the map T (i.e. $T_\# \mu(A) := \mu(T^{-1}(A))$, for any $A \subset \mathbb{R}^p$).

Variational approaches consist in looking for ρ as the minimizer of a certain cost that takes into account the various data of the problem. The novelty of our approach is to consider, instead of more familiar L^2 criteria, optimal transport costs (for the quadratic cost eventhough other choices are possible without adding much difficulty to the problem). The variational problem we introduce below is very close to the Wasserstein barycenter problem studied in [2] for which we refer for certain proofs that can be adapted directly. Here, we shall rather emphasize the differences with respect to [2] both from a theoretical and numerical point of view due to the fact that the π_i's are one-dimensional.

3 The variational problem and its dual

3.1 Primal formulation

In the sequel, $\mathcal{P}_2(\mathbb{R}^d)$ (respectively $\mathcal{P}_2(\mathbb{R})$) will denote the set of probability measures on $\mathbb{R}^d$ ($\mathbb{R}$) with finite second moment. For μ and ν in $\mathcal{P}_2(\mathbb{R}^d)$, the squared-2-Wasserstein distance between μ and ν is by definition

$$W^2(\mu, \nu) := \inf_{\gamma \in \Gamma(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |\mathbf{x} - \mathbf{y}|^2 \gamma(d\mathbf{x}, d\mathbf{y})$$

where $\Gamma(\mu, \nu)$ denotes the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ having μ and ν as marginals. The fact that the previous infimum is attained is classical, also, we recall a useful dual formula due to Kantorovich (see for instance [28]) that enables one to express W^2 as

$$W^2(\mu, \nu) = \sup_{(f,g) \in C_b(\mathbb{R}^d) \times C_b(\mathbb{R}^d)} \left\{ \int_{\mathbb{R}^d} f d\mu + \int_{\mathbb{R}^d} g d\nu \ : \ f(\mathbf{x}) + g(\mathbf{y}) \leq |\mathbf{x} - \mathbf{y}|^2 \right\}$$

Slightly abusing notations, we shall also use the notations W^2 for the squared 2-Wasserstein distance between probability measures on the real line and the notation $\Gamma(\mu, \nu)$ for the set of probability measures having μ and ν as marginals *even if μ and ν are probability measures on spaces with different dimensions.*

Given positive weights $\lambda_0, \lambda_1, \ldots, \lambda_k$, $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$, and $\pi_i \in \mathcal{P}_2(\mathbb{R})$ for $i = 1, \ldots, k$, we consider as cost the weighted sum of squared 2-Wasserstein distances i.e.

$$J(\rho) := \frac{\lambda_0}{2} W_2^2(\rho_0, \rho) + \frac{1}{2} \sum_{i=1}^k \lambda_i W^2(\pi_i, \Pi_{\mathbf{d}_i \#} \rho).$$

For further use, let us remark that it is easy to see that one may express the one-dimensional squared 2-Wasserstein distance between π_i and $\Pi_{\mathbf{d}_i \#} \rho$ equivalently either as

$$W^2(\pi_i, \Pi_{\mathbf{d}_i \#} \rho) = \inf_{\gamma_i \in \Gamma(\pi_i, \Pi_{\mathbf{d}_i \#} \rho)} \int_{\mathbb{R}} (x_i - y_i)^2 \gamma_i(dx_i, dy_i)$$

or

$$W^2(\pi_i, \Pi_{\mathbf{d}_i \#} \rho) = \inf_{\eta_i \in \Gamma(\pi_i, \rho)} \int_{\mathbb{R}} (x_i - \mathbf{x} \cdot \mathbf{d}_i)^2 \eta_i(dx_i, d\mathbf{x}) \tag{3.1}$$

Our aim is to study the following (convex) minimization problem

$$(\mathcal{P}) \quad \inf_{\rho \in \mathcal{P}_2(\mathbb{R}^d)} J(\rho). \tag{3.2}$$

First, it is easy to see that the direct method of the calculus of variations applies (see Prop. 2.3 in [2] for details) so that

Theorem 3.1. $(\mathcal{P})$ *admits at least a minimizer.*

3.2 Dual formulation

To address uniqueness of the minimizer and further characterize it, we shall use a dual formulation (see [11]) as in [8] (compact case) and [2] ($\mathcal{P}_2$ framework as in the present paper). Let f_0 be a real-valued function defined on $\mathbb{R}^d$; we then define $f_0^{\lambda_0}$ by the infimal convolution formula:

$$f_0^{\lambda_0}(\mathbf{x_0}) := \inf_{\mathbf{x} \in \mathbb{R}^d} \left\{ \frac{\lambda_0}{2} |\mathbf{x_0} - \mathbf{x}|^2 - f_0(\mathbf{x}) \right\}, \ \forall \mathbf{x_0} \in \mathbb{R}^d. \tag{3.3}$$

Note as soon as it is finite, this infimum define a semiconcave function of $\mathbf{x_0}$. In a similar way, for $f_i \colon \mathbb{R} \to \mathbb{R}$, we define

$$f_i^{\lambda_i}(x_i) := \inf_{y_i \in \mathbb{R}^d} \left\{ \frac{\lambda_i}{2} (x_i - y_i)^2 - f_i(y_i) \right\}, \ \forall x_i \in \mathbb{R}.$$

By construction, one has

$$f_0(\mathbf{x}) + f_0^{\lambda_0}(\mathbf{x_0}) \leq \frac{\lambda_0}{2} |\mathbf{x_0} - \mathbf{x}|^2, \ f_i(y_i) + f_i^{\lambda_i}(x_i) \leq \frac{\lambda_i}{2} (x_i - y_i)^2. \tag{3.4}$$

Then define

$$F(f_0, f_1, \ldots, f_k) := \int_{\mathbb{R}^d} f_0^{\lambda_0}(\mathbf{x_0}) \rho_0(d\mathbf{x_0}) + \sum_{i=1}^k \int_{\mathbb{R}} f_i^{\lambda_i}(x_i) \pi_i(dx_i)$$

and consider the (concave) maximization problem

$$\sup_{(f_0, \ldots, f_k) \in K} F(f_0, f_1, \ldots, f_k) \tag{3.5}$$

where K consists of continuous functions that have at most quadratic growth at infinity and such that

$$f_0(\mathbf{x}) + \sum_{i=1}^{k} f_i(\mathbf{d}_i \cdot \mathbf{x}) = 0, \ \forall \mathbf{x} \in \mathbb{R}^d. \tag{3.6}$$

The fact that $\sup(\mathcal{P}^*) \leq \inf(\mathcal{P})$ is easy to see. Indeed, take $(f_0, \ldots, f_k) \in K$, $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, $\gamma_0 \in \Gamma(\rho_0, \rho)$ and $\eta_i \in \Gamma(\pi_i, \rho)$, and using (3.6), we have

$$\frac{\lambda_0}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} |\mathbf{x_0} - \mathbf{x}|^2 \gamma_0(d\mathbf{x_0}, d\mathbf{x}) + \sum_{i=1}^{k} \frac{\lambda_i}{2} \int_{\mathbb{R} \times \mathbb{R}} (x_i - \mathbf{x} \cdot \mathbf{d}_i)^2 \eta_i(dx_i, d\mathbf{x})$$

$$\geq \int_{\mathbb{R}^d} (f_0(\mathbf{x}) + \sum_{i=1}^{k} f_i(\mathbf{d}_i \cdot \mathbf{x})) \rho(d\mathbf{x}) + F(f_0, f_1, \ldots, f_k) = F(f_0, f_1, \ldots, f_k)$$

minimizing the left-hand side with respect to $\gamma_0 \in \Gamma(\rho_0, \rho)$ and $\eta_i \in \Gamma(\pi_i, \rho)$ immediately gives $\sup(\mathcal{P}^*) \leq \inf(\mathcal{P})$. The fact that there is no duality gap and that $(\mathcal{P}^*)$ has a maximizer requires some extra work for which we refer to [2]:

Theorem 3.2. *The following duality relation holds*

$$\inf(\mathcal{P}) = \sup(\mathcal{P}^*).$$

Moreover, $(\mathcal{P}^*)$ admits a solution $(f_0, f_1, \ldots, f_1)$ that can be chosen in such a way that the functions $v_1, \ldots, v_k$ defined by

$$v_i(t) := \frac{\lambda_i}{2} t^2 - f_i(t), \ t \in \mathbb{R}, \ i = 1, \ldots, k \tag{3.7}$$

are convex (which in particular implies that the functions $f_1, \ldots, f_k$ can be chosen semiconcave on $\mathbb{R}$ and f_0 semiconvex on $\mathbb{R}^d$).

From the previous duality result we may deduce (as usual with convex duality) the following optimality conditions. Let ρ solve $(\mathcal{P})$, $\gamma_0 \in \Gamma(\rho_0, \rho)$ be an optimal transport plan for $W^2(\rho_0, \rho)$ and $\eta_i \in \Gamma(\pi_i, \rho)$ be optimal in (3.1), we then have

$$J(\rho) = \frac{\lambda_0}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} |\mathbf{x_0} - \mathbf{x}|^2 \gamma_0(d\mathbf{x_0}, d\mathbf{x}) + \sum_{i=1}^{k} \frac{\lambda_i}{2} \int_{\mathbb{R} \times \mathbb{R}} (x_i - \mathbf{x} \cdot \mathbf{d}_i)^2 \eta_i(dx_i, d\mathbf{x})$$

$$= F(f_0, f_1 \ldots, f_k) = \int_{\mathbb{R}^d \times \mathbb{R}^d} (f_0^{\lambda_0}(\mathbf{x_0}) + f_0(\mathbf{x})) \gamma_0(d\mathbf{x_0}, d\mathbf{x}) +$$

$$\sum_{i=1}^{k} \int_{\mathbb{R} \times \mathbb{R}} (f_i^{\lambda_i}(x_i) + f_i(\mathbf{x} \cdot x_i)) \eta_i(dx_i, d\mathbf{x})$$

so that for γ_0-a.e. $(\mathbf{x_0}, \mathbf{x})$ one has

$$f_0^{\lambda_0}(\mathbf{x_0}) = \frac{\lambda_0}{2} |\mathbf{x_0} - \mathbf{x}|^2 - f_0(\mathbf{x}) = \inf_{\mathbf{z} \in \mathbb{R}^d} \left\{ \frac{\lambda_0}{2} |\mathbf{x_0} - \mathbf{z}|^2 - f_0(\mathbf{z}) \right\} \tag{3.8}$$

and, similarly for η_i-a.e. $(x_i, \mathbf{x})$,

$$f_i^{\lambda_i}(x_i) = \frac{\lambda_i}{2}(x_i - \mathbf{x} \cdot \mathbf{d}_i)^2 - f_i(\mathbf{x} \cdot x_i) = \inf_{z \in \mathbb{R}} \left\{ \frac{\lambda_i}{2}(x_i - z)^2 - f_i(z) \right\}. \tag{3.9}$$

If $f_0^{\lambda_0}$ is differentiable at $\mathbf{x_0}$, it is immediate to deduce from (3.8) that

$$\nabla f_0^{\lambda_0}(\mathbf{x_0}) = \lambda_0(\mathbf{x_0} - \mathbf{x}), \text{ i.e. } \mathbf{x} = \mathbf{x_0} - \frac{1}{\lambda_0}\nabla f_0^{\lambda_0}(\mathbf{x_0}). \tag{3.10}$$

Similarly, if $f_i^{\lambda_i}$ is differentiable at x_i, one deduces from (3.9) that

$$(f_i^{\lambda_i})'(x_i) = \lambda_i(x_i - \mathbf{x} \cdot \mathbf{d}_i), \text{ i.e. } \mathbf{x} \cdot \mathbf{d}_i = x_i - \frac{1}{\lambda_i}(f_i^{\lambda_i})'(x_i). \tag{3.11}$$

Definition 3.3. *A probability measure μ on $\mathbb{R}^d$ is said to vanish on small sets if and only if $\mu(A) = 0$ for every Borel set A of $\mathbb{R}^d$, having Hausdorff dimension less than or equal to $d-1$.*

Since semiconvex functions are differentiable on the complement on a small set, one immediately deduces from the duality theorem 3.2 the following uniqueness result:

Corollary 3.4. *If ρ_0 vanishes on small sets, problem $(\mathcal{P})$ admits a unique solution ρ. More precisely, if $(f_0, f_1, \ldots, f_k)$ solves $(\mathcal{P}^*)$, then $\rho = T_{0\#}\rho_0$ where*

$$T_0(\mathbf{x_0}) := \mathbf{x_0} - \frac{1}{\lambda_0}\nabla f_0^{\lambda_0}(\mathbf{x_0}), \ \forall \mathbf{x_0} \in \mathbb{R}^d.$$

Proof. The proof is a straightforward consequence of the fact that $f_0^{\lambda_0}$ being semiconvex, it is differentiable ρ_0 a.e. and formula (3.10). $\square$

Remark 3.5. Assume that ρ_0 vanishes on small sets, once one knows a solution of $(\mathcal{P}^*)$, one deduces the minimizer ρ of $(\mathcal{P})$ by the previous corollary. Now we claim that $(\mathcal{P}^*)$ is in principle easier to solve since it only involves the k potentials of one variable, $f_1, \ldots, f_k$ (from which one deduces f_0 by (3.6) and then $f_0^{\lambda_0}$ and thus finally ρ by the previous corollary).

4 Multi-marginal reformulation

Our aim now is to give an equivalent linear reformulation that takes the form of a multi-marginal optimal transport problem. For $\overline{\mathbf{x}} := (\mathbf{x_0}, x_1, \ldots, x_k) \in \mathbb{R}^d \times \mathbb{R}^k$, let us define

$$\overline{c}(\overline{\mathbf{x}}) := \inf_{\mathbf{x} \in \mathbb{R}^d} \left\{ \frac{\lambda_0}{2}|\mathbf{x_0} - \mathbf{x}|^2 + \sum_{i=1}^{k} \frac{\lambda_i}{2}(x_i - \mathbf{x} \cdot \mathbf{d}_i)^2 \right\}. \tag{4.1}$$

This quadratic problem has a unique minimizer that we denote $\overline{T}(\overline{\mathbf{x}})$; its expression is easy to compute and reads as

$$\overline{T}(\overline{\mathbf{x}}) = \left(\lambda_0 \operatorname{id} + \sum_{i=1}^{k} \lambda_i \mathbf{d}_i \otimes \mathbf{d}_i \right)^{-1} \left(\lambda_0 \mathbf{x_0} + \sum_{i=1}^{k} \lambda_i x_i \mathbf{d}_i \right). \tag{4.2}$$

Replacing and developing the squares then gives

$$\overline{c}(\overline{\mathbf{x}}) = \frac{\lambda_0}{2}|\mathbf{x_0}|^2 + \sum_{i=1}^{k} \frac{\lambda_i}{2}x_i^2 - \frac{1}{2}\overline{T}(\overline{\mathbf{x}}) \cdot \left(\lambda_0 \mathbf{x_0} + \sum_{i=1}^{k} \lambda_i x_i \mathbf{d}_i \right)$$

which we can rewrite in a more concise way as

$$\overline{c}(\overline{\mathbf{x}}) = \frac{\lambda_0}{2}|\mathbf{x_0}|^2 + \sum_{i=1}^{k} \frac{\lambda_i}{2} x_i^2 - \frac{1}{2} A z(\overline{\mathbf{x}}) \cdot z(\overline{\mathbf{x}}) \tag{4.3}$$

with

$$A := \left(\lambda_0 \operatorname{id} + \sum_{i=1}^{k} \lambda_i \mathbf{d}_i \otimes \mathbf{d}_i \right)^{-1} \text{ and } z(\overline{\mathbf{x}}) := \lambda_0 \mathbf{x_0} + \sum_{i=1}^{k} \lambda_i x_i \mathbf{d}_i. \tag{4.4}$$

The multi-marginal optimal transport problem then reads

$$\inf_{\gamma \in \Gamma(\rho_0, \pi_1, \ldots, \pi_k)} \int_{\mathbb{R}^d \times \mathbb{R}^k} \overline{c}(\overline{\mathbf{x}}) \gamma(d\overline{\mathbf{x}}) \tag{4.5}$$

where $\Gamma(\rho_0, \pi_1, \ldots, \pi_k)$ denotes the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^k$ having $\rho_0, \pi_1, \ldots, \pi_k$ as marginals and $\overline{c}$ is the cost computed above. Again the existence of an optimal measure for (4.5) is easy to prove.

4.1 Equivalence

The connection between $(\mathcal{P})$ and (4.5) is then given by

Proposition 4.1. *If $\overline{\gamma}$ solves (4.5) then $\overline{\rho} := \overline{T}_{\#} \overline{\gamma}$ (where $\overline{T}$ is given by (4.2)) solves $(\mathcal{P})$.*

Proof. Let $\overline{\gamma}$ be a solution of (4.5). Let $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, $\gamma_0 \in \Gamma(\rho_0, \rho)$ be an optimal transport plan for $W^2(\rho_0, \rho)$ and $\eta_i \in \Gamma(\pi_i, \rho)$ be optimal in (3.1). Using the disintegration theorem (see for instance [10]) we may write

$$\gamma_0 = \gamma_0^{\mathbf{x}} \otimes \rho, \ \eta_i = \eta_i^{\mathbf{x}} \otimes \rho.$$

Then define $\xi \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^k)$ by

$$\xi := \left(\gamma_0^{\mathbf{x}} \bigotimes_{i=1}^{k} \eta_i^{\mathbf{x}} \right) \otimes \rho$$

i.e. for every $\varphi \in C_c(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^k)$

$$\int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^k} \varphi d\xi = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d \times \mathbb{R}^k} \varphi(\mathbf{x_0}, \mathbf{x}, x_1, \ldots, x_k) \gamma_0^x(d\mathbf{x_0}) \eta_1^x(dx_1) \cdots \eta_k^x(dx_k) \right) \rho(d\mathbf{x}).$$

Let then $\gamma \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^k)$ be defined by:

$$\int_{\mathbb{R}^d \times \mathbb{R}^k} \psi d\gamma = \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^k} \psi(\mathbf{x_0}, x_1, \ldots, x_k) \xi(d\mathbf{x_0}, d\mathbf{x}, dx_1, \ldots, dx_k).$$

for every $\psi \in C_c(\mathbb{R}^d \times \mathbb{R}^k)$. By construction, the projection of ξ on the $(\mathbf{x_0}, \mathbf{x})$ variables (respectively $(x_i, \mathbf{x})$ variables) is γ_0 (respectively η_i), this implies in particular that $\gamma \in \Gamma(\rho_0, \pi_1, \ldots, \pi_k)$. We then have

$$J(\rho) = \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^k} \left(\frac{\lambda_0}{2}|\mathbf{x_0} - \mathbf{x}|^2 + \sum_{i=1}^{k} \frac{\lambda_i}{2}(x_i - \mathbf{x} \cdot \mathbf{d}_i)^2 \right) \xi(d\mathbf{x_0}, d\mathbf{x}, dx_1, \ldots, dx_k)$$

$$\geq \int_{\mathbb{R}^d \times \mathbb{R}^k} \overline{c}(\overline{\mathbf{x}}) \gamma(d\overline{\mathbf{x}}) \geq \int_{\mathbb{R}^d \times \mathbb{R}^k} \overline{c}(\overline{\mathbf{x}}) \overline{\gamma}(d\overline{\mathbf{x}})$$

(again using the concise notation $\overline{\mathbf{x}} := (\mathbf{x_0}, x_1, \ldots, x_k)$). Now for $\overline{\rho} := \overline{T}_\# \overline{\gamma}$, we have

$$J(\overline{\rho}) \leq \int_{\mathbb{R}^d \times \times \mathbb{R}^k} \Big(\frac{\lambda_0}{2} |\mathbf{x_0} - \overline{T}(\overline{\mathbf{x}})|^2 + \sum_{i=1}^k \frac{\lambda_i}{2} (x_i - \overline{T}(\overline{\mathbf{x}}) \cdot \mathbf{d}_i)^2 \Big) \overline{\gamma}(d\overline{\mathbf{x}})$$
$$= \int_{\mathbb{R}^d \times \mathbb{R}^k} \overline{c}(\overline{\mathbf{x}}) \overline{\gamma}(d\overline{\mathbf{x}})$$
.

Which proves the optimality of $\overline{\rho}$ in $(\mathcal{P})$.

$\square$

Remark 4.2. It directly follows from the previous proposition and the fact that $\overline{T}$ is linear that $\overline{\rho}$ has compact support as soon as ρ_0 and $\pi_1, \ldots, \pi_k$ are compactly supported, more precisely, recalling (4.2)-(4.4), we have

$$\mathrm{spt}(\overline{\rho}) \subset A(\lambda_0 \, \mathrm{spt}(\rho_0) + \sum_{i=1}^k \lambda_i \, \mathrm{spt}(\pi_i) \mathbf{d}_i).$$

4.2 Duality

It follows from the expression (4.3) that (4.5) is equivalent to

$$\sup_{\gamma \in \Gamma(\rho_0, \pi_1, \ldots, \pi_k)} \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^k} Az(\overline{\mathbf{x}}) \cdot z(\overline{\mathbf{x}}) \gamma(d\overline{\mathbf{x}}) \qquad (4.6)$$

where the symmetric positive definite matrix A and the linear map z are defined in (4.4). It is well known (see for instance [14]) that this linear problem admits as dual formulation

$$\inf \left\{ \int_{\mathbb{R}^d} \mathbf{u}_0 \rho_0 + \sum_{i=1}^k \int_{\mathbb{R}} u_i \pi_i \ : \ \mathbf{u}_0(\mathbf{x_0}) + \sum_{i=1}^k u_i(x_i) \geq \frac{1}{2} Az(\overline{\mathbf{x}}) \cdot z(\overline{\mathbf{x}}), \ \forall \overline{\mathbf{x}} \in \mathbb{R}^d \times \mathbb{R}^k \right\} \qquad (4.7)$$

and that the latter admits a minimizer which satisfies

$$\mathbf{u}_0(\mathbf{x_0}) = \sup_{(x_1, \ldots x_k)} \left\{ \frac{1}{2} A(\lambda_0 \mathbf{x_0} + \sum_{i=1}^k \lambda_i x_i \mathbf{d}_i) \cdot (\lambda_0 \mathbf{x_0} + \sum_{i=1}^k \lambda_i x_i \mathbf{d}_i) - \sum_{i=1}^k u_i(x_i) \right\},$$

as well as

$$u_j(x_j) = \sup_{(\mathbf{x_0}, x_{i \neq j})} \left\{ \frac{1}{2} A(\lambda_0 \mathbf{x_0} + \sum_{i=1}^k \lambda_i x_i \mathbf{d}_i) \cdot (\lambda_0 \mathbf{x_0} + \sum_{i=1}^k \lambda_i x_i \mathbf{d}_i)) - \sum_{i=1, \, i \neq j}^k u_i(x_i) - \mathbf{u}_0(\mathbf{x_0}) \right\}$$

which is easily seen to imply that $\mathbf{x_0} \mapsto \mathbf{u}_0(\mathbf{x_0}) - \frac{\lambda_0^2}{2} A\mathbf{x_0} \cdot \mathbf{x_0}$ and $x_i \mapsto u_i(x_i) - \frac{\lambda_i^2}{2} (A\mathbf{d}_i \cdot \mathbf{d}_i) x_i^2$ are convex so that the potentials $\mathbf{u}_0, u_1, \ldots, u_k$ are strongly convex (i.e. have an Hessian that is bounded from below away from zero). By duality, if γ is optimal for (4.6) and $(\mathbf{u}_0, u_1, \ldots, u_k)$ solves (4.7), then for γ a.e. $\overline{\mathbf{x}} = (\mathbf{x_0}, x_1, \ldots, x_k)$ one has

$$\mathbf{u}_0(\mathbf{x_0}) + \sum_{i=1}^k u_i(x_i) = \frac{1}{2} Az(\overline{\mathbf{x}}) \cdot z(\overline{\mathbf{x}}).$$

So that if, in addition, $\mathbf{u_0}$ is differentiable at $\mathbf{x_0}$ and u_i is differentiable at x_i one has:

$$\nabla \mathbf{u_0}(\mathbf{x_0}) = \lambda_0 A z(\overline{\mathbf{x}}) \Rightarrow z(\overline{\mathbf{x}}) = \frac{A^{-1}\nabla \mathbf{u_0}(\mathbf{x_0})}{\lambda_0} \tag{4.8}$$

and (using the fact that u_i' is injective since u_i is strongly convex) we also have

$$u_i'(x_i) = \lambda_i A z(\overline{\mathbf{x}}) \cdot \mathbf{d}_i = \frac{\lambda_i}{\lambda_0}\nabla \mathbf{u_0}(\mathbf{x_0}) \cdot \mathbf{d}_i \Rightarrow x_i = (u_i')^{-1}\Big(\frac{\lambda_i}{\lambda_0}\nabla \mathbf{u_0}(\mathbf{x_0}) \cdot \mathbf{d}_i\Big). \tag{4.9}$$

4.3 Characterization and regularity

From the previous duality relations we deduce the following result on uniqueness of the optimal measure γ for (4.6) and that it is of Monge type (i.e. supported by a graph over the $\mathbf{x_0}$ variable):

Theorem 4.3. *Assume that ρ_0 vanishes on small sets at that π_i does not charge points for $i = 1, \ldots, k$ then (4.6) admits a unique solution γ that is of Monge-Type (i.e. induced by a map) and given by*

$$\gamma := \Psi_{\#}\rho_0$$

where

$$\Psi(\mathbf{x_0}) := \Big(\mathbf{x_0}, (u_1')^{-1}\Big(\frac{\lambda_1}{\lambda_0}\nabla \mathbf{u_0}(\mathbf{x_0}) \cdot \mathbf{d}_1\Big), \ldots, (u_k')^{-1}\Big(\frac{\lambda_k}{\lambda_0}\nabla \mathbf{u_0}(\mathbf{x_0}) \cdot \mathbf{d}_k\Big)\Big)$$

and the strongly convex potentials $\mathbf{u_0}, u_1, \ldots, u_k$ solve the dual problem (4.7).

Proof. The proof follows at once from the fact that $\mathbf{u_0}, u_1, \ldots, u_k$ are differentiable γ-almost everywhere (since these are convex potentials) and the relations (4.8)-(4.9). $\qquad \square$

Combining the previous with proposition 4.1 we immediately get

Corollary 4.4. *Under the assumptions of theorem 4.3, the solution $\overline{\rho}$ of $(\mathcal{P})$ is of Monge type and given by*

$$\overline{\rho} = \overline{F}_{\#}\rho_0$$

where

$$\overline{F}(\mathbf{x_0}) := A\Big(\lambda_0\mathbf{x_0} + \sum_{i=1}^{} \lambda_i(u_i')^{-1}\Big(\frac{\lambda_i}{\lambda_0}\nabla \mathbf{u_0}(\mathbf{x_0}) \cdot \mathbf{d}_i\Big)\mathbf{d}_i\Big). \tag{4.10}$$

In the sequel (especially for numerical tests) we use a formulation derived from equations (4.9) and (4.10) as $\overline{F}(\mathbf{x_0}) = \mathbf{x_0} + \delta\mathbf{x_0}$, where $\delta\mathbf{x_0}$ stands for the displacement. Indeed, these two equations yield

$$A^{-1}(\overline{F}(\mathbf{x_0})) = A^{-1}(\mathbf{x_0} + \delta\mathbf{x_0}) = \lambda_0\mathbf{x_0} + \sum_{i=1}^{k} \lambda_i x_i\, \mathbf{d}_i\ .$$

With (4.4), we get

$$A^{-1}(\delta\mathbf{x_0}) + \sum_{i=1}^{k} \lambda_i(\mathbf{d}_i \otimes \mathbf{d}_i)\mathbf{x_0} = \sum_{i=1}^{k} \lambda_i x_i\, \mathbf{d}_i$$

that is

$$A^{-1}(\delta\mathbf{x_0}) = \sum_{i=1}^{k} \lambda_i(x_i - \mathbf{x_0} \cdot \mathbf{d}_i)\mathbf{d}_i\ .$$

Finally

$$\overline{F}(\mathbf{x_0}) = \mathbf{x_0} + A\left(\sum_{i=1}^{k} \lambda_i(x_i - \mathbf{x_0} \cdot \mathbf{d}_i)\mathbf{d}_i\right) . \tag{4.11}$$

Proceeding as in [2], we deduce the following regularity result:

Theorem 4.5. *If, in addition to the assumptions of theorem 4.3, ρ_0 belongs to $L^\infty(\mathbb{R}^d)$ then $\overline{\rho}$ belongs to $L^\infty(\mathbb{R}^d)$ as well.*

Proof. Recalling formula (4.10) and using the fact that A is nonsingular, we see that it is enough to prove that $\nu := \overline{G}_\#\rho_0$ is L^∞ where $\overline{G}$ is defined by

$$\overline{G}(\mathbf{x_0}) := \lambda_0\mathbf{x_0} + \sum_{i=1}^{k} \lambda_i(u_i')^{-1}\left(\frac{\lambda_i}{\lambda_0}\nabla\mathbf{u_0}(\mathbf{x_0}) \cdot \mathbf{d}_i\right)$$

which, setting

$$\varphi_i(p) := \lambda_0 u_i^*\left(\frac{\lambda_i}{\lambda_0}p \cdot \mathbf{d}_i\right), \ \forall p \in \mathbb{R}^d$$

can be rewritten as

$$\overline{G} = \lambda_0 \,\text{id} + \sum_{i=1}^{k} \nabla\varphi_i \circ \nabla\mathbf{u_0}.$$

If the convex functions $\mathbf{u_0}$ and φ_i were smooth we could then write

$$D\overline{G} = \lambda_0 \,\text{id} + S_1 S_2$$

where S_1 and S_2 are symmetric positive definite which would imply that $\det(D\overline{G}) \geq \lambda_0^d$ hence that

$$\rho_0(\mathbf{x_0}) = \det(D\overline{G}(\mathbf{x_0}))\nu(\overline{G}(\mathbf{x_0})) \geq \lambda_0^d\nu(\overline{G}(\mathbf{x_0}))$$

which gives the desired L^∞ bound on ν. We refer to [2], proof of proposition 5.1, for a detailed regularization to recover this inequality in general.

$\square$

5 Numerical experiments

In this section, we give preliminary results. Indeed, the numerical realization is quite delicate and deserves a finer study (by comparing different points of view) that will be perfomed in a forthcoming paper. What follows is rather an validation of the model than a complete numerical investigation.

Let us describe the numerical process to solve problem (4.7)

$$\inf\left\{\int_{\mathbb{R}^d} \mathbf{u_0}\,\rho_0 + \sum_{i=1}^{k} \int_{\mathbb{R}} u_i\pi_i \ \Big| \ \mathbf{u_0}(\mathbf{x_0}) + \sum_{i=1}^{k} u_i(x_i) \geq \frac{1}{2}Az(\overline{\mathbf{x}}) \cdot z(\overline{\mathbf{x}}), \ \forall\overline{\mathbf{x}} \in \mathbb{R}^d \times \mathbb{R}^k\right\} .$$

where $\mathbf{u} = (u_1, \cdots, u_k)$, $\mathbf{x} = (x_1, \cdots, x_k) \in \mathbb{R}^k$, $\overline{\mathbf{x}} = (\mathbf{x_0}, \mathbf{x}) \in \mathbb{R}^d \times \mathbb{R}^k$, A, and $z(\overline{\mathbf{x}})$ are defined by equations (4.3) and (4.4).

The above constraint can be equivalently written as

$$\forall\mathbf{x_0} \in \mathbb{R}^d \qquad \mathbf{u_0}(\mathbf{x_0}) = -\min_{\mathbf{x}\in\mathbb{R}^k} \sum_{i=1}^{k} u_i(x_i) - \frac{1}{2}Az(\mathbf{x_0}, \mathbf{x}) \cdot z(\mathbf{x_0}, \mathbf{x}) , \tag{5.1}$$

so that problem (4.7) writes

$$\inf_{\mathbf{u}} \left\{ \sum_{i=1}^{k} \int_{\mathbb{R}} u_i \pi_i - \int_{\mathbb{R}^d} \min_{\mathbf{x} \in \mathbb{R}^k} \left(\sum_{i=1}^{k} u_i(x_i) - \frac{1}{2} Az(\mathbf{x_0}, \mathbf{x}) \cdot z(\mathbf{x_0}, \mathbf{x}) \right) \rho_0 \, d\mathbf{x_0} \right\} . \tag{5.2}$$

We set

$$H(\mathbf{u}, \mathbf{x_0}, \mathbf{x}) = \sum_{i=1}^{k} u_i(x_i) - \frac{1}{2} Az(\mathbf{x_0}, \mathbf{x}) \cdot z(\mathbf{x_0}, \mathbf{x})$$

and compute, if possible, a solution to

$$\min_{\mathbf{x} \in \mathbb{R}^k} H(\mathbf{u}, \mathbf{x_0}, \mathbf{x}) . \tag{5.3}$$

Note that the existence and uniqueness of solutions is not a priori ensured. It depends on the behaviour of the functions $\mathbf{u}$ with respect to the quadratic form $Az(\mathbf{x_0}, \mathbf{x}) \cdot z(\mathbf{x_0}, \mathbf{x})$. Indeed we know by the theory that the solution $\mathbf{u}$ is strongly convex. We must ensure coercivity for H. This will be the case if the strong convexity constant of $\mathbf{u}$ is greater that the one of the quadratic form. The latter is driven by the parameters $\lambda_i, i = 0, \cdots, k$. As the different directions $\mathbf{d}_i$ play the same role it is consistent to choose $\lambda_i = 1, i = 1, \cdots, k$. The only parameter to tune is $\lambda = \lambda_0$. Roughly speaking the spectral radius of A is driven by $1/\lambda$. Therefore, if λ is large then we should get coercivity for H (and uniqueness somehow, via strict convexity) and problem (5.3) has a unique solution : this corresponds to the case where displacements are small and so is the transportation. On the other hand, if λ is too small they may be no solution to (5.3). This remark is important from a numerical point of view to set the inital guess of an iterative process : functions u_i have to be quite large at infinity.

Assuming that problem (5.3) has at least a solution $\mathbf{x}^*$ we may use Euler equation and set $\nabla_{\mathbf{x}} H(\mathbf{u}, \mathbf{x_0}, \mathbf{x}^*) = 0$. A short computation gives :

$$\forall i = 1, \cdots, k \qquad \frac{\partial H}{\partial x_i}(\mathbf{u}, \mathbf{x_0}, \mathbf{x}) = u_i'(x_i) - \lambda_i Az(\mathbf{x_0}, \mathbf{x}) \cdot \mathbf{d}_i .$$

Therefore the solution $\mathbf{x}^*(\mathbf{u}, \mathbf{x_0})$ is implicitely given by the following system :

$$\forall i = 1, \cdots, k \quad \frac{\partial H}{\partial x_i}(\mathbf{u}, \mathbf{x_0}, \mathbf{x}^*) = u_i'(x_i^*) - \lambda_i Az(\mathbf{x_0}, x_1^*, \cdots, x_k^*) \cdot \mathbf{d}_i = 0 . \tag{5.4}$$

With (4.4) relation (5.4) writes:

$$\forall i = 1, \cdots, k \qquad u_i'(x_i^*) = \lambda_i \lambda_0 A\mathbf{x_0} \cdot \mathbf{d}_i + \sum_{p=1}^{k} \alpha_{i,p} x_p^* ,$$

where we have set $\alpha_{i,p} = \lambda_p (A\mathbf{d}_p \cdot \mathbf{d}_i)$ for $i, p = 1, \cdots, k$. Setting $\mathbb{A} = (a_{i,p})_{i,p=1,\cdots,k}$ gives that $\mathbf{x}^*(\mathbf{u}, \mathbf{x_0})$ is solution of the following system

$$\forall i = 1, \cdots, k \quad u_i'(x_i^*) = \lambda_0 A\mathbf{x_0} \cdot (\lambda_i \mathbf{d}_i) + (\mathbb{A} x^*)_i . \tag{5.5}$$

Once $\mathbf{x}^*(\mathbf{u}, \mathbf{x_0})$ is computed, the cost functional in problem (4.7) reads

$$\Phi(\mathbf{u}) = \sum_{i=1}^{k} \int_{\mathbb{R}} u_i \pi_i - \int_{\mathbb{R}^d} H(\mathbf{u}, \mathbf{x_0}, \mathbf{x}^*(\mathbf{u}, \mathbf{x_0})) \rho_0 \, d\mathbf{x_0} .$$

To use a numerical method to minimize Φ we need to compute $\nabla \Phi(\mathbf{u})$.
Let be $\varphi : \mathbb{R} \to \mathbb{R}$ and $i \in \{1, \cdots, k\}$.

$$\frac{\partial \Phi}{\partial u_i}(\mathbf{u}) \cdot \varphi = \int_{\mathbb{R}} \varphi \pi_i - \int_{\mathbb{R}^d} \left(\frac{\partial H}{\partial u_i}((\mathbf{u}, \mathbf{x_0}, \mathbf{x}^*(\mathbf{u}, \mathbf{x_0})) \cdot \varphi \right) \rho_0 \, d\mathbf{x_0} \ .$$

The computation of $\dfrac{\partial H}{\partial u_i}((\mathbf{u}, \mathbf{x_0}, \mathbf{x}^*) \cdot \varphi$ gives

$$\frac{\partial H}{\partial u_i}(\mathbf{u}, \mathbf{x_0}, \mathbf{x}^*) \cdot \varphi = \varphi(x_i^*) + \sum_{j=1}^{k} \frac{\partial H}{\partial x_j}(\mathbf{u}, \mathbf{x_0}, \mathbf{x}^*) \frac{\partial x_j^*}{\partial u_i}(\mathbf{u}, \mathbf{x_0}, \mathbf{x}^*) \cdot \varphi \ .$$

Therefore

$$\forall i = 1, \cdots, k \quad \frac{\partial \Phi}{\partial u_i}(\mathbf{u}) \cdot \varphi = \int_{\mathbb{R}} \varphi \pi_i - \int_{\mathbb{R}^d} \varphi(x_i^*(\mathbf{u}, \mathbf{x_0})) \rho_0 \, d\mathbf{x_0} \ , \tag{5.6}$$

where $\mathbf{x}^*(\mathbf{u}, \mathbf{x_0})$ satisfies (5.5). We decided to use a Galerkin type method to approximate the solution. More precisely, we choose a suitable basis (FEM, spectral or spline) to decompose the function. Here, we decided to use a spline basis, so that u is described by very few scalar coefficients. In addition, such an approach allows to compute the integral quantities once at the begining of the process. The algorithm writes :

Algorithm 1

Given λ_i, $i = 0, \cdots, k$, $\mathbf{d}_i$, $i = 1, \cdots, k$. Compute integrals ($\mathbf{x_0}$ and φ are known), A and $\mathbb{A}$.

1. Choose $\mathbf{u}_0$

2. Iteration k : $\mathbf{u}_k$ has been computed

   (a) Compute $\mathbf{x}_k^*(\mathbf{u}_k, \mathbf{x_0})$ solving for every $i = 1, \cdots, k$

   $$u_{k,i}'(x_{k,i}^*) - (\mathbb{A}x^*)_i = \lambda_0 A \mathbf{x_0} \cdot (\lambda_i \mathbf{d}_i)$$

   (b) Compute $\mathbf{u}_{k+1}$

   $$\mathbf{u}_{k+1} = \mathbf{u}_k - \tau_k \nabla \Phi(\mathbf{u}_k).$$

We present here below two academic examples obtained with two views in orthogonal directions: $k = 2$, $\mathbf{d}_1 \perp \mathbf{d}_2$. In both cases $\lambda_i = 1$, $i = 1, 2$.

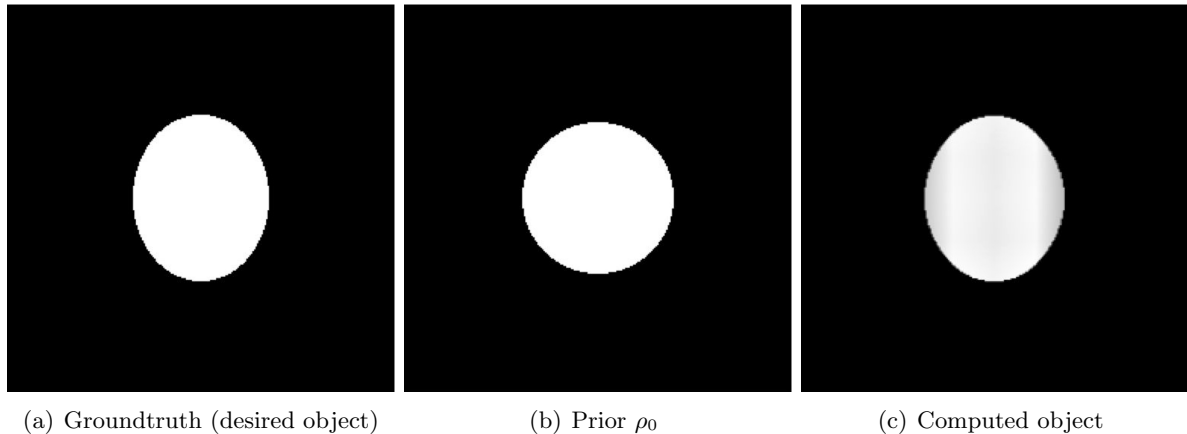(a) Groundtruth (desired object)   (b) Prior ρ_0 (c) Computed object

Figure 5.1: Ellipse example

The method is promising but preliminary tests show a great lack of numerical stability . Numerical realization appears to be quite challenging and need further investigation.

6 Conclusion

We proposed an original model based on optimal transport theory that allows to perform tomography reconstruction of an object once a prior is given, in the case there are few radiographs. The main novelty lies both in the use of optimal transport to model the problem and the fact that marginal data belong to different spaces.

The numerical realization is not straightforward and has to be further investigated. This will be done in a future work.

References

[1] I. Abraham, R. Abraham, M. Bergounioux, *A variational method for tomographic reconstruction with few views* (2012), http://hal.archives-ouvertes.fr/hal-00697218

[2] M. Agueh, G. Carlier, *Barycenters in the Wasserstein space*, SIAM J. Math. An., **43**, no. 2, pp. 904–924 (2011).

[3] R. Abraham, M. Bergounioux, and E. Trélat, *A penalization approach for tomographic reconstruction of binary axially symmetric objects*, Applied Mathematics and Optimization (2008), no. 58:345-371.

[4] K.J. Batenburg, *Reconstructing binary images from discrete x-rays*, Tech. report, Leiden University, The Netherlands, 2004.

[5] K.J. Batenburg, *A network flow algorithm for binary image reconstruction from few projections*, (2006), no. 4245:6–97.

[6] R.H.T. Bates, K.L. Garden, and T.M. Peters, *Overview of computerized tomography with emphasis on future developments*, Proeedings IEEE, vol. 71, 1983, pp. 256–297.

[7] M. Bergounioux and E. Trélat, *A variational method using fractional order hilbert spaces for tomographic reconstruction of blurred and noised binary images*, Journal of Functional Analysis (2010), no. 259: 2296–2332.

[8] G. Carlier, I. Ekeland, *Matching for teams*, Econ. Theory, **42**, no. 2, pp. 397–418 (2010).

[9] M. Courdurier, F. Noo, M. Defrise, and H. Kudo, *Solving the interior problem of computed tomography using a priori knowledge*, Inverse Problems (2008), no. 24.

[10] C. Dellacherie, P.-A. Meyer, *Probabilities and Potential*, Mathematical Studies **29**, North-Holland, 1978.

[11] I. Ekeland, R. Temam, *Convex Analysis and Variational Problems*, Classics in Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, (1999).

[12] J.-M.Dinten, *Tomographie à partir d'un nombre limité de projections : régularisation par des champs markoviens*, Ph.D. thesis, Université de Paris–Sud, 1990.

[13] N.J. Dusaussoy, *Image reconstruction from projections*, SPIE's international symposium on optics, imaging and instrumentation, San Diego, 1984.

[14] W. Gangbo, A. Święch, *Optimal maps for the multidimensional Monge-Kantorovich problem*, Comm. Pure Appl. Math., **51** , no. 1, pp. 23–45 (1998).

[15] G. Herman, *Image reconstruction from projections: the fundamentals of computerized tomography*, Academic Press, 1980.

[16] Y. L. Hstao, G.T. Herman, and T. Gabor, *A coordinate ascent approach to tomographic reconstruction of label images from a few projections*, (2005), no. 151:184-197.

[17] Y. L. Hstao, G.T. Herman, and T. Gabor, *Discrete tomography withj a very few views, using gibbs priors and a marginal posterior mode approach*, (2005), no. 20:399-418.

[18] A. Kirsch, *An introduction to the mathematical theory of inverse problems*, Springer-verlag, New York, 1996.

[19] E. Klann, *A Mumford-Shah-Like Method for Limited Data Tomography with an Application to Electron Tomography*, SIAM J. Imaging Science, (2011), Vol. 4, No. 4, pp. 1029–1048.

[20] A.K. Louis, *Ghosts in tomography - the null space of Radon transform*, Mathematical Models in the Applied Sciences, 3 (1), pp. 1–10, 1981.

[21] F. Natterer, *The mathematics of computerized tomography*, SIAM, 2001.

[22] F. Natterer and F. Wübeling, *Mathematical methods in image reconstruction*, SIAM, 2001.

[23] F. Noo, R. Clackdoyle, and J.D. Pack, *A two-step hilbert transform method for 2d image reconstruction*, Physics in Medicine and Biology (2004), no. 49(17):3903-3923.

[24] D. Partouche-Sebban and I. Abraham, *Scintillateur pour dispositif d'imagerie, module scintillateur, dispositif d'imagerie avec un tel scintillateur et procédé de fabrication d'un scintillateur*, French patent no 2922319, April 2009.

[25] D. Partouche-Sebban, I. Abraham, S. Laurio and C.Missault, *Multi-mev flash radiography in shock physics experiments: Specific assemblages of monolithic scintillating crystals for use in ccd-based imagers*, X-Ray Optics and Instrumentation (2010), Article ID 156984–9p.

[26] E.T. Quinto, *Singularities of the x-ray transform and limited data tomography in $\mathbb{R}^2$ and $\mathbb{R}^3$*, SIAM J. Math. Anal. (1993), no. 24:1215-1225.

[27] R. Ramlau and W. Ring, *A Mumford-Shah level-set approach for the inversion and segmentation of X-ray tomography data*, Journal of Computational Physics 221 (2007) 539-557

[28] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, 58, American Mathematical Society, Providence, RI, 2003.