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

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

April 21, 2015

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. For this purpose, X-ray radiographs are performed, and the density of the object must then be reconstructed using a tomographic approach (see Figure 1.1).

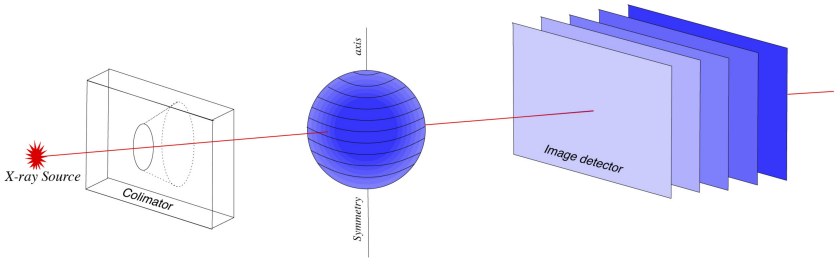


Figure 1.1: Tomography experiment

[*]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`.

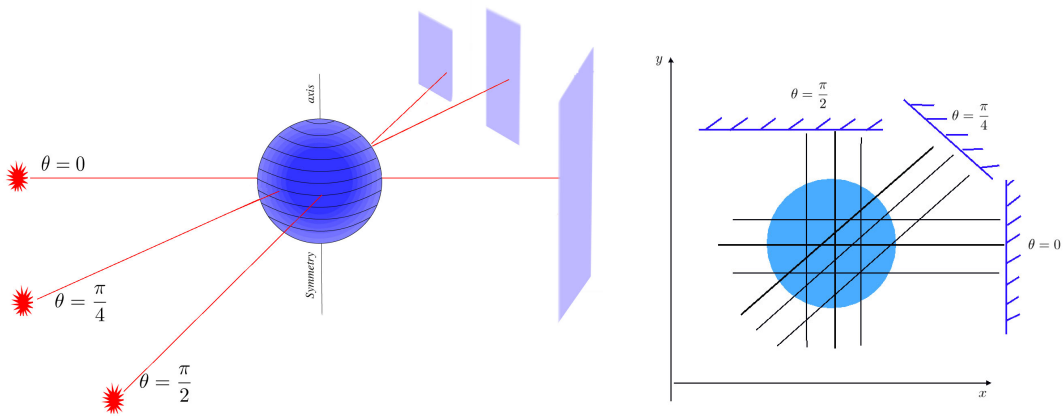


Figure 1.2: Different projections around the tomography axis

From a theoretical point of view, all the radiographs for a range of angles $\theta \in [0, \pi]$ are needed to obtain exact formulas to compute the solution (see [16, 17] or [13]). In practice, a finite (large) number of such radiographs is enough to get a good approximation of the solution (see ([17], chapter 6 or [20]). Unfortunately, due to the experimental setup, only very few radiographs (two or three) are available in our case (Figure 1.2).

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 [11]. In the case where we deal with specific objects there exists methods selecting a solution with respect to some prior : in [15], [14] the authors use a bayesian model while an optimization approach is used in [4],[3] where the problem is modeled as a minimal cost flow problem.

In [2] 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 based on the minimization of a penalized least squares cost function ([5, 2]).

In the present paper, we abandon the axisymmetry assumption and we investigate a different modeling using an optimal transport approach (for an overview of optimal transport theory, see the textbooks of Villani [23] or Santambrogio [22]). To our knowledge, this point of view is new in the present tomography reconstruction context.

We first suppose that the studied object is described by a function (with compact support) $f : \mathbb{R}^d \longmapsto \mathbb{R}$ that gives the attenuation coefficient of the material at the current point, and we set $\rho(d\mathbf{x}) = cf(\mathbf{x})d\mathbf{x}$ where c is a normalizing constant so that ρ is a probability measure on $\mathbb{R}^d$ (we denote by bold letters the vectors and by regular ones the real numbers). In what follows, we want to recover ρ from the data and some prior, and we will search ρ in the set of probability measures on $\mathbb{R}^d$ (with finite second moments).

The data of tomographic reconstruction problem are

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

The probability π_i represents the X-ray measurement on a line whose direction is $\mathbf{d_i}$. Let us denote by $\Pi_{\mathbf{d_i}}$ the projection operator on the line directed by $\mathbf{d_i}$ and by $T_{\#}\mu$ the push-forward of the measure μ through the map T: T is a map from $\mathbb{R}^n$ to $\mathbb{R}^p$, such that for every Borel subset A of $\mathbb{R}^p$ $T_{\#}\mu(A) = \mu(T^{-1}(A))$.Then, without noise perturbation, , we have

$$\pi_i = \Pi_{\mathbf{d}_i \#}\rho$$

or equivalently, for every continuous function φ on $\mathbb{R}$ with compact support

$$\int_{\mathbb{R}} \varphi(t)\pi_i(dt) = \int_{\mathbb{R}^d} \varphi(\mathbf{d}_i \cdot \mathbf{x})\rho(d\mathbf{x}).$$

In practice, 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 [19, 18]. Therefore, these theoretical projections are perturbed by some noise that we do not need to model in our approach.

We focus here on one-dimensional data π_i since we suppose that the incident X-rays are parallel. As a consequence, the horizontal planes are independent and we can reconstruct each of them separately (hence considering a two-dimensional object and one-dimensional data, see Figure 1.3).

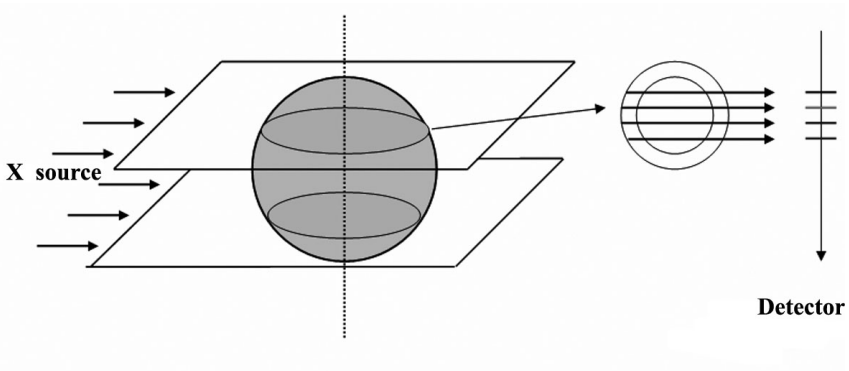


Figure 1.3: Parallel X-rays : the information along a detector segment depends on a planar slice of the object.

This simplification would not be valid anymore if we add a regularization term in our assumptions that would link the different slices of the object together.

As said before, the measure ρ is not uniquely determined by these few projections and we must add some additional assumption to be able to perform the desired reconstruction. Here, we suppose that we have a prior ρ_0 (which is a theoretical object that can be for

instance obtained by numerical simulations) and that the real object ρ is *close* to this prior.

The novelty of our approach is to consider, instead of more familiar L^2 criteria, optimal transport costs. We are looking for a probability measure ρ whose transportation cost from ρ_0 is small and whose projections match the data π_i. As these data are noisy, this matching cannot be perfect and we introduce a variational problem that contains the squared Wasserstein distance between the object ρ and the prior ρ_0 as well as the squared Wasserstein distances between the projections of ρ and the data π_i.

This variational problem (see Section 2 for a precise formulation) is very close to the Wasserstein barycenter problem studied in [1] for which we refer for certain proofs that can be adapted directly. Here, we shall rather emphasize the differences with respect to [1] both from a theoretical and numerical point of view due to the fact that the π_i's are one-dimensional. Another related problem, which has applications in texture mixing is the so-called sliced-Wasserstein distance (see [21] and [7]). Due to the difficulty of computing Wasserstein distances in several dimensions, the sliced-Wasserstein distance, which roughly speaking is an average over directions of the corresponding one-dimensional Wasserstein distance was introduced in [21]. The sliced Wasserstein distance is easy to compute (unidimensional optimal transport is basically a sorting problem) and differentiate; so it can be used as a numerically tractable surrogate for multidimensional optimal transport. However, because of the distance to the multi-dimensional prior ρ_0 our variational problem cannot be brought down only to one-dimensional Wasserstein distances computations.

The paper is organized as follows. We first make the optimal transport model precise, give a dual formulation and prove existence and uniqueness of the solution. However, the numerical computation of the problem may costful (many inf-convolution process to compute). So, we give an equivalent multi-marginal formulation of the same problem in Section 3 and the associated dual problem, which seem easier to handle. In the last section, we present numerical hints based on the dual formulation of the multi-marginal problem and some preliminary results.

2 The variational problem and its dual

2.1 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 [23], [22]) 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.

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

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

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

or

$$W^2(\pi_i, \Pi_{\mathbf{d}_i \#}\rho) = \inf_{\eta_i \in \Gamma(\pi_i, \rho)} \int_{\mathbb{R}\times\mathbb{R}^d} (x_i - \mathbf{x}\cdot\mathbf{d}_i)^2 \eta_i(dx_i, d\mathbf{x}). \tag{2.1}$$

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

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

The direct method of the calculus of variations applies (see Prop. 2.3 in [1] for details) so that

Theorem 2.1. *($\mathcal{P}$) admits at least a minimizer.*

Remark 2.2. We chose to take into account the constraints that $\Pi_{\mathbf{d}_i \#}\rho$ should be close to π_i via penalized terms with parameters λ_i in the functional J. An alternative modelling would be to consider

$$\inf \left\{ W^2(\rho_0, \rho) \ : \ \rho \in \mathcal{P}_2(\mathbb{R}^d), \ \ W_2(\pi_i, \Pi_{\mathbf{d}_i \#}\rho) \leq \sigma_i, i = 1, \cdots, k \right\}$$

where $\sigma_i, i = 1, \cdots, k$ is a relaxation parameter (that may represent a noise level for example). Interpreting the weights λ_i as multipliers associated to the previous constraints, both formulations are actually equivalent. For the sake of the exposition, we have prefered to work with the penalized problem rather than with the constrained one.

2.2 Dual formulation

To address uniqueness of the minimizer and further characterize it, we shall use a dual formulation (see [10]) as in [8] (compact case) and [1] ($\mathcal{P}_2$ 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:

$$f_0^{\lambda_0}(\mathbf{x_0}) := \inf_{\mathbf{x}\in\mathbb{R}^d} \left\{ \frac{\lambda_0}{2}|\mathbf{x_0}-\mathbf{x}|^2 - f_0(\mathbf{x}) \right\}, \ \forall \mathbf{x_0} \in \mathbb{R}^d \ . \tag{2.2}$$

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

$$f_0(\mathbf{x}) + f_0^{\lambda_0}(\mathbf{x_0}) \leq \frac{\lambda_0}{2}|\mathbf{x_0}-\mathbf{x}|^2, \ f_i(y_i) + f_i^{\lambda_i}(x_i) \leq \frac{\lambda_i}{2}(x_i-y_i)^2. \tag{2.3}$$

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

$$\sup_{(f_0,\ldots,f_k)\in K} F(f_0, f_1, \ldots, f_k) \tag{$\mathcal{P}^*$}$$

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

$$f_0(\mathbf{x}) + \sum_{i=1}^k f_i(\mathbf{d}_i\cdot\mathbf{x}) = 0, \ \forall \mathbf{x} \in \mathbb{R}^d. \tag{2.4}$$

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 (2.4), we have

$$\begin{aligned}\frac{\lambda_0}{2}\int_{\mathbb{R}^d\times\mathbb{R}^d} |\mathbf{x_0}-\mathbf{x}|^2\gamma_0(d\mathbf{x_0}, d\mathbf{x}) + \sum_{i=1}^k \frac{\lambda_i}{2}\int_{\mathbb{R}\times\mathbb{R}} (x_i - \mathbf{x}\cdot\mathbf{d}_i)^2\eta_i(dx_i, d\mathbf{x}) \\ \geq \int_{\mathbb{R}^d} (f_0(\mathbf{x}) + \sum_{i=1}^k f_i(\mathbf{d}_i\cdot\mathbf{x}))\rho(d\mathbf{x}) + F(f_0, f_1, \ldots, f_k) = F(f_0, f_1, \ldots, f_k).\end{aligned}$$

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 [1]:

Theorem 2.3. *The following duality relation holds*

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

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

$$v_i(t) := \frac{\lambda_i}{2} t^2 - f_i(t),\ t \in \mathbb{R},\ i = 1, \dots, k \tag{2.5}$$

are convex (which in particular implies that the functions $f_1, \dots, f_k$ *can be chosen semi-concave 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 (2.1), we then have

$$\begin{aligned} J(\rho) &= \frac{\lambda_0}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} |\mathbf{x_0} - \mathbf{x}|^2 \gamma_0(d\mathbf{x_0}, d\mathbf{x}) + \sum_{i=1}^k \frac{\lambda_i}{2} \int_{\mathbb{R}\times\mathbb{R}} (x_i - \mathbf{x} \cdot \mathbf{d}_i)^2 \eta_i(dx_i, d\mathbf{x}) \\ &= F(f_0, f_1 \dots, f_k) \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} (f_0^{\lambda_0}(\mathbf{x_0}) + f_0(\mathbf{x})) \gamma_0(d\mathbf{x_0}, d\mathbf{x}) + \sum_{i=1}^k \int_{\mathbb{R}\times\mathbb{R}} (f_i^{\lambda_i}(x_i) + f_i(\mathbf{x} \cdot x_i)) \eta_i(dx_i, d\mathbf{x}), \end{aligned}$$

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

$$f_0^{\lambda_0}(\mathbf{x_0}) = \frac{\lambda_0}{2} |\mathbf{x_0} - \mathbf{x}|^2 - f_0(\mathbf{x}) = \inf_{\mathbf{z} \in \mathbb{R}^d} \left\{ \frac{\lambda_0}{2} |\mathbf{x_0} - \mathbf{z}|^2 - f_0(\mathbf{z}) \right\} \tag{2.6}$$

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

$$f_i^{\lambda_i}(x_i) = \frac{\lambda_i}{2} (x_i - \mathbf{x} \cdot \mathbf{d}_i)^2 - f_i(\mathbf{x} \cdot x_i) = \inf_{z \in \mathbb{R}} \left\{ \frac{\lambda_i}{2} (x_i - z)^2 - f_i(z) \right\}. \tag{2.7}$$

If $f_0^{\lambda_0}$ is differentiable at $\mathbf{x_0}$, we deduce from (2.6) that

$$\nabla f_0^{\lambda_0}(\mathbf{x_0}) = \lambda_0(\mathbf{x_0} - \mathbf{x}),\ \text{i.e. } \mathbf{x} = \mathbf{x_0} - \frac{1}{\lambda_0} \nabla f_0^{\lambda_0}(\mathbf{x_0}). \tag{2.8}$$

Similarly, if $f_i^{\lambda_i}$ is differentiable at x_i, one gets from (2.7) that

$$(f_i^{\lambda_i})'(x_i) = \lambda_i(x_i - \mathbf{x} \cdot \mathbf{d}_i),\ \text{i.e. } \mathbf{x} \cdot \mathbf{d}_i = x_i - \frac{1}{\lambda_i} (f_i^{\lambda_i})'(x_i). \tag{2.9}$$

Definition 2.4. *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 2.3 the following uniqueness result:

Corollary 2.5. *If ρ_0 vanishes on small sets, problem $(\mathcal{P})$ admits a unique solution ρ. More precisely, if $(f_0, f_1, \dots, f_k)$ 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 (2.8). □

Remark 2.6. When ρ_0 vanishes on small sets, it is well-known that $\rho \mapsto W^2(\rho, \rho_0)$ is strictly convex (see for instance proposition 7.19 in [22]), and so is J as well. This can be used to give another uniqueness proof.

Assume that ρ_0 vanishes on small sets, as soon as one knows a solution of $(\mathcal{P}^*)$, the minimizer ρ of $(\mathcal{P})$ may be deduced by the previous corollary . At first glance, $(\mathcal{P}^*)$ looks simpler than $(\mathcal{P})$ because it only involves k functions of one variable (f_0 being deduced from $f_1, \dots, f_k$ by (2.4)). However, it is not easy to solve in practice since it involves the computation of infimal convolutions for all the potentials $f_0, f_1, \dots, f_k$. We propose in the next section a multi-marginal reformulation whose dual requires only one such infimum computation.

3 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 any $\overline{\mathbf{x}} := (\mathbf{x_0}, x_1, \dots, x_k) \in \mathbb{R}^d \times \mathbb{R}^k$, let us define

$$\overline{c}(\overline{\mathbf{x}}) := \inf_{\mathbf{x}\in\mathbb{R}^d} \left\{ \frac{\lambda_0}{2}|\mathbf{x_0} - \mathbf{x}|^2 + \sum_{i=1}^k \frac{\lambda_i}{2}(x_i - \mathbf{x}\cdot\mathbf{d}_i)^2 \right\}. \tag{3.1}$$

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

$$\overline{T}(\overline{\mathbf{x}}) = \left(\lambda_0\,\mathrm{id} + \sum_{i=1}^k \lambda_i \mathbf{d}_i \otimes \mathbf{d}_i\right)^{-1} (\lambda_0\mathbf{x_0} + \sum_{i=1}^k \lambda_i x_i \mathbf{d}_i). \tag{3.2}$$

Replacing and developing the squares then gives

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

which we can rewrite in a more concise way as

$$\overline{c}(\overline{\mathbf{x}}) = \frac{\lambda_0}{2}|\mathbf{x_0}|^2 + \sum_{i=1}^k \frac{\lambda_i}{2}x_i^2 - \frac{1}{2}Az(\overline{\mathbf{x}})\cdot z(\overline{\mathbf{x}}) \tag{3.3}$$

with

$$A := \left(\lambda_0\,\mathrm{id} + \sum_{i=1}^k \lambda_i \mathbf{d}_i \otimes \mathbf{d}_i\right)^{-1} \text{ and } z(\overline{\mathbf{x}}) := \lambda_0\mathbf{x_0} + \sum_{i=1}^k \lambda_i x_i \mathbf{d}_i. \tag{3.4}$$

The multi-marginal optimal transport problem then reads

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

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

3.1 Equivalence

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

Proposition 3.1. *If $\overline{\gamma}$ solves ($\mathcal{P}_m$) then $\overline{\rho} := \overline{T}_\#\overline{\gamma}$ (where $\overline{T}$ is given by (3.2)) solves ($\mathcal{P}$).*

Proof. Let $\overline{\gamma}$ be a solution of ($\mathcal{P}_m$). 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 (2.1). Using the disintegration theorem (see for instance [9]) 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 $\varphi\in C_c(\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^k)$

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

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

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

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 $\gamma\in\Gamma(\rho_0,\pi_1,\dots,\pi_k)$ and we have

$$\begin{aligned}J(\rho) &= \int_{\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^k}\Big(\frac{\lambda_0}{2}|\mathbf{x_0}-\mathbf{x}|^2+\sum_{i=1}^{k}\frac{\lambda_i}{2}(x_i-\mathbf{x}\cdot\mathbf{d}_i)^2\Big)\xi(d\mathbf{x_0},d\mathbf{x},dx_1,\dots,dx_k)\\ &\geq \int_{\mathbb{R}^d\times\mathbb{R}^k}\overline{c}(\overline{\mathbf{x}})\gamma(d\overline{\mathbf{x}}) \geq \int_{\mathbb{R}^d\times\mathbb{R}^k}\overline{c}(\overline{\mathbf{x}})\overline{\gamma}(d\overline{\mathbf{x}})\end{aligned}$$

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

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

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

□

Remark 3.2. It follows from the previous proposition and the fact that A is linear that $\overline{\rho}$ has compact support as soon as ρ_0 and $\pi_1, \dots, \pi_k$ are compactly supported. More precisely, recalling (3.2)-(3.4), we have

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

3.2 Duality

Thank to expression (3.3), problem $(\mathcal{P}_m)$ is equivalent to

$$\sup_{\gamma \in \Gamma(\rho_0, \pi_1, \dots, \pi_k)} \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^k} Az(\overline{\mathbf{x}}) \cdot z(\overline{\mathbf{x}}) \gamma(d\overline{\mathbf{x}}) \tag{3.5}$$

where the symmetric positive definite matrix A and the linear map z are defined in (3.4). It is well known (see for instance [12]) that this linear problem admits the dual formulation

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

and that the latter admits a minimizer which satisfies

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

as well as

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

It is easy seen 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 $\mathbf{u}_0,\ u_1, \dots, u_k$ are strongly convex (i.e. have an Hessian that is bounded from below away from zero). By duality, if γ is optimal for (3.5) and $(\mathbf{u}_0, u_1, \dots, u_k)$ solves $(\mathcal{P}_m^*)$, then for γ a.e. $\overline{\mathbf{x}} = (\mathbf{x_0}, x_1, \dots, x_k)$ one has

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

If, in addition, $\mathbf{u}_0$ is differentiable at $\mathbf{x_0}$ and u_i is differentiable at x_i one has:

$$\nabla \mathbf{u}_0(\mathbf{x_0}) = \lambda_0 A z(\overline{\mathbf{x}}) \Rightarrow z(\overline{\mathbf{x}}) = \frac{A^{-1}\nabla \mathbf{u}_0(\mathbf{x_0})}{\lambda_0} \tag{3.6}$$

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

$$u_i'(x_i) = \lambda_i A z(\overline{\mathbf{x}}) \cdot \mathbf{d}_i = \frac{\lambda_i}{\lambda_0}\nabla \mathbf{u}_0(\mathbf{x_0}) \cdot \mathbf{d}_i \Rightarrow x_i = (u_i')^{-1}\Big(\frac{\lambda_i}{\lambda_0}\nabla \mathbf{u}_0(\mathbf{x_0}) \cdot \mathbf{d}_i\Big). \tag{3.7}$$

3.3 Characterization and regularity

From the previous duality relations we deduce the following result on uniqueness of the optimal measure γ for (3.5) and that it is of Monge type (i.e. supported by a graph over the $\mathbf{x_0}$ variable):

Theorem 3.3. *Assume that ρ_0 vanishes on small sets and that π_i does not charge points for $i = 1, \dots, k$. Then (3.5) admits a unique solution γ that is of Monge-Type (i.e. induced by a map) and given by*

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

where

$$\Psi(\mathbf{x_0}) := \Big(\mathbf{x_0}, (u_1')^{-1}\Big(\frac{\lambda_1}{\lambda_0}\nabla \mathbf{u}_0(\mathbf{x_0}) \cdot \mathbf{d}_1\Big), \dots, (u_k')^{-1}\Big(\frac{\lambda_k}{\lambda_0}\nabla \mathbf{u}_0(\mathbf{x_0}) \cdot \mathbf{d}_k\Big)\Big)$$

and the strongly convex potentials $\mathbf{u}_0$, $u_1, \dots, u_k$ solve the dual problem ($\mathcal{P}_m^*$).

Proof. The proof comes from the fact that $\mathbf{u}_0$, $u_1, \dots, u_k$ are differentiable γ-almost everywhere (since these are convex potentials) and the relations (3.6)-(3.7). □

Combining the previous with proposition 3.1 we obtain

Corollary 3.4. *Under the assumptions of theorem 3.3, the solution $\overline{\rho}$ of $(\mathcal{P})$ is of Monge type and given by*

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

where

$$\overline{F}(\mathbf{x_0}) := A\Big(\lambda_0 \mathbf{x_0} + \sum_{i=1} \lambda_i (u_i')^{-1}\Big(\frac{\lambda_i}{\lambda_0}\nabla \mathbf{u}_0(\mathbf{x_0}) \cdot \mathbf{d}_i\Big)\mathbf{d}_i\Big). \tag{3.8}$$

In the sequel (especially for numerical tests) we use a formulation derived from equations (3.7) and(3.8) 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 (3.4), 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

$$\overline{F}(\mathbf{x_0}) = \mathbf{x_0} + A\left(\sum_{i=1}^{k} \lambda_i (x_i - \mathbf{x_0} \cdot \mathbf{d}_i)\mathbf{d}_i\right) \ . \tag{3.9}$$

Proceeding as in [1], we deduce the following regularity result:

Theorem 3.5. *If, in addition to the assumptions of Theorem 3.3, ρ_0 belongs to $L^\infty(\mathbb{R}^d)$ then $\overline{\rho}$ belongs to $L^\infty(\mathbb{R}^d)$ as well.*

Proof. Recalling formula (3.8) 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

$$\overline{G}(\mathbf{x_0}) := \lambda_0 \mathbf{x_0} + \sum_{i=1}^{k} \lambda_i (u_i')^{-1}\Big(\frac{\lambda_i}{\lambda_0}\nabla \mathbf{u}_0(\mathbf{x_0}) \cdot \mathbf{d}_i\Big)$$

which, setting

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

can be rewritten as

$$\overline{G} = \lambda_0 \,\mathrm{id} + \sum_{i=1}^{k} \nabla \varphi_i \circ \nabla \mathbf{u}_0.$$

If the convex functions $\mathbf{u}_0$ and φ_i were smooth we could then write

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

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 [1], proof of proposition 5.1, for a detailed regularization to recover this inequality in general.

□

4 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 performed in a forthcoming paper. What follows is rather a validation of the model than a complete numerical investigation.

Let us describe the numerical process to solve problem ($\mathcal{P}_m^*$) (which involves the computation of only one infimal convolution, whereas ($\mathcal{P}^*$) involves k of them)

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

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 (3.3) and (3.4).

The above constraint can be equivalently written as

$$\forall\mathbf{x_0}\in\mathbb{R}^d\qquad\mathbf{u_0}(\mathbf{x_0})=-\min_{\mathbf{x}\in\mathbb{R}^k}\sum_{i=1}^k u_i(x_i)-\frac{1}{2}Az(\mathbf{x_0},\mathbf{x})\cdot z(\mathbf{x_0},\mathbf{x})\ ,\tag{4.1}$$

so that problem ($\mathcal{P}_m^*$) writes

$$\inf_{\mathbf{u}}\left\{\sum_{i=1}^k\int_{\mathbb{R}}u_i\pi_i-\int_{\mathbb{R}^d}\min_{\mathbf{x}\in\mathbb{R}^k}\left(\sum_{i=1}^k u_i(x_i)-\frac{1}{2}Az(\mathbf{x_0},\mathbf{x})\cdot z(\mathbf{x_0},\mathbf{x})\right)\rho_0(d\mathbf{x_0})\right\}.\tag{4.2}$$

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

$$\min_{\mathbf{x}\in\mathbb{R}^k}H(\mathbf{u},\mathbf{x_0},\mathbf{x})\ .\tag{4.3}$$

Note that the existence and uniqueness of solutions is not a priori ensured. It depends on the behavior 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 $\lambda=\lambda_0$. We have chosen λ small so that the weight of the prior is small, which is better since the prior information is less *faithful* in some sense. For numerics we set $\lambda=10^{-4}$.

Assuming that problem (4.3) 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 implicitly given by the following system :

$$\forall i=1,\cdots,k\quad\frac{\partial H}{\partial x_i}(\mathbf{u},\mathbf{x_0},\mathbf{x}^*)=u_i'(x_i^*)-\lambda_i Az(\mathbf{x_0},x_1^*,\cdots,x_k^*)\cdot\mathbf{d}_i=0\ .\tag{4.4}$$

With (3.4) relation (4.4) 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

$$\forall i = 1, \cdots, k \quad u_i'(x_i^*) = \lambda_0 A\mathbf{x_0} \cdot (\lambda_i \mathbf{d}_i) + (\mathbb{A}x^*)_i \ . \tag{4.5}$$

Once $\mathbf{x}^*(\mathbf{u}, \mathbf{x_0})$ is computed, the cost functional in problem ($\mathcal{P}_m^*$) 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 have 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

$$\forall i = 1, \cdots, k \quad \frac{\partial \Phi}{\partial u_i}(\mathbf{u}) \cdot \varphi = \int_{\mathbb{R}} \varphi \pi_i - \int_{\mathbb{R}^d} \varphi(x_i^*(\mathbf{u}, \mathbf{x_0})) \rho_0(d\mathbf{x_0}) \ , \tag{4.6}$$

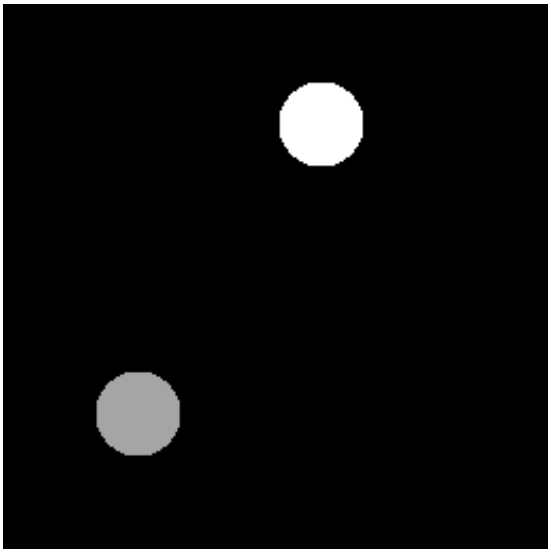
where $\mathbf{x}^*(\mathbf{u}, \mathbf{x_0})$ satisfies (4.5). We decided to use a Galerkin type method to approximate the solution. More precisely, we choose a suitable basis (FEM, spectral or spline) to write the function u. Here, we decided to use a spline basis, so that u is described by very few scalar coefficients. In addition, such an approach allows to compute the integral quantities once at the beginning of the process. The algorithm writes :

Algorithm 1

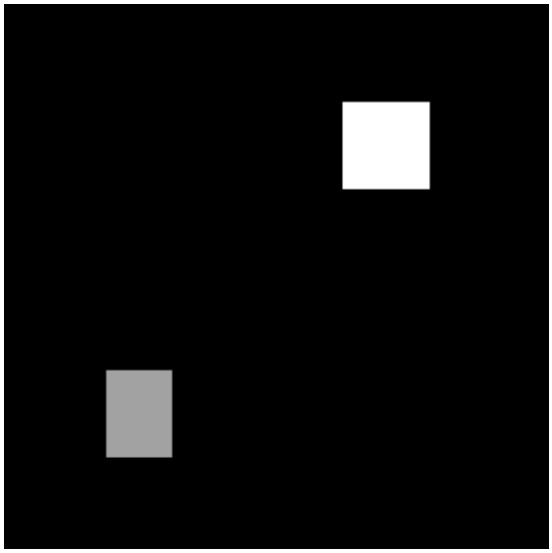
Given λ_i, $i = 0, \cdots, k$, $\mathbf{d}_i$, $i = 1, \cdots, k$. Compute integrals ($\mathbf{x_0}$ and φ are known), A and $\mathbb{A}$.

1. Choose $\mathbf{u}^0_{i=1..k}$
2. Iteration n : $\mathbf{u}^n_{i=1..k}$ has been computed
   (a) Compute $\mathbf{x}^{*n}_{i=1..k}(\mathbf{u}^n_{i=1..k}, \mathbf{x_0})$
   $$(u^n_{k,i})'(x^{*n}_{k,i}) - (\mathbb{A}x^{*n})_i = \lambda_0 A\mathbf{x_0} \cdot (\lambda_i \mathbf{d}_i);$$
   (b) Compute $\nabla\Phi(\mathbf{u}^n_{i=1..k})$ with
   $$\frac{\partial \Phi}{\partial u_j}(\mathbf{u}^n_{i=1..k}) = \pi_j - T_{j,k}\#\rho_0$$
   where $T^n_{i=1..k}(\mathbf{x_0}) := x^{n*}_{i=1..k}(\mathbf{u}^n_{i=1..k}, \mathbf{x_0})$.
   (c) Compute $\mathbf{u}^{n+1}_{i=1..k}$
   $$\mathbf{u}^{n+1}_{i=1..k} = \mathbf{u}^n_{i=1..k} - \tau_k \nabla\Phi(\mathbf{u}^n_{i=1..k}).$$

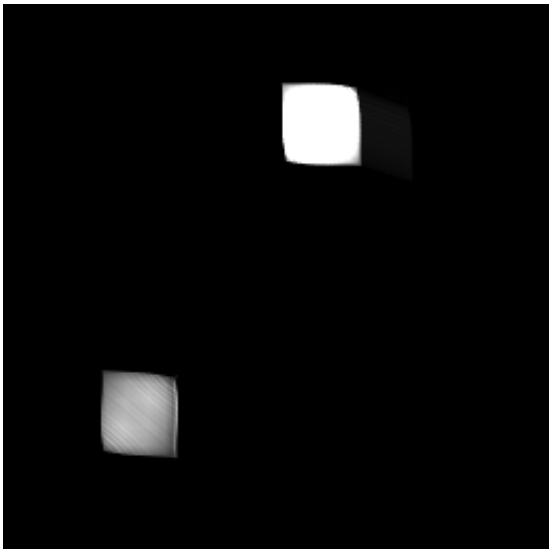
We present here below two academic examples obtained with three views with directions $\mathbf{d}_1 = (1, 0), \mathbf{d}_2 = (0, 1)$, $\mathbf{d}_3 = (1, -1)/\sqrt{2}$. In both cases $\lambda_i = 1$, $i = 1, 2, 3$ and $\lambda = 10^{-4}$. These examples seem similar. The main difference is that the material density is not zero in example 2. So the induced mass will be transported as well : this is an undesirable effect that we see on Figure 4.2.



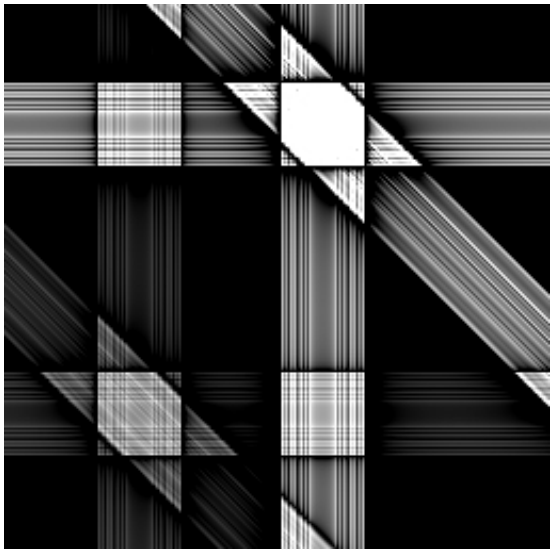
(a) Groundtruth (desired object)



(b) Prior ρ_0: the lower left square is located at the same place as the desired disk. The upper right one is not.

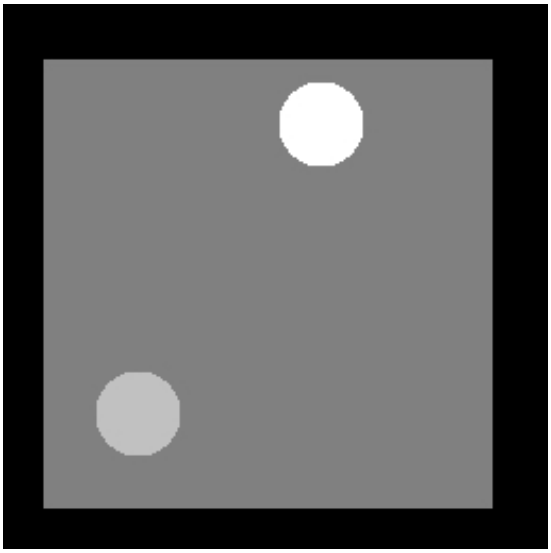


(c) Computed object : the squares have been moved at the right place (the centers of the disk and square are the same). However, the shapes do not fit well.

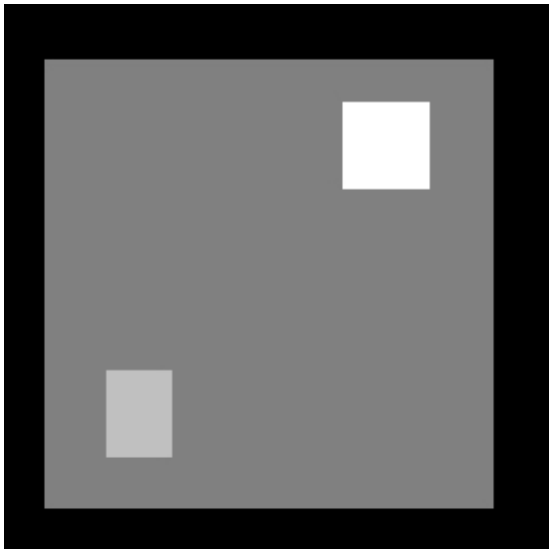


(d) Computed object with the standard back-filtered projection algorithm

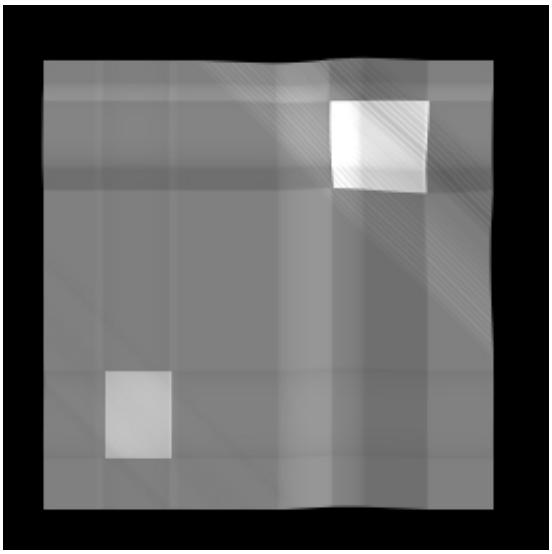
Figure 4.1: Example 1: no mass outside the two disks.



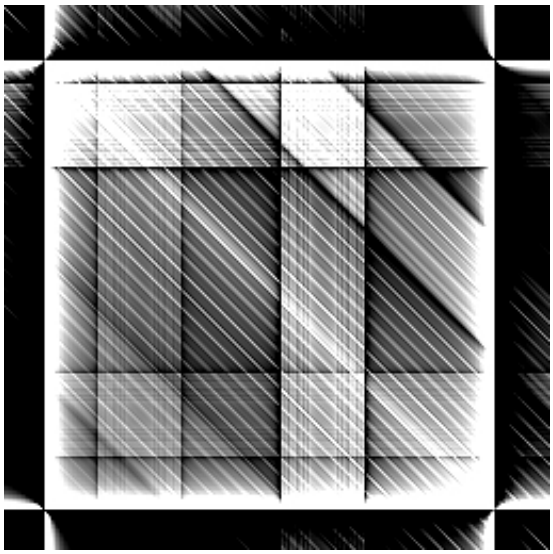
(a) Groundtruth (desired object)



(b) Prior ρ_0



(c) Computed object: the whole band containing the square has been transported and not the square only. So, the upper right square is not located at it should be.



(d) Computed object with the standard backfiltered projection algorithm

Figure 4.2: Example 2 : there is mass outside the two disks.

The method is promising but preliminary tests show a great lack of numerical stability and the numerical realization needs further investigation. Moreover, the example in Figure 4.2 shows that the model we propose has to be improved to avoid the *diffusion effect* we observe, due to the transportation of mass that should not be transported. It seems important to add geometric features that allow transport while preserving topology.

5 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. Let us mention that the entropic regularization of the problem has been explored recently in [6]. There is still a lot of challenging modelling and numerical analysis of potential interest (3D tomography, second-order methods ...) left for future research.

References

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

[2] 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.

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

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

[5] 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.

[6] J.-D. Benamou, G. Carlier, M. Cuturi, G.Peyré and L. Nenna, *Iterative Bregman projections for regularized transportation problems*, to appear in SIAM J. Sci. Comp.

[7] N. Bonnotte, *Unidimensional and Evolution Methods for Optimal Transportation*, PhD Thesis, Univ. Paris Sud Orsay, available at `cvgmt.sns.it` (2014).

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

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

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

[11] 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.

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

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

[14] 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.

[15] 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.

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

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

[18] 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.

[19] 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.

[20] 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.

[21] J. Rabin, G. Peyré, J. Delon, and M. Bernot, *Wasserstein Barycenter and Its Application to Texture Mixing*, In Scale Space and Variational Methods in Computer Vision, LNCS, vol. 6667, pp. 435-446, Springer, 2012.

[22] F. Santambrogio, *Optimal Transport for Applied Mathematicians*, Birkhäuser Verlag, Basel (2015).

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