Is the mailing Gilbert-Steiner problem convex?

Gershon Wolansky
Department of Mathematics, Technion, Haifa 32000, Israel
February 8, 2019

Abstract

A convexification of the mailing version of the finite Gilbert problem for optimal networks is introduced. It is a convex functional on the set of probability measures subject to the Wasserstein $p-$ metric. The minimizer of this convex functional is a measure supported in a graph. If this graph is a tree (i.e. contains no cycles) then this tree is also a minimum of the corresponding mailing Gilbert problem. The convexification of the Steiner problem is the limit of these convexified Gilbert’s problems in the limit $p \to \infty$. A numerical algorithm for the implementation of the convexified Gilbert-mailing problem is also suggested, based on entropic regularization.

1 Introduction

Optimal branched transportation is a variant of the theory of Monge-Kantorovich [11, 14] on optimal transportation. The classical Kantorovich cost of transporting a probability measure $f^+(dx)$ to another probability measure $f^-(dy)$, where the cost of transporting one unit from $x$ to $y$ is $c(x,y)$, is given by

$$\min_{\pi \in \Pi(f^+, f^-)} \int \int c(x,y) \pi(dx,dy)$$

(1)

where $\Pi(f^+, f^-)$ stands for the set of all transport plans $\pi$, which are 2-points probability measures whose marginals are $f^\pm$ [16, 14]. In contrast to the classical Monge-Kantorovich theory where each of the "mass particles" is transported independently of the others, in the branched transport (as well as a congestion transport) the particles are assumed to interact with each other while moving from a source to a target distribution. Thus, while Monge-Kantorovich optimal transport [1] is, basically, a linear programming in the affine space of transport plans given by probability measures (even though it is highly nonlinear in the domain of deterministic transport plans), a branched (and congested [6, 7]) transport is a minimizer of nonlinear functional even on the affine space of probability measures. Moreover, while the
The cost function of a branched transport induces ramifications in a natural way. It simulates many natural phenomena such as roots systems of trees, leaf ribs and the nervous system, as well as supply-demand distribution networks such as irrigation networks and electric power supply. The common principle behind all these cases is that the cost functions privilege large flows over diffusive ones, which causes the orbits of the transport to concentrate on 1-dimensional currents.

One of first models for branched transport was introduced by Gilbert [10] (see also [23]). Given source $f^+$ and target $f^-$ probability measures on $\mathbb{R}^d$ supported each in a finite set of points, he considered optimization over finite directed tree $T = (E, V)$ whose vertices $V$ contain $\text{supp}(f^+) \cup \text{supp}(f^-)$. Gilbert suggested to minimize

$$\sum_{e \in E(T)} w^\sigma(e)|e|$$

over all such trees $T$ and weight functions $w : E(T) \to \mathbb{R}^+$ representing the fluxes at the edges. Here $|e|$ is the length of an edge $e$ and $w$ satisfies

$$\sum_{e \in E^+(v)} w(e) = f^+(v), v \in \text{supp}(f^+) \quad \sum_{e \in E^-(v)} w(e) = \sum_{e \in E^+(v)} w(e) \quad v \in V - (\text{supp}(f^+) \cup \text{supp}(f^-))$$

where $E^+(v)$ ($E^-(v)$) stands for all edges outgoing from (ingoing to) $v$. The exponent $\sigma$ is chosen to be a number in the interval $[0, 1)$, reflecting the preference for concentration of the flow, due to the inequality $w^1_1(e) + w^\sigma_2(e) \geq (w_1(e) + w_2(e))^{\sigma}$. This choice of $\sigma$ is the source of non-convexity (and non-uniqueness) of this problem. In particular, the case $\sigma = 0$ corresponds to the Steiner problem of minimal graphs.

In 2003 Xia [20] (see also [21] and references therein) extended this model to a continuous framework using Radon vector-valued measures $u$ in on $M$ which satisfy the condition $\nabla \cdot u = f^+ - f^-$ in sense of distributions. This is a weak formulation of (3). The optimal transport network $u$ is obtained by minimizing

$$\mathcal{M}^\sigma(u) := \int_M \theta^\sigma(u) dH^1$$

over all $u$ satisfying the above constraint, where $\theta(u)(x)$ is the corresponding occupation density of $u$ at $x$ (which is a weak formulation of (2)). Such an extension makes sense for general Borel probability measures, and is reduced to Gilbert’s formulation (2, 3) if $f^\pm$ are supported on a finite set of points. One of the fundamental results of Xia is the condition $\sigma > 1 - 1/d$ which
guarantees a finite value of the transport cost for any Borel measures \( f^\pm \) in \( \mathbb{R}^d \) representing the source and sink distributions (see also \([8, 9]\)).

Another approach \([2, 12]\) represents traffic plans as measures on the set of Lipschitz paths connecting the source and sinks. The functional then acts on probability measures on this set. This approach is equivalent (in Euler representation) to Xia formulation (see \([2, 8]\)), and it can also accommodate the mailing problem.

In another approach, (see \([5]\) and references therein), a transportation network is modeled as a connected set \( \Sigma \subset \mathbb{R}^n \), and the linear Monge-Kantorovich problem is introduced to the metric \( d_\Sigma(x, y) = d(x, y) \wedge (\text{dist}(x, \Sigma) + \text{dist}(y, \Sigma)) \) as follows: Given source and target distributions \( f^\pm \), the problem is reduced to minimizing the Kantorovich cost

\[
\Sigma \Rightarrow \min_{\pi} \int \int d_\Sigma(x, y) \pi(dx, dy)
\]

among all transport plans \( \pi \) whose marginals are prescribed by \( f^\pm \), and all connected sets \( \Sigma \) whose length \( H_1(\Sigma) \) are bounded by a prescribed value \( L > 0 \). The non-linearity (and non-convexity) is conveyed in the dependence on \( \Sigma \). This non-linearity persists also in the mailing version, i.e. for minimizing \( \int \int d_\Sigma(x, y) \pi(dx, dy) \) with respect to \( \Sigma \), where a transport plan \( \pi \) is prescribed (rather than its marginals \( f^\pm \)).

Another approach (see e.g.\([4]\)) extend Benamou Brenier \([1]\) kinetic formulation of optimal transport to this setting. This formulation turns out to be equivalent to Xia’s approach. It seems, however, that the mailing problem cannot be accommodated in this setting.

Still another approach using a limit theorem was introduced the author in \([19]\). All in all, these approaches share the non-convex structure and, as a result, do not guarantee a unique solution, in general.

2 Formulation of the mailing problem

In the mailing problem (or ”who goes where” situation \([21, 2, 3, 9]\)), a transport plan \( \pi \) is prescribed, determining the ”mailing address” in the support of \( f^- \) for each point in the support of \( f^+ \).

Let us consider first the discrete version of the Gilbert mailing problem. Suppose \( A = \text{supp}(f^+) \), \( B = \text{supp}(f^-) \) be finite sets in \( \mathbb{R}^d \). The mailing program from \( f^+ \) to \( f^- \) is a non-negative function \( \pi(x, y) \) on \( A \times B \) which satisfies

\[
\sum_{x \in A} \pi(x, y) = f^-(y), \quad \sum_{y \in B} \pi(x, y) = f^+(x)
\]  

(4)

We view \( \pi(x, y) \) as the flux of mass from \( x \in A \) to \( y \in B \).

A network \( T \) in the class \( (A, B, \pi) \) is a tree embedded in \( \mathbb{R}^d \). This tree is a union of a finite number of segments \( e \), called edges.
If \( x \in A, y \in B \) and \( \pi(x, y) > 0 \) then there exists a connected component \( o_T(x, y) \subset T \) such that \( x, y \in o_T(x, y) \). We refer to \( o_T(x, y) \) as the orbit from \( x \) to \( y \) along the tree \( T \). In particular we may view \( o_T(x, y) \) as a subset of the edges \( e \) composing \( T \). By definition of a tree, there exists at most one connected orbit connecting each \( x, y \in T \). In particular, there exists a unique orbit connecting \( x \in A \) to \( y \in B \) provided \( \pi(x, y) > 0 \).

For any edge \( e \in T \) let
\[
    w_\pi(e) := \sum_{x \in A, y \in B; o_T(x, y) \ni e} \pi(x, y).
\] (5)

The length of an edge \( e \in T \) is denoted by \(|e|\).

**Remark 1.** \( w_\pi \) satisfies the Kirchhoff’s condition (3) on the tree \( T \).

**Problem 1.** Let \( \sigma \in [0, 1) \). The Gilbert mailing problem associated with \((A, B, \pi)\) is:

Minimize
\[
    G(T; \pi) := \sum_{e \in T} w_\pi(e)^\sigma |e|
\]

among all trees \( T \) in the class \((A, B, \pi)\).

**Proposition 2.1.** The discrete Gilbert’s problem (3) is obtained by minimizing \( G(\pi) := \min_T G(T, \pi) \) over all plans \( \pi \) satisfying (4).

We now introduce an equivalent formulation of the Gilbert’s mailing problem:

Let \( s(e) \geq 0 \) represents the cost of construction the edge \( e \) per unit length. Thus, the cost of construction of \( e \) is just \(|e|s(e)\). We impose the limit budget of constructing the network \( T \) by
\[
    \sum_{e \in T} |e|s(e) \leq 1.
\] (6)

The cost of transporting a unit of mass along the edge depends on the cost of construction of this edge in a monotone decreasing way: the higher the investment in the segment, the easier (and cheaper) is the cost of the transport trough this edge. We presume that the cost of transporting a unit mass per unit length in an edge \( e \) is \( s^{-\alpha} \) where \( \alpha > 0 \). Thus, the cost of transporting a unit of mass from point \( x \in A \) to \( y \in B \) trough the network \( T \) is 
\[
    \sum_{e \in o_T(x,y)} |e|s(e)^{-\alpha}.
\]

We define the network for the mailing problem as follows:

**Problem 2.** Minimize
\[
    H(T, s) := \sum_{x \in A} \sum_{y \in B} \pi(x, y) \sum_{e \in o_T(x,y)} |e|s(e)^{-\alpha}
\]

over all trees \( T \) in class \((A, B, \pi)\), and functions \( s = s(e) \) on \( T \) satisfying the constraint (6).
Lemma 2.1. The mailing Gilbert’s problem 1 and Problem 2 are equivalent, provided $\sigma = \frac{1}{\alpha + 1}$.

Proof. By (5) we rewrite $H(T, s)$ as

$$H(T, s) = \sum_{e \in T} w_{\pi}(e)|e|s^{-\alpha}(e)$$

Let

$$H(T) = \min_s H(T, s)$$

where the minimum is subjected to the constraint (6). For a given tree $T$ this is a strictly convex function of $s$. There is a unique global minimizer under the constraint (6). Since $\alpha > 0$ this minimizer is obtained at positive $s$ for any edge $e$. Let $\lambda$ be the Lagrange multiplied due to the constraint (6). Then

$$-\alpha w_{\pi}(e)s^{-\alpha-1}(e) + \lambda = 0$$

for any $e \in T$, thus $s(e) = (\lambda/\alpha)^{-1/(\alpha+1)}w_{\pi}(\alpha)^{1/(\alpha+1)}$. Substituting this in the constraint (which must holds with equality) implies

$$\left(\frac{\lambda}{\alpha}\right)^{\alpha/(\alpha+1)} = \left(\sum_{e \in T} |e|w_{\pi}(e)^{1/(\alpha+1)}\right)^{\alpha}.$$ 

Substitute $s(e)$ in $H$ we get $H(T) = (\alpha/\lambda)^{\alpha/(\alpha+1)} \sum_{e \in T} w_{\pi}(e)^{1/(1+\alpha)}|e|$. Thus

$$H(T) = \left(\sum_{e \in T} |e|w_{\pi}(e)^{1/(\alpha+1)}\right)^{1+\alpha}$$

which is the $1 + \alpha$ power of the Gilbert-mailing cost $G(T, \pi)$ under $\sigma = 1/(\alpha + 1)$. \qed

3 Conditional Wasserstein metric

We consider now $\nu_\pm$ as a pair of Borel probability measures on $\mathbb{R}^d$. Recall the Wasserstein metric on the set of probability Borel measures $\mathcal{P}(\mathbb{R}^d)$

$$W_p(\nu_+, \nu_-) = \left(\min_{\pi \in \Pi(\nu_+, \nu_-)} \int \int |x - y|^p \pi(dx, dy)\right)^{1/p}$$

where $\Pi(\nu_+, \nu_-)$ is the continuum version of (4)

$$\Pi(\nu_+, \nu_-) = \left\{ \pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d); \int_{y \in \mathbb{R}^d} \pi(dx, dy) = \nu_+(dx), \quad \int_{x \in \mathbb{R}^d} \pi(dx, dy) = \nu_-(dy) \right\}$$

It is known that $W_p$ is a metric on the set of Borel probability measures on $\mathbb{R}^d$ having a finite $p-$moment, if $p \geq 1$ (see, e.g. [15]). Let us define the conditional Wasserstein $p-$metric $W_p(\nu_+, \nu_- \| \mu)$, where $\mu \in \mathcal{P}(\mathbb{R}^d)$:

$$W_p(\nu_+, \nu_- \| \mu) := \lim_{\varepsilon \to 0} \varepsilon^{-1}W_p(\mu + \varepsilon \nu_+; \mu + \varepsilon \nu_-).$$
Theorem 1. \cite{7,18} 
\begin{align*}
W_p(\nu_+, \nu_-|\mu) = \sup \left\{ \int \phi d(\nu_+ - \nu_-), \quad \phi \in C^1(\mathbb{R}^d), \quad \int |\nabla \phi|^{p'} d\mu \leq 1 \right\}
\end{align*}
where $p' = \frac{p}{p-1}$. In particular,
\begin{align*}
\mu \rightarrow W_p^p(\nu_+, \nu_-|\mu) = p \sup_{\phi \in C^1} \int \phi d(\nu_+ - \nu_-) - \frac{p}{p'} \int |\nabla \phi|^{p'} d\mu
\end{align*}
is a convex function in $\mu \in \mathcal{P}$ for fixed $\nu_+$.

Let us now substitute $\delta_x$ for $\nu_+$ and $\delta_y$ for $\nu_-$. 
\begin{align*}
W_p(x, y|\mu) := W_p(\delta_x, \delta_y|\mu) \equiv \sup_{\phi \in C^1, \phi \neq 0} \frac{\phi(x) - \phi(y)}{\int |\nabla \phi|^{p'} d\mu} . \tag{8}
\end{align*}
It follows that

**Lemma 3.1.** $\mu \rightarrow W_p^p(x, y|\mu) \in R_+ \cup \{\infty\}$ is a convex function in $\mu \in \mathcal{P}$ for any $x, y \in \mathbb{R}^d$. In addition, $W_p(\cdot; \cdot|\mu)$ is an extended metric on $\mathbb{R}^d$ (i.e. it may attains infinite values), for fixed $\mu \in \mathcal{P}$.

### 4 Relaxation of the Gilbert-mailing problem

Let now $T$ be a tree in class $(A, B, s)$. We associated with this tree a probability measure $\mu_T$ which is supported on this tree and such that $\mu_T(e) = |e|s(e)$ for any $e \in T$.

**Lemma 4.1.** If $x \in A, y \in B$ then $W_p^p(x, y|\mu_T) \geq \sum_{e \in \partial_T(x, y)} |e|s^{1-p}(e)$. The equality holds iff $\mu_T$ is a uniform measure on each edge $e$.

**Proof.** Let $\bigcup_{i=1}^l e_i = \partial_T(x, y)$, $|\partial_T(x, y)| := \sum_{e \in \partial_T(x, y)} |e|$. $q : [0, |\partial_T(x, y)|] \rightarrow \partial_T(x, y)$ be an arc-length parameterization of the orbit, and $\rho$ a density of a positive measure on $[0, |\partial_T(x, y)|]$ such that $q_#(\rho d\tau) = \mu_T|_{\partial_T(x, y)}$, that is, for any test function $\phi \in C^1_0(\mathbb{R}^d)$:
\begin{align*}
\int_{\partial_T(x, y)} \phi d\mu_T = \int_0^{|\partial_T(x, y)|} \phi(q(\tau)) \rho(\tau) d\tau .
\end{align*}
Let $\tau_i = q^{-1}(e_i \cap e_{i+1})$ for $i = 1 \ldots l - 1$. We set $\tau_0 = q^{-1}(x) = 0$ and $\tau_l = q^{-1}(y) = |\partial_T(x, y)|$.

Since $q$ is an arc-length parameterization and $\mu_T(e_i) = |e_i|s(e_i)$, we get
\begin{align*}
\int_{\tau_i}^{\tau_{i+1}} \rho d\tau = |e_i|s(e_i) \quad \text{and} \quad \tau_{i+1} - \tau_i = |e_i| . \tag{9}
\end{align*}

Let $\phi \in C^1_0(\mathbb{R}^d)$ and $\psi(\tau) = \phi(q(\tau))$. Then
\begin{align*}
\int_{\partial_T(x, y)} |\nabla \phi|^{p'} d\mu_T \geq \int_0^{|\partial_T(x, y)|} |\psi'|^{p'} \rho(\tau) d\tau , \quad \phi(y) - \phi(x) = \psi(|\partial_T(x, y)|) - \psi(0) .
\end{align*}
A direct calculation yields
\[
\int_0^{\|o_T(x,y)\|} \rho |\psi'|^p \, d\tau \geq |\psi_i(\|o_T(x,y)\|) - \psi_i(0)|^p \left( \int_0^{\|o_T(x,y)\|} \rho^{1-p} \right)^{1/(1-p)},
\]
hence (since \( p = p'/(p' - 1) \))
\[
\left( \frac{\psi(\|o_T(x,y)\|) - \psi(0)}{\int_0^{\|o_T(x,y)\|} \rho |\psi'|^p} \right)^p \leq \int_0^{\|o_T(x,y)\|} \rho^{1-p}
\]
holds for any smooth function \( \psi \) on \([0, 1]\). Moreover, we can find a sequence of such smooth functions \( \psi_n \) for which the limit above holds with an equality, or the left hand side blows to infinity, if the integral on the right hand side diverges. Each such smooth \( \psi \) can be lifted to a Lipschitz function on the orbit \( o_T(x,y) \) via \( \psi \circ q^{-1} \), and we can extend \( \psi \circ q^{-1} \) to a function \( \phi \) the entire tree \( T \) such that \( \phi \) is a constant = \( \psi(\tau_i) \) on each sub-tree rooted at \( e_{i+1} \cap e_i \). Since all these sub-tree are disjoint, we obtain from \( \psi \) a Lipschitz function \( \phi \) which can be extended to \( \mathbb{R}^d \) such that
\[
\left( \frac{\phi(x) - \phi(y)}{\int_{\mathbb{R}^d} |\nabla \phi|^p \, d\mu} \right)^p \leq \int_0^{\|o_T(x,y)\|} \rho^{1-p}.
\]
Since the supremum on the left yields the equality it follows by (9) that \( W_p^\rho(x,y)\|\mu_T\| = \int_0^{\|o_T(x,y)\|} \rho^{1-p} \).

By (9) and the Jensen’s inequality it follows that
\[
\int_0^{\|o_T(x,y)\|} \rho^{1-p} \geq \sum_{e \subseteq o_T(x,y)} |e| s^{1-p}(e)
\]
and the equality is attained at \( \rho(\tau) = s(e) \) on \( \tau \in (\tau_i, \tau_{i+1}) \) (i.e \( \mu_T \) is a the uniform measure \( s(e) \) on each edge \( e \in o_T(x,y) \)).

By (8) it follows that
\[
\rho(\tau) = s(e) \quad \text{on} \quad \tau \in (\tau_i, \tau_{i+1})
\]
and the equality is attained at \( \rho(\tau) = s(e) \) on \( \tau \in (\tau_i, \tau_{i+1}) \) (i.e \( \mu_T \) is a the uniform measure \( s(e) \) on each edge \( e \in o_T(x,y) \)).

### The convexified Gilbert-mailing problem

The convexified Gilbert-mailing problem for a given \( \sigma \in (0, 1) \) takes the form:

Minimize \( H_p(\mu) := \sum_{x \in A} \sum_{y \in B} \pi(x, y) W_p^\rho(x, y)\|\mu\|, \) over \( \mu \in P \), where \( p = 1/\sigma \).

Since \( H_p \) is a convex functional on \( P \) we obtained

**Theorem 2.** There exists a unique solution to the Gilbert-mailing problem, supported in a finite graph. If the support is a tree then this tree is a solution of the mailing Gilbert problem.

**Remark 2.** The possibility of closed cycles cannot be excluded. Suppose there are two orbits \( o_1(x, y), \ldots, o_k(x, y) \) connecting \( x \) and \( y \). Let \( \mu \) be a positive measure supported on \( \cup o_i(x, y) \), \( \rho_i \)
being the density on \( o_i(x, y) \). By a computation similar to Lemma \[4.1\] we obtain

\[
W_p^p(x, y\|\mu) = \left( \sum_{i=1}^{k} \left( \int_{o_i(x, y)} \rho_i^{1-p} \right)^{1/(1-p)} \right)^{1-p}.
\]

**Remark 3.** \( \sigma = 0 \) (the Steiner mailing case) corresponds to the limit \( p = \infty \). In particular, the mailing Steiner problem is a limit of convex optimization problem. Note that the mailing Steiner problem is equivalent to the Steiner problem itself if \( \pi(x, y) > 0 \) for any \( x \in A, y \in B \).

## 5 Numerical implementation and entropic relaxation of the mailing problem

Let us consider a finite grid \( Z \subset \mathbb{R}^d \). Let \( A := \{x_1, \ldots, x_m\}, B = \{y_1, \ldots, y_n\} \) subsets of \( Z \).

On the grid \( Z \) we assign weights \( m \) in the simples

\[
S(Z) := \left\{ m : Z \to \mathbb{R}_+; \ m(z) \geq 0, \sum_{z \in Z} m(z) = 1 \right\}.
\]

This is a discretization of the probability measures \( \mu \) on \( Z \).

For \( x_i, y_j \in A, B \), the discrete version of \( W_p^p(x_i, y_j\|\mu) \) is the supremum over \( \phi \in \mathbb{R}^Z \) of

\[
W_p^p(m, \phi; x_i, y_j) = -\frac{p}{p'} \sum_{z \in Z} \sum_{z' \in N(z)} m(z) \frac{1}{|N(z)|} |\phi(z) - \phi(z')|^{p'} + p(\phi(x_i) - \phi(y_j))
\]

(c.f. Theorem \[1\]). Here \( N(z) \) stands for the neighbors in \( Z \) of the grid point \( z \) and \( |N(z)| \) the cardinality of \( N(z) \) (recall \( x_i, y_j \in Z \) as well). Our object is to find

\[
\min_{m \in S(Z)} \sum_{x_i \in A, y_j \in B} \pi(i, j) \max_{\phi \in \mathbb{R}^Z} W_p^p(m, \phi; x_i, y_j)
\]

where \( \pi(i, j) := \pi(x_i, y_j) \) is the transport plan from \( A \) to \( B \). The support of the minimizer \( m \) in \( S(Z) \) is the discrete approximation of the optimal tree solving the mailing Gilbert problem for \( \sigma = 1/p \).

Let

\[
\phi_{i,j}(m) := \arg \max W_p^p(m, \phi; x_i, y_j).
\]

The entropic regularization of \[10\] is

\[
\min_{m \in S(Z)} \sum_{x_i \in A, y_j \in B} \pi(i, j) W_p^p(m, \phi_{i,j}(m); x_i, y_j) + \left( \frac{\varepsilon p}{p'} \right) \sum_{z \in Z} m(z) \ln(m(z)).
\]

The minimal \( m \in S(Z) \) is

\[
m(z) = \frac{\exp \left( \sum_{i,j} \pi(i, j) D\phi_{i,j}(z)/2\varepsilon \right)}{\sum_{z' \in Z} \exp \left( \sum_{i,j} D\phi_{i,j}(z')/2\varepsilon \right)}.
\]
where

$$D\phi(z) := \sum_{z' \in N(z)} \frac{1}{|N(z)|}|\phi(z) - \phi(z')|^{p'}$$  \hspace{1cm} (13)$$

Substitute this in (11) and using MinMax Theorem we obtain the equivalent problem: Maximize over \( \{\phi\} \in \mathbb{R}^{Z \times A \times B} \):

$$H^\varepsilon(\{\phi\}) = -\varepsilon \ln \left( \sum_{z \in Z} \exp \left( \frac{\sum_{i,j} \pi(i,j) D\phi_{i,j}(z)}{2\varepsilon} \right) \right) + p \sum_{i,j} \pi(i,j) (\phi_{i,j}(x_i) - \phi_{i,j}(y_j)) .$$

where \( D\phi \) defined by (13).

References

[1] J.D.Benamou, Y. Brenier: A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, Numer.Math., 84 (2000), 375-393.

[2] Bernot, M; Caselles. V, and Morel. J-M: Optimal transportation networks. Models and theory, Lecture Notes in Mathematics, 1955. Springer, Berlin, 2009

[3] Bernot, M; Caselles, V; Morel, J.M: Traffic plans, Publ. Mat. 49 (2005), no. 2, 417-451

[4] Brasco, L; G. Buttazzo, G and F. Santambrogio, F: A Benamou-Brenier approach to branched transport, SIAM J. Math. Anal., 43(2):1023-1040, 2011

[5] Buttazzo, G, Pratelli, A, Stepanov, S and Solimini, S: Optimal Urban Networks via Mass Transportation, Lec. notes in Mathematics 1961, Springer , 2009

[6] Buttazzo, G; Jimenez, C and Oudet, E: An optimization problem for mass transportation with congested dynamics, SIAM J. Control Optim., 48 (2009), pp. 1961-1976.

[7] Carrillo, J, Patacchini, F., Sternberg, P. and Wolansky, G.: Convergence of a particle method for diffusive gradient flows in one dimension, SIAM J. Math. Anal. 48, no. 6, 3708-3741, 2016

[8] Colombo,M; De Rosa, A and Marchese, A: Improved stability of optimal traffic paths, Calc. Var. Partial Differential Equations. (2018), Volume 57, Issue 28.

[9] Colombo,M; De Rosa, A and Marchese, A: Stability for the mailing problem, ArXiv: 1801.05624

[10] Gilbert, E.N: Minimum cost communication networks, Bell System Tech. J., 46: 2209-2227, 1967
[11] Kantorovich, L: *On the translocation of masses*, C.R (Doclady) Acad. Sci. URSS (N.S), 37, (1942), 199-201

[12] Maddalena, F; Solimini, S and Morel, J.M: *A variational model of irrigation patterns*. Interfaces and Free Boundaries, 5:391-416, 2003

[13] Oudet, E, Santambrogio, F: *A Modica-Mortola approximation for branched transport and applications* Arch. Ration. Mech. Anal. 201, no. 1, 115?142, 2011

[14] Villani C.: *Optimal Transport: Old and New*, Grundlehren der mathematischen Wissenschaften, (2008)

[15] W. Gangbo and R.J. McCann: *The geometry of optimal transportation*, Acta Math., 177 (1996), 113-161

[16] Villani, C: *Topics in Optimal Transportation*, Graduate Studies in Mathematics, Vol. 58, AMS

[17] Wolansky, G: *Limit theorems for optimal mass transportation*, Calc. Var. Partial Differential Equations 42 (2011), no. 3-4, 487?516.

[18] Wolansky, G: *From optimal transportation to optimal teleportation* Ann. Inst. H. Poincaré Anal. Non Linéaire 34 (2017), no. 7, 1669-1685.

[19] Wolansky, G: *Limit theorems for optimal mass transportation and applications to networks*. Discrete Contin. Dyn. Syst. 30, no. 1, 365-374, 2011

[20] Xia, Q: *Optimal paths related to transport problems* Commun. Contemp. Math., 5: 251-279, 2003.

[21] Xia, Q and Vershynina, A: *On the transport dimension of measures*, SIAM J. MATH. ANAL. Vol. 41, No. 6, pp. 2407-2430, 2010

[22] Yan, S, Juang, D Chen, C and Lai, W:*Global and local search algorithms for concave cost transshipment problems* J. Global optimization, 33(1) 123-156, 2005

[23] Zangwill, W.I:*Minimum concave cost flows in certain networks*, Management science 14(7), 429-450, 1968