On a solution to the Monge transport problem on the real line arising from the strictly concave case

Nicolas Juillet

Abstract

It is well-known that the optimal transport problem on the real line for the classical distance cost may not have a unique solution. In this paper we recover uniqueness by considering the transport problems where the costs are a power smaller than one of the distance, and letting this parameter tend to one. A complete construction of this solution that we call excursion coupling is given. This is reminiscent to the one in the convex case. It is also characterized as the solution of secondary transport problems. Moreover, a combinatoric/geometric characterization of the routes used for this transport plan is provided.

We first introduce the mass transport problem and quickly arrive to our Main Theorem. The reason why it can appear so fast is that it concerns the most basic setting – together with the finite discrete setting – where the optimal transport problem can be introduced, namely the real line for a power cost.

Definition 0.1. Let $(\mathcal{X}, \mathrm{d})$ be a metric space and p be a positive real number. We denote by $\mathcal{P}_p(\mathcal{X})$ the set of Borel probability measures $\mu \in \mathcal{P}(\mathcal{X})$ such that $\int \mathrm{d}(x_0, x)^p \mathrm{d}\mu(x) < +\infty$ for some (and in fact any) $x_0 \in \mathcal{X}$ and say that μ has finite moment of order p. The set $\mathrm{Marg}(\mu, \nu)$ is the (convex) set of measures $\pi \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$ with first marginal μ and second marginal ν, i.e $(\mathrm{proj}^1)_{\#}\pi = \mu$ and $(\mathrm{proj}^2)_{\#}\pi = \nu$ where proj^i is the i-th coordinate function. These definitions naturally extend to positive measures π, μ, ν of finite positive mass $\mu(\mathcal{X})$.

For μ and ν two measures with $\mu(\mathcal{X}) = \nu(\mathcal{X})$ we call "L^p transport problem" (of Monge and Kantorovich), the problem to minimize

$$ C_p : \pi \in \mathrm{Marg}(\mu, \nu) \mapsto \iint \mathrm{d}(x, y)^p \, \mathrm{d}\pi(x, y) \in \mathbf{R} \cup \{\infty\}. $$

We denote by $\mathrm{Marg}_p^*(\mu, \nu)$ the (convex) set of solutions to this problem.

The space $\mathrm{Marg}(\mu, \nu)$ is equipped with the weak convergence topology and it is compact according to Prokhorov Theorem, see [Vil03, §1.1.7]. Moreover C_p is continuous on $\mathrm{Marg}(\mu, \nu)$, so that $\mathrm{Marg}_p^*(\mu, \nu)$ is not empty. One may have a look at page 23 where this basic continuity result and a more evolved one are stated and proved. Note that if μ and ν have finite moment of order p the

value of the problem is finite, with $\min_{\pi \in \mathrm{Marg}(\mu,\nu)} C_p(\pi) \leq C_p(\mu \otimes \nu) < +\infty$. Hence, we will assume that μ and ν have finite moment of order 1 permitting a finite value for all the L^p transport problems, for $p \in]0,1]$. Note moreover that for every $p \in]0,1]$ since C_p is continuous and $\mathrm{Marg}(\mu,\nu)$ compact, the set $\mathrm{Marg}_p^*(\mu,\nu)$ is not empty.

Let us review some of the known facts concerning the solutions of the L^p transport problem on the real line $\mathcal{X} = \mathbf{R}$. See also Remark 5.8 for the connexion of the L^1 transport problem on $\mathbf{R}$ with the L^1 transport problem on geodesic spaces.

- For $p > 1$, except from degeneracy occurring when $\min_{\pi \in \mathrm{Marg}(\mu,\nu)} C_p = \infty$ the set of solutions $\mathrm{Marg}_p^*(\mu,\nu)$ is reduced to a single element, known as *commonotonic coupling*, *quantile coupling*, *monotone rearrangement*, *Hoeffding-Fréchet transport* or any combination of these vocables. The universality and usefulness of the notion is reflected by this number of names. If the atomic part of μ is zero, i.e $\mu(x) = 0$ for every $x \in \mathbf{R}$ the monotone rearrangement π writes $\pi = (\mathrm{id}, T)_{\#}\mu$ where T is a non increasing map namely $F_\nu \circ F_\mu^{-1} : F_\mu(\mathbf{R}) \subset \mathbf{R} \to \mathbf{R}$ in the notation of Section 1. Such a T is called an *optimal transport map* since it minimizes $S \mapsto \int d^p(x, S(x)) \, \mathrm{d}\mu(x)$ among all *transport maps*, i.e the maps S that satisfy $S_{\#}\mu = \nu$.

- The value $p = 1$ is the one in the original problem of Monge of 1781 [Mon81] (except that Monge considers $\mathcal{X} = \mathbf{R}^2$ or $\mathbf{R}^3$). For this parameter the quantile transport is still one of the solutions but it may not be the unique one. There is actually a broad class of pairs (μ, ν) such that the set of solutions $\mathrm{Marg}_1^*(\mu, \nu)$ is the whole set of transport plans $\mathrm{Marg}(\mu, \nu)$, see Remark 5.3. Note also that the solutions of the Monge problem in many geodesic space $\mathcal{X}$ can be disintegrated along transport rays, i.e parts of $\mathcal{X}$ isometric to intervals where the obtained measures are solution of the Monge problem on the real line. This disintegration is the fundament for many powerful recent applications, see Remark 5.8

- The problem for the values $p \in]0,1[$ may be less known and understood as $p \geq 1$. However, Gangbo-McCann [GM96] and McCann [McC01] thoroughly explored this range of the parameter p and discovered a paradoxal hierarchical behavior. They say: "[...] the concavity of the cost function favours a long trip and a shorter trip over two trips of average length; as a result, it can be efficient for two trucks carrying the same commodity to pass each other travelling opposite directions on the highway [...]. In optimal solutions, such 'pathologies" may nest on many scales, leading to a natural hierarchy among the regions [...]." Unlike the monotone rearrangement the orientation an optimal transport plan can reverse the order of two points, typically making it locally decreasing if it is a map. However, the orientation does not have to be globally reversed. In particular the antimonotonic coupling is not but for very special pairs (μ, ν).

For a class of absolutely continuous measures μ and ν, McCann found an algorithm to restrict the search for the solution – it is unique under the absolute continuity assumption – to a finite number of classes, where regions of $\mathbf{R}$ concerning μ are mapped onto other regions of $\mathbf{R}$ concerning ν, the frontiers between the regions having to be determined. Note that for general values of (μ, ν), the solution space $\mathrm{Marg}_p^*(\mu, \nu)$ may not be reduced to a single value. Moreover it is not constant as a function of $p \in\,]0, 1[$. See Example 2.2 about these two facts. See also Lemma 2.4: the maximal amount of mass that can stay must also stay on place.

A few recent more works also deal with the 'concave costs': a corollary of [PPS15] for instance yields that if the atomic part of μ is reduced to zero and μ and ν are mutually singular there exists a unique optimal transport plan π and this one can be represented by a transport plan, i.e $\pi = (\mathrm{id}, T)_{\#}\mu$. This improves $\mu(\mathrm{Spt}(\nu)) = 0$ that was the sufficient condition in Gangbo–McCann's work. On the discrete side, an algorithm given in [DSS12] provides the optimal assignment in the case of finitely many atoms of same mass.

- Our paper being concerned with the value $p = 1$ and its limits $p \to 1^{\pm}$, we interrupt here the description of the power costs. It is continued in Remarks 5.1 and 5.2 where we give an account on $p = 0$ and, beyond, $p < 0$.

The novelty of this paper is that for the critical and historical parameter $p = 1$ we distinguish a special element of $\mathrm{Marg}_p^*(\mu, \nu)$ that we call the *excursion coupling*. This element may be seen as a solution for $p \to 1^-$, the counterpart of the quantile coupling that would be the solution for $p \to 1^+$.

Main Theorem. *Let μ and ν be two probability mesures in $\mathcal{P}_1(\mathbf{R})$, $\pi \in \mathrm{Marg}(\mu, \nu)$ and $q \in\,]0, 1[$. The following assumptions are equivalent.*

1. *(Solution of the L^{1^-} limit transport problem) There exists a sequence $(\pi_n, p_n)_n \in \mathbf{N}$ with $p_n < 1$ and $\pi_n \in \mathrm{Marg}_{p_n}^*(\mu, \nu)$ for every n, such that $(\pi_n, p_n) \to_n (\pi, 1)$.*

2. *(Solution of the $L^{1,q}$ secondary transport problem) π is in $\mathrm{Marg}_1^*(\mu, \nu)$ and in this space it minimizes $\gamma \in \mathrm{Marg}_1^*(\mu, \nu) \mapsto \iint |y - x|^q \, \mathrm{d}\pi(x, y)$.*

3. *(Monotonicity) π is concentrated on a set $S \subset \mathbf{R} \times \mathbf{R}$ whose arches do not cross, do not connect and have the same orientation when they are nested (see Definition 0.3 and Figure 4).*

4. *(Excursion coupling) π is the excursion coupling of μ and ν as defined in Section 1.*

In Section 4 we will prove the Main Theorem as well as the following corollary based on the uniqueness of a coupling satisfying *1* or *2* in the Main Theorem.

Corollary 0.2. *Let μ and ν be two probability mesures in $\mathcal{P}_1(\mathbf{R})$. The excursion coupling $\pi \in \mathrm{Marg}(\mu, \nu)$ satisfies the two following statements.*

1'. For $(\pi_n, p_n)_{n \in \mathbf{N}}$ satisfying $p_n \to 1$ (with $p_n < 1$) and $\pi_n \in \mathrm{Marg}^*_{p_n}(\mu, \nu)$ it holds $\pi_n \to \pi$.

2'. For every $q' < 1$, the map $\gamma \in \mathrm{Marg}^*_1(\mu, \nu) \mapsto \iint |y - x|^{q'} \, \mathrm{d}\pi(x, y)$ is minimized by π.

The paper is organized as follow: To give the Main Theorem a complete meaning we first briefly make *3.* more precise in Definition 0.3 and define in Section 1 the excursion coupling attached to a pair (μ, ν). Note that this definition relies on several facts concerning functions with bounded variations that will be recalled and established in §3.1. Then we prove step by step the following implications: We prove *1 ⇒ 3* and *2 ⇒ 3* in Section 2. We prove *3 ⇒ 4* in Section 3. The fact that there exists at least one solution to the L^{1^-} problem and the $L^{1,q}$ secondary problem completes the proof (see Section 4). In Section 5 we provide some more comments.

Definition 0.3 (Arches in the Main Theorem)**.** Let S be a subset $\mathbf{R} \times \mathbf{R}$, whose elements we call *transport routes*. In what follows $[a, b]$ means $[\min(a, b), \max(a, b)]$ and the routes are seen as arches (half-circles) over the real line.

- The routes of S are said *non-crossing arches* if for every $(x, y) \in S$ and $(x', y') \in S$ at least one of the three happens.

  - $[x, y] \cap [x', y'] = \emptyset$
  - $[x, y] \cap [x', y'] = \{z\}$ for some $z \in \mathbf{R}$
  - One of the two, $[x, y]$ or $[x', y']$, is included in the other.

- The arches *do not connect* means that if $(x, y) \in S$, $(x', y') \in S$ and $\min(|y - x|, |y' - x'|) > 0$ then $y \neq x'$.

- The *nested arches have the same orientation* if for every $(x, y) \in S$ and $(x', y') \in S$ such that $[x', y'] \subset]x, y[$ we have $(y - x)(y' - x') \geq 0$.

The first property can be summed up saying that in the plane the two circles of diameter $|y - x|$ and $|y' - x'|$ linking x and y, and x' and y', respectively, do not cross. The second property states that a point can not be at the same time a starting and arriving point. In the last property is stated that transporting x to y, and x' to y' must be done in the same direction when one of the arches is included in the other. See Figure 4 for a representation of the forbidden configuration and the authorized rerouting – where (x, y') and (x', y) are the new routes.

1   Definition of the excursion coupling

Given two probability measures μ and ν, we consider the signed measure $\sigma = \mu - \nu$ and its cumulative distribution function

$$F_\sigma : x \in \mathbf{R} \mapsto \sigma]-\infty, x] = F_\mu(x) - F_\nu(x). \tag{1}$$

Since it is the difference of F_μ and F_ν, the function F_σ is càdlàg (right continuous and with left limits at any point) and has limits 0 in $\pm\infty$. The graph $\mathrm{Graph}(F_\sigma) = \{(x, F_\sigma(x)) \in \mathbf{R}^2 : x \in \mathbf{R})\}$ of F_σ can be completed at discontinuity points x by vertical segments $[(x, F_\sigma(x^-)), (x, F_\sigma(x))] \subset \mathbf{R}^2$ where $F_\sigma(x^-) = \lim\limits_{\varepsilon \to 0,\, \varepsilon > 0} F_\sigma(x - \varepsilon)$ denotes the left limit at x. We denote by F_σ^* the multivalued map defined by $F_\sigma^*(x) = [F_\sigma(x^-), F_\sigma(x)]$ at discontinuity points x and $F_\sigma^*(x) = \{F_\sigma(x)\}$ at continuity points. Its graph is

$$\mathrm{Graph}(F_\sigma^*) = \{(x, h) \in \mathbf{R}^2 : h \in F_\sigma^*(x)\}$$

and we may also denote it by $\mathrm{Graph}^*(F_\sigma)$. If $(x, h) \in \mathrm{Graph}(F_\sigma^*)$ the real x is called a generalized solution of $F_\sigma = h$.

We further introduce the following subsets of $\mathrm{Graph}^*(F_\sigma)$: $\mathrm{Graph}^{*,+}(F_\sigma)$ is the set of increasing points (x, h), i.e such that in a neighborhood U_x of x any point $(x', h') \in \mathrm{Graph}^*(F_\sigma)$ with $x' \in U_x \setminus \{x\}$ satisfies $(h' - h)(x' - x) > 0$. $\mathrm{Graph}^{*,-}(F_\sigma)$ is the set of decreasing points (x, y), i.e such that in a neighborhood U_x of x any point $(x', h') \in \mathrm{Graph}^*(F_\sigma)$ with $x' \in U_x \setminus \{x\}$ satisfies $(h' - h)(x' - x) < 0$.

Theorem 1.1 (Excursion couplings can be defined). *Let μ and ν be probability measures and $\sigma = \mu - \nu$, F_σ and $\mathrm{Graph}(F_\sigma^*)$ be defined as above. One can define a transport of $\mathrm{Marg}(\mu, \nu)$ as described in what follows, the results implicitly stated during this construction (see Remark 1.2 for a list) are correct and we call* excursion coupling *the resulting coupling.*

We distinguish two cases (the first one is a special case of the second).

1. *Assume that μ and ν are singular measures ($\mu \perp \nu$). We define a coupling (X, Y) where $X \sim \mu$ and $Y \sim \nu$. Let θ be the measure with density $h \mapsto i^\pm(h) = \mathrm{card}[(\mathbf{R} \times \{h\}) \cap \mathrm{Graph}^{*,+}(F_\sigma)] + \mathrm{card}[(\mathbf{R} \times \{h\}) \cap \mathrm{Graph}^{*,-}(F_\sigma)]$. Let H be a random variable with law $\theta/2$ and note $(\mathbf{R} \times \{H\}) \cap \mathrm{Graph}^*(F_\sigma) = \{x_1, \ldots, x_N\} \times \{H\}$ where $N = i^\pm(H)$ and $x_1 < \cdots < x_N$. Conditionally on H and $H > 0$, the random vector (X, Y) is defined to be uniform on $\{(x_1, x_2), (x_3, x_4), \ldots, (x_{N-1}, x_N)\}$. Conditionally on H and $H < 0$ it is uniform on $\{(x_2, x_1), (x_4, x_3), \ldots, (x_N, x_{N-1})\}$.*

2. *If $\mu = \eta + \mu_0$ and $\nu = \eta + \nu_0$ with $\mu_0 \perp \nu_0$, with probability $\eta(\mathbf{R})$ the random vector (X, Y) satisfies $X = Y$ and $X \sim \eta$. On the complementary event, with probability $1 - \eta(\mathbf{R}) = \mu_0(\mathbf{R})$ it is distributed as the coupling of the singular probability measures $\mu_0(\mathbf{R})^{-1}\mu_0$ and $\nu_0(\mathbf{R})^{-1}\nu_0$ defined in the first item.*

Remark 1.2. To make Theorem 1.1 a rigorous definition of the excursion coupling we will have to prove that $\theta/2(\mathrm{d}h)$ is a probability density, that N is almost surely both finite and even with $(\mathbf{R} \times \{h\}) \cap \mathrm{Graph}^*(F_\sigma) = (\mathbf{R} \times \{h\}) \cap (\mathrm{Graph}^{*,-}(F_\sigma) \cup \mathrm{Graph}^{*,+}(F_\sigma))$. Moreover, we must prove that the laws of X and Y are μ and ν, respectively.

To illustrate the definition of the excursion coupling given in Theorem 1.1 we provide a few naive examples that we encourage the reader to test by drawing. They should in particular permit her to understand the coupling in terms of transport maps. Transport maps are often more intuitive than plans. However, they are less symmetric and general. The excursion coupling can in particular unfortunately not always be represented by a transport map.

Example 1.3. Let μ be the uniform measure on $[0, 1] \cup [2, 3]$ and ν uniform on $[1, 2] \cup [2, 4]$. Then F_σ is continuous and defined piecewise by

$$
F_\sigma(t) = \begin{cases}
t/2 & \text{if } t \in [0, 1] \\
1 - t/2 & \text{if } t \in [1, 2] \\
-1 + t/2 & \text{if } t \in [2, 3] \\
2 - t/2 & \text{if } t \in [3, 4] \\
0 & \text{otherwise.}
\end{cases}
$$

For (X, Y, H) the coupling in Theorem 1.1 we have H uniform on $[0, 1/2]$ and $H = F_\sigma(X)$. Moreover, the excursion coupling is induced by a map T, i.e $\pi = (\mathrm{id}, T)_{\#}\mu = \mathrm{Law}(X, T(X))$ with $\mathrm{Law}(X) = \mu$. The map T is defined by

$$
T(x) = \begin{cases}
2 - x & \text{if } x \in [0, 1] \\
6 - x & \text{if } x \in [2, 3].
\end{cases}
$$

This map is locally non increasing on each of the two intervals but not globally. Finally, one can check that $Y = T(X)$ has law $\nu = T_{\#}\mu$.

Example 1.4. More generally, let a be in $[0, 1)$ and assume that μ is uniform on $[0, 1 + a] \cup [2, 3 - a]$ and ν on $[1 + a, 2] \cup [3 - a, 4]$. We have

$$
F_\sigma(t) = \begin{cases}
t/2 & \text{if } t \in [0, 1 + a] \\
1 + a - t/2 & \text{if } t \in [1 + a, 2] \\
a - 1 + t/2 & \text{if } t \in [2, 3 - a] \\
2 - t/2 & \text{if } t \in [3 - a, 4] \\
0 & \text{otherwise.}
\end{cases}
$$

Again the excursion coupling is $\pi = (\mathrm{id}, T)_{\#}\mu = \mathrm{Law}(X, T(X))$. The random variable H has density 1 on $[0, a]$, and 2 on $[a, (1 + a)/2]$. The map T is defined by

$$
T(x) = \begin{cases}
4 - x & \text{if } x \in [0, 2a) \\
2 + 2a - x & \text{if } x \in [2a, 1 + a] \\
6 - 2a - x & \text{if } t \in [2, 3 - a]
\end{cases}
$$

The three following examples illustrate that the excursion coupling may fail to be represented by a transport map T from μ to ν. This occurs in particular when μ possesses a non trivial atomic part (see Examples 1.5 and 1.6) or when μ and ν are not mutually singular as in Example 1.7.

Example 1.5. Let $\mu = \delta_0$ be the Dirac mass in zero and μ be uniform on $[-1, 1]$. Therefore F_σ is given by

$$F_\sigma(t) = \begin{cases} -t/2 & \text{if } t \in [-1, 0[\\ (1-t)/2 & \text{if } t \in [0, 1] \\ 0 & \text{otherwise.} \end{cases}$$

This is completed be the vertical segment between $-1/2$ and $1/2$ on the y-axis. The excursion coupling is not given by a map as before. Instead of that, we uniformly choose h in $[-1/2, 1/2]$ on the y-axis and map the mass at 0 to $T_h(0) = -1 - 2h$ if $h < 0$ and $T_h(0) = 1 - 2h$ if $h \geq 0$. Therefore, we have $X = 0$ and $Y = T_H(X)$. The transport map T of the previous examples is replaced by a random map T_H.

Example 1.6. For a more complex excursion coupling with atoms we consider $\mu = (1/3)(\delta_{-1} + 1_{[-1,0]}dx + \delta_0)$ and ν uniform on $[0, 1]$. In this case $T_H(X) = (1-X)/3$ if $X \in (-1, 0)$ which correspond to a deterministic transport (notice $H = (2+X)/3$ in this case). However, conditioned on $\{X = -1\}$, the height H is uniform on $[0, 1/3]$. It is uniform on $[2/3, 1]$ if $X = 0$ and we have $T_H(X) = 1 - H$ both for $X = 0$ and $X = 1$. This correspond to a randomized multivalued map (probabilists say a kernel).

Note that in this example the excursion coupling can be represented by a transport map from ν to μ –in place of from μ to ν.

Example 1.7. Let now μ be uniform on $[0, 1]$ and ν uniform on $[0, 3]$. Then the common measure $\mu \wedge \nu = \frac{1}{3}1_{[0,1]}dx$ represents the mass that stays during the transport. The information on the moving mass is contained in the function

$$F_\sigma(t) = \begin{cases} 2t/3 & \text{if } t \in [0, 1[\\ 1 - t/3 & \text{if } t \in [1, 3] \\ 0 & \text{otherwise.} \end{cases}$$

that permits one to consider the map $T_{\text{move}} : x \in [0, 1] \mapsto 3 - 2x$. Finally the excursion transport is multivalued. With probability $1/3$ the mass is staying $T_{\text{stay}}(x) = x$, with probability $2/3$ it is moving with T_{move}. This writes $\pi = (\text{id}, T_{\text{stay}})_\#(\mu \wedge \nu) + (\text{id}, T_{\text{move}})_\#(\mu - (\mu \wedge \nu))$.

Remark 1.8. Note that if $\mu \perp \nu = 0$ and μ has atomic part zero, almost surely H simply writes $H = F_\sigma(X)$ and since the arches corresponding to the height H are almost surely properly separated, Y is a function of X. This behavior that was illustrated in Examples 1.3 and 1.4 corresponds to the one observed in the strictly concave case, as proved in [PPS15]. If $\mu \perp \nu \neq 0$ and μ (or $\mu - \mu \wedge \nu$) still has atomic part zero we are in the situation of Example 1.7 where the excursion coupling writes $\pi = (\text{id}, T_{\text{stay}})_\#(\mu \wedge \nu) + (\text{id}, T_{\text{move}})_\#(\mu - (\mu \wedge \nu))$. This is again the shape obtained in [GM96, PPS15].

2   Monotonicity of the solutions of the $L^{1,q}$ and L^{1^-} transport problems

In this section we prove the implications *1 ⇒ 3* and *2 ⇒ 3* of the Main Theorem. The title "(cyclical-)monotonicity" is a generic name in Optimal Transport that in this section is represented by property *3*. (Cyclical-)monotonicity results are variations of the following simple swapping lemma that concerns the L^p transport problem for cycles of length two. For the L^{1^-} transport problem we will firstly interpret Lemma 2.1 for $p < 1$ and secondly let p go to 1. For the $L^{1,q}$ transport problem will need a more deep result than Lemma 2.1: It will be Lemma 2.5 on page 13. Wheras the first lemma is classical and could be skipped, the second is not. We state and prove both results for the sake of completeness and because the proof of Lemma 2.1 is required in Lemma 2.5.

Lemma 2.1 (Swapping lemma)**.** *Let p be positive and π be in $\mathrm{Marg}_p^*(\mu, \nu)$. Assume moreover $C_p(\pi) < +\infty$. Consider a set $S \subset \mathbf{R}^2$ such that $\pi(S) = 1$ and for $(x, y) \in S$ any neighborhood of (x, y) has positive measure (Notice that the support of π satisfies these conditions, so that we may choose $S = \mathrm{Spt}(\pi)$). For any (x, y) and (x', y') in S, the following holds:*

$$|y - x|^p + |y' - x'|^p \leq |y - x'|^p + |y' - x|^p. \tag{2}$$

Proof. Striking for a contradiction, suppose that the opposite identity holds. Then for some $(x, y), (x', y') \in S$ there exists $\varepsilon > 0$ such that for every (a, b, a', b') with $\max(|a-x|, |a'-x'|, |b-y|, |b'-y'|) \leq \varepsilon$ one has $|b-a'|^p + |b'-a|^p < |b-a|^p + |b'-a'|^p$. The two balls (in the ∞-norm) of radius ε centered in (x, y) and (x', y') have positive measure. We choose ε small enough to make their intersection the empty set. Then it is easily possible to replace π by a competitor π' that coincide with it outside the balls $\mathcal{B}_\infty((a, b), \varepsilon)$, $\mathcal{B}_\infty((a', b'), \varepsilon)$, $\mathcal{B}_\infty((a, b'), \varepsilon)$ and $\mathcal{B}_\infty((a', b), \varepsilon)$, has marginals μ and ν and satisfies $C_p(\pi') < C_p(\pi)$ (see e.g [GM96, page 129] or the end of the proof of Lemma 2.5), a contradiction.   □

This simple principle allows for a complete characterization of π in the case $p > 1$. We provide it now for the sake of completeness and for further comparison with the L^{1^-} limit and $L^{1,q}$ secondary problems.

The L^p transport problem for $p > 1$ In this case the identity $|y - x|^p + |y' - x'|^p \leq |y - x'|^p + |y' - x|^p$ obtained by applying Lemma 2.1 is equivalent to $(y' - y)(x' - x) \geq 0$. This may be graphically represented by the condition that segments in $\mathbf{R}^2$ connecting for every $(x, y) \in S$ the point $(x, 1)$ to $(y, 0)$ are not allowed to cross each other, see Figure 1. These segments may be interpreted as transport routes between μ concentrated on a line parallel to the x-axis and ν concentrated on the x-axis. The striking fact is that, given μ and ν, the elements π of $\mathrm{Marg}(\mu, \nu)$ being concentrated on such a set S are in fact reduced to a single transport plan, called the *quantile transport plan*. The latter is the law of (G_μ, G_ν) on the probability space $(]0, 1[, \lambda)$ of quantile levels, where λ
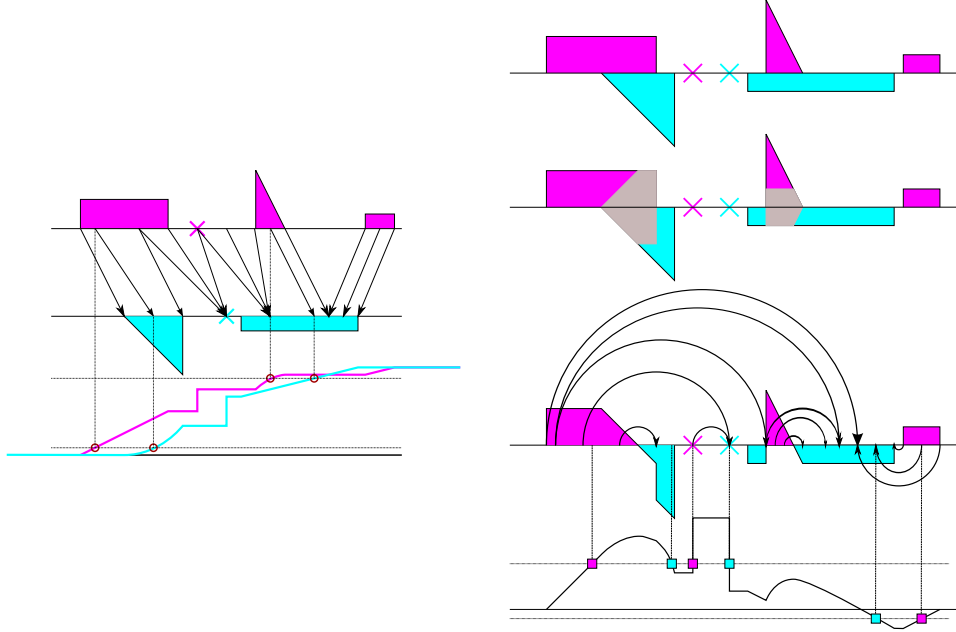
Fig. 1: Left: the quantile transport is based on simultaneous simulation through the cumulative distribution functions F_μ and F_ν. The dashed horizontal lines (below on the figure) are chosen uniformly in $]0,1[$. The transport routes are materialized by arrows (above on the figure) that do not cross.

Fig. 2: Right: the moving mass of the excursion coupling is simulated through the intersections with $F_\nu - F_\nu$; the commun mass $\mu \wedge \nu$ stays on the on the same place. Given the horizontal line the excursion is chosen uniformly (among two or one pairs for the lines on the figure). The increasing intersection corresponds to μ and the decreasing one to ν. The transport routes are materialized by arches that do not cross.

is the Lebesgue measure and for any a real probability measure η, the quantile function G_η is defined as a pseudo-inverse of $F_\eta : x \mapsto \eta(]-\infty, x])$, namely

$$G_\eta(\alpha) = \inf\{x \in \mathbf{R} : F_\eta(x) \geq \alpha\}$$

(this infimum is a minimum).

2.1   The L^p transport problem for $p < 1$ and the L^{1^-} limit transport problem

Important preliminary comparaison to the convex case $p > 1$ In the previous section we recalled that for $p > 1$, provided the L^p transport problem admits a
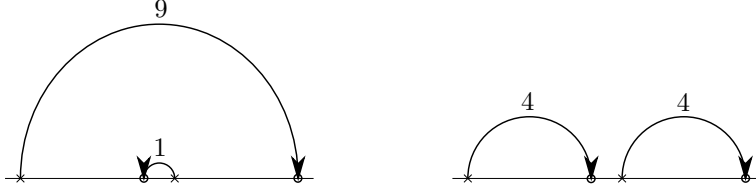
Fig. 3: For transporting the two crosses to the two circles, each symbol being of mass 1/2, the left pattern is optimal for $p \leq 1/2$, the right is for $p \geq 1/2$.

solution π_p with $C_p(\pi_p) < \infty$, this transport plan π_p is the *unique solution* and it is the quantile coupling. In particular it is *independent of the value of p*.

The two last assumptions are *both false* in the case $p < 1$, which we prove in the following example.

Example 2.2. Set $\mu = \frac{1}{2}(\delta_0 + \delta_5)$ and $\nu = \frac{1}{2}(\delta_4 + \delta_9)$. The set of transport plans is easily described as

$$\mathrm{Marg}(\mu, \nu) = \{\pi^\lambda = \lambda\pi'' + (1-\lambda)\pi' : \lambda \in [0,1]\}$$

where $\pi' = \frac{1}{2}(\delta_{0,4} + \delta_{5,9})$ and $\pi'' = \frac{1}{2}(\delta_{0,9} + \delta_{5,4})$. In the case $p = 1/2$ every transport plan π gives the same global cost $C_p(\pi^\lambda) = \lambda\frac{1}{2}(\sqrt{4} + \sqrt{4}) + (1-\lambda)\frac{1}{2}(\sqrt{9} + \sqrt{1}) = 2$. This proves the *non-uniqueness*; $\mathrm{Marg}^*_{1/2}(\mu, \nu) = \mathrm{Marg}(\mu, \nu)$. Moreover, with the same marginals, depending whether $p < 1/2$ or $p > 1/2$ the measure $\pi' = \pi^0$ or $\pi'' = \pi^1$, respectively, is the unique optimal transport plan. Hence the solution *depends on the value of p*.

Example 2.3. With this example strongly inspired by Example 1.1 in [McC01] we again illustrate that the solution (here it is unique) depends on $p \in (0,1)$. Consider as before (see Example 1.3) μ the uniform measure on $[0,1] \cup [2,3]$ and ν uniform on $[2,3] \cup [3,4]$. In this case the algorithm developed in [McC01] shows that the optimal transport plan is induced by a map, namely the map

$$T(x) = \begin{cases} 4 - x & \text{if } x \in [0, a_p) \\ 2 - x & \text{if } x \in [a_p, 1] \\ 4 - x & \text{if } t \in [2, 2 + a_p) \\ 6 - x & \text{if } t \in [2 + a_p, 3]. \end{cases}$$

As we announced this map depends on p through the solution $a_p > 0$ of the equation $(4 - 2a_p)^p + (2a_p)^p = 2(2 - 2a_p)^p$. It is not a surprise that a_p goes to zero as p tends to 1 since the formula for $a_p = 0$ corresponds to the excursion coupling. Note that on each interval the map is decreasing.

In the two next paragraphs we prove that in the L^p transport problem (for $p < 1$) arches do not connect and do not cross. In the third next paragraph this will be transmitted to the L^{1^-} (limit) transport problem where we also prove that nested arches have the same orientation, completing the proof of *1 ⇒ 3* in the Main Theorem.

Conclusion concerning coinciding points for $p < 1$ Let S satisfying (2) in the swapping lemma, Lemma 2.1 for $p < 1$ and (x, y, x', y') be in $S \times S$. Suppose moreover that y and x' coincide meaning that one can take a route from x to y and a second one from $y = x'$ to y'. Then studying the variations of $t \in [0, +\infty[\mapsto (t+h)^p - t^p$ and considering $h = |y - y'|$ we see that this can only occur in one of the degenerate case $x = y$ or $x' = y'$ – compare with the two most right subfigures in Figure 4. With respect to the terminology of Definition 0.3 we have proved that a solution π of the L^p transport problem ($p < 1$) is concentrated on a set whose *arches do not connect*. This well-known situation has been studied be Gangbo and McCann in [GM96, Proposition 2.9] more than 20 years ago. It yields that π can be decomposed as the sum $\pi = \pi_\Delta + \pi_0$ where $\pi_\Delta = (\mathrm{id} \times \mathrm{id})_\# (\mu \wedge \nu)$. Let us remind the argument for the sake of completeness.

Lemma 2.4. *If $\pi \in \mathrm{Marg}(\mu, \nu)$ is concentrated on a set S whose arches do not connect, it can be decomposed as follows: $\pi = \pi_\Delta + \pi_0$ where $\pi_\Delta = (\mathrm{id} \times \mathrm{id})_\# (\mu \wedge \nu)$ and, for $\Delta = \{(x, y) \in \mathbf{R}^2 : y = x\}$ denoting the diagonal, it holds $\pi_0(\Delta) = 0$. Consequently, π_0 is a transport plan with the two marginals singular with respect to each other.*

Proof. Let us write $\pi = \pi_\Delta + \pi_0$ where π_0 is concentrated on $S \setminus \Delta$ and π_Δ is concentrated on Δ. Therefore π_Δ writes $(\mathrm{id} \times \mathrm{id})_\# \eta$ and the marginals of π_0 are $\mu - \eta$ and $\nu - \eta$. They are respectively concentrated on the two projections $\{x' \in \mathbf{R} : \exists y' \neq x', (x', y') \in S \setminus \Delta\}$ and $\{y \in \mathbf{R} : \exists x \neq y, (x, y) \in S \setminus \Delta\}$ of the set $S \setminus \Delta$. The fact that these two sets do not intersect is the direct consequence of our assumption. It follows $\mu - \eta \perp \nu - \eta$ which is equivalent to $\eta = \mu \wedge \nu$. $\square$

Interpretation of the swapping lemma for $p < 1$ **and non coinciding points**
For $p < 1$ the use of the swapping lemma furnishes, compared to $(y'-y)(x'-x) \geq 0$, less direct information. Equation (2) may be seen, similarly as in Example 2.2, as a competition of two transport plans, each transporting two points in two other points, in one or the other way. For this reduced transport problem if $|y - x|^p + |y' - x'|^p \neq |y' - x|^p + |y - x'|^p$ at most one of the two is true $(x, y, x', y') \in S \times S$ or $(x, y', x', y) \in S \times S$. Unlike the situation studied for $p > 1$, determining which transport is better does not only depend on the relative positions of x and y with respect to x' and y', respectively, but on the 4! relative positions of the four points. Moreover, even though this ranking of the four point is important and yields the conclusion in configurations $xxyy$ and $xyyx$ we know since Example 2.2 and Figure 3 that it does not permit to conclude in configuration $xyxy$.

Notation. The notation $xyxy$ denotes the configurations of two routes (x, y) and (x', y') where $x < y < x' < y'$ or $x < y' < x' < y$ or $x' < y < x < y'$ or $x' < y' < x < y$, the different alternatives being signed as $xyx'y'$, $xy'x'y$, $xy'xy'$ and $x'yxy'$, respectively.

Since the L^p problem in $\mathrm{Marg}(\mu, \nu)$ is in bijection with the one in $\mathrm{Marg}(\nu, \mu)$, when we study configuration $xyxy$ we in fact also study $yxyx$. In order to consider all configurations without coinciding points we finally only have to

look at $xxyy$, $xyyx$ and $xyxy$. Here is the conclusion for these three cases. They are also illustrated on Figure 4 (corresponding, in the same order, to the first three pairs of patterns, from the left).

- In the first case $xx'y'y$ is allowed and $xx'yy'$ forbidden. To see that, one can study the variations of $t \in [0, +\infty[\mapsto (t+h)^p - t^p$ where $h > 0$ and apply it to $h = |y - y'|$.

- In the second case $xyy'x'$ is allowed and $xy'yx'$ forbidden. This is a simple consequence of the fact that $d \in \mathbf{R}^+ \mapsto d^p$ is increasing.

- In the last case, as observed in Example 2.2, it depends on p whether $xyx'y'$ or $xy'x'y$ is forbidden or authorized.

With the two first points we have proved that the arches of S *do not cross*. Let us give another, less descriptive argument. When four points bound three intervals of length a, b and c, in this order, the cost for a crossing corresponds to $(a+b)^p + (b+c)^p$. It is easy to check that it smaller than $a^p + c^p$ or $(a+b+c)^p + b^p$, that is the cost of the swapped (and uncrossed) configuration.

From the L^p to the L^{1^-} transport problem In this paragraph we prove that the solutions of the L^{1^-} transport problem have arches that do not connect, do not cross and have the same orientation when they are nested. This is implication *1 ⇒ 3* of Main Theorem.

Consider π such that there exists a sequence $(\pi_n)_{n \in \mathbf{N}}$ weakly converging to π where $\pi_n \in \mathrm{Marg}^*_{p_n}(\mu, \nu)$ for $p_n \to 1^-$. It also holds $\pi_n \otimes \pi_n \to \pi \otimes \pi$, actually an equivalent fact. Hence, if $F \subset \mathbf{R}^4$ is a closed set and $\pi_n \otimes \pi_n(F) = 1$, the equation goes to the limit. In particular this holds for the complementary set F_1 of the open set $\{(x, y, x', y') \in \mathbf{R}^4 : x < x' < y < y' \text{ or } x' < x < y' < y\}$ that encodes the condition on non-intersecting arches in configuration $xxyy$ (first condition of Definition 0.3). As it is satisfied by π_n it is also satisfied by π. The same argument applies to the other configurations: $yyxx$, $xyyx$ and $yxxy$. Therefore, π is concentrated on a set whose arches *do not cross*.

Suppose by contradiction that the set $F_3^c \subset \mathbf{R}^4$ of pairs of nested arches that do not have the same orientation has positive measure for $\pi \otimes \pi$ and without loss of generality we assume that it happens for the configuration $xy'x'y$. Due to the countable additivity of $\pi \otimes \pi$ there exists rational numbers $x < y' < x' < y$ and $\varepsilon \in \mathbb{Q}^+$ with $\min(|y' - x|, |x' - y'|, |y - x'|) > \varepsilon$ such that $U =]x \pm \varepsilon/10[\times]y \pm \varepsilon/10[\times]x' \pm \varepsilon/10[\times]y' \pm \varepsilon/10[$ has positive measure. Since

$$[(y'-x)+2\varepsilon/10]^{p_n}+[(y-x')+2\varepsilon/10]^{p_n} \leq [(y-x)-2\varepsilon/10]^{p_n}+[(x'-y')-2\varepsilon/10]^{p_n}$$

for p_n close enough to 1, the open set U has measure zero for the corresponding measures $\pi_n \otimes \pi_n$. Therefore, it has measure zero for $\pi \otimes \pi$ as well, a contradiction. We conclude that π concentrated on a set whose *nested arches have the same orientation*.

The implication *1 ⇒ 3* finally amounts to prove that π is concentrated on a set satisfying the second condition of Definition 0.3: *the arches do not connect*.

Recall from Lemma 11 that $\pi_n = (\mathrm{id} \times \mathrm{id})_\#(\mu \wedge \nu) + \pi_n^0$ where π_n^0 has marginals $\mu^0 = \mu - (\mu \wedge \nu)$, $\nu^0 = \nu - (\mu \wedge \nu)$ and $\mu^0 \perp \nu^0$. Thus π is the limit of the sequence if and only if it can be written $(\mathrm{id} \times \mathrm{id})_\#(\mu \wedge \nu) + \pi^0$ where π^0 is the limit of $(\pi_n^0)_{n \in \mathbf{N}}$. Therefore, π^0 has marginals μ_0 and ν_0. Thus there exists A and $B = A^c$ such that π is concentrated on $E = [(A \times \mathbf{R}) \cap (\mathbf{R} \times B)] \cup \Delta$. This exactly means that it is concentrated on a set whose arches do not connect.

2.2   The $L^{1,q}$ secondary transport problem

Swapping lemma for the $L^{1,q}$ secondary transport problem and interpretation
We furnish now a swapping lemma corresponding to the $L^{1,q}$ transport problem.

Lemma 2.5 (Swapping lemma for the secondary problem)**.** *Let π be a transport plan such that $C_1(\pi) < +\infty$ and $\pi \in \mathrm{Marg}_{1,q}^{**}(\mu, \nu)$ for some fixed $q < 1$. There exists a set $S \subset \mathbf{R}^2$ with $\pi(S) = 1$ such that for any (x,y) and (x',y') in S it holds*

$$|y - x| + |y' - x'| \le |y - x'| + |y' - x| \tag{3}$$

and if $|y - x| + |y' - x'| = |y - x'| + |y - x'|$, the following holds:

$$|y - x|^q + |y' - x'|^q \le |y - x'|^q + |y' - x|^q. \tag{4}$$

Proof. Let π be an optimal transport plan for the L^1 primary and the $L^{1,q}$ secondary transport problem. Let $N_{y,-}$ be $\{(x,y) \in \mathrm{Spt}(\pi) : \exists \varepsilon > 0, \pi([x - \varepsilon, x + \varepsilon] \times [y - \varepsilon, y]) = 0\}$. It is the set of routes in the support of π such that at some size $\varepsilon > 0$ no mass is transported from the ball of center x and radius ε to the left of y at distance less than ε. We have $\pi(N_{y,-}) = 0$. The proof of this result is postponed to Lemma 2.6. The symmetrically defined sets $N_{y,+}$, $N_{x,-}$ and $N_{x,+}$ have measure zero too (for instance $N_{x,+}$ here denotes the set $\{(x,y) \in \mathrm{Spt}(\pi) : \exists \varepsilon > 0, \pi([x, x + \varepsilon] \times [y - \varepsilon, y + \varepsilon]) = 0\}$). We define now S as $S_0 \setminus (N_{y,-} \cup N_{y,+} \cup N_{x,-} \cup N_{x,+})$ where S_0 is any set, as for instance $\mathrm{Spt}(\pi)$, satisfying the condition in Lemma 2.1. Observe that $\pi(S) = 1$.

We are now ready for the proof. It is almost identical to that of Lemma 2.1 that we invite the reader to read again: the principle is that π is compared to a competitor π' defined rerouting part of the mass around (x,y) and (x',y') to mass aroud (x,y') and (x',y). From Lemma 2.1 we note that (3) is satisfied for any $(x,y,x',y') \in S \times S$. Aiming for a contradiction, suppose that for some routes (x,y) and (x',y') in S it holds at the same time $|y-x|+|y'-x'| = |y-x'|+|y'-x|$ and $|y'-x|^q+|y-x'|^q < |y-x|^q+|y'-x'|^q$. Until the end of the paragraph let us see that without loss of generality we can assume $x < x' \le y < y'$: The first equation induces 'max$(x,x') \le$ min(y,y') or max$(y,y') \le$ min(x,x')'. Without loss of generality we can assume the first case. The second inequality implies $x \ne x'$ and $y \ne y'$ and without loss of generality we assume $x < x'$. This finally implies $y < y'$.

Apparently the swapping method used in the proof of Lemma 2.1 that consists in picking mass in equal quantity around the points x, x', y and y' can be

applied without problem providing a competitor $\pi' \in \mathrm{Marg}(\mu, \nu)$ with $C_q(\pi') < C_q(\pi)$, a contradiction. However, this argument only works if $x' < y$ and not directly if $x' = y$ because $\pi' \in \mathrm{Marg}_1^*(\mu, \nu)$ can not be certified: during the swap the routes (x, y) and (x', y') with $x < x' = y < y'$ have a priori in their neighborhood routes $(\tilde{x}, \tilde{y})$ and $(\tilde{x}', \tilde{y}')$ with $\tilde{x} < \tilde{y} < \tilde{x}' < \tilde{y}'$ that do no longer satisfy the preliminary condition for (4) (we mean $|y-x|+|y'-x'| = |y-x'|+|y-x'|$), so that $C_1(\pi') > C_1(\pi)$ is made possible. The relation $C_q(\pi') < C_q(\pi)$ is no longer a contradiction because $\pi' \notin \mathrm{Marg}_1^*(\mu, \nu)$, i.e $C_1(\pi') = C_1(\pi)$ is false. In this critical situation let us call z the point $x' = y$. As (x, y) and (x', y') are not in $N_{y,-} \cup N_{y,+} \cup N_{x,-} \cup N_{x,+}$ we can swap selecting the mass on the proper side of z. More precisely for π there is some mass traveling from a neighborhood (as small as we want) of x to a small right neighborhood of z. There is the same mass on a small left neighborhood of z transported to a small neighborhood of y'. We swap for defining π': the mass around x is transported around y' and the mass directly on the left of z is transported directly to the right of z. We obtain $C_q(\pi') < C_q(\pi)$ and keep $C_1(\pi') = C_1(\pi)$, a contradiction. Let us reformulate this argument in a less descriptive manner. There exists $\varepsilon > 0$ as well as two positive measures γ and γ' with the same mass that satisfies a bunch of properties. Due to our assumptions we can in fact assume the following: $\gamma \leq \pi$ and γ is concentrated on $[x - \varepsilon, x + \varepsilon] \times [z, z + \varepsilon]$; in a parallel way $\gamma' \leq \pi$ and γ' is concentrated on $[z - \varepsilon, z] \times [y' - \varepsilon, y' + \varepsilon_1]$. Moreover, this can be done for every ε close enough to zero. Setting $m = \gamma'(\mathbf{R}^2) = \gamma'(\mathbf{R}^2)$ we note that $\zeta' = (1/m)(\mathrm{proj}_\#^1 \gamma \times \mathrm{proj}_\#^2 \gamma') + (1/m)(\mathrm{proj}_\#^1 \gamma' \times \mathrm{proj}_\#^2 \gamma)$ has the same marginals as $\zeta = \gamma + \gamma'$. Let us now denote $\pi - \zeta + \zeta'$ by π'. One can easily check that for ε small enough it holds at the same time $C_1(\pi') = C_1(\pi)$ and $C_q(\pi') > C_q(\pi)$.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Lemma 2.6. *Let π be a solution to the $L^{1,q}$ secondary optimal transport and $N_{y,-}$ be $\{(x, y) \in \mathrm{Spt}(\pi) : \exists \varepsilon > 0, \pi([x - \varepsilon, x + \varepsilon] \times [y - \varepsilon, y]) = 0\}$. Then $\pi(N_{y,-}) = 0$.*

Proof. Fix $a < b$ two rational numbers. Consider the set $A_{a,b}$ of points $(x, y) \in \mathrm{Spt}(\pi|_{]a,b[\times\mathbf{R}})$ such that there exists $\varepsilon > 0$ with $\pi([x - \varepsilon, x + \varepsilon] \times [y - \varepsilon, y]) = 0$ and moreover $x \in]a, b[\subset [x - \varepsilon, x + \varepsilon]$. Then every point $y \in \mathrm{proj}^2(A_{a,b})$ is in the support of the measure $\nu_{a,b}$ defined by $(\mathrm{proj}^2)_\# \pi|_{]a,b[\times\mathbf{R}}$, but it is also isolated on the left in the sense that $\nu_{a,b}([y - h, y]) = 0$ for h small enough. One can easily convince that there are countably many such points y and they are not atoms. Therefore, $\nu(\mathrm{proj}^2(A_{a,b})) = 0$ and $\pi(A_{a,b}) = \pi|_{]a,b[\times\mathbf{R}}(A_{a,b}) \leq \nu(\mathrm{proj}^y(A_{a,b})) = 0$. Finally, since $N_{y,-} = \bigcup_{a<b\in\mathbb{Q}} A_{a,b}$ we find $\pi(N_{y,-}) = 0$. $\square$

We can now give the geometric meaning for the routes of sets S as in Lemma 2.5. As for the L^{1^-} problem we compare the different configurations with their alternative – recall the list on page 12 – and again these comparaisons are illustrated on the three most left patterns of Figure 4.
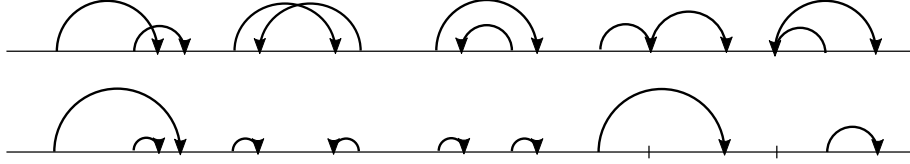
Fig. 4: Forbidden (first line) versus authorized (second line) configurations in the $L^{1,q}$ secondary transport problem, where $p < 1$.

- Pattern $xx'y'y$ is allowed and $xx'yy'$ forbidden. To see that, since $|y' - x'| + |y - x| = |y' - x| + |y - x'|$ we have to look at the secondary problem.

- Pattern $xyy'x'$ is allowed and $xy'yx'$ forbidden because as in the L^p problem, (3) is a strict inequality.

- Finally, pattern $xyx'y'$ is allowed and $xy'x'y$ is forbidden because as in the L^p problem, (3) is a strict inequality.

Some of the points x, y, x', y' may be equal. Swapping does not change the cost if $x = x'$ or $y = y'$. We have only to look at $x = y'$ and $x' \neq y$ and see that it is never better than $x = y$ and $x' \neq y'$ (see on Figure 4 the two patterns on the right).

- For the pattern $x(yx)y$, meaning $x < y = x < y$, we have $|y'-x'|+|y-x| = |y' - x| + |y - x'|$ but the secondary problem tells us to choose $xx'y'y$ in place of $xyx'y'$.

- For $(xy)xy$, from the primary problem we choose $x = y < x' < y'$ in place of $x = y' < x' < y$.

Finally as for the L^{1^-} limit transport problem, we have proved that if π is a solution of the $L^{1,q}$ secondary transport problem it is concentrated on a set S whose arches do not cross, do not connected, and have the same orientation when they are nested.

3   Transport plans concentrated on monotone sets of arches are the excursion coupling

3.1   Proof of Theorem 1.1 defining the excursion coupling

We need to explain why Theorem 1.1 can define the excursion coupling. We only need to investigate the construction in case *1* where μ and ν are singular ($\mu \perp \nu$), which we thus assume in the present subsection.

With the following lemma we will be able to handle with F_σ as if it were a continuous function.

Lemma 3.1 (Generalized intermediate value theorem)**.** *For any càdlàg function F, any $x_0, x_1 \in \mathbf{R}$ and $h \in \mathbf{R}$ such that $x_0 < x_1$ and $(F(x_0)-h)(F(x_1)-h) < 0$, there exists $x \in]x_0, x_1]$ such that $(x, h) \in \mathrm{Graph}^*(F)$.*

Proof. Let F, x_0, x_1 be as in the statement. Without loss of generality we assume $h = 0$, $F(x_0) < 0$ and $F(x_1) > 0$. Let x be the infimum of $A = \{x \in [x_0, x_1] : F(x) \geq 0\}$. As $A \ni x_1$ it is a not empty set. As moreover F is right continuous we have $x \in A$ and $x \neq x_0$. Due to the definitions of A and x, any $x' < x$ satisfies $F(x') < 0$. If F is left-continuous at x we have $F(x) = 0$. If it is not, as $F(x-) < 0$ we also have $(x, 0) \in \mathrm{Graph}(F^*)$. $\qquad\square$

A result by Bertoin and Yor establishes a relation between the occupation measure in a set $B \subset \mathbf{R}$ of a function F of finite variation and its variations when its values are in B, see Remark 5.6 for details. In particular we can apply their Theorem 1 in [BY14] (see also their §5) to $F = F_\sigma$ and $B = \mathbf{R}$, and relate the total variation (without its saltus part) with the number of solutions of the equations $F_\sigma = h$, where h goes over $\mathbf{R}$: For any points $s \leq t$ in $\overline{\mathbf{R}}$

$$TV_s^t(F_\sigma) - \sum_{x \in]s,t]} |\sigma(x)| = \int_{\mathbf{R}} \mathrm{card}\{x \in [s,t] : (x, h) \in \mathrm{Graph}(F_\sigma)\} \, \mathrm{d}h \qquad (5)$$

$$= \int_{\mathbf{R}} \mathrm{card}[(\mathbf{R} \times \{h\}) \cap \mathrm{Graph}(F_\sigma)] \, \mathrm{d}h$$

$$= \int_{\mathbf{R}} i_{]s,t]}(h) \, \mathrm{d}h$$

where

$$i_A : h \in \mathbf{R} \mapsto i_A(h) = \mathrm{card}\{x \in A : F_\sigma(x) = h\}$$
$$= \mathrm{card}\{x \in A : (x, h) \in \mathrm{Graph}(F_\sigma)\}$$

is the so-called Banach indicatrix, after [Ban25]. Notice that in [BY14] the result is stated for $s = 0$ and $t \in [0, \infty[$. Our statement on a general interval $]s, t]$ is a trivial generalization. Another difference is that we only apply the formula to the occupation measure of F_σ in B when $B = \mathbf{R}$. It is hardly more than a simple exercise to rewrite (5) for generalized solutions of $F_\sigma = h$, which permits at the same time to forget about the saltus part. For this purpose we introduce the generalized Banach indicatrix. We set

$$i_A^* : h \in \mathbf{R} \mapsto i_A^*(h) = \mathrm{card}\{x \in A : h \in F_\sigma^*(x)\}$$
$$= \mathrm{card}\{x \in A : (x, h) \in \mathrm{Graph}(F_\sigma^*)\} \in \mathbf{N} \cup \{\infty\}.$$

Therefore, (5) yields

$$TV_s^t(F_\sigma) = \int i_{]s,t]}^*(h) \, \mathrm{d}h. \qquad (6)$$

A consequence of Theorem 1 in [BY14] specifies that not all intersections with the graph need to be considered. Recall first that $\mathrm{Graph}^+(F_\sigma^*)$ and $\mathrm{Graph}^-(F_\sigma^*)$ have been defined in Section 1. Bertoin and Yor proved that almost surely for h, the (generalized) Banach indicatrix $i_{]s,t]}^*$ equals $i_{]s,t]}^{*,\pm} := i_{]s,t]}^{*,+} + i_{]s,t]}^{*,-}$ where

$$i_{]s,t]}^{*,+} : h \in \mathbf{R} \mapsto i^+(h) = \mathrm{card}\{x \in \mathbf{R} : (x,h) \in \mathrm{Graph}^+(F_\sigma^*)\} \in \mathbf{N} \cup \{\infty\},$$

$$i_{]s,t]}^{*,-} : h \in \mathbf{R} \mapsto i^-(h) = \mathrm{card}\{x \in \mathbf{R} : (x,h) \in \mathrm{Graph}^-(F_\sigma^*)\} \in \mathbf{N} \cup \{\infty\}$$

(with obvious notation Bertoin and Yor in fact proved the equality $i = i^\pm$ for the non yet generalized indicatrix $i = i^+ + i^-$). This also comes from Theorem 1 in [BY14] (where $n^x(t)dx = \lambda^x(t)dx$):

$$TV_s^t(F_\sigma) = \int i_{]s,t]}^{*,\pm}(h)\,\mathrm{d}h \leq \int i_{]s,t]}^*(h)\,\mathrm{d}h = TV_s^t(F_\sigma). \tag{7}$$

With the next result we go further in the analysis.

Proposition 3.2. *For almost every $h \in \mathbf{R}$, the set $(\mathbf{R} \times \{h\}) \cap \mathrm{Graph}^* F_\sigma$ has cardinality a finite and even integer. In fact for almost every $h \in \mathbf{R}$ it holds $i^*(h) = i^{*,+}(h) + i^{*,-}(h)$ with $i^{*,+}(h) = i^{*,-}(h)$. Moreover if $h \neq 0$ the generalized solutions of $F_\sigma = h$ are alternatively crossing positively and negatively, starting with the most-left solution $(x,h) \in \mathrm{Graph}^{*,+}(F_\sigma)$ if $h > 0$ and $(x,h) \in \mathrm{Graph}^{*,-}(F_\sigma)$ if $h < 0$.*

Proof. Due to (6) the generalized Banach indicatrix is finite for almost every $h \in \mathbf{R}$ and with (7) for almost every h we have $(\mathbf{R} \times \{h\}) \cap \mathrm{Graph}^* F_\sigma = [(\mathbf{R} \times \{h\}) \cap \mathrm{Graph}^{*,+} F_\sigma] \cup [(\mathbf{R} \times \{h\}) \cap \mathrm{Graph}^{*,-} F_\sigma]$. With the generalized intermediate value theorem (Lemma 3.1) and as $\lim_\infty F_\sigma = \lim_{-\infty} F_\sigma = 0$ we conclude that $i^*(h)$ is an even number for almost every h. Let us clarify this. Let h be as before and $\bar{x}_0 < \bar{x}_1$ satisfy $(\bar{x}_0, h) \in \mathrm{Graph}^*(F_\sigma)$ and $(\bar{x}_1, h) \in \mathrm{Graph}^*(F_\sigma)$. We moreover assume

$$(x,h) \notin \mathrm{Graph}^*(F_\sigma) \text{ for every } x \in (\bar{x}_0, \bar{x}_1). \tag{8}$$

Therefore $F_\sigma > h$ or $F_\sigma < h$ on all $(\bar{x}_0, \bar{x}_1)$. If not the generalized intermediate value theorem applied to the proper interval $[x, x'] \subset (\bar{x}_0, \bar{x}_1)$ furnishes a contradiction with (8). Since now $F_\sigma > h$ on $(\bar{x}_0, \bar{x}_1)$ – the opposite assumption $F_\sigma > h$ yields the symmetric conclusion – and $(\mathbf{R} \times \{h\}) \cap \mathrm{Graph}^* F_\sigma = [(\mathbf{R} \times \{h\}) \cap \mathrm{Graph}^{*,+} F_\sigma] \cup [(\mathbf{R} \times \{h\}) \cap \mathrm{Graph}^{*,-} F_\sigma]$ we find $(\bar{x}_0, h) \in \mathrm{Graph}^{*,+} F_\sigma$ and $(\bar{x}_1, h) \in \mathrm{Graph}^{*,-} F_\sigma$. A similar argument leads to the fact that $h - F_\sigma$ has a unique sign in the neighborhood of $-\infty$ and in the neighborhood of $+\infty$. It is the sign of h (where we assume $h \neq 0$). Finally $h - F_\sigma$ can only change an even number of times its sign. More precisely, due to the generalized intermediate value theorem applied to F_σ and reminding that this function has limit zero in ∞ and $+\infty$ the points $x_1 < x_1' < x_2 < \cdots < x_{n-1}' < x_n < x_n'$ of $\{x \in \mathbf{R} : (x,y) \in \mathrm{Graph}^*(F_\sigma)\}$ are ordered with $(x_1, h), \ldots, (x_n, h) \in \mathrm{Graph}^{*,+}(F_\sigma)$ and $(x_1', h), \ldots, (x_n', h) \in \mathrm{Graph}^{*,-}(F_\sigma)$ if $h > 0$, $(x_1, h), \ldots, (x_n, h) \in \mathrm{Graph}^{*,-}(F_\sigma)$ and $(x_1', h), \ldots, (x_n', h) \in \mathrm{Graph}^{*,+}(F_\sigma)$ if $h < 0$. $\quad\square$

Coming back to (6), another direct generalization of Bertoin and Yor's study is

$$F_\sigma(t) - F_\sigma(s) = \int_{\mathbf{R}} i^{*,+}_{]s,t]} - i^{*,-}_{]s,t]} \, \mathrm{d}h. \tag{9}$$

It will be useful for recovering F_μ and F_ν from F_σ.

For any measurable set $G \subset \mathbf{R}^2$, let ζ_G denote the following positive measure

$$\zeta_G : E \in \mathcal{T}(\mathbf{R}^2) \mapsto \zeta_G(E) = \int_{-\infty}^{+\infty} \mathrm{card}(x \in \mathbf{R} : (x, h) \in E \cap G) \, \mathrm{d}h.$$

We consider in particular $G = \mathrm{Graph}^*(F_\sigma)$ and $G_+ = \mathrm{Graph}^{*,+}(F_\sigma)$ and $G_- = \mathrm{Graph}^{*,-}(F_\sigma)$ and call ζ, ζ_+ and ζ_- the corresponding measures. As a consequence of Proposition 3.2 and of the concerned definitions we can already state $\zeta = \zeta_+ + \zeta_-$ and $\mathrm{proj}^2_\# \zeta_+ = \mathrm{proj}^2_\# \zeta_- = i^*_{\mathbf{R}}(h)/2 \, \mathrm{d}h$. The next result indicates the other projections.

Proposition 3.3. *The measures ζ_+ and ζ_- defined as in the previous paragraph satisfy $\mathrm{proj}^1_\# \zeta_+ = \mu$ and $\mathrm{proj}^1_\# \zeta_- = \nu$.*

Proof. From (6) and (9), computing the half sum and the half difference we already have

$$(TV_+)^t_s(F_\sigma) = \int_{\mathbf{R}} i^{*,+}_{]s,t]}(h) \, \mathrm{d}h \quad \text{and} \quad (TV_-)^t_s(F_\sigma) = \int_{\mathbf{R}} i^{*,-}_{]s,t]}(h) \, \mathrm{d}h. \tag{10}$$

Therefore,

$$(TV_+)^t_s(F_\sigma) = \int_{\mathbf{R}} \mathrm{card}(x \in]s, t] : (x, h) \in \mathrm{Graph}^+(F^*_\sigma)) \, \mathrm{d}h \tag{11}$$

$$= \int_{-\infty}^{+\infty} \mathrm{card}(x \in \mathbf{R} : (x, h) \in (]s, t] \times \mathbf{R}) \cap \mathrm{Graph}^+(F^*_\sigma)) \, \mathrm{d}h \tag{12}$$

$$= \mathrm{proj}^1_\# \zeta_+(]s, t]) \tag{13}$$

and $(TV_-)^t_s(F_\sigma) = \mathrm{proj}^1_\# \zeta_-(]s, t])$. Recall moreover the definitions

$$\begin{cases} TV^t_s(F) = \sup_{m \in \mathbf{N}, \, s = r_0 < \ldots < r_{m+1} = t} \sum_{j=0}^{m} |F(r_{k+1}) - F(r_k)| \\[2ex] (TV_+)^t_s(F) = \sup_{m \in \mathbf{N}, \, s = r_0 < \ldots < r_{m+1} = t} \sum_{j=0}^{m} [F(r_{k+1}) - F(r_k)]^+ \\[2ex] (TV_-)^t_s(F) = \sup_{m \in \mathbf{N}, \, s = r_0 < \ldots < r_{m+1} = t} \sum_{j=0}^{m} [F(r_{k+1}) - F(r_k)]^- \end{cases} \tag{14}$$

where $[u]^+ = \frac{1}{2}(|u| + u)$ and $[u]^- = \frac{1}{2}(|u| - u)$ are the positive and negative parts of u. As $\mu \perp \nu$, there exists P with $\mu(P) = \mu(\mathbf{R})$ and $\nu(P) = 0$. By outer

regularity for every ε there exists an open set $U \supset P$ such that $|\mu(P) - \mu(\mathbf{R})| = 0$ and $\nu(U) < \varepsilon/2$.

Finally from the σ-additivity of measures and the fact $\mu(]a, b - 1/n]) \to \mu(]a, b[)$ for every $a < b$ there exists a finite union $J = \bigcup_i]a_i, b_i] \subset U$ of semi open intervals such that $\nu(J) < \varepsilon$ and $|\mu(J) - \mu(\mathbf{R})| < \varepsilon$. The partition of $\mathbf{R}$ given by J permits us to check that the positive total variation of F_σ is greater than $\sum [\sigma(]a_i, b_i])]^+$ which is greater than $\sum_i \sigma(]a_i, b_i]) = \mu(J) - \nu(J) \geq \mu(\mathbf{R}) - 2\varepsilon$. This holds for every $\varepsilon \geq 0$ so that $TV_+(F_\sigma) \geq TV(F_\mu) = \mu(\mathbf{R})$. Symmetrically $TV_+(F_\sigma) \geq TV(F_\nu) = \nu(\mathbf{R})$. Finally as $TV_+(F_\sigma) + TV_-(F_\sigma) = TV(F_\sigma) = TV(F_\mu - F_\nu) \leq TV(F_\mu) + TV(F_\nu) = (\mu + \nu)(\mathbf{R}) = |\sigma|(\mathbf{R})$ we conclude with $TV_+(F_\sigma) = TV(F_\mu)$ and $TV_-(F_\sigma) = TV(F_\nu)$. The same identity is correct on $]s, t]$ with the same proof (for instance $(TV_+)_s^t(F_\sigma) = (TV)_s^t(F_\mu) = \mu(]s, t]) = \mathrm{proj}^1_\# \zeta_+(]s, t])$. □

Proof of Theorem 1.1. We proved that $i_\mathbf{R}$ is almost surely an even integer and $\int_{-\infty}^{+\infty} i_\mathbf{R}(h) \, \mathrm{d}h = TV(F_\sigma) = TV(F_\mu) + TV(F_\nu) = 2$. Due to Proposition 3.2 (X, H) is concentrated on $\mathrm{Graph}^+(F_\sigma^*)$ and is distributed as ζ_+. Similarly (Y, H) is concentrated on $\mathrm{Graph}^-(F_\sigma^*)$ and distributed as ζ_-. Finally the law of X is μ and the law of Y is ν. □

3.2 Proof that a transport plan concentrated on a monotone set is the excursion coupling
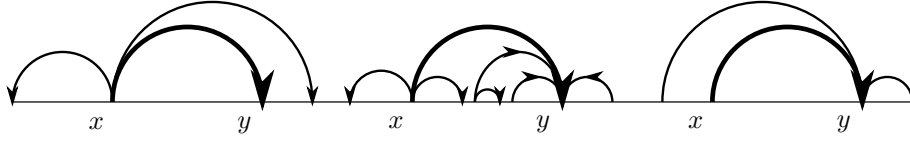
In this subsection we call monotone a set S with arches that do not cross, do not connect and have same orientation when they are nested. In what follows we prove *3 ⇒ 4* of Main Theorem, i.e. that measures concentrated on a monotone set are the excursion coupling of their marginals (this is correct even though these measures do not have finite first moment). Firstly, we prove this in the case $\mu \perp \nu$. Secondly, we prove the general case on page 22.

Proof of *3 ⇒ 4* for measures $\mu \perp \nu$ The generalized $\mathrm{Graph}^*(F_\sigma)$, π and the related objects are still defined as above and we still assume $\mu \perp \nu$. We define Γ (that depends on σ) as

$$\Gamma = \bigcup_{h > 0, h \in C} \bigcup_{k=1}^{i_\mathbf{R}^*(h)/2} \{(x_{2k-1}^h, x_{2k}^h)\} \cup \bigcup_{h < 0, h \in C} \bigcup_{k=1}^{i_\mathbf{R}^*(h)/2} \{(x_{2k}^h, x_{2k-1}^h)\} \quad (15)$$

where C is the set of levels h such that $i_\mathbf{R}^*(h) = i_\mathbf{R}^{*,\pm}(h)$ and $x_1^h < x_2^h < \cdots < x_{i_\mathbf{R}^*-1}^h < x_{i_\mathbf{R}^*}^h$ are the points of $\{x \in \mathbf{R} : (x, h) \in \mathrm{Graph}^*(F_\sigma)\}$. We stated in Proposition 3.2 that C has full measure with respect to $i_\mathbf{R}^*(\mathrm{d}h)$. Therefore, with respect to the definition of the excursion coupling, we have $\pi(\Gamma) = 1$.

Remark 3.4. We could prove that Γ is monotone (the arches of Γ do not cross, do not connect and have the same orientation when they are nested). Since $\pi(\Gamma) = 1$, this would correspond to the implication *4 ⇒ 3*. This is correct and can be proved directly but our proof of Main Theorem goes *4 ⇒ 1 ⇒ 2 ⇒ 3*.

Fig. 5: Transport routes inside and outside $]x, y[$ and at points x and y

Proposition 3.5. *Let μ and ν be mutually singular measures of $\mathcal{P}_1(\mathbf{R})$ and γ be a monotone transport plan in $\mathrm{Marg}(\mu, \nu)$. Let S be a monotone set with $\gamma(S) = 1$. Then $\Gamma \cap S$ is still monotone and satisfies $\gamma(\Gamma \cap S) = 1$.*

Proof. It is not a priori known that $\gamma(\Gamma) = 1$ and this statement is in fact clearly equivalent to the proposition result. Let S be as in the statement. We will define $S \supset S_1 \supset S_2 \supset S_3$ such that $S_k \setminus S_{k-1}$ has measure zero for γ and $S_3 \subset \Gamma$. Hence we will have $\gamma(\Gamma) \geq \gamma(\Gamma \cap S) \geq \gamma(S_3 \cap S) = \gamma(S) = 1$.

Let γ and S be as in the statement. Let Δ be the diagonal $\{(x, y) \in \mathbf{R}^2 : x = y\}$ and $A = \mathrm{proj}^1(\Gamma)$ and $B = \mathrm{proj}^2(\Gamma)$. Let S_1 be $S \cap (\mathbf{R}^2 \setminus \Delta)$ and $S_2 = S_1 \cap (A \times \mathbf{R}) \cap (\mathbf{R} \times B)$. Finally $S_3 = S_2 \setminus \{(x, y) \in \mathbf{R}^2 : \gamma(x, y) = 0, \mu(x) > 0, \nu(y) > 0\}$.

Proof of $\gamma(S_3) = \gamma(S)$. We already recalled in Lemma 2.4 that transport plans $\gamma \in \mathrm{Marg}(\mu, \nu)$, whose arches do not connect takes the form $\gamma = (\mathrm{id} \times \mathrm{id})_{\#}(\mu \wedge \nu) + \gamma_0$ where $\gamma_0(\Delta) = 0$. As $\mu \perp \nu$ we have $\gamma(\Delta) = 0$ and $\gamma(S_1) = \gamma(S)$. We have $\gamma(A \times \mathbf{R}) = \mu(A) = \pi(\Gamma) = 1$. Similarly, $\gamma(\mathbf{R} \times B) = 1$ so that $\gamma(S_2) = \gamma(S_1) = 1$. Finally there are countably many pairs (x, y) with $\mu(x) > 0$ and $\nu(y) > 0$. Thus $\gamma(\{(x, y) \in \mathbf{R}^2 : \gamma(x, y) = 0, \mu(x) > 0, \nu(y) > 0\}) = 0$ and $\gamma(S_3) = \gamma(S_2) = 1$.

Proof of $S_3 \subset \Gamma$. Let (x, y) be in S_3. Without loss of generality we can assume $x < y$ ($x = y$ became impossible as S was replaced by S_1). Let us first prove

$$\gamma((\mathbf{R} \setminus [x, y]) \times]x, y[) = 0 \qquad \text{and} \qquad \gamma(]x, y[\times (\mathbf{R} \setminus [x, y])) = 0. \qquad (16)$$

If not there exists $(x', y') \in \mathrm{Spt}(\gamma)$ with $x' \notin [x, y]$ and $y' \in]x, y[$ (or the same property inverting the role of x' and y'). As S is dense in $\mathrm{Spt}(\gamma)$ the same is correct for a point $(x'', y'') \in S$, which leads to a contradiction with the monotonicity of S, whose arches should not cross.

Case 1 in the proof of $S_3 \subset \Gamma$: Assume $\mu(x) = \nu(y) = 0$; the complementary case is considered further in case 2. Then $F_\sigma(y) - F_\sigma(x) = \mu([x, y]) - \nu([x, y]) = \gamma([x, y] \times \mathbf{R}) - \gamma(\mathbf{R} \times [x, y]) = \gamma([x, y]^2) - \gamma([x, y]^2) = 0$. Moreover, a similar argument as for (16) shows that for any $y' \in]x, y]$ we have

$$\nu(]x, y']) = \gamma([x, y'] \times]x, y']) \leq \mu(]x, y']). \qquad (17)$$

In fact $\gamma((\mathbf{R} \setminus [x, y']) \times]x, y']) \neq 0$ would imply that there exists some $(x'', y'') \in S$ with $y'' \in]x, y'] \subset]x, y[$ and $x'' \notin [x, y']$. If $x'' < x$ this contradicts the fact that

arches (x, y) and (x'', y'') do not cross. If $x'' > y'$ the latter fact or the one that nested arches have the same orientation is violated. From (17) we find $F_\sigma(y') \geq F_\sigma(x) = F_\sigma(y)$ for every $y' \in]x, y[$. As $\mu(x) = 0$ and $\nu(x) = 0$ ($\nu(x) > 0$ is not possible because x is in A so that it can not be an abscise of the decreasing part of $\mathrm{Graph}^-(F_\sigma)$) the multivalued function F_σ^* is single valued at x. Hence, from the definition of A involved in S_2 it follows that $h = F_\sigma(x)$ must be an element of C. Therefore, since the level h cuts $\mathrm{Graph}(F_\sigma^*)$ in points of $\mathrm{Graph}^+(F_\sigma^*)$ or $\mathrm{Graph}^-(F_\sigma^*)$ it is not possible to have $F_\sigma(x) = F_\sigma(y') = F_\sigma(y)$ for $y' \in]x, y[$ (In a neighborhood of y' we would have $F_\sigma \geq h$). It follows that x and y are consecutive zeros of $F_\sigma^* = h$. Thus $(x, y) \in \Gamma$.

Case 2 in the proof of $S_3 \subset \Gamma$: We want to finalize the inclusion $S_3 \subset \Gamma$ looking at the pairs $(x, y) \in S_3 \subset S$ where x or y is an atom of σ. Since the arches of S do not cross (x, y) at least one of the two is true: *i)* $] -\infty, x[\times \{y\}$ has empty intersection with S), or *ii)* $\{x\} \times]y, \infty[$ has empty intersection with S (recall that here $x < y$), i.e the arches of the left and right patterns of Figure 5 can not all be in S. Without loss of generality we will assume that $] -\infty, x[\times \{y\}$ is empty. This corresponds to the two first patterns on Figure 5. Adapting the argument of case 1 we find $F_\sigma(y) - F_\sigma(x) = \mu(]x, y]) - \nu(]x, y]) \leq 0$ and $F_\sigma(y) - F_\sigma(x^-) = \mu([x, y]) - \nu([x, y]) \geq 0$ in place of $F_\sigma(y) - F_\sigma(x) = 0$. Thus $F_\sigma(y)$ is in $[F_\sigma(x^-), F_\sigma(x)]$. Let x' be in $]x, y]$. We have $F_\sigma(y) - F_\sigma(x') = \mu(]x', y]) - \nu(]x', y]) \leq 0$. This is due to the fact that the arches starting from $]x', y]$ have the same orientation as (x, y) and do not cross it. Moreover, for ν some additional mass in y could arrive from $]y, +\infty[$. We want to prove that $F_\sigma(y) - F_\sigma(x') < 0$ is also true for any $x' \in]x, y[$. If y is not an atom we can proceed as before starting with $F_\sigma(y) \in C$ and the arch (x, y) with $F_\sigma^*(x) \ni F_\sigma(y)$. Therefore, we assume that y is an atom of ν. If $F_\sigma(y) \in]F_\sigma(x^-), F_\sigma(x)[$ we will be able to conclude easily that there exists $h \in C \cap]F_\sigma(x^-), F_\sigma(x)[\cap]F_\sigma(y^-), F_\sigma(y)[$ and we conclude as we did twice before (on $[x, y]$ the generalized function F_σ^* can not only touch the level h but it must cut it, which is not possible because $(x, y) \in \Gamma$ for the level h). In the other case $F_\sigma(x^-) < F_\sigma(x) = F_\sigma(y)$ and all the mass arriving in y comes from $]x, y]$. Therefore, (x, y) may be an element of S_2 but not of S_3, a contradiction. This case can not happen and we proved $(x, y) \in \Gamma$ in all the other cases. Finally we proved $S_3 \subset \Gamma$.

$\square$

Proposition 3.6. *Let μ and ν be singular measures and $\pi \in \mathrm{Marg}(\mu, \nu)$ the excursion coupling and Γ as defined in (15). Let $\pi' \in \mathrm{Marg}(\mu, \nu)$ be another coupling concentrated on Γ. Then $\pi' = \pi$.*

Proof. Let A be the set of atomic points of μ and B the set of atomic points of ν. From the definition of Γ we see that the set $\Gamma \setminus (A \times \mathbf{R})$ is contained in the graph of a function from $\mathbf{R} \setminus A$ to $\mathbf{R}$. The same is true, inverting the coordinates for $\Gamma \setminus (\mathbf{R} \times B)$. Hence the measures π and π' coincide on the set $\mathbf{R}^2 \setminus (A \times B) = [(\mathbf{R} \setminus A) \times \mathbf{R}] \cup [\mathbf{R} \times (\mathbf{R} \setminus B)]$ that we denote by E, i.e

$$\pi(\cdot \cap E) = \pi'(\cdot \cap E).$$

Therefore, we aim at proving that π and π' coincide on the countable set $E^c = A \times B$. We will in fact prove $\pi\{(a,b)\} = \pi'\{(a,b)\}$ for every $(a,b) \in \Gamma$. Let (a,b) be in Γ. Let us assume without loss of generality $a < b$. We further assume $F_\sigma(a^-) \geq F_\sigma(b)$ so that according to the definition of Γ the transport plans π and π' are concentrated on it and the mass of μ on $[a,b[$ must be transported onto $]a,b]$. This writes $\mu([a,b[) = \pi([a,b[\times]a,b])$ and we have also $\mu([a,b[) = \pi'([a,b[\times]a,b])$. Reasoning similarly we obtain $\nu(]a,b[) = \pi([a,b[\times]a,b[)$. Therefore,

$$\pi([a,b[\times\{b\}) = \mu([a,b[) - \nu(]a,b[). \tag{18}$$

The construction of Γ associates the route $(a,b) \in \Gamma$ with $a < b$ to some level $h > 0$. The generalized intermediate value theorem, Lemma 3.1, permits us to derive $F_\sigma(a') > h$ for every $a' \in]a,b[$ so that $F_\sigma(a'^-) \geq F_\sigma(b)$ also holds. Therefore, (18) can be written for any a' such that $(a',b) \in \Gamma$ in place of a. Recall that it also holds for π' in place of π. We will be done if we can prove $\pi(]a,b[\times\{b\}) = \pi'(]a,b[\times\{b\})$. This is in fact correct because

$$\begin{aligned}
\pi(]a,b[\times\{b\}) &= \pi((]a,b[\times\{b\}) \cap \Gamma) \\
&= \sup\{\pi(([a',b[\times\{b\}) \cap \Gamma)) : a' \in \mathbf{R}, (a',b) \in \Gamma, a < a' < b\} \\
&= \sup\{\pi'(([a',b[\times\{b\}) \cap \Gamma)) : a' \in \mathbf{R}, (a',b) \in \Gamma, a < a' < b\} \\
&= \pi'((]a,b[\times\{b\}) \cap \Gamma) = \pi'(]a,b[\times\{b\}).
\end{aligned}$$

$\square$

Proof of $3 \Rightarrow 4$ for general measures μ in ν   We no longer assume $\mu \perp \nu$. In this case we know from Lemma 11 that any γ satisfying 3 in Main Theorem can be written in the form $(\mathrm{id}\times\mathrm{id})_\#(\mu\wedge\nu) + \gamma_0$ where $\gamma_0 \in \mathrm{Marg}(\mu-(\mu\wedge\nu), \nu-(\mu\wedge\nu))$ satisfies $\gamma_0(\Delta) = 0$. We are in the situation of the last paragraph because $\mu - (\mu \wedge \nu) \perp \nu - (\mu \wedge \nu)$ and γ_0 satisfies 3: since we know that there is a monotone set S with $\gamma(\mathbf{R}^2 \setminus S) = 0$, equality $\gamma_0(\mathbf{R}^2 \setminus S) = 0$ also holds. From the discussion above we obtain that γ_0 is the excursion coupling of $\mu - (\mu \wedge \nu)$ and $\nu - (\mu \wedge \nu)$. This exactly implies that γ is the excursion coupling of μ and ν.

4   Final elements of proof of Main Theorem and its corollary

Proof of Main Theorem. The structure of the proof is the following:

- The set of measures satisfying 1 (the solutions to the L^{1^-} problem) is not empty.

- The set of measures satisfying 2 (the solutions to the $L^{1,q}$ problem) is not empty.

- Assumption 1 implies 3 and assumption 2 implies 3 (see Section 2).

- Assumption *3* implies *4* (see Section 3).

- There is a unique and well-defined coupling satisfying *4* (see Theorem 1.1).

Therefore, if π satisfies *4* it equals any coupling satisfying *1*, respectively *2*. As these sets are not empty, if π satisfies *4* it also satisfies *1*, respectively *2*.

We proved everything except the two first existence statements. They will be obtained as consequences of Lemma 4.1 that is proved in this section. $\square$

Il order to prove that there exists a solution to the L^{1^-} limit transport problem (property *1*) it suffices to remind of two elementary facts. First, any sequence $(\pi_n)_{n \in \mathbb{N}}$ in $\mathrm{Marg}(\mu, \nu)$ admits cluster points. The set $\mathrm{Marg}(\mu, \nu)$ is indeed a compact set for the weak topology, as a simple consequence of Prokhorov Theorem and of the fact that it is closed. Second, as we recalled in the introduction, the set $\mathrm{Marg}_p^*(\mu, \nu)$ is not empty for every $p \in]0, 1[$. These two facts permit us to conclude that there exists at least one element $\pi \in \mathrm{Marg}(\mu, \nu)$ that satisfies *1*. However, the non emptiness of $\mathrm{Marg}_p^*(\mu, \nu)$ relied on the following Lemma 4.1 through the standard argument of optimization. We prove it now for the sake of completeness and in order to prepare the proof of Lemma 4.3.

Lemma 4.1. *Let p be in $]0, 1[$ and μ, ν be two mesures with finite moment of order p. The function*

$$C_p : \pi \in \mathrm{Marg}(\mu, \nu) \to \iint |y - x|^p \, \mathrm{d}\pi(x, y)$$

is continuous on $\mathrm{Marg}(\mu, \nu)$, endowed with the weak topology of laws on $\mathbf{R}^2$.

Proof. Let p, μ and ν be as in the statement. For every $n \in \mathbf{N}^*$ we decompose $(x, y) \to |y - x|^p$ as the sum of a bounded continuous function and a reminder function:

$$|y - x|^p = \underbrace{(|y - x|^p \wedge n)}_{:= f_{n,p}(x,y)} + \underbrace{(|y - x|^p - n)^+}_{:= h_{n,p}(x,y)}$$

where $(u)^+ = \frac{1}{2}(u + |u|)$ denotes the positive part of u. Since $n \geq 1$ we have

$$h_n(x, y)) \leq (|x|^p - n/2)^+ + (|y|^p - n/2)^+ \leq |x|^p \mathbb{1}_{|x| \geq n/2} + |y|^p \mathbb{1}_{|y| \geq n/2}$$

so that $\iint h_n(x, y) \mathrm{d}\pi(x, y) \leq \int_{|x|^p \geq n/2} |x|^p \mathrm{d}\mu(x) + \int_{|y|^p \geq n/2} |y|^p \mathrm{d}\nu(y)$ for every $\pi \in \mathrm{Marg}(\mu, \nu)$. Thus $s_{n,p} := \sup\{\iint h_n \mathrm{d}\pi : \pi \in \mathrm{Marg}(\mu, \nu)\} \to_{n \to \infty} 0$. Let us finish proving that C_p is continuous at π. Let ε be a positive real number and n large enough to make $s_{n,p}$ smaller that $\varepsilon/4$. Now

$$|C_p(\pi) - C_p(\pi')| \leq \left| \iint f_{n,p} \, \mathrm{d}(\pi - \pi') \right| + \left| \iint h_{n,p} \, \mathrm{d}(\pi - \pi') \right|$$

$$\leq \left| \iint f_{n,p} \, \mathrm{d}(\pi - \pi') \right| + 2 s_{n,p}.$$

As f_n is continuous and bounded this estimate proves that there exists a neighborhood U of π such that for every $\pi' \in U$ it holds $|C_p(\pi') - C_p(\pi)| \leq \varepsilon$. $\square$

Remark 4.2. Until now, even though it is clear from the context and a consequence of Main Theorem (implication $2 \Rightarrow 1$) we never proved that a so-called solution to the L^{1^-} limit transport problem is a solution to the L^1 transport problem. The next result may be invoked to prove it directly.

Lemma 4.3. *The function* $(p, \pi) \in]0, 1] \times \mathrm{Marg}(\mu, \nu) \to F_p(\pi)$ *is continuous.*

Proof. The proof is similar to the one of Lemma 4.1 from which we keep the notation $s_{n,1}$. Let us fix (p, π). For another pair (p', π') we have

$$|C_p(\pi) - C_{p'}(\pi')| \leq |C_p(\pi) - C_{p'}(\pi)| + |C_{p'}(\pi) - C_{p'}(\pi')|$$

$$\leq 2s_{n,1} + |C_p(\pi) - C_{p'}(\pi)| + \left| \iint [(1 + |y - x|) \wedge n] \, \mathrm{d}(\pi - \pi') \right|.$$

Let $\varepsilon > 0$. Due to the dominated convergence theorem the term $|C_p(\pi) - C_{p'}(\pi)|$ is smaller than $\varepsilon/4$ for $p' \in]0, 1]$ in a neighborhood of p. For n large enough $s_{n,p}$ is smaller that $\varepsilon/4$. For this n fixed we see that the last term is smaller than $\varepsilon/4$ when π' is in a certain neighborhood of π. □

Let us now prove that the set of solutions to the $L^{1,q}$ secondary transport problem is not empty. With what precedes including Lemma 4.1 we indeed know that $\mathrm{Marg}_1^*(\mu, \nu)$ is not empty and that it is closed. Let $(\pi_n)_{n \in \mathbf{N}}$ be a minimizing sequence for $\pi \mapsto \iint |y - x|^q \mathrm{d}\pi(x, y)$ on $\mathrm{Marg}_1^*(\mu, \nu)$. This is a continuous function so that $\mathrm{Marg}_{1,q}^{**}$ is not empty.

We finish the section with the proof that seemingly stronger results are equivalent to *1* and *2* in Main Theorem.

Proof of Corollary 0.2. Main Theorem $4 \Rightarrow 1$ applies to any subsequence $(p_{\varphi(n)})_{n \in \mathbf{N}}$ that is increasing. Therefore, in the compact set $\mathrm{Marg}(\mu, \nu)$ any subsequence of $(\pi_n)_{n \in \mathbf{N}}$ possesses an increasing subsequence converging to π. This proves *1'*.

Let q' be an exponent smaller than 1. Then Main Theorem $4 \Rightarrow 2$ applies to $q' = q$. □

5   Concluding bibliographic remarks and perspectives

Remark 5.1 (the L^0 transport problem). For $p = 0$ the cost function is $d^0(x, y) = \mathbb{1}_{\{x \neq y\}}$ where d is the distance on $\mathbf{R}$. This corresponds to the coupling problem that defines the total variation of μ and ν. A coupling π is a solution if and only if it writes $(\mathrm{id} \times \mathrm{id})_\# (\mu \wedge \nu) + \pi_0$ and $\min_{\pi \in \mathrm{Marg}(\mu, \nu)} C_0(\pi) = \max |\mu(A) - \nu(A)| = (1/2)|\mu - \nu|(\mathbf{R})$.

Remark 5.2 (the L^p transport problem for $p < 0$). For $p < 0$ the cost function d^p where d is the distance on $\mathbf{R}$ is singular on the diagonal $\Delta = \{(x, y) \in \mathbf{R}^2 : x = y\}$ where it takes the value $+\infty$. Notice that it writes $\ell(|y - x|)$ where $\ell : [0, \infty) \to \mathbf{R}^+ \cup \{\infty\}$ is decreasing, convex and has limits ∞ and 0. This type of cost, including the Coulomb cost $c : (x, y) \mapsto |y - x|^{-1}$ has been thoroughly studied by Cotar, Friesecke and Klüppelberg in [CFK13] with the purpose of

determining the joint distribution of electronic particles on their orbitals. Their Theorem 3.1 states an existence and uniqueness result for measures admitting a density. In §4.1 they conduct a precise study of the one dimensional case in the spirit of [GM96, McC01], the same geometric spirit that is also inspiring us in the present paper. Concerning the Coulomb type costs, notice that the assumption $\mu = \nu$ is the natural one for the chemical application. In their Theorem 4.8 the authors completely characterize the optimal transport for an absolutely continuous measures $\mu = \nu$ with positive density. As for $p > 0$ this solution does not depend on the particular value of $p < 0$ (or of l).

However, there is no uniqueness of the optimal transport plan in general, as can be seen for instance with the example $\mu = \nu = \frac{1}{3}(\delta_{-1} + \delta_0 + \delta_1)$. Moreover the example $\mu = (1/2)(\delta_0 + \delta_{1+\Phi})$, $\nu = (1/2)(\delta_{-\Phi} + \delta_1)$ with $\Phi = (1 + \sqrt{5})/2$ is similar to Example 2.2: For $p < -1$ the problem admits a unique solution π^0, for $p > -1$ another transport plan π^1 is the unique solution and for $p = -1$ the solutions are the plans $(\pi^\lambda)_{\lambda \in [0,1]}$ defined by $\pi^\lambda = \lambda \pi^1 + (1 - \lambda)\pi^0$.

Remark 5.3 (On the solution of the Monge problem selected in [DML18]). It is clear that if there exists $a \in \mathbf{R}$ such that μ and ν are concentrated in $[a, +\infty[$ and $]-\infty, a]$ respectively, then $\mathrm{Marg}_1^*(\mu, \nu)$ coincide with the set of all transport plans between μ and ν. In fact for measures with finite first moment, identity $\mathrm{Marg}_1^*(\mu, \nu) = \mathrm{Marg}(\mu, \nu)$ holds if and only if there exists such a real $a \in \mathbf{R}$ splitting the supports of μ and ν (but the symmetric situation $\mathrm{Spt}(\mu) \in [a, +\infty[$ is possible). For general measures $\mu, \nu \in \mathcal{P}_1(\mathbf{R})$ the set $\mathrm{Marg}_1^*(\mu, \nu)$ has recently be described by Di Marino and Louet in [DML18]. This recent paper concerns another way to select a special element of $\mathrm{Marg}_1^*(\mu, \nu)$ when the entropy parameter in the entropic regularized Monge problem tends to zero. Note that the resulting coupling is different from ours. If the measure μ, ν are as in the beginning of this remark, the plan of Di Marino and Louet is $\mu \times \nu$. We obtain the decreasing rearrangement, i.e the law of $(F_\mu, 1 - F_\nu)$ seen as a random vector on $([0, 1], \lambda)$.

Remark 5.4 (On the Skorkhod problem for unbiased Brownian motions). One motivation to our paper was to better understand a work by Last, Mörters and Thorisson [LMT14] and reformulate their construction in the framework of the optimal transport theory. In this paper eternal Brownian motions $(B_t)_{t \in \mathbf{R}}$ starting in μ are embedded onto ν with non-negative random time T such that $(B_{t-T})_{t \in \mathbf{R}}$ is an eternal Brownian motion independent of T. The authors define a coupling similar to our excursion coupling but for two σ-finite random measures on $\mathbf{R}$. This solution minimizes $\mathbb{E}(\varphi(T))$ for any concave function $\varphi : \mathbf{R}^+ \to \mathbf{R}^+$. Comparing with our case that concerns deterministic measures $\mu, \nu \in \mathcal{P}_1(\mathbf{R})$ one can conjecture that among couplings (X, Y) with law in $\mathrm{Marg}(\mu, \nu)$ that satisfy the constraint $X \leq Y$, the excursion coupling is the one minimizing $\mathbb{E}(Y - X)^p$ for every $p \in]0, 1[$.

Remark 5.5 (Discrete Setting). When μ and ν are the sum of finitely many distinct atoms of same mass, the excursion coupling reduces to the *Dyck matching* described in the physics paper [CDES20]. Here points are matched in the same way as one associates opening and closing parentheses in a formal word

of parentheses. In this paper the authors consider the asymptotic transport cost for increasingly many points picked at random. A connection to the same problem for power costs $p < 1$ (as in [DSS12]) is evoked.

Remark 5.6. Bertoin and Yor [BY14] established their deterministic formulae in relation with an important chapter of Stochastic Calculus. The occupation measure of (the continuous part of) a real semimartingale turns out to be a random absolutely countinuous measure with density expressed in terms of local times, that are quantities described by the Meyer–Tanaka formula. The work of Bertoin and Yor provides analogue results in the deterministic word of functions with finite variation.

Remark 5.7 (Sharpness of the assumptions). The assumptions in Main Theorem are by no mean claimed to be sharp. For property *1*, a rough analysis of the proof seems to indicate that the family of costs $(c_\varepsilon)_{\varepsilon \in]0,1[}$ defined by $c_\varepsilon(x,y) = |y-x|^{1-\varepsilon}$ can be replaced by any family $(c'_\varepsilon)_{\varepsilon \in]0,1[}$ of type $c'_\varepsilon(x,y) = \ell_\varepsilon(|y-x|)$ where it is assumed $\ell_\varepsilon(d) \to_{\varepsilon \to 0^+} d$ for every $d \geq 0$ and ℓ_ε is increasing and strictly concave. Concerning property *2*, any c' of the same type as before should play the same role as $|y-x|^q$. Finally, the fact that μ and ν have a finite first moment should not be necessary to state the equivalence between *3* and *4* (see also Remark 3.4).

Remark 5.8 (L^{1-} limit transport problem in Euclidean spaces). The transport problem for the Monge distance cost in Euclidean and further in some more general geodesic spaces is a research stream with a rich history. It recently culminated with the optimal transport proof by Cavalletti and Mondino of the Lévy-Gromov inequality [CM17]. It is intimately connected to the Monge problem on the real line because under appropriate assumptions on the space and the marginal measures, any optimal transport plan can be disintegrated as a mixture of one dimensional transport plans concentrated on disjoint geodesic rays.

   A natural question with respect to the present paper is the existence and uniqueness of a solution to the L^{1^-} limit problem in Euclidean spaces. One can moreover conjecture that such a cluster solution can be disintegrated in a way that the transport along the geodesic rays always is the excursion coupling. A similar result has been proved in [AP03] for the L^{1^+} limit problem and the quantile coupling.

Remark 5.9 (Generating π by picking a random point on a tree). A popular construction in probability is to associate a random tree with a random function (or process) f. Typically in the construction of the continuum random Brownian tree [Ald93] a random tree is associated to a random excursion. However this topological construction is purely deterministic. Let f be function defined on an interval I. We write $x \sim x'$ if and only if $f(x) = f(x')$ and $y \in]x, x'[\Rightarrow f(y) \geq f(x)$. The tree is the quotient space $I/\sim$. The fact that we conducted a similar operation with our function $F_\mu - F_\nu$ yields an appealing interpretation. In place of choosing a random point h on the y-axis with density i^* and continue selecting uniformly an excursion among $i^*(h) \in \mathbf{N}$ possible ones we could directly choose randomly a point on the associated tree according to the length measure, i.e the

Hausdorff measure of dimension 1. Doing this we come closer to the classical simulation of the quantile coupling where a point h is chosen on the tree $[0, 1]$ (a segment) according to the length measure (the Lebesgue measure). While it is clear that the measure on the tree is the correct one if μ and ν are simple, e.g finite sums of atoms, or such that $F = F_\mu - F_\nu$ is of class $\mathcal{C}^1$ with finitely many changes of monotonicity, it is not is the general case. We leave it as a conjecture that the quotient measure on the tree corresponding to our generalized function $F^* = (F_\mu - F_\nu)^*$ always is the length measure.

References

[Ald93] D. Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.

[AP03] L. Ambrosio and A. Pratelli. Existence and stability results in the L^1 theory of optimal transportation. In *Optimal transportation and applications (Martina Franca, 2001)*, volume 1813 of *Lecture Notes in Math.*, pages 123–160. Springer, Berlin, 2003.

[Ban25] S. Banach. Sur les lignes rectifiables et les surfaces dont l'aire est finie. *Fundam. Math.*, 7:225–236, 1925.

[BY14] J. Bertoin and M. Yor. Local times for functions with finite variation: two versions of Stieltjes change-of-variables formula. *Bull. Lond. Math. Soc.*, 46(3):553–560, 2014.

[CDES20] S. Caracciolo, M. P. D'Achille, V. Erba, and A. Sportiello. The Dyck bound in the concave 1-dimensional random assignment model. *J. Phys. A*, 53(6):064001, 25, 2020.

[CFK13] C. Cotar, G. Friesecke, and C. Klüppelberg. Density functional theory and optimal transportation with Coulomb cost. *Communications on Pure and Applied Mathematics*, 66(4):548–599, 2013.

[CM17] F. Cavalletti and A. Mondino. Sharp and rigid isoperimetric inequalities in metric-measure spaces with lower Ricci curvature bounds. *Invent. Math.*, 208(3):803–849, 2017.

[DML18] S. Di Marino and J. Louet. The entropic regularization of the Monge problem on the real line. *SIAM J. Math. Anal.*, 50(4):3451–3477, 2018.

[DSS12] J. Delon, J. Salomon, and A. Sobolevski. Local matching indicators for transport problems with concave costs. *SIAM J. Discrete Math.*, 26(2):801–827, 2012.

[GM96] W. Gangbo and R. J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.

[LMT14] G. Last, P. Mörters, and H. Thorisson. Unbiased shifts of Brownian motion. *The Annals of Probability*, 42(2):431–463, 2014.

[McC01] R. J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.

[Mon81] G. Monge. Mémoire sur la théorie des déblais et des remblais. *Histoire de l'académie Royale des Sciences de Paris*, 1781.

[PPS15] P. Pegon, D. Piazzoli, and F. Santambrogio. Full characterization of optimal transport plans for concave costs. *Discrete Contin. Dyn. Syst.*, 35(12):6113–6132, 2015.

[Vil03] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.