

# Greedy matching in optimal transport with concave cost

ANDREA OTTOLINI and STEFAN STEINERBERGER

**Abstract.** We consider the optimal transport problem between a set of  $n$  red points and a set of  $n$  blue points subject to a concave cost function such as  $c(x, y) = \|x - y\|^p$  for  $0 < p < 1$ . Our focus is on a particularly simple matching algorithm: match the closest red and blue point, remove them both and repeat. We prove that it provides good results in any metric space  $(X, d)$  when the cost function is  $c(x, y) = d(x, y)^p$  with  $0 < p < 1/2$ . Empirically, the algorithm produces results that are *remarkably* close to optimal—especially as the cost function gets more concave; this suggests that greedy matching may be a good toy model for optimal transport for very concave transport cost.

## Ahne paritus parhaana massansiirtona koveran kustannuksen tapauksessa

**Tiivistelmä.** Tässä työssä tarkastellaan parasta massansiirtoa yhtä monen punaisen ja sinisen pisteen välillä, kun kustannusfunktio on kovera, kuten  $c(x, y) = \|x - y\|^p$ , missä  $0 < p < 1$ . Työssä keskitytään erityisen yksinkertaiseen algoritmiin: paritetaan lähin punainen ja sininen piste, poistetaan molemmat ja toistetaan. Osoittautuu, että tämä antaa hyvän tuloksen missä tahansa metrisessä avaruudessa  $(X, d)$ , kun kustannusfunktio on  $c(x, y) = d(x, y)^p$  ja  $0 < p < 1/2$ . Kokeellisesti tämä menetelmä tuottaa tuloksia, jotka ovat *huomattavan* lähellä parasta mahdollista – erityisesti kustannusfunktion muuttuessa entistä koverammaksi. Tämä viittaa siihen, että ahne paritus voi tarjota hyvän pelkistetyn mallin parhaalle massansiirrolle erittäin koveran siirtokustannuksen tapauksessa.

## 1. Introduction

**1.1. The problem.** The original motivation behind this paper is to understand the geometry of optimal transport with concave cost. Perhaps the easiest instance of this problem is the following: let  $X = \{x_1, \dots, x_n\}$  and  $Y = \{y_1, \dots, y_n\}$  be two sets of real numbers, what can be said about the optimal transport cost

$$W_p^p(X, Y) = \min_{\pi \in S_n} \sum_{i=1}^n |x_i - y_{\pi(i)}|^p,$$

where  $\pi: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$  ranges over all permutations? The answer is trivial when  $p \geq 1$ : order both sets in increasing order and send the  $i$ -th largest element  $x_i$  to the  $i$ -th largest element  $y_i$ . After ordering the points, the optimal permutation is the identity permutation  $\pi(i) = i$ . The problem becomes highly nontrivial when the cost function is concave.

As was already pointed out by Gangbo and McCann [16]: “For concave functions of the distance, the picture which emerges is rather different. Here the optimal maps will not be smooth, but display an intricate structure which—for us—was unexpected;

---

<https://doi.org/10.54330/afm.173113>

2020 Mathematics Subject Classification: Primary 82B44, 90B80.

Key words: Optimal transport, Euclidean random assignment problems.

© 2025 The Finnish Mathematical Society

it seems equally fascinating from the mathematical and the economic point of view. [...] To describe one effect in economic terms: the concavity of the cost function favors a long trip and a short trip over two trips of average length [...] it can be efficient for two trucks carrying the same commodity to pass each other traveling opposite directions on the highway: one truck must be a local supplier, the other on a longer haul.”

The problem has received increased attention in recent years, we refer to results of Bobkov and Ledoux [4, 5], Boerma, Tsyvinski, Wang and Zhang [6], Caracciolo, D’Achille, Erba and Sportiello [8], Caracciolo, Erba and Sportiello [9, 10], Delon, Salomon and Sobolevski [12, 13], Juillet [19] and McCann [20].

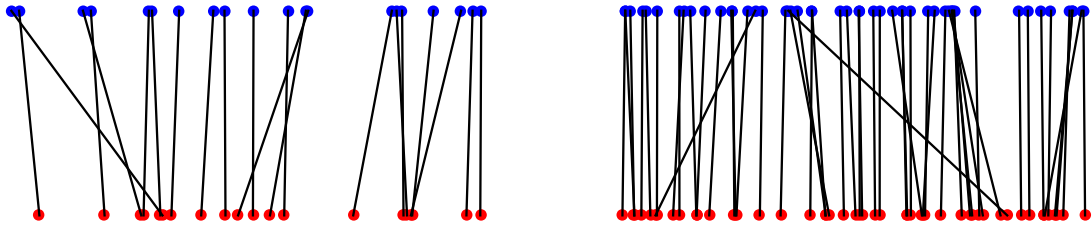


Figure 1. 20 (right: 50) red points on  $\mathbb{R}$  being optimally matched to 20 (right: 50) blue points on  $\mathbb{R}$  (shown displaced to illustrate the matching) and subject to cost  $c(x, y) = |x - y|^{1/2}$ .

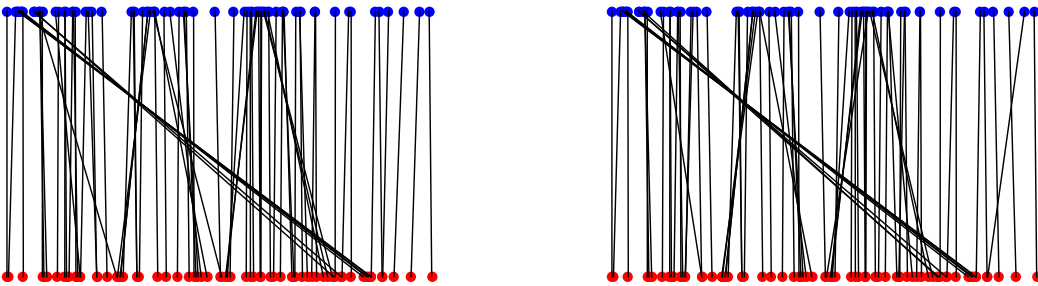


Figure 2. Left: 50 red points on  $\mathbb{R}$  being matched to 50 blue points on  $\mathbb{R}$  when  $c(x, y) = \log|x - y|$ . Right: same points and same cost function matched with the greedy algorithm. The two matchings are very similar: why?

A reason why the problem is interesting is illustrated in Figure 2: as suggested by Gangbo–McCann, there is a very curious dichotomy where most points get matched to points that are very close with a few exceptional points being transported a great distance. It is somewhat clear, in a qualitative sense, that this is to be expected (considering, for example, the Jensen inequality for concave functions). However, on a more quantitative level, the non-locality poses considerable difficulties.

**1.2. Dyck and greedy matching.** If the cost function is given by  $c(x, y) = h(|x - y|)$  with  $h$  concave, it appears there exist two natural toy models that are effective in different regimes: the Dyck matching and the greedy matching.



Figure 3. A collection of  $n = 5$  red and blue points, the function  $g(x)$  (left, rescaled for clarity) and the Dyck matching (right).

The first such model is the *Dyck matching* of Caracciolo–D’Achille–Erba–Sportiello [8]: their idea is to introduce  $g: [0, 1] \rightarrow \mathbb{Z}$

$$g(x) = \# \{1 \leq i \leq n: x_i \leq x\} - \# \{1 \leq i \leq n: y_i \leq x\}$$

The function is increasing whenever  $x$  crosses a new element of  $X$  while it decreases every time it crosses an element of  $Y$ . The Dyck matching is then obtained by matching across level sets of the function  $g$  (see Fig. 3). The Dyck matching is independent of the cost function. It is shown numerically in [8] (and reproduced in §2.6) that the Dyck matching produces a nearly optimal matching whose costs exceeds the optimal cost by very little. The second toy model is given by a simple greedy matching which works in general metric spaces.

### Greedy matching.

- (1) Determine

$$m = \min_{1 \leq i, j \leq n} c(x_i, y_j).$$

- (2) Find a pair  $(x_i, y_j)$  with  $c(x_i, y_j) = m$  and set  $\pi(i) = j$ .
- (3) Remove  $x_i$  from  $X$  and  $y_j$  from  $Y$  and repeat.

If the cost function is strictly monotonically increasing in the distance  $c(x, y) = h(|x - y|)$ , this greedy matching is, like the Dyck matching, independent of the cost function. This algorithm leads to mediocre results when the cost function is convex. This was already observed in the PhD thesis of D’Achille [11, Section 1.4] who explicitly considers the algorithm when  $c(x, y) = |x - y|^p$  and  $p \geq 1$  and shows that the results are not particularly good. One of the main points of our paper is to point out that the greedy matching is *very* good for very concave cost functions.

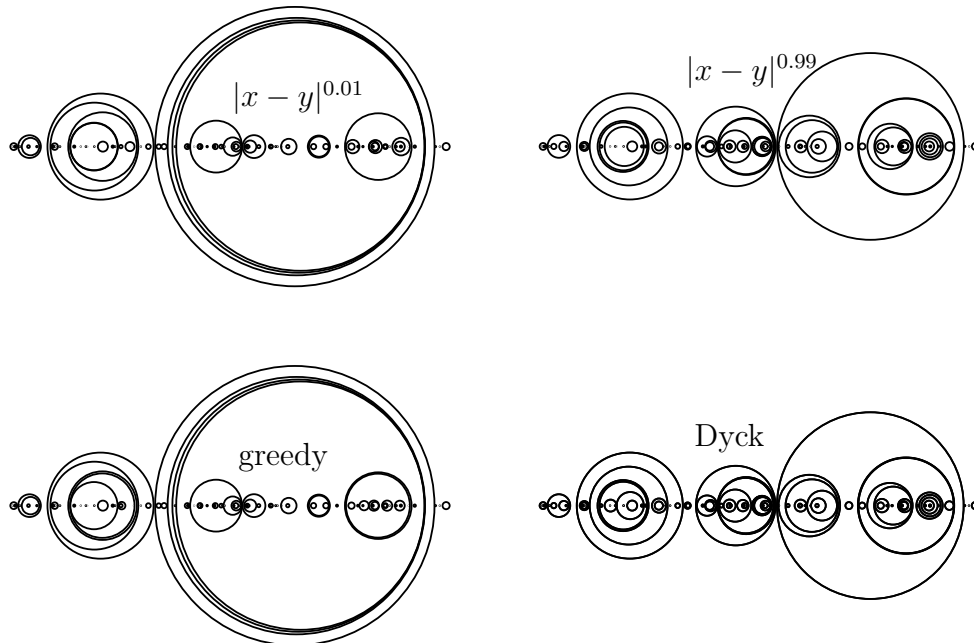


Figure 4. Four different matchings of the exact same set of  $n = 100$  points illustrated using McCann circles (§2.2): the optimal matching for  $c(x, y) = |x - y|^{0.01}$  is very similar to the greedy matching, for  $c(x, y) = |x - y|^{0.99}$  it is close to the Dyck matching.

## 2. Results

**2.1. Main result.** It is clear that a greedy matching will initially perform well since it is matching points that are very close to each other. The main concern is that the greedy matching ends up maneuvering itself into a situation where all remaining choices are bad. We will now show that, in a suitable sense, this does not happen. The result will be phrased in terms of the Wasserstein distance  $W_1$  between the sets  $X$  and  $Y$  where the sets will be assumed to lie in a metric space

$$W_1(X, Y) = \inf_{\pi \in S_n} \sum_{i=1}^n d(x_i, y_{\pi(i)}).$$

When  $0 < p < 1$  it follows from Hölder's inequality with coefficients  $1/p$  and  $1/(1-p)$  that we can bound the size of the optimal matching by

$$W_p^p(X, Y) = \inf_{\pi \in S_n} \sum_{i=1}^n d(x_i, y_{\pi(i)})^p \leq W_1(X, Y)^p \cdot n^{1-p}.$$

Our main result shows that, in any arbitrary metric space, the greedy matching, proceeding at each step blindly and without any foresight into the future, achieves the same rate up to a constant when  $0 < p < 1/2$ . Note that the greedy matching will usually be very different from the one that minimizes  $W_1$ .

**Theorem 1.** (Main result) *For any  $0 < p < 1$  there is a constant  $c_p > 0$  such that for any two sets of  $n$  points  $X, Y$  in any metric space, the greedy matching produces matching satisfying, with respect to the cost function  $c(x, y) = d(x, y)^p$ ,*

$$\text{Greedy}_p(X, Y) \leq c_p \cdot W_1(X, Y)^p \cdot \begin{cases} n^{1-p} & \text{if } 0 < p < \frac{1}{2}, \\ \sqrt{n} \cdot \log n & \text{if } p = \frac{1}{2}, \\ n^p & \text{if } \frac{1}{2} < p < 1. \end{cases}$$

The result is optimal up to constants when  $0 < p < 1/2$ . We illustrate this on  $[0, 1]$  equipped with the Euclidean distance (see Fig. 5). If the blue points are close to 0 and the red points are close to 1, then the greedy algorithm takes  $\text{Greedy}_p(X, Y) \sim n$  while  $W_1(X, Y) \sim n$ . If all the points alternate in an equispaced way,  $\text{Greedy}_p(X, Y) \sim n^{1-p}$  and  $W_1(X, Y) \sim 1$ . The proof shows that  $c_p \sim 1 + 2p$  for  $p$  small. There is a clear change around  $p = 1/2$ , the greedy matching becomes less effective (see also Fig. 8). We also note a result of Bobkov–Ledoux [5] in one dimension: the optimal transport cost for  $d(x, y)^p$  and  $0 < p < 1$  is, in expectation, smaller than the one induced by the ordered optimal  $p = 1$  matching.



Figure 5. Examples where the Theorem is sharp ( $0 < p < 1/2$ ).

**2.2. Greedy is non-crossing.** It is known (see McCann [20]) that if we match points  $\{x_1, \dots, x_n\}$  to  $\{y_1, \dots, y_n\}$  under a cost function  $c(x, y) = h(|x - y|)$  with  $h \geq 0$  concave, then the optimal matching satisfies a non-crossing condition which can be described as follows: if the optimal matching sends  $x_i$  to  $y_{\pi(i)}$ , then the  $n$  circles  $C_i$  that go tangentially through  $x_i$  and  $y_{\pi(i)}$ , the circles

$$C_i = \left\{ z \in \mathbb{R}^2 : \left\| z - \frac{x_i + y_{\pi(i)}}{2} \right\| = \frac{|y_{\pi(i)} - x_i|}{2} \right\},$$

do not intersect. The greedy matching has this desirable property for trivial reasons.

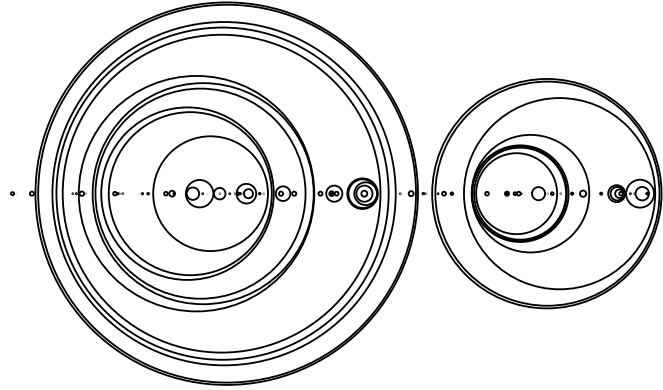


Figure 6. The circles  $C_i$  for an example of  $n = 100$  points greedily matched with another  $n$  points under  $c(x, y) = |x - y|^{0.01}$ .

**Proposition 1.** *The greedy matching is non-crossing.*

*Proof.* The argument is very simple. Suppose  $i < j$  and circle  $C_i$  intersects circle  $C_j$ .  $C_i$  intersects  $x_i$  and  $y_{\pi(i)}$  while  $C_j$  intersects  $x_j$  and  $x_{\pi(j)}$ . Suppose w.l.o.g.  $x_i < y_{\pi(i)}$  (otherwise relabel  $X$  and  $Y$ ). Since  $C_i$  and  $C_j$  intersect we have to either have  $x_i < x_j < y_{\pi(i)}$  or  $x_i < y_{\pi(j)} < y_{\pi(i)}$  but either of these cases leads to a contradiction at stage  $i$  of the greedy algorithm.  $\square$

**2.3. Random points I.** One of the most interesting cases is naturally that of matching random points to random points. Here, we can show that the greedy matching leads to nontrivial results.

**Corollary 1.** *Let  $X, Y$  be two sets of  $n$  uniform i.i.d. random variables on  $[0, 1]^d$ . The greedy matching subject to  $c(x, y) = \|x - y\|^p$  for  $0 < p < 1/2$  satisfies*

$$\mathbb{E} \text{ Greedy}_p(X, Y) \leq c_p \cdot \begin{cases} n^{1-p} & \text{when } d = 1 \text{ and } p < 1/4, \\ n^{1-p/2} & \text{when } d = 1 \text{ and } 1/4 \leq p < 1/2, \\ n^{1-p/2} & \text{when } d = 2, \\ n^{1-p/d} & \text{when } d \geq 3. \end{cases}$$

Of these, the first inequality ( $d = 1$  and  $0 < p < 1/4$ ) is optimal up to constants. It seems likely to assume that, when  $d = 1$  and  $p < 1/2$ , we have  $\mathbb{E} \text{ Greedy}_p(X, Y) \leq c_p \cdot n^{1-p}$ . Even stronger results seem to be true: the greedy matching appears to be *remarkably* close to the cost function, even at the pointwise level. The Dyck matching is known to produce an optimal rate.

**Theorem 2.** (Caracciolo–D’Achille–Erba–Sportiello [8]) *For various different models of  $n$  random points on  $[0, 1]$ , we have*

$$\mathbb{E} (\text{Dyck}_n) \sim \begin{cases} n^{1-p} & \text{if } 0 < p < \frac{1}{2}, \\ n^{1/2} \log n & \text{if } p = \frac{1}{2}, \\ n^{1/2} & \text{if } \frac{1}{2} < p \leq 1. \end{cases}$$

The Dyck matching is an a priori global construction that is naturally connected to the structure of a Brownian bridge, a fact that allows for a variety of tools to be applied. Conversely, the greedy algorithm is a ‘blind’ local algorithm: it is difficult to predict what it will do without running it. As such, proving a result along the

lines of Caracciolo–D’Achille–Erba–Sportiello [8] for the greedy matching appears to be difficult and in need of new ideas.

**2.4. Random points II.** More can be said if we assume that the points are chosen uniformly at random with respect to a fixed probability measure  $\mu$  since one would then expect a certain limiting behavior to arise. This is indeed the case.

**Theorem 3.** (Barthe–Bordenave [2], special case) *Let  $0 < p < d/2$  and  $\mu$  be the uniform measure on a bounded set  $\Omega \subset \mathbb{R}^d$ . If  $X, Y$  are i.i.d. copies from  $\mu$ , then*

$$\lim_{n \rightarrow \infty} n^{\frac{p}{d}-1} \cdot W_p^p(X, Y) = \beta_p(d) \cdot |\Omega|^{\frac{p}{d}}.$$

When  $d = 1$ , the result has recently been extended by Goldman–Trevisan [17] to also allow for randomness with respect to variable absolutely continuous density.

**Theorem 4.** (Goldman–Trevisan [17], special case) *Let  $0 < p < 1/2$ . There exists  $c_p > 0$  such that for compactly supported  $\mu$  with a.c. density  $f(x)dx$*

$$\lim_{n \rightarrow \infty} n^{p-1} \cdot W_p^p(X, Y) = c_p \int_{\mathbb{R}} f(x)^{1-p} dx.$$

The restriction  $p < 1/2$  is necessary, the behavior is strictly different at  $p = 1/2$  (this is the scale where the fluctuations of the empirical distribution of the points starts to come into play). Since the uniform measure is a special case of a measure of the form  $f(x)dx$ , the implicit (universal) constants are the same as in the result of Barthe–Bordenave in the sense of  $c_p = \beta_p(1)$ . We provide an easy explicit lower bound on these constants.

**Proposition 2.** *Let  $0 < p < d/2$ . Then*

$$\beta_p(d) \geq \omega_d^{-\frac{p}{d}} \cdot \Gamma\left(1 + \frac{p}{d}\right),$$

where  $\omega_d$  is the volume of the unit ball in  $\mathbb{R}^d$ .

The argument is not terribly difficult and somewhat standard for these types of problems, see [21, 24]. The bound seems to be rather accurate for small values of  $p$ . As a consequence, we obtain two-sided bounds in the one-dimensional setting when  $p$  is close to 0. There is an interesting heuristic: we expect that the optimal matching will send most points distance  $\sim c \cdot n^{-1}$  with  $c \sim 1$  ranging over some different values but being approximately at order  $\sim 1$  and thus the cost should be close to  $\sim c^p \cdot n^{1-p}$ . However, since  $p \sim 0+$ , one expects  $c^p \sim 1$  and thus that the optimal transport cost is perhaps given by  $(1 + \mathcal{O}(p)) \cdot n^{1-p}$ .

**Corollary 2.** (Prop. 2 and [8]) *For  $0 < p < 1/2$ ,*

$$2^{-p} \cdot \Gamma(1 + p) \leq \beta_1(p) \leq \frac{1}{1 - 2p} \frac{2^p}{\Gamma(1 - p)}.$$

In particular, we conclude that  $\beta_1(p) = 1 + \mathcal{O}(p)$  when  $p \sim 0$  is small.

**2.5. Extreme concave matching.** The cost  $c(x, y) = |x - y|^p$  is particularly natural. There is, in a suitable sense, a canonical limit as  $p \rightarrow 0^+$  since

$$\lim_{p \rightarrow 0^+} \frac{|x - y|^p - 1}{p} = \log |x - y|.$$

As suggested by Fig. 2, the greedy algorithm performs very well in this setting. We prove a basic result suggesting that this is not a coincidence.

**Proposition 3.** *Let  $X, Y$  be two sets of  $n$  distinct points in a metric space such that  $d(x_i, y_j) \leq 1$  and assume that all pairwise distances are unique. Then there exists  $K \in \mathbb{N}$  such that for all  $k \geq K$  the solution of the optimal matching problem*

$$\min_{\pi \in S_n} \sum_{i=1}^n (\log d(x_i, y_{\pi(i)}))^{2k+1} \quad \text{is given by the greedy matching.}$$

The argument is quite simple and there is nothing particularly special about the logarithm, similar results could be attained with many other cost functions that are dramatically different across different length scales. The main point of this simple Proposition is to illustrate that the effectiveness of the greedy matching in the setting of very concave cost functions is perhaps not entirely surprising: the dramatic separation of scales puts a heavy reward on having matching with very small distance which, coupled with the separation of scales, then suggests the greedy algorithm as a natural object.

**2.6. Numerics.** The purpose of this section is to consider the behavior of the greedy algorithm and the Dyck matching when matching  $n$  i.i.d. random points on  $[0, 1]$  given the cost function  $c(x, y) = |x - y|^p$  for  $0 \leq p \leq 1/2$ . The results suggest that, at least for random points, the greedy matching leads to results that are remarkably close to the ground truth: this effect becomes more pronounced when  $p$  becomes smaller (also at least partially suggested by the Proposition).

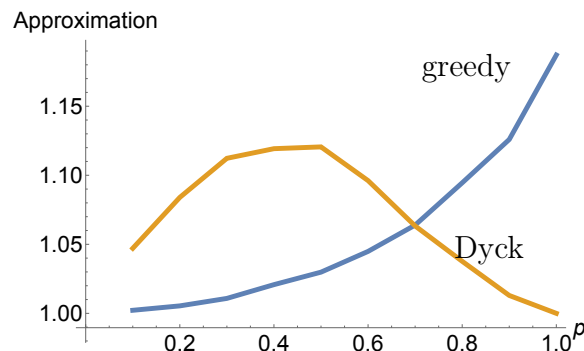
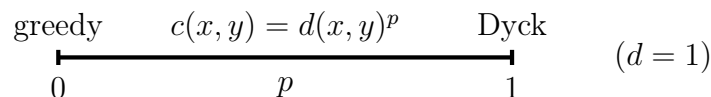


Figure 7. Average ratio of cost divided by the minimal cost for both the Dyck matching and the greedy matching for  $n = 250$  iid points subject to the cost function  $c(x, y) = |x - y|^p$ .

The effectiveness of the greedy algorithm is strictly restricted to the region  $0 < p < 1/2$ : for  $p \geq 1/2$ , the greedy algorithm starts to scale differently and becomes less effective. This can already be observed for small values of  $n$  (and is pointed out by d'Achille [11, Section 1.4] for  $p \geq 1$ ). We also observe that, as  $p$  becomes smaller, the effectiveness of the greedy matching increases dramatically while, for  $p$  close to 1, the effectiveness of the Dyck matching increases dramatically (see Fig. 8). The Dyck matching is optimal when  $p = 1$  (see [9]).

Fig. 7 is well suited to iterate a main contribution of our paper: when  $d = 1$  and considering matchings with  $c(x, y) = d(x, y)^p$  and  $0 < p < 1$ , there exists a natural dichotomy depending on whether  $p$  is close to 0 or whether it is close to 1.



A natural question is, for example, whether one can identify the precise value  $0 < p^* < 1$  where the effectiveness of the two algorithms undergoes a phase transition

and the matching problem for  $n$  iid random points is better approximated by the Dyck matching or the greedy matching, respectively.

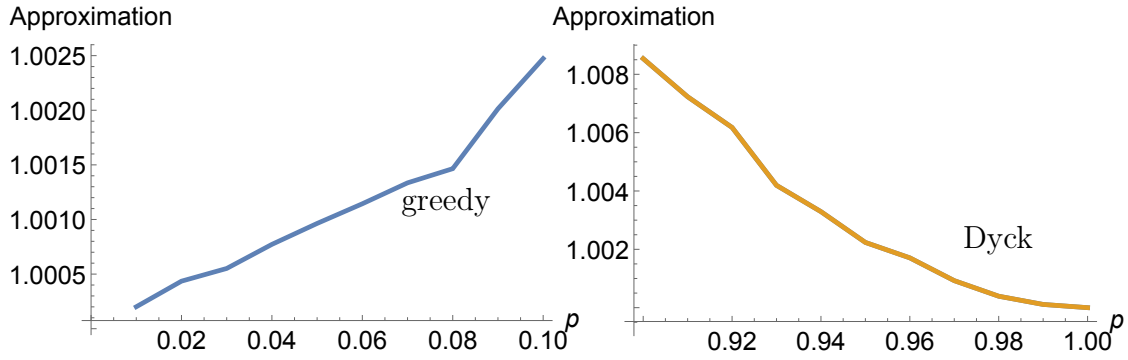


Figure 8. Ratio of obtained matching cost divided by the minimal cost for  $n = 100$  random points. Left: the greedy matching when  $p \sim 0$ . Right: the Dyck matching when  $p \sim 1$ .

### 3. Proof of Theorem 1

*Proof.* We assume the sets of points are  $X = \{x_1, \dots, x_n\}$  and  $Y = \{y_1, \dots, y_n\}$  and we will use  $X_k, Y_k$  to denote the set of points after  $k-1$  points have been removed following the greedy algorithm. In particular,  $X = X_1$  and  $Y = Y_1$ . We will also abbreviate the cost at stage  $k$  via

$$c_k = \inf_{x \in X_k, y \in Y_k} d(x, y).$$

Our goal is to estimate

$$\text{the cost of the greedy algorithm} \quad \sum_{k=1}^n c_k^p.$$

Recalling that, by definition,  $X_k$  and  $Y_k$  have  $n-k+1$  elements each, the Wasserstein distance  $W_1$  can be written as

$$W_1(X_k, Y_k) = \min_{\substack{\pi: X_k \rightarrow Y_k \\ \pi \text{ bijective}}} \sum_{x \in X_k} d(x, \pi(x))$$

where the minimum ranges over all bijections. At this point, we remark that the definition of  $W_1$  is slightly more comprehensive (the infimum ranges over all ways of splitting and rearranging points): at this point, we employ the celebrated result of Birkhoff [3] and von Neumann [23] ensuring that in the case of  $n$  equal masses being transported to  $n$  equal masses, the optimal solution can be realized by a permutation (no mass is ‘split’), we refer to [18] for a generalization.

Our first observation uses averaging. Let us use  $\pi$  to denote the permutation achieving the optimal  $W^1$  transport cost. Then the distance achieved by the greedy matching in the  $k$ -th step can be bounded from above by

$$c_k = \inf_{x \in X_k, y \in Y_k} d(x, y) \leq \frac{1}{n-k+1} \sum_{x \in X_k} d(x, \pi(x)) = \frac{W_1(X_k, Y_k)}{n-k+1}.$$

The second ingredient will be to show that  $W_1(X_{k+1}, Y_{k+1})$  cannot be much larger than  $W_1(X_k, Y_k)$ . For this purpose, let us assume that  $X_k$  is given by the points



$x_1, x_2, \dots, x_{n-k+1}$  and, similarly,  $Y_k$  is given by  $y_1, y_2, \dots, y_{n-k+1}$ . Then, after possibly relabeling the points, we have that

$$W_1(X_k, Y_k) = \sum_{i=1}^{n-k+1} d(x_i, y_i).$$

Let us now assume that the greedy matching at this point matches  $x_i$  and  $y_j$ , meaning that  $d(x_i, y_j)$  is the smallest pairwise distance in  $X_k \times Y_k$ . Then the greedy matching is going to match these two points up and the remaining sets of points are given by

$$X_{k+1} = X_k \setminus \{x_i\} \quad \text{and} \quad Y_{k+1} = Y_k \setminus \{y_j\}.$$

We will provide an upper bound on  $W_1(X_{k+1}, Y_{k+1})$  by taking the original matching between  $X_k$  and  $Y_k$  and then modify it a little to obtain a matching for  $X_{k+1}$  with  $Y_{k+1}$ . This is done by a simple modification: the point  $x_j$  is now mapped to  $y_i$  (note that  $x_i$  has been mapped to  $y_j$  and both have been removed from the set). We preserve all other matchings. Then

$$W_1(X_{k+1}, Y_{k+1}) \leq W_1(X_k, Y_k) + d(x_j, y_i) - d(x_i, y_i) - d(x_j, y_j).$$

At this point, we invoke the triangle inequality and argue that

$$\begin{aligned} d(x_j, y_i) &\leq d(x_j, y_j) + d(y_j, y_i) \\ &\leq d(x_j, y_j) + d(y_j, x_i) + d(x_i, y_i) \end{aligned}$$

Combining the last two inequalities, we realize that we can bound the increase in the  $W_1$ -distance between the two sets in terms of the cost function of the greedy matching at the  $k$ -th step by

$$\begin{aligned} W_1(X_{k+1}, Y_{k+1}) &\leq W_1(X_k, Y_k) + d(x_j, y_i) - d(x_i, y_i) - d(x_j, y_j) \\ &\leq W_1(X_k, Y_k) + d(x_i, y_j) \\ &= W_1(X_k, Y_k) + c_k. \end{aligned}$$

Combining these two ingredients, we arrive

$$\begin{aligned} W_1(X_{k+1}, Y_{k+1}) &\leq W_1(X_k, Y_k) + c_k \\ &\leq W_1(X_k, Y_k) + \frac{W_1(X_k, Y_k)}{n-k+1} \\ &= W_1(X_k, Y_k) \left(1 + \frac{1}{n-k+1}\right). \end{aligned}$$

By induction, we obtain

$$W_1(X_k, Y_k) \leq W_1(X, Y) \cdot \prod_{\ell=1}^{k-1} \left(1 + \frac{1}{n-\ell+1}\right).$$

Observe that

$$\prod_{\ell=1}^{k-1} \left(1 + \frac{1}{n-\ell+1}\right) = \prod_{\ell=1}^{k-1} \frac{n-\ell+2}{n-\ell+1} = \frac{n+1}{n-k+2}.$$

Therefore

$$W_1(X_k, Y_k) \leq \frac{n+1}{n-k+2} \cdot W_1(X, Y).$$

Applying the pigeonhole principle one more time, we see that

$$c_k \leq \frac{W_1(X_k, Y_k)}{n - k + 1} \leq \frac{n + 1}{(n - k + 1)^2} \cdot W_1(X, Y).$$

Thus

$$\sum_{k=1}^n c_k^p \leq W_1(X, Y)^p \cdot (n + 1)^p \cdot \sum_{k=1}^n \frac{1}{(n - k + 1)^{2p}}.$$

We have

$$\sum_{k=1}^n \frac{1}{(n - k + 1)^{2p}} = \sum_{k=1}^n \frac{1}{k^{2p}}.$$

When  $0 < p < 1/2$ , we have

$$\sum_{k=1}^n \frac{1}{k^{2p}} \leq 1 + \int_1^{n+1} \frac{1}{x^{2p}} dx \leq 1 + \frac{(n + 1)^{1-2p}}{1 - 2p}.$$

Dealing with the remaining cases in the usual fashion, we obtain

$$\sum_{k=1}^n c_k^p \leq c_p \cdot W_1(X, Y)^p \cdot \begin{cases} n^{1-p} & \text{if } 0 < p < \frac{1}{2}, \\ \sqrt{n} \cdot \log n & \text{if } p = \frac{1}{2}, \\ n^p & \text{if } \frac{1}{2} < p < 1. \end{cases}$$

The argument also shows that, for  $p < 1/2$  we have

$$\text{Greedy}_p(X, Y) \leq \left( \frac{1}{1 - 2p} + o(1) \right) \cdot n^{1-p} \cdot W_1(X, Y).$$

In particular, for  $p$  close to 0, we have that  $1/(1 - 2p) \sim 1 + 2p$  and the implicit constant is close to 1.  $\square$

### 3.1. Proof of Corollary 1.

*Proof.* Corollary 1 follows immediately from the Theorem. The missing ingredient is a good estimate on  $W_1(X, Y)$  where  $X$  and  $Y$  are two sets of  $n$  i.i.d. uniformly distributed points in  $[0, 1]^d$ . The case  $d = 2$  is arguably the most famous, the celebrated result of Ajtai, Komlos and Tusnady [1] ensures that

$$c_1 \sqrt{n \log n} \leq \min_{\pi \in S_n} \sum_{i=1}^n \|x_i - y_{\pi(i)}\| \leq c_2 \sqrt{n \log n}.$$

with high probability. The one-dimensional case is a bit simpler and one has (see, for example, [4]), with high probability,

$$\min_{\pi \in S_n} \sum_{i=1}^n \|x_i - y_{\pi(i)}\| \leq c \sqrt{n}.$$

The case  $d \geq 3$  where

$$\min_{\pi \in S_n} \sum_{i=1}^n |x_i - y_{\pi(i)}| \leq c \cdot n^{1-1/d}$$

was already remarked by Ajtai, Komlos, Tusnady [1]. A modern treatment of a much more general case is given in [15]. This proves that, for  $X$  and  $Y$  i.i.d. random points

and  $p < 1/2$

$$\mathbb{E} \text{Greedy}_p(X, Y) \leq c_p \cdot \begin{cases} n^{1-p/2} & \text{when } d = 1, \\ n^{1-p/2} \cdot (\log n)^{p/2} & \text{when } d = 2, \\ n^{1-p/d} & \text{when } d \geq 3. \end{cases}$$

There is a slight improvement that was pointed out to us by an anonymous referee. For any  $0 < \alpha < 1$ , we can write

$$\sum_{i=1}^n \|x_i - y_{\pi(i)}\|^p = \sum_{i=1}^n (\|x_i - y_{\pi(i)}\|^\alpha)^{p/\alpha}.$$

If  $d(x, y)$  is a metric, then  $d(x, y)^\alpha$  is also a metric for every  $0 < \alpha < 1$ . Moreover, since  $x \rightarrow x^\alpha$  is monotonically increasing, the greedy algorithm with respect to the metric  $d(x, y)^\alpha$  picks the exact same matching as the greedy algorithm with respect to  $d(x, y)$ . We use  $W_{1,\alpha}$  to denote the Wasserstein-1 distance with respect to the metric  $d_\alpha(x, y) = \|x - y\|^\alpha$ . Applying Theorem 1 to this new metric, we get that, assuming  $x_i \leftrightarrow y_i$  to be the greedy matching and  $0 < p/\alpha < 1$  instead of  $p$ , then

$$\sum_{i=1}^n \|x_i - y_i\|^p \leq c_{p,\alpha} (W_{1,\alpha}(X, Y))^{p/\alpha} \cdot \begin{cases} n^{1-p/\alpha} & \text{if } 0 < p/\alpha < \frac{1}{2}, \\ \sqrt{n} \cdot \log n & \text{if } p/\alpha = \frac{1}{2}, \\ n^{p/\alpha} & \text{if } \frac{1}{2} < p/\alpha < 1, \end{cases}$$

where

$$W_{1,\alpha}(X, Y) = \inf_{\pi \in S_n} \sum_{i=1}^n \|x_i - y_{\pi(i)}\|^\alpha.$$

The result of Barthe–Bordenave [2] implies that as long as  $\alpha < d/2$  and the  $X, Y$  are i.i.d. random variables with respect to the uniform measure, then

$$\inf_{\pi \in S_n} \sum_{i=1}^n \|x_i - y_{\pi(i)}\|^\alpha \lesssim n^{1-\alpha/d}.$$

Combining all these results, assuming that  $0 < \alpha < 1$  as well as  $0 < p < \alpha$  and  $\alpha < d/2$ , then

$$\sum_{i=1}^n \|x_i - y_i\|^p \leq c_{p,\alpha} \begin{cases} n^{1-p/d} & \text{if } 0 < p/\alpha < \frac{1}{2}, \\ n^{1-p/d} \cdot \log n & \text{if } p/\alpha = \frac{1}{2}, \\ n^{p(2/\alpha-d)} & \text{if } \frac{1}{2} < p/\alpha < 1. \end{cases}$$

The first case improves on the previous estimate when  $d = 1$ , this requires  $\alpha < 1/2$  and thus  $p < 1/4$ . When  $d = 2$ , then for any  $p < 1/2$  we can find  $p < \alpha < 1$  such that  $p/\alpha < 1/2$  which implies the upper bound  $n^{1-p/2}$  without logarithmic corrections. Altogether, this gives, for  $0 < p < 1/2$

$$\mathbb{E} \text{Greedy}_p(X, Y) \leq c_p \cdot \begin{cases} n^{1-p} & \text{when } d = 1 \text{ and } p < 1/4, \\ n^{1-p/2} & \text{when } d = 1 \text{ and } 1/4 \leq p < 1, \\ n^{1-p/2} & \text{when } d = 2, \\ n^{1-p/d} & \text{when } d \geq 3. \end{cases} \quad \square$$

**3.2. Proof of Proposition 2.** Since the constant is independent of the domain, it will be enough to derive an upper bound in a fixed, arbitrary domain. We choose the unit cube  $[0, 1]^d$  for convenience, however, we emphasize that there is nothing particularly special about the unit cube.

**Lemma.** Given  $n$  i.i.d. uniform points  $X_1, \dots, X_n$  in  $[0, 1]^d$  and an independent uniform point  $Y$ , we have for all sufficiently small  $0 < \varepsilon < \varepsilon_0$  that whenever  $Y$  is sufficiently far from the boundary of the unit cube  $d(Y, \partial[0, 1]^d) \geq \varepsilon$ , then

$$\mathbb{P}(\min_{1 \leq i \leq n} |Y - X_i| \geq \varepsilon) = (1 - \omega_d \varepsilon^d)^n,$$

where  $\omega_d$  is the volume of the unit ball in  $\mathbb{R}^d$ .

*Proof.* Conditional on  $Y$ , since the  $X_i$ 's are all independent, we obtain

$$\mathbb{P}(\min_{1 \leq i \leq n} |Y - X_i| \geq \varepsilon | Y) = (1 - |B_Y(\varepsilon) \cap [0, 1]^d|)^n.$$

Since  $d(Y, \partial[0, 1]^d) \geq \varepsilon$ , we have

$$(1 - |B_Y(\varepsilon) \cap [0, 1]^d|)^n = (1 - \omega_d \varepsilon^d)^n.$$

A simple monotonicity argument guarantees that for all  $\varepsilon$  the conclusion still holds, up to replacing  $=$  with  $\geq$ .  $\square$

*Proof of Proposition 2.* We assume the sets of points  $\{X_1, \dots, X_n\} \subset [0, 1]^d$  and  $\{Y_1, \dots, Y_n\} \subset [0, 1]^d$  are both sets of independent uniformly distributed random variables in  $[0, 1]^d$ . The main idea is the use of the trivial bound

$$\mathbb{E} \inf_{\pi \in S_n} \sum_{i=1}^n d(x_i, y_{\pi(i)})^p \geq \sum_{i=1}^n \mathbb{E} \inf_{x \in X} |Y_i - x|^p = n \mathbb{E} \min_{1 \leq i \leq n} |X_i - Y|^p$$

We introduce a change of coordinates  $\varepsilon = c^p / n^{p/d}$ . Then, for all sufficiently small  $\varepsilon$  (i.e., for fixed  $c, p, d$  and  $n \rightarrow +\infty$ ), we have

$$\begin{aligned} \mathbb{P}\left(\min_{1 \leq j \leq n} |X_i - Y|^p \geq \varepsilon\right) &= \mathbb{P}\left(\min_{1 \leq j \leq n} |X_i - Y| \geq \frac{c}{n^{1/d}}\right) \\ &= \left(1 - \frac{\omega_d c^d}{n}\right)^n \rightarrow e^{-\omega_d c^d}. \end{aligned}$$

As we remarked earlier, for all  $\varepsilon \geq 0$  (i.e., for all  $c$ ) we have

$$\mathbb{P}\left(\min_{1 \leq j \leq n} |X_i - Y|^p \geq \varepsilon\right) \geq e^{-\omega_d c^p},$$

so that we obtain, as  $n \rightarrow +\infty$ ,

$$\begin{aligned} n \cdot \mathbb{E} \inf_{1 \leq i \leq n} |X_i - Y|^p &= n \int_0^\infty \mathbb{P}\left(\min_{1 \leq j \leq n} |X_i - Y|^p \geq \varepsilon\right) d\varepsilon \\ &= \int_0^{+\infty} p c^{p-1} e^{-\omega_d c^d} dc \\ &= n^{1-\frac{p}{d}} \cdot \omega_d^{-\frac{p}{d}} \cdot \Gamma\left(1 + \frac{p}{d}\right). \end{aligned}$$

We note that the argument slightly improves when  $X_i$  is close to the boundary (since then the volume of a neighborhood intersected with  $[0, 1]^d$  has smaller volume). However, since the number of points distance  $\sim n^{-1/d}$  close to the boundary is  $\sim n^{(d-1)/d} \ll n$ , exploiting this fact would not lead to better asymptotics.  $\square$

### 3.3. Proof of Proposition 3.

*Proof.* Since  $0 < d(x_i, y_j) < 1$ , all the summands are negative and can use the trivial estimate

$$\begin{aligned} n \min_{1 \leq i \leq n} (\log d(x_i, y_{\pi(i)}))^{2k+1} &\leq \sum_{i=1}^n (\log d(x_i, y_{\pi(i)}))^{2k+1} \\ &\leq \min_{1 \leq i \leq n} (\log d(x_i, y_{\pi(i)}))^{2k+1}. \end{aligned}$$

Suppose now that

$$\min_{1 \leq i \leq n} d(x_i, y_{\pi(i)}) > \min_{1 \leq i, j \leq n} d(x_i, y_j).$$

Then, for all  $k \geq K_1$  sufficiently large,

$$n \min_{1 \leq i \leq n} (\log d(x_i, y_{\pi(i)}))^{2k+1} > \min_{1 \leq i, j \leq n} (\log d(x_i, y_j))^{2k+1}$$

since one grows exponentially larger than the other. This shows that any optimal matching has to at least coincide with the greedy matching in the first step. We remove the closest pair of points  $(x_i, y_j)$  and repeat the procedure on the remaining set of points. We see that for all  $k \geq K_2$ , the next step has to coincide with that of the greedy matching. Repeating the procedure, we see that for all  $K \geq \max(K_1, \dots, K_{n-1})$  the minimum can only be given by the greedy matching.  $\square$

*Acknowledgment.* The authors gratefully acknowledge support from the Kantorovich Initiative. A.O. is supported by an AMS-Simons Travel Grant. S.S. was supported by the NSF (DMS-2123224). The authors are indebted to Aleh Tsyvinski for drawing their attention to the problem and to an anonymous referee for suggesting improved rates for Corollary 1.

## References

- [1] AJTAI, M., J. KOMLÓS, and G. TUSNÁDY: On optimal matchings. - *Combinatorica* 4, 1984, 259–264.
- [2] BARTHE, F., and C. BORDENAVE: Combinatorial optimization over two random point sets. - In: *Séminaire de probabilités XLV*, Springer, Cham, 2013, 483–535.
- [3] BIRKHOFF, G.: Tres observaciones sobre el álgebra lineal. - *Universidad Nacional de Tucumán, Revista Series A* 5, 147–151, 1946.
- [4] BOBKOV, S., and M. LEDOUX: One-dimensional empirical measures, order statistics, and Kantorovich transport distances. - *Mem. Amer. Math. Soc.* 261:1259, Amer. Math. Soc., 2019.
- [5] BOBKOV, S., and M. LEDOUX: Transport inequalities on Euclidean spaces for non-Euclidean metrics. - *J. Fourier Anal. Appl.* 26, 2020, 60.
- [6] BOERMA, J., A. TSYVINSKI, R. WANG, and Z. ZHANG: Composite sorting. - arXiv:2303.06701 [econ.GN], 2023.
- [7] BROWN, L., and S. STEINERBERGER: On the Wasserstein distance between classical sequences and the Lebesgue measure. - *Trans. Amer. Math. Soc.* 373, 2020, 8943–8962.
- [8] CARACCILO, S., M. D’ACHILLE, V. ERBA, and A. SPORTIELLO: The Dyck bound in the concave 1-dimensional random assignment model. - *J. Phys. A* 53, 2020, 064001.
- [9] CARACCILO, S., V. ERBA, and A. SPORTIELLO: The number of optimal matchings for Euclidean Assignment on the line. - *J. Stat. Phys.* 183, 2021, 1–27.
- [10] CARACCILO, S., V. ERBA, and A. SPORTIELLO: The  $p$ -Airy distribution. - arXiv:2010.14468 [math.CO], 2020.

- [11] D'ACHILLE, M.: Statistical properties of the Euclidean random assignment problem. - PhD Thesis, Université Paris-Saclay, 2020.
- [12] DELON, J., J. SALOMON, and A. SOBOLEVSKI: Local matching indicators for concave transport costs. - *C. R. Math. Acad. Sci. Paris* 348, 2010, 901–905.
- [13] DELON, J., J. SALOMON, and A. SOBOLEVSKI: Local matching indicators for transport problems with concave costs. - *SIAM J. Discrete Math.* 26, 2012, 801–827.
- [14] DIACONIS, P.: Group representations in probability and statistics. - *IMS Lecture Notes Monogr. Ser.* 11, 1988.
- [15] FOURNIER, N., and A. GUILLIN: On the rate of convergence in Wasserstein distance of the empirical measure. - *Probab. Theory Related Fields* 162, 2015, 707–738.
- [16] GANGBO, W., and R. MCCANN: The geometry of optimal transportation. - *Acta Math.* 177, 1996, 113–161.
- [17] GOLDMAN, M., and D. TREVISAN: On the concave one-dimensional random assignment problem and Young integration theory. - *arXiv:2305.09234 [math.PR]*, 2023.
- [18] HOSSEINI, B., and S. STEINERBERGER: Intrinsic Sparsity of Kantorovich solutions. - *C. R. Math. Acad. Sci. Paris* 360:G10, 2022, 1173–1175.
- [19] JUILLET, N.: On a solution to the Monge transport problem on the real line arising from the strictly concave case. - *SIAM J. Math. Anal.* 52, 2020, 4783–4805.
- [20] MCCANN, R.: Exact solutions to the transportation problem on the line. - *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 455, 1999, 1341–1380.
- [21] STEELE, J. M.: Probability theory and combinatorial optimization. - *Society for Industrial and Applied Mathematics*, 1997.
- [22] STEINERBERGER, S.: A Wasserstein inequality and minimal Green energy on compact manifolds. - *J. Funct. Anal.* 281, 2021, 109076.
- [23] VON NEUMANN, J.: A certain zero-sum two-person game equivalent to an optimal assignment problem. - *Ann. of Math. Stud.* 28, 1953, 5–12.
- [24] YUKICH, J. E.: Probability theory of classical Euclidean optimization problems. - *Springer*, 2006.
- [25] ZINTERHOF, P.: Über einige Abschätzungen bei der Approximation von Funktionen mit Gleichverteilungsmethoden. - *Österreich. Akad. Wiss. Math.-Naturwiss. Kl. S.-B. II* 185:1-3, 1976, 121–132.

Received 17 April 2025 • Revision received 26 August 2025 • Accepted 26 August 2025

Published online 18 September 2025

Andrea Ottolini  
Yale University  
Cowles Foundation  
New Haven, CT 06520-8281, USA  
andrea.ottolini@yale.edu

Stefan Steinerberger  
University of Washington  
Department of Mathematics  
Seattle, WA 98195, USA  
steinerb@uw.edu