Electronic Journal of Probability

Electron. J. Probab. **29** (2024), article no. 106, 1–20.
ISSN: 1083-6489 https://doi.org/10.1214/24-EJP1171

There is no stationary p-cyclically monotone Poisson matching in 2d*

Martin Huesmann† Francesco Mattesini‡ Felix Otto§

Abstract

We show that for $p > 1$ there is no p-cyclically monotone stationary matching of two independent Poisson processes in dimension $d = 2$. The proof combines the p-harmonic approximation result from [15, Theorem 1.1] with local asymptotics for the two-dimensional matching problem. Moreover, we prove a.s. local upper bounds of the correct order in the case $p > 1$, which, to the best of our knowledge, are not readily available in the current literature.

Keywords: matching; allocation; cyclical monotonicity; stationarity; invariance; optimal transport; p-minimal.
MSC2020 subject classifications: 60D05; 60G55; 52A22; 49Q22.
Submitted to EJP on December 5, 2023, final version accepted on July 3, 2024.

1 Introduction

Let $\{X\}, \{Y\} \subset \mathbb{R}^d$ be two locally finite[1] random point sets. We consider their matching, that is a (random) bijection from $\{X\}$ to $\{Y\}$. More precisely we will focus on dimension $d = 2$ and we are primarily interested in the case where the two random point sets are given by two independent Poisson point process of unit intensity and the map T is p-locally optimal for $p \geq 1$, meaning that for any other bijection $\tilde{T}$ that differs from T only on a finite number of points

$$\sum_X (|T(X) - X|^p - |\tilde{T}(X) - X|^p) \leq 0. \tag{1.1}$$

*All authors are supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the SPP 2265 *Random Geometric Systems*. MH and FM have been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2044−390685587, Mathematics Münster: Dynamics–Geometry–Structure. FM has been funded by the Max Planck Institute for Mathematics in the Sciences.
†Universität Münster, Germany. E-mail: martin.huesmann@uni-muenster.de
‡Universität Münster & MPI Leipzig, Germany. E-mail: francesco.mattesini@uni-muenster.de
§MPI Leipzig, Germany. E-mail: felix.otto@mis.mpg.de
[1] And thus countable.

Since the sum in (1.1) is finite, the latter provides a natural connection to the optimal transport problem between the measures

$$\mu := \sum_X \delta_X \quad \text{and} \quad \nu := \sum_Y \delta_Y \tag{1.2}$$

related via $T_\# \mu = \nu$. However, note that the map T cannot be viewed as a usual minimizer in the optimal transport problem due to the (typically) infinite number of points.

Let us now be more specific on the random setting we consider. We assume that the σ-algebra generated by $(\{X\}, \{Y\}, T)$ is rich enough so that the numbers of matched pairs $(X, Y) \in U \times V$ of any two Lebesgue-measurable sets $U, V \subset \mathbb{R}^d$ (with U or V having finite Lebesgue measure[2])

$$N_{U,V} := \#\{(X, Y) \in U \times V \mid Y = T(X)\} \in \{0, 1, \dots\}$$

are measurable. Moreover, we assume that the law of the triple $(\{X\}, \{Y\}, T)$ is stationary, that means it is invariant under the action of the additive group $\mathbb{Z}^d$

$$(\{X\}, \{Y\}, T) \mapsto (\{\bar{x} + X\}, \{\bar{x} + Y\}, T(\cdot - \bar{x}) + \bar{x}) \quad \text{for } \bar{x} \in \mathbb{Z}^d. \tag{1.3}$$

Note that stationarity is a structural assumption which will allow us to say that for any shift vector $\bar{x}$, the random natural numbers $N_{\bar{x}+U, \bar{x}+V}$ and $N_{U,V}$ have the same distribution. Furthermore, we make the assumption that the action (1.3) is ergodic.

The aim of this paper is to explore the geometric properties of the matching T between two independent Poisson point processes in dimension 2. A matching in $\mathbb{R}^2$ is called *planar* if for any choice of points X, X', the line segments connecting X to $T(X)$ and X' to $T(X')$ do not intersect. In 2002 the following question was proposed by Peres in [11].

Question 1.1. For two independent Poisson processes of intensity one does there exist a stationary planar matching?

It has been shown by Holroyd in [9], that there is no translation-invariant planar matching on the strip $\mathbb{R} \times [0, 1)$. Yet, Question 1.1 is still unsolved in $\mathbb{R}^2$ and it is far from clear what its answer should be. Again in [9], it was observed by Holroyd, that by the triangle inequality the 1-local optimality condition (1.1) implies planarity. Hence, it is natural to consider the following modification of Question 1.1, which has been proposed in [10].

Question 1.2. For two independent Poisson processes of intensity one does there exist a stationary and p-locally optimal matching?

Question 1.2 is well understood for dimension $d = 1$ and $d \geq 3$, e.g. see [10, Theorem 2 and Theorem 7] or [12]. Nevertheless, the two dimensional setting is partially unsolved. In [10, Theorem 7] it has been shown that stationary p-locally optimal matchings exist for $p < 1$. Our result, together with the one obtained in [13], complements the one of [10] in the regime $p > 1$ and $d = 2$, leaving unsolved the case $p = 1$ and of course the question on planarity.

Theorem 1.3. *For $d = 2$ and $p > 1$, there exists no stationary and ergodic ensemble of $(\{X\}, \{Y\}, T)$, where $\{X\}, \{Y\}$ are independent Poisson point processes and T is a p-cyclically monotone bijection of $\{X\}$ and $\{Y\}$.*

Before commenting on the proof of Theorem 1.3 let us add some remarks on extensions and variants of Theorem 1.3.

[2] So that the following number is finite.

Remark 1.4. Theorem 1.3 remains true if we replace the bijection T by the a priori more general object of a stationary coupling Q. This can be seen for instance by directly writing the proof in terms of couplings which essentially only requires notational changes.

Remark 1.5. Natural variants of stationary matchings are given by stationary allocations of a point process $\{X\}$, i.e. a stationary map $T : \mathbb{R}^d \to \{X\}$ such that $\mathsf{Leb}(T^{-1}(X))$ equals $\mathbb{E}[\#\{X \in (0,1)^d\}]^{-1}$, e.g. see [8, 4, 16, 14].

Mimickicking the proof of Theorem 1.3, one can show that in $d = 2$ there is no locally p-optimal stationary allocation to a Poisson process. The only place which will require minor changes is the L^∞ estimate Lemma 2.4.

Remark 1.6. Since by ergodicity and stationarity we can argue on a pathwise level via the p-harmonic approximation Theorem (cf. Section 2.3), we do not use many particular features of the Poisson point processes μ and ν in the proof of Theorem 1.3.

We mainly use two properties: The first property is concentration around the mean. The second property is more involved. Denote by W_p the L^p Wasserstein distance. We use that[3] $\frac{1}{R^d} W_p(\mu \llcorner B_R, \frac{\mu(B_R)}{|B_R|}\mathsf{Leb})$ diverges at the same rate for $R \to \infty$ as $\frac{1}{R^d} W_{p-\epsilon}(\mu \llcorner B_R, \frac{\mu(B_R)}{|B_R|}\mathsf{Leb})$ for some $\epsilon > 0$.

The proof of Theorem 1.3 goes along the same lines as in [13, Theorem 1.1], see also [13, Section 1.1]. We already remark here that there are two new ingredients: The p-harmonic approximation theorem and almost sure upper asymptotics for the matching cost. The former, already shown in [15, Theorem 1.1], states that the displacement $T(X) - X$ is close (in the p-norm distance) to a p'-harmonic[4] gradient field $|\nabla\Phi|^{p'-2}\nabla\Phi$ provided that we are in a perturbative regime, which is quantified in terms of smallnes of the local energy

$$E_p(R) := \frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X|^p \tag{1.4}$$

and of the data term, that is the distance of $\mu \llcorner B_R$ and $\nu \llcorner B_R$ to the Lebesgue measure on the ball B_R of radius R[5]

$$D_p(R) := \frac{1}{R^d} W_{p,B_R}^p(\mu, n_\mu) + \frac{R^p}{n_\mu}(n_\mu - 1)^p + \frac{1}{R^d} W_{p,B_R}^p(\nu, n_\nu) + \frac{R^p}{n_\nu}(n_\nu - 1)^p, \tag{1.5}$$

where $n_\mu = \frac{\#\{X \in B_R\}}{|B_R|}$ and $n_\nu = \frac{\#\{Y \in B_R\}}{|B_R|}$, and $W_{p,\Gamma}(\mu,\nu) = W_p(\mu \llcorner \Gamma, \nu \llcorner \Gamma)$ for a Borel set $\Gamma \subset \mathbb{R}^d$. The latter, which we state here, is our second main result and concerns concentration properties of the matching cost.

Theorem 1.7. *Let μ, ν denote two independent Poisson point processes in $\mathbb{R}^d$ of unit intensity. There exists a constant C, and a random radius $r_* < \infty$ a. s. such that for a (random) sequence of approximately dyadic radii[6] $R \geq r_*$*

$$D_p(R) \leq C \begin{cases} \ln^{\frac{p}{2}} R & \textit{if } d = 2, \\ 1 & \textit{if } d \geq 3. \end{cases} \tag{1.6}$$

We remark here that by the annealed (i. e. in expectation, see for instance [1], [2]) results for the matching problem in dimension $d = 2$ and by the concentration properties of the Poisson process we may expect that $D_p(R) \leq O(\ln^{\frac{p}{2}} R)$. However, the standard

[3] Given a measure μ on a $\mathbb{R}^d$ and a subset $\Gamma \subseteq \mathbb{R}^d$ we denote its restriction to Γ by $\mu \llcorner \Gamma(\,\cdot\,) := \mu(\,\cdot\, \cap \Gamma)$.

[4] We denote by p' the conjugate exponent of p, i. e. $\frac{1}{p} + \frac{1}{p'} = 1$.

[5] We tacitly identify the (random) number density n_μ with the uniform measure $n_\mu \mathrm{d}x$.

[6] We say that a radius R is approximately dyadic if there exists a dyadic radius R' and a constant $C \in \left(\frac{1}{2}, 2\right)$ such that $R = CR'$.

arguments based on concentration of measures to improve the annealed result already available in the literature to an almost sure one fail whenever $p > d$, see also [6, Remark 6.5] for a discussion of the problem in the setting of strong convergence of asymptotic costs. In order to prove Theorem 1.7 we make use of the dynamical formulation of optimal transport, which allows us to combine PDE arguments together with the already existing concentration arguments for the Poisson point process, see Section 2.1.

1.1 Main steps in the proof of Theorem 1.3

We describe here the main steps in the proof of Theorem 1.3. For the detailed proofs we refer the reader to Section 2.

Following the arguments of the proof of [13, Theorem 1.1], we argue by contradiction. We show that for a locally p-optimal stationary matching T between $\{X\}$ and $\{Y\}$ we have the upper bound

$$\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X| \leq o(\ln^{\frac{1}{2}} R), \tag{1.7}$$

and the lower bound (see [13, Lemma 2.4])

$$\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X| \geq \Omega(\ln^{\frac{1}{2}} R),$$

implying the desired contradiction.

We now describe the main steps and the main differences between the proof of the upper bound (1.7) in the general case $p > 1$ and in the quadratic case. The first common step is the observation that by stationarity and ergodicity the number of Poisson points which are transported by a far distance is small in volume fraction, i. e. the following L^0-estimate on the displacement holds

$$\#\{X \in (-R, R)^d \ : \ |T(X) - X| \gg 1\} \leq o(R^d), \tag{1.8}$$

see [13, Lemma 2.1] for a precise statement. *As in the quadratic case this will be the only place where stationarity and ergodicity enter.* The next step consists on improving the ergodic estimate (1.8) to a uniform bound. As opposed to [13, Lemma 2.2] we cannot rely on the monotonicity of the map T. However, the local p-optimality of the matching T allows us to exploit the geometry of its support to improve (1.8) to

$$|T(X) - X| \leq o(R) \text{ provided that } X \in (-R, R)^d, \tag{1.9}$$

see Lemma 2.4. By concentration properties of the Poisson process we may assume that $\frac{\#\{X \in B_R\}}{|B_R|} \in [\frac{1}{2}, 2]$ for $R \gg 1$. Summing (1.9) over B_R we obtain

$$\frac{1}{R^d} \sum_{X \in B_R} |T(X) - X|^p \leq o(R^p). \tag{1.10}$$

The bound (1.10) will let us run the harmonic approximation argument already employed in [13, Lemma 2.3]. Nevertheless, in the current setting we need a *p-cost version* of the latter. By the p-harmonic approximation theorem [15, Theorem 1.1] we know that if the energy term (1.4) and the data term (1.5) are small then the displacement $T(X) - X$ is close to a p'-harmonic gradient field. By (1.10) and by exchanging the roles of $\{X\}$ and $\{Y\}$ in (1.10) we have $E_p(R) \leq o(R^p)$. On the other hand Theorem 1.7 ensures that $D_p(R) \leq O(\ln^{\frac{p}{2}} R)$. Hence, we are in a position to iteratively exploit the p-harmonic approximation result on an increasing sequence of scales to obtain that the local energy inherits the asymptotic of the data term D_p:

$$\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X|^p \leq O(\ln^{\frac{p}{2}} R), \tag{1.11}$$

see Lemma 2.6. Combining this with the L^0-estimate as in [13, Lemma 2.4] yields

$$\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X| \leq o(\ln^{\frac{1}{2}} R),$$

see Lemma 2.7.

2 Proofs

2.1 Upper bound

In this section we establish the upper bound asymptotics for the data term (1.5). The proof of Theorem 1.7 will follow from the upper bound asymptotics of the distance between the Poisson point process on a torus $[0,R)^d$ and the Lebesgue measure. Given two measures μ, ν on $\mathbb{R}^d$, we consider their projection on the torus $[0,R)^d$ and we denote by $\tilde{W}_{[0,R)^d;p}(\mu,\nu)$ the p-Wasserstein distance on the torus $[0,R)^d$ between them. Given a Borel set $\Gamma \subset [0,R)^d$, we denote by $\tilde{W}_{[0,R)^d;p,\Gamma}(\mu,\nu) = \tilde{W}_{[0,R)^d;p}(\mu \llcorner \Gamma, \nu \llcorner \Gamma)$ its restriction to Γ.

Lemma 2.1. *Let μ be a Poisson point process on the torus $[0,R)^d$ of unit intensity. There exists a constant C, a random radius $r_* < \infty$ a. s. such that for any dyadic radii $R \geq r_*$ and any $p \geq 1$*

$$\tilde{W}^p_{[0,R)^d;p}(\mu, n) \leq CR^d \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2, \\ 1 & \text{if } d \geq 3, \end{cases} \tag{2.1}$$

where $n = \frac{\mu([0,R)^d)}{R^d}$ is the (random) number density.

We shall derive Theorem 1.7 combining Lemma 2.1 with a restriction result for the data term, which will allow us to exchange the periodic Wasserstein distance with the Euclidean one.

Lemma 2.2. *For any positive measure μ on the torus $[-2R,2R)^d$ there exists a constant $C > 0$ such that*

$$\int_{R-\frac{1}{2}}^{R+\frac{1}{2}} \left(\tilde{W}^p_{[-4R,4R)^d;p,B_{\bar{R}}}(\mu, n_{\bar{R}}) + \frac{(n_{\bar{R}} - 1)^p}{n_{\bar{R}}} \right) d\bar{R} \leq C\tilde{D},$$

provided

$$\tilde{D} := \tilde{W}^p_{[-4R,4R)^d;p}(\mu, n) + \frac{(n-1)^p}{n} \ll 1,$$

where $n_{\bar{R}} = \frac{\mu(B_{\bar{R}})}{|B_{\bar{R}}|}$ and $n = \frac{\mu([-4R,4R)^d)}{(8R)^d}$ are the (random) number densities.

The latter can be derived combining the proof of [5, Lemma 2.10] and [15, Lemma 6.1].

Proof of Theorem 1.7. Step 1. Let $R \geq 1$ be an increasing sequence of approximately dyadic radii. We claim that there exists a constant C and a random radius $r_* < \infty$ a. s. such that for the fixed sequence of dyadic radii $R \geq r_*$

$$\frac{R^p}{n_\mu} |n_\mu - 1|^p \leq C \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2, \\ 1 & \text{if } d \geq 3, \end{cases} \tag{2.2}$$

and

$$\frac{R^p}{n_\nu}|n_\nu - 1|^p \le C \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2, \\ 1 & \text{if } d \ge 3. \end{cases} \tag{2.3}$$

W.l.o.g. we focus on (2.2). Indeed, (2.3) will follow from (2.4) exchanging the role of μ and ν and taking the maximum of this radius and the one pertaining ν. Since for large $R \gg 1$ we have $\frac{\ln^{\frac{p}{2}} R}{R^p} \ll 1$ (2.2) is equivalent to [7]

$$R^p|n_\mu - 1|^p \lesssim \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2, \\ 1 & \text{if } d \ge 3. \end{cases}$$

Since $n_\mu|B_R|$ is Poisson distributed with parameter $|B_R|$ by Cramér-Chernoff's bounds [3, Theorem 1] we get for $d = 2$

$$\mathbb{P}(R^p|n_\mu - 1|^p > \ln^{\frac{p}{2}} R) = \mathbb{P}(|n_\mu|B_R| - |B_R|| > CR\ln^{\frac{1}{2}} R) \lesssim \exp(-C\ln R)$$

and

$$\mathbb{P}(R^p|n_\mu - 1|^p > C) = \mathbb{P}(|n_\mu|B_R| - |B_R|| > CR^{d-1}) \lesssim \exp(-CR^{d-2}),$$

for $d \ge 3$. Finally, by a Borel-Cantelli argument (2.2) holds for the fixed sequence of approximately dyadic radii $R \ge r_*$. Step 2. We claim that there exist a constant C and a random radius $r_* < \infty$ a. s. such that for a (random) sequence of approximately dyadic radii $R \ge r_*$

$$\frac{1}{R^d}W^p_{p,B_R}(\mu, n_\mu) + \frac{1}{R^d}W^p_{p,B_R}(\nu, n_\nu) \le C \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2, \\ 1 & \text{if } d \ge 3. \end{cases} \tag{2.4}$$

By Lemma 2.1 we may assume that (2.1) holds with $[0, R)^d$ replaced by $[0, 8R)^d$. Moreover, by stationarity of the Poisson point process we may assume that there exists a random radius $r_* < \infty$ a. s. such that for any dyadic $R \ge r_*$ (2.1) holds in the form

$$\frac{1}{R^d}\tilde{W}^p_{[-4R,4R)^d;p}(\mu, \tilde{n}_\mu) \le C \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2, \\ 1 & \text{if } d \ge 3, \end{cases} \tag{2.5}$$

where $\tilde{n}_\mu = \frac{\mu([-4R,4R)^d)}{(8R)^d}$ is the (random) number density. Arguing in the same manner for ν we can deduce (possibly enlarging r_*) that

$$\frac{1}{R^d}\tilde{W}^p_{[-4R,4R)^d;p}(\nu, \tilde{n}_\nu) \le C \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2, \\ 1 & \text{if } d \ge 3, \end{cases} \tag{2.6}$$

where $\tilde{n}_\nu = \frac{\nu([-4R,4R)^d)}{(8R)^d}$ is the (random) number density. By the restriction property of Lemma 2.2 we may deduce that there exists a radius $R' \sim R$ such that

$$\begin{aligned} &W^p_{p,B_{R'}}(\mu, n_\mu) + W^p_{p,B_{R'}}(\nu, n_\nu) = \tilde{W}^p_{[-4R,4R)^d;p,B_{R'}}(\mu, n_\mu) + \tilde{W}^p_{[-4R,4R)^d;p,B_{R'}}(\nu, n_\nu) \\ &\lesssim \tilde{W}^p_{[-4R,4R)^d;p}(\mu, \tilde{n}_\mu) + \frac{R^p}{\tilde{n}_\mu}(\tilde{n}_\mu - 1)^p + \tilde{W}^p_{[-4R,4R)^d;p}(\nu, \tilde{n}_\nu) + \frac{R^p}{\tilde{n}_\nu}(\tilde{n}_\nu - 1)^p. \end{aligned} \tag{2.7}$$

Moreover, by the same argument as in Step 1 we may assume that (2.2) holds with n_μ and n_ν replaced by $\tilde{n}_\mu$ and $\tilde{n}_\nu$. Combining the latter with (2.7) and (2.5) and relabeling R' yields (2.4).

Step 3. Conclusion. Combining (2.2), (2.3) and (2.4) yields (1.6). □

[7] We use the notation $A \lesssim B$ if there exists a global constant $C > 0$, which may only depend on d, such that $A \le CB$. We write $A \sim B$ if both $A \lesssim B$ and $B \lesssim A$ hold.

Let us now turn to Lemma 2.1. In view of the dynamical formulation of optimal transport we investigate the Moser coupling [20, Appendix p. 16] between μ and n_μ. Let $(\cdot)_1$ denote a mollification on scale 1, say the convolution[8] with the standard Gaussian. Let ϕ denote the solution of the Poisson problem on $[0,R)^d$

$$-\Delta\phi = \mu_1 - \fint_{[0,R)^d} \mu_1 = \mu_1 - n. \tag{2.8}$$

We are interested in the spatially averaged p-moment of its gradient

$$F := \fint_{[0,R)^d} \frac{1}{p}|\nabla\phi|^p \quad \text{for any} \quad p < \infty. \tag{2.9}$$

We shall establish that thanks to the spatial averaging, F has good concentration properties around its expectation $\mathbb{E}F$. As we establish, the latter is $O(\ln^{\frac{p}{2}} R)$ for $d = 2$ and $O(1)$ for $d \geq 3$. What matters to us is that the probability that $F \gg \ln^{\frac{p}{2}} R$ if $d = 2$ and $F \gg 1$ if $d \geq 3$ is very small.

Lemma 2.3. *There exists a constant $C < \infty$ such that for $R \geq C$ and any $p \geq 1$,*

$$\mathbb{P}(F \geq C\ln^{\frac{p}{2}} R) \leq \frac{C}{\ln^2 R} \quad \textit{if } d = 2. \tag{2.10}$$

and

$$\mathbb{P}(F \geq C) \leq \frac{C}{R^{2d-4}} \quad \textit{if } d \geq 3. \tag{2.11}$$

An inspection of the proof reveals that the exponent 2 if $d = 2$, $2d - 4$ if $d \geq 3$, on the r. h. s. could be replaced by any exponent $< \infty$ (on which C will depend). However, it is sufficient for our purposes that the r. h. s. of (2.10) and (2.11) is summable over dyadic R, which holds for any exponent > 1.

The proof of Lemma 2.1 is a direct consequence of Lemma 2.3.

Proof of Lemma 2.1. Step 1. Definition of r_*. By Lemma 2.3 and a Borel-Cantelli argument we can deduce that there exists a constant C and a random radius $r_* < \infty$ a. s. such that for any dyadic radii $R \geq r_*$

$$\int_{[0,R)^d} |\nabla\phi|^p \leq CR^d \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2, \\ 1 & \text{if } d \geq 3, \end{cases} \tag{2.12}$$

where ϕ solves (2.8). Moreover, arguing as for (2.2) we may assume that r_* is large enough so that

$$\frac{\mu([0,R)^d)}{R^d} \in \left[\frac{1}{2}, 2\right] \quad \text{for } R \geq r_*. \tag{2.13}$$

Step 2. Proof of (2.1). By the triangle inequality and the semigroup contraction property of the Wasserstein distance we may write

$$\tilde W^p_{[0,R)^d;p}(\mu, n) \lesssim \tilde W^p_{[0,R)^d;p}(\mu, \mu_1) + \tilde W^p_{[0,R)^d;p}(\mu_1, n) \lesssim 1 + \tilde W^p_{[0,R)^d;p}(\mu_1, n). \tag{2.14}$$

We now turn to estimate the second item of the r. h. s. of (2.14). By (2.13) we have the lower bound $n \geq \frac{1}{2}$ which together with (2.14) implies (2.1) by the inequalities

$$\tilde W^p_{[0,R)^d,p}(\mu_1, n) \leq 2^{p-1}p^p \int_{[0,R)^d} |\nabla\phi|^p \overset{(2.12)}{\leq} CR^d \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2, \\ 1 & \text{if } d \geq 3. \end{cases} \tag{2.15}$$

[8] In the torus $[0,R)^d$.

The first inequality in (2.15) can be derived from the dynamical description of the transport distance, namely the Benamou-Brenier formulation of the transport distance

$$\tilde{W}_{[0,R)^2;p}(\mu_1, n) = \min\left\{ \int_0^1 \int_{[0,R)^d} \frac{|j|^p}{\rho^{p-1}} \ : \ \partial_t \rho + \nabla \cdot j = 0, \rho_0 = \mu_1, \rho_1 = n \right\}, \tag{2.16}$$

where ρ, j have to be understood as distributions on $[0,1] \times [0,R)^d$ (see for instance [17, Theorem 5.28]). Indeed, the couple (ρ, j), with $\rho = (1-t)^p \mu_1 + (1-(1-t)^p)n$ and $j = -p(1-t)^{p-1}\nabla\phi$ is an admissible candidate for (2.16) and by (2.13) we have the lower bound $\rho \geq (1-t)^p n \geq \frac{1}{2}(1-t)^p$. □

Proof of Lemma 2.3. By Jensen's inequality we can restrict ourselves to the case $p \geq 2$. By Chebyshev's inequality, it is enough to establish

$$\mathbb{E}(F - C\ln^{\frac{p}{2}} R)_+^4 \lesssim \ln^{2(p-1)} R \quad \text{if } d = 2, \tag{2.17}$$

and

$$\mathbb{E}(F - C)_+^4 \lesssim R^{4-2d} \quad \text{if } d \geq 3, \tag{2.18}$$

for some constant C. We start by ignoring the spatial averaging in (2.9) by considering

$$G = G(\mu) := \nabla\phi(X) \quad \text{for some fixed point } X$$

and establish its concentration. In fact, we shall derive a mixture of exponential and Gaussian concentration:

$$-\ln \mathbb{P}(|G - \mathbb{E}G| \geq M) \gtrsim \left\{ \begin{array}{ll} \frac{M^2}{\ln R} & \text{for } d = 2 \text{ and } M \ll \ln R \\ M & \text{otherwise} \end{array} \right\}, \tag{2.19}$$

which by the layer cake representation implies the L^p estimate in probability

$$\mathbb{E}^{\frac{1}{p}}|G - \mathbb{E}G|^p \lesssim \begin{cases} \ln^{\frac{1}{2}} R & \text{if } d = 2 \\ 1 & \text{if } d \geq 3. \end{cases} \tag{2.20}$$

By invariance of the ensemble and covariance of ϕ under reflection w. r. t. to the d Cartesian hyper-planes crossing X, we have $\mathbb{E}G = 0$, so that (2.20) sharpens to

$$\mathbb{E}|G|^p \lesssim \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2 \\ 1 & \text{if } d \geq 3. \end{cases} \tag{2.21}$$

We now turn to the argument for (2.19). The concentration principle in [21, Proposition 3.1] already applied in [13, Lemma 2.5] monitors the change $D_{X_0}G := G(\mu + \delta_{X_0}) - G(\mu)$ of G arising from adding a point at position X_0 to the point cluster. In view of (2.8), the effect is given by

$$D_{X_0}G = \nabla D_{X_0}\phi(X) \quad \text{where} \quad -\Delta D_{X_0}\phi = (\delta_{X_0})_1 - R^{-d}. \tag{2.22}$$

We learn that $D_{X_0}\phi$ is the mollified potential function for a periodic charge distribution at $X_0 + (R\mathbb{Z})^d$ with a constant background charge making the distribution overall neutral. This object is well-defined on the level of its gradient and satisfies[9]

$$|D_{X_0}G| = |\nabla D_{X_0}\phi(X)| \lesssim \big(\text{dist}(X, X_0 + (R\mathbb{Z})^d) + 1\big)^{1-d}, \tag{2.23}$$

[9] We denote by $\text{dist}(x, y + (R\mathbb{Z})^d) := \min_{k \in (R\mathbb{Z})^d} |x - y - k|$ the periodic distance on the torus $[0,R)^d$.

from which we learn, in the notation of [21, Proposition 3.1],

$$\beta := \sup_{X_0} |D_{X_0}G| \lesssim 1 \quad \text{and} \quad \alpha^2 := \int_{[0,R)^d} dX_0 |D_{X_0}G|^2 \lesssim \begin{cases} \ln R & \text{if } d=2, \\ 1 & \text{if } d \geq 3. \end{cases}$$

This implies

$$\mathbb{P}(|G - \mathbb{E}G| \geq M) \leq 2\exp\biggl(-\frac{M}{2\beta}\ln\biggl(1+\frac{\beta M}{\alpha^2}\biggr)\biggr),$$

which is easily seen to imply (2.19).

It is convenient to use a different concentration principle for (2.17). While for (2.20), we used the concentration principle for the "grand canonical ensemble" of the Poisson point process on $[0,R)^2$, for (2.17) is convenient to disintegrate this grand canonical ensemble into the "canonical ensemble" $\mathbb{E}_N$ of N i. i. d. points uniformly distributed. Note that (2.8) assumes the form

$$-\Delta\phi = \mu_1 - N \quad \text{on } [0,R)^d \quad \text{where} \quad \mu = \sum_{n=1}^{N} \delta_{X_n}, \tag{2.24}$$

where also the convolution refers to the torus $[0,R)^d$. The advantage is that we have an easy spectral gap estimate, which is in fact just the tensorization of the standard Poincaré inequality with mean value zero on $[0,R)^d$: For any suitable function $F = F(X_1, \cdots, X_N)$ we have

$$\mathbb{E}_N(F - \mathbb{E}_N F)^2 \lesssim R^2 \mathbb{E}_N \sum_{n=1}^{N} |\nabla_{X_n} F|^2.$$

Applying this inequality with F^2 playing the role F, and appealing to the Cauchy-Schwarz inequality in probability, we may upgrade this standard version to the exponent 4, which will be sufficient for our purposes:

$$\mathbb{E}_N(F - \mathbb{E}_N F)^4 \lesssim R^4 \mathbb{E}_N \biggl(\sum_{n=1}^{N} |\nabla_{X_n} F|^2\biggr)^2.$$

By the Cauchy-Schwarz inequality in N, and for our F that is invariant under permuting its argument, this yields

$$\mathbb{E}_N(F - \mathbb{E}_N F)^4 \lesssim R^4 N^2 \mathbb{E}_N |\nabla_{X_N} F|^4. \tag{2.25}$$

We now derive a suitable representation for $\nabla_{X_n} F$. From (2.9) we obtain for the partial derivative $\nabla_{X_n} F$ for our F

$$\nabla_{X_n} F = R^{-d} \int_{[0,R)^d} |\nabla\phi|^{p-2}\nabla\phi \cdot (\nabla\nabla_{X_n}\bar{\phi})_1, \qquad \text{where} \quad -\Delta\nabla_{X_n}\bar{\phi} = \nabla\delta_{X_n} \quad \text{on } [0,R)^d. \tag{2.26}$$

From the latter, we learn that $-\nabla\nabla_{X_n}\bar{\phi}(x)$ is the translation-invariant kernel, evaluated at $x - X_n$, of the Helmholtz projection, i. e. the $L^2([0,R)^d)$-orthogonal projection onto gradient fields (see [18, Theorem 2.4.9] and the discussion below). Since the latter operator is symmetric, and commutes with mollification, (2.26) can be reformulated as

$$\nabla_{X_n} F = R^{-d} \nabla u_1(X_n), \tag{2.27}$$

where $-\nabla u$ is the Helmholtz projection of $|\nabla\phi|^{p-2}\nabla\phi$, that is

$$-\Delta u = \nabla \cdot |\nabla\phi|^{p-2}\nabla\phi. \tag{2.28}$$

Inserting (2.27) into (2.25), we obtain

$$\mathbb{E}_N(F-\mathbb{E}_N F)^4 \lesssim \left(\frac{N}{R^d}\right)^2 R^{4-2d}\mathbb{E}_N|\nabla u_1(X_N)|^4. \tag{2.29}$$

In view of the Calderón-Zygmund estimate (see [19, Chapter I, II and III] for a reference on classical Calderón-Zygmund's theory) for (2.28)

$$\fint_{[0,R)^d}|\nabla u|^4 \le \fint_{[0,R)^d}|\nabla\phi|^{4(p-1)}, \tag{2.30}$$

the plan now is to pass from $\mathbb{E}_N|\nabla u_1(X_N)|^4$ to $\mathbb{E}_N \fint_{[0,R)^d}|\nabla u_1|^4$. To this purpose, we will consider ϕ' defined in (2.24) with N replaced by $N-1$, and u' defined like u in (2.28) with ϕ replaced by ϕ'. Hence provided we can

$$\text{estimate} \quad \mathbb{E}_N|\nabla u_1(X_N)|^4 \quad \text{by} \quad \mathbb{E}_{N-1}|\nabla u_1'(X_N)|^4, \tag{2.31}$$

we may proceed to capitalize on the (stochastically) independence of u' of the uniformly distributed X_N to the effect of

$$\mathbb{E}_N|\nabla u_1'(X_N)|^4 = \mathbb{E}_{N-1}\fint_{[0,R)^d}|\nabla u_1'|^4.$$

Using that $(\cdot)_1$ contracts the norm and (2.30) we obtain by shift-covariance of $\nabla\phi'$

$$\mathbb{E}_N|\nabla u_1'(X_N)|^4 \lesssim \mathbb{E}_{N-1}|\nabla\phi'|^{4(p-1)}. \tag{2.32}$$

Hence provided we may

$$\text{estimate} \quad \mathbb{E}_{N-1}|\nabla\phi'|^{4(p-1)} \quad \text{by} \quad \mathbb{E}_N|\nabla\phi|^{4(p-1)} \tag{2.33}$$

we hope to obtain from (2.29) that

$$\mathbb{E}_N(F-\mathbb{E}_N F)^4 \lesssim \left(\frac{N}{R^d}\right)^2 R^{4-2d}(\mathbb{E}_N|\nabla\phi|^{4(p-1)}+1). \tag{2.34}$$

Applying $\mathbb{E}$ to (2.34), which just means applying the Poisson distribution with mean R^d to N, and using the Cauchy-Schwarz inequality on the latter, and its good concentration property (on the level of the fourth moment), we obtain

$$\mathbb{E}(F-\mathbb{E}_N F)^4 \lesssim R^{4-2d}\big(\mathbb{E}|\nabla\phi|^{8(p-1)}+1\big)^{\frac{1}{2}}.$$

Inserting (2.21) (with p replaced by $8(p-1)$) we obtain

$$\mathbb{E}(F-\mathbb{E}_N F)^4 \lesssim \begin{cases} \ln^{2(p-1)} R & \text{if } d=2, \\ R^{4-2d} & \text{if } d\ge 3, \end{cases} \tag{2.35}$$

the major step towards (2.17).

We now turn to (2.31) and (2.33). Momentarily introducing $j(z) := |z|^{p-2}z$ we note that

$$|j(z)-j(z')| \lesssim |z'|^{p-2}|z-z'| + |z-z'|^{p-1}.$$

From (2.28) we deduce a representation of $\nabla(u-u')$ as the Helmholtz projection of $j(\nabla\phi)-j(\nabla\phi')$ on $[0,R)^d$. In view of the mollification we obtain

$$|\nabla(u-u')_1(X_N)| \lesssim \int_{[0,R)^d} \big(\mathrm{dist}(\cdot, X_N+(R\mathbb{Z})^d)+1\big)^{-d}|j(\nabla\phi)-j(\nabla\phi')|.$$

From (2.24) we obtain, cf. (2.23),

$$|\nabla(\phi-\phi')| \lesssim \big(\mathrm{dist}(\cdot, X_N+(R\mathbb{Z})^d)+1\big)^{1-d} \lesssim 1.$$

The combination of these yields (using $p-1\geq 1$)

$$\begin{aligned}&|\nabla(u-u')_1(X_N)|\\ &\lesssim \int_{[0,R)^d} \big(\mathrm{dist}(\cdot, X_N+(R\mathbb{Z})^d)+1\big)^{1-2d}(|\nabla\phi'|^{p-2}+1).\end{aligned}$$

Since by $1-2d<-d$ we have

$$\int_{[0,R)^d} \big(\mathrm{dist}(\cdot, X_N+(R\mathbb{Z})^d)+1\big)^{1-2d} \lesssim 1, \tag{2.36}$$

this implies

$$\begin{aligned}&|\nabla(u-u')_1(X_N)|^4\\ &\lesssim \int_{[0,R)^d} \big(\mathrm{dist}(\cdot, X_N+(R\mathbb{Z})^d)+1\big)^{1-2d}(|\nabla\phi'|^{4(p-2)}+1).\end{aligned}$$

Applying $\mathbb{E}_N$, using the shift covariance on the level $\mathbb{E}_{N-1}$, and once more (2.36) gives

$$\mathbb{E}_N|\nabla(u-u')_1(X_N)|^4 \lesssim \mathbb{E}_{N-1}|\nabla\phi'|^{4(p-2)}+1. \tag{2.37}$$

By the triangle inequality, (2.37) deals with (2.31); by (2.21) the additional error term is of higher order of the one already present on the r. h. s. of (2.32), and the $+1$ is of higher order. The argument for (2.33) is easier. In fact, for later purpose, we shall establish the monotonicity

$$f(N) := \mathbb{E}_N\frac{1}{p}|\nabla\phi|^p = \mathbb{E}_N F \quad \text{satisfies} \quad f(N-1)\leq f(N). \tag{2.38}$$

Appealing to the monotonicity with p replaced by $4(p-1)$ we obtain (2.33). Here comes the argument for (2.38). By convexity of $z\mapsto \frac{1}{p}|z|^p$ we have $\frac{1}{p}|z|^p \geq \frac{1}{p}|z'|^p + |z'|^{p-2}z'\cdot(z-z')$, which we use in form of

$$\frac{1}{p}|\nabla\phi|^p \geq \frac{1}{p}|\nabla\phi'|^p + |\nabla\phi'|^{p-2}\nabla\phi'\cdot\nabla(\phi-\phi').$$

Since $\nabla\phi'$ is independent of $\nabla(\phi-\phi')$, and since the expectation of the latter vanishes as we discussed above based on reflection symmetry, this implies (2.38). Hence we have completed the argument for (2.34) and thus (2.35).

It remains to post-process (2.35) to (2.17). To this purpose, we prove a partial reverse of (2.38), namely

$$f(N)\lesssim f(N') \quad \text{provided } N\leq 2N'. \tag{2.39}$$

Indeed, now we start from $\frac{1}{p}|z|^p \leq \frac{1}{p}|z'|^p + |z|^{p-2}z\cdot(z-z')$, in form of

$$\frac{1}{p}|\nabla\phi|^p \leq \frac{1}{p}|\nabla\phi'|^p + |\nabla\phi|^{p-1}|\nabla(\phi-\phi')|.$$

We now apply Hölder's inequality in probability to the effect of

$$\begin{aligned} f(N) &\le f(N') + (pf(N))^{\frac{p-1}{p}}(pf(N-N'))^{\frac{1}{p}} \\ &\overset{(2.38)}{\le} f(N') + (pf(N))^{\frac{p-1}{p}}(pf(N'))^{\frac{1}{p}} \quad \text{provided } N \le 2N', \end{aligned}$$

so that (2.39) follows from Young's inequality.

Equipped with (2.38) and (2.39) we now may pass from (2.35) to (2.17), where we use that by (2.21) and shift-covariance of $|\nabla\phi|^p$ we have

$$\mathbb{E}F \le \mathbb{E}^{\frac{1}{8}}F^8 \lesssim \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2, \\ 1 & \text{if } d \ge 3. \end{cases} \tag{2.40}$$

We fix an $N_0 \in \mathbb{N}$ with

$$N_0 \approx 2R^d,$$

so that by the concentration properties of the Poisson distribution we have thanks to the factor of 2

$$\mathbb{P}(N \ge N_0) \quad \text{is sub-algebraic in } R \text{ while} \quad \mathbb{P}\left(N \le \frac{N_0}{2}\right) \sim 1. \tag{2.41}$$

By (2.40), the first item in (2.41) transmits to

$$\begin{aligned} &\mathbb{E}I(N \ge N_0)(F - f(N_0))_+^4 \\ &\le \left(\mathbb{P}(N \ge N_0)\mathbb{E}F^8\right)^{\frac{1}{2}} \quad \text{is sub-algebraic in } R. \end{aligned} \tag{2.42}$$

Once more by (2.40), the second item in (2.41) implies by Chebyshev

$$f(N_0) \overset{(2.39)}{\lesssim} f\left(\frac{N_0}{2}\right) \lesssim \mathbb{E}[f(N)] = \mathbb{E}F \lesssim \begin{cases} \ln^{\frac{p}{2}} R & \text{if } d = 2 \\ 1 & \text{if } d \ge 3. \end{cases} \tag{2.43}$$

Now for the complementary portion to (2.42), we may appeal to (2.38) in order to connect to (2.35):

$$\mathbb{E}I(N \le N_0)(F - f(N_0))_+^4 \le \mathbb{E}(F - f(N))^4 \lesssim \begin{cases} \ln^{2(p-1)} R & \text{if } d = 2, \\ R^{4-2d} & \text{if } d \ge 3. \end{cases} \tag{2.44}$$

The desired (2.17) now follows from combining (2.42) with (2.44) and inserting (2.43). □

2.2 A L^∞-estimate

In this section we improve the L^0 estimate [13, Lemma 2.1] into a L^∞ estimate for the displacement $|T(X) - X|$. The proof of Lemma 2.4 is in the spirit of [15, Lemma 3.1], see also [7] for an L^∞ estimate in the regime $p \ge 2$ using different techniques.

Lemma 2.4. *For every $\epsilon > 0$ there exists a random radius $r_* < \infty$ a. s. such that for every $R \ge r_*$*

$$|T(X) - X| \le \epsilon R \quad \textit{provided that } X \in (-R, R)^d. \tag{2.45}$$

The proof is very similar to [13, Lemma 2.2]. As opposed to the quadratic case the support of T is not monotonic for general $p \ge 1$. To overcome this additional difficulty we need to argue by p-cyclically monotonicity to exploit the geometry of the support of the matching. To be more precise we consider a toy case. Let 0 denote the origin in $\mathbb{R}^d$ and let $e_1 = (1, 0, \dots, 0)$. We now think of the origin playing the role of a Poisson point, and e_1 being a Poisson point, close to 0, which is transported by a moderate distance, in particular we may suppose $T(e_1) = e_1$. By p-cyclically monotonicity we may write

$$|T(0)|^p \leq |T(0) - e_1|^p + 1.$$

The latter defines a constraint for $T(0)$ in which the origin is transported. Our aim is to understand the not admissible set in which 0 is transported. In particular, we show in the next lemma that the region in which the origin is not transported contains a convex set, i. e. a cone.

Lemma 2.5. *Let $1 < p < \infty$ and $\alpha \in (0, \pi)$. Define the domain $U \subset \mathbb{R}^d$*

$$U := \{x = (x_1, x') \in \mathbb{R} \times \mathbb{R}^{d-1} \mid |x|^p \geq ((x_1 - 1)^2 + |x'|^2)^{\frac{p}{2}} + 2\}.$$

Then there exists a cone with vertex $v = (\bar{x}(p, \alpha), 0)$, aperture α and axis $(1, 0, \ldots, 0)$ which lies in D.

Proof. We denote $y = |x'|$ and introduce the function $F(x, y) := (x^2+y^2)^{\frac{p}{2}} - ((x-1)^2+y^2)^{\frac{p}{2}}$.

Case $p \geq 2$. We claim that there exists a constant $c_0 > 0$ such that for $c \geq c_0$ the half-space

$$C_p = \{x : \ x_1 \geq c\} \quad \text{is contained in } U. \tag{2.46}$$

We start by noticing that by a direct calculation since $p \geq 2$, for $x \in C_p$ and $c \geq \frac{1}{2}$ we have $\partial_y F(x, y) \geq 0$. Thus for $c \geq \frac{1}{2}$, in order to ensure (2.46) it is enough to show $F(x_1, 0) \geq 2$. We note that for $x \in C_p$ and $c \geq 1$, the inequality $F(x_1, 0) \geq 2$ is satisfied if the following holds true

$$\left(1 + \frac{1}{c-1}\right)^p - 1 \geq \frac{2}{(c-1)^p}.$$

By Bernoulli's inequality, (i. e. $(1+u)^n \geq 1 + nu$ for $n \geq 0, u > -1$) the latter is satisfied if $c \geq 1 + \left(\frac{2}{p}\right)^{\frac{1}{p-1}}$. Thus, choosing $c_0 = 1 + \left(\frac{2}{p}\right)^{\frac{1}{p-1}}$ yields (2.46).

Case $1 < p \leq 2$. We claim that there exists a constant $c_0 > 0$ such that for $c \geq c_0$ the cone

$$C_p = \{x : \ y \leq \alpha(x_1 - c)\} \quad \text{is contained in } U. \tag{2.47}$$

We start by noticing that by a direct calculation since $1 < p \leq 2$, for $x \in C_p$ and $c \geq \frac{1}{2}$ we have $\partial_y F(x, y) \leq 0$. Thus it suffices to show that for $x_1 \geq c$ it holds $F(x_1, \alpha(x_1 - c)) \geq 2$. Let us denote by $g(z)$ the function $g(z) := (z^2 + \alpha^2(x_1 - c)^2)^{\frac{p}{2}}$. By the mean value theorem, there is $\xi \in [x_1 - 1, x_1]$ such that, for $1 < p \leq 2$

$$\begin{aligned}
F(x_1, \alpha(x_1 - c)) &= g(x_1) - g(x_1 - 1) \\
&= p \frac{\xi}{(\xi^2 + \alpha^2(x_1 - c)^2)^{\frac{2-p}{2}}} \\
&\geq p \frac{\xi}{(\xi^{2-p} + \alpha^{2-p}(x_1 - c)^{2-p})} \\
&\geq p \frac{x_1 - 1}{(1 + \alpha^{2-p})x_1^{2-p} + \alpha^{2-p}c^{2-p}} =: h(x_1).
\end{aligned}$$

Note that since $1 < p \leq 2$ we have that $h(x_1) \to \infty$ for $x_1 \to \infty$ and is increasing for x_1 sufficiently large, hence we can choose $c_0 < \infty$ such that $h(c_0) \geq 2$. Finally, noting that for $x \in C_p$ we have $x_1 \geq \frac{y}{\alpha} + c \geq c_0$ yields (2.47). ☐

Proof of Lemma 2.4. Step 1. Definition of $r_* = r_*(\epsilon)$ given $0 < \epsilon \ll 1$ as the maximum of three r_*'s. First, by [13, Lemma 2.1], there exists a (deterministic) length $L < \infty$ and the (random) length $r_* < \infty$ such that for $4R \geq r_*$, the number density of the Poisson points in $(-2R, 2R)^d$ transported further than the "moderate distance" L is small in the sense of

$$\#\{\, X \in (-2R, 2R)^d \,|\, |T(X) - X| > L \,\} \leq (\epsilon 4R)^d. \tag{2.48}$$

Second, by Lemma 1.7 we may also assume that r_* is so large that for $R \geq r_*$, the non-dimensionalized transportation distance of μ to its number density n is small, and that by the concentration properties of the Poisson point process $n \approx 1$, in the sense of

$$W^p_{p,(-2R,2R)^d}(\mu, n) + \frac{(4R)^{d+p}}{n}(n-1)^p \leq (\epsilon 4R)^{d+p}. \tag{2.49}$$

Third, w. l. o. g. we may assume that r_* is so large that

$$L \leq \epsilon r_*. \tag{2.50}$$

We now fix a realization and $R \geq r_*$.

Step 2.

There are enough Poisson points on mesoscopic scales. We claim that for any cube $Q \subset (-2R, 2R)^d$ of "mesoscopic" side length

$$r \gg \epsilon R \tag{2.51}$$

we have

$$\#\{\, X \in Q \,\} \gtrsim r^d. \tag{2.52}$$

Indeed, it follows from the definition of $W_{p,(-2R,2R)^d}(\mu, n)$ that for any Lipschitz function η with support in Q we have

$$\left| \int \eta d\mu - \int \eta n dy \right| \leq (\mathrm{Lip}\eta) \Big(\int_Q d\mu + n|Q| \Big)^{\frac{1}{p'}} W_{p,(-2R,2R)^d}(\mu, n).$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. Indeed, by Hölder's inequality

$$\begin{aligned} \left| \int \eta d\mu - \int \eta n dy \right| &= \left| \int (\eta(x) - \eta(y))\, d\pi(x, y) \right| \\ &\leq (\mathrm{Lip}\eta) \int |x - y| d\pi \\ &\leq (\mathrm{Lip}\eta) \Big(\int_Q d\mu + n|Q| \Big)^{\frac{1}{p'}} W_{p,(-2R,2R)^d}(\mu, n). \end{aligned}$$

We now specify to an $\eta \leq 1$ supported in Q, to the effect of $\int \eta d\mu \leq \int d\mu = \#\{\, X \in Q \,\}$, so that by Young's inequality and the trivial inequality $(x + y)^{\frac{1}{p}} \leq x^{\frac{1}{p}} + y^{\frac{1}{p}}$ for $x, y > 0, p > 1$ we have

$$\begin{aligned} \int \eta n dy \lesssim \#\{\, X \in Q \,\} &+ (\mathrm{Lip}\eta)^p W^p_{p,(-2R,2R)^d}(\mu, n) \\ &+ (\mathrm{Lip}\eta)\big(n|Q|\big)^{\frac{1}{p'}} W_{p,(-2R,2R)^d}(\mu, n). \end{aligned} \tag{2.53}$$

At the same time, we may ensure $\int_{(-2R,2R)^d} \eta \gtrsim r^d$ and $\mathrm{Lip}\eta \lesssim r^{-1}$, so that by (2.49), which in particular ensures $n \approx 1$, (2.53) turns into

$$r^d \lesssim \#\{\, X \in Q \,\} + r^{-p}(\epsilon R)^{d+p} + r^{\frac{d}{p'} - 1}(\epsilon R)^{\frac{d}{p} + 1}.$$

Thanks to assumption (2.51) we obtain (2.52).

Step 3. Iteration. At mesoscopic distance around a given point $X \in (-R,R)^d$, there are sufficiently many Poisson points that are transported only over a moderate distance in any direction. More precisely, we claim that for any cube $Q \subset (-2R,2R)^d$ of side-length satisfying (2.51) we have

$$\text{there exists } X \in Q \text{ with } |T(X) - X| \le L. \tag{2.54}$$

We suppose that (2.54) were violated for some cube Q. By (2.52), there are $\gtrsim r^d$ of such points. By assumption (2.51), there are thus $\gg (\epsilon R)^d$ Poisson points in $(-2R,2R)^d$ that get transported by a distance $> L$, which contradicts (2.48).

Step 4. Building barriers. We show that for any Poisson point X and any unitary vector e, if we are given a cube Q^X with barycenter X of side-length satisfying (2.51) and $X' \in Q + 2re$ such that

$$|T(X') - X'| \le L$$

there exists a cone $C_{X,X'}$ with vertex $X + r\rho\frac{X'-X}{|X'-X|}$ for some finite constant $\rho = \rho(p,d) > 0$, aperture 1 and axis $\frac{X'-X}{|X'-X|}$ such that $T(X) \notin C_{X,X'}$. Indeed, by p-cyclically monotonicity of T we get

$$|T(X) - X|^p \le |T(X) - X|^p + |T(X') - X'|^p \le |T(X') - X|^p + |T(X) - X'|^p.$$

By a change of coordinates we may assume that $X = 0$ and $X' = (\tau, 0)$ with $\tau \in (\frac{3}{2}r, \frac{\sqrt{26}}{2}r)$. In particular, writing $T(X) = (y_0, y_1)$ with $y_0 \in \mathbb{R}$ and $y_1 \in \mathbb{R}^{d-1}$ we get

$$|T(X)|^p \le (L+\tau)^p + (|y_0 - \tau|^2 + |y_1|^2)^{\frac{p}{2}}.$$

Introduce $(\tilde{y}_0, \tilde{y}_1) = \frac{1}{\tau}(y_0, y_1)$. Then we have

$$\begin{aligned} |\tilde{y}|^p &\le (|\tilde{y}_0 - 1|^2 + |\tilde{y}_1|^2)^{\frac{p}{2}} + \left(\frac{L}{\tau} + 1\right)^p \\ &\overset{(2.50),\,(2.51)}{\le} (|\tilde{y}_0 - 1|^2 + |\tilde{y}_1|^2)^{\frac{p}{2}} + 2. \end{aligned}$$

In particular by Lemma 2.5 there exists (going back to the original coordinates) a cone $C_{X,X'}$ with vertex $X + r\rho\frac{X'-X}{|X'-X|}$, aperture 1 and axis $\frac{X'-X}{|X'-X|}$ such that $T(X) \notin C_{X,X'}$.

Step 5. All Poisson points are transported over distances $\ll R$. We claim that for all Poisson points X

$$|T(X) - X| \lesssim \epsilon R \quad \text{provided } X \in (-R,R)^d. \tag{2.55}$$

Choosing $c(d,p)$ directions e_i, we get points $\{X_i\}_{i=1}^{c(d)}$ with $X_i \in Q^{X_i} + 2re_i$ for some finite constants $\rho_i > 0$, cones C_{X,X_i} with vertexes $X + r\rho_i\frac{X_i - X}{|X_i - X|}$, aperture 1 and axes $\frac{X_i - X}{|X_i - X|}$ and a finite constant $\rho_i \le \bar{\rho} < \infty$ for every i such that

$$T(X) \notin \bigcup_{i=1}^{c(d)} C_{X,X_i} \quad \text{and} \quad \mathbb{R}^d \setminus B_{\bar{\rho}r}(X) \subset \bigcup_{i=1}^{c(d)} C_{X,X_i}.$$

In particular,

$$|T(X) - X| \le \bar{\rho}r \lesssim r.$$

Since (2.51) was the only constraint on r, we obtain (2.55). □

2.3 Application of the p-harmonic approximation theorem

Lemma 2.6. *Let $p > 1$. There exist a constant C and a random radius $r_* < \infty$ a. s. such that for every $R \ge r_*$ we have*

$$\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X|^p \le C \ln^{\frac{p}{2}} R. \tag{2.56}$$

Proof. The proof relies on the p-harmonic approximation result [15, Theorem 1.1]. This result establishes that for any $0 < \tau \ll 1$, there exists an $\epsilon > 0$, a constant $C > 0$ (which does not depend on τ) and a constant $C_\tau < \infty$ such that provided for some R

$$\frac{1}{R^p} E_p(4R) + \frac{1}{R^p} D_p(4R) \le \epsilon \tag{2.57}$$

there exists a p-harmonic gradient field $\nabla \Phi$ such that

$$\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} \left| T(X) - X - |\nabla\Phi(X)|^{p'-2} \nabla\Phi(X) \right|^p \le \tau E_p(4R) + C_\tau D_p(4R), \tag{2.58}$$

and[10]

$$\sup_{B_{2R}} |\nabla \Phi|^{\frac{p}{p-1}} \le C \left(E_p(4R) + D_p(4R) \right), \tag{2.59}$$

where p' is the conjugate exponent of p. The fraction τ will be chosen at the end of the proof.

Step 1. Definition of r_* depending on τ. For $0 < \tau \ll 1$ let $\epsilon = \epsilon(\tau)$ be as above. By Theorem 1.7 we may assume that r_* is large enough so that for a (random) sequence of approximately dyadic radii $R \ge r_*$

$$D_p(R) \le C \ln^{\frac{p}{2}} R. \tag{2.60}$$

Moreover, by Theorem 1.7 possibly enlarging r_* and we may assume that for a (random) sequence of approximately dyadic radii[11]

$$\frac{D_p(4R)}{R^p} \le \frac{\epsilon}{2}. \tag{2.61}$$

Note that only the bound (2.60) is specific to $d = 2$. Moreover, the estimate (2.61) is not sharp, but it is enough for our purpose. From now on, we restrict ourselves to the sequence of approximately dyadic radii R coming from (2.60) and (2.61), which we may do w. l. o. g. for (2.56). Note that by the bound on $D_p(4R)$ in (2.61) and the second and fourth term in the definition of $D_p(R)$ in (1.5)

$$\#\left(\{X \in B_R\} \cup \{T(X) \in B_R\} \right) \le C R^d. \tag{2.62}$$

Moreover, we may assume that r_* is large enough so that (2.45) holds. Since $B_{4R} \subset (-4R, 4R)^d$ we may sum (2.45) over B_R to obtain for $R \ge r_*$

$$\frac{1}{(4R)^d} \sum_{X \in B_{4R}} |T(X) - X|^p \le \frac{\epsilon}{4} R^p.$$

[10] Note that in [13, (2.21)] the constants on the right hand side were both labeled C_τ. However, the constant which controls $\sup_{B_{2R}} |\nabla\Phi|^2$ in [13, (2.21)] does not depend on τ. This is important in the proof of (2.68). To avoid confusion we labeled the constants in (2.58) and (2.59) differently.

[11] By the footnote of [15, Theorem 1.1] it suffices that (2.57) is satisfied for a sequence of radii $4R' \sim 4R$, thus we do not rename the sequence in (2.61).

By symmetry, potentially enlarging r_*, we may also assume that (2.45) holds with X replaced by $T(X)$ so that for $R \geq r_*$ both

$$T(X) \in B_R \Rightarrow X \in B_{2R} \tag{2.63}$$

and

$$\frac{1}{(4R)^d} \sum_{T(X) \in B_{4R}} |T(X) - X|^p \leq \frac{\epsilon}{4} R^p,$$

thus

$$E_p(4R) = \frac{1}{(4R)^d} \sum_{X \in B_{4R} \text{ or } T(X) \in B_{4R}} |T(X) - X|^p \leq \frac{\epsilon}{2} R^p, \tag{2.64}$$

and in particular (2.57) holds. Finally by [13, Lemma 2.1] we may assume, possibly enlarging r_*, that there exists a deterministic constant L_τ and for $R \geq r_*$ we both have

$$\#\left(\{X \in Q_R \mid |T(X) - X| > L_\tau\} \cup \{T(X) \in Q_R \mid |T(X) - X| > L_\tau\}\right) \leq \tau R^d. \tag{2.65}$$

and

$$L_\tau^p \leq \ln^{\frac{p}{2}} R. \tag{2.66}$$

Step 2. Application of harmonic approximation. For all $R \geq r_*$

$$E_p(R) \leq \tau E_p(32R) + C_\tau \ln^{\frac{p}{2}} R. \tag{2.67}$$

We start by showing that for the (random) sequence of approximately dyadic radii of Step 1 it holds for $R \geq r_*$

$$E_p(R) \leq \tau E_p(4R) + C_\tau \ln^{\frac{p}{2}} R. \tag{2.68}$$

We split the sum according to whether the transportation distance is moderate or large. On the latter we use the harmonic approximation:

$$\begin{aligned}
&\frac{1}{R^d} \sum_{(X \in B_R \text{ or } T(X) \in B_R) \text{ and } |T(X)-X| > L_\tau} |T(X) - X|^p \\
&\leq \frac{2^p}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X - |\nabla\Phi(X)|^{p'-2} \nabla\Phi(X)|^p \\
&+ \frac{2^p}{R^d} \sum_{(X \in B_R \text{ or } T(X) \in B_R) \text{ and } |T(X)-X| > L_\tau} |\nabla\Phi(X)|^{\frac{p}{p-1}} \\
&\overset{(2.65)}{\leq} 2^p \left(\tau E_p(4R) + C_\tau D_p(4R)\right) + 2^p \tau \sup_{B_R} |\nabla\Phi|^2 \\
&\overset{(2.58),\,(2.59),\,(2.63)}{\leq} 2^p \tau E_p(4R) + 2^p C_\tau D(4R) + 2^p \tau C (E_p(4R) + D_p(4R)) \\
&= 2^p \tau (1 + C) E_p(4R) + 2^p (C_\tau + \tau C) D_p(4R).
\end{aligned}$$

The last estimate combines to

$$\begin{aligned}
&\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} |T(X) - X|^p \\
&= \frac{1}{R^d} \sum_{(X \in B_R \text{ or } T(X) \in B_R) \text{ and } |T(X)-X| \leq L_\tau} |T(X) - X|^p \\
&+ \frac{1}{R^d} \sum_{(X \in B_R \text{ or } T(X) \in B_R) \text{ and } |T(X)-X| > L_\tau} |T(X) - X|^p \\
&\overset{(2.62)}{\leq} C L_\tau^p + 2^p \tau (1 + C) E_p(4R) + 2^p (C_\tau + \tau C) D_p(4R) \\
&\overset{(2.60),\,(2.66)}{\leq} 2\tau (1 + C) E_p(4R) + (2^p (C_\tau + \tau C) + C) \ln^{\frac{p}{2}} R.
\end{aligned}$$

Relabeling τ and C_τ, this implies (2.68). Let R be any radius $R \geq r_*$. There exists a dyadic radius $R' \geq R$ satisfying (2.68) so that[12]

$$E_p(R) \leq CE_p(R') \leq C\tau E_p(4R') + CC_\tau \ln^{\frac{p}{2}} R' \leq C\tau E_p(32R) + CC_\tau \ln^{\frac{p}{2}} R,$$

for some constant $C > 0$. Relabeling τ and C_τ yields (2.67).

Step 3. Iteration. Iterating (2.67), we obtain for any $k \geq 1$

$$\begin{aligned} E_p(R) &\leq \tau E_p(32R) + C_\tau \ln^{\frac{p}{2}} R \\ &\leq \tau^2 E_p(32^2 R) + \tau C_\tau \ln^{\frac{p}{2}} R + C_\tau \ln^{\frac{p}{2}} R \\ &\leq \tau^k E_p(32^k R) + C_\tau \sum_{l=0}^{k-1} \tau^l \ln^{\frac{p}{2}} R \\ &\overset{(2.57)}{\leq} \epsilon \left(32^p \tau\right)^k R^p + C_\tau \sum_{l=0}^{k-1} \tau^l \ln^{\frac{p}{2}} R. \end{aligned}$$

We now fix τ such that $32^p \tau < 1$ to the effect of

$$E_p\left(R\right) \leq C \sum_{l=0}^{\infty} \tau^l \ln^{\frac{p}{2}} R \leq C \ln^{\frac{p}{2}} R. \qquad \square$$

2.4 Trading integrability against asymptotics

Lemma 2.7. *Let $p > 1$. For every $\epsilon > 0$ there exists a random radius $r_* < \infty$ a. s. such that*

$$\frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} \left|T\left(X\right) - X\right| \leq \epsilon \ln^{\frac{1}{2}} R.$$

Proof. By Lemma 2.6, we know that there exists a random radius r_* such that for $R \geq r_*$ we have

$$E_p\left(R\right) = \frac{1}{R^d} \sum_{X \in B_R \text{ or } T(X) \in B_R} \left|T\left(X\right) - X\right|^p \leq C \ln^{\frac{p}{2}} R. \tag{2.69}$$

Let $0 < \epsilon \ll 1$. Possibly enlarging r_*, we may also assume by Lemma [13, Lemma 2.1] that there exists a deterministic constant L such that for $R \geq r_*$

$$\#\left(\left\{X \in B_R \;\middle|\; \left|T\left(X\right) - X\right| > L\right\} \cup \left\{T(X) \in B_R \;\middle|\; \left|T\left(X\right) - X\right| > L\right\}\right) \leq \epsilon R^d. \tag{2.70}$$

Furthermore, note that by Lemma 1.7 and the second and fourth term in the definition of $D_p\left(R\right)$ in (1.5) we may also assume possibly enlarging r_* again that for $R \geq r_*$ (2.62) holds. Finally, we may also assume possibly enlarging r_* that for $R \geq r_*$

$$L \leq \epsilon^{\frac{1}{p'}} \ln^{\frac{1}{2}} R. \tag{2.71}$$

We split again the sum into moderate and large transportation distance and apply

[12] Note that if $R_1 < R_2$ are two consecutive approximately dyadic radii, by definition there exists a dyadic radius R such that $R_1 \sim R$ and $R_2 \sim 2R$. In particular, $\frac{R_2}{R_1} < 8$.

Hölder's inequality:

$$\begin{aligned}\frac{1}{R^d}\sum_{X\in B_R \text{ or } T(X)\in B_R}|T(X)-X| &\leq \frac{1}{R^d}\sum_{(X\in B_R \text{ or } T(X)\in B_R) \text{ and } |T(X)-X|\leq L}|T(X)-X| \\ &+\frac{1}{R^d}\sum_{(X\in B_R \text{ or } T(X)\in B_R) \text{ and } |T(X)-X|> L}|T(X)-X| \\ &\overset{(2.62),\ (2.69),\ (2.70)}{\leq} CL+\epsilon^{\frac{1}{p'}}E_p(R)^{\frac{1}{p}} \\ &\overset{(2.71)}{\leq} C\epsilon^{\frac{1}{p'}}\ln^{\frac{1}{2}}R.\end{aligned}$$

Relabeling ϵ proves the claim. □

References

[1] M. Ajtai, J. Komlós, and G. Tusnády. On optimal matchings. *Combinatorica*, 4:259–264, 1984. MR0779885

[2] F. Barthe and C. Bordenave. Combinatorial optimization over two random point sets. In *Séminaire de probabilités XLV*, pages 483–535. Cham: Springer, 2013. MR3185927

[3] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. MR3185193

[4] S. Chatterjee, R. Peled, Y. Peres, and D. Romik. Gravitational allocation to Poisson points. *Ann. Math.*, pages 617–671, 2010. MR2680428

[5] M. Goldman, M. Huesmann, and F. Otto. Quantitative linearization results for the Monge-Ampère equation. *Com. Pure and Appl. Math*, 2021. MR4373162

[6] M. Goldman and D. Trevisan. Convergence of asymptotic costs for random Euclidean matching problems. *Probab. Math. Phys.*, 2020. MR4408015

[7] C. E. Gutiérrez and A. Montanari. L^∞-estimates in optimal transport for non quadratic costs. *Calc. Var. Partial Differ. Equ.*, 61(5):163, Jun 2022. MR4444176

[8] C. Hoffman, A. E. Holroyd, and Y. Peres. A stable marriage of Poisson and Lebesgue. *Ann. Probab.*, 34(4):1241–1272, 2006. MR2257646

[9] A. E. Holroyd. Geometric properties of Poisson matchings. *Probab. Theory Relat. Fields*, 150(3):511–527, 2011. MR2824865

[10] A. E. Holroyd, S. Janson, and J. Wästlund. Minimal matchings of point processes. *Probab. Theory Relat. Fields*, 184(1):571–611, Oct 2022. MR4498517

[11] A. E. Holroyd, R. Pemantle, Y. Peres, and O. Schramm. Poisson matching. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 45(1):266–287, 2009. MR2500239

[12] M. Huesmann. Optimal transport between random measures. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 52(1):196 – 232, 2016. MR3449301

[13] M. Huesmann, F. Mattesini, and F. Otto. There is no stationary cyclically monotone Poisson matching in 2d. *Probab. Theory Relat. Fields*, August 2023. MR4664582

[14] M. Huesmann and K.-T. Sturm. Optimal transport from Lebesgue to Poisson. *Ann. Probab.*, 41(4):2426–2478, 2013. MR3112922

[15] L. Koch. Geometric regularisation for optimal transport with strongly p-convex cost *Calc. Var. Partial Differ. Equ.* , 63(4): 87, 2024 MR4728221

[16] R. Markó and Á. Timár. A Poisson allocation of optimal tail. *Ann. Probab.*, 44(2):1285–1307, 2016. MR3474472

[17] F. Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling. MR3409718

[18] G. Schwarz. *Hodge Decomposition – A Method for Solving Boundary Value Problems*. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2006. MR1367287

[19] E. M. Stein. *Singular Integrals and Differentiability Properties of Functions (PMS-30)*. Princeton University Press, 1970. MR0290095

[20] C. Villani. *Optimal transport: old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2009. MR2459454

[21] L. Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probab. Theory Relat. Fields*, 118(3):427–438, 2000. MR1800540