



**HAL**  
open science

# On the Wasserstein distance between a hyperuniform point process and its mean

Raphael Butez, Sandrine Dallaporta, David García-Zelada

► **To cite this version:**

Raphael Butez, Sandrine Dallaporta, David García-Zelada. On the Wasserstein distance between a hyperuniform point process and its mean. 2024. hal-04544006v2

**HAL Id: hal-04544006**

**<https://hal.science/hal-04544006v2>**

Preprint submitted on 19 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the Wasserstein distance between a hyperuniform point process and its mean

Raphael Butez, Sandrine Dallaporta, David García-Zelada

July 18, 2024

## Abstract

We study the average  $p$ -Wasserstein distance between a finite sample of an infinite hyperuniform point process on  $\mathbb{R}^2$  and its mean for any  $p \geq 1$ . The average Wasserstein transport cost is shown to be bounded from above and from below by some multiples of the number of points. More generally, we give a control on the  $p$ -Wasserstein distance in function of a control on the  $L^p$  norm of the difference of the point process and its mean. We also obtain the  $d$ -dimensional version of this result.

## 1 Introduction

Evaluating the transport cost between a point process and its mean, or between two independent copies of the same point process, which are two very closely related questions, have a long and rich history. The notion of transport cost involves the Wasserstein distance (Kantorovich-Wasserstein distance)  $W_p$  whose definition is recalled in the next section. One of the seminal papers on this question is the work of Ajtai, Komlós and Tusnády [AKT84] in the unit square whose generalization for  $d$ -dimensional cubes says that, for a positive integer  $d$  and for a fixed  $p \geq 1$ , there exists a constant  $C_{d,p} > 0$  such that, if  $X_1, \dots, X_N$  is an independent sequence of uniformly distributed random variables on  $[0, 1]^d$  and if  $\mu_N = \frac{1}{N} \sum_{i=1}^N \delta_{X_i}$  is its empirical measure and  $\mu$  is the uniform measure on  $[0, 1]^d$ , then

$$\mathbb{E}[W_p^p(\mu_N, \mu)]^{1/p} \leq C_{d,p} \begin{cases} N^{-1/2} & \text{if } d = 1 \\ N^{-1/2} \sqrt{\log N} & \text{if } d = 2 \\ N^{-1/d} & \text{if } d \geq 3 \end{cases} \quad (1)$$

and the order is optimal, i.e., there exists a matching lower bound. This statement can be found in the work of Ledoux [Led19, Equation (7)]. The same holds for uniformly distributed independent samples on compact Riemannian manifolds [Led19, Section 4]. Generalizations to other i.i.d. samples and cost functions include [Tal94, DY95, BG14, FG15, Led19]. The first terms of an asymptotic expansion have been obtained in [AST19, GT21, AGT22, GH018].

By decomposing the unit square as a square grid and considering a point at each vertex, we obtain the optimal order of convergence towards the uniform measure. It is quite natural to consider this square grid decomposition as a typical “uniform” set of points. From this point of view, the above paragraph tells us that uniformly distributed i.i.d. points are not so “uniform” after all. The question arises of why is that so. It seems that some interaction between the points is needed for them to be closer to this “uniform” grid.

Can we find conditions on two-dimensional point processes for which the Wasserstein distance is of the optimal order? When the points are not independent and identically distributed, only few results about convergence rates are available. To our knowledge, the first result available is from Meckes and Meckes [MM15] who obtained bounds on the expected  $W_p$  distance for the empirical measure of the Ginibre ensemble. In the case of two-dimensional Coulomb gases, ideas from Chafaï, Hardy and Maïda [CHM18] combined with the asymptotic expansion of the partition function stated in the work of Sandier and Serfaty [SS15] give a bound

$$\mathbb{E}[W_1(\mu_N, \mu)] \leq \frac{C}{\sqrt{N}} \quad (2)$$

where  $\mu_N$  is the empirical measure of the Coulomb gas and  $\mu$  its equilibrium measure. The same bound was obtained by Carroll, Marzo, Massaneda and Ortega-Cerdà in [CMMOC18] for a class of  $\beta$ -ensembles on compact manifolds. In his PhD thesis, Prod'homme [Pro21] obtained a similar result for the  $W_2$  distance for the Ginibre ensemble. Furthermore, Jallowy [Jal23] gave general bounds for the  $W_p$  distance with a family of random matrices improving the bounds given by O'Rourke and Williams [OW23].

Recently, Lachièze-Rey and Yogeshwaran [LRY24] studied the transportation problem for hyperuniform point processes. They obtain optimal transport cost for  $p$ -Wasserstein distances for translation invariant point processes with some conditions on the pair correlation function and a control of the integrable reduced pair correlation function. This project was carried independently to ours and, despite the overlap of the results, use totally different techniques. In addition, Leblé and Huesmann obtained in [HL24] the almost equivalence between hyperuniformity and the existence of Wasserstein bounds in dimension 2, to the cost of a much more technical approach.

## 2 Background

### 2.1 Space of measures and point processes

The information of a configuration of points on  $\mathbb{R}^d$  will be captured by its counting measure. More precisely, the information of a locally finite multiset  $X \subset \mathbb{R}^d$  is contained in the locally finite measure  $\mu = \sum_{x \in X} \delta_x$  and the observables of interest are the number of points on sets  $B$  which can be nicely described as  $\mu(B)$ . This motivates the following definition.

Consider the set  $\mathcal{M}$  of locally finite measures on  $\mathbb{R}^d$  and define, for every bounded measurable set  $B \subset \mathbb{R}^d$ , the function  $\pi_B : \mathcal{M} \rightarrow [0, \infty)$  by  $\pi_B(\mu) = \mu(B)$ . We endow the set  $\mathcal{M}$  with the  $\sigma$ -algebra generated by the functions  $\pi_B$  for every bounded measurable  $B$ . Moreover, the set of point configurations  $\mathcal{C}$  is defined as the measurable set

$$\mathcal{C} = \{X \in \mathcal{M} : X(B) \in \mathbb{N} \text{ for every bounded measurable } B\}.$$

In this way, the map that associates a locally finite multiset  $\tilde{X} \subset \mathbb{R}^d$  to the point configuration  $X = \sum_{x \in \tilde{X}} \delta_x \in \mathcal{C}$  is a bijection. For a measurable subset  $A \subset \mathbb{R}^d$ , the set of locally finite measures  $\mathcal{M}_A$  on  $A$  can be seen as a measurable subset of  $\mathcal{M}$  while the set of point configurations  $\mathcal{C}_A$  on  $A$  can be seen as a measurable subset of  $\mathcal{C}$ . A random measure on  $A$  is a random element of  $\mathcal{M}_A$  and a point process on  $A$  is a random element of  $\mathcal{C}_A$ . If  $\mu$  is a random measure on  $A$ , we can define its expected value  $\mathbb{E}[\mu]$  as the measure on  $A$  that satisfies

$$\mathbb{E}[\mu](B) = \mathbb{E}[\mu(B)] \text{ for every measurable } B \subset A.$$

A particular kind of point processes used in Proposition 8 is the Hermitian determinantal point processes. We define here the particular case of such processes on  $\mathbb{R}^d$  with respect to Lebesgue measure. Let  $K : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{C}$  be a measurable function such that  $K(x, y) = \overline{K(y, x)}$  for every  $x, y \in \mathbb{R}^d$ . We say that the point process  $X$  on  $\mathbb{R}^d$  is determinantal with kernel  $K$  if  $X(\{x\}) \leq 1$  for all  $x \in \mathbb{R}^d$  and for every  $B_1, \dots, B_k \subset \mathbb{R}^d$  pairwise disjoint measurable subsets

$$\mathbb{E}[X(B_1) \dots X(B_k)] = \int_{B_1 \times \dots \times B_k} \det(K(x_i, x_j)_{1 \leq i, j \leq k}) dx_1 \dots dx_k.$$

## 2.2 Wasserstein distance

For  $p \geq 1$  we define the Wasserstein distance (or Kantorovich distance as remarked, for instance, in [Cha16] or as explained in [Ver13, Page 8]) between two finite positive measures  $\mu$  and  $\nu$  on  $\mathbb{R}^d$  of the same mass by

$$W_p(\mu, \nu) = \left( \inf_{\Pi} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|^p d\Pi(x, y) \right)^{1/p} \in [0, \infty],$$

where the infimum is taken over all positive measures  $\Pi$  on  $\mathbb{R}^d \times \mathbb{R}^d$  with first marginal  $\mu$  and second marginal  $\nu$  (couplings of  $\mu$  and  $\nu$ ) or, more precisely, such that  $\Pi(A \times \mathbb{R}^d) = \mu(A)$  and  $\Pi(\mathbb{R}^d \times A) = \nu(A)$  for every measurable subset  $A \subset \mathbb{R}^d$ . Since those  $\Pi$  must have the same mass as  $\mu$  and  $\nu$ , the requirement of  $\mu$  and  $\nu$  having the same mass is essential for this definition to make sense. In this way,  $W_p$  defines a distance on the set of finite positive measures of a fixed mass. Moreover, we may notice that  $W_p$  is compatible with the measurable structure on the measurable set of finite measures  $\mathcal{M}_f \subset \mathcal{M}$ , i.e.,  $W_p$  is a measurable function on the measurable set  $\{(\mu, \nu) \in \mathcal{M}_f \times \mathcal{M}_f : \mu(\mathbb{R}^d) = \nu(\mathbb{R}^d)\}$ .

Our main goal in this work is to use some hyperuniformity conditions to obtain optimal rates for the expected value of the Wasserstein distance. Since the notion of hyperuniformity (defined below) typically applies to point processes and not to sequences of empirical measures, we would like to translate the result stated in the introduction to a different scale. To achieve this, we first notice the following. Consider two measures  $\mu$  and  $\nu$  on  $\mathbb{R}^d$  with the same mass and two positive constants  $\lambda, N > 0$ . Define  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  by  $f(x) = \lambda x$ . Then, by using that  $\Pi$  is a coupling of  $\mu$  and  $\nu$  if and only if  $N(f \times f)_* \Pi$  is a coupling of  $Nf_* \mu$  and  $Nf_* \nu$ , we get

$$W_p^p(Nf_* \mu, Nf_* \nu) = N \lambda^p W_p^p(\mu, \nu).$$

With this in mind, Equation (1) would read as follows. For a positive integer  $d$  and for a fixed real number  $p \geq 1$ , there is a constant  $C_{d,p} > 0$  such that, if  $X_1, \dots, X_N$  is an independent sequence of uniformly distributed random variables on  $[0, N^{1/d}]^d$ , and if we let  $\mu_N = \sum_{i=1}^N \delta_{X_i}$  and consider the Lebesgue measure  $\mu$  on  $[0, N^{1/d}]^d$ ,

$$\mathbb{E}[W_p^p(\mu_N, \mu)] \leq C_{d,p} \begin{cases} N^{1+p/2} & \text{if } d = 1 \\ N(\log N)^{p/2} & \text{if } d = 2. \\ N & \text{if } d \geq 3 \end{cases} \quad (3)$$

This setting, where the important object is a point process (a ‘‘counting’’ measure), can be seen as the *microscopic* setting while the rescaled version considered in the introduction, where the empirical measure is the main object, can be thought of as the *macroscopic* setting. We have decided to write our main result Theorem 2 in the microscopic setting in which we see a linear upper bound  $\mathbb{E}[W_p^p(\mu_N, \mu)] \leq CN$ .

A standard argument gives a lower bound for the distance between a uniform measure on a finite set and an absolutely continuous measure with respect to Lebesgue measure; see the following lemma. For the sake of completeness, we include the proof in the appendix.

**Lemma 1** (Deterministic lower bound). *There exists a constant  $\alpha_d > 0$  such that, for any  $n \geq 1$ , any  $x_1, \dots, x_n \in \mathbb{R}^d$  and any measure  $\mu$  on  $\mathbb{R}^d$  of mass  $n$  with density with respect to Lebesgue measure bounded from above by a constant  $A > 0$ , we have that for every  $p \geq 1$*

$$W_p^p\left(\sum_{i=1}^n \delta_{x_i}, \mu\right) \geq (\alpha_d)^p \frac{n}{A^{p/d}}.$$

## 2.3 Hyperuniformity

In this subsection we give the standard notion of hyperuniformity which is related to the hypotheses of our results in dimension 2. Although there is a standard definition of hyperuniformity in the literature for general dimension, it does not seem to be the right notion to keep in mind for Theorem 4 where much less than the standard hyperuniformity is needed. So, we will only focus on dimension 2 and refer to [Tor18, Equation (97)] for general dimension. A hyperuniform point process on  $\mathbb{R}^2$  is usually defined as a point process  $X$  on  $\mathbb{R}^2$  whose law is invariant under translations and for which

$$\lim_{R \rightarrow +\infty} \frac{\text{Var}(X(D(0, R)))}{|D(0, R)|} = 0.$$

Here  $D(0, R)$  denotes the disk of radius  $R$  centered at 0. There is a vast literature on hyperuniform point processes and we refer to the surveys of Coste [Cos21] and Torquato [Tor18] on this topic. The usual motivation for this definition is to think of them as “more uniform” or more ordered than stationary Poisson point processes for which  $\text{Var}(X(D(0, R)))$  is proportional to  $|D(0, R)|$ . The Ginibre point process or the zeros of the flat Gaussian analytic function are both hyperuniform. Intuitively, the points of a hyperuniform point process are well spread out in space, with no cluster or empty spots, which makes those processes easier to match to a continuous measure and good candidates to match the optimal transport cost. There are three classes of specific hyperuniform point processes which depend on the decay rate of  $\text{Var}(X(D(0, R)))$ .

$$\begin{cases} \text{Type I} & \text{if: } \text{Var}(X(D(0, R))) \lesssim |D(0, R)|^{1/2}, \\ \text{Type II} & \text{if: } \text{Var}(X(D(0, R))) \lesssim |D(0, R)|^{1/2} \log R, \\ \text{Type III} & \text{if: } \text{Var}(X(D(0, R))) \lesssim |D(0, R)|^{1-\varepsilon}. \end{cases}$$

Here,  $\lesssim$  means that the left-hand side is bounded by a constant times the right-hand side.

There exist hyperuniform point processes which do not fall in any of those categories, as explained by [DFHL24, Theorem 3], but these three classes give at least an idea of what is expected for many models.

In this work, we will consider squares instead of disks when taking the variance. There exist some point processes which are hyperuniform with disks and not hyperuniform with squares, but there exist general conditions which ensure that a point process is hyperuniform no matter the form of the window, see [Cos21]. On the other hand, our point processes do not need to be defined on the whole plane but only on squares, so that invariance under translations is not really a meaningful condition here.

### 3 Results

We start by stating a result on random measures on dimension 2 and then we give its  $d$ -dimensional version. Later, we give a corollary for a  $p$ -th moment version of type III hyperuniform point process. Finally, we see that usual type III hyperuniform Hermitian determinantal point processes satisfy our hypotheses for every finite  $p$ .

**Theorem 2.** Fix  $p \geq 1$  and consider a real number  $N > 0$ . Let  $\mu_N$  be a random measure on the square  $Q_N = [0, \sqrt{N}]^2$  such that the following conditions hold.

1. There exists a monotone function (non-decreasing or non-increasing)  $g : [1, \infty) \rightarrow (0, \infty)$  such that, for any square  $B \subset Q_N$  that satisfies  $|B| \geq 1$ , it holds that

$$\mathbb{E}[|\mu_N(B) - \mathbb{E}[\mu_N(B)]|^p] \leq |B|^{p/2} g(|B|)^p. \quad (4)$$

2. There exist two constants  $a \in (0, 1]$  and  $A \in [1, \infty)$  such that for any square  $B \subset Q_N$

$$a|B| \leq \mathbb{E}[\mu_N(B)] \leq A|B|. \quad (5)$$

This is equivalent to requiring that  $\mathbb{E}[\mu_N]$  has a density with respect to the Lebesgue measure bounded from above by  $A$  and from below by  $a$ .

Let  $M_N = \mu_N(Q_N)$  and define the random measure  $\bar{\mu}_N = \frac{M_N}{\mathbb{E}[M_N]} \mathbb{E}[\mu_N]$  of total mass  $M_N$ . Then there exists a constant  $C_p$  (it only depends on  $p$ ) such that

$$\mathbb{E}[W_p^p(\mu_N, \bar{\mu}_N)] \leq (C_p a^{1-2p} A^p) N \left( 1 + \int_1^{\max(1, N)} g(y) \frac{dy}{y} \right)^p.$$

Notice that this result is not interesting for  $N$  small since, if we use the crude upper bound  $W_p^p(\mu_N, \bar{\mu}_N) \leq \text{diam}(Q_N)^p M_N$ , we obtain  $\mathbb{E}[W_p^p(\mu_N, \bar{\mu}_N)] \leq \text{diam}(Q_N)^p A N = 2^{p/2} A N^{1+p/2}$ . For large  $N$  this grows much faster than linear but for small  $N$  this is much smaller than linear.

For  $N$  i.i.d. uniformly distributed points on  $Q_N$  and  $p \geq 1$  fixed we can choose  $g$  as a constant. The bound obtained in this case is proportional to  $N(\log N)^p$ , which is not the optimal bound  $N(\log N)^{p/2}$  from Equation (3). On the other hand, if  $g(y) = \frac{1}{(1+\log y)^r}$  with  $r > 1$ ,  $y^{-1}g(y)$  is integrable so that we obtain a linear bound in that case.

**Remark 3.** If we are only interested in the  $p = 1$  case we could change Hypothesis 1 to

$$\mathbb{E} \left[ \left| \frac{\mu_N(B)}{\mathbb{E}[\mu_N(B)]} - 1 \right| \right] \leq |B|^{-1/2} g(|B|) \text{ for every square } B \text{ that satisfies } |B| \geq 1$$

and eliminate Hypothesis 2 to obtain a universal constant  $\alpha$  such that

$$\mathbb{E}[W_1(\mu_N, \bar{\mu}_N)] \leq \alpha \mathbb{E}[\mu_N(Q_N)] \left( 1 + \int_1^{\max(1, N)} g(y) \frac{dy}{y} \right).$$

The previous theorem is stated in dimension 2 as it is the most studied case. Nevertheless, our proof applies to any dimension and we also state a general theorem, whose proof is exactly the same up to minor modifications.

**Theorem 4.** Fix  $p \geq 1$  and consider a real number  $N > 0$ . Let  $\mu_N$  be a random measure on the square  $Q_N = [0, N^{1/d}]^d$  such that the following conditions hold.

1. There exists a monotone function (non-decreasing or non-increasing)  $g : [1, \infty) \rightarrow (0, \infty)$  such that, for any square  $B \subset Q_N$  that satisfies  $|B| \geq 1$ , it holds that

$$\mathbb{E}[|\mu_N(B) - \mathbb{E}[\mu_N(B)]|^p] \leq |B|^{p\frac{(d-1)}{d}} g(|B|)^p. \quad (6)$$

2. There exist two constants  $a \in (0, 1]$  and  $A \in [1, \infty)$  such that for any square  $B \subset Q_N$

$$a|B| \leq \mathbb{E}[\mu_N(B)] \leq A|B|. \quad (7)$$

Let  $M_N = \mu_N(Q_N)$  and  $\bar{\mu}_N = \frac{M_N}{\mathbb{E}[M_N]} \mathbb{E}[\mu_N]$ . Then there exists a constant  $C_{p,d}$  such that

$$\mathbb{E}[W_p^p(\mu_N, \bar{\mu}_N)] \leq (C_{p,d} a^{1-2p} A^p) N \left(1 + \int_1^{\max(1,N)} g(y) \frac{dy}{y}\right)^p.$$

Since  $aN \leq \mathbb{E}[\mu_N(Q_N)] \leq AN$ , we should think of  $N$  as being the mass of  $\mu_N(Q_N)$  or, if  $\mu_N$  is a point process, as the number of particles of  $\mu_N$ . As explained in Subsection 2.2, we may translate this theorem to one about measures on the unit square  $Q_1$  by scaling everything so that (6) becomes  $\mathbb{E}[|\mu_N(B) - \mathbb{E}[\mu_N(B)]|^p] \leq N^{p\frac{(d-1)}{d}} |B|^{p\frac{(d-1)}{d}} g(N|B|)^p$  and (7) becomes  $aN|B| \leq \mathbb{E}[\mu_N(B)] \leq AN|B|$ . Since  $N$  may be thought of as the number of particles,  $\mu_N/N$  can be thought of as the empirical measure and (7) would be related to bounds on the density of  $\lim_{N \rightarrow \infty} \mu_N/N$ , the so-called equilibrium measure. In this way, an hypothesis like (7) makes sense on metric spaces with different classes of shapes. An example of such bounds for some sequences of point processes on manifolds can be found in [CE20, Theorem 1.6]. Moreover, we may replace  $Q_1$  by a flat  $d$ -dimensional torus and Theorem 4 still holds in this case since the distance in the torus is smaller than the distance in the cube.

**Remark 5** (Bound for uniform i.i.d. and  $d \neq 2$ ). *For the point process  $\mu_N$  of  $N$  i.i.d. uniformly distributed points on the  $d$ -cube  $Q_N$ , we may choose  $g(x)$  proportional to  $x^{\frac{1}{d}-\frac{1}{2}}$  (with a constant that does not depend on  $N$ ). We obtain the known upper bound (3) of optimal order if  $d \neq 2$ .*

**Remark 6** (On the first hypothesis of Theorem 4). *At first glance, Condition (6) seems somewhat peculiar. It becomes less mysterious if we think  $|B|^{\frac{(d-1)}{d}}$  as essentially being the area  $|\partial B|$  of the boundary  $\partial B$ . The case  $\|\mu_N(B) - \mathbb{E}[\mu_N(B)]\|_p \leq C|\partial B|$  seems naturally associated to perturbed lattices. More precisely, if we consider a family of random variables  $(X_z)_{z \in \mathbb{Z}^d}$  bounded by a constant  $R > 0$ , then the point process formed by  $(z + X_z)_{z \in \mathbb{Z}^d}$  (a perturbed lattice) satisfy Condition (6) with  $g$  constant. For this point process we have a bound  $\mathbb{E}[W_p^p(\mu_N, \bar{\mu}_N)] \leq \tilde{C}N$  so that the previous theorem does not give us the optimal rate, but the following interpretation may be suggested. If the fluctuation of the number of particles on  $B$  is much smaller than that of a perturbed lattice then  $\mathbb{E}[W_p^p(\mu_N, \bar{\mu}_N)]$  is bounded by a linear function of  $N$ . If the variables  $(X_z)_{z \in \mathbb{Z}^d}$  are independent, then we have an even better bound  $\|\mu_N(B) - \mathbb{E}[\mu_N(B)]\|_2 \leq C\sqrt{|\partial B|}$  and this condition is sometimes considered as the hyperuniformity of type I in general dimension. An example of the latter “type I” kind of bound for domains with smooth boundary on manifolds can be found in [CE20, Theorem 1.6]. We may also see [DFHL24] for connections between hyperuniformity and perturbed lattices.*

A particular interesting case is the one for a  $p$ -th moment version of type III hyperuniform point process that, since we have only defined type III for  $\mathbb{R}^2$ , is stated in this setting.

**Corollary 7.** *Fix  $p \geq 1$  and let  $X$  be a point process on  $\mathbb{R}^2$  and let  $\bar{X}$  be its expected value. Suppose that the following properties hold.*

1. There exist  $\varepsilon \in (0, 1)$  and  $\gamma > 0$  such that, for any square  $B \subset \mathbb{R}^2$  with  $|B| \geq 1$ ,

$$\mathbb{E}[|X(B) - \bar{X}(B)|^p] \leq \gamma |B|^{(1-\varepsilon)p/2}. \quad (8)$$

2. There exist  $a, A > 0$  such that for any square  $B \subset \mathbb{R}^2$ ,

$$a|B| \leq \bar{X}(B) \leq A|B|. \quad (9)$$

This is equivalent to requiring that  $\mathbb{E}[X]$  has a density bounded from above and from below with respect to the Lebesgue measure on  $\mathbb{R}^2$ .

Then there exists  $\alpha, \tilde{\alpha} > 0$  such that for any  $N > 0$ , if  $C_N = [-\sqrt{N}/2, \sqrt{N}/2]^2$ ,

$$\tilde{\alpha}N \leq \mathbb{E} \left[ W_p^p \left( X|_{C_N}, \frac{X(C_N)}{\bar{X}(C_N)} \bar{X}|_{C_N} \right) \right] \leq \alpha N.$$

The  $d$ -dimensional version would ask for a bound  $\mathbb{E}[|X(B) - \bar{X}(B)|^p]^{1/p} \leq \gamma |B|^{(1-\varepsilon)\frac{(d-1)}{d}}$ , which is weaker than the usual type III hyperuniformity found in the literature whose  $p$ -th moment version would be  $\mathbb{E}[|X(B) - \bar{X}(B)|^p]^{1/p} \leq \gamma |B|^{(1-\varepsilon)/2}$ .

For Hermitian determinantal point processes, the  $L^2$ -norm of the deviation essentially bounds its  $L^p$ -norm so that the first hypothesis of Theorem 4 for  $p = 2$  implies the same hypothesis for general finite  $p$ . In particular, for dimension 2, type III hyperuniform Hermitian determinantal point processes satisfy the first hypothesis of Corollary 7. This is the content of the next proposition.

**Proposition 8.** Fix  $p \geq 1$ . There exists  $\gamma_p > 0$  such that, for any Hermitian determinantal point process  $X$  and any measurable set  $B$  such that  $\mathbb{E}[X(B)]$  is finite, we have

$$\mathbb{E}[|X(B) - \mathbb{E}[X(B)]|^p] \leq \gamma_p (\text{Var}(X(B))^{p/2} + 1).$$

In particular, a Hermitian determinantal point process  $X$  on  $\mathbb{R}^2$  satisfying the first hypothesis of Corollary 7 for  $p = 2$  also satisfies it for every  $p \in [1, \infty)$  and the same  $\varepsilon$ .

## 4 Perspectives, related results and comments

In [LRY24, Theorem 3], Lachièze-Rey and Yogeshwaran obtained a theorem which deals with the same question, but with a very different approach and in the setting of stationary point processes. Comparing their hypothesis on the integrability of the reduced pair correlation measure and the first hypothesis of Theorem 2 does not seem obvious.

Two new works [HL24] and [DFHL24] have very recently given a different perspective on the main result in the setting of stationary point processes. On the one hand, Theorems 1 and 4 of [HL24] together say that a condition which is slightly weaker than the hypothesis of Theorem 2 for  $p = 2$  but stronger than classical hyperuniformity implies 2-Wasserstein linear (in  $N$ ) bounds. On the other hand, we may find in [DFHL24, Theorem 1] that an  $L^2$ -perturbed lattice in dimension 2 is hyperuniform which can be restated, by using an adapted version of [LRY24, Proposition 2], as saying that linear 2-Wasserstein bounds imply hyperuniformity. These implications nearly give the full picture for the relation between hyperuniformity and 2-Wasserstein linear bounds for 2-dimensional stationary point processes.

In dimension greater than 2, there exist perturbed lattices which are not hyperuniform with the usual definition of hyperuniformity given by comparing the process with a stationary Poisson



point process. The connection between hyperuniformity, Wasserstein bounds and perturbed lattices is now almost clear in dimension 2, but our result tells us that the more pertinent bound is the one obtained by comparing the process with a perturbed lattice as explained in Remark 6 above. This is an interesting direction for further investigations.

To conclude, this paper gives two main contributions to this topic. First we provide bounds which are valid for general random measures, not only stationary point processes, any dimension and any finite  $p \geq 1$ . Second, the proof is short and not very technical. One direction of improvement for the main result of this paper would be to relax the shape dependency for the moments estimate, i.e., to consider, for instance, disks instead of squares in the first hypotheses of Theorem 2, or to replace  $Q_N$  by other shapes or more general metric spaces.

## Acknowledgements

We thank Michael Goldman, Jonas Jalowy, Raphaël Lachièze-Rey and Thomas Leblé for useful discussions and insights. This work was supported by the CNRS via the PEPS funding “*Vitesse de convergence de mesures spectrales empiriques*”. We are also glad to acknowledge support from the GdR *Matrices et Graphes Aléatoires* (now included in the RT *Mathématiques et Physique*) which made this collaboration possible.

## 5 Proofs

We prove all results of Section 3 by order of apparition.

### 5.1 Proof of Theorem 2

The proof strategy is very close to the approach of Prod’homme for the Ginibre ensemble [Pro21]. It relies on a multi-scale argument which was already used on this problem by several authors, including [AKT84].

The main idea is to interpolate between  $\mu_N$  and  $\bar{\mu}_N$  by starting with  $\bar{\mu}_N$  on  $[0, \sqrt{N}]^2$  and dividing the square in four equal squares. Consider a measure  $\nu_1$  that on each of those new squares  $B$  is proportional to  $\bar{\mu}_N|_B$  but whose mass is  $\mu_N(B)$ . Then, divide each of the new squares in four equal squares  $B'$  and consider again a measure  $\nu_2$  proportional to  $\bar{\mu}_N|_{B'}$  but whose mass is  $\mu_N(B')$ . Repeat this procedure until some  $K$ -th subdivision where the smallest squares have size of order 1. Now, the point is that to compare  $\nu_k$  with  $\nu_{k+1}$  we may restrict ourselves to the squares in the  $k$ -th subdivision, and our task becomes a comparison between some measure and a “subdivision” of it in four equal squares. Here the key is [GT21, Lemma 3.4] which is written here under the name of Lemma 9. Finally, we compare the  $K$ -th step measure  $\nu_K$  with the measure  $\mu_N$  by comparing again  $\nu_K$  and  $\mu_N$  on each of the squares of the  $K$ -th subdivision but now using that the smallest squares have size of order 1.

For future use, notice first that, for any square  $B \subset [0, \sqrt{N}]^2$ ,

$$\mathbb{E} \left[ \left| \frac{\mu_N(B)}{\mathbb{E}[\mu_N(B)]} - 1 \right|^p \right] = \frac{1}{\mathbb{E}[\mu_N(B)]^p} \mathbb{E} \left[ \left| \mu_N(B) - \mathbb{E}[\mu_N(B)] \right|^p \right] \leq \frac{1}{a^p} |B|^{-p/2} g(|B|)^p \quad (10)$$

provided that  $|B| \geq 1$ . We shall suppose  $g$  non-increasing, the non-decreasing case works in a similar way.

**Step 1 – Interpolation:** We start by defining a finite sequence of intermediate measures which interpolate between  $\bar{\mu}_N$  and  $\mu_N$ . Let us denote  $K = \lfloor \log_4(N) \rfloor = \lfloor \log_2(\sqrt{N}) \rfloor$  and let us

consider, for each  $k \in \{0, \dots, K\}$ , the set  $\mathcal{B}_k$  of  $4^k$  squares which form a partition of  $[0, \sqrt{N}]^2$  obtained by recursively dividing each square into 4 equal squares. Equivalently, let us partition  $[0, \sqrt{N}]$  in  $2^k$  intervals of length  $\sqrt{N}2^{-k}$  as

$$\mathcal{I}_k = \left\{ \left[ \frac{\sqrt{N}(i-1)}{2^k}, \frac{\sqrt{N}i}{2^k} \right) : i \in \{0, \dots, 2^k - 1\} \right\}$$

and use this to partition the square  $[0, \sqrt{N}]^2$  in the  $4^k$  squares given by products of intervals from  $\mathcal{I}_k$ , i.e.,  $\mathcal{B}_k = \{I_1 \times I_2 : I_1, I_2 \in \mathcal{I}_k\}$ . Notice that the smallest square obtained in this way has length  $\sqrt{N}2^{-K} \geq \sqrt{N}2^{-\log_2(\sqrt{N})} = 1$ . For any  $k \in \{1, \dots, K\}$ , we define the measure

$$\nu_k = \sum_{B \in \mathcal{B}_k} \frac{\mu_N(B)}{\bar{\mu}_N(B)} \bar{\mu}_N|_B = \sum_{B \in \mathcal{B}_k} \frac{\mu_N(B)}{\mathbb{E}[\mu_N(B)]} \mathbb{E}[\mu_N]|_B.$$

In particular,  $\nu_0 = \bar{\mu}_N$ . Let us use the notation  $\nu_{K+1} = \mu_N$ . Since the  $L^p$  norm of a metric satisfies the triangle inequality, we have that

$$\mathbb{E}[W_p^p(\mu_N, \bar{\mu}_N)]^{1/p} \leq \sum_{k=0}^K \mathbb{E}[W_p^p(\nu_k, \nu_{k+1})]^{1/p}.$$

Indeed, the triangle inequality for  $W_p$  gives us

$$W_p(\mu_N, \bar{\mu}_N) \leq \sum_{k=0}^K W_p(\nu_k, \nu_{k+1})$$

so that, taking the  $L^p$  norm  $\|X\|_{L^p} = \mathbb{E}[|X|^p]^{1/p}$  at both sides we get

$$\|W_p(\mu_N, \bar{\mu}_N)\|_{L^p} \leq \left\| \sum_{k=0}^K W_p(\nu_k, \nu_{k+1}) \right\|_{L^p} \leq \sum_{k=0}^K \|W_p(\nu_k, \nu_{k+1})\|_{L^p}$$

which is what we wanted.

**Step 2 – Comparison between  $\nu_k$  and  $\nu_{k+1}$ :** We are now left to bound from above the quantity  $\mathbb{E}[W_p^p(\nu_k, \nu_{k+1})]$ , which will be done by controlling the transport cost on each of the squares of  $\mathcal{B}_k$  and then gluing together the squares. For a given level  $k \in \{0, \dots, K-1\}$  and a given square  $B \in \mathcal{B}_k$ , define the good event  $G_B$  by

$$G_B = \{\mu_N(B) \geq 0.5\mathbb{E}[\mu_N(B)]\}.$$

What is important from the number 0.5 in the definition of  $G_B$  is that it is in the interval  $(0, 1)$ , we may take whichever we prefer. We start by writing, for  $B \in \mathcal{B}_k$ ,

$$\mathbb{E}[W_p^p(\nu_k|_B, \nu_{k+1}|_B)] = \mathbb{E}[W_p^p(\nu_k|_B, \nu_{k+1}|_B)1_{G_B}] + \mathbb{E}[W_p^p(\nu_k|_B, \nu_{k+1}|_B)1_{G_B^c}] \quad (11)$$

and we bound each term separately.

**Substep 2.1 – Inside the event  $G_B$ :** For the first term, we notice that, on the event  $G_B$ , the density of  $\nu_k|_B$  with respect to Lebesgue measure is bounded from below by  $0.5a$ , where  $a$  is given in Hypothesis 2. Hence we can use the following lemma from [GT21, Lemma 3.4].

**Lemma 9.** *Let  $R$  be a square and let  $\mu$  and  $\lambda$  be two measures on  $R$  with equal mass, both absolutely continuous with respect to Lebesgue and such that  $\inf_R \lambda > 0$ . Then for every  $p \geq 1$ ,*

$$W_p^p(\mu, \lambda) \leq \theta_p \frac{\text{diam}(R)^p}{(\inf_R \lambda)^{p-1}} \int_R |\mu - \lambda|^p,$$

where the constant  $\theta_p$  only depends on  $p$ .

Notice that the  $p = 1$  case can be obtained by the identity  $W_1(\mu, \lambda) = \sup_{\|f\|_{\text{Lip}} \leq 1} \int f(\mu - \lambda)$  and does not need a lower bound for the density. For  $B \in \mathcal{B}_k$ , there are four squares in  $\mathcal{B}_{k+1}$  that are contained in  $B$ . Using Lemma 9, we can write

$$\begin{aligned} & \mathbb{E}[W_p^p(\nu_k|_B, \nu_{k+1}|_B) 1_{G_B}] \\ & \leq \theta_p (0.5a)^{1-p} \text{diam}(B)^p \mathbb{E} \left[ \int_B |\nu_k|_B - \nu_{k+1}|_B|^p \right] \\ & = \theta_p (0.5a)^{1-p} \text{diam}(B)^p \sum_{\tilde{B} \in \mathcal{B}_{k+1}, \tilde{B} \subset B} \mathbb{E} \left[ \left| \frac{\mu_N(B)}{\mathbb{E}[\mu_N(B)]} - \frac{\mu_N(\tilde{B})}{\mathbb{E}[\mu_N(\tilde{B})]} \right|^p \right] \int_{\tilde{B}} \mathbb{E}[\mu_N](x)^p dx, \end{aligned}$$

where  $\mathbb{E}[\mu_N](\cdot)$  denotes the probability density function of  $\mathbb{E}[\mu_N]$  with respect to Lebesgue measure. Use (10) and that  $|\tilde{B}| = |B|/4$  to obtain

$$\begin{aligned} \mathbb{E} \left[ \left| \frac{\mu_N(B)}{\mathbb{E}[\mu_N(B)]} - \frac{\mu_N(\tilde{B})}{\mathbb{E}[\mu_N(\tilde{B})]} \right|^p \right] & = \mathbb{E} \left[ \left| \frac{\mu_N(B)}{\mathbb{E}[\mu_N(B)]} - 1 + 1 - \frac{\mu_N(\tilde{B})}{\mathbb{E}[\mu_N(\tilde{B})]} \right|^p \right] \\ & \leq 2^{p-1} \left( \mathbb{E} \left[ \left| \frac{\mu_N(B)}{\mathbb{E}[\mu_N(B)]} - 1 \right|^p \right] + \mathbb{E} \left[ \left| 1 - \frac{\mu_N(\tilde{B})}{\mathbb{E}[\mu_N(\tilde{B})]} \right|^p \right] \right) \\ & \leq \frac{2^{p-1}}{a^p} (|B|^{-p/2} g(|B|)^p + |\tilde{B}|^{-p/2} g(|\tilde{B}|)^p) \\ & \leq \frac{2^{p-1}}{a^p} (1 + 2^p) |B|^{-p/2} g(|B|/4)^p. \end{aligned}$$

By defining  $\alpha_1 = \theta_p (0.5a)^{1-p} 2^{p-1} a^{-p} (1 + 2^p)$ , we have obtained

$$\mathbb{E}[W_p^p(\nu_k|_B, \nu_{k+1}|_B) 1_{G_B}] \leq \alpha_1 \text{diam}(B)^p |B|^{-p/2} g(|B|/4)^p \int_B \mathbb{E}[\mu_N](x)^p dx.$$

**Substep 2.2 – Outside of the event  $G_B$ :** Let us now consider  $\mathbb{E}[W_p^p(\nu_k|_B, \nu_{k+1}|_B) 1_{G_B^c}]$ . We have the crude bound

$$\mathbb{E}[W_p^p(\nu_k|_B, \nu_{k+1}|_B) 1_{G_B^c}] \leq \text{diam}(B)^p 0.5 \mathbb{E}[\mu_N(B)] \mathbb{P}(G_B^c).$$

For  $\mathbb{P}(G_B^c)$  we may use Markov's inequality to get that

$$\mathbb{P}(G_B^c) = \mathbb{P} \left( 1 - \frac{\mu_N(B)}{\mathbb{E}[\mu_N(B)]} > 0.5 \right) \leq \mathbb{P} \left( \left| 1 - \frac{\mu_N(B)}{\mathbb{E}[\mu_N(B)]} \right|^p > 0.5^p \right) \leq 0.5^{-p} \mathbb{E} \left[ \left| 1 - \frac{\mu_N(B)}{\mathbb{E}[\mu_N(B)]} \right|^p \right].$$

Define  $\alpha_2 = (0.5)(0.5)^{-p} a^{-p}$ . So, by using that  $g$  is non-increasing and (10), we have obtained

$$\mathbb{E}[W_p^p(\nu_k|_B, \nu_{k+1}|_B) 1_{G_B^c}] \leq \alpha_2 \text{diam}(B)^p |B|^{-p/2} g(|B|/4)^p \mathbb{E}[\mu_N(B)].$$

**Step 3 – Gluing the squares:** Notice that for  $B \in \mathcal{B}_k$  we have  $\text{diam}(B)^p |B|^{-p/2} = 2^{p/2}$  and  $|B| = N/4^k$ . By using that  $\sum_{B \in \mathcal{B}_k} \nu_k|_B = \nu_k$  and  $\sum_{B \in \mathcal{B}_k} \nu_{k+1}|_B = \nu_{k+1}$ , we have

$$\mathbb{E}[W_p^p(\nu_k, \nu_{k+1})] \leq 2^{p/2} g \left( \frac{N}{4^{k+1}} \right)^p \left( \alpha_1 \int_{Q_N} \mathbb{E}[\mu_N](x)^p dx + \alpha_2 \mathbb{E}[\mu_N(Q_N)] \right)$$

for  $k + 1 \leq K$ . Notice that, by the definition of  $K$ , we have that  $\frac{N}{4} \leq 4^K \leq N$ . Then,

$$\begin{aligned} \mathbb{E}[W_p^p(\nu_0, \nu_K)]^{1/p} &\leq \sum_{k=0}^{K-1} \mathbb{E}[W_p^p(\nu_k, \nu_{k+1})]^{1/p} \\ &\leq \sqrt{2} \left( \alpha_1 \int_{Q_N} \mathbb{E}[\mu_N](x)^p dx + \alpha_2 \mathbb{E}[\mu_N(Q_N)] \right)^{1/p} \sum_{k=0}^{K-1} g\left(\frac{N}{4^{k+1}}\right) \\ &\leq \sqrt{2} \left( \alpha_1 \int_{Q_N} \mathbb{E}[\mu_N](x)^p dx + \alpha_2 \mathbb{E}[\mu_N(Q_N)] \right)^{1/p} \int_0^K g\left(\frac{N}{4^{t+1}}\right) dt. \end{aligned}$$

By a change of variables we obtain

$$\int_0^K g\left(\frac{N}{4^{t+1}}\right) dt = \frac{1}{\ln 4} \int_{\frac{N}{4^{K+1}}}^{\frac{N}{4}} g(y) \frac{dy}{y} \leq \frac{1}{\ln 4} \int_1^N g(y) \frac{dy}{y}.$$

We have one term left to understand, namely  $\mathbb{E}[W_p^p(\nu_K, \mu_N)]$ . We can use that

$$W_p^p(\nu_K, \mu_N) \leq \sum_{B \in \mathcal{B}_K} W_p^p(\nu_K|_B, \mu_N|_B) \leq \sum_{B \in \mathcal{B}_K} \text{diam}(B)^p \mu_N(B)$$

We now use that the diameter of each  $B \in \mathcal{B}_K$  is  $\frac{\sqrt{2N}}{2^K} = \sqrt{\frac{2N}{4^K}} \leq 2\sqrt{2}$  and that  $\sum_{B \in \mathcal{B}_K} \mu_N(B)$  is  $\mu_N(Q_N)$ . Here we are using that  $\mu_N(\partial Q_N) = 0$ , consequence of  $\mathbb{E}[\mu_N(\partial Q_N)] = 0$ . Then,

$$\mathbb{E}[W_p^p(\nu_K, \mu_N)] \leq 2^{3p/2} \mathbb{E}[\mu_N(Q_N)].$$

We conclude the inequality

$$\begin{aligned} \mathbb{E}[W_p^p(\mu_N, \bar{\mu}_N)]^{1/p} &\leq \frac{\sqrt{2}}{\ln 4} \left( \alpha_1 \int_{Q_N} \mathbb{E}[\mu_N](x)^p dx + \alpha_2 \mathbb{E}[\mu_N(Q_N)] \right)^{1/p} \int_1^N g(y) \frac{dy}{y} \\ &\quad + 2^{3/2} \mathbb{E}[\mu_N(Q_N)]^{1/p}. \end{aligned}$$

By using that  $\mathbb{E}[\mu_N](x) \leq A$  and defining  $\alpha = 2^{p-1}(\sqrt{2}/\ln 4)^p(\alpha_1 A^p + (\alpha_2 + 2^{3p/2})A)$  we get

$$\mathbb{E}[W_p^p(\mu_N, \bar{\mu}_N)] \leq \alpha N \left( 1 + \int_1^N g(y) \frac{dy}{y} \right)^p.$$

Since  $A \leq A^p$  and  $a^{-p} \leq a^{1-2p}$ , we can bound  $\alpha$  by a  $a^{1-2p} A^p$  times a constant  $C_p$  that only depends on  $p$  (for instance,  $C_p = 2^{5p}(1 + \theta_p)$  works).

## 5.2 Proof of Theorem 4

The proof of Theorem 4 follows the proof of Theorem 2 line by line. One has to divide the cube  $[0, N^{1/d}]^d$  in  $2^{dk}$  sub-cubes of size  $2^{-k} N^{1/d}$ . The key [GT21, Lemma 3.4] (stated here as Lemma 9) is valid in any dimension which allows us to generalize the proof.

## 5.3 Proof of Corollary 7

Theorem 2 gives the upper bound if we define  $g(x) = \gamma x^{-\varepsilon/2}$ ,

$$\mu_N = X|_{C_N} \quad \text{and} \quad \bar{\mu}_N = \frac{X(C_N)}{\bar{X}(C_N)} \bar{X}|_{C_N}.$$

For the lower bound consider  $\delta > 0$  and define  $A_\delta = \{(1 - \delta)N \leq X(C_N) \leq (1 + \delta)N\}$ . In  $A_\delta$ , the density of  $\bar{\mu}_N$  is bounded from above by  $(1 + \delta)a^{-1}A$  so that, by Lemma 1, in  $A_\delta$

$$W_p^p(\mu_N, \bar{\mu}_N) \geq \frac{\alpha_2^p X(C_N)}{(1 + \delta)^{p/2} a^{-p/2}} \geq c_\delta N,$$

where  $c_\delta = \alpha_2^p (1 - \delta)(1 + \delta)^{-p/2} a^{p/2}$ . Then,  $\mathbb{E}[W_p^p(\mu_N, \bar{\mu}_N)] \geq c_\delta N \mathbb{P}(A_\delta)$ . By noticing that

$$\mathbb{P}(A_\delta^c) = \mathbb{P}\left(\left|\frac{X(C_N)}{N} - 1\right| > \delta\right) \leq \frac{\mathbb{E}\left[\left|\frac{X(C_N)}{N} - 1\right|\right]}{\delta} \leq \frac{A + 1}{\delta},$$

we may choose  $\delta = 2(A + 1)$  and obtain the inequality  $\mathbb{E}[W_p^p(\mu_N, \bar{\mu}_N)] \geq \frac{c_\delta}{2} N$ .

## 5.4 Proof of Proposition 8

Consider a Hermitian determinantal point process  $X$  on  $\mathbb{R}^d$  (or any space). Then, for any measurable  $B \subset \mathbb{R}^d$ , we have  $X(B) = \sum_{i=1}^n \xi_i$ , where  $n \in \mathbb{N} \cup \{+\infty\}$  and the  $\xi_i$ 's are independent Bernoulli variables [HKPV09, Theorem 4.5.3]. Bernstein's inequality [BLB04, Theorem 3] applied to the bounded centered independent random variables  $\xi_i - \mathbb{E}[\xi_i]$  yields that

$$\mathbb{P}(|X(B) - \mathbb{E}[X(B)]| \geq t) \leq 2 \exp\left(-\frac{t^2}{2 \text{Var}(X(B)) + t/3}\right)$$

for every  $t > 0$ . Then,

$$\begin{aligned} & \mathbb{E}[|X(B) - \mathbb{E}[X(B)]|^p] \\ &= \int_0^{+\infty} p t^{p-1} \mathbb{P}(|X(B) - \mathbb{E}[X(B)]| > t) dt \\ &\leq 2 \int_0^{+\infty} p t^{p-1} e^{-t^2/(2 \text{Var}(X(B)) + t/3)} dt \\ &\leq 2 \int_0^{6 \text{Var}(X(B))} p t^{p-1} e^{-t^2/(4 \text{Var}(X(B)))} dt + 2 \int_{6 \text{Var}(X(B))}^{+\infty} p t^{p-1} e^{-3t/2} dt \\ &\leq 2^{p+1} \text{Var}(X(B))^{p/2} \int_0^{3\sqrt{\text{Var}(X(B))}} p u^{p-1} e^{-u^2} dt \\ &\quad + 2e^{-3 \text{Var}(X(B))/4} \int_{6 \text{Var}(X(B))}^{+\infty} p t^{p-1} e^{-3t/4} dt \\ &\leq 2^{p+1} \text{Var}(X(B))^{p/2} \int_0^{+\infty} p u^{p-1} e^{-u^2} dt + 2e^{-3 \text{Var}(X(B))/4} \int_0^{+\infty} p t^{p-1} e^{-3t/4} dt \\ &\leq \gamma_p (\text{Var}(X(B))^{p/2} + e^{-3 \text{Var}(X(B))/4}) \\ &\leq \gamma_p (\text{Var}(X(B))^{p/2} + 1) \end{aligned}$$

where  $\gamma_p$  is the maximum of both integrals times  $2^{p+1}$ .

## 6 Appendix: Proof of the lower bound

*Proof of Lemma 1.* Let us begin with  $p = 1$ . Define the 1-Lipschitz function  $f : \mathbb{R}^d \rightarrow [0, \infty[$ ,

$$f(x) = \max\{(c - |x - x_1|)_+, \dots, (c - |x - x_n|)_+\},$$

so that  $f(x_i) = c$ . Define  $\mu_n = \sum_{i=1}^n \delta_{x_i}$  and notice that  $\int_{\mathbb{R}^d} f d\mu_n = cn$ . Since  $\mu$  has a density bounded from above by  $A$ ,

$$\int_{\mathbb{R}^d} f d\mu \leq A \sum_{i=1}^n \int_{\mathbb{R}^d} 1_{\{|x-x_i| \leq c\}} (c - |x - x_i|) dx.$$

Since

$$\int_{\mathbb{R}^d} 1_{\{|x-y| \leq c\}} (c - |x - y|) dx = c|B_1|c^d - |\partial B_1| \int_0^c r^d dr = \frac{c^{d+1}|\partial B_1|}{d(d+1)},$$

we get that

$$\int_{\mathbb{R}^d} f d\mu_n - \int_{\mathbb{R}^d} f d\mu \geq n \left( c - A \frac{c^{d+1}|\partial B_1|}{d(d+1)} \right).$$

Taking  $c$  that maximizes  $c \mapsto c - A \frac{c^{d+1}|\partial B_1|}{d(d+1)}$ , i.e.,  $c = \left( \frac{d}{A|\partial B_1|} \right)^{1/d}$  gives

$$W_1(\mu_n, \mu) \geq n \left( \frac{d}{A|\partial B_1|} \right)^{1/d} \left( \frac{d}{d+1} \right).$$

By Hölder's inequality,  $W_1(\mu, \nu) \leq n^{1-\frac{1}{p}} W_p(\mu, \nu)$ , so that

$$W_p(\mu_n, \mu) \geq n^{1/p} \left( \frac{d}{A|\partial B_1|} \right)^{1/d} \left( \frac{d}{d+1} \right).$$

□

## References

- [AGT22] Luigi Ambrosio, Michael Goldman, and Dario Trevisan. On the quadratic random matching problem in two-dimensional domains. *Electronic Journal of Probability*, 27:1–35, 2022.
- [AKT84] Miklós Ajtai, János Komlós, and Gábor Tusnády. On optimal matchings. *Combinatorica*, 4:259–264, 1984.
- [AST19] Luigi Ambrosio, Federico Stra, and Dario Trevisan. A PDE approach to a 2-dimensional matching problem. *Probability Theory and Related Fields*, 173:433–477, 2019.
- [BG14] Emmanuel Boissard and Thibaut Le Gouic. On the mean speed of convergence of empirical and occupation measures in Wasserstein distance. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 50(2):539 – 563, 2014.
- [BLB04] Stéphane Boucheron, Gábor Lugosi, and Olivier Bousquet. *Concentration Inequalities*, pages 208–240. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.
- [CE20] Laurent Charles and Benoit Estienne. Entanglement entropy and Berezin-Toeplitz operators. *Comm. Math. Phys.*, 376(1):521–554, 2020.
- [Cha16] Djilil Chafaï. <https://djilil.chafai.net/blog/2016/10/19/kantorovich-invented-wasserstein-distances/>, 2016.

- [CHM18] Djali Chafaï, Adrien Hardy, and Mylène Maïda. Concentration for Coulomb gases and Coulomb transport inequalities. *Journal of Functional Analysis*, 275(6):1447–1483, 2018.
- [CMMOC18] Tom Carroll, Jordi Marzo, Xavier Massaneda, and Joaquim Ortega-Cerdà. Equidistribution and  $\beta$ -ensembles. *Ann. Fac. Sci. Toulouse, Math. (6)*, 27(2):377–387, 2018.
- [Cos21] Simon Coste. Order, fluctuations, rigidities. [https://scoste.fr/assets/survey\\_hyperuniformity.pdf](https://scoste.fr/assets/survey_hyperuniformity.pdf), 2021.
- [DFHL24] David Dereudre, Daniela Flimmel, Martin Huesmann, and Thomas Leblé. (Non)-hyperuniformity of perturbed lattices. *arXiv preprint arXiv:2405.19881*, 2024.
- [DY95] Vladimir Dobrić and Joshep E. Yukich. Asymptotics for transportation cost in high dimensions. *Journal of Theoretical Probability*, 8(1):97–118, 1995.
- [FG15] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probability theory and related fields*, 162(3):707–738, 2015.
- [GHO18] Michael Goldman, Martin Huesmann, and Felix Otto. A large-scale regularity theory for the Monge-Ampère equation with rough data and application to the optimal matching problem. *arXiv preprint arXiv:1808.09250*, 2018.
- [GT21] Michael Goldman and Dario Trevisan. Convergence of asymptotic costs for random Euclidean matching problems. *Probability and Mathematical Physics*, 2(2):341–362, 2021.
- [HKPV09] J. Ben Hough, Manjunath Krishnapur, Yuval Peres, and Bálint Virág. *Zeros of Gaussian analytic functions and determinantal point processes*, volume 51 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2009.
- [HL24] Martin Huesmann and Thomas Leblé. The link between hyperuniformity, Coulomb energy, and Wasserstein distance to Lebesgue for two-dimensional point processes. *arXiv preprint arXiv:2404.18588*, 2024.
- [Jal23] Jonas Jalowy. The Wasserstein distance to the circular law. In *Annales de l’Institut Henri Poincaré (B) Probabilités et statistiques*, volume 59, pages 2285–2307. Institut Henri Poincaré, 2023.
- [Led19] Michel Ledoux. On optimal matching of Gaussian samples. *Journal of Mathematical Sciences*, 238(4):495–522, 2019.
- [LRY24] Raphaël Lachièze-Rey and Dhandapani Yogeshwaran. Hyperuniformity and optimal transport of point processes. *arXiv preprint arXiv:2402.13705*, 2024.
- [MM15] Elizabeth S. Meckes and Mark W. Meckes. A rate of convergence for the circular law for the complex Ginibre ensemble. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 24, pages 93–117, 2015.
- [OW23] Sean O’Rourke and Noah Williams. Partial linear eigenvalue statistics for non-Hermitian random matrices. *Theory Probab. Appl.*, 67(4):613–632, 2023. Reprint of *Teor. Veroyatn. Primen.* **67** (2022), 768–791.

- [Pro21] Maxime Prod'Homme. *Contributions au problème du transport optimal et à sa régularité*. PhD thesis, Toulouse 3, 2021.
- [SS15] Etienne Sandier and Sylvia Serfaty. 2D Coulomb gases and the renormalized energy. *Ann. Probab.*, 43(4):2026–2083, 2015.
- [Tal94] Michel Talagrand. The transportation cost from the uniform measure to the empirical measure in dimension  $\geq 3$ . *The Annals of Probability*, 22(2):919 – 959, 1994.
- [Tor18] Salvatore Torquato. Hyperuniform states of matter. *Physics Reports*, 745:1–95, 2018.
- [Ver13] A. M. Vershik. Long history of the Monge-Kantorovich transportation problem. *The Mathematical Intelligencer*, 35(4):1–9, 2013.