



# Annales de la Faculté des Sciences de Toulouse

MATHÉMATIQUES

EMANUELE CAGLIOTI, MICHAEL GOLDMAN, FRANCESCA PIERONI AND DARIO  
TREVISAN

*Subadditivity and optimal matching of unbounded samples*

Tome XXXV, n° 3 (2026), p. 697–742.

<https://doi.org/10.5802/afst.1861>

© les auteurs, 2026.

Les articles des *Annales de la Faculté des Sciences de Toulouse* sont mis  
à disposition sous la license Creative Commons Attribution (CC-BY) 4.0  
<http://creativecommons.org/licenses/by/4.0/>



Publication membre du centre  
Mersenne pour l'édition scientifique ouverte  
<http://www.centre-mersenne.org/>  
e-ISSN : 2258-7519

## Subadditivity and optimal matching of unbounded samples <sup>(\*)</sup>

EMANUELE CAGLIOTI <sup>(1)</sup>, MICHAEL GOLDMAN <sup>(2)</sup>,  
FRANCESCA PIERONI <sup>(3)</sup> AND DARIO TREVISAN <sup>(4)</sup>

---

**ABSTRACT.** — We obtain new bounds for the optimal matching cost for empirical measures with unbounded support. For a large class of radially symmetric and rapidly decaying probability laws, we prove for the first time the asymptotic rate of convergence for the whole range of power exponents  $p$  and dimensions  $d$ . Moreover we identify the exact prefactor when  $p \leq d$ . We cover in particular the Gaussian case, going far beyond the currently known bounds. Our proof technique is based on approximate sub- and super-additivity bounds along a geometric decomposition adapted to some features of the density, such as its radial symmetry and its decay at infinity.

---

<sup>(\*)</sup> Reçu le 12 juillet 2024, accepté le 5 janvier 2026.

**Keywords:** bipartite matching problem, optimal transport, geometric probability.

**2020 Mathematics Subject Classification:** 60D05, 90C05, 39B62, 60F25, 49Q22.

<sup>(1)</sup> Sapienza University of Rome, Department of Mathematics, Piazzale Aldo Moro, 5, 00185, Rome, Italy — emanuele.caglioti@uniroma1.it

<sup>(2)</sup> École polytechnique, Institut Polytechnique de Paris, CMAP, CNRS, 91120 Palaiseau, France — michael.goldman@cnrs.fr

<sup>(3)</sup> Sapienza University of Rome, Department of Mathematics, Piazzale Aldo Moro, 5, 00185, Rome, Italy — francesca.pieroni@uniroma1.it

<sup>(4)</sup> University of Pisa, Department of Mathematics, Largo Bruno Pontecorvo, 5, 56127, Pisa, Italy — dario.trevisan@unipi.it

F.P., M.G. and D.T. acknowledges the project G24-202 “Variational methods for geometric and optimal matching problems” funded by Università Italo Francese. E.C. and D.T. acknowledge partial support by PNRR - MUR project PE00000013 - “FAIR - Future Artificial Intelligence Research”. D.T. acknowledges the MUR Excellence Department Project awarded to the Department of Mathematics, University of Pisa, CUP I57G22000700001, the HPC Italian National Centre for HPC, Big Data and Quantum Computing - CUP I53C22000690001, the PRIN 2022 Italian grant 2022WHZ5XH - “understanding the Learning process of Quantum Neural networks (LeQun)”, CUP J53D23003890006, the INdAM-GNAMPa project 2024 “Tecniche analitiche e probabilistiche in informazione quantistica”. E.C. and F.P. acknowledge the grant “Progetti di ricerca d’Ateneo 2024” by Sapienza University, Rome and GNFM - INdAM.

Article proposé par François Delarue.

**RÉSUMÉ.** — Nous obtenons de nouvelles bornes pour le coût d'appariement optimal de certaines mesures empiriques à support non borné. Pour une large classe de lois de probabilité à symétrie radiale et à décroissance rapide, nous établissons pour la première fois le taux asymptotique de convergence pour toute la gamme des exposants de puissance  $p$  et des dimensions  $d$ . De plus, nous identifions le facteur multiplicatif exact lorsque  $p \leq d$ . Nous couvrons en particulier le cas gaussien, en allant bien au-delà des bornes actuellement connues. Notre technique de preuve repose sur des bornes d'additivité approchée inférieure et supérieure, fondées sur une décomposition géométrique adaptée à certaines propriétés de la densité, telles que sa symétrie radiale et sa décroissance à l'infini.

## 1. Introduction

In this paper we investigate the behavior of the transportation cost between exponentially decaying distributions and their empirical measure obtained from a sample of  $n$  independent observations. To a large extent, we complete the program initiated in [10, 19, 20, 23].

In order to state our main result let us introduce some notation. Letting  $d \geq 1$  and  $(X_i)_{i=1}^\infty$  be i.i.d. random variables uniformly distributed on the unit cube  $Q := (0, 1)^d$  and then  $\mu_n := \sum_{i=1}^n \delta_{X_i}$  be the (non-normalized) empirical measure, the optimal matching problem concerns the study of the asymptotic behavior of

$$\mathbb{E}[W_p^p(\mu_n, n\mathbb{1}_Q)],$$

where  $W_p^p$  denotes the  $p$ -Wasserstein distance. Letting for  $p, d \geq 1$ ,

$$\eta_n^{p,d} := \begin{cases} n^{1-\frac{p}{2}} & \text{if } d = 1, \\ n^{1-\frac{p}{2}}(\log n)^{\frac{p}{2}} & \text{if } d = 2, \\ n^{1-\frac{p}{d}} & \text{if } d \geq 3, \end{cases} \quad (1.1)$$

we define

$$\bar{\mathbf{c}}_{p,d} := \limsup_{n \rightarrow \infty} \frac{1}{\eta_n^{p,d}} \mathbb{E}[W_p^p(\mu_n, n\mathbb{1}_Q)]. \quad (1.2)$$

To simplify notation we write  $\eta_n$  for  $\eta_n^{p,d}$  when  $p$  and  $d$  are clear from the context. For the case  $d = 1$ , we refer to [9] for a thorough discussion and will focus on the case  $d \geq 2$  from now on. Since the work [1], see also [19, Section 4], it is known that for every  $p \geq 1$  and every  $d \geq 2$ ,  $\bar{\mathbf{c}}_{p,d} \in (0, \infty)$ .

*Remark 1.1.* — Let us point out that if  $d \geq 3$  or  $p = d = 2$ , this limsup is actually a limit, see [3, 5, 11, 12, 16] but this does not affect our analysis. It is still an open problem to prove the convergence when  $d = 2$  and  $p \neq 2$ . See however [4] for recent progress in this direction.

Following classical ideas in the matching problem, see e.g. [2, 5, 12, 25] we also consider a variant of the problem where mass can be transported to and from the boundary. We denote by  $Wb_Q^p$  the boundary  $p$ -Wasserstein distance (see (2.5) for the definition) and set

$$\underline{\mathfrak{c}}^{p,d} := \liminf_{n \rightarrow \infty} \frac{1}{\eta_n} \mathbb{E} \left[ Wb_Q^p(\mu_n, n\mathbb{1}_Q) \right]. \quad (1.3)$$

The point of introducing such a limiting constant is that the boundary Wasserstein distance is superadditive, contrary to the classical Wasserstein distance, which is subadditive. Hence, it turns out to be very useful for the lower bound.

While it is probably common knowledge in the community that also  $\underline{\mathfrak{c}}^{p,d} \in (0, \infty)$  (the proof for  $d \geq 3$  can be easily adapted from [14]), we provide a proof in Proposition A.3 in the appendix.

*Remark 1.2.* — As for  $\bar{\mathfrak{c}}_{p,d}$ , when  $p = d = 2$  or  $d \geq 3$ , the  $\liminf$  is actually a limit, see [2, 17] or a very simple proof in the case  $d \geq 3$  in Proposition A.3. Again, this will play no role in our analysis.

*Remark 1.3.* — While it is conjectured that  $\underline{\mathfrak{c}}^{p,d} = \bar{\mathfrak{c}}_{p,d}$  for every  $p, d \geq 1$  it is currently only known for  $p = 2$ , see [2, 15]. See also [24] for the case  $d = 1$  and  $p \in (0, 1/2)$ .

For  $q > 0$  we define the rate function

$$\tau_n := \begin{cases} \eta_n & \text{if } p \in [1, d), \\ (\log n)^{1+\frac{2}{q}} & \text{if } p = d = 2, \\ (\log n)^{\frac{d}{q}} & \text{if } d \geq 3 \text{ and } p = d, \\ (\log n)^{d-1-p(1-\frac{1}{q})} & \text{if } d \geq 2 \text{ and } p > d. \end{cases} \quad (1.4)$$

Let now  $V$  be a radial and  $C^1$  function satisfying for some  $C > 0$  and every  $t \geq 1$  (with a slight abuse of notation we identify  $V(x)$  and  $V(|x|)$ ),

$$\frac{1}{C} t^{q-1} \leq V'(t) \leq C t^{q-1}. \quad (1.5)$$

We assume moreover that  $V(0)$  is chosen such that  $\rho(x) := \exp(-V(x))$  is a probability density on  $\mathbb{R}^d$ . Let  $R_n$  be the unique solution to

$$V(R_n) = \log n \quad (1.6)$$

and set

$$\begin{aligned} \bar{\ell} &:= \limsup_{n \rightarrow \infty} \frac{R_n^d}{(\log n)^{\frac{d}{q}}} \int_{B_1} \left( 1 - \mathbb{1}_{d=2} \frac{V(R_n x)}{\log n} \right) \\ \text{and } \underline{\ell} &:= \liminf_{n \rightarrow \infty} \frac{R_n^d}{(\log n)^{\frac{d}{q}}} \int_{B_1} \left( 1 - \mathbb{1}_{d=2} \frac{V(R_n x)}{\log n} \right). \end{aligned} \quad (1.7)$$

*Remark 1.4.* — Under hypothesis (1.5) it is not hard to check that  $0 < \underline{\ell} \leq \bar{\ell} < \infty$ . Moreover, it is easy to construct examples of potentials  $V$  satisfying (1.5) but for which  $\underline{\ell} < \bar{\ell}$ . Let us also notice that the value of  $\underline{\ell}, \bar{\ell}$  does not change if instead of (1.6) we use

$$V(R_n) + \alpha \log R_n = \log n$$

for some  $\alpha \in \mathbb{R}$ .

*Remark 1.5.* — If  $V(t) = V(0) + t^q/q$  (the case  $q = 2$  corresponding to the Gaussian) then to leading order,

$$R_n = (q \log n)^{\frac{1}{q}} \quad \text{and} \quad \underline{\ell} = \bar{\ell} = q^{\frac{d}{q}} \int_{B_1} (1 - \mathbb{1}_{d=2}|x|^q).$$

Our main result is the following:

**THEOREM 1.6.** — *Let  $(X_i)_{i=1}^\infty$  be i.i.d. random variables of law  $\rho$ . Setting  $\mu_n := \sum_{i=1}^n \delta_{X_i}$ , we have for  $p \leq d$ ,*

$$\begin{aligned} \bar{\mathfrak{c}}_{p,d} \left( \int_{\mathbb{R}^d} \rho^{1-\frac{p}{d}} \mathbb{1}_{p < d} + \bar{\ell} \mathbb{1}_{p=d} \right) &\geq \limsup_{n \rightarrow \infty} \frac{1}{\tau_n} \mathbb{E}[W_p^p(\mu_n, n\rho)] \\ &\geq \liminf_{n \rightarrow \infty} \frac{1}{\tau_n} \mathbb{E}[W_p^p(\mu_n, n\rho)] \geq \underline{\mathfrak{c}}^{p,d} \left( \int_{\mathbb{R}^d} \rho^{1-\frac{p}{d}} \mathbb{1}_{p < d} + \underline{\ell} \mathbb{1}_{p=d} \right) \end{aligned} \quad (1.8)$$

and for  $p > d$ ,

$$\mathbb{E}[W_p^p(\mu_n, n\rho)] \simeq \tau_n. \quad (1.9)$$

Theorem 1.6 is obtained as a combination of Proposition 3.1 for the lower bound and Proposition 4.1 for the upper bound.

*Remark 1.7.* — On the one hand, when  $p = d = q = 2$  (where  $\bar{\mathfrak{c}}_{p,d} = \underline{\mathfrak{c}}^{p,d} = 1/(4\pi)$ ) we get from Remark 1.5,

$$\underline{\mathfrak{c}}^{p,d} \underline{\ell} = \bar{\mathfrak{c}}_{p,d} \bar{\ell} = \frac{1}{4},$$

recovering the result from [10]. On the other hand, for  $p < d/2$  with  $d \geq 3$ , Theorem 1.6 is covered by the results in [5, 12].

In Theorem 8.1, we complement the convergence in expectation with concentration estimates when  $p \leq d$  and  $\rho$  satisfies a Poincaré inequality. Similar concentration results were obtained in this range of parameters (actually  $p \leq 2$  if  $d = 2$  and  $p < d$  if  $d \geq 3$ ) in the case of the matching to the uniform measure on the cube in [3, Remark 4.7], see also [16, Remark 6.5] or [17, Proposition 5.3].

## Previous results, heuristics and strategy of proof

As understood in [5, 12], while the matching problem is quite robust when  $p < \frac{d}{2}$ ,  $d \geq 3$ , the situation becomes more complicated when  $p \geq \frac{d}{2}$  (or  $d = 2$ ). While the case of general densities on bounded domains was essentially treated in [2, 6, 16] (see also Theorem A.2) the unbounded case remains challenging. Motivated by this question, the study of matching for Gaussian samples was initiated in [19], where using semi-group techniques, Ledoux proved the correct upper bound for the scaling law in the case  $p = d = 2$  but only a sub-optimal local bound. Soon after the correct lower bound (in terms of scaling laws) was obtained in [23]. In that paper, Talagrand proves the correct lower bound  $(\log n)^{\frac{d}{2}}$  when  $p = d \geq 3$  and suggests that his proofs should also apply to the family of densities  $\rho(x) = \exp(-|x|^q)$  with  $q > 0$ . Still at the level of scaling law, the case  $p < d$  with  $d \geq 3$  was covered in [20]. Using subadditivity techniques in the spirit of [2, 6, 16], the optimal constant was recently obtained (still in the Gaussian case) for the case  $p = d = 2$  in [10].

Here are the main contributions of this paper:

- In the Gaussian case we improve on [19, 20, 23] by obtaining the optimal prefactors. Moreover, we are able to treat (this time only at the level of scaling laws) also the case  $p > d$  where no conjecture was even made. A surprising feature of this regime is that the extra logarithmic in dimension  $d = 2$  disappears, see Remark 1.8.
- We cover a large family of radial densities thus proving that there is nothing special about the Gaussian measure. In particular [19, 20] strongly relies on the properties of the Ornstein–Uhlenbeck semi-group while [10] used crucially the tensorization properties of the Gaussian measure. Notice that for  $q < 1$ , measures satisfying (1.5) cannot be log-concave (see also Remark 8.3) which shows that also this natural property is not essential.
- We show that obtaining (sharp) lower bounds in the matching problem is much easier than obtaining the corresponding upper bounds, see Remark 1.10. In particular we include in Appendix A a short proof of the convergence of the boundary transport problem in the case of the uniform measure on the cube as well as the lower bound on bounded sets. The main new ingredient is the elementary estimate (2.7) of Lemma 2.5.
- For the upper bound we use a general subadditivity property of the Wasserstein distance, see Lemma 2.2. This goes beyond the classical application of subadditivity in the context of matching problems where usually the measures are required to have disjoint support, see also the discussion below.

As in [2, 6, 10, 16, 17, 21], see also [5, 12] in the case  $p < d/2$ , our proof, which we now sketch, relies on subadditivity/superadditivity arguments. We focus on the case  $d \geq 3$  to simplify the notation. The idea is to subdivide  $\mathbb{R}^d$  in cubes  $Q_j = z_j + (0, h_j)^d$  for  $h_j > 0$  and  $z_j \in \mathbb{R}^d$ . Assuming for a moment that for every  $j$  (which of course is never the case),

$$\mu_n(Q_j) = n\rho(Q_j), \quad (1.10)$$

we can write by subadditivity (see (2.3) for the notation  $W_\Omega$ ),

$$W_p^p(\mu_n, n\rho) \leq \sum_j W_{Q_j}^p(\mu_n, n\rho).$$

Now provided  $V$  does not oscillate too much on  $Q_j$  i.e.

$$h_j \sup_{Q_j} |\nabla V| \simeq h_j |z_j|^{q-1} \ll 1, \quad (1.11)$$

we may consider that in each  $Q_j$  the points  $X_i$  follow a uniform distribution. Finally, provided there are in expectation many points in each of the cubes i.e.

$$\mathbb{E}[\mu_n(Q_j)] = n\rho(Q_j) \simeq nh_j^d \rho(z_j) \gg 1, \quad (1.12)$$

we can use the definition of  $\bar{\mathbf{c}}_{p,d}$  to infer

$$\begin{aligned} \mathbb{E}\left[W_{Q_j}^p(\mu_n, n\rho)\right] &\leq (1 + o(1)) \bar{\mathbf{c}}_{p,d} h_j^p (n\rho(Q_j))^{1-\frac{p}{d}} \\ &= (1 + o(1)) \bar{\mathbf{c}}_{p,d} n^{1-\frac{p}{d}} \int_{Q_j} \rho^{1-\frac{p}{d}}. \end{aligned}$$

This estimate is then summed over all cubes satisfying (1.11) and (1.12). Let  $R_n := \max |z_j|$  and set  $h = R_n^{1-q}$  so that (1.11) is almost satisfied. From (1.12) we see that we must impose

$$nR_n^{d(1-q)} \exp(-V(R_n)) \gg 1.$$

This motivates the choice of  $R_n$  as the unique solution to

$$V(R_n) + d(q-1) \log R_n = \log n. \quad (1.13)$$

Notice that with this definition, we have

$$R_n \simeq (\log n)^{\frac{1}{q}}.$$

Since (see Remark 1.4 and the computation (4.14))

$$\begin{aligned} n^{1-\frac{p}{d}} \int_{B_{R_n}} \rho^{1-\frac{p}{d}} &= (1 + o(n)) \tau_n \\ &\times \left( \mathbb{1}_{p < d} \int_{\mathbb{R}^d} \rho^{1-\frac{p}{d}} + \mathbb{1}_{p=d} \frac{R_n^d}{(\log n)^{\frac{d}{q}}} \int_{B_1} \left( 1 - \mathbb{1}_{d=2} \frac{V(R_n x)}{\log n} \right) + C \mathbb{1}_{p > d} \right) \end{aligned}$$

we can then conclude the proof of (1.8)–(1.9).

*Remark 1.8.* — In the case  $d = 2 < p$ , the integral  $\int_{B_{R_n}} \rho^{1-\frac{p}{2}}$  concentrates around  $\partial B_{R_n}$  where in each cube  $Q_j$  the number of points is of order one. This explains why there is no additional logarithms with respect to the case  $d \geq 3$ .

One simple but important observation of this work is that the sketch of proof above can be made rigorous for the lower bound without many changes, see Section 3. Indeed, for the boundary functional we do not need to impose condition (1.10) and can locally modify the mass thanks to Lemma 2.5. This leads to much simpler proofs than in earlier works on the subject where the boundary and standard Wasserstein distances were treated in a similar way, see [2, 10, 17]. While this is mostly of aesthetic nature in the case  $p \leq d$ , it turns out to be an important observation to treat the case  $p > d$ , see Remark 1.10 below.

For the proof of the upper bound, many changes need to be made to the sketch of proof above. First, as already understood in [19], in order to hope for (1.11)–(1.12) to hold we need to truncate the measure  $\rho$  on the ball  $B_{R_n}$ . While the construction of [19] works for  $p < d$  (and also for  $p = d = 2$  thanks to the extra logarithm, see [10]), it yields a transport cost which is too large for  $p \geq d$ . As alternative we essentially project the points outside  $B_{R_n}$  on  $\partial B_{R_n}$ , see Lemma 5.1. In particular, this truncation imposes the choice

$$V(R_n) + \alpha \log R_n = \beta \log n$$

where (compare with (1.13))

$$\begin{aligned} \beta & \begin{cases} \in (\frac{p}{d}, 1) & \text{if } p < d \\ = 1 & \text{if } p \geq d \end{cases} \\ \text{and } \alpha & \begin{cases} = 0 & \text{if } p < d \\ \in (d(q-1), 2q + d(q-1)) & \text{if } p = d = 2 \\ \in (d(q-1), q + d(q-1)) & \text{if } p = d \geq 3 \\ = d(q-1) & \text{if } p > d. \end{cases} \end{aligned} \quad (1.14)$$

The second and most important modification is due to the condition (1.10). Following [16] a way around it is to use triangle inequality and Young inequality, see also Lemma 2.2, and write for  $\varepsilon \in (0, 1)$ ,

$$\begin{aligned} \mathbb{E}[W_p^p(\mu_n, n\rho)] & \leq (1 + \varepsilon) \sum_j \mathbb{E} \left[ W_{Q_j}^p \left( \mu_n, \frac{\mu_n(Q_j)}{\rho(Q_j)} \rho \right) \right] \\ & \quad + \frac{C}{\varepsilon^{p-1}} \mathbb{E} \left[ W_p^p \left( \sum_j \frac{\mu_n(Q_j)}{\rho(Q_j)} \mathbb{1}_{Q_j} \rho, n\rho \right) \right]. \end{aligned} \quad (1.15)$$

The first term, which we sometimes call the local term, may be estimated following now exactly the sketch of proof while the difficulty is to prove that the second term, sometimes called the global term, is negligible. To prove that, one takes advantage on the radial symmetry. Indeed, thanks to this assumption, it is possible to split the estimate into a radial one and an angular one, as we are going to clarify below.

Since the estimate of the global term is usually delicate, see [2, 17, 21] and since we need to be very precise in particular for  $p \geq d$ , we had to resolve to modify the partition in order to take advantage of the radial symmetry of the measure.<sup>(1)</sup> A crucial requirement for the strategy to work is that all the elements of the partition have the same shape so that the rates of convergence in Theorem A.1 are uniform. To this aim we relax the condition of being a partition, see Lemma 2.2, and use instead coverings by sets of the form

$$\Delta_r^{\delta,z} = r\{1 \leq |x| \leq 1 + \delta : \hat{x} \in D_\delta(z)\}$$

where  $\hat{x} = x/|x|$  and  $D_\delta(z)$  denotes the image through the exponential map on  $\mathbb{S}^{d-1}$  of a  $(d-1)$ -dimensional cube on the tangent (the precise reason for such choice is technical, since it is common in the literature to argue with cubes). The estimate of the corresponding global term in Proposition 7.5 is divided into a radial, see Section 7.1 and an angular, see Section 7.2, transport. While for most values of  $p$  and  $d$  the order of subdivision (first radial and then angular or vice-versa) does not matter, for certain values it seems instead crucial (see Remark 7.6).

*Remark 1.9.* — Let us point out that a similar strategy for the upper bound was suggested in [23] for the case  $p = d = 2$ .

*Remark 1.10.* — In the case of the lower bound, one could argue as in the upper bound and use triangle inequality together with Young inequality to obtain the analogue of (1.15). The main difference would be the minus sign in front of the global term. Using that  $Wb \leq W$ , one could then use the estimates for the upper bounds to conclude. This would however not work when  $p > d$  since the local and the global terms are both of the order of  $\tau_n$ , see Proposition 7.5.

---

<sup>(1)</sup> Let us point out that for  $q < 1$  there is anyway the issue that in order to satisfy (1.11)–(1.12) at the same time it is not possible to take  $h = h_j$  independent of  $Q_j$ .

## 2. Notation and preliminary results

### 2.1. Notation

We identify measures which are absolutely continuous with respect to the Lebesgue measure with their density. Moreover, when integrating against the Lebesgue measure, and when it is clear from the context, we omit both integration variables and the measure in the integrals. We write  $|\Omega|$  for the Lebesgue measure of a set  $\Omega$ . We write  $\mathcal{H}^k$  for the Hausdorff measure of dimension  $k$ . We denote by  $\mathcal{M}(\Omega)$  the set of signed measures on  $\Omega$  and by  $\mathcal{M}^+(\Omega)$  the set of the positive measures on  $\Omega$ .

The notation  $A \lesssim B$ , which we use in output statements, means that there exists a universal (meaning depending possibly on  $d$  and  $p$ ) constant  $C > 0$  such that  $A \leq CB$ . We write  $A \lesssim_\varepsilon B$  if the implicit constant depends on the parameter  $\varepsilon$ . We write  $A \simeq B$  if  $A \lesssim B$  and  $B \lesssim A$ . We denote by  $\omega(t)$  a rate function which may vary from line to line such that  $\lim_{t \rightarrow 0} \omega(t) = \lim_{t \rightarrow \infty} \omega(t) = 0$ . We write  $\omega_\varepsilon$  to indicate that the rate function also depends on  $\varepsilon$ .

For  $d \geq 2$  we write  $\mathbb{S}^{d-1} = \partial B_1$  for the unit sphere and denote by

$$\sigma := \frac{1}{\mathcal{H}^{d-1}(\mathbb{S}^{d-1})} \mathcal{H}^{d-1} \llcorner \mathbb{S}^{d-1}$$

the uniform probability measure on  $\mathbb{S}^{d-1}$ . Moreover  $\mathbb{S}^{d-1}$  is endowed with the (Riemannian) spherical distance  $d_{\mathbb{S}^{d-1}}$ . We write  $\tilde{\mathbb{S}}^{d-1} := \mathbb{S}^{d-1} \times O(d-1)$ , where  $O(d-1)$  denotes the orthogonal group on  $\mathbb{R}^{d-1}$ . For  $z = (y, U) \in \tilde{\mathbb{S}}^{d-1}$  where  $p \in \mathbb{S}^{d-1}$ ,  $U \in O(d-1)$ , we define

$$D_\delta(z) := \exp_y(U(-\delta/2, \delta/2)^{d-1}) \subseteq \mathbb{S}^{d-1}. \quad (2.1)$$

where  $\exp_y$  denotes the exponential map on the sphere  $\mathbb{S}^{d-1}$  at  $y$ . For  $z = (e_1, \text{Id}_{d-1})$  we simply write  $D_\delta$ . Notice that we are implicitly identifying the tangent at  $y$  with  $\mathbb{R}^{d-1}$ , which can be done e.g. by viewing  $\mathbb{S}^{d-1}$  as a submanifold of  $\mathbb{R}^d$ . We also write  $\tilde{\sigma}$  for the uniform (Haar) probability measure on  $\tilde{\mathbb{S}}$ .

### 2.2. The Wasserstein distance

For  $\mu, \lambda \in \mathcal{M}^+(\mathbb{R}^d)$  with  $\mu(\mathbb{R}^d) = \lambda(\mathbb{R}^d)$ , we define the  $p$ -Wasserstein distance between  $\mu$  and  $\lambda$  as

$$W_p^p(\mu, \lambda) := \min \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\pi(x, y) \mid \pi_1 = \mu, \pi_2 = \lambda \right\}. \quad (2.2)$$

For a set  $\Omega$  we write

$$W_\Omega^p(\mu, \lambda) := W_p^p(\mu \llcorner \Omega, \lambda \llcorner \Omega). \quad (2.3)$$

We recall the following classical consequence of the Benamou–Brenier formula, see e.g. [22] or [16, Lemma 3.4].

**PROPOSITION 2.1.** — *Let  $\Omega$  be an open set and  $\mu, \lambda \in \mathcal{M}^+(\Omega)$  be such that  $\mu(\Omega) = \lambda(\Omega)$ . If  $\phi$  is a solution of*

$$\Delta\phi = \mu - \lambda \quad \text{in } \Omega \quad \text{and} \quad \nabla\phi \cdot \nu = 0 \quad \text{on } \partial\Omega,$$

where  $\nu$  denotes the exterior normal to  $\partial\Omega$ , then

$$W_p^p(\mu, \lambda) \lesssim \int_\Omega \frac{|\nabla\phi|^p}{\mu^{p-1}}.$$

We now state a simple subadditivity result for the Wasserstein distance whose proof is immediate. It extends more classical subadditivity results to the case of coverings which are not necessarily partitions. The last statement of the lemma comes from combining the second one with the property  $(a+b)^p \leq (1+\varepsilon)a^p + \frac{c}{\varepsilon^{p-1}}b^p$ .

**LEMMA 2.2.** — *Let  $(E, \mathcal{E}, \zeta)$  be a measure space and  $(\mu_z, \lambda_z)_{z \in E}$  a collection of measures on  $\Omega$  such that  $\mu_z(\Omega) = \lambda_z(\Omega)$ . Then*

$$W_p^p\left(\int_E \mu_z d\zeta(z), \int_E \lambda_z d\zeta(z)\right) \leq \int_E W_p^p(\mu_z, \lambda_z) d\zeta(z).$$

In particular, for  $\Omega \subseteq \mathbb{R}^d$ , if  $(\Omega_z)_{z \in E}$  is a collection of Borel subsets of  $\Omega$ , such that

$$\int_E \mathbb{1}_{\Omega_z} d\zeta(z) = 1 \quad \lambda + \mu - a.e. \quad \text{and} \quad \mu(\Omega_z) = \lambda(\Omega_z) \quad \text{for } \zeta \text{ a.e. } z,$$

then

$$W_p^p(\mu, \lambda) \leq \int_E W_{\Omega_z}^p(\mu, \lambda) d\zeta(z).$$

Therefore, letting  $\kappa := \mu(\Omega)/\lambda(\Omega)$  and  $\kappa_z := \mu(\Omega_z)/\lambda(\Omega_z)$ , there exists  $c = c(p, d) > 0$  such that for every  $\varepsilon \in (0, 1)$

$$\begin{aligned} W_p^p(\mu, \kappa\lambda) &\leq (1+\varepsilon) \int_E W_\Omega^p(\mu, \kappa_z\lambda) d\zeta(z) \\ &\quad + \frac{c}{\varepsilon^{p-1}} W_p^p\left(\int_E \kappa_z \lambda \mathbb{1}_{\Omega_z} d\zeta(z), \kappa\lambda\right). \end{aligned} \quad (2.4)$$

**Remark 2.3.** — To recover classical applications of subadditivity, see e.g. [16, Lemma 3.1], it is enough to apply the lemma with  $E$  finite and  $\zeta = \sum_z \delta_z$  and obtain

$$W_p^p\left(\sum_{z \in E} \mu_z, \sum_{z \in E} \lambda_z\right) \leq \sum_{z \in E} W_p^p(\mu_z, \lambda_z).$$

### 2.3. The boundary Wasserstein distance

For the lower bound we will use the boundary Wasserstein distance, that has been defined in [13] between  $\mu$  and  $\lambda$  measures on  $\Omega$  not necessarily of the same mass. It is defined as

$$Wb_{\Omega}^p(\mu, \lambda) := \min \left\{ \int_{\bar{\Omega} \times \bar{\Omega}} |x - y|^p d\pi(x, y) \mid \pi_1 \lfloor \Omega = \mu, \pi_2 \lfloor \Omega = \lambda \right\}. \quad (2.5)$$

As for  $W_{\Omega}$  we write for  $\Omega' \subset \Omega$ ,

$$Wb_{\Omega'}^p(\mu, \lambda) := Wb_{\Omega'}^p(\mu \lfloor \Omega', \lambda \lfloor \Omega').$$

We recall that  $Wb$  is superadditive, see e.g. [2] or [5, Lemma 25]. Let us point out that while the subadditivity of the classical Wasserstein distance is understood with respect to the measure arguments, the superadditivity of the boundary Wasserstein distance is understood with respect to the underlying domain.

LEMMA 2.4. — *Let  $\Omega \subset \mathbb{R}^d$  be an open set and let  $\Omega_j \subset \Omega$  be pairwise disjoint open sets then*

$$Wb_{\Omega}^p(\mu, \lambda) \geq \sum_j Wb_{\Omega_j}^p(\mu, \lambda). \quad (2.6)$$

We conclude with a simple lemma which will allow us to adjust the densities when working with  $Wb$ .

LEMMA 2.5. — *Let  $\Omega$  be a bounded set. Then for every  $m, m' > 0$*

$$Wb_{\Omega}^p(m, m') \lesssim \text{diam}(\Omega)^{p+d} \frac{|m - m'|^p}{\max(m, m')^{p-1}}. \quad (2.7)$$

*Proof.* — Since  $Wb_{\Omega} \leq Wb_{\text{Conv } \Omega}$  where  $\text{Conv } \Omega$  denotes the convex envelope of  $\Omega$ , we may assume without loss of generality that  $\Omega$  is convex. Let us first treat the one dimensional case. We may assume that  $\Omega = (0, L)$  for some  $L > 0$  and  $m \geq m'$ . We then set for  $x \in (0, L)$ ,

$$T(x) := \min \left( \frac{m}{m'} x, L \right)$$

so that  $T$  sends  $m \mathbb{1}_{(0,L)}$  on  $m' \mathbb{1}_{(0,L)} + (m - m')/L \delta_L$  and is thus  $(\text{Id} \times T)_{\#}$  ( $m \mathbb{1}_{(0,L)}$ ) is an admissible transport plan for  $Wb_{\Omega}^p(m, m')$ . We then compute

$$Wb_{\Omega}^p(m, m') \leq m \int_0^L |T - x|^p = m \int_0^{\frac{m'}{m} L} \left| \frac{m}{m'} - 1 \right|^p x^p + m \int_{\frac{m'}{m} L}^L (L - x)^p.$$

An elementary computation then yields the claimed

$$Wb_{\Omega}^p(m, m') \lesssim L^{p+1} \frac{|m - m'|^p}{m^{p-1}}. \quad (2.8)$$

Let us now conclude the proof in the case  $d \geq 2$ . For  $x' \in \mathbb{R}^{d-1}$  let

$$\Omega_{x'} = \{t \in \mathbb{R} : (t, x') \in \Omega\}$$

be the intersection of  $\Omega$  with the line  $\mathbb{R}e_1 + x'$ . Since  $\Omega$  is convex this is a segment. Letting  $\Omega' := \{x' \in \mathbb{R}^{d-1} : \Omega_{x'} \neq \emptyset\}$  and for  $x' \in \Omega_{x'}$ ,  $\pi_{x'}$  be an admissible transport plan for  $Wb_{\Omega_{x'}}^p(m, m')$  we see that the plan defined by

$$\int_{\bar{\Omega} \times \bar{\Omega}} \zeta d\pi := \int_{\Omega'} \int_{\bar{\Omega}_{x'} \times \bar{\Omega}_{x'}} \zeta(x' + te_1, x' + se_1) d\pi_{x'}(t, s) dx'$$

is admissible for  $Wb_{\Omega}^p(m, m')$  and thus

$$\begin{aligned} Wb_{\Omega}^p(m, m') &\leq \int_{\Omega'} Wb_{\Omega_{x'}}^p(m, m') dx' \stackrel{(2.8)}{\lesssim} \int_{\Omega'} \text{diam}(\Omega_{x'})^{p+1} \frac{|m - m'|^p}{m^{p-1}} \\ &\lesssim \text{diam}(\Omega)^{p+d} \frac{|m - m'|^p}{m^{p-1}}. \end{aligned}$$

This concludes the proof of (2.7).  $\square$

### 3. Proof of the lower bound

In this section we prove the lower bound part of Theorem 1.6. Let us recall the definition (1.4) of  $\tau_n$  and (1.7) of  $\underline{\ell}$ .

**PROPOSITION 3.1.** — *Let  $(X_i)_{i=1}^{\infty}$  be i.i.d. random variables of law  $\rho$ . Setting  $\mu_n := \sum_{i=1}^n \delta_{X_i}$ , we have for  $p \leq d$ ,*

$$\liminf_{n \rightarrow \infty} \frac{1}{\tau_n} \mathbb{E}[W_p^p(\mu_n, n\rho)] \geq \underline{\ell}^{p,d} \left( \int_{\mathbb{R}^d} \rho^{1-\frac{p}{d}} \mathbb{1}_{p < d} + \underline{\ell} \mathbb{1}_{p=d} \right) \quad (3.1)$$

and for  $p > d$ ,

$$\mathbb{E}[W_p^p(\mu_n, n\rho)] \gtrsim \tau_n. \quad (3.2)$$

*Proof.*

*Step 1.* — Let us first set some notation. For  $h > 0$  and  $z \in h\mathbb{Z}^d$ , we denote  $Q_z := z + (0, h)^d$ . For any  $Z \subset h\mathbb{Z}^d$ , we set

$$\Omega_Z := \bigcup_{z \in Z} Q_z. \quad (3.3)$$

Let  $\varepsilon > 0$  be fixed. We first claim that provided there exists  $m = m_{\varepsilon}$  large enough such that for  $z \in Z$ ,

$$hV'(|z|) \lesssim \varepsilon \quad (3.4)$$

and

$$nh^d \rho(|z|) \gtrsim_{\varepsilon} \exp(C\varepsilon m) \quad (3.5)$$

then

$$\mathbb{E}[W_p^p(\mu_n, n\rho)] \geq \left(1 - \omega(\varepsilon) - \max_{z \in Z} \omega(n\rho(Q_z))\right) \underline{\mathbf{c}}^{p,d} n^{1-\frac{p}{d}} \sum_{z \in Z} \int_{Q_z} \rho^{1-\frac{p}{d}} \left(1 + \mathbb{1}_{d=2}(\log n\rho(Q_z))^{\frac{p}{2}}\right). \quad (3.6)$$

Let us first prove the claim. To this aim we first use superadditivity, see Lemma 2.4, to obtain

$$\mathbb{E}[W_p^p(\mu_n, n\rho)] \geq \sum_{z \in Z} \mathbb{E}[Wb_{Q_z}^p(\mu_n, n\rho)].$$

Fix  $z \in Z$  and set  $N := \mu_n(Q_z)$ . Let  $\hat{X}_i$  be i.i.d. random variables of law  $\hat{\lambda} := \rho \mathbb{1}_{Q_z} / \rho(Q_z)$ . In this way  $\mu_n \llcorner Q_z$  has the same law as  $\hat{\mu}_N := \sum_{i=1}^N \delta_{\hat{X}_i}$ , and therefore we have

$$\mathbb{E}[Wb_{Q_z}^p(\mu_n, n\rho)] = \mathbb{E}\left[\mathbb{E}\left[Wb_{Q_z}^p(\hat{\mu}_k, n\rho) \mid N = k\right]\right].$$

By (3.4), we can use either the Knothe map as in [6, Lemma 1] or [2, Proposition 2.4] (in both cases the proof easily extend to general dimensions) to find a Lipschitz map  $T : Q_z \rightarrow Q_z$  such that  $T_{\#}\hat{\lambda} = \frac{1}{|Q_z|}$  and such that  $\text{Lip } T \leq 1 + \omega(\varepsilon)$ . Letting  $\mu'_k := \sum_{i=1}^k \delta_{T(\hat{X}_i)}$  we find for every  $k$ ,

$$\mathbb{E}[Wb_{Q_z}^p(\hat{\mu}_k, n\rho)] \geq (1 - \omega(\varepsilon)) \mathbb{E}\left[Wb_{Q_z}^p\left(\mu'_k, \frac{n\rho(Q_z)}{|Q_z|}\right)\right].$$

Using triangle inequality and (2.7) from Lemma 2.5, we find

$$\begin{aligned} Wb_{Q_z}^p(\hat{\mu}_k, n\rho) &\geq (1 - \omega(\varepsilon)) Wb_{Q_z}^p\left(\mu'_k, \frac{k}{|Q_z|}\right) - \frac{C}{\varepsilon^{p-1}} Wb_{Q_z}^p\left(\frac{k}{|Q_z|}, \frac{n\rho(Q_z)}{|Q_z|}\right) \\ &\geq (1 - \omega(\varepsilon)) Wb_{Q_z}^p\left(\mu'_k, \frac{k}{|Q_z|}\right) - \frac{C}{\varepsilon^{p-1}} |Q_z|^{\frac{p}{d}} \frac{|k - n\rho(Q_z)|^p}{(n\rho(Q_z))^{p-1}}. \end{aligned}$$

Taking expectation and using the definition (1.3) of  $\underline{\mathbf{c}}^{p,d}$  (recall also the definition (1.1) of  $\eta_n$ ) we find

$$\mathbb{E}[Wb_{Q_z}^p(\hat{\mu}_k, n\rho)] \geq (1 - \omega(\varepsilon) - \omega(k)) \underline{\mathbf{c}}^{p,d} |Q_z|^{\frac{p}{d}} \eta_k - \frac{C}{\varepsilon^{p-1}} |Q_z|^{\frac{p}{d}} \frac{|k - n\rho(Q_z)|^p}{(n\rho(Q_z))^{p-1}}.$$

Taking the expectation over  $N$  and using that  $\mathbb{E}[|N - n\rho(Q_z)|^p] \lesssim n\rho(Q_z)^{p/2}$  (recall that (3.5) holds and thus  $\mathbb{E}[N] = n\rho(Q_z) \gg 1$  provided  $m$  is large

enough) we get

$$\begin{aligned}
 & \mathbb{E}\left[Wb_{Q_z}^p(\mu_n, n\rho)\right] \\
 & \geq (1 - \omega(\varepsilon) - \omega(n\rho(Q_z))) \underline{\mathbf{c}}^{p,d} |Q_z|^{\frac{p}{d}} \mathbb{E}[\eta_N] - \frac{C}{\varepsilon^{p-1}} |Q_z|^{\frac{p}{d}} (n\rho(Q_z))^{1-\frac{p}{2}} \\
 & = (1 - \omega(\varepsilon) - \omega(n\rho(Q_z))) \underline{\mathbf{c}}^{p,d} |Q_z|^{\frac{p}{d}} \eta_{n\rho(Q_z)} - \frac{C}{\varepsilon^{p-1}} |Q_z|^{\frac{p}{d}} (n\rho(Q_z))^{1-\frac{p}{2}}.
 \end{aligned}$$

Recalling that

$$\eta_{n\rho(Q_z)} = (n\rho(Q_z))^{1-\frac{p}{d}} \left(1 + \mathbb{1}_{d=2}(\log n\rho(Q_z))^{\frac{p}{2}}\right)$$

we see that provided  $m$  is large enough (depending on  $\varepsilon$ ) we have

$$\frac{1}{\varepsilon^{p-1}} (n\rho(Q_z))^{1-\frac{p}{2}} \ll \eta_{n\rho(Q_z)}$$

so that

$$\begin{aligned}
 & \mathbb{E}\left[Wb_{Q_z}^p(\mu_n, n\rho)\right] \\
 & \geq (1 - \omega(\varepsilon) - \omega(n\rho(Q_z))) \underline{\mathbf{c}}^{p,d} |Q_z|^{\frac{p}{d}} \eta_{n\rho(Q_z)} \\
 & \geq (1 - \omega(\varepsilon) - \omega(n\rho(Q_z))) \underline{\mathbf{c}}^{p,d} n^{1-\frac{p}{d}} \int_{Q_z} \rho^{1-\frac{p}{d}} \left(1 + \mathbb{1}_{d=2}(\log(n\rho(Q_z)))^{\frac{p}{2}}\right).
 \end{aligned}$$

After summation this proves (3.6).

*Step 2.* — We now consider separately the cases  $p < d$ ,  $p = d$  and  $p > d$ .

*Step 2.1: The case  $p < d$ .* — Let  $R > 0$  be fixed, set

$$h = \varepsilon \min(1, R^{1-q}) \quad \text{and} \quad Z := \{z \in h\mathbb{R}^d : |z| \leq R\}.$$

In this case (3.4) holds and

$$\min_Z (nh^d \rho(|z|)) \gtrsim_R n \quad \text{and} \quad \log(n\rho(Q_z)) \geq \log n - C_R$$

so that (3.5) holds. Applying (3.6) we find (recall the definition (1.4) of  $\tau_n$ )

$$\mathbb{E}[W_p^p(\mu_n, n\rho)] \geq (1 - \omega(\varepsilon) - \omega_{\varepsilon,R}(n)) \underline{\mathbf{c}}^{p,d} \tau_n \int_{\Omega_Z} \rho^{1-\frac{p}{d}}.$$

Dividing by  $\tau_n$  and sending  $n \rightarrow \infty$  we find

$$\liminf_{n \rightarrow \infty} \frac{1}{\tau_n} \mathbb{E}[W_p^p(\mu_n, n\rho)] \geq (1 - \omega(\varepsilon)) \int_{B_R} \rho^{1-\frac{p}{d}}.$$

Sending  $\varepsilon \rightarrow 0$  and  $R \rightarrow \infty$  concludes the proof of (3.1) in this case.

*Step 2.2.* — In the case  $p = d$  we choose  $\alpha > \max(d(q-1), 0)$  and set  $R_n$  to be the unique solution to

$$V(R_n) + \alpha \log R_n = \log n.$$

We then set

$$h := \varepsilon \min(1, R_n^{1-q}) \quad \text{and} \quad Z := \{z \in h\mathbb{Z}^d : Q_z \cap B_{R_n} \neq \emptyset\}.$$

By definition of  $h$  and (1.5), we see that (3.4) holds. Let us check that also (3.5) holds. For this we see that for  $z \in Z$ ,

$$\begin{aligned} nh^d \rho(|z|) &\gtrsim nh^d \rho(R_n) = \varepsilon n \min(1, R_n^{d(1-q)}) \exp(-V(R_n)) \\ &= \varepsilon \min(1, R_n^{d(1-q)}) R_n^\alpha \end{aligned}$$

so that by our choice of  $\alpha$  we get that (3.5) holds and so that

$$\max_{z \in Z} \omega(n\rho(Q_z)) = \omega_\varepsilon(n).$$

By (3.6) and the fact that for  $d = 2$ ,

$$\sum_{z \in Z} \int_{Q_z} |\log|Q_z|| \lesssim |\Omega_Z| \log R_n \lesssim |B_{R_n}| \log R_n \ll \tau_n$$

we have

$$\begin{aligned} \mathbb{E}[W_p^p(\mu_n, n\rho)] &\geq (1 - \omega(\varepsilon) - \omega_\varepsilon(n)) \underline{\mathbf{c}}^{p,d} \int_{\Omega_Z} (1 + \mathbb{1}_{d=2} \log(n\rho)) \\ &\geq (1 - \omega(\varepsilon) - \omega_\varepsilon(n)) \underline{\mathbf{c}}^{p,d} \int_{B_{R_n}} (1 + \mathbb{1}_{d=2} \log(n\rho)). \end{aligned}$$

If  $d \geq 3$ , since  $\tau_n = (\log n)^{\frac{d}{q}}$  this concludes the proof of (3.1). If instead  $d = 2$ , we write

$$\int_{B_{R_n}} \log(n\rho) = \int_{B_{R_n}} (\log n - V(x)) = R_n^d \log n \int_{B_1} \left(1 - \frac{V(R_n x)}{\log n}\right)$$

concluding similarly the proof of (3.1).

*Step 2.3: The case  $p > d$ .* — We finally choose  $R_n$  to be the unique solution to

$$V(R_n) + d(q-1) \log R_n = \log n \tag{3.7}$$

and set for  $m \gg 1$ ,

$$h := \varepsilon R_n^{1-q} \quad \text{and} \quad Z := \{z \in h\mathbb{Z}^d : |z| \in [R_n - 4mh, R_n - mh]\}.$$

Since for  $z \in Z$ ,  $|z| \simeq R_n$  as above (3.4) is easily seen to hold. Let us check that also (3.5) is satisfied. For  $z \in Z$  we have

$$\begin{aligned} nh^d \rho(|z|) &\geq nh^d \rho(R_n - mh) = nh^d \exp(-V(R_n - mh)) \\ &= nh^d \exp(-V(R_n)) \exp\left(\int_{R_n - mh}^{R_n} V'(t)\right) \simeq_\varepsilon \exp\left(\int_{R_n - mh}^{R_n} V'(t)\right) \\ &\geq \exp(C(R_n^q - (R_n - mh)^q)) \geq \exp(Cm\varepsilon). \end{aligned}$$

Thus (3.5) holds. Applying (3.6) yields

$$\begin{aligned}
 \mathbb{E}[W_p^p(\mu_n, n\rho)] &\gtrsim_\varepsilon n^{1-\frac{p}{d}} \int_{\Omega_Z} \rho^{1-\frac{p}{d}} \\
 &\geq n^{1-\frac{p}{d}} \int_{R_n-3mh}^{R_n-2mh} r^{d-1} \rho^{1-\frac{p}{d}}(r) \gtrsim_m n^{1-\frac{p}{d}} R_n^{d-q} \rho(R_n)^{1-\frac{p}{d}} \\
 &= n^{1-\frac{p}{d}} R_n^{d-q} \exp\left(\left(\frac{p}{d}-1\right)V(R_n)\right) \stackrel{(3.7)}{=} R_n^{d-q-(q-1)(p-d)} \simeq \tau_n.
 \end{aligned}$$

This proves (3.2).  $\square$

## 4. Proof of the upper bound

### 4.1. The geometric setup

Motivated by the heuristics from the introduction, we set  $r_1 \geq 1$  be arbitrarily chosen such that in  $[r_1, \infty)$ , the function  $t^{d-q} \exp(-V(t))$  is decreasing and then for  $j \geq 1$ , set

$$r_j := (1 + \delta_j) r_{j-1} \quad (4.1)$$

where for  $\varepsilon \in (0, 1)$ ,

$$\delta_j := \varepsilon \begin{cases} r_{j-1}^{-q} & \text{if } d = 2 \text{ and } p \in [1, 2) \\ (\log n)^{-1} & \text{otherwise.} \end{cases} \quad (4.2)$$

For the distinction between the cases see Remark 7.6 in Section 7. It is not hard to check that in both cases  $r_j \rightarrow \infty$  as  $j \rightarrow \infty$ . When  $\delta_j = \varepsilon(\log n)^{-1}$  we simply write  $\delta$  for  $\delta_j$ . Let  $\bar{R}_n$  be the unique solution to

$$V(\bar{R}_n) + \alpha \log \bar{R}_n = \beta \log n \quad (4.3)$$

where we recall that  $\alpha$  and  $\beta$  are defined in (1.14). Notice that by (1.5), we have

$$\bar{R}_n \simeq (\log n)^{\frac{1}{q}}.$$

Let  $j_n$  be the first index such that

$$r_{j_n} \geq \bar{R}_n.$$

We then set

$$R_n := r_{j_n}, \quad \Omega_n := B_{R_n}, \quad R'_n := r_{j_n-1}, \quad \text{and} \quad \Omega'_n := B_{R'_n}. \quad (4.4)$$

It is an elementary computation to check that

$$|R_n - \bar{R}_n| \lesssim \varepsilon \bar{R}_n^{1-q} \quad \text{and} \quad \delta_{j_n} \simeq \varepsilon (\log n)^{-1} \simeq \varepsilon R_n^{-q}. \quad (4.5)$$

We define the truncated probability density  $\rho_n$  by

$$\rho_n := \left( \mathbb{1}_{\Omega'_n} + \frac{1 - \rho(\Omega'_n)}{\rho(\Omega_n \setminus \Omega'_n)} \mathbb{1}_{\Omega_n \setminus \Omega'_n} \right) \rho. \quad (4.6)$$

Let us observe that since  $(R_n - R'_n)R_n^{q-1} = \delta_{j_n} R_n^q \stackrel{(4.5)}{\simeq} \varepsilon$ , (1.11) is satisfied and thus

$$\frac{1 - \rho(\Omega'_n)}{\rho(\Omega_n \setminus \Omega'_n)} = \frac{\int_{R'_n}^{\infty} r^{d-1} \rho(r) dr}{\int_{R'_n}^{R_n} r^{d-1} \rho(r) dr} \simeq \frac{R_n^{d-q} \rho(R_n)}{R_n^{d-1} \rho(R_n) R_n \delta_{j_n}} = \frac{1}{R_n \delta_{j_n}} \simeq \frac{1}{\varepsilon}. \quad (4.7)$$

For  $0 \leq r < R \leq \infty$ ,  $z \in \tilde{\mathbb{S}}^{d-1}$  and  $\delta > 0$ , we define the set (recall the definition (2.1) of  $D_\delta(z)$ )

$$\Lambda_{r,R}^{z,\delta} = \{r \leq |x| \leq R : \hat{x} \in D_\delta(z)\} \quad (4.8)$$

where  $\hat{x} = x/|x|$ . When  $\delta \geq 2\pi$  so that  $D_\delta(z) = \mathbb{S}^{d-1}$ , we simply write

$$A_{r,R} := \{r \leq |x| \leq R\} = \Lambda_{r,R}^{z,\delta}. \quad (4.9)$$

Finally for  $z \in \tilde{\mathbb{S}}^{d-1}$  and  $2 \leq j \leq j_n$ , we define the quasi-cylindrical domains

$$\begin{aligned} \Delta_j^z &= \Lambda_{r_{j-1}, r_j}^{z, \delta_j} \\ &= \{x \in \mathbb{R}^d : r_{j-1} \leq |x| \leq (1 + \delta_j)r_{j-1} \text{ and } \hat{x} \in D_{\delta_j}(z)\}. \end{aligned} \quad (4.10)$$

## 4.2. The upper bound

In this section we prove the upper bound part of Theorem 1.6. We recall the definition (1.4) of  $\tau_n$  and (1.7) of  $\bar{\ell}$ .

**PROPOSITION 4.1.** — *Let  $(X_i)_{i=1}^\infty$  be i.i.d. random variables of law  $\rho$ . Setting  $\mu_n := \sum_{i=1}^n \delta_{X_i}$ , we have for  $p \leq d$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{\tau_n} \mathbb{E}[W_p^p(\mu_n, n\rho)] \leq \bar{\mathbf{c}}_{p,d} \left( \int_{\mathbb{R}^d} \rho^{1-\frac{p}{d}} \mathbb{1}_{p < d} + \bar{\ell} \mathbb{1}_{p=d} \right) \quad (4.11)$$

and for  $p > d$ ,

$$\mathbb{E}[W_p^p(\mu_n, n\rho)] \lesssim \tau_n. \quad (4.12)$$

*Proof.* — Let  $T$  be the optimal transport map for  $W_p(\rho, \rho_n)$ . Let  $\hat{X}_i := T(X_i)$  and then  $\hat{\mu}_n = T_\# \mu_n = \sum_{i=1}^n \delta_{\hat{X}_i}$ . Notice that  $\hat{X}_i$  are i.i.d. random variables with law  $\rho_n$  and  $\hat{\mu}_n \ll \Omega'_n = \mu_n \ll \Omega_n$  (recall (4.4)). Since

$$\mathbb{E}[W_p^p(\mu_n, \hat{\mu}_n)] \leq n W_p^p(\rho, \rho_n),$$

we have by triangle inequality and Young that for  $\varepsilon \in (0, 1)$ ,

$$\mathbb{E}[W_p(\mu_n, n\rho)] \leq (1 + \varepsilon) \mathbb{E}[W_p^p(\hat{\mu}_n, n\rho_n)] + \frac{C}{\varepsilon^{p-1}} n W_p^p(\rho, \rho_n).$$

Since  $nW_p^p(\rho, \rho_n) = o(\tau_n) + \mathbb{1}_{p>d}O(\tau_n)$  by Lemma 5.1 in Section 5, we are only left with estimating  $\mathbb{E}[W_p^p(\widehat{\mu}_n, n\rho_n)]$ . To this aim, we set

$$\kappa_1 := \frac{\widehat{\mu}_n(B_{r_1})}{\rho_n(B_{r_1})} \quad \text{and} \quad \kappa_j^z := \frac{\widehat{\mu}_n(\Delta_j^z)}{\rho_n(\Delta_j^z)}. \quad (4.13)$$

Using the subadditivity inequality (2.4) from Lemma 2.2 on  $E = \mathbb{N} \times \widetilde{\mathbb{S}}^{d-1}$  with

$$\zeta := \sum_{j=1}^{j_n} \frac{\mathbb{1}_j \otimes d\widetilde{\sigma}(z)}{\widetilde{\sigma}_j}, \quad \text{where} \quad \widetilde{\sigma}_j = \int_{\widetilde{\mathbb{S}}^{d-1}} \mathbb{1}_{D_{\delta_j}(z)}(e_1) d\widetilde{\sigma}(z)$$

and

$$\Omega_{1,z} := B_{r_1}, \quad \text{and} \quad \Omega_{j,z} := \Delta_j^z \quad \text{for } j \geq 2$$

we find

$$\begin{aligned} \mathbb{E}[W_p^p(\widehat{\mu}_n, n\rho_n)] &\leq (1 + \varepsilon) \\ &\times \left( \mathbb{E}[W_{B_{r_1}}^p(\widehat{\mu}_n, \kappa_1 \rho_n)] + \sum_{j=2}^{j_n} \frac{1}{\widetilde{\sigma}_j} \int_{\widetilde{\mathbb{S}}^{d-1}} \mathbb{E}[W_{\Delta_j^z}^p(\widehat{\mu}_n, \kappa_j^z \rho_n)] d\widetilde{\sigma}(z) \right) \\ &+ \frac{C}{\varepsilon^{p-1}} \mathbb{E} \left[ W_p^p \left( \kappa_1 \mathbb{1}_{B_{r_1}} \rho_n + \sum_{j=2}^{j_n} \frac{1}{\widetilde{\sigma}_j} \int_{\widetilde{\mathbb{S}}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\widetilde{\sigma}(z), n\rho_n \right) \right]. \end{aligned}$$

Let us first consider the case  $p \leq d$ . In this case, applying (6.6) from Proposition 6.2 (in Section 6) and (7.17) from Proposition 7.5 (in Section 7) we obtain (recall the definition (1.1) of  $\eta_n$ )

$$\begin{aligned} \mathbb{E}[W_p^p(\widehat{\mu}_n, n\rho_n)] &\leq (1 + \omega(\varepsilon) + \omega(n)) \bar{\mathbf{c}}_{p,d} \eta_n \\ &\int_{\Omega_n} \rho_n^{1-\frac{p}{d}} \left( 1 + \mathbb{1}_{d=2} \left( \frac{\log n \rho_n}{\log n} \right)^{\frac{p}{2}} \right) + o_\varepsilon(\tau_n). \end{aligned}$$

Using that for  $p \leq d$ ,

$$\begin{aligned} \eta_n \int_{\Omega_n} \rho_n^{1-\frac{p}{d}} \left( 1 + \mathbb{1}_{d=2} \left( \frac{\log n \rho_n}{\log n} \right)^{\frac{p}{2}} \right) \\ \leq (1 + \omega_\varepsilon(n)) \tau_n \int_{\mathbb{R}^d} \rho^{1-\frac{p}{d}} (\mathbb{1}_{p<d} + \bar{\ell} \mathbb{1}_{p=d}) \end{aligned}$$

this concludes the proof of (4.11) sending first  $n \rightarrow \infty$  and then  $\varepsilon \rightarrow 0$ .

In the case  $p > d$ , we use (6.7) from Proposition 6.2 (in Section 6) and (7.17) from Proposition 7.5 (in Section 7) to obtain

$$\mathbb{E}[W_p^p(\widehat{\mu}_n, n\rho_n)] \lesssim_\varepsilon n^{1-\frac{p}{d}} \int_{\Omega_n} \rho_n^{1-\frac{p}{d}} + \tau_n.$$

Since

$$\begin{aligned}
 n^{1-\frac{p}{d}} \int_{\Omega_n} \rho_n^{1-\frac{p}{d}} &\simeq_\varepsilon n^{1-\frac{p}{d}} \int_0^{R_n} r^{d-1} \rho(r)^{1-\frac{p}{d}} \\
 &= n^{1-\frac{p}{d}} \rho(R_n)^{1-\frac{p}{d}} \int_0^{R_n} r^{d-1} \exp\left(-\left(\frac{p}{d}-1\right) \int_r^{R_n} V'(t) dt\right) dr \\
 &\stackrel{(1.5)}{\simeq} (n\rho(R_n))^{1-\frac{p}{d}} \int_0^{R_n} r^{d-1} \exp(-C(R_n^q - r^q)) dr \\
 &\simeq R_n^{d-q} (n\rho(R_n))^{1-\frac{p}{d}} \stackrel{(4.5)}{\simeq} \tau_n,
 \end{aligned} \tag{4.14}$$

this concludes the proof of (4.12).  $\square$

## 5. Cost of the cut-off

In this section we estimate the distance between  $\rho$  and  $\rho_n$ , with  $\rho_n$  defined in (4.6). We also recall the definition (1.4) of  $\tau_n$ .

LEMMA 5.1. — *Let  $\rho_n$  be the probability measure defined in (4.6). We have*

$$nW_p^p(\rho, \rho_n) = o(\tau_n) + \mathbb{1}_{p>d}O(\tau_n). \tag{5.1}$$

*Proof.* — By subadditivity,

$$W_p^p(\rho, \rho_n) \leq W_p^p\left(\mathbb{1}_{(\Omega'_n)^c} \rho, \frac{1 - \rho(\Omega'_n)}{\rho(\Omega_n \setminus \Omega'_n)} \mathbb{1}_{\Omega_n \setminus \Omega'_n} \rho\right).$$

Let

$$\hat{\rho}_n = \frac{1 - \rho(\Omega'_n)}{\mathcal{H}^{d-1}(\partial\Omega'_n)} \mathcal{H}^{d-1} \llcorner \partial\Omega'_n.$$

Then by triangle inequality,

$$\begin{aligned}
 W_p^p\left(\mathbb{1}_{(\Omega'_n)^c} \rho, \frac{1 - \rho(\Omega'_n)}{\rho(\Omega_n \setminus \Omega'_n)} \mathbb{1}_{\Omega_n \setminus \Omega'_n} \rho\right) \\
 \lesssim W_p^p(\mathbb{1}_{(\Omega'_n)^c} \rho, \hat{\rho}_n) + W_p^p\left(\hat{\rho}_n, \frac{1 - \rho(\Omega'_n)}{\rho(\Omega_n \setminus \Omega'_n)} \mathbb{1}_{\Omega_n \setminus \Omega'_n} \rho\right).
 \end{aligned}$$

We now estimate both terms separately using that by uniqueness of the optimal map and rotational invariance of the measures, the optimal transport map is also radially symmetric. For the first term we have

$$W_p^p(\mathbb{1}_{(\Omega'_n)^c} \rho, \hat{\rho}_n) \lesssim \int_{R'_n}^\infty r^{d-1} |R'_n - r|^p \rho(r) \stackrel{(4.5)}{\lesssim} \int_{R_n}^\infty r^{d-1} |R_n - r|^p \rho(r).$$

We then estimate

$$\begin{aligned}
 & \int_{R_n}^{\infty} |R_n - r|^p r^{d-1} \rho(r) dr \\
 &= \exp(-V(R_n)) \int_{R_n}^{\infty} |R_n - r|^p r^{d-1} \exp\left(-\int_{R_n}^r V'(t) dt\right) dr \\
 &\leq \rho(R_n) \int_{R_n}^{\infty} |R_n - r|^p r^{d-1} \exp(-C(r^q - R_n^q)) \\
 &\lesssim R_n^{d-q+p(1-q)} \rho(R_n).
 \end{aligned}$$

For the second term we may brutally estimate the transport distance by  $R_n - R'_n = \delta_{j_n} R'_n$  (recall (4.1)) so that

$$\begin{aligned}
 W_p^p\left(\hat{\rho}_n, \frac{1 - \rho(\Omega'_n)}{\rho(\Omega_n \setminus \Omega'_n)} \mathbb{1}_{\Omega_n \setminus \Omega'_n} \rho\right) &\lesssim (\delta_{j_n} R'_n)^p (1 - \rho(\Omega'_n)) \\
 &\stackrel{(4.5)}{\lesssim} R_n^{p(1-q)} \int_{R_n}^{\infty} r^{d-1} \rho(r) \\
 &\lesssim R_n^{d-q+p(1-q)} \rho(R_n).
 \end{aligned}$$

Recalling the definition (1.14) of  $\alpha, \beta$  as well as (4.3) and (4.5), we thus conclude that

$$nW_p^p(\rho, \rho_n) \lesssim nR_n^{d-q+p(1-q)} \rho(R_n) \lesssim \begin{cases} (\log n)^{\frac{d+p}{q}-1-p} n^{1-\beta} & \text{if } p < d \\ (\log n)^{\frac{d}{q}-p(1-\frac{1}{q})-1+\frac{\alpha}{q}} & \text{if } p = d \\ (\log n)^{d-1-p(1-\frac{1}{q})} & \text{if } p > d. \end{cases}$$

Recalling the definition (1.4) of  $\tau_n$  and the range (1.14) of  $\alpha$  in the case  $p = d$  we conclude the proof of (5.1).  $\square$

## 6. The local terms

For  $r, \delta > 0$  we set (recall the definition (2.1) of  $D_\delta$ )

$$\Delta_r^\delta := r\{1 \leq |x| \leq 1 + \delta : \hat{x} \in D_\delta\}$$

where  $\hat{x} = x/|x|$ . Recall the definitions (1.2) of  $\bar{\mathbf{c}}_{p,d}$  and (1.3) of  $\underline{\mathbf{c}}^{p,d}$  as well as the definition (1.1) of  $\eta_n$ . Our main building block is a variant of Theorem A.1 with a more explicit dependence on the parameters. Let us however stress that the main point in Proposition 6.1 is that  $\Delta_r^\delta$  is almost isometric to a cube while the difficulty in Theorem A.1 is to cover arbitrary domains.

PROPOSITION 6.1. — *There exists  $\omega : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  with  $\lim_{t \rightarrow 0} \omega(t) = \lim_{t \rightarrow \infty} \omega(t) = 0$  with the following property. Let  $r \geq 1$ ,  $\delta \in (0, 1)$  and  $\lambda$  be a Lipschitz continuous probability density on  $\Delta_r^\delta$  such that*

$$\ell r \delta \leq 1,$$

where  $\ell := \|\lambda\|_{\text{Lip}} |\Delta_r^\delta|$ . For every  $n \geq 2$ , if  $(X_i)_{i=1}^n$  are i.i.d. random variables on  $\Delta_r^\delta$  with law  $\lambda$ , denoting  $\mu_n = \sum_{i=1}^n \delta_{X_i}$ , we have

$$\mathbb{E} \left[ W_{\Delta_r^\delta}^p(\mu_n, n\lambda) \right] \leq (1 + \omega(\ell r \delta) + \omega(\delta) + \omega(n)) \bar{c}_{p,d} \eta_n \int_{\Delta_r^\delta} \lambda^{1-\frac{p}{d}}. \quad (6.1)$$

*Proof.* — Setting  $\lambda_r(x) := r^d \lambda(rx)$  we see by scaling that it is enough to prove (6.1) with  $r = 1$ . To lighten notation we set  $\Delta^\delta := \Delta_1^\delta$ . The map

$$\Phi : D_\delta \times (0, \delta) \longrightarrow \mathbb{R}^d, \quad (z, s) \longmapsto (1+s)z \in \Delta^\delta$$

satisfies  $\|\nabla \Phi - \text{Id}_{\mathbb{R}^d}\| \lesssim \delta$ . Moreover, by the very definition of  $D_\delta$ , the exponential map on  $\mathbb{S}^{d-1}$  provides a diffeomorphism

$$\exp_{e_1} : (-\delta/2, \delta/2)^{d-1} \subseteq \mathbb{R}^{d-1} \longrightarrow D_\delta$$

such that  $\|\nabla \exp_{e_1} - \text{Id}_{\mathbb{R}^d}\| \lesssim \delta$ . Letting

$$C_\delta := (-\delta/2, \delta/2)^{d-1} \times (0, \delta),$$

it follows that the map  $F : \Delta^\delta \mapsto C_\delta$  defined through

$$F^{-1}(y, s) := \Phi(\exp_{e_1}(y), s) = (1+s)\exp_{e_1}(y)$$

is a diffeomorphism. Moreover, it satisfies

$$\|\nabla F - \text{Id}_{\mathbb{R}^d}\|_\infty + \|\nabla^2 F\|_\infty \lesssim \delta$$

and the same estimate also holds for  $F^{-1}$ . In particular

$$|\Delta^\delta| - |C_\delta| \leq \int_{C_\delta} |1 - \det \nabla F^{-1}| \lesssim \delta |C_\delta|. \quad (6.2)$$

Now we define

$$\widehat{\lambda} := F_\# \lambda \quad \text{and} \quad \widehat{\mu}_n := F_\# \mu_n.$$

Notice that  $\widehat{\mu}_n = \sum_{i=1}^n \delta_{\widehat{X}_i}$  where  $\widehat{X}_i = F(X_i)$  are i.i.d. random variables of law  $\widehat{\lambda}$ . We have

$$\widehat{\ell} := \|\widehat{\lambda}\|_{\text{Lip}} |C_\delta| \lesssim \ell + \delta \quad \text{and} \quad (6.3)$$

$$\begin{aligned} \mathbb{E}[W_{\Delta^\delta}^p(\mu_n, n\lambda)] &\leq \|F^{-1}\|_{\text{Lip}}^p \mathbb{E}[W_{C_\delta}^p(\widehat{\mu}_n, n\widehat{\lambda})] \\ &\leq (1 + C\delta) \mathbb{E}[W_{C_\delta}^p(\widehat{\mu}_n, n\widehat{\lambda})]. \end{aligned} \quad (6.4)$$

We then set for  $x \in C_1$ ,  $\lambda'(x) = \widehat{\lambda}(\delta x)\delta^d$  and  $\mu'_n = \sum_{i=1}^n \delta_{X'_i}$  where  $X'_i = \delta^{-1}\widehat{X}_i$  are i.i.d. random variables of law  $\lambda'$ . By scaling we have

$$\begin{aligned} \ell' &:= \|\lambda'\|_{\text{Lip}}|C_1| = \delta\widehat{\ell} \lesssim \delta\ell + \delta^2 \\ \text{and} \quad \mathbb{E}\left[W_{C_\delta}^p(\widehat{\mu}_n, n\widehat{\lambda})\right] &= \delta^p \mathbb{E}\left[W_{C_1}^p(\mu'_n, n\lambda')\right]. \end{aligned} \quad (6.5)$$

Since  $C_1$  is a cube, we can use either the Knothe map as in [6, Lemma 1] or [2, Proposition 2.4] to obtain a map  $T$  transporting  $\lambda'$  on the uniform measure on  $C_1$  with

$$\|T\|_{\text{Lip}}, \|T^{-1}\|_{\text{Lip}} \leq 1 + O(\ell').$$

Arguing exactly as above and setting  $\widetilde{\mu}_n = \sum_{i=1}^n \delta_{T(X'_i)}$  (so that  $T(X'_i)$  are i.i.d. uniformly distributed in  $C_1$ ), we find using the definition of  $\eta_n$ ,

$$\begin{aligned} \mathbb{E}\left[W_{C_1}^p(\mu'_n, n\lambda')\right] &\leq (1 + O(\ell')) \mathbb{E}\left[W_{C_1}^p\left(\widetilde{\mu}_n, \frac{n}{|C_1|}\right)\right] \\ &\leq (1 + \omega(\ell') + \omega(n)) \bar{\mathbf{c}}_{p,d} \eta_n |C_1|^{\frac{p}{d}}. \end{aligned}$$

Combining this with (6.3), (6.5) and (6.2), we obtain

$$\mathbb{E}\left[W_{\Delta_\delta}^p(\mu_n, n\lambda)\right] \leq (1 + \omega(\delta\ell) + \omega(\delta) + \omega(n)) \bar{\mathbf{c}}_{p,d} \eta_n |\Delta^\delta|^{\frac{p}{d}}.$$

Using that  $\lambda|\Delta^\delta|$  is  $\ell$ -Lipschitz continuous, we can replace  $|\Delta^\delta|^{p/d}$  by  $\int_{\Delta^\delta} \lambda^{1-p/d}$  at a cost of order  $\omega(\delta\ell)$  in the previous inequality thus concluding the proof of (6.1).  $\square$

We recall the following notation from the proof of Proposition 4.1. For  $\widehat{X}_i$  i.i.d. with common law  $\rho_n$ , we set  $\widehat{\mu}_n := \sum_{i=1}^n \delta_{\widehat{X}_i}$  and then

$$\kappa_1 := \frac{\widehat{\mu}_n(B_{r_1})}{\rho_n(B_{r_1})} \quad \text{and} \quad \kappa_j^z := \frac{\widehat{\mu}_n(\Delta_j^z)}{\rho_n(\Delta_j^z)}.$$

PROPOSITION 6.2. — *We have for  $p \leq d$*

$$\begin{aligned} &\mathbb{E}\left[W_{B_{r_1}}^p(\widehat{\mu}_n, \kappa_1 \rho_n)\right] + \sum_{j=2}^{j_n} \frac{1}{\bar{\sigma}_j} \int_{\mathbb{S}^{d-1}} \mathbb{E}\left[W_{\Delta_j^z}^p(\widehat{\mu}_n, \kappa_j^z \rho_n)\right] d\bar{\sigma}(z) \\ &\leq (1 + \omega(\varepsilon) + \omega(n)) \bar{\mathbf{c}}_{p,d} \eta_n \int_{\Omega_n} \rho_n^{1-\frac{p}{d}} \left(1 + \mathbb{1}_{d=2} \left(\frac{\log n \rho_n}{\log n}\right)^{\frac{p}{2}}\right) \quad (6.6) \end{aligned}$$

and for  $p > d$

$$\begin{aligned} \mathbb{E}\left[W_{B_{r_1}}^p(\widehat{\mu}_n, \kappa_1 \rho_n)\right] + \sum_{j=2}^{j_n} \frac{1}{\bar{\sigma}_j} \int_{\mathbb{S}^{d-1}} \mathbb{E}\left[W_{\Delta_j^z}^p(\widehat{\mu}_n, \kappa_j^z \rho_n)\right] d\tilde{\sigma}(z) \\ \lesssim_\varepsilon n^{1-\frac{p}{d}} \int_{\Omega_n} \rho_n^{1-\frac{p}{d}}. \end{aligned} \quad (6.7)$$

*Proof.* — We set

$$n_1 := \widehat{\mu}_n(B_{r_1}) \quad \text{and} \quad n_j^z := \widehat{\mu}_n(\Delta_j^z).$$

In  $B_{r_1}$  we use Theorem A.2 to obtain

$$\begin{aligned} \mathbb{E}\left[W_{B_{r_1}}^p(\widehat{\mu}_n, \kappa_1 \rho_n)\right] &= \mathbb{E}\left[W_{B_{r_1}}^p\left(\widehat{\mu}_n, n_1 \frac{\rho}{\rho(B_{r_1})}\right)\right] \\ &\leq \mathbb{E}\left[(1 + \omega(n_1)) \left(\frac{n_1}{\rho(B_{r_1})}\right)^{1-\frac{p}{d}} \left(1 + \mathbb{1}_{d=2}(\log n_1)^{\frac{p}{2}}\right)\right] \bar{\mathbf{c}}_{p,d} \int_{B_{r_1}} \rho_n^{1-\frac{p}{d}}. \end{aligned}$$

Since  $n_1$  is a binomial random variable with exponential concentration around its expectation  $\mathbb{E}[n_1] = n\rho(B_{r_1})$ ,

$$\begin{aligned} \mathbb{E}\left[W_{B_{r_1}}^p(\widehat{\mu}_n, \kappa_1 \rho_n)\right] &\leq (1 + \omega(n)) \bar{\mathbf{c}}_{p,d} \eta_n \int_{B_{r_1}} \rho_n^{1-\frac{p}{d}} \\ &\leq (1 + \omega(n)) \bar{\mathbf{c}}_{p,d} n^{1-\frac{p}{d}} \int_{B_{r_1}} \rho_n^{1-\frac{p}{d}} \left(1 + \mathbb{1}_{d=2}(\log n \rho_n)^{\frac{p}{2}}\right). \end{aligned} \quad (6.8)$$

We now consider a fixed set  $\Delta_j^z$ . Since  $\lambda_j^z := \rho_n / \rho_n(\Delta_j^z)$  is Lipschitz continuous in  $\Delta_j^z$  with

$$\ell_j^z := \|\lambda_j^z\|_{\text{Lip}} |\Delta_j^z| \lesssim \sup_{\Delta_j^z} |\nabla V| \lesssim r_j^{q-1}$$

we find

$$\ell_j^z r_j \delta_j \lesssim r_j^q \delta_j \stackrel{(4.2) \& (4.5)}{\lesssim} \varepsilon. \quad (6.9)$$

We may thus apply (6.1) from Proposition 6.1 and obtain

$$\begin{aligned} \mathbb{E}\left[W_{\Delta_j^z}^p(\widehat{\mu}_n, \kappa_j^z \rho_n)\right] \\ \leq \bar{\mathbf{c}}_{p,d} \mathbb{E}\left[\left(1 + \omega(\varepsilon) + \omega(n_j^z)\right) \left(\frac{n_j^z}{\rho_n(\Delta_j^z)}\right)^{1-\frac{p}{d}} \left(1 + \mathbb{1}_{d=2}(\log(1 + n_j^z))^{\frac{p}{2}}\right) \mathbb{1}_{n_j^z \geq 1}\right] \int_{\Delta_j^z} \rho_n^{1-\frac{p}{d}}. \end{aligned}$$

We now observe that  $n_j^z$  is a binomial random variable with

$$\mathbb{E}[n_j^z] = n\rho_n(\Delta_j^z) \simeq n\rho_n(r_j)(r_j \delta_j)^d \stackrel{(4.2)}{\gtrsim_\varepsilon} n\rho(R_n) R_n^{d(1-q)}. \quad (6.10)$$

From (4.5) and (4.3) we get (recall the definition (1.14) of  $\alpha$  and  $\beta$ )

$$n\rho(R_n)R_n^{d(1-q)} \simeq \begin{cases} n^{1-\beta}(\log n)^{\frac{d}{q}(1-q)} & \text{if } p < d \\ (\log n)^{\frac{1}{q}(\alpha-d(q-1))} & \text{if } p = d \\ 1 & \text{if } p > d. \end{cases}$$

We now argue separately in the cases  $p \leq d$  and  $p > d$ . The case  $d = 2$  requires a little more care so we first consider the case  $d \geq 3$ . Since  $p \leq d$ , we have  $\lim_{n \rightarrow \infty} \inf_{j,z} \mathbb{E}[n_j^z] = \infty$  and by the good concentration properties of binomial random variables we have

$$\mathbb{E} \left[ \left( 1 + \omega(\varepsilon) + \omega(n_j^z) \right) \left( \frac{n_j^z}{\rho(\Delta_j^z)} \right)^{1-\frac{p}{d}} \middle| n_j^z \geq 1 \right] \leq (1 + \omega(\varepsilon) + \omega_\varepsilon(n)) n^{1-\frac{p}{d}}.$$

This proves that for  $p \leq d$ ,

$$\mathbb{E} \left[ W_{\Delta_j^z}^p(\hat{\mu}_n, \kappa_j^z \rho_n) \right] \leq \bar{c}_{p,d} (1 + \omega(\varepsilon) + \omega_\varepsilon(n)) n^{1-\frac{p}{d}} \int_{\Delta_j^z} \rho_n^{1-\frac{p}{d}}. \quad (6.11)$$

After summation and integration, combining (6.8) and (6.11) concludes the proof of (6.6). When  $p < d = 2$ , since  $n\rho_n(\Delta_j^z) = C_\varepsilon n^{1-\beta}$  for  $\Delta_j^z$  close to  $\partial B_{R_n}$  we need to be a bit more careful. Letting  $\hat{R}_n := \log \log n$  (this choice is somewhat arbitrary) we have for  $\Delta_j^z \subset B_{\hat{R}_n}$ ,  $\log(n\rho_n(\Delta_j^z)) = (1 + \omega_\varepsilon(n)) \log n$  while for  $\Delta_j^z \cap A_{\hat{R}_n, R_n} \neq \emptyset$  (recall definition (4.9)),  $\log(n\rho_n(\Delta_j^z)) = O_\varepsilon(\log n)$  so that the same argument as above gives

$$\begin{aligned} \mathbb{E} \left[ W_{B_{r_1}}^p(\hat{\mu}_n, \kappa_1 \rho_n) \right] &+ \sum_{j=2}^{j_n} \frac{1}{\tilde{\sigma}_j} \int_{\mathbb{S}^1} \mathbb{E} \left[ W_{\Delta_j^z}^p(\hat{\mu}_n, \kappa_j^z \rho_n) \right] d\tilde{\sigma}(z) \\ &\leq (1 + \omega(\varepsilon) + \omega(n)) \bar{c}_{p,d} n^{1-\frac{p}{2}} (\log n)^{\frac{p}{2}} \left( \int_{B_{\hat{R}_n}} \rho_n^{1-\frac{p}{2}} + C_\varepsilon \int_{A_{\hat{R}_n, R_n}} \rho_n^{1-\frac{p}{2}} \right) \\ &\leq (1 + \omega(\varepsilon) + \omega(n)) \bar{c}_{p,d} n^{1-\frac{p}{2}} \int_{\Omega_n} \rho_n^{1-\frac{p}{2}} (\log n \rho_n)^{\frac{p}{2}}. \end{aligned}$$

Thus (6.6) also holds in this case. If  $p = d = 2$  we have

$$\mathbb{E} \left[ W_{\Delta_j^z}^2(\hat{\mu}_n, \kappa_j^z \rho_n) \right] \leq \bar{c}_{p,d} (1 + \omega(\varepsilon) + \omega_\varepsilon(n)) \int_{\Delta_j^z} \log(n\rho_n |\Delta_j^z|).$$

Since it is readily checked that

$$\sum_{j=2}^{j_n} \frac{1}{\tilde{\sigma}_j} \int_{\mathbb{S}^1} |\Delta_j^z| \log |\Delta_j^z| d\tilde{\sigma}(z) = o(\eta_n),$$

after summation we conclude again the proof of (6.6).

In the case  $p > d$ , since  $\inf_{j,z} \mathbb{E}[n_j^z] \gtrsim_\varepsilon 1$ , we can still conclude that

$$\begin{aligned} \mathbb{E} \left[ \left( 1 + \omega(\varepsilon) + \omega(n_j^z) \right) \left( \frac{n_j^z}{\rho(\Delta_j^z)} \right)^{1-\frac{p}{d}} \left( 1 + \mathbb{1}_{d=2} (\log(1 + n_j^z))^{\frac{p}{2}} \right) \middle| n_j^z \geq 1 \right] \\ \lesssim_\varepsilon n^{1-\frac{p}{d}} \left( 1 + \mathbb{1}_{d=2} (\log(1 + n\rho_n(\Delta_j^z)))^{\frac{p}{2}} \right). \end{aligned}$$

Therefore

$$\mathbb{E} \left[ W_{\Delta_j^z}^p(\hat{\mu}_n, \kappa_j^z \rho_n) \right] \lesssim_\varepsilon n^{1-\frac{p}{d}} \left( 1 + \mathbb{1}_{d=2} (\log(1 + n\rho_n(\Delta_j^z)))^{\frac{p}{2}} \right) \int_{\Delta_j^z} \rho_n^{1-\frac{p}{d}}. \quad (6.12)$$

In the case  $d > 2$ , after summation and integration, combining (6.8) and (6.12) concludes the proof of (6.7). We are left with the case  $p > d = 2$ . Using that  $\rho_n(\Delta_j^z) \simeq_\varepsilon \rho(r_j) r_j^d (\log n)^{-d}$  we have

$$\begin{aligned} \sum_{j=2}^{j_n} \frac{1}{\tilde{\sigma}_j} \int_{\tilde{\mathbb{S}}^{d-1}} (\log(1 + n\rho_n(\Delta_j^z)))^{\frac{p}{2}} \int_{\Delta_j^z} \rho_n^{1-\frac{p}{d}} d\tilde{\sigma}(z) \\ \lesssim_\varepsilon \int_{\Omega_n \setminus B_{r_1}} \left( \log \left( 1 + n \left( \frac{|x|}{\log n} \right)^d \rho_n \right) \right)^{\frac{p}{2}} \rho_n^{1-\frac{p}{d}}. \end{aligned}$$

Before starting the estimate let us recall that by (4.14)

$$\int_{\Omega_n} \rho_n^{1-\frac{p}{d}} \simeq_\varepsilon \rho(R_n)^{1-\frac{p}{d}} R_n^{d-q}. \quad (6.13)$$

We first estimate in  $B_{R_n/2} \setminus B_{r_1}$ ,

$$\begin{aligned} \int_{B_{R_n/2} \setminus B_{r_1}} \left( \log \left( 1 + n \left( \frac{|x|}{\log n} \right)^d \rho_n \right) \right)^{\frac{p}{2}} \rho_n^{1-\frac{p}{d}} \\ \lesssim (\log n)^{\frac{p}{2}} \int_{B_{R_n/2}} \rho^{1-\frac{p}{d}} \\ \lesssim (\log n)^{\frac{p}{2}} R_n^{d-q} \rho(R_n/2)^{1-\frac{p}{d}} \\ \lesssim \rho(R_n)^{1-\frac{p}{d}} R_n^{d-q} \left( (\log n)^{\frac{p}{2}} \exp(-CR_n^q) \right) \\ \stackrel{(6.13)}{\ll} \int_{\Omega_n} \rho_n^{1-\frac{p}{d}}. \end{aligned}$$

Here we used that by (1.5), we have for some  $C > 0$ ,

$$V(R_n/2) \leq V(R_n) - CR_n^q.$$

In  $\Omega_n \setminus B_{R_n/2}$  we use that since  $n(|x|/\log n)^d \rho_n \gtrsim_\varepsilon 1$  we have for some  $0 < c_\varepsilon \ll 1$ ,

$$\begin{aligned} \left(1 + n \left(\frac{|x|}{\log n}\right)^d \rho_n\right) &\leq 1 + c_\varepsilon \log \left(n \left(\frac{|x|}{\log n}\right)^d \rho_n\right) \\ &\lesssim 1 + c_\varepsilon \log \left(n R_n^{d(1-q)} \rho_n\right). \end{aligned}$$

Since  $n R_n^{d(1-q)} \rho(R_n) \simeq_\varepsilon 1$  we can write

$$n R_n^{d(1-q)} \rho_n(x) \simeq_\varepsilon \frac{\rho_n(x)}{\rho_n(R_n)} = \exp(V(R_n) - V(x)).$$

Thus

$$\begin{aligned} \int_{\Omega_n \setminus B_{R_n/2}} \left( \log \left( 1 + n \left( \frac{|x|}{\log n} \right)^d \rho_n \right) \right)^{\frac{p}{2}} \rho_n^{1-\frac{p}{d}} &\lesssim_\varepsilon \int_{\Omega_n \setminus B_{R_n/2}} (1 + (V(R_n) - V(x)))^{\frac{p}{2}} \rho_n^{1-\frac{p}{d}} \\ &\stackrel{(1.5)}{\lesssim} R_n^{d-1} \int_{R_n/2}^{R_n} (1 + C(R_n^q - r^q))^{\frac{p}{2}} \rho_n^{1-\frac{p}{d}}(r) \\ &\lesssim R_n^{d-q} \rho_n^{1-\frac{p}{d}}(R_n) \stackrel{(6.13)}{\lesssim} \int_{\Omega_n} \rho_n^{1-\frac{p}{d}}. \end{aligned}$$

In conclusion we find

$$\int_{\Omega_n \setminus B_{r_1}} \left( \log \left( 1 + n \left( \frac{|x|}{\log n} \right)^d \rho_n \right) \right)^{\frac{p}{2}} \rho_n^{1-\frac{p}{d}} \lesssim_\varepsilon \int_{\Omega_n} \rho_n^{1-\frac{p}{d}}.$$

Since by (6.13) again,

$$(\log n)^{\frac{p}{2}} \int_{B_{r_1}} \rho_n^{1-\frac{p}{d}} \ll \int_{\Omega_n} \rho_n^{1-\frac{p}{d}},$$

this concludes the proof of (6.7) also in the case  $d = 2$ . □

## 7. The global terms

### 7.1. The main radial estimates

**DEFINITION 7.1.** — *Let  $\gamma \in \mathbb{R}$ . We say that a probability distribution  $\lambda$  on  $(0, R)$  is of type  $\gamma$  if there exist  $\bar{R} \in [0, R]$  such that there exist  $C_\pm$*

positive with

$$\int_0^r \lambda \leq C_- r \lambda(r) \quad \text{for } r \in (0, \bar{R}) \quad (7.1)$$

and

$$\int_r^R \lambda \leq C_+ r^\gamma \lambda(r) \quad \text{for } r \in (\bar{R}, R). \quad (7.2)$$

We write  $A \lesssim_\pm B$  if there is a constant  $C > 0$  depending only on  $p, d, \gamma$  and  $C_\pm$  such that  $A \leq CB$ .

We now prove two very similar results. It should be possible to combine them in a single statement but we decided for clarity to keep them separate.

**PROPOSITION 7.2.** — *Let  $\lambda$  be a distribution of type  $\gamma$  and assume that  $(0, R) = (0, \bar{R}) \cup [\bar{R}, R) = I_0 \cup I_1$ . For  $n \in \mathbb{N}$ , let  $(n_i)_{i \in \{0,1\}}$  be a multinomial distribution of parameters  $(n\lambda(I_i))_{i \in \{0,1\}}$  and set  $\kappa_i = n_i/\lambda(I_i)$ . Assume that*

$$n\lambda(I_i) \geq 1 \quad \text{for } i = 0, 1. \quad (7.3)$$

Then,

$$\begin{aligned} \mathbb{E} \left[ W_p^p \left( \sum_i \kappa_i \lambda \mathbb{1}_{I_i}, n\lambda \right) \right] \\ \lesssim_\pm n^{1-\frac{p}{2}} \left( \bar{R}^p \lambda(I_0)^{1-\frac{p}{2}} + \lambda(I_1)^{-\frac{p}{2}} \int_{I_1} r^{\gamma p} \lambda(r) \right). \end{aligned} \quad (7.4)$$

*Proof.* — Let  $\phi$  be the solution to

$$\phi'' = \sum_i (\kappa_i - n) \lambda \mathbb{1}_{I_i}, \quad \phi'(0) = \phi'(R) = 0$$

we then have by Proposition 2.1,

$$\mathbb{E} \left[ W_p^p \left( \sum_i \kappa_i \lambda \mathbb{1}_{I_i}, n\lambda \right) \right] \lesssim n^{1-p} \int_0^R \frac{1}{\lambda^{p-1}} \mathbb{E}[|\phi'|^p].$$

We claim that

$$n^{1-p} \int_{I_0} \frac{1}{\lambda^{p-1}} \mathbb{E}[|\phi'|^p] \lesssim_\pm n^{1-\frac{p}{2}} \bar{R}^p \lambda(I_0)^{1-\frac{p}{2}}. \quad (7.5)$$

We write for  $r \in I_0$ ,

$$|\phi'(r)| = \left| \frac{n_0 - n\lambda(I_0)}{\lambda(I_0)} \int_0^r \lambda \right| \stackrel{(7.1)}{\lesssim_\pm} \frac{|n_0 - n\lambda(I_0)|}{\lambda(I_0)} r \lambda(r)$$

so that using (7.3) we get

$$\mathbb{E}[|\phi'(r)|^p] \lesssim_\pm n^{\frac{p}{2}} \lambda(I_0)^{-\frac{p}{2}} r^p \lambda(r)^p.$$

After integration this yields

$$n^{1-p} \int_0^{\bar{R}} \frac{1}{\lambda^{p-1}} \mathbb{E}[|\phi'|^p] \lesssim_{\pm} n^{1-\frac{p}{2}} \lambda(I_0)^{-\frac{p}{2}} \int_0^{\bar{R}} r^p \lambda \lesssim_{\pm} n^{1-\frac{p}{2}} \bar{R}^p \lambda(I_0)^{1-\frac{p}{2}}.$$

This proves (7.5). We now claim that

$$n^{1-p} \int_{I_1} \frac{1}{\lambda^{p-1}} \mathbb{E}[|\phi'|^p] \lesssim_{\pm} n^{1-\frac{p}{2}} \lambda(I_1)^{-\frac{p}{2}} \int_{I_1} r^{\gamma p} \lambda(r). \quad (7.6)$$

Let  $r \in I_1 = [\bar{R}, R)$ . We write

$$|\phi'(r)| = \left| \int_r^R (\kappa_1 - n) \lambda \right| = \frac{|n_1 - n \lambda(I_1)|}{\lambda(I_1)} \int_r^R \lambda.$$

Therefore, using (7.3) we get

$$\mathbb{E}[|\phi'|^p] \lesssim_{\pm} n^{\frac{p}{2}} \lambda(I_1)^{-\frac{p}{2}} \left( \int_r^R \lambda \right)^p \stackrel{(7.2)}{\lesssim_{\pm}} n^{\frac{p}{2}} \lambda(I_1)^{-\frac{p}{2}} r^{\gamma p} \lambda(r)^p.$$

After integration this proves (7.6). Combining (7.5) and (7.6) together proves (7.4).  $\square$

We now prove a second version of this estimate when  $(\bar{R}, R)$  is decomposed into sub-intervals.

**PROPOSITION 7.3.** — *Let  $\lambda$  be a distribution of type  $\gamma$  with  $\lambda = 0$  on  $(0, \bar{R})$ . Assume that  $(\bar{R}, R) = \bigcup I_i$  where  $I_i$  are disjoint intervals indexed in increasing order. For  $n \in \mathbb{N}$ , let  $(n_i)_i$  be a multinomial distribution of parameters  $(n \lambda(I_i))_i$  and set  $\kappa_i = n_i / \lambda(I_i)$ . Assume that for some  $\varepsilon > 0$ ,*

$$n \lambda(I_i) \geq \varepsilon \quad \text{for every } i \geq 1. \quad (7.7)$$

*Assume moreover that  $r \mapsto r^{\gamma} \lambda(r)$  is decreasing in each  $I_i$ . Then,*

$$\mathbb{E} \left[ W_p^p \left( \sum_i \kappa_i \lambda \mathbb{1}_{I_i}, n \lambda \right) \right] \lesssim_{\pm, \varepsilon} n^{1-\frac{p}{2}} \int_{\bar{R}}^R r^{\gamma \frac{p}{2}} \lambda(r)^{1-\frac{p}{2}} dr. \quad (7.8)$$

*Proof.* — As above we consider  $\phi$  the solution to

$$\phi'' = \sum_i (\kappa_i - n) \lambda \mathbb{1}_{I_i}, \quad \phi'(\bar{R}) = \phi'(R) = 0$$

so that as above,

$$\mathbb{E} \left[ W_p^p \left( \sum_i \kappa_i \lambda \mathbb{1}_{I_i}, n \lambda \right) \right] \lesssim n^{1-p} \int_{\bar{R}}^R \frac{1}{\lambda^{p-1}} \mathbb{E}[|\phi'|^p].$$

Let  $r \in I_i = (r_{i-1}, r_i)$  for some  $i \geq 1$ . We write

$$\phi'(r) = - \int_r^{r_i} (\kappa_i - n) \lambda - \int_{r_i}^R \sum_{j>i} (\kappa_j - n) \lambda. \quad (7.9)$$

Regarding the first term we have

$$\left| \int_r^{r_i} (\kappa_i - n) \lambda \right| = \frac{|n_i - n\lambda(I_i)|}{\lambda(I_i)} \int_r^{r_i} \lambda.$$

For the second term we observe that

$$\begin{aligned} \int_{r_i}^R \sum_{j>i} (\kappa_j - n) \lambda &= \sum_{j>i} \lambda(I_j) (\kappa_j - n) = \left( \sum_{j>i} n_j \right) - n \left( \sum_{j>i} \lambda(I_j) \right) \\ &= n_i^+ - n\lambda(I_i^+) \end{aligned}$$

where  $n_i^+ = \sum_{j>i} n_j$  and  $I_i^+ = (r_i, R) = \bigcup_{j>i} I_j$  so that  $\sum_{j>i} \lambda(I_j) = \lambda(I_i^+)$ . Observe that  $n_i^+$  is a binomial random variable of parameters  $(n, \lambda(I_i^+))$ . Plugging this into (7.9) we find thanks to (7.7),

$$\begin{aligned} \mathbb{E}[|\phi'|^p] &\lesssim \mathbb{E}[|n_i - n\lambda(I_i)|^p] + \mathbb{E}[|n_i^+ - n\lambda(I_i^+)|^p] \\ &\lesssim_\varepsilon n^{\frac{p}{2}} \lambda(I_i)^{-\frac{p}{2}} \left( \int_r^{r_i} \lambda \right)^p + n^{\frac{p}{2}} \lambda(I_i^+)^{\frac{p}{2}}. \end{aligned}$$

Using that

$$\lambda(I_i)^{-\frac{p}{2}} \left( \int_r^{r_i} \lambda \right)^p \leq \left( \int_r^{r_i} \lambda \right)^{\frac{p}{2}} \leq \lambda(I_i^+)^{\frac{p}{2}}$$

this simplifies to

$$\mathbb{E}[|\phi'|^p] \lesssim_\varepsilon n^{\frac{p}{2}} \lambda(I_i^+)^{\frac{p}{2}} \stackrel{(7.2)}{\lesssim_\pm} n^{\frac{p}{2}} r_i^{\gamma \frac{p}{2}} \lambda(r_i)^{\frac{p}{2}} \leq n^{\frac{p}{2}} r^{\gamma \frac{p}{2}} \lambda(r)^{\frac{p}{2}}$$

where in the last inequality we used that  $r \mapsto r^\gamma \lambda(r)$  is decreasing in  $I_i$ . After integration this gives

$$n^{1-p} \int_{\bar{R}}^R \frac{1}{\lambda^{p-1}} \mathbb{E}[|\phi'|^p] \lesssim_{\pm, \varepsilon} n^{1-\frac{p}{2}} \int_{\bar{R}}^R r^{\gamma \frac{p}{2}} \lambda(r)^{1-\frac{p}{2}}$$

concluding the proof of (7.8).  $\square$

## 7.2. The main angular estimate

Recalling the definition (4.8) of  $\Lambda_{r,R}^{z,\delta}$  we set  $\Lambda^{z,\delta} := \Lambda_{0,\infty}^{z,\delta}$ .

**PROPOSITION 7.4.** — *Let  $\lambda$  be a radially symmetric probability density on  $\mathbb{R}^d$ ,  $n \in \mathbb{N}$  and  $\delta \in (0, \pi]$ . Let*

$$M_p := \int_{\mathbb{R}^d} |x|^p d\lambda \quad \text{and} \quad \tilde{\sigma}_\delta := \int_{\tilde{\mathbb{S}}^{d-1}} \mathbb{1}_{D_\delta(z)}(e_1) d\tilde{\sigma}(z).$$

For  $(X_i)_{i=1}^n$  i.i.d. random variables of law  $\lambda$ , let  $\mu_n := \sum_{i=1}^n \delta_{X_i}$  and for  $z \in \mathbb{S}^{d-1}$

$$\kappa_z := \frac{\mu_n(\Lambda^{z,\delta})}{\lambda(\Lambda^{z,\delta})}.$$

Then, it holds (recall the definition (1.1) of  $\eta_n$ )

$$\mathbb{E} \left[ W_p^p \left( \frac{1}{\tilde{\sigma}_\delta} \int_{\mathbb{S}^{d-1}} \kappa_z \mathbb{1}_{\Lambda^{z,\delta}} \lambda d\tilde{\sigma}(z), n\lambda \right) \right] \lesssim M_p \eta_n^{p,d-1}. \quad (7.10)$$

*Proof.* — We use polar coordinates to write for  $r > 0$  and  $z = (y, U) \in \mathbb{S}^{d-1}$ ,

$$\lambda(ry) = \lambda_{\text{rad}}(r) dr \otimes d\sigma(y).$$

Noticing that for  $y' \in \mathbb{S}^{d-1}$  it holds  $\mathbb{1}_{\Lambda^{z,\delta}}(ry') = \mathbb{1}_{D_\delta(z)}(y')$ , we can write

$$\left( \frac{1}{\tilde{\sigma}_d} \int_{\mathbb{S}^{d-1}} \kappa_z \mathbb{1}_{\Lambda^{z,\delta}} \lambda d\tilde{\sigma}(z) \right)(ry') = \lambda_{\text{rad}}(r) dr \otimes \nu(y') d\sigma(y')$$

where

$$\nu(y') := \frac{1}{\tilde{\sigma}_\delta} \int_{\mathbb{S}^{d-1}} \kappa_z \mathbb{1}_{D_\delta(z)}(y') d\tilde{\sigma}(z).$$

If  $T_{\text{ang}}$  is the optimal transport map between  $n\sigma$  and  $\nu\sigma$  (for the Euclidean cost in  $\mathbb{R}^d$ ), we may use  $T(ry) := rT_{\text{ang}}(y)$  as a competitor to estimate

$$\begin{aligned} W_p^p \left( \frac{1}{\tilde{\sigma}_\delta} \int_{\mathbb{S}^{d-1}} \kappa_z \mathbb{1}_{\Lambda^{z,\delta}} \lambda d\tilde{\sigma}(z), n\lambda \right) &\leq \int_0^\infty \int_{\mathbb{S}^{d-1}} |T(ry) - ry|^p \lambda_{\text{rad}}(r) dr n d\sigma(y) \\ &= \int_0^\infty r^p \lambda_{\text{rad}}(r) dr \int_{\mathbb{S}^{d-1}} |T_{\text{ang}}(y) - y|^p n d\sigma(y) \\ &= M_p W_p^p(\nu\sigma, n\sigma) \leq M_p W_{\mathbb{S}^{d-1}}^p(\nu\sigma, n\sigma) \end{aligned} \quad (7.11)$$

where with a slight abuse of notation we write  $W_{\mathbb{S}^{d-1}}^p$  for the Wasserstein distance on the sphere with respect to the geodesic distance  $d_{\mathbb{S}^{d-1}}$ . Let now  $\hat{X}_i := X_i/|X_i|$  be the projections of  $X_i$  on the sphere. Since  $\lambda$  is radially symmetric,  $\hat{X}_i$  are i.i.d. uniformly distributed on  $\mathbb{S}^{d-1}$  i.e. their common law is  $\sigma$ . If  $\hat{\mu}_n := \sum_{i=1}^n \delta_{\hat{X}_i}$ , using that  $\lambda(\Lambda^{z,\delta}) = \tilde{\sigma}_j$ , we have

$$\kappa_z = \frac{\hat{\mu}_n(D_\delta(z))}{\sigma(D_\delta)}. \quad (7.12)$$

We define the Markov kernel  $k_\delta$  on  $\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$  as

$$k_\delta(x, y) := \int_{\mathbb{S}^{d-1}} \frac{\mathbb{1}_{D_\delta(z)}(x) \mathbb{1}_{D_\delta(z)}(y)}{\tilde{\sigma}_\delta \sigma(D_\delta)} d\tilde{\sigma}(z). \quad (7.13)$$

Given a measure  $\mu$  on  $\mathbb{S}^{d-1}$ , we then define  $K_\delta(\mu)$  as the measure on  $\mathbb{S}^{d-1}$  whose density with respect to  $\sigma$  is  $\int_{\mathbb{S}^{d-1}} k_\delta(x, \cdot) d\mu(x)$ , i.e.

$$K_\delta(\mu)(E) := \int_{\mathbb{S}^{d-1}} \mathbb{1}_E(x) \int_{\mathbb{S}^{d-1}} k_\delta(x, y) d\mu(y) d\sigma(x). \quad (7.14)$$

In particular, by (7.12) we have

$$\nu\sigma = K_\delta(\widehat{\mu}_n) \quad \text{and} \quad n\sigma = K_\delta(n\sigma). \quad (7.15)$$

For  $x, y \in \mathbb{S}^{d-1}$ , let  $\eta_y(x) := k_\delta(x, y)$ . We claim (and will prove below) that for every  $y, y' \in \mathbb{S}^{d-1}$  we have (identifying  $\eta_y\sigma$  with  $\eta_y$ )

$$W_{\mathbb{S}^{d-1}}^p(\eta_y, \eta_{y'}) \leq d_{\mathbb{S}^{d-1}}^p(y, y'). \quad (7.16)$$

Postponing the proof of (7.16) we now estimate  $W_{\mathbb{S}^{d-1}}^p(K_\delta(\widehat{\mu}_n), K_\delta(n\sigma))$ . Let  $\widehat{\pi}$  be an optimal transport plan for  $W_{\mathbb{S}^{d-1}}^p(\widehat{\mu}_n, n\sigma)$  and for  $y, y' \in \mathbb{S}^{d-1}$ , let  $\pi_{y, y'}$  be an optimal transport plan for  $W_{\mathbb{S}^{d-1}}^p(\eta_y, \eta_{y'})$  we define the transport plan  $\pi$  by its action on  $\zeta \in C^0(\mathbb{S}^{d-1} \times \mathbb{S}^{d-1})$ :

$$\begin{aligned} & \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} \zeta(x, x') d\pi(x, x') \\ &:= \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} \left[ \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} \zeta(x, x') d\pi_{y, y'}(x, x') \right] d\widehat{\pi}(y, y'). \end{aligned}$$

Let us prove that the first marginal of  $\pi$  is  $K_\delta(\widehat{\mu}_n)$ . Indeed, for every  $\zeta \in C^0(\mathbb{S}^{d-1})$ , we have

$$\begin{aligned} & \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} \zeta(x) d\pi(x, x') \\ &= \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} \left[ \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} \zeta(x) d\pi_{y, y'}(x, x') \right] d\widehat{\pi}(y, y') \\ &= \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} \left[ \int_{\mathbb{S}^{d-1}} \zeta(x) \eta_y(x) d\sigma(x) \right] d\widehat{\pi}(y, y') \\ &= \int_{\mathbb{S}^{d-1}} \left[ \int_{\mathbb{S}^{d-1}} \zeta(x) k_\delta(x, y) d\sigma(x) \right] d\widehat{\mu}_n(y) \\ &= \int_{\mathbb{S}^{d-1}} \zeta(x) dK_\delta(\widehat{\mu}_n)(x). \end{aligned}$$

Similarly the second marginal of  $\pi$  is  $K_\delta(n\sigma)$ . Therefore,

$$\begin{aligned}
 W_{\mathbb{S}^{d-1}}^p(K_\delta(\widehat{\mu}_n), K_\delta(n\sigma)) &\leq \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} \left[ \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} \mathbf{d}_{\mathbb{S}^{d-1}}^p(x, x') d\pi_{y, y'}(x, x') \right] d\widehat{\pi}(y, y') \\
 &= \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} W_{\mathbb{S}^{d-1}}^p(\eta_y, \eta_{y'}) d\widehat{\pi}(y, y') \\
 &\stackrel{(7.16)}{\leq} \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} \mathbf{d}_{\mathbb{S}^{d-1}}^p(y, y') d\widehat{\pi}(y, y') \\
 &= W_{\mathbb{S}^{d-1}}^p(\widehat{\mu}_n, n\sigma).
 \end{aligned}$$

Combining this with (7.11), (7.15) together with the scaling of the optimal matching problem on the sphere, see e.g. [19, Section 4], we conclude that

$$\begin{aligned}
 \mathbb{E} \left[ W_p^p \left( \frac{1}{\widetilde{\sigma}_j} \int_{\widetilde{\mathbb{S}}^{d-1}} \kappa_z \mathbb{1}_{\Lambda^{z, \delta}} \lambda d\widetilde{\sigma}(z), n\lambda \right) \right] &\lesssim M_p \mathbb{E} [W_{\mathbb{S}^{d-1}}^p(\widehat{\mu}_n, n\sigma)] \\
 &\lesssim M_p \eta_n^{p, d-1}.
 \end{aligned}$$

Therefore in order to conclude the proof of (7.10) we are left with the proof of (7.16). Let  $y, y' \in \mathbb{S}^{d-1}$  and let  $\Pi$  be the two-dimensional plane containing the great circle of  $\mathbb{S}^{d-1}$  containing  $y$  and  $y'$ . We then let  $R_\theta$  be the rotation of angle  $\theta \in (0, \pi]$  such that  $R_\theta = \text{Id}$  on  $\Pi^\perp$  and  $R_\theta(y) = y'$  so that  $\mathbf{d}_{\mathbb{S}^{d-1}}(y, y') = \theta$ . Since  $R_\theta$  is an isometry of  $\mathbb{S}^{d-1}$ , it induces an orthogonal transformation  $U_\theta$  between the tangents at  $y$  and  $y'$ , i.e., we can write

$$R_\theta(D_\delta(y, U)) = D_\delta(y', U_\theta U)$$

Since in the definition of  $\eta_y(x) = k_\delta(x, y)$  we integrate (also) with respect to the Haar measure on  $O(d-1)$ , which is left invariant, we have that  $(R_\theta)_\# \eta_y = \eta_{y'}$ . Finally, for every  $x \in \mathbb{S}^{d-1}$  if  $r \leq 1$  is the radius of the circle given by the intersection between  $\mathbb{S}^{d-1}$  and the plane parallel to  $\Pi$  containing  $x$  we have

$$\mathbf{d}_{\mathbb{S}^{d-1}}(x, R_\theta(x)) \leq r\theta \leq \mathbf{d}_{\mathbb{S}^{d-1}}(y, y')$$

so that

$$W_{\mathbb{S}^{d-1}}^p(\eta_y, \eta_{y'}) \leq \int_{\mathbb{S}^{d-1}} \mathbf{d}_{\mathbb{S}^{d-1}}^p(x, R_\theta(x)) \eta_y(x) d\sigma(x) \leq \mathbf{d}_{\mathbb{S}^{d-1}}^p(y, y').$$

This concludes the proof of (7.16).  $\square$

### 7.3. Estimate of the global terms

We may now estimate the global cost. Let us recall that if  $\widehat{\mu}_n = \sum_{i=1}^n \delta_{\widehat{X}_i}$  with  $\widehat{X}_i$  i.i.d. random variables with law  $\rho_n$ , we have set in (4.13) (recall the

definition (4.10) of  $\Delta_j^z$ ,

$$\kappa_1 := \frac{\widehat{\mu}_n(B_{r_1})}{\rho_n(B_{r_1})} \quad \text{and} \quad \kappa_j^z := \frac{\widehat{\mu}_n(\Delta_j^z)}{\rho_n(\Delta_j^z)}.$$

We recall also the definition (1.4) of  $\tau_n$ .

PROPOSITION 7.5. — *We have*

$$\mathbb{E} \left[ W_p^p \left( \kappa_1 \mathbb{1}_{B_{r_1}} \rho_n + \sum_{j=2}^{j_n} \frac{1}{\widetilde{\sigma}_j} \int_{\widetilde{\mathbb{S}}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\widetilde{\sigma}(z), n\rho_n \right) \right] = o_\varepsilon(\tau_n) + \mathbb{1}_{p>d} O_\varepsilon(\tau_n). \quad (7.17)$$

*Proof.*

*Step 1.* — We first separate the ball  $B_{r_1}$  from the annulus  $A := A_{r_1, R_n}$  (recall (4.9)). We set

$$n_A := \widehat{\mu}_n(A) \quad \text{and then} \quad \kappa_A := \frac{n_A}{\rho_n(A)}$$

and estimate thanks to triangle inequality and subadditivity,

$$\begin{aligned} W_p^p \left( \kappa_1 \mathbb{1}_{B_{r_1}} \rho_n + \sum_{j=2}^{j_n} \frac{1}{\widetilde{\sigma}_j} \int_{\widetilde{\mathbb{S}}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\widetilde{\sigma}(z), n\rho_n \right) \\ \lesssim W_p^p \left( \kappa_1 \mathbb{1}_{B_{r_1}} \rho_n + \kappa_A \mathbb{1}_A \rho_n, n\rho_n \right) \\ + W_A^p \left( \sum_{j=2}^{j_n} \frac{1}{\widetilde{\sigma}_j} \int_{\widetilde{\mathbb{S}}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\widetilde{\sigma}(z), \kappa_A \rho_n \right). \end{aligned} \quad (7.18)$$

To estimate the first term we use Proposition 7.2 with  $\lambda(r) = r^{d-1} \rho_n(r)$ , which is of type  $\gamma = 1 - q$ ,  $n_0 = \widehat{\mu}_n(B_{r_1})$  and  $n_1 = n_A$  so that

$$n\lambda(I_i) \gtrsim n \gg 1.$$

By (7.4),

$$\mathbb{E} [W_p^p(\kappa_1 \mathbb{1}_{B_{r_1}} \rho_n + \kappa_A \mathbb{1}_A \rho_n, n\rho_n)] \lesssim_\varepsilon n^{1-\frac{p}{2}} = o_\varepsilon(\tau_n). \quad (7.19)$$

In the next two steps we estimate the second right-hand side term in (7.18). We claim that

$$\mathbb{E} \left[ W_A^p \left( \sum_{j=2}^{j_n} \frac{1}{\widetilde{\sigma}_j} \int_{\widetilde{\mathbb{S}}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\widetilde{\sigma}(z), \kappa_A \rho_n \right) \right] = o_\varepsilon(\tau_n) + \mathbb{1}_{p>d} O_\varepsilon(\tau_n). \quad (7.20)$$

Combined with (7.18) and (7.19), this would conclude the proof of (7.17).

*Step 2: The case  $d \geq 3$  or  $d = 2$  and  $p \geq 2$ .* — Let us recall that by (4.2),  $\delta_j = \delta$  for every  $j \geq 2$  in this case. Set  $\tilde{\sigma} := \tilde{\sigma}_j = \int_{\tilde{\mathbb{S}}^{d-1}} \mathbb{1}_{D_\delta(z)}(e_1) d\tilde{\sigma}(z)$ . Here, we first apply the angular estimate from Proposition 7.4 to pass from  $A$  to the angular sectors  $\Lambda^z := \Lambda_{r_1, R_n}^{z, \delta}$  and then decompose each angular sector using the radial estimate from Proposition 7.3. We thus define

$$n_z := \hat{\mu}_n(\Lambda^z) \quad \text{and then} \quad \kappa_z := \frac{n_z}{\rho_n(\Lambda^z)}. \quad (7.21)$$

Using triangle inequality we estimate

$$\begin{aligned} & W_A^p \left( \sum_{j=2}^{j_n} \frac{1}{\sigma_j} \int_{\tilde{\mathbb{S}}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\tilde{\sigma}(z), \kappa_A \rho_n \right) \\ & \lesssim W_A^p \left( \frac{1}{\tilde{\sigma}} \int_{\tilde{\mathbb{S}}^{d-1}} \kappa_z \mathbb{1}_{\Lambda^z} \rho_n d\tilde{\sigma}(z), \kappa_A \rho_n \right) \\ & \quad + W_A^p \left( \sum_{j=2}^{j_n} \frac{1}{\tilde{\sigma}} \int_{\tilde{\mathbb{S}}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\tilde{\sigma}(z), \frac{1}{\tilde{\sigma}} \int_{\tilde{\mathbb{S}}^{d-1}} \kappa_z \mathbb{1}_{\Lambda^z} \rho_n d\tilde{\sigma}(z) \right). \end{aligned} \quad (7.22)$$

*Step 2.1.* — To estimate the first term in (7.22), we apply Proposition 7.4 with  $\lambda = (\rho_n/\rho_n(A)) \mathbb{1}_A$  and  $n = n_A$  (notice that  $\mathbb{E}[n_A] = n\rho_n(A) \simeq n$  and in this case  $M_p \lesssim 1$ ) to get

$$\begin{aligned} & \mathbb{E} \left[ W_A^p \left( \frac{1}{\tilde{\sigma}} \int_{\tilde{\mathbb{S}}^{d-1}} \kappa_z \mathbb{1}_{\Lambda^z} \rho_n d\tilde{\sigma}(z), \kappa_A \rho_n \right) \right] \\ & \lesssim \mathbb{E}[n_A]^{1-\frac{p}{d-1}} \mathbb{1}_{d \geq 4} + \mathbb{E}[n_A]^{1-\frac{p}{2}} \left[ \mathbb{1}_{d=2} + (\log \mathbb{E}[n_A])^{\frac{p}{2}} \mathbb{1}_{d=3} \right] \\ & \lesssim \eta_n^{p, d-1} \stackrel{(1.4)}{=} o(\tau_n). \end{aligned} \quad (7.23)$$

*Step 2.2.* — To estimate the second term in (7.22), we first notice that for every  $z \in \tilde{\mathbb{S}}^{d-1}$ ,

$$\sum_j \kappa_j^z \rho_n(\Delta_j^z) \stackrel{(4.13)}{=} \hat{\mu}_n(\Lambda^z) \stackrel{(7.21)}{=} \kappa_z \rho_n(\Lambda^z)$$

so that we may apply Lemma 2.2 and get

$$\begin{aligned} & W_A^p \left( \sum_{j=2}^{j_n} \frac{1}{\tilde{\sigma}} \int_{\tilde{\mathbb{S}}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\tilde{\sigma}(z), \frac{1}{\tilde{\sigma}} \int_{\tilde{\mathbb{S}}^{d-1}} \kappa_z \mathbb{1}_{\Lambda^z} \rho_n d\tilde{\sigma}(z) \right) \\ & \leq \frac{1}{\tilde{\sigma}} \int_{\tilde{\mathbb{S}}^{d-1}} W_{\Lambda^z}^p \left( \sum_{j=2}^{j_n} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n, \kappa_z \rho_n \right) d\tilde{\sigma}(z). \end{aligned} \quad (7.24)$$

Fix  $z \in \widetilde{\mathbb{S}}^{d-1}$ . Let

$$I := (r_1, R_n), \quad I_j := (r_{j-1}, r_j) \quad (7.25)$$

and then

$$\lambda(r) := \frac{\bar{\sigma}}{\rho_n(\Lambda^z)} r^{d-1} \rho_n(r) \mathbb{1}_I, \quad n_j := \widehat{\mu}_n(\Delta_j^z), \quad \text{and} \quad \kappa_j := \frac{n_j}{\lambda(I_j)}. \quad (7.26)$$

By (4.13) and (7.21), we have

$$\bar{\sigma} \kappa_j^z r^{d-1} \rho_n(r) = \kappa_j \lambda(r) \quad \text{and} \quad \bar{\sigma} \kappa_z r^{d-1} \rho_n(r) = n_z \lambda(r)$$

so that by radial symmetry of  $\rho_n$ ,

$$W_{\Lambda^z}^p \left( \sum_{j=2}^{j_n} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n, \kappa_z \rho_n \right) = W_p^p \left( \sum_{j=2}^{j_n} \kappa_j \mathbb{1}_{I_j} \lambda, n_z \lambda \right).$$

Let us notice that  $n_z$  is a binomial random variable with

$$\mathbb{E}[n_z] = n \rho_n(\Lambda^z) \simeq n \delta^{d-1} \gg 1$$

so that with overwhelming probability,  $n_z \geq n \rho_n(\Lambda^z)/2$ . In that event, for every  $j \in [2, j_n]$ ,

$$n_z \lambda(I_j) \gtrsim n \rho_n(\Delta_j^z) \stackrel{(6.10)}{\gtrsim_\varepsilon} 1.$$

Therefore, since  $\lambda$  is easily seen to be of type  $\gamma = 1 - q$  with constants depending only on  $\varepsilon$  and since when conditioned on  $n_z$ ,  $(n_j)_{j \in [2, j_n]}$  are multinomial with parameters  $(n_z \lambda(I_j))_{j \in [2, j_n]}$ , we may apply Proposition 7.3 to get by definition (7.26) of  $\lambda$ ,

$$\begin{aligned} & \mathbb{E} \left[ W_{\Lambda^z}^p \left( \sum_{j=2}^{j_n} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n, \kappa_z \rho_n \right) \right] \\ & \lesssim_\varepsilon \mathbb{E}[n_z]^{1-\frac{p}{2}} \int_I r^{\frac{p(1-q)}{2}} \lambda(r)^{1-\frac{p}{2}} dr \\ & \lesssim_\varepsilon (n \delta^{d-1})^{1-\frac{p}{2}} \begin{cases} 1 & \text{if } p < 2 \\ 1 + R_n^{2-q} + \mathbb{1}_{q=2} \log R_n & \text{if } p = 2 \\ (R_n^{d+q-2} \rho(R_n))^{1-\frac{p}{2}} R_n^{2(1-q)} & \text{otherwise.} \end{cases} \end{aligned}$$

Plugging this back into (7.24) we find

$$\begin{aligned} & \mathbb{E} \left[ W_A^p \left( \sum_{j=2}^{j_n} \frac{1}{\bar{\sigma}} \int_{\widetilde{\mathbb{S}}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\tilde{\sigma}(z), \frac{1}{\bar{\sigma}} \int_{\widetilde{\mathbb{S}}^{d-1}} \kappa_z \mathbb{1}_{\Lambda^z} \rho_n d\tilde{\sigma}(z) \right) \right] \\ & \lesssim_\varepsilon n^{1-\frac{p}{2}} \delta^{-(d-1)\frac{p}{2}} \begin{cases} 1 & \text{if } p < 2 \\ 1 + R_n^{2-q} + \mathbb{1}_{q=2} \log R_n & \text{if } p = 2 \\ (R_n^{d+q-2} \rho(R_n))^{1-\frac{p}{2}} R_n^{2(1-q)} & \text{otherwise.} \end{cases} \quad (7.27) \end{aligned}$$

Let us check that for  $d \geq 3$  or  $p \geq d$ , the right-hand side is  $o_\varepsilon(\tau_n) + \mathbb{1}_{p>d}O_\varepsilon(\tau_n)$ . Combining this with (7.22)–(7.23) would prove (7.20) in this case. If  $1 \leq p \leq 2 < d$ , since  $\delta$  is logarithmically small in  $n$  and  $n^{1-\frac{p}{d}} = \tau_n \gg n^{1-\frac{p}{2}}$ , the statement is clear. If  $p = d = 2$ , the right-hand side of (7.27) is of the order of

$$(\log n) \left( 1 + (\log n)^{\frac{2}{q}-1} + \log \log n \right) \ll (\log n)^{1+\frac{2}{q}} = \tau_n.$$

We may thus focus on the case  $p > 2$ . Recalling that  $\delta \simeq \varepsilon R_n^{-q}$ , (4.5) and the definition (1.14) of  $\alpha, \beta$ , there exists  $\gamma \in \mathbb{R}$  such that the right-hand side of (7.27) is of the order of

$$\begin{aligned} R_n^{(d-1)q-p(q-1)} \left( R_n^{-d(q-1)} n \rho(R_n) \right)^{1-\frac{p}{2}} \\ \lesssim_\varepsilon \begin{cases} (\log n)^\gamma n^{-(1-\beta)(\frac{p}{2}-1)} & \text{if } 2 < p < d \\ (\log n)^{\frac{d}{q}-1-\frac{1}{q}(\alpha-d(q-1))(\frac{d}{2}-1)} & \text{if } p = d \geq 3 \\ (\log n)^{d-1-p(1-\frac{1}{q})} & \text{if } p > d. \end{cases} \end{aligned}$$

Using that when  $2 < p < d$ ,  $\beta < 1$  and for  $p = d$ ,  $\alpha > d(q-1)$ , and recalling the definition (1.4) of  $\tau_n$ , the claim follows.

*Step 3.* — We finally consider the case  $d = 2$  and  $p < 2$ . In this case we will first split the annulus  $A$  into the annuli  $A_j := A_{r_{j-1}, r_j}$  and then split each annulus into the corresponding  $\Delta_j^z$ . We introduce the notation

$$n_j := \widehat{\mu}_n(A_j) \quad \text{and then} \quad \kappa_{A_j} := \frac{n_j}{\rho_n(A_j)}. \quad (7.28)$$

By triangle inequality and subadditivity,

$$\begin{aligned} W_A^p \left( \sum_{j=2}^{j_n} \frac{1}{\widetilde{\sigma}_j} \int_{\mathbb{S}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\widetilde{\sigma}(z), \kappa_A \rho_n \right) \\ \lesssim W_A^p \left( \sum_{j=2}^{j_n} \kappa_{A_j} \mathbb{1}_{A_j} \rho_n, \kappa_A \rho_n \right) \\ + \sum_{j=2}^{j_n} W_{A_j}^p \left( \frac{1}{\widetilde{\sigma}_j} \int_{\mathbb{S}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\widetilde{\sigma}(z), \kappa_{A_j} \rho_n \right). \end{aligned} \quad (7.29)$$

*Step 3.1.* — In this step we estimate the first right-hand side term of (7.29). We recall (7.25) and set

$$\lambda(r) := \frac{1}{\rho_n(A)} r^{d-1} \rho_n(r) \mathbb{1}_I \quad \text{and} \quad \kappa_j := \frac{n_j}{\lambda(I_j)} = \kappa_{A_j} \rho_n(A).$$

We then have by radial symmetry of  $\rho_n$ ,

$$W_A^p \left( \sum_{j=2}^{j_n} \kappa_{A_j} \mathbb{1}_{A_j} \rho_n, \kappa_A \rho_n \right) = W_p^p \left( \sum_{j=2}^{j_n} \kappa_j \mathbb{1}_{I_j} \lambda, n_A \lambda \right).$$

As above we notice that with overwhelming probability,

$$n_A \geq \mathbb{E}[n_A]/2 \gtrsim n$$

and in this event, for every  $j \geq 2$  (recall that  $d = 2$  and that  $r \mapsto r^{2-q}\rho(r)$  is decreasing in  $[r_1, +\infty)$ ),

$$\begin{aligned} n_A \lambda(I_j) &\gtrsim n \rho_n(A_j) \gtrsim_\varepsilon n r_j (r_j - r_{j-1}) \rho(r_j) \\ &\stackrel{(4.1) \& (4.2)}{\gtrsim_\varepsilon} n r_j^{2-q} \rho(r_j) \geq n R_n^{2-q} \rho(R_n) \stackrel{(4.5) \& (4.3)}{\gtrsim} (\log n)^{\frac{2}{q}-1} n^{1-\beta} \gg 1. \end{aligned}$$

Since  $\lambda$  is of type  $\gamma = 1 - q$  with constants depending only on  $\varepsilon$ , we find from Proposition 7.3,

$$\begin{aligned} &\mathbb{E} \left[ W_A^p \left( \sum_{j=2}^{j_n} \kappa_{A_j} \mathbb{1}_{A_j} \rho_n, \kappa_A \rho_n \right) \right] \\ &\lesssim_\varepsilon \mathbb{E}[n_A]^{1-\frac{p}{2}} \int_I r^{p(\frac{1-q}{2})} \lambda(r)^{1-\frac{p}{2}} dr \lesssim n^{1-\frac{p}{2}} = o_\varepsilon(\tau_n). \end{aligned} \quad (7.30)$$

*Step 3.2.* — In this final step we estimate the second right-hand side term in (7.29). For fixed  $j$  we apply Proposition 7.4 with  $\delta = \pi$ ,  $\lambda = \rho_n/\rho_n(A_j) \mathbb{1}_{A_j}$  and  $n = n_j$  to get (recall that  $d = 2$ )

$$\begin{aligned} \mathbb{E} \left[ W_{A_j}^p \left( \frac{1}{\tilde{\sigma}_j} \int_{\tilde{\mathbb{S}}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\tilde{\sigma}(z), \kappa_{A_j} \rho_n \right) \right] &\lesssim \mathbb{E}[n_j]^{1-\frac{p}{2}} \int_{A_j} |x|^p d\lambda \\ &\lesssim r_j^p (n \rho_n(A_j))^{1-\frac{p}{2}}. \end{aligned}$$

Writing that by (4.1) and (4.2),

$$r_j^p (n \rho_n(A_j))^{1-\frac{p}{2}} \simeq_\varepsilon r_j^{1+\frac{pq}{2}} (n \rho(r_j))^{1-\frac{p}{2}} (r_j - r_{j-1})$$

we have after summation,

$$\begin{aligned} &\mathbb{E} \left[ \sum_{j=2}^{j_n} W_{A_j}^p \left( \frac{1}{\tilde{\sigma}_j} \int_{\tilde{\mathbb{S}}^{d-1}} \kappa_j^z \mathbb{1}_{\Delta_j^z} \rho_n d\tilde{\sigma}(z), \kappa_{A_j} \rho_n \right) \right] \\ &\lesssim_\varepsilon n^{1-\frac{p}{2}} \sum_{j=2}^{j_n} r_j^{1+\frac{pq}{2}} \rho(r_j)^{1-\frac{p}{2}} (r_j - r_{j-1}) \\ &\lesssim_\varepsilon n^{1-\frac{p}{2}} \int_{r_1}^{R_n} r^{1+\frac{pq}{2}} \rho(r)^{1-\frac{p}{2}} dr \lesssim n^{1-\frac{p}{2}} = o_\varepsilon(\tau_n). \end{aligned} \quad (7.31)$$

Combining (7.29), (7.30) and (7.31) together concludes the proof of (7.20) also in this case.  $\square$

*Remark 7.6.* — Let us point out that if we tried to use the argument from Step 2 in the case  $1 \leq p < 2$  and  $d = 2$  we would have an issue with the error coming from the radial part of the transport. Indeed, the error term coming from (7.27) would be much larger than  $\tau_n = n^{1-p/2}(\log n)^{p/2}$ . Moreover, if we kept  $\delta_j = \delta$  we would have also an issue with the term (7.31). Vice-versa, if we used the argument from Step 3 in the case  $p > d$  we would have an issue with the error coming from the angular part of the transport when  $d = 2$  or  $d = 3$  (see Step 3.2 and (7.10)). In the remaining cases, both arguments give similar conclusions.

## 8. Concentration inequalities

In this section we briefly prove some concentration inequalities. For  $p \leq d$  and  $\varepsilon > 0$ , let

$$\xi_n := \begin{cases} n^{\frac{1}{p}-\frac{1}{2}} \tau_n^{-\frac{1}{p}} & \text{if } p \leq 2 \\ \tau_n^{\frac{\varepsilon-1}{2}} \mathbb{1}_{p < d} + \tau_n^{-\frac{1}{d}+\varepsilon} \mathbb{1}_{p=d} & \text{otherwise.} \end{cases} \quad (8.1)$$

Notice that  $\xi_n \rightarrow 0$  as  $n \rightarrow \infty$ .

**THEOREM 8.1.** — *Let  $\{X_i\}_{i=1}^\infty$  be iid random variables with common distribution  $\rho$  satisfying a Poincaré inequality and  $\mu_n := \sum_{i=1}^n \delta_{X_i}$ . Then for every  $1 \leq p \leq d$ ,  $r \geq 2$  and  $t > 0$ , we have*

$$\mathbb{P}(\tau_n^{-1} |W_p^p(\mu_n, n\rho) - \mathbb{E}[W_p^p(\mu_n, n\rho)]| > t)^{\frac{1}{r}} \lesssim_{r,\varepsilon} t^{-1} \xi_n.$$

In particular for any  $\delta > 0$

$$\sum_{n=1}^{\infty} \mathbb{P}(|W_p^p(\mu_n, n\rho) - \mathbb{E}[W_p^p(\mu_n, n\rho)]| \geq \delta) < \infty,$$

that is, we have complete (hence  $\mathbb{P}$  a.s.) convergence for  $p < d$ .

*Proof.* — By Chebyshev inequality it is enough to prove that for every  $r \geq 2$ ,

$$\mathbb{E}[|W_p^p(\mu_n, n\rho) - \mathbb{E}[W_p^p(\mu_n, n\rho)]|^r]^{\frac{1}{r}} \lesssim_{\varepsilon,r} \tau_n \xi_n. \quad (8.2)$$

We start by recording the following a result from [16, Proposition 5.3]: if  $\pi$  is the optimal transport plan between  $\mu_n$  and  $n\rho$  and  $r \geq 2$ , we have

$$\begin{aligned} \mathbb{E}[|W_p^p(\mu_n, n\rho) - \mathbb{E}[W_p^p(\mu_n, n\rho)]|^r]^{\frac{1}{r}} \\ \lesssim \mathbb{E}\left[\left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^{2(p-1)} d\pi(x, y)\right)^{\frac{r}{2}}\right]^{\frac{1}{r}}. \end{aligned} \quad (8.3)$$

*Step 1: The case  $p \leq 2$ .* In this case, since  $2(p-1) \leq p$ , we may use Hölder inequality to upgrade (8.3) into

$$\begin{aligned} \mathbb{E}[|W_p^p(\mu_n, n\rho) - \mathbb{E}[W_p^p(\mu_n, n\rho)]|^r]^{\frac{1}{r}} \\ \lesssim n^{\frac{1}{p} - \frac{1}{2}} \mathbb{E}\left[(W_p^p(\mu_n, n\rho))^{\frac{r(p-1)}{p}}\right]^{\frac{1}{r}}. \end{aligned} \quad (8.4)$$

We now prove by induction that for every  $r \in \mathbb{N}$  with  $r \geq 2$ , we have

$$\mathbb{E}[|W_p^p(\mu_n, n\rho) - \mathbb{E}[W_p^p(\mu_n, n\rho)]|^r]^{\frac{1}{r}} \lesssim_r n^{\frac{1}{p} - \frac{1}{2}} \tau_n^{1 - \frac{1}{p}}. \quad (8.5)$$

By definition (8.1) of  $\xi_n$  this would conclude the proof of (8.2) in this case. Before embarking on the proof, let us point out that as a consequence of (8.5),  $\xi_n \rightarrow 0$  and triangle inequality we have

$$\mathbb{E}[(W_p^p(\mu_n, n\rho))^r]^{\frac{1}{r}} \lesssim_r \tau_n. \quad (8.6)$$

We start by proving (8.5) in the case  $r = 2$ . In this case, since  $2(p-1)/p \leq 1$  it is a direct consequence of (8.4), Hölder inequality and  $\mathbb{E}[W_p^p(\mu_n, n\rho)] = \tau_n$  by definition.

We now assume that (8.5) holds up to  $r - 1$ . Let us show that it holds also for  $r$ . Since

$$r \frac{p-1}{p} \leq \frac{r}{2} < r - 1,$$

We have by (8.6) with  $r - 1$  instead of  $r$  and Hölder inequality,

$$\mathbb{E}\left[(W_p^p(\mu_n, n\rho))^{\frac{r(p-1)}{p}}\right]^{\frac{1}{r}} \leq \mathbb{E}\left[(W_p^p(\mu_n, n\rho))^{r-1}\right]^{\frac{p-1}{p(r-1)}} \lesssim_r \tau_n^{1 - \frac{1}{p}}$$

so that by (8.4), we get that (8.5) also holds for  $r$ .

*Step 2: the case  $p > 2$ .* — Notice that in particular we are then only concerned with  $d \geq 3$ . For  $0 < \eta < 1$  we write  $2(p-1) = (1-\eta)p + \eta m$  where

$$m = \frac{2(p-1) - (1-\eta)p}{\eta}. \quad (8.7)$$

Using triangle inequality in the form

$$|a - b|^m \leq (|a| + |b|)^m \leq 2^m \max(|a|^m, |b|^m) \leq 2^m (|a|^m + |b|^m)$$

we have

$$\mathbb{E} \left[ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^m d\pi \right] \lesssim 2^m n \int_{\mathbb{R}^d} |y|^m d\rho \lesssim (2^q m)^{\frac{m}{q}} n.$$

In the last line we used that if  $\rho = \exp(-V)$  with  $V$  satisfying (1.5), then  $(\int |x|^M d\rho)^{\frac{1}{M}} \simeq M^{1/q}$ . Indeed, this follows from  $\rho \lesssim \exp(-C|x|^q)$  and a change of variables to reduce to the case  $q = 1$  where the computation is explicit. Combining this with Hölder inequality we can derive from (8.3),

$$\begin{aligned} \mathbb{E} \left[ |W_p^p(\mu_n, n\rho) - \mathbb{E}[W_p^p(\mu_n, n\rho)]|^r \right]^{\frac{1}{r}} \\ \lesssim (2^q m)^{\frac{m\eta}{2q}} n^{\frac{\eta}{2}} \mathbb{E} \left[ (W_p^p(\mu_n, n\rho))^{\frac{(1-\eta)r}{2}} \right]^{\frac{1}{r}}. \end{aligned} \quad (8.8)$$

We claim that for every  $\varepsilon \ll 1$  and  $r \in \mathbb{N}$  with  $r \geq 2$ ,

$$\mathbb{E} \left[ |W_p^p(\mu_n, n\rho) - \mathbb{E}[W_p^p(\mu_n, n\rho)]|^r \right]^{\frac{1}{r}} \lesssim_{r,\varepsilon} \tau_n^{\frac{1+\varepsilon}{2}} \mathbb{1}_{p < d} + \tau_n^{1-\frac{1}{d}+\varepsilon} \mathbb{1}_{p=d}. \quad (8.9)$$

As in the case  $p \leq 2$  this would conclude the proof of (8.2) by definition (8.1) of  $\xi_n$ . Notice that if (8.5) holds for some  $r$ , then (8.6) also holds for the same  $r$  by triangle inequality. We fix  $1 \gg \varepsilon > 0$ ,  $0 < c < 2\varepsilon d/q$  and choose

$$\eta := \begin{cases} \varepsilon \left( \frac{d}{p} - 1 \right) & \text{if } p < d \\ c \frac{\log \log n}{\log n} & \text{if } p = d. \end{cases}$$

Notice that this choice of  $\eta$  is such that

$$(2^q m)^{\frac{m\eta}{2q}} n^{\frac{\eta}{2}} \tau_n^{\frac{1-\eta}{2}} \lesssim_{\varepsilon} \tau_n^{\frac{1+\varepsilon}{2}} \mathbb{1}_{p < d} + \tau_n^{1-\frac{1}{d}+\varepsilon} \mathbb{1}_{p=d}. \quad (8.10)$$

We prove (8.9) by induction and start with  $r = 2$ . In this case, from (8.8) and Hölder inequality we have

$$\begin{aligned} \mathbb{E} \left[ |W_p^p(\mu_n, n\rho) - \mathbb{E}[W_p^p(\mu_n, n\rho)]|^r \right]^{\frac{1}{r}} &\lesssim (4m)^{\frac{m\eta}{4}} n^{\frac{\eta}{2}} \mathbb{E} [W_p^p(\mu_n, n\rho)]^{\frac{1-\eta}{2}} \\ &= (4m)^{\frac{m\eta}{4}} n^{\frac{\eta}{2}} \tau_n^{\frac{1-\eta}{2}} \\ &\stackrel{(8.10)}{\lesssim_{\varepsilon}} \tau_n^{\frac{1+\varepsilon}{2}} \mathbb{1}_{p < d} + \tau_n^{1-\frac{1}{d}+\varepsilon} \mathbb{1}_{p=d}. \end{aligned}$$

This proves (8.9) when  $r = 2$ .

Assume now that (8.9) holds for  $r - 1$ . We then have

$$\frac{(1-\eta)r}{2} \leq \frac{r}{2} \leq r - 1$$

so that by (8.6) with  $r - 1$  and Hölder we have

$$\mathbb{E} \left[ (W_p^p(\mu_n, n\rho))^{\frac{(1-\eta)r}{2}} \right]^{\frac{1}{r}} \leq \mathbb{E} \left[ (W_p^p(\mu_n, n\rho))^{r-1} \right]^{\frac{1-\eta}{2(r-1)}} \lesssim_r \tau_n^{\frac{1-\eta}{2}}.$$

Plugging this into (8.8) and recalling (8.10) concludes the proof of (8.9) also for  $r$ .  $\square$

*Remark 8.2.* — The proof above fails when  $p > d$ . This may be easily seen from the fact that we have a gap in the rate, that is  $\tau_n = (\log n)^{d-1-\frac{p}{2}}$  (that does not coincide with the correct rate when  $p = d$ ), and this prevents us from finding a suitable  $\eta = \eta(n)$  as we did before.

*Remark 8.3.* — On the one hand, all log-concave distributions satisfy a Poincaré inequality by [7]. On the other hand, if  $\rho$  satisfies a Poincaré inequality then it must have exponentially decaying tails by [8]. Therefore, Theorem 8.1 mostly applies for potentials satisfying (1.5) with  $q \geq 1$ .

## Appendix A. Bounds for the matching problem in bounded domains

We start by recording that arguing verbatim as in [21, Proposition B.1], we have the following de-Poissonization result.

PROPOSITION A.1. — *Let  $\Omega$  be a bounded and connected set,  $\lambda$  be a probability measure on  $\Omega$ ,  $\{X_i\}_{i=1}^\infty$  be i.i.d. random variables of law  $\lambda$  and for every  $n$ , let  $N_n$  be Poisson random variables of parameter  $n$ . Letting  $\mu_n = \sum_{i=1}^n \delta_{X_i}$  and then for  $W_*^p \in \{W_\Omega^p, Wb_\Omega^p\}$ ,*

$$f(n) := \mathbb{E}[W_*^p(\mu_n, n\lambda)],$$

*we have for every  $\alpha, \beta \in \mathbb{R}$ ,*

$$\liminf_{n \rightarrow \infty} n^\alpha (\log n)^\beta f(n) \geq \liminf_{n \rightarrow \infty} n^\alpha (\log n)^\beta \mathbb{E}[f(N_n)]$$

*and*

$$\limsup_{n \rightarrow \infty} n^\alpha (\log n)^\beta f(n) \leq \limsup_{n \rightarrow \infty} n^\alpha (\log n)^\beta \mathbb{E}[f(N_n)].$$

We start by discussing upper bounds. Let us recall that if  $X_i$  are i.i.d. random variables uniformly distributed in  $Q_1 = (0, 1)^d$  and if  $\mu_n = \sum_{i=1}^n \delta_{X_i}$  we have set

$$\bar{c}_{p,d} = \limsup_{n \rightarrow \infty} \frac{1}{\eta_n} \mathbb{E}[W_{Q_1}^p(\mu_n, n)]$$

where

$$\eta_n = n^{1-\frac{p}{d}} \left(1 + \mathbb{1}_{d=2} (\log n)^{\frac{p}{2}}\right).$$

It is now well-established, see e.g. [19, Section 4] that  $\bar{c}_{p,d} \in (0, \infty)$  for every  $d \geq 2$  and every  $p \geq 1$ . Notice that actually by [3, 5, 12, 16], the limsup is a limit when  $p = d = 2$  (with  $\bar{c}_{p,d} = 1/(4\pi)$ ) or when  $d \geq 3$ . Combining the

proof of [21] where the case of uniform densities is treated together with [2, Proposition 2.4] we have (see also [5, 12, 17]).

**THEOREM A.2.** — *Let  $\Omega \subset \mathbb{R}^d$  be a bounded and connected open set either  $C^2$  or convex and let  $\lambda$  be a Hölder continuous probability density on  $\bar{\Omega}$  bounded away from 0. Let  $X_i$  be i.i.d. random variables with common law  $\lambda$ . then*

$$\limsup_{n \rightarrow \infty} \frac{1}{\eta_n} \mathbb{E}[W_\Omega^p(\mu_n, n\lambda)] \leq \bar{\mathbf{c}}_{p,d} \int_\Omega \lambda^{1-\frac{p}{d}}. \quad (\text{A.1})$$

We now turn to the lower bounds and recall that

$$\underline{\mathbf{c}}^{p,d} = \liminf_{n \rightarrow \infty} \frac{1}{\eta_n} \mathbb{E}[W_{Q_1}^p(\mu_n, n)].$$

We first prove that  $\underline{\mathbf{c}}^{p,d} \in (0, \infty)$ .

**PROPOSITION A.3.** — *For every  $d \geq 2$  and every  $p \geq 1$ , we have  $\underline{\mathbf{c}}^{p,d} \in (0, \infty)$ . Moreover, the liminf in the definition of  $\underline{\mathbf{c}}^{p,d}$  is a limit at least if  $p = d = 2$  or if  $d \geq 3$ .*

*Proof.* — Notice first that  $\underline{\mathbf{c}}^{p,d} \leq \bar{\mathbf{c}}_{p,d} < \infty$ . The fact that the limit exists and is equal to  $1/(4\pi)$  in the case  $p = d = 2$  was proven in [2, Proposition 3.1] so we focus on the remaining cases. By scaling, Proposition A.1 and Hölder inequality, it is enough to prove the following Poisson version of the statement: let  $\mu$  be a Poisson point process of intensity one on  $\mathbb{R}^d$  and set

$$f_{p,d}(L) := \mathbb{E} \left[ \frac{1}{|Q_L|} W_{Q_L}^p(\mu, \kappa) \right]$$

where  $\kappa := \mu(Q_L)/|Q_L|$  then

$$f_{1,d}(L) \gtrsim (\log L)^{\frac{1}{2}} \quad (\text{A.2})$$

and for  $d \geq 3, p \geq 1$ , (the existence of the limit is part of statement)

$$\lim_{L \rightarrow \infty} f_{p,d}(L) = \underline{\mathbf{c}}^{p,d} > 0. \quad (\text{A.3})$$

For (A.2) we use [18, Lemma 2.7] to infer that

$$\begin{aligned} & f_{1,d}(L) \\ &= \mathbb{E} \left[ \frac{1}{L^2} \sup \left\{ \int_{Q_L} \zeta d(\mu - \kappa) : |\nabla \zeta| \leq 1, \zeta = 0 \text{ on } \partial Q_L \right\} \right] \\ &\geq \mathbb{E} \left[ \frac{1}{L^2} \sup \left\{ \int_{Q_L} \zeta d(\mu - \kappa) : |\nabla \zeta| \leq 1, \zeta = 0 \text{ on } \partial Q_L \text{ \& } \int_{Q_L} \zeta = 0 \right\} \right] \\ &= \mathbb{E} \left[ \frac{1}{L^2} \sup \left\{ \int_{Q_L} \zeta d\mu : |\nabla \zeta| \leq 1, \zeta = 0 \text{ on } \partial Q_L \text{ \& } \int_{Q_L} \zeta = 0 \right\} \right] \\ &\gtrsim (\log L)^{\frac{1}{2}}. \end{aligned}$$

We now turn to (A.3). Since for  $\varepsilon \in (0, 1)$ ,

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{|Q_L|} Wb_{Q_L}^p(\mu, \kappa) \right] \\ \geq (1 - \varepsilon) \mathbb{E} \left[ \frac{1}{|Q_L|} Wb_{Q_L}^p(\mu, 1) \right] - \frac{C}{\varepsilon^{p-1}} \mathbb{E} \left[ \frac{1}{|Q_L|} Wb_{Q_L}^p(1, \kappa) \right] \end{aligned}$$

and by Lemma 2.5,

$$\mathbb{E} \left[ \frac{1}{|Q_L|} Wb_{Q_L}^p(1, \kappa) \right] \lesssim L^p \mathbb{E}[|1 - \kappa|^p] \lesssim L^{p(1 - \frac{d}{2})}, \quad (\text{A.4})$$

in order to prove (A.3) it is enough to prove it for

$$\widehat{f}_{p,d}(L) := \mathbb{E} \left[ \frac{1}{|Q_L|} Wb_{Q_L}^p(\mu, 1) \right].$$

Let  $L > 0$  and  $m \in \mathbb{N}$ . We divide  $Q_{mL}$  into  $m^d$  cubes  $Q_i$  of sidelength  $L$ . By superadditivity, see Lemma 2.4, and stationarity of  $\mu$ ,

$$\begin{aligned} \widehat{f}_{p,d}(mL) &\geq \sum_{i=1}^{m^d} \mathbb{E} \left[ \frac{1}{|Q_{mL}|} Wb_{Q_i}^p(\mu, 1) \right] \\ &= \sum_{i=1}^{m^d} \frac{|Q_L|}{|Q_{mL}|} \mathbb{E} \left[ \frac{1}{|Q_L|} Wb_{Q_L}^p(\mu, 1) \right] = \widehat{f}_{p,d}(L). \end{aligned}$$

Since  $\widehat{f}_{p,d}$  is a continuous function the statement follows.  $\square$

*Remark A.4.* — Notice that with the notation from the proof of Proposition A.3, we have using the fact that  $\widehat{f}_{p,d}(mL)$  is increasing in  $m$ , triangle inequality and (A.4),

$$f(L) \leq \underline{c}^{p,d} + \frac{C}{L^{\frac{d-2}{2}}}.$$

See also [16, Theorem 4.1] for a similar estimate regarding  $\bar{c}_{p,d}$ .

We may now prove the counterpart of Theorem A.2.

**THEOREM A.5.** — *Let  $\Omega \subset \mathbb{R}^d$  be a bounded open set let  $\lambda$  a Hölder continuous probability density on  $\bar{\Omega}$  bounded away from 0. Let  $X_i$  be i.i.d. random variables with common law  $\lambda$ . then*

$$\liminf_{n \rightarrow \infty} \frac{1}{\eta_n} \mathbb{E}[Wb_{\Omega}^p(\mu_n, n\lambda)] \geq \underline{c}^{p,d} \int_{\Omega} \lambda^{1 - \frac{p}{d}}. \quad (\text{A.5})$$

*Proof.* — Let  $1 \gg \varepsilon > 0$  be fixed. And for  $z \in \mathbb{Z}^d$  let  $Q_z := \varepsilon(z + Q_1)$ . Setting

$$Z := \{z \in \mathbb{Z}^d : Q_z \subset \Omega\} \quad \text{and} \quad \Omega_{\varepsilon} := \bigcup_{z \in Z} Q_z,$$

we have by superadditivity, see Lemma 2.4,

$$\mathbb{E}[Wb_{\Omega}^p(\mu_n, n\lambda)] \geq \sum_{z \in Z} \mathbb{E}[Wb_{Q_z}^p(\mu_n, n\lambda)].$$

We claim that for every  $z \in Z$ ,

$$\liminf_{n \rightarrow \infty} \frac{1}{\eta_n} \mathbb{E}[Wb_{Q_z}^p(\mu_n, n\lambda)] \geq (1 - \omega(\varepsilon)) \underline{\mathbf{c}}^{p,d} \int_{Q_z} \lambda^{1-\frac{p}{d}}. \quad (\text{A.6})$$

Summing over  $z$  and letting then  $\varepsilon \rightarrow 0$ , this would conclude the proof of (A.5). We now prove (A.6) and set for simplicity  $Q := Q_z$ . Letting  $N := \mu_n(Q)$  and for  $(\hat{X}_i)_{i=1}^{\infty}$  i.i.d. random variables of law  $\hat{\lambda} := \lambda \llcorner Q/\lambda(Q)$ ,  $\hat{\mu}_N := \sum_{i=1}^N \delta_{\hat{X}_i}$ ,  $\mu_n \llcorner Q$  and  $\hat{\mu}_N$  have the same law thus

$$\mathbb{E}[Wb_Q^p(\mu_n, n\lambda)] = \mathbb{E}[\mathbb{E}[Wb_Q^p(\hat{\mu}_k, n\lambda) \mid N = k]].$$

By [2, Proposition 2.4], there is a transport map  $T : Q \mapsto Q$  with  $T_{\#}\hat{\lambda} = 1/|Q|$  and such that  $\text{Lip } T \leq 1 + C\omega(\varepsilon)$ . Letting  $\mu'_n := \sum_{i=1}^n \delta_{T(\hat{X}_i)}$  we then have for every  $k \in \mathbb{N}$ ,

$$\begin{aligned} & \mathbb{E}[Wb_Q^p(\hat{\mu}_k, n\lambda)] \\ & \geq (1 - \omega(\varepsilon)) \mathbb{E}\left[Wb_Q^p\left(\mu'_k, \frac{n\lambda(Q)}{|Q|}\right)\right] \\ & \geq (1 - \omega(\varepsilon)) \mathbb{E}\left[Wb_Q^p\left(\mu'_k, \frac{k}{|Q|}\right)\right] - \frac{C}{\varepsilon^{p-1}} \mathbb{E}\left[Wb_Q^p\left(\frac{k}{|Q|}, \frac{n\lambda(Q)}{|Q|}\right)\right] \\ & \stackrel{(2.7)}{\geq} (1 - \omega(\varepsilon)) \mathbb{E}\left[Wb_Q^p\left(\mu'_k, \frac{k}{|Q|}\right)\right] - C_{\varepsilon} \frac{|n\lambda(Q) - k|^p}{(n\lambda(Q))^{p-1}}. \end{aligned}$$

Taking the expectation we find

$$\begin{aligned} \mathbb{E}[Wb_Q^p(\mu_n, n\lambda)] &= (1 - \omega(\varepsilon)) \mathbb{E}\left[Wb_Q^p\left(\mu'_N, \frac{N}{|Q|}\right)\right] - C_{\varepsilon} \mathbb{E}\left[\frac{|n\lambda(Q) - N|^p}{(n\lambda(Q))^{p-1}}\right] \\ &\geq (1 - \omega(\varepsilon)) \underline{\mathbf{c}}^{p,d} |Q|^{\frac{p}{d}} \mathbb{E}[\eta_N^{p,d}] - C_{\varepsilon} n^{1-\frac{p}{2}} \end{aligned}$$

where in the last line we used the definition of  $\underline{\mathbf{c}}^{p,d}$  and the fact that  $N$  is a binomial random variable with  $\mathbb{E}[N] = n\lambda(Q)$ . Since  $n^{1-\frac{p}{2}} = o(\eta_n)$  and

$$\lim_{n \rightarrow \infty} \mathbb{E}\left[\frac{\eta_N^{p,d}}{\eta_n}\right] = \lambda(Q)^{1-\frac{p}{d}},$$

we find

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{\eta_n} \mathbb{E} \left[ Wb_{Q_z}^p(\mu_n, n\lambda) \right] &\geq (1 - \omega(\varepsilon)) \underline{c}^{p,d} |Q|^{\frac{p}{d}} \lambda(Q)^{1-\frac{p}{d}} \\ &\geq (1 - \omega(\varepsilon)) \underline{c}^{p,d} \int_Q \lambda^{1-\frac{p}{d}}. \end{aligned}$$

This concludes the proof of (A.6).  $\square$

## Bibliography

- [1] M. AJTAI, J. KOMLÓS & G. TUSNÁDY, “On optimal matchings”, *Combinatorica* **4** (1984), p. 259-264.
- [2] L. AMBROSIO, M. GOLDMAN & D. TREVISAN, “On the quadratic random matching problem in two-dimensional domains”, *Electron. J. Probab.* **27** (2022), article no. 54 (35 pages).
- [3] L. AMBROSIO, F. STRA & D. TREVISAN, “A PDE approach to a 2-dimensional matching problem”, *Probab. Theory Relat. Fields* **173** (2019), no. 1-2, p. 433-477.
- [4] L. AMBROSIO, F. VITILLARO & D. TREVISAN, “Sharp PDE estimates for random two-dimensional bipartite matching with power cost function”, 2024, <https://arxiv.org/abs/2405.09397>.
- [5] F. BARTHE & C. BORDENAVE, “Combinatorial optimization over two random point sets”, in *Séminaire de probabilités XLV*, Springer, 2013, p. 483-535.
- [6] D. BENEDETTO & E. CAGLIOTI, “Euclidean random matching in 2D for non-constant densities”, *J. Stat. Phys.* **181** (2020), no. 3, p. 854-869.
- [7] S. G. BOBKOV, “Isoperimetric and analytic inequalities for log-concave probability measures”, *Ann. Probab.* **27** (1999), no. 4, p. 1903-1921.
- [8] S. G. BOBKOV & M. LEDOUX, “Poincaré’s inequalities and Talagrand’s concentration phenomenon for the exponential distribution”, *Probab. Theory Relat. Fields* **107** (1997), no. 3, p. 383-400.
- [9] ———, *One-dimensional empirical measures, order statistics, and Kantorovich transport distances*, Memoirs of the American Mathematical Society, vol. 1259, American Mathematical Society, 2019.
- [10] E. CAGLIOTI & F. PIERONI, “Random Matching in 2D with Exponent 2 for Gaussian Densities”, *J. Stat. Phys.* **191** (2024), no. 5, article no. 62 (34 pages).
- [11] S. CARACCILO, C. LUCIBELLO, G. PARISI & G. SICURO, “Scaling hypothesis for the Euclidean bipartite matching problem”, *Phys. Rev. E* **90** (2014), no. 1, article no. 012118 (9 pages).
- [12] S. DEREICH, M. SCHEUTZOW & R. SCHOTTSTEDT, “Constructive quantization: approximation by empirical measures”, *Ann. Inst. Henri Poincaré, Probab. Stat.* **49** (2013), no. 4, p. 1183-1203.
- [13] A. FIGALLI & N. GIGLI, “A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions”, *J. Math. Pures Appl. (9)* **94** (2010), no. 2, p. 107-130.
- [14] N. FOURNIER & A. GUILLIN, “On the rate of convergence in Wasserstein distance of the empirical measure”, *Probab. Theory Relat. Fields* **162** (2015), no. 3-4, p. 707-738.
- [15] M. GOLDMAN, M. HUESMANN & D. TREVISAN, “Asymptotics for Random Quadratic Transportation Costs”, 2025, <https://arxiv.org/abs/2409.08612>.
- [16] M. GOLDMAN & D. TREVISAN, “Convergence of asymptotic costs for random Euclidean matching problems”, *Probab. Math. Phys.* **2** (2021), no. 2, p. 341-362.

- [17] ———, “Optimal transport methods for combinatorial optimization over two random point sets”, *Probab. Theory Relat. Fields* **188** (2024), no. 3-4, p. 1315-1384.
- [18] M. HUESMANN, F. MATTESINI & F. OTTO, “There is no stationary cyclically monotone Poisson matching in 2d”, *Probab. Theory Relat. Fields* **187** (2023), no. 3-4, p. 629-656.
- [19] M. LEDOUX, “On optimal matching of Gaussian samples”, *Zap. Nauchn. Semin. (POMI)* **457** (2017), no. 25, p. 226-264.
- [20] M. LEDOUX & J.-X. ZHU, “On optimal matching of Gaussian samples III”, *Probab. Math. Stat.* **41** (2021), no. 2, p. 237-265.
- [21] M. MARIANI & D. TREVISAN, “Wasserstein Asymptotics for Brownian Motion on the Flat Torus and Brownian Interlacements”, 2023, <https://arxiv.org/abs/2307.10325>.
- [22] R. PEYRE, “Comparison between  $W_2$  distance and  $\dot{H}^{-1}$  norm, and localization of Wasserstein distance”, *ESAIM, Control Optim. Calc. Var.* **24** (2018), no. 4, p. 1489-1501.
- [23] M. TALAGRAND, “Scaling and non-standard matching theorems”, *Comptes Rendus. Mathématique* **356** (2018), no. 6, p. 692-695.
- [24] D. TREVISAN & M. GOLDMAN, “On the concave one-dimensional random assignment problem and Young integration theory”, *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (5)* **27** (2026), no. 1, p. 589-612.
- [25] J. E. YUKICH, *Probability theory of classical Euclidean optimization problems*, Lecture Notes in Mathematics, vol. 1675, Springer, 2006, eBook.