Rend. Lincei Mat. Appl. 35 (2024), 323–342
DOI 10.4171/RLM/1042

Mathematical Analysis, Probability Theory. – *Sharp PDE estimates for random two-dimensional bipartite matching with power cost function*, by Luigi Ambrosio, Federico Vitillaro and Dario Trevisan, communicated on 8 November 2024.

Abstract. – We investigate the random bipartite optimal matching problem on a flat torus in two dimensions, considering general strictly convex power costs of the distance. We extend the successful ansatz first introduced by Caracciolo et al. for the quadratic case, involving a linear Poisson equation, to a non-linear equation of q-Poisson type, allowing for a more comprehensive analysis of the optimal transport cost. Our results establish new asymptotic connections between the energy of the solution to the PDE and the optimal transport cost, providing insights on their asymptotic behavior.

Keywords. – optimal transport, q-Laplacian, Hopf–Lax semigroup.

Mathematics Subject Classification 2020. – 60D05 (primary); 49J55, 60H15 (secondary).

1. Introduction and main result

Let $(X_1, \ldots, X_n)$, $(Y_1, \ldots, Y_n)$ be two sets of n random points independent and uniformly distributed on the flat torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, i.e., with common law given by the Lebesgue measure $\mathfrak{m}$ on $\mathbb{T}^2$. The random bipartite optimal matching problem concerns the study of the optimal coupling (with respect to a certain cost function) of these points, that is, the optimal transport from the empirical measure $\mu^n := \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ to $\nu_n := \frac{1}{n}\sum_{i=1}^n \delta_{Y_i}$, in particular in the limit $n \gg 1$.

For the quadratic cost $c(x, y) := \mathbf{d}(x, y)^2$, where $\mathbf{d}$ is the quotient (flat) distance in $\mathbb{T}^2$, the seminal paper [11] gave a very appealing PDE ansatz on the asymptotic of the expectation of the optimal transport cost, based on a linearization of the Monge–Ampère equation. While it was already known in the literature that, for the cost $c = \mathbf{d}^p$ in dimension $d = 2$, the expectation of the optimal transport cost behaves like $(n^{-1}\ln n)^{p/2}$ (since [1]), in [11], they managed to predict the limit coefficient as $1/(2\pi)$ in the case $p = 2$, exploiting Fourier analysis and some renormalization procedure. This prediction was then rigorously proven in [7], together with a new PDE proof of the classical bounds in [1].

Since then, several works have been using such PDE ansatz to estimate with different degrees of sharpness the asymptotics of random optimal matching costs and their solutions, in several settings. Focusing only on the two-dimensional case, but possibly including more general manifolds than $\mathbb{T}^2$, we mention here the rigorous results [4–6, 10, 12, 14–16, 18] as well as further intriguing predictions from the physical literature [8, 9] and refer e.g. to the contribution [25] for a more general overview on the subject.

The aim of the present work is to establish new asymptotic connections between the solution of a "linearized PDE" and the expectation of the optimal transport cost, on $\mathbb{T}^2$, for general $p > 1$, extending the main results in [4, 7]. Let us mention here that recently other works focused on two-dimensional random optimal matching problems, beyond the quadratic cost, in particular [22], where the quantitative harmonic approximation techniques – originally in [15], see also the exposition [23] – are extended to any $p > 1$, and the preprint [19], where the existence of a p-cyclically monotone stationary matching from a Poisson point process to the Lebesgue measure is ruled out for any $p > 1$ – the quadratic case is covered in [18].

In order to describe here informally our results, we may treat the empirical measures $\mu^n = \rho_0 \mathfrak{m}$, $\nu^n = \rho_1 \mathfrak{m}$ as absolutely continuous with respect to $\mathfrak{m}$. This will be made rigorous by a regularization with the heat kernel P_t on $\mathbb{T}^2$, as performed in [7]. We first recall (see e.g., [2, Remark 5.3]) that the Kantorovich potential ϕ is related to the optimal transport map T by the identity

$$T(x) = x + \left|\nabla\phi(x)\right|^{q-2}\nabla\phi(x),$$

where, throughout the paper, $q = p/(p-1)$ denotes the dual exponent of p. Then, the Monge–Ampère equation takes the form

$$\rho_1\big(x + \left|\nabla\phi(x)\right|^{q-2}\nabla\phi(x)\big)\det\big(\nabla\big(x + \left|\nabla\phi(x)\right|^{q-2}\nabla\phi(x)\big)\big) = \rho_0(x).$$

This PDE contains three non-linearities: the determinant, the dependence of ρ_1 on $\nabla\phi$, and finally, when $p \neq 2$, the nonlinear term $|\nabla\phi|^{q-2}\nabla\phi$. Our main result shows that in order to obtain a good first-order approximation of the expected value of the transport cost, it is sufficient to remove only the first two non-linearities, keeping the third one. This invokes the "linearized" (but still nonlinear!) PDE of q-Poisson type

$$-\operatorname{div}\big(|\nabla\phi|^{q-2}\nabla\phi\big) = \rho_1 - \rho_0, \quad \phi \in H^{1,q}(\mathbb{T}^2) \tag{1.1}$$

in the sense of distributions, namely,

$$\int_{\mathbb{T}^2} |\nabla\phi|^{q-2}\langle\nabla\phi, \nabla\eta\rangle\, d\mathfrak{m} = \int_{\mathbb{T}^2}(\rho_1 - \rho_0)\eta\, d\mathfrak{m} \quad \forall \eta \in H^{1,q}(\mathbb{T}^2), \tag{1.2}$$

where we always assume, just to ensure uniqueness, that $\int_{\mathbb{T}^2}\phi\, d\mathfrak{m} = 0$, yielding the approximation

$$\left|T(x) - x\right|^p \approx \left|\left|\nabla\phi(x)\right|^{q-2}\nabla\phi(x)\right|^p = \left|\nabla\phi(x)\right|^q. \tag{1.3}$$

Our main result makes precise such approximation (see Section 2 for more details on the notation).

THEOREM 1.1 (Main result). *If* $(X_i)_{i=1}^{\infty}$ *and* $(Y_i)_{i=1}^{\infty}$ *are independent and identically distributed random variables with law* $\mathfrak{m}$ *on* $\mathbb{T}^2$, *then*

$$\lim_{n\to\infty}\left(\frac{n}{\ln n}\right)^{p/2}\left|\mathbb{E}\left[W_p^p(\mu^n,\nu^n)\right]-\mathbb{E}\left[\int_{\mathbb{T}^2}|\nabla\phi_n|^q\,\mathrm{d}\mathfrak{m}\right]\right|=0$$

where

$$\mu^n=\frac{1}{n}\sum_{i=1}^n\delta_{X_i},\qquad \nu^n=\frac{1}{n}\sum_{i=1}^n\delta_{Y_i},\tag{1.4}$$

and ϕ_n *is the solution to* (1.2) *with random right-hand side* $\rho_1-\rho_0=\rho_{1,n}-\rho_{0,n}$ *and*

$$\rho_{0,n}\mathfrak{m}=P_{t_n}\mu^n,\qquad \rho_{1,n}\mathfrak{m}=P_{t_n}\nu^n\tag{1.5}$$

provided $t_n\gg n^{-1}\ln n$ *and* $\ln(nt_n)\ll\ln n$.

For instance, a good choice of the intermediate regularization scale t_n in the main result would be $t_n=n^{-1}(\ln n)^{\beta}$ with $\beta>1$. Thanks to this result, the existence of the limit

$$\lim_{n\to\infty}\frac{\mathbb{E}\left[W_p^p(\mu^n,\nu^n)\right]}{\left((\ln n)/n\right)^{p/2}}$$

is equivalent to the existence of the limit when, in the numerator, $\mathbb{E}[W_p^p(\mu^n,\nu^n)]$ is replaced by $\mathbb{E}[\int_{\mathbb{T}^2}|\nabla\phi_n|^q\,\mathrm{d}\mathfrak{m}]$, with ϕ_n solutions to the PDE (1.2) with a random right-hand side (1.5). It would be interesting to prove or disprove the existence of the limit thanks to this reduction to a stochastic PDE.

In order to prove Theorem 1.1, the only probabilistic ingredients (see Section 2.4) will consist in checking that as $n\to\infty$, with high probability, the densities $\rho_{i,n}$ in (1.5), for $i=0,1$, are both sufficiently close to the constant density (Proposition 2.6), as well as not too far from μ^n and ν^n in the Wasserstein sense (Proposition 2.5), collecting and slightly extending some results from [4, 7]. Then, in Sections 4.1 and 4.2, we will focus our efforts on showing the following deterministic result.

THEOREM 1.2. *Let* $p>1$, *let* ϕ *be a solution of* (1.2), *and let*

$$c:=2\max_{i=0,1}\|\rho_i-1\|_{L^\infty(\mathbb{T}^2)}.$$

Then, there exist $\underline{\delta}=\underline{\delta}(c,p)$ *and* $\overline{\delta}=\overline{\delta}(c,p)$ *such that* $\underline{\delta}+\overline{\delta}\to0$ *as* $c\to0$ *and*

$$(1-\underline{\delta})\int_{\mathbb{T}^2}|\nabla\phi|^q\,\mathrm{d}\mathfrak{m}\le W_p^p(\rho_0\mathfrak{m},\rho_1\mathfrak{m})\le(1+\overline{\delta})\int_{\mathbb{T}^2}|\nabla\phi|^q\,\mathrm{d}\mathfrak{m}.$$

This result actually holds, with the same proof, on any d-dimensional torus. The extension to the setting of compact Riemannian manifolds (along the lines of [7]) possibly with boundary is beyond the scope of this note and requires in particular the understanding in that more general setting of the stability of the estimates from above for the Riemannian analogous of the operator $\operatorname{div}(|\nabla\phi|^{q-1}\nabla\phi)$ under the action of the Hopf–Lax semigroup, even after shocks.

2. Preliminaries

2.1. The Wasserstein distance

Given probability measures μ, ν on $\mathbb{T}^2$ and $p \geq 1$, we define the p-Wasserstein distance between μ and ν as

$$W_p(\mu,\nu) := \min\left\{\left(\int_{\mathbb{T}^2\times\mathbb{T}^2} \mathbf{d}(x,y)^p \, d\pi(x,y)\right)^{1/p} \,\Big|\, \pi_1 = \mu,\ \pi_2 = \nu\right\}.$$

We refer to [2] for an introduction to the subject. In particular, we will use throughout that W_p enjoys the triangle inequality. Moreover, we recall here for later use the following consequence of the Benamou–Brenier formula; see e.g. [27], [24, Theorem 2] or [17, Lemma 3.4].

Proposition 2.1. *Let $\mu = \rho_0 \mathfrak{m}$, $\nu = \rho_1 \mathfrak{m}$ be absolutely continuous with respect to $\mathfrak{m}$ and let ϕ be a solution to (1.2) with $q = 2$. Then, for every $p \geq 1$, there exists a constant*

$$C = C(\mathbb{T}^2, p) < \infty$$

such that

$$W_p^p(\mu,\nu) \leq C(\text{ess-inf}\,\rho_1)^{1-p} \int_{\mathbb{T}^2} |\nabla\phi|^p \, d\mathfrak{m}. \tag{2.1}$$

We notice that the bound above is asymmetric in the roles of μ and ν since only ρ_1 is required to be (essentially) bounded from below. In some sense, our work aims to sharpen (2.1) by replacing the linear Poisson equation with the non-linear q-Poisson one, and indeed Proposition 4.1 below is proved using a similar argument. However, (2.1) is useful as one can combine it with harmonic analysis tools, as done e.g. in [7, 24]. For example, for any $p > 1$, by the classical boundedness of the Riesz transform operator $\nabla(-\Delta)^{-1/2}$ on $\mathbb{T}^2$, where $(-\Delta)^{-1/2}$ is defined as a Fourier multiplier, one can further bound from above

$$\int_{\mathbb{T}^2} |\nabla\phi|^p \, d\mathfrak{m} \leq C \int_{\mathbb{T}^2} \left|(-\Delta)^{-1/2}(\rho_1 - \rho_0)\right|^p d\mathfrak{m},$$

where $C = C(\mathbb{T}^2, p) < \infty$. Hence, from (2.1), we further deduce the upper bound

$$W_p^p(\mu, \nu) \le C(\text{ess-inf}\,\rho_1)^{1-p} \int_{\mathbb{T}^2} \left|(-\Delta)^{-1/2}(\rho_1 - \rho_0)\right|^p \mathrm{d}\mathfrak{m}, \tag{2.2}$$

where again $C = C(\mathbb{T}^2, p) < \infty$.

2.2. *Viscosity solutions*

Viscosity solutions are designed to give a suitable notion of solution (with good properties such as uniqueness, stability, and comparison principles) for general nonlinear equations for which the distributional point of view does not make sense, as fully nonlinear PDEs. However, this notion reveals to be useful also for PDEs having a distributional formulation. This is the case of the q-Laplace (also called q-Poisson) equation considered in this paper, associated with the differential operator

$$-\Delta_q u := -\operatorname{div}\left(|\nabla u|^{q-2}\nabla u\right).$$

Actually, we will just deal with supersolutions.

DEFINITION 2.2. Let $g : \mathbb{T}^2 \to \mathbb{R}$. We say that a function

$$f : \mathbb{T}^2 \to (-\infty, +\infty]$$

is a *viscosity supersolution* for the equation $-\Delta_q u + g = 0$, and we write

$$-\Delta_q u + g \ge 0 \quad \text{in the viscosity sense} \tag{2.3}$$

if the following conditions hold:

(i) f is lower semicontinuous, $f \not\equiv +\infty$, and

(ii) whenever $x_0 \in \mathbb{T}^2$ and $\varphi \in C^2(\mathbb{T}^2)$ are such that $f - \varphi$ has a local minimum at x_0 and $\nabla\varphi(x_0) \neq 0$, we have

$$-\Delta_q\varphi(x_0) + g(x_0) \ge 0.$$

Definition 2.2 is adapted to the special form of the q-Laplace PDE. Indeed, the additional requirement $\nabla\varphi(x_0) \neq 0$ (not present in the general theory of viscosity solutions, see for instance [13]) is due to the fact that the expression

$$\Delta_q\varphi = |\nabla\varphi|^{q-4}\left[|\nabla\varphi|^2\Delta\varphi + (q-2)\sum_{i,j=1}^n \frac{\partial\varphi}{\partial x_i}\frac{\partial\varphi}{\partial x_j}\frac{\partial^2\varphi}{\partial x_i \partial x_j}\right] \tag{2.4}$$

is singular at the critical points of φ, when $1 < q < 2$.

Remark 2.3. With this convention, any $f \in C^2(\mathbb{T}^2)$ satisfying $-\Delta_q f + g \geq 0$ in the pointwise sense is also a viscosity supersolution. This follows from the fact that if we call

$$F_q(v, S) : \big(\mathbb{R}^2 \setminus \{0\}\big) \times \mathrm{Sym}^{2\times 2}(\mathbb{R}) \to \mathbb{R}$$

the differential operator such that $F_q(\nabla u, \nabla^2 u) = -\Delta_q u$, then F is non-increasing with respect to S(just look at (2.4)). It follows that if $f - \varphi$ has a local minimum at x_0 with $\nabla\varphi(x_0) \neq 0$, then $\nabla f(x_0) = \nabla\varphi(x_0) \neq 0$ and

$$F_q\big(\nabla\varphi(x_0), \nabla^2\varphi(x_0)\big) + g(x_0) \geq F_q\big(\nabla f(x_0), \nabla^2 f(x_0)\big) + g(x_0) \geq 0$$

as $\nabla^2 f(x_0) \geq \nabla^2\varphi(x_0)$.

2.3. *Hopf–Lax semigroup*

Given $f : \mathbb{T}^2 \to \mathbb{R}$ lower semicontinuous, let $u = Q_t f$ be the *Hopf–Lax semigroup* associated with the Hamilton–Jacobi equation

$$\partial_t u + \frac{|\nabla u|^q}{q} = 0; \tag{2.5}$$

that is,

$$(Q_t f)(x) = \min_{y \in \mathbb{T}^2} \left\{ f(y) + \frac{\mathbf{d}^p(x, y)}{p t^{p-1}} \right\}. \tag{2.6}$$

The following properties of the semigroup $Q_t f$, with $Q_0 f = f$, are well known; see for instance [3, Proposition 3.3] for a detailed proof.

Proposition 2.4. *Let $f : \mathbb{T}^2 \to \mathbb{R}$ be Lipschitz. Then, the functions $Q_t f$ are Lipschitz, uniformly with respect to $t \in [0, 1]$, $t \mapsto Q_t f$ is Lipschitz from $[0, 1]$ to $C(\mathbb{T}^2)$, and the PDE (2.5) is satisfied almost everywhere in $(0, 1) \times \mathbb{T}^2$.*

2.4. *Heat kernel on $\mathbb{T}^2$*

We recall that the heat kernel on the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ is given by

$$p_t(x) := \sum_{\mathbf{n} \in \mathbb{Z}^2} \bar{p}_t(x + \mathbf{n}), \tag{2.7}$$

where $\bar{p}_t(x) = \frac{1}{4\pi t} e^{-\frac{|x|^2}{4t}}$, $x \in \mathbb{R}^2$, is the Euclidean heat kernel. Given a probability measure μ in $\mathbb{T}^2$, we denote by $P_t\mu \ll \mathfrak{m}$ the probability measure having density

$$\rho(x) = \int_{\mathbb{T}^2} p_t(x - y)\, \mathrm{d}\mu(y).$$

Let us recall that $(P_t)_{t\geq 0}$ defines a symmetric Markov (convolution) semigroup, $P_{s+t} = P_s \circ P_t$ with (unique) invariant measures $\mathfrak{m}$ and generator given by the (distributional) Laplacian. Let us recall the following deterministic dispersion bound, directly coming from the coupling

$$\Sigma = \int_{\mathbb{T}^2} p_t(z)\Sigma_z \,\mathrm{d}\mathfrak{m}(z) \quad \text{with } \Sigma_z = (\mathrm{Id}\times\tau_z)_\#\mu$$

between μ and $P_t\mu$ (where τ_z is the shift map):

$$W_p(\mu, P_t\mu) \leq C_0\sqrt{t} \quad \forall t > 0, \tag{2.8}$$

with $C_0 = C_0(\mathbb{T}^2) = (\int_{\mathbb{T}^2} |z|^p p_1(z)\,\mathrm{d}\mathfrak{m}(z))^{1/p}$ for any probability measure μ in $\mathbb{T}^2$.

A remarkable fact, first noticed in [4, Theorem 5.2], is that the dispersion bound above can be significantly improved (in average) when applied to empirical measures μ^n as in (1.4).

PROPOSITION 2.5. *For every $p \geq 1$, there exists positive constant $C_1(\mathbb{T}^2, p)$, $C_2(\mathbb{T}^2, p)$ such that the following holds. If $t = \alpha/n \leq \frac{1}{2}$ with $\alpha \geq C_1(\mathbb{T}^2, p)\ln n$, then*

$$\mathbb{E}\left[W_p^p(\mu^n, P_t\mu^n)\right] \leq C_2(\mathbb{T}^2, p)\left(\frac{\ln\alpha}{n}\right)^{p/2}.$$

PROOF. The case $p = 2$ is established in [4, Theorem 5.2], and by the Hölder inequality, it entails the thesis for every $1 \leq p < 2$:

$$\mathbb{E}\left[W_p^p(\mu^n, P_t\mu^n)\right] \leq (C_2)^{p/2}\frac{(\ln\alpha)^{p/2}}{n^{p/2}}, \quad t = \frac{\alpha}{n},\ \alpha \geq C_1\ln n. \tag{2.9}$$

Therefore, it is sufficient to consider the case $p \geq 2$. To this aim, we combine the argument from [4] with the application of Rosenthal's inequality, from [24], where the upper bounds for the random bipartite matching cost are proved for any $p \geq 2$. By the triangle inequality and the elementary bound

$$|x+y|^p \leq 2^{p-1}\big(|x|^p + |y|^p\big) \tag{2.10}$$

for some $C = C(p) < \infty$, we find

$$\begin{aligned} &\mathbb{E}\left[W_p^p(\mu^n, P_t\mu^n)\right] \\ &\quad\leq 2^{p-1}\big(\mathbb{E}\left[W_p^p(\mu^n, P_{1/n}\mu^n)\right] + \mathbb{E}\left[W_p^p(P_{1/n}\mu^n, P_t\mu^n)\right]\big) \\ &\quad\leq 2^{p-1}\big(C_0 n^{-p/2} + \mathbb{E}\left[W_p^p(P_{1/n}\mu^n, P_t\mu^n)\right]\big), \end{aligned} \tag{2.11}$$

having used (2.8) in the second inequality. Thus, we are reduced to bound from above the expectation of $W_p^p(P_{1/n}\mu^n, P_t\mu^n)$. Since this random variable is always bounded from above by $\mathrm{diam}(\mathbb{T}^2)^p$, by choosing e.g. $d = 1/2$ in (2.15) of Proposition 2.6 below, we see that, if we pick $C_1 = (\ln a)^{-1}K$ sufficiently large – precisely such that $5 - Kd^2 < p/2$, we can safely reduce ourselves to argue on the event $\|\rho_{t,n} - 1\| \le 1/2$, so that $P_t\mu^n = \rho_{t,n}\mathfrak{m}$ has a density uniformly bounded from below by $1/2$. On such event, we use (2.2) (with $\mu = P_{1/n}\mu^n$ and $\nu = P_t\mu^n$), and we find

$$W_p^p(P_{1/n}\mu^n, P_t\mu^n) \le C\int_{\mathbb{T}^2}\left|(-\Delta)^{-1/2}(\rho_{1/n,n} - \rho_{t,n})\right|^p \mathrm{d}\mathfrak{m}, \tag{2.12}$$

where $C = C(\mathbb{T}^2, p) < \infty$. By the linearity of the operator $(-\Delta)^{-1/2}$, we collect the identity

$$(-\Delta)^{-1/2}(\rho_{1/n,n} - \rho_{t,n})(x) = \frac{1}{n}\sum_{i=1}^n \left[(-\Delta)^{-1/2}(p_{1/n} - p_t)\right](X_i - x),$$

and notice that, for each $x \in \mathbb{T}^2$, the random variables

$$\varphi_i(x) := \left[(-\Delta)^{-1/2}(p_{1/n} - p_t)\right](X_i - x), \quad \text{for } i = 1, \dots, n,$$

are independent and centered. After taking expectation in (2.12), we see that the thesis amounts to bound from above the quantity

$$\int_{\mathbb{T}^2}\mathbb{E}\left[\left|\frac{1}{n}\sum_{i=1}^n \varphi_i(x)\right|^p\right]\mathrm{d}\mathfrak{m}(x),$$

where we recognize, for every x, the p-th moment of a sum of independent centered random variables. By Rosenthal's inequality, [28], we have for some constant $C = C(p) < \infty$,

$$\mathbb{E}\left[\left|\frac{1}{n}\sum_{i=1}^n \varphi_i(x)\right|^p\right] \le C\left[\frac{1}{n^{p-1}}\mathbb{E}\left[|\varphi(x)|^p\right] + \frac{1}{n^{p/2}}\mathbb{E}\left[|\varphi(x)|^2\right]^{p/2}\right], \tag{2.13}$$

where we write $\varphi := [(-\Delta)^{-1/2}(p_{1/n} - p_t)](X_1 - x)$. To conclude, we follow very closely the argument in [24, (34)] onwards (in the case $d = 2$), so we omit some details. We collect first the uniform bound, valid for $0 < s \le 1/2$:

$$\sup_{z\in\mathbb{T}^2}\left|(-\Delta)^{-1/2}(p_s - 1)(z)\right| \le \frac{C}{s^{1/2}},$$

which we apply in particular to $s \in \{1/n, t\}$, yielding

$$\begin{aligned}\sup_{z\in\mathbb{T}^2}|\varphi(z)| &\le \sup_{z\in\mathbb{T}^2}\left|(-\Delta)^{-1/2}(p_{1/n} - 1)(z)\right| + \sup_{z\in\mathbb{T}^2}\left|(-\Delta)^{-1/2}(p_t - 1)(z)\right| \\ &\le Cn^{1/2}.\end{aligned}$$

Then, by the representation $(-\Delta)^{-1} = \int_0^\infty P_s ds$, we find for any $f \in L^2(\mathbb{T}^2)$ with $\int_{\mathbb{T}^2} f \, \mathrm{d}\mathfrak{m} = 0$ that

$$\int_{\mathbb{T}^2} \left[(-\Delta)^{-1/2} f\right]^2 \mathrm{d}\mathfrak{m} = \int_{\mathbb{T}^2} f(-\Delta)^{-1} f \, \mathrm{d}\mathfrak{m} = \int_0^\infty \int_{\mathbb{T}^2} f P_s f \, \mathrm{d}\mathfrak{m} \, \mathrm{d}s$$
$$= \int_0^\infty \int_{\mathbb{T}^2} (P_{s/2} f)^2 \, \mathrm{d}\mathfrak{m} \, \mathrm{d}s.$$

We use this identity in our case, i.e., with $f = p_{1/n} - p_t$, yielding

$$\begin{aligned}
\mathbb{E}\left[|\varphi(x)|^2\right] &= \int_{\mathbb{T}^2} \left[(-\Delta)^{-1/2}(p_{1/n} - p_t)\right]^2 (y - x) \, \mathrm{d}\mathfrak{m}(y) \\
&= \int_0^\infty \int_{\mathbb{T}^2} \left(p_{s/2+1/n}(y - x) - p_{s/2+t}(y - x)\right)^2 \mathrm{d}\mathfrak{m}(y) \, \mathrm{d}s \\
&= \int_0^\infty \left[p_{s+2/n}(0) + p_{s+2t}(0) - 2p_{s+t+1/n}(0)\right] \mathrm{d}s \\
&= O\big(-\log(2/n) - \log(2t) + 2\log(t + 1/n) + 1\big) = O(\ln \alpha),
\end{aligned}$$

where in developing the square we invoked the semigroup property (so that, for any $t_1, t_2 > 0$, $\int_{\mathbb{T}^2} p_{t_1}(y - x) p_{t_2}(y - x) \, \mathrm{d}\mathfrak{m}(y) = P_{t_1+t_2}\delta_x(x) = p_{t_1+t_2}(0)$), and the final asymptotics can be computed directly from (2.7).

Combining these bounds, we find

$$\mathbb{E}\left[|\varphi(x)|^p\right] \le \sup_{z \in \mathbb{T}^2} |\varphi(z)|^{p-2} \, \mathbb{E}\left[|\varphi(x)|^2\right] \le C n^{(p-2)/2} \ln \alpha,$$

and therefore we bound from above the right-hand side in (2.13) with

$$\left[\frac{1}{n^{p-1}} \mathbb{E}\left[|\varphi(x)|^p\right] + \frac{1}{n^{p/2}} \mathbb{E}\left[|\varphi(x)|^2\right]^{p/2}\right] \le C \frac{\ln \alpha}{n^{p/2}} + C \left(\frac{\ln \alpha}{n}\right)^{p/2} \tag{2.14}$$

and the thesis follows. ∎

In the proof above, we used a regularizing property of the heat semigroup, when acting on empirical measures, as established in [4] (see Theorem 3.3 and Remark 3.17 therein), that we report here.

PROPOSITION 2.6. *If μ^n are as in (1.4) and $P_t \mu^n = \rho_{t,n} \mathfrak{m}$, then*

$$\mathbb{P}\big(\{\|\rho_{t,n} - 1\|_\infty > d\}\big) \le \frac{C_3(\mathbb{T}^2)}{d^2 t^3} a^{-ntd^2} \text{ for some } C_3(\mathbb{T}^2) > 0 \text{ and } a = a(\mathbb{T}^2) > 1.$$

In particular, if $d \ge n^{-1}$ and $t = (\ln a)^{-1} K n^{-1} \ln n$ with $K \ge 1$, then

$$\mathbb{P}\big(\{\|\rho_{t,n} - 1\|_\infty > d\}\big) \le C_3(\mathbb{T}^2)(\ln a)^3 n^{5-Kd^2}. \tag{2.15}$$

Our strategy for proving Theorem 1.1 will be to adjust the parameters $K = K_n \to \infty$ and $d = d_n \to 0$ in such a way that Kd is sufficiently large, so that the probability of the deviation from the constant density 1 will have the power like decay we need with respect to n.

We will also need L^p estimates on $\rho_{t,n}$, provided by the following proposition.

PROPOSITION 2.7. *Let t_n be as in Theorem 1.1, and $K = K_n$ related to t_n as in Proposition 2.6. Fixing $k > 0$, take $c_n \to 0^+$ such that*

$$\liminf_n K_n c_n^2 > k + 5. \tag{2.16}$$

Then,

$$\sup \left\{ n^k \, \mathbb{E} \left[\mathbf{1}_{\{\|\rho_{t_n,n}-1\|_\infty > c_n\}} \int_{\mathbb{T}^2} |\rho_{t_n,n} - 1|^p \, \mathrm{d}\mathfrak{m} \right] : \ n \geq 2 \right\} < \infty.$$

PROOF. In this proof, C denotes a positive constant, depending only on $\mathbb{T}^2$. Arguing as in [4, proof of Theorem 3.3], the bounds

$$\mathbb{E}[Y^2] \leq \frac{C}{t}, \quad |Y| \leq \frac{C}{t}, \quad t \in (0,1)$$

for the random variables $Y = Y_i = p_t(X_i, y) - 1$, together with Bernstein's inequality yield

$$\mathbb{P}\big(\{|\rho_{t,n}(y) - 1| > \xi\}\big) \leq C \exp(-nct\xi) \quad \forall t \in (0,1), \ \xi > 1$$

for all $y \in \mathbb{T}^2$. For our choice of $t = t_n$, Fubini's theorem and Cavalieri's formula yield

$$n^k \, \mathbb{E} \left[\int_{\{|\rho_{t_n,n}-1|>1\}} \left|\rho_{t_n,n} - 1\right|^p \mathrm{d}\mathfrak{m} \right] \leq C \int_1^\infty n^{k-c(\ln a)^{-1} K_n \xi} \cdot \xi^{p-1} \, \mathrm{d}\xi.$$

Thus, for $n \gg 1$,

$$\begin{aligned} &n^k \, \mathbb{E} \left[\int_{\{|\rho_{t_n,n}-1|>1\}} \left|\rho_{t_n,n} - 1\right|^p \mathrm{d}\mathfrak{m} \right] \\ &\quad \leq C \int_1^\infty n^{\xi(k-c(\ln a)^{-1} K_n)} \xi^{p-1} d\xi \leq C \int_1^\infty 2^{-\xi} \xi^{p-1} \, \mathrm{d}\xi < \infty. \end{aligned}$$

On the other hand, exploiting Proposition 2.6 along with (2.16), we get

$$\begin{aligned} &n^k \, \mathbb{E} \left[\mathbf{1}_{\{\|\rho_{t_n,n}-1\|_\infty > c_n\}} \int_{\{|\rho_{t_n,n}-1|\leq 1\}} \left|\rho_{t_n,n} - 1\right|^p \mathrm{d}\mathfrak{m} \right] \\ &\quad \leq n^k \mathbb{P}\big(\{\|\rho_{t_n,n} - 1\|_\infty > c_n\}\big) \leq C_3(\mathbb{T}^2)(\ln a)^3 n^{k+5-K_n c_n^2} \to 0. \qquad \blacksquare \end{aligned}$$

3. Propagation of q-Laplacian estimates and differentiation of $\int |\nabla Q_t\phi|^q \, \mathrm{dm}$

Recalling the definition of $c \geq 0$ in Theorem 1.2, $\phi \in H^{1,q}(\mathbb{T}^2)$ satisfies in a distributional sense the inequality

$$-\operatorname{div}\left(|\nabla\phi|^{q-2}\nabla\phi\right) + c \geq 0. \tag{3.1}$$

Namely, for every non-negative $\eta \in C^\infty(\mathbb{T}^2)$, we have

$$\int_{\mathbb{T}^2} |\nabla\phi|^{q-2}\langle\nabla\phi, \nabla\eta\rangle \, \mathrm{dm} + c\int_{\mathbb{T}^2} \eta \, \mathrm{dm} \geq 0. \tag{3.2}$$

In order to control the time derivative of $\int_{\mathbb{T}^2} |\nabla Q_t\phi|^q \, \mathrm{dm}$, we would like to show that (3.2) propagates with the Hopf–Lax semigroup; that is, it is satisfied also by $Q_t\phi$ for any $t \in (0, 1)$. The proof of this stability property becomes much easier if we understand (3.1) in the viscosity sense; this is possible thanks to the following result (see [20, 21] for the homogeneous case $g = 0$ and Remark 3.3 below).

THEOREM 3.1. *Let $f \in H^{1,q}(\mathbb{T}^2)$ and $g : \mathbb{T}^2 \to \mathbb{R}$ be continuous. Then, $-\Delta_q f + g \geq 0$ in the viscosity sense, according to Definition 2.2, if and only if $-\Delta_q f + g \geq 0$ in the sense of distributions.*

We are going to use Theorem 3.1 both ways: first we pass from the distributional sense for ϕ, granted by (1.1), to the viscosity sense, and then we pass from the viscosity sense to the distributional sense for $Q_s\phi$ in the proof of Lemma 3.4.

Then, let us show the propagation of the estimate $-\Delta_q\phi + c \geq 0$ to $Q_t\phi$ in the viscosity sense. Actually, it will be useful to prove this property for the Hopf–Lax semigroup associated with any power $r > 1$. We provide a direct proof, even though the statement could directly follow by the general fact that viscosity supersolutions to $-\Delta_q + c \geq 0$ are stable under translations in the dependent and independent variables, and infimum.

PROPOSITION 3.2. *Let $f : \mathbb{T}^2 \to \mathbb{R}$ be lower semicontinuous and satisfying $-\Delta_q f + c \geq 0$ in the viscosity sense and $r \in (1, \infty)$. Then, for all $t > 0$, the function*

$$f_t(x) := \min_{y\in\mathbb{T}^2}\left\{f(y) + \frac{\mathbf{d}^r(x, y)}{rt^{r-1}}\right\}$$

still satisfies $-\Delta_q f_t + c \geq 0$ in the viscosity sense.

PROOF. Given $t > 0$ and $x_0 \in \mathbb{T}^2$, let $y_0 \in \mathbb{T}^2$ be a point where the minimum in the definition of f_t is attained, so that

$$f_t(x_0) = f(y_0) + \frac{\mathbf{d}^r(x_0, y_0)}{rt^{r-1}}.$$

Consider $\varphi \in C^2(\mathbb{T}^2)$ such that $f_t - \varphi$ has a local minimum in x_0 and, with no loss of generality, assume that the minimum is global and $f_t(x_0) = \varphi(x_0)$.

If we set $\psi(x) := \varphi(x - y_0 + x_0)$, we claim that $\phi - \psi$ has a minimum in y_0, equal to $-\mathbf{d}^r(x_0, y_0)/(rt^{r-1})$. From this we would obtain

$$F_q\big(\nabla\psi(y_0), \nabla^2\psi(y_0)\big) \le c$$

and thus

$$F_q\big(\nabla\varphi(x_0), \nabla^2\varphi(x_0)\big) \le c.$$

To prove the claim, we notice that

$$\phi(y_0) - \psi(y_0) = \phi(y_0) - \varphi(x_0) = \phi(y_0) - f_t(x_0) = -\frac{1}{rt^{r-1}}\mathbf{d}^r(x_0, y_0),$$

while on the other hand, $f_t(x) \ge \varphi(x)$ implies

$$\phi(y) + \frac{1}{rt^{r-1}}\mathbf{d}^r(x, y) \ge \varphi(x) \quad \forall x, y.$$

Choosing $y = x - x_0 + y_0$ (understanding the sum modulo $\mathbb{Z}^2$), we obtain

$$\phi(y) - \psi(y) \ge -\frac{1}{rt^{r-1}}\mathbf{d}^r(x_0, y_0) \quad \forall y,$$

as desired. ■

REMARK 3.3. We can use Proposition 3.2 to provide a sketchy proof of the implication from viscous to distributional granted, also in the converse direction, by Theorem 3.1. Indeed, we can use the Hopf–Lax semigroup with power $r = 2$ to obtain that $f_s = Q_s f$ still satisfy $-\Delta_q f_s + c \ge 0$ in the viscosity sense and $C^{1,1}$ regularity of f_s. Since $f_s \to f$ in $H^{1,q}(\mathbb{T}^2)$ as $s \to 0^+$, it is then sufficient to show that $-\Delta_q f_s + c \ge 0$ in the sense of distributions. Here, we can use the $C^{1,1}$ regularity of f_s to build appropriate test functions ϕ, of the form

$$\phi(x) = f_s(x_0) + \big\langle \nabla f_s(x_0), x - x_0\big\rangle + \frac{1}{2}\big\langle \nabla^2 f_s(x_0)(x - x_0), (x - x_0)\big\rangle - \varepsilon|x - x_0|^2$$

at any point $x_0 \in \mathbb{T}^2$ where $\nabla f_s(x_0) \ne 0$ and $\nabla^2 f_s(x_0)$ exists. This leads to the validity of $-\Delta_q f_s + c \ge 0$ almost everywhere in the open set $\Omega_s = \{|\nabla f_s| \ne 0\}$. Then, one obtains the validity of the inequality in the sense of distributions first in Ω_s and then on the whole of $\mathbb{T}^2$, using the fact that the flux of the continuous vector field $|\nabla f_s|^{q-2}\nabla f_s$ is null on the boundary (because $q > 1$). If Ω_s is not smooth, one can perform a further approximation since

$$\frac{1}{\varepsilon}\int_0^\varepsilon \int_{\{|\nabla f_s| = \tau\}} |\nabla f_s|^{q-1}\, \mathrm{d}\mathscr{H}^1\, \mathrm{d}\tau = \int_{\{0 < |\nabla f_s| < \varepsilon\}} |\nabla f_s|^q\, \mathrm{d}\mathfrak{m}$$

tends to 0 as $\varepsilon \to 0$.

Now, we apply Proposition 3.2 with $f = \phi$ and $r = p$ in order to estimate the variation in time of $\int_{\mathbb{T}^2} |\nabla Q_t \phi|^q \,\mathrm{d}\mathfrak{m}$.

LEMMA 3.4. *Let $\Lambda(t) := \int_{\mathbb{T}^2} |\nabla Q_t \phi|^q \,\mathrm{d}\mathfrak{m}$ with ϕ as in (1.2) and $c = \|\rho_1 - \rho_0\|_\infty$. Then, Λ is Lipschitz in $[0, 1]$ and $\frac{\mathrm{d}}{\mathrm{d}t}\Lambda(t) \le c\Lambda(t)$ for almost every $t \in (0, 1)$. In particular,*

$$\int_{\mathbb{T}^2} |\nabla Q_t \phi|^q \,\mathrm{d}\mathfrak{m} \le e^{ct} \int_{\mathbb{T}^2} |\nabla \phi|^q \,\mathrm{d}\mathfrak{m} \quad \forall t \in [0, 1]. \tag{3.3}$$

PROOF. Thanks to Proposition 3.2, $f_t = Q_t \phi$ satisfy $-\Delta_q f_t + c \ge 0$ in the viscosity sense. Therefore, Theorem 3.1 grants this property also in the sense of distributions; namely (notice that the improvement from $C^\infty(\mathbb{T}^2)$ to $H^{1,q}(\mathbb{T}^2)$ follows by density and L^p integrability of $|\nabla f_t|^{q-2}\nabla f_t$),

$$\int_{\mathbb{T}^2} |\nabla f_t|^{q-2} \langle \nabla f_t, \nabla \eta \rangle \,\mathrm{d}\mathfrak{m} + c \int_{\mathbb{T}^2} \eta \,\mathrm{d}\mathfrak{m} \ge 0 \quad \forall \eta \in H^{1,q}(\mathbb{T}^2),\ \eta \ge 0. \tag{3.4}$$

First, we note that, by (3.4), the distribution $T := -\operatorname{div}(|\nabla f_t|^{q-2}\nabla f_t) + c$ is non-negative. Thus, if $\eta \in C^\infty(\mathbb{T}^2)$,

$$\langle T, \eta \rangle \le \big\langle T, \|\eta\|_\infty 1 \big\rangle = \|\eta\|_\infty c$$

and then T is represented by a non-negative finite measure with mass less than or equal to c (here we used that $\langle \operatorname{div}(|\nabla f_t|^{q-2}\nabla f_t), 1\rangle = 0$ and therefore $\langle T, 1 \rangle = c$). It follows that $\mu_t := \operatorname{div}(|\nabla f_t|^{q-2}\nabla f_t)$ is a signed measure with $\|\mu_t\| \le 2c$.

By the convexity of $y \mapsto |y|^q$, we then infer

$$\begin{aligned} \Lambda(t) - \Lambda(s) &\ge q \int_{\mathbb{T}^2} |\nabla f_s|^{q-2} \big\langle \nabla f_s, \nabla (f_t - f_s) \big\rangle \,\mathrm{d}\mathfrak{m} = q \int_{\mathbb{T}^2} (f_s - f_t) \,\mathrm{d}\mu_s \\ &\ge -2cq \|f_t - f_s\|_\infty \end{aligned} \tag{3.5}$$

for every $s, t \in [0, 1]$. From the Lipschitz regularity of the initial datum ϕ (which follows by [29, Theorem 2.1]) and Proposition 2.4, we deduce that the map $t \mapsto f_t$ is Lipschitz with respect to the sup norm, let us say with constant L. Hence, exchanging the roles of t and s, we conclude that

$$\big|\Lambda(t) - \Lambda(s)\big| \le 2cqL|t - s|,$$

as we desired.

Now, we can refine (3.5) as follows. Let $t \in (0, 1)$ be a differentiability point for Λ such that $-q\frac{\mathrm{d}}{\mathrm{d}t} f_t = |\nabla f_t|^q$ a.e. in $\mathbb{T}^2$. Thanks to Rademacher's theorem and Proposition 2.4, both properties are satisfied for a.e. $t \in (0, 1)$. For $s \ge t$, using the inequality $f_t \geq f_s$ granted directly from the definition (2.6), as well as the inequality $-\operatorname{div}(|\nabla f_s|^{q-2}\nabla f_s) + c \geq 0$ in the sense of distributions, we get

$$\begin{aligned}\Lambda(s) - \Lambda(t) &\leq -q \int_{\mathbb{T}^2} |\nabla f_s|^{q-2}\big\langle \nabla f_s, \nabla(f_t - f_s)\big\rangle \,\mathrm{d}\mathfrak{m} \\ &= q \int_{\mathbb{T}^2} \operatorname{div}\big(|\nabla f_s|^{q-2}\nabla f_s\big)(f_t - f_s)\,\mathrm{d}\mathfrak{m} \\ &\leq cq \int_{\mathbb{T}^2} (f_t - f_s)\,\mathrm{d}\mathfrak{m},\end{aligned}$$

so that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Lambda(t) &= \lim_{s\to t+} \frac{\Lambda(s) - \Lambda(t)}{s-t} \leq \lim_{s\to t+} cq \int_{\mathbb{T}^2} \frac{f_t - f_s}{s-t}\,\mathrm{d}\mathfrak{m} \\ &= -cq \int_{\mathbb{T}^2} \frac{\mathrm{d}}{\mathrm{d}t} f_t \,\mathrm{d}\mathfrak{m} = c\int_{\mathbb{T}^2} |\nabla f_t|^q \,\mathrm{d}\mathfrak{m},\end{aligned}$$

which proves that $\Lambda'(t) \leq c\Lambda(t)$. Finally, the validity of (3.3) follows by Gronwall's lemma. ∎

4. Proof of Theorem 1.2

In this section, ϕ, c are as in Theorem 1.2.

4.1. Upper bound

The upper bound in Theorem 1.2 can be obtained immediately by repeating the argument in [7, Proposition 2.3], involving duality and the Hopf–Lax formula. We still give the proof here for the sake of completeness.

Since $\mathbb{T}^2$ is compact, the duality formula for W_p^p can be written in the form

$$\begin{aligned}\frac{1}{p} W_p^p(\rho_0\mathfrak{m}, \rho_1\mathfrak{m}) \\ = \sup\left\{ -\int_{\mathbb{T}^2} f\rho_0 \,\mathrm{d}\mathfrak{m} + \int_{\mathbb{T}^2} (Q_1 f)\rho_1 \,\mathrm{d}\mathfrak{m} : f : \mathbb{T}^2 \to \mathbb{R} \text{ Lipschitz}\right\}.\end{aligned} \tag{4.1}$$

Proposition 4.1 (Upper bound). *There exists $\overline{\delta}(c,p)$ such that $\overline{\delta}(c,p) \to 0$ as $c \to 0$ and*

$$W_p^p(\rho_0\mathfrak{m}, \rho_1\mathfrak{m}) \leq \big(1 + \overline{\delta}(c,p)\big) \int_{\mathbb{T}^2} |\nabla\phi|^q \,\mathrm{d}\mathfrak{m}.$$

PROOF. Let us bound uniformly the argument of the supremum in (4.1), for $f:\mathbb{T}^2\to\mathbb{R}$ Lipschitz, exploiting the PDE (1.2) satisfied by ϕ, the fact that $Q_t f$ solves (2.5) almost everywhere in $(0,1)\times\mathbb{T}^2$ and dominated convergence to put $\frac{\mathrm{d}}{\mathrm{d}s}$ under the integral sign. If we set $\rho_t := t\rho_1 + (1-t)\rho_0$ per $t\in(0,1)$, then

$$
\begin{aligned}
(4.2)\qquad & \int_{\mathbb{T}^2} (\rho_1 Q_1 f - \rho_0 f)\,\mathrm{d}\mathfrak{m} \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\int_{\mathbb{T}^2} \rho_s Q_s f \,\mathrm{d}\mathfrak{m}\,\mathrm{d}s \\
&= \int_0^1 \int_{\mathbb{T}^2} \left(\rho_s \frac{\mathrm{d}}{\mathrm{d}s} Q_s f + (\rho_1-\rho_0) Q_s f\right)\mathrm{d}\mathfrak{m}\,\mathrm{d}s \\
&= \int_0^1 \int_{\mathbb{T}^2} \left(-\frac{1}{q}|\nabla Q_s f|^q \rho_s + |\nabla\phi|^{q-2}\langle\nabla\phi,\nabla Q_s f\rangle\right)\mathrm{d}\mathfrak{m}\,\mathrm{d}s \\
&\le \int_0^1 \int_{\mathbb{T}^2} \left(-\frac{1}{q}|\nabla\phi|^q \rho_s^{-\frac{q}{q-1}}\rho_s + |\nabla\phi|^q \rho_s^{-\frac{1}{q-1}}\right)\mathrm{d}\mathfrak{m}\,\mathrm{d}s \\
&= \frac{1}{p}\int_{\mathbb{T}^2} \left(\int_0^1 \rho_s^{-\frac{1}{q-1}}\,\mathrm{d}s\right)|\nabla\phi|^q\,\mathrm{d}\mathfrak{m},
\end{aligned}
$$

where for the inequality we used that $v=\rho_s^{-\frac{1}{q-1}}\nabla\phi$ minimizes

$$v\mapsto \frac{1}{q}|v|^q\rho_s - |\nabla\phi|^{q-2}\langle\nabla\phi, v\rangle.$$

In conclusion,

$$W_p^p(\rho_0\mathfrak{m},\rho_1\mathfrak{m}) \le \int_{\mathbb{T}^2} M_q(\rho_0,\rho_1)|\nabla\phi|^q\,\mathrm{d}\mathfrak{m},$$

where $M_q(\rho_0,\rho_1)(x)=\int_0^1\rho_s(x)^{-\frac{1}{q-1}}\,\mathrm{d}s \lesssim 1$ as $c\to 0$. More precisely, since

$$\|\rho_i - 1\|_\infty \le c/2, \quad \text{for } c<2,$$

one has

$$M_q(\rho_0,\rho_1)(x)\le 1+\overline{\delta}(c,p)$$

with $\overline{\delta}(c,p) = (1-c/2)^{-\frac{1}{q-1}} - 1$. ∎

4.2. *Lower bound*

PROPOSITION 4.2 (Lower bound). *There exists* $\underline{\delta}(c,p)$ *such that* $\underline{\delta}(c,p)\to 0$ *as* $c\to 0$ *and*

$$W_p^p(\rho_0\mathfrak{m},\rho_1\mathfrak{m}) \ge \left(1-\underline{\delta}(c,p)\right)\int_{\mathbb{T}^2}|\nabla\phi|^q\,\mathrm{d}\mathfrak{m}.$$

PROOF. From the duality formula, with an integration by parts and Fubini's theorem, we get

$$\begin{aligned}\frac{1}{p}W_p^p(\rho_1\mathfrak{m},\rho_0\mathfrak{m}) &\geq -\int_{\mathbb{T}^2}\phi\rho_0\,\mathrm{d}\mathfrak{m}+\int_{\mathbb{T}^2}(Q_1\phi)\rho_1\,\mathrm{d}\mathfrak{m}\\ &=\int_{\mathbb{T}^2}\phi(\rho_1-\rho_0)\,\mathrm{d}\mathfrak{m}+\int_{\mathbb{T}^2}(Q_1\phi-\phi)\rho_1\,\mathrm{d}\mathfrak{m}\\ &=\int_{\mathbb{T}^2}|\nabla\phi|^q\,\mathrm{d}\mathfrak{m}-\frac{1}{q}\int_0^1\int_{\mathbb{T}^2}|\nabla Q_s\phi|^q\rho_1\,\mathrm{d}\mathfrak{m}\,\mathrm{d}s.\end{aligned}$$

Now, we can use first the inequality $\|\rho_1-1\|_\infty\leq c/2$ to replace ρ_1 with 1 and then Lemma 3.4 with $c\geq\|\rho_1-\rho_0\|_\infty$ to estimate

$$\frac{1}{p}W_p^p(\rho_1\mathfrak{m},\rho_0\mathfrak{m})\geq\frac{1}{p}\int_{\mathbb{T}^2}|\nabla\phi|^q\,\mathrm{d}\mathfrak{m}-\frac{ce^c/2+e^c-1}{q}\int_{\mathbb{T}^2}|\nabla\phi|^q\,\mathrm{d}\mathfrak{m},$$

so that $\underline{\delta}(c,p)=(p-1)(ce^c/2+e^c-1)$. ∎

5. PROOF OF THEOREM 1.1

In this section, we adopt the notation in the statement of Theorem 1.1.

5.1. Upper bound

Since $\ln(nt_n)\ll\ln n$, using Proposition 2.5 and the triangle inequality for W_p, arguing as in [7], the proof of the upper bound reduces to the following estimate:

$$\limsup_{n\to\infty}\left(\frac{n}{\ln n}\right)^{p/2}\left(\mathbb{E}\left[W_p^p(P_{t_n}\mu^n,P_{t_n}\nu^n)\right]-\mathbb{E}\left[\int_{\mathbb{T}^2}|\nabla\phi|^q\,\mathrm{d}\mathfrak{m}\right]\right)\leq 0 \tag{5.1}$$

where ϕ is the solution to (1.2) with the right-hand side

$$\rho_{0,n}\mathfrak{m}=P_{t_n}\mu^n,\quad \rho_{1,n}\mathfrak{m}=P_{t_n}\nu^n.$$

Now, since $t_n\gg n^{-1}\ln n$, we can use Proposition 2.6 to write t_n as $(\ln a)^{-1}K_nn^{-1}\ln n$ with $K_n\geq 1$ and $c_n\to 0$ in such a way that $K_nc_n^2>2p+10$, so that

$$\mathbb{P}\left(\left\{\|\rho_{i,n}-1\|_\infty>\frac{c_n}{2}\right\}\right)\leq C_3(\mathbb{T}^2)(\ln a)^3n^{5-K_nc_n^2/4}=O(n^{-p/2}),\quad i=0,1.$$

Since $W_p(\mu,\nu)\leq\operatorname{diam}(\mathbb{T}^2)$ for any pair of probability measures μ, ν, it follows that the contribution to (5.1) of the event $\{\max_i\|\rho_{i,n}-1\|_\infty>\frac{c_n}{2}\}$ is null, and in the complementary event, we can use Theorem 1.2 to conclude.

5.2. *Lower bound*

Recall that the semigroup P_t is contractive in $\mathbb{T}^2$ with respect to any W_p distance; this can be easily proved taking any coupling Σ between μ and ν and considering the average $\bar{\Sigma} = \int \Sigma_z p_t(z)\, d\mathfrak{m}(z)$ of the shifted couplings

$$\Sigma_z := (\tau_z \times \tau_z)_\# \Sigma \quad \text{with } \tau_z(x) = x + z,$$

which provides a coupling between $P_t\mu$ and $P_t\nu$ with the same cost. Therefore, the lower bound

$$\liminf_{n\to\infty} \left(\frac{n}{\ln n}\right)^{p/2} \left(\mathbb{E}\left[W_p^p(\mu^n, \nu^n)\right] - \mathbb{E}\left[\int_{\mathbb{T}^2} |\nabla\phi|^q \, d\mathfrak{m}\right]\right) \geq 0 \tag{5.2}$$

can be deduced from

$$\liminf_{n\to\infty} \left(\frac{n}{\ln n}\right)^{p/2} \left(\mathbb{E}\left[W_p^p(\rho_{0,n}\mathfrak{m}, \rho_{1,n})\mathfrak{m}\right] - \mathbb{E}\left[\int_{\mathbb{T}^2} |\nabla\phi|^q \, d\mathfrak{m}\right]\right) \geq 0. \tag{5.3}$$

Now, recall that the solution ϕ to (1.1) is the unique minimizer of the functional

$$\Lambda_q(f) := \int_{\mathbb{T}^2} \frac{1}{q}|\nabla f|^q - f(\rho_1 - \rho_0)\, d\mathfrak{m} = \int_{\mathbb{T}^2} \frac{1}{q}|\nabla f|^q - (f - \bar{f})(\rho_1 - \rho_0)\, d\mathfrak{m}$$

whose minimum value is non-positive. Hence, from the Sobolev embedding, we obtain

$$\frac{1}{q}\int_{\mathbb{T}^2} |\nabla\phi|^q \, d\mathfrak{m} \leq \|\rho_1 - \rho_0\|_p \|\phi\|_q \leq c_S \|\rho_1 - \rho_0\|_p \left(\int_{\mathbb{T}^2} |\nabla\phi|^q \, d\mathfrak{m}\right)^{1/q}$$

and then the deterministic upper bound

$$\int_{\mathbb{T}^2} |\nabla\phi|^q \, d\mathfrak{m} \leq \left(\|\rho_1 - \rho_0\|_p c_S q\right)^p. \tag{5.4}$$

As in the proof of the upper bound, since $nt_n \gg \ln n$, we can use this time Proposition 2.7 that provides an estimate in expectation on $\|\rho_{i,n} - 1\|_p^p$ to show that the contribution to (5.3) of the event $\{\max_i \|\rho_{i,n} - 1\|_\infty > \frac{c_n}{2}\}$ is null (if we also require $K_n c_n^2 > 2p + 20$ in order to satisfy (2.16) with $\frac{c_n}{2}$ and $k = \frac{p}{2}$), and in the complementary event, we can use Theorem 1.2 to conclude.

Acknowledgments. – The authors thank N. Gigli for having pointed out to them [26, Section 4.3] where, on Riemannian manifolds, via the theory of characteristics, the preservation of upper bounds on the q-Laplacian under the action of the p-Hopf–Lax semigroup is shown, even in the nonlinear case $p \neq 2$, before shocks. Eventually, in the case of the flat space $\mathbb{T}^2$, the proof we gave in Proposition 3.2 does not use this computation and works even beyond shocks, using solutions in the viscosity sense.

Funding. – L. A. and F. V. acknowledge the PRIN Italian grant 202244A7YL "Gradient Flows and Non-Smooth Geometric Structures with Applications to Optimization and Machine Learning". D. T. acknowledges the MUR Excellence Department Project awarded to the Department of Mathematics, University of Pisa, CUP I57G22000700001, the Italian National Centre for HPC, Big Data and Quantum Computing - Proposal code CN1 CN00000013, CUP I53C22000690001, the PRIN Italian grant 2022WHZ5XH - "understanding the LEarning process of QUantum Neural networks (LeQun)", CUP J53D23003890006, the INdAM-GNAMPA project 2023 "Teoremi Limite per Dinamiche di Discesa Gradiente Stocastica: Convergenza e Generalizzazione", INdAM-GNAMPA project 2024 "Tecniche analitiche e probabilistiche in informazione quantistica" and the project G24-202 "Variational methods for geometric and optimal matching problems" funded by Università Italo Francese. Research also partly funded by PNRR - M4C2 - Investimento 1.3, Partenariato Esteso PE00000013 - "FAIR - Future Artificial Intelligence Research" - Spoke 1 "Human-centered AI", funded by the European Commission under the Next Generation EU programme.

References

[1] M. Ajtai – J. Komlós – G. Tusnády, On optimal matchings. *Combinatorica* **4** (1984), no. 4, 259–264. Zbl 0562.60012 MR 0779885

[2] L. Ambrosio – E. Brué – D. Semola, *Lectures on optimal transport*. Unitext 130, Springer, Cham, 2021. Zbl 1485.49001 MR 4294651

[3] L. Ambrosio – N. Gigli – G. Savaré, Density of Lipschitz functions and equivalence of weak gradients in metric measure spaces. *Rev. Mat. Iberoam.* **29** (2013), no. 3, 969–996. Zbl 1287.46027 MR 3090143

[4] L. Ambrosio – F. Glaudo, Finer estimates on the 2-dimensional matching problem. *J. Éc. polytech. Math.* **6** (2019), 737–765. Zbl 1434.60054 MR 4014635

[5] L. Ambrosio – F. Glaudo – D. Trevisan, On the optimal map in the 2-dimensional random matching problem. *Discrete Contin. Dyn. Syst.* **39** (2019), no. 12, 7291–7308. Zbl 1476.60020 MR 4026190

[6] L. Ambrosio – M. Goldman – D. Trevisan, On the quadratic random matching problem in two-dimensional domains. *Electron. J. Probab.* **27** (2022), article no. 54. Zbl 1487.60017 MR 4416128

[7] L. Ambrosio – F. Stra – D. Trevisan, A PDE approach to a 2-dimensional matching problem. *Probab. Theory Related Fields* **173** (2019), no. 1-2, 433–477. Zbl 1480.60017 MR 3916111

[8] D. Benedetto – E. Caglioti, Euclidean random matching in 2D for non-constant densities. *J. Stat. Phys.* **181** (2020), no. 3, 854–869. Zbl 1458.60020 MR 4160913

[9] D. Benedetto – E. Caglioti – S. Caracciolo – M. D'Achille – G. Sicuro – A. Sportiello, Random assignment problems on $2d$ manifolds. *J. Stat. Phys.* **183** (2021), no. 2, article no. 34. Zbl 1470.60035 MR 4257835

[10] S. G. Bobkov – M. Ledoux, A simple Fourier analytic proof of the AKT optimal matching theorem. *Ann. Appl. Probab.* **31** (2021), no. 6, 2567–2584. Zbl 1482.60011 MR 4350968

[11] S. Caracciolo – C. Lucibello – G. Parisi – G. Sicuro, Scaling hypothesis for the Euclidean bipartite matching problem. *Phys. Rev. E* **90** (2014), no. 1, article no. 012118.

[12] N. Clozeau – F. Mattesini, Annealed quantitative estimates for the quadratic 2D-discrete random matching problem. *Probab. Theory Related Fields* **190** (2024), no. 1-2, 485–541. Zbl 07924626 MR 4797374

[13] M. G. Crandall – H. Ishii – P.-L. Lions, User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)* **27** (1992), no. 1, 1–67. Zbl 0755.35015 MR 1118699

[14] M. Goldman – M. Huesmann, A fluctuation result for the displacement in the optimal matching problem. *Ann. Probab.* **50** (2022), no. 4, 1446–1477. Zbl 1491.35137 MR 4420424

[15] M. Goldman – M. Huesmann – F. Otto, Quantitative linearization results for the Monge-Ampère equation. *Comm. Pure Appl. Math.* **74** (2021), no. 12, 2483–2560. Zbl 1480.35082 MR 4373162

[16] M. Goldman – M. Huesmann – F. Otto, Almost sharp rates of convergence for the average cost and displacement in the optimal matching problem. 2023, arXiv:2312.07995v1.

[17] M. Goldman – D. Trevisan, Convergence of asymptotic costs for random Euclidean matching problems. *Probab. Math. Phys.* **2** (2021), no. 2, 341–362. Zbl 1491.35138 MR 4408015

[18] M. Huesmann – F. Mattesini – F. Otto, There is no stationary cyclically monotone Poisson matching in 2d. *Probab. Theory Related Fields* **187** (2023), no. 3-4, 629–656. Zbl 1527.60016 MR 4664582

[19] M. Huesmann – F. Mattesini – F. Otto, There is no stationary p-cyclically monotone Poisson matching in 2d. *Electron. J. Probab.* **29** (2024), article no. 1. Zbl 1527.60016 MR 4779877

[20] V. Julin – P. Juutinen, A new proof for the equivalence of weak and viscosity solutions for the p-Laplace equation. *Comm. Partial Differential Equations* **37** (2012), no. 5, 934–946. Zbl 1260.35069 MR 2915869

[21] P. Juutinen – P. Lindqvist – J. J. Manfredi, On the equivalence of viscosity solutions and weak solutions for a quasi-linear equation. *SIAM J. Math. Anal.* **33** (2001), no. 3, 699–717. Zbl 0997.35022 MR 1871417

[22] L. Koch, Geometric linearisation for optimal transport with strongly p-convex cost. *Calc. Var. Partial Differential Equations* **63** (2024), no. 4, article no. 87. Zbl 1540.49055 MR 4728221

[23] L. Koch – F. Otto, Lecture notes on the harmonic approximation to quadratic optimal transport. 2023, arXiv:2303.14462v1.

[24] M. Ledoux, On optimal matching of Gaussian samples. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **457** (2017), 226–264. Zbl 1478.60080 MR 3723584

[25] M. Ledoux, Optimal matching of random samples and rates of convergence of empirical measures. In *Mathematics going forward—collected mathematical brushstrokes*, pp. 615–627, Lecture Notes in Math. 2313, Springer, Cham, 2023. Zbl 1536.60015 MR 4627987

[26] S.-i. Ohta, On the curvature and heat flow on Hamiltonian systems. *Anal. Geom. Metr. Spaces* **2** (2014), no. 1, 81–114. Zbl 1295.53029 MR 3208069

[27] R. Peyre, Comparison between W_2 distance and $\dot{\mathrm{H}}^{-1}$ norm, and localization of Wasserstein distance. *ESAIM Control Optim. Calc. Var.* **24** (2018), no. 4, 1489–1501. Zbl 1415.49031 MR 3922440

[28] H. P. Rosenthal, On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.* **8** (1970), 273–303. Zbl 0213.19303 MR 0271721

[29] E. V. Teixeira, Regularity for quasilinear equations on degenerate singular sets. *Math. Ann.* **358** (2014), no. 1-2, 241–256. Zbl 1286.35119 MR 3157997

Received 16 May 2024

Luigi Ambrosio
Scuola Normale Superiore
Piazza dei Cavalieri 7, 56126 Pisa, Italy
luigi.ambrosio@sns.it

Federico Vitillaro
Scuola Normale Superiore
Piazza dei Cavalieri 7, 56126 Pisa, Italy
federico.vitillaro@sns.it

Dario Trevisan
Dipartimento di Matematica, Università di Pisa
Largo Bruno Pontecorvo 5, 56127 Pisa, Italy
dario.trevisan@unipi.it