On optimal matching of Gaussian samples II

Michel Ledoux

University of Toulouse, France

Abstract

This note is a short addition to the paper [4]. Given $X_1, \ldots, X_n$ independent random variables with common distribution the standard Gaussian measure μ on $\mathbb{R}^2$, and $\mu_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ the associated empirical measure, it holds true that

$$\mathbb{E}\big(\mathrm{W}_2^2(\mu_n, \mu)\big) \approx \frac{(\log n)^2}{n}$$

where W_2 is the quadratic Kantorovich metric. The upper bound has been obtained in [4] by the pde and mass transportation approach developed by L. Ambrosio, F. Stra and D. Trevisan in a compact setting, and the lower bound was achieved recently by M. Talagrand using a scaling argument and ideas from the original Ajtai-Komlós-Tusnády theorem.

We note here that the pde and mass transportation approach may actually also be used to reach the lower bound. In addition, we sharpen the limit obtained by L. Ambrosio, F. Stra and D. Trevisan on a 2-dimensional compact Riemannian manifold in the spirit of the conjecture of S. Caracciolo, C. Lucibello, G. Parisi and G. Sicuro.

The context and methodology of this note is based on the investigation [1] by L. Ambrosio, F. Stra and D. Trevisan. The framework and the notation are taken from the reference [4], and we do not reproduce them here.

The main purpose of the note (achieved in Section 2) is to provide an alternate proof based on the pde method of [1] of the lower bound

$$\mathbb{E}\big(\mathrm{W}_2^2(\mu_n, \mu)\big) \geq c\,\frac{(\log n)^2}{n} \tag{1}$$

which has been established recently in [6]. We will actually deal simultaneously with the one and two-dimensional cases, and also recover in dimension one the lower bound

$$\mathbb{E}\big(\mathrm{W}_2^2(\mu_n, \mu)\big) \geq c\,\frac{\log\log n}{n} \tag{2}$$

proved in [2] (by specific one-dimensional tools).

To this task, we develop first a two-sided bound on the Kantorovich metric W_2 in terms of Sobolev norms. We then address in Section 2 the lower bound for the Gaussian sample. In Section 3, we refine the limit of [1] on a 2-dimensional compact Riemannian manifold in the direction of the conjecture of [3].

1 A two-sided bound on W_2

The context here is the one of a weighted Riemannian manifold M with weigthed probability measure μ, under the curvature condition $CD(K, \infty)$ for some $K \in \mathbb{R}$ as described in [4]. The following two statements provide general bounds on the Kantorovich metric W_2 in terms of a Sobolev norm. The lower bound is taken from [1].

Theorem 1. *Let $d\nu = f d\mu$ and $f = 1 + g$, and let $0 < c \leq 1$. If $g \geq -c$, then*

$$\mathrm{W}_2^2(\nu, \mu) \leq \frac{4}{c^2} \left[1 - \sqrt{1-c}\,\right]^2 \int_M g(-\mathrm{L})^{-1} g \, d\mu \tag{3}$$

(where g is assumed to belong to the suitable domain so that the left-hand side makes sense).

Theorem 2. *Assume that the curvature condition $CD(0, \infty)$ holds. Let $d\nu = f d\mu$ and $f = 1+g$. Then, whenever g and $h : M \to \mathbb{R}$ belong to the suitable domain and h is such that $\int_M h d\mu = 0$ and $h \leq c$ uniformly for some $c > 0$,*

$$\mathrm{W}_2^2(\nu, \mu) \geq 2 \int_M g(-\mathrm{L})^{-1} h \, d\mu - \frac{e^c - 1}{c} \int_M h(-\mathrm{L})^{-1} h \, d\mu. \tag{4}$$

In particular, if $g \leq c$,

$$\mathrm{W}_2^2(\nu, \mu) \geq \left(2 - \frac{e^c - 1}{c}\right) \int_M g(-\mathrm{L})^{-1} g \, d\mu. \tag{5}$$

Recall that by integration by parts

$$\int_M g(-\mathrm{L})^{-1} g \, d\mu = \int_M \left|\nabla((-\mathrm{L})^{-1} g)\right|^2 d\mu$$

which is the Sobolev norm alluded to above (see [4]). Note also that as $c \to 0$,

$$\frac{4}{c^2} \left[1 - \sqrt{1-c}\,\right]^2 \sim 1 + \frac{c}{2} \qquad \text{and} \qquad 2 - \frac{e^c - 1}{c} \sim 1 - \frac{c}{2}$$

so that the bounds (3) and (5) are sharp in this regime.

Proof of Theorem 1. It is shown in [4] that for every (smooth) increasing $\theta : [0,1] \to [0,1]$ with $\theta(0) = 0$, $\theta(1) = 1$,

$$\mathrm{W}_2^2(\nu, \mu) \leq \int_M \left|\nabla((-\mathrm{L})^{-1} g)\right|^2 \int_0^1 \frac{\theta'(s)^2}{1 + \theta(s) g} \, ds \, d\mu.$$

Using that $g \geq -c$,

$$\mathrm{W}_2^2(\nu,\mu) \,\leq\, \int_0^1 \frac{\theta'(s)^2}{1-\theta(s)c}\, ds \int_M \big|\nabla((-\mathrm{L})^{-1}g)\big|^2 d\mu.$$

The claim (3) follows from the (optimal) choice

$$\theta(s) \,=\, \frac{1-\sqrt{1-c}}{c}\Big(2s - \big[1-\sqrt{1-c}\,\big]s^2\Big), \quad s \in [0,1].$$

□

Proof of Theorem 2. As announced, we follow [1]. By the Kantorovich dual description, for any (smooth) $\varphi : M \to \mathbb{R}$,

$$\begin{aligned} \frac{1}{2}\,\mathrm{W}_2^2(\nu,\mu) \,&\geq\, \int_M \varphi\, f d\mu - \int_M \widehat{Q}_1\varphi\, d\mu \\ &=\, \int_M \varphi\, g d\mu - \bigg(\int_M \widehat{Q}_1\varphi\, d\mu - \int_M \varphi\, d\mu\bigg) \end{aligned}$$

where $\widehat{Q}_1$ is the supremum convolution

$$\widehat{Q}_1\varphi(x) \,=\, \sup_{y\in M}\Big[\varphi(y) - \frac{1}{2}\, d(x,y)^2\Big].$$

Choose then $\varphi = (-\mathrm{L})^{-1}h$. Now

$$\int_M \widehat{Q}_1\varphi\, d\mu - \int_M \varphi\, d\mu \,=\, \frac{1}{2}\int_0^1\int_M |\nabla\widehat{Q}_s\varphi|^2 d\mu\, ds.$$

It is shown in [1] that since $-\mathrm{L}\varphi = h \leq c$ uniformly, under a $CD(0,\infty)$ curvature condition,

$$\int_M |\nabla\widehat{Q}_s\varphi|^2 d\mu \,\leq\, e^{cs}\int_M |\nabla\varphi|^2 d\mu, \quad 0 \leq s \leq 1.$$

Therefore

$$\int_M \widehat{Q}_1\varphi\, d\mu - \int_M \varphi\, d\mu \,\leq\, \frac{e^c - 1}{c}\int_M |\nabla\varphi|^2 d\mu.$$

Since

$$\int_M |\nabla\varphi|^2 d\mu \,=\, \int_M \varphi(-\mathrm{L}\varphi) d\mu \,=\, \int_{\mathbb{R}^2} h(-\mathrm{L})^{-1}h\, d\mu,$$

the assertion (4) follows. □

We note from [1] that Theorem 2 admits a version under a curvature condition $CD(K,\infty)$ for some $K \in \mathbb{R}$, which on a compact manifold essentially leads to similar conclusions. We freely use this comment in Section 3 below.

By elementary algebra, the inequality (4) of Theorem 2 immediately leads to

$$
\begin{aligned}
&\mathrm{W}_2^2(\nu,\mu) \\
&\geq 2\int_M g(-\mathrm{L})^{-1}g\,d\mu - 2\int_M (g-h)(-\mathrm{L})^{-1}g\,d\mu \\
&\quad - \frac{e^c-1}{c}\bigg(\int_M g(-\mathrm{L})^{-1}g\,d\mu + \int_M (g-h)(-\mathrm{L})^{-1}(g-h)\,d\mu - 2\int_M (g-h)(-\mathrm{L})^{-1}g\,d\mu\bigg)
\end{aligned}
$$

so that, if $c \leq 1$ for example,

$$
\begin{aligned}
\mathrm{W}_2^2(\nu,\mu) \,\geq\, & \Big(2 - \frac{e^c-1}{c}\Big)\int_M g(-\mathrm{L})^{-1}g\,d\mu \\
& - 2\int_M (g-h)(-\mathrm{L})^{-1}(g-h)\,d\mu - 6\bigg|\int_M (g-h)(-\mathrm{L})^{-1}g\,d\mu\bigg|.
\end{aligned}
\tag{6}
$$

The correction terms in (6) are typically handled with a spectral gap hypothesis, with constant $\lambda > 0$. Indeed, since $(-\mathrm{L})^{-1} = \int_0^\infty P_s ds$, by the spectral gap inequality

$$
\int_M (g-h)(-\mathrm{L})^{-1}(g-h)d\mu \,=\, 2\int_0^\infty \big\|P_s(g-h)\big\|_2^2\,ds \,\leq\, \frac{1}{\lambda}\,\|g-h\|_2^2.
$$

In the same way,

$$
\bigg|\int_M (g-h)(-\mathrm{L})^{-1}g d\mu\bigg| \,\leq\, \frac{1}{\sqrt{\lambda}}\,\|g\|_2\|g-h\|_2.
$$

so that we may reformulate (6) as

$$
\mathrm{W}_2^2(\nu,\mu) \,\geq\, \Big(2 - \frac{e^c-1}{c}\Big)\int_M g(-\mathrm{L})^{-1}g\,d\mu - \frac{2}{\lambda}\,\|g-h\|_2^2 - \frac{6}{\sqrt{\lambda}}\,\|g\|_2\|g-h\|_2. \tag{7}
$$

2 Lower bound for the Gaussian sample

In this section, we deal with the lower bounds (1) and (2) in the Gaussian case, simultaneously in dimensions $d = 1$ and $d = 2$. We make use of the general Theorem 2.

The first step is the Kantorovich contraction property under a $CD(0,\infty)$ curvature condition, which holds in Gauss space for the Mehler kernel $p_t(x,y)$,

$$
\mathrm{W}_2^2(\mu_n,\mu) \,\geq\, \mathrm{W}_2^2(\mu_n^t,\mu) \tag{8}
$$

where we recall that $d\mu_n^t = f d\mu$, $f(y) = 1 + g(y)$, $g = g(y) = \frac{1}{n}\sum_{i=1}^n [p_t(X_i,y) - 1]$, $t > 0$.

For $R > 0$, recall $d\mu^R = \frac{1}{\mu(B_R)}\,\mathbb{1}_{B_R} d\mu$ where $B_R = B(0,R)$ is the Euclidean ball centered at 0 with radius R in $\mathbb{R}^d$, and the independent random variables X_i^R, $i = 1,\ldots,n$, with common distribution μ^R defined by

$$
X_i^R \,=\, \begin{cases} X_i & \text{if} \quad X_i \in B_R, \\ Z_i & \text{if} \quad X_i \notin B_R, \end{cases}
$$

where $Z_1, \ldots, Z_n$ are independent with distribution μ^R, independent of the X_i's. Let

$$\tilde{g} \,=\, \tilde{g}(y) \,=\, \frac{1}{n} \sum_{i=1}^{n} \big[p_t(X_i^R, y) - \mathbb{E}\big(p_t(X_i^R, y)\big)\big],$$

and, for $c > 0$,

$$\tilde{g}_c \,=\, (\tilde{g} \wedge c) \vee (-c) - \int_{\mathbb{R}^d} \big[(\tilde{g} \wedge c) \vee (-c)\big] d\mu$$

(so that $|\tilde{g}_c| \leq 2c$ and $\int_{\mathbb{R}^d} \tilde{g}_c d\mu = 0$).

Now, in (4) of Theorem 2, we choose $h = \tilde{g}_c$ and perform some minor modifications on (6) and (7). It holds that

$$\begin{aligned}\int_{\mathbb{R}^d} \tilde{g}_c (-\mathrm{L})^{-1} \tilde{g}_c d\mu \,=\, & \int_{\mathbb{R}^d} \tilde{g} (-\mathrm{L})^{-1} \tilde{g} d\mu + \int_{\mathbb{R}^d} (\tilde{g} - \tilde{g}_c)(-\mathrm{L})^{-1}(\tilde{g} - \tilde{g}_c) d\mu \\ & - 2 \int_{\mathbb{R}^d} (\tilde{g} - \tilde{g}_c)(-\mathrm{L})^{-1} \tilde{g} d\mu.\end{aligned}$$

Therefore, with $c \leq \frac{1}{2}$ for example,

$$\begin{aligned}\mathrm{W}_2^2(\mu_t^n, \mu) \,\geq\, & 2 \int_{\mathbb{R}^d} \tilde{g}(-\mathrm{L})^{-1} g d\mu - \frac{e^{2c} - 1}{2c} \int_{\mathbb{R}^d} \tilde{g}(-\mathrm{L})^{-1} \tilde{g} d\mu \\ & - 2 \int_{\mathbb{R}^d} (\tilde{g} - \tilde{g}_c)(-\mathrm{L})^{-1} g d\mu \\ & - 2 \int_{\mathbb{R}^d} (\tilde{g} - \tilde{g}_c)(-\mathrm{L})^{-1}(\tilde{g} - \tilde{g}_c) d\mu - 4 \bigg| \int_{\mathbb{R}^d} (\tilde{g} - \tilde{g}_c)(-\mathrm{L})^{-1} \tilde{g} d\mu \bigg|.\end{aligned} \tag{9}$$

We handle the error terms in (9) by the spectral gap inequality. Namely

$$\int_{\mathbb{R}^d} (\tilde{g} - \tilde{g}_c)(-\mathrm{L})^{-1}(\tilde{g} - \tilde{g}_c) d\mu \,=\, 2 \int_0^\infty \big\| P_s(\tilde{g} - \tilde{g}_c) \big\|_2^2 \, ds \,\leq\, \|\tilde{g} - \tilde{g}_c\|_2^2.$$

In the same way,

$$\int_{\mathbb{R}^d} (\tilde{g} - \tilde{g}_c)(-\mathrm{L})^{-1} g d\mu \,\leq\, \|g\|_2 \|\tilde{g} - \tilde{g}_c\|_2.$$

and

$$\bigg| \int_{\mathbb{R}^d} (\tilde{g} - \tilde{g}_c)(-\mathrm{L})^{-1} \tilde{g} d\mu \bigg| \,\leq\, \|\tilde{g}\|_2 \|\tilde{g} - \tilde{g}_c\|_2.$$

Putting things together, and since

$$|\tilde{g} - \tilde{g}_c| \,\leq\, |\tilde{g}| \mathbb{1}_{\{|\tilde{g}| \geq c\}} + \int_{\mathbb{R}^2} |\tilde{g}| \mathbb{1}_{\{|\tilde{g}| \geq c\}} d\mu,$$

we deduce from (8) and (9) that for every $0 < c \leq \frac{1}{2}$,

$$\begin{aligned}\mathrm{W}_2^2(\mu^n, \mu) \,\geq\, \mathrm{W}_2^2(\mu_t^n, \mu) \,\geq\, & 2 \int_{\mathbb{R}^d} \tilde{g}(-\mathrm{L})^{-1} g d\mu - \frac{e^{2c} - 1}{2c} \int_{\mathbb{R}^d} \tilde{g}(-\mathrm{L})^{-1} \tilde{g} d\mu \\ & - 8 \int_{\{|\tilde{g}| \geq c\}} |\tilde{g}|^2 d\mu - 8 \big(\|g\|_2 + \|\tilde{g}\|_2 \big) \bigg(\int_{\{|\tilde{g}| \geq c\}} |\tilde{g}|^2 d\mu \bigg)^{1/2}.\end{aligned} \tag{10}$$

Next, we integrate over the samples $X_1, \ldots, X_n$ and $X_1^R, \ldots, X_n^R$ the first two terms on the right-hand side of (10). If we recall the definitions of g and $\tilde{g}$, by independence and identical distribution,

$$\mathbb{E}\bigg(\int_{\mathbb{R}^d} \tilde{g}(-\mathrm{L})^{-1} g d\mu \bigg) = \frac{1}{n} \int_t^\infty \int_{\mathbb{R}^d} \mathbb{E}\Big(\big[p_t(X_1^R, y) - \mathbb{E}\big(p_t(X_1^R, y)\big) \big] p_s(X_1, y) \Big) d\mu(y) ds.$$

By definition of X_1^R,

$$\begin{aligned}
\mathbb{E}\Big(\big[p_t(X_1^R, y) - \mathbb{E}\big(p_t(X_1^R, y)\big) \big] p_s(X_1, y) \Big) \\
&= \mathbb{E}\Big(\mathbb{1}_{\{X_1 \in B_R\}} \big[p_t(X_1, y) - \mathbb{E}\big(p_t(X_1^R, y)\big) \big] p_s(X_1, y) \Big) \\
&\quad + \mathbb{E}\Big(\mathbb{1}_{\{X_1 \notin B_R\}} \big[p_t(Z_1, y) - \mathbb{E}\big(p_t(X_1^R, y)\big) \big] p_s(X_1, y) \Big) \\
&= \mathbb{E}\Big(\mathbb{1}_{\{X_1 \in B_R\}} \big[p_t(X_1, y) - \mathbb{E}\big(p_t(X_1^R, y)\big) \big] p_s(X_1, y) \Big)
\end{aligned}$$

since Z_1 is independent of X_1 and with the same law as X_1^R. Hence, after integration in $d\mu(y)$ and the semigroup property,

$$\begin{aligned}
\mathbb{E}\bigg(\int_{\mathbb{R}^d} \tilde{g}(-\mathrm{L})^{-1} g d\mu \bigg) \\
&= \frac{\mu(B)}{n} \int_t^\infty \bigg[\int_{\mathbb{R}^d} p_{t+s}(x, x) d\mu^R(x) - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} p_{t+s}(x, x') d\mu^R(x) d\mu^R(x') \bigg] ds \\
&= \frac{\mu(B)}{n} \int_{2t}^\infty \bigg[\int_{\mathbb{R}^d} p_s(x, x) d\mu^R(x) - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} p_s(x, x') d\mu^R(x) d\mu^R(x') \bigg] ds
\end{aligned}$$

In the same way,

$$\begin{aligned}
\mathbb{E}\bigg(\int_{\mathbb{R}^d} \tilde{g}(-\mathrm{L})^{-1} \tilde{g} d\mu \bigg) \\
&= \frac{1}{n} \int_t^\infty \int_{\mathbb{R}^d} \mathbb{E}\Big(\big[p_t(X_1^R, y) - \mathbb{E}\big(p_t(X_1^R, y)\big) \big] p_s(X_1^R, y) \Big) d\mu(y) ds \\
&= \frac{1}{n} \int_{2t}^\infty \bigg[\int_{\mathbb{R}^d} p_s(x, x) d\mu^R(x) - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} p_s(x, x') d\mu^R(x) d\mu^R(x') \bigg] ds.
\end{aligned}$$

As a consequence, given $0 < \eta < 1$, if $c > 0$ is small enough and if $\mu(B)$ is close to 1 (in terms of η),

$$\begin{aligned}
\mathbb{E}\bigg(2 \int_{\mathbb{R}^d} \tilde{g}(-\mathrm{L})^{-1} g d\mu - \frac{e^{2c} - 1}{2c} \int_{\mathbb{R}^d} \tilde{g}(-\mathrm{L})^{-1} \tilde{g} d\mu \bigg) \\
&\geq (1 - \eta) \bigg(\frac{1}{n} \int_{2t}^\infty \bigg[\int_{\mathbb{R}^d} p_s(x, x) d\mu^R(x) - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} p_s(x, x') d\mu^R(x) d\mu^R(x') \bigg] ds \bigg).
\end{aligned}$$

Also, by the spectral gap inequality,

$$
\begin{aligned}
\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} p_s(x,x')d\mu^R(x)d\mu^R(x') &= \frac{1}{\mu(B_R)^2}\int_{\mathbb{R}^d} \mathbb{1}_{B_R} P_s(\mathbb{1}_{B_R})d\mu \\
&= 1 + \frac{1}{\mu(B_R)^2}\int_{\mathbb{R}^d} \mathbb{1}_{B_R} P_s\big(\mathbb{1}_{B_R} - \mu(B_R)\big)d\mu \\
&\leq 1 + \frac{1-\mu(B_R)}{\mu(B_R)}\, e^{-s} \\
&\leq 1 + 2\, e^{-s}
\end{aligned}
$$

provided that $\mu(B_R) \geq \frac{1}{2}$. As a conclusion at this stage,

$$
\begin{aligned}
\mathbb{E}\big(\mathrm{W}_2^2(\mu^n,\mu)\big) \;\geq\; &(1-\eta)\,\frac{1}{n}\int_{2t}^{\infty}\int_{\mathbb{R}^d}\big[p_s(x,x)-1\big]d\mu^R(x)ds - \frac{2}{n} \\
&- 8\,\mathbb{E}\bigg(\int_{\{|\tilde g|\geq c\}} |\tilde g|^2 d\mu\bigg) - 8\,\mathbb{E}\bigg(\big(\|g\|_2 + \|\tilde g\|_2\big)\bigg(\int_{\{|\tilde g|\geq c\}} |\tilde g|^2 d\mu\bigg)^{1/2}\bigg).
\end{aligned}
\tag{11}
$$

The final part of the proof will be to take care of the correction terms on the right-hand side of the preceding (11). To this task, we develop some (crude) bounds on the Mehler kernel. Recall the Mehler kernel

$$
p_t(x,y) \;=\; \frac{1}{(1-a^2)^{d/2}}\,\exp\bigg(-\frac{a^2}{2(1-a^2)}\Big[|x|^2+|y|^2-\frac{2}{a}\,x\cdot y\Big]\bigg)
$$

where $a = e^{-t}$, $t > 0$, $x, y \in \mathbb{R}^d$. Consider for each $y \in \mathbb{R}^d$ and $q \geq 1$,

$$
\int_{B_R} p_t(x,y)^q d\mu(x)
$$

After translation and a change of variable,

$$
\int_{B_R} p_t(x,y)^q d\mu(x) \;=\; \frac{1}{(1-a^2)^{qd/2}}\, e^{\frac{q(q-1)a^2}{2(1+(q-1)a^2)}\,|y|^2}\int_{B(-\kappa y,R)} \exp\bigg(-\frac{1+(q-1)a^2}{2(1-a^2)}\,|x|^2\bigg)\frac{dx}{(2\pi)^{d/2}}
$$

where $\kappa = \frac{qa}{1+(q-1)a^2}$.

Note that

$$
\frac{q(q-1)a^2}{2(1+(q-1)a^2)} \;\leq\; \frac{q}{2}
$$

and that $\kappa \geq \frac{1}{2}$, at least provided that a is close to one which we may assume. Then, if $|y| \geq 4R$ and $x \in B(-\kappa y, R)$, we have $|x| \geq \frac{|y|}{4}$. Hence, whenever $|y| \geq 4R$,

$$
\int_{B_R} p_t(x,y)^q d\mu(x) \;\leq\; \frac{1}{(1-a^2)^{(q-1)d/2}}\, e^{-\left(\frac{1}{32(1-a^2)}-\frac{q}{2}\right)|y|^2}.
$$

Otherwise, that is when $|y| \leq 4R$,

$$
\int_{B_R} p_t(x,y)^q d\mu(x) \;\leq\; \frac{1}{(1-a^2)^{(q-1)d/2}}\, e^{\frac{q}{2}|y|^2}.
$$

Recall

$$\tilde{g} \,=\, \tilde{g}(y) \,=\, \frac{1}{n}\sum_{i=1}^{n}\big[p_t(X_i^R,y) - \mathbb{E}\big(p_t(X_i^R,y)\big)\big].$$

By Rosenthal's inequality for sums of independent (identically distributed) random variables [5], for any $q \geq 2$ there exists $C_q > 0$ only depending on q such that

$$\begin{aligned}
\mathbb{E}\big(|\tilde{g}(y)|^q\big) \,&\leq\, C_q\bigg(\frac{1}{n^{q-1}}\,\mathbb{E}\big(p_t(X_1^R,y)^q\big) + \frac{1}{n^{q/2}}\,\Big[\mathbb{E}\big(p_t(X_1^R,y)^2\big)\Big]^{q/2}\bigg)\\
&\leq\, 2^{\frac{q}{2}}C_q\bigg(\frac{1}{n^{q-1}}\,\int_{B_R} p_t(x,y)^q d\mu(x) + \frac{1}{n^{q/2}}\,\bigg[\int_{B_R} p_t(x,y)^2 d\mu(x)\bigg]^{q/2}\bigg)
\end{aligned}$$

where we assumed that $\mu(B_R) \geq \frac{1}{2}$.

In the following $q \geq 2$ is fixed. Then $t > 0$ may be chosen small enough (in terms of q but independently of n) such that $\kappa \geq \frac{1}{2}$ and $\frac{1}{32(1-a^2)} - \frac{q}{2} \geq 0$ (for example). By the previous step,

$$\int_{\mathbb{R}^d}\bigg(\int_{B_R} p_t(x,y)^q d\mu(x)\bigg)d\mu(y) \,\leq\, \frac{1}{(1-a^2)^{(q-1)d/2}}\,\big(1+e^{8qR^2}\big).$$

In the same way,

$$\int_{\mathbb{R}^d}\bigg(\int_{B_R} p_t(x,y)^2 d\mu(x)\bigg)^{q/2} d\mu(y) \,\leq\, \frac{1}{(1-a^2)^{qd/4}}\,\big(1+e^{4qR^2}\big).$$

For simplicity (in order not to carry the two preceding expressions with $q-1$ and $\frac{q}{2}$), assume in the following that $(1-a^2)^{d/2}n \geq 1$. Therefore, using that $q-1 \geq \frac{q}{2}$,

$$\int_{\mathbb{R}^2}\mathbb{E}\big(|\tilde{g}(y)|^q\big)d\mu(y) \,\leq\, \frac{2^q C_q}{[(1-a^2)^{d/2}n]^{q/2}}\,\big(1+e^{8qR^2}\big). \tag{12}$$

We use the preceding bounds to control the error term

$$\mathbb{E}\bigg(\int_{\{|\tilde{g}|\geq c\}}|\tilde{g}|^2 d\mu\bigg) + \mathbb{E}\bigg(\big(\|g\|_2 + \|\tilde{g}\|_2\big)\bigg(\int_{\{|\tilde{g}|\geq c\}}|\tilde{g}|^2 d\mu\bigg)^{1/2}\bigg)$$

of (11). By repeated use of the Young and Hölder inequalities, the latter is bounded from above for any $\delta > 0$ and $\alpha > 1$ by

$$\delta\,\Big[\mathbb{E}\big(\|g\|_2^2\big) + \mathbb{E}\big(\|\tilde{g}\|_2^2\big)\Big] + \frac{1+2\delta}{2c^{2(\alpha-1)}\delta}\,\int_{\mathbb{R}^d}\mathbb{E}\big(|\tilde{g}(y)|^{2\alpha}d\mu(y).$$

Since $p_t(x,x) = \frac{1}{1-a^2}\,e^{\frac{a}{1+a}|x|^2}$, again with $\mu(B_R) \geq \frac{1}{2}$,

$$\begin{aligned}
\mathbb{E}\big(\|\tilde{g}\|_2^2\big) \,&=\, \frac{1}{n}\int_{\mathbb{R}^d}\Big[\mathbb{E}\big(p_t(X_1^R,y)^2\big) - \mathbb{E}\big(p_t(X_1^R,y)\big)^2\Big]d\mu(y)\\
&\leq\, \frac{1}{n}\int_{\mathbb{R}^d} p_t(x,x)d\mu^R(x)\\
&\leq\, \frac{1}{n\mu(B_R)(1-a)^d}\\
&\leq\, \frac{2}{n(1-a)^d}\,.
\end{aligned}$$

Similarly

$$\mathbb{E}\big(\|g\|_2^2\big) \,\leq\, \frac{1}{n}\int_{\mathbb{R}^2} p_t(x,x)d\mu(x) \,\leq\, \frac{1}{n(1-a)^d}\,.$$

On the other hand, (12) with $q = 2\alpha$ yields

$$\int_{\mathbb{R}^d} \mathbb{E}\big(\big|\tilde{g}(y)\big|^{2\alpha}\big)d\mu(y) \,\leq\, \frac{4^\alpha C_{2\alpha}}{[(1-a^2)^{d/2}n]^\alpha}\big(1+e^{16\alpha R^2}\big).$$

Hence, for any $0 < \delta \leq 1$,

$$\begin{aligned}\mathbb{E}\bigg(\int_{\{|\tilde{g}|\geq c\}} |\tilde{g}|^2 d\mu\bigg) + \mathbb{E}\bigg(\big(\|g\|_2 + \|\tilde{g}\|_2\big)\bigg(\int_{\{|\tilde{g}|\geq c\}} |\tilde{g}|^2 d\mu\bigg)^{1/2}\bigg)& \\ \leq\, \frac{3\delta}{n(1-a)^d} + \frac{4^{\alpha+1}C_{2\alpha}}{c^{2(\alpha-1)}\delta[(1-a^2)^{d/2}n]^\alpha}\big(1+e^{16\alpha R^2}\big).&\end{aligned} \tag{13}$$

In this last step, we fix the various parameters involved in the previous analysis. Basically, $t \sim \frac{1}{n^\varepsilon}$ and $R \sim \varepsilon\sqrt{\log n}$ for some small $\varepsilon > 0$, and $\alpha > 1$ is chosen large enough. Take for example $t = \frac{1}{n^{1/d}}$ and $R^2 = \frac{1}{64}\log n$. Then, for n large enough, the necessary conditions on $a = e^{-t}$ or $\mu(B_R)$ are fulfilled. After some details, the choice of $\delta = \frac{1}{n}$ and $\alpha = 8$ in (13) yields that

$$\mathbb{E}\bigg(\int_{\{|\tilde{g}|\geq c\}} |\tilde{g}|^2 d\mu\bigg) + \mathbb{E}\bigg(\big(\|g\|_2 + \|\tilde{g}\|_2\big)\bigg(\int_{\{|\tilde{g}|\geq c\}} |\tilde{g}|^2 d\mu\bigg)^{1/2}\bigg) \,=\, O\Big(\frac{1}{n}\Big).$$

Collecting this bound with (11) yields

$$\mathbb{E}\big(\mathrm{W}_2^2(\mu^n,\mu)\big) \,\geq\, (1-\eta)\,\frac{1}{n}\int_{2t}^\infty \int_{\mathbb{R}^d} \big[p_s(x,x) - 1\big]d\mu^R(x)ds - O\Big(\frac{1}{n}\Big).$$

From the analysis of the upper bound (cf. [4]), it is known that as $t << \frac{1}{R^2}$, for some $c > 0$,

$$\int_{2t}^\infty \int_{\mathbb{R}^2} \big[p_s(x,x) - 1\big]d\mu^R(x)ds \,\geq\, cR^2 \log\Big(\frac{1}{t}\Big)$$

when $d = 2$ and

$$\int_{2t}^\infty \int_{\mathbb{R}} \big[p_s(x,x) - 1\big]d\mu^R(x)ds \,\geq\, c\log(R^2)$$

when $d = 1$. For the preceding choices of t and R, this establishes the claims (1) and (2).

3 On the limit in the compact case

It has been shown in [1] that when μ is the normalized Riemannian measure on a compact (smooth) 2-dimensional Riemannian manifold M without boundary, then

$$\lim_{n\to\infty} \frac{n}{\log n}\,\mathbb{E}\big(\mathrm{W}_2^2(\mu_n,\mu)\big) \,=\, \frac{\mathrm{vol}(M)}{4\pi}\,, \tag{14}$$

where here $\mu_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$ with the X_i's, $i = 1, \ldots, n$, independent and distributed as μ. Since the statement is invariant under rescaling of the measure, assume in the following that $\mathrm{vol}(M) = 1$. The further conjecture from [3] is that

$$\lim_{n\to\infty} \bigg(\frac{n}{\log n}\, \mathbb{E}\big(\mathrm{W}_2^2(\mu_n,\mu)\big) - \frac{1}{4\pi}\bigg) \log n \,=\, \xi \in \mathbb{R}. \tag{15}$$

In this section, we will show that the arguments of [1] might be somewhat tightened so to yield

$$\liminf_{n\to\infty} \bigg(\frac{n}{\log n}\, \mathbb{E}\big(\mathrm{W}_2^2(\mu_n,\mu)\big) - \frac{1}{4\pi}\bigg) \frac{\log n}{\log\log n} \,>\, -\infty, \tag{16}$$

and

$$\limsup_{n\to\infty} \bigg(\frac{n}{\log n}\, \mathbb{E}\big(\mathrm{W}_2^2(\mu_n,\mu)\big) - \frac{1}{4\pi}\bigg) \frac{(\log n)^{\frac{1}{4}}}{\log\log n} \,<\, \infty. \tag{17}$$

We start with the proof of the liminf (16) which follows the developments of Section 1. For simplicity we deal with a manifold with non-negative curvature, but the arguments may be modified to cover the general case (cf. [1]). Recall first the contraction property under a $CD(0,\infty)$ curvature condition, for any $t > 0$,

$$\mathrm{W}_2^2(\mu_n,\mu) \,\geq\, \mathrm{W}_2^2(\mu_n^t,\mu)$$

where $d\mu_n^t = f d\mu$, $f(y) = 1 + g(y)$, $g = g(y) = \frac{1}{n}\sum_{i=1}^n [p_t(X_i,y) - 1]$, $t > 0$. We next apply Theorem 2 and (7) to

$$g \,=\, g(y) \,=\, \frac{1}{n}\sum_{i=1}^n \big[p_t(X_i,y) - 1\big].$$

and

$$h \,=\, g_c \,=\, (g\wedge c)\vee(-c) - \int_{\mathbb{R}^d} \big[(g\wedge c)\vee(-c)\big] d\mu$$

with $0 < c \leq \frac{1}{2}$. Therefore, after integration along the sample $X_1, \ldots, X_n$,

$$\begin{aligned}\mathbb{E}\big(\mathrm{W}_2^2(\mu_n,\mu)\big) \,\geq\, &\Big(2 - \frac{e^{2c}-1}{2c}\Big)\, \mathbb{E}\bigg(\int_M g(-\mathrm{L})^{-1} g\, d\mu\bigg) \\ &- \frac{4}{\lambda}\, \mathbb{E}\bigg(\int_{\{|g|\geq c\}} g^2 d\mu\bigg) - \frac{12}{\sqrt{\lambda}}\, \mathbb{E}\big(\|g\|_2^2\big)^{1/2}\, \mathbb{E}\bigg(\int_{\{|g|\geq c\}} g^2 d\mu\bigg)^{1/2}\end{aligned}$$

since

$$|g - h| \,=\, |g - g_c| \,\leq\, |g|\mathbb{1}_{\{|g|\geq c\}} + \int_{\{|g|\geq c\}} |g| d\mu.$$

For every $y \in M$,

$$\mathbb{E}\big(g(y)^2\big) \,=\, \frac{1}{n}\, p_{2t}(y,y) \,\leq\, \frac{C}{nt}\,.$$

By standard exponential inequalities for sums of independent (bounded) random variables (cf. [1]), for every $u > 0$,

$$\mathbb{P}\big(|g(y)| \geq u\big) \,\leq\, Ce^{-nt\min(u,u^2)/C}$$

for some $C > 0$ possibly varying from line to line. It is then an easy task to show that for $c = \frac{1}{\log n}$ and $t = \frac{(\log n)^\kappa}{n}$ where $\kappa > 0$ is large enough,

$$\mathbb{E}\bigg(\int_{\{|g| \geq c\}} g^2 d\mu \bigg) \, = \, O\Big(\frac{1}{n^3} \Big).$$

The trace asymptotics (cf. [1]) shows that as $t = t_n \to 0$

$$\mathbb{E}\bigg(\int_M g(-\mathrm{L})^{-1} g \, d\mu \bigg) \, = \, \frac{1}{4\pi} \frac{\log \frac{1}{t_n}}{n} + O\Big(\frac{1}{n} \Big). \tag{18}$$

After some details, we conclude that

$$\mathbb{E}\big(\mathrm{W}_2^2(\mu_n^t, \mu) \big) \, \geq \, \frac{1}{4\pi} \frac{\log n}{n} - \frac{C \log \log n}{n}$$

which is the announced lower bound (16).

Next we turn to the limsup upper bound (17). The first step is the standard regularization procedure. For every $0 < \alpha \leq 1$, and $t > 0$,

$$\mathbb{E}\big(\mathrm{W}_2^2(\mu_n, \mu) \big) \, \leq \, (1 + \alpha) \, \mathbb{E}\big(\mathrm{W}_2^2(\mu_n^t, \mu) \big) + \frac{Ct}{\alpha} \, .$$

Then, we slightly modify the proof of Theorem 1. Namely, for every $\theta : [0,1] \to [0,1]$ increasing, $\theta(0) = 0$, $\theta(1) = 1$,

$$\mathrm{W}_2^2(\mu_n^t, \mu) \, \leq \, \int_M \big| \nabla \big((-\mathrm{L})^{-1} g \big) \big|^2 \int_0^1 \frac{\theta'(s)^2}{1 + \theta(s) g} \, ds \, d\mu.$$

Given $0 < c < 1$, choose $\theta(s) = (1+c)s$ if $s \in [0, 1-c]$ and $\theta(s) = 2s - s^2$ if $s \in [1-c, 1]$. Then

$$\begin{aligned} \int_0^1 \frac{\theta'(s)^2}{1 + \theta(s) g} \, ds \, &= \, \int_0^{1-c} \frac{(1+c)^2}{1 + (1+c)sg} \, ds + 4 \int_{1-c}^1 \frac{(1-s)^2}{1 + (2s - s^2) g} \, ds \\ &\leq \, \frac{1+c}{g} \, \log \big(1 + (1 - c^2) g \big) + 4c \end{aligned}$$

where we used that $g \geq -1$ in the last step. Observe next that

$$\frac{1}{g} \log \big(1 + (1 - c^2) g \big) \, \leq \, 1 + 2|g|$$

if $-\frac{1}{2} \leq g$ while

$$\frac{1}{g} \log \big(1 + (1 - c^2) g \big) \, \leq \, 2 \log \frac{1}{c^2}$$

if $g \leq -\frac{1}{2}$ so that in any case

$$\frac{1}{g} \log \big(1 + (1 - c^2) g \big) \, \leq \, 1 + 2|g| \Big(1 + 2 \log \frac{1}{c^2} \Big).$$

As a consequence,

$$\mathrm{W}_2^2(\mu_n^t,\mu) \,\leq\, (1+5c)\int_M \big|\nabla((-\mathrm{L})^{-1}g)\big|^2 d\mu + 4\Big(1+\log\frac{1}{c^2}\Big)\int_M |g|\big|\nabla(-\mathrm{L})^{-1}g\big|^2 d\mu.$$

By the Cauchy-Schwarz inequality,

$$\mathbb{E}\bigg(\int_M |g|\big|\nabla(-\mathrm{L})^{-1}g\big|^2 d\mu\bigg) \,\leq\, \mathbb{E}\big(\|g\|_2^2\big)^{1/2}\,\mathbb{E}\bigg(\int_M \big|\nabla(-\mathrm{L})^{-1}g\big|^4 d\mu\bigg)^{1/2}$$

Together with the Riesz transform bounds, it is shown in [1] that

$$\mathbb{E}\bigg(\int_M \big|\nabla(-\mathrm{L})^{-1}g\big|^4 d\mu\bigg) \,\leq\, C\left(\frac{\log n}{n}\right)^2.$$

On the other hand $\mathbb{E}(\|g\|_2^2) \leq \frac{C}{nt}$.

At this level, we have thus obtained that

$$\begin{aligned}\mathbb{E}\big(\mathrm{W}_2^2(\mu_n,\mu)\big) \,\leq\, (1+5c)(1+\alpha)\,\mathbb{E}\bigg(\int_M \big|\nabla((-\mathrm{L})^{-1}g)\big|^2 d\mu\bigg)\\ + \frac{Ct}{\alpha} + C\Big(1+2\log\frac{1}{c^2}\Big)\frac{\log n}{n\sqrt{nt}}\,.\end{aligned}$$

Together with the trace asymptotics (18), take

$$c \,=\, \frac{1}{\log n}\,, \quad t \,=\, \frac{(\log n)^{\frac{1}{4}}}{n}\,, \quad \alpha \,=\, \frac{1}{(\log n)^{\frac{1}{4}}}$$

to get that

$$\limsup_{n\to\infty}\left(\frac{n}{\log n}\,\mathbb{E}\big(\mathrm{W}_2^2(\mu_n,\mu)\big) - \frac{1}{4\pi}\right)\frac{(\log n)^{\frac{1}{4}}}{\log\log n} \,<\, \infty.$$

(With a few more efforts, it is possible to get rid of the $\log\log n$ factor.)

References

[1] L. Ambrosio, F. Stra, D. Trevisan. A PDE approach to a 2-dimensional matching problem (2016). *Probab. Theory Related Fields* 173, 433-478 (2019).

[2] S. Bobkov, M. Ledoux. One-dimensional empirical measures, order statistics, and Kantorovich transport distances (2016). To appear in *Memoirs Amer. Math. Soc.*

[3] S. Caracciolo, C. Lucibello, G. Parisi, G. Sicuro. Scaling hypothesis for the Euclidean bipartite matching problem. *Physical Review E* 90, 012118 (2014).

[4] M. Ledoux. On optimal matching of Gaussian samples. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 457, Veroyatnost' i Statistika.* 25, 226–264 (2017).

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

[6] M. Talagrand. Scaling and non-standard matching theorems. *Comptes Rendus Acad. Sciences Paris, Mathématique* 356, 692–695 (2018).

Institut de Mathématiques de Toulouse
Université de Toulouse – Paul-Sabatier, F-31062 Toulouse, France
& Institut Universitaire de France
ledoux@math.univ-toulouse.fr