A fluctuation result in dual Sobolev norm for the optimal matching problem

Michel Ledoux

University of Toulouse, France

Abstract

We discuss a fluctuation result in a dual Sobolev norm for the optimal matching problem. The principle is based on the heat kernel regularization procedure put forward by L. Ambrosio, F. Stra and D. Trevisan in the proof of the exact limit in the two-dimensional matching problem, and is coherent with the known one-dimensional case. The arguments are based on the classical central limit theorem in Hilbert space together with heat kernel bounds, and the limiting distribution is identified as the suitable renormalization of a time-space Gaussian Free Field.

1 Introduction and main results

The purpose of this paper is to study some fluctuation results in the optimal matching problem in the dual Sobolev norm, linearization of the Monge-Kantorovich metric W_2 .

The quadratic Monge-Kantorovich distance (cf. e.g. [11]) between two probability measures ν and μ on the Borel sets of a metric space (M, ρ) with a finite second moment is defined by

$$\mathcal{W}_2(\nu,\mu) \;=\; \left(\; \int_{M\times M} \rho(x,y)^2 d\pi(x,y) \right)^{1/2}$$

where the infimum is taken over all couplings π on $M \times M$ with respective marginals ν and μ .

When M is a smooth compact Riemannian manifold with its Riemannian metric ρ and normalized Riemannian volume element μ , it is a classical result from optimal transport (cf. e.g. [11], p. 588) that for $d\nu_{\varepsilon} = fd\mu$ where $f = 1 + \varepsilon g$,

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon^2} \operatorname{W}_2^2(\nu_{\varepsilon}, \mu) = \|g\|_{\mathrm{H}^{-1}(\mu)}^2$$
(1)

where

$$\|g\|_{\mathbf{H}^{-1}(\mu)} = \left(\int_{M} \left|\nabla((-\Delta)^{-1}g)\right|^{2} d\mu\right)^{1/2}$$
(2)

is the dual Sobolev norm of g (provided it is well-defined – see below for further details on this definition). With some abuse, we write $\|\nu - \mu\|_{\mathrm{H}^{-1}(\mu)} = \|f - 1\|_{\mathrm{H}^{-1}(\mu)}$ whenever $d\nu = f d\mu$. The Sobolev norm $\|\cdot\|_{\mathrm{H}^{-1}(\mu)}$ thus appears as a linearization of the Monge-Kantorovich metric W_2^2 . In addition to the infinitesimal behaviour (1), it also holds true that

$$W_2^2(\nu,\mu) \le 4 \|\nu - \mu\|_{H^{-1}(\mu)}^2$$
(3)

for every $\nu \ll \mu$ (cf. [9]).

Our interest in the dual Sobolev norm $\mathrm{H}^{-1}(\mu)$ is motivated by one illustration in optimal matching. In the following, μ is the Riemannian measure on a smooth closed *d*-dimensional Riemannian manifold M with volume one (the statements are invariant under rescaling of the measure). Given $X_1, \ldots, X_n, n \geq 1$, independent random variables with common distribution μ , let $\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ be the associated empirical measure. It has been shown by L. Ambrosio, F. Stra and D. Trevisan [1], that in dimension d = 2,

$$\lim_{n \to \infty} \frac{n}{\log n} \mathbb{E} \left(W_2^2(\mu_n, \mu) \right) = \frac{1}{4\pi} \,. \tag{4}$$

This result answers a conjecture by S. Caracciolo, C. Lucibello, G. Parisi and G. Sicuro [6], who expect moreover that

$$\lim_{n \to \infty} \left(\frac{n}{\log n} \mathbb{E} \left(W_2^2(\mu_n, \mu) \right) - \frac{1}{4\pi} \right) \log n = \xi$$
(5)

for some $\xi \in \mathbb{R}$ (a value is conjectured in [6]).

A further conjecture in this framework would be that

$$n\Big[\mathrm{W}_{2}^{2}(\mu_{n},\mu) - \mathbb{E}\big(\mathrm{W}_{2}^{2}(\mu_{n},\mu)\big)\Big] \to \chi$$
(6)

in distribution where χ is some centered random variable with an explicit distribution. Provided that the conjecture (5) holds true,

$$n\left[\mathrm{W}_{2}^{2}(\mu_{n},\mu)-\frac{1}{4\pi}\log n\right] \rightarrow \xi+\chi$$

in distribution, which would be the ultimate description of the limiting behaviour of $W_2^2(\mu_n, \mu)$ (provided the limiting value ξ is identified). For the matter of comparison, it may be mentioned that in dimension d = 1, for μ the Lebesgue measure on [0,1], $\mathbb{E}(W_2^2(\mu_n,\mu)) = \frac{1}{6n}$ while $n W_2^2(\mu_n,\mu)$ converges in law to $\int_0^1 B(t)^2 dt$ with B a Brownian bridge on [0,1] (in particular $\mathbb{E}(\int_0^1 B(t)^2 dt) = \frac{1}{6}$) [3].

The conjectures in [6] are based on a very appealing pde ansatz which however ignores some necessary regularization effect since the empirical measures μ_n are of course discrete. To address this issue, L. Ambrosio, F. Stra and D. Trevisan [1] propose a regularization procedure by the heat kernel $p_t(x, y), t > 0, x, y \in M$, on the manifold, replacing μ_n by the probability measure μ_n^t with density

$$f_n^t = \frac{1}{n} \sum_{i=1}^n p_t(X_i, \cdot)$$

with respect to μ . By the law of large numbers, the random densities f_n^t are close to 1 as $n \to \infty$ so that the approximation (1) may be used to identify the limit (4). More accurately, by the central limit theorem,

$$f_n^t = \frac{1}{n} \sum_{i=1}^n p_t(X_i, \cdot) \sim 1 + \frac{G(t)}{\sqrt{n}}$$

where $G(t) = \{G(t, y); y \in M\}$ is a Gaussian process with covariance

$$\mathbb{E}(G(t,y)G(t,y')) = \mathbb{E}([p_t(X_1,y)-1][p_t(X_1,y')-1])$$

= $\int_M [p_t(x,y)-1][p_t(x,y')-1]d\mu(x)$
= $p_{2t}(y,y')-1, \quad y,y' \in M,$

thereby suggesting a fluctuation result at the level of the W_2^2 metric. The delicate issue is however to balance the regularization in $t \to 0$ with the limit in n by a suitable choice $t = t(n) \to 0$. In particular, the validity of this procedure in dimension $d \ge 3$ is still conjectural.

To somewhat turn around this difficulty, we suggest here, as a further ansatz, to replace the Monge-Kantorovich metric W_2^2 by the norm in the dual Sobolev space $H^{-1}(\mu)$. To this task, note that simple computations developed in Section 3 show that

$$\mathbb{E}(\|\mu_{n}^{t} - \mu\|_{\mathbf{H}^{-1}(\mu)}^{2}) = \frac{1}{n}\gamma(t)$$

where

$$\gamma(t) = 2 \int_0^\infty \int_M \left[p_{2(t+s)}(x,x) - 1 \right] d\mu(x) \, ds.$$

Unless d = 1, it is not possible to make sense of $\|\mu_n - \mu\|_{\mathrm{H}^{-1}(\mu)}^2$ as the limit as $t \to 0$ of $\|\mu_n^t - \mu\|_{\mathrm{H}^{-1}(\mu)}^2$ since $\gamma(t) \to \infty$ as $t \to 0$. However, whenever $d \leq 3$, for each $n \geq 1$, the renormalization

$$n \|\mu_n^t - \mu\|_{\mathbf{H}^{-1}(\mu)}^2 - \gamma(t) \tag{7}$$

may be shown to converge in probability as $t \to 0$ to a random variable χ_n .

In a sense, the random variable χ_n would play the role of the left-hand side of (6) in the dual Sobolev norm, for which we establish the following fluctuation theorem. Set

$$\chi = 2 \int_0^\infty \left[\int_M G(s,y)^2 d\mu(y) - \int_M \mathbb{E} \left(G(s,y)^2 \right) d\mu(y) \right] ds.$$

It will be part of the result that χ is a well-defined random variable.

Theorem 1. Let $d \leq 3$. Under the preceding notation, $\chi_n \to \chi$ in distribution.

The proof of Theorem 1 is based on the central limit theorem in Hilbert space, together with appropriate equicontinuity properties derived from the heat kernel behaviour. In particular, these are only relevant for $d \leq 3$. The random variable χ may be formally understood in the following way. The quantity

$$\int_0^\infty\!\int_M G(s,y)^2 d\mu(y) ds$$

may be interpreted, in distribution, as the $H^{-1}(\mu)$ -norm of the Gaussian noise G(0) which, although not making sense as a function on M, may be described as the Gaussian field indexed by $L^{2}(\mu)$ such that for any $\phi, \psi \in L^{2}(\mu)$,

$$\mathbb{E}\big(\langle\phi, G(0)\rangle\langle\psi, G(0)\rangle\big) = \langle\phi, \psi\rangle - \langle\phi\rangle\langle\psi\rangle$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in $L^2(\mu)$ and $\langle \phi \rangle = \langle \phi, \mathbb{1} \rangle$. The time-space process $G(t) = \{G(t, y); y \in M\}$ is the heat kernel regularization of G(0) and its $H^{-1}(\mu)$ -norm is given by

$$\int_0^\infty \int_M G(t+s,y)^2 d\mu(y) ds$$

which is well-defined for any t > 0. Unless d = 1 (see below), it is not possible to let $t \to 0$, and χ is then obtained by renormalization by centering as in (7).

The drawback of Theorem 1 is of course that it is not expressed in terms of the Monge-Kantorovich metric W₂. Recent developments of [2, 8] allow for some approximate statements. Namely, it is shown there that there exist sequences $t = t(n) \rightarrow 0$ such that if $\hat{\mu}_n^t$ is the push-forward of μ by the exponential map $\exp(\nabla((-\Delta)^{-1}(f_n^t - 1))))$, then

$$W_2^2(\hat{\mu}_n^t,\mu) = \|\mu_n^t - \mu\|_{H^{-1}(\mu)}^2$$

with high probability. Furthermore, as developed in [2], provided that $t = t(n) \ge \frac{(\log n)^{\kappa}}{n^{1/(d-1)}}$ for some $\kappa > 0$ large enough $(t = t(n) \ge \frac{(\log n)^{\kappa}}{n^2}$ if d = 1),

$$\mathbb{E}\Big(\left| \mathrm{W}_2^2(\mu_n^t, \mu) \right) - \mathrm{W}_2^2(\hat{\mu}_n^t, \mu) \Big) \Big| \Big) \leq \frac{1}{n^2} \,.$$

It is thus an easy task to deduce the following consequence to Theorem 1.

Corollary 2. Let $d \leq 3$. There are sequences $t = t(n) \rightarrow 0$ such that

$$n \Big[\mathbf{W}_2^2(\mu_n^t, \mu) - \mathbb{E} \big(\mathbf{W}_2^2(\mu_n^t, \mu) \big) \Big] \rightarrow \chi$$

in distribution as $n \to \infty$.

The difficulty is then to replace $W_2^2(\mu_n^t, \mu)$ by $W_2^2(\mu_n, \mu)$, but unfortunately the known regularization results do not allow for such a conclusion for the admissible sequences $t = t(n) \to 0$. They are indeed unable to control the difference

$$\mathbb{E}\Big(\big|\mathbf{W}_2^2(\mu_n,\mu)\Big)-\mathbf{W}_2^2(\mu_n^t,\mu)\Big)\Big|\Big)$$

which is typically of order t.

The conclusion is however good enough in dimension d = 1, more specifically for the Lebesgue measure μ on [0, 1], since then the admissible sequence $t = t(n) \geq \frac{(\log n)^{\kappa}}{n^2}$ decays faster than $\frac{1}{n}$. Hence

$$n\left[\mathbf{W}_{2}^{2}(\mu_{n},\mu)-\mathbb{E}\left(\mathbf{W}_{2}^{2}(\mu_{n},\mu)\right)\right] \rightarrow \chi$$

weakly. Now, in this setting, on the one hand,

$$\gamma(t) = 2 \int_0^\infty \int_M \mathbb{E} \left(G(t+s,y)^2 \right) d\mu(y) ds \to \gamma(0) = \frac{1}{6}$$

while on the other

$$2\int_0^\infty \int_M G(s,y)^2 d\mu(y) ds$$

may be seen to have the same distribution as $\int_0^1 B(t)^2 dt$ where B is a Brownian bridge on [0, 1]. Therefore χ has the distribution of $\int_0^1 B(t)^2 dt - \frac{1}{6}$, and since $\mathbb{E}(nW_2^2(\mu_n, \mu)) = \frac{1}{6}$ (cf. e.g. [5]), we recover that

$$n \operatorname{W}_{2}^{2}(\mu_{n},\mu) \to \int_{0}^{1} B(t)^{2} dt$$
 (8)

in distribution, showing the coherence of Theorem 1 and Corollary 2 with the conjecture (6). But of course, the whole investigation here is only a minor step towards (6).

In the next Section 2, we give a brief account on the time-space Gaussian field $G = \{G(s, y); s > 0, y \in M\}$. Then, in Section 3, we prove Theorem 1 on the basis of the classical central limit theorem in Hilbert space. We work throughout this work with a closed manifold M to safely use the classical heat kernel asymptotics, although the case with boundary should be handled similarly (and we actually perform some comparison with the example of [0, 1], assuming implicitly the validity of the conclusion in this case).

2 A space-time Gaussian Free Field

This section provides a description of the Gaussian field $G = \{G(s, y); s > 0, y \in M\}$ presented in the introduction, and of some of its properties. Recall first the heat kernel $p_t(x, y), t > 0$, $x, y \in M$, on the closed *d*-dimensional Riemannian manifold M, and the associated semigroup $(P_t)_{t>0}$, invariant and symmetric with respect to μ , defined, on any $\phi \in L^2(\mu)$, by

$$P_t\phi(x) = \int_M \phi(y)p_t(x,y)d\mu(y).$$

Recall also the convolution and symmetry properties of the heat kernel in the form of

$$\int_{M} p_t(x,z) p_{t'}(z,y) d\mu(z) = p_{t+t'}(x,y) = p_{t+t'}(y,x)$$

for all $t, t' > 0, x, y \in M$.

If $\sigma(-\Delta)$ denotes the (countable) spectrum of the (inverse) Laplace operator $-\Delta$, with the eigenvalues counted with their multiplicities,

$$p_t(x,y) = \sum_{\lambda \in \sigma(-\Delta)} e^{-\lambda t} u_\lambda(x) u_\lambda(y)$$
(9)

where $\{u_{\lambda}\}_{\lambda\in\sigma(-\Delta)}$ is an $L^{2}(\mu)$ orthonormal basis of eigenvectors of $-\Delta$.

The semigroup $(P_t)_{t\geq 0}$ may be used in a simplified description of the dual Sobolev norm

$$\|g\|_{\mathbf{H}^{-1}(\mu)}^{2} = \int_{M} \left|\nabla((-\Delta)^{-1}g)\right|^{2} d\mu$$

where $g: M \to \mathbb{R}$ has mean zero and belongs to the suitable domain. Namely, by integration by parts,

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

By the spectral representation $(-\Delta)^{-1} = \int_0^\infty P_s ds$ and symmetry,

$$\|g\|_{\mathrm{H}^{-1}(\mu)}^2 = 2 \int_0^\infty \int_M (P_s g)^2 d\mu \, ds.$$
 (10)

For any t > 0, $G(t) = \{G(t, y); y \in M\}$ is therefore the Gaussian process indexed by M with covariance

$$\mathbb{E}\big(G(t,y)G(t,y')\big) = p_{2t}(y,y') - 1, \quad y,y' \in M$$

According to the eigenfunction expansion (9), G(t) may be represented as

$$G(t,y) = \sum_{\lambda \in \sigma(-\Delta) \setminus \{0\}} e^{-\lambda t} h_{\lambda} u_{\lambda}(y)$$
(11)

where $\{h_{\lambda}\}_{\lambda\in\sigma(-\Delta)}$ is a family of independent standard normal variables. In particular $P_s(G(t)) = G(t+s)$ for all s > 0.

We may observe that, for every t > 0, the map $G(t) : y \mapsto G(t, y)$ belongs almost surely to $H^{-1}(\mu)$. To this task, note first that $\int_M G(t)d\mu = 0$ almost surely by the expansion (11). Next, in the formulation (10),

$$\|G(t)\|_{\mathrm{H}^{-1}(\mu)}^2 = 2\int_0^\infty \int_M P_s (G(t))^2 d\mu \, ds = 2\int_0^\infty \int_M G(t+s)^2 d\mu \, ds.$$

Hence

$$\mathbb{E}(\|G(t)\|_{\mathrm{H}^{-1}(\mu)}^2) = 2\int_0^\infty \int_M \left[p_{2(t+s)}(x,x) - 1\right] d\mu(x) ds.$$

On a compact d-dimensional Riemannian manifold, it holds true that (see (17) below)

$$\sup_{x,y \in M} p_s(x,y) \le \frac{C}{s^{d/2}}, \quad 0 < s \le 1,$$
(12)

which takes care of the small time heat kernel behaviour. On the other hand, the spectral gap $\lambda_1 > 0$ ensures an exponential decay of convergence to equilibrium in the sense that for any $\varphi : M \to \mathbb{R}$ with mean zero and any u > 0,

$$\int_{M} (P_u \varphi)^2 d\mu \leq e^{-2\lambda_1 u} \int_{M} \varphi^2 d\mu$$

Apply this to $\varphi(y) = p_v(x, y) - 1$, v > 0, $y \in M$, for $x \in M$ fixed. Since then $P_u\varphi(y) = p_{v+u}(x, y) - 1$,

$$\int_{M} \left[p_{v+u}(x,y) - 1 \right]^2 d\mu(y) \le e^{-2\lambda_1 u} \left(\int_{M} \left[p_v(x,y) - 1 \right]^2 d\mu(y) \right).$$

By heat kernel convolution,

$$p_{2(v+u)}(x,x) - 1 \le e^{-2\lambda_1 u} [p_{2v}(x,x) - 1],$$
 (13)

for any $x \in M$ and v, u > 0, which is useful for the long time behaviour estimates.

As a consequence therefore of (13),

$$\mathbb{E}\left(\|G(t)\|_{H^{-1}(\mu)}^{2}\right) = 2\int_{0}^{\infty}\int_{M}\left[p_{2(t+s)}(x,x) - 1\right]d\mu(x)ds = \gamma(t) < \infty$$
(14)

for every t > 0, so that the Gaussian process $G(t) = \{G(t, y); y \in M\}$ indeed belongs almost surely to the dual Sobolev space $\mathrm{H}^{-1}(\mu)$.

Unless d = 1, it follows from the heat kernel behaviour (17) at $t \to 0$ that $\gamma(t) \to \infty$ as $t \to 0$, and therefore the $\mathrm{H}^{-1}(\mu)$ -norm of G(0) does not make sense. However, it may be shown that

$$||G(t)||^2_{\mathrm{H}^{-1}(\mu)} - \gamma(t)|$$

converges as $t \to 0$ to a well-defined random variable, and this for every $d \leq 3$. To reach this claim, we start with the following elementary lemma.

Lemma 3. Let U be a real square integrable random variable. Then

$$\mathbb{E}\left(|U^2 - \mathbb{E}(U^2)|\right) \leq 2\sqrt{2} \left(\operatorname{Var}(U)\right)^{1/2} \left(\mathbb{E}(U^2)\right)^{1/2}$$

Proof. If V is an independent copy of U, by independence and Jensen's inequality

$$\mathbb{E}(|U^2 - \mathbb{E}(U^2)|) \leq \mathbb{E}(|U^2 - V^2|).$$

Now, from the Cauchy-Schwarz inequality,

$$\mathbb{E}(|U^{2} - V^{2}|) \leq \mathbb{E}(|U - V|^{2})^{1/2} \mathbb{E}(|U + V|^{2})^{1/2}$$

from which the conclusion immediately follows.

Denote by $||G(t)||^2 = \int_M G(t, y)^2 d\mu(y)$ the $L^2(\mu)$ -norm of G(t). The preceding lemma may be combined with the concentration properties of norms of Gaussian vectors (see [10]) which express in particular that

$$\operatorname{Var}(\|G(t)\|) \leq \sigma^2$$

where

$$\sigma^2 = \sigma^2 (G(t)) = \sup \mathbb{E} (\langle \phi, G(t) \rangle^2)$$

the supremum being running over all unit vectors ϕ in $L^2(\mu)$. Now, again by (11),

$$\mathbb{E}(\langle \phi, G(t) \rangle^2) = \sum_{\lambda \in \sigma(-\Delta) \setminus 0\}} e^{-2\lambda t} \langle \phi, u_\lambda \rangle^2 \le \|\phi\|_{\mathrm{L}^2(\mu)}^2.$$

Hence $\sigma^2(G(t)) \leq 1$ and, by the lemma,

$$\mathbb{E}\Big(\big|\|G(t)\|^2 - \mathbb{E}\big(\|G(t)\|^2\big)\big|\Big) \le 2\sqrt{2}\,\mathbb{E}\big(\big\|G(t)\big\|^2\big)^{1/2}$$

On the other hand,

$$\mathbb{E}\left(\|G(t)\|^2\right) = \int_M \mathbb{E}\left(G(t,y)^2\right) d\mu(y) = \int_M \left[p_{2t}(y,y) - 1\right] d\mu(y)$$

which, as discussed above in (13) and (12), decreases exponentially as $t \to \infty$ while of order $t^{-2/d}$ as $t \to 0$. Therefore

$$\int_0^\infty \mathbb{E}\Big(\big|\|G(t)\|^2 - \mathbb{E}\big(\|G(t)\|^2\big)\big|\Big)dt < \infty$$

as soon as $d \leq 3$.

For the further purposes, observe in addition that

$$\int_{0}^{t} \mathbb{E}\Big(\big|\|G(s)\|^{2} - \mathbb{E}\big(\|G(s)\|^{2}\big)\big|\Big) ds \leq C t^{1-(d/4)}, \quad 0 < t \leq \frac{1}{2}.$$
(15)

The following statement summarizes the prior analysis.

Proposition 4. In the preceding notation, when $d \leq 3$, the random variable

$$\chi = 2 \int_0^\infty \left[\|G(s)\|^2 - \mathbb{E} \left(\|G(s)\|^2 \right) \right] ds$$

is well-defined.

Although formally we prefer to work on closed manifolds, we may check the coincidence in law in dimension one of $\int_0^1 B(t)^2 dt$ where B is a Brownian bridge on [0, 1] and of the L²-norm

$$\|G(0)\|_{\mathcal{H}^{-1}(\lambda)}^2 = 2 \int_0^\infty \int_{[0,1]} G(s,y)^2 d\mu(y) ds$$

where μ is Lebesgue measure on [0, 1]. The comparison is actually exact on the circle, but the one-dimensional optimal matching problem is classically studied and formulated on the interval [0, 1]. Following the discussion in the introduction, the random variable χ of Proposition 4 thus has the distribution of $\int_0^1 B(t)^2 dt - \frac{1}{6}$. To this task, recall first that the Brownian bridge on [0, 1] may be represented by the series expansion

$$B(t) = \sqrt{2} \sum_{k=1}^{\infty} h_k \frac{\sin(k\pi t)}{k\pi}, \quad t \in [0, 1],$$

where the h_k 's are independent standard normal variables. In particular,

$$\int_0^1 B(t)^2 dt = \sum_{k=1}^\infty \frac{h_k^2}{k^2 \pi^2}$$

On the other hand, since

$$G(s,y) = \sum_{k=1}^{\infty} e^{-k^2 \pi^2 s} h_k u_k(y), \quad t > 0,$$

where u_k are the eigenfunctions orthonormal for the Lebesgue measure on [0, 1],

$$\int_{[0,1]} G(s,y)^2 d\lambda(y) = \sum_{k=1}^{\infty} e^{-2k^2 \pi^2 s} h_k^2$$

from which the claim follows after integration in $s \in (0, \infty)$.

3 Proof of Theorem 1

Recall $X_1, \ldots, X_n, n \ge 1$, independent random variables with common distribution μ , and $\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ the associated empirical measure. For each t > 0, the random measure $d\mu_n^t = f_n^t d\mu$ where $f_n^t = \frac{1}{n} \sum_{i=1}^n p_t(X_i, \cdot)$ is obtained by regularization with the heat kernel.

To start with, according to (10),

$$\begin{aligned} \|\mu_n^t - \mu\|_{\mathrm{H}^{-1}(\mu)}^2 &= \|f_n^t - 1\|_{\mathrm{H}^{-1}(\mu)}^2 = 2\int_0^\infty \int_M [P_s f_n^t - 1]^2 d\mu \, ds \\ &= 2\int_0^\infty \int_M \left[\frac{1}{n} \sum_{i=1}^n \left[p_{t+s}(X_i, \cdot) - 1\right]\right]^2 d\mu \, ds. \end{aligned}$$

Averaging over the sample (X_1, \ldots, X_n) , by independence and identical distribution,

$$\mathbb{E}(\|\mu_n^t - \mu\|_{\mathrm{H}^{-1}(\mu)}^2) = \frac{2}{n} \int_0^\infty \int_M \mathbb{E}([p_{t+s}(X_i, \cdot) - 1]^2) d\mu \, ds$$
$$= \frac{2}{n} \int_0^\infty \int_M [p_{2(t+s)}(x, x) - 1] d\mu(x) ds = \frac{1}{n} \gamma(t)$$

since, for each $y \in M$,

$$\mathbb{E}(\left[p_{t+s}(X_i, y) - 1\right]^2) = \int_M \left[p_{t+s}(x, y) - 1\right]^2 d\mu(x) = p_{2(t+s)}(y, y) - 1$$

by the convolution properties of the heat kernel.

Next, we develop tools from the study of the central limit theorem in Hilbert space towards the goal. For each fixed t > 0, and $i \ge 1$, introduce the independent identically distributed, centered, random vectors

$$Z_i = Z_i(t+s,y) = p_{t+s}(X_i,y) - 1, \quad s \ge t, \ y \in M,$$

with values in the Hilbert space $L^2(ds \otimes \mu) = L^2((0, \infty) \times M, ds \otimes \mu)$. Denoting also (for simplicity) by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ the scalar product and norm in $L^2(ds \otimes \mu)$,

$$\chi_n^t = n \left\| \mu_n^t - \mu \right\|_{\mathrm{H}^{-1}(\mu)}^2 - \gamma(t) = 2 \left(\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \right\|^2 - \mathbb{E} \left(\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \right\|^2 \right) \right).$$
(16)

It holds that

$$\mathbb{E}\left(\|Z_i\|^2\right) = \int_0^\infty \int_M \mathbb{E}\left(\left[p_{t+s}(X_i, y) - 1\right]^2\right) d\mu(y) ds = \frac{1}{2}\gamma(t).$$

Since $\gamma(t) < \infty$, by the classical central limit theorem in Hilbert space (cf. [10]), the sequence

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}Z_{i}, \quad n \ge 1,$$

converges in distribution in $L^2(ds \otimes \mu)$ to the Gaussian vector $G^t = \{G(t+s, y); s > 0, y \in M\}$.

In addition, it may also be noted that

$$\mathbb{E}\left(\left\|\sum_{i=1}^{n} Z_{i}\right\|^{4}\right) = \sum_{i,j,k,\ell=1}^{n} \mathbb{E}\left(\langle Z_{i}, Z_{j}\rangle\langle Z_{k}, Z_{\ell}\rangle\right)$$
$$= n \mathbb{E}\left(\|Z_{1}\|^{4}\right) + n(n-1) \mathbb{E}\left(\|Z_{1}\|^{2}\right)^{2} + 2n(n-1) \mathbb{E}\left(\langle Z_{1}, Z_{2}\rangle^{2}\right)$$

since by independence and centering, only the indices such that $i = j = k = \ell$, $i = j \neq k = \ell$ or $i = k \neq j = \ell$, $i = \ell \neq j = k$ are contributing to the sum. Now, by the heat kernel convolution properties,

$$\mathbb{E}(||Z_1||^4) = \int_M \left(\int_0^\infty \left[p_{2(t+s)}(x,x) - 1\right] ds\right)^2 d\mu(x),$$
$$\mathbb{E}(||Z_1||^2)^2 = \left(\int_0^\infty \int_M \left[p_{2(t+s)}(x,x) - 1\right] d\mu(x) ds\right)^2$$

and

$$\mathbb{E}(\langle Z_1, Z_2 \rangle^2) = \int_0^\infty \int_M \int_0^\infty \int_M \mathbb{E}(Z_1(t+s, y)Z_1(t+s', y'))^2 ds ds' d\mu(y) d\mu(y')$$

= $\int_0^\infty \int_M \int_0^\infty \int_M \left[p_{2t+s+s'}(y, y') - 1 \right]^2 ds ds' d\mu(y) d\mu(y')$
= $\int_0^\infty \int_0^\infty \int_M \left[p_{2(2t+s+s')}(x, x) - 1 \right] d\mu(x) ds ds'.$

Together with (13) for, it therefore follows that

$$\sup_{n\geq 1} \mathbb{E}\left(\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_{i}\right\|^{4}\right) < \infty.$$

Hence, by convergence of moments, we also have that

$$\lim_{n \to \infty} \mathbb{E}\left(\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i \right\|^2 \right) = \mathbb{E}\left(\|G^t\|^2 \right).$$

We may summarize this analysis in the following statement. Define $\chi^t = 2(\|G^t\|^2 - \mathbb{E}(\|G^t\|^2))$ where

$$||G^t||^2 = \int_0^\infty \int_M G(t+s,x)^2 d\mu(x) ds$$

which is well defined for any t > 0.

Proposition 5. For every t > 0, as $n \to \infty$,

$$\chi_n^t = n \|\mu_n^t - \mu\|_{\mathbf{H}^{-1}(\mu)}^2 - \gamma(t) \to \chi^t$$

in distribution.

The next step investigates the behaviour of χ_n^t as $t \to 0$ (for each fixed n). The same Hilbert space computations as above actually show that for each fixed s > 0,

$$\operatorname{Var}\left(\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}Z_{i}(s)\right\|^{2}\right) = \frac{1}{n}\mathbb{E}\left(\left\|Z_{1}(s)\right\|^{4}\right) - \frac{1}{n}\mathbb{E}\left(\left\|Z_{1}(s)\right\|^{2}\right)^{2} + 2\left(1 - \frac{1}{n}\right)\mathbb{E}\left(\left\langle Z_{1}(s), Z_{2}(s)\right\rangle^{2}\right).$$

Here, the independent and identically distributed random variables $Z_i(s) = \{p_s(X_i, y) - 1; y \in M\}, i = 1, ..., n$, take their values in the Hilbert space $L^2(\mu)$, and the scalar product and norm are understood in this space. Again by the heat kernel properties,

$$\mathbb{E}(||Z_1(s)||^4) = \int_M [p_{2s}(x,x) - 1]^2 d\mu(x),$$
$$\mathbb{E}(||Z_1(s)||^2)^2 = \left(\int_M [p_{2s}(x,x) - 1] d\mu(x)\right)^2$$

and

$$\mathbb{E}\big(\langle Z_1(s), Z_2(s)\rangle^2\big) = \int_M \big[p_{4s}(x, x) - 1\big]d\mu(x).$$

The classical Minakshisundaram-Pleijel formula in Riemannian geometry (cf. e.g. [4, 7]) ensures that the heat kernel $p_u(x, x)$ admits, at each $x \in M$, a complete asymptotic expansion

$$p_u(x,x) \sim u^{-d/2} (a_0(x) + a_1(x)u + a_2(x)u^2 + \cdots), \quad u \to 0,$$
 (17)

where the functions $a_j(x)$ are smooth and determined by the metric and its derivatives at x. In particular, a_0 is constant equal to $\frac{1}{(4\pi)^{d/2}}$. The asymptotic expansion (17) holds uniformly in x, so that one can integrate it over $x \in M$, and together with the spectral representation of $p_u(x, y)$ and under the normalization of the volume element,

$$\int_{M} \left[p_u(x,x) - 1 \right] d\mu(x) \sim u^{-d/2} \left(\alpha_0 + \alpha_1 u + \alpha_2 u^2 + \cdots \right), \quad u \to 0.$$
 (18)

Now, since (cf. (16))

$$\chi_{n}^{t} = 2 \int_{t}^{\infty} \left[\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i}(s) \right\|^{2} - \mathbb{E} \left(\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_{i}(s) \right\|^{2} \right) \right] ds,$$

for 0 < t < t',

$$\chi_n^t - \chi_n^{t'} = 2 \int_t^{t'} \left[\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i(s) \right\|^2 - \mathbb{E} \left(\left\| \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i(s) \right\|^2 \right) \right] ds.$$

By the previous variance expansion,

$$\mathbb{E}\left(|\chi_{n}^{t}-\chi_{n}^{t'}|\right) \leq \frac{2}{\sqrt{n}} \int_{t}^{t'} \left(\int_{M} \left[p_{2s}(x,x) - \int_{M} p_{2s}(y,y)d\mu(y)\right]^{2} d\mu(x)\right)^{1/2} ds + 2\sqrt{2} \int_{t}^{t'} \left(\int_{M} \left[p_{4s}(x,x) - 1\right] d\mu(x)\right)^{1/2} ds$$

Making use of (17), (18) and (12), for every $\varepsilon > 0$ there exists $t_0 = t_0(\varepsilon) > 0$ such that, for $0 < t < t' < t_0$,

$$\mathbb{E}\left(|\chi_n^t - \chi_n^{t'}|\right) \le \varepsilon + \int_t^{t'} \frac{\sqrt{C}}{s^{d/4}} \, ds \le 2\varepsilon$$

uniformly in $n \ge 1$ (since $d \le 3$). By a Cauchy-type argument, we may therefore define, for each $n \ge 1$, the random variable $\chi_n = \lim_{t\to 0} \chi_n^t$, such that, moreover, for every $0 < t < t_0$,

$$\sup_{n\geq 1} \mathbb{E}(|\chi_n^t - \chi_n|) \leq 2\varepsilon.$$
(19)

We conclude the proof of Theorem 1 and show the weak convergence of the sequence χ_n , $n \geq 1$, to the random variable χ of Proposition 4. Let $\psi : \mathbb{R} \to \mathbb{R}$ be a bounded 1-Lipschitz function. By the triangle inequality, for each $n \geq 1$ and t > 0,

$$\begin{aligned} \left| \mathbb{E} \big(\psi(\chi_n) \big) - \mathbb{E} \big(\psi(\chi) \big) \right| &\leq \left| \mathbb{E} \big(\psi(\chi_n) \big) - \mathbb{E} \big(\psi(\chi_n^t) \big) \right| + \left| \mathbb{E} \big(\psi(\chi^t) \big) - \mathbb{E} \big(\psi(\chi^t) \big) \right| \\ &+ \left| \mathbb{E} \big(\psi(\chi^t) \big) - \mathbb{E} \big(\psi(\chi) \big) \right| \\ &\leq \mathbb{E} \big(\left| \chi_n - \chi_n^t \right) \right| \big) + \left| \mathbb{E} \big(\psi(\chi_n^t) \big) - \mathbb{E} \big(\psi(\chi^t) \big) \right| \\ &+ \mathbb{E} \big(\left| \chi^t - \chi \right) \right| \big). \end{aligned}$$

Given $\varepsilon > 0$, choose $0 < t < t_0(\varepsilon)$ small enough so that, by (15) and (19), $\mathbb{E}(|\chi^t - \chi)|) \leq \varepsilon$ and $\mathbb{E}(|\chi_n - \chi_n^t)|) \leq 2\varepsilon$ uniformly in $n \geq 1$. It remains to apply Proposition 5 to this value of t. The proof of Theorem 1 is therefore complete.

References

- L. Ambrosio, F. Stra, D. Trevisan. A PDE approach to a 2-dimensional matching problem. Probab. Theory Related Fields 173, 433-478 (2019).
- [2] L. Ambrosio, F. Glaudo. Finer estimates on the 2-dimensional matching problem (2018).
- [3] E. del Barrio, E. Giné, F. Utzet. Asymptotics for L_2 functionals of the empirical quantile process, with applications to tests of fit based on weighted Wasserstein distances. *Bernoulli* 1, 131–189 (2005).
- [4] M.Berger, P. Gauduchon, E. Mazet. Le spectre d'une variété riemannienne. Lecture Notes in Mathematics 194. Springer (1971).
- [5] S. Bobkov, M. Ledoux. One-dimensional empirical measures, order statistics, and Kantorovich transport distances (2016). To appear in *Memoirs Amer. Math. Soc.*
- [6] S. Caracciolo, C. Lucibello, G. Parisi, G. Sicuro. Scaling hypothesis for the Euclidean bipartite matching problem. *Physical Review E* 90, 012118 (2014).
- [7] I. Chavel. *Eigenvalues in Riemannian geometry*. Pure and Applied Mathematics 115. Academic Press (1984).
- [8] F. Glaudo. On the *c*-concavity with respect to the quadratic cost on a manifold. *Nonlinear Analysis* 178, 145–151 (2019).
- 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).
- [10] M. Ledoux, M. Talagrand. Probability in Banach spaces (Isoperimetry and processes). Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer (1991).
- [11] C. Villani. *Optimal transport. Old and new.* Grundlehren der mathematischen Wissenschaften 338. Springer (2009).

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