New estimates on the matching problem[1]

Luigi Ambrosio

Scuola Normale Superiore, Pisa
luigi.ambrosio@sns.it
http://cvgmt.sns.it

Paris, 5.07.2019

[1] Joint works with Federico Stra, Federico Glaudo and Dario Trevisan

Outline

Matching problems

Generally speaking, matching problems deal with families of random M points, independent and identically distributed in a given d-dimensional domain D.

The problem is then to estimate (since exact computations are basically impossible, except in some 1-d cases) the cost, for M large, of the optimal matching (optimal transport).

Large literature in Statistical Mechanics, Computer Science, Probability, both from the theoretical and computational points of view.

The results depend in a very sensitive way on the dimension d and on the power p of the cost function $c = \mathrm{dist}^p$. Typical domains: $D = [0, 1]^d$, $D = \mathbb{T}^d$.

Our result, based on semigroup techniques, covers also more general domains, with $d = 1,\ 2$.

Two particular cases of optimal matching

$$W_p^p\Big(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}, \frac{1}{N}\sum_{i=1}^N \delta_{Y_i}\Big) := \min_{T\in\mathcal{S}\{1,\ldots,N\}} \frac{1}{N}\sum_{i=1}^N |Y_{T(i)} - X_i|^p$$

$$W_p^p\Big(\boldsymbol{m}, \frac{1}{N}\sum_{i=1}^N \delta_{X_i}\Big) := \min_{T:D\to\{X_1,\ldots,X_N\},\ \boldsymbol{m}(\{T=X_i\})=1/N} \int_D |T(x)-x|^p \, d\boldsymbol{m}(x)$$

In particular for $p = 2$ the level sets $\{T = X_i\}$ are convex (Voronoi cells).

The dual formulation

The optimal transport cost can also be written as a supremum:

$$W_p^p(\mu,\nu) = \sup_{f:X\to\mathbb{R}} -\int f\,d\mu + \int Q_1 f\,d\nu$$

with

$$Q_t f(y) = \inf_{x\in X} \left\{ f(x) + \frac{1}{pt^{p-1}}|x-y|^p \right\}.$$

In the case $p=1$ the formula reduces to (Kantorovich-Rubinstein duality)

$$W_1(\mu,\nu) = \sup\left\{ -\int f\,d\mu + \int f\,d\nu :\ \mathrm{Lip}(f) \le 1 \right\}.$$

Matching problems

• **Bipartite problem:** $M = 2N$, with N blue points, N red points, *all i.i.d.*, and we want to match each red point to a blue point, so that the problem is about the rate of convergence to 0 of

$$\mathbb{E}\Big(W_p^p\big(\tfrac{1}{N}\sum_i \delta_{X_i}, \tfrac{1}{N}\sum_i \delta_{Y_i}\big)\Big).$$

• **Monopartite problem:** $M = 2N$, but the points are not coloured (or coloured, but free to marry another point with the same colour).

• **Semi discrete problem.** If $\boldsymbol{m}$ is the common law of the X_i, we want to know the rate of convergence to 0 of

$$\mathbb{E}\Big(W_p^p\big(\tfrac{1}{N}\sum_i \delta_{X_i}, \boldsymbol{m}\big)\Big).$$

• **Grid matching problem.** Given a deterministic grid of "equally spaced" points, $Y_1, \ldots, Y_N$, estimate

$$\mathbb{E}\Big(W_p^p\big(\tfrac{1}{N}\sum_i \delta_{X_i}, \tfrac{1}{N}\sum_i \delta_{Y_i}\big)\Big).$$

Monotone rearrangements and why the 1d case is special

If $X_i,\ Y_i$ are $[0,1]$-valued, then assuming wlog that the X_i' are ordered, one has

$$W_p^p(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}, \frac{1}{N}\sum_{i=1}^N \delta_{Y_i}) = \frac{1}{N}\sum_{i=1}^N |Z_i - X_i|^p,$$

where $Z_1,\ldots,Z_N$ the *monotone rearrangement* of Y_i. Since the statistics of Z_i is known (β statistics $B(i,n+1-i)$), this leads to many explicit computations, for instance:

$$\mathbb{E}\big(W_2^2(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}, \frac{1}{N}\sum_{i=1}^N \delta_{Y_i})\big) = \frac{1}{3(N+1)} \qquad \text{for any } N, \text{ if } D=[0,1].$$

In higher dimensions, gradient monotone maps correspond to gradients of convex functions, and are much harder to compute.

Theorem. (Brenier) *If $\mu_0=\rho_0\mathscr{L}^d,\ \mu_1=\rho_1\mathscr{L}^d\in\mathscr{P}_2(\mathbb{R}^d)$, the optimal map T from μ_0 to μ_1 is unique and it is the gradient of a convex function ψ solving the Monge-Ampére equation* $\det\nabla^2\psi(x)=\rho_0(x)/\rho_1(\nabla\psi(x))$.

Three level of investigation

(1) Find *tight* upper and lower bounds:

$$C^{-1}\phi_{p,d}(N) \leq \mathbb{E}(W_p^p) \leq C\phi_{p,d}(N);$$

(2) Prove the existence of the limit of renormalized expectations, possibly computing/characterizing the limit:

$$\exists \ell_{p,d} := \lim_{N\to\infty} \frac{\mathbb{E}(W_p^p)}{\phi_{p,d}(N)};$$

(3) Find the second term in the expansion:

$$\mathbb{E}(W_p^p) \sim \ell_{p,d}\phi_{p,d}(N) + \phi^*_{p,d}(N) + o\big(\phi^*_{p,d}(N)\big).$$

Heuristics

Since we have N points in a d-dimensional domain, say $(0,1)^d$, we expect an average distance $\sim N^{-1/d}$, and so the naive guess is

$$\mathbb{E}\big(W_p^p(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}, \boldsymbol{m})\big) \sim \frac{1}{N^{p/d}}.$$

Using the random 1-Lipschitz function $\phi(z) := \min_i |z - X_i|$, Kantorovich duality with Hölder inequality give indeed

$$\mathbb{E}\big(W_p^p(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}, \boldsymbol{m})\big) \geq \big[\mathbb{E}\big(W_1(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}, \boldsymbol{m})\big)\big]^p \gtrsim \frac{1}{N^{p/d}}.$$

Heuristics

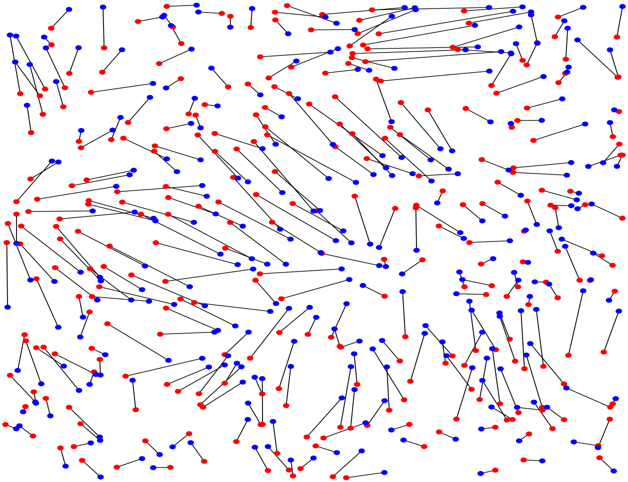
However, this lower bound is tight for $d > 2$, but *not* tight for $d = 2$, where a logarithmic correction appears:

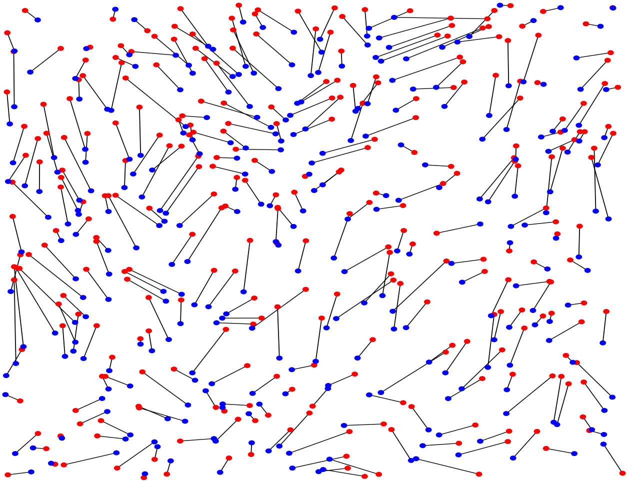
Theorem. (Ajtai-Komlós-Tusnády, Combinatorica, 1984) *For* $D = (0,1)^2$, $\boldsymbol{m}$ *the uniform measure and all* $p \geq 1$, *there exists* $c_p \in (0,\infty)$ *such that*

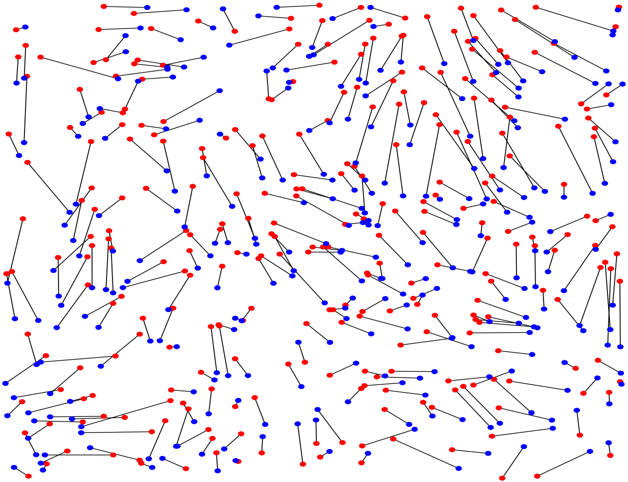
$$c_p^{-1}\frac{(\log N)^{p/2}}{N^{p/2}} \leq \mathbb{E}\big(W_p^p(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}, \boldsymbol{m})\big) \leq c_p \frac{(\log N)^{p/2}}{N^{p/2}}.$$

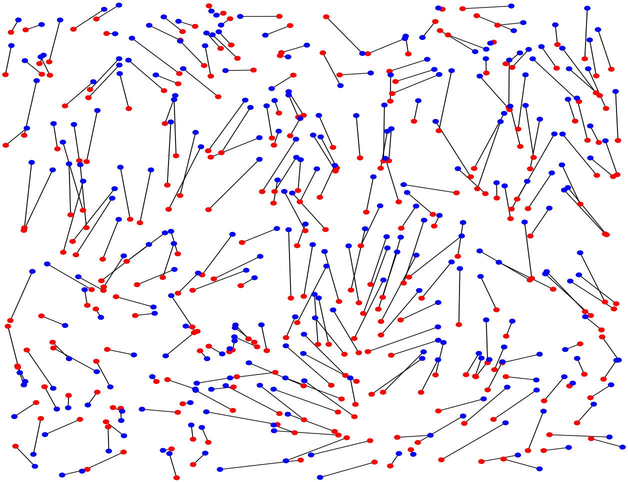
In physicist's words, *this is due to the fluctuations in the number of points, in small regions, which imply the necessity of "long distance pairings".*

If $d = 1$, as we have seen, we have even a larger deviation: $N^p\mathbb{E}(W_p^p) \sim N^{p/2}$.









Convergence of empirical measures

By the law of large numbers, for any continuous test function f one has

$$\frac{1}{N}\sum_i f(X_i) - \int f\,d\boldsymbol{m} \to 0 \quad \text{almost surely,}$$

which means that $\frac{1}{N}\sum_i \delta_{X_i} \to \boldsymbol{m}$ weakly as $N \to \infty$. Obviously we need a quantitative version of this fact, for instance the central limit theorem tells that

$$\sqrt{N}\Big(\frac{1}{N}\sum_i f(X_i) - \int f\,d\boldsymbol{m}\Big) \quad \text{weakly converge to a centered Gaussian, } \forall f.$$

Another information comes from Sanov's theorem, which gives

$$\mathbb{E}\Big(W_p\big(\frac{1}{N}\sum_i \delta_{X_i}, \boldsymbol{m}\big) > \epsilon\Big) \sim e^{-N\alpha(\epsilon)} \quad \alpha(\epsilon) := \inf\Big\{\mathrm{Ent}_{\boldsymbol{m}}(\nu):\ W_p(\nu, \boldsymbol{m}) \geq \epsilon\Big\}.$$

However these estimates are valid, for $\epsilon > 0$ fixed, for $N \geq N(\epsilon)$, and therefore useless to estimate $\mathbb{E}\big(W_p^p(\frac{1}{N}\sum_i \delta_{X_i}, \boldsymbol{m})\big)$.

Some results

Theorem. (Talagrand, Annals Appl. Prob., 1992) For $D=[0,1]^d$ and $d \geq 3$,

$$\limsup_{N\to\infty} N^{1/d}\mathbb{E}\Big(W_1\big(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}, \boldsymbol{m}\big)\Big) \leq \omega_d^{-1/d}\Big(1+K\frac{\log d}{d}\Big).$$

Theorem. (Dobric-Yukich, J. Th. Prob., 1995) *If $d \geq 3$, $D=(0,1)^d$ and $\boldsymbol{m}=\rho\mathscr{L}^d$, then*

$$\lim_{N\to\infty} N^{1/d}\,\mathbb{E}\Big(W_1\big(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}, \boldsymbol{m}\big)\Big) = \beta(d)\int_D \rho^{1-1/d}\,dx$$

for some constant $\beta(d)$.

Theorem. (Barthe-Bordenave, LNM, 2013) *If $D=[0,1]^d$ and $2p<d$, then*

$$\lim_{N\to\infty} N^{p/d}\,\mathbb{E}\Big(W_p^p\big(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}, \boldsymbol{m}\big)\Big) = \tilde{\beta}(d).$$

These results do not cover the case $d=2$, $p\geq 1$.

More probabilistic techniques

This topic is well illustrated in the 2014 monograph "Upper and lower bounds for stochastic processes" by Talagrand, particularly in the case $p = 1$.

The general idea, first developed in the Gaussian setting, is to estimate the expectation of the supremum

$$V := \sup_{u \in U} Z_u$$

of a centered stochastic process $\{Z_u\}_{u \in U}$ knowing the law of the random variables Z_u and the "metric" information

$$\big(\mathbb{E}(|Z_u - Z_v|^2)\big)^{1/2} \leq \rho(u, v).$$

This leads to bounds of the form (Dudley)

$$\mathbb{E}\Big(\sup_{v \in B_\delta(u)} |Z_v - Z_u| \Big) \leq C \int_0^\delta \sqrt{\log n(U, \rho, \epsilon)}\, d\epsilon \qquad \forall \delta > 0,$$

where $n(U, \rho, \epsilon)$ is the minimum number n of balls with radius ϵ needed to cover U, so the geometry of the space of parameters (U, ρ) comes into play.

More probabilistic techniques

Using Kantorovich duality, this technique can be applied with $U = \mathrm{Lip}_1(D)$, and

$$Z_u(\omega) := \int_D u\, d\boldsymbol{m} - \sum_{i=1}^{N} \frac{u(X_i(\omega))}{N}.$$

This technique is very general and powerful, but it does not seem to provide more than tight upper and lower bounds. Indeed, Talagrand raises (Research problem 4.3.3) the question about the existence of the limit

$$\lim_{N\to\infty} \sqrt{\frac{N}{\log N}}\, \mathbb{E}\big(W_1(\sum_{i=1}^{N} \frac{1}{N}\delta_{X_i}, \boldsymbol{m})\big)$$

in the 2-d case.

Moreover, when we consider W_2, we are forced to consider, as space of parameters U, the space of $d^2/2$-concave functions, and these arguments do not seem to be applicable, because the “geometry” of this space is harder.

The Caracciolo-Parisi ansatz

In a recent work (*Scaling hypothesis for the Euclidean bipartite matching problem*, Physical Review E, 2014), Caracciolo-Lucibello-Parisi-Sicuro used a specific ansatz to make predictions on the expansion of $\mathbb{E}\big(W_p^p(\rho_0,\rho_1)\big)$, in the case $D=\mathbb{T}^d$.

Predictions:

$$\frac{\mathbb{E}\big(W_p^p(\rho_0,\rho_1)\big)}{N^{-p/d}} \sim \begin{cases} \text{for } d=1,\ O(N^{p/2}) \text{ and } \frac{N}{6} \text{ for } p=2; \\ \text{for } d=2,\ O((\log N)^{p/2}),\ \frac{1}{2\pi}\log N + o(\log N) \text{ for } p=2; \\ \text{for } d>2,\ O(N^{(2-d)/d}); \\ \text{for } d>2 \text{ and } p=2,\ \frac{\zeta_d(1)}{2\pi^2}N^{(2-d)/d} + o\big(N^{(2-d)/d}\big). \end{cases}$$

The correctedness of the constants in blue, relative to W_2 and computed with the ansatz, has also been validated numerically (ζ_d is the so-called Epstein function).

The Caracciolo-Parisi ansatz

These predictions are obtained by linearizing *in C^1 topology* the Monge-Ampére equation

$$\det \nabla^2 \psi = \frac{\rho_0}{\rho_1 \circ \nabla \psi}$$

(which describes the optimal transport $\nabla \psi$ map from ρ_0 to ρ_1) around $\rho_0 = \rho_1 = 1$, thus writing $\nabla \psi = Id + \nabla \phi$ one obtains

$$-\Delta \phi = \rho_1 - \rho_0.$$

The ansatz says that $\nabla \phi$ should be "close" to the optimal displacement map and the predictions come from the computation of $\mathbb{E}(|\nabla \phi|^2)$, in discrete Fourier variables:

$$\mathbb{E}\Big(\int |\nabla \phi|^2\Big) = \mathbb{E}\Bigg(\sum_{\mathbf{n} \in \mathbb{Z}^d \setminus \{0\}} \frac{|\rho_{1,\mathbf{n}} - \rho_{0,\mathbf{n}}|^2}{4\pi^2 |\mathbf{n}|^2}\Bigg).$$

But, the empirical measures ρ_0 and ρ_1 *do not* belong to $H^{-1}(\mathbb{T}^d)$ as soon as $d > 1$, hence this energy is infinite for every ω!

The Caracciolo-Parisi ansatz

This ansatz is appealing, as it calls for PDE techniques in a random combinatorial optimization problem.

But, by mathematical standards, the proof of these predictions is not rigorous, first of all because of the appearence of divergent quantities, but also because in any case the ansatz does not provide a coupling between ρ_0 and ρ_1, only an approximate one, in some sense.

In any case, even if this were an exact coupling, the necessity of *lower* bounds (or the necessity to estimate how close it is to being optimal) remains.

Main result

Theorem. *Let D be a smooth, compact 2-dimensional Riemannian manifold. Then, if $\tilde{\boldsymbol{m}}$ is the normalization of Riemannian volume measure, one has*

$$\lim_{N\to\infty} \frac{N}{\log N}\, \mathbb{E}\Big(W_2^2\big(\sum_{i=1}^N \frac{1}{N}\delta_{X_i}, \tilde{\boldsymbol{m}}\big)\Big) = \frac{\boldsymbol{m}(D)}{4\pi}.$$

An analogous result is proved in the 1-d case. Our "PDE" proof use semigroup techniques and spectral analysis, for this reason it works for general domains.

Standard, by now, techniques related to the phenomenon of concentration of measure (Lévy, Milman, Gromov, in the more modern form Gaussian concentration is due to lower bounds on the Ricci tensor) then give also that

$$\frac{N}{\log N} W_2^2\big(\sum_{i=1}^N \frac{1}{N}\delta_{X_i}, \tilde{\boldsymbol{m}}\big) \quad \text{converge in law to} \quad (4\pi)^{-1}\boldsymbol{m}(D).$$

Main result

We are not yet able to attack Talagrand's problem, replacing $p = 2$ by $p = 1$ (more later). Nevertheless, our method provides a new "PDE" proof of the AKT result, namely

$$c_p^{-1}\frac{(\log N)^{p/2}}{N^{p/2}} \le \mathbb{E}\Big(W_p^p\big(\frac{1}{N}\sum_{i=1}^N \delta_{X_i}, \boldsymbol{m}\big)\Big) \le c_p\frac{(\log N)^{p/2}}{N^{p/2}}.$$

We are unfortunately very far from justifying all the predictions of the paper by Caracciolo-Lucibello-Parisi-Sicuro, this seems to require a much more refined analysis.

Ideas from the proof: upper bound

The heuristic idea is very natural. Since we know that $\mathbb{E}(W_2^2) \sim N^{-1}\log N$ exceeds the square N^{-1} of the "natural" length scale $\ell_N \sim N^{-1/2}$, we may hope to regularize just a bit the random densities $\rho \mapsto P_t\rho$, with $t = t_N = o(\frac{\log N}{N})$, so that one can apply the "dispersion" estimate

$$W_2^2(P_t\rho, \rho) \le C_D t = o\Big(\frac{\log N}{N}\Big).$$

Then, we can try to find an *exact* coupling between the regularized densities $P_t\rho_0$ and $P_t\rho_1$ and use the triangle inequality

$$W_2(\rho_0, \rho_1) \le W_2(\rho_0, P_t\rho_0) + W_2(P_t\rho_0, P_t\rho_1) + W_2(P_t\rho_1, \rho_1)$$

to get a good upper bound on $\mathbb{E}(W_2^2(\rho_0, \rho_1))$.

In order to provide a good coupling between $P_t\rho_0$ and $P_t\rho_1$ we use the Dacorogna-Moser interpolation. The estimates are quite delicate because $t_N \to 0$, so that in the limit the measures are concentrated.

Dacorogna-Moser interpolation

Given "nice" probability densities $\rho_0,\ \rho_1$, one can find a transport map T from ρ_0 to ρ_1 as the solution at $t = 1$ of the ODE

$$\frac{d}{dt}\boldsymbol{X}(t,x) = \boldsymbol{b}_t\big(\boldsymbol{X}(t,x)\big), \qquad \boldsymbol{X}(0,x) = x,$$

where the vector field $\boldsymbol{b}_t$ is $\rho_t^{-1}\nabla\phi$ and ϕ can be found solving the elliptic PDE

$$-\mathrm{div}\big(\nabla\phi\big) = \rho_1 - \rho_0 = \frac{d}{dt}\rho_t \quad \text{(with Neumann b.c.)} \tag{$*$}$$

with $\rho_t = (1-t)\rho_0 + t\rho_1$.

The reason (and the link with Benamou-Brenier's formula in optimal transport) is that, since $\boldsymbol{b}_t\rho_t = -\nabla\phi$, Poisson's equation (*) above can be written in the form of continuity equation:

$$\frac{d}{dt}\rho_t + \mathrm{div}\,(\boldsymbol{b}_t\rho_t) = 0.$$

Dacorogna-Moser interpolation

One has then, with simple computations,

$$
\begin{aligned}
W_2^2(\rho_0,\rho_1) &\leq \int |T(x)-x|^2\rho_0(x)\,d\boldsymbol{m}(x) \leq \int_0^1 \left(\int \frac{|\nabla\phi|^2}{\rho_t}\,d\boldsymbol{m} \right) dt \\
&= \int \int_0^1 \frac{1}{(1-t)\rho_0+t\rho_1}\,dt\,|\nabla\phi|^2\,d\boldsymbol{m} = \int \frac{|\nabla\phi|^2}{M(\rho_0,\rho_1)}\,d\boldsymbol{m}.
\end{aligned}
$$

The quantity $M(a,b)=(a-b)/(\log a-\log b)$ above is the so-called *logarithmic mean* of a and b.

Ideas from the proof: upper bound

Eventually, with some computations based on semigroup techniques we find:

$$
\begin{aligned}
\frac{N}{\log N}\mathbb{E}(W_2^2(\rho_0,\rho_1)) &\le \frac{N}{\log N}E\Big(\int \frac{|\nabla\phi|^2}{M(\rho_0,\rho_1)}\,d\boldsymbol{m}\Big)\\
&\sim \frac{N}{\log N}E\Big(\int |\nabla\phi|^2\,d\boldsymbol{m}\Big)\\
&= \frac{2}{\log N}\int_{1/N}^{\infty}\Big(\int p_{2t}(x,x)\,d\boldsymbol{m}(x)-1\Big)\,dt.
\end{aligned}
$$

The crucial quantity in this formula is $T(s) := \int p_s(x,x)\,d\boldsymbol{m}(x)$, which is related to the spectrum $\sigma(\Delta)$ of Δ by the *trace formula*

$$T(s) = \sum_{\lambda\in\sigma(\Delta)} e^{\lambda s}$$

(it is sufficient to write $p_t(x,y) = \sum_\lambda e^{\lambda t} f_\lambda(x) f_\lambda(y)$ with $y = x$ and integrate).

Ideas from the proof: upper bound

It turns out that the relevant limit is

$$\lim_N \frac{2}{\log N} \int_{1/N}^{\infty} \Big(\int p_{2t}(x,x)\, d\boldsymbol{m}(x) - 1 \Big)\, dt = \lim_N \frac{2}{\log N} \int_{1/N}^{\infty} \sum_{\lambda \in \sigma(\Delta)\setminus\{0\}} e^{2\lambda t}\, dt.$$

One can then use the asymptotic formula (McKean, Brown)

$$T(s) = \frac{1}{4\pi s} \Big(\boldsymbol{m}(D) - \frac{\sqrt{\pi s}}{2} \mathscr{H}^1(\partial D) + o(\sqrt{s}) \Big) \quad \text{as } s \to 0$$

to compute the limit (we are assuming $\boldsymbol{m}(D) = 1$), getting

$$\limsup_{N\to\infty} \frac{N}{\log N} \mathbb{E}\Big(W_2^2 \big(\sum_{i=1}^{N} \frac{1}{N} \delta_{X_i}, \boldsymbol{m} \big) \Big) \le \frac{1}{4\pi}$$

and similarly

$$\limsup_{N\to\infty} \frac{N}{\log N} \mathbb{E}\Big(W_2^2 \big(\sum_{i=1}^{N} \frac{1}{N} \delta_{X_i}, \sum_{i=1}^{N} \frac{1}{N} \delta_{Y_i} \big) \Big) \le \frac{1}{2\pi},$$

in agreement with the constants found numerically.

Ideas from the proof: lower bound

In the proof of the lower bound we use that D has no boundary, and let's assume that one of the densities, say ρ_1, is 1, we set $\rho_0 = \rho$.

For the lower bound it is natural to use Kantorovich duality: for any map ϕ one has

$$\frac{1}{2} W_2^2(\rho, 1) \geq -\int \phi\rho \, d\boldsymbol{m} + \int Q_1\phi \, d\boldsymbol{m},$$

where $Q_t\phi$ is given by the Hopf-Lax formula

$$Q_t\phi(y) := \inf_{x\in D} \phi(x) + \frac{1}{2t} d^2(x,y) \quad \text{solving} \quad \frac{d}{dt} Q_t\phi + \frac{1}{2}|\nabla Q_t\phi|^2 = 0.$$

If we choose ϕ with the ansatz, namely $-\Delta\phi = 1 - \rho$, let us try to estimate the lower bound from below, getting the term we had in the upper bound:

$$\begin{aligned} -\int \phi\rho \, d\boldsymbol{m} + \int Q_1\phi \, d\boldsymbol{m} &= \int \phi(1-\rho) \, d\boldsymbol{m} + \int (Q_1\phi - \phi) \, d\boldsymbol{m} \\ &= \int |\nabla\phi|^2 \, d\boldsymbol{m} - \frac{1}{2}\int_0^1 \int |\nabla Q_s\phi|^2 \, d\boldsymbol{m} \, ds. \end{aligned}$$

Ideas from the proof: lower bound

$$
\begin{aligned}
-\int \phi\rho\, d\boldsymbol{m} + \int Q_1\phi\, d\boldsymbol{m} &= \int \phi(1-\rho)\, d\boldsymbol{m} + \int (Q_1\phi - \phi)\, d\boldsymbol{m} \\
&= \int |\nabla\phi|^2\, d\boldsymbol{m} - \frac{1}{2}\int_0^1 \int |\nabla Q_s\phi|^2\, d\boldsymbol{m}\, ds \\
&\gtrsim \int |\nabla\phi|^2\, d\boldsymbol{m} - \frac{1}{2}\int |\nabla\phi|^2\, d\boldsymbol{m} \\
&= \frac{1}{2}\int |\nabla\phi|^2\, d\boldsymbol{m},
\end{aligned}
$$

where the last step is justified by the estimate

$$
\int |\nabla Q_s\phi|^2\, d\boldsymbol{m} \le (1 + O(\|\Delta\phi\|_\infty)) \int |\nabla\phi|^2\, d\boldsymbol{m}.
$$

For instance, if D has nonnegative Ricci curvature, it comes from

$$
\int |\nabla Q_s\phi|^2\, d\boldsymbol{m} \le e^{s\|\Delta\phi\|_\infty} \int |\nabla\phi|^2\, d\boldsymbol{m}.
$$

Ideas from the proof: lower bound

Recalling that ϕ solves the random PDE

$$-\Delta\phi = 1 - \rho,$$

we need to show that $1 - \rho$ is small in L^∞ with high probability (recall that ρ arises from the regularization with P_t, $t = o(N^{-1}\log N)$, of a random empirical measure).

This is the hardest part of the proof that prevents, for instance, a straightforward extension to Gaussian spaces.

Ideas from the proof: lower bound

The actual proof is a bit different, because $Q_t\phi$ is not so smooth. Hence, to prove the apriori estimates above on $\int |\nabla Q_s\phi|^2\, d\boldsymbol{m}$ we use the regularized HJ equation

$$\frac{d}{dt}f_t + \frac{1}{2}|\nabla f_t|^2 = \sigma\Delta f_t, \qquad f_0 = f$$

whose solution is explicitly given by the Hopf-Cole transform

$$f_t = -\sigma\log\left(P_{\sigma t}e^{-f/\sigma}\right),$$

and let $\sigma \to 0^+$ (here we need that D has no boundary).

In the case of manifolds with boundary, a new strategy introduced in a joint work with Glaudo bypasses the use of duality and uses instead the fact that $Id+\nabla\phi$ is an optimal map with very high probability and a refined contractivity estimate (allowing for a regularization scale $t(N) \gg N^{-1}\log N$).

The bipartite case

In the case of bipartite matching (N blue points, N red points) we expect

$$\mathbb{E}\big(W_2^2(\tfrac{1}{N}\sum_i \delta_{X_i}, \tfrac{1}{N}\sum_i \delta_{Y_i})\big) \sim 2\mathbb{E}\big(W_2^2(\tfrac{1}{N}\sum_i \delta_{X_i}, \boldsymbol{m})\big).$$

The heuristic argument is that on small scales $\mathscr{P}_2(D)$ is "Riemannian", so that

$$|a+b|^2 \sim |a|^2 + |b|^2 + 2|a||b|\cos\theta$$

and, since the "vectors" a and b pointing respectively from $\boldsymbol{m}$ to the random measures $\frac{1}{N}\sum_i \delta_{X_i}$, $\frac{1}{N}\sum_i \delta_{Y_i}$ are independent, on average the cosine term should give a null contribution.

We have been able to turn this heuristic argument into a proof (the inequality $\gtrsim$ is the hardest one).

Open problems: the case $p = 1$

This is the problem raised in Talagrand's book. If we want to attack even this one by PDE methods, we could go back to the PDE formulation of optimal transport (Evans-Gangbo)

$$\begin{cases} -\operatorname{div}(a\nabla u) = \rho_1 - \rho_0 \\ |\nabla u| \le 1, \ a(1 - |\nabla u|) = 0 \end{cases}$$

where $a \ge 0$ is the transport density, and to its q-Laplacian approximation, $q \to \infty$:

$$\begin{cases} -\operatorname{div}(|\nabla u|^{q-2}\nabla u) = \rho_1 - \rho_0 \\ \dfrac{\partial u}{\partial n} = 0. \end{cases}$$

Can we get sufficiently sharp estimates?

Does independence in the r.h.s. of this random PDE lead to convergence of the expectations? This is (much) less clear, for the moment.

More open problems

Even in the case $p = 2$, $D = \mathbb{T}^d$, there are many more open mathematical questions:

- Higher order expansion, for $d = 2$:

$$\lim_{N\to\infty} \left[\frac{N}{\log N} \mathbb{E}\Big(W_2^2\big(\tfrac{1}{N} \sum_i \delta_{X_i}, \tfrac{1}{N} \sum_i \delta_{Y_i}\big)\Big) - \frac{1}{2\pi} \right] \log N \in \mathbb{R}.$$

In a recent work with Glaudo, using a refined contractivity estimate, we proved that $[\cdots] = O(\sqrt{\log\log N / \log N})$.

- For $d > 2$, prove with this method existence of the limit

$$\lim_{N\to\infty} N^{d/2} \mathbb{E}\Big(W_2^2\big(\frac{1}{N} \sum_i \delta_{X_i}, \boldsymbol{m}\big)\Big).$$

In this case we lose the extra room granted in 2-d by the logarithmic correction, nevertheless Ledoux adapted the PDE technique to get at least tight bounds. We know from Barthe-Bordenave that the limit exists for $d \geq 5$.

More open problems

• What can be say about the asymptotic distribution of the $L^2(D;\mathbb{R}^d)$ random variables

$$\omega \mapsto \frac{T_{\boldsymbol{m}}^{\mu_N(\omega)} - Id}{\sqrt{\frac{\log N}{N}}} \ ?$$

In a recent work with Trevisan and Glaudo, we proved that the CPLS ansatz is correct also at the level of the maps. This is also related to the recent work of Goldman-Huesmann-Otto on the problem of the Lebesgue-to-Poisson coupling.

• What happens in the Gaussian case, i.e. when $\boldsymbol{m}$ is the standard Gaussian in $\mathbb{R}^d$? For $d=1$, one has

$$N^{-1}\log\log N \lesssim \mathbb{E}(W_2^2) \lesssim N^{-1}\log\log N$$

(Bobkov-Ledoux), while Talagrand recently proved

$$\mathbb{E}\Big(W_2^2\big(\frac{1}{N}\sum_i \delta_{X_i}, \boldsymbol{m}\big)\Big) \sim \frac{\log^2 N}{N} \qquad \text{if } d=2.$$

Again, the PDE technique can be used to provide even this result (Ledoux).

Thank you for the attention!

Slides available upon request