Quantitative stability for Optimal Transport

- a complex geometric approach

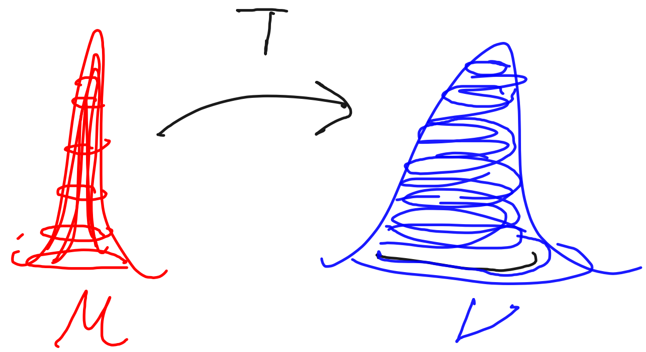
(Inra, 9/2-2022)

Robert Berman

OT setup

Assume given two probability measures on $\mathbb{R}^n$ μ and ν which have

- compact support
- densities wrt dx



A map T from $\mathbb{R}^n$ to $\mathbb{R}^n$ is called a *transport map* if

$$T_*\mu = \nu, \quad T \in L^\infty_{loc}$$

The *optimal transport map* minimizes

$$\min_T \int_{\mathbb{R}^n} |T(x) - x|^2 \mu =: \mathcal{W}_2(\mu, \nu)^2$$

It is unique and denoted by $T_{\mu,\nu}$.

Stability

The optimal transport map $T_{\mu,\nu}$ is stable wrt variations of the source μ :

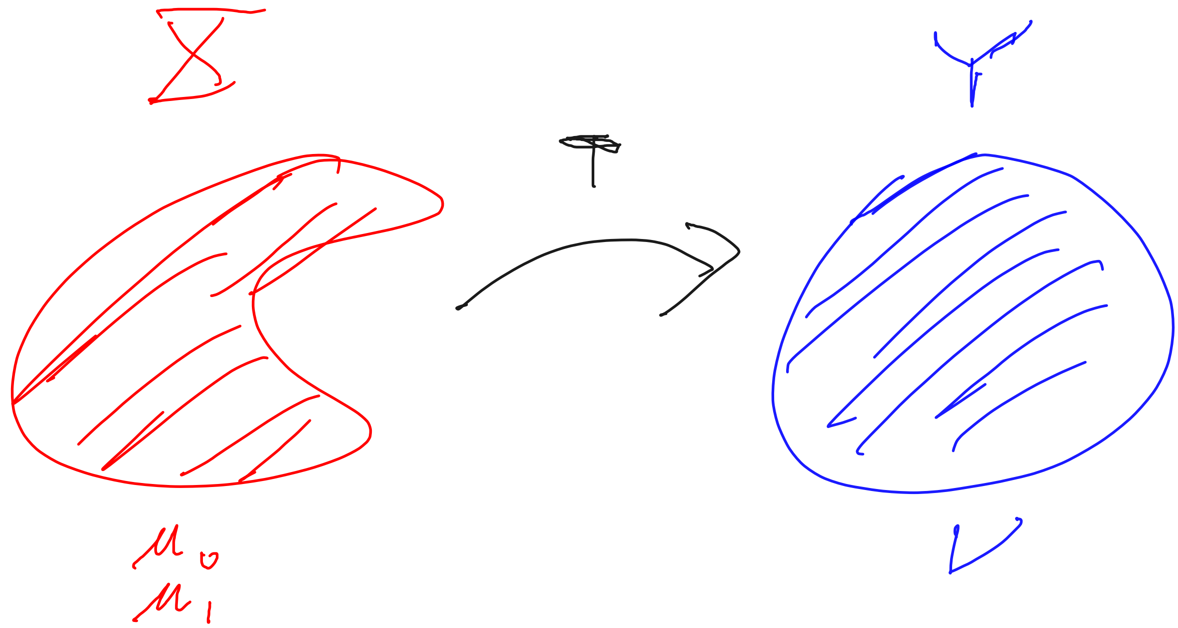
$$\mu_0 - \mu_1 \to 0 \implies T_{\mu_0,\nu} - T_{\mu_1,\nu} \to 0$$

(in the weak topology).

Can we make this stability *quantitative* (with a rate):

$$\|T_{\mu_0,\nu} - T_{\mu_1,\nu}\|_{L^2(X,dx)} \preceq \mathcal{W}_1(\mu_0,\mu_1)^\alpha$$

for some $\alpha > 0$?



Throughout the talk we will assume

- "sources" μ_0 and μ_1 are supported in the same bounded domain X
- The support of the "target" ν is a convex body Y
- $\inf_Y \frac{\nu}{dy} > 0$

Result (regular case)

There exists a constant c_0 such that

$$\|T_{\mu_0,\nu} - T_{\mu_1,\nu}\|_{L^2(X,dx)} \leq c_0 \mathcal{W}_1(\mu_0, \mu_1)^{1/2}$$

if μ_0 is "regular" (defined later on).

The general case

There exists a constant c such that

$$\|T_{\mu_0,\nu} - T_{\mu_1,\nu}\|_{L^2(X,dx)} \leq c\mathcal{W}_1(\mu_0,\mu_1)^{1/2^n}$$

The constant c only depends on

- upper bounds on the the *diameters* of X and Y

- positive lower bounds on the *volume* of Y and $\delta(:= \sup_Y(\nu/dy))$

Comparison with other results

- Previous result by Ambrosio 2011 (but variations wrt the *target* in a regular setting):

- Subsequent general result by Delalande-Mérigot 2021: universal exponent independent of n (!)(but worse when n is small)

- ...

Monge-Ampère formulation

Recall: the exists a *convex* function ϕ on $\mathbb{R}^n$ such that

$$T_{\mu,\nu} = \nabla\phi, \qquad \overline{(\nabla\phi)(\mathbb{R}^n)} = Y$$

It is uniquely determined (mod $\mathbb{R}$) by the transport condition:

$$(\nabla\phi)^{-1}\nu = \mu$$

Thus ϕ solves the following *MA-equation* on $\mathbb{R}^n$:

$$MA_\nu(\phi) = \mu, \qquad \overline{(\nabla\phi)(\mathbb{R}^n)} = Y$$

$$MA_\nu(\phi) = \mu, \qquad \overline{(\nabla\phi)(\mathbb{R}^n)} = Y$$

- Now μ need *not* be absolute continuous wrt dx.

- The quantitative stability now becomes

$$\|\nabla\phi_0 - \nabla\phi_1\|_{L^2(X,dx)} \preceq \mathcal{W}_1(\mu_0,\mu_1)^\alpha$$

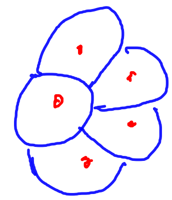
α $MA_\nu(\varphi_0)$

- The regularity assumption on μ_0 is:

$$\nabla^2\phi_0 \geq \epsilon I, \quad \epsilon > 0$$

(in the weak sense)

Application to discretizations



$$MA_\nu(\phi) = \mu, \qquad \overline{(\nabla\phi)(\mathbb{R}^n)} = Y$$

Discretize μ with a point cloud: and dual tesselation C_i

$$\mu_n \to \mu \qquad \mu_n = \cancel{\frac{1}{N}} \sum_{i=1}^{N} f_i \delta_{x_i} \to \mu \quad \text{(weakly)} \qquad f_i = \mu(C_i)$$

Denote by h the "spatial resolution"

$$h := \max_{i<N} \operatorname{diam}(C_i)$$

for the cells C_i in a dual tesselation of X and by ϕ_h the corresponding solution. Then

$$\|\nabla\phi - \nabla\phi_h\|_{L^2(X,dx)} \preceq h^{1/2}$$

in the "regular case" and in general

$$\|\nabla\phi - \nabla\phi_h\|_{L^2(X,dx)} \preceq h^{1/2^n}$$

The key new analytic inequalities

Henceforth, consider only smooth convex functions ϕ on $\mathbb{R}^n$ such that

$$\overline{(\nabla\phi)(\mathbb{R}^n)} = Y \quad (:= \text{support of } \nu)$$

In the "regular case", $\nabla^2\phi_0 \geq \epsilon I$,

$$\int_X |\nabla\phi_0 - \nabla\phi_1|^2\,dx \leq c_0 \int (\phi_1 - \phi_0)\,(MA_\nu(\phi_0) - MA_\nu(\phi_1))$$

In general, need to raise the lhs to 2^{n-1}.

- In the rhs can integrate over X or $\mathbb{R}^n$

- OT stability follows from

$$\left|\int_{\mathbb{R}^n} (\phi_1 - \phi_0)\,(\mu_0 - \mu_1)\right| \leq \mathrm{diam}(Y)\mathcal{W}_1(\mu_0, \mu_1)$$

using Kantorovich duality formula for $\mathcal{W}_1(\mu_0, \mu_1)$

$\|\nabla\phi_i\|_{L^\infty} \leq C$

Proof using complex geometry

Why?

- Can leverage the exterior algebra of complex forms (= multi-linearity)

- Can "compactify" $\mathbb{R}^n$ to avoid boundary terms in integration by parts

Starting point

It is enough to consider the case when

$$\nu = 1_Y dx$$

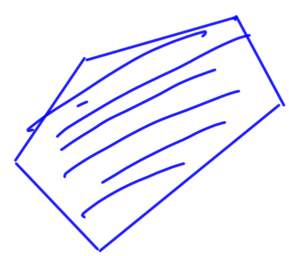
(not explained today...). Then

$$MA_\nu(\phi) = \det\left(\nabla^2\phi\right) dx$$

Moreover, may assume:

Y is a convex *polytope* with rational vertices

(by approximation)



Warm-up: $n = 1$

$$Y = [0,1], \quad MA(\phi) = \left(\nabla^2 \phi\right) dx.$$

In this case the key inequality is an *equality:*

$$\int_{\mathbb{R}} |\nabla\phi_0 - \nabla\phi_1|^2 \, dx = \int_{\mathbb{R}} (\phi_1 - \phi_0) \left(\nabla^2\phi_0 - \nabla^2\phi_1\right) dx.$$

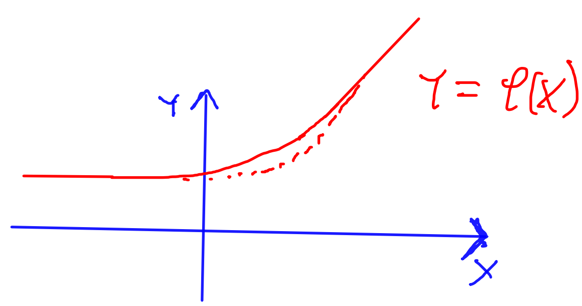
Indeed, integration by parts gives "boundary terms":

$$(\phi_1 - \phi_0)\,(\nabla\phi_0 - \nabla\phi_1)\,(\pm\infty).$$

But, by assumption,

$$\overline{(\nabla\phi)(\mathbb{R}^n)} = [0,1] \implies (\nabla\phi_0 - \nabla\phi_1)\,(\pm\infty) = 0$$

and $(\phi_1 - \phi_0) = O(1)$.



The "regular case" wher
$n > 1$

We will focus on the "regular case" $\nabla^2\phi_0 \geq \epsilon I$ and show

$$\|\nabla\phi_0 - \nabla\phi_1\|^2_{L^2(X,dx)} \leq c_0 \int (\phi_0 - \phi_1)\,(MA(\phi_1) - MA(\phi_0))$$

Start with the rhs.

- Will leverage complex exterior algebra:

$$MA(\phi) := \det\left(\nabla^2\phi\right) dx \sim (\partial\bar{\partial}\phi)^n.$$

"Hessian"

- Will first proceed formally

- To simplify notation will consider $\mathbb{R}^2$

Caveat

To simplify the notation will leave out all factors of 2,π etc!

$a^2 - b^2$

$= (a-b)(a+b)$

$$MA(\phi_1) - MA(\phi_0) = (\partial\bar{\partial}\phi_1)^2 - (\partial\bar{\partial}\phi_0)^2 =$$

$$= \left(\partial\bar{\partial}\phi_1 - \partial\bar{\partial}\phi_0\right) \wedge \left(\partial\bar{\partial}\phi_1 + \partial\bar{\partial}\phi_0\right).$$

Hence,

$$\int_{\mathbb{R}^2} (\phi_0 - \phi_1)\,(MA(\phi_1) - MA(\phi_0)) =$$

$$\int (\phi_0 - \phi_1)\left(\partial\bar{\partial}\phi_1 - \partial\bar{\partial}\phi_0\right) \wedge \left(\partial\bar{\partial}\phi_1 + \partial\bar{\partial}\phi_0\right) =$$

$$\int \partial(\phi_0 - \phi_1) \wedge \bar{\partial}\,(\phi_0 - \phi_1) \wedge \left(\underbrace{\cancel{\partial\bar{\partial}\phi_1}}_{\geq 0} + \partial\bar{\partial}\phi_0\right) \geq$$

$$\int \partial(\phi_0 - \phi_1) \wedge \bar{\partial}\,(\phi_0 - \phi_1) \wedge \underbrace{\partial\bar{\partial}\phi_0}_{\geq \varepsilon} \geq$$

$$\epsilon\,\|\nabla(\phi_0 - \phi_1)\|^2_{L^2(X,dx)}$$

Details on complexification

$$z := x + iy \in \mathbb{C}^n := \mathbb{R}^n + i\mathbb{R}^n$$
$\bar{z} = x - iy$

Decompose real exterior derivative d wrt dz_i and $d\bar{z}_i$:

$$d = \partial + \bar{\partial},$$

on $\mathbb{R}^n \times \mathbb{R}^n$

$$\partial := \sum_{i=1}^{n} \frac{\partial \cdot}{\partial z} dz \wedge, \quad \bar{\partial} := \sum_{i=1}^{n} \frac{\partial \cdot}{\partial \bar{z}} d\bar{z} \wedge$$

Hence,

$$\partial\bar{\partial}\phi(z) = \sum_{i=1}^{n} \frac{\partial^2 \phi}{\partial z_j \partial \bar{z}_j} dz_i \wedge d\bar{z}_j$$

In particular, if $\phi = \phi(x)$, then

$$\partial\bar{\partial}\phi = \sum_{i=1}^{n} \frac{\partial^2 \phi}{\partial x_j \partial x_j} dz_i \wedge d\bar{z}_j$$

$\left(\Leftrightarrow \nabla_x^2 \phi \right)$

As a consequence,

$\partial^2 f = 0$

$$\partial\bar{\partial}\phi = \sum_{i=1}^{n} \frac{\partial^2\phi}{\partial x_j \partial \bar{x}_j} dz_i \wedge d\bar{z}_j \implies$$

$\mathbb{C}^n$

$$(\partial\bar{\partial}\phi)^n = \det\left(\frac{\partial^2\phi}{\partial x_j \partial \bar{x}_j}\right) dx \wedge dy = MA(\phi) \wedge dy$$

Moreover, if $u = u(x)$ and $\phi_0 = |x|^2/2$, then

$$\partial u \wedge \bar{\partial} u \wedge (\partial\bar{\partial}\phi_0)^{n-1} = \left(|\nabla u|^2 dx\right) \wedge dy$$

$u = \varphi_0 - \varphi_1$

This justifies the formal proof, except the integration by parts.

It is handled by a compactification argument:

$$\int_{\mathbb{R}^n} (\phi_0 - \phi_1)\,(MA(\phi_1) - MA(\phi_0)) =$$

$$= \int_{X_Y} (\phi_0 - \phi_1)\left((\partial\bar{\partial}\phi_1)^n - (\partial\bar{\partial}\phi_0)^n\right)$$

on a certain *compact* complex manifold X_Y (without boundary!) such that

$$(\partial\bar{\partial}\phi_i) \quad \text{and} \quad \phi_0 - \phi_1$$

extend to X_Y.

Compactification

Recall: starting with $x \in \mathbb{R}^n$, we consider $z = x + iy \in \mathbb{C}^n$.

Step 1: "compactify" the $y-$direction by replacing

$$i\mathbb{R}^n \rightsquigarrow i(\mathbb{R}/2\pi\mathbb{Z})^n.$$

In complex terms: set

$$z_i = \log w_i \in \mathbb{C}, \quad w \in (\mathbb{C}^*)^n$$

Hence,

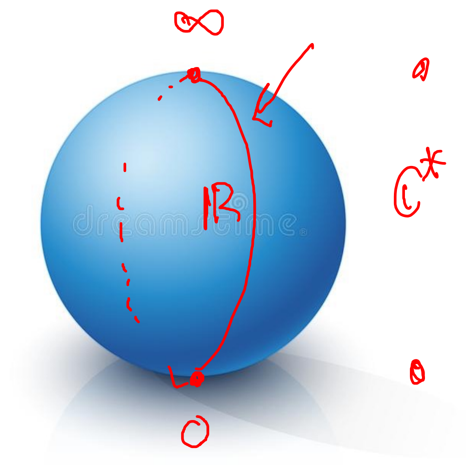
$$x_i = \log |w_i|$$

and

$$\int_{\mathbb{R}^n} (\phi_0 - \phi_1)\,(MA(\phi_1) - MA(\phi_0)) =$$

$$\int_{(\mathbb{C}^*)^n} (\phi_0 - \phi_1)\left((\partial\bar{\partial}\phi_1)^n - (\partial\bar{\partial}\phi_0)^n\right)$$

using that $\int_{\mathbb{R}/2\pi\mathbb{Z}} dy = 1/(2\pi)^n$.



Step 2: compactify $(\mathbb{C}^*)^n$

Motivation $(n = 1)$:

$$\mathbb{C}^* \hookrightarrow \mathbb{P}^1_{\mathbb{C}} (:= \mathbb{C}^* \bigcup \{0, \infty\},$$

where $\mathbb{P}^1_{\mathbb{C}}$ is a complex compact manifold (the Riemann sphere).

If

$$(\nabla\phi)(\mathbb{R}) = [0, 1],$$

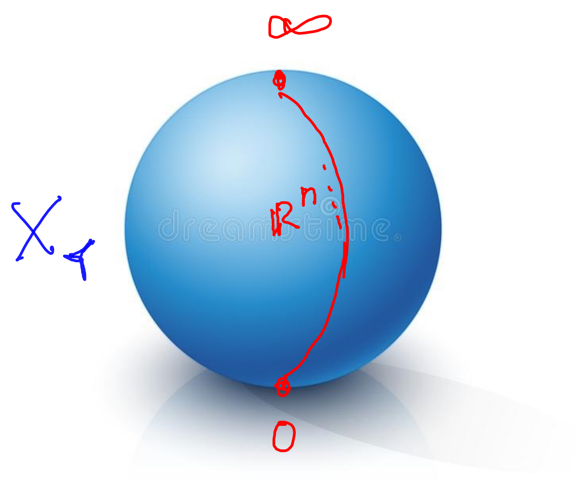
then

$\mathbb{P}^1_{\mathbb{C}}$

$$1 = \int_{\mathbb{R}} \nabla^2 \phi dx = \frac{i}{2\pi} \int_{\mathbb{C}^*} \partial\bar{\partial}\phi = \frac{i}{2\pi}\int_{\mathbb{P}^1} \partial\bar{\partial}\phi$$

Moreover,

$$\int_{\mathbb{R}} |\nabla(\phi_1 - \phi_0)|^2 \, dx = \frac{i}{2\pi} \int_{\mathbb{P}^1_{\mathbb{C}}} \partial(\phi_1 - \phi_0) \wedge \bar{\partial}(\phi_1 - \phi_0)$$



In higher dimensions, if

$$(\nabla\phi)(\mathbb{R}^n) = Y$$

for a convex bounded rational *polytope* Y. Then

$$(\mathbb{C}^*)^n \hookrightarrow X_Y \quad \text{(dense image)}$$

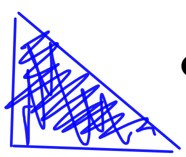
for a compact complex manifold (without boundary!) such that

$\partial\bar{\partial}\phi$ extends from $(\mathbb{C}^*)^n$ to X_Y

and likewise for differences $\phi_0 - \phi_1$

Ex:

- for $Y = [0,1]^2 \subset \mathbb{R}^2$ get $X_Y = \mathbb{P}^1_{\mathbb{C}} \times \mathbb{P}^1_{\mathbb{C}}$.



- for Y a simplex in $\mathbb{R}^n$ get $X_Y = \mathbb{P}^n_{\mathbb{C}}$

In general, X_Y is a *toric variety*, i.e. an equivariant compacfication of the complex torus $\mathbb{C}^{*n}$.

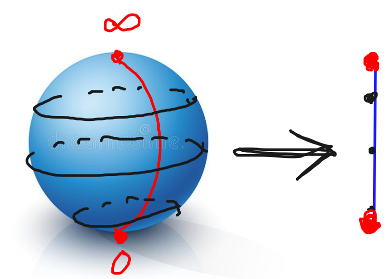
This proves the regular case.

Questions?

General case from Blocki's ineq.
(Cauchy-Schwartz)

Symplectic intepretation

$$\begin{tikzcd} T \arrow[r] & X_Y \arrow[r, two heads] & Y \\ \nabla\phi : & \mathbb{R}^n \arrow[u] \arrow[r, two heads] & \mathring{Y} \arrow[u] \end{tikzcd}$$

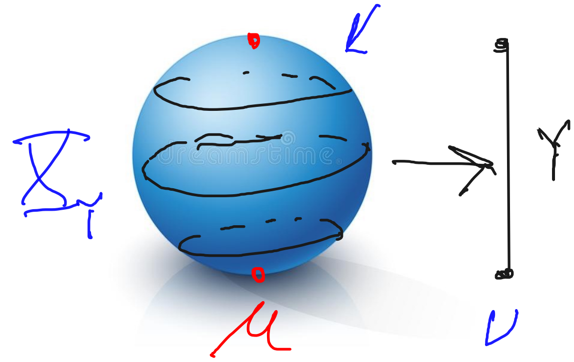


- dim $X_Y = 2n$

- The fibers of X_Y over $\mathring{Y}$ are $n-$dimensional tori T^n

- Some of the torus-dimensions shrink over the boundary of Y.

$$\boxed{\omega = i\partial\bar{\partial}\phi}$$



- The torus T^n acts on X_Y

- Can identify μ with a T^n-invariant measure on X_Y

- Finding $\nabla\phi$ amounts to finding a symplectic two-form ω on X_Y such that

$$\omega^n/n! = \mu.$$

- Then $\nabla\phi$ lifts to the *moment map* for the T^n-action on (X, ω) :

$$X_Y \twoheadrightarrow Y \subset \text{Lie}(T^n)^* = \mathbb{R}^n$$ (action/angle coord.)

- The moment map pushes forward $\left(\frac{\omega^n}{n!}\right)$ to dy on Y :

$$(\nabla\phi)_*\left(\frac{\omega^n}{n!}\right) = (\nabla\phi)_*\mu = 1_Y dy,$$

$$\underline{\omega = i\partial\bar{\partial}\phi}$$

- In complex notation:

$$\omega = i\partial\bar{\partial}\phi$$