Mini course GSSI

Felix Otto, MPI-MiS Leipzig

A variational regularity theory for optimal transportation, and applications to the matching problem

version June 25th 2025

I. Approximating optimal transportation by electrostatics

Kantorowicz' formulation of Monge's optimal transportation;
direct method of calculus of variations

Kantorowicz potential and Brenier's map; convex duality

Eulerian perspective: trajectories X and flux q

Entering and exiting times σ, τ and measures f, g for a ball $\bar{B}$

Electrostatics: the Helmholtz projection ∇u of q on B;
some regularity theory

Relating the Eulerian flux q
to the Lagrangian displacement $(y-x)d\pi$, locally

The flux q is close to its Helmholtz projection ∇u;
almost in total variation norm

Kantorowicz' formulation of Monge's optimal transp., direct method of calculus of variations

Given: (locally finite Borel) measures
$\lambda \geq 0$ on $\mathbb{R}^d \ni x$ and $\mu \geq 0$ on $\mathbb{R}^d \ni y$.
A measure $\pi \geq 0$ on $\mathbb{R}^d \times \mathbb{R}^d \ni (x, y)$ is "admissible"
iff it has marginals λ and μ:

$$\int \zeta(x) d\pi = \int \zeta d\lambda \quad \text{and} \quad \int \zeta(y) d\pi = \int \zeta d\mu.$$

Provided mass is finite and equal

$$\lambda(\mathbb{R}^d) = \mu(\mathbb{R}^d) \in (0, \infty),$$

the product measure $\pi = \frac{1}{\lambda(\mathbb{R}^d)} \lambda \otimes \mu = \frac{1}{\mu(\mathbb{R}^d)} \lambda \otimes \mu$ is admissible.
Note that for any admissible π

$$\pi(\mathbb{R}^d \times \mathbb{R}^d) = \lambda(\mathbb{R}^d) = \mu(\mathbb{R}^d) < \infty.$$

Consider squared transport distance and

$$\text{minimize} \quad \int |y-x|^2 d\pi \quad \text{among all } \pi \text{ admissible.}$$

Provided λ, μ have finite second moments,

$$\int |x|^2 d\lambda < \infty \quad \text{and} \quad \int |y|^2 d\mu < \infty,$$

any admissible π satisfies (monotone convergence)

$$\frac{1}{2}\int |x-y|^2 d\pi \le \int |x|^2 + |y|^2 d\pi = \int |x|^2 d\lambda + \int |y|^2 d\mu < \infty.$$

In particular, infimum $\in [0, \infty)$, and any minimizing sequence of π's is tight, so that marginals are preserved in the limit. Since functional is lower semi-continuous (Fatou), get minimizer by direct method. We fix a minimizer π.

Kantorowicz potential and Brenier's map; convex duality

By convex duality $\exists$convex function $\psi\colon \mathbb{R}^d \to (-\infty, +\infty]$ (not $\equiv +\infty$) such that

$$\text{supp}\pi \subset \partial\psi,$$

where the subgradient $\partial\psi \subset \mathbb{R}^d \times \mathbb{R}^d$ is defined by

$$(x, y) \in \partial\psi \quad \Longleftrightarrow \quad \forall\ x' \in \mathbb{R}^d\ \psi(x') \geq \psi(x) + (x' - x) \cdot y.$$

Informally suppπ is d-dimensional, as opposed to $2d$-dimensional for product measure. In particular, we have

$$\text{supp}\lambda \subset \{\, x \,|\, \exists y\ (x, y) \in \partial\psi \,\} =: \mathcal{D}(\psi).$$

Suppose $\exists$ an (open) ball $B \subset \mathbb{R}^d$ such that

$$B \subset \mathrm{supp}\lambda.$$

Then have $B \subset \mathcal{D}(\psi) \subset \{\, x \,|\, \psi(x) < \infty \,\}$. As convex function, ψ is locally bounded and locally Lipschitz on B. As locally Lipschitz function, ψ is Lebesgue-almost everywhere differentiable on B. If ψ is differentiable in $x \in B$, then by definition $\{\, y \,|\, (x, y) \in \partial\psi \,\} = \{\nabla\psi(x)\}$. Hence there exists a Lebesgue null set $N \subset B$ such that

$$(x, y) \in \partial\psi \text{ and } x \in B - N \quad \Longrightarrow \quad y = \nabla\psi(x).$$

If we suppose in addition

$$\lambda \ll dx \text{ on } B$$

then we obtain

$$\int_{B\times\mathbb{R}^d} \zeta(x, y)d\pi = \int_B \zeta(x, \nabla\psi(x))d\lambda.$$

Eulerian perspective: trajectories X and flux q

We identify pairs $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$
with straight trajectories $X\colon [0,1] \to \mathbb{R}^d$ via (the Borel map)

$$X_t = ty + (1-t)x \quad \text{so that} \quad \dot{X} = y - x.$$

Let the vectorial Borel measure q be defined through

$$\int \xi \cdot dq = \int \int_0^1 \xi(X_t) \cdot \dot{X} dt d\pi,$$

where ξ is a bounded smooth vector field on $\mathbb{R}^d$. Note q has finite total variation since integrand $\leq \sup|\xi|$ times $|y-x| \leq |x| + |y|$ $\leq 1 + \frac{1}{2}|x|^2 + \frac{1}{2}|y|^2$.

Applying definition to gradient fields $\xi = \nabla\zeta$, appealing to the chain rule $\nabla\zeta(X_t) \cdot \dot{X} = \frac{d}{dt}\zeta(X_t)$, and to $\int_0^1 \nabla\zeta(X_t) \cdot \dot{X} dt = \zeta(y) - \zeta(x)$, we obtain by the admissibility of π

$$\int \nabla\zeta \cdot dq = \int \zeta(d\mu - d\lambda).$$

Incidentally, this means

$$-\nabla \cdot dq = d\mu - d\lambda \quad \text{distributionally.}$$

In view of this we think of q as a flux.

Entering and exiting times σ, τ and measures f, g for a ball

Given a closed ball $\bar{B} \subset \mathbb{R}^d$,
define $\Omega_{\bar{B}}$ to be the set of trajectories that spend time in $\bar{B}$:

$$\Omega_{\bar{B}} \overset{\text{short}}{=} \Omega := \{\, X = (x, y) \,|\, \exists t \in [0,1]\; X_t \in \bar{B} \,\}.$$

Define the two Borel functions $\sigma_{\bar{B}}, \tau_{\bar{B}}$ or short $\sigma, \tau \colon \Omega \to [0,1]$ to be the times X enters/exits $\bar{B}$:

$$\begin{aligned} \sigma(X) &:= \min\{t \in [0,1] | X_t \in \bar{B}\} \\ &\le \max\{t \in [0,1] | X_t \in \bar{B}\} =: \tau(X). \end{aligned}$$

Define the two Borel measures $f_{\bar{B}}, g_{\bar{B}} \geq 0$, or short f, g, where the trajectories enter or exit:

$$\int \zeta df = \int_{\Omega \cap \{\sigma > 0\}} \zeta(X_\sigma) d\pi \quad \text{and} \quad \int \zeta dg = \int_{\Omega \cap \{\tau < 1\}} \zeta(X_\tau) d\pi;$$

well-defined because of $\pi(\mathbb{R}^d \times \mathbb{R}^d) < \infty$.

Since by definition,

$$\left\{ \begin{array}{l} \sigma(X) > 0 \iff X_{\sigma(X)} \in \partial B \\ \tau(X) < 1 \iff X_{\tau(X)} \in \partial B \end{array} \right\}$$

we have

$$f, g \quad \text{are supported on} \quad \partial B.$$

Claim

$$\int_{\bar{B}} \nabla\zeta \cdot dq = \int_{\bar{B}} \zeta(d\mu - d\lambda) + \int_{\partial B} \zeta(dg - df).$$

Apply definition of q to $\xi = I(\bar{B})\nabla\zeta$,
use $\int_0^1 \xi(X_t)\cdot\dot{X}dt = \int_\sigma^\tau \nabla\zeta(X_t)\cdot\dot{X}dt = \zeta(X_{\tau(X)}) - \zeta(X_{\sigma(X)})$. Since

$$\sigma(X) = 0 \iff x \in \bar{B} \quad \text{and} \quad \tau(X) = 1 \iff y \in \bar{B},$$

we get

$$\int_0^1 \xi(X_t)\cdot\dot{X}dt = I(y \in \bar{B})\zeta(y) - I(x \in \bar{B})\zeta(x) \\ + I(\tau(X) < 1)\zeta(X_{\tau(X)}) - I(\sigma(X) > 0)\zeta(X_{\sigma(X)}).$$

Integrate against π, use admissibility of π and definition of f, g.

Incidentally,

normal trace of q on $\partial B \;=\; g - f$
provided $\quad |q|(\partial B) = \lambda(\partial B) = \mu(\partial B) = 0.$

Electrostatics: the Helmholtz projection ∇u of q on B; some regularity theory

Helmholtz projection $\mathcal{H}_B = \mathcal{H}$ on B is $L^2(B, \mathbb{R}^d)$-orthogonal projection onto closed subspace of gradient fields. By singular integral theory, if ξ is smooth on $\bar{B}$, then $\mathcal{H}\xi$ is smooth on $\bar{B}$, and the $C^k(\bar{B})$-norm of $\mathcal{H}\xi$ is controlled by the $C^{k+1}(\bar{B})$-norm of ξ. Moreover, $\mathcal{H}$ is characterized by how it acts on smooth fields, namely

$$\mathcal{H}\nabla\zeta = \nabla\zeta \quad \text{for smooth } \zeta \text{ on } \bar{B},$$

$$\mathcal{H}\xi = 0 \quad \text{for smooth divergence-free } \xi \text{ supported in } B.$$

Hence to every distribution f on $\bar{B}$, we can associate its Helmholtz projection $\mathcal{H}f$ by duality via $\mathcal{H}f.\xi = f.\mathcal{H}\xi$. It is characterized by

$$\begin{cases} \mathcal{H}f.\nabla\zeta = f.\nabla\zeta \quad \text{for smooth } \zeta \text{ on } \bar{B}, \\ \mathcal{H}f.\xi = 0 \quad \text{for smooth divergence-free } \xi \text{ supported in } B. \end{cases}$$

As finite measure, $f = q\lfloor \bar{B}$ is a distribution.
Claim: $\mathcal{H}f$ is absolutely continuous w. r. t. Lebesgue:

$$\mathcal{H}q\lfloor \bar{B} \ \ll \ dx\lfloor B.$$

Enough to construct a $u_B = u \in H^{1,1}(B)$ such that

$$\int_B \nabla\zeta \cdot \nabla u dx = \int_{\bar{B}} \nabla\zeta \cdot dq;$$

then we have $\mathcal{H}q\lfloor \bar{B} = \nabla u dx\lfloor B$. Enough to establish

$$\int_B \nabla\zeta \cdot \nabla u dx = \int_{\bar{B}} \zeta(d\mu - d\lambda) + \int_{\partial B} \zeta(dg - df).$$

Consider $\int_{\bar{B}} \zeta(d\mu - d\lambda) + \int_{\partial B} \zeta(dg - df)$ as a linear form in ζ. It is bounded w. r. t. $\sup_{\bar{B}} |\zeta|$; it vanishes for constant ζ. By Sobolev embedding

$$\sup_{x,y\in\bar{B}} \frac{|\zeta(y) - \zeta(x)|}{|y - x|^\alpha} \lesssim \Big(\int_B |\nabla\zeta|^p dx \Big)^{\frac{1}{p}}$$

form is bounded w. r. t. $\nabla\zeta \in L^p(B, \mathbb{R}^d)$ for $p \in (d, \infty)$. By duality it can be represented by $\int_B \nabla\zeta \cdot \tilde{q} dx$ for some $\tilde{q} \in L^{p'}(B, \mathbb{R}^d)$ with $p' \in (1, \frac{d}{d-1})$. Then ∇u is the Helmholtz projection of $\tilde{q}$, which by singular integral theory is bounded in $L^{p'}(B, \mathbb{R}^d)$. In particular $u \in H^{1,p'}(B) \subset H^{1,1}(B)$.

Incidentally, u satisfies the Poisson equation with Neumann b. c.:

$-\triangle u = \mu - \lambda$ in B and $\nu \cdot \nabla u = \nu \cdot q$ on ∂B in a weak sense.

Relating the Eulerian flux q to the Lagrangian displacement $(y-x)d\pi$, locally

From definition of q

$$\begin{aligned}
&\int \xi(x)\cdot(dq-(y-x)d\pi) \\
&= \int_0^1 dt \int \Big(\xi(ty+(1-t)x)-\xi(x)\Big)\cdot(y-x)d\pi,
\end{aligned}$$

we obtain the inequality

$$\Big|\int \xi(x)\cdot(dq-(y-x)d\pi)\Big| \leq \sup|\nabla\xi| \int_0^1 dt \int t|y-x|^2 d\pi,$$

which entails

$$\Big|\int \xi(x)\cdot(dq-(y-x)d\pi)\Big| \leq \frac{1}{2}\sup|\nabla\xi| \int |y-x|^2 d\pi.$$

Seek version with transportation cost localized to a ball B;

$$E(B) := \int_{\Omega(B)} |y-x|^2 d\pi.$$

Replace ξ by $I(\bar{B})\xi$ in definition of q, split difference into

$$\begin{aligned} I(X_t \in \bar{B})\xi(X_t) - I(x \in \bar{B})\xi(x) &= I(X_t \in \bar{B})I(x \in \bar{B})(\xi(X_t) - \xi(x)) \\ &+ I(X_t \in \bar{B}, x \notin \bar{B})\xi(X_t) - I(X_t \notin \bar{B}, x \in \bar{B})\xi(x). \end{aligned}$$

First contribution as before:

$$\begin{aligned} &\Big| \int_0^1 dt \int I(X_t \in \bar{B})I(x \in \bar{B})(\xi(X_t) - \xi(x)) \cdot (y-x) d\pi \Big| \\ &\le \sup_{\bar{B}} |\nabla \xi| \int_0^1 dt \int I(X_t \in \bar{B}) |X_t - x||y-x| d\pi \le \sup_{\bar{B}} |\nabla \xi| \frac{1}{2} E(\bar{B}). \end{aligned}$$

Second contribution:

$$\Big|\int_0^1 dt \int \Big(I(X_t \in \bar{B}, x \notin \bar{B})\xi(X_t)$$
$$- I(X_t \notin \bar{B}, x \in \bar{B})\xi(x)\Big) \cdot (y-x)d\pi\Big|$$
$$\leq \sup_B |\xi| \int_0^1 dt \int |I(X_t \in \bar{B}) - I(x \in \bar{B})||y-x|d\pi.$$

Specify to a ball $\bar{B} = \bar{B}_R$ with radius R and write $|I(X_t \in \bar{B}) - I(x \in \bar{B})| = |I(R \geq |X_t|) - I(R \geq |x|)|$. Hence integral in R is estimated by $||X_t| - |x|| \leq |X_t - x|$ to the effect of

$$\int_0^{\bar{R}} dR \sup_\xi \frac{1}{\sup_{\bar{B}_R} |\xi|} \Big|\int_0^1 dt \int \Big(I(X_t \in \bar{B}_R, x \notin \bar{B}_R)\xi(X_t)$$
$$- I(X_t \notin \bar{B}_R, x \in \bar{B}_R)\xi(x)\Big) \cdot (y-x)d\pi\Big| \leq \frac{1}{2}E(B_{\bar{R}}).$$

We summarize these findings on the average-in-R estimate of a dual norm of $dq - (y-x)d\pi$ in

Lemma 1.

$$\int_0^{\bar R} dR \sup_\xi \frac{\left|\int_{\bar B_R} \xi(x)\cdot(dq-(y-x)d\pi)\right|}{\max\{\sup_{\bar B_R}|\xi|, \bar R \sup_{\bar B_R}|\nabla\xi|\}} \le E(B_{\bar R}).$$

We now comment on the regime in which Lemma 1 is not vacuous. Note that the l. h. s. compares $dq\lfloor \bar B_R$ to the marginal in x of $(y-x)d\pi\lfloor(\bar B_R\times\mathbb{R}^d)$, in a norm that scales like the total variation (but is weaker more like the flat norm). Hence Lemma 1 is meaningful if and only if $\int_0^{\bar R} dR \int_{\bar B_R\times\mathbb{R}^d}|y-x|d\pi$ is small compared to the r. h. s. that by definition dominates $\int_{B_{\bar R}\times\mathbb{R}^d}|y-x|^2 d\pi$. This is the case if

$$|y-x| \ll \bar R \quad \text{on average w. r. t. } \pi\lfloor(B_{\bar R}\times\mathbb{R}^d).$$

Loosely speaking, this means

$$\text{transportation distance} \ll \text{localization scale.}$$

The flux q is close to its Helmholtz projection ∇u; almost in total variation norm

From now on we need

$$\lambda = dx \quad \text{in } B_{\bar{R}}.$$

In this case

$$\int_{\bar{B}\times\mathbb{R}^d} \zeta(x,y)d\pi = \int_B \zeta(x,\nabla\psi(x))dx.$$

Hence expression in Lemma 1 turns into

$$\int_{\bar{B}\times\mathbb{R}^d} \xi(x)\cdot(dq-(y-x)d\pi) = \int_{\bar{B}} \xi(x)\cdot(dq-(\nabla\psi(x)-x)dx).$$

Note that by definition of Helmholtz projection on B (on $L^2(B,\mathbb{R}^d)$) we have $\mathcal{H}(\nabla\psi - \mathrm{id}) = \nabla\psi - \mathrm{id}$. Together with $\nabla u dx\lfloor B = \mathcal{H}q\lfloor\bar{B}$ we have in terms of the Helmholtz projection (on distributions)

$$dq\lfloor\bar{B} - \nabla u dx\lfloor B = (\mathrm{id}-\mathcal{H})(dq\lfloor\bar{B} - (\nabla\psi - \mathrm{id})dx\lfloor B).$$

Note that like $\mathcal{H}$, the "Leray projection" $\mathrm{id} - \mathcal{H}$ is bounded in the Hölder space $C^{1,\alpha}(\bar{B}, \mathbb{R}^d)$ for $\alpha \in (0,1)$; more precisely, it is uniformly in B bounded w. r. t. the norm

$$\sup_{\bar{B}} |\xi| + R^{1+\alpha} \sup_{x,y \in \bar{B}} \frac{|\nabla\xi(x) - \nabla\xi(y)|}{|x-y|^\alpha},$$

where R is the radius of B. We appeal to the embeddings

$$\sup_{\bar{B}_R} |\xi| + R \sup_{\bar{B}_R} |\nabla\xi| \lesssim \text{above norm on } B_R \lesssim \sup_{\bar{B}_R} |\xi| + R^2 \sup_{\bar{B}_R} |\nabla^2\xi|.$$

Corollary 1. of Lemma 1

$$\int_{\frac{\bar{R}}{2}}^{\bar{R}} dR \sup_{\xi} \frac{\left| \int_{\bar{B}_R} \xi \cdot (dq - \nabla u_R dx) \right|}{\sup_{\bar{B}_R} |\xi| + R^2 \sup_{\bar{B}_R} |\nabla^2\xi|} \lesssim E(B_{\bar{R}}).$$

Corollary 1 expresses closeness in a norm that is weaker than the total variation norm; it is even weaker than the flat norm.

In particular, we cannot take $\xi = I(\hat{B})e$ some some unit vector $e \in \mathbb{R}^d$ and some ball $\hat{B}$. However, we will obtain an estimate as if we had control in the total variation norm, provided we average in the radius r of such a ball $\hat{B}_r$. This follows from a more subtle statement on the boundedness of the Leray projection:

$$\xi_{rR} := \text{Leray projection of } I(\hat{B}_r)e \text{ in } B_R$$

can be (not quite uniquely) written in form of

$$\xi_{rR} = I(\hat{B}_r)\xi_{rR}^{in} + I(B_R)\xi_{rR}^{out},$$

where both $\xi_{rR}^{in/out}$ are smooth provided $\hat{B}_r$ is compactly contained in B_R. This allows us to apply Lemma 1 to

$$\int_{\hat{B}_r} e \cdot (dq - \nabla u_R dx) = \int_{B_R} \xi_{rR} \cdot (dq - (\nabla\psi(x) - x)dx)$$
$$= \int_{B_r} \xi_{rR}^{in} \cdot (dq - (\nabla\psi(x) - x)dx) + \int_{B_R} \xi_{rR}^{out} \cdot (dq - (\nabla\psi(x) - x)dx).$$

In order to quantify smoothness, fix center of $\hat{B}_r \in B_{\frac{\bar{R}}{8}}$; then

$$\mathrm{dist}(\hat{B}_r, B_R^c) \geq \frac{R}{4} \quad \text{as} \quad r \leq \frac{\bar{R}}{8} \text{ and } \frac{\bar{R}}{2} \leq R \leq \bar{R}.$$

By translation invariance, center of $\hat{B}_r$ fixed; by scaling invariance, $r = 1$. Then $\xi_{R,r=1}^{in/out}$ converge as $R \uparrow \infty$; hence smoothness is uniform in R. This (informally) establishes the estimates

$$\left.\begin{array}{c} \max\{\sup_{B_r} |\xi_{rR}^{in}|, r \sup_{B_r} |\nabla \xi_{rR}^{in}|\} \\ \max\{\sup_{B_R} |\xi_{rR}^{out}|, r \sup_{B_R} |\nabla \xi_{rR}^{out}|\} \end{array}\right\} \lesssim 1.$$

Proposition 1.

$$\int_{\frac{\bar{R}}{2}}^{\bar{R}} dR \int_0^{\frac{\bar{R}}{8}} dr \left| \int_{\hat{B}_r} (dq - \nabla u_R dx) \right| \lesssim \bar{R} E(B_{\bar{R}}).$$

II. Optimal semidiscrete matching, heuristics, main result

Matching a law λ to its empirical measure μ

Scaling of mean-square Wasserstein distance $W(\lambda, \mu)$
by Ajtai-Komlòs-Tusnàdy

Approximation by Helmholtz projection, small-scale divergence.

A cut-off on scales $\ll$ particle distance

Implementation by Ambrosio-Stra-Trevisan, on macroscopic scales

Heuristics by Carraciolo-Lucibello-Parisi-Sicuro,
on mesoscopic scales

Comparison of the Parisi-et-al. heuristics to ours

Heuristics made rigorous by Goldman-Huesmann-O.,
on mesoscopic scales

Matching a law λ to its empirical measure μ

Specify to $\lambda(\mathbb{R}^d) = 1$, i. e. to a probability measure.
Given $N \in \mathbb{N}$, draw $Y_1, \cdots, Y_N \in \mathbb{R}^d$ be N independent samples distributed according to λ.
Consider $\mu := \frac{1}{N}\sum_{n=1}^{N} \delta_{Y_n}$, "empirical measure".
The probability measure μ on $\mathbb{R}^d$ is random.

As $N \uparrow \infty$, μ weakly converges to λ, almost surely.
Monitor the (squared) Wasserstein distance
$W^2(\lambda, \mu) := \inf\{ \int |y - x|^2 d\pi \,|\, \pi$ admissible for $\lambda, \mu \}$.
"Semi-discrete matching".

Scaling of mean-square Wasserstein distance $W(\lambda, \mu)$ by Ajtai-Komlòs-Tusnàdy

Simplest case:
$\lambda =$ uniform distribution on a cube Q_L of side length L.
Ignore probability normalization: $\lambda = dx \lfloor Q_L$;
use number density normalization: $N = L^d \in \mathbb{N}$ and $\mu = \sum_{n=1}^{N} \delta_{Y_n}$.

Monitor $\sqrt{\mathbb{E}\frac{1}{N} W^2(\lambda, \mu)}$
$=$ (mean-square) expected transportation distance per point.

Theorem 1 (Ajtai, Komlós, Tusnády '84)**.**

$$\sqrt{\mathbb{E}\frac{1}{N} W^2(\lambda, \mu)} \sim \left\{ \begin{array}{cl} 1 & \textit{for } d > 2, \\ \sqrt{\ln N} & \textit{for } d = 2, \\ \sqrt{N} & \textit{for } d = 1 \end{array} \right\}$$

Hence transportation distance $\ll$ system size ($= L$) for all d,
but transportation distance $\sim$ particle distance ($= 1$) iff $d > 2$.
Hence $d = 2$ is the critical dimension.

Approximation by Helmholtz projection

Consider the distributional Helmholtz projection on Q_L of $\mu - dx$; given by $\nabla u dx \lfloor Q_L$ characterized through

$$\int_{Q_L} \nabla\zeta \cdot \nabla u dx = \int_{Q_L} \zeta(d\mu - dx).$$

Informally, ∇u is solution of Neumann-Poisson problem

$$-\triangle u = \mu - dx \text{ in } Q_L \quad \text{and} \quad \nu \cdot \nabla u = 0 \text{ in } \partial Q_L.$$

By Section 1 $\int_{B\times\mathbb{R}^d}(y-x)d\pi \approx \int_B \nabla u dx$ for most balls $B \subset Q_L$ of (localization) radius $R \gg$ transportation distance $\sim \sqrt{\ln N}$. Ignoring contribution of scales $\lesssim \sqrt{\ln N}$ to macroscopic output naively expect $W^2(\lambda,\mu) = \int |y-x|^2 d\pi \approx \int_{Q_L} |\nabla u|^2 dx$; use in averaged form of $\frac{1}{N}W^2(\lambda,\mu) \approx \frac{1}{|Q_L|}\int_{Q_L} |\nabla u|^2 dx$.

Small scale divergence in $d \geq 2$

However, since points have capacity zero in $d \geq 2$, meaning that Dirac $\delta \notin H^{-1}(Q_L)$, we have $\int_{Q_L} |\nabla u|^2 dx =: \int_{Q_L} ||\nabla|^{-1}(\mu - dx)|^2 dx = +\infty$.

Need to cut off scales $\lesssim \sqrt{\ln N}$; via spectral implementation:
$L^2(Q_L)$-normalized eigenfunctions/-values of Neumann-Laplacian:
$e_m(x) := (\frac{2}{L})^{\frac{d}{2}} \prod_{i=1}^d \cos(\frac{\pi m_i x_i}{L}),\ \lambda_m = (\frac{\pi |m|}{L})^2$ for $m \in \mathbb{N}_0^d - \{0\}$.
Plancherel: $\int_{Q_L} |\nabla u|^2 = \sum_{m \neq 0} \frac{1}{\lambda_m} \Big(\int_{Q_L} e_m(d\mu - dx) \Big)^2$.
Second moments of shot noise $\mu - dx$ as if it were white noise:
$\mathbb{E}(\int_{Q_L} e_m(d\mu - dx))^2 = \int_{Q_L} e_m^2 dx = 1$.
We recover $\mathbb{E} \int_{Q_L} |\nabla u|^2 = \sum_{m \neq 0} (\frac{L}{\pi |m|})^2 = +\infty$ iff $d \geq 2$.

Implementation by Ambrosio-Stra-Trevisan, on macroscopic scale

Define the cut-off version $\nabla\bar{u}$ of ∇u
by removing the wavelengths $\frac{\pi|m|}{L} > \sqrt{\ln N}$,
i. e. by projecting on the wave numbers $|m| \leq \frac{L\sqrt{\ln N}}{\pi}$.

Get $\quad \mathbb{E}\frac{1}{|Q_L|}\int_{Q_L}|\nabla\bar{u}|^2 dx = \sum_{m\in\mathbb{N}_0^2,\, 0<|m|\leq\frac{L\sqrt{\ln N}}{\pi}}(\frac{1}{\pi|m|})^2$

$\approx \frac{1}{4}\frac{2\pi}{\pi^2}\ln\frac{L\sqrt{\ln N}}{\pi} \approx \frac{1}{4\pi}\ln N$ since $L = \sqrt{N}$.

Theorem 2 (Ambrosio, Stra, Trevisan '19)**.** *For* $d = 2$

$$\mathbb{E}\frac{1}{N}W^2(\lambda,\mu) \approx \frac{1}{4\pi}\ln N \quad \textit{for } N \gg 1.$$

Heuristics by Carraciolo-Lucibello-Parisi-Sicuro '14, on mesoscopic level

Recall convex duality from Section 1:
$\exists$ convex $\phi\colon \mathbb{R}^d \to (-\infty, \infty]$ such that supp$\pi \subset \partial\phi$.

Assume momentarily that supp$\mu = Q_L$ and $\mu \ll dy$.
Then $\int_{\mathbb{R}^d \times Q_L} \zeta(x,y) d\pi = \int_{Q_L} \zeta(\nabla\phi(y), y) d\mu$,
by admissibility of π $\int_{Q_L} \zeta(x) dx = \int_{Q_L} \zeta(\nabla\phi(y)) d\mu$.

Assume momentarily $\nabla\phi$ is diffeomorphism of Q_L.
Then $\int_{Q_L} \zeta(\nabla\phi(y)) \det D^2\phi(y) dy = \int_{Q_L} \zeta(\nabla\phi(y)) \frac{d\mu}{dy} dy$.
Get Monge-Ampère equation $\det D^2\phi = \frac{d\mu}{dy}$.

Expect Section 1: $\nabla\phi \approx$ id
when averaged over scales $\gg$ transportation distance.
Writing $\nabla v := \nabla\phi -$ id, naively expect $\det D^2\phi \approx 1 + \mathrm{tr} D^2 v$
when averaged over scales $\gg$ transportation distance.
To leading order, ∇v would be characterized by $\mathrm{tr} D^2 v = \frac{d\mu}{dy} - 1$.

Comparison of the Parisi-et-al. heuristics to ours

Parisi et al.'s heuristics predicts $(x-y)d\pi \approx \nabla v$
when averaged on scales $\gg$ transportation distance
where $\triangle v dy = \mu - dy$.
Our heuristics predicts $(y-x)d\pi \approx q \approx \nabla u$
when averaged on scales $\gg$ transportation distance
where $-\triangle u dx = \mu - dx$.
The two predictions are identical.

The two heuristics reflect the two faces of the Laplace operator:
divergence-form $-\nabla \cdot \nabla u$ vs. non-divergence form tr$D^2 v$.

Also reflect the two faces of the Monge-Ampère equation:
Euler-Lagrange equation of inf$\{\int |y-x|^2 d\pi \,|\, \pi$admissible$\}$
vs. fully non-linear equation det$D^2\phi = \frac{d\mu}{dy}$ (Caffarelli, Figalli).

Heuristics made rigorous by Goldman-Huesmann-O.

Mesoscopic vs. macroscopic closeness of $(y-x)d\pi$ to ∇u.
"Mesoscopic" means $\gg$ particle distance,
"macroscopic" means $\ll$ system size.

As before, $\lambda = dx\lfloor Q_L$. More generally as before:
(deterministic) μ with supp$\mu \subset Q_L$ and $\mu(Q_L) = |Q_L|$

Monitor closeness of μ to Lebesgue measure dy on set $B \subset \mathbb{R}^d$:

$$D(B) := W^2(\mu\lfloor B, \frac{\mu(B)}{|B|}dy\lfloor B) + |B|R^2(\frac{\mu(B)}{|B|} - 1)^2.$$

Back to $\mu = \sum_{n=1}^{N} \delta_{Y_n}$ with $N = L^d$ and Y_n's indep. and uniform in Q_L.
Consider all concentric balls $B = B_R \subset Q_L$, center = origin. Have
for $d = 2$ by Ajtai et al. with overwhelming probability

$$D(Q_L) \sim |Q_L| \ln L \text{ and } D(B_R) \sim |B_R| \ln R \text{ for } 1 \ll R.$$

Convenient to introduce a rate functional $T\colon [0,\infty) \to [0,\infty)$, called sublinear iff $\forall 0 < \theta \ll 1$ $\forall R$ $\theta T(R) \ll T(\theta R)$.

Recall: π is the minimizer of $\{\int |y-x|^2 d\pi \,|\, \pi$ admissible for $\lambda, \mu\}$, $\nabla u dx \lfloor Q_L$ is distributional Helmholtz projection on Q_L of $\mu - dx\lfloor Q_L$.

Theorem 3 (Goldman, Huesmann, O. '21)**.** *If μ is such that there exists a sublinear rate function T and a radius $\bar{r} \le \frac{L}{4}$ with*

$$D(Q_L) \le |Q_L| T^2(L) \ \textit{ and } \ D(B) \le |B_R| T^2(R) \ \textit{ for } \bar{r} \le R \le \frac{L}{4}$$

then π and ∇u are related by

$$\frac{1}{\bar{r}} \int_{\frac{\bar{r}}{2}}^{\bar{r}} dr \Big| \frac{1}{\pi(\mathbb{R}^d \times B_r)} \int_{\mathbb{R}^d \times B_r} (y-x) d\pi - \frac{1}{|B_r|} \int_{B_r} \nabla u dx \Big| \lesssim \int_{\bar{r}}^{L} \frac{dr}{r} \frac{T^2(r)}{r}.$$

Theorem 4 (Goldman, Huesmann, O. '21). *If μ is such that there exists a sublinear rate function T and a radius $\bar{r} \leq \frac{L}{4}$ with*

$$D(Q_L) \leq |Q_L| T^2(L) \ \text{ and } \ D(B) \leq |B_R| T^2(R) \ \text{ for } \bar{r} \leq R \leq \frac{L}{4}$$

then π and ∇u are related by

$$\frac{1}{\bar{r}} \int_{\frac{\bar{r}}{2}}^{\bar{r}} dr \Big| \frac{1}{\pi(\mathbb{R}^d \times B_r)} \int_{\mathbb{R}^d \times B_r} (y-x) d\pi - \frac{1}{|B_r|} \int_{B_r} \nabla u dx \Big| \lesssim \int_{\bar{r}}^{L} \frac{dr}{r} \frac{T^2(r)}{r}.$$

Displacement $y - x$ π-averaged over $\mathbb{R}^d \times B_r$

$\approx$ electrostatic field ∇u dx-averaged over $x \in B_r$.

R. h. s. is average over radii $r \sim \bar{r}$, cf. Proposition 1.

For empirical measure in $d = 2$:

$\bar{r} \sim 1$ and $T(r) \sim \sqrt{\ln r}$ (which clearly is sublinear),

so that l. h. s. $\lesssim 1 =$ particle distance.

Hence on scales $\sim$ particle distance, have

displacement $=$ electrostatic field $+ O$(particle distance).

Best possible closeness.

III. Proof of Theorem 4

Most involved ingredient: a large-scale regularity result for π expressed in terms of shifts b

Relating a shift b to an average flux increment $d(q-\widehat{q})$:
Lemma 1 $\Longrightarrow$ Lemma 2

Optimal shifts b_k at geometrically decreasing radii $\bar{R}_k$

Fluxes q_k; relate shift increments $b_k - b_{k-1}$
to flux increments $q_{k-1} - q_k$ by Lemma 2

Fields ∇u_k; relate fluxes q_k to fields ∇u_k by Proposition 1

Usage of mean-value property to rewrite
field increments $\nabla u_{k-1} - \nabla u_k$ amenable to telescoping

Telescoping to relate shift b_K to field increment $\nabla u - \nabla u_K$

Last meter: relate field ∇u_K to displacement $(y-x)d\pi - b_K$

Most involved ingredient: a large-scale regularity result for π expressed in terms of shifts b

To every shift vector $b \in \mathbb{R}^d$ we associate the plan π_b

$$\pi_b \quad \text{is push forward of} \quad \pi \quad \text{under} \quad (\hat{x}, \hat{y}) = (x + b, y).$$

On level of trajectories $X =$ shear of the graph $\{(t, X_t)\}$.

Involved ingredient for Theorem 4 = large-scale regularity result:
For sublinear rate T, bound on μ transmit to π modulo shift:

Theorem 5 (GHO '21). *If $\exists$ sublinear T and a radius $\bar{r} \le \frac{L}{4}$ with*

$$D(Q_L) \le |Q_L| T^2(L) \;\; \textit{and} \;\; D(B) \le |B_R| T^2(R) \;\; \textit{for} \; \bar{r} \le R \le \frac{L}{4}$$

then

$$\inf_{b \in \mathbb{R}^d} \int_{\Omega(B_R)} |\hat{y} - \hat{x}|^2 d\pi_b \lesssim |B_R| T^2(R) \quad \textit{for} \quad \bar{r} \le R \le \frac{L}{4}.$$

Relating a shift b to an average flux increment $d(q-\widehat{q})$: Lemma 1 $\Longrightarrow$ Lemma 2

Recall Lagrangian displacement $(y-x)d\pi \approx$ Eulerian flux q.
Lemma 1 was established for general λ and μ.
Hence we may exchange the roles of x and y.
Both displacement and flux change sign under, so that

$$\int_0^{\bar{r}} dr \Big| \int_{B_r} dq - \int_{\mathbb{R}^d \times B_r} (y-x) d\pi \Big| \leq E(B_{\bar{r}}).$$

Assume that $\widehat{\pi}$ is related to original π by a shift $b \in \mathbb{R}^d$:

$$\widehat{\pi} \text{ is push forward of } \pi \text{ under } (\widehat{x}, \widehat{y}) = (x+b, y).$$

Assume that we control both $\widehat{E}(B_{\bar{r}})$ and $E(B_{\bar{r}})$; analogously

$$\int_0^{\bar{r}} dr \Big| \int_{B_r} d\widehat{q} - \int_{\mathbb{R}^d \times B_r} (\widehat{y}-\widehat{x}) d\widehat{\pi} \Big| \leq \widehat{E}(B_{\bar{r}}).$$

By the triangle inequality

$$\int_0^{\bar r} dr \Big| \int_{B_r} d(q-\widehat{q}) - \int_{\mathbb{R}^d \times B_r} b d\pi \Big| \le (E+\hat{E})(B_{\bar r}).$$

By the admissibility of π

$$\int_0^{\bar r} dr \Big| \int_{B_r} d(q-\widehat{q}) - \mu(B_r) b \Big| \le (E+\hat{E})(B_{\bar r}).$$

By $\mu(B_r) \approx |B_r|$ informally

$$b \approx \frac{1}{|B_r|} \int_{B_r} d(q-\widehat{q}) \quad \text{on average in } r \le \bar r$$

in the rigorous sense of

Lemma 2. *Provided* $D(B_{\bar r}) \ll \bar r^{d+2}$,

$$\int_0^{\bar r} dr \Big| |B_r| b - \int_{B_r} d(q-\widehat{q}) \Big| \lesssim (E+\hat{E})(B_{\bar r}).$$

Optimal shifts b_k at geometrically decreasing radii $\bar{R}_k$

Fix (sufficiently spaced) geometrically decreasing radii $\{\bar{R}_k\}_{k=1,\cdots,K}$ connecting macroscopic $\bar{R}_1 \sim L$ to microscopic $\bar{R}_K = \bar{r}$.

Theorem 5: $\forall\, k \in \{1, \cdots, K\}\ \exists\, b_k \in \mathbb{R}^d$ s. t.

$$\pi_k := \pi_{b_k} \quad \text{satisfies} \quad E_k := \int_{\Omega(B_k)} |y-x|^2 d\pi_k \lesssim |B_k| T^2(R_k).$$

Convenient to extend these definitions to $k = 0$ by setting

$$\bar{R}_0 := L,\ \ B_0 := Q_L,\ \ b_0 = 0,\ \ \pi_0 := \pi,\ \ E_0 := \int |y-x|^2 d\pi.$$

Trivially have the analogous estimate

$$E_0 = W^2(\lambda, \mu) \le D(Q_L) \le |Q_L| T^2(L) = |B_0| T^2(\bar{R}_0).$$

Fluxes q_k; relate shift increments $b_k - b_{k-1}$ to flux increments $q_{k-1} - q_k$ by Lemma 2

For $k = 0, \cdots, K$, associate fluxes q_k to plans π_k:

$$\int \xi \cdot dq_k = \int \xi(x) \cdot (y - x) d\pi_k.$$

For $k = 1, \cdots, K$, shift increments $\approx$ averages of flux increments:

$$b_k - b_{k-1} \approx \frac{1}{|B_r|} \int_{B_r} d(q_{k-1} - q_k) \quad \text{on average in } r \leq \bar{R}_k.$$

By this informal statement we mean the estimate

$$\frac{2}{\bar{R}_k} \int_{\frac{\bar{R}_k}{2}}^{\bar{R}_k} dr \Big| (b_k - b_{k-1}) - \frac{1}{|B_r|} \int_{B_r} d(q_{k-1} - q_k) \Big| \lesssim \frac{T^2(\bar{R}_{k-1})}{R_{k-1}} + \frac{T^2(\bar{R}_k)}{\bar{R}_k}.$$

This relies on definition of π_k in incremental version of

π_k is push forward of π_{k-1} under $(\hat{x}, \hat{y}) = (x + b_k - b_{k-1}, y)$,

to which we apply Lemma 2; estimate of r. h. s. by Theorem 5.

Fields ∇u_k; relate fluxes q_k to fields ∇u_k by Proposition 1

For any $k = 1, \cdots, K$, introduce electrostatic fields $\nabla u_{k,r}$:

$\nabla u_{k,r} dx \lfloor B_r :=$ distributional Helmholtz projection on $\bar{B}_r$ of $q_k \lfloor \bar{B}_r$.

Automatically,

$$\int_{B_r} dq_k = \int_{B_r} \nabla u_{k,r} dx \quad \text{for a. e. } \; r \le \bar{R}_k.$$

Claim separately for $k = 2, \cdots, K$ and $k = 1$

$$\int_{B_r} dq_{k-1} \approx \int_{B_r} \nabla u_{k-1,R} dx \quad \text{on average in } r \le \bar{R}_k, \; R \le \bar{R}_{k-1},$$

$$\int_{B_r} dq_0 \approx \int_{B_r} \nabla u dx \quad \text{on average in } r \le \bar{R}_1,$$

where we recall

$$\nabla u dx \lfloor Q_L := \text{ distributional Helmholtz projection on } Q_L \text{ of } q_0.$$

Again, the informal

$$\int_{B_r} dq_{k-1} \approx \int_{B_r} \nabla u_{k-1,R} dx \quad \text{on average in } r \leq \bar{R}_k,\ R \leq \bar{R}_{k-1},$$
$$\int_{B_r} dq_0 \approx \int_{B_r} \nabla u dx \quad \text{on average in } r \leq \bar{R}_1.$$

are supposed to mean

$$\frac{2}{\bar{R}_{k-1}} \int_{\frac{\bar{R}_{k-1}}{2}}^{\bar{R}_{k-1}} dR \frac{2}{\bar{R}_k} \int_{\frac{R_k}{2}}^{\bar{R}_k} dr \Big| \frac{1}{|B_r|} \int_{B_r} (dq_{k-1} - \nabla u_{k-1,R} dx) \Big| \lesssim \frac{T^2(\bar{R}_{k-1})}{\bar{R}_{k-1}},$$
$$\frac{2}{\bar{R}_1} \int_{\frac{\bar{R}_1}{2}}^{\bar{R}_1} dr \Big| \frac{1}{|B_r|} \int_{B_r} (dq_0 - \nabla u dx) \Big| \lesssim \frac{T^2(\bar{R}_0)}{\bar{R}_0}.$$

Usage of mean-value property to rewrite field increments $\nabla u_{k-1} - \nabla u_k$ amenable to telescoping

Crucial identity for telescoping:

$$\begin{aligned}
&\frac{1}{|B_r|}\int_{B_r}(\nabla u_{k-1,R} - \nabla u_{k,r})dx \\
&= \frac{1}{|B_r|}\int_{B_r}(\nabla u - \nabla u_{k,r})dx - \frac{1}{|B_r|}\int_{B_r}(\nabla u - \nabla u_{k-1,R})dx \\
&\overset{!}{=} \frac{1}{|B_r|}\int_{B_r}(\nabla u - \nabla u_{k,r})dx - \frac{1}{|B_R|}\int_{B_R}(\nabla u - \nabla u_{k-1,R})dx
\end{aligned}$$

Definitions of ∇u and $\nabla u_{k-1,R}$

$\Longrightarrow$ $\nabla u - \nabla u_{k-1,R}$ is (distributionally) divergence-free in B_R.

Hence $\nabla u - \nabla u_{k-1,R}$ is (componentwise) harmonic

$\Longrightarrow$ satisfies mean-value property:

$$\frac{1}{|B_r|}\int_{B_r}(\nabla u - \nabla u_{k-1,R})dx = \frac{1}{|B_R|}\int_{B_R}(\nabla u - \nabla u_{k-1,R})dx.$$

Telescoping to relate shift b_K to field increment $\nabla u - \nabla u_K$

Claim

$$b_K \approx \frac{1}{|B_r|}\int_{B_r}(\nabla u - \nabla u_{K,r})dx \quad \text{on average in } r \le R_K$$

$$\text{in sense of} \quad \frac{2}{\bar r}\int_{\frac{\bar r}{2}}^{\bar r} dr\Big|b_K - \frac{1}{|B_r|}\int_{B_r}(\nabla u - \nabla u_{K,r})dx\Big| \lesssim \sum_{k=0}^{K}\frac{T^2(\bar R_k)}{\bar R_k}.$$

Insert the two flux-field relations

$$\int_{B_r} dq_k = \int_{B_r}\nabla u_{k,r}dx \quad \text{and} \quad \int_{B_r} dq_{k-1} \approx \int_{B_r}\nabla u_{k-1,R}dx$$

(all on average in $r \le \bar R_k,\ R \le \bar R_{k-1}$) into incremental shift-flux relation

$$b_k - b_{k-1} \approx \frac{1}{|B_r|}\int_{B_r} d(q_{k-1} - q_k),$$

and use mean-field identity to obtain telescoping relation

$$b_k - b_{k-1} \approx \frac{1}{|B_r|}\int_{B_r}(\nabla u - \nabla u_{k,r})dx - \frac{1}{|B_R|}\int_{B_R}(\nabla u - \nabla u_{k-1,R})dx.$$

Telescoping relation holds for $k = 2, \cdots, K$.
For $k = 1$, appeal instead to field-flux relation

$$\int_{B_r} dq_0 \approx \int_{B_r} \nabla u dx \quad \text{on average in } r \leq \bar{R}_1,$$

and to $b_0 = 0$ to get

$$b_1 \approx \frac{1}{|B_r|} \int_{B_r} (\nabla u - \nabla u_{1,r}) dx \quad \text{on average in } r \leq \bar{R}_1.$$

Insert this into sum over telescoping relation

$$b_k - b_{k-1} \approx \frac{1}{|B_r|} \int_{B_r} (\nabla u - \nabla u_{k,r}) dx - \frac{1}{|B_R|} \int_{B_R} (\nabla u - \nabla u_{k-1,R}) dx$$
$$\text{on average in} \quad r \leq \bar{R}_k, \; R \leq \bar{R}_{k-1}$$

to get desired

$$b_K \approx \frac{1}{|B_r|} \int_{B_r} (\nabla u - \nabla u_{K,r}) dx \quad \text{on average in } r \leq R_K$$

Last meter: relate field ∇u_K to displacement $(y-x)d\pi - b_K$

Remains to show the displacement-field relation on the "last meter"

$$\frac{1}{\pi(\mathbb{R}^d \times B_r)}\int_{\mathbb{R}^d\times B_r}(y-x)d\pi - b_K \approx \frac{1}{|B_r|}\int_{B_r}\nabla u_{K,r}dx \quad \text{on average in } r \le R_K$$

since together with the previously established shift-field relation

$$b_K \approx \frac{1}{|B_r|}\int_{B_r}(\nabla u - \nabla u_{K,r})dx \quad \text{on average in } r \le R_K$$

this implies the desired displacement-field relation

$$\frac{1}{\pi(\mathbb{R}^d \times B_r)}\int_{\mathbb{R}^d\times B_r}(y-x)d\pi \approx \frac{1}{|B_r|}\int_{B_r}\nabla u dx \quad \text{on average in } r \le R_K.$$

This is informal for the statement in Theorem 4

$$\frac{2}{\bar r}\int_{\frac{\bar r}{2}}^{\bar r} dr\Big|\frac{1}{\pi(\mathbb{R}^d \times B_r)}\int_{\mathbb{R}^d\times B_r}(y-x)d\pi - \frac{1}{|B_r|}\int_{B_r}\nabla u dx\Big| \lesssim \sum_{k=0}^{K}\frac{T^2(\bar R_k)}{\bar R_k}.$$

By definition of $\widehat{\pi} := \pi_K$,

$$\frac{1}{\pi(\mathbb{R}^d \times B_r)} \int_{\mathbb{R}^d \times B_r} (y-x)d\pi + b_K = \frac{1}{\pi_K(\mathbb{R}^d \times B_r)} \int_{\mathbb{R}^d \times B_r} (\widehat{y}-\widehat{x})d\widehat{\pi}.$$

Lemma 1: Lagrangian displacements $\approx$ Eulerian fluxes, integral of fluxes = integral of fields $\nabla \widehat{u}_r := \nabla u_{K,r}$:

$$\int_0^{\bar{r}} dr \Big| \int_{\mathbb{R}^d \times B_r} (\widehat{y}-\widehat{x})d\widehat{\pi} - \int_{B_r} \nabla \widehat{u}_r \Big| \le \widehat{E}(B_{\bar{r}}).$$

Need to divide the first integrand contribution by $\pi_K(\mathbb{R}^d \times B_r)$ and the second one by $|B_r|$.

This relies on

$$r|\widehat{\pi}(\mathbb{R}^d \times B_r) - |B_r|| \lessapprox \sqrt{|B_{\bar{r}}|D(B_{\bar{r}})} \quad \text{on average in } r \le \bar{r}$$

and

$$\Big| \int_{\mathbb{R}^d \times B_r} (\widehat{y}-\widehat{x})d\widehat{\pi} \Big| \le \int_{\Omega(B_r)} |\widehat{y}-\widehat{x}|d\widehat{\pi} \lessapprox \sqrt{|B_{\bar{r}}|\widehat{E}(B_{\bar{r}})} \quad \text{on average in } r \le \bar{r}.$$