Last time:

$$\mathbb{E} W_p(\mu_n, \mu) \gtrsim \begin{cases} n^{-1/2} & \text{always.} \\ n^{-1/2p} & \text{if } supp(\mu) \text{ is disconnected} \\ n^{-1/d} & \text{if } \mu \text{ has a bounded Lebesgue density.} \end{cases}$$

Theorem:

$$\sup_{\mu \in \mathcal{P}(\Omega)} \mathbb{E} W_p(\mu_n, \mu) \lesssim \begin{cases} n^{-\frac{1}{2p}} & ; \; d < 2p \\ n^{-\frac{1}{2p}} (\log n)^{1/p} & ; \; d = 2p \\ n^{-1/d} & ; \; d > 2p \end{cases}.$$

Lecture 2.

- Let $Q_j = \{$Natural partition of Ω consisting of cubes of side-length $2^{-j}\}$

- **Proposition** ("**Dyadic Upper Bound**"): For all $p \geq 1$, $J \geq 0$, $\mu, \nu \in \mathcal{P}(\Omega)$,

$$W_p^p(\mu, \nu) \leq \frac{d^{p/2}}{2^{Jp}} + d^{p/2} \sum_{j=0}^{J-1} 2^{-jp} \sum_{S \in Q_{j+1}} |\mu(S) - \nu(S)|$$

Proof Sketch:

Idea: Decompose $\mu = \sum_{j=0}^{J} \mu_j$ & $\nu = \sum_{j=0}^{J} \nu_j$ recursively.

$$\mu_j(\Omega) = \nu_j(\Omega) \quad \forall j$$

Step J: Define $\forall S \in Q_J$:

$$\mu_J|_S = \left(\frac{\mu(S) \wedge \nu(S)}{\mu(S)}\right) \mu|_S$$

$$\nu_J|_S = \left(\frac{\mu(S) \wedge \nu(S)}{\nu(S)}\right) \nu|_S .$$

Steps $j = J-1, \ldots, 1$: Set:

$$\mu_j' = \mu - \sum_{k=j+1}^{J} \mu_k \qquad \nu_j' = \nu - \sum_{k=j+1}^{J} \nu_k .$$

$\forall S \in Q_j$: $$\mu_j|_S = \left(\frac{\mu_j'(S) \wedge \nu_j'(S)}{\mu_j'(S)}\right) \mu_j'|_S$$

$$\nu_j|_S = \left(\frac{\mu_j'(S) \wedge \nu_j'(S)}{\nu_j'(S)}\right) \nu_j'|_S .$$

Step 0: Set $\mu_0 = \mu - \sum_{j=1}^{J} \mu_j$ & $\nu_0 = \nu - \sum_{j=1}^{J} \nu_j$.

Some Facts:

1) $\mu_j, \nu_j \geq 0 \quad \forall j$

2) $\mu_j(S) = \nu_j(S) \quad \forall S \in Q_j, \; j = 0, \ldots, J.$

3) $\mu_j(\Omega) \leq \sum_{S \in Q_{j+1}} |\mu(S) - \nu(S)| \qquad j = 0, \ldots, J-1$

Now, let $\gamma_j \in \Pi(\mu_j, \nu_j)$ be any coupling such that:

$$\gamma_j(S \times R) = 0 \qquad \forall S \neq R, \quad S, R \in Q_j.$$

$$\Longrightarrow \gamma \in \Pi(\mu, \nu) \qquad \gamma = \sum_{j=0}^{J} \gamma_j$$

$$
\begin{aligned}
W_p^p(\mu, \nu) &\leq \int \|x-y\|^p \, d\gamma(x,y). \\
&= \sum_{j=0}^{J} \int \|x-y\|^p \, d\gamma_j(x,y). \\
&\leq \sum_{j=0}^{J} \int (\sqrt{d}\, 2^{-j})^p \, d\gamma_j(x,y) \\
&= \sum_{j=0}^{J} (\sqrt{d}\, 2^{-j})^p \mu_j(\Omega) \\
&\leq (\sqrt{d}\, 2^{-J})^p + \sum_{j=0}^{J-1} (\sqrt{d}\, 2^{-j})^p \sum_{S \in Q_{j+1}} |\mu(S) - \nu(S)|.
\end{aligned}
$$

$\square$

Proof of Theorem.

$$\mathbb{E}\left[W_p^p(\mu_n,\mu)\right]$$

$$\mu_n(S) = \frac{1}{n}\sum_{i=1}^n I(X_i \in S).$$

$$\lesssim 2^{-Jp} + \sum_{j=0}^{J-1} 2^{-jp} \sum_{S\in Q_{j+1}} \mathbb{E}|\mu_n(S)-\mu(S)|$$

$$\leq 2^{-Jp} + \sum_{j=0}^{J-1} 2^{-jp} \sum_{S\in Q_{j+1}} \sqrt{\underbrace{\mathrm{Var}[\mu_n(S)]}}$$

$$= \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}[I(X_i\in S)]$$

$$\leq \frac{1}{n}\mathbb{E}\left[I(X_1\in S)\right]$$

$$= \frac{1}{n}\mathbb{P}(X_1\in S) = \frac{\mu(S)}{n}.$$

$$\leq 2^{-Jp} + \frac{1}{\sqrt{n}}\sum_{j=0}^{J-1} 2^{-jp} \underbrace{\sum_{S\in Q_{j+1}} \sqrt{\mu(S)}}_{\|v\|_1}. \qquad v = \left(\sqrt{\mu(S)}\right)_{S\in Q_j}$$

$$\|v\|_2 = 1$$

$$\leq \sqrt{|Q_{j+1}|}\cdot\|v\|_2 = \sqrt{|Q_{j+1}|}$$

$$\leq 2^{-Jp} + \frac{1}{\sqrt{n}}\sum_{j=0}^{J-1} 2^{-jp}\, 2^{jd/2}$$

$$= 2^{-Jp} + \frac{1}{\sqrt{n}}\sum_{j=0}^{J-1} 2^{j\left(\frac{d}{2}-p\right)} \lesssim 2^{-Jp} + \frac{1}{\sqrt{n}}\begin{cases} 2^{J\left(\frac{d}{2}-p\right)} & ;\ \frac{d}{2} > p \\ J & ;\ d/2 = p. \\ 1 & ;\ \frac{d}{2} < p \ \square \end{cases}$$

We have proved:

$$\sup_{\mu \in \mathcal{P}(\Omega)} \mathbb{E} W_p(\mu_n, \mu) \lesssim \begin{cases} n^{-\frac{1}{2p}} & ; d < 2p \quad (A) \\ n^{-\frac{1}{2p}} (\log n)^{1/p} & ; d = 2p \\ n^{-1/d} & ; d > 2p \quad (B) \end{cases}$$

<u>Remarks</u>:

(B): Can be improved to "$n^{-\frac{1}{k+\varepsilon}}$" if μ is intrinsically "k-dimensional".

$\longrightarrow$ Weed-Bach (Bernoulli '19)

(A): Ledoux-Zhu '19, get $n^{-1/2}$ upper bound if the support of μ is connected.

<u>See Also</u>:

- Bobkov-Ledoux '19 Monograph — Sharp treatment of $d=1$ case.

- AKT Theorem, 1984

  ↳ <u>Modern proof</u>: Bobkov-Ledoux (AoAP, 2021)

  ↳ <u>Refinements</u>: Ambrosio, Stra, Trevisan (PTRF, 2019)

Sharp log factors when $d=2$ and measures have connected support.

II. Estimation of Wasserstein Distances.

[Chizat et al. '20]

- **Theorem**: For all $d \geqslant 5$, if $W_2(\mu, \nu) \geq \delta > 0$, then:

$$\mathbb{E}\left| W_2(\mu_n, \nu_n) - W_2(\mu, \nu) \right| \lesssim n^{-2/d}.$$

First, some reductions :

1. Simplification: Work with $\mathbb{E}\left| W_2(\mu, \nu_n) - W_2(\mu, \nu) \right|$.

2. Note that when $W_2(\mu, \nu) \geq \delta$,

$$\left| W_2(\mu, \nu_n) - W_2(\mu, \nu) \right| \asymp \left| W_2^2(\mu, \nu_n) - W_2^2(\mu, \nu) \right|.$$

3. Bias-variance Decomposition:

$$\mathbb{E}\left| W_2^2(\mu, \nu_n) - W_2^2(\mu, \nu) \right|$$

$$\leq \underbrace{\mathbb{E}\left| W_2^2(\mu, \nu_n) - \mathbb{E} W_2^2(\mu, \nu_n) \right|}_{\leq \sqrt{\mathrm{Var}[W_2^2(\mu, \nu_n)]}} + \underbrace{\left| \mathbb{E} W_2^2(\mu, \nu_n) - W_2^2(\mu, \nu) \right|}_{\text{Bias}}.$$

4. Variance:

Lemma: $\mathrm{Var}[W_2^2(\mu, \nu_n)] \lesssim \frac{1}{n}$.

↳ Next time.

(5) Bias.

Lemma: $\mathbb{E} W_2^2(\mu, \nu_n) \geq W_2^2(\mu, \nu)$. $\rightarrow$ upward bias!

Proof: $\mathbb{E}\left[\sup_{f,g} \int f \, d\mu + \int g \, d\nu_n\right]$

$$\geq \sup_{f,g} \mathbb{E}\left[\int f \, d\mu + \int g \, d\nu_n\right]$$

$$= \sup_{f,g} \int f \, d\mu + \int g \, d\nu = W_2^2(\mu, \nu). \quad \square$$