Wasserstein metric between a discrete probability measure and a continuous one

Authors: Yang Weihua, Wang Xia, Yang Weihua, Wang Xia

Date: 2021-12-16T00:00:00+00:00

Abstract

This paper considers the Wasserstein metric between the empirical probability measure of n discrete random variables and a continuous uniform distribution on the d-dimensional ball, and provides an asymptotic estimate of its expectation as $n \to \infty$. Furthermore, we consider the aforementioned problem for a mixed process, i.e., n discrete random variables are generated by a Poisson process.

Full Text

Preamble

Wasserstein Metric Between a Discrete Probability Measure and a Continuous One

Weihua Yang and Xia Wang
December 16, 2021

Abstract

This paper investigates the Wasserstein metric between the empirical probability measure of n discrete random variables and a continuous uniform measure on the d-dimensional ball, providing asymptotic estimates for their expectation as $n \to \infty$. Furthermore, we consider the aforementioned problem for a mixed process, where the n discrete random variables are generated by a Poisson process.

Keywords: Wasserstein metric; optimal matching; from discrete to continuous; random variable; Poisson process.

1 Introduction

Paper [1] studied the Ollivier curvature of random geometric graphs, where a key step involves establishing estimates for the Wasserstein metric between the empirical probability measure of n discrete random variables and a continuous uniform measure on the d-dimensional ball. The authors applied results from [5], but those results were established on $[0,1]^d$, whereas Ollivier curvature is defined on balls. Consequently, the proof approach in [1] appears both tedious and problematic. On the other hand, lattice methods in statistical mechanical approaches [6] often involve similar notation and convergence from discrete physical quantities to continuous ones, which should have close connections with the convergence from discrete probability to continuous probability. These considerations motivate us to study the approximation problem between discrete and continuous probability measures. In this article, we select the Wasserstein metric as our measure of approximation, through which we hope to build a bridge for investigating the relationship between discrete and continuous quantities in broader contexts.

2 Preliminary Estimation

Definition 1. Let $X_1, X_2, \cdots, X_n, Y_1, Y_2, \cdots, Y_n$ be independent uniformly distributed random variables on the d-dimensional ball $B(0;1) = \{x \in \mathbb{R}^d, \|x\| \leq 1\}$, where $d \geq 2$ and $\|\cdot\|$ is a norm on $\mathbb{R}^d$. The random variable $M_n^d = \inf_\sigma \sum_{i=1}^n \|X_i - Y_{\sigma(i)}\|$ denotes the optimal matching between $X_1, X_2, \cdots, X_n$ and $Y_1, Y_2, \cdots, Y_n$, where σ ranges over all permutations of $\{1, 2, \cdots, n\}$. By the dual principle [2, 3], we have $M_n^d = \sup_{f\in\mathcal{L}1} \sum_{i=1}^n (f(X_i) - f(Y_i)) = \sup_{f\in\mathcal{L}1} \left|\sum_{i=1}^n (f(X_i) - f(Y_i))\right|$, where the set of Lipschitz functions is $\mathcal{L}_1 = \{f : B(0;1) \to \mathbb{R}; |f(x) - f(y)| \leq \|x - y\|, \forall x, y \in B(0;1), f(0) = 0\}$. Note that every Lipschitz function in $\mathcal{L}_1$ can be extended to a function in $\mathcal{L} = \{f : \mathbb{R}^d \to \mathbb{R}; |f(x) - f(y)| \leq \|x - y\|, \forall x, y \in \mathbb{R}^d, f(0) = 0, \|f\|_{L^\infty} \leq 1\}$, so $\mathcal{L}_1 = \mathcal{L}|_{B(0,1)}$. The following Lemma 1 provides upper and lower bound estimates for the expectation $\mathbb{E}(M_n^d)$.

Lemma 1 (Optimal Matching). For the above optimal matching problem, we have for dimension $d \geq 2$:

$$\frac{1}{d+1} \leq \liminf_{n\to\infty} \mathbb{E}(M_n^d) n^{1-\frac{1}{d}} \leq \limsup_{n\to\infty} \mathbb{E}(M_n^d) n^{1-\frac{1}{d}} \leq 2d + 5 + 8.$$

Proof. We place the proof in the appendix. The method is essentially from [5], with improvements and modifications to extend it to random variables on balls.

3 Main Results and Proofs

The following results concern the Wasserstein metric between empirical and uniform measures on the d-dimensional ball. In general, the Wasserstein metric between two probability measures μ_1, μ_2 is given by the following definition.

Definition 2. Let μ_1 and μ_2 be Borel probability measures on a compact metric space $(\mathcal{X}, d)$ and let $\Gamma(\mu_1, \mu_2)$ denote the set of joint probability measures μ on $\mathcal{X} \times \mathcal{X}$ with marginals μ_1 and μ_2 respectively. The Wasserstein metric is defined by

$$W(\mu_1, \mu_2) = \inf_{\mu\in\Gamma(\mu_1,\mu_2)} \int_{\mathcal{X}\times\mathcal{X}} d(x, y)\, d\mu(x, y).$$

By the duality principle (Kantorovich Dual Theorem) [7], the Wasserstein metric can be expressed as

$$W(\mu_1, \mu_2) = \sup_{f \in \mathcal{L}1(\mathcal{X})} \left(\int_{\mathcal{X}} f(x)\, d\mu_1(x) - \int_{\mathcal{X}} f(y)\, d\mu_2(y) \right),$$

where $\mathcal{L}_1(\mathcal{X})$ denotes the set of 1-Lipschitz functions on $\mathcal{X}$ with respect to the metric d. From the above duality formula, we may further assume that any $f \in \mathcal{L}_1(\mathcal{X})$ satisfies $f(0) = 0$.

Note: In what follows, all metrics are induced by some norm. To maintain consistent mathematical notation with most original articles, we still use 'd' to denote the metric in a space, which should not cause confusion with the dimension notation.

Theorem 1. Let $X_1, X_2, \cdots, X_n$ be independent uniformly distributed random variables on the d-dimensional ball $B(0;1)$, let m_n^d denote the empirical measure $m_n^d(y) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{y=X_i}$ and μ_d the uniform measure on $B(0;1)$. Then as $n \to \infty$,

$$\mathbb{E}[W_d(m_n^d, \mu_d)] = O(n^{-\frac{1}{d}}), \quad d \geq 2.$$

Proof. By the definition of the Wasserstein metric and the Kantorovich duality theorem,

$$W_d(m_n^d, \mu_d) = \inf_{\mu \in \Gamma(m_n^d, \mu_d)} \int_{B(0;1)\times B(0;1)} d(x,y)\, d\mu(x,y) = \sup_{f \in \mathcal{L}1(B(0;1))} \left(\int_{B(0;1)} f(x)\, dm_n^d(x) - \int_{B(0;1)} f(y)\, d\mu_d(y) \right)$$

Let $Y_1, Y_2, \cdots, Y_n$ be independent uniformly distributed random variables on $B(0;1)$. Then

$$\int_{B(0;1)} f(y)\, d\mu_d(y) = \mathbb{E}[f(Y_i)], \quad i = 1, \cdots, n.$$

Therefore,

$$W_d(m_n^d, \mu_d) = \sup_{f \in \mathcal{L}1(B(0;1))} \left(\frac{1}{n} \sum_{i=1}^n f(X_i) - \mathbb{E}[f(Y_i)] \right) = \sup_{f \in \mathcal{L}1(B(0;1))} \frac{1}{n} \sum_{i=1}^n (f(X_i) - \mathbb{E}[f(Y_i)|X_i]).$$

Conditioning on $X = (X_1, \cdots, X_n)$, we have

$$W_d(m_n^d, \mu_d) = \sup_{f \in \mathcal{L}1(B(0;1))} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^n (f(X_i) - f(Y_i)) \middle| X \right] \leq \frac{1}{n} \mathbb{E}\left[M_n^d \middle| X \right].$$

Hence, from Lemma 1, we obtain

$$\mathbb{E}[W_d(m_n^d, \mu_d)] \leq \frac{1}{n} \mathbb{E}[M_n^d] = O(n^{-\frac{1}{d}}).$$

Corollary 1. More generally, let $X_1, X_2, \cdots, X_n$ be independent uniformly distributed random variables on the d-dimensional ball $B(0,\delta)$ with radius $\delta > 0$, and let m_n^d denote the empirical measure and μ_d the uniform measure on $B(0,\delta)$. Then

$$\mathbb{E}[W_d(m_n^d, \mu_d)] = \delta O(n^{-\frac{1}{d}}), \quad d \geq 2.$$

Proof. Let $y = \delta t$ and define $\bar{m}_n^d(t) = \delta^d m_n^d(\delta t) = \frac{\delta^d}{n} \sum_{i=1}^n \mathbf{1}_{\delta t = X_i}$ and $\bar{\mu}_d(t) = \delta^d \mu_d(\delta t)$. Then $\bar{m}_n^d$ and $\bar{\mu}_d$ are respectively the empirical and uniform measures for the variable t on $B(0;1)$. In particular,

$$W_d(m_n^d, \mu_d) = \inf_{\mu \in \Gamma(m_n^d, \mu_d)} \int_{B(0,\delta)\times B(0,\delta)} d(x,y)\, \mu(x,y)\, dx\, dy = \inf_{\mu \in \Gamma(m_n^d, \mu_d)} \int_{B(0;1)\times B(0;1)} d(\delta t, \delta\tau)\, \mu(\delta t, \delta\tau)\, \delta^{2d}\, dt\, d\tau$$

Therefore,

$$\mathbb{E}[W_d(m_n^d, \mu_d)] = \delta O(n^{-\frac{1}{d}}), \quad d \geq 2.$$

Theorem 2. Consider a Poisson process $\mathcal{P}$ with intensity measure $(1+\alpha_n)\frac{n}{|B(0;1)|} dV_d$ on the d-dimensional ball $B(0;1)$, for some sequence $0 \leq \alpha_n \to 0$. Let $m_{\mathcal{P}}^d$ denote the empirical random measure with respect to $\mathcal{P}$, defined by $m_{\mathcal{P}}^d(y) = \frac{1}{|\mathcal{P}|} \sum_{p \in \mathcal{P}} \mathbf{1}_{y=p}$, and let μ_d be the uniform measure on the ball. Then

$$\mathbb{E}[W(m_{\mathcal{P}}^d, \mu_d)] = O(n^{-\frac{1}{d}}).$$

Proof. Since the number of random variables in $m_{\mathcal{P}}^d$ is generated by the mixed random process $\mathcal{P}$, we have

$$\mathbb{E}[W(m_{\mathcal{P}}^d, \mu_d)] = \sum_{k=0}^{\infty} \mathbb{E}[W(m_{\mathcal{P}}^d, \mu_d) \mid |\mathcal{P}| = k] \mathbb{P}(\mathrm{Po}((1+\alpha_n)n) = k).$$

From Theorem 1, we know that

$$\mathbb{E}[W(m_{\mathcal{P}}^d, \mu_d) \mid |\mathcal{P}| = k] = O(k^{-\frac{1}{d}}).$$

On the other hand, from Lemma 1.2 in [4] we have for any constant c:

$$\mathbb{P}(\mathrm{Po}((1+\alpha_n)n) > (1+\alpha_n)n + c\sqrt{(1+\alpha_n)n\log n}) \leq e^{-(1+\alpha_n)n\log n + \cdots} = O(n^{-c^2/2}),$$

and

$$\mathbb{P}(\mathrm{Po}((1+\alpha_n)n) < (1+\alpha_n)n - c\sqrt{(1+\alpha_n)n\log n}) \leq e^{-(1+\alpha_n)n\log n + \cdots} = O(n^{-c^2/2}).$$

Define $a_n^{\pm} = [(1+\alpha_n)n \pm c\sqrt{(1+\alpha_n)n\log n}]$. Then

$$\mathbb{E}[W(m_{\mathcal{P}}^d, \mu_d)] = \sum_{k<a_n^-} + \sum_{a_n^- \leq k \leq a_n^+} + \sum_{k>a_n^+} \mathbb{E}[W(m_{\mathcal{P}}^d, \mu_d) \mid |\mathcal{P}| = k] \mathbb{P}(\mathrm{Po}((1+\alpha_n)n) = k) := I_1 + I_2 + I_3.$$

For I_1 and I_3, we have

$$I_1 \leq 2\mathbb{P}(\mathrm{Po}((1+\alpha_n)n) < a_n^-) = O(n^{-c^2/2}), \quad I_3 \leq 2\mathbb{P}(\mathrm{Po}((1+\alpha_n)n) > a_n^+) = O(n^{-c^2/2}).$$

For the main term I_2, we have

$$I_2 = \sum_{k=a_n^-}^{a_n^+} \mathbb{E}[W(m_{\mathcal{P}}^d, \mu_d) \mid |\mathcal{P}| = k]\mathbb{P}(\mathrm{Po}((1+\alpha_n)n) = k) \leq \sum_{k=a_n^-}^{a_n^+} O(k^{-\frac{1}{d}})\mathbb{P}(\mathrm{Po}((1+\alpha_n)n) = k) = O\Big(((1+\alpha_n)n - $$

Finally, for an appropriately chosen constant c, we obtain

$$\mathbb{E}[W(m_{\mathcal{P}}^d, \mu_d)] = O(n^{-\frac{1}{d}}).$$

Corollary 2. Let $0 \leq \alpha_n \to 0$, $x \in \mathbb{R}^d$, and consider a Poisson process $\mathcal{P}$ with intensity measure $(1+\alpha_n)\frac{n}{|B(0;1)|}dV_d$ on the d-dimensional ball $B(x;\delta)$ with radius $\delta > 0$. Let $m_{\mathcal{P},x}^d$ denote the empirical measure with respect to $\mathcal{P}$, i.e., $m_{\mathcal{P},x}^d(y) = \frac{1}{|\mathcal{P}|}\sum_{p\in\mathcal{P}} \mathbf{1}_{y=p}$, and let $\mu_{d,x}$ be the uniform measure on $B(x;\delta)$. Then

$$\mathbb{E}[W_d(m_{\mathcal{P},x}^d, \mu_{d,x})] = O(n^{-\frac{1}{d}}).$$

Proof. As in the proof of Theorem 2, first note that $|\mathcal{P}|$ follows a Poisson distribution on $B(x;\delta)$ with mean $(1+\alpha_n)n\delta^d$. Therefore,

$$\mathbb{E}[W_d(m_{\mathcal{P},x}^d, \mu_{d,x})] = \mathbb{E}[\delta W_d(\bar{m}_{\mathcal{P},x}^d, \bar{\mu}_{d,x})] = \delta\mathbb{E}[W_d(\bar{m}_{\mathcal{P},x}^d, \bar{\mu}_{d,x})] = \delta O((n\delta^d)^{-\frac{1}{d}}) = O(n^{-\frac{1}{d}}).$$

Conclusion

Corollary 2 may be directly applied to obtain the result in Appendix A.3 of [1]. Next, we hope to apply our method to analyze electrostatic approximation problems.

Appendix A: The Lower Bound of $\mathbb{E}(M_n^d)$

Since $M_n^d = \inf_\sigma \sum_{i=1}^n \|X_i - Y_{\sigma(i)}\| \geq \inf_\sigma \sum_{i=1}^n \min_{j\leq n} \|X_i - Y_j\| = \sum_{i=1}^n \min_{j\leq n} \|X_i - Y_j\|$, we have

$$\mathbb{E}(M_n^d \mid X) \geq \sum_{i=1}^n \mathbb{E}(\min_{j\leq n} \|X_i - Y_j\| \mid X) \geq n \min_{x\in B(0;1)} \mathbb{E}(\min_{j\leq n} \|x - Y_j\|).$$

Let $B(x,t) = \{y \in \mathbb{R}^d : \|x - y\| \leq t\}$. Then for $t < 1$, $\frac{|B(x,t)\cap B(0;1)|}{|B(0;1)|} \leq \min\{t^d, 1\}$ and

$$\mathbb{P}(\min_{j\leq n} \|x - Y_j\| \geq t) \geq (1 - t^d)^n.$$

Therefore,

$$\mathbb{E}(\min_{j\le n}\|x-Y_j\|) = \int_0^\infty \mathbb{P}(\min_{j\le n}\|x-Y_j\| \ge t)\,dt \ge \int_0^1 (1-t^d)^n\,dt = n^{-1/d}\int_0^{n^{1/d}} (1-u^d/n)^n\,du.$$

Finally, by Fatou’ s lemma, we have

$$\liminf_{n\to\infty} \mathbb{E}(M_n^d)n^{1-\frac{1}{d}} \ge \frac{1}{d+1}.$$

Appendix B: The Upper Bound of $\mathbb{E}(M_n^d)$

Set $r = \frac{1}{n\omega_d}$ where $\omega_d = |B(0;1)|$, so that $r \to 0$ as $n \to \infty$ and $|B(x,r)| = \frac{1}{n}$.
Define $u(i,j) = \mathbf{1}_{\{\|X_i-Y_j\|\le r\}}$ and

$$b(x) = \frac{|B(x,r)\cap B(0;1)|}{|B(0;1)|},$$

so that $b(x) \le \frac{1}{n} < 1$ if $n > 1$. First decompose $\sum_{i\le n}(f(X_i)-f(Y_i))$ as follows:

$$\sum_{i=1}^n (f(X_i)-f(Y_i)) = \sum_{i=1}^n\sum_{j=1}^n u(i,j)(f(X_i)-f(Y_j)) + \sum_{i=1}^n f(X_i)\left(1-\sum_{j=1}^n u(i,j)\right) - \sum_{j=1}^n f(Y_j)\left(1-\sum_{i=1}^n u(i,j)\right).$$

Thus,

$$\left|\sum_{i=1}^n (f(X_i)-f(Y_i))\right| \le \left|\sum_{i=1}^n\sum_{j=1}^n u(i,j)(f(X_i)-f(Y_j))\right| + \left|\sum_{i=1}^n f(X_i)\left(1-\sum_{j=1}^n u(i,j)\right) - \sum_{j=1}^n f(Y_j)\left(1-\sum_{i=1}^n u(i,j\right.$$

We will estimate $\mathbb{E}\sup_{f\in\mathcal{L}1} I_1$ and $\mathbb{E}\sup_{f\in\mathcal{L}1} I_2$ separately, and finally obtain

$$\mathbb{E}(M_n^d) = \mathbb{E}\sup_{f\in\mathcal{L}1}\left|\sum_{i=1}^n (f(X_i)-f(Y_i))\right| \le \mathbb{E}\sup_{f\in\mathcal{L}1} I_1 + \mathbb{E}\sup_{f\in\mathcal{L}1} I_2 \le n^{1-\frac{1}{d}} + 2dn^{1-\frac{1}{d}} + 4n^{1-\frac{1}{d}} + 8.$$

Therefore,

$$\limsup_{n\to\infty} \mathbb{E}(M_n^d)n^{1-\frac{1}{d}} \le 1 + 2d + 4 + 8.$$

B.1 The Estimation of $\mathbb{E}\sup_{f\in\mathcal{L}1} I_1$

We have

$$I_1 = \left|\sum_{i=1}^n\sum_{j=1}^n u(i,j)(f(X_i)-f(Y_j))\right| \le r\sum_{i=1}^n\sum_{j=1}^n u(i,j).$$

Since f is Lipschitz,

$$\mathbb{E}\sup_{f\in\mathcal{L}1} I_1 \le r\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}[u(i,j)] = r\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}[\mathbb{E}(u(i,j)\mid X_i)] = r\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}[b(X_i)] \le r\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{n} = n^{1-\frac{1}{d}}.$$

Note that the estimation of I_1 could be further optimized; see [5].

B.2 The Estimation of $\mathbb{E}\sup_{f\in\mathcal{L}1} I_2$

Decompose I_2 as follows:

$$I_2 = \left|\sum_{i=1}^{n} f(X_i)\left(1-nb(X_i)\right) - \sum_{j=1}^{n} f(Y_j)\left(1-nb(Y_j)\right) + \sum_{i=1}^{n} f(X_i)\left(nb(X_i)-\sum_{j=1}^{n}u(i,j)\right) - \sum_{j=1}^{n} f(Y_j)\left(nb(Y_j)\right.\right.$$

We will estimate these three parts and finally obtain

$$\mathbb{E}\sup_{f\in\mathcal{L}1} I_2 \le \mathbb{E}\sup_{f\in\mathcal{L}1} I_{21} + \mathbb{E}\sup_{f\in\mathcal{L}1} I_{22} + \mathbb{E}\sup_{f\in\mathcal{L}1} I_{23} \le 2dn^{1-\frac{1}{d}} + 4n^{1-\frac{1}{d}} + 8.$$

B.2.1 The Estimation of $\mathbb{E}\sup_{f\in\mathcal{L}1} I_{21}$ Note that $\|f\|_{L^\infty[B(0;1)]} \le 1$ and the value of $b(X_i)$ on $B(0;1)$. We have

$$\mathbb{E}\sup_{f\in\mathcal{L}1}\left|\sum_{i=1}^{n} f(X_i)(1-nb(X_i))\right| \le \mathbb{E}\sup_{f\in\mathcal{L}1}\sum_{i=1}^{n}|f(X_i)||1-nb(X_i)| \le \sum_{i=1}^{n}\mathbb{E}|1-nb(X_i)| \le ndr = dn^{1-\frac{1}{d}}.$$

Thus,

$$\mathbb{E}\sup_{f\in\mathcal{L}1} I_{21} = \mathbb{E}\sup_{f\in\mathcal{L}1}\left|\sum_{i=1}^{n} f(X_i)(1-nb(X_i)) - \sum_{j=1}^{n} f(Y_j)(1-nb(Y_j))\right| \le 2dn^{1-\frac{1}{d}}.$$

B.2.2 The Estimation of $\mathbb{E}\sup_{f\in\mathcal{L}1} I_{22}$ and $\mathbb{E}\sup_{f\in\mathcal{L}1} I_{23}$ This part of the estimation is difficult and requires convolution decomposition to restrict f to small regions. The following estimate holds:

$$\mathbb{E}\sup_{f\in\mathcal{L}1} I_{22} = \mathbb{E}\sup_{f\in\mathcal{L}1} I_{23} = \mathbb{E}\sup_{f\in\mathcal{L}1}\left|\sum_{i=1}^{n} f(X_i)\left(nb(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right| \le 2n^{1-\frac{1}{d}}+4.$$

First, assume f is an indicator function on some set A and estimate

$$\mathbb{E}\left|\sum_{i\le n}\mathbf{1}_A(X_i)\left(nb(X_i)-\sum_{j\le n}u(i,j)\right)\right|^2.$$

We have

$$\mathbb{E}\left|\sum_{i=1}^{n}\mathbf{1}_A(X_i)\left(nb(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right|^2=\mathbb{E}\left[\sum_{i,i'\leq n}\sum_{j,j'\leq n}\mathbf{1}_A(X_i)\left(b(X_i)-u(i,j)\right)\mathbf{1}_A(X_{i'})\left(b(X_{i'})-u(i',j')\right)\right].$$

Considering different cases for i,i',j,j', we obtain

$$\mathbb{E}\left|\sum_{i=1}^{n}\mathbf{1}_A(X_i)\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right|^2\leq\frac{n^2(n-1)}{|B(0;1)|^2}+\frac{n^2}{|B(0;1)|}.$$

Next, consider convolution. Let g be a bounded Lipschitz function on $\mathbb{R}^d$, h a bounded support function on $\mathbb{R}^d$, and define $g*h(x)=\int_{\mathbb{R}^d}g(t)h(x-t)\,dt$. To estimate the expectation of

$$\left|\sum_{i=1}^{n}g*h(X_i)\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right|,$$

we write

$$\left|\sum_{i=1}^{n}g*h(X_i)\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right|=\left|\int\sum_{i=1}^{n}\sum_{j=1}^{n}h(X_i-t)\left(b(X_i)-u(i,j)\right)dt\right|.$$

Thus,

$$\mathbb{E}\left|\sum_{i=1}^{n}g*h(X_i)\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right|\leq\|g\|_{L^\infty}\int\mathbb{E}\left|\sum_{i=1}^{n}\sum_{j=1}^{n}h(X_i-t)\left(b(X_i)-u(i,j)\right)\right|dt.$$

Further setting $h(x)=c_0\mathbf{1}_A(x)$, we have by the previous formula

$$\int\mathbb{E}\left|\sum_{i=1}^{n}\sum_{j=1}^{n}h(X_i-t)\left(b(X_i)-u(i,j)\right)\right|^2dt\leq 2nc_0^2|A|.$$

Hence, we obtain the following estimate for the convolution with a characteristic function:

$$\mathbb{E}\left|\sum_{i=1}^{n}g*(c_0\mathbf{1}_A)(X_i)\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right|\leq c_0\sqrt{2|A|n}\|g\|_{L^\infty}.$$

Finally, we decompose f into a sum of well-defined convolutions to estimate these parts. Since a Lipschitz function f on $B(0;1)\subset\mathbb{R}^d$ with $f(0)=0$ can be extended to a Lipschitz function on $\mathbb{R}^d$ with $\|f\|_{L^\infty}\leq 1$, we may consider $f\in\mathcal{L}$

and decompose $f = \sum_{l=1}^{q+1} f_l$, where the decomposition is constructed using a sequence of approximations. Let q be such that $2^{q+1}r \geq 1 > 2^q r$.

Then

$$\mathbb{E}\sup_{f\in\mathcal{L}1} I_{22} = \mathbb{E}\sup_{f\in\mathcal{L}1}\left|\sum_{i=1}^{n}\sum_{l=1}^{q+1} f_l(X_i)\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right| \leq \sum_{l=1}^{q+1}\mathbb{E}\sup_{f\in\mathcal{L}1}\left|\sum_{i=1}^{n} f_l(X_i)\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right|.$$

For the first part, we have

$$\mathbb{E}\sup_{f\in\mathcal{L}1}\left|\sum_{i=1}^{n} f_1(X_i)\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right| \leq \|f-f*h_1\|_{L^\infty}\mathbb{E}\left|\sum_{i=1}^{n}\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right| \leq n^2 r,$$

where we omit the computation showing $\mathbb{E}\left|\sum_{i=1}^{n}\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right|^2 \leq 1$.

For the expectation involving $f_l = (f - f*h_l)*h_1*\cdots*h_{l-1}$ with $2 \leq l \leq q$, we have from the previous estimate

$$\mathbb{E}\left|\sum_{i=1}^{n} f_l(X_i)\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right| \leq 2|B(0;1)|^{-1}(2^{l-1}r)^{-d}|\mathrm{supp}(h_{l-1})|n^{1/2}\|(f-f*h_l)*h_1*\cdots*h_{l-2}\|_{L^\infty} \leq \frac{2}{2^l r}.$$

For the last part involving $f_{q+1} = f*h_1*\cdots*h_q$, we have similarly

$$\mathbb{E}\left|\sum_{i=1}^{n} f_{q+1}(X_i)\left(b(X_i)-\sum_{j=1}^{n}u(i,j)\right)\right| \leq \frac{2}{2^{q+1}r}.$$

Summing all part estimations, we obtain

$$\mathbb{E}\sup_{f\in\mathcal{L}1} I_{22} \leq n^2 r + \left(\frac{2}{2^2 r}+\cdots+\frac{2}{2^{q+1}r}\right) \leq 2n^{1-\frac{1}{d}} + 4.$$

References

[1] Hoorn P. van der, Lippner G., Trugenberger C., Krioukov D. Ollivier-Ricci curvature convergence in random geometric graphs. *Physical Review Research* 3, 013211 (2021).

[2] Kuhn H. W. The Hungarian method for the assignment problem. *Naval Res. Logistics Quarterly*, Volume 2, Issue 1-2, March 1955, Pages 83-97.

[3] Papadimitriou C. H., Steiglitz K. *Combinatorial Optimization, Algorithms and Complexity.* Prentice-Hall, Englewood Cliffs, N.J. (1982).

[4] Penrose M. *Random geometric graphs.* Oxford university press, 2003.

[5] Talagrand M. Matching random samples in many dimensions. *The Annals of Applied Probability* 1992, Vol. 2, No. 4, 846-856.

[6] Veronika Kralj-Iglič, Aleš Iglič, A Simple Statistical Mechanical Approach to the free Energy of the Electric Double Layer Including the Excluded Volume Effect. *Journal de Physique II*, Volume 6, Issue 4, April 1996, pp. 477-491.

[7] Villani C., *Topics in optimal transportation.* Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, RI, 2003.

Note: Figure translations are in progress. See original paper for figures.