ASYMPTOTICS OF KANTOROVICH DISTANCE FOR EMPIRICAL MEASURES OF THE LAGUERRE MODEL

HUAIQIAN LI AND BINGYAO WU

(Communicated by Zhen-Qing Chen)

ABSTRACT. We estimate the rate of convergence for the Kantorovich (or Wasserstein) distance between empirical measures of i.i.d. random variables associated with the Laguerre model of order α on $(0, \infty)^N$ and their common law, which is not compactly supported and has no rotational symmetry. Compared with the Gaussian case, our result is sharp provided the parameter α and the dimension N are chosen in a specified regime.

1. INTRODUCTION

Let $\mathbb{M}$ be a Polish space and let $(Z_n)_{n \in \mathbb{N}}$ be a sequence of i.i.d. $\mathbb{M}$-valued random variables with common distribution $\mathfrak{m}$. Define the empirical measures associated with $(Z_n)_{n \in \mathbb{N}}$ as

$$\mathfrak{m}_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{Z_k}, \quad n \in \mathbb{N},$$

where δ stands for the Dirac measure. It is well known that, by the Glivenko–Cantelli theorem (see e.g. Theorem 7.1 on page 53 of [20]), almost surely, the sequence of random probability measures $(\mathfrak{m}_n)_{n \in \mathbb{N}}$ converges weakly to $\mathfrak{m}$ as $n \to \infty$. So, what about the rate of convergence and in what sense?

Let d be the metric on $\mathbb{M}$ inducing the given topology of $\mathbb{M}$ and let $\mathcal{P}(\mathbb{M})$ be the class of all Borel probability measures on $\mathbb{M}$. For every $p \in [1, \infty)$, define the (pseudo) Kantorovich (or Wasserstein) distance $W_p : \mathcal{P}(\mathbb{M}) \times \mathcal{P}(\mathbb{M}) \to [0, \infty]$ as the infimum

$$W_p(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \left(\int_{\mathbb{M} \times \mathbb{M}} d(x, y)^p \, \pi(\mathrm{d}x, \mathrm{d}y) \right)^{1/p}, \quad \mu, \nu \in \mathcal{P}(\mathbb{M}),$$

where $\Pi(\mu, \nu)$ is the class of all Borel probability measures on the product space $\mathbb{M} \times \mathbb{M}$ with marginals μ and ν, respectively. For $p \in [1, \infty)$, let

$$\mathcal{P}_p(\mathbb{M}) = \left\{ \mu \in \mathcal{P}(\mathbb{M}) : \int_{\mathbb{M}} d(x, o)^p \, \mu(\mathrm{d}x) < \infty \right\},$$

where o is an arbitrary point in $\mathbb{M}$. It is well known that, for each $p \in [1, \infty)$, convergence in the Kantorovich distance W_p implies the weak convergence of sequences from $\mathcal{P}_p(\mathbb{M})$ in general, and if $\mathbb{M}$ is compact, then W_p metrizes the weak topology; see e.g. [24, Theorem 6.9 and Corollary 6.13]. See also [24] for more details on the Kantorovich distance and its connection to the optimal transport theory.

So, it is natural to consider the (often challenging) problem of quantifying the rate of convergence of $W_p(\mathfrak{m}_n, \mathfrak{m})$ as $n \to \infty$.

It is reasonable to see that a solution to this problem involves the distribution $\mathfrak{m}$ and also the geometry and topology of the space $(\mathbb{M}, d)$. Being a classic subject with numerous applications, this kind of problem has attracted intensive investigation; see e.g. [1, 2, 4, 7, 8] for the i.i.d. case, [8, 22] for the Markov chain case, also [13–15, 25, 26, 28, 29] for the case of continuous-time (subordinated) diffusion processes on compact or noncompact Riemannian manifolds with or without boundary and [27] for the non-symmetric case, and even [9] for the case of (non-Markovian) fractional Brownian motions on the flat torus and [16] for the subordinated case.

The aim of the present work is to quantify the Kantorovich distance between empirical measures of i.i.d. random variables associated with the Laguerre model and their common distribution. The motivation comes from recent studies on the i.i.d. case associated with the Ornstein–Uhlenbeck (abbrev. OU or Hermite model [11, 12] and the Jacobi model [30], respectively. The OU case corresponding to the standard Gaussian distribution which is rotationally symmetric and the Jacobi case corresponding to the Beta distribution which has compact support. However, the Laguerre case corresponding to the exponential distribution which has neither the compact support nor the rotational symmetry.

We should mention that, the Laguerre model is naturally associated with the Laguerre polynomial, which is one of the four classes of classical orthogonal polynomials and plays an important role in physics and mathematics; see [10]. Moreover, being a typical example of the Sturm–Liouville operator, the Laguerre model has been studied intensively in analysis and probability; see e.g. [3, 18, 19, 21] and references therein.

In Section 2, we recall the Laguerre model and introduce the main results, and in Section 3, we present proofs of our main results.

2. Preliminaries and main results

We begin with the introduction of the Laguerre model; see [3, 18, 19] for more details.

Let $N \in \mathbb{N}$ and $\mathbb{R}_> = (0, \infty)$. Let $\alpha = (\alpha_1, \cdots, \alpha_N)$ be a multi-index from $(-1, \infty)^N$. For each $x \in \mathbb{R}_>^N$, we always write $x = (x_1, \cdots, x_N)$. Consider the probability measure given by

$$\mu^\alpha(\mathrm{d}x) = \prod_{i=1}^N \frac{x_i^{\alpha_i} e^{-x_i}}{\Gamma(1 + \alpha_i)} \, \mathrm{d}x,$$

where $\Gamma(\cdot)$ stands for the Gamma function and dx denotes the Lebesgue measure on $\mathbb{R}^N$. The Laguerre (differential) operator of type α, denoted by

$$\mathcal{L}^\alpha = \sum_{i=1}^N \left[x_i \partial_{x_i}^2 + (\alpha_i + 1 - x_i)\partial_{x_i} \right],$$

is non-positive and symmetric in $L^2(\mathbb{R}_>^N, \mu^\alpha)$, and $\mathcal{L}^\alpha$ has a self-adjoint extension in $L^2(\mathbb{R}_>^N, \mu^\alpha)$ which is denoted by the same notation.

Let $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$. Consider the N-dimensional Laguerre polynomials of type α, i.e.,

$$L_n^\alpha(x) = \prod_{i=1}^N L_{n_i}^{\alpha_i}(x_i),$$

for every $n = (n_1, \cdots, n_N) \in \mathbb{N}_0^N$ and every $x = (x_1, \cdots, x_N) \in \mathbb{R}_>^N$, where for each $i = 1, \cdots, N$, $L_{n_i}^{\alpha_i}(x_i)$ is a one-dimensional Laguerre polynomial with degree n_i of type α_i given by

$$L_{n_i}^\alpha(x_i) = \frac{1}{n_i!} x_i^{-\alpha_i} e^{x_i} \frac{d^{n_i}}{dx_i^{n_i}} \left(x_i^{n_i + \alpha_i} e^{-x_i} \right).$$

For every $n \in \mathbb{N}_0^N$ with $n = (n_1, \cdots, n_N)$, set $n! := \prod_{i=1}^N n_i!$ and $|n| := n_1 + \cdots + n_N$. Let

$$l_n^\alpha(x) = \sqrt{n!} \prod_{i=1}^N \frac{L_{n_i}^{\alpha_i}(x_i)}{\Gamma(1 + \alpha_i + n_i)}, \quad x = (x_1, \cdots, x_N) \in \mathbb{R}_>^N,$$

be the normalized Laguerre polynomials. It is well known that the family $\{l_n^\alpha : n \in \mathbb{N}_0^N\}$ is an orthonormal basis of $L^2(\mathbb{R}_>^N, \mu^\alpha)$, and each l_n^α is an eigenfunction of $-\mathcal{L}^\alpha$ with corresponding eigenvalue $|n|$, i.e.,

$$-\mathcal{L}^\alpha l_n^\alpha = |n| l_n^\alpha, \quad n \in \mathbb{N}_0^N.$$

The Laguerre semigroup generated by $\mathcal{L}^\alpha$ is denoted by $\{P_t^\alpha\}_{t \geq 0}$, which is known to be a symmetric diffusion semigroup in the sense of Stein [23, page 65] (see .e.g. [18]) and can be represented by

$$P_t^\alpha f(x) = \int_{\mathbb{R}_>^N} p_t^\alpha(x, y) f(y) \, \mu^\alpha(dy), \quad f \in L^p(\mathbb{R}_>^N, \mu^\alpha), \, t > 0,$$

for all $p \in [1, \infty]$, where $\{p_t^\alpha\}_{t>0}$ is the Laguerre kernel given by

$$p_t^\alpha(x, y) = \sum_{k \in \mathbb{N}_0^N} l_k^\alpha(x) l_k^\alpha(y) e^{-|k|t}, \quad x, y \in \mathbb{R}_>^N, \, t > 0.$$

Indeed, $p_t^\alpha(x, y)$ is strictly positive and infinitely differentiable w.r.t. (t, x, y) in $(0, \infty) \times \mathbb{R}_>^N \times \mathbb{R}_>^N$, and more precisely (see e.g. [10, page 78]),

$$p_t^\alpha(x, y) = \prod_{k=1}^N \Gamma(1 + \alpha_k) \frac{\exp\left[-\frac{e^{-t}}{1-e^{-t}}(x_k + y_k) \right]}{(1 - e^{-t})(e^{-t} x_k y_k)^{\alpha_k/2}} I_{\alpha_k}\left(\frac{2\sqrt{e^{-t} x_k y_k}}{1 - e^{-t}} \right)$$

$$=: \prod_{k=1}^N p_t^{\alpha_k}(x_k, y_k), \quad (t, x, y) \in \mathbb{R}_>^{1+2N},$$

where for each $\lambda \in \mathbb{R}$, I_λ is the modified Bessel function of the first kind (also called Bessel function of imaginary argument) with order λ, defined as

$$I_\lambda(x) = \sum_{j=0}^{\infty} \frac{1}{\Gamma(j+1)\Gamma(\lambda+j+1)} \left(\frac{x}{2}\right)^{\lambda+2j}, \quad x \in \mathbb{R}_>;$$

see e.g. [10, Section 5.7]. Moreover, by [17, (4.5) and (4.6)], for each $k = 1, \cdots, N$, there exists constants $C_k, c_k > 0$ depending only on α_k such that

$$(2.1) \qquad c_k \phi_t(x_k, y_k) \le p_t^{\alpha_k}(x_k, y_k) \le C_k \phi_t(x_k, y_k), \quad (t, x_k, y_k) \in \mathbb{R}_>^3,$$

where
(2.2)

$$\phi_t(x, y) := \begin{cases} \frac{1}{(1-e^{-t})^{\alpha+1}} \exp\left(-\frac{e^{-t}(x+y)}{1-e^{-t}}\right), & 0 < xy < \frac{(1-e^{-t})^2}{4e^{-t}}, \\ \frac{(4e^{-t}xy)^{-\alpha/2-1/4}}{e(1-e^{-t})^{1/2}} \exp\left(-\frac{e^{-t}(x+y)-2(e^{-t}xy)^{1/2}}{1-e^{-t}}\right), & xy \ge \frac{(1-e^{-t})^2}{4e^{-t}}. \end{cases}$$

Indeed, (2.2) comes from standard estimates on the modified Bessel function of the first kind.

Now consider the case when $N = 1$. Let $\alpha \in (-1, \infty)$. It is easy to see that, for every $f \in C^2(\mathbb{R}_>)$, the *carré du champ* associated with $\mathcal{L}^\alpha$ is given as follows, i.e.,

$$\mathrm{G}^\alpha(f)(x) := \frac{1}{2}\left[\mathcal{L}^\alpha(f^2) - 2f\mathcal{L}^\alpha f\right](x) = x\left[f'(x)\right]^2, \quad x \in \mathbb{R}_>.$$

Then $(\mathbb{R}_>, \mu^\alpha, \mathrm{G}^\alpha)$ is a Markov Triple; see [3] for more details. It is easy to see that the intrinsic metric induced by the Markov Triple $(\mathbb{R}_>, \mu^\alpha, \mathrm{G}^\alpha)$ is given by

$$\varrho(x, y) = \left|\int_x^y \frac{1}{\sqrt{t}}\, \mathrm{d}t\right| = 2|\sqrt{x} - \sqrt{y}|, \quad x, y \in \mathbb{R}_>.$$

Let $p \in [1, \infty)$ and let $\mathcal{P}(\mathbb{R}_>^N)$ be the class of all Borel probability measures on $\mathbb{R}_>^N$. On the product space $\mathbb{R}_>^N$, we endow the metric

$$\varrho_N(x, y) = \left(\sum_{j=1}^N \varrho(x_j, y_j)^2\right)^{1/2}, \quad x = (x_1, \cdots, x_N) \in \mathbb{R}_>^N, \, y = (y_1, \cdots, y_N) \in \mathbb{R}_>^N.$$

We use $\varrho_N(0, \cdot)$ to denote the distance to the boundary $\{0\}$ of $\mathbb{R}_>^N$. Obviously, $\varrho_N = \varrho$ when $N = 1$. For every $\mu, \nu \in \mathcal{P}(\mathbb{R}_>^N)$, recall the (pseudo) Kantorovich distance of order p between μ and ν, i.e.,

$$\mathrm{W}_p(\mu, \nu) = \left(\inf_{\gamma \in \Pi(\mu, \nu)} \int_{\mathbb{R}_>^N \times \mathbb{R}_>^N} \varrho_N(x, y)^p\, \gamma(\mathrm{d}x, \mathrm{d}y)\right)^{1/p} \in [0, \infty],$$

where $\Pi(\mu, \nu)$ is the set of all couplings of μ and ν, i.e., the set of Borel probability measures γ on $\mathbb{R}_>^N \times \mathbb{R}_>^N$ with marginals $\gamma(A \times \mathbb{R}_>^N) = \mu(A)$ and $\gamma(\mathbb{R}_>^N \times B) = \nu(B)$ for all Borel subsets A, B of $\mathbb{R}_>^N$. In what follows, we concentrate on the quadratic Kantorovich distance W_2.

The main result is presented in the next theorem. In what follows, we employ the notation $a \preceq b$, which means that $a \le cb$ for some positive constant $c > 0$, and c may be either numerical or depend only on the parameters N, α but never on n. And we write $a \asymp b$ if both $a \preceq b$ and $b \preceq a$ hold. For every $\alpha = (\alpha_1, \cdots, \alpha_N) \in \mathbb{R}^N$, set $\alpha_\star := \alpha_1 + \cdots + \alpha_N$.

Theorem 2.1. *Let $n \in \mathbb{N}$, $\alpha \in [-1/2, \infty)^N$ and $X_1, \cdots, X_n$ be independent random variables on $\mathbb{R}_>^N$ with common distribution μ^α. Set*

$$\mu_n = \frac{1}{n} \sum_{j=1}^{n} \delta_{X_j}.$$

Then, for every large enough n,

$$(2.3) \qquad \mathbb{E}\big[W_2(\mu_n, \mu^\alpha)^2\big] \preceq \begin{cases} \frac{\log n}{n^{1/(\alpha_\star + N)}}, & \alpha_\star \in (1 - N, \infty), \\ \frac{(\log n)^2}{n}, & \alpha_\star = 1 - N, \\ \frac{(\log n)^{2(\alpha_\star + N) - 1}}{n}, & \alpha_\star \in (1/2 - N, 1 - N), \\ \frac{\log \log n}{n}, & \alpha_\star = 1/2 - N. \end{cases}$$

Some remarks are in order.

Remark 2.2. Compared with the main result in [11], for the case when $N = 1$ and $\alpha = -1/2$, the rate $(\log \log n)/n$ in (2.3) coincides with the one-dimensional OU case, and for the case when $N = 2$ and $\alpha = (\alpha_1, \alpha_2) = (-1/2, -1/2)$, the rate $(\log n)^2/n$ in (2.3) coincides with the two-dimensional OU case. We should mention that the rate $(\log n)^2/n$ is sharp for the OU case in the critical two-dimension situation which is obtained recently in [6]; refer also to the aforementioned paper for the precise renormalized limit. However, when $N \geq 3$, the result in (2.3) contains an extra factor $\log n$ which is worse than the OU case in [12, Theorem 1.1].

Remark 2.3. The ultra-contractivity, which plays an important role in [2, 26, 28–30], fails for $\{P_t^\alpha\}_{t > 0}$. Indeed, for every $t > 0$,

$$\|P_t^\alpha\|_{L^1(\mathbb{R}_>^N, \mu^\alpha) \to L^\infty(\mathbb{R}_>^N, \mu^\alpha)} := \sup_{\|f\|_{L^1(\mathbb{R}_>^N, \mu^\alpha)} \leq 1} \|P_t^\alpha f\|_{L^\infty(\mathbb{R}_>^N, \mu^\alpha)}$$

$$= \sup_{(x,y) \in \mathbb{R}_>^{2N}} p_t^\alpha(x, y) = \infty.$$

Consider the $N = 1$ case. Applying (2.2), we have for any $x \geq \frac{1 - e^{-t}}{2e^{-t/2}}$,

$$\phi_t(x, x) = \frac{(4e^{-t}x^2)^{-\alpha/2 - 1/4}}{e(1 - e^{-t})^{1/2}} \exp\Big(\frac{2e^{-t/2}x(1 - e^{-t/2})}{1 - e^{-t}}\Big),$$

and hence

$$\sup_{(x,y) \in \mathbb{R}_>^2} \phi_t(x, y) \geq \sup_{x \geq \frac{1 - e^{-t}}{2e^{-t/2}}} \phi_t(x, x) = \infty,$$

which together with (2.1) clearly implies that $\sup_{(x,y) \in \mathbb{R}_>^2} p_t^\alpha(x, y) = \infty$.

Due to technical issues, we are not able to obtain reasonable upper bound estimates on $\mathbb{E}[W_p(\mu_n, \mu^\alpha)^p]$ for other p's at the moment. The problem seems much more complicated than the OU case. So we leave it for future study.

3. Proofs of the main result

In this section, we present the proof of the main theorem. The proof is based on the standard truncation argument and the PDE approach introduced in [5, 11] and [2], respectively.

Proof of Theorem 2.1. Let $R > 0$ and let $B_R = B_R^1 \times \cdots \times B_R^N$, where for each $k = 1, \cdots, N$, $B_R^k := \{y \in \mathbb{R}_> : \ \varrho(0, y) < R\}$, i.e., $B_R^k = \{y \in \mathbb{R}_> : \ y < R^2/4\}$. Set

$$\mu_R^\alpha = \frac{\mathbb{1}_{B_R}}{\mu^\alpha(B_R)} \mu^\alpha.$$

Let $Y_1, \cdots, Y_n$ be independent random variables with common distribution μ_R^α and be also independent of $X_1, \cdots, X_n$. For each $j = 1, \cdots, n$, define

$$X_{j,R} = \begin{cases} X_j, & X_j \in B_R, \\ Y_j, & X_j \in \mathbb{R}_>^N \setminus B_R. \end{cases}$$

Then $X_{1,R}, \cdots, X_{n,R}$ are independent with common distribution μ_R^α. Set

$$\mu_{n,R} = \frac{1}{n} \sum_{j=1}^n \delta_{X_{j,R}}.$$

Let

$$f_{n,R,t}(\cdot) = \frac{1}{n} \sum_{j=1}^n p_t^\alpha(X_{j,R}, \cdot),$$

and denote $\mu_{n,R,t} = f_{n,R,t} \mu^\alpha$.

In order to estimate $\mathbb{E}\big[\mathrm{W}_2(\mu_n, \mu^\alpha)^2\big]$, the idea is simply to write

(3.1)
$$\mathbb{E}\big[\mathrm{W}_2(\mu_n, \mu^\alpha)^2\big] \preceq \mathbb{E}\big[\mathrm{W}_2(\mu_n, \mu_{n,R})^2\big] + \mathbb{E}\big[\mathrm{W}_2(\mu_{n,R}, \mu_{n,R,t})^2\big] + \mathbb{E}\big[\mathrm{W}_2(\mu_{n,R,t}, \mu^\alpha)^2\big],$$

by applying the triangular inequality, and then estimate each term in the right hand side of (3.1) which is presented respectively in PART (1)–(3).

Part (1). Estimate $\mathbb{E}\big[\mathrm{W}_2(\mu_n, \mu_{n,R})^2\big]$. By the convexity of W_2^2 (see e.g. [24, Theorem 4.8]) or by noticing that $\frac{1}{n} \sum_{j=1}^n \delta_{X_j}(\mathrm{d}x)\delta_{X_{j,R}}(\mathrm{d}y) \in \Pi(\mu_n, \mu_{n,R})$, we have

$$\begin{aligned} \mathrm{W}_2(\mu_n, \mu_{n,R})^2 &\le \frac{1}{n} \sum_{j=1}^n \varrho_N(X_j, X_{j,R})^2 \\ &= \frac{1}{n} \sum_{j=1}^n \varrho_N(X_j, Y_j)^2 \mathbb{1}_{\mathbb{R}_>^N \setminus B_R}(X_j) \\ &\le \frac{2}{n} \sum_{j=1}^n \big[\varrho_N(0, X_j)^2 + \varrho_N(0, Y_j)^2\big] \mathbb{1}_{\mathbb{R}_>^N \setminus B_R}(X_j). \end{aligned}$$

Then

$$\begin{aligned} \mathbb{E}\big[\mathrm{W}_2(\mu_n, \mu_{n,R})^2\big] &\le 4 \int_{\mathbb{R}_>^N \setminus B_R} \varrho_N(0, x)^2 \, \mu^\alpha(\mathrm{d}x) \\ &= 16 \int_{\mathbb{R}_>^N \setminus B_R} (x_1 + \cdots + x_N) \prod_{j=1}^N \frac{x_j^{\alpha_j} e^{-x_j}}{\Gamma(1 + \alpha_j)} \, \mathrm{d}x_1 \cdots \mathrm{d}x_N \\ &=: f(R), \quad R > 0. \end{aligned}$$

Let

$$h(R) = \prod_{j=1}^N \int_{R^2/4}^\infty \frac{x_j^{\alpha_j+1} e^{-x_j}}{\Gamma(1 + \alpha_j)} \, \mathrm{d}x_j, \quad R > 0.$$

Noting that $B_R = B_R^1 \times \cdots \times B_R^N$, we may easily check that, as $R \to \infty$, both functions $R \mapsto f(R)$ and $R \mapsto h(R)$ have the same order. For each $j \in \{1, \cdots, N\}$, it is easy to see that the integral

$$\int_{R^2/4}^{\infty} x_j^{\alpha_j+1} e^{-x_j} \, \mathrm{d}x_j$$

has the order of R as $R^{2(\alpha_j+1)} e^{-R^2/4}$ when R tends to ∞. Hence, as a function of R, $\mathbb{E}\big[W_2(\mu_n, \mu_{n,R})^2\big]$ has at most the order $R^{2(\alpha_\star+N)} e^{-R^2/4}$ as $R \to \infty$, where $\alpha_\star := \alpha_1 + \cdots + \alpha_N$ as before.

Now choosing $R = 2(C \log n)^{1/2}$ for some large enough constant $C > 0$, we derive that there exists a constant $c > 1$ such that

$$(3.2) \qquad \mathbb{E}\big[W_2(\mu_n, \mu_{n,R})^2\big] \preceq \frac{1}{n^c},$$

for every large enough n.

Part (2). Estimate $\mathbb{E}\big[W_2(\mu_{n,R}, \mu_{n,R,t})^2\big]$. Let

$$\hat{\pi}(\mathrm{d}x, \mathrm{d}y) = \frac{1}{n} \sum_{j=1}^{n} \delta_{X_{j,R}}(\mathrm{d}x) p_t^\alpha(X_{j,R}, y) \mu^\alpha(\mathrm{d}y), \quad t > 0.$$

It is easy to check that $\hat{\pi} \in \Pi(\mu_{n,R}, \mu_{n,R,t})$ for every $t > 0$. Then

$$W_2(\mu_{n,R}, \mu_{n,R,t})^2 \leq \frac{1}{n} \sum_{j=1}^{n} W_2\big(\delta_{X_{j,R}}, p_t^\alpha(X_{j,R}, \cdot)\mu^\alpha\big)^2$$

$$\leq \frac{1}{n} \sum_{j=1}^{n} \int_{\mathbb{R}_>^N} \varrho_N(X_{j,R}, y)^2 p_t^\alpha(X_{j,R}, y) \, \mu^\alpha(\mathrm{d}y).$$

Hence

$$\mathbb{E}\big[W_2(\mu_{n,R}, \mu_{n,R,t})^2\big]$$

$$\leq \frac{1}{n} \sum_{j=1}^{n} \int_{\mathbb{R}_>^N} \int_{\mathbb{R}_>^N} \varrho_N(x, y)^2 p_t^\alpha(x, y) \, \mu_R^\alpha(\mathrm{d}x)\mu^\alpha(\mathrm{d}y)$$

$$\leq \frac{1}{\mu^\alpha(B_R)} \int_{\mathbb{R}_>^N} \int_{\mathbb{R}_>^N} \varrho_N(x, y)^2 p_t^\alpha(x, y) \, \mu^\alpha(\mathrm{d}x)\mu^\alpha(\mathrm{d}y)$$

$$= \frac{4}{\mu^\alpha(B_R)} \sum_{j=1}^{N} \int_{\mathbb{R}_>} \int_{\mathbb{R}_>} |\sqrt{x_j} - \sqrt{y_j}|^2 p_t^{\alpha_j}(x_j, y_j) \, \mu^{\alpha_j}(\mathrm{d}x_j)\mu^{\alpha_j}(\mathrm{d}y_j)$$

$$=: \frac{4}{\mu^\alpha(B_R)} \sum_{j=1}^{N} I_j(t).$$

For convenience, we omit the subscript j in the expression of I_j and write

$$I(t) = \int_{\mathbb{R}_>} \int_{\mathbb{R}_>} |\sqrt{x} - \sqrt{y}|^2 p_t^\alpha(x, y) \, \mu^\alpha(\mathrm{d}x)\mu^\alpha(\mathrm{d}y).$$

Let $\phi(x) = \sqrt{x}$, $x \in \mathbb{R}_>$. Then, for every $t > 0$,

$$
\begin{aligned}
\mathrm{I}(t) &= \int_{\mathbb{R}_>} \int_{\mathbb{R}_>} |\phi(x) - \phi(y)|^2 p_t^\alpha(x,y)\, \mu^\alpha(\mathrm{d}x)\mu^\alpha(\mathrm{d}y) \\
&= 2 \int_{\mathbb{R}_>} \int_{\mathbb{R}_>} \big[\phi^2(x) - \phi(x)\phi(y)\big] p_t^\alpha(x,y)\, \mu^\alpha(\mathrm{d}x)\mu^\alpha(\mathrm{d}y) \\
&= 2\Big[\int_{\mathbb{R}_>} P_t^\alpha(\phi^2)\, \mathrm{d}\mu^\alpha - \int_{\mathbb{R}_>} (P_{t/2}^\alpha \phi)^2\, \mathrm{d}\mu^\alpha \Big] \\
&= 2\Big[\int_{\mathbb{R}_>} \phi^2\, \mathrm{d}\mu^\alpha - \int_{\mathbb{R}_>} (P_{t/2}^\alpha \phi)^2\, \mathrm{d}\mu^\alpha \Big] \\
&= -2 \int_0^{t/2} \Big(\frac{\mathrm{d}}{\mathrm{d}s} \int_{\mathbb{R}_>} (P_s^\alpha \phi)^2\, \mathrm{d}\mu^\alpha \Big)\, \mathrm{d}s \\
&= -4 \int_0^{t/2} \int_{\mathbb{R}_>} (P_s^\alpha \phi)\mathcal{L}^\alpha(P_s^\alpha \phi)\, \mathrm{d}\mu^\alpha \mathrm{d}s \\
&= 2 \int_0^{t/2} \int_{\mathbb{R}_>} \mathrm{G}^\alpha(P_s^\alpha \phi)\, \mathrm{d}\mu^\alpha \mathrm{d}s \\
&\leq 2 \int_0^{t/2} \int_{\mathbb{R}_>} e^{-s} P_s^\alpha \mathrm{G}^\alpha(\phi)\, \mathrm{d}\mu^\alpha \mathrm{d}s \\
&= 2 \int_0^{t/2} e^{-s}\, \mathrm{d}s \int_{\mathbb{R}_>} \mathrm{G}^\alpha(\phi)\, \mathrm{d}\mu^\alpha \\
&\leq 1 - e^{-t/2} \leq \frac{t}{2},
\end{aligned}
$$

where in the second equality we used the symmetry $p_t^\alpha(x,y) = p_t^\alpha(y,x)$, in the third equality we used the semigroup property of $(P_t^\alpha)_{t>0}$, in the fourth equality we used the invariance of P_t^α w.r.t. μ^α, and in the first inequality we employed the Bakry–Ledoux gradient estimate, i.e.,

$$
\mathrm{G}^\alpha(P_s^\alpha \phi) \leq e^{-s} P_s^\alpha \mathrm{G}^\alpha(\phi), \quad s > 0,
$$

since $\mathcal{L}^\alpha$ satisfies the curvature-dimension condition $\mathrm{CD}(1/2, \infty)$ when $\alpha \in [-1/2, \infty)$ (see e.g. [3]).

Thus, for every $t > 0$ and for all large enough $R > 0$ such that $\mu^\alpha(B_R) \geq 1/2$,

$$
(3.3) \qquad \mathbb{E}\big[\mathrm{W}_2(\mu_{n,R}, \mu_{n,R,t})^2\big] \leq \frac{2Nt}{\mu^\alpha(B_R)} \leq 4Nt.
$$

Part (3). Estimate $\mathbb{E}\big[\mathrm{W}_2(\mu_{n,R,t}, \mu^\alpha)^2\big]$. This part is more technical. By [11, Theorem 2] (see also [2] and [26]), i.e.,

$$
\mathrm{W}_2(h\mu^\alpha, \mu^\alpha)^2 \leq 4 \int_{\mathbb{R}_>^N} \mathrm{G}^\alpha\big((-\mathcal{L}^\alpha)^{-1}(h-1)\big)\, \mathrm{d}\mu^\alpha, \quad h \geq 0,\, \mu^\alpha(h) = 1,
$$

we have

$$(3.4) \quad \mathbb{E}\big[W_2(\mu_{n,R,t}, \mu^\alpha)^2\big] \leq 4\mathbb{E}\Big[\int_{\mathbb{R}_>^N} G^\alpha\big((-\mathcal{L}^\alpha)^{-1}[f_{n,R,t} - 1]\big)(y)\, \mu^\alpha(\mathrm{d}y)\Big]$$

$$= 4\mathbb{E}\Big[\int_{\mathbb{R}_>^N} (f_{n,R,t} - 1)\int_0^\infty P_s^\alpha(f_{n,R,t} - 1)\, \mathrm{d}s\, \mathrm{d}\mu^\alpha\Big]$$

$$= 4\mathbb{E}\int_0^\infty \int_{\mathbb{R}_>^N} \big[P_{s/2}^\alpha(f_{n,R,t} - 1)\big]^2 \, \mathrm{d}\mu^\alpha\, \mathrm{d}s$$

$$= 8\mathbb{E}\int_0^\infty \int_{\mathbb{R}_>^N} \big[P_s^\alpha(f_{n,R,t} - 1)\big]^2 \, \mathrm{d}\mu^\alpha\, \mathrm{d}s.$$

Let $y \in \mathbb{R}_>^N$. Set

$$g(y) := \frac{1}{n}\sum_{j=1}^n \big(p_t^\alpha(X_{j,R}, y) - \mathbb{E}[p_t^\alpha(X_{j,R}, y)]\big), \quad b(y) := \mathbb{E}[p_t^\alpha(X_{j,R}, y)] - 1.$$

Recalling that

$$f_{n,R,t} = \frac{1}{n}\sum_{j=1}^n p_t^\alpha(X_{j,R}, \cdot),$$

we have

$$(3.5) \qquad\qquad\qquad f_{n,R,t}(y) - 1 = g(y) + b(y).$$

Since

$$(3.6) \qquad\qquad b(y) = \int_{\mathbb{R}_>^N} p_t^\alpha(x, y)\, \mu_R^\alpha(\mathrm{d}x) - 1$$

$$(3.7) \qquad\qquad = \frac{1}{\mu^\alpha(B_R)}\int_{\mathbb{R}_>^N} [\mathbb{1}_{B_R} - \mu^\alpha(B_R)]p_t^\alpha(x, y)\, \mu^\alpha(\mathrm{d}x)$$

$$(3.8) \qquad\qquad = \frac{1}{\mu^\alpha(B_R)} P_t^\alpha[\mathbb{1}_{B_R} - \mu^\alpha(B_R)](y),$$

we have

$$P_s^\alpha b(y) = \frac{1}{\mu^\alpha(B_R)} P_{t+s}^\alpha[\mathbb{1}_{B_R} - \mu^\alpha(B_R)](y).$$

By the invariance of $(P_t^\alpha)_{t>0}$ w.r.t. μ^α, it is immediate to see

$$\mu^\alpha(P_s^\alpha b) = \frac{1}{\mu^\alpha(B_R)}\int_{\mathbb{R}_>^N} P_{t+s}^\alpha[\mathbb{1}_{B_R} - \mu^\alpha(B_R)]\, \mathrm{d}\mu^\alpha = 0.$$

Hence, by this and (3.6), we have

$$(3.9) \qquad\quad \mu^\alpha[(P_{t+s}^\alpha \mathbb{1}_{B_R})^2] = \int_{\mathbb{R}_>^N} \mu^\alpha(B_R)^2(1 + P_s^\alpha b)^2 \, \mathrm{d}\mu^\alpha$$

$$(3.10) \qquad\qquad\qquad = \mu^\alpha(B_R)^2\Big(1 + \int_{\mathbb{R}_>^N} (P_s^\alpha b)^2 \, \mathrm{d}\mu^\alpha\Big).$$

Since $X_1, \cdots, X_n$ are independent and have the common distribution μ^α, we derive

$$\mathbb{E}\big[(P_s^\alpha g(y))^2\big] = \mathbb{E}\Big\{\Big(\frac{1}{n}\sum_{j=1}^n \big[p_{t+s}^\alpha(X_{j,R}, y) - \mathbb{E}p_{t+s}^\alpha(X_{j,R}, y)\big]\Big)^2\Big\}$$

$$= \frac{1}{n}\mathbb{E}\big([p_{t+s}^\alpha(X_{1,R}, y) - \mathbb{E}p_{t+s}^\alpha(X_{1,R}, y)]^2\big)$$

$$= \frac{1}{n}\Big[\int_{\mathbb{R}_>^N} p_{t+s}^\alpha(x, y)^2\, \mu_R^\alpha(\mathrm{d}x) - \Big(\int_{\mathbb{R}_>^N} p_{t+s}^\alpha(x, y)\, \mu_R^\alpha(\mathrm{d}x)\Big)^2\Big]$$

$$= \frac{1}{n}\Big[\int_{\mathbb{R}_>^N} p_{t+s}^\alpha(x, y)^2\, \mu_R^\alpha(\mathrm{d}x) - \frac{1}{\mu^\alpha(B_R)^2}\big(P_{t+s}^\alpha \mathbb{1}_{B_R}(y)\big)^2\Big].$$

Hence, by (3.9), the symmetry and the semigroup property, we deduce that

(3.11)
$$\int_{\mathbb{R}_>^N} \mathbb{E}\big[(P_s^\alpha g(y))^2\big]\, \mu^\alpha(\mathrm{d}y)$$

(3.12)
$$= \frac{1}{n}\int_{\mathbb{R}_>^N}\int_{\mathbb{R}_>^N} p_{t+s}^\alpha(x, y)^2\, \mu_R^\alpha(\mathrm{d}x)\, \mu^\alpha(\mathrm{d}y) - \frac{1}{n}\int_{\mathbb{R}_>^N}\Big(1 + \int_{\mathbb{R}_>^N}(P_s^\alpha b)^2\, \mathrm{d}\mu^\alpha\Big)\, \mathrm{d}\mu^\alpha$$

(3.13)
$$\leq \frac{1}{n}\int_{\mathbb{R}_>^N} [p_{2(t+s)}^\alpha(x, x) - 1]\, \mu_R^\alpha(\mathrm{d}x) - \frac{1}{n}\int_{\mathbb{R}_>^N}(P_s^\alpha b)^2\, \mathrm{d}\mu^\alpha.$$

Combining (3.5) and (3.11), we obtain

(3.14)
$$\int_{\mathbb{R}_>^N} \mathbb{E}\big[P_s^\alpha(f_{n,R,t} - 1)(y)\big]^2\, \mu^\alpha(\mathrm{d}y)$$

(3.15)
$$= \int_{\mathbb{R}_>^N} \mathbb{E}\big[(P_s^\alpha g(y))^2 + (P_s^\alpha b(y))^2\big]\, \mu^\alpha(\mathrm{d}y)$$

(3.16)
$$\leq \frac{1}{n}\int_{\mathbb{R}_>^N} \big(p_{2(t+s)}^\alpha(x, x) - 1\big)\, \mu_R^\alpha(\mathrm{d}x) + \Big(1 - \frac{1}{n}\Big)\int_{\mathbb{R}_>^N}(P_s^\alpha b)^2\, \mathrm{d}\mu^\alpha.$$

Since b has mean zero, by the Poincaré inequality implied by the curvature-dimension condition $\mathrm{CD}(1/2, \infty)$, we have

$$\int_{\mathbb{R}_>^N}(P_s^\alpha b)^2\, \mathrm{d}\mu^\alpha \leq e^{-s}\int_{\mathbb{R}_>^N} b^2\, \mathrm{d}\mu^\alpha, \quad s > 0.$$

Note that, by (3.6),

$$\int_{\mathbb{R}_>^N} b^2\, \mathrm{d}\mu^\alpha = \frac{1}{\mu^\alpha(B_R)^2}\int_{\mathbb{R}_>^N} \big(P_t^\alpha[\mathbb{1}_{B_R} - \mu^\alpha(B_R)]\big)^2\, \mathrm{d}\mu^\alpha$$

$$\leq \frac{1}{\mu^\alpha(B_R)^2}\int_{\mathbb{R}_>^N} \big(\mathbb{1}_{B_R} - \mu^\alpha(B_R)\big)^2\, \mathrm{d}\mu^\alpha$$

$$= \frac{1 - \mu^\alpha(B_R)}{\mu^\alpha(B_R)}.$$

Hence

$$\int_{\mathbb{R}_>^N}(P_s^\alpha b)^2\, \mathrm{d}\mu^\alpha \leq e^{-s}\frac{1 - \mu^\alpha(B_R)}{\mu^\alpha(B_R)}, \quad s > 0.$$

As above, letting $R = 2\sqrt{C \log n}$, we have some constant $c > 1$ such that

$$1 - \mu^\alpha(B_R) \preceq \frac{1}{n^c},$$

for every large enough n. Hence, together with (3.14), there exists a constant $c > 1$ such that

$$(3.17) \qquad \int_{\mathbb{R}_>^N} \mathbb{E}\big[P_s^\alpha(f_{n,R,t} - 1)(y)\big]^2 \mu^\alpha(\mathrm{d}y)$$

$$(3.18) \qquad \preceq \frac{1}{n} \int_{\mathbb{R}_>^N} \big(p_{2(t+s)}^\alpha(x,x) - 1\big) \mu_R^\alpha(\mathrm{d}x) + \frac{1}{n^c} e^{-s}, \quad t,s > 0,$$

for every large enough n.

Thus, combining with (3.4) and (3.17), we have some constants $C > 0$, $c > 1$ such that

$$\mathbb{E}\big[\mathrm{W}_2(\mu_{n,R,t}, \mu^\alpha)^2\big] \preceq \frac{1}{n} \int_0^\infty \int_{\mathbb{R}_>^N} \big(p_{2(t+s)}^\alpha(x,x) - 1\big) \mu_R^\alpha(\mathrm{d}x)\, \mathrm{d}s + \frac{1}{n^c}$$

$$(3.19)$$

$$= \frac{1}{2n} \int_{2t}^\infty \int_{\mathbb{R}_>^N} \big(p_s^\alpha(x,x) - 1\big) \mu_R^\alpha(\mathrm{d}x)\, \mathrm{d}s + \frac{1}{n^c}, \quad t > 0,$$

for $R = 2\sqrt{C \log n}$ and every large enough n.

Now we shall estimate

$$\int_{2t}^\infty \int_{\mathbb{R}_>^N} \big(p_s^\alpha(x,x) - 1\big) \mu_R^\alpha(\mathrm{d}x)\, \mathrm{d}s.$$

Let $x = (x_1, \cdots, x_N) \in \mathbb{R}_>^N$. By (2.1) and (2.2), it is easy to see that, for each $j = 1, \cdots, N$,

$$(3.20)$$

$$p_t^{\alpha_j}(x_j, x_j) \preceq \begin{cases} \frac{1}{(1-e^{-t})^{\alpha_j+1}} \exp\Big(-\frac{2e^{-t}x_j}{1-e^{-t}}\Big), & 0 < x_j \le \frac{1-e^{-t}}{2e^{-t/2}}, \\ \frac{(4e^{-t}x_j^2)^{-\alpha_j/2-1/4}}{e(1-e^{-t})^{1/2}} \exp\Big(-\frac{2e^{-t}x_j - 2(e^{-t}x_j^2)^{1/2}}{1-e^{-t}}\Big), & x_j > \frac{1-e^{-t}}{2e^{-t/2}}. \end{cases}$$

Recalling that

$$p_t^\alpha(x,x) = \prod_{j=1}^N p_t^{\alpha_j}(x_j, x_j),$$

we have, for any large enough $R > 0$,

$$\int_{B_R} p_t^\alpha(x,x)\mathrm{d}\mu^\alpha(x)$$

$$\preceq \int_{B_R^1 \times \cdots \times B_R^N} \Big(\prod_{j=1}^N p_t^{\alpha_j}(x_j, x_j) x_j^{\alpha_j} e^{-x_j}\Big)\, \mathrm{d}x_1 \cdots \mathrm{d}x_N$$

$$(3.21)$$

$$= \prod_{j=1}^N \int_{B_R^j} p_t^{\alpha_j}(x_j, x_j) x^{\alpha_j} e^{-x_j}\, \mathrm{d}x_j$$

$$= \prod_{j=1}^N \Big(\int_0^{\frac{1-e^{-t}}{2e^{-t/2}}} p_t^{\alpha_j}(x_j, x_j) x^{\alpha_j} e^{-x_j}\, \mathrm{d}x_j + \int_{\frac{1-e^{-t}}{2e^{-t/2}}}^{\frac{R^2}{4}} p_t^{\alpha_j}(x_j, x_j) x_j^{\alpha_j} e^{-x_j} \mathrm{d}x_j\Big).$$

HUAIQIAN LI AND BINGYAO WU

We need to estimate the two terms in the parentheses of (3.21), and, for simplicity, we omit the subscript j here. According to (3.20), on the one hand, it is easy to get that

$$\int_0^{\frac{1-e^{-t}}{2e^{-t/2}}} p_t^\alpha(x,x)x^\alpha e^{-x}\,\mathrm{d}x$$

(3.22)
$$\preceq \frac{1}{(1-e^{-t})^{\alpha+1}}\int_0^{\frac{1-e^{-t}}{2e^{-t/2}}} \exp\Big(-\frac{2e^{-t}x}{1-e^{-t}}\Big)x^\alpha e^{-x}\,\mathrm{d}x$$

$$= \frac{1}{(1-e^{-t})^{\alpha+1}}\int_0^{\frac{1-e^{-t}}{2e^{-t/2}}} x^\alpha \exp\Big(-\frac{(1+e^{-t})x}{1-e^{-t}}\Big)\,\mathrm{d}x,$$

and on the other hand, when $\alpha \in [-\frac{1}{2},\infty)$, we have

(3.23)
$$\int_{\frac{1-e^{-t}}{2e^{-t/2}}}^{\frac{R^2}{4}} p_t^\alpha(x,x)x^\alpha e^{-x}\,\mathrm{d}x$$

$$\preceq \int_{\frac{1-e^{-t}}{2e^{-t/2}}}^{\frac{R^2}{4}} \frac{(4e^{-t}x^2)^{-\alpha/2-1/4}}{e(1-e^{-t})^{1/2}} \exp\Big(-\frac{2e^{-t}x-2(e^{-t}x^2)^{1/2}}{1-e^{-t}}\Big)x^\alpha e^{-x}\,\mathrm{d}x$$

$$\preceq \frac{e^{t(\alpha/2+1/4)}}{(1-e^{-t})^{1/2}}\Big(\frac{1-e^{-t}}{e^{-t/2}}\Big)^{-\alpha-1/2}\int_{\frac{1-e^{-t}}{2e^{-t/2}}}^{\frac{R^2}{4}} x^\alpha \exp\Big(-\frac{(1+e^{-t}-2e^{-t/2})x}{1-e^{-t}}\Big)\,\mathrm{d}x$$

$$\preceq \frac{1}{(1-e^{-t})^{\alpha+1}}\int_{\frac{1-e^{-t}}{2e^{-t/2}}}^{\frac{R^2}{4}} x^\alpha \exp\Big(-\frac{(1+e^{-t}-2e^{-t/2})x}{1-e^{-t}}\Big)\,\mathrm{d}x.$$

Combining (3.22) and (3.23) with (3.21), for $\alpha \in [-1/2,\infty)^N$, we have the following estimate, i.e.,

$$\int_{B_R} p_t^\alpha(x,x)\,\mu(\mathrm{d}x)$$

$$\preceq \prod_{j=1}^N \frac{1}{(1-e^{-t})^{\alpha_j+1}}\int_0^{\frac{R^2}{4}} x_j^{\alpha_j} \exp\Big(-\frac{(1+e^{-t}-2e^{-t/2})x_j}{1-e^{-t}}\Big)\,\mathrm{d}x_j$$

$$= \prod_{j=1}^N \frac{1}{(1-e^{-t})^{\alpha_j+1}}\int_0^{\frac{1+e^{-t}-2e^{-t/2}}{1-e^{-t}}\frac{R^2}{4}} \Big(\frac{1-e^{-t}}{1+e^{-t}-2e^{-t/2}}\Big)^{\alpha_j+1} x_j^{\alpha_j} e^{-x_j}\,\mathrm{d}x_j$$

$$\preceq \prod_{j=1}^N \frac{1}{(1-e^{-t})^{\alpha_j+1}\theta^{2(\alpha_j+1)}}\mu^{\alpha_j}(B_{\theta R}^j)$$

$$= \frac{1}{(1-e^{-t})^{\alpha_\star+N}\theta^{2(\alpha_\star+N)}}\mu^\alpha(B_{\theta R}),$$

where $\theta := \frac{1-e^{-t/2}}{\sqrt{1-e^{-t}}}$ and it obviously belongs to $(0,1)$ for every $t > 0$. Thus,

(3.24)
$$\int_{\mathbb{R}^N_>} [p_t^\alpha(x,x) - 1]\,\mu_R^\alpha(\mathrm{d}x) = \frac{1}{\mu^\alpha(B_R)} \int_{B_R} p_t^\alpha(x,x)\,\mu^\alpha(\mathrm{d}x) - 1$$
$$\preceq \frac{1}{(1-e^{-t})^{\alpha_\star+N}\theta^{2(\alpha_\star+N)}} \cdot \frac{\mu^\alpha(B_{\theta R})}{\mu^\alpha(B_R)} - 1.$$

If $\theta R \leq 1$, then $t \leq 2\log(\frac{R^2+1}{R^2-1}) \leq \frac{8}{R^2}$ for large enough $R > 0$. Then

(3.25)
$$\frac{1}{(1-e^{-t})^{\alpha_\star+N}\theta^{2(\alpha_\star+N)}} \frac{\mu^\alpha(B_{\theta R})}{\mu^\alpha(B_R)} - 1 \preceq \frac{(\theta R)^{2(\alpha_\star+N)}}{(1-e^{-t})^{\alpha_\star+N}\theta^{2(\alpha_\star+N)}}$$
$$\preceq \frac{R^{2(\alpha_\star+N)}}{t^{\alpha_\star+N}},$$

where we used the fact that $\mu^\alpha(B_R) \geq 1/2$ for large enough $R > 0$ and

$$\mu^\alpha(B_{\theta R}) = \prod_{j=1}^{N} \int_0^{(\theta R)^2/4} \frac{x_j^{\alpha_j}e^{-x_j}}{\Gamma(1+\alpha_j)}\,\mathrm{d}x_j$$
$$\preceq \prod_{j=1}^{N} (\theta R)^{2(\alpha_j+1)} = (\theta R)^{2(\alpha_\star+N)}.$$

If $\theta R > 1$, then $e^{-t/2} < \frac{R^2-1}{R^2+1} \in (0,1)$, $t > 0$. Hence

$$\mu^\alpha(B_{\theta R}) - (1-e^{-t/2})^{2(\alpha_\star+N)}\mu^\alpha(B_R)$$
$$= \left[\mu^\alpha(B_{\theta R}) - \mu^\alpha(B_R)\right] + \left[1 - (1-e^{-t/2})^{2(\alpha_\star+N)}\right]\mu^\alpha(B_R)$$
$$\leq \mu^\alpha(B_R)\left[1 - (1-e^{-t/2})^{2(\alpha_\star+N)}\right]$$
$$\leq 2(\alpha_\star + N)\mu^\alpha(B_R)e^{-t/2}.$$

Then

(3.26)
$$\frac{1}{(1-e^{-t})^{\alpha_\star+N}\theta^{2(\alpha_\star+N)}} \frac{\mu^\alpha(B_{\theta R})}{\mu^\alpha(B_R)} - 1$$
$$= \frac{1}{(1-e^{-t/2})^{2(\alpha_\star+N)}} \frac{\mu^\alpha(B_{\theta R})}{\mu^\alpha(B_R)} - 1$$
$$= \frac{\mu^\alpha(B_{\theta R}) - (1-e^{-t/2})^{2(\alpha_\star+N)}\mu^\alpha(B_R)}{(1-e^{-t/2})^{2(\alpha_\star+N)}\mu^\alpha(B_R)}$$
$$\leq \frac{2(\alpha_\star + N)e^{-t/2}}{(1-e^{-t/2})^{2(\alpha_\star+N)}}.$$

Combining (3.24), (3.25) and (3.26) with $T = 2\log(\frac{R^2+1}{R^2-1}) \leq \frac{8}{R^2}$ for large enough $R > 0$ and $T \gg t$, we obtain that, when $\alpha_\star + N > 1$,

$$
\begin{aligned}
\mathrm{J} &:= \int_{2t}^{\infty} \int_{\mathbb{R}_>^N} [p_s(x,x) - 1] \, \mathrm{d}\mu_R^\alpha(x) \mathrm{d}s \\
&= \int_{2t}^{T} \int_{\mathbb{R}_>^N} [p_s(x,x) - 1] \, \mathrm{d}\mu_R^\alpha(x) \mathrm{d}s + \int_T^{\infty} \int_{\mathbb{R}_>} [p_s(x,x) - 1] \, \mathrm{d}\mu_R^\alpha(x) \mathrm{d}s \\
&\preceq \int_{2t}^{T} \frac{R^{2(\alpha_\star+N)}}{s^{\alpha_\star+N}} \, \mathrm{d}s + \int_T^{\infty} \frac{e^{-s/2}}{(1 - e^{-s/2})^{2(\alpha_\star+N)}} \, \mathrm{d}s \\
&\preceq \frac{R^{2(\alpha_\star+N)}}{t^{\alpha_\star+N-1}} + R^{4(\alpha_\star+N)-2};
\end{aligned}
$$

when $\alpha_\star + N = 1$,

$$
\mathrm{J} \preceq R^2 \log\left(\frac{1}{t}\right) + R^2;
$$

when $\alpha_\star + N \in (1/2, 1)$,

$$
\mathrm{J} \preceq R^{2(\alpha_\star+N)} T^{1-\alpha_\star-N} + R^{4(\alpha_\star+N)-2} \preceq R^{4(\alpha_\star+N)-2};
$$

when $\alpha_\star + N = 1/2$,

$$
\mathrm{J} \preceq R\sqrt{T} + \log\left(\frac{1}{T}\right) \preceq 1 + \log R^2.
$$

Thus, together with (3.19), we deduce that

$$
\begin{aligned}
\mathbb{E}\left[\mathrm{W}_2(\mu_{n,R,t}, \mu^\alpha)^2\right] &\preceq \frac{1}{n}\mathrm{J} + \frac{1}{n^c} \\
(3.27) \qquad &\preceq
\begin{cases}
\frac{1}{n^c} + \frac{1}{n}\left[\frac{R^{2(\alpha_\star+N)}}{t^{\alpha_\star+N-1}} + R^{4(\alpha_\star+N)-2}\right], & \alpha_\star \in (1-N, \infty), \\
\frac{1}{n^c} + \frac{1}{n}\left[R^2 \log\left(\frac{1}{t}\right) + R^2\right], & \alpha_\star = 1-N, \\
\frac{1}{n^c} + \frac{1}{n}R^{4(\alpha_\star+N)-2}, & \alpha_\star \in (1/2-N, 1-N), \\
\frac{1}{n^c} + \frac{1}{n}\left(1 + \log R^2\right), & \alpha_\star = 1/2-N,
\end{cases}
\end{aligned}
$$

which finishes the proof of PART (3).

Finally, combining (3.1) with the estimates in (3.2), (3.3) and (3.27) together, we arrive at
(3.28)

$$
\mathbb{E}\left[\mathrm{W}_2(\mu_n, \mu^\alpha)^2\right] \preceq
\begin{cases}
\frac{1}{n^c} + t + \frac{1}{n}\left[\frac{R^{2(\alpha_\star+N)}}{t^{\alpha_\star+N-1}} + R^{4(\alpha_\star+N)-2}\right], & \alpha_\star \in (1-N, \infty), \\
\frac{1}{n^c} + t + \frac{1}{n}\left[R^2 \log\left(\frac{1}{t}\right) + R^2\right], & \alpha_\star = 1-N, \\
\frac{1}{n^c} + t + \frac{1}{n}R^{4(\alpha_\star+N)-2}, & \alpha_\star \in (1/2-N, 1-N), \\
\frac{1}{n^c} + t + \frac{1}{n}\left(1 + \log R^2\right), & \alpha_\star = 1/2-N,
\end{cases}
$$

for any $R \asymp \sqrt{\log n}$ and large n. Optimizing (3.28) in $t > 0$, we complete the proof of (2.3) for every large enough n. □

ACKNOWLEDGMENT

References

[1] Miklós Ajtai, János Komlós, and Gábor Tusnády, *On optimal matchings*, Combinatorica **4** (1984), no. 4, 259–264, DOI 10.1007/BF02579135. MR779885

[2] Luigi Ambrosio, Federico Stra, and Dario Trevisan, *A PDE approach to a 2-dimensional matching problem*, Probab. Theory Related Fields **173** (2019), no. 1-2, 433–477, DOI 10.1007/s00440-018-0837-x. MR3916111

[3] Dominique Bakry, Ivan Gentil, and Michel Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, 2014, DOI 10.1007/978-3-319-00227-9. MR3155209

[4] Emmanuel Boissard and Thibaut Le Gouic, *On the mean speed of convergence of empirical and occupation measures in Wasserstein distance* (English, with English and French summaries), Ann. Inst. Henri Poincaré Probab. Stat. **50** (2014), no. 2, 539–563, DOI 10.1214/12-AIHP517. MR3189084

[5] François Bolley, Arnaud Guillin, and Cédric Villani, *Quantitative concentration inequalities for empirical measures on non-compact spaces*, Probab. Theory Related Fields **137** (2007), no. 3-4, 541–593, DOI 10.1007/s00440-006-0004-7. MR2280433

[6] Emanuele Caglioti and Francesca Pieroni, *Random matching in 2D with exponent 2 for densities defined on unbounded sets*, Preprint, `arXiv:2302.02602`, 2023.

[7] Steffen Dereich, Michael Scheutzow, and Reik Schottstedt, *Constructive quantization: approximation by empirical measures* (English, with English and French summaries), Ann. Inst. Henri Poincaré Probab. Stat. **49** (2013), no. 4, 1183–1203, DOI 10.1214/12-AIHP489. MR3127919

[8] Nicolas Fournier and Arnaud Guillin, *On the rate of convergence in Wasserstein distance of the empirical measure*, Probab. Theory Related Fields **162** (2015), no. 3-4, 707–738, DOI 10.1007/s00440-014-0583-7. MR3383341

[9] Martin Huesmann, Francesco Mattesini, and Dario Trevisan, *Wasserstein asymptotics for the empirical measure of fractional Brownian motion on a flat torus*, Stochastic Process. Appl. **155** (2023), 1–26, DOI 10.1016/j.spa.2022.09.008. MR4499874

[10] N. N. Lebedev, *Special functions and their applications*, Dover Publications, Inc., New York, 1972. Revised edition, translated from the Russian and edited by Richard A. Silverman; Unabridged and corrected republication. MR350075

[11] Michel Ledoux, *On optimal matching of Gaussian samples*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **457** (2017), no. Veroyatnost′i Statistika. 25, 226–264; English transl., J. Math. Sci. (N.Y.) **238** (2019), no. 4, 495–522. MR3723584

[12] Michel Ledoux and Jie-Xiang Zhu, *On optimal matching of Gaussian samples III*, Probab. Math. Statist. **41** (2021), no. 2, 237–265, DOI 10.37190/0208-4147.41.2.3. MR4359822

[13] Huaiqian Li and Bingyao Wu, *Wasserstein convergence rates for empirical measures of subordinated processes on noncompact manifolds*, J. Theoret. Probab. **36** (2023), no. 2, 1243–1268, DOI 10.1007/s10959-022-01196-8. MR4591873

[14] Huaiqian Li and Bingyao Wu, *Wasserstein convergence for conditional empirical measures of subordinated Dirichlet diffusions on Riemannian manifolds*, Commun. Pure Appl. Anal. **23** (2024), no. 4, 546–576, DOI 10.3934/cpaa.2024024. MR4737038

[15] Huaiqian Li and Bingyao Wu, *Wasserstein convergence for empirical measures of subordinated dirichlet diffusions on Riemannian manifolds*, preprint (2022), `arXiv:2206.03901`.

[16] Huaiqian Li and Bingyao Wu, *Wasserstein convergence for empirical measures of subordinated fractional Brownian motions on the flat torus*, preprint (2023), `arXiv:2305.01228`.

[17] Benjamin Muckenhoupt, *Poisson integrals for Hermite and Laguerre expansions*, Trans. Amer. Math. Soc. **139** (1969), 231–242, DOI 10.2307/1995316. MR249917

[18] Adam Nowak, *On Riesz transforms for Laguerre expansions*, J. Funct. Anal. **215** (2004), no. 1, 217–240, DOI 10.1016/j.jfa.2003.11.012. MR2085116

[19] Adam Nowak, Peter Sjögren, and Tomasz Z. Szarek, *Maximal operators of exotic and non-exotic Laguerre and other semigroups associated with classical orthogonal expansions*, Adv. Math. **318** (2017), 307–354, DOI 10.1016/j.aim.2017.07.026. MR3689743

[20] Kalyanapuram Rangachari Parthasarathy, *Probability measures on metric spaces*, Probability and Mathematical Statistics, No. 3, Academic Press, Inc., New York-London, 1967. MR226684

[21] Pierre Patie and Mladen Savov, *Spectral expansions of non-self-adjoint generalized Laguerre semigroups*, Mem. Amer. Math. Soc. **272** (2021), no. 1336, vii+182, DOI 10.1090/memo/1336. MR4320772

[22] Adrian Riekert, *Convergence rates for empirical measures of Markov chains in dual and Wasserstein distances*, Statist. Probab. Lett. **189** (2022), Paper No. 109605, 8, DOI 10.1016/j.spl.2022.109605. MR4452774

[23] Elias M. Stein, *Topics in harmonic analysis related to the Littlewood-Paley theory*, Annals of Mathematics Studies, No. 63, Princeton University Press, Princeton, NJ; University of Tokyo Press, Tokyo, 1970. MR252961

[24] Cédric Villani, *Optimal transport*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009. Old and new, DOI 10.1007/978-3-540-71050-9. MR2459454

[25] Feng-Yu Wang, *Precise limit in Wasserstein distance for conditional empirical measures of Dirichlet diffusion processes*, J. Funct. Anal. **280** (2021), no. 11, Paper No. 108998, 23, DOI 10.1016/j.jfa.2021.108998. MR4232671

[26] Feng-Yu Wang, *Wasserstein convergence rate for empirical measures on noncompact manifolds*, Stochastic Process. Appl. **144** (2022), 271–287, DOI 10.1016/j.spa.2021.11.006. MR4347493

[27] Feng-Yu Wang, *Convergence in Wasserstein distance for empirical measures of non-symmetric subordinated diffusion processes*, preprint (2023), `arXiv:2301.08420v2`.

[28] Feng-Yu Wang and Bingyao Wu, *Wasserstein convergence for empirical measures of subordinated diffusions on Riemannian manifolds*, Potential Anal. **59** (2023), no. 3, 933–954, DOI 10.1007/s11118-022-09989-6.

[29] Feng-Yu Wang and Jie-Xiang Zhu, *Limit theorems in Wasserstein distance for empirical measures of diffusion processes on Riemannian manifolds* (English, with English and French summaries), Ann. Inst. Henri Poincaré Probab. Stat. **59** (2023), no. 1, 437–475, DOI 10.1214/22-aihp1251. MR4533736

[30] Jie-Xiang Zhu, *Some results on the optimal matching problem for the Jacobi model*, Potential Anal. **55** (2021), no. 1, 167–188, DOI 10.1007/s11118-020-09852-6. MR4261307

Center for Applied Mathematics, Tianjin University, Tianjin 300072, People's Republic of China

Email address: `huaiqianlee@gmail.com`

School of Mathematics and Statistics, Fujian Normal University, Fuzhou 350007, People's Republic of China

Email address: `bingyaowu@163.com`