Submitted to the Annals of Applied Probability

CONVERGENCE IN WASSERSTEIN DISTANCE FOR EMPIRICAL MEASURES OF SEMILINEAR SPDES

BY FENG-YU WANG[1],

[1] *Center for Applied Mathematics, Tianjin University, Tianjin 300072, China*
Department of Mathematics, Swansea University, Bay Campus, Swansea, SA1 8EN, United Kingdom wangfy@tju.edu.cn

The convergence rate in Wasserstein distance is estimated for the empirical measures of symmetric semilinear SPDEs. Unlike in the finite-dimensional case that the convergence is of algebraic order in time, in the present situation the convergence is of log order with a power given by eigenvalues of the underlying linear operator.

1. Introduction. As the continuous Markov process counterpart of Wasserstein matching problem for i.i.d. samples studied in [2, 5] and references within, in [11, 13, 12, 14] we have estimated the convergence rate in Wasserstein distance for empirical measures of symmetric diffusion processes.

Let $V \in C^2(M)$ for a d-dimensional compact connected Riemannian manifold M, let X_t be the diffusion process generated by $L := \Delta + \nabla V$ on M with reflecting boundary if exists, and let $\mathbb{W}_2$ be the L^2-Wasserstein distance induced by the Riemannian metric. According to [14], the empirical measure $\mu_t := \frac{1}{t}\int_0^t \delta_{X_s} \mathrm{d}s$ satisfies

$$\lim_{t\to\infty} t\mathbb{E}[\mathbb{W}_2(\mu_t,\mu)^2] = \sum_{i=1}^\infty \frac{2}{\delta_i},$$

where $\{\delta_i\}_{i\geq 1}$ are all non-trivial eigenvalues of $-L$ in $L^2(\mu)$ counting multiplicities, with Neumann condition if the boundary exists. Since $\sum_{i=1}^\infty \frac{2}{\delta_i} < \infty$ if and only if $d \leq 3$, so that when $t \to \infty$

$$\mathbb{E}[\mathbb{W}_2(\mu_t,\mu)^2] \approx \frac{1}{t},\ \ d \leq 3,$$

where we write $a(t) \approx b(t)$ for two positive functions a and b on $(0,\infty)$, if there exists a constant $C > 1$ such that $C^{-1}a(t) \leq b(t) \leq Ca(t)$ holds for large $t > 0$. Moreover, we have proved in [14] that

$$\mathbb{E}[\mathbb{W}_2(\mu_t,\mu)^2] \approx \begin{cases} \frac{1}{t}\log t, & \text{if } d = 4,\\ t^{-\frac{2}{d-2}}, & \text{if } d \geq 5. \end{cases}$$

These results were then extended in [11, 13] for the empirical measure μ_t of conditional Dirichlet diffusion processes not reaching the boundary before time t, and in [12] for diffusion processes on non-compact complete Riemannian manifolds.

In this paper, we investigate the problem for semilinear SPDEs, whose solutions provide a fundamental class of infinite-dimensional diffusion processes, see [3, 4] for details. It turns out that for this kind of infinite-dimensional processes the convergence of empirical measures becomes log order with a power determined by eigenvalues of the underlying linear operator.

Consider the following SDE on a separable Hilbert space $\mathbb{H}$:

$$\mathrm{d}X_t = \big\{\nabla V(X_t) - AX_t\big\}\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t, \tag{1}$$

MSC2020 subject classifications: Primary 00X00, 00X00; secondary 00X00.
Keywords and phrases: First keyword, second keyword.

where W_t is the cylindrical Brownian motion on $\mathbb{H}$, i.e.

$$W_t = \sum_{i=1}^{\infty} B_t^i e_i, \quad t \geq 0$$

for an orthonormal basis $\{e_i\}_{i\geq 1}$ of $\mathbb{H}$ and a sequence of independent one-dimensional Brownian motions $\{B_t^i\}_{i\geq 1}$, $(A, \mathcal{D}(A))$ is a positive definite self-adjoint operator and $V \in C^1(\mathbb{H})$ satisfying the following assumption.

(H_1) A has discrete spectrum with eigenvalues $\{\lambda_i > 0\}_{i\geq 1}$ listed in the increasing order counting multiplicities satisfying $\sum_{i=1}^d \lambda_i^{-\delta} < \infty$ for some constant $\delta \in (0,1)$, and $V \in C^1(\mathbb{H})$, ∇V is Lipschitz continuous in $\mathbb{H}$ such that

$$\langle \nabla V(x) - \nabla V(y), x - y \rangle \leq (K + \lambda_1)|x-y|^2, \quad x, y \in \mathbb{H} \tag{2}$$

holds for some constant $K \in \mathbb{R}$. Moreover, $Z_V := \mu_0(\mathrm{e}^V) < \infty$, where μ_0 is the centered Gaussian measure on $\mathbb{H}$ with covariance operator A^{-1}.

Under this condition, for any $\mathcal{F}_0$-measurable random variable X_0 on $\mathbb{H}$, (1) has a unique mild solution, and there exists an increasing function $\psi : [0,\infty) \to (0,\infty)$ such that

$$\mathbb{E}[|X_t|^2] \leq \psi(t)\big(1 + \mathbb{E}[|X_0|^2]\big), \quad t \geq 0, \tag{3}$$

see for instance [10, Theorem 3.1.1], or the earlier monographs [3, 4].

Let P_t be the associated Markov semigroup, i.e.

$$P_t f(x) := \mathbb{E}^x[f(X_t)], \quad t \geq 0, f \in \mathcal{B}_b(\mathbb{H}), \quad x \in \mathbb{H},$$

where $\mathcal{B}_b(\mathbb{H})$ is the class of all bounded measurable functions on $\mathbb{H}$, and $\mathbb{E}^x$ is the expectation for the solution X_t of (1) with $X_0 = x$. In general, for a probability measure ν on $\mathbb{H}$, let $\mathbb{E}^\nu$ be the expectation for X_t with initial distribution ν.

By (H_1), we define the probability measure

$$\mu(\mathrm{d}x) := Z_V^{-1} \mathrm{e}^{V(x)} \mu_0(\mathrm{d}x).$$

Then P_t is symmetric in $L^2(\mu)$. For any $p \geq 1$, the L^p-Wasserstein distance is given by

$$\mathbb{W}_p(\mu_1, \mu_2) := \inf_{\pi \in \mathcal{C}(\mu_1,\mu_2)} \left(\int_{\mathbb{H}\times\mathbb{H}} |x-y|^p \pi(\mathrm{d}x, \mathrm{d}y) \right)^{\frac{1}{p}}, \quad \mu_1, \mu_2 \in \mathcal{P}(\mathbb{H}),$$

where $\mathcal{P}(\mathbb{H})$ is the set of all probability measures on $\mathbb{H}$ and $\mathcal{C}(\mu_1, \mu_2)$ is the class of all couplings of μ_1 and μ_2.

In the following two sections, we investigate the upper bound and lower bound estimates on $\mathbb{W}_p(\mu_t, \mu)$ for the empirical measures

$$\mu_t := \frac{1}{t} \int_0^t \delta_{X_s} \mathrm{d}s, \quad t > 0$$

of solutions to (1), where δ_x stands for the Dirac measure at point x. Concrete examples are given to illustrate the resulting estimates, which show that in the present setting the convergence rate is of log order in t with a power given by the growth of λ_i as $i \to \infty$. In particular, when $|V(x)| \leq c(1+|x|)$ for some constant $c > 0$ and all $x \in \mathbb{H}$, and $\lambda_i \approx i^p$ for some $p > 1$ and large i, Example 2.1 and Example 3.1 below imply

$$c_1(\log t)^{1-p\wedge 3} \leq \mathbb{E}^\mu\big[\mathbb{W}_2(\mu_t, \mu)^2\big] \leq c_2(\log t)^{\frac{1}{p}-1}$$

for some constants $c_1, c_2 > 0$ and large $t > 0$.

To conclude this section, we compare the present study with the corresponding ones in [11]-[14] where finite-dimensional diffusion processes are investigated. Due to the lack of sharp estimates on heat kernel and eigenvalues for the generator, new techniques have been developed for the present infinite-dimensional setting. In particular, for the upper bound estimate we apply the dimension-free Harnack inequality established by the author (Section 2), while for the lower bound estimate we present a general result on the Wasserstein distance for discrete measures which applies to infinite-dimensions (Section 3). To derive a reasonable convergence rate, optimization methods are applied for Wasserstein distances of regularized and discretized approximations of empirical measures.

2. Upper bound estimate. We first observe that (H_1) implies the following dimension-free Harnack inequality:

$$(P_t f(x))^p \le (P_t f^p(y)) \exp\Big[\frac{pK|x-y|^2}{2(p-1)(1-e^{-2Kt})}\Big], \ \ t>0, x,y\in\mathbb{H}, f\in\mathcal{B}^+(\mathbb{H}), \tag{4}$$

where $\mathcal{B}^+(\mathbb{H})$ is the class of all nonnegative measurable functions on $\mathbb{H}$. Indeed, by (H_1), the operator $(\lambda_1 - A, \mathcal{D}(A))$ satisfies **(A3.3)** in [10], while $b(s,x) := \nabla V(x) - \lambda_1 x$ and $S(t) := e^{(\lambda_1 - A)t}$ satisfy **(A3.1)** and **(A3.2)** in [10]. So, (4) follows from [10, Theorem 3.2.1].

Next, according to [10, Theorem 1.4.1(6)], (4) implies that P_t has a (symmetric) heat kernel $p_t(x,y)$ with respect to μ such that

$$\begin{aligned}\mu\big(p_t(x,\cdot)^{\frac{p}{p-1}}\big)^{p-1} &= \sup_{\mu(|f|^p)\le 1} (P_t f(x))^p\\ &\le \bigg(\int_{\mathbb{H}} e^{-\frac{pK|x-y|^2}{(p-1)(1-e^{-2Kt})}}\mu(dy)\bigg)^{-1}, \ \ x\in\mathbb{H}, t>0, p>1.\end{aligned} \tag{5}$$

In particular, by taking $p=2$ we obtain

$$p_{2t}(x,x) \le c(t,x) := \bigg(\int_{\mathbb{H}} e^{-\frac{2K|x-y|^2}{1-e^{-2Kt}}}\mu(dy)\bigg)^{-1} < \infty, \ \ t>0, x\in\mathbb{H}. \tag{6}$$

We assume that for any $t>0$,

$$\begin{aligned}\alpha(t) &:= \mathbb{E}^\mu\big[|X_0 - X_t|^2\big] = \int_{\mathbb{H}\times\mathbb{H}} |x-y|^2 p_t(x,y)\mu(dx)\mu(dy) < \infty,\\ \beta(t) &:= \int_{\mathbb{H}} p_{2t}(x,x)\mu(dx) = \int_{\mathbb{H}\times\mathbb{H}} p_t(x,y)^2\mu(dx)\mu(dy) < \infty, \ \ t>0.\end{aligned} \tag{7}$$

In particular, $\beta(t)<\infty$ implies the uniform integrability of P_t in $L^2(\mu)$, so that by [6, Lemma 3.1], P_t is compact in $L^2(\mu)$ and the generator L has discrete spectrum. Since the associated Dirichlet form is irreducible, this implies that L has a spectral gap $\lambda_0>0$, such that

$$\mu(|P_t f - \mu(f)|^2) \le e^{-2\lambda_0 t}\mu(|f-\mu(f)|^2), \ \ t\ge 0, f\in L^2(\mu). \tag{8}$$

In the following theorem, we use α and β to estimate the convergence rate of $\mathbb{E}[\mathbb{W}_2(\mu_t,\mu)^2]$ as $t\to\infty$.

THEOREM 2.1. *Assume (H_1) and (7), and let $c(t,x)$ be in (6). We have*

$$\mathbb{E}^\mu\big[\mathbb{W}_2(\mu_t,\mu)^2\big] \le \inf_{\varepsilon\in(0,1)}\Big\{\frac{16\beta(\varepsilon)}{\lambda_0^2 t} + 2\alpha(\varepsilon)\Big\} =: \xi_t, \ \ t>0. \tag{9}$$

Consequently, for any $x\in\mathbb{H}$,

$$\big(\mathbb{E}^x[\mathbb{W}_2(\mu_t,\mu)]\big)^2 \le \inf_{r>0}\Big\{\frac{8r}{t}\sup_{s\ge 0}\mathbb{E}^x|X_s|^2 + 2c(r,x)\xi_t\Big\}, \ \ t>0. \tag{10}$$

PROOF. (a) We will use the following inequality due to [8, Theorem 2]:

$$\mathbb{W}_2(f\mu,\mu)^2 \le 4\mu(|\nabla(-L)^{-1}(f-1)|^2),\ \ f\ge 0, \mu(f)=1. \tag{11}$$

This estimate was proved using the Kantorovich dual formula and the Hamilton-Jacobi equations, see [2] for an alternative estimate.

To apply (11), we consider the modified empirical measures

$$\mu_{\varepsilon,t} := \mu_t P_\varepsilon = f_{\varepsilon,t}\mu,\ \ \varepsilon>0, t>0, \tag{12}$$

where

$$f_{\varepsilon,t} := \frac{1}{t}\int_0^t p_\varepsilon(X_s,\cdot)\mathrm{d}s. \tag{13}$$

Noting that

$$P_s\{p_\varepsilon(x,\cdot)\}(y) = p_{s+\varepsilon}(x,y),\ \ x,y\in\mathbb{H}, s\ge 0,$$

by the spectral representation we obtain

$$\begin{aligned} \mu(|\nabla(-L)^{-1}(f_{\varepsilon,t}-1)|^2) &= \int_0^\infty \mu(|P_{s/2}(f_{\varepsilon,t}-1)|^2)\mathrm{d}s \\ &= \int_0^\infty \mathrm{d}s\int_{\mathbb{H}}\left(\frac{1}{t}\int_0^t \big(p_{\varepsilon+s/2}(X_u,\cdot)-1\big)\mathrm{d}u\right)^2\mathrm{d}\mu \\ &= \frac{2}{t^2}\int_0^\infty \mathrm{d}s\int_0^t \mathrm{d}s_1\int_{s_1}^t \mu\big(\{p_{\varepsilon+s/2}(X_{s_1},\cdot)-1\}\cdot\{p_{\varepsilon+s/2}(X_{s_2},\cdot)-1\}\big)\mathrm{d}s_2 \\ &= \frac{2}{t^2}\int_0^\infty \mathrm{d}s\int_0^t \mathrm{d}s_1\int_{s_1}^t \big\{p_{2\varepsilon+s}(X_{s_1},X_{s_2})-1\big\}\mathrm{d}s_2. \end{aligned} \tag{14}$$

Next, by (8) we have

$$p_{r+s}(x,x)-1 = \mu\big(|P_{\frac{s}{2}}\{p_{\frac{r}{2}}(x,\cdot)\}-1|^2\big) \le \mathrm{e}^{-\lambda_0 s}\big\{p_r(x,x)-1\big\},\ \ s,r>0. \tag{15}$$

Combining this with the Markov property we derive

$$\begin{aligned} \mathbb{E}^\mu\big\{p_{2\varepsilon+s}(X_{s_1},X_{s_2})-1\big\} &= \int_{\mathbb{H}} P_{s_2-s_1}\big\{p_{2\varepsilon+s}(x,\cdot)-1\big\}(x)\mu(\mathrm{d}x) \\ &= \int_{\mathbb{H}}\big\{p_{2\varepsilon+s+s_2-s_1}(x,x)-1\big\}\mu(\mathrm{d}x) \le \mathrm{e}^{-\lambda_0(s+s_2-s_1)}\beta(\varepsilon). \end{aligned}$$

Therefore, (11) for $f := f_{\varepsilon,t}$ and (14) imply

$$\begin{aligned} \mathbb{E}^\mu\big[\mathbb{W}_2(\mu_{\varepsilon,t},\mu)^2\big] &\le \frac{8\beta(\varepsilon)}{t^2}\int_0^\infty \mathrm{d}s\int_0^t \mathrm{d}s_1\int_{s_1}^t \mathrm{e}^{-\lambda_0(s+s_2-s_1)}\mathrm{d}s_2 \\ &= \frac{8\beta(\varepsilon)}{t^2}\int_0^\infty \mathrm{d}s\int_0^t \frac{\mathrm{e}^{-\lambda_0 s}-\mathrm{e}^{-\lambda_0(s+t-s_1)}}{\lambda_0}\mathrm{d}s_1 \le \frac{8\beta(\varepsilon)}{t^2}\int_0^\infty \mathrm{d}s\int_0^t \frac{\mathrm{e}^{-\lambda_0 s}}{\lambda_0}\mathrm{d}s_1 \\ &= \frac{8\beta(\varepsilon)}{t\lambda_0}\int_0^\infty \mathrm{e}^{-\lambda_0 s}\mathrm{d}s = \frac{8\beta(\varepsilon)}{t\lambda_0^2},\ \ t,\varepsilon>0. \end{aligned} \tag{16}$$

On the other hand, by Jensen's inequality and that $\delta_x P_\varepsilon = \mathcal{L}_{X_\varepsilon}$ for $X_0 = x$, we obtain

$$\mathbb{W}_2(\mu_{\varepsilon,t},\mu_t)^2 \le \left(\frac{1}{t}\int_0^t \mathbb{W}_2(\delta_{X_s},\delta_{X_s}P_\varepsilon)\mathrm{d}s\right)^2 \le \frac{1}{t}\int_0^t \big\{\mathbb{E}^x\big[|x-X_\varepsilon|^2\big]\big\}\big|_{x=X_s}\mathrm{d}s.$$

Since $\mathcal{L}_{X_s} = \mu$ for $\mathcal{L}_{X_0} = \mu$, this implies

$$\mathbb{E}^\mu\big[\mathbb{W}_2(\mu_{\varepsilon,t},\mu_t)^2\big] \le \mathbb{E}^\mu\big[|X_\varepsilon - X_0|^2\big] = \alpha(\varepsilon).$$

Combining with (16), we derive

$$\mathbb{E}^\mu\big[\mathbb{W}_2(\mu_t,\mu)^2\big] \le 2\mathbb{E}^\mu[\mathbb{W}_2(\mu_{\varepsilon,t},\mu)^2] + 2\mathbb{E}^\mu[\mathbb{W}_2(\mu_{\varepsilon,t},\mu)^2] \le \frac{16\beta(\varepsilon)}{t\lambda_0} + 2\alpha(\varepsilon),\ \ \varepsilon\in(0,1).$$

Therefore, (9) holds.

(b) For any $r>0$, let

$$\mu_t^{(r)} := \frac{1}{t}\int_r^{t+r}\delta_{X_s}\mathrm{d}s,\ \ t>0.$$

By the Markov property, Schwarz inequality, (6) and (9), we obtain

$$\begin{aligned}\big(\mathbb{E}^x[\mathbb{W}_2(\mu_t^{(r)},\mu)]\big)^2 &= \bigg(\int_{\mathbb{H}}\mathbb{E}^y[\mathbb{W}_2(\mu_t,\mu)]p_r(x,y)\mu(\mathrm{d}y)\bigg)^2\\ &\le p_{2r}(x,x)\int_{\mathbb{H}}\mathbb{E}^y\big[\mathbb{W}_2(\mu_t,\mu)^2\big]\mu(\mathrm{d}y)\\ &\le c(x,r)\mathbb{E}^\mu\big[\mathbb{W}_2(\mu_t,\mu)^2\big] \le c(x,r)\xi_t,\ \ t>0.\end{aligned}\tag{17}$$

On the other hand, it is easy to see that

$$\pi_t := \frac{1}{t}\int_0^{r\wedge t}\delta_{(X_s,X_{s+\frac{tr}{t\wedge r}})}\mathrm{d}s + \frac{1}{t}\int_{r\wedge t}^t\delta_{(X_s,X_s)}\mathrm{d}s \in \mathcal{C}(\mu_t,\mu_t^{(r)}),$$

so that

$$\begin{aligned}\mathbb{E}^x\big[\mathbb{W}_2(\mu_t,\mu_t^{(r)})^2\big] &\le \mathbb{E}^x\int_{\mathbb{R}^d\times\mathbb{R}^d}|y-z|^2\pi_t(\mathrm{d}y,\mathrm{d}z)\\ &= \frac{1}{t}\int_0^{t\wedge r}\mathbb{E}^x|X_s - X_{r+\frac{ts}{r\wedge t}}|^2\mathrm{d}s \le \frac{4r}{t}\sup_{s\ge0}\mathbb{E}^x|X_s|^2.\end{aligned}$$

This together with (17) and the triangle inequality for $\mathbb{W}_2$, we prove (10). □

Since the heat kernel $p_t(x,y)$ is usually unknown, the estimate presented in Theorem 2.1 is not explicit. To derive explicit estimates, we make the following assumption.

(H_2) There exists an increasing function $\gamma:(0,\infty)\to[0,\infty)$ such that

$$|V(x)| \le \frac{1}{2}\big(\gamma(\varepsilon^{-1}) + \varepsilon|x|^2\big),\ \ x\in\mathbb{H},\varepsilon>0.$$

(H_3) There exist constants $c>0$ and $\theta\in[0,\lambda_1)$

$$|\nabla V(x)| \le c + \theta|x|,\ \ x\in\mathbb{H}.$$

COROLLARY 2.2. *Assume (H_1) and (H_2). Then:*

(1) *There exists a constant $c_0>0$ such that*

$$\mathbb{E}^\mu\big[\mathbb{W}_2(\mu_t,\mu)^2\big] \le c_0\inf_{\varepsilon\in(0,1)}\bigg(\frac{1}{t}\mathrm{e}^{k\varepsilon^{-1}+\gamma(k\varepsilon^{-1})} + \sum_{i=1}^\infty\frac{1-\mathrm{e}^{-2\lambda_i\varepsilon}}{\lambda_i}\bigg) =: \eta_t,\ \ t>0.\tag{18}$$

(2) *If (H_3) holds, then for any $k>K^+$ there exists a constant $c(k)>0$ such that*

$$\big(\mathbb{E}^x[\mathbb{W}_2(\mu_t,\mu)]\big)^2 \le c(k)\mathrm{e}^{k|x|^2}\eta_t,\ \ t\ge1.\tag{19}$$

To prove this result, we need the following two lemmas.

LEMMA 2.3. *Assume* (H_1) *and* (H_3). *There exists a constant* $k > 0$ *such that*

$$\sup_{t\geq 0} \mathbb{E}^x[|X_t|^2] \leq k(1+|x|^2), \;\; x \in \mathbb{H}.$$

PROOF. For $X_0 = x$ we have

$$X_t = \mathrm{e}^{-At}x + \int_0^t \mathrm{e}^{-A(t-s)}\nabla V(X_s)\mathrm{d}s + \sqrt{2}\int_0^t \mathrm{e}^{-A(t-s)}\mathrm{d}W_s, \;\; t \geq 0.$$

By (H_1) and (H_3), we obtain

$$\begin{aligned}
\mathbb{E}|X_t|^2 &\leq (1+\varepsilon^{-1})\mathbb{E}^x\left|\mathrm{e}^{-At}x + \sqrt{2}\int_0^t \mathrm{e}^{A(t-s)}\mathrm{d}W_s\right|^2 \\
&\quad + (1+\varepsilon)\mathbb{E}^x\left|\int_0^t \mathrm{e}^{-A(t-s)}\nabla V(X_s)\mathrm{d}s\right|^2 \\
&\leq 2(1+\varepsilon^{-1})\left(\left|\mathrm{e}^{-At}x\right|^2 + 2\mathbb{E}^x\left|\int_0^t \mathrm{e}^{-A(t-s)}\mathrm{d}W_s\right|^2\right) \\
&\quad + (1+\varepsilon)\left(\int_0^t \left|\mathrm{e}^{-A(t-s)}\nabla V(X_s)\right|\mathrm{d}s\right)^2 \\
&\leq 2(1+\varepsilon^{-1})\left(\mathrm{e}^{-2\lambda_1 t}|x|^2 + 2\sum_{i=1}^\infty \int_0^t \mathrm{e}^{-2\lambda_i(t-s)}\mathrm{d}s\right) \\
&\quad + (1+\varepsilon)\mathbb{E}^x\left(\int_0^t \mathrm{e}^{-\lambda_1(t-s)}\big(c+\theta|X_s|\big)\mathrm{d}s\right)^2, \;\; \varepsilon > 0, t \geq 0.
\end{aligned} \tag{20}$$

Noting that (H_1) implies

$$\sum_{i=1}^\infty \int_0^t \mathrm{e}^{-2\lambda_i(t-s)}\mathrm{d}s \leq \sum_{i=1}^\infty \frac{1}{2\lambda_i} < \infty,$$

and that by Jensen's inequality

$$\begin{aligned}
&\mathbb{E}^x\left(\int_0^t \mathrm{e}^{-\lambda_1(t-s)}\big(c+\theta|X_s|\big)\mathrm{d}s\right)^2 \leq \frac{1}{\lambda_1}\mathbb{E}^x\int_0^t \mathrm{e}^{-\lambda_1(t-s)}\big(c+\theta|X_s|\big)^2\mathrm{d}s \\
&\leq \frac{(1+\varepsilon^{-1})c^2}{\lambda_1} + \frac{(1+\varepsilon)\theta^2}{\lambda_1}\int_0^t \mathrm{e}^{-\lambda_1(t-s)}\mathbb{E}^x[|X_s|^2]\mathrm{d}s,
\end{aligned}$$

for any $\varepsilon > 0$ we find a constant $C(\varepsilon) > 0$ such that (20) yields

$$\mathbb{E}|X_t|^2 \leq C(\varepsilon)(1+|x|^2) + \frac{(1+\varepsilon)^2\theta^2}{2\lambda_1}\int_0^t \mathrm{e}^{-\lambda_1(t-s)}\mathbb{E}[|X_s|^2]\mathrm{d}s, \;\; \varepsilon > 0, t \geq 0.$$

Since $\theta < \lambda_1$, we may take $\varepsilon > 0$ such that $\lambda_\varepsilon := \frac{(1+\varepsilon)^2\theta^2}{\lambda_1} < \lambda_1$, so that

$$\mathrm{e}^{\lambda_1 t}\mathbb{E}^x[|X_t|^2] \leq C(\varepsilon)(1+|x|^2)\mathrm{e}^{\lambda_1 t} + \lambda_\varepsilon\int_0^t \mathrm{e}^{\lambda_1 s}\mathbb{E}|X_s|^2\mathrm{d}s, \;\; t \geq 0.$$

By Gronwall's lemma, we obtain

$$\begin{aligned}\mathrm{e}^{\lambda_1 t}\mathbb{E}^x[|X_t|^2] &\le C(\varepsilon)(1+|x|^2)\mathrm{e}^{\lambda_1 t} + \lambda_\varepsilon\int_0^t C(\varepsilon)(1+|x|^2)\mathrm{e}^{\lambda_1 s}\mathrm{e}^{\lambda_\varepsilon(t-s)}\mathrm{d}s\\ &\le \frac{\lambda_1 C(\varepsilon)}{\lambda_1-\lambda_\varepsilon}\mathrm{e}^{\lambda_1 t}(1+|x|^2),\ \ x\in\mathbb{H},\ \ t\ge 0.\end{aligned}$$

Therefore, the proof is finished. □

LEMMA 2.4. *Under (H_1) and (H_2), there exists a constant $k>0$ such that*

$$\int_{\mathbb{H}}\frac{\mu(\mathrm{d}x)}{\int_{\mathbb{H}}\mathrm{e}^{-\lambda|x-y|^2}\mu(\mathrm{d}y)} \le \mathrm{e}^{\gamma(k\lambda)}\prod_{i=1}^\infty\frac{\lambda_i+k\lambda}{\sqrt{\lambda_i^2-\frac{1}{2}\lambda_1^2}},\ \ \lambda\ge 1. \tag{21}$$

PROOF. Let $\{e_i\}_{i\ge 1}$ be the eigen-basis of A, i.e. it is an orthonormal basis of $\mathbb{H}$ such that

$$Ae_i=\lambda_i e_i,\ \ i\ge 1.$$

Each $x\in\mathbb{H}$ is corresponding to an eigen-coordinate

$$(x_i)_{i\ge1}:=(\langle x,e_i\rangle)_{i\ge 1}\in\ell^2:=\Big\{(r_i)_{i\ge1}\subset\mathbb{R}^\infty:\sum_{i=1}^\infty r_i^2<\infty\Big\}.$$

Under this coordinate we have

$$\mu_0(\mathrm{d}x)=\prod_{i=1}^\infty\frac{\sqrt{\lambda_i}}{\sqrt{2\pi}}\mathrm{e}^{-\frac{\lambda_i x_i^2}{2}}\mathrm{d}x_i. \tag{22}$$

Combining this with (H_2) and $\mu(\mathrm{d}x)=Z_V^{-1}\mathrm{e}^{V(x)}\mu_0(\mathrm{d}x)$, we obtain

$$I:=\int_{\mathbb{H}}\frac{\mu(\mathrm{d}x)}{\int_{\mathbb{H}}\mathrm{e}^{-\lambda|x-y|^2}\mu(\mathrm{d}y)}\le \mathrm{e}^{\gamma(\varepsilon^{-1})}\prod_{i=1}^\infty\Big\{\int_{\mathbb{R}}\frac{\mathrm{e}^{-\frac{\lambda_i-\varepsilon}{2}x_i^2}}{\int_{\mathbb{R}}\mathrm{e}^{-\lambda|x_i-y_i|^2-\frac{\lambda_i+\varepsilon}{2}y_i^2}\mathrm{d}y_i}\Big\}\mathrm{d}x_i,\ \ \varepsilon>0.$$

Noting that

$$\lambda|x_i-y_i|^2+\frac{\lambda_i+\varepsilon}{2}y_i^2=\frac{2\lambda+\lambda_i+\varepsilon}{2}\Big(y_i-\frac{2\lambda x_i}{2\lambda+\lambda_i+\varepsilon}\Big)^2+\frac{\lambda(\lambda_i+\varepsilon)x_i^2}{2\lambda+\lambda_i+\varepsilon},$$

we derive

$$\int_{\mathbb{R}}\mathrm{e}^{-\lambda|x_i-y_i|^2-\frac{\lambda_i+\varepsilon}{2}x_i^2}\mathrm{d}y_i=\Big(\frac{2\pi}{2\lambda+\lambda_i+\varepsilon}\Big)^{\frac12}\mathrm{e}^{\frac{-\lambda(\lambda_i+\varepsilon)x_i^2}{2\lambda+\lambda_i+\varepsilon}},\ \ \varepsilon>0. \tag{23}$$

So, for $\varepsilon>0$ such that $\lambda_1^2-4\lambda\varepsilon-\varepsilon^2>0$, we have

$$\begin{aligned}I&\le \mathrm{e}^{\gamma(\varepsilon^{-1})}\prod_{i=1}^\infty\Big\{\Big(\frac{2\lambda+\lambda_i+\varepsilon}{4\pi}\Big)^{\frac12}\int_{\mathbb{R}}\mathrm{e}^{-(\frac{\lambda_i-\varepsilon}{2}-\frac{\lambda\lambda_i+\lambda\varepsilon}{2\lambda+\lambda_i+\varepsilon})x_i^2}\mathrm{d}x_i\Big\}\\ &=\mathrm{e}^{\gamma(\varepsilon^{-1})}\prod_{i=1}^\infty\Big\{\Big(\frac{2\lambda+\lambda_i+\varepsilon}{4\pi}\Big)^{\frac12}\Big(\frac{4\pi(2\lambda+\lambda_i+\varepsilon)}{\lambda_i^2-4\lambda\varepsilon-\varepsilon^2}\Big)^{\frac12}\Big\},\ \ \varepsilon>0.\end{aligned} \tag{24}$$

Taking $\varepsilon:=\sqrt{4\lambda^2+\frac12\lambda_1^2}-2\lambda$, we have

$$\begin{aligned}&\varepsilon\in\Big(\frac{\lambda_1^2}{2\sqrt{16\lambda^2+2\lambda_1^2}},\ \frac{\lambda_1^2}{8\lambda}\Big),\\ &\lambda_i^2-4\lambda\varepsilon-\varepsilon^2\ge\lambda_i^2-\frac12\lambda_1^2>0,\ \ i\ge 1.\end{aligned}$$

Then for any $\lambda\ge 1$,

$$\varepsilon^{-1}\le \frac{2\sqrt{16+2\lambda_1^2}}{\lambda_1^2}\lambda,\quad \Big(\frac{2\lambda+\lambda_i+\varepsilon}{4\pi}\Big)^{\frac{1}{2}}\Big(\frac{4\pi(2\lambda+\lambda_i+\varepsilon)}{\lambda_i^2-4\lambda\varepsilon-\varepsilon^2}\Big)^{\frac{1}{2}}\le \frac{\lambda_i+(2+\frac{\lambda_1^2}{8})\lambda}{\sqrt{\lambda_i^2-\frac{1}{2}\lambda_1^2}}.$$

Therefore, (24) implies (21) for some constant $k>0$. □

PROOF OF COROLLARY 2.2. (1) By (6) and the second formula in (7), we find a constant $c_1>1$ such that

$$\beta(\varepsilon)\le \int_{\mathbb{H}}\frac{\mu(\mathrm{d}x)\mu(\mathrm{d}y)}{\int_{\mathbb{H}}\mathrm{e}^{-c_1\varepsilon^{-1}|x-y|^2}},\ \ \varepsilon\in(0,1).$$

Combining this with Lemma 2.4, we find constants $c_2,c_3,c_4>0$ such that

$$\begin{aligned}\beta(\varepsilon)&\le \mathrm{e}^{\gamma(c_2\varepsilon^{-1})}\exp\bigg[\sum_{i=1}^\infty\log\bigg(1+\frac{\lambda_i+c_2\varepsilon^{-1}-(\lambda_i^2-\frac{1}{2}\lambda_1^2)^{\frac{1}{2}}}{(\lambda_i^2-\frac{1}{2}\lambda_1^2)^{\frac{1}{2}}}\bigg)\bigg]\\ &\le \mathrm{e}^{\gamma(c_2\varepsilon^{-1})}\exp\bigg[c_3\varepsilon^{-1}\sum_{i=1}^\infty\frac{1}{\lambda_i}\bigg]\le \mathrm{e}^{\gamma(c_2\varepsilon^{-1})+c_4\varepsilon^{-1}},\ \ \varepsilon\in(0,c_1).\end{aligned}$$

Noting that $\beta(\varepsilon)$ is decreasing in ε, we find a constant $k>0$ such that

$$\beta(\varepsilon)\le \mathrm{e}^{\gamma(k\varepsilon^{-1})+k\varepsilon^{-1}},\ \ \varepsilon\in(0,1). \tag{25}$$

On the other hand, by the definition of the mild solution and that of α in (7), we have

$$\begin{aligned}\alpha(\varepsilon)&=\mathbb{E}^\mu\big[|X_\varepsilon-X_0|^2\big]\\ &=\mathbb{E}^\mu\bigg[\bigg|\mathrm{e}^{-A\varepsilon}X_0-X_0+\int_0^\varepsilon \mathrm{e}^{-A(\varepsilon-s)}\nabla V(X_s)\mathrm{d}s+\sqrt{2}\int_0^\varepsilon \mathrm{e}^{-A(\varepsilon-s)}\mathrm{d}W_s\bigg|^2\bigg],\\ &\le 3\mathbb{E}^\mu\big[|\mathrm{e}^{-A\varepsilon}X_0-X_0|^2\big]+3\varepsilon\int_0^\varepsilon\mathbb{E}^\mu\big[|\nabla V(X_s)|^2\big]\mathrm{d}s+6\int_0^\varepsilon\|\mathrm{e}^{-A(\varepsilon-s)}\|_{HS}^2\mathrm{d}s.\end{aligned} \tag{26}$$

Moreover, by (H_1) and (H_2), $\nabla V(x)$ is Lipschitz continuous hence has a linear growth in $|x|$, and $\mu(|\cdot|^2)<\infty$. So, (3) implies $\sup_{s\in[0,1]}\mathbb{E}^\mu[|\nabla V(X_s)|^2]<\infty$. Thus, by (H_1) and (H_2) which imply

$$\mathbb{E}^\mu[\langle X_0,e_i\rangle^2]=\mu(|x_i|^2)\le\frac{c}{\lambda_i},\ \ i\ge 1$$

for some constant $c>0$, we find constants $c_5,c_6>0$ such that

$$\begin{aligned}&\varepsilon\int_0^\varepsilon\mathbb{E}[|\nabla V(X_s)|^2]\mathrm{d}s+\mathbb{E}^\mu[|\mathrm{e}^{-A\varepsilon}X_0-X_0|^2]+\int_0^\varepsilon\|\mathrm{e}^{-A(\varepsilon-s)}\|_{HS}^2\mathrm{d}s\\ &=c_5\varepsilon^2+\sum_{i=1}^\infty\bigg(\frac{(1-\mathrm{e}^{-\lambda_i\varepsilon})^2}{\lambda_i}+\int_0^\varepsilon\mathrm{e}^{-2\lambda_i(\varepsilon-s)}\mathrm{d}s\bigg)\le c_6\sum_{i=1}^\infty\frac{1-\mathrm{e}^{-2\lambda_i\varepsilon}}{\lambda_i},\ \ \varepsilon\in(0,1).\end{aligned}$$

Substituting into (26), we find a constant $k>0$ such that

$$\alpha(\varepsilon)\le k\sum_{i=1}^\infty\frac{1-\mathrm{e}^{-2\lambda_i\varepsilon}}{\lambda_i},\ \ \varepsilon\in(0,1). \tag{27}$$

Combining this with (25) and applying Theorem 2.1, we prove (1).

(2) According to the first assertion and (10), it suffices to show that for any $k > K^+$ there exist constants $r, k_r > 0$ such that

$$c(r,x) \le k_r e^{k_r |x|^2}, \quad x \in \mathbb{H},$$

which follows from (H_2) and (23) with $\lambda = \frac{2K}{1-e^{-2Kr}}$ and $\varepsilon = 0$.

□

2.0.0.1. Example 2.1.. Let ∇V be Lipschitz continuous, $\lambda_i \ge c_0 i^p$ for some constants $c_0 > 0$ and $p > 1$, and there exists a constant $c > 0$ such that

$$|V(x)| \le c(1+|x|), \quad x \in \mathbb{H} \tag{28}$$

holds. Then there exists a a constant $\kappa > 0$ such that

$$\mathbb{E}^\mu[\mathbb{W}_2(\mu_t,\mu)^2] \le \kappa (\log t)^{p^{-1}-1}, \quad t \ge 2. \tag{29}$$

If moreover (H_3) holds, then for any $k > K^+$ there exists a constant $c(k) > 0$ such that

$$\big(\mathbb{E}^\mu[\mathbb{W}_2(\mu_t,\mu)]\big)^2 \le c(k) e^{k|x|^2} (\log t)^{p^{-1}-1}, \quad t \ge 2, \quad x \in \mathbb{H}. \tag{30}$$

PROOF. Let

$$h(\varepsilon) = \sum_{i=1}^\infty \frac{1-e^{-2\varepsilon\lambda_i}}{\lambda_i}, \quad \varepsilon \in [0,1]. \tag{31}$$

When $\lambda_i \ge c i^p$ for some constants $c > 0$ and $p > 1$, we find a constant $c_1 > 0$ such that

$$h'(\varepsilon) = \sum_{i=1}^\infty 2e^{-2\varepsilon\lambda_i} \le 2 + 2\int_1^\infty e^{-2c\varepsilon s^p} ds \le c_1 \varepsilon^{-p^{-1}}, \quad \varepsilon \in (0,1].$$

Thus, there exists a constant $c_2 > 0$ such that

$$\sum_{i=1}^\infty \frac{1-e^{-2\varepsilon\lambda_i}}{\lambda_i} = \int_0^\varepsilon h'(s) ds \le c_2 \varepsilon^{1-p^{-1}}, \quad \varepsilon \in (0,1]. \tag{32}$$

On the other hand, (28) implies (H_2) with

$$\gamma(s) = c_3 s, \quad s \ge 1$$

for some constant $c_3 > 0$. Then by taking $\varepsilon = \frac{2(c_3+k)}{\log t}$, we find constants $c_4, c_5 > 0$ such that

$$\begin{aligned}
&\inf_{\varepsilon\in(0,1)} \left\{ \frac{1}{t} e^{k\varepsilon^{-1} + \gamma(k\varepsilon^{-1})} + \sum_{i=1}^\infty \frac{1-e^{-2\lambda_i\varepsilon}}{\lambda_i} \right\} \\
&\le c_4 \inf_{\varepsilon\in(0,1)} \left\{ \frac{1}{t} e^{(k+c_3)\varepsilon^{-1}} + c_2 \varepsilon^{1-p^{-1}} \right\} \\
&\le c_5 (\log t)^{p^{-1}-1}, \quad t \ge 2.
\end{aligned}$$

Therefore, the desired assertions follow from Corollary 2.2. □

2.0.0.2. Example 2.2.. Let ∇V be Lipschitz continuous, $\lambda_i \geq c\mathrm{e}^{i^p}$ for some constant $c > 0$ and $p > 0$, and (28) holds for some constant $c > 0$. Then there exists a a constant $\kappa > 0$ such that

$$\mathbb{E}^\mu[\mathbb{W}_2(\mu_t, \mu)^2] \leq \kappa(\log t)^{-1} \log\log t, \quad t \geq 4. \tag{33}$$

If moreover (H_3) holds, then for any $k > K^+$ there exists a constant $c(k) > 0$ such that

$$\mathbb{E}^\mu[\mathbb{W}_2(\mu_t, \mu)^2] \leq c(k)\mathrm{e}^{k|x|^2}(\log t)^{-1} \log\log t, \quad t \geq 4, \quad x \in \mathbb{H}. \tag{34}$$

PROOF. Let h be in (31). When $\lambda_i \geq c\mathrm{e}^{ci^p}$ for some constant $c > 0$ and $p > 0$, by using the integral transform $r = \varepsilon \mathrm{e}^{cs^p}$, we find constants $c_1, c_2 > 0$ such that

$$\begin{aligned}
h'(\varepsilon) &= 2\sum_{i=1}^\infty 2\mathrm{e}^{-2\varepsilon\lambda_i} \leq 2\int_0^\infty \exp\big[-2c\varepsilon \mathrm{e}^{cs^p}\big]\mathrm{d}s \\
&= 2\int_\varepsilon^\infty \mathrm{e}^{-2cr}\frac{\mathrm{d}}{\mathrm{d}r}\big\{c^{-1}\log[r\varepsilon^{-1}]\big\}^{\frac{1}{p}}\,\mathrm{d}r \\
&\leq c_1\int_\varepsilon^1 \big\{\log r + \log\varepsilon^{-1}\big\}^{\frac{1}{p}-1}\mathrm{d}\log r + c_1\big\{\log(1+\varepsilon^{-1})\big\}^{\frac{1}{p}-1} \\
&= c_1 \log(1+\varepsilon^{-1})^{\frac{1}{p}-1} + c_0\int_{\log\varepsilon}^0 \big\{u + \log\varepsilon^{-1}\big\}^{\frac{1}{p}-1}\mathrm{d}u \\
&\leq c_2\log(1+\varepsilon^{-1})^{\frac{1}{p}}, \quad \varepsilon \in (0,1].
\end{aligned}$$

Thus, there exists a constant $c_3 > 0$ such that

$$\sum_{i=1}^\infty \frac{1-\mathrm{e}^{-2\varepsilon\lambda_i}}{\lambda_i} = \int_0^\varepsilon h'(s)\mathrm{d}s \leq c_3\varepsilon\log(1+\varepsilon^{-1})^{\frac{1}{p}}, \quad \varepsilon \in (0,1].$$

So, as in the proof of Example 2.1, we find constants $c_4, c_5 > 0$ such that

$$\begin{aligned}
&\inf_{\varepsilon\in(0,1)}\left\{\frac{1}{t}\mathrm{e}^{k\varepsilon^{-1}+\gamma(k\varepsilon^{-1})} + \sum_{i=1}^\infty \frac{1-\mathrm{e}^{-2\lambda_i\varepsilon}}{\lambda_i}\right\} \\
&\leq c_4\inf_{\varepsilon\in(0,1)}\left\{\frac{1}{t}\mathrm{e}^{(k+c_4)\varepsilon^{-1}} + c_3\varepsilon\log(1+\varepsilon^{-1})^{\frac{1}{p}}\right\} \\
&\leq c_5(\log t)^{-1}(\log\log t)^{\frac{1}{p}}, \quad t \geq 4.
\end{aligned}$$

Therefore, the desired assertions follow from Corollary 2.2. □

3. Lower bound estimate.

We first present a lower bound estimate on

$$\mathbb{W}_p(\mu,\nu) := \inf_{\pi\in\mathcal{C}(\mu,\nu)}\left\{\int_{E\times E}\rho(x,y)^p\pi(\mathrm{d}x,\mathrm{d}y)\right\}^{\frac{1}{p}}, \quad p > 0, \mu,\nu \in \mathcal{P}(E) \tag{35}$$

for a metric space (E,ρ), where $\mathcal{P}(E)$ is the set of all probability measures on E. As a generalization to [7, Proposition 4.2] which essentially works for the finite-dimensional setting, we have the following result which also applies to infinite dimensions.

LEMMA 3.1. *Let $\mu \in \mathcal{P}(E)$ such that*

$$\sup_{x\in E}\mu(B(x,r)) \leq \psi(r), \quad r \geq 0 \tag{36}$$

holds for some increasing function ψ, *where* $B(x,r) := \{y \in E : \rho(x,y) < r\}$. *Then for any* $N \geq 1$ *and any probability measure* μ_N *supported on a set of* N *points in* E,

$$\mathbb{W}_p(\mu_N, \mu) \geq 2^{-\frac{1}{p}} \psi^{-1}\Big(\frac{1}{2N}\Big), \tag{37}$$

where $\psi^{-1}(s) := \sup\{r \geq 0 : \psi(r) \leq s\}, s \geq 0$.

PROOF. Let $D = \mathrm{supp}\mu_N$ which contains N many points, so that from (36) we conclude that $D_r := \cup_{x\in D} B(x,r)$ satisfies

$$\mu(D_r) \leq \sum_{x\in D} \mu(B(x,r)) \leq N\psi(r), \quad r \geq 0.$$

Therefore, for any $\pi \in \mathcal{C}(\mu_N, \mu)$, we get

$$\int_{E\times E} \rho(x,y)^p \pi(\mathrm{d}x, \mathrm{d}y) \geq \int_{D\times D_r^c} r^p \pi(\mathrm{d}x, \mathrm{d}y) = r^p \mu(D_r^c) \geq r^p\{1 - N\psi(r)\}, \quad r \geq 0.$$

Combining this with (35) we obtain

$$\mathbb{W}_p(\mu, \nu)^p \geq \sup_{r\geq 0} r^p[1 - N\psi(r)] \geq \frac{1}{2}\{\psi^{-1}(1/(2N))\}^p.$$

□

Now, let $E = \mathbb{H}$ and consider

$$\tilde{\mathbb{W}}_1(\mu, \nu) = \inf_{\pi\in\mathcal{C}(\mu,\nu)} \int_{\mathbb{H}\times\mathbb{H}} \{|x - y| \wedge 1\} \pi(\mathrm{d}x, \mathrm{d}y), \quad \mu, \nu \in \mathcal{P}.$$

THEOREM 3.2. *Assume* (H_1). *Then there exists a constant* $k > 0$ *such that*

$$\mathbb{E}^\mu[\tilde{\mathbb{W}}_1(\mu_t, \mu)] \geq \sup_{N\in\mathbb{N}} \left\{\frac{1}{2}\psi^{-1}\big((2N)^{-1}\big) - \Big(k\sum_{i=1}^\infty \frac{1 - \mathrm{e}^{-2\lambda_i t/N}}{\lambda_i}\Big)^{\frac{1}{2}}\right\}, \quad t \geq 1. \tag{38}$$

If moreover (H_2) *holds, then there exists a constant* $k > 0$ *such that for any* $x \in \mathbb{H}$,

$$\mathbb{E}^x[\tilde{\mathbb{W}}_1(\mu_t, \mu)] \geq \sup_{N\in\mathbb{N}} \left\{\frac{1}{2}\psi^{-1}\big((2N)^{-1}\big) - \Big(k(1+|x|^2)\sum_{i=1}^\infty \frac{1 - \mathrm{e}^{-2\lambda_i t/N}}{\lambda_i}\Big)^{\frac{1}{2}}\right\}, \quad t > 0. \tag{39}$$

PROOF. For any $t > 0$ and $N \in \mathbb{N}$, let

$$t_i = \frac{(i-1)t}{N}, \quad 1 \leq i \leq N+1.$$

Take

$$\mu_{t,N} = \frac{1}{N}\sum_{i=1}^N \delta_{X_{t_i}} = \frac{1}{t}\sum_{i=1}^N \int_{t_i}^{t_{i+1}} \delta_{X_s} \mathrm{d}s.$$

Noting that

$$\pi(\mathrm{d}x, \mathrm{d}y) := \frac{1}{t}\sum_{i=1}^N \int_{t_i}^{t_{i+1}} \delta_{X_{t_i}}(\mathrm{d}x)\delta_{X_{t_i}}(\mathrm{d}y)\mathrm{d}s \in \mathcal{C}(\mu_{t,N}, \mu_t),$$

we obtain

$$
\begin{aligned}
\mathbb{E}^{\mu}\big[\tilde{\mathbb{W}}_1(\mu_{t,N},\mu_t)\big] &\le \mathbb{E}^{\mu}\bigg[\frac{1}{t}\sum_{i=1}^{N}\int_{t_i}^{t_{i+1}}\big\{|X_s-X_{t_i}|\wedge 1\big\}\,\mathrm{d}s\bigg] \\
&\le \frac{1}{t}\sum_{i=1}^{N}\int_{t_i}^{t_{i+1}}\mathbb{E}^{\mu}[|X_0-X_{s-t_i}|]\mathrm{d}s \le \sup_{s\in[0,t/N]}\big(\mathbb{E}^{\mu}[|X_0-X_s|^2]\big)^{\frac{1}{2}}.
\end{aligned}
\tag{40}
$$

This together with (27) implies

$$
\mathbb{E}^{\mu}\big[\tilde{\mathbb{W}}_1(\mu_{t,N},\mu_t)\big] \le \bigg(k\sum_{i=1}^{\infty}\frac{1-\mathrm{e}^{-2\lambda_i t/N}}{\lambda_i}\bigg)^{\frac{1}{2}}.
$$

Therefore, by combining with (37) for $E=\mathbb{H}$ and $\rho(x,y)=|x-y|\wedge 1$, we arrive at

$$
\begin{aligned}
\mathbb{E}^{\mu}\big[\tilde{\mathbb{W}}_1(\mu,\mu_t)\big] &\ge \mathbb{E}^{\mu}\big[\tilde{\mathbb{W}}_1(\mu_{t,N},\mu)\big]-\mathbb{E}^{\mu}\big[\tilde{\mathbb{W}}_1(\mu_{t,N},\mu_t)\big] \\
&\ge \frac{1}{2}\psi^{-1}\Big(\frac{1}{2N}\Big)-\bigg(k\sum_{i=1}^{\infty}\frac{1-\mathrm{e}^{-2\lambda_i t/N}}{\lambda_i}\bigg)^{\frac{1}{2}},\quad N\in\mathbb{N}.
\end{aligned}
\tag{41}
$$

Then (38) holds.

On the other hand, by Lemma 2.3 and the linear growth of $|\nabla V|$, we find constants $c_1,c_2>0$ such that for any $s\in[t_i,t_{i+1}]$,

$$
\begin{aligned}
\mathbb{E}^x|X_s-X_{t_i}| &= \mathbb{E}^x\bigg|\int_{t_i}^{s}\mathrm{e}^{-A(r-t_i)}\nabla V(X_r)\mathrm{d}r+\sqrt{2}\int_{t_i}^{s}\mathrm{e}^{-A(r-t_i)}\mathrm{d}W_r\bigg| \\
&\le \int_{t_i}^{s}\mathrm{e}^{-\lambda_1(r-t_i)}\mathbb{E}^x[|\nabla V(X_r)|]\mathrm{d}r+\sqrt{2}\bigg(\sum_{j=1}^{\infty}\int_{t_i}^{s}\mathrm{e}^{-2\lambda_j(r-t_i)}\mathrm{d}r\bigg)^{\frac{1}{2}} \\
&\le \frac{c_1(1-\mathrm{e}^{-\lambda_1 N/t})}{\lambda_1}(1+|x|)+\sqrt{2}\bigg(\sum_{j=1}^{\infty}\frac{1-\mathrm{e}^{-2\lambda_j t/N}}{2\lambda_j}\bigg)^{\frac{1}{2}} \\
&\le c_2(1+|x|)\bigg(\sum_{j=1}^{\infty}\frac{1-\mathrm{e}^{-2\lambda_j t/N}}{\lambda_j}\bigg)^{\frac{1}{2}}.
\end{aligned}
\tag{42}
$$

where the last step follows from the fact that

$$
\bigg(\sum_{j=1}^{\infty}\frac{1-\mathrm{e}^{-2\lambda_j t/N}}{\lambda_j}\bigg)^{\frac{1}{2}} \ge \bigg(\frac{1-\mathrm{e}^{-2\lambda_1 t/N}}{\lambda_1}\bigg)^{\frac{1}{2}} \ge \frac{1-\mathrm{e}^{-2\lambda_1 t/N}}{\sqrt{\lambda_1}}.
$$

Moreover, by the same reason leading to (40) we have

$$
\mathbb{W}_1(\mu_{N,t},\mu_t) \le \frac{1}{t}\sum_{i=1}^{N}\int_{t_i}^{t_{i+1}}|X_s-X_{t_i}|\mathrm{d}s.
$$

This together with (42) yields

$$
\begin{aligned}
\mathbb{E}^x[\mathbb{W}_1(\mu_{t,N},\mu_t)] &\le \frac{1}{t}\sum_{i=1}^{N}\int_{t_i}^{t_{i+1}}\mathbb{E}^x[|X_s-X_{t_i}|]\mathrm{d}s \\
&\le c_2(1+|x|)\bigg(\sum_{j=1}^{\infty}\frac{1-\mathrm{e}^{-2\lambda_j t/N}}{\lambda_j}\bigg)^{\frac{1}{2}},\quad t\ge 1, N\in\mathbb{N}.
\end{aligned}
$$

Therefore, as in (41) we prove (39) for some constant $k > 0$.

□

3.0.0.3. Example 3.1.. Assume $(H_1), (H_2)$. If there exist constants $p \geq q > 1$ and $k_1, k_2 > 0$ such that

$$k_1 i^q \leq \lambda_i \leq k_2 i^p, \ \ i \geq 1, \tag{43}$$

then there exists a constant $c > 0$ such that for large $t > 1$,

$$\mathbb{E}^\mu[\tilde{\mathbb{W}}_1(\mu_t, \mu)] \geq c\{\log t\}^{-(\frac{p-1}{2}\wedge 1)}. \tag{44}$$

Moreover, for any $x \in \mathbb{H}$ there exist constants $c(x), t(x) > 0$ such that

$$\mathbb{E}^x[\tilde{\mathbb{W}}_1(\mu_t, \mu)] \geq c(x)\{\log t\}^{-\frac{p-1}{2}\wedge 1}, \ \ t \geq t(x). \tag{45}$$

PROOF. (a) We first consider $p > 2$. By (H_2) with $\varepsilon = \frac{\lambda_1}{2}$ and (22), we find a constant $c_1 > 0$ such that

$$\psi(r) := \sup_{x\in\mathbb{H}} \mu(B(x,r)) \leq c_1 \prod_{i=1}^\infty \frac{\sqrt{\lambda_i}}{\sqrt{2\pi}} \int_0^r \mathrm{e}^{-\frac{(\lambda_i-\varepsilon)s^2}{2}} \mathrm{d}s$$
$$= c_1 \prod_{i=1}^\infty \frac{\sqrt{\lambda_i}}{\sqrt{\lambda_i-\varepsilon}} \Big(1 - \frac{\sqrt{\lambda_i-\varepsilon}}{\sqrt{2\pi}} \int_r^\infty \mathrm{e}^{-\frac{(\lambda_i-\varepsilon)s^2}{2}} \mathrm{d}s\Big).$$

Since $\lambda_i \geq k_2 i^p$ and $p > 2$, and $\varepsilon = \frac{\lambda_1}{2} \leq \frac{\lambda_i}{2}$, there exist constants $c_2, c_3, c_4 > 0$ such that

$$\log\Big[\frac{\psi(r)}{c_1}\Big] \leq \sum_{i=1}^\infty \Big\{\frac{1}{2}\log\Big(1 + \frac{\varepsilon}{\lambda_i - \varepsilon}\Big) + \log\Big(1 - \frac{\sqrt{\lambda_i-\varepsilon}}{\sqrt{2\pi}} \int_r^\infty \mathrm{e}^{-\frac{(\lambda_i-\varepsilon)s^2}{2}} \mathrm{d}s\Big)\Big\}$$
$$\leq \sum_{i=1}^\infty \Big\{\frac{\varepsilon}{2(\lambda_i-\varepsilon)} - \frac{\sqrt{\lambda_i-\varepsilon}}{\sqrt{2\pi}} \int_r^\infty \mathrm{e}^{-\frac{(\lambda_i-\varepsilon)s^2}{2}} \mathrm{d}s\Big\}$$
$$\leq \sum_{i=1}^\infty \Big(\frac{1}{\lambda_i} - \frac{c_2}{\sqrt{\lambda_i}}\Big) \leq c_3 - c_4 r^{-1}, \ \ r > 0.$$

Since $\psi(r) \leq 1$, this implies that for some constant $c_5 > 0$,

$$\psi(r) \leq c_5 \mathrm{e}^{-c_4 r^{-1}}, \ \ r > 0.$$

Therefore, there exist constants $c_6 > 0$ such that

$$\psi^{-1}(1/(2N)) \geq c_6\{\log(2N)\}^{-1}, \ \ N \geq 1. \tag{46}$$

On the other hand, since $\lambda_i \geq k_1 i^q$ for some $q > 1$, (32) holds for q replacing p, i.e. there exists a constant $k > 0$ such that

$$\sum_{i=1}^\infty \frac{1 - \mathrm{e}^{-2\varepsilon\lambda_i}}{\lambda_i} \leq k\varepsilon^{1-q^{-1}}, \ \ \varepsilon \in (0,1]. \tag{47}$$

Combining this with (46) and (38) with $N = 1 + \lceil t\rceil^2$, where $\lceil t\rceil$ is the integer part of t, we find a constant $c > 0$ such that for large t

$$\mathbb{E}^\mu[\tilde{\mathbb{W}}_1(\mu_t, \mu)] \geq c\{\log t\}^{-1}. \tag{48}$$

Similarly, for any $x \in \mathbb{H}$ there exist constants $c(x), t(x) > 0$ such that

$$\mathbb{E}^x[\tilde{\mathbb{W}}_1(\mu_t, \mu)] \geq c_1\{\log t\}^{-1}, \ \ t \geq t(x). \tag{49}$$

(b) Take $\varepsilon = 1$ in (H_2), we find a constant $C_1 > 0$ such that for any $R > 0$,

$$\psi(r) \le C_1 \int_{B(0,r)} \mathrm{e}^{\frac{|x|^2}{2}} \mu_0(\mathrm{d}x)$$
$$\le C_1 \mathrm{e}^{\frac{(R+1)r^2}{2}} \int_{\mathbb{H}} \mathrm{e}^{-\frac{R}{2}|x|^2} \mu_0(\mathrm{d}x) = C_1 \mathrm{e}^{\frac{(R+1)r^2}{2}} \prod_{i=1}^{\infty} \frac{\sqrt{\lambda_i}}{\sqrt{\lambda_i + R}}.$$

Since $\lambda_i \le k_2 i^p$ for $i \ge 1$, this implies

$$\log\Big[\frac{\psi(r)}{C_1}\Big] \le \frac{(R+1)r^2}{2} + \frac{1}{2}\log\Big(1 - \frac{R}{\lambda_i + R}\Big) \le \frac{Rr^2}{2} - \frac{1}{2}\sum_{i=1}^{\infty} \frac{R}{\lambda_i + R}$$
$$\le \frac{(R+1)r^2}{2} - \frac{R}{2}\int_1^{\infty} \frac{\mathrm{d}s}{k_2 s^p + R} \le \frac{(R+1)r^2}{2} - c_1 R(1 + R^{\frac{1}{p}})^{1-p}$$
$$\le Rr^2 - c_2 R^{\frac{1}{p}}, \quad r > 0, R \ge 1$$

for some constants $c_1, c_2 > 0$. By taking $r_0 := \varepsilon^{\frac{p-1}{2p}}$ and $R = \varepsilon r^{-\frac{2p}{p-1}}$ for small enough $\varepsilon > 0$, we find a constant $c_3 > 0$ such that

$$\log\Big[\frac{\psi(r)}{C_1}\Big] \le Rr^2 - c_2 R^{\frac{1}{p}} \le -c_3 r^{-\frac{2}{p-1}}, \quad r \in (0, r_0].$$

Combining this with $\psi(r) \le 1$ for all $r \ge 0$, we find a constant $c_4 > 0$ such that

$$\psi(r) \le c_4 \mathrm{e}^{-c_3 r^{-\frac{2}{p-1}}}, \quad r > 0.$$

This implies

$$\psi^{-1}(1/(2N)) \ge c_5 \{\log(2N)\}^{-\frac{p-1}{2}}, \quad N \in \mathbb{N}$$

for some constant $c_5 > 0$. Combining this with (38), (47) and taking $N = 1 + \lceil t \rceil^2$ for large $t > 0$, we find a constant $c_6 > 0$ such that

$$\mathbb{E}[\tilde{\mathbb{W}}_1(\mu_t, \mu)] \ge c_6 \{\log t\}^{-\frac{p-1}{2}}$$

holds for large $t > 0$. This together with (48) implies (44). Similarly, (45) holds for any $x \in \mathbb{H}$ and some constants $c(x), t(x) > 0$. $\square$

3.0.0.4. Acknowledgement.. The author would like to thank the referee for helpful comments and corrections.

REFERENCES

[1]

[2] AMBROSIO, L., STRA, F. and TREVISAN, D. (2019). A PDE approach to a 2-dimensional matching problem. *Probab. Theory Relat. Fields* **173** 433–477.

[3] DA PRATO, G. and ZABACZYK, J. (1992). *Stochastic Equations in Infinite Dimensions*. Cambridge Univ. Press.

[4] DA PRATO, G. and ZABACZYK, J. (1996). *Ergodicity for Infinite-Dimensional Systems*. Cambridge Univ. Press.

[5] FOURNIER, N. and GUILLIN, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Relat. Fields* **162** 707–738.

[6] GONG, F.-Z. and WANG, F.-Y. (2002). Functional inequalities for uniformly integrable semigroups and application to essential spectrum. *Forum Math.* **14** 293–313.

[7] KLOECKNER, B. (2012). Approximation by finitely supported measures. *ESAIM Control Optim. Calc. Var.* **18** 343–359.

[8] LEDOUX, M. (2017). On optimal matching of Gaussian samples. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 457, Veroyatnost' i Statistika.* **25** 226–264.

[9] WANG, F.-Y. (1999). Existence of the spectral gap for elliptic operators. *Arkiv för Math.* **37** 395–407.

[10] WANG, F.-Y. (2013). *Harnack Inequality for Stochastic Partial Differential Equations*. Math. Brief. Springer.

[11] WANG, F.-Y. (2021). Precise limit in Wasserstein distance for conditional empirical measures of Dirichlet diffusion processes. *J. Funct. Anal.* **280** 108998, 23pp.

[12] WANG, F.-Y. (2022). Wasserstein convergence rate for empirical measures on noncompact manifolds. *Stoch. Proc. Appl.* **144** 271–287.

[13] WANG, F.-Y. (2022). Convergence in Wasserstein distance for empirical measures of Dirichlet diffusion processes on manifolds. *J. Eur. Math. Soc.* to appear.

[14] WANG, F.-Y. and ZHU, J.-X. (2022). Limit theorems in Wasserstein distance for empirical measures of diffusion processes on Riemannian manifolds. *Ann. l'Inst. H. Poin. Probab. Stat.* to appear.