Wasserstein Asymptotics for Empirical Measures of Subordinate Killed Diffusions on Compact Riemannian Manifolds

Huaiqian Li

(*Joint work with* Bingyao Wu)

Center for Applied Mathematics
Tianjin University
Tianjin 300072, P. R. China

The 17th Workshop on Markov Processes and Related Topics
2022.11.25–27.

Outline

Outline

The Question

Given a Markov process $X := (X_t)_{t\geq 0}$ on a Polish space (M, ρ) with a Borel probability measure μ, we expect that

$$\frac{1}{t}\int_0^t T_s f \,\mathrm{d}s \overset{t\to\infty}{\to} \int_M f \,\mathrm{d}\mu =: \mu(f), \quad f \in C_b(M), \tag{E}$$

where $T_t f(x) := \mathbb{E}^x[f(X_t)]$.

Denote the empirical measures associated with X as

$$\mu_t = \frac{1}{t}\int_0^t \delta_{X_u}\,\mathrm{d}u, \quad t > 0,$$

where $\delta_\cdot$ is the Dirac measure.

CLASSIC: if μ is the unique invariant probability measure of X, then for every $x \in M$, $\mathbb{P}^x$-a.s., as $t \to \infty$,

$$\mu_t \overset{w}{\to} \mu, \qquad \text{i.e.,} \qquad \mu_t(f) \to \mu(f),\ f \in C_b(M).$$

which in particular implies (E). See e.g. Da Prato–Zabczyk (1996).

What about the rate of convergence??

The Question

Given a Markov process $X := (X_t)_{t\geq 0}$ on a Polish space (M, ρ) with a Borel probability measure μ, we expect that

$$\frac{1}{t}\int_0^t T_s f \,\mathrm{d}s \overset{t\to\infty}{\to} \int_M f \,\mathrm{d}\mu =: \mu(f), \quad f \in C_b(M), \tag{E}$$

where $T_t f(x) := \mathbb{E}^x[f(X_t)]$.

Denote the empirical measures associated with X as

$$\mu_t = \frac{1}{t}\int_0^t \delta_{X_u}\,\mathrm{d}u, \quad t > 0,$$

where $\delta_\cdot$ is the Dirac measure.

CLASSIC: if μ is the unique invariant probability measure of X, then for every $x \in M$, $\mathbb{P}^x$-a.s., as $t \to \infty$,

$$\mu_t \xrightarrow{w} \mu, \qquad \text{i.e.,} \qquad \mu_t(f) \to \mu(f),\ f \in C_b(M).$$

which in particular implies (E). See e.g. Da Prato–Zabczyk (1996).

What about the rate of convergence??

The Question

Given a Markov process $X := (X_t)_{t\geq 0}$ on a Polish space (M, ρ) with a Borel probability measure μ, we expect that

$$\frac{1}{t}\int_0^t T_s f\,\mathrm{d}s \overset{t\to\infty}{\to} \int_M f\,\mathrm{d}\mu =: \mu(f), \quad f \in C_b(M), \tag{E}$$

where $T_t f(x) := \mathbb{E}^x[f(X_t)]$.

Denote the empirical measures associated with X as

$$\mu_t = \frac{1}{t}\int_0^t \delta_{X_u}\,\mathrm{d}u, \quad t > 0,$$

where $\delta_\cdot$ is the Dirac measure.

CLASSIC: if μ is the unique invariant probability measure of X, then for every $x \in M$, $\mathbb{P}^x$-a.s., as $t \to \infty$,

$$\mu_t \overset{w}{\to} \mu, \qquad \text{i.e.,} \qquad \mu_t(f) \to \mu(f),\, f \in C_b(M).$$

which in particular implies (E). See e.g. Da Prato–Zabczyk (1996).

What about the rate of convergence??

The Question

Given a Markov process $X := (X_t)_{t\geq 0}$ on a Polish space (M, ρ) with a Borel probability measure μ, we expect that

$$\frac{1}{t}\int_0^t T_s f \, \mathrm{d}s \overset{t\to\infty}{\to} \int_M f \, \mathrm{d}\mu =: \mu(f), \quad f \in C_b(M), \tag{E}$$

where $T_t f(x) := \mathbb{E}^x[f(X_t)]$.

Denote the empirical measures associated with X as

$$\mu_t = \frac{1}{t}\int_0^t \delta_{X_u} \, \mathrm{d}u, \quad t > 0,$$

where $\delta_\cdot$ is the Dirac measure.
CLASSIC: if μ is the unique invariant probability measure of X, then for every $x \in M$, $\mathbb{P}^x$-a.s., as $t \to \infty$,

$$\mu_t \overset{w}{\to} \mu, \qquad \text{i.e.,} \qquad \mu_t(f) \to \mu(f),\ f \in C_b(M).$$

which in particular implies (E). See e.g. Da Prato–Zabczyk (1996).

What about the rate of convergence??

The Problem

Let $p \in [1, \infty)$. Define the (pseudo) Wasserstein (or Kantorovich) distance as

$$W_p(\nu_1, \nu_2) = \Big(\inf_{\pi \in \mathcal{C}(\nu_1, \nu_2)} \int_{M \times M} \rho(x, y)^p \, \mathrm{d}\pi(x, y) \Big)^{1/p}, \quad \nu_1, \nu_2 \in \mathcal{P}(M),$$

where $\mathcal{C}(\nu_1, \nu_2)$ stands for the class of couplings of ν_1 and ν_2, $\mathcal{P}(M)$ is the class of probability measures on M.

The problem is to study the behavior for large t of

$$\mathbb{E}\big[W_p(\mu_t, \mu)^p\big].$$

Diffusion processes

Consider the triple (M, ρ, μ), where

- M d-dimensional compact connected Riemannian manifold with smooth boundary ∂M,
- ρ geodesic distance on M,
- μ Borel probability measure defined by

$$\mu(\mathrm{d}x) = e^{U(x)}\mathrm{vol}(\mathrm{d}x),$$

where $U \in C^2(M)$ and vol is the volume measure.

Let X_t be the diffusion process generated by

$$\mathcal{L} = \Delta + \nabla U,$$

with hitting time

$$\tau = \inf\{t \geq 0 : \ X_t \in \partial M\}.$$

Diffusion processes

Consider the triple (M, ρ, μ), where

- M d-dimensional compact connected Riemannian manifold with smooth boundary ∂M,
- ρ geodesic distance on M,
- μ Borel probability measure defined by

$$\mu(\mathrm{d}x) = e^{U(x)}\mathrm{vol}(\mathrm{d}x),$$

where $U \in C^2(M)$ and vol is the volume measure.

Let X_t be the diffusion process generated by

$$\mathcal{L} = \Delta + \nabla U,$$

with hitting time

$$\tau = \inf\{t \geq 0 : X_t \in \partial M\}.$$

Spectral properties of $-\mathcal{L}$

Let $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$. It is well known that the spectrum of $-\mathcal{L}$ is discrete, whose eigenvalues are listed in an ascending order counting multiplicities

$$0 < \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \to \infty,$$

and the eigenfunctions ϕ_m, $m \in \mathbb{N}_0$, satisfying the Dirichlet boundary condition, form an ONB in $L^2(\mu)$.

WLOG, assume $\phi_0 > 0$ on $\mathring{M} := M \setminus \partial M$.

Moreover, there exists a constant $\kappa \geq 1$ such that

$$\kappa^{-1} m^{\frac{2}{d}} \leq \lambda_m - \lambda_0 \leq \kappa m^{\frac{2}{d}}, \quad \|\phi_m\|_\infty \leq \kappa \sqrt{m}, \qquad m \in \mathbb{N}.$$

Subordinate killed diffusion processes

Let

$$\mathbf{B} = \{B: \ B \text{ is a Bernstein function with } B(0) = 0, \ B'(0) > 0\},$$

where recall that B is a Bernstein function if

$$B \in C^\infty\big((0,\infty);[0,\infty)\big) \cap C\big([0,\infty);[0,\infty)\big),$$

and for each $n \in \mathbb{N}$,

$$(-1)^{n-1}\frac{\mathrm{d}^n}{\mathrm{d}r^n}B(r) \geq 0, \quad r > 0.$$

WELL KNOWN: $\forall\, B \in \mathbf{B}$, $\exists!$ subordinator $(S_t^B)_{t\geq 0}$ such that

$$\mathbb{E}e^{-\lambda S_t^B} = e^{-tB(\lambda)}, \quad t, \lambda \geq 0.$$

From now on, let $B \in \mathbf{B}$ and $(S_t^B)_{t\geq 0}$ be the subordinator *independent* of $(X_t)_{t\geq 0}$. Define the subordinate killed diffusion process $(X_t^B)_{t\geq 0}$ as

$$X_t^B = X_{S_t^B \wedge \tau}, \quad t \geq 0.$$

Subordinate killed diffusion processes

Let

$$\mathbf{B} = \{B: \ B \text{ is a Bernstein function with } B(0) = 0, \ B'(0) > 0\},$$

where recall that B is a Bernstein function if

$$B \in C^\infty\big((0,\infty);[0,\infty)\big) \cap C\big([0,\infty);[0,\infty)\big),$$

and for each $n \in \mathbb{N}$,

$$(-1)^{n-1}\frac{\mathrm{d}^n}{\mathrm{d}r^n}B(r) \geq 0, \quad r > 0.$$

WELL KNOWN: $\forall\, B \in \mathbf{B}$, $\exists!$ subordinator $(S_t^B)_{t\geq 0}$ such that

$$\mathbb{E}e^{-\lambda S_t^B} = e^{-tB(\lambda)}, \quad t, \lambda \geq 0.$$

From now on, let $B \in \mathbf{B}$ and $(S_t^B)_{t\geq 0}$ be the subordinator *independent* of $(X_t)_{t\geq 0}$. Define the subordinate killed diffusion process $(X_t^B)_{t\geq 0}$ as

$$X_t^B = X_{S_t^B \wedge \tau}, \quad t \geq 0.$$

Subordinate killed diffusion processes

Let

$$\mathbf{B} = \{B : \ B \text{ is a Bernstein function with } B(0) = 0,\ B'(0) > 0\},$$

where recall that B is a Bernstein function if

$$B \in C^\infty\big((0,\infty);[0,\infty)\big) \cap C\big([0,\infty);[0,\infty)\big),$$

and for each $n \in \mathbb{N}$,

$$(-1)^{n-1}\frac{\mathrm{d}^n}{\mathrm{d}r^n}B(r) \geq 0, \quad r > 0.$$

WELL KNOWN: $\forall\, B \in \mathbf{B}$, $\exists!$ subordinator $(S_t^B)_{t\geq 0}$ such that

$$\mathbb{E}e^{-\lambda S_t^B} = e^{-tB(\lambda)}, \quad t, \lambda \geq 0.$$

From now on, let $B \in \mathbf{B}$ and $(S_t^B)_{t\geq 0}$ be the subordinator *independent* of $(X_t)_{t\geq 0}$. Define the subordinate killed diffusion process $(X_t^B)_{t\geq 0}$ as

$$X_t^B = X_{S_t^B \wedge \tau}, \quad t \geq 0.$$

Conditional empirical measures

Let

$$\sigma_\tau^B := \inf\{t \geq 0 : \ S_t^B > \tau\}.$$

Define the conditional empirical measures associated with $(X_t^B)_{t\geq 0}$ by

$$\mu_t^{B,\nu} = \mathbb{E}^\nu\Big(\frac{1}{t}\int_0^t \delta_{X_s^B}\,\mathrm{d}s\Big|\sigma_\tau^B > t\Big), \quad t > 0,\ \nu \in \mathcal{P}(M),$$

where $\mathcal{P}(M)$ stands for the set of all Borel probability measures on M.

NOTE: to avoid the situation that $\mathbb{P}^\nu(\sigma_\tau^B > t) = 0$, we should consider

$$\mathcal{P}_0(M) := \{\nu \in \mathcal{P}(M) : \ \nu(\mathring{M}) > 0\}.$$

Let

$$\mu_0 := \phi_0^2\mu,$$

which turns out to be the unique *quasi-ergodic distribution* of $(X_t^B)_{t\geq 0}$ for "nice" B, i.e., for every $\nu \in \mathcal{P}(M)$ and every Borel set $E \subset M$,

$$\lim_{t\to\infty} \mathbb{E}^\nu\Big[\frac{1}{t}\int_0^t \mathbb{1}_E(X_s^B)\,\mathrm{d}s\Big|\sigma_\tau^B > t\Big] = \mu_0(E).$$

Conditional empirical measures

Let

$$\sigma_\tau^B := \inf\{t \geq 0 : \; S_t^B > \tau\}.$$

Define the conditional empirical measures associated with $(X_t^B)_{t\geq 0}$ by

$$\mu_t^{B,\nu} = \mathbb{E}^\nu\left(\frac{1}{t}\int_0^t \delta_{X_s^B}\,\mathrm{d}s \,\middle|\, \sigma_\tau^B > t\right), \quad t > 0,\, \nu \in \mathcal{P}(M),$$

where $\mathcal{P}(M)$ stands for the set of all Borel probability measures on M.

NOTE: to avoid the situation that $\mathbb{P}^\nu(\sigma_\tau^B > t) = 0$, we should consider

$$\mathcal{P}_0(M) := \{\nu \in \mathcal{P}(M) : \; \nu(\mathring{M}) > 0\}.$$

Let

$$\mu_0 := \phi_0^2 \mu,$$

which turns out to be the unique *quasi-ergodic distribution* of $(X_t^B)_{t\geq 0}$ for "nice" B, i.e., for every $\nu \in \mathcal{P}(M)$ and every Borel set $E \subset M$,

$$\lim_{t\to\infty} \mathbb{E}^\nu\left[\frac{1}{t}\int_0^t \mathbb{1}_E(X_s^B)\,\mathrm{d}s \,\middle|\, \sigma_\tau^B > t\right] = \mu_0(E).$$

The aim

Recall that, for every $p \in [1, \infty)$,

$$W_p(\mu, \nu) = \Big(\inf_{\pi \in \mathcal{C}(\mu,\nu)} \int_{M\times M} \rho(x,y)^p \, \mathrm{d}\pi(x,y) \Big)^{1/p}, \quad \mu, \nu \in \mathcal{P}(M).$$

The aim is, *for every $\nu \in \mathcal{P}_0(M)$, as $t \to \infty$, to study the rate of convergence of $\mu_t^{B,\nu}$ to μ_0 under the quadratic Wasserstein distance*

$$W_2(\mu_t^{B,\nu}, \mu_0).$$

In particular, when $B(t) = t$, see F.-Y. Wang (2021).

The aim

Recall that, for every $p \in [1, \infty)$,

$$W_p(\mu, \nu) = \Big(\inf_{\pi \in \mathcal{C}(\mu,\nu)} \int_{M \times M} \rho(x, y)^p \, \mathrm{d}\pi(x, y) \Big)^{1/p}, \quad \mu, \nu \in \mathcal{P}(M).$$

The aim is, *for every $\nu \in \mathcal{P}_0(M)$, as $t \to \infty$, to study the rate of convergence of $\mu_t^{B,\nu}$ to μ_0 under the quadratic Wasserstein distance*

$$W_2(\mu_t^{B,\nu}, \mu_0).$$

In particular, when $B(t) = t$, see F.-Y. Wang (2021).

Outline

Some notations

Let $\alpha \in (0, 1]$ and

$$\mathbf{B}^{\alpha} := \left\{ B \in \mathbf{B} : \liminf_{\lambda \to \infty} \frac{B(\lambda)}{\lambda^{\alpha}} > 0 \right\},$$

$$\mathbf{B}_{\alpha} := \left\{ B \in \mathbf{B} : \limsup_{\lambda \to \infty} \frac{B(\lambda)}{\lambda^{\alpha}} < \infty \right\}.$$

Typical example:

$$B(t) = t^{\alpha}, \quad \alpha \in (0, 1].$$

For other examples, refer to Schilling–Song–Vondraček (2012).

RECALL: $\lambda_m, \phi_m, m \in \mathbb{N}_0$, are eigenvalues and eigenfunctions with Dirichlet boundary condition of the operator $-\mathcal{L}$ in $L^2(\mu)$.

Some notations

Let $\alpha \in (0, 1]$ and

$$\mathbf{B}^\alpha := \left\{ B \in \mathbf{B} : \liminf_{\lambda \to \infty} \frac{B(\lambda)}{\lambda^\alpha} > 0 \right\},$$

$$\mathbf{B}_\alpha := \left\{ B \in \mathbf{B} : \limsup_{\lambda \to \infty} \frac{B(\lambda)}{\lambda^\alpha} < \infty \right\}.$$

Typical example:

$$B(t) = t^\alpha, \quad \alpha \in (0, 1].$$

For other examples, refer to Schilling–Song–Vondraček (2012).

RECALL: $\lambda_m, \phi_m, m \in \mathbb{N}_0$, are eigenvalues and eigenfunctions with Dirichlet boundary condition of the operator $-\mathcal{L}$ in $L^2(\mu)$.

Upper bounds

Recall $\mu = e^U \text{vol}$. Let $B \in \mathbf{B}$ and $\nu \in \mathcal{P}_0(M)$. Set

$$\mathrm{I} := \frac{1}{\{\mu(\phi_0)\nu(\phi_0)\}^2} \sum_{m=1}^{\infty} \frac{\{\nu(\phi_0)\mu(\phi_m) + \mu(\phi_0)\nu(\phi_m)\}^2}{(\lambda_m - \lambda_0)[B(\lambda_m) - B(\lambda_0)]^2}.$$

Theorem (L.–Bingyao Wu)

Let $\alpha \in (0,1]$ and $\nu \in \mathcal{P}_0(M)$. If $B \in \mathbf{B}^\alpha$, then

$$\limsup_{t\to\infty}\{t^2 W_2(\mu_t^{B,\nu}, \mu_0)^2\} \le 4\mathrm{I} \in (0,\infty].$$

Moreover, if $B \in \mathbf{B}^\alpha \cap \mathbf{B}_\alpha$, then I is finite in either of the two cases:

(1) $d \le 2(1+2\alpha)$,

(2) $d > 2(1+2\alpha)$ *and* $\nu = h\mu$ *with* $h \in L^{2d/(d+2+4\alpha)}(\mu)$.

The rate of convergence is sharp!

The exact limit

Theorem (L.–Bingyao Wu)

Let $\alpha \in (0,1]$ *and* $B \in \mathbf{B}^\alpha$. *Then, for any* $\nu = h\mu \in \mathcal{P}_0(M)$ *with* $h\phi_0^{-1} \in L^p(\mu_0)$ *for some* $p \in (p_0, \infty]$,

$$\lim_{t\to\infty} \{t^2 W_2(\mu_t^{B,\nu}, \mu_0)^2\} = \mathrm{I},$$

where

$$p_0 := \frac{6(d+2)}{d+2+12} \vee \frac{3}{2}.$$

In the particular $B(t) = t$ case, F.-Y. Wang (2021) prove the limit for all $\nu \in \mathcal{P}_0(M)$. So, *how to drop the addition condition on* ν*?*

The exact limit

Theorem (L.–Bingyao Wu)

Let $\alpha \in (0,1]$ *and* $B \in \mathbf{B}^\alpha$. *Then, for any* $\nu = h\mu \in \mathcal{P}_0(M)$ *with* $h\phi_0^{-1} \in L^p(\mu_0)$ *for some* $p \in (p_0, \infty]$,

$$\lim_{t\to\infty}\{t^2 W_2(\mu_t^{B,\nu}, \mu_0)^2\} = \mathrm{I},$$

where

$$p_0 := \frac{6(d+2)}{d+2+12} \vee \frac{3}{2}.$$

In the particular $B(t) = t$ case, F.-Y. Wang (2021) prove the limit for all $\nu \in \mathcal{P}_0(M)$. So, *how to drop the addition condition on* ν*?*

Outline

The target

From now on, $B \in \mathbf{B}^\alpha$ for some $\alpha \in (0,1]$, which implies that

$$B(r) \gtrsim r^\alpha, \quad r > 1.$$

With some efforts (NOT TRIVIAL), we reduce to prove the following.

Proposition

For every $\nu \in \mathcal{P}_0(M)$ satisfying that $\nu = h\mu$ and $\|h\phi_0^{-1}\|_\infty < \infty$,

$$\limsup_{t\to\infty}\{t^2 W_2(\mu_t^{B,\nu}, \mu_0)^2\} \le 4\mathrm{I}.$$

Recall

$$\mathrm{I} = \frac{1}{\{\mu(\phi_0)\nu(\phi_0)\}^2} \sum_{m=1}^{\infty} \frac{\{\nu(\phi_0)\mu(\phi_m) + \mu(\phi_0)\nu(\phi_m)\}^2}{(\lambda_m - \lambda_0)[B(\lambda_m) - B(\lambda_0)]^2}.$$

The overall idea

Consider the Doob transform of $\mathcal{L}$, denoted by

$$\mathcal{L}_0 = \mathcal{L} + 2\nabla \log \phi_0,$$

and the corresponding semigroup $(P_t^0)_{t\geq 0}$.

FACT: $\{\phi_m \phi_0^{-1}\}_{m\in\mathbb{N}_0}$ is an eigenbasis of $-\mathcal{L}_0$ in $L^2(\mu_0)$.

To estimate $W_2(\mu_t^{B,\nu}, \mu_0)^2$, we apply the inequality (F.-Y. Wang 2022 or Ambrosio–Stra–Trevisan 2019 or Ledoux 2017): for every $f \geq 0$ with $\mu_0(f) = 1$,

$$W_2(f\mu_0, \mu_0)^2 \leq 4 \int_M |\nabla(-\mathcal{L}_0)^{-1}(f-1)|^2 \, \mathrm{d}\mu_0.$$

The overall idea

Consider the Doob transform of $\mathcal{L}$, denoted by

$$\mathcal{L}_0 = \mathcal{L} + 2\nabla \log \phi_0,$$

and the corresponding semigroup $(P_t^0)_{t\geq 0}$.

FACT: $\{\phi_m \phi_0^{-1}\}_{m\in\mathbb{N}_0}$ is an eigenbasis of $-\mathcal{L}_0$ in $L^2(\mu_0)$.

To estimate $W_2(\mu_t^{B,\nu}, \mu_0)^2$, we apply the inequality (F.-Y. Wang 2022 or Ambrosio–Stra–Trevisan 2019 or Ledoux 2017): for every $f \geq 0$ with $\mu_0(f) = 1$,

$$W_2(f\mu_0, \mu_0)^2 \leq 4 \int_M |\nabla(-\mathcal{L}_0)^{-1}(f-1)|^2 \, \mathrm{d}\mu_0.$$

Step 1: calculate the Radon–Nikodym derivative

$$\frac{\mathrm{d}\mu_t^{B,\nu}}{\mathrm{d}\mu_0} = \rho_t^{B,\nu} - \mathrm{A}_t + \frac{1}{\mathbb{P}^\nu(t < \sigma_\tau^B)} \int_0^t \xi_s \, \mathrm{d}s + 1,$$

where

$$\rho_t^{B,\nu} := \frac{e^{-B(\lambda_0)t}}{t\mathbb{P}^\nu(t < \sigma_\tau^B)} \sum_{m=1}^{\infty} \frac{\mu(\phi_0)\nu(\phi_m) + \nu(\phi_0)\mu(\phi_m)}{B(\lambda_m) - B(\lambda_0)} \phi_m \phi_0^{-1},$$

$$\mathrm{A}_t := \frac{1}{t\mathbb{P}^\nu(t < \sigma_\tau^B)} \sum_{m=1}^{\infty} \frac{\mu(\phi_0)\nu(\phi_m) + \nu(\phi_0)\mu(\phi_m)}{B(\lambda_m) - B(\lambda_0)} e^{-B(\lambda_m)t} \phi_m \phi_0^{-1},$$

$$\xi_s := \left(\sum_{m=1}^{\infty} e^{-B(\lambda_m)s} \nu(\phi_m) \phi_m \phi_0^{-1} \right) \left(\sum_{n=1}^{\infty} e^{-B(\lambda_n)(t-s)} \mu(\phi_n) \phi_n \phi_0^{-1} \right) - \sum_{m=1}^{\infty} e^{-B(\lambda_m)t} \mu(\phi_m)\nu(\phi_m), \quad 0 < s \le t.$$

Step 2: apply Ledoux's inequality

For any $\epsilon > 0$, we have

$$\begin{aligned}
W_2(\mu_t^{B,\nu}, \mu_0)^2 &\leq 4 \int_M \left| \nabla(-\mathcal{L}_0)^{-1} \left(\frac{\mathrm{d}\mu_t^{B,\nu}}{\mathrm{d}\mu_0} \right) \right|^2 \mathrm{d}\mu_0 \\
&\leq 4(1+\epsilon) \int_M |\nabla(-\mathcal{L}_0)^{-1} \rho_t^{B,\nu}|^2 \, \mathrm{d}\mu_0 \\
&\quad + 8(1+\epsilon^{-1}) \int_M |\nabla(-\mathcal{L}_0)^{-1} \mathrm{A}_t|^2 \, \mathrm{d}\mu_0 \\
&\quad + 8(1+\epsilon^{-1}) \int_M \left| \nabla(-\mathcal{L}_0)^{-1} \frac{1}{t\mathbb{P}^\nu(t < \sigma_\tau^B)} \int_0^t \xi_s \, \mathrm{d}s \right|^2 \mathrm{d}\mu_0 \\
&=: 4(1+\epsilon) J_1 + 8(1+\epsilon^{-1}) J_2 + 8(1+\epsilon^{-1}) J_3,
\end{aligned}$$

where the triangle inequality of $\| \cdot \|_{L^2(\mu_0)}$ was also employed.

Step 3: calculate J_1 and J_2

Since

$$-\mathcal{L}_0(\phi_m\phi_0^{-1}) = (\lambda_m - \lambda_0)\phi_m\phi_0^{-1}, \ \ \|\phi_m\phi_0^{-1}\|_{L^2(\mu_0)} = 1, \quad m \in \mathbb{N},$$

by the integration-by-parts formula, we have

$$\int_M |\nabla(-\mathcal{L}_0)^{-1}(\phi_m\phi_0^{-1})|^2 \, \mathrm{d}\mu_0 = \frac{1}{\lambda_m - \lambda_0}, \quad m \in \mathbb{N}.$$

Then

$$J_1 = \frac{e^{-2B(\lambda_0)t}}{t^2\mathbb{P}^\nu(t < \sigma_\tau^B)^2} \sum_{m=1}^{\infty} \frac{[\mu(\phi_0)\nu(\phi_m) + \nu(\phi_0)\mu(\phi_m)]^2}{(\lambda_m - \lambda_0)[B(\lambda_m) - B(\lambda_0)]^2}, \ \ (\textit{dominant term})$$

$$J_2 = \frac{1}{t^2\mathbb{P}^\nu(t < \sigma_\tau^B)^2} \sum_{m=1}^{\infty} \frac{[\mu(\phi_0)\nu(\phi_m) + \nu(\phi_0)\mu(\phi_m)]^2}{(\lambda_m - \lambda_0)[B(\lambda_m) - B(\lambda_0)]^2} e^{-2B(\lambda_m)t}.$$

Step 4: estimate J_3

Since $(-\mathcal{L}_0)^{-\frac{1}{2}} = \frac{2}{\sqrt{\pi}} \int_0^\infty P^0_{s^2}\,\mathrm{d}s$, by Minkowski's inequality, we get

$$J_3 = \frac{1}{t^2 \mathbb{P}^\nu(t < \sigma_\tau^B)^2} \int_M \left| \frac{2}{\sqrt{\pi}} \int_0^\infty \int_0^t P^0_{r^2} \xi_s \,\mathrm{d}s\mathrm{d}r \right|^2 \mathrm{d}\mu_0$$

$$\leq \frac{4}{\pi t^2 \mathbb{P}^\nu(t < \sigma_\tau^B)^2} \left(\int_0^\infty \int_0^t \|P^0_{r^2}\xi_s\|_{L^2(\mu_0)} \,\mathrm{d}s\mathrm{d}r \right)^2, \quad t > 0.$$

Applying the fact that, there exists a constant $\eta > 0$ such that, for every $t \geq 0$, $p \in [1, \infty]$ and $f \in L^p(\mu_0)$,

$$\|(P_t^0 - \mu_0)f\|_{L^p(\mu_0)} \leq \eta e^{-(\lambda_1 - \lambda_0)t} \|f\|_{L^p(\mu_0)},$$

we have

$$\|P^0_{r^2}\xi_s\|_{L^2(\mu_0)} \leq 2\eta \|h\phi_0^{-1}\|_\infty e^{-B(\lambda_1)t} e^{-(\lambda_1 - \lambda_0)r^2}, \quad r, s > 0.$$

Thus, by $B(r) \gtrsim r^\alpha$ for $r > 1$, there exists a constant $c > 0$ such that

$$J_3 \leq c \|h\phi_0^{-1}\|_\infty^2 e^{-2[B(\lambda_1) - B(\lambda_0)]t}, \quad t > 97.$$

Final step

Putting these estimates together, we find a constants $c > 0$ such that

$$t^2 W_2(\mu_t^{B,\nu}, \mu_0)^2 \leq \frac{4(1+\epsilon)e^{-2B(\lambda_0)t}}{\mathbb{P}^\nu(t < \sigma_\tau^B)^2} \sum_{m=1}^{\infty} \frac{[\mu(\phi_0)\nu(\phi_m) + \nu(\phi_0)\mu(\phi_m)]^2}{(\lambda_m - \lambda_0)[B(\lambda_m) - B(\lambda_0)]^2}$$

$$+ c(1+\epsilon^{-1}) \sum_{m=1}^{\infty} \frac{[\mu(\phi_0)\nu(\phi_m) + \nu(\phi_0)\mu(\phi_m)]^2}{(\lambda_m - \lambda_0)[B(\lambda_m) - B(\lambda_0)]^2} e^{-2[B(\lambda_m) - B(\lambda_0)]t}$$

$$+ c(1+\epsilon^{-1}) \|h\phi_0^{-1}\|_\infty^2 e^{-2[B(\lambda_1) - B(\lambda_0)]t}, \quad t > 997,\ \epsilon > 0.$$

Due to that

$$\lim_{t\to\infty} \{e^{B(\lambda_0)t} \mathbb{P}^\nu(t < \sigma_\tau^B)\} = \mu(\phi_0)\nu(\phi_0),$$

letting $t \to \infty$ first and then $\epsilon \to 0^+$, we finally arrive at

$$\limsup_{t\to\infty} \{t^2 W_2(\mu_t^{B,\nu}, \mu_0)^2\} \leq 4\mathrm{I}.$$

Final step

Putting these estimates together, we find a constants $c > 0$ such that

$$t^2 W_2(\mu_t^{B,\nu}, \mu_0)^2 \leq \frac{4(1+\epsilon)e^{-2B(\lambda_0)t}}{\mathbb{P}^\nu(t < \sigma_\tau^B)^2} \sum_{m=1}^{\infty} \frac{[\mu(\phi_0)\nu(\phi_m) + \nu(\phi_0)\mu(\phi_m)]^2}{(\lambda_m - \lambda_0)[B(\lambda_m) - B(\lambda_0)]^2}$$

$$+ c(1+\epsilon^{-1}) \sum_{m=1}^{\infty} \frac{[\mu(\phi_0)\nu(\phi_m) + \nu(\phi_0)\mu(\phi_m)]^2}{(\lambda_m - \lambda_0)[B(\lambda_m) - B(\lambda_0)]^2} e^{-2[B(\lambda_m) - B(\lambda_0)]t}$$

$$+ c(1+\epsilon^{-1})\|h\phi_0^{-1}\|_\infty^2 e^{-2[B(\lambda_1) - B(\lambda_0)]t}, \quad t > 997,\ \epsilon > 0.$$

Due to that

$$\lim_{t\to\infty} \{e^{B(\lambda_0)t}\mathbb{P}^\nu(t < \sigma_\tau^B)\} = \mu(\phi_0)\nu(\phi_0),$$

letting $t \to \infty$ first and then $\epsilon \to 0^+$, we finally arrive at

$$\limsup_{t\to\infty} \{t^2 W_2(\mu_t^{B,\nu}, \mu_0)^2\} \leq 4\mathrm{I}.$$

" Thanks! "