Analysis on free Riemannian path spaces

Shizan Fang a, Feng-Yu Wang b, *

a I.M.B. B.P. 47870, Université de Bourgogne, Dijon, France
b School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China

Received 15 October 2004; accepted 8 November 2004
Available online 13 January 2005

Abstract

The gradient operator is defined on the free path space with reference measure $P_{\mu}$, the law of the Brownian motion on the base manifold with initial distribution $\mu$, where $\mu$ has strictly positive density w.r.t. the volume measure. The formula of integration by parts is established for the underlying directional derivatives, which implies the closability of the gradient operator so that it induces a conservative Dirichlet form on the free path space. The log-Sobolev inequality for this Dirichlet form is established and, consequently, the transportation cost inequality is obtained for the associated intrinsic distance.

Keywords: Log-Sobolev inequality; Transportation cost inequality; Free path space; Integration by parts formula

1. Introduction

Let $(M, \langle \cdot, \cdot \rangle)$ be a $d$-dimensional connected complete Riemannian manifold, $TM$ the bundle of tangent spaces of $M$, and $\mu$ a probability measure on $M$ having strictly positive density w.r.t. the volume measure. For any smooth, strictly positive definite map-
ping $A: TM \to TM$ such that the induced Riemannian metric $\langle X, Y \rangle_A := \langle A^{-1}X, Y \rangle$, $X, Y \in T_x M$, $x \in M$ is complete, the quadratic form
\[
\delta(f, g) := \int_M \langle A\nabla f, \nabla g \rangle d\mu, \quad f, g \in C^1_0(M),
\] (1.1)
is closable in $L^2(M, \mu)$ (due to the integration by parts formula) and the closure $(\mathcal{E}, D(\mathcal{E}))$ is a conservative Dirichlet form on $L^2(M, \mu)$ (due to the completeness of the induced metric, see e.g. [19]), where $\nabla$ is the gradient operator on $M$. Moreover, the Riemannian distance induced by the metric $\langle ., \rangle_A$ coincides with the intrinsic distance induced by the Dirichlet form, that is, the Riemannian distance has the representation
\[
\rho_A(x, y) = \sup \{\|f(x) - f(y)\|: f \in C^1_b(M), \langle \nabla_A f, \nabla_A f \rangle_A := \langle A\nabla f, \nabla f \rangle \leq 1\},
\] (1.2)
where $\nabla_A$ is the gradient operator associated to the metric $\langle ., \rangle_A$. This follows from the mean-value theorem and the fact that the Riemannian distance function has unit gradient.

Suppose that $\mu$ satisfies the following logarithmic Sobolev inequality
\[
\mu(f^2 \log f^2) \leq C \mu(\langle A\nabla f, \nabla f \rangle), \quad f \in C^1_b(M), \quad \mu(f^2) = 1,
\] (1.3)
where and in the sequel, $\mu(f) := \int_M f d\mu$ for a $\mu$-integrable function $f$. Then by using the evolution along the heat semi-group as in [16] or in [18], as well as the evolution along Hamilton–Jacobi semi-group as in [3], the inequality (1.3) implies the following quadratic transportation cost inequality
\[
W^2_{2, \rho_A}(f \mu, \mu) \leq C \mu(f \log f), \quad f \geq 0, \quad \mu(f) = 1,
\] (1.4)
where
\[
W^2_{2, \rho_A}(f \mu, \mu) := \inf_{\tilde{\mu} \in \mathcal{C}(f \mu, \mu)} \int_{M \times M} \rho^2_A(x, y) \tilde{\mu}(dx, dy)
\] (1.5)
for $\mathcal{C}(f \mu, \mu)$ the totality of Borel probability measures on $M \times M$ having $f \mu$ and $\mu$ as marginal laws. When $A = \text{Id}$ and $\mu$ is the standard Gaussian measure on $\mathbb{R}^d$, (1.4) is originally due to [17]. Its generalization to Wiener spaces has been done in [8,9].

The situation is however quite different when the path space
\[
P_x(M) := \{\gamma \in C([0,T]: M) : \gamma(0) = x\}
\]
is considered. Firstly the definition of "Riemannian distance" on $P_x(M)$ is not possible and the intrinsic distance associated to the canonical Dirichlet form on $P_x(M)$ may take values $+\infty$. Secondly neither Radmancher theorem nor semi-group techniques for the Ornstein–Uhlenbeck semi-group on $P_x(M)$ are available. So, the direct passage from log-Sobolev inequalities to transportation cost inequalities on $P_x(M)$ is not yet possible. Nevertheless, by using a passage throughout finite-dimensional manifolds, this difficulty has been overcome in [18].

The main purpose of this work is to discuss these problems on the free Riemannian path space
\[
P(M) := C([0, T]; M)
\]
endowed with a diffusion law $P_\mu := \int_M P_x \, d\mu(x)$, where $\mu$ is a probability measure on $M$. It turns out that in this case known results on $P_x(M)$ do not apply automatically, since the differential structure on $P_x(M)$ is dependent of the law $P_x$ (due to Itô stochastic parallel transport) and $P_x$ is singular to $P_y$ for $x \neq y$. Therefore, due to the freedom of the starting point, the gradient operator as well as the integration by parts formula have to be reformulated.

The organization of the paper is as follows. In Section 2 we first introduce the gradient operator on $P(M)$. When the manifold $M$ is not parallelized, the free path space $P(M)$ should be not parallelized, contrary to the case of based path spaces. To prove the closability of the gradient operator, a formula of integration by parts is established. When the initial law $\mu$ is the Riemannian measure, such a formula of integration by parts was proved in [13]. Instead of using chaos expansion as in [13], we shall make use of the Girsanov theorem to construct quasi-invariant transformations on $P(M)$ using ideas in [2,7]. We refer to [12] for invariant Sobolev calculus on free loop spaces. In Section 3, we prove the log-Sobolev inequality for $P_\mu$ under the hypothesis that the initial law $\mu$ satisfies a log-Sobolev inequality on $M$. This hypothesis is indeed necessary. In the last section, we follow the idea of [18,19] to derive the transportation cost inequality on $P(M)$ for the associated intrinsic distance.

2. Gradient operator and integration by parts

From now on, we assume that the Ricci curvature of $(M, \langle \cdot, \cdot \rangle)$ is bounded, which in particular implies the stochastic completeness of the manifold. Let $O(M)$ be the bundle of orthonormal frames of $M$, that is, an element $r \in O(M)$ is an isometry from $\mathbb{R}^d$ onto the tangent space $T_{\pi(r)}M$ where $\pi : O(M) \to M$ is the natural projection. For a vector field $X$ on $M$, we denote by $\tilde{X}$ the horizontal lift (with respect to Levi-Civita connection) of $X$ to $O(M)$ which satisfies the relation $\pi'(r)\tilde{X}(r) = X(\pi(r))$.

For fixed $T > 0$, let $W_0(\mathbb{R}^d)$ be the space of continuous functions from $[0, T]$ into $\mathbb{R}^d$, vanishing at the origin. Let $P^W$ be the Wiener measure on $W_0(\mathbb{R}^d)$. Next, let $\mu$ be a fixed probability measure on $M$ such that

$$d\mu = v(x) \, dx \quad \text{for } v \text{ strictly positive with } |\nabla \log v| \in L^2_{\text{loc}}(\mu), \tag{2.1}$$

where $dx$ is the Riemannian volume measure on $M$. For $\tilde{r}$ a fixed Borel section of $\pi$, that is, $\tilde{r}(x) \in O_x(M)$ is Borel measurable in $x$.

We now consider the space $\Omega := W_0(\mathbb{R}^d) \times M$ endowed with the product Borel $\sigma$-field $\mathcal{F}$ and the product measure $P = P^W \times \mu$. Define $b_t : \Omega \to \mathbb{R}^d$ by $b_t(\omega) = w_t$ and $x_0 : \Omega \to M$ by $x_0(\omega) = x$, where $\omega = (w, x)$. Then $(b_t(\omega))_{t \geq 0}$ is a Brownian motion independent of the random variable $x_0$. Let $\tilde{\mathcal{F}}_t$ be the complete $\sigma$-field on $\Omega$ generated by $\{b_s(\cdot) : s \leq t\}$ and $x_0$. Then $(b_t)_{t \geq 0}$ is a $\tilde{\mathcal{F}}_t$-Brownian motion. In the sequel, always this probability space $(\Omega, P, \mathcal{F}, \mathcal{F}_t)$ and the initial random variable $r_0 = \tilde{r}(x_0)$ are considered.
Let \( \{H_1, \ldots, H_d\} \) be the canonical horizontal vector fields on \( O(M) \). Consider the following Stratanovich stochastic differential equation

\[
dr_t = \sum_{i=1}^{d} H_i(r_t) \circ db^i_t, \quad r_0 = \bar{r}(x_0),
\]

(2.2)

where \( x_0 \) is the above defined random variable on \( M \) with distribution \( \mu \). Let \( x_t(\omega) := \pi (r_t(\omega)) \).

(2.3)

Then \( x_t \) is a Brownian motion on \( M \) with initial distribution \( \mu \). We denote by \( P_\mu \) the law of \( x_\cdot \) on the path space \( \mathcal{P}(M) \).

On the other hand, since \( r_t \) is the horizontal lift of \( x_t \), that is,

\[
dr_t = H(r_t) \circ dx_t, \quad r_0 = \bar{r}(x_0),
\]

where \( H \) is the horizontal lift operator, we conclude that \( r_t \) is measurable with respect to \( \mathcal{F}^x_t := \sigma(x_s; s \leq t) \).

To define the gradient operator on the Brownian path space, let us first recall the procedure made for the case with fixed initial point \( x \).

Let \( H_0 := \{h \in C([0, T]; \mathbb{R}^d) : h(0) = 0, |h|_{H_0} := \int_0^T |\dot{h}(s)|^2 ds < +\infty \} \) be the Cameron–Martin space on the flat path space. For a given Brownian path \( x_\cdot \) with horizontal lift \( r_\cdot \) and for any \( h_0 \in H_0 \), define the geodesic flow

\[
x_{h_0 \cdot}^t := \exp_{x_t} \left[ \varepsilon r_t h_0(t) \right], \quad \varepsilon \geq 0, \quad t \in [0, T].
\]

Then the directional derivative along \( h_0 \) of a function \( F \) on \( \mathcal{P}_x(M) \) is defined by

\[
D_{h_0}^0 F(x_\cdot)(t) := \left. \frac{d}{d\varepsilon} F(x_{h_0 \cdot}^t) \right|_{\varepsilon=0}
\]

as soon as the right-hand side exists. In particular, if \( F \) is a cylinder function (denoted by \( F \in \mathcal{F}C_0^\infty \)), that is, if there exist \( 0 \leq t_1 < \cdots < t_N \leq T \) and \( f \in C_0^\infty (M^N) \) such that

\[
F(x_\cdot) = f(x_{t_1}, \ldots, x_{t_N}), \quad x_\cdot \in \mathcal{P}_x(M),
\]

(2.4)

then

\[
D_{h_0}^0 F(x_\cdot) = \sum_{i=1}^{N} \left[ \nabla_i f, r_{t_i} h_0(t_i) \right], \quad x_\cdot \in \mathcal{P}_x(M),
\]

where \( \nabla_i \) is the gradient operator on the \( i \)-th component. Consequently, the gradient \( D^0 F(x_\cdot) \) is an \( H_0 \)-valued random variable defined through

\[
\langle D^0 F(x_\cdot), h_0 \rangle_{H_0} = D_{h_0}^0 F(x_\cdot), \quad h_0 \in H_0.
\]

Thus, we have the following representation

\[
D^0 F(x_\cdot)(t) = \sum_{i=1}^{N} (t \wedge t_i)^{-1} \nabla_i f.
\]

(2.5)
Note that the first term in the summation vanishes if \( t_1 = 0 \). Furthermore, the integration by parts formula established in [5] for \( D^0 \) ensures the closability of \((D^0, \mathcal{F}C^1_0)\) on \( L^2(\Omega \to H_0; P) \), so that it provides the canonical Dirichlet form on \( L^2(P_x(M); P_x) \), where \( P_x \) is the law of the Brownian motion on \( M \) starting from \( x \).

Now, we consider the free path space. Due to the freedom of the initial point, it is natural for us to make use of the following Cameron–Martin space:

\[
\mathcal{H} := \left\{ h \in C([0, T]; \mathbb{R}^d) : \int_0^T |\dot{h}(s)|^2 \, ds < \infty \right\},
\]

(2.6)

which is a Hilbert space under the inner product

\[
\langle h_1, h_2 \rangle_{\mathcal{H}} := \langle h_1(0), h_2(0) \rangle_{\mathbb{R}^d} + \langle h_1 - h_1(0), h_2 - h_2(0) \rangle_{H_0}.
\]

For given \( h \in \mathcal{H} \), the corresponding directional derivative of a good function \( F \) along \( h \) is defined by

\[
D_h F(x \cdot) := \left\{ \frac{d}{d\varepsilon} F(\exp_{x} \left[ \varepsilon r_t(h(t)) \right]) \right\}_{\varepsilon = 0}.
\]

Thus, the gradient \( DF(x \cdot) \) can be fixed as an \( \mathcal{H} \)-valued random variable through

\[
\langle DF(x \cdot), h \rangle_{\mathcal{H}} = D_h F(x \cdot), \quad h \in \mathcal{H}.
\]

In particular, for \( F \in \mathcal{F}C^\infty_0 \) given by (2.4), one has

\[
D_h F(x \cdot) = \sum_{i=1}^{N} [\nabla_i f, r_t h(t_i)], \quad h \in \mathcal{H},
\]

and hence,

\[
DF(x)(t) = \sum_{i=1}^{N} \left( 1 + (t \wedge t_i) \right) r_t^{-1} \nabla_i f = \sum_{i=1}^{N} r_t^{-1} \nabla_i f + D^0 F(x)(t).
\]

(2.7)

Since \( r_t \) is \( \mathcal{F}_{s_t}^N \)-measurable, \( DF \in L^2(P(M) \to \mathcal{H}; P_{\mu}) \) with

\[
\delta(F, F) := \int_{P(M)} |DF|^2 \, dP_{\mu} = \mathbb{E} \left[ \sum_{i=1}^{N} r_t^{-1} \nabla_i f \right]^2 + |D^0 F|^2_{H_0} < \infty.
\]

(2.8)

The fact that we can see the gradient operator \( DF \) as a \( \mathcal{H} \)-valued random variable is due to the choice of a Borel section \( \tilde{r}_\alpha \) of \( \pi : O(M) \to M \). Nevertheless when \( M \) is not parallelized, we can not see \( \mathcal{H} \) as tangent spaces of \( P(M) \), in contrast of based path spaces. The reason is that the directions \( h \in \mathcal{H} \) are not admissible for the measure \( P_{\mu} \); this means that the integration by parts formula does not hold for \( D_h F \) when \( h \in \mathcal{H} \). Let’s introduce

\[
TP(M) := \left\{ \psi \in C([0, T] \times P(M); T_M) : \psi(t) \in T_{\pi \cdot M} \text{ is } \mathcal{F}_{s_t}^N \text{-adapted}, \right.

\[
\left. r_t^{-1} \psi \in L^2(P(M) \to \mathcal{H}; P_{\mu}) \right\}.
\]

We shall establish the integration by parts formula for directional derivatives along a large enough class of vectors in \( TP(M) \).
Let \( \mathcal{H}_0(M) \) be the space of smooth vector fields on \( M \) with compact supports, and let \( \tilde{\mathcal{H}}_0 := \{ h_0 \in L^2(\Omega \to H_0; \mathcal{P}); h_0(t) \text{ is } \mathcal{F}_t^{\mathcal{P}} \text{-adapted} \} \).

To \( (h_0, X) \in \tilde{\mathcal{H}}_0 \times \mathcal{H}_0(M) \), we associate
\[
\psi_{h_0,X}(t) := h_0(t) + r_0^{-1} X(x_0).
\]
We have \( r_0 \psi_{h_0,X} \in T P(M) \).

Let \( \Omega_r \) be the curvature tensor of \( M \) under the frame \( r \in O(M) \). Define
\[
\Gamma(t) := \int_0^t \Omega_r (\psi_{h_0,X}(s), \circ dB_s).
\]
(2.9)

Let \( \hat{h}(t) := \dot{\psi}_{h_0,X}(t) + \frac{1}{2} \text{Ric}_r(t) \psi_{h_0,X}(t), \hat{h}(0) = 0 \), where \( \text{Ric}_r \) is the Ricci tensor at the frame \( r \). For any \( \varepsilon > 0 \), let
\[
M_t^{h_0,X,\varepsilon} := \exp \left\{-\varepsilon \int_0^t \langle e^{-\varepsilon \Gamma(s)} \hat{h}(s), dB_s \rangle - \frac{\varepsilon^2}{2} \int_0^t |\hat{h}(s)|^2 \, ds \right\}.
\]
(2.10)

Since \( \Omega_r \) is skew symmetric, \( e^{\alpha \Gamma(s)} \) is orthogonal for any \( \alpha \in \mathbb{R} \). Then \( M_t^{h_0,X,\varepsilon} \) is a \( \mathcal{F}_t \)-martingale. By Girsanov theorem,
\[
B_t^{h_0,X,\varepsilon} := \int_0^t e^{\varepsilon \Gamma(s)} dB_s + \varepsilon \hat{h}(t)
\]
defines a Brownian motion under the probability measure
\[
Q^{h_0,X,\varepsilon} := M_T^{h_0,X,\varepsilon} P.
\]
(2.11)

Consider the Stratanovich stochastic differential equation on \( O(M) \):
\[
dt r_t^{h_0,X,\varepsilon} = \sum_{i=1}^d H_i (r_t^{h_0,X,\varepsilon}) \circ dB_t^{h_0,X,\varepsilon,i}, \quad r_0^{h_0,X,\varepsilon} = \tilde{\varphi}_\varepsilon (r_0),
\]
(2.12)

where \( \tilde{\varphi}_\varepsilon : O(M) \to O(M) \) is the horizontal flow such that \( \frac{d\tilde{\varphi}_\varepsilon}{dt} = \tilde{X}(\tilde{\varphi}_\varepsilon) \) and \( \tilde{X} \) is the horizontal lift of \( X \) to \( O(M) \). Let
\[
x_t^{h_0,X,\varepsilon} := \pi (r_t^{h_0,X,\varepsilon}).
\]

**Proposition 2.1.** Let \( \mu_\varepsilon \) be the induced measure on \( M \) of \( Q^{h_0,X,\varepsilon} \) by the map \( \omega \mapsto x_0^{h_0,X,\varepsilon}(\omega) \). Then
\[
\mu_\varepsilon = (\varphi_\varepsilon)^* \mu
\]
(2.13)

where \( \varphi_\varepsilon : M \to M \) is the flow associated to \( X \).
Proof. Remark first
\[ \varphi_\varepsilon(x_0) = x_0^{h_0,X,\varepsilon} = \pi(\tilde{\varphi}_\varepsilon(r_0)) . \] (2.14)

In fact, we have
\[ \frac{d}{d\varepsilon} \pi(\tilde{\varphi}_\varepsilon(r_0)) = \pi'(\tilde{\varphi}_\varepsilon(r_0)) \cdot \tilde{X}(\tilde{\varphi}_\varepsilon(r_0)) = X(\pi(\tilde{\varphi}_\varepsilon(r_0))) \]
and \( \pi(\tilde{\varphi}_\varepsilon(r_0)) = \pi(r_0) = x_0 \). By the uniqueness of solutions, we obtain (2.14). For any \( f \in C_b(M) \), since \( M^{h_0,X,\varepsilon}_T \) is a martingale, we have
\[ \int_M f \, d\mu_\varepsilon = \mathbb{E}(f(\varphi_\varepsilon(x_0)) M^{h_0,X,\varepsilon}_T) = \mathbb{E}(f(\varphi_\varepsilon(x_0)) \mathbb{E}_0(M^{h_0,X,\varepsilon}_T)) \]
\[ = \mathbb{E}(f(\varphi_\varepsilon(x_0))) = \int_M f(\varphi_\varepsilon) \, d\mu . \]
Therefore we get (2.13).

Proposition 2.2. Let \( Q^{h_0,X,\varepsilon}_\mu \) be the law of \( x_0^{h_0,X,\varepsilon} \) on \( \mathbb{P}(M) \) under \( Q^{h_0,X,\varepsilon} \). Then \( Q^{h_0,X,\varepsilon}_\mu \) is absolutely continuous with respect to \( P_\mu \).

Proof. Let \( \{ P_x : x \in M \} \) be the diffusion system associated to \( x \), and \( \{ \tilde{P}_x : x \in M \} \) associated to \( x^{h_0,X,\varepsilon} \), but under the probability \( Q^{h_0,X,\varepsilon} \). By the uniqueness of laws, \( P_x = \tilde{P}_x \) for all \( x \in M \). Now let \( A \subset \mathbb{P}(M) \) such that \( P_\mu(A) = 0 \). This means that \( \int_M P_x(A) \, d\mu(x) = 0 \). Then there exists \( U \subset M \) of \( \mu(U) = 0 \) such that
\[ P_x(A) = 0, \quad x \notin U. \]
By (2.13) and (2.1), we see that \( \mu_\varepsilon \) is absolutely continuous with respect to \( \mu \). Then \( \mu_\varepsilon(U) = 0 \). Therefore,
\[ Q^{h_0,X,\varepsilon}_\mu(A) = \int_M \tilde{P}_x(A) \, d\mu_\varepsilon(x) = \int_M P_x(A) \, d\mu_\varepsilon(x) = 0. \]
The proof is complete.

Theorem 2.3. Let \( P_{\mu_\varepsilon} \) be the law of \( x_0^{h_0,X,\varepsilon} \) under \( P \). Then \( P_{\mu_\varepsilon} \) is absolutely continuous with respect to \( P \).

Proof. Let \( A \subset \mathbb{P}(M) \) such that \( P_{\mu_\varepsilon}(A) = 0 \). By Proposition 2.2,
\[ Q^{h_0,X,\varepsilon}_\mu(A) = \int_\Omega 1_A(x_0^{h_0,X,\varepsilon}) M^{h_0,X,\varepsilon}_T \, dP = 0. \]
It follows that \( 1_A(x_0^{h_0,X,\varepsilon}) = 0 \) for \( P \)-a.s. Hence \( P_{\mu_\varepsilon}(A) = \int_\Omega 1_A(x_0^{h_0,X,\varepsilon}) \, dP = 0. \)
Theorem 2.4. For \( h_0 \in \tilde{H}_0 \) and \( X \in \mathcal{D}_0(M) \),
\[
\left\{ \frac{d}{d\varepsilon} \lambda_{t}^{h_0,X,\varepsilon} \right\}_{\varepsilon = 0} = r_t \psi_{h_0,X}(t). \quad (2.15)
\]

**Proof.** Let \((\theta, \Theta)\) be the absolute parallelism on \( O(M) \). This means that \((\theta, \Theta)\) is \( \mathbb{R}^d \times \text{so}(d) \)-valued 1-differential form on \( O(M) \), which satisfies the following structure equation
\[
\begin{align*}
    d\theta &= -\Theta \wedge \theta, \\
    d\Theta &= -\Theta \wedge \Theta + \Omega,
\end{align*}
\]
where \( \Omega \) denotes the curvature tensor. Recall that \( r_t^{h_0,X,\varepsilon} \) is the solution of stochastic differential equation (2.12). Let
\[
\beta(t) = \langle \theta, \left\{ \frac{d}{d\varepsilon} r_t^{h_0,X,\varepsilon} \right\}_{\varepsilon = 0} \rangle, \\
q(t) = \langle \Theta, \left\{ \frac{d}{d\varepsilon} r_t^{h_0,X,\varepsilon} \right\}_{\varepsilon = 0} \rangle.
\]
We have \( \left\{ \frac{d}{d\varepsilon} r_t^{h_0,X,\varepsilon} \right\}_{\varepsilon = 0} = \left\{ \frac{d}{d\varepsilon} \tilde{\psi}_{t}(r_0) \right\}_{\varepsilon = 0} = \tilde{X}(r_0) \). Therefore,
\[
\beta(0) = r_0^{-1}X(x_0), \quad q(0) = 0.
\]
Now proceeded similarly as in [2,5,7], \((\beta(t), q(t))\) satisfies the following relations
\[
\begin{align*}
    d\beta(t) &= \dot{\beta}(t) dt - \frac{1}{2} \text{Ric}_t \beta(t) dt, \\
    dq(t) &= -\Omega_{r_t}(\psi_{h_0,X}(t), \circ db_t).
\end{align*}
\]
From the first equation, we see that \( \dot{\beta}(t) = \hat{h}(t) \). Let \( u(t) = \beta(t) - \psi_{h_0,X}(t) \). Then \( u \) satisfies the linear differential equation
\[
\dot{u}(t) + \frac{1}{2} \text{Ric}_t u(t) = 0, \quad u(0) = 0,
\]
which implies that \( u = 0 \). Therefore \( \beta(t) = \psi_{h_0,X}(t) \) and hence,
\[
\pi'(r_t) \left\{ \frac{d}{d\varepsilon} r_t^{h_0,X,\varepsilon} \right\}_{\varepsilon = 0} = r_t \beta(t) = r_t h_0(t) + r_t (r_0^{-1}X(x_0)).
\]
The proof of (2.15) is complete. \( \square \)

Now let \( F \) be a cylinder function on \( \mathbf{P}(M) \). We denote
\[
D_{\psi_{h_0,X}} F = \left\{ \frac{d}{d\varepsilon} F(\lambda_{t}^{h_0,X,\varepsilon}) \right\}_{\varepsilon = 0}.
\]
To state the formula of integration by parts, let us recall the notion of divergence \( \text{div}_\mu(X) \) of a smooth vector field \( X \) with respect to the measure \( \mu \). Let \( \varphi_\varepsilon \) be the flow associated to \( X \). By (2.1), \( (\varphi_\varepsilon)_* \mu \) is absolutely continuous with respect to \( \mu \). Let \( k_\varepsilon \) be the density. Then
\[
\text{div}_\mu(X) := -\left\{ \frac{dk_\varepsilon}{d\varepsilon} \right\}_{\varepsilon = 0} = \text{div}(X) + \langle \nabla \log v, X \rangle.
\]
exists in $L^2(\mu)$ according to (2.1). Thus, the integration by parts formula on $M$ reads
\[ \int_M \langle \nabla f, X \rangle d\mu = -\int_M f \text{ div}_\mu(X) d\mu, \quad f \in C^1_0(M). \]

We define the divergence of $\psi_{h_0,X}$ by
\[ \text{div}(\psi_{h_0,X})(x) := \int_0^T \left( \dot{\psi}_{h_0,X} + \frac{1}{2} \text{Ric}_s \psi_{h_0,X}, db_s \right) - \text{div}_\mu(X)(x_0), \quad (2.16) \]
which is in $L^2(P\mu)$ since Ric is bounded and $|h_0|_{H^0} \in L^2(P\mu)$ and $X \in X_0(M)$.

**Theorem 2.5.** For any $h_0 \in \tilde{H}_0$ and $X \in \mathcal{X}_0(M)$, define
\[ D^*_\psi h_0,X F := -D\psi h_0,X F - F \text{ div}(\psi h_0,X), \quad F \in \mathcal{F} C_0^\infty. \quad (2.17) \]
We have
\[ E(G D\psi h_0,X F) = E(F D^*_\psi h_0,X G), \quad F, G \in \mathcal{F} C_0^\infty. \quad (2.18) \]
Consequently, $(D\psi h_0,X, \mathcal{F} C_0^\infty)$ is a closable operator on $L^2(P\mu)$.

**Proof.** It suffices to prove (2.18). Using Proposition 2.1, we have
\[ E(G(x h_0,X,\epsilon)Mh_0,X,\epsilon T) = \int_M FG dQ h_0,X,\epsilon \mu = \int_M E_x(FG) d\mu \epsilon = \int_M E_x(FG) k \epsilon \ d\mu, \]
where $k \epsilon := \frac{d\mu}{d\mu} = E_x(F) := \int_M F d\mu$. Now taking the derivative in the above equality with respect to $\epsilon$ and at $\epsilon = 0$ and according to (2.10) we get
\[ E(D\psi h_0,X (FG) - FG \int_0^T \dot{h}(s), db_s) \]
\[ = -\int_M E_x(FG) \text{ div}_\mu(X) d\mu = -E[FG \text{ div}_\mu(X)]. \]
Thus, using the fact that $D\psi h_0,X (FG) = D\psi h_0,X F G + F D\psi h_0,X G$ we get (2.18).

To prove the closability of the gradient operator from the above integration by parts formula, we need to represent $DF$ using $\psi h_0,X$ for $h_0 \in \tilde{H}_0$ and $X \in \mathcal{X}_0(M)$.

**Lemma 2.6.** For any $F \in \mathcal{F} C_0^\infty$, there exist sequences of $F_i \in L^\infty(P\mu)$, $h_i \in \tilde{H}_0$, and $X_i \in \mathcal{X}_0(M)$ such that
\[ DF = \sum_i F_i \psi_{h_i,X_i} \]
the series being convergent in $L^2(\mu)$.

**Proof.** By Nash’s embedding theorem, we can embed $M$ isometrically into a Euclidean space $\mathbb{R}^p$ by some $p > d$. Let $X(x) : \mathbb{R}^p \to T_x M$ be the orthogonal projection and $Y(x) : T_x M \to \mathbb{R}^p$ the embedding map. We have $X(x)Y(x) = \text{Id}_{T_x M}$. Having these preparations, we obtain

$$Y_{x_0}(r_0r_1^{-1}\nabla_i f) = \sum_{k=1}^p [Y_{x_0}(r_0r_1^{-1}\nabla_i f), e_k]X_k(x_0)$$

for $\{e_1, \ldots, e_p\}$ an orthonormal basis of $\mathbb{R}^p$. Let $X_k(x) := X(x)e_k$. Then $\{X_k\}_{k=1}^p$ are smooth vector fields on $M$ such that

$$Y_{x_0}(r_0r_1^{-1}\nabla_i f)(x) = \sum_{n=1}^\infty \sum_{k=1}^p [Y_{x_0}(r_0r_1^{-1}\nabla_i f)(x), e_k]X_k(x_0).$$

(2.19)

Let $\{\alpha_n; n \geq 1\}$ be a partition of unity on $M$, that is $\alpha_n \in C^\infty_0(M)$, $\alpha_n \geq 0$ and $\sum_n \alpha_n = 1$. It is easy to see that the series

$$Y_{x_0}(r_0r_1^{-1}\nabla_i f)(x) = \sum_{n=1}^\infty \alpha_n(x_0)(r_0r_1^{-1}\nabla_i f)(x)$$

converges in $L^2(\mu)$. According to (2.19), we obtain

$$Y_{x_0}(r_0r_1^{-1}\nabla_i f)(x) = \sum_{n=1}^\infty \sum_{k=1}^p [Y_{x_0}(r_0r_1^{-1}\nabla_i f)(x), e_k]\alpha_n(x_0)X_k(x_0).$$

(2.20)

Let $F_{i\delta} = \langle Y_{x_0}(r_0r_1^{-1}\nabla_i f)(x), e_k \rangle$ which are bounded and $X_{nk} = \alpha_nX_k \in \mathcal{F}_0^\infty(M)$ and $h_{i\delta}(t) = (t \land t_i)r_0^{-1}X_{nk}$, we obtain the result from expressions (2.5) and (2.7).

---

**Theorem 2.7.** The gradient operator $(D, \mathcal{F}^\infty_0) : L^2(\mu) \to L^2(\mathcal{P}(M) \to \mathcal{H}; \mu)$ is closable. Consequently, the quadratic from $(\mathcal{E}, \mathcal{F}^\infty_0)$ is closable in $L^2(\mu)$ and the closure $(\mathcal{E}, \mathcal{F}^\infty_0)$ is a conservative Dirichlet form.

**Proof.** It suffices to prove the first assertion. Let $F_n \in \mathcal{F}^\infty_0$ be a sequence of cylinder functions such that $\lim_{n \to +\infty} F_n = 0$ in $L^2(\mathcal{P}(M))$ and $Z := \lim_{n \to +\infty} DF_n$ in $L^2(\mathcal{P}(M) \to \mathcal{H}; \mu)$. We intend to prove that $Z = 0$. By Lemma 2.6, we only need to show that

$$\mathbb{E}(Z, F_{\psi_{h_0}, X}) = 0$$

(2.21)

for any $F \in L^\infty(\mu)$, $h_0 \in \mathcal{H}_0$ and $X \in \mathcal{F}_0^\infty(M)$. Since $\mathcal{F}^\infty_0$ is dense in $L^2(\mu)$, we may assume that $F \in \mathcal{F}^\infty_0$. Then by Theorem 2.5 we have

$$\mathbb{E}(Z, F_{\psi_{h_0}, X}) = \lim_{n \to +\infty} \mathbb{E}(DF_n, F_{\psi_{h_0}, X}) = \lim_{n \to +\infty} \mathbb{E}[F_n D^*_{\psi_{h_0}, X} F] = 0$$

since $F_n \to 0$ in $L^2(\mu)$. Thus, (2.21) holds. □
3. Logarithmic Sobolev inequalities on $P(M)$

Logarithmic Sobolev inequalities on $P_x(M)$ were established independently by E.P. Hsu [10] and by S. Aida and D. Elworthy [1].

**Theorem 3.1.** Let $K := \sup_{r \in O(M)} \| \text{Ric}_r \|$. Then

$$E_x(F^2 \log F^2) \leq 2e^{KT} E_x(\|D^0 F\|^2_{H_0}) + E_x(F^2) \log E_x(F^2), \quad F \in \mathcal{F}C^1_b. \quad (3.1)$$

For proof of this result, we refer to [11] or [4]. The main result of this section is the following.

**Theorem 3.2.** Let $d\mu = v dx$ be a probability measure on $M$ satisfying (2.1). Suppose that

$$\mu(\phi^2 \log \phi^2) \leq C_M \mu(\|\nabla \phi\|^2) + \mu(\phi^2) \log \mu(\phi^2), \quad \phi \in C^1_b(M). \quad (3.2)$$

Then

$$P_\mu(F^2 \log F^2) \leq \alpha \mathcal{E}(F, F) + P_\mu(F^2) \log P_\mu(F^2), \quad F \in \mathcal{F}C^1_b, \quad (3.3)$$

where

$$\alpha := 2e^{KT} + C_M \left(1 + \frac{K}{4}(e^{KT} - 1)\right). \quad (3.4)$$

We remark that condition (3.2) is essential for the log-Sobolev inequality (3.3) to hold. Indeed, taking $F(x) := \phi(x_0)$, (3.3) implies (3.2) for some constant. To prove this theorem, we need the following lemma essentially due to [10].

**Lemma 3.3.** Let $F$ be a cylinder function given in (2.4). Then the gradient of the function $x \mapsto E_x(F)$ on $M$ is given by

$$\nabla_x E_x(F) = \sum_{i=1}^N E_x(r_0 R_0 r_i^{-1} \partial_i f) \quad (3.5)$$

where $R_i$ is the solution of the following resolvent equation

$$\frac{dR_i}{dt} = -\frac{1}{2} R_i \text{Ric}_{r_i}, \quad R_0 = \text{Id}. \quad (3.6)$$

**Proof.** When $F$ is dependent of a single time, the above formula (3.5) can be deduced from Weitzenböck formula (see [2,14]). The general case can be obtained by induction (see [6,11]). □

**Proof of Theorem 3.2.** Let $\psi(x) := (\int_{P(M)} F^2 dP_x)^{1/2}$. By integrating two sides of (3.1), we have

$$\int_{P(M)} F^2 \log F^2 dP_\mu \leq 2e^{KT} \int_{P(M)} \|D^0 F\|^2_{H_0} dP_\mu + \int_M \psi^2 \log \psi^2 d\mu. \quad (3.7)$$
Let $F$ be given by (2.4) with $t_1 = 0$. Applying (3.5),

$$\nabla_x E_x \left( f^2(x, \gamma(t_2), \ldots, \gamma(t_N)) \right) = 2E_x \left( F \cdot \sum_{i=1}^N r_0 R_i r_i^{-1} (\partial_i f) \right).$$

By Cauchy–Schwartz inequality,

$$| \nabla_x E_x \left( f^2(x, \gamma(t_2), \ldots, \gamma(t_N)) \right) |^2 \leq 4E_x (F^2) \cdot E_x \left( \left| \sum_{i=1}^N r_0 R_i r_i^{-1} (\partial_i f) \right|^2 \right). \tag{3.8}$$

Now let $\xi_i \in \mathbb{R}^d$. By (3.6),

$$R_i \xi_i = \xi_i - \frac{1}{2} \int_0^{t_i} R_s \text{Ric}_s \xi_i \, ds. \tag{3.9}$$

Then

$$\left| \sum_{i=1}^N R_i \xi_i \right|^2 \leq (1 + \varepsilon) \left| \sum_{i=1}^N \xi_i \right|^2 + \frac{1 + \varepsilon - 1}{4} \sum_{i=2}^N \int_0^{t_i} R_s \text{Ric}_s \xi_i \, ds, \quad \varepsilon > 0.$$ \tag{3.10}

But the second term on the right hand side is equal to

$$\frac{1 + \varepsilon - 1}{4} \int_0^T \left| \sum_{i=2}^N \text{Ric}_s \left( \sum_{i=1}^N \mathbf{1}_{(s \leq t_i)} \xi_i \right) \right|^2 \, ds,$$

which is dominated by

$$\frac{1 + \varepsilon - 1}{4} \left( \int_0^T \| R_s \text{Ric}_s \|^2 \, ds \right) \left( \int_0^T \left| \sum_{i=2}^N \mathbf{1}_{(s \leq t_i)} \xi_i \right|^2 \, ds \right).$$

Since $\text{Ric} \geq -K$, by resolvent equation (3.6), $\| R_s \| \leq e^{Kt/2}$. Therefore,

$$\int_0^T \| R_s \text{Ric}_s \|^2 \, ds \leq \frac{1}{K} (e^{KT} - 1) \cdot K^2 = K(e^{KT} - 1).$$

Hence

$$\left| \sum_{i=1}^N R_i \xi_i \right|^2 \leq (1 + \varepsilon) \left| \sum_{i=1}^N \xi_i \right|^2 + \frac{1 + \varepsilon - 1}{4} K(e^{KT} - 1) \left( \int_0^T \left| \sum_{i=2}^N \mathbf{1}_{(s \leq t_i)} \xi_i \right|^2 \, ds \right), \quad \varepsilon > 0. \tag{3.10}$$

Replacing $\xi_i = r_i^{-1} (\partial_i f)$ in (3.10) and according to (3.8) and (2.6), we get
\[ |\nabla x E_x \left( f^2(x, \gamma(t_1), \ldots, \gamma(t_N)) \right)|^2 \]
\[ \leq 4 E_x(f^2) \left[ (1 + \varepsilon) E_x \left( \left( \sum_{i=1}^N r_i^{-1}(\partial_i f) \right)^2 \right) \right. \]
\[ + \frac{1}{4} K(e^{KT} - 1) E_x \left( \int_0^T \sum_{i=2}^N \mathbf{1}_{t_i < t} r_i^{-1}(\partial_i f) \, ds \right) \]
\[ \leq 4 \left( (1 + \varepsilon) \vee \left[ \frac{1}{4} K(e^{KT} - 1) \right] \right) E_x(F^2) E_x(|DF|_H). \]

Taking \( \varepsilon = \frac{K}{4} (e^{KT} - 1) \), we arrive at
\[ |\nabla x \phi(x)|^2 \leq \frac{|\nabla x E_x(f^2, \gamma(t_1), \ldots, \gamma(t_N))|^2}{4 E_x(F^2)} \]
\[ \leq \left( 1 + \frac{K}{4} (e^{KT} - 1) \right) E_x(|DF|_H). \]

Thus, the proof is completed by combining (3.2) and (3.7). \( \square \)

4. Transportation cost inequalities on \( P(M) \)

The difficulty to establish transportation cost inequalities in infinite dimensional situation is to handle the intrinsic distance. Once we have established the logarithmic Sobolev inequality (3.3) on the free path space \( P(M) \), we can proceed as in [18,19] to obtain the transportation cost inequality on \( P(M) \).

Define the intrinsic distance on \( P(M) \) by
\[ \rho_H(\gamma_1, \gamma_2) := \sup \{|F(\gamma_1) - F(\gamma_2)| : F \in \mathcal{F}C^1_b, |DF|_H^2 \leq 1 \text{ P}\mu\text{-a.s.} \}. \]

It is clear by expression (4.1) that \( (\gamma_1, \gamma_2) \mapsto \rho_H(\gamma_1, \gamma_2) \) is semi lower-continuous on \( P(M) \times P(M) \). Let \( F \) be a positive measurable function on \( P(M) \) such that \( P\mu(F) = 1 \).

Introduce the quadratic Wasserstein distance \( W_{2,\rho_H} \) by
\[ W_{2,\rho_H}(FP\mu, P\mu) = \inf_{\hat{\pi} \in \mathcal{C}(FP\mu, \mu)} \int_{P(M) \times P(M)} \rho_H^2(\gamma_1, \gamma_2) \hat{\pi}(d\gamma_1, d\gamma_2), \]
where \( \mathcal{C}(FP\mu, \mu) \) denotes the set of couplings for \( FP\mu \) and \( P\mu \), that is, the set of Borel probability measures on \( P(M) \times P(M) \) having \( FP\mu \) and \( P\mu \) as marginal laws. The main result of this section is the following.

**Theorem 4.1.** Let \( \alpha \) be given in Theorem 3.2. Then
\[ W_{2,\rho_H}(FP\mu, P\mu) \leq \alpha P\mu(F \log F), \quad F \geq 0, \quad P\mu(F) = 1. \]
Proof. It suffice to prove for cylinder functions. Let \( I = \{0 \leq t_1 < \cdots < t_N \leq T\} \) be a partition of the interval \([0, T]\). Define the projection \( A_I : \mathcal{P}(M) \to M^I \) by
\[
\gamma \mapsto A_I(\gamma) = (\gamma(t_1), \ldots, \gamma(t_N)).
\]
Let \( \mu_I = (A_I)_* P_\mu \) be the induced measure on \( M^I \). Then \( \mu_I \) has the expression
\[
d\mu_I = p_t(x, x) p_{t_1}^{-1}(x_1, x) \cdots p_{t_{N-1}}^{-1}(x_{N-1}, x_N) \nu(x) \, dx \, dx_1 \cdots dx_N,
\]
where \( p_t(x, y) \) is the heat kernel on \( M \). Let \( F \) be a cylinder function in the form \( F = f(A_I) \). Define
\[
q_{ij} := 1 + t_i \land t_j, \quad i, j \geq 1. \tag{4.4}
\]
According to (2.6), we have the expression
\[
|DF|^2_H = \sum_{i, j=1}^N q_{ij} r_{t_i}^{-1}(\nabla_i f, \nabla_j f), \tag{4.5}
\]
Define for \( z \in M^I \), \( A_{ij}^I(z) = q_{ij} E(r_{t_i} r_{t_j}^{-1} | A_I = z) \). Then we have the equality
\[
E(|DF|^2_H) = \int_{M^I} \sum_{i, j=1}^N [A_{ij}^I(z)(\partial_i f)(z), (\partial_j f)(z)] d\mu_I(z). \tag{4.6}
\]
For a vector field \( X \) on \( M^I \), let \( X_i \) be the \( i \)th-component of \( X \) on \( TM^I \). Define
\[
(A^I X)_j = \sum_{i=1}^N A_{ij}^I X_i.
\]
Since Itô parallel transports are smooth functionals in Malliavin calculus, \( z \mapsto A_{ij}^I(z) \) is a smooth mapping (see [15]). By (4.5), we see that \( A^I : TM^I \to TM^I \) is a positive definite symmetric operator, which defines a Riemannian metric \( d_I \) on \( M^I \). Since \( A^I \) is bounded and the original metric \( \langle \cdot, \cdot \rangle \) is complete, the induced metric \( \langle \cdot, \cdot \rangle_A \) is complete too.

Set \( \nabla_{M^I} f := (\nabla_1 f, \ldots, \nabla_N f) \in TM^I \). Then
\[
\langle A^I \nabla_{M^I} f, \nabla_{M^I} f \rangle(z) = E(|D(f \circ A_I)|^2_H | A_I = z). \tag{4.7}
\]
and
\[
E(|DF|^2_H) = \int_{M^I} \langle A^I \nabla_{M^I} f, \nabla_{M^I} f \rangle d\mu_I =: \mathcal{E}_I(f, f). \tag{4.8}
\]
By Theorem 3.2, the log-Sobolev inequality holds on \( M^I \) for the Dirichlet form \( \mathcal{E}_I(f, f) \) with constant \( \alpha \). Using (1.4), we have
\[
W_{2, d_I}(f \mu_I, \mu_I) \leq \alpha \mu_I(f \log f), \quad f \geq 0, \quad \mu_I(f) = 1. \tag{4.9}
\]
According to (1.2), the distance \( d_I \) has the expression
\[
d_I(z_1, z_2) = \sup\{|f(z_1) - f(z_2)| : f \in C^1_b(M^I), \langle A \nabla_{M^I} f, \nabla_{M^I} f \rangle \leq 1\}.
\]
Define the cylinder distance
\[ \rho_I(\gamma_1, \gamma_2) = \sup \{ |f \circ \Lambda_I(\gamma_1) - f \circ \Lambda_I(\gamma_2)| : f \in C^1_b(M^I), \nabla(f \circ \Lambda_I) \leq 1 \}. \] (4.10)

By (4.7) we have
\[ d_I(\Lambda_I(\gamma_1), \Lambda_I(\gamma_2)) \geq \rho_I(\gamma_1, \gamma_2). \] (4.11)

Suppose that \( F = f \circ \Lambda_I \) for \( f \in C_b(M^I) \), and take a sequence of partition \( I_n \) finer and finer such that \( \bigcup_{n \geq 0} I_n \) is dense in \([0, T]\). By (4.10), it is easy to see that
\[ \rho_{I_n} \leq \rho_{I_{n+1}} \quad \text{and} \quad \rho_H(\gamma_1, \gamma_2) = \sup_{n \geq 0} \rho_{I_n}(\gamma_1, \gamma_2). \] (4.12)

Since \( I_{n+1} \supset I_n \), we have \( \mu_{I_n}(f \log f) = P_{\mu}(F \log F) \) for all \( n \geq 0 \). By (4.9), for any \( m \geq 1 \) there exists a coupling measure \( \tilde{\pi}_m \in \mathcal{C}(\mu_{I_m}, \mu_{I_m}) \) such that
\[ \int_{P^{2, d_{I_n}}(M^I \times P(M))} \rho^2_{I_n}(\gamma_1, \gamma_2) d\tilde{\pi}_m(\gamma_1, \gamma_2) \leq W^2_{2, d_{I_n}}(f \mu_{I_n}, \mu_{I_n}) + \frac{1}{m} \leq \alpha P_{\mu}(F \log F) + \frac{1}{m}. \] (4.13)

Now define
\[ \hat{\pi}_m(d\gamma_1, d\gamma_2) := \tilde{\pi}_m(dz_1, dz_2)(F P_{\mu})(d\gamma_1 | \Lambda_{I_n}(\gamma_1) = z_1) P_{\mu}(d\gamma_2 | \Lambda_{I_n}(\gamma_2) = z_2) \]
where \( P_{\mu}(\cdot | \Lambda_{I_n}(\gamma) = z) \) (resp. \( F P_{\mu}(\cdot | \Lambda_{I_n}(\gamma) = z) \)) is the regular conditional distributions of \( P_{\mu} \) (resp. \( F P_{\mu} \)) given \( \Lambda_{I_n}(\gamma) = z \). Since the regular conditional distributions \( P_{\mu}(\cdot | \Lambda_{I_n}(\gamma) = z) \) and \( F P_{\mu}(\cdot | \Lambda_{I_n}(\gamma) = z) \) are transition probability measures, \( \hat{\pi}_m \) is a well-defined probability measure on \( P(M) \times P(M) \). Moreover, since \( \mu_{I_n} \) and \( f \mu_{I_n} \) are marginal distributions of \( P_{\mu} \) and \( F P_{\mu} \) on \( M^{I_n} \) respectively, by the definition of the regular conditional distribution we have
\[ P_{\mu}(d\gamma) = \mu_{I_n}(dz) P_{\mu}(d\gamma | \Lambda_{I_n}(\gamma) = z), \]
\[ (F P_{\mu})(d\gamma) = (f \mu_{I_n})(dz) (F P_{\mu})(d\gamma | \Lambda_{I_n}(\gamma) = z). \]

Therefore, \( \hat{\pi}_m \in \mathcal{C}(F P_{\mu}, P_{\mu}) \).

It is now standard to observe the tightness of \( \mathcal{C}(F P_{\mu}, P_{\mu}) \). Indeed, for any \( \varepsilon > 0 \) let \( K \subset P(M) \) be compact such that \( P_{\mu}(K^\varepsilon) + (F P_{\mu})(K^\varepsilon) < \varepsilon \), then for any \( \pi \in \mathcal{C}(F P_{\mu}, P_{\mu}) \) one has
\[ \hat{\pi}(K \times K^\varepsilon) \leq \hat{\pi}(K^\varepsilon \times P(M)) + \hat{\pi}(P(M) \times K^\varepsilon) = P_{\mu}(K^\varepsilon) + (F P_{\mu})(K^\varepsilon) < \varepsilon. \]

Thus, up to subsequence, \( \hat{\pi}_m \) converges to some \( \hat{\pi} \) weakly. It is easy to see that \( \hat{\pi} \in \mathcal{C}(F P_{\mu}, P_{\mu}) \). Now for \( m \geq n \), we have, according to (4.11) and (4.13),
\[ \int_{P(M) \times P(M)} \rho^2_{I_n}(\gamma_1, \gamma_2) \hat{\pi}_m(d\gamma_1, d\gamma_2) \]
\[
\leq \int_{\mathcal{P}(M) \times \mathcal{P}(M)} \rho_{\mu}^2(\gamma_1, \gamma_2) \tilde{\pi}_m(d\gamma_1, d\gamma_2)
\leq \int_{\mathcal{P}(M) \times \mathcal{P}(M)} d_{\mu}^2(\Lambda_I(\gamma_1), \Lambda_I(\gamma_2)) \tilde{\pi}_m(d\gamma_1, d\gamma_2)
= \int_{\mathcal{M}^{lm} \times \mathcal{M}^{lm}} d_{\mu}^2(z_1, z_2) \tilde{\pi}_m(dz_1, dz_2)
\leq \alpha P_{\mu}(F \log F) + \frac{1}{m}.
\]

Letting \( m \to +\infty \) in the above inequality, we get

\[
\int_{\mathcal{P}(M) \times \mathcal{P}(M)} \rho_{\mu}^2(\gamma_1, \gamma_2) \tilde{\pi}(d\gamma_1, d\gamma_2) \leq \alpha P_{\mu}(F \log F).
\]

According to (4.12), we complete the proof of (4.3) by letting \( n \to +\infty \). \( \square \)

Acknowledgements

The first named author would like to thank the Stochastic Center of Beijing Normal University for the support and the hospitality during his stay where this work was being prepared.

References