In this article, exponential contraction in Wasserstein distance for heat semigroups of diffusion processes on Riemannian manifolds is established under curvature conditions where Ricci curvature is not necessarily required to be non-negative. Compared to the results of Wang (2016), we focus on explicit estimates for the exponential contraction rate. Moreover, we show that our results extend to manifolds evolving under a geometric flow. As application, for the time-inhomogeneous semigroups, we obtain a gradient estimate with an exponential contraction rate under weak curvature conditions, as well as uniqueness of the corresponding evolution system of measures.

AMS 2010 Subject Classification: 60J60, 58J65, 53C44.

Key words: Wasserstein distance, diffusion process, coupling, Ricci curvature, Ricci flow, exponential contraction.

1. INTRODUCTION

Let $M$ be a $d$-dimensional connected complete Riemannian manifold and consider the operator $L = \Delta + Z$ where $\Delta$ is the Laplace-Beltrami operator and $Z$ a $C^1$-vector field on $M$. We denote by $X_t$ the diffusion process with generator $L$, which is characterized by the property that for any test function $f$ on $M$, the relation

$$df(X_t) - Lf(X_t)\,dt = 0,$$

modulo differentials of martingals,

holds in the Itô sense. Throughout the paper we assume that the $L$-diffusion process is non-explosive. This holds true, in particular, when the Bakry-Émery Ricci curvature of $M$ is bounded from below, that is, for some real constant $K$,

$$(1.1) \quad \text{Ric}^Z(X,X) := \text{Ric}(X,X) - \langle \nabla_X Z, X \rangle \geq K|X|^2, \quad X \in T_x M, \ x \in M.$$

Let $P_t$ be the Markov transition semigroup associated to $X_t$ and $\mu P_t$ be the law of $X_t$ with initial distribution $\mu$. It is well known that there are various functional inequalities on $P_t$ which all give conditions equivalent to the curvature condition (1.1), see [4, 15].
In this article, we investigate $L^q$-Wasserstein contraction inequalities ($q \geq 1$) for $\mu P_t$. Denote by $\mathcal{P}(M)$ the set of probability measures on $M$. On $\mathcal{P}(M)$ the $L^q$-Wasserstein distance is defined as

$$W_q(\mu_1, \mu_2) := \inf_{\pi \in \mathcal{C}(\mu_1, \mu_2)} \left( \int \int_{M \times M} \rho(x,y)^q \, d\pi(x,y) \right)^{1/q}, \quad \mu_1, \mu_2 \in \mathcal{P}(M),$$

where $\rho$ denotes the Riemannian distance on $M$ and $\mathcal{C}(\mu_1, \mu_2)$ consists of all couplings of $\mu_1$ and $\mu_2$. The Wasserstein distance has various characterizations and plays an important role in the study of SDEs, partial differential equations, optimal transportation problem, etc. For more background, one may consult [13, 10, 15] and the references therein.

The $L^q$-Wasserstein distance $W_q$ on $\mathcal{P}(M)$ will be used to quantify the time evolution of $(\mu P_t)_{t \geq 0, \mu \in \mathcal{P}(M)}$. A typical phenomenon of interest for the system $(\mu P_t)_{t \geq 0, \mu \in \mathcal{P}(M)}$ is exponential contraction in the Wasserstein distance, i.e.

$$W_q(\mu_1 P_t, \mu_2 P_t) \leq c e^{-\kappa t} W_q(\mu_1, \mu_2), \quad t \geq 0, \quad q \geq 1,$$

with positive constants $c$ and $\kappa$. We refer the reader to [8, 9, 12, 17] for work in this direction on the Euclidean space $M = \mathbb{R}^d$. When $M$ is a Riemannian manifold, for instance, under the curvature condition

$$\text{Ric}^Z(X,X) \geq \kappa |X|^2$$

with $\kappa \geq 0$, the exponential contraction (1.2) holds with $c = 1$ and $\kappa$ the curvature bound in (1.3). Moreover, it is well-known that inequality (1.2) with $c = 1$ is actually equivalent to the lower curvature bound (1.3). For certain cases, when $\text{Ric}^Z$ is not bounded from below by zero, Wang [16] showed that the following inequality holds: for any $q \geq 1$,

$$W_q(\delta_x P_t, \delta_y P_t) \leq c e^{-\lambda t} \left( \rho(x,y) \vee \rho(x,x)^{1/q} \right)$$

for some constant $c > 1$ and $\lambda > 0$.

In order to weaken condition (1.3) as in [16], let us first recall the definition of the index:

$$I^Z(x,y) = I(x,y) + \langle Z, \nabla \rho(\cdot,y) \rangle + \langle Z, \nabla \rho(\cdot,x) \rangle, \quad x, y \in M,$$

where

$$I(x,y) = \int_0^{\rho(x,y)} \sum_{i=1}^{d-1} \left\{ |\nabla \gamma_i|^2 - \langle R(\gamma, J_i) \gamma, J_i \rangle \right\} (\gamma_s) \, ds.$$

Here $\rho$ is the distance function, $R$ the Riemann curvature tensor, $\gamma$: $[0, \rho(x,y)] \rightarrow M$ the minimal geodesic from $x$ to $y$ with unit speed, $(J_i)_{i=1,\ldots,d-1}$ are Jacobi fields along $\gamma$ such that

$$J_i(y) = P_{x,y} J_i(x), \quad i = 1, \ldots, d-1,$$
for the parallel transport $P_{x,y}: T_xM \to T_yM$ along the geodesic $\gamma$, and
\[
\{ \dot{\gamma}(s), J_i(s) : 1 \leq i \leq d-1 \}, \quad s = 0, \ \rho(x,y),
\]
is an orthonormal basis of the tangent space at $x$, respectively $y$. Note that when $(x,y) \in \text{Cut}(M)$, that is if $x$ is in the cut-locus of $y$, the minimal geodesic may be not unique. As it is a common convention in the literature, all conditions on the index $I^Z$ are supposed to hold outside of $\text{Cut}(M)$. If there exist positive constants $K_1$ and $K_2$ such that
\[
I^Z(x,y) \leq \left( (K_1 + K_2) \mathbb{1}_{\{\rho(x,y) \leq r_0\}} - K_2 \right) \rho(x,y)
\]
for some $r_0 > 0$ and if $\text{Ric}^Z$ is bounded below, then (1.2) holds with $\kappa > 0$ and $c > 1$ for any $q \geq 1$, see [16]. This is the case, for instance, when $\text{Ric}^Z$ is bounded below by a positive constant outside a compact set. It is crucial that the exponential rate $\lambda$ is independent of $p$. Due to the equivalence of (1.2) with $c = 1$ and (1.3), in the negative curvature case it is essential that $c > 1$.

In this paper, we give quantitative estimates of $\kappa$ and $c$ by constructing a suitable auxiliary function. We begin the discussion with a more general condition (see Assumption (A1) below) which includes situation (1.5). Actually, we rewrite condition (1.5) as follows:
\[
I^Z(x,y) \leq k_1 - k_2 \rho(x,y),
\]
for some constants $k_1 \geq 0$ and $k_2 > 0$. Then, for $p > 1$, $t \geq 0$, and $x, y \in M$, we obtain (see Corollary 2.5 below) that
\[
W_p(\delta_x P_t, \delta_y P_t) \leq \left( 1 + \frac{2k_1}{k_2} \right)^{(p-1)/p} \exp \left( \frac{k_1^2}{pk_2} - \frac{k_2}{2p e^{k_1^2/k_2} t} \right) \left( \rho(x,y) \lor \rho(x,y)^{1/p} \right).
\]
Note that the constant $k_2/(2 e^{k_1^2/k_2})$ is independent of $p$.

Our approach to exponential contraction results relies in a crucial way on coupling arguments for Brownian motions, resp. $L$-diffusion processes. To derive exponential contraction of the Wasserstein distance we construct a coupling of $L$-diffusions where we use coupling by reflection for short distance and coupling by parallel displacement for long distance. Intuitively speaking, coupling by reflection is very powerful; in particular when curvature is negative it prevents the coupled processes $X_t$ and $Y_t$ from moving too far away from each other. On the other hand, parallel coupling has the advantage that it leads to simpler calculations, since the martingale part of the distance process $\rho(X_t, Y_t)$ vanishes. This coupling works well under stronger lower curvature bounds, for instance, if there is a positive lower bound of the Ricci curvature. We will see in Section 2 that under our assumptions a mixture of the two couplings is favorable and that from the constructed distance process sharp transportation-cost inequalities can be derived.
In the second part of the paper, we extend the results from Riemannian manifolds to the differentiable manifolds carrying a geometric flow of complete Riemannian metrics. More precisely, for some $T_c \in (-\infty, \infty]$, we consider the situation of a $d$-dimensional differentiable manifold $M$ equipped with a $C^1$ family of complete Riemannian metrics $(g_t)_{t \in (-\infty, T_c)}$. Let $L_t = \Delta_t + Z_t$, where $\Delta_t$ is the Laplace-Beltrami operator associated with the metric $g_t$ and $(Z_t)_{t \in [0,T_c]}$ is a $C^1$-family of vector fields on $M$. Assume that the diffusion process $(X_t)$ generated by $L_t$ is non-explosive before time $T_c$ (see [1] for detailed construction). Let $P_{s,t}$ be the corresponding time-inhomogeneous semigroup.

It is shown in [2, Theorem 4.1 (b)] that if

$$\left( \text{Ric}_t^Z - \frac{1}{2} \partial_t g_t \right)(X,X)(x) \geq \kappa |X|^2_t(x)$$

for some positive constant $\kappa$, where $\text{Ric}_t^Z$ is defined as in (1.1) for the manifold $(M, g_t)$, then exponential contraction in $L^p$-Wasserstein distance holds with respect to the $g_t$-Riemannian distance $\rho_t$, see also [5].

In this paper, we consider situations where $\text{Ric}_t^Z - \frac{1}{2} \partial_t g_t$ is not necessarily bounded below by zero. More precisely, assuming that there exists a real-valued function $k$ such that $\liminf_{r \to \infty} k(r) > 0$ and

$$\left( \text{Ric}_t^Z - \frac{1}{2} \partial_t g_t \right)(X,X)(x) \geq k(\rho_t(x)) |X|^2_t(x),$$

we prove that

$$(1.6) \quad \tilde{W}_{p,s}(\mu_1 P_{s,t}, \mu_2 P_{s,t}) \leq c e^{-\frac{1}{p} \lambda (t-s)} \tilde{W}_{p,t}(\mu_1, \mu_2), \quad t \geq s, \ p \geq 1,$$

holds for some positive constants $c$ and $\lambda$, where

$$\tilde{W}_{p,t}(\mu_1, \mu_2) = \inf_{\pi \in \mathcal{P} \mu_1, \mu_2} \left( \int_{M \times M} \rho_t(x,y)^p \vee \rho_t(x,y) \pi(dx,dy) \right)^{1/p}.$$ 

Moreover, in Theorem 3.1 we give estimates for the constants $c$ and $\lambda$ and apply these results to estimates of the semigroup.

Furthermore, we use the $W_{1,t}$-contraction property to prove uniqueness of the evolution system of measures. It is well known that invariant measure provide important tools in the study of the long behavior of diffusion processes. When it comes to time-inhomogeneous diffusions, the evolution system of measures plays a role similar to the invariant measure. In [6], the first two authors investigated existence and uniqueness of evolution systems of measures. In particular, they found that $W_{1}$-contraction of the distance helps to prove uniqueness properties (see [6] for details). Since now the $W_{1}$-contraction is established even in cases when the lower bound of the curvature may be negative, this allows to improve the result in [6] where a uniform lower curvature bound had been imposed for each time. Inspired by this, in Section
4, we consider uniqueness of the evolution system of measures under a new relaxed curvature condition which allows a lower bound of curvature depending on the radial distance (see Theorem 3.5). It is surprising that under this new condition, a type of dimension-free Harnack inequality can be derived which then may be used to obtain supercontractivity of the semigroup $P_{s,t}$ (see Theorem 3.7).

The paper is structured as follows. In Section 2, we investigate (1.4) by constructing a suitable coupling $(X,Y)$ and using a new auxiliary function to measure the distance of $X$ and $Y$. Our result in this section can be applied to the time-inhomogeneous diffusion process on manifolds carrying geometric flows in Section 3. Section 4 is devoted to the study of existence of evolution system of measures under the new kind of curvature condition. Finally, supercontractivity of the semigroup $P_{s,t}$ with respect to the evolution system of measure is studied by establishing dimension-free Harnack inequalities.

2. EXPONENTIAL CONTRACTION IN WASSERSTEIN DISTANCE

We begin this section by specifying our assumptions.

ASSUMPTION (A1). There exist a non-negative continuous function $k_1$ on $(0,\infty)$, a positive constant $k_2$ and a constant $\theta \geq 0$ such that

\[
I^Z(x,y) \leq k_1(\rho(x,y)) - k_2\rho(x,y)^{1+\theta}
\]

and such that for some positive constants $r_0$ and $k_3$ (with $k_3 < k_2$) the following two conditions hold:

1. $k_1(r) - k_2r^{1+\theta} \leq -k_3r^{1+\theta}$, for $r \geq r_0$,
2. $\int_0^r k_1(v) \, dv < \infty$, for each $r > 0$.

Remark 2.1. We denote by $\rho(x)$ the distance of $x \in M$ to an arbitrary base point. Note that if $\text{Ric}^Z(x) \geq k(\rho(x))$ for all $x$ with $\liminf_{r \to \infty} k(r) > 0$, then there exist constants $k_1$ and $k_2$ such that

\[
I^Z(x,y) \leq k_1 - k_2\rho(x,y).
\]

In this case, Assumption (A1) is satisfied with $k_1$ a non-negative constant and $\theta = 0$.

We now state some exponential contraction inequalities for the Wasserstein distance with explicit estimates of the decay rate.

THEOREM 2.2. Suppose that Assumption (A1) holds. Then,
(i) for $p > 1$, $t \geq 0$, and $x, y \in M$, we have
\[
W_p(\delta_x P_t, \delta_y P_t) \leq c_p e^{-\lambda t/p} (\rho(x, y) \vee \rho(x, y)^{1/p}),
\]
where
\[
c_p = (1 + r_0)^{(p-1)/p} \exp \left( \frac{1}{4p} \int_0^{r_0} k_1(r) \, dr + \frac{k_2}{8p r_0^{2+\theta}} \right)
\]
and
\[
\lambda = k_3 r_0^\theta \exp \left( - \frac{1}{4} \int_0^{r_0} k_1(r) \, dr - \frac{k_2}{8 r_0^{2+\theta}} \right);
\]
(ii) for $t \geq 0$, $\mu_1, \mu_2 \in \mathcal{P}(M)$ and $p > 1$, we have
\[
\tilde{W}_p(\mu_1 P_t, \mu_2 P_t) \leq c_p e^{-\lambda t/p} \tilde{W}_p(\mu_1, \mu_2),
\]
where
\[
\tilde{W}_p(\mu_1, \mu_2) = \inf_{\pi \in \mathcal{C}(\mu_1, \mu_2)} \left( \int_{M \times M} \rho(x, y)^p \vee \rho(x, y) \pi(dx, dy) \right)^{1/p};
\]
(iii) for $t \geq 0$ and $\mu_1, \mu_2 \in \mathcal{P}(M)$, we have
\[
W_1(\mu_1 P_t, \mu_2 P_t) \leq c_1 e^{-\lambda t} W_1(\mu_1, \mu_2).
\]

Remark 2.3. Since $r_0$ and $k_3$ are independent of $p$, the constant $\lambda$ in Theorem 2.2 also does not depend on $p$. Moreover, although $c_p$ depends on $p$, it can be controlled by a constant independent of $p$:
\[
c_p = (1 + r_0)^{(p-1)/p} \exp \left( \frac{1}{4p} \int_0^{r_0} k_1(r) \, dr + \frac{k_2}{8p r_0^{2+\theta}} \right)
\leq (1 + r_0) \exp \left( \frac{1}{4} \int_0^{r_0} k_1(r) \, dr + \frac{k_2}{8 r_0^{2+\theta}} \right).
\]

For the proof of Theorem 2.2, the function $\psi$ defined below and its properties will be crucial. First let $\sigma \in C^1([0, \infty))$ be a function satisfying $0 < \sigma \leq 1$ for $r \in (r_0, r_0 + 1)$, $\sigma \equiv 1$ for $r \leq r_0$ and $\sigma \equiv 0$ for $r \geq r_0 + 1$. Furthermore, define
\[
\ell_0(r) = 4p^2 r^{2(p-1)/p} \sigma(r^{1/p})^2,
\]
\[
\ell_1(r) = p r^{1-1/p} k_1(r^{1/p}) - pk_2 r^{1+\theta/p} + 4p(p - 1) r^{1-2/p} \sigma(r^{1/p})^2,
\]
and let
\[
\ell(r) = pk_2 r_0^\theta r 1_{[0, r_0]} + \left( \frac{p - 1}{p} \frac{\ell_0(r)}{r} - \ell_1(r) \right) 1_{(r_0, \infty)}.
\]
Since $\theta \geq 0$, it is obvious that for $r \in (0, r_0)$,
\[
k_1(r) - k_2 r^{\theta+1} > -k_2 r_0^\theta r.
\]
We thus have $\ell_1 + \ell > 0$, according to the definitions of $\ell_1$ and $\ell$. Next, consider the function

$$
\psi(r) = \int_0^r \exp \left( -\int_{r_0}^u \frac{\ell_1(v) + \ell(v)}{\ell_0(v)} dv \right) du.
$$

The following lemma collects properties of $\psi$.

**Lemma 2.4.** Let $k_1, k_2, k_3, \theta$ and $r_0$ be given by Assumption (A1). The function $\psi$ in (2.4) is well defined, twice differentiable on $(0, \infty)$, and satisfies $\psi' > 0$ and $\psi'' < 0$. In addition,

(i) for $r > 0$, we have

$$\ell_1(r) \psi'(r) + \ell_0(r) \psi''(r) = -\ell(r) \psi'(r);$$

(ii) there exist positive constants $\tilde{c}_1$ and $\tilde{c}_2$ such that

$$\tilde{c}_1 r^{1/p} \leq \psi(r) \leq \tilde{c}_2 r^{1/p}$$

where

$$\tilde{c}_1 = p r_0^{p-1} \quad \text{and} \quad \tilde{c}_2 = p r_0^{p-1} \exp \left( \frac{1}{4} \int_0^{r_0} k_1(r) dr + \frac{k_2}{8} r_0^{2+\theta} \right);$$

(iii) for any $r > 0$,

$$\ell(r) \psi'(r) \geq \lambda \psi(r),$$

where

$$\lambda = k_3 r_0^\theta \exp \left( -\frac{1}{4} \int_0^{r_0} k_1(r) dr - \frac{k_2}{8} r_0^{2+\theta} \right).$$

**Proof.** The first assertion is immediate from the definition of $\psi$. For $0 < r < r_0^p$, we have $\sigma(r^{1/p}) = 1$ and then

$$
\int_{r_0^p}^r \frac{\ell(v) + \ell_1(v)}{\ell_0(v)} dv = \int_{r_0^p}^r \left( \frac{k_1(v^{1/p})}{4pv^{1-1/p}} + \frac{p-1}{p} v^{-1} - \frac{k_2 v^{-1+2+\theta}}{4p} + \frac{k_2 r_0^\theta v^{-1+2+\theta}}{4p} \right) dv.
$$

As $k_1, k_2$ satisfy (2.3), we find

$$
\int_{r_0^p}^r \frac{\ell(v) + \ell_1(v)}{\ell_0(v)} dv \leq \ln r^{(p-1)/p} - \ln r_0^{p-1}
$$

and

$$
\int_{r_0^p}^r \frac{\ell(v) + \ell_1(v)}{\ell_0(v)} dv \geq \ln r^{(p-1)/p} - \ln r_0^{p-1} - \int_0^{r_0^p} \frac{k_1(v^{1/p})}{4pv^{1-1/p}} dv - \int_0^{r_0^p} \frac{k v^{-1+2+\theta}}{4p} dv.
$$
Combining (2.5) and (2.6), we conclude that

\[ r_0^{p-1} r^{(1-p)/p} \leq \psi'(r) \leq \exp \left( \frac{1}{4} \int_0^{r_0} k_1(r) dr + \frac{k}{8 r_0^2} \right) r_0^{p-1} r^{(1-p)/p}. \]

This implies

\[ (2.7) \quad \tilde{c}_1 r^{1/p} \leq \psi(r) \leq \tilde{c}_2 r^{1/p}, \quad 0 < r < r_0^p, \]

where

\[ \tilde{c}_1 := pr_0^{p-1} \quad \text{and} \quad \tilde{c}_2 := p \exp \left( \frac{1}{4} \int_0^{r_0} k_1(r) dr + \frac{k}{8 r_0^2} \right) r_0^{p-1}. \]

On the other hand, for \( r \geq r_0^p \), we have

\[ \int_{r_0^p}^{r} \frac{\ell(u)}{\ell_0(u)} du = \frac{p-1}{p} \int_{r_0^p}^{r} \frac{1}{u} du = \frac{p-1}{p} (\ln r - p \ln r_0) \]

which gives

\[
\psi(r) = \psi(r_0^p) + \int_{r_0^p}^{r} \exp \left( - \int_{r_0^p}^{u} \frac{\ell_1(v)}{\ell_0(v)} dv \right) du \\
= \psi(r_0^p) + p \left( r^{1/p} r_0^{p-1} - r_0^p \right).
\]

Moreover,

\[ \psi'(r) = r_0^{p-1} r^{(1-p)/p}, \quad r \geq r_0^p. \]

In particular, \( \psi \) is well defined. Combining this with (2.7), we obtain, for all \( r > 0 \),

\[ \tilde{c}_1 r^{1/p} \leq \psi(r) \leq \tilde{c}_2 r^{1/p} \]

where

\[ \tilde{c}_1 = pr_0^{p-1}, \]
\[ \tilde{c}_2 = p \exp \left( \frac{1}{4} \int_0^{r_0} k_1(r) dr + \frac{k}{8 r_0^2} \right) r_0^{p-1}. \]

Using the condition \( k_1(r) - k_2 r^{1+\theta} \leq -k_3 r^{1+\theta} \) on \([r_0^p, \infty)\) and the above estimates for \( \psi \) and \( \psi' \), we arrive at

\[
\ell(r) \psi'(r) \geq \left( k_2 r_0^\theta p r \mathbb{1}_{[0, r_0^p)} + k_3 pr^{1+\theta/p} \mathbb{1}_{[r_0^p, \infty)} \right) \psi'(r) \\
\geq pr_0^{p-1} \left( k_2 r_0^\theta \mathbb{1}_{[0, r_0^p)} + k_3 r_0^\theta \mathbb{1}_{[r_0^p, \infty)} \right) r_0^{1/p} \\
\geq \min\{k_2, k_3\} r_0^\theta \exp \left( - \frac{1}{4} \int_0^{r_0} k_1(r) dr - \frac{k}{8 r_0^2} \right) \psi(r) \\
= \lambda \psi(r).
\]
Proof of Theorem 2.2. Consider the operator $L = \Delta + Z$ where $Z$ is a vector field on $M$. Let $d_t$ denote the Itô differential on $M$. Then the $L$-diffusion process $X_t$ is obtained as solution to the Itô-SDE

$$(2.8) \quad d_tX_t = \sqrt{2} u_t dB_t + Z(X_t) \, dt, \quad X_0 = x,$$

where $(B_t)_{t \geq 0}$ is a $d$-dimensional standard Brownian motion on $\mathbb{R}^d$ and $(u_t)_{t \geq 0}$ a horizontal lift of $(X_t)_{t \geq 0}$ to the orthonormal frame bundle over $M$. As explained in the Introduction, our approach is to construct the coupling for short distance by reflection and for long distance by parallel displacement. To this end, we choose a cut-off function $\sigma \in C^1([0, \infty))$ as before, that is a function $\sigma \in C^1([0, \infty))$ satisfying $0 < \sigma \leq 1$ when $r \in (r_0, r_0 + 1)$, and $\sigma \equiv 0$ when $r \geq r_0 + 1$ and $\sigma \equiv 1$ when $r \leq r_0$. For $(x, y) \notin \text{Cut}(M)$, let

$$M_{x,y} : T_x M \to T_y M, \quad v \mapsto P_{x,y} v - 2 \langle \gamma (x) \gamma (y) \rangle$$

be the mirror reflection, where $\gamma$ is the minimal geodesic from $x$ to $y$. We rewrite SDE (2.8) as

$$d_tX_t = \sqrt{2} \left( \sigma(\rho(X_t, Y_t)) u_t dB'_t + \sqrt{1 - \sigma(\rho(X_t, Y_t))^2} u_t dB''_t \right) + Z(X_t) \, dt,$$

where $B'_t$ and $B''_t$ are two independent Brownian motions. Now define $Y_t$ as solution to the following SDE on $M$ with initial condition $Y_0 = y$:

$$(2.9) \quad d_tY_t = \sqrt{2} \left( \sigma(\rho(X_t, Y_t)) M_{X_t, Y_t} u_t dB'_t + \sqrt{1 - \sigma(\rho(X_t, Y_t))^2} P_{X_t, Y_t} u_t dB''_t \right) + Z(Y_t) \, dt.$$

Since the coefficients of the SDE are at least $C^1$ outside the diagonal $\{(z, z) : z \in M\}$, there is a unique solution up to the coupling time

$$T := \inf \{ t \geq 0 : X_t = Y_t \}.$$

As usual, we let $X_t = Y_t$ for $t \geq T$. We ignore here technical difficulties related to a possibly non-empty cut-locus $\text{Cut}(M)$. It is well known how to deal with these issues, see for instance [14, Chapt. 2] or [3, Sect. 3] for details. The presence of a cut-locus actually facilitates the coupling; it decreases the distance of the two marginal processes.

Next, we have by Itô’s formula,

$$d \rho(X_t, Y_t) \leq 2 \sqrt{2} \sigma(\rho(X_t, Y_t)) \, db_t + I^2(X_t, Y_t) \, dt$$

$$\leq 2 \sqrt{2} \sigma(\rho(X_t, Y_t)) \, db_t + \left( k_1(\rho(X_t, Y_t)) - k_2 \rho(X_t, Y_t)^{1+\theta} \right) \, dt, \quad t \leq T,$$

where $b_t$ is a one-dimensional Brownian motion on $\mathbb{R}$. Thus,

$$d \rho(X_t, Y_t)^p \leq p \rho(X_t, Y_t)^{p-1} \, d \rho(X_t, Y_t) + \frac{1}{2} p(p-1) \rho(X_t, Y_t)^{p-2} \, d\langle \rho \rangle_t,$$
\[ \leq p \rho(X_t, Y_t)^{p-1} \left\{ 2\sqrt{2} \sigma(\rho(X_t, Y_t)) \, dt + \left(k_1(\rho(X_t, Y_t)) - k_2(\rho(X_t, Y_t)^{1+\theta}) \right) \, dt \right\} + 4p(p-1)\sigma(\rho(X_t, Y_t))^2 \rho(X_t, Y_t)^{p-2} \, dt, \quad t \leq T, \]

where \( \langle \rho \rangle_t \) denote the quadratic variation of \( \rho(X_t, Y_t) \).

Taking this calculation into account, our next step is to look at properties of the process \( \psi(\rho(X_t, Y_t)^p) \). First of all, by Itô’s formula, we have

\[
d\psi(\rho(X_t, Y_t)^p) \\
\leq \psi'(\rho(X_t, Y_t)^p) \left( 2\sqrt{2}p \rho(X_t, Y_t)^{p-1} \sigma(\rho(X_t, Y_t)) \, dt + \ell_1(\rho(X_t, Y_t)^p) \, dt \right) + \psi''(\rho(X_t, Y_t)^p) \ell_0(\rho(X_t, Y_t)^p) \, dt \\
= dM_t - \ell(\rho(X_t, Y_t)^p) \psi'(\rho(X_t, Y_t)^p) \, dt, \quad t \leq T,
\]

where

\[
dM_t = 2\sqrt{2}p \psi'(\rho(X_t, Y_t)^p) \rho(X_t, Y_t)^{p-1} \sigma(\rho(X_t, Y_t)) \, db_t.
\]

By means of Lemma 2.4 (iii), we get

\[
d\psi(\rho(X_t, Y_t)^p) \leq dM_t - \lambda \psi(\rho(X_t, Y_t)^p) \, dt, \quad t \leq T.
\]

Let \( \tau_n = \{ t \geq 0 : \rho(X_t, Y_t) \notin [1/n, n] \} \). Then \( \tau_n \uparrow T \) as \( n \to \infty \), and for \( s \leq t, \)

\[
\mathbb{E}\psi(\rho^p(X_{t\wedge \tau_n}, Y_{t\wedge \tau_n})) \\
\leq \mathbb{E}\psi(\rho^p(X_{s\wedge \tau_n}, Y_{s\wedge \tau_n})) - \lambda \int_s^t \mathbb{E}\psi(\rho^p(X_{r\wedge \tau_n}, Y_{r\wedge \tau_n})) \, dr.
\]

(2.10)

From now on, for the sake of brevity, we simply write \( \rho_t^p := \rho(X_t, Y_t)^p \). Since \( \psi(0) = 0 \) and \( X_t = Y_t \) for \( t \geq T \), we have

\[
\mathbb{E}\psi(\rho_t^p) = \mathbb{E}\left[ \psi(\rho_t^p) \mathbb{1}_{\{t < T\}} \right] + \mathbb{E}\left[ \psi(\rho_T^p) \mathbb{1}_{\{t \geq T\}} \right] = \mathbb{E}\psi(\rho_T^p).
\]

(2.11)

Letting \( n \to \infty \) of (2.10) and using (2.11), we conclude that

\[
\mathbb{E}\psi(\rho_t^p) \leq \mathbb{E}\psi(\rho_S^p) - \lambda \int_s^t \mathbb{E}\psi(\rho_r^p) \, dr.
\]

Thus, letting

\[
f(t) = \mathbb{E}\psi(\rho_t^p),
\]

we obtain

\[
f(t) \leq f(s) - \lambda \int_s^t f(r) \, dr.
\]

For the function

\[
U(t) = e^{-\lambda t} \psi(\rho(x, y)^p),
\]
it is immediate that
\[ U(t) = U(s) - \lambda \int_s^t U(r) \, dr, \quad U(0) = \psi(\rho(x,y)^p). \]
This implies
\[ f(t) \leq U(t), \quad t \geq 0. \]
Actually assume that there exists \( t_0 > 0 \) such that \( f(t_0) \geq U(t_0) \) and let
\[ t_1 = \sup\{ s \leq t_0 : f(s) \leq U(s) \}. \]
By the continuity of \( f \) and \( U \), we obtain \( f(t_1) = U(t_1) \) and \( f(r) > U(r) \) for \( r \in (t_1, t_0) \). From this we conclude that
\[ f(t) \leq f(t_1) - \lambda \int_{t_1}^t f(r) \, dr < U(t_1) - \lambda \int_{t_1}^t U(r) \, dr = U(t), \quad t \in (t_1, t_0), \]
and hence \( t_0 = t_1 \). Thus we have
\[(2.12) \quad \mathbb{E} \psi(\rho(X_t, Y_t)^p) \leq e^{-\lambda t} \psi(\rho(x,y)^p). \]
From Lemma 2.10 (ii) that there exist two constants \( \tilde{c}_1 \) and \( \tilde{c}_2 \) such that
\[(2.13) \quad \tilde{c}_1 r^{1/p} \leq \psi(r) \leq \tilde{c}_2 r^{1/p}. \]
Combining (2.13) with (2.12) we obtain the estimate:
\[(2.14) \quad \mathbb{E} \rho(X_t, Y_t) \leq \frac{1}{\tilde{c}_1} \mathbb{E} \psi(\rho(X_t, Y_t)^p) \leq \frac{\tilde{c}_2}{\tilde{c}_1} e^{-\lambda t} \rho(x,y). \]
Recall that
\[ d \psi(\rho(X_t, Y_t)^p) \leq dM_t - \ell(\rho(X_t, Y_t)^p) \psi'(\rho(X_t, Y_t)^p) \, dt. \]
Since \( \sigma(\rho(X_t, Y_t)) = 0 \) for \( \rho(X_t, Y_t) \geq r_0 + 1 \) while \( d \psi(\rho(X_t, Y_t)^p) < 0 \) if \( \rho(X_t, Y_t) \geq r_0 + 1 \), we have
\[ \psi(\rho(X_t, Y_t)^p) \leq \psi((r_0 + 1)^p \lor \rho^p(x,y)), \]
which together with the fact that \( \psi' > 0 \) implies
\[ \rho(X_t, Y_t) \leq (r_0 + 1) \lor \rho(x,y). \]
Combined with (2.14) this implies
\[ \mathbb{E}^{(x,y)}[\rho(X_t, Y_t)^p] \leq ((1 + r_0) \lor \rho(x,y))^{p-1} \mathbb{E}[\rho(X_t, Y_t)] \]
\[ \leq \frac{\tilde{c}_2}{\tilde{c}_1} e^{-\lambda t} (1 + r_0)^{p-1} \rho(x,y) \lor \rho(x,y)^p. \]
According to Remark 2.1, under the assumption that
\[ \liminf_{\rho(x) \to \infty} \text{Ric}^Z(x) > 0, \]
we can find positive constants $k_1$ and $k_2$ such that

$$I^Z(x, y) \leq k_1 - k_2 \rho(x, y),$$

and then by Theorem 2.2, there exist constants $c$ and $\lambda$ such that (1.4) holds. More precisely, we have now the following results with explicit values for $c$ and $\lambda$.

**Corollary 2.5.** Assume that

$$(2.15) \quad I^Z(x, y) \leq k_1 - k_2 \rho(x, y),$$

for some constants $k_1 \geq 0$ and $k_2 > 0$. Then,

(i) for $p > 1$, $t \geq 0$, and $x, y \in M$,

$$W_p(\delta_x Y_t, \delta_y Y_t) \leq \left(1 + \frac{2k_1}{k_2}\right)^{(p-1)/p} \exp\left(\frac{k_1^2}{pk_2} - \frac{k_2}{2pe^{k_1^2/k_2}t}\right) (\rho(x, y) \vee \rho(x, y)^{1/p});$$

(ii) for $p > 1$, $t \geq 0$ and $\mu_1, \mu_2 \in \mathcal{P}(M)$,

$$\tilde{W}_p(\mu_1 P_t, \mu_2 P_t) \leq \left(1 + \frac{2k_1}{k_2}\right)^{(p-1)/p} \exp\left(\frac{k_1^2}{pk_2} - \frac{k_2}{2pe^{k_1^2/k_2}t}\right) \tilde{W}_p(\mu_1, \mu_2),$$

where

$$\tilde{W}_p(\mu_1, \mu_2) = \inf_{\pi \in \mathcal{C}(\mu_1, \mu_2)} \left(\int_{M \times M} \rho(x, y)^p \vee \rho(x, y) \pi(dx, dy)\right)^{1/p};$$

(iii) in particular, for $t \geq 0$,

$$\mathbb{E}^{(x, y)} \rho(X_t, Y_t) \leq \exp\left(\frac{k_1^2}{2k_2} - \frac{k_2}{2e^{k_1^2/k_2}t}\right) \rho(x, y).$$

**Proof.** By assumption, we have

$$I^Z(x, y) \leq k_1 - k_2 \rho(x, y).$$

Let $r_0 = 2k_1/k_2$. Then, for $r \geq r_0$, we have $k_1 \leq k_2 r/2$, or equivalently,

$$k_1 - k_2 r \leq -\frac{1}{2} k_2 r.$$

Thus, we find $k_3 = k_2/2$ and

$$\lambda = k_3 \exp\left(-\frac{1}{4} \int_0^{r_0} k_1 dr - \frac{k_2}{8} r_0^2\right) = \frac{k_2}{2} \exp\left(-\frac{k_1^2}{k_2}\right).$$

Substituting the explicit constants in the results of Theorem 2.2, we complete the proof. □
Corollary 2.6. Keeping the assumptions as in Theorem 2.2, we have for any \( t \geq 0 \) and any \( f \in C^\infty_0(M) \),

\[
|\nabla P_t f| \leq c_1 e^{-\lambda t} \|\nabla f\|_\infty
\]

where the constants \( c_1 \) and \( \lambda \) are given in Theorem 2.2.

Proof. For \( f \in C^\infty_0(M) \), according to the definition of \( |\nabla P_t f| \), we have

\[
|\nabla P_t f|(x) = \lim_{\rho(x,y) \to 0} \left| \frac{P_t f(x) - P_t f(y)}{\rho(x,y)} \right|
\]

\[
= \lim_{\rho(x,y) \to 0} \mathbb{E}^{(x,y)} \left[ \frac{f(X_t) - f(Y_t)}{\rho(X_t, Y_t)} \frac{\rho(X_t, Y_t)}{\rho(x,y)} \right]
\]

\[
\leq c_1 e^{-\lambda t} \|\nabla f\|_\infty
\]

for \( t \geq 0 \). \( \square \)

3. Exponential Contraction in Wasserstein Distance on Evolving Manifolds

In this section, we deal with the case that the underlying manifold carries a geometric flow of complete Riemannian metrics. More precisely, we consider a \( d \)-dimensional differentiable manifold \( M \) equipped with a \( C^1 \) family of complete Riemannian metrics \((g_t)_{t \in (-\infty, T_c)}\) for some \( T_c \in (-\infty, \infty] \). We denote the interval \((-\infty, T_c)\) by \( I \).

We first give some quantitative results concerning exponential contraction in Wasserstein distance over evolving manifolds. As application, we use the \( W_1 \)-contraction inequality to derive a gradient inequality and uniqueness for the evolution system of measure.

3.1. Main results

Let \( \nabla^t \) be the Levi-Civita connection and \( \Delta_t \) the Laplace-Beltrami operator associated with the Riemannian metric \( g_t \). In addition, let \((Z_t)_{t \in [0, T_c]}\) be a \( C^1 \)-family of vector fields on \( M \). We set

\[
I^Z(t, x, y) = I(t, x, y) + \langle Z_t, \nabla^t \rho_t(\cdot, y) \rangle_t + \langle Z_t, \nabla^t \rho_t(\cdot, x) \rangle_t
\]

where

\[
I(t, x, y) = \int_0^{\rho_t(x,y)} \sum_{i=1}^{d-1} \left\{ |\nabla_{\gamma}^t J_i^t|^2 - \langle R_t(\gamma, J_i^t) \gamma, J_i^t \rangle_t \right\} (\gamma_s) + \partial_t g_t(\gamma, \gamma)(\gamma_s) ds.
\]
Now \( \rho_t \) is the Riemannian distance, \( R_t \) the Riemannian tensor, and \( \gamma : [0, \rho_t(x, y)] \to M \) the minimal geodesic from \( x \) to \( y \) with unit speed, everything taken with respect to the Riemannian metric \( g_t \); in addition, \( \{ J^i_t \}_{i=1}^{d-1} \) are Jacobi fields along \( \gamma \) such that

\[
J^i_t(y) = P^i_{x,y}[J^i_t(x)], \quad i = 1, \ldots, d - 1,
\]

in terms of the parallel transport \( P^i_{x,y} : T_xM \to T_yM \) along the geodesic \( \gamma \), and such that

\[
\{ \gamma(s), J^i_t(s) : 1 \leq i \leq d - 1 \}, \quad s = 0, \rho_t(x, y)
\]

are orthonormal bases of the tangent spaces \( T_xM \), respectively \( T_yM \), with respect to \( g_t \).

We first give a precise formulation of our assumptions in the time-dependent case.

**Assumption (A2).** There exist a non-negative continuous function \( k_1 \in C(0, \infty) \), a positive constant \( k_2 \) and a constant \( \theta \geq 0 \) such that

\[
I^Z(t,x,y) \leq k_1(\rho_t(x,y)) - k_2 \rho_t(x,y)^{1+\theta}
\]

and such that there exist positive constants \( k_3 \) (\( k_3 < k_2 \)) and \( r_0 \) with the property:

\[
k_1(r) - k_2 r^{1+\theta} \leq -k_3 r^{1+\theta}, \quad r \geq r_0,
\]

and \( \int_0^r k_1(v) \, dv < \infty \) for each \( r > 0 \).

Consider the operator \( L_t = \Delta_t + Z_t \) where \( Z_t \) is a family of vector fields which is \( C^1 \) in \( t \). Let \( (X_t) \) be the diffusion process generated by \( L_t \) which is assumed to be non-explosive up to time \( T_c \), and let \( P_{s,t} \) be the corresponding time-inhomogeneous semigroup.

**Theorem 3.1.** Assume that Assumption (A2) holds. Then

(i) for \( x, y \in M, p \geq 1 \) and \( s \leq t < T_c \),

\[
W_{p,t}(\delta_x P_{s,t}, \delta_y P_{s,t}) \leq c_p e^{-\lambda(t-s)/p} (\rho_s(x,y) \vee \rho_s(x,y)^{1/p}),
\]

where

\[
c_p = (1 + r_0)^{(p-1)/p} \exp \left( \frac{1}{4p} \int_0^{r_0} k_1(r) \, dr + \frac{k_2}{8p} r_0^{2+\theta} \right),
\]

\[
\lambda = k_3 r_0^\theta \exp \left( -\frac{1}{4} \int_0^{r_0} k_1(r) \, dr - \frac{k_2}{8} r_0^{2+\theta} \right);
\]

(ii) for \( s \leq t < T_c, p > 1 \) and \( \mu_1, \mu_2 \in \mathcal{P}(M) \), we have

\[
\tilde{W}_{p,t}(\mu_1 P_{s,t}, \mu_2 P_{s,t}) \leq c_p e^{-\lambda(t-s)/p} \tilde{W}_{p,s}(\mu_1, \mu_2),
\]

where

\[
\tilde{W}_{p,t}(\mu_1, \mu_2) = \inf_{\pi \in \mathcal{C}(\mu_1, \mu_2)} \left( \int_{M \times M} \rho_t(x,y)^p \vee \rho_t(x,y) \pi(dx, dy) \right)^{1/p}.
\]
(iii) for \( s \leq t < T_c \) and \( \mu_1, \mu_2 \in \mathcal{P}(M) \),
\[
W_{1,t}(\mu_1 P_{s,t}, \mu_2 P_{s,t}) \leq c_1 e^{-\lambda(t-s)} W_{1,s}(\mu_1, \mu_2).
\]

**Proof.** Let \( X_t \) be the \( L_t \)-diffusion process, which we assume to be non-explosive. It is well known that the process \( X_t \) solves the following SDE:
\[
d_tX_t = \sqrt{2}u_t dB_t + Z_t(X_t)dt, \quad X_s = x,
\]
where \((B_t)_{t \geq s} \) is a \( d \)-dimensional Brownian motion on \( \mathbb{R}^d \). Here \((u_t)_{t \geq s} \) is a horizontal lift of \((X_t)_{t \geq s} \) to the frame bundle over \( M \) such that the parallel transport
\[
u_t u_s^{-1} : (T_s M, g_s) \to (T_X M, g_t)
\]
along \( X_t \) is isometric. We may rewrite SDE (3.4) as
\[
d_tX_t = \sqrt{2} \left( \sigma(\rho_t(X_t, Y_t)) u_t dB'_t + \sqrt{1 - \sigma(\rho_t(X_t, Y_t))^2} u_t dB''_t \right) + Z_t(X_t)dt,
\]
where \( B'_t \) and \( B''_t \) are two independent Brownian motion on \( \mathbb{R}^d \). Recall that \( \sigma \in C^1([0, \infty)) \) is a function satisfying \( 0 < \sigma \leq 1 \) when \( r \in (r_0, r_0 + 1) \), and \( \sigma \equiv 0 \) when \( r \geq r_0 + 1 \) and \( \sigma \equiv 1 \) when \( r \leq r_0 \). Let \( Y_t \) solve the following SDE on \( M \) (with initial condition \( Y_s = y \)):
\[
d_tY_t = \sqrt{2} \left( \sigma(\rho_t(X_t, Y_t)) M^t_{X_t, Y_t} u_t dB'_t + \sqrt{1 - \sigma(\rho_t(X_t, Y_t))^2} P^t_{X_t, Y_t} u_t dB''_t \right) + Z_t(Y_t)dt,
\]
where \( P^t_{X_t, Y_t} \) and \( M^t_{X_t, Y_t} \) denote respectively the parallel transport and the mirror reflection along the \( g_t \)-geodesic connecting \( X_t \) and \( Y_t \) with respect to the metric \( g_t \). Since the coefficients of the SDE are at least \( C^1 \) outside the diagonal \( \{(z, z) : z \in M\} \), it has a unique solution up to the coupling time
\[
T := \inf\{t \geq s : X_t = Y_t\}.
\]
Let \( X_t = Y_t \) for \( t \geq T \) as usual. Then, by Itô’s formula (see [5]), we have
\[
d\rho_t(X_t, Y_t) \leq 2\sqrt{2} db_t + I^Z(t, X_t, Y_t)dt
\leq 2\sqrt{2} db_t + (k_1(\rho_t(X_t, Y_t)) - k_2\rho_t(X_t, Y_t)^{1+\theta})dt, \quad t \leq T,
\]
where \( b_t \) is a one-dimensional Brownian motion on \( \mathbb{R} \). Moreover,
\[
d \rho_t(X_t, Y_t)^p \leq p \rho_t(X_t, Y_t)^{p-1} d \rho_t(X_t, Y_t) + \frac{1}{2} p(p-1) \rho_t(X_t, Y_t)^{p-2} d\langle \rho \rangle_t
\leq p \rho_t(X_t, Y_t)^{p-1} \left\{ 2\sqrt{2} db_t + \left( k_1(\rho_t(X_t, Y_t)) - k_2\rho_t(X_t, Y_t)^{1+\theta} \right) dt \right\}
+ 4p(p-1) \rho_t(X_t, Y_t)^{p-2} dt.
\]
Then, by the Itô formula for \( \psi(\rho_t(X_t, Y_t)^p) \), we have
\[
d \psi(\rho_t(X_t, Y_t)^p) \leq \psi' \rho_t(X_t, Y_t)^p \left( 2\sqrt{2} p \rho_t(X_t, Y_t)^{p-1} db_t + \ell_1(\rho_t(X_t, Y_t)^p) dt \right)
\]
\[
\begin{align*}
\psi''(\rho_t(X_t, Y_t)^p) & \ell_0(\rho_t(X_t, Y_t)^p) dt \\
= dM_t - \ell(\rho_t(X_t, Y_t)^p) \psi'(\rho_t(X_t, Y_t)^p) dt
\end{align*}
\]

where
\[
dM_t = 2\sqrt{2}p \psi'(\rho_t(X_t, Y_t)^p) \rho_t(X_t, Y_t)^{p-1} db_t.
\]

The remaining steps are the same as in the proof of Theorem 2.2. We skip the details.

\[\square\]

**Remark 3.2.** It is natural to ask whether contraction in Wasserstein distance still holds when the curvature condition (3.1) is weakened as follows: there exist non-negative functions \(k_1, k_2 \in C^1(I)\) and \(\phi \in C([0, \infty))\) such that

\[
I^Z(t, x, y) \leq k_1(t) \phi(\rho_t(x, y)) - k_2(t) \rho_t(x, y)^{1+\theta}.
\]

A possible way to deal with this case is to prove the result for each interval \([s, t]\). Assume that for an interval \([s, t] \subset I\),

\[
I^Z(u, x, y) \leq k_1(s, t) \phi(\rho_u(x, y)) - k_2(s, t) \rho_u(x, y)^{1+\theta}, \quad u \in [s, t],
\]

and there exist \(k_3(s, t)\) and \(r_0(s, t)\) such that

\[
k_1(s, t) \phi(r) - k_2(s, t) r^{1+\theta} \leq -k_3(s, t) r^{1+\theta}, \quad r \geq r_0(s, t)
\]

and \(\int_0^r \phi(u) du < \infty\) for \(r > 0\). Then, by an analogous procedure as in the proof of Theorem 3.1, we get

\[
W_{p,t}(\delta_x P_{s,t}, \delta_y P_{s,t}) \leq c_p(s, t) e^{-\lambda(s,t)(t-s)/p} \left( \rho_s(x, y) \vee \rho_s(x, y)^{1/p} \right).
\]

Hence, if the coefficient \(c_p(s, t) e^{-\lambda(s,t)(t-s)/p}\) converges to 0, as \(t-s \to \infty\), we still have contraction of the Wasserstein distance \(\tilde{W}_{p,t}\).

Assume that \(\text{Ric}^Z_t \geq k(\rho_t)\) and \(\liminf_{r \to \infty} k(r) > 0\). Then there exist positive constants \(k_1\) and \(k_2\) such that

\[
I(t, x, y) \leq k_1 - k_2 \rho_t(x, y).
\]

In this case, the following corollary follows directly from Theorem 3.1.

**Corollary 3.3.** Suppose that

\[
I^Z(t, x, y) \leq k_1 - k_2 \rho_t(x, y), \quad t \in I
\]

for some non-negative constant \(k_1\) and positive constant \(k_2\). Then,

1. for \(p > 1\), \(s \leq t < T_c\), and \(x, y \in M\),

\[
W_{p,t}(\delta_x P_{s,t}, \delta_y P_{s,t}) \leq \left( 1 + \frac{2k_1}{k_2} \right)^{(p-1)/p} \exp \left( \frac{k_1^2}{pk_2} - \frac{k_2}{2p} e^{k_1^2/k_2}(t-s) \right) \times \left( \rho_s(x, y) \vee \rho_s(x, y)^{1/p} \right);
\]
(ii) for $s \leq t < T_c$, $p > 1$ and $\mu_1, \mu_2 \in \mathcal{P}(M)$,

$$\bar{W}_{p,t}(\mu_1 P_{s,t}, \mu_2 P_{s,t}) \leq \left(1 + \frac{2k_1}{k_2}\right)^{(p-1)/p} \exp\left(\frac{k_1^2}{pk_2} - \frac{k_2}{2p\exp^{k_1^2/k_2}}(t-s)\right) \bar{W}_{p,s}(\mu_1, \mu_2)$$

where

$$\bar{W}_{p,s}(\mu_1, \mu_2) = \inf_{\pi \in \mathcal{C}(\mu_1, \mu_2)} \left(\int_{M \times M} \rho_s(x,y)^p \vee \rho_s(x,y) \pi(dx,dy)\right)^{1/p};$$

(iii) in particular, for $s \leq t < T_c$,

$$\mathbb{E}^{(x,y)}(\rho_t(X_t, Y_t) \leq \exp\left(\frac{k_1^2}{k_2} - \frac{k_2}{2 \exp^{k_1^2/k_2}}(t-s)\right) \rho_s(x,y).$$

We now apply Theorem 3.1 (iii) to derive gradient estimates for the 2-parameter semigroup $P_{s,t}$.

**Corollary 3.4.** Under the same conditions as in Theorem 3.1, we have

$$|\nabla^s P_{s,t} f|_s \leq c_1 e^{-\lambda(t-s)} \|\nabla^t f\|_t$$

for any $s \leq t$ and any $f \in C^\infty_0(M)$, where $c_1$ and $\lambda$ are defined as in (3.2) and (3.3) respectively.

**Proof.** For $f \in C^\infty_0(M)$, according to the definition of $\nabla^s P_{s,t} f$, we have for $s \leq t$,

$$|\nabla^s P_{s,t} f|_s(x) = \lim_{\rho_s(x,y) \to 0} \frac{|P_{s,t} f(x) - P_{s,t} f(y)|}{\rho_s(x,y)} = \lim_{\rho_s(x,y) \to 0} \mathbb{E}^{(s,(x,y))}\left(\frac{f(X_t) - f(Y_t)}{\rho_t(X_t, Y_t)} \frac{\rho_t(X_t, Y_t)}{\rho_s(x,y)}\right) \leq c_1 e^{-\lambda(t-s)} \|\nabla^t f\|_t.$$  

\[\square\]

### 3.2. Applications

Let us first recall the notion of an evolution system of measures for a 2-parameter semigroup. A family of Borel probability measures $(\mu_t)_{t \in I}$ on $M$ is called an evolution system of measures for $P_{s,t}$ (see [7]) if

$$\int_M P_{s,t} \phi d\mu_s = \int \phi d\mu_t, \quad \phi \in \mathcal{B}(M)$$

for $s \leq t < T_c$. In [6], we investigated existence and uniqueness of evolution systems of measures. The condition for uniqueness (H3) in [6, Theorem 2.3] requires that the lower bound of $\text{Ric}_t^Z - \frac{1}{2} \partial_t g_t$ depends only on time $t$ and satisfies an integrability condition. Here we give another condition in terms of lower bounds on $\text{Ric}_t^Z - \frac{1}{2} \partial_t g_t$ depending on the radial distance $\rho_t$. 
Theorem 3.5. Suppose that there exists a function \( k \in C([0, \infty)) \) with 
\[
\liminf_{s \to \infty} k(s) > 0
\]
such that
\[
\text{Ric}_t^Z - \frac{1}{2} \partial_t g_t \geq k(\rho_t)
\]
and that there exist \( \varepsilon > 0 \) and \( x_0 \in M \) such that for some constant \( C \),
\[
3k_{\varepsilon}(t)\varepsilon + |Z_t|_t(x_0) \leq C, \quad t \in I,
\]
where
\[
k_{\varepsilon}(t) := \sup\{\text{Ric}_t(x) : \rho_t(x_0, x) \leq \varepsilon\}.
\]
Then there exists a unique evolution system of measures \((\mu_s)_{s \in I}\) for \( P_{s,t} \).

Proof. First of all, by [11, Lemma 9], we have
\[
(L_t + \partial_t)\rho_t^2 = 2\rho_t(L_t + \partial_t)\rho_t + 2
\]
\[
\leq 2 \left( F_t(\rho_t) - \int_0^{\rho_t} k(\rho_t(\gamma(s))) \, ds + |Z_t(x_0)|_t \right) \rho_t + 2,
\]
where
\[
F_t(s) = \sqrt{k_{\varepsilon}(t)(d-1)} \coth \left( \frac{\sqrt{k_{\varepsilon}(t)(d-1)(s \wedge \varepsilon)}}{s \wedge \varepsilon} \right) + k_{\varepsilon}(t)(s \wedge \varepsilon)
\]
and \( k_{\varepsilon}(t) \) is given by Eq. (3.8). There exists a positive constant \( c \) such that
\[
- \int_0^{\rho_t} k(\rho_t(\gamma(s))) \, ds = - \int_0^{\rho_t} k(\rho_t(\gamma(s))) \, ds - \int_{r_0}^{\rho_t} k(\rho_t(\gamma(s))) \, ds
\]
\[
\leq -\sigma(\rho_t - r_0) - \int_{r_0}^{\rho_t} k(\rho_t(\gamma(s))) \, ds
\]
\[
\leq -\sigma \rho_t + \sigma r_0 - \int_0^{\rho_t} k(\rho_t(\gamma(s))) \, ds
\]
\[
\leq -\sigma \rho_t + c.
\]
As \( \text{Ric}_t^Z - \frac{1}{2} \partial_t g_t \geq k(\rho_t) \) and \( \liminf_{s \to \infty} k(s) > 0 \), the function \( k \) is bounded below and there exist constants \( r_0 > 0 \) and \( \kappa > 0 \) such that for \( r \geq r_0 \),
\[
k(r) \geq \kappa > 0.
\]
By straightforward estimates, using the obvious inequality \( \coth(x) \leq 1 + \frac{1}{x} \), we obtain
\[
(L_t + \partial_t)\rho_t^2 \leq 2d + (3k_{\varepsilon}(t)\varepsilon + |Z_t|_t(x_0) + 3(d-1)\varepsilon^{-1}) \rho_t + c \rho_t - 2k_{\rho_t^2}.
\]
Thus, if \( 3k_{\varepsilon}(t)\varepsilon + |Z_t|_t(x_0) \leq C \), we can find constants \( C_1 \) and \( C_2 \) such that
\[
(L_t + \partial_t)\rho_t^2 \leq C_1 - C_2 \rho_t^2.
\]
Therefore, by [6, Theorem 2.3], there exists an evolution system of measures \((\mu_s)\) such that

\[
\sup_{s \in (-\infty, t]} \mu_s(p_s^2) \leq \frac{C_1}{C_2} < \infty.
\]

Recall that \(\text{Ric}^Z_t - \frac{1}{2} \partial_t g_t \geq k(\rho_t)\) with \(k(r) > \kappa > 0\) for \(r_0 > 0\). Moreover, given condition (3.7), there exist positive constants \(k_1\) and \(k_2\) such that

\[
I(t, x, y) \leq - \int_0^1 \rho_t(x, y) \left( \text{Ric}^Z_t - \frac{1}{2} \partial_t g_t \right) (\gamma(s), \dot{\gamma}(s)) ds \leq k_1 - k_2 \rho_t(x, y).
\]

Hence condition (3.6) in Corollary 3.3 is satisfied, and by this corollary, there exist constants \(c_1\) and \(\lambda\) depending on \(k_1\) and \(k_2\) such that

\[
|P_{s,t}f(x_0) - \mu_t(f)| = \left| \int (P_{s,t}f(x_0) - P_{s,t}f(y)) \mu_s(dy) \right|
\]

\[
= \left| \int \mathbb{E}^{(s, (x_0, y))} \left[ \frac{f(X_t) - f(Y_t)}{\rho_t(X_t, Y_t)} \rho_t(X_t, Y_t) \right] \mu_s(dy) \right|
\]

\[
\leq \| |\nabla^t f|_t \|_\infty \int \mathbb{E}^{(s, (x_0, y))} [\rho_t(X_t, Y_t)] \mu_s(dy)
\]

\[
\leq c_1 e^{-\lambda(t-s)} \| |\nabla^t f|_t \|_\infty \mu_s(\rho_s)
\]

\[
\leq c_1 e^{-\lambda(t-s)} \| |\nabla^t f|_t \|_\infty \sqrt{C_1/C_2},
\]

which implies

\[
\lim_{s \to -\infty} |P_{s,t}f(x_0) - \mu_t(f)| = 0.
\]

If there is another evolution system of measures \(\nu_t\), then

\[
|\mu_t(f) - \nu_t(f)| \leq \lim_{s \to -\infty} (|P_{s,t}f(x_0) - \mu_t(f)| + |P_{s,t}f(x_0) - \nu_t(f)|) = 0,
\]

i.e. \(\mu_t \equiv \nu_t\). This finishes the proof. \(\square\)

**Remark 3.6.** Comparing the above conditions to [6, Theorem 2.3], we note that the function \(k(r)\) is only required to be positive outside a compact set. If \(k(r)\) is not bounded below by zero, the situation is not covered by [6, Theorem 2.3].

It is well known that evolution systems of measures play a similar role in the inhomogeneous setting as invariant measures for homogeneous semigroups \(P_t\). Inspired by this, we take the system \((\mu_s)_{s \in I}\) as reference measures and study the contraction properties of the two-parameter semigroup \(P_{s,t}\).

**Theorem 3.7.** We keep the assumptions of Theorem 3.5 and assume that \(\mu_s(e^{\varepsilon \rho_s}) < \infty\) for any \(\varepsilon > 0\). Then the semigroup \(P_{s,t}\) is supercontractive.

The idea is to first establish a dimension-free Harnack inequality under assumption (3.9) below.
Lemma 3.8. Suppose that there exist constants $k_1, k_2$ such that

$$I^Z(t, x, y) \leq k_1 - k_2 \rho_t(x, y). \tag{3.9}$$

Then, for any $p > 1$, the following dimension-free Harnack inequality holds:

$$(P_{s,t} f)^p(x) \leq P_{s,t} (f^p)(y) \exp \left( \frac{p}{4(p-1)} \left( k_1^2 (t-s) + 4k_1 \rho_s(x, y) \frac{2k_2 \rho_s(x, y)^2}{e^{k_2(t-s)} + 1} \right) \right)$$

for any non-negative function $f \in \mathcal{B}_b(M)$ and $s \leq t < T_c$.

**Proof.** Let $X_t$ solve the stochastic differential equation

$$d_t X_t = \sqrt{2} u_t dB_t + Z_t(X_t) dt, \quad X_s = x,$$

and let $Y_t$ solve the stochastic differential equation

$$d_t Y_t = \sqrt{2} P_{X_t,Y_t} u_t dB_t + (Z_t(Y_t) + \xi(t, X_t, Y_t)) dt, \quad Y_s = y,$$

where the function $\xi \in C^1(I \times M \times M)$ will be specified later. Since the coefficients of the coupled SDE are at least $C^1$ outside the diagonal $\{(z, z) : z \in M\}$, the coupled SDE has a unique solution up to the coupling time

$$\tau := \inf\{t \geq s : X_t = Y_t\}.$$

Let $X_t = Y_t$ for $t \geq \tau$ as usual. By Itô’s formula, we have

$$d \rho_t(X_t, Y_t) \leq I^Z(t, X_t, Y_t) dt - \xi_t dt \leq (k_1 - k_2 \rho_t(X_t, Y_t) - \xi_t) dt, \quad t \leq \tau,$$

where $\xi_t := \xi(t, X_t, Y_t)$. Now, for a fixed constant $T \in (s, T_c)$, let

$$\xi_t = k_1 + \frac{2k_2 e^{k_2(t-s)} \rho_s(x, y)}{e^{2k_2(T-s)} - 1}, \quad t \geq s.$$

Then

$$\int_s^T (k_1 - \xi_t) e^{k_2(t-s)} dt = -\rho_s(x, y),$$

and

$$\rho_T(X_T, Y_T) - \rho_s(x, y) \leq \int_s^T (k_1 - \xi_t) e^{k_2(t-s)} dt - \int_s^T \rho_t(X_t, Y_t) dt \leq -\rho_s(x, y) - \int_s^T \rho_t(X_t, Y_t) dt.$$  

From this, it is easy to see that $\tau \leq T$ and hence $X_T = Y_T$.

Now due to Girsanov’s theorem, $Y$ is generated by $L_t$ under the weighted probability measure $R \mathbb{P}$ where the density $R$ is given by

$$R = \exp \left( \frac{1}{\sqrt{2}} \int_s^\tau \langle \xi_t \nabla' \rho_t(X_t, \cdot)(Y_t), P_{X_t,Y_t} u_t dB \rangle_t - \frac{1}{4} \int_s^\tau \xi_t^2 dt \right).$$
Thus,

\[(P_{s,T}f(y))^p \leq (P_{s,T}f^p(x)) \left( \mathbb{E}R^{p/(p-1)} \right)^{p-1}.\]

Since \( \tau \leq T \) and

\[N_t := \exp \left( \frac{p}{\sqrt{2}(p-1)} \int_s^t \langle \xi_r \nabla_r \rho_r(X_r, \cdot), P_{X_r,Y_r,ur}dB_r \rangle_r - \frac{p^2}{4(p-1)^2} \int_s^t \xi_r^2 dr \right)\]

is a martingale, we have \( \mathbb{E}N_T = 1 \) and hence,

\[\mathbb{E}R^{p/(p-1)} = \mathbb{E} \left[ N_T \exp \left( \frac{p}{4(p-1)^2} \int_s^T \xi_r^2 dr \right) \right] \]

\[\leq \exp \left( \frac{p}{4(p-1)^2} \int_s^T \xi_r^2 dr \right) \]

\[= \exp \left( \frac{p}{4(p-1)^2} \left( k_1^2(T - s) + \frac{4k_1 \rho_s(x,y)}{e^{k_2(t-s) + 1}} + \frac{2k_2 \rho_s(x,y)^2}{e^{2k_2(t-s) - 1}} \right) \right).\]

It follows that

\[(P_{s,T}f(x))^p \leq (P_{s,T}f^p(y)) \exp \left( \frac{p}{4(p-1)^2} \left( k_1^2(T - s) + \frac{4k_1 \rho_s(x,y)}{e^{k_2(t-s) + 1}} + \frac{2k_2 \rho_s(x,y)^2}{e^{2k_2(t-s) - 1}} \right) \right),\]

as claimed. \( \square \)

**Proof of Theorem 3.7.** As explained in the proof of Theorem 3.5, there exist positive constants \( k_1 \) and \( k_2 \) such that (3.9) holds. Noting that \( (\mu_s) \) is the evolution system of measures and using Lemma 3.8, we have

\[1 = \int_M P_{s,t} |f|^p \mu_s(dy) \]

\[\geq |P_{s,t}f|^p(x) \int_M \exp \left( - \frac{p}{4(p-1)} \left( k_1^2(t-s) + \frac{4k_1 \rho_s(x,y)}{e^{k_2(t-s) + 1}} + \frac{2k_2 \rho_s(x,y)^2}{e^{2k_2(t-s) - 1}} \right) \right) \mu_s(dy) \]

\[\geq |P_{s,t}f|^p(x) \int_{B_s(x_0,1)} \exp \left( - \frac{p}{4(p-1)} \left( k_1^2(t-s) + \frac{4k_1(\rho_s(x) + 1)}{e^{k_2(t-s) + 1}} + \frac{2k_2(\rho_s(x) + 1)^2}{e^{2k_2(t-s) - 1}} \right) \right) \mu_s(dy) \]

\[\geq |P_{s,t}f|^p(x) \mu_s(B_s(x_0,1)) \exp \left( -pC(t-s, p, k_1, k_2) \right. \]

\[\left. - \frac{p(2k_1 e^{k_2(t-s) + k_2 - 2k_1})}{(p-1)(e^{2k_2(t-s) - 1})} \rho_s(x)^2 \right),\]

where \( B_s(x_0, 1) = \{ x \in M : \rho_s(x) \leq 1 \} \) is the unit geodesic ball (with respect to \( g_s \)) centered at \( x_0 \) and \( C(t-s, p, k_1, k_2) \) is a constant depending on \( t-s, p, k_1 \) and \( k_2 \).
Letting
\[ \lambda = \frac{2k_1 e^{k_2(t-s)} + k_2 - 2k_1}{(p-1)(e^{2k_2(t-s)} - 1)}, \]
we get
\[ |P_{s,t}f|(x) \leq \frac{\exp(C(t-s,p,k_1,k_2))}{\mu_s(B_s(x_0,1))^{1/p}} e^{\lambda p^2_s} < \infty, \quad \mu_t(|f|^p) = 1. \]

Therefore
\[ \mu_s(|P_{s,t}f|^q)^{1/q} \leq \frac{\exp(C(t-s,p,k_1,k_2))}{\mu_s(B_s(x_0,1))^{1/p}} (\mu_s(e^{\lambda qp^2_s}))^{1/q}. \]

Thus if \( \mu_s(e^{\lambda qp^2_s}) < \infty \) for some \( s \in I \), then \( P_{s,t} \) is supercontractive, i.e.,
\[ \|P_{s,t}\|_{(p,t) \to (q,s)} < \infty \]
for any \( 1 < p < q < \infty \). \( \square \)

Acknowledgments. This work has been partly supported by Fonds National de la Recherche Luxembourg (Open project O14/7628746 GEOMREV). Shao-Qin Zhang has been supported by the National Natural Science Foundation of China (Grant No. 11901604, 11771326).

REFERENCES

[1] M. Arnaudon, K. A. Coulibaly, and A. Thalmaier, Brownian motion with respect to a metric depending on time: definition, existence and applications to Ricci flow, C. R. Math. Acad. Sci. Paris 346 (2008), 13-14, 773–778.

[2] M. Arnaudon, K. A. Coulibaly, and A. Thalmaier, Horizontal diffusion in \( C^1 \) path space. In: Séminaire de Probabilités XLIII, Lecture Notes in Math., vol. 2006, Springer, Berlin, 2011, pp. 73–94.

[3] M. Arnaudon, A. Thalmaier, and F.-Y. Wang, Harnack inequality and heat kernel estimates on manifolds with curvature unbounded below. Bull. Sci. Math. 130 (2006), 3, 223–233.

[4] D. Bakry, I. Gentil, and M. Ledoux, Analysis and geometry of Markov diffusion operators. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, 2014.

[5] L.-J. Cheng, Diffusion semigroup on manifolds with time-dependent metrics. Forum Math. 29 (2017), 4, 775–798.

[6] L.-J. Cheng and A. Thalmaier, Evolution systems of measures and semigroup properties on evolving manifolds. Electron. J. Probab. 23 (2018), Paper No. 20, 27.

[7] G. Da Prato and M. Röckner, A note on evolution systems of measures for time-dependent stochastic differential equations. In Seminar on Stochastic Analysis, Random Fields and Applications V, Progr. Probab., vol. 59, Birkhäuser, Basel, 2008, pp. 115–122.

[8] A. Eberle, Reflection coupling and Wasserstein contractivity without convexity. C. R. Math. Acad. Sci. Paris 349 (2011), 19-20, 1101–1104.

[9] A. Eberle, Reflection couplings and contraction rates for diffusions. Probab. Theory Related Fields 166 (2016), 3-4, 851–886.
[10] K. Kuwada, *Gradient estimate for Markov kernels, Wasserstein control and Hopf-Lax formula.* In: *Potential theory and its related fields*, RIMS Kôkyûroku Bessatsu, B43, Res. Inst. Math. Sci. (RIMS), Kyoto, 2013, pp. 61–80.

[11] K. Kuwada and R. Philipowski, *Non-explosion of diffusion processes on manifolds with time-dependent metric.* Math. Z. **268** (2011), 3-4, 979–991.

[12] D. Luo and J. Wang, *Exponential convergence in L^p-Wasserstein distance for diffusion processes without uniformly dissipative drift.* Math. Nachr. **289** (2016), 14-15, 1909–1926.

[13] M.-K. von Renesse and K.-T. Sturm, *Transport inequalities, gradient estimates, entropy, and Ricci curvature.* Comm. Pure Appl. Math. **58** (2005), 7, 923–940.

[14] F.-Y. Wang, *Functional inequalities, Markov semigroups and spectral theory.* Science Press, Beijing, 2005.

[15] F.-Y. Wang, *Analysis for diffusion processes on Riemannian manifolds.* Advanced Series on Statistical Science & Applied Probability, vol. 18, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2014.

[16] F.-Y. Wang, *Exponential contraction in Wasserstein distances for diffusion semigroups with negative curvature.* Preprint arXiv:1603.05749 (2016).

[17] S.-Q. Zhang, *Exponential convergence in Wasserstein distance for diffusion semigroups with irregular drifts.* Preprint arXiv:1812.10190v3 (2018).

*University of Luxembourg, Department of Mathematics*
*Maison du Nombre, 4364 Esch-sur-Alzette, Luxembourg*

*Zhejiang University of Technology*
*Department of Applied Mathematics*
*Hangzhou 310023, The People’s Republic of China*
*lijuan.cheng@uni.lu and chenglj@zjut.edu.cn*

*University of Luxembourg, Department of Mathematics*
*Maison du Nombre, 4364 Esch-sur-Alzette, Luxembourg*
*anton.thalmaier@uni.lu*

*Central University of Finance and Economics*
*School of Statistics and Mathematics*
*Beijing 100081, The People’s Republic of China*
*zhangsq@cufe.edu.cn*