Horizontal Diffusion in C^1 Path Space

Marc Arnaudon, Koléhè Abdoulaye Coulibaly, and Anton Thalmaier

Abstract We define horizontal diffusion in C^1 path space over a Riemannian manifold and prove its existence. If the metric on the manifold is developing under the forward Ricci flow, horizontal diffusion along Brownian motion turns out to be length preserving. As application, we prove contraction properties in the Monge–Kantorovich minimization problem for probability measures evolving along the heat flow. For constant rank diffusions, differentiating a family of coupled diffusions gives a derivative process with a covariant derivative of finite variation. This construction provides an alternative method to filtering out redundant noise.

Keywords Brownian motion · Damped parallel transport · Horizontal diffusion · Monge–Kantorovich problem · Ricci curvature

1 Preliminaries

The main concern of this paper is to answer the following question: Given a second order differential operator L without constant term on a manifold M and a C^1 path $u \mapsto \varphi(u)$ taking values in M, is it possible to construct a one parameter family $X_t(u)$ of diffusions with generator L and starting point $X_0(u) = \varphi(u)$, such that the derivative with respect to u is locally uniformly bounded?

If the manifold is $\mathbb{R}^n$ and the generator L a constant coefficient differential operator, there is an obvious solution: the family $X_t(u) = \varphi(u) + Y_t$, where Y_t is

M. Arnaudon (✉) and K.A. Coulibaly
Laboratoire de Mathématiques et Applications CNRS: UMR 6086, Université de Poitiers, Téléport 2 - BP 30179, 86962 Futuroscope Chasseneuil Cedex, France
e-mail: marc.arnaudon@math.univ-poitiers.fr; abdoulaye.coulibaly@math.univ-poitiers.fr

K.A. Coulibaly and A. Thalmaier
Unité de Recherche en Mathématiques, FSTC, Université du Luxembourg,
6, rue Richard Coudenhove-Kalergi, 1359 Luxembourg, Grand-Duchy of Luxembourg
e-mail: abdoulaye.coulibaly@uni.lu; anton.thalmaier@uni.lu

C. Donati-Martin et al. (eds.), *Séminaire de Probabilités XLIII*, Lecture Notes in Mathematics 73
2006, DOI 10.1007/978-3-642-15217-7_2,

an L-diffusion starting at 0, has the required properties. But already on $\mathbb{R}^n$ with a non-constant generator, the question becomes difficult.

In this paper we give a positive answer for elliptic operators L on general manifolds; the result also covers time-dependent elliptic generators $L = L(t)$.

It turns out that the constructed family of diffusions solves the ordinary differential equation in the space of semimartingales:

$$\partial_u X_t(u) = W(X(u))_t(\dot{\varphi}(u)), \tag{1}$$

where $W(X(u))$ is the so-called deformed parallel translation along the semimartingale $X(u)$.

The problem is similar to finding flows associated to derivative processes as studied in [7–10, 12–15]. However it is transversal in the sense that in these papers diffusions with the same starting point are deformed along a drift which vanishes at time 0. In contrast, we want to move the starting point but to keep the generator. See Stroock [22], Chap. 10, for a related construction.

Our strategy of proof consists in iterating parallel couplings for closer and closer diffusions. In the limit, the solution may be considered as an infinite number of infinitesimally coupled diffusions. We call it horizontal L-diffusion in C^1 path space.

If the generator L is degenerate, we are able to solve (1) only in the constant rank case; by parallel coupling we construct a family of diffusions satisfying (1) at $u = 0$. In particular, the derivative of $X_t(u)$ at $u = 0$ has finite variation compared to parallel transport.

Note that our construction requires only a connection on the fiber bundle generated by the "carré du champ" operator. In the previous approach of [11], a stochastic differential equation is needed and ∇ has to be the Le Jan-Watanabe connection associated to the SDE.

The construction of families of $L(t)$-diffusions $X_\bullet(u)$ with $\partial_u X_\bullet(u)$ locally uniformly bounded has a variety of applications. In Stochastic Analysis, for instance, it allows to deduce Bismut type formulas without filtering redundant noise. If only the derivative with respect to u at $u = 0$ is needed, parallel coupling as constructed in [5, 6] would be a sufficient tool. The horizontal diffusion however is much more intrinsic by yielding a flow with the deformed parallel translation as derivative, well-suited to applications in the analysis of path space. Moreover for any u, the diffusion $X_\bullet(u)$ generates the same filtration as $X_\bullet(0)$, and has the same lifetime if the manifold is complete.

In Sect. 4 we use the horizontal diffusion to establish a contraction property for the Monge–Kantorovich optimal transport between probability measures evolving under the heat flow. We only assume that the cost function is a non-decreasing function of distance. This includes all Wasserstein distances with respect to the time-dependent Riemannian metric generated by the symbol of the generator $L(t)$. For a generator which is independent of time, the proof could be achieved using simple parallel coupling. The time-dependent case however requires horizontal diffusion as a tool.

2 Horizontal Diffusion in C^1 Path Space

Let M be a complete Riemannian manifold with ρ its Riemannian distance. The Levi–Civita connection on M will be denoted by ∇.

Given a continuous semimartingale X taking values in M, we denote by $\mathrm{d}^{\nabla}X = \mathrm{d}X$ its Itô differential and by $\mathrm{d}_m X$ the martingale part of $\mathrm{d}X$. In local coordinates,

$$\mathrm{d}^{\nabla}X \equiv \mathrm{d}X = \left(\mathrm{d}X^i + \frac{1}{2}\,\Gamma^i_{jk}(X)\,\mathrm{d}<X^j, X^k>\right)\frac{\partial}{\partial x^i} \tag{2}$$

where Γ^i_{jk} are the Christoffel symbols of the Levi–Civita connection on M. In addition, if

$$\mathrm{d}X^i = \mathrm{d}M^i + \mathrm{d}A^i$$

where M^i is a local martingale and A^i a finite variation process, then

$$\mathrm{d}_m X = \mathrm{d}M^i \frac{\partial}{\partial x^i}.$$

Alternatively, if

$$P_t(X) \equiv P_t^M(X) : T_{X_0}M \to T_{X_t}M$$

denotes parallel translation along X, then

$$\mathrm{d}X_t = P_t(X)\,\mathrm{d}\left(\int_0^{\bullet} P_s(X)^{-1}\delta X_s\right)_t$$

and

$$\mathrm{d}_m X_t = P_t(X)\,\mathrm{d}N_t$$

where N_t is the martingale part of the Stratonovich integral

$$\int_0^t P(X)_s^{-1}\delta X_s.$$

If X is a diffusion with generator L, we denote by $W(X)$ the so-called deformed parallel translation along X. Recall that $W(X)_t$ is a linear map $T_{X_0}M \to T_{X_t}M$, determined by the initial condition $W(X)_0 = \mathrm{Id}_{T_{X_0}M}$ and the covariant Itô stochastic differential equation:

$$DW(X)_t = -\frac{1}{2}\,\mathrm{Ric}^{\sharp}(W(X)_t)\,\mathrm{d}t + \nabla_{W(X)_t}Z\,\mathrm{d}t. \tag{3}$$

By definition we have

$$DW(X)_t = P_t(X)\mathrm{d}\left(P_{\bullet}(X)^{-1}W(X)\right)_t. \tag{4}$$

Note that the Itô differential (2) and the parallel translation require only a connection ∇ on M. For the deformed parallel translation (3) however the connection has to be adapted to a metric.

In this section the connection and the metric are independent of time. We shall see in Sect. 3 how these notions can be extended to time-dependent connections and metrics.

Theorem 2.1. *Let $\mathbb{R} \to M$, $u \mapsto \varphi(u)$, be a C^1 path in M and let Z be a vector field on M. Further let X^0 be a diffusion with generator*

$$L = \Delta/2 + Z,$$

starting at $\varphi(0)$, and lifetime ξ. There exists a unique family

$$u \mapsto (X_t(u))_{t\in[0,\xi[}$$

of diffusions with generator L, almost surely continuous in (t, u) and C^1 in u, satisfying $X(0) = X^0$, $X_0(u) = \varphi(u)$ and

$$\partial_u X_t(u) = W(X(u))_t(\dot{\varphi}(u)). \tag{5}$$

Furthermore, the process $X(u)$ satisfies the Itô stochastic differential equation

$$\mathrm{d}X_t(u) = P_{0,u}^{X_t(\cdot)}\,\mathrm{d}_m X_t^0 + Z_{X_t(u)}\,\mathrm{d}t, \tag{6}$$

where $P_{0,u}^{X_t(\cdot)} : T_{X_t^0}M \to T_{X_t(u)}M$ denotes parallel transport along the C^1 curve

$$[0, u] \to M, \quad v \mapsto X_t(v).$$

Definition 2.2. We call $t \mapsto (X_t(u))_{u\in\mathbb{R}}$ the horizontal L-diffusion in C^1 path space $C^1(\mathbb{R}, M)$ over X^0, starting at φ.

Remark 2.3. Given an elliptic generator L, we can always choose a metric g on M such that

$$L = \Delta/2 + Z$$

for some vector field Z where Δ is the Laplacian with respect to g. Assuming that M is complete with respect to this metric, the assumptions of Theorem 2.1 are fulfilled. In the non-complete case, a similar result holds with the only difference that the lifetime of $X_\bullet(u)$ then possibly depends on u.

Remark 2.4. Even if $L = \Delta/2$, the solution we are looking for is not the flow of a Cameron–Martin vector field: firstly the starting point here is not fixed and secondly the vector field would have to depend on the parameter u. Consequently one cannot apply for instance Theorem 3.2 in [15]. An adaptation of the proof of the cited result would be possible, but we prefer to give a proof using infinitesimal parallel coupling which is more adapted to our situation.

Proof (Proof of Theorem 2.1).

Without loss of generality we may restrict ourselves to the case $u \geq 0$.

A. *Existence.* Under the assumption that a solution $X_t(u)$ exists, we have for any stopping time T,

$$W_{T+t}(X(u))(\dot{\varphi}(u)) = W_t(X_{T+\cdot}(u))\,(\partial X_T(u)),$$

for $t \in [0, \xi(\omega) - T(\omega)[$ and $\omega \in \{T < \xi\}$. Here $\partial X_T := (\partial X)_T$ denotes the derivative process ∂X with respect to u, stopped at the random time T; note that by (5),

$$(\partial X_T)(u) = W(X(u))_T(\dot{\varphi}(u)).$$

Consequently we may localize and replace the time interval $[0, \xi[$ by $[0, \tau \wedge t_0]$ for some $t_0 > 0$, where τ is the first exit time of X from a relatively compact open subset U of M with smooth boundary.

We may also assume that U is sufficiently small and included in the domain of a local chart; moreover we can choose $u_0 \in]0, 1]$ with $\int_0^{u_0} \|\dot{\varphi}(u)\|\,\mathrm{d}u$ small enough such that the processes constructed for $u \in [0, u_0]$ stay in the domain U of the chart. At this point we use the uniform boundedness of W on $[0, \tau \wedge t_0]$.

For $\alpha > 0$, we define by induction a family of processes $(X_t^\alpha(u))_{t \geq 0}$ indexed by $u \geq 0$ as follows: $X^\alpha(0) = X^0$, $X_0^\alpha(u) = \varphi(u)$, and if $u \in]n\alpha, (n+1)\alpha]$ for some integer $n \geq 0$, $X^\alpha(u)$ satisfies the Itô equation

$$\mathrm{d}X_t^\alpha(u) = P_{X_t^\alpha(n\alpha), X_t^\alpha(u)} \mathrm{d}_m X_t^\alpha(n\alpha) + Z_{X_t^\alpha(u)}\,\mathrm{d}t, \tag{7}$$

where $P_{x,y}$ denotes parallel translation along the minimal geodesic from x to y. We choose α sufficiently small so that all the minimizing geodesics are uniquely determined and depend smoothly of the endpoints: since $X^\alpha(u)$ is constructed from $X^\alpha(n\alpha)$ via parallel coupling (7), there exists a constant $C > 0$ such that

$$\rho(X_t^\alpha(u), X_t^\alpha(n\alpha)) \leq \rho(X_0^\alpha(u), X_0^\alpha(n\alpha))\,\mathrm{e}^{Ct} \leq \|\dot{\varphi}\|_\infty\, \alpha \mathrm{e}^{Ct_0} \tag{8}$$

(see e.g. [16]).

The process $\partial X^\alpha(u)$ satisfies the covariant Itô stochastic differential equation

$$\begin{aligned} D\partial X^\alpha(u) = {} & \nabla_{\partial X^\alpha(u)} P_{X^\alpha(n\alpha), \cdot} \mathrm{d}_m X_t^\alpha(n\alpha) \\ & + \nabla_{\partial X^\alpha(u)} Z\,\mathrm{d}t - \frac{1}{2}\,\mathrm{Ric}^\sharp(\partial X^\alpha(u))\,\mathrm{d}t, \end{aligned} \tag{9}$$

(see [3] (4.7), along with Theorem 2.2).

Step 1 We prove that if X and Y are two L-diffusions stopped at $\tau_0 := \tau \wedge t_0$ and living in U, then there exists a constant C such that

$$\mathbb{E}\left[\sup_{t \leq \tau_0} \|W(X)_t - W(Y)_t\|^2\right] \leq C\,\mathbb{E}\left[\sup_{t \leq \tau_0} \|X_t - Y_t\|^2\right]. \tag{10}$$

Here we use the Euclidean norm defined by the chart.

Write

$$L = a^{ij}\partial_{ij} + b^j\partial_j$$

with $a^{ij} = a^{ji}$ for $i, j \in \{1, \dots, \dim M\}$.

For L-diffusions X and Y taking values in U, we denote by N^X, respectively N^Y, their martingale parts in the chart U. Then Itô's formula yields

$$\begin{aligned}
&\left\langle (N^X)^k - (N^Y)^k, (N^X)^k - (N^Y)^k \right\rangle_t \\
&\quad = (X_t^k - Y_t^k)^2 - (X_0^k - Y_0^k)^2 \\
&\qquad - 2\int_0^t (X_s^k - Y_s^k)\,\mathrm{d}((N_s^X)^k - (N_s^Y)^k) \\
&\qquad - 2\int_0^t (X_s^k - Y_s^k)\left(b^k(X_s) - b^k(Y_s)\right)\mathrm{d}s.
\end{aligned}$$

Thus, for U sufficiently small, denoting by $\langle N^X - N^Y | N^X - N^Y \rangle$ the corresponding Riemannian quadratic variation, there exists a constant $C > 0$ (possibly changing from line to line) such that

$$\begin{aligned}
&\mathbb{E}\left[\langle N^X - N^Y | N^X - N^Y \rangle_{\tau_0}\right] \\
&\quad \le C\,\mathbb{E}\left[\sup_{t\le\tau_0} \|X_t - Y_t\|^2\right] + C\sum_k \mathbb{E}\left[\int_0^{\tau_0} |X_t^k - Y_t^k|\,|b^k(X_t) - b^k(Y_t)|\,\mathrm{d}t\right] \\
&\quad \le C\,\mathbb{E}\left[\sup_{t\le\tau_0} \|X_t - Y_t\|^2\right] + C\int_0^{t_0} \mathbb{E}\left[\sup_{s\le\tau_0} \|X_s - Y_s\|^2\right]\mathrm{d}t \\
&\quad \le C(1 + t_0)\,\mathbb{E}\left[\sup_{t\le\tau_0} \|X_t - Y_t\|^2\right].
\end{aligned}$$

Finally, again changing C, we obtain

$$\mathbb{E}\left[\langle N^X - N^Y | N^X - N^Y \rangle_{\tau_0}\right] \le C\,\mathbb{E}\left[\sup_{t\le\tau_0} \|X_t - Y_t\|^2\right]. \tag{11}$$

Writing $W(X) = P(X)\left(P(X)^{-1}W(X)\right)$, a straightforward calculation shows that in the local chart

$$\begin{aligned}
dW(X) = &-\Gamma(X)(\mathrm{d}X, W(X)) \\
&- \frac{1}{2}(d\Gamma)(X)(\mathrm{d}X)(\mathrm{d}X, W(X)) \\
&+ \frac{1}{2}\Gamma(X)(\mathrm{d}X, \Gamma(X)(\mathrm{d}X, W(X))) \\
&- \frac{1}{2}\,\mathrm{Ric}^\sharp(W(X))\,\mathrm{d}t \\
&+ \nabla_{W(X)}Z\,\mathrm{d}t.
\end{aligned} \tag{12}$$

We are going to use (12) to evaluate the difference $W(Y) - W(X)$. Along with the already established bound (11), taking into account that $W(X)$, $W(Y)$ and the derivatives of the brackets of X and Y are bounded in U, we are able to get a bound for

$$F(t) := \mathbb{E}\left[\sup_{s\le t\wedge\tau} \|W(Y) - W(X)\|^2\right].$$

Indeed, first an estimate of the type

$$F(t) \le C_1\,\mathbb{E}\left[\sup_{s\le\tau_0} \|X_s - Y_s\|^2\right] + C_2\int_0^t F(s)\,\mathrm{d}s, \quad 0 \le t \le t_0,$$

is derived which then by Gronwall's lemma leads to

$$F(t) \le C_1\,\mathrm{e}^{C_2 t}\,\mathbb{E}\left[\sup_{t\le\tau_0} \|X_t - Y_t\|^2\right]. \tag{13}$$

Letting $t = t_0$ in (13) we obtain the desired bound (10).

Step 2 We prove that there exists $C > 0$ such that for all $u \in [0, u_0]$,

$$\mathbb{E}\left[\sup_{t\le\tau_0} \rho^2\left(X_t^\alpha(u), X_t^{\alpha'}(u)\right)\right] \le C(\alpha + \alpha')^2. \tag{14}$$

From the covariant equation (9) for $\partial X_t^\alpha(v)$ and the definition of deformed parallel translation (3),

$$DW(X)_t^{-1} = \frac{1}{2}\,\mathrm{Ric}^\sharp(W(X)_t^{-1})\,\mathrm{d}t - \nabla_{W(X)_t^{-1}} Z\,\mathrm{d}t,$$

we have for $(t, v) \in [0, \tau_0] \times [0, u_0]$,

$$W(X^\alpha(v))_t^{-1}\,\partial X_t^\alpha(v) = \dot\varphi(v) + \int_0^t W(X^\alpha(v))_s^{-1}\nabla_{\partial X_s^\alpha(v)} P_{X_s^\alpha(v_\alpha),\bullet}\,\mathrm{d}_m X_s^\alpha(v_\alpha),$$

or equivalently,

$$\begin{aligned}\partial X_t^\alpha(v) &= W(X^\alpha(v))_t\,\dot\varphi(v)\\ &\quad + W(X^\alpha(v))_t\int_0^t W(X^\alpha(v))_s^{-1}\,\nabla_{\partial X_s^\alpha(v)} P_{X_s^\alpha(v_\alpha),\bullet}\,\mathrm{d}_m X_s^\alpha(v_\alpha)\end{aligned} \tag{15}$$

with $v_\alpha = n\alpha$, where the integer n is determined by $n\alpha < v \le (n+1)\alpha$. Consequently, we obtain

$$
\begin{aligned}
&\rho(X_t^{\alpha}(u), X_t^{\alpha'}(u)) \\
&= \int_0^u \left\langle \mathrm{d}\rho, \left(\partial X_t^{\alpha}(v), \partial X_t^{\alpha'}(v)\right)\right\rangle \mathrm{d}v \\
&= \int_0^u \left\langle \mathrm{d}\rho, \left(W(X^{\alpha}(v))_t \dot{\varphi}(v), W(X^{\alpha'}(v))_t \dot{\varphi}(v)\right)\right\rangle \mathrm{d}v \\
&\quad + \int_0^u \left\langle \mathrm{d}\rho, \left(W(X^{\alpha}(v))_t \int_0^t W(X^{\alpha}(v))_s^{-1} \nabla_{\partial X_s^{\alpha}(v)} P_{X_s^{\alpha}(v_\alpha),\bullet} \mathrm{d}_m X_s^{\alpha}(v_\alpha), 0\right)\right\rangle \mathrm{d}v \\
&\quad + \int_0^u \left\langle \mathrm{d}\rho, \left(0, W(X^{\alpha'}(v))_t \int_0^t W(X^{\alpha'}(v))_s^{-1} \nabla_{\partial X_s^{\alpha'}(v)} P_{X_s^{\alpha'}(v_{\alpha'}),\bullet} \mathrm{d}_m X_s^{\alpha'}(v_{a'})\right)\right\rangle \mathrm{d}v.
\end{aligned}
$$

This yields, by means of boundedness of $\mathrm{d}\rho$ and deformed parallel translation, together with (13) and the Burkholder–Davis–Gundy inequalities,

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\le\tau_0} \rho^2\left(X_t^{\alpha}(u), X_t^{\alpha'}(u)\right)\right] &\le C\int_0^u \mathbb{E}\left[\sup_{t\le\tau_0} \rho^2\left(X_t^{\alpha}(v), X_t^{\alpha'}(v)\right)\right]\mathrm{d}v \\
&\quad + C\int_0^u \mathbb{E}\left[\int_0^{\tau_0} \left\|\nabla_{\partial X_s^{\alpha}(v)} P_{X_s^{\alpha}(v_\alpha),\bullet}\right\|^2 \mathrm{d}s\right]\mathrm{d}v \\
&\quad + C\int_0^u \mathbb{E}\left[\int_0^{\tau_0} \left\|\nabla_{\partial X_s^{\alpha'}(v)} P_{X_s^{\alpha'}(v_{\alpha'}),\bullet}\right\|^2 \mathrm{d}s\right]\mathrm{d}v.
\end{aligned}
$$

From here we obtain

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\le\tau_0} \rho^2\left(X_t^{\alpha}(u), X_t^{\alpha'}(u)\right)\right] &\le C\int_0^u \mathbb{E}\left[\sup_{t\le\tau_0} \rho^2\left(X_t^{\alpha}(v), X_t^{\alpha'}(v)\right)\right]\mathrm{d}v \\
&\quad + C\alpha^2\int_0^u \mathbb{E}\left[\int_0^{\tau_0} \left\|\partial X_s^{\alpha}(v)\right\|^2 \mathrm{d}s\right]\mathrm{d}v \\
&\quad + C\alpha'^2\int_0^u \mathbb{E}\left[\int_0^{\tau_0} \left\|\partial X_s^{\alpha'}(v)\right\|^2 \mathrm{d}s\right]\mathrm{d}v,
\end{aligned}
$$

where we used the fact that for $v \in T_xM$, $\nabla_v P_{x,\bullet} = 0$, together with

$$
\rho(X_s^{\beta}(v), X_s^{\beta}(v_\beta)) \le C\beta, \quad \beta = \alpha, \alpha',
$$

see estimate (8).

Now, by (9) for $D\partial X^{\beta}$, there exists a constant $C' > 0$ such that for all $v \in [0, u_0]$,

$$
\mathbb{E}\left[\int_0^{\tau_0} \left\|\partial X_s^{\beta}(v)\right\|^2 \mathrm{d}s\right] < C'.
$$

Consequently,

$$\mathbb{E}\left[\sup_{t\le\tau_0}\rho^2\left(X_t^\alpha(u),X_t^{\alpha'}(u)\right)\right]\le C\int_0^u\mathbb{E}\left[\sup_{t\le\tau_0}\rho^2\left(X_t^\alpha(v),X_t^{\alpha'}(v)\right)\right]\mathrm{d}v + 2CC'(\alpha+\alpha')^2,$$

which by Gronwall lemma yields

$$\mathbb{E}\left[\sup_{t\le\tau_0}\rho^2\left(X_t^\alpha(u),X_t^{\alpha'}(u)\right)\right]\le C\,(\alpha+\alpha')^2$$

for some constant $C>0$. This is the desired inequality.

Step 3 Recall that

$$L=a^{ij}\partial_{ij}+b^j\partial_j.$$

Denoting by (a_{ij}) the inverse of (a^{ij}), we let ∇' be the connection with Christoffel symbols

$$(\Gamma')_{ij}^k=-\frac{1}{2}(a_{ik}+a_{jk})b^k. \tag{16}$$

We are going to prove that all L-diffusions are ∇'-martingales:

(i) On one hand, ∇'-martingales are characterized by the fact that for any k,

$$\mathrm{d}X^k+\frac{1}{2}(\Gamma')_{ij}^k\,\mathrm{d}\langle X^i,X^j\rangle \quad \text{is the differential of a local martingale.} \tag{17}$$

(ii) On the other hand, L-diffusions satisfy the following two conditions:

$$\mathrm{d}X^k-b^k(X)\,\mathrm{d}t \quad \text{is the differential of a local martingale,} \tag{18}$$

and

$$\mathrm{d}\langle X^i,X^j\rangle=(a^{ij}(X)+a^{ji}(X))\,\mathrm{d}t. \tag{19}$$

From this it is clear that (16), (18) together with (19) imply (17).

From inequality (14) we deduce that there exists a limiting process

$$(X_t(u))_{0\le t\le\tau_0,\ 0\le u\le u_0}$$

such that for all $u\in[0,u_0]$ and $\alpha>0$,

$$\mathbb{E}\left[\sup_{t\le\tau_0}\rho^2\left(X_t^\alpha(u),X_t(u)\right)\right]\le C\alpha^2. \tag{20}$$

In other words, for any fixed $u\in[0,u_0]$, the process $(X_t^\alpha(u))_{t\in[0,\tau_0]}$ converges to $(X_t(u))_{t\in[0,\tau_0]}$ uniformly in L^2 as α tends to 0. Since these processes are ∇'-martingales, convergence also holds in the topology of semimartingales ([4], Proposition 2.10). This implies in particular that for any $u\in[0,u_0]$, the process $(X_t(u))_{t\in[0,\tau_0]}$ is a diffusion with generator L, stopped at time τ_0.

Extracting a subsequence $(\alpha_k)_{k\geq 0}$ convergent to 0, we may assume that almost surely, for all dyadic $u \in [0, u_0]$,

$$\sup_{t\leq\tau_0} \rho\left(X_t^{\alpha}(u), X_t(u)\right)$$

converges to 0. Moreover we can choose $(\alpha_k)_{k\geq 0}$ of the form $\alpha_k = 2^{-n_k}$ with $(n_k)_{k\geq 0}$ an increasing sequence of positive integers. Due to (8), we can take a version of the processes $(t, u) \mapsto X_t^{\alpha_k}(u)$ such that

$$u \mapsto X_t^{\alpha_k}(u)$$

is uniformly Lipschitz in $u \in \mathbb{N}\alpha_k \cap [0, u_0]$ with a Lipschitz constant independent of k and t. Passing to the limit, we obtain that a.s for any $t \in [0, \tau_0]$, the map

$$u \mapsto X_t(u)$$

is uniformly Lipschitz in $u \in \mathscr{D} \cap [0, u_0]$ with a Lipschitz constant independent of t, where $\mathscr{D}$ is the set of dyadic numbers. Finally we can choose a version of

$$(t, u) \mapsto X_t(u)$$

which is a.s. continuous in $(t, u) \in [0, \tau_0] \times [0, u_0]$, and hence uniformly Lipschitz in $u \in [0, u_0]$.

Step 4 We prove that almost surely, $X_t(u)$ is differentiable in u with derivative

$$W(X(u))_t(\dot{\varphi}(u)).$$

More precisely, we show that in local coordinates, almost surely, for all $t \in [0, \tau_0]$, $u \in [0, u_0]$,

$$X_t(u) = X_t^0 + \int_0^u W(X(v))_t(\dot{\varphi}(v))\,\mathrm{d}v. \tag{21}$$

From the construction it is clear that almost surely, for all $t \in [0, \tau_0]$, $u \in [0, u_0]$,

$$\begin{aligned} X_t^{\alpha_k}(u) &= X_t^0 + \int_0^u W(X^{\alpha_k}(v))_t(\dot{\varphi}(v))\,\mathrm{d}v \\ &\quad + \int_0^u \left(W(X^{\alpha_k}(v))_t \int_0^t W(X^{\alpha_k}(v))_s^{-1} \nabla_{\partial X_s^{\alpha_k}(v)} P_{X_s^{\alpha_k}(v_{\alpha_k}),\bullet}\, \mathrm{d}_m X_s^{\alpha_k}(v_{\alpha_k}) \right) \mathrm{d}v. \end{aligned}$$

This yields

$$\begin{aligned} &X_t(u) - X_t^0 - \int_0^u W(X(v))_t(\dot{\varphi}(v))\,\mathrm{d}v \\ &\quad = X_t(u) - X_t^{\alpha_k}(u) + \int_0^u (W(X^{\alpha_k}(v))_t - W(X(v))_t)\,\dot{\varphi}(v)\,\mathrm{d}v \\ &\qquad + \int_0^u \left(W(X^{\alpha_k}(v))_t \int_0^t W(X^{\alpha_k}(v))_s^{-1} \nabla_{\partial X_s^{\alpha_k}(v)} P_{X_s^{\alpha_k}(v_{\alpha_k}),\bullet}\, \mathrm{d}_m X_s^{\alpha_k}(v_{\alpha_k}) \right) \mathrm{d}v. \end{aligned}$$

The terms of right-hand-side are easily estimated, where in the estimates the constant C may change from one line to another. First observe that

$$\mathbb{E}\left[\sup_{t\le\tau_0}\left\|X_t(u)-X_t^{\alpha_k}(u)\right\|^2\right]\le C\alpha_k^2.$$

Using (10) and (20) we have

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\le\tau_0}\left\|\int_0^u (W(X^{\alpha_k}(v))_t - W(X(v))_t)\,\mathrm{d}v\right\|^2\right]\\
&\quad\le \mathbb{E}\left[\sup_{t\le\tau_0}\int_0^u \|W(X^{\alpha_k}(v))_t - W(X(v))_t\|^2\,\mathrm{d}v\right]\\
&\quad= \int_0^u \mathbb{E}\left[\sup_{t\le\tau_0}\|W(X^{\alpha_k}(v))_t - W(X(v))_t\|^2\right]\mathrm{d}v\\
&\quad\le C\alpha_k^2,
\end{aligned}$$

and finally

$$\begin{aligned}
&\mathbb{E}\left[\sup_{t\le\tau_0}\left\|\int_0^u W(X^{\alpha_k}(v))_t\right.\right.\\
&\qquad\left.\left.\times\left(\int_0^t W(X^{\alpha_k}(v))_s^{-1}\nabla_{\partial X_s^{\alpha_k}(v)}P_{X_s^{\alpha_k}(v_{\alpha_k}),\bullet}\,\mathrm{d}_m X_s^{\alpha_k}(v_{\alpha_k})\right)\mathrm{d}v\right\|^2\right]\\
&\le C\int_0^u \mathbb{E}\left[\sup_{t\le\tau_0}\left\|\int_0^t W(X^{\alpha_k}(v))_s^{-1}\nabla_{\partial X_s^{\alpha_k}(v)}P_{X_s^{\alpha_k}(v_{\alpha_k}),\bullet}\,\mathrm{d}_m X_s^{\alpha_k}(v_{\alpha_k})\right\|^2\right]\mathrm{d}v\\
&\le C\int_0^u \mathbb{E}\left[\int_0^{\tau_0}\left\|\nabla_{\partial X_s^{\alpha_k}(v)}P_{X_s^{\alpha_k}(v_{\alpha_k}),\bullet}\right\|^2\mathrm{d}s\right]\mathrm{d}v\quad\text{(since } W^{-1}\text{ is bounded)}\\
&\le C\alpha_k^2\int_0^u \mathbb{E}\left[\int_0^{\tau_0}\|\partial X_s^{\alpha_k}(v)\|^2\,\mathrm{d}s\right]\mathrm{d}v\\
&\le C\alpha_k^2.
\end{aligned}$$

where in the last but one inequality we used $\nabla_v P_{x,\bullet}=0$ for any $v\in T_xM$ which implies

$$\left\|\nabla_v P_{y,\bullet}\right\|^2\le C\,\rho(x,y)^2\|v\|^2,$$

and the last inequality is a consequence of (9).

We deduce that

$$\mathbb{E}\left[\sup_{t\le\tau_0}\left\|X_t(u)-X_t^0-\int_0^u W(X(v))_t(\dot\varphi(v))\,\mathrm{d}v\right\|^2\right]\le C\alpha_k^2.$$

Since this is true for any α_k, using continuity in u of $X_t(u)$, we finally get almost surely for all t, u,

$$X_t(u) = X_t^0 + \int_0^u W(X(v))_t(\dot{\varphi}(v))\,\mathrm{d}v.$$

Step 5 Finally we are able to prove (6):

$$\mathrm{d}X_t(u) = P_{0,u}^{X_t(\cdot)}\,\mathrm{d}_m X_t^0 + Z_{X_t(u)}\,\mathrm{d}t.$$

Since a.s. the mapping $(t, u) \mapsto \partial X_t(u)$ is continuous, the map $u \mapsto \partial X(u)$ is continuous in the topology of uniform convergence in probability. We want to prove that $u \mapsto \partial X(u)$ is continuous in the topology of semimartingales.

Since for a given connection on a manifold, the topology of uniform convergence in probability and the topology of semimartingale coincide on the set of martingales (Proposition 2.10 of [4]), it is sufficient to find a connection on TM for which $\partial X(u)$ is a martingale for any u. Again we can localize in the domain of a chart. Recall that for all u, the process $X(u)$ is a ∇'-martingale where ∇' is defined in step 1. Then by [1], Theorem 3.3, this implies that the derivative with respect to u with values in TM, denoted here by $\partial X(u)$, is a $(\nabla')^c$-martingale with respect to the complete lift $(\nabla')^c$ of ∇'. This proves that $u \mapsto \partial X(u)$ is continuous in the topology of semimartingales.

Remark 2.5. Alternatively, one could have used that given a generator L', the topologies of uniform convergence in probability on compact sets and the topology of semimartingales coincide on the space of L'-diffusions. Since the processes $\partial X(u)$ are diffusions with the same generator, the result could be derived as well.

As a consequence, Itô integrals commute with derivatives with respect to u (see e.g. [4], Corollary 3.18 and Lemma 3.15). We write it formally as

$$D\partial X = \nabla_u \mathrm{d}X - \frac{1}{2} R(\partial X, \mathrm{d}X)\mathrm{d}X. \tag{22}$$

Since

$$\mathrm{d}X(u) \otimes \mathrm{d}X(u) = g^{-1}(X(u))\,\mathrm{d}t$$

where g is the metric tensor, (22) becomes

$$D\partial X = \nabla_u \mathrm{d}X - \frac{1}{2}\,\mathrm{Ric}^\sharp(\partial X)\,\mathrm{d}t.$$

On the other hand, (5) and (3) for W yield

$$D\partial X = -\frac{1}{2}\,\mathrm{Ric}^\sharp(\partial X)\,\mathrm{d}t + \nabla_{\partial X} Z\,\mathrm{d}t.$$

From the last two equations we obtain

$$\nabla_u \mathrm{d}X = \nabla_{\partial X} Z\,\mathrm{d}t.$$

This along with the original equation

$$\mathrm{d}X^0 = \mathrm{d}_m X^0 + Z_{X^0}\,\mathrm{d}t$$

gives

$$\mathrm{d}X_t(u) = P_{0,u}^{X_t(\cdot)}\,\mathrm{d}_m X_t^0 + Z_{X_t(u)}\,\mathrm{d}t,$$

where

$$P_{0,u}^{X_t(\cdot)}\colon T_{X_t}M \to T_{X_t(u)}M$$

denotes parallel transport along the C^1 curve $v \mapsto X_t(v)$.

B. *Uniqueness.* Again we may localize in the domain of a chart U. Letting $X(u)$ and $Y(u)$ be two solutions of (5), then for $(t,u) \in [0,\tau_0[\times[0,u_0]$ we find in local coordinates,

$$Y_t(u) - X_t(u) = \int_0^u \big(W(Y(v))_t - W(X(v))_t\big)(\dot{\varphi}(v))\,\mathrm{d}v. \tag{23}$$

On the other hand, using (10) we have

$$\mathbb{E}\left[\sup_{t\le\tau_0}\|Y_t(u) - X_t(u)\|^2\right] \le C\int_0^u \mathbb{E}\left[\sup_{t\le\tau_0}\|Y_t(v) - X_t(v)\|^2\right]\mathrm{d}v \tag{24}$$

from which we deduce that almost surely, for all $t \in [0,\tau_0]$, $X_t(u) = Y_t(u)$. Consequently, exploiting the fact that the two processes are continuous in (t,u), they must be indistinguishable. □

3 Horizontal Diffusion Along Non-Homogeneous Diffusion

In this section we assume that the elliptic generator is a C^1 function of time: $L = L(t)$ for $t \ge 0$. Let $g(t)$ be the metric on M such that

$$L(t) = \frac{1}{2}\Delta^t + Z(t)$$

where Δ^t is the $g(t)$-Laplacian and $Z(t)$ a vector field on M.

Let (X_t) be an inhomogeneous diffusion with generator $L(t)$. Parallel transport $P^t(X)_t$ along the $L(t)$-diffusion X_t is defined analogously to [2] as the linear map

$$P^t(X)_t\colon T_{X_0}M \to T_{X_t}M$$

which satisfies

$$D^t P^t(X)_t = -\frac{1}{2}\,\dot{g}^\sharp(P^t(X)_t)\,\mathrm{d}t \tag{25}$$

where $\dot{g}$ denotes the derivative of g with respect to time; the covariant differential D^t is defined in local coordinates by the same formulas as D, with the only difference that Christoffel symbols now depend on t.

Alternatively, if J is a semimartingale over X, the covariant differential $D^t J$ may be defined as $\tilde{D}(0, J) = (0, D^t J)$, where $(0, J)$ is a semimartingale along (t, X_t) in $\tilde{M} = [0, T] \times M$ endowed with the connection $\tilde{\nabla}$ defined as follows: if

$$s \mapsto \tilde{\varphi}(s) = (f(s), \varphi(s))$$

is a C^1 path in $\tilde{M}$ and $s \mapsto \tilde{u}(s) = (\alpha(s), u(s)) \in T\tilde{M}$ is C^1 path over $\tilde{\varphi}$, then

$$\tilde{\nabla}\tilde{u}(s) = \left(\dot{\alpha}(s), \left(\nabla^{f(s)} u\right)(s)\right)$$

where ∇^t denotes the Levi–Civita connection associated to $g(t)$. It is proven in [2] that $P^t(X)_t$ is an isometry from $(T_{X_0}M, g(0, X_0))$ to $(T_{X_t}M, g(t, X_t))$.

The damped parallel translation $W^t(X)_t$ along X_t is the linear map

$$W^t(X)_t : T_{X_0}M \to T_{X_t}M$$

satisfying

$$D^t W^t(X)_t = \left(\nabla^t_{W^t(X)_t} Z(t, \cdot) - \frac{1}{2}(\mathrm{Ric}^t)^\sharp(W^t(X)_t)\right) \mathrm{d}t. \tag{26}$$

If $Z \equiv 0$ and $g(t)$ is solution to the backward Ricci flow:

$$\dot{g} = \mathrm{Ric}, \tag{27}$$

then damped parallel translation coincides with the usual parallel translation:

$$P^t(X) = W^t(X),$$

(see [2], Theorem 2.3).

The Itô differential $\mathrm{d}^\nabla Y = \mathrm{d}^{\nabla^t} Y$ of an M-valued semimartingale Y is defined by formula (2), with the only difference that the Christoffel symbols depend on time.

Theorem 3.1. *Keeping the assumptions of this section, let*

$$\mathbb{R} \to M, \quad u \mapsto \varphi(u),$$

be a C^1 path in M and let X^0 be an $L(t)$-diffusion with starting point $\varphi(0)$ and lifetime ξ. Assume that $(M, g(t))$ is complete for every t. There exists a unique family

$$u \mapsto (X_t(u))_{t \in [0, \xi[}$$

of $L(t)$-diffusions, which is a.s. continuous in (t, u) and C^1 in u, satisfying

$$X(0) = X^0 \text{ and } X_0(u) = \varphi(u),$$

and solving the equation

$$\partial X_t(u) = W^t(X(u))_t(\dot{\varphi}(u)). \tag{28}$$

Furthermore, $X(u)$ solves the Itô stochastic differential equation

$$\mathrm{d}^{\nabla} X_t(u) = P_{0,u}^{t,X_t(\cdot)}\, \mathrm{d}^{\nabla(t)} X_t + Z(t, X_t(u))\, \mathrm{d}t, \tag{29}$$

where

$$P_{0,u}^{t,X_t(\cdot)} : T_{X_t^0} M \to T_{X_t(u)} M$$

denotes parallel transport along the C^1 curve

$$[0, u] \to M, \quad v \mapsto X_t(v),$$

with respect to the metric $g(t)$.

If $Z \equiv 0$ and if $g(t)$ is given as solution to the backward Ricci flow equation, then almost surely for all t,

$$\|\partial X_t(u)\|_{g(t)} = \|\dot{\varphi}(u)\|_{g(0)}. \tag{30}$$

Definition 3.2. We call

$$t \mapsto (X_t(u))_{u \in \mathbb{R}}$$

the horizontal $L(t)$-diffusion in C^1 path space $C^1(\mathbb{R}, M)$ over X^0, started at φ.

Remark 3.3. Equation (30) says that if $Z \equiv 0$ and if g is solution to the backward Ricci flow equation, then the horizontal $g(t)$-Brownian motion is length preserving (with respect to the moving metric).

Remark 3.4. Again if the manifold $(M, g(t))$ is not necessarily complete for all t, a similar result holds with the lifetime of $X_\cdot(u)$ possibly depending on u.

Proof (Proof of Theorem 3.1). The proof is similar to the one of Theorem 2.1. We restrict ourselves to explaining the differences.

The localization procedure carries over immediately; we work on the time interval $[0, \tau \wedge t_0]$.

For $\alpha > 0$, we define the approximating process $X_t^\alpha(u)$ by induction as

$$X_t^\alpha(0) = X_t^0, \quad X_0^\alpha(u) = \varphi(u),$$

and if $u \in]n\alpha, (n+1)\alpha]$ for some integer $n \geq 0$, then $X^\alpha(u)$ solves the Itô equation

$$\mathrm{d}^{\nabla} X_t^\alpha(u) = P^t_{X_t^\alpha(n\alpha), X_t^\alpha(u)} \mathrm{d}_m X_t^\alpha(n\alpha) + Z(t, X_t(u))\, \mathrm{d}t \tag{31}$$

where $P^t_{x,y}$ is the parallel transport along the minimal geodesic from x to y, for the connection ∇^t.

Alternatively, letting $\tilde{X}_t^\alpha = (t, X_t^\alpha)$, we may write (31) as

$$\mathrm{d}^{\tilde{\nabla}} \tilde{X}_t^\alpha(u) = \tilde{P}_{\tilde{X}_t^\alpha(n\alpha), \tilde{X}_t^\alpha(u)} \mathrm{d}_m \tilde{X}_t^\alpha(n\alpha) + Z(\tilde{X}_t^\alpha(u))\,\mathrm{d}t \tag{32}$$

where $\tilde{P}_{\tilde{x},\tilde{y}}$ denotes parallel translation along the minimal geodesic from $\tilde{x}$ to $\tilde{y}$ for the connection $\tilde{\nabla}$.

Denoting by $\rho(t, x, y)$ the distance from x to y with respect to the metric $g(t)$, Itô's formula shows that the process $\rho\big(t, X_t^\alpha(u), X_t^\alpha(n\alpha)\big)$ has locally bounded variation. Moreover since locally $\partial_t \rho(t, x, y) \le C\rho(t, x, y)$ for $x \neq y$, we find similarly to (8),

$$\rho(t, X_t^\alpha(u), X_t^\alpha(n\alpha)) \le \rho(0, X_0^\alpha(u), X_0^\alpha(n\alpha))\,\mathrm{e}^{Ct} \le \|\dot{\varphi}\|_\infty\, \alpha\, \mathrm{e}^{Ct_0}.$$

Since all Riemannian distances are locally equivalent, this implies

$$\rho(X_t^\alpha(u), X_t^\alpha(n\alpha)) \le \rho(X_0^\alpha(u), X_0^\alpha(n\alpha))\,\mathrm{e}^{Ct} \le \|\dot{\varphi}\|_\infty\, \alpha\, \mathrm{e}^{Ct_0} \tag{33}$$

where $\rho = \rho(0, \cdot, \cdot)$.

Next, differentiating (32) yields

$$\begin{aligned}\tilde{D}\partial_u \tilde{X}_t^\alpha(u) = {} & \tilde{\nabla}_{\partial_u \tilde{X}_t^\alpha(u)} \tilde{P}_{\tilde{X}_t^\alpha(n\alpha), \cdot} \mathrm{d}_m \tilde{X}_t^\alpha(n\alpha) \\ & + \tilde{\nabla}_{\partial_u \tilde{X}_t^\alpha(u)} Z\,\mathrm{d}t - \frac{1}{2}\tilde{R}\big(\partial_u \tilde{X}_t^\alpha(u), d\,\tilde{X}_t^\alpha(u)\big) d\,\tilde{X}_t^\alpha(u).\end{aligned}$$

Using the fact that the first component of $\tilde{X}_t^\alpha(u)$ has finite variation, a careful computation of $\tilde{R}$ leads to the equation

$$\begin{aligned}D^t \partial_u X_t^\alpha(u) = {} & \nabla^t_{\partial_u X_t^\alpha(u)} P^t_{X_t^\alpha(n\alpha), \cdot} \mathrm{d}_m X_t^\alpha(n\alpha) \\ & + \nabla^t_{\partial_u X_t^\alpha(u)} Z(t, \cdot) - \frac{1}{2}(\mathrm{Ric}^t)^\sharp\big(\partial_u X_t^\alpha(u)\big)\,\mathrm{d}t.\end{aligned}$$

To finish the proof, it is sufficient to remark that in step 1, (11) still holds true for X and Y $g(t)$-Brownian motions living in a small open set U, and that in step 5, the map $u \mapsto \partial X(u)$ is continuous in the topology of semimartingales. This last point is due to the fact that all $\partial X(u)$ are inhomogeneous diffusions with the same generator, say L', and the fact that the topology of uniform convergence on compact sets and the topology of semimartingales coincide on L'-diffusions. □

4 Application to Optimal Transport

In this section we assume again that the elliptic generator $L(t)$ is a C^1 function of time with associated metric $g(t)$:

$$L(t) = \frac{1}{2}\Delta^t + Z(t), \quad t \in [0, T],$$

where Δ^t is the Laplacian associated to $g(t)$ and $Z(t)$ is a vector field. We assume further that for any t, the Riemannian manifold $(M, g(t))$ is metrically complete, and $L(t)$ diffusions have lifetime T.

Letting $\varphi\colon [0,T] \to \mathbb{R}_+$ be a non-decreasing function, we define a cost function

$$c(t,x,y) = \varphi(\rho(t,x,y)) \tag{34}$$

where $\rho(t,\cdot,\cdot)$ denotes distance with respect to $g(t)$.

To the cost function c we associate the Monge–Kantorovich minimization between two probability measures on M

$$\mathscr{W}_{c,t}(\mu,\nu) = \inf_{\eta\in\Pi(\mu,\nu)} \int_{M\times M} c(t,x,y)\,\mathrm{d}\eta(x,y) \tag{35}$$

where $\Pi(\mu,\nu)$ is the set of all probability measures on $M\times M$ with marginals μ and ν. We denote

$$\mathscr{W}_{p,t}(\mu,\nu) = \left(\mathscr{W}_{\rho^p,t}(\mu,\nu)\right)^{1/p} \tag{36}$$

the Wasserstein distance associated to $p>0$. For a probability measure μ on M, the solution of the heat flow equation associated to $L(t)$ will be denoted by μP_t.

Define a section $(\nabla^t Z)^\flat \in \Gamma(T^*M \odot T^*M)$ as follows: for any $x\in M$ and $u,v\in T_xM$,

$$(\nabla^t Z)^\flat(u,v) = \frac{1}{2}\left(g(t)(\nabla^t_u Z, v) + g(t)(u, \nabla^t_v Z)\right).$$

In case the metric is independent of t and $Z = \operatorname{grad} V$ for some C^2 function V on M, then

$$(\nabla^t Z)^\flat(u,v) = \nabla dV(u,v).$$

Theorem 4.1. *We keep notation and assumptions from above.*

(a) *Assume*

$$\mathrm{Ric}^t - \dot g - 2(\nabla^t Z)^\flat \geq 0, \quad t\in[0,T]. \tag{37}$$

Then the function

$$t \mapsto \mathscr{W}_{c,t}(\mu P_t, \nu P_t)$$

is non-increasing on $[0,T]$.

(b) *If for some* $k\in\mathbb{R}$,

$$\mathrm{Ric}^t - \dot g - 2(\nabla^t Z)^\flat \geq kg, \quad t\in[0,T], \tag{38}$$

then we have for all $p>0$

$$\mathscr{W}_{p,t}(\mu P_t, \nu P_t) \leq \mathrm{e}^{-kt/2}\,\mathscr{W}_{p,0}(\mu,\nu), \quad t\in[0,T].$$

Remark 4.2. Before turning to the proof of Theorem 4.1, let us mention that in the case $Z = 0$, g constant, $p = 2$ and $k = 0$, item (b) is due to [21] and [20]. In the case where g is a backward Ricci flow solution, $Z = 0$ and $p = 2$, statement (b) for M compact is due to Lott [18] and McCann–Topping [19]. For extensions about $\mathcal{L}$-transportation, see [23].

Proof (Proof of Theorem 4.1). (a) Assume that $\mathrm{Ric}^t - \dot{g} - 2(\nabla^t Z)^\flat \geq 0$. Then for any $L(t)$-diffusion (X_t), we have

$$\begin{aligned}
&d\big(g(t)(W(X)_t, W(X)_t)\big) \\
&\quad = \dot{g}(t)\big(W(X)_t, W(X)_t\big)\,\mathrm{d}t + 2g(t)\left(D^t W(X)_t, W(X)_t\right) \\
&\quad = \dot{g}(t)\big(W(X)_t, W(X)_t\big)\,\mathrm{d}t \\
&\qquad + 2g(t)\left(\nabla^t_{W(X)_t} Z(t,\cdot) - \frac{1}{2}(\mathrm{Ric}^t)^\sharp(W(X)_t), W(X)_t\right)\mathrm{d}t \\
&\quad = \left(\dot{g} + 2(\nabla^t Z)^\flat - \mathrm{Ric}^t\right)\left(W(X)_t, W(X)_t\right)\mathrm{d}t \leq 0.
\end{aligned}$$

Consequently, for any $t \geq 0$,

$$\|W(X)_t\|_t \leq \|W(X)_0\|_0 = 1. \tag{39}$$

For $x, y \in M$, let $u \mapsto \gamma(x,y)(u)$ be a minimal $g(0)$-geodesic from x to y in time 1: $\gamma(x,y)(0) = x$ and $\gamma(x,y)(1) = y$. Denote by $X^{x,y}(u)$ a horizontal $L(t)$-diffusion with initial condition $\gamma(x,y)$.

For $\eta \in \Pi(\mu,\nu)$, define the measure η_t on $M \times M$ by

$$\eta_t(A \times B) = \int_{M\times M} \mathbb{P}\big\{X_t^{x,y}(0) \in A,\ X_t^{x,y}(1) \in B\big\}\,\mathrm{d}\eta(x,y),$$

where A and B are Borel subsets of M. Then η_t has marginals μP_t and νP_t. Consequently it is sufficient to prove that for any such η,

$$\int_{M\times M} \mathbb{E}\big[c(t, X_t^{x,y}(0), X_t^{x,y}(1))\big]\,\mathrm{d}\eta(x,y) \leq \int_{M\times M} c(0,x,y)\,\mathrm{d}\eta(x,y). \tag{40}$$

On the other hand, we have a.s.,

$$\begin{aligned}
\rho(t, X_t^{x,y}(0), X_t^{x,y}(1)) &\leq \int_0^1 \left\|\partial_u X_t^{x,y}(u)\right\|_t \mathrm{d}u \\
&= \int_0^1 \left\|W(X^{x,y}(u))_t\,\dot{\gamma}(x,y)(u)\right\|_t \mathrm{d}u \\
&\leq \int_0^1 \left\|\dot{\gamma}(x,y)(u)\right\|_0 \mathrm{d}u \\
&= \rho(0,x,y),
\end{aligned}$$

and this clearly implies

$$c\big(t, X_t^{x,y}(0), X_t^{x,y}(1)\big) \le c(0,x,y) \quad \text{a.s.},$$

and then (40).

(b) Under condition (38), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\, g(t)\big(W(X)_t, W(X)_t\big) \le -k\, g(t)\big(W(X)_t, W(X)_t\big),$$

which implies

$$\|W(X)_t\|_t \le \mathrm{e}^{-kt/2},$$

and then

$$\rho\big(t, X_t^{x,y}(0), X_t^{x,y}(1)\big) \le \mathrm{e}^{-kt/2}\rho(0,x,y).$$

The result follows. □

5 Derivative Process Along Constant Rank Diffusion

In this section we consider a generator L of constant rank: the image E of the "carré du champ" operator $\Gamma(L) \in \Gamma(TM \otimes TM)$ defines a subbundle of TM. In E we then have an intrinsic metric given by

$$g(x) = (\Gamma(L)|E(x))^{-1}, \quad x \in M.$$

Let ∇ be a connection on E with preserves g, and denote by ∇' the associated semi-connection: if $U \in \Gamma(TM)$ is a vector field, $\nabla'_v U$ is defined only if $v \in E$ and satisfies

$$\nabla'_v U = \nabla_{U_{x_0}} V + [V,U]_{x_0}$$

where $V \in \Gamma(E)$ is such that $V_{x_0} = v$ (see [11], Sect. 1.3). We denote by $Z(x)$ the drift of L with respect to the connection ∇.

For the construction of a flow of L-diffusions we will use an extension of ∇ to TM denoted by $\tilde{\nabla}$. Then the associated semi-connection ∇' is the restriction of the classical adjoint of $\tilde{\nabla}$ (see [11], Proposition 1.3.1).

Remark 5.1. It is proven in [11] that a connection ∇ always exists, for instance, we may take the Le Jan-Watanabe connection associated to a well chosen vector bundle homomorphism from a trivial bundle $M \times H$ to E where H is a Hilbert space.

If X_t is an L-diffusion, the parallel transport

$$P(X)_t : E_{X_0} \to E_{X_t}$$

along X_t (with respect to the connection $\tilde{\nabla}$) depends only on ∇. The same applies for the Itô differential $\mathrm{d}X_t = \mathrm{d}^{\nabla} X_t$. We still denote by $\mathrm{d}_m X_t$ its martingale part.

We denote by

$$\tilde{P}'(X)_t : T_{X_0}M \to T_{X_t}M$$

the parallel transport along X_t for the adjoint connection $(\tilde{\nabla})'$, and by $\tilde{D}'J$ the covariant differential (with respect to $(\tilde{\nabla})'$) of a semimartingale $J \in TM$ above X; compare (4) for the definition.

Theorem 5.2. *We keep the notation and assumptions from above. Let x_0 be a fixed point in M and $X_t(x_0)$ an L-diffusion starting at x_0. For $x \in M$ close to x_0, we define the L-diffusion $X_t(x)$, started at x, by*

$$\mathrm{d}X_t(x) = \tilde{P}_{X_t(x_0),X_t(x)}\, \mathrm{d}_m X_t(x_0) + Z(X_t(x))\, \mathrm{d}t \tag{41}$$

where $\tilde{P}_{x,y}$ denotes parallel transport (with respect to $\tilde{\nabla}$) along the unique $\tilde{\nabla}$-geodesic from x to y. Then

$$\tilde{D}'T_{x_0}X = \tilde{\nabla}_{T_{x_0}X} Z\, \mathrm{d}t - \frac{1}{2}\operatorname{Ric}^{\sharp}(T_{x_0}X)\, \mathrm{d}t \tag{42}$$

where

$$\operatorname{Ric}^{\sharp}(u) = \sum_{i=1}^{d} \tilde{R}(u, e_i)e_i, \quad u \in T_xM,$$

and $(e_i)_{i=1,\dots,d}$ an orthonormal basis of E_x for the metric g.

Under the additional assumption that $Z \in \Gamma(E)$, the differential $\tilde{D}'T_{x_0}X$ does not depend on the extension $\tilde{\nabla}$, and we have

$$\tilde{D}'T_{x_0}X = \nabla_{T_{x_0}X} Z\, \mathrm{d}t - \frac{1}{2}\operatorname{Ric}^{\sharp}(T_{x_0}X)\, \mathrm{d}t. \tag{43}$$

Proof. From [3, eq. (7.4)] we have

$$\begin{aligned}\tilde{D}'T_{x_0}X = {} & \tilde{\nabla}_{T_{x_0}X} \tilde{P}_{X_t(x_0),\bullet}\, \mathrm{d}_m X_t(x_0) + \tilde{\nabla}_{T_{x_0}X} Z\, \mathrm{d}t \\ & - \frac{1}{2}\left(\tilde{R}'(T_{x_0}X, \mathrm{d}X(x_0))\, \mathrm{d}X(x_0) + \tilde{\nabla}'\tilde{T}'(\mathrm{d}X(x_0), T_{x_0}X, \mathrm{d}X(x_0))\right) \\ & - \frac{1}{2}\tilde{T}'(\tilde{D}'T_{x_0}X, \mathrm{d}X)\end{aligned}$$

where $\tilde{T}'$ denotes the torsion tensor of $\tilde{\nabla}'$. Since for all $x \in M$, $\tilde{\nabla}_v \tilde{P}_{x,\bullet} = 0$ if $v \in T_xM$, the first term in the right vanishes. As a consequence, $\tilde{D}'T_{x_0}X$ has finite variation, and $T'(\tilde{D}'T_{x_0}X, \mathrm{d}X) = 0$. Then using the identity

$$\tilde{R}'(v,u)u + \tilde{\nabla}'\tilde{T}'(u,v,u) = \tilde{R}(v,u)u, \quad u, v \in T_xM,$$

which is a particular case of identity (C.17) in [11], we obtain

$$\tilde{D}' T_{x_0} X = \tilde{\nabla}_{T_{x_0} X} Z \, \mathrm{d}t - \frac{1}{2} \tilde{R}(T_{x_0} X, \mathrm{d}X(x_0))\mathrm{d}X(x_0).$$

Finally writing

$$\tilde{R}(T_{x_0} X, \mathrm{d}X(x_0))\mathrm{d}X(x_0) = \mathrm{Ric}^\sharp(T_{x_0} X)\, \mathrm{d}t$$

yields the result. □

Remark 5.3. In the non-degenerate case, ∇ is the Levi–Civita connection associated to the metric generated by L, and we are in the situation of Sect. 2. In the degenerate case, in general, ∇ does not extend to a metric connection on M. However conditions are given in [11] (1.3.C) under which $P'(X)$ is adapted to some metric, and in this case $T_{x_0} X$ is bounded with respect to the metric.

One would like to extend Theorem 2.1 to degenerate diffusions of constant rank, by solving the equation

$$\partial_u X(u) = \tilde{\nabla}_{\partial_u X(u)} Z \, \mathrm{d}t - \frac{1}{2} \mathrm{Ric}^\sharp(\partial_u X(u))\, \mathrm{d}t.$$

Our proof does not work in this situation for two reasons. The first one is that in general $\tilde{P}'(X)$ is not adapted to a metric. The second one is the lack of an inequality of the type (8) since ∇ does not have an extension $\tilde{\nabla}$ which is the Levi–Civita connection of some metric.

Remark 5.4. When M is a Lie group and L is left invariant, then $\tilde{\nabla}$ can be chosen as the left invariant connection. In this case $(\tilde{\nabla})'$ is the right invariant connection, which is metric.

Acknowledgements The first named author wishes to thank the University of Luxembourg for support.

Note added in proof Using recent results of Kuwada and Philipowski [17], the condition at the beginning of Sect. 4 that $L(t)$ diffusions have lifetime T is automatically satisfied in the case of a family of metrics $g(t)$ evolving by backward Ricci flow on a $g(0)$-complete manifold M. Thus our Theorem 4.1 extends in particular the result of McCann–Topping [19] and Topping [23] from compact to complete manifolds.

References

1. Arnaudon, M.: Differentiable and analytic families of continuous martingales in manifolds with connections. Probab. Theor. Relat. Field **108**(3), 219–257 (1997)
2. Arnaudon, M., Coulibaly, K.A., Thalmaier, A.: Brownian motion with respect to a metric depending on time; definition, existence and applications to Ricci flow. C. R. Math. Acad. Sci. Paris **346**(13–14), 773–778 (2008)

3. Arnaudon, M., Thalmaier, A.: Horizontal martingales in vector bundles, Séminaire de Probabilités XXXVI. Lecture Notes in Mathematics, vol. 1801, pp. 419–456. Springer, Berlin (2003)
4. Arnaudon, M., Thalmaier, A.: Stability of stochastic differential equations in manifolds, Séminaire de Probabilités XXXII. Lecture Notes in Mathematics, vol. 1686, pp. 188–214. Springer, Berlin (1998)
5. Arnaudon, M., Thalmaier, A., Wang, F.-Y.: Gradient estimate and Harnack inequality on non-compact Riemannian manifolds. Stoch. Process. Appl. **119**(10), 3653–3670 (2009)
6. Arnaudon, M., Thalmaier, A., Wang, F.-Y.: Harnack inequality and heat kernel estimates on manifolds with curvature unbounded below. Bull. Sci. Math. **130**, 223–233 (2006)
7. Cipriano, F., Cruzeiro, A.B.: Flows associated with irregular $\mathbb{R}^d$ vector fields. J. Diff. Equ. **210**(1), 183–201 (2005)
8. Cipriano, F., Cruzeiro, A.B.: Flows associated to tangent processes on the Wiener space. J. Funct. Anal. **166**(2), 310–331 (1999)
9. Cruzeiro, A.B.: Equations différentielles sur l'espace de Wiener et formules de Cameron-Martin non linéaires. J. Funct. Anal. **54**(2), 206–227 (1983)
10. Driver, B.K.: A Cameron-Martin type quasi-invariance theorem for Brownian motion on a compact manifold. J. Funct. Anal. **110**, 272–376 (1992)
11. Elworthy, K.D., Le Jan, Y., Li, X.-M.: On the geometry of diffusion operators and stochastic flows. Lecture Notes in Mathematics, vol. 1720, p. 118. Springer, Berlin (1999)
12. Fang, S., Luo, D.: Quasi-invariant flows associated to tangent processes on the Wiener space, Bull. Sci. Math. **134**(3), 295–328 (2010)
13. Gong, F., Zhang, J.: Flows associated to adapted vector fields on the Wiener space, J. Funct. Anal. **253**(2), 647–674 (2007)
14. Hsu, E.P.: Quasi-invariance of the Wiener measure on path spaces: noncompact case, J. Funct. Anal. **193**(2), 278–290 (2002)
15. Hsu, E.P.: Quasi-invariance of the Wiener measure on the path space over a compact Riemannian manifold. J. Funct. Anal. **134**(2), 417–450 (1995)
16. Kendall, W.: Nonnegative Ricci curvature and the Brownian coupling property. Stochastics **19**(1–2), 111–129 (1986)
17. Kuwada, K., Philipowski, R.: Non-explosion of diffusion processes on manifolds with time-dependent metric, To appear in Math. Z. arXiv:0910.1730
18. Lott, J.: Optimal transport and Ricci curvature for metric-measure spaces. Surveys in differential geometry, vol. 11, pp. 229–257. International Press, Somerville, MA (2007)
19. McCann, R.J., Topping, P.M.: Ricci flow, entropy and optimal transportation. Amer. J. Math. **132**(3), 711–730 (2010)
20. Otto, F., Westdickenberg, M.: Eulerian calculus for the contraction in the Wasserstein distance, SIAM J. Math. Anal. **37**(4), 1227–1255 (electronic) (2005)
21. von Renesse, M.K., Sturm, K.T.: Transport inequalities, gradient estimates, entropy, and Ricci curvature, Comm. Pure Appl. Math. **58**(7), 923–940 (2005)
22. Stroock, D.W.: An introduction to the analysis of paths on a Riemannian manifold, Mathematical Surveys and Monographs, vol. 74. American Mathematical Society, Providence, RI (2000)
23. Topping, P.: $\mathcal{L}$-optimal transportation for Ricci flow, J. Reine Angew. Math. **636**:93–122 (2009)

http://www.springer.com/978-3-642-15216-0

Séminaire de Probabilités XLIII
(Eds.)C. Donati-Martin; A. Lejay; A. Rouault
2011, XI, 503 p. 14 illus., Softcover
ISBN: 978-3-642-15216-0