

# COUPLINGS IN $L^p$ DISTANCE OF TWO BROWNIAN MOTIONS AND THEIR LÉVY AREA

MICHEL BONNEFONT AND NICOLAS JUILLET

ABSTRACT. We study co-adapted couplings of (canonical hypoelliptic) diffusions on the (subRiemannian) Heisenberg group, that we call (Heisenberg) Brownian motions and are the joint laws of a planar Brownian motion with its Lévy area. We show that contrary to the situation observed on Riemannian manifolds of non-negative Ricci curvature, for any co-adapted coupling, two Heisenberg Brownian motions starting at two given points can not stay at bounded distance for all time  $t \geq 0$ . Actually, we prove the stronger result that they can not stay bounded in  $L^p$  for  $p \geq 2$ . We also study the coupling by reflection, and show that it stays bounded in  $L^p$  for  $0 \leq p < 1$ . Finally, we explain how the results generalise to the Heisenberg groups of higher dimension

## 1. INTRODUCTION

1.1.  $L^\infty$  **control**. The first goal of this paper is to answer a question, concerning heat diffusion on the Heisenberg group, that is implicitly raised by Kuwada in [15, Remark 4.4]. Let us start with some background that motivates the problem. All definitions will be given later. In the literature,  $L^\infty$ -Wasserstein control for a diffusion has been used to deduce  $L^1$ -gradient estimates of its associated semigroup (see for example [23] and the references therein). Kuwada extends this result to  $L^p$ -Wasserstein control and  $L^q$ -gradient estimates for all  $p, q \geq 1$  with  $\frac{1}{p} + \frac{1}{q} = 1$  and, using Kantorovich duality, proves that, conversely,  $L^q$ -gradient estimates allow one to obtain  $L^p$ -Wasserstein control for the diffusion.

On the Heisenberg group  $\mathbb{H}$ , the following  $L^1$ -gradient bound was established by H.Q Li [17] (see also [1]) generalising [8]

$$\forall f \in C_c^\infty(\mathbb{H}_1), \forall t \geq 0, \forall a \in \mathbb{H}, |\nabla_{\mathbb{H}} P_t f(a)| \leq C P_t(|\nabla_{\mathbb{H}} f|)(a),$$

where  $C > 1$  is constant,  $P_t$  denotes the heat semigroup associated to half the sub-Laplacian and  $\nabla_{\mathbb{H}}$  the horizontal gradient (see Section 2.1 for the definitions). Consequently, Kuwada's result implies that the heat diffusion of the Heisenberg group possesses a  $L^\infty$ -Wasserstein control at any time. Therefore, for each  $a, a' \in \mathbb{H}$

---

*Date:* Version of January 11, 2018.

*2010 Mathematics Subject Classification.* 60H10, 60J60, 60J65, 53C17, 22E25.

*Key words and phrases.* Heisenberg group; Co-adapted coupling; Wasserstein distance; Hypoelliptic diffusion.

and each  $t \geq 0$ , there exists a coupling  $(\mathbf{B}_t^a, \mathbf{B}_t^{a'})_{t \geq 0}$  of two Heisenberg Brownian, that is Brownian motions of  $\mathbb{H}$ , starting respectively in  $a, a'$  such that

$$d_{\mathbb{H}}(\mathbf{B}_t^a, \mathbf{B}_t^{a'}) \leq C d_{\mathbb{H}}(a, a') \quad \text{almost surely.} \quad (1)$$

Please note: Firstly the time  $t \geq 0$  is fixed; secondly  $\mathbf{B}_t^a$  and  $\mathbf{B}_t^{a'}$  are conveniently defined on the same probability space. Kuwada's problem is precisely on inverting the quantifiers  $\forall$  and  $\exists$ , namely, he asks whether it is possible to define a coupling of the two Heisenberg Brownian motions  $(\mathbf{B}_t^a)_{t \geq 0}$  and  $(\mathbf{B}_t^{a'})_{t \geq 0}$  such that (1) holds for all  $t \geq 0$ .

In this paper we answer negatively and show that (1) can not hold for all  $t \geq 0$  for *co-adapted* couplings (see Definition 2.1), probably the most usual couplings in the literature for our type of problem, (see, e.g. [4, 11, 12, 13, 6, 21, 16, 19]). Our results hold for the Heisenberg groups of higher dimension, as explained in Section 5, but we only prove them thoroughly in the first Heisenberg group where all the significative ideas are present and the notation is lighter.

**Theorem 1.1.** *Let  $(\mathbf{B}_t)_t$  and  $(\mathbf{B}'_t)_t$  be any two co-adapted Heisenberg Brownian motions starting respectively in  $a = (x, y, z)$  and  $a' = (x', y', z')$  with  $(x' - x)^2 + (y' - y)^2 > 0$ . Then, for every  $C > 0$ ,*

$$\mathbb{P}(\forall t \geq 0, d_{\mathbb{H}}(\mathbf{B}_t, \mathbf{B}'_t) \leq C) \neq 1.$$

**1.2. Comparison with the Riemannian case.** This result shows a significative difference with the Riemannian case. Indeed, on a Riemannian manifold  $M$ , it is well known (see e.g. [23] and [22]) that if the Ricci curvature is bounded from below by  $k \in \mathbb{R}$ , there exists a Markovian coupling of two Brownian motions such that almost surely

$$d(\mathbf{B}_t^a, \mathbf{B}_t^{a'}) \leq e^{-(k/2)t} d(a, a') \quad \text{for all } t \geq 0, a \in M, a' \in M. \quad (2)$$

Here we call Brownian motions the diffusion processes starting at  $a$  and  $a'$  respectively, having generator half the Laplace–Beltrami operator. We make clear that Markovian coupling is a type of co-adapted coupling. Note moreover that the motivation for proving (2) is exactly to provide estimates on the heat semi-group (see, e.g. [6, 7]), so that the historical  $L^p$ -Wasserstein controls have been established for co-adaptive processes whereas  $L^p$ -Wasserstein controls at fixed time may first appear unusual from a stochastic perspective.

We note further that

- the Heisenberg group can be thought as the first sub-elliptic model space of curvature 0 (e.g. [18]) but, as we will see, its behaviour with respect to coupling co-adapted Brownian motions is completely different from the case of Riemannian manifolds with curvature bounded below by  $k = 0$ .
- the Heisenberg group is also classically presented as the limit space for a sequence of Riemannian metrics on the Lie group, the optimal lower bound on the Ricci curvature of which tends to  $-\infty$ . On this topic see [10, 3].

This fact is coherent with the interpretation of Theorem 1.1 as a special case of (2) where the best bound for the  $L^\infty$  control is  $C = e^{-kt}$  with  $k = -\infty$ : There is no possible control for  $t > 0$ .

1.3.  **$L^p$  control for  $p < \infty$ .** To go further, given two diffusion processes  $(\mathbf{B}_t)_{t \geq 0}$  and  $(\mathbf{B}'_t)_{t \geq 0}$  on a metric space  $(M, d)$ , we shall consider the function

$$t \in [0, \infty) \rightarrow \mathbb{E} [d^p(\mathbf{B}_t, \mathbf{B}'_t)]^{\frac{1}{p}} \in [0, \infty].$$

and try to bound it from above uniformly in time for some well-chosen co-adapted coupling. Informally, a coupling is co-adapted if the interaction in the coupling only depends on the common past of the process  $(\mathbf{B}_t, \mathbf{B}'_t)$ . If we denote by  $\mu_t$  and  $\nu_t$  the law of the processes  $(\mathbf{B}_t)_{t \geq 0}$  and  $(\mathbf{B}'_t)_{t \geq 0}$ , we clearly have for each  $t \geq 0$ :

$$\mathbb{E} [d^p(\mathbf{B}_t, \mathbf{B}'_t)]^{\frac{1}{p}} \geq \mathcal{W}_p(\mu_t, \nu_t),$$

where we recall that the  $L^p$ -Wasserstein distance between two probability measures  $\mu$  and  $\nu$  is given by

$$\mathcal{W}_p(\mu, \nu) = \left( \inf_{\pi \in \Pi(\mu, \nu)} \int \int d(x, y)^p d\pi(x, y) \right)^{1/p}. \quad (3)$$

Here  $\Pi(\mu, \nu)$  is the set of probability measures on  $M \times M$  with marginals  $\mu$  et  $\nu$ .

On the Heisenberg group, we will prove the result stronger than Theorem 1.1 that any co-adapted coupling  $(\mathbf{B}_t)_t$  and  $(\mathbf{B}'_t)_t$  of Brownian motions do not stay bounded in  $L^2$ :

**Theorem 1.2.** *Let  $(\mathbf{B}_t)_{t \geq 0}$  and  $(\mathbf{B}'_t)_{t \geq 0}$  be any two co-adapted Heisenberg Brownian motion starting respectively in  $a = (x, y, z)$  and  $a' = (x', y', z')$  such that  $(x' - x)^2 + (y' - y)^2 > 0$ . Then,*

$$\limsup_{t \rightarrow +\infty} \mathbb{E} [d_{\mathbb{H}}^2(\mathbf{B}_t, \mathbf{B}'_t)] \rightarrow +\infty.$$

To complete the picture, we provide a positive result. We show that the coupling by reflection on the Heisenberg group stays bounded in  $L^p$  for  $0 < p < 1$ . Note that for  $0 < p < 1$ , the quantity  $\mathbb{E} [d^p(\mathbf{B}_t, \mathbf{B}'_t)]^{\frac{1}{p}}$  is not a distance, but a quasidistance, in the sense that the triangle inequality only holds up to a multiplicative constant.

**Theorem 1.3.** *Let  $(\mathbf{B}_t)_{t \geq 0}$  and  $(\mathbf{B}'_t)_{t \geq 0}$  be a coupling by reflection of two Heisenberg Brownian motions starting in  $(x, y, z)$  and  $(x', y', z')$ . Then, for every  $p \in (0, 1)$ ,*

$$\sup_{t \geq 0} \mathbb{E} [d_{\mathbb{H}}^p(\mathbf{B}_t, \mathbf{B}'_t)] < +\infty. \quad (4)$$

Unfortunately, we were not able to obtain any results for  $p \in [1, 2)$  and we ignore whether it is possible to find co-adapted couplings such that (4) holds or not for  $p \in [1, 2)$ . One difficulty in this study is to obtain estimates for the expectation of nonnegative (nonconvex) functionals of martingales as typically  $x \mapsto |x|^{1/2}$ , see Remark 3.3.

**Remark 1.4.** *On a Riemannian manifold, for  $p \geq 1$ , the  $L^p$  version of the  $L^\infty$  control (2), namely*

$$\mathbb{E} \left[ d^p(\mathbf{B}_t^a, \mathbf{B}_t^{a'}) \right]^{1/p} \leq e^{-kt/2} d(a, a')$$

*is satisfied (for some appropriate co-adapted coupling) if and only if the Ricci curvature is bounded from below by  $k$  (see [22] and [15, Remark 2.3]). Therefore all the above  $L^p$ -Wasserstein controls, for  $p \in [1, \infty]$  are equivalent and only depend on the Ricci curvature lower bound.*

We mention the interesting recent work by S. Banerjee, M. Gordina and P. Mariano [2] where the authors use non co-adapted couplings to study the decay in total variation for the laws of Heisenberg Brownian motions and obtain gradient estimates for harmonic functions. This work and our work seems to deliver a common message namely that co-adapted couplings are not the unique relevant couplings, what concerns obtaining gradient estimates.

The paper is organised as follows. In Section 2.1, we recall the notion of co-adapted coupling and describe quickly the geometry of the Heisenberg group, its associated Brownian motions and their coupling. We also discuss some classical couplings. The proofs of the main theorems on the non-existence of co-adapted Heisenberg Brownian motion which stay at bounded distance are given Section 3. The reflection coupling on  $\mathbb{H}$  is studied in Section 4 and the results are generalised to the Heisenberg groups of higher dimension in the final section.

## 2. CO-ADAPTED COUPLINGS ON THE HEISENBERG GROUP

**2.1. The Heisenberg group.** The Heisenberg group can be identified with  $\mathbb{R}^3$  equipped with the law:

$$(x, y, z) \cdot (x', y', z') = \left( x + x', y + y', z + z' + \frac{1}{2}(xy' - yx') \right).$$

The left invariant vector fields are given by

$$\begin{cases} X(f)(x, y, z) = \frac{d}{dt}|_{t=0} f((x, y, z) \cdot (t, 0, 0)) = \left( \partial_x - \frac{y}{2} \partial_z \right) f(x, y, z) \\ Y(f)(x, y, z) = \frac{d}{dt}|_{t=0} f((x, y, z) \cdot (0, t, 0)) = \left( \partial_y + \frac{x}{2} \partial_z \right) f(x, y, z) \\ Z(f)(x, y, z) = \frac{d}{dt}|_{t=0} f((x, y, z) \cdot (t, 0, 0)) = \partial_z f(x, y, z). \end{cases}$$

Note that  $[X, Y] = Z$  and that  $Z$  commutes with  $X$  and  $Y$ .

We are interested in half the sub-Laplacian  $L = \frac{1}{2}(X^2 + Y^2)$ . This is a diffusion operator that satisfies the Hörmander bracket condition and thus the associated heat semigroup  $P_t = e^{tL}$  admits a  $\mathcal{C}^\infty$  positive kernel  $p_t$ .

From a probabilistic point of view,  $L$  is the generator of the following stochastic process starting in  $(x, y, z)$ :

$$\mathbf{B}_t^{(x,y,z)} := (x, y, z) \cdot \left( B_t^1, B_t^2, \frac{1}{2} \left( \int_0^t B_s^1 dB_s^2 - \int_0^t B_s^2 dB_s^1 \right) \right)$$

where  $(B_t^1)_{t \geq 0}$  and  $(B_t^2)_{t \geq 0}$  are two standard independent 1-dimensional Brownian motions. The quantity  $\int_0^t B_s^1 dB_s^2 - \int_0^t B_s^2 dB_s^1$  that we denote by  $A_t$  is one of the first stochastic integral ever considered. It is the Lévy area of the 2-dimensional Brownian motion  $(B_t)_{t \geq 0} := (B_t^1, B_t^2)_{t \geq 0}$ .

It is easily seen that  $(\mathbf{B}_t)_{t \geq 0}$  is a continuous process with independent and stationary increments. We simply call it the Heisenberg Brownian motion.

The sub-Laplacian  $L$  is strongly related to the following subRiemmanian distance (also called Carnot-Carathéodory) on  $\mathbb{H}$ :

$$d_{\mathbb{H}}(a, a') = \inf_{\gamma} \int_0^1 |\dot{\gamma}(t)|_{\mathfrak{h}} dt$$

where  $\gamma$  ranges over the horizontal curves connecting  $\gamma(0) = a$  and  $\gamma(1) = a'$ . We remind the reader of the fact that a curve is said horizontal if it is absolutely continuous and  $\dot{\gamma}(t) \in \text{Vect}(X(\gamma(t)), Y(\gamma(t)))$  almost surely holds. The horizontal norm  $|\cdot|_{\mathfrak{h}}$  is a Euclidean norm on  $\text{Vect}(X, Y)$  obtained by asserting that  $(X, Y)$  is an orthonormal basis of  $\text{Vect}(X(a), Y(a))$  at each point  $a \in \mathbb{H}$ . Finally the horizontal gradient  $\nabla_{\mathfrak{h}} f$  is  $(Xf)X + (Yf)Y$ .

The Heisenberg group admits homogeneous dilations adapted both to the distance and the group structure. They are given by

$$\text{dil}_{\lambda}(x, y, z) = (\lambda x, \lambda y, \lambda^2 z)$$

for  $\lambda > 0$  and satisfies in law the following equality:

$$\text{dil}_{\frac{1}{\sqrt{t}}}\left(B_t^1, B_t^2, \frac{1}{2}\left(\int_0^t B_s^1 dB_s^2 - \int_0^t B_s^2 dB_s^1\right)\right) \stackrel{\text{Law}}{=} \left(B_1^1, B_1^2, \frac{1}{2}\left(\int_0^1 B_s^1 dB_s^2 - \int_0^1 B_s^2 dB_s^1\right)\right).$$

The distance is clearly left-invariant:

$$d_{\mathbb{H}}(a, a) = d_{\mathbb{H}}(e, a^{-1}a')$$

with  $e = (0, 0, 0)$ . Since its explicit expression is not so easy, it is often simpler to work with a homogenous quasinorm (still in the sense that the triangle inequality only holds up to a multiplicative constant). We will use

$$H : a = (x, y, z) \in \mathbb{H} \mapsto \sqrt{x^2 + y^2 + |z|} \in \mathbb{R},$$

and the attached homogeneous quasidistance  $d_H(a, a') = H(a^{-1}a')$ . It satisfies

$$cd_H(a, a') \leq d_{\mathbb{H}}(a, a') \leq Cd_H(a, a')$$

for some constants  $c, C > 0$ .

**2.2. Co-adapted couplings.** We first recall the notion of *co-adapted* coupling of two processes. Indeed, in this study, we only want to consider couplings build solely knowing the past of the two processes. The definition below is taken from [13, Definition 1.1].

**Definition 2.1.** Given two continuous-time Markov processes  $(X_t^{(1)})_{t \geq 0}$ ,  $(X_t^{(2)})_{t \geq 0}$ , we say that  $(\tilde{X}_t^{(1)}, \tilde{X}_t^{(2)})_{t \geq 0}$  is a co-adapted coupling of  $(X_t^{(1)})_{t \geq 0}$  and  $(X_t^{(2)})_{t \geq 0}$  if  $\tilde{X}^{(1)}$  and  $\tilde{X}^{(2)}$  are defined on the same filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , satisfy  $\text{Law}(X_t^{(i)})_{t \geq 0} = \text{Law}(\tilde{X}_t^{(i)})_{t \geq 0}$  for  $i = 1, 2$ , and

$$\tilde{P}_t^{(i)} f : z \mapsto \mathbb{E} \left[ f(\tilde{X}_{t+s}^{(i)}) \mid \mathcal{F}_s, \tilde{X}_s^{(i)} = z \right]$$

equals

$$P_t^{(i)} f : z \mapsto \mathbb{E} \left[ f(X_{t+s}^{(i)}) \mid X_s^{(i)} = z \right], \quad \text{Law}(X_s^{(i)})\text{-almost surely}$$

for  $i = 1, 2$ , for each bounded measurable function  $f$ , each  $z$ , each  $s, t \geq 0$ .

If we moreover assume that the full process  $(\tilde{X}_t^{(1)}, \tilde{X}_t^{(2)})_{t \geq 0}$  is Markovian, we say that the co-adapted coupling is Markovian.

The next lemma describes more explicitly co-adapted couplings in the case of Brownian motion in  $\mathbb{R}^2$  (see [12, Lemma 6]).

**Lemma 2.2.** Let  $(B_t)_t$  and  $(B'_t)_t$  be two co-adapted Brownian motions on  $\mathbb{R}^2 \times \mathbb{R}^2$  defined on some filtered probability space  $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ . Then, enriching the filtration if necessary, there exists a Brownian motion  $(\hat{B}_t)_{t \geq 0}$  defined on the same filtration  $(\mathcal{F}_t)_{t \geq 0}$  and independent from  $(B_t)_{t \geq 0}$  such that

$$dB'(t) = J(t)dB_t + \hat{J}(t)d\hat{B}_t \tag{5}$$

where  $(J_t)_{t \geq 0} = ((J_t^{i,j})_{1 \leq i, j \leq 2})_{t \geq 0}$  and  $\hat{J}$  are matrices satisfying

$$JJ^T + \hat{J}\hat{J}^T = I_2 \tag{6}$$

and  $J(t), \hat{J}(t) \in \mathcal{F}_t$ .

In the following  $\|\cdot\|$  may denote the operator norm of a matrix attached to the Euclidean norm, or the Euclidean norm of a vector.

**Lemma 2.3.** Let  $J$  be a  $2 \times 2$  real matrix  $J = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ . Then

$$0 \leq J^T J \leq I_2 \iff \|J\| \leq 1 \iff 0 \leq JJ^T \leq I_2,$$

where  $\leq$  is the ordering of symmetric matrices. In particular

- $a^2 + b^2$ ,  $a^2 + c^2$ ,  $c^2 + d^2$  and  $b^2 + d^2$  are smaller or equal to 1,
- all the four entries of  $J$  are in  $[-1, 1]$ .

*Proof.* Let  $\mathbb{S}^1 = \{(\cos(\theta), \sin(\theta)) \in \mathbb{R}^2 : \theta \in \mathbb{R}\}$  be the Euclidean sphere of  $\mathbb{R}^2$  and  $Q : x \in \mathbb{S}^1 \mapsto (x, J^T J x) = \|Jx\|^2$ . Therefore,  $Q$  is bounded by 1 if and only if  $\|Jx\| \leq 1$ , for all  $x \in \mathbb{S}^1$ . The bound  $0 \leq J^T J$  is trivially satisfied. The proof is completed by  $\|J\| = \sup_{x, y \in \mathbb{S}^1} (Jx, y) = \|J^T\|$ .  $\square$

**Remark 2.4.** A necessary and sufficient condition can be found considering  $\lambda$ , the greatest eigenvalue of  $J^T J$ . It writes

$$2\lambda = (a^2 + b^2 + c^2 + d^2) + \sqrt{(a^2 + b^2 + c^2 + d^2)^2 - 4(bc - ad)^2} \leq 2 \times 1.$$

Paradoxically, it not easy to deduce  $|a|, |b|, |c|, |d| \leq 1$  from this condition.

**2.3. Co-adapted couplings on  $\mathbb{H}$ .** We now describe co-adapted Heisenberg Brownian motions. As seen before, a Brownian motion  $\mathbf{B}$  is entirely determined by its two first coordinates  $(B_t)_t = (B_t^1, B_t^2)_t$ ; the third one being  $(A_t)_t$  the Lévy area swept by this 2-dimensional process  $(B_t)_t$ .

Thus two Heisenberg Brownian motions  $(\mathbf{B}_t)_t = (B_t^1, B_t^2, A_t)_t$  and  $(\mathbf{B}'_t)_t = (B_t'^1, B_t'^2, A'_t)_t$  on  $\mathbb{H}$  are co-adapted if and only if  $B = (B_t^1, B_t^2)_t$  and  $B' = (B_t'^1, B_t'^2)_t$  are two co-adapted Brownian motion on  $\mathbb{R}^2$  and if moreover their third coordinates satisfy

$$dA_t = \frac{1}{2} (B_t^1 dB_t^2 - B_t^2 dB_t^1)$$

and

$$dA'_t = \frac{1}{2} (B_t'^1 dB_t'^2 - B_t'^2 dB_t'^1).$$

For the following, we denote by  $J$  and  $\hat{J}$  the matrices appearing in Lemma 2.2. A computation gives:

$$\mathbf{B}'_t{}^{-1} \mathbf{B}_t = \left( B_t^1 - B_t'^1, B_t^2 - B_t'^2, B_t^3 - B_t'^3 - \frac{1}{2} (B_t^1 B_t'^2 - B_t^2 B_t'^1) \right)$$

and thus:

$$d(\mathbf{B}'_t{}^{-1} \mathbf{B}_t) = \left( \begin{array}{c} dB_t^1 - dB_t'^1 \\ dB_t^2 - dB_t'^2 \\ (B_t^1 - B_t'^1) \left( \frac{dB_t'^2 + dB_t^2}{2} \right) - (B_t^2 - B_t'^2) \left( \frac{dB_t'^1 + dB_t^1}{2} \right) - \frac{1}{2} (d\langle B_t^1, B_t'^2 \rangle - d\langle B_t^2, B_t'^1 \rangle) \end{array} \right),$$

where we used:

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + d\langle X_t, Y_t \rangle.$$

We denote by  $R_t$  the horizontal distance between the two Brownian motions  $B_t$  and  $B'_t$  in  $\mathbb{R}^2$ , that is  $R_t^2 = (B_t^1 - B_t'^1)^2 + (B_t^2 - B_t'^2)^2$  and by  $Z_t$  the third coordinate, the relative Lévy area. Hence  $Z_t = (\mathbf{B}'_t{}^{-1} \mathbf{B}_t)_3$ .

The homogeneous distance  $d_H(\mathbf{B}_t, \mathbf{B}'_t)$  is thus given by

$$\sqrt{R_t^2 + |Z_t|}.$$

In the following, when  $R_t > 0$ , we choose to work in the direct orthonormal (random moving) frame  $(v_1, v_2)$  defined by taking  $v_1(t)$  the normalised vector of  $\mathbb{R}^2$  directed by  $B_t - B'_t$ . Let  $Q_t$  be the matrix whose columns are respectively  $v_1(t)$

and  $v_2(t)$ . In this new basis, for  $(\alpha, \beta) \in \mathbb{R}^2$  and  $(\cdot | \cdot)$  the usual scalar product on  $\mathbb{R}^2$ , we have:

$$\begin{aligned} (\alpha v_1 + \beta v_2 | dB'_t) &= \left( Q_t \begin{pmatrix} \alpha \\ \beta \end{pmatrix} | J_t dB_t + \hat{J}_t d\hat{B}_t \right) \\ &= \left( \begin{pmatrix} \alpha \\ \beta \end{pmatrix} | (Q_t^T J_t Q_t) Q_t^T dB_t + (Q_t^T \hat{J}_t Q_t) Q_t^T d\hat{B}_t \right) \\ &= \left( \begin{pmatrix} \alpha \\ \beta \end{pmatrix} | K_t dW_t + \hat{K}_t d\hat{W}_t \right) \end{aligned}$$

for  $K_t = Q_t^T J_t Q_t$  and  $\hat{K}_t = Q_t^T \hat{J}_t Q_t$ , and where  $W$  and  $\hat{W}$  are the two standard independent 2-dimensional Brownian motions defined by

$$dW_t = Q_t^T dB_t, \quad d\hat{W}_t = Q_t^T d\hat{B}_t.$$

This can be summed up as follows:

$$Q_t^T dB'_t = \underbrace{(Q_t^T J_t Q_t)}_{K_t} Q_t^T dB_t + \underbrace{(Q_t^T \hat{J}_t Q_t)}_{\hat{K}_t} Q_t^T d\hat{B}_t.$$

The next easy lemma describes the relation between the matrices  $J$  and  $K$ .

**Lemma 2.5.** *With the above notation, when  $R_t > 0$ ,*

- *Equation (6) is satisfied for  $(K, \hat{K})$  if and only if it is satisfied for  $(J, \hat{J})$ .*
- $\text{tr } K = \text{tr } J$ .
- $K^{1,2} - K^{2,1} = J^{1,2} - J^{2,1}$ .

*Proof.* The first two relations follow from the fact that  $Q$  is an orthogonal matrix. For the last relation, one can note that  $J^{1,2} - J^{2,1} = \text{tr}(JM) = \text{tr}(KQ^T M Q)$  with  $M$  the matrix  $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ . Now a computation gives  $Q^T M Q = (\det Q)M$  and the last relation follows from the fact  $Q$  is actually a rotation matrix.  $\square$

The stochastic processes  $R_t^2$  and  $Z_t$  are semimartingales defined for all time  $t \geq 0$ . In the next statement, we provide stochastic differential equations for their evolution.

**Lemma 2.6.** *With the above notation, when  $R_t \neq 0$ , the processes  $R_t^2$  and  $Z_t$  solve the stochastic differential equation:*

$$\begin{cases} d(R_t^2) = 2R_t \sqrt{2(1 - K^{1,1})} dC_t + (2(1 - K^{1,1}) + 2(1 - K^{2,2})) dt \\ dZ_t = \frac{R_t}{2} \sqrt{2(1 + K^{2,2})} d\tilde{C}_t + \frac{1}{2}(K^{1,2} - K^{2,1}) dt \end{cases}$$

where  $(C_t)_{t \geq 0}$  and  $(\tilde{C}_t)_{t \geq 0}$  are some 1-dimensional Brownian motions whose co-variation satisfies:

$$\langle \sqrt{2(1 - K^{1,1})} dC_t, \sqrt{2(1 + K^{2,2})} d\tilde{C}_t \rangle = (K^{1,2} - K^{2,1}) dt. \quad (7)$$



**Remark 2.7.** *Actually the stochastic process  $(R_t^2, Z_t)_{t \geq 0}$  is perfectly defined for all  $t \geq 0$  (even when  $R_t = 0$ ). The technical problem in Lemma 2.6 is that the matrix  $Q_t$  and thus the matrix  $K_t$  are only defined for  $R_t \neq 0$ . However, the matrix  $J_t$  is defined for every value of  $R_t$  and we have:*

$$\begin{cases} d(R_t^2) = \sigma_R(B_t, B'_t, J_t) dC_t + \left(2(1 - J^{1,1}) + 2(1 - J^{2,2})\right) dt \\ dZ_t = \sigma_Z(B_t, B'_t, J_t) d\tilde{C}_t + \frac{1}{2}(J^{1,2} - J^{2,1})dt \end{cases}$$

where  $\sigma_R$  and  $\sigma_Z$  are defined by:

$$\sigma_R(B_t, B'_t, J_t) = \begin{cases} 0 & \text{if } B_t = B'_t \\ 2R_t \sqrt{2(1 - (P_t^T J_t P_t)^{1,1})} & \text{if } B_t \neq B'_t \end{cases}$$

and

$$\sigma_Z(B_t, B'_t, J_t) = \begin{cases} 0 & \text{if } B_t = B'_t \\ \frac{R_t}{2} \sqrt{2(1 + (P_t^T J_t P_t)^{2,2})} & \text{if } B_t \neq B'_t. \end{cases}$$

Note finally that the fact that  $\sigma_R$  and  $\sigma_Z$  vanish for  $R_t = 0$  is rather clear from their expressions in Lemma 2.6.

*Proof of Lemma 2.6.* The computations are done in [11] but we repeat them for the sake of completeness.

First by Itô formula and with the previous notation:

$$\begin{aligned} dR_t^2 &= d\left((B_t^1 - B_t'^1)^2 + (B_t^2 - B_t'^2)^2\right) \\ &= 2R_t(v_1 \mid (dB_t - dB_t')) + d\langle (B_t^1 - B_t'^1), (B_t^1 - B_t'^1) \rangle + d\langle (B_t^2 - B_t'^2), (B_t^2 - B_t'^2) \rangle. \end{aligned}$$

We turn to the martingale part and write

$$\begin{aligned} (v_1 \mid (dB_t - dB_t')) &= \left((K^{1,1} - 1)dW_t^1 + K^{1,2}dW_t^2 + \hat{K}^{1,1}d\hat{W}_t^1 + \hat{K}^{1,2}d\hat{W}_t^2\right) \\ &= \sqrt{2(1 - K^{1,1})}dC_t \end{aligned}$$

for some 1-dimensional Brownian motion  $(C_t)_t$  where we used Lemma 2.5 for

$$(K^{1,1})^2 + (K^{1,2})^2 + (\hat{K}^{1,1})^2 + (\hat{K}^{1,2})^2 = 1.$$

The quadratic variation writes

$$d\langle (B_t^1 - B_t'^1), (B_t^1 - B_t'^1) \rangle = (J^{1,1} - 1)^2 + (J^{1,2})^2 + (\hat{J}^{1,1})^2 + (\hat{J}^{1,2})^2 = 2 - 2J^{1,1}$$

and similarly

$$d\langle (B_t^2 - B_t'^2), (B_t^2 - B_t'^2) \rangle = (J^{2,1})^2 + (J^{2,2} - 1)^2 + (\hat{J}^{2,1})^2 + (\hat{J}^{2,2})^2 = 2 - 2J^{2,2},$$

thus

$$d\langle (B_t^1 - B_t'^1), (B_t^1 - B_t'^1) \rangle + d\langle (B_t^2 - B_t'^2), (B_t^2 - B_t'^2) \rangle = 2\text{tr}(I - J) = 2\text{tr}(I - K).$$

We turn now to  $Z_t$ . Using the basis  $(v_1, v_2)$ , we can rewrite

$$dZ_t = \frac{R_t}{2} (v_2 \mid (dB'_t + dB_t)) - \frac{1}{2} (d\langle B_t^1, B_t'^2 \rangle - d\langle B_t^2, B_t'^1 \rangle).$$

As before, we get:

$$\begin{aligned} (v_2 \mid (dB'_t + dB_t)) &= K^{2,1} dW_t^1 + (K^{2,2} + 1) dW_t^2 + \hat{K}^{2,1} d\hat{W}_t^1 + \hat{K}^{2,2} d\hat{W}_t^2 \\ &= \sqrt{2(1 + K^{2,2})} d\tilde{C}_t \end{aligned}$$

for some 1-dimensional Brownian motion  $(\tilde{C}_t)_t$ . Moreover:

$$d\langle B_t^1, B_t'^2 \rangle - d\langle B_t^2, B_t'^1 \rangle = (J^{2,1} - J^{1,2}) dt = (K^{2,1} - K^{1,2}) dt.$$

The equation on the covariation (7) follows since by (6) and Lemma 2.5,

$$K^{1,1} K^{2,1} + K^{1,2} K^{2,2} + \hat{K}^{1,1} \hat{K}^{2,1} + \hat{K}^{1,2} \hat{K}^{2,2} = 0.$$

□

**Remark 2.8.** *Since we will use them in the following we also write stochastic differential equations satisfied by  $R_t$ ,  $R_t^4$  and  $Z_t^2$ , obtained for  $R_t \neq 0$  using Itô's formula in Lemma 2.6:*

$$\left\{ \begin{array}{l} dR_t = \sqrt{2(1 - K^{1,1})} dC_t + \frac{1 - K^{2,2}}{R_t} dt, \\ dR_t^4 = 4R_t^3 \sqrt{2(1 - K^{1,1})} dC_t + (4R_t^2(1 - K^{2,2})) dt + 12R_t^2(1 - K^{1,1}) dt, \\ d(Z_t^2) = Z_t R_t \sqrt{2(1 + K^{2,2})} d\tilde{C}_t + \left( Z_t (K^{1,2} - K^{2,1}) + \frac{R_t^2}{2} (1 + K^{2,2}) \right) dt. \end{array} \right.$$

As in Remark 2.7, an expression for  $dZ_t^2$  is possible in the canonical basis with the matrix  $J$  in place of  $K$ . According to Lemma 2.5,  $K^{1,2} - K^{2,1}$  is replaced by  $J^{1,2} - J^{2,1}$  and the factor  $R_t$  make the undefined terms vanish when  $R_t = 0$ . The same holds for the semimartingale  $(R_t^4)_{t \geq 0}$ . On the contrary  $R_t$  is not a semimartingale, but only locally when  $R_t > 0$ .

**2.4. Description of some couplings.** In this section, we describe some interesting couplings. We assume for simplicity that  $R_0 = 1$  and  $Z_0 = 0$ .

2.4.1. *The synchronous coupling* ( $K^{1,1} = 1$ ,  $K^{2,2} = 1$ ,  $K^{1,2} = K^{2,1} = 0$ .) The coupling is called synchronous because the planar trajectories  $(B_t)_{t \geq 0}$  and  $(B'_t)_{t \geq 0}$  are parallel. Here  $R_t \equiv R_0$  and  $Z_t = Z_0 + W_{R_0 t}$  with  $W$  a Brownian motion.

2.4.2. *The reflection coupling* ( $K^{1,1} = -1$ ,  $K^{2,2} = 1$ ,  $K^{1,2} = K^{2,1} = 0$ .) For the reflection coupling the planar trajectories of  $(B_t)_{t \geq 0}$  and  $(B'_t)_{t \geq 0}$  evolve symmetrically with respect to the bisector of line segment  $[B_0, B'_0]$ . We stop the coupling when  $R_t$  hits 0 and continue synchronously with  $J^{1,1} = J^{2,2} = 1$ . Denote by  $\tau$  this hitting time. One thus has

$$R_t = R_0 + 2C_{t \wedge \tau} \text{ and } Z_t = Z_0 + \int_0^{t \wedge \tau} (R_0 + 2C_{s \wedge \tau}) d\tilde{C}_s$$

where  $(C_s)_s$  and  $(\tilde{C}_s)_s$  are two independent Brownian motions (starting in 0) and with  $\tau = \inf\{s \geq 0, 2C_s = -R_0\}$ . This coupling is studied in Section 4. On Euclidean and Riemannian manifolds, the efficiency of reflection coupling has been studied in [9, 14].

2.4.3. *Kendall's coupling*: ( $K^{1,1} = \pm 1$ ,  $K^{1,2} = K^{2,1} = 0$  and  $K^{2,2} = 1$ ). In [11], Kendall describes a coupling which alternates between synchronous coupling and reflection coupling. In order to avoid the use of local times the strategy of Kendall is defined with hysteresis. The regime swaps when the process  $(R_t, |Z_t|)$  hits a certain parabola, not the same depending for the synchronous or the reflection coupling. Thus the process is not Markovian, but it is co-adapted. The author proves that this coupling is successful: this means  $T := \inf\{s \geq 0, \mathbf{B}_s = \mathbf{B}'_s\}$  is almost surely finite, or, equivalently, the process  $(R_t, Z_t)$  hits almost surely  $(0, 0)$  in finite time.

2.4.4. *The perverse coupling*:  $K^{1,1} = 1$ ,  $K^{2,2} = -1$ ,  $K^{1,2} = K^{2,1} = 0$ . It satisfies,

$$dR(t) = \frac{2}{R_t} dt \text{ and } dZ_t = 0.$$

Thus the distance  $R_t$  and  $Z_t$  are deterministic and given by:

$$R_t = \sqrt{R_0^2 + 4t} \text{ and } Z_t = Z_0.$$

The name perverse coupling is given by Kendall in [12] as a generic name for a repulsive coupling. Here, the planar components of  $(\mathbf{B}_t)_t$  and  $(\mathbf{B}'_t)_t$  are coupled in a perverse way. This particular method to produce a perverse coupling appears in [19, Section 5] in a Riemannian setting.

### 3. NON-EXISTENCE OF CO-ADAPTED HEISENBERG BROWNIAN MOTIONS AT BOUNDED DISTANCE.

We now turn to the proofs of Theorems 1.1 and 1.2.

3.1. **Proof of Theorem 1.1.** Theorem 1.1 is clearly a corollary of Theorem 1.2 but we can prove it more easily and it already shows a clear difference with the Riemannian case. Hence, we first present a proof of this result.

*Proof of Theorem 1.1.* Assume that  $(\mathbf{B}_t)_t$  and  $(\mathbf{B}'_t)_t$  are two co-adapted Heisenberg Brownian motions starting in  $a = (x, y, z)$  and  $a' = (x', y', z')$  with  $R_0 = \sqrt{(x' - x)^2 + (y' - y)^2} > 0$ . Striving for a contradiction we suppose that  $t \mapsto d_{\mathbb{H}}(\mathbf{B}_t, \mathbf{B}'_t)$  is almost surely and uniformly bounded. More precisely for some  $C > 0$  we assume  $|R_t| + |Z_t|^{1/2} \leq C$  for every  $t \geq 0$ .

Using Lemmas 2.5 and 2.6 (or simply Remark 2.7), we have

$$\mathbb{E}[R_t^2] = R_0^2 + \mathbb{E} \left[ 2 \int_0^t (1 - J^{1,1}(s)) + (1 - J^{2,2}(s)) ds \right]$$

and  $R_t \leq C$  gives  $\mathbb{E}[R_t^2] \leq C^2$  and

$$\mathbb{E} \left[ 2 \int_0^t (1 - J^{1,1}(s)) + (1 - J^{2,2}(s)) ds \right] \leq C^2. \quad (8)$$

Recall from Lemma 2.5 that  $K^{1,1} + K^{2,2} = J^{1,1} + J^{2,2}$  and from Lemma 2.3 that the matrix entries are  $\geq -1$ . Therefore

$$\max_{i \in \{1,2\}} \mathbb{E} \left[ \int_0^t \left( (1 - K^{i,i}(s)) \right) \mathbf{1}_{\{R_s > 0\}} ds \right] \leq C^2/2 \quad (9)$$

and  $R_t \leq C$ , again, gives

$$\mathbb{E} \left[ \int_0^t (1 - K^{2,2}(s)) \frac{R_s^2}{2} ds \right] \leq C^4/4. \quad (10)$$

Until now we have used  $\mathbb{E}(R_t^2) \leq C^2$  and  $R_t \leq C$ . We turn to exploit  $\mathbb{E}(Z_t^2) \leq C^4$ . Lemma 2.5 and Remark 2.8 give

$$\mathbb{E}[Z_t^2] = Z_0^2 + \mathbb{E} \left[ 2 \int_0^t Z_s (J^{1,2}(s) - J^{2,1}(s)) ds + \int_0^t \frac{R_s^2}{2} (1 + K^{2,2}(s)) ds \right]$$

which does not have to be nonnegative. However,

$$\mathbb{E} \left[ 2 \int_0^t Z_s (J^{1,2}(s) - J^{2,1}(s)) ds + \int_0^t \frac{R_s^2}{2} (1 + K^{2,2}(s)) ds \right] \leq C^4. \quad (11)$$

Adding (10) and (11) we obtain

$$\mathbb{E} \left[ 2 \int_0^t Z_t (J^{1,2}(s) - J^{2,1}(s)) ds + \int_0^t R_s^2 ds \right] \leq (1 + 1/4)C^4. \quad (12)$$

Next, we aim to compare  $\mathbb{E} \left[ 2 \int_0^t Z_s (J^{1,2}(s) - J^{2,1}(s)) ds \right]$  with  $\mathbb{E} \left[ \int_0^t R_s^2 ds \right]$  that both appear in (12). On the one hand, since  $Z_t$  stays bounded  $\mathbb{E}[Z_t^2]$  is also bounded. Hence by Cauchy-Schwarz inequality (for the product measure on  $\Omega \times$

$[0, t]$ ):

$$\begin{aligned}
 \left| \mathbb{E} \left[ \int_0^t Z_s (J^{1,2}(s) - J^{2,1}(s)) ds \right] \right| &\leq \left( \int_0^t \mathbb{E}[Z_t^2] ds \right)^{1/2} \left( \int_0^t \mathbb{E}(J^{1,2} - J^{2,1})^2 ds \right)^{1/2} \\
 &\leq \left( \int_0^t \mathbb{E}[Z_t^2] ds \right)^{1/2} \left( 2 \int_0^t \mathbb{E}((J^{1,2})^2 + (J^{2,1})^2) ds \right)^{1/2} \\
 &\leq \sqrt{C^4 t} \left( 2 \int_0^t \mathbb{E}(J^{1,2})^2 + \mathbb{E}(J^{2,1})^2 ds \right)^{1/2} \\
 &\leq \sqrt{C^4 t} \cdot \sqrt{2C^2} = C^3 \sqrt{2t}. \tag{13}
 \end{aligned}$$

The last estimate follows from Lemma 2.3 (the rows and columns of  $J$  have  $L^2$ -norm smaller than 1),  $1 - J_{i,i}^2 \leq (1 - J_{i,i})(1 + J_{i,i})$  and (8):

$$\begin{aligned}
 \int_0^t \mathbb{E}(J^{1,2})^2 + \mathbb{E}(J^{2,1})^2 ds &\leq \int_0^t \mathbb{E}[(1 - (J^{1,1})^2) + (1 - (J^{2,2})^2)] ds \\
 &\leq \int_0^t 2\mathbb{E}[(1 - J^{1,1}) + (1 - J^{2,2})] ds \leq C^2.
 \end{aligned}$$

On the other hand, since  $(R_t^2)_{t \geq 0}$  is a submartingale,  $\mathbb{E}[R_s^2] \geq R_0^2$  and

$$\int_0^t \mathbb{E}[R_s^2] ds \geq R_0^2 t. \tag{14}$$

Since  $R_0 > 0$ , (13) and (14) provide a contradiction in (12) as  $t$  goes to infinity.  $\square$

**Remark 3.1.** *The proof of Theorem 1.1 can be improved to show that any co-adapted Heisenberg Brownian motions can not stay bounded in  $L^4$ . In this proof, the only place where we fully use the fact that  $R_t$  is uniformly bounded almost surely is (10). At the other places we merely need that  $\mathbb{E}[R_t^2]$  and  $\mathbb{E}[Z_t^2]$  are bounded. But  $\mathbb{E}[R_t^4] \leq C^4$  for every  $t \geq 0$  is a sufficient assumption for (10) and, hence, for the proof.*

Indeed, by Lemma 2.8 one has

$$\mathbb{E}[R_t^4] = R_0^4 + \mathbb{E} \left[ \int_0^t 12R_s^2(1 - K^{1,1}(s)) + 4R_s^2(1 - K^{2,2}(s)) ds \right].$$

*This quantity is uniformly bounded by  $C^4$  for every  $t \geq 0$  so that (10) holds. (the bound in (10) can even be divided by two:  $C^4/8$  in place of  $C^4/4$ ).*

**3.2. Proof of Theorem 1.2.** To go beyond Theorem 1.1, we conduct a precise study of the expected total variation (or length in  $L^1$ ) of the martingale part and of the drift part of  $(Z_t)_{t \geq 0}$ , the relative Lévy area. We need (13) from the proof of Theorem 1.1, and Lemma 3.2, the statement and the proof of which are postponed at the end of the section.

*Proof of Theorem 1.2.* Let  $C$  be  $\sup_{t \geq 0} \sqrt{\max(\mathbb{E}(R_t^2), \mathbb{E}(|Z_t|))}$  and as before, assume  $C < +\infty$  by contradiction. First recall

$$\mathbb{E}(R_t^2) = R_0^2 + \mathbb{E}\left(\int_0^t 2[(1 - J^{1,1}) + (1 - J^{2,2})]ds\right) \geq 0,$$

which gives

$$\mathbb{E}\left(\int_0^{+\infty} [(1 - J^{1,1}) + (1 - J^{2,2})]ds\right) \leq \frac{C^2}{2} < +\infty. \quad (15)$$

Let  $T := \inf\{t \geq 0, R_t = \frac{R_0}{2}\}$  be the hitting time of  $\frac{R_0}{2}$ . We show that we can assume  $\mathbb{P}(T = +\infty) > 0$ . Suppose for the rest of this paragraph  $\mathbb{P}(T < +\infty) = 1$  and let  $S$  the finite random variable defined by

$$S = \int_0^T 2(1 - K^{1,1})ds.$$

Because of the nonnegative drift in the stochastic differential equation of  $R_t$ , using the Dambins theorem, the random variable  $S$  is greater in stochastic order than the hitting time of  $\frac{R_0}{2}$  for a Brownian motion starting in  $R_0$ . This hitting time is almost surely finite but nonintegrable. Thus  $\mathbb{E}(S) = +\infty$  which contradicts (15) (Recall from Lemma 2.5 that  $J^{1,1} + J^{2,2} = K^{1,1} + K^{2,2}$  and from Lemma 2.3 that these quantities are  $\geq -1$ ).

Now, let us decompose the semimartingale  $(Z_t)_t = M_t - A_t$  into its martingale  $M_t$  and bounded variation part  $-A_t$ . From Lemma 2.6, we recall:

$$-A_t = \frac{1}{2} \int_0^t (J^{1,2} - J^{2,1})ds. \quad (16)$$

Applying Cauchy–Schwarz inequality and following the same track as for (13) we obtain

$$\mathbb{E} \int_0^t |J^{1,2} - J^{2,1}|ds \leq \sqrt{t} \sqrt{\mathbb{E} \int_0^t (J^{1,2} - J^{2,1})^2 ds} \leq \sqrt{t} \sqrt{2C^2}. \quad (17)$$

Remark now that the quantity on the left hand side is two times the expected total variation of  $A_t$  on  $[0, t]$ .

We postpone the proof of the following result until the end of the (present) proof. It occurs as an application of Lemma 3.2: there exists  $h > 0$  such that for every  $t > 0$ ,

$$\mathbb{E}(|M_{t+h} - M_t|) \geq 10C^2. \quad (18)$$

Since, we have assumed  $E(|M_t - A_t|) \leq C^2$  for every  $t \geq 0$ , the triangle inequality implies  $\mathbb{E}(|A_{t+h} - A_t|) \geq 8C^2$  for every  $t \geq 0$ . The control of the expected total variation of  $(A_t)$  expressed in (17) and the lower estimate just proved give

$$8C^2 n \leq \sum_{k=1}^n \mathbb{E}(|A_{kh} - A_{(k-1)h}|) \leq \frac{1}{2} \int_0^{nh} |J^{1,2} - J^{2,1}| ds \leq \sqrt{\frac{C^2 hn}{2}},$$

which, as  $n$  tends to  $\infty$ , provides a contradiction with our initial assumption that was  $\sup_{t \geq 0} \sqrt{\max(\mathbb{E}(R_t^2), \mathbb{E}(|Z_t|))} \leq C$ . We are left with the proof of (18) (under the assumption of the  $L^2$  boundedness).

Recall  $T = \inf\{t \geq 0 : R_t = \frac{R_0}{2}\}$  and set  $q = \mathbb{P}(T = +\infty)$ . We have already proved  $q > 0$ . We shall show further that for  $h \geq \frac{5C^2}{q}$ ,

$$\mathbb{P}\left(\langle M \rangle_{t+h} - \langle M \rangle_t \geq \frac{R_0^2 h}{8}\right) \geq q - \frac{C^2}{2h} \geq \frac{9q}{10}. \quad (19)$$

We hence obtain (18) taking  $h$  great enough in (19) and applying Lemma 3.2 to  $(M_{t+h} - M_t)_{h \geq 0}$ . Note that the proof of Lemma 3.2 is independent from the rest of the present proof.

Proof of (19): considering only the event  $\{T = +\infty\}$  for the martingale part of  $Z_t$  described in Lemma 2.6, one has

$$\langle M \rangle_{t+h} - \langle M \rangle_t \geq \mathbf{1}_{\{T=+\infty\}} \left(\frac{R_0}{2}\right)^2 \int_t^{t+h} \frac{1 + K^{2,2}}{2} ds. \quad (20)$$

Now, since by (15) it holds

$$\mathbb{E}\left[\int_t^{t+h} \mathbf{1}_{\{T=+\infty\}} \frac{1 - K^{2,2}}{2} ds\right] \leq C^2/4,$$

taking the complementary set of  $\{\int_t^{t+h} \frac{1+K^{2,2}}{2} ds \geq \frac{h}{2}\}$  in  $\{T = +\infty\}$  and using Markov inequality, one obtains:

$$\begin{aligned} & \mathbb{P}\left(\mathbf{1}_{\{T=+\infty\}} \int_t^{t+h} \frac{1 + K^{2,2}}{2} ds \geq \frac{h}{2}\right) \\ &= q - \mathbb{P}\left(\mathbf{1}_{\{T=+\infty\}} \int_t^{t+h} \frac{1 - K^{2,2}}{2} ds > \frac{h}{2}\right) \geq q - \frac{C^2}{4} \cdot \frac{2}{h}. \end{aligned} \quad (21)$$

Hence, in (20) we consider the probability that the right-hand side is greater than  $(R_0/2)^2 \cdot (h/2)$ , which, with (21), gives the wanted estimate (19) for every  $h \geq \frac{5C^2}{q}$ .  $\square$

**Lemma 3.2.** *Let  $(N_t)_{t \geq 0}$  be a continuous martingale with  $N_0 = 0$  and  $p$  be in  $]0, 1[$ . Then there exists  $a_p > 0$  such that for every positive real numbers  $\beta$  and  $h$ , the estimate*

$$\mathbb{P}(\langle N \rangle_h \geq \beta) \geq p$$

*implies*

$$\mathbb{E}[|N_h|] \geq a_p \sqrt{\beta}.$$

*Proof.* Let  $\phi$  be the quadratic variation of  $N$

$$\phi(t) = \langle N \rangle_t,$$

and consider the hitting time  $\tau = \inf\{t \geq 0 : \phi(t) \geq \beta\}$ . Set  $\psi(t) = \phi(t) \wedge \beta$ . The Dambins theorem shows that there exists a standard Brownian motion  $(W_t)_{t \geq 0}$  such that for every  $t \geq 0$ ,

$$\mathbb{E}[|N_t|] \geq \mathbb{E}[|N_{t \wedge \tau}|] = \mathbb{E}[|W_{\psi(t)}|]. \quad (22)$$

Let now  $A$  be the event  $\{\omega \in \Omega : \phi(h) \geq \beta\}$  and recall the assumption  $\mathbb{P}(A) \geq p$ . One has

$$\mathbb{E}[|W_{\psi(h)}|] \geq \mathbb{E}[|W_{\psi(h)}| \cdot \mathbf{1}_A] = \mathbb{E}[|W_\beta| \cdot \mathbf{1}_A] \geq a_p \sqrt{\beta} \quad (23)$$

where the constant  $a_p$  is given by

$$a_p = \mathbb{E}\left[|G| \mathbf{1}_{\{|G| \leq \Phi^{-1}(\frac{1+p}{2})\}}\right] = \int_{\Phi^{-1}(\frac{1-p}{2})}^{\Phi^{-1}(\frac{1+p}{2})} |x| \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx.$$

with  $G$  a standard normal random variable and  $\Phi$  its cumulative distribution function. The lower bound in (23) is obtained for  $\mathbb{P}(A) = p$  and the normal random variable  $W_\beta$  of variance  $\beta$  concentrated as much as possible close to zero on event  $A$ . Equation (22) for  $t = h$  and (23) finally provide the wanted estimate.  $\square$

**Remark 3.3.** *The major constraint for generalising Theorem 1.2 and its proof to a  $L^p$ -Wasserstein control for  $p < 2$  is that  $\mathbb{E}(|Z_t|)$  is replaced by  $\mathbb{E}(|Z_t|^{p/2})$ . Here  $x \mapsto |x|^{p/2}$  is not convex when  $p < 2$  and Jensen's inequality does not apply.*

#### 4. COUPLING BY REFLECTION

In this section, we study precisely the coupling by reflection. We recall that  $(\mathbf{B}_t)_{t \geq 0}$  and  $(\mathbf{B}'_t)_{t \geq 0}$  are two Heisenberg Brownian motions coupled by reflection if and only if  $(B_t)_{t \geq 0}$  and  $(B'_t)_{t \geq 0}$  are two Brownian motions on  $\mathbb{R}^2$  coupled by reflection. This means that the coupling matrices are given by  $K^{1,1} = -1$ ,  $K^{2,2} = 1$ ,  $K^{1,2} = K^{2,1} = 0$  for  $t < \tau$  and by the matrix  $J = \text{Id}_2$  for  $t \geq \tau$  where  $\tau = \inf\{s \geq 0 : R_s = 0\}$  is the hitting time of 0 for  $(R_t)_{t \geq 0}$ . We recall

$$R_t = R_0 + 2C_{t \wedge \tau} \quad \text{and} \quad Z_t = Z_0 + \int_0^{t \wedge \tau} (R_0 + 2C_{s \wedge \tau}) d\tilde{C}_s$$

where  $(C_s)_s$  and  $(\tilde{C}_s)_s$  are two independent Brownian motions (starting in 0) and with  $\tau = \inf\{s \geq 0 : C_s = -R_0/2\}$ .

For simplicity, in the following we only consider the case  $R_0 > 0$  and  $Z_0 = 0$ .

**Proposition 4.1.** *With the above notation, assume  $R_0 > 0$  and  $Z_0 = 0$ , Let  $p > 0$ , then there exists some constants  $C_p, C'_p, C''_p > 0$  such that*

$$\mathbb{E}[R_t] = R_0$$

and

$$\begin{cases} \mathbb{E}[|Z_t|^p] \sim_{t \rightarrow \infty} C_p R_0 t^{p-\frac{1}{2}} & \text{if } p > \frac{1}{2} \\ \mathbb{E}[|Z_t|^p] \sim_{t \rightarrow \infty} C'_p R_0 \ln t & \text{if } p = \frac{1}{2} \\ \mathbb{E}[|Z_t|^p] \rightarrow_{t \rightarrow +\infty} C''_p R_0^{2p} & \text{if } 0 < p < \frac{1}{2}. \end{cases}$$



**Remark 4.2.** *In particular for  $0 < \alpha < 1$ , the upper bound*

$$\sup_{t \geq 0} \mathbb{E} [d_{\mathbb{H}}(\mathbf{B}_t, \mathbf{B}'_t)^\alpha] < +\infty$$

*is satisfied by the coupling by reflection. This is obtained by recalling that  $d_{\mathbb{H}}(\mathbf{B}_t, \mathbf{B}'_t)$  is equivalent to the homogeneous distance  $\sqrt{R_t^2 + |Z_t|}$  using Proposition 4.1 for  $p = \alpha/2$  and Jensen's inequality  $\mathbb{E}(R_t)^\alpha \leq \mathbb{E}(R_t)^\alpha$ .*

*Proof.* We assume  $R_0 = 1$ . Let  $t > 0$  be fixed. By the Dambins theorem,  $Z$  is a changed time Brownian motion:

$$Z_t = W_{T(t)} \text{ with } T(t) = \int_0^t R_s^2 ds$$

with  $W$  a Brownian motion independent from  $(R_t)_{t \geq 0}$ . Set  $\tau = \inf\{s \geq 0 : R_s = 0\}$ . As  $(R_s/2)_{s \geq 0}$  is a Brownian motion starting in  $R_0/2$  and stopped in 0, it is known that  $\tau$  is almost surely finite and that its density  $f_\tau$  is given by

$$f_\tau(u) = \frac{R_0/2}{\sqrt{2\pi}u^{3/2}} e^{-\frac{R_0^2}{4u}}, \quad u \geq 0. \quad (24)$$

Using  $\tau$ , we compute

$$\begin{aligned} E[|Z_t|^p] &= \mathbb{E}(|W_{T(t)}|^p) \\ &= \int_0^{+\infty} \mathbb{E}(|W_{T(t)}|^p | \tau = u) f_\tau(u) du \\ &= \underbrace{\int_0^t \mathbb{E}(|W_{T(t)}|^p | \tau = u) f_\tau(u) du}_{h_1(t)} + \underbrace{\int_t^{+\infty} \mathbb{E}(|W_{T(t)}|^p | \tau = u) f_\tau(u) du}_{h_2(t)}. \end{aligned} \quad (25)$$

In the last line, we split the integral between the trajectories of  $R$  that have hit 0 before  $t$  and those which will hit 0 after  $t$ .

Let us estimate  $h_1(t)$ , the first integral in the decomposition (25). Hence we set  $u \leq t$ . Since  $W$  and  $R$  are independent, with  $c_p = \mathbb{E}(|W_1|^p)$ , one has:

$$\begin{aligned} \mathbb{E}(|W_{T(t)}|^p | \tau = u) &= c_p \mathbb{E}(|T(t)|^{p/2} | \tau = u) \\ &= c_p \mathbb{E} \left( \left( \int_0^u R_s^2 ds \right)^{p/2} | \tau = u \right) \\ &= c_p 2^p u^p \mathbb{E} \left[ \left( \int_0^1 \tilde{R}_\lambda^2 d\lambda \right)^{p/2} | \tau = u \right] \end{aligned} \quad (26)$$

where we have introduced the normalised process  $(\tilde{R}_\lambda)_{\lambda \in [0,1]}$  (defined almost surely, since the hitting time  $\tau$  is almost surely finite) in such a way it hits 0 at time 1:

$$\tilde{R}_\lambda = \frac{1}{2\sqrt{\tau}} R_{\tau\lambda}, \quad \lambda \in [0, 1].$$

Note that  $\tilde{R}_0 = \frac{R_0}{2\sqrt{\tau}}$ .

It is then well-known that, conditioned on  $\tau = u$ , the entire process  $(\tilde{R}_\lambda)_{\lambda \in [0,1]}$  converges in law when  $u \rightarrow \infty$  to a normal positive Brownian excursion  $(X_s)_{s \in [0,1]}$ . Moreover, as proven in Lemma 4.3, when  $u \rightarrow \infty$ ,

$$\mathbb{E} \left[ \left( \int_0^1 \tilde{R}_\lambda^2 d\lambda \right)^{p/2} \mid \tau = u \right] \rightarrow \mathbb{E} \left[ \left( \int_0^1 X_s^2 ds \right)^{p/2} \right]. \quad (27)$$

Finally with (26) denoting the limit in (27) by  $E_p$ , the first integral in (25) satisfies the following equivalence:

$$h_1(t) = \int_0^t \mathbb{E}(|W_{T(t)}|^p \mid \tau = u) f_\tau(u) du \sim_{t \rightarrow \infty} 2^p c_p E_p \int_0^t u^p f_\tau(u) du$$

From the density estimate of  $f_\tau$  in (24) we have  $u^p f_\tau(u) \sim_{+\infty} \frac{R_0}{2\sqrt{2\pi}} u^{p-3/2}$ . Therefore:

- If  $p > 1/2$ , the function  $h_1(t)$  is equivalent to  $\frac{2^{p-3/2} R_0 c_p E_p}{(p-1/2)\sqrt{\pi} u^{3/2}} t^{p-1/2}$  at  $+\infty$ ,
- if  $p = 1/2$ , it is equivalent to  $\frac{R_0 c_p E_p}{2\sqrt{\pi} u^{3/2}} \ln t$ ,
- if  $0 < p < 1/2$ , it converges to a positive constant.

We now turn to  $h_2$ . As before,

$$\begin{aligned} h_2(t) &= \int_t^{+\infty} \mathbb{E}(|W_{T(t)}|^p \mid \tau = u) f_\tau(u) du \\ &= c_p \int_t^{+\infty} \mathbb{E} \left( \left( \int_0^t R_s^2 ds \right)^{p/2} \mid \tau = u \right) f_\tau(u) du \\ &= 2^p c_p \int_t^{+\infty} u^p \mathbb{E} \left[ \left( \int_0^{\frac{t}{u}} \tilde{R}_\lambda^2 d\lambda \right)^{p/2} \mid \tau = u \right] f_\tau(u) du \\ &= 2^p c_p t^{p+1} \int_1^{+\infty} v^p \mathbb{E} \left[ \left( \int_0^{\frac{1}{v}} \tilde{R}_\lambda^2 d\lambda \right)^{p/2} \mid \tau = tv \right] f_\tau(tv) dv \\ &= \frac{2^{p-3/2} c_p R_0 t^{p-3/2}}{\sqrt{\pi}} \int_1^{+\infty} v^{p-3/2} \mathbb{E} \left[ \left( \int_0^{\frac{1}{v}} \tilde{R}_\lambda^2 d\lambda \right)^{p/2} \mid \tau = tv \right] e^{-\frac{R_0^2}{tv}} dv \end{aligned}$$

where, as above,  $s = u\lambda$  and  $\tilde{R}_\lambda = R_{\tau\lambda}/\sqrt{\tau}$  and where we set the change of variable  $u = tv$  in the next to last line.

Now, Lemma 4.3 and the dominated convergence, which is completely justified by Lemma 4.4, give as  $t \rightarrow +\infty$ ,

$$\int_1^{+\infty} v^{p-3/2} \mathbb{E} \left[ \left( \int_0^{\frac{1}{v}} \tilde{R}_\lambda^2 d\lambda \right)^{p/2} \middle| \tau = tv \right] e^{-\frac{R_0^2}{4tv}} dv \rightarrow \int_1^{+\infty} v^{p-3/2} \mathbb{E} \left[ \left( \int_0^{\frac{1}{v}} X_s^2 ds \right)^{p/2} \right] dv.$$

As a consequence, denoting by  $I_p$  the last integral,

$$h_2(t) \sim_{t \rightarrow \infty} \frac{2^{p-3/2} c_p R_0}{\sqrt{\pi}} t^{p-1/2} I_p.$$

This with the treatment of  $h_1$  above gives the complete result in case  $R_0 = 1$ . Next, if  $R_0 > 0$  one infers  $\mathbb{E}_{R_0}[|Z_t|^p] = R_0^{2p} \mathbb{E}_{\{R_0=1\}}[|Z_{t/R_0^2}|^p]$  from the classical dilations of Subsection 2.1, so that the general case follows.  $\square$

The two next lemmas complete the proof of Proposition 4.1.

**Lemma 4.3.** *Let  $(\tilde{R}_t)_{t \in [0,1]}$  be a Brownian motion starting in  $r_0 > 0$  conditioned to hit 0 for the first time at time 1. Let  $\alpha \in [0, 1]$  and  $p$  be positive. As  $r_0 \rightarrow 0$ ,*

$$\mathbb{E} \left[ \left( \int_0^\alpha \tilde{R}_\lambda^2 d\lambda \right)^{p/2} \right] \rightarrow \mathbb{E} \left[ \left( \int_0^\alpha X_s^2 ds \right)^{p/2} \right] \quad (28)$$

where  $(X_s)_{s \in [0,1]}$  is a Brownian excursion.

*Proof.* The process  $(\tilde{R}_t)_{t \in [0,1]}$  converges in law to the Brownian excursion  $(X_s)_{s \in [0,1]}$ . To obtain the convergence of the moments of  $\int_0^\alpha \tilde{R}_\lambda^2 d\lambda$ , we use a uniform integrability property. Let  $0 \leq \alpha \leq 1$ . We bound

$$\mathbb{P} \left( \left( \int_0^\alpha \tilde{R}_\lambda^2 d\lambda \right) \geq y \middle| \tau = u \right) \leq \mathbb{P} \left( \sup_{0 \leq s \leq 1} W_s \geq \sqrt{y} \middle| T_0 = 1 \right)$$

where  $(W_s)_{s \geq 0}$  is a Brownian motion starting in  $r_0$  and  $T_0$  its hitting time of 0. Next, by [5, Formula 2.1.4 (1) p.198], still for  $W$  starting in  $r_0$ , we have for every  $t > 0$

$$\mathbb{P} \left( \sup_{0 \leq s \leq T_0} W_s < y \middle| T_0 = t \right) = \sum_{k=-\infty}^{+\infty} \frac{(r_0 + 2ky)}{\sqrt{2\pi t^{3/2}}} \exp \left( -\frac{(r_0 + 2ky)^2}{2t} \right) \frac{\sqrt{2\pi t^{3/2}}}{r_0} \exp \left( \frac{r_0^2}{2t} \right),$$

In particular, reorganising the terms,

$$\mathbb{P} \left( \sup_{0 \leq s \leq 1} W_s < y \middle| T_0 = 1 \right) = 1 - 2 \sum_{k=1}^{+\infty} \left( 4k^2 y^2 \frac{\sinh(2kyr_0)}{2kyr_0} - \cosh(2kyr_0) \right) \exp(-2k^2 y^2),$$

and, since for  $u \geq 0$ ,  $\frac{\sinh u}{u} \leq e^u$ , uniformly on  $0 < r_0 \leq 1$  and  $y \geq 1$ ,

$$\mathbb{P} \left( \sup_{0 \leq s \leq 1} W_s \geq y \middle| T_0 = 1 \right) \leq 8 \sum_{k=1}^{+\infty} k^2 y^2 \exp(2ky) \exp(-2k^2 y^2) \leq a \exp(-by^2)$$

for some  $a, b > 0$ .

Thus, for all  $0 < r_0 \leq 1$ , the random variables  $\left(\int_0^\alpha \tilde{R}_\lambda^2 d\lambda\right)_{r_0 \leq 1}$  admit some uniformly bounded exponential moment. As a consequence, the corresponding  $\left(\int_0^\alpha \tilde{R}_\lambda^2 d\lambda\right)^{\frac{p}{2}}$  are uniformly integrable and the desired convergence follows.  $\square$

**Lemma 4.4.** *With the above notation, there exists a coupling of  $(\tilde{R}_t)_{t \in [0,1]}$  and of a 3-Bessel process  $(V_t)_{t \in [0,1]}$  on the same probability space such that both start in  $r_0 \geq 0$  and such almost surely:*

$$\tilde{R}_t \leq V_t, \text{ for all } 0 \leq t \leq 1.$$

In particular, for  $R_0$  fixed and  $p > 0$  there exists a constant  $D_p > 0$  such that for all  $v \geq 1$  and  $r_0 \leq \frac{R_0}{\sqrt{v}}$  it holds

$$v^p \mathbb{E} \left[ \left( \int_0^{\frac{1}{v}} \tilde{R}_\lambda^2 d\lambda \right)^{p/2} \right] \leq D_p$$

where the process  $\tilde{R}_\lambda$  starts in  $r_0$ .

*Proof.* The process  $\tilde{R}_t$  can be thought as a Bessel bridge (see e.g. [20, Chapter XIII]) and thus satisfies

$$d\tilde{R}_t = dW_t + \left( \frac{1}{\tilde{R}_t} - \frac{\tilde{R}_t}{1-t} \right) dt$$

for some Brownian motion  $(W_t)_{t \in [0,1]}$ . The coupling is obtained by considering the same Brownian motion in the stochastic differential equation defining  $(V_t)_{t \in [0,1]}$ :

$$dV_t = dW_t + \frac{1}{V_t} dt$$

The other conclusion follows since the 3-Bessel process shares the same scaling property as the Brownian motion:  $(V_{\lambda t}/\sqrt{\lambda})_t$  has the same law as the 3-Bessel process starting in  $r_0/\sqrt{\lambda}$ .  $\square$

## 5. GENERALISATION OF THEOREMS 1.1 1.2 AND 1.3 TO THE HEISENBERG GROUPS OF HIGHER DIMENSION.

In this section, we prove that Theorems 1.1 1.2 and 1.3 also hold in the case of the Heisenberg groups of higher dimension. For  $n \geq 1$  the Heisenberg group  $\mathbb{H}_n$  can be identified with  $\mathbb{R}^{2n+1}$  equipped with the law:

$$((x_i, y_i)_{i=1, \dots, n}, z) \cdot ((x'_i, y'_i)_{i=1, \dots, n}, z') = \left( (x_i + x'_i, y_i + y'_i)_{i=1, \dots, n}, z + z' + \frac{1}{2} \sum_{i=1}^n (x_i y'_i - y_i x'_i) \right).$$

The corresponding Brownian motion starting from  $0_{\mathbb{R}^{2n+1}}$  is given by:

$$\mathbf{B}_t := \left( (B_{t,i}^1, B_{t,i}^2)_{i=1, \dots, n}, \frac{1}{2} \sum_{i=1}^n \left( \int_0^t B_{s,i}^1 dB_{s,i}^2 - \int_0^t B_{s,i}^2 dB_{s,i}^1 \right) \right)$$

where  $B_t := (B_{t,i}^1, B_{t,i}^2)_{i=1,\dots,n}$  is a  $2n$ -dimensional standard Brownian motion.

Lemma 2.2 can directly be generalised for describing co-adapted Heisenberg Brownian motions  $(\mathbf{B}_t, \mathbf{B}'_t)$  but with matrices  $J, \hat{J} \in \mathcal{M}_{2n}(\mathbb{R})$ . As above, we denote by  $R_t$  the Euclidean norm of  $B'_t - B_t \in \mathbb{R}^{2n}$  and by  $Z_t$  the last coordinate of  $\mathbf{B}'_t^{-1}\mathbf{B}_t$  that we still call the relative Lévy area. The quantity

$$d_H(\mathbf{B}_t, \mathbf{B}'_t) := \sqrt{R_t^2 + |Z_t|}$$

is still a homogenous distance on  $\mathbb{H}_n$  and is equivalent to the Carnot-Carathéodory distance  $d_{\mathbb{H}}$ .

When  $R_t > 0$  we introduce the following basis: let  $e_1$  be  $\frac{1}{R_t}(B'_t - B_t) \in \mathbb{R}^{2n}$ , write  $e_1 = (a_1, \dots, a_n)$ ,  $a_j \in \mathbb{C} = \mathbb{R}^2$ , set  $e_2 = (ia_1, \dots, ia_n)$  and complete  $(e_1, e_2)$  into a direct orthonormal basis of  $\mathbb{R}^{2n}$ . This basis is well adapted for studying couplings in  $\mathbb{H}_n$ . Indeed, with  $L$  and  $\hat{L}$  being the coupling matrices in this new basis in place of  $J, \hat{J}$  in the canonical basis, a computation gives:

**Lemma 5.1.** *With the above notation, if  $R_t > 0$ , then*

$$\begin{cases} d(R_t^2) = 2R_t \sqrt{2(1 - L^{1,1})} dC_t + 2 \operatorname{tr}(I_{2n} - J) dt \\ dZ_t = \frac{1}{2} R_t \sqrt{2(1 + L^{2,2})} d\tilde{C}_t + \frac{1}{2} \sum_{i=1}^n (J^{2i-1,2i} - J^{2i,2i-1}) dt \end{cases}$$

where  $(C_t)_{t \geq 0}$  and  $(\tilde{C}_t)_{t \geq 0}$  are some 1-dimensional (possibly correlated) standard Brownian motions.

Now, since  $\operatorname{tr}L = \operatorname{tr}J$ , and since each  $|L^{i,i}| \leq 1$  for  $1 \leq i \leq n$ ,

$$1 - L^{2,2} \leq \operatorname{tr}(I - L) = \operatorname{tr}(I - J)$$

and one can directly adapt the proof of Theorems 1.1 and 1.2 to this setting.

The coupling by reflection can also be done on  $\mathbb{H}_n$ . It corresponds to the matrix  $L$  defined by

$$\begin{cases} L^{1,1} = -1 \\ L^{i,i} = 1 \text{ for } 2 \leq i \leq n \\ L^{i,j} = 0 \text{ for } i \neq j. \end{cases}$$

In this case, a computation easily gives that  $C_t$  and  $\tilde{C}_t$  are independent. Moreover since  $L$  is symmetric,  $J$  is also symmetric and  $\sum_{i=1}^n (J^{2i-1,2i} - J^{2i,2i-1}) = 0$ . As a consequence,  $R_t^2$  and  $Z_t$  satisfy the same stochastic differential system as in the case of  $\mathbb{H}_1$ . Thus Proposition 4.1 holds in  $\mathbb{H}_n$  with constants  $C_p, C'_p$  and  $C''_p$  independent of the dimension and Theorem 1.3 also holds for  $\mathbb{H}_n$ .

## REFERENCES

- [1] D. Bakry, F. Baudoin, M. Bonnefont, and D. Chafaï. On gradient bounds for the heat kernel on the Heisenberg group. *Journal of Functional Analysis*, 255:1905–1938, 2008.

- [2] S. Banerjee, M. Gordina, and P. Mariano. Coupling in the Heisenberg group and its applications to gradient estimates. *ArXiv e-prints*, Oct. 2016.
- [3] F. Baudoin and N. Garofalo. Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries. *J. Eur. Math. Soc. (JEMS)*, 19(1):151–219, 2017.
- [4] G. Ben Arous, M. Cranston, and W. S. Kendall. Coupling constructions for hypoelliptic diffusions: two examples. In *Stochastic analysis (Ithaca, NY, 1993)*, volume 57 of *Proc. Sympos. Pure Math.*, pages 193–212. Amer. Math. Soc., Providence, RI, 1995.
- [5] A. N. Borodin and P. Salminen. *Handbook of Brownian motion—facts and formulae*. Probability and its Applications. Birkhäuser Verlag, Basel, second edition, 2002.
- [6] M. Cranston. Gradient estimates on manifolds using coupling. *J. Funct. Anal.*, 99(1):110–124, 1991.
- [7] M. Cranston. A probabilistic approach to gradient estimates. *Canad. Math. Bull.*, 35(1):46–55, 1992.
- [8] B. K. Driver and T. Melcher. Hypoelliptic heat kernel inequalities on the Heisenberg group. *J. Funct. Anal.*, 221(2):340–365, 2005.
- [9] E. P. Hsu and K.-T. Sturm. Maximal coupling of Euclidean Brownian motions. *Commun. Math. Stat.*, 1(1):93–104, 2013.
- [10] N. Juillet. Geometric inequalities and generalized Ricci bounds in the Heisenberg group. *International Mathematical Research Notices*, 2009, 2009.
- [11] W. S. Kendall. Coupling all the Lévy stochastic areas of multidimensional Brownian motion. *Ann. Probab.*, 35(3):935–953, 2007.
- [12] W. S. Kendall. Brownian couplings, convexity, and shy-ness. *Electron. Commun. Probab.*, 14:66–80, 2009.
- [13] W. S. Kendall. Coupling time distribution asymptotics for some couplings of the Lévy stochastic area. In *Probability and mathematical genetics*, volume 378 of *London Math. Soc. Lecture Note Ser.*, pages 446–463. Cambridge Univ. Press, Cambridge, 2010.
- [14] K. Kuwada. On uniqueness of maximal coupling for diffusion processes with a reflection. *J. Theoret. Probab.*, 20(4):935–957, 2007.
- [15] K. Kuwada. Duality on gradient estimates and Wasserstein controls. *J. Funct. Anal.*, 258(11):3758–3774, 2010.
- [16] K. Kuwada and K.-T. Sturm. A counterexample for the optimality of Kendall-Cranston coupling. *Electron. Comm. Probab.*, 12:66–72, 2007.
- [17] H.-Q. Li. Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg. *J. Funct. Anal.*, 236(2):369–394, 2006.
- [18] R. Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.
- [19] M. N. Pascu and I. Popescu. Shy and fixed-distance couplings of Brownian motions on manifolds. *Stochastic Process. Appl.*, 126(2):628–650, 2016.
- [20] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.
- [21] M.-K. von Renesse. Intrinsic coupling on Riemannian manifolds and polyhedra. *Electron. J. Probab.*, 9:no. 14, 411–435, 2004.
- [22] M.-K. von Renesse and K.-T. Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.
- [23] F.-Y. Wang. On estimation of the logarithmic Sobolev constant and gradient estimates of heat semigroups. *Probab. Theory Related Fields*, 108(1):87–101, 1997.

INSTITUT DE MATHÉMATIQUES DE BORDEAUX, UNIVERSITÉ DE BORDEAUX, 351, COURS DE  
LA LIBÉRATION, 33405 TALENCE CEDEX, FRANCE

*E-mail address:* michel.bonnefont@math.u-bordeaux.fr

INSTITUT DE RECHERCHE MATHÉMATIQUE AVANCÉE, UNIVERSITÉ DE STRASBOURG, 7 RUE  
RENÉ-DESCARTES 67084 STRASBOURG CEDEX FRANCE

*E-mail address:* nicolas.juillet@math.unistra.fr