Couplings in L^p distance of two Brownian motions and their Lévy area

Michel Bonnefont, Nicolas Juillet

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 prove two positive results. We first study the coupling by reflection and show that it stays bounded in L^p for $0 \leq p < 1$. Secondly, we construct an explicit static (and in particular non co-adapted) coupling between the laws of two Brownian motions, which provides L^1-Wasserstein control uniformly in time.

Finally, we explain how the results generalise to the Heisenberg groups of higher dimension.

1. Introduction

1.1. L^∞ control. The motivation for this paper is a question, concerning heat diffusion on the Heisenberg group, that is implicitly raised by Kuwada in [17, Remark 4.4], and that we reproduce at page 2 after Theorem 1.1. Before we reach this question let us start with some background and a few definitions. All remaining material will be introduced later in the paper. In the literature, L^∞-Wasserstein control for a diffusion has been used to deduce L^1-gradient estimates of its associated semigroup (see for instance [25] 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.

Date: Version of July 11, 2018.

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

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

We recall that, on a metric space (M,d), for $p \in (0,\infty]$, the L^p-Wasserstein distance between two probability measures μ and ν is given by

$$\mathcal{W}_p(\mu,\nu) = \left(\inf_{\pi\in\Pi(\mu,\nu)} \iint d(x,y)^p d\pi(x,y) \right)^{1/p} = \inf_{X\sim\mu,\, Y\sim\nu} \|d_{\mathbb{H}}(X,Y)\|_p. \quad (1)$$

Here $\Pi(\mu,\nu)$ is the set of probability measures on $M\times M$ with marginals μ and ν. For $p=\infty$ the first expression is replaced by the essential supremum of d. Note that $\mathcal{W}_p$ is a distance only for $p\geq 1$. For $0<p<1$, it is only a quasidistance, in the sense that the triangle inequality only holds up to a multiplicative constant. Using Hölder inequality, it is clear that $\mathcal{W}_p(\mu,\nu)\leq \mathcal{W}_q(\mu,\nu)$ if $0<p\leq q$.

On the Heisenberg group $\mathbb{H}$, the following L^1-gradient bound was established by H.Q. Li [19] (see also [1]) generalising [8]

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

where $C>1$ is constant, P_t denotes the heat semigroup associated to half the sub-Laplacian and ∇_{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^∞-Wasserstein control for its diffusion:

Theorem 1.1 (H.Q. Li, Kuwada)**.** *There exists $C>0$ such that for every $t\geq 0$ and $a,a'\in\mathbb{H}$*

$$\mathcal{W}_\infty(\mu_t^a,\mu_t^{a'}) \leq C d_{\mathbb{H}}(a,a'), \quad (2)$$

where $\mu_t^a = \mathrm{Law}(\mathbf{B}_t^a)$ and $\mu_t^{a'} = \mathrm{Law}(\mathbf{B}_t^{a'})$ and $(\mathbf{B}_s^a)_{s\geq 0}, (\mathbf{B}_s^{a'})_{s\geq 0}$ are two Heisenberg Brownian motions, starting respectively in a,a'. Moreover,

$$\mathcal{W}_p(\mu_t^a,\mu_t^{a'}) \leq C d_{\mathbb{H}}(a,a') \quad (3)$$

holds for every $p<\infty$.

In other words, for each $a,a'\in\mathbb{H}$ and each $t\geq 0$, there exists a coupling $(\mathbf{B}_s^a,\mathbf{B}_s^{a'})_{s\geq 0}$ of the two Heisenberg Brownian motions such that

$$d_{\mathbb{H}}(\mathbf{B}_t^a,\mathbf{B}_t^{a'}) \leq C d_{\mathbb{H}}(a,a') \quad \text{almost surely.} \quad (4)$$

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 (and the remaining random variables $(\mathbf{B}_s^a)_{s\neq t}$ and $(\mathbf{B}_s^{a'})_{s\neq t}$ of the Heisenberg Brownian motions are defined without paying attention to their correlation). 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 (4) holds for all $t\geq 0$.

In this paper we answer negatively and show that (4) 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, 13, 14, 15, 6, 23, 18, 21]). Informally, a coupling of two processes $(\mathbf{B}_t)$ and $(\mathbf{B}'_t)$ is said co-adapted if the interaction in the coupling only depends on the common past of the process $(\mathbf{B}_t, \mathbf{B}'_t)$. See Definition 2.1 fo the rigorous definition.

Our results hold for the Heisenberg groups of higher dimension, as explained in Section 6, but we only prove them thoroughly in the first Heisenberg group where all the significative ideas are present and the notation is lighter.

Theorem 1.2. *For every $T > 0$ and every $C > 0$ there exists two points $a, a' \in \mathbb{H}$ with $a \neq a'$ such that for every co-adapted coupling $(\mathbf{B}_t^a, \mathbf{B}_t^{a'})_{0 \le t \le T}$, there exists $t \le T$ such that*

$$\operatorname{essup}_{(\Omega,\mathbb{P})} d_{\mathbb{H}}(\mathbf{B}_t^a, \mathbf{B}_t^{a'}) > C d_{\mathbb{H}}(a, a').$$

Remark 1.3. *Another way to state Theorem 1.2 is as follows: Let $T > 0$, then*

$$\sup_{a \neq a' \in \mathbb{H}} \inf_{\mathcal{A}_T^{(a,a')}} \sup_{0 \le t \le T} \frac{\operatorname{essup}_{(\Omega,\mathbb{P})} d_{\mathbb{H}}(\mathbf{B}_t^a, \mathbf{B}_t^{a'})}{d_{\mathbb{H}}(a, a')} = +\infty$$

where $\mathcal{A}_T^{(a,a')}$ denotes the set of all co-adapted couplings $(\mathbf{B}_t^a, \mathbf{B}_t^{a'})_{0 \le t \le T}$ of two Heisenberg Brownian motions starting respectively in a and a'.

The proof will be based on the following result and the use of homogeneous dilations (defined in Section 2.1):

Theorem 1.4. *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 $C > 0$,*

$$\mathbb{P}\left(\forall t \ge 0,\ d_{\mathbb{H}}(\mathbf{B}_t, \mathbf{B}'_t) \le C\right) \neq 1.$$

1.2. **Comparison with the Riemannian case.** These results show a significative difference with the Riemannian case. Indeed, on a Riemannian manifold M, it is well known (see e.g. [25] and [24]) 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

$$d(\mathbf{B}_t^a, \mathbf{B}_t^{a'}) \le e^{-(k/2)t} d(a, a') \quad \text{for all } t \ge 0,\ a \in M,\ a' \in M. \tag{5}$$

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 (5) 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. [20]) but, its behaviour with respect to couplings of co-adapted Brownian motions is therefore completely different from the case of Riemannian manifolds with curvature bounded from 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 [11, 3]. This fact is coherent with the interpretation of Theorem 1.4 as a special case of (5) where the best bound for the L^∞ control is $C = e^{-kt}$ with $k = -\infty$: There is no possible control for $t > 0$.

1.3. L^p **control for** $0 < p < \infty$**.** 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

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

and try to bound it from above uniformly in time for some well-chosen co-adapted coupling. If we denote by μ_t and ν_t the law of the processes $(\mathbf{B}_t)_{t\geq 0}$ and $(\mathbf{B}'_t)_{t\geq 0}$, we clearly have for each $t \geq 0$:

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

On the Heisenberg group, we will prove the result stronger than Theorem 1.2 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 1.5. *Let $p \geq 2$. For every $T > 0$ and every $C > 0$ there exists two points $a, a' \in \mathbb{H}$ with $a \neq a'$ such that for every co-adapted coupling $(\mathbf{B}^a_t, \mathbf{B}^{a'}_t)_{0\leq t\leq T}$, there exists $t \leq T$ such that*

$$\mathbb{E}\left[d^p_{\mathbb{H}}(\mathbf{B}^a_t, \mathbf{B}^{a'}_t)\right] > C d_{\mathbb{H}}(a, a')^p.$$

Remark 1.6. *Equivalently Theorem 1.5 can be stated as follows: Let $p \geq 2$. Let $T > 0$, then*

$$\sup_{a\neq a'\in\mathbb{H}} \inf_{\mathcal{A}_T^{(a,a')}} \sup_{0\leq t\leq T} \frac{\mathbb{E}[d^p_{\mathbb{H}}(\mathbf{B}^a_t, \mathbf{B}^{a'}_t)]^{\frac{1}{p}}}{d_{\mathbb{H}}(a, a')} = +\infty$$

where $\mathcal{A}_T^{(a,a')}$ denotes the set of all co-adapted coupling $(\mathbf{B}^a_t, \mathbf{B}^{a'}_t)_{0\leq t\leq T}$ of two Heisenberg Brownian motions starting respectively in a and a'.

As for $p = \infty$, the proof will be based on the following result and the use of dilations.

Theorem 1.7. *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,*

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

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

Theorem 1.8. *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 $p \in (0, 1)$,*

$$\sup_{t\geq 0} \mathbb{E}\left[d_{\mathbb{H}}(\mathbf{B}_t, \mathbf{B}'_t)^p\right] < +\infty. \tag{6}$$

Moreover, for the coupling by reflection, for every $p \in (0, 1)$, we also have:

$$\sup_{a\neq a'\in\mathbb{H}} \sup_{t\geq 0} \frac{\mathbb{E}[d_{\mathbb{H}}(\mathbf{B}^a_t, \mathbf{B}^{a'}_t)^p]}{d_{\mathbb{H}}(a, a')^p} < +\infty \tag{7}$$

Unfortunately, the above result is false for the reflection coupling for $p \geq 1$ (as a close look at Proposition 4.1 shows).

In the general context of co-adapted coupling, we were not able to obtain any results for $p \in [1, 2)$: we ignore whether there exist co-adapted couplings satisfying (6) or (7), 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.

The next remark recalls the situation of L^p-Wasserstein control in the case of Riemannian manifolds.

Remark 1.9. *On a Riemannian manifold M, for $p \geq 1$, the L^p version of the L^∞ control* (2)*, namely*

$$\mathcal{W}_p(\mu^a_t, \mu^{a'}_t) \leq e^{-kt/2} d(a, a') \text{ for all } t \geq 0,\ a \in M,\ a' \in M;$$

is satisfied if and only if the Ricci curvature is bounded from below by k (see [24] *and* [17, 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. Moreover in this situation, as said before, one has the existence of some appropriate Markovian coupling such that almost surely:*

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

We now turn to the second positive result. We propose an explicit coupling of the laws $\mu^a_t = \mathrm{Law}(\mathbf{B}^a_t)$ and $\mu^{a'}_t = \mathrm{Law}(\mathbf{B}^{a'}_t)$. This coupling is not at all dynamical and is made at a given fixed time t. This coupling has thus no interpretation in terms of co-adapted coupling. It provides a new proof of the case $p = 1$ in Theorem 1.1. This is the weakest result in the spectrum of L^p-Wasserstein controls ; the strongest result is for $p = \infty$. However, we stress that, apparently, our static coupling provides the first direct proof (that is, not obtained by duality) of the case $p = 1$.

Theorem 1.10. *There exists $C > 0$ such that for every $t \geq 0$ and $a = (x, y, z), a' = (x', y', z') \in \mathbb{H}$, there is a random vector (X, Y, Z, X', Y', Z') of marginals $\mu_t^a = \mathrm{Law}(\mathbf{B}_t^a) = \mathrm{Law}(X, Y, Z)$ and $\mu_t^{a'} = \mathrm{Law}(\mathbf{B}_t^{a'}) = \mathrm{Law}(X', Y', Z')$ such that*

$$\mathcal{W}_1(\mu_t^a, \mu_t^{a'}) \leq \mathbb{E}(d_{\mathbb{H}}((X, Y, Z), (X', Y', Z'))) \leq C d_{\mathbb{H}}(a, a')$$

and

$$\begin{cases} X' = X + (x' - x) & \textit{almost surely,} \\ Y' = Y + (y' - y) & \textit{almost surely.} \end{cases}$$

The idea of the coupling is the following. The goal is to construct a random vector $((X, Y, Z); (X', Y', Z'))$ of marginals μ_t^a and $\mu_t^{a'}$. First we perform a coupling of the horizontal parts (X, Y) and (X', Y') of the two Brownian laws by a simple translation. It appears that the conditional laws of the last coordinates $\mathcal{L}(Z|(X, Y))$ and $\mathcal{L}(Z'|(X', Y'))$ differ also only by a translation; but which depends on the value of (X, Y). Eventually, we use a coupling of the last coordinates which is well adapted to optimal transport for the cost $(z, z') \mapsto \sqrt{|z - z'|}$ on the real line, and is better than the simple translation. Note that our proof requires at the end an analytic estimate on the heat kernel, see (38).

We mention the interesting recent work by S. Banerjee, M. Gordina and P. Mariano [2] where the authors also 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 motions which stay at bounded distance are given Section 3. The reflection coupling on $\mathbb{H}$ is studied in Section 4. The construction of the static coupling between the Brownian laws is given in Section 5. The results are then 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

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

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 integrals 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 $\mathbb{H}$:

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

where γ 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 $\dot\gamma(t) \in \mathrm{Vect}(X(\gamma(t)), Y(\gamma(t)))$ almost surely holds. The horizontal norm $|\cdot|_{\mathrm{h}}$ is a Euclidean norm on $\mathrm{Vect}(X,Y)$ obtained by asserting that (X,Y) is an orthonormal basis of $\mathrm{Vect}(X(a),Y(a))$ at each point $a \in \mathbb{H}$. Finally the horizontal gradient $\nabla_{\mathrm{h}} f$ is $(Xf)X + (Yf)Y$.

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

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

for $\lambda > 0$. They satisfy $d_{\mathbb{H}}(\mathrm{dil}_\lambda(a), \mathrm{dil}_\lambda(a')) = \lambda d_{\mathbb{H}}(a,a')$ and, in law:

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

The distance is clearly left-invariant so that $\mathrm{trans}_p : q \in \mathbb{H} \mapsto p.q$ is an isometry for every $p \in \mathbb{H}$. In particular

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

with $e = (0,0,0)$. Another isometry is the rotation $\mathrm{rot}_\theta : (x+iy, z) \in \mathbb{C}\times\mathbb{R} \equiv \mathbb{H} \mapsto (\mathrm{e}^{i\theta}(x+iy), z)$, for every $\theta \in \mathbb{R}$. Since the explicit expression of $d_{\mathbb{H}}$ 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

$$c^{-1} d_H(a,a') \le d_{\mathbb{H}}(a,a') \le c d_H(a,a') \tag{8}$$

for some constant $c > 1$. We finally mention $d_{\mathbb{H}}((0,0,0),(x,y,0)) = \sqrt{x^2+y^2}$ and $d_{\mathbb{H}}((x,y,z),(x,y,z+h)) = 2\sqrt{\pi|h|}$.

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 built solely knowing the past of the two processes. The definition below is taken from [15, Definition 1.1.].

Definition 2.1. *Given two continuous-time Markov processes $(X^{(1)}_t)_{t\ge0}$, $(X^{(2)}_t)_{t\ge0}$, we say that $(\tilde{X}^{(1)}_t, \tilde{X}^{(2)}_t)_{t\ge0}$ is a* co-adapted coupling *of $(X^{(1)}_t)_{t\ge0}$ and $(X^{(2)}_t)_{t\ge0}$ if $\tilde{X}^{(1)}$ and $\tilde{X}^{(2)}$ are defined on the same filtered probability space $(\Omega, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$, satisfy* $\mathrm{Law}(X^{(i)}_t)_{t\ge0} = \mathrm{Law}(\tilde{X}^{(i)}_t)_{t\ge0}$ *for $i = 1,2$, and*

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

equals

$$P^{(i)}_t f : z \mapsto \mathbb{E}\left[f(X^{(i)}_{t+s}) |\, X^{(i)}_s = z\right], \quad \mathrm{Law}(X^{(i)}_s)\text{-almost surely}$$

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

If we moreover assume that the full process $(\tilde{X}^{(1)}_t, \tilde{X}^{(2)}_t)_{t\ge0}$ 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 [14, 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 $(\Omega, (\mathcal{F}_t)_{t\ge0}, \mathbb{P})$. Then, enriching the filtration if necessary, there exists a Brownian motion $(\hat{B}_t)_{t\ge0}$ defined on the same filtration $(\mathcal{F}_t)_{t\ge0}$ and independent of $(B_t)_{t\ge0}$ such that*

$$dB'(t) = J(t) dB_t + \hat{J}(t) d\hat{B}_t \tag{9}$$

where $(J_t)_{t\ge0} = ((J^{i,j}_t)_{1\le i,j\le 2})_{t\ge0}$ and $\hat{J}$ are matrices satisfying

$$JJ^T + \hat{J}\hat{J}^T = I_2 \tag{10}$$

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 2×2 real matrix $J = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)$. Then*

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

where $\le$ 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 $Q : x \in \mathbb{S}^1 \mapsto (x, J^T Jx) = \|Jx\|^2$. Therefore, Q is bounded by 1 if and only if $\|Jx\| \le 1$, for all $x \in \mathbb{S}^1$. The bound $0 \le J^T J$ is trivially satisfied. The proof is completed by $\|J\| = \sup_{x,y\in\mathbb{S}^1}(Jx, y) = \|J^T\|$. ☐

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

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

Paradoxically, it not easy to deduce $|a|, |b|, |c|, |d| \le 1$ 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 motions on $\mathbb{R}^2$ and if moreover their third coordinates satisfy

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

and

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

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

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

and thus:

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

where we used:

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

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 $(\cdot \mid \cdot)$ the usual scalar product on $\mathbb{R}^2$, we have:

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

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

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

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

Proof. The first two relations follow from the fact that Q is an orthogonal matrix. For the last relation, one can note that $J^{1,2} - J^{2,1} = \operatorname{tr}(JM) = \operatorname{tr}(KQ^T MQ)$ with M the matrix $\left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$. Now a computation gives $Q^T MQ = (\det Q)M$ 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:*

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

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

$$\langle \sqrt{2(1-K^{1,1})}dC_t, \sqrt{2(1+K^{2,2})}d\tilde{C}_t\rangle = (K^{1,2} - K^{2,1})dt. \tag{11}$$

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:*

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

where σ_R and σ_Z are defined by:

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

and

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

Note finally that the fact that σ_R and σ_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 [13] 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

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

and similarly

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

thus

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

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 (11) follows since by (10) 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:*

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

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.

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 τ 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 [10, 16].

2.4.3. *Kendall's coupling: ($K^{1,1} = \pm 1$, $K^{1,2} = K^{2,1} = 0$ and $K^{2,2} = 1$).* In [13], 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 $\{8Z_t^2 = \kappa^2 R_t^4\}$ or $\{8Z_t^2 = (\kappa - \varepsilon)^2 R_t^4\}$ (see [13, Theorem 4]), 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: $K^{1,1} = 1$, $K^{2,2} = -1$, $K^{1,2} = K^{2,1} = 0$.* We assume $R_0 > 0$. 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 [14] 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 [21, 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 1.4 and 1.7.

3.1. **Proof of Theorem 1.4.**

Theorem 1.4 is clearly a corollary of Theorem 1.7 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 1.4. 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

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

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 ≥ -1. Therefore

$$\max_{i \in \{1,2\}} \mathbb{E}\left[\int_0^t \Big((1 - K^{i,i}(s))\Big)\mathbf{1}_{\{R_s > 0\}} ds\right] \leq C^2/2 \tag{13}$$

and $R_t \leq C$, again, gives

$$\mathbb{E}\left[\int_0^t (1 - K^{2,2}(s))\frac{R_s^2}{2} ds\right] \leq C^4/4. \tag{14}$$

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

$$\mathbb{E}[Z_t^2] = Z_0^2 + \mathbb{E}\left[2\int_0^t Z_s(J^{1,2}(s) - J^{2,1}(s))ds + \int_0^t \frac{R_s^2}{2}(1 + K^{2,2}(s))ds\right] \tag{15}$$

Adding (14) and (15), and using $\mathbb{E}(Z_t^2) \leq C^4$, we obtain

$$\mathbb{E}\left[2\int_0^t Z_t(J^{1,2}(s) - J^{2,1}(s))ds + \int_0^t R_s^2 ds\right] \leq (1 + 1/4)C^4. \tag{16}$$

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 (16). 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 $\Omega \times [0,t]$):

$$\begin{aligned}\left|\mathbb{E}\left[\int_0^t Z_s(J^{1,2}(s) - J^{2,1}(s))ds\right]\right| &\leq \left(\int_0^t \mathbb{E}[Z_t^2]ds\right)^{1/2}\left(\int_0^t \mathbb{E}(J^{1,2} - J^{2,1})^2 ds\right)^{1/2} \\ &\leq \left(\int_0^t \mathbb{E}[Z_t^2]ds\right)^{1/2}\left(2\int_0^t \mathbb{E}((J^{1,2})^2 + (J^{2,1})^2)ds\right)^{1/2} \\ &\leq \sqrt{C^4 t}\left(2\int_0^t \mathbb{E}(J^{1,2})^2 + \mathbb{E}(J^{2,1})^2 ds\right)^{1/2} \\ &\leq \sqrt{C^4 t}\cdot\sqrt{2C^2} = C^3\sqrt{2t}. \end{aligned} \tag{17}$$

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 (12):

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

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

$$\int_0^t \mathbb{E}[R_s^2]ds \geq R_0^2 t. \tag{18}$$

Since $R_0 > 0$, (17) and (18) provide a contradiction in (16) as t goes to infinity. ☐

Remark 3.1. *The proof of Theorem 1.4 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* (14)*. 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* (14) *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 (14) *holds. (the bound in* (14) *can even be divided by two: $C^4/8$ in place of $C^4/4$).*

3.2. **Proof of Theorem 1.7.** To go beyond Theorem 1.4, 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. As before, the proof will be by contradiction. The principle is the following. We derive an upper bound for the drift part of Z_t similar to (17) from the proof of Theorem 1.4; and using Lemma 3.2 below, we provide a lower bound for the martingale part of Z_t.

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 β and h, the estimate*

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

implies

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

The proof of Lemma 3.2 is postponed at the end of the section.

Proof of Theorem 1.7. 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

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

which gives

$$\mathbb{E}\left(\int_0^{+\infty}[(1-J^{1,1}) + (1-J^{2,2})]ds\right) \leq \frac{C^2}{2} < +\infty. \tag{19}$$

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 non-negative drift in the stochastic differential equation of R_t, using the Dambins-Dubins-Schwarz theorem (see e.g. [22]), 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 (19) (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 ≥ -1).

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

$$-A_t = \frac{1}{2}\int_0^t (J^{1,2} - J^{2,1})ds. \tag{20}$$

Applying Cauchy–Schwarz inequality and following the same track as for (17) we obtain

$$\mathbb{E}\int_0^t |J^{1,2} - J^{2,1}|ds \leq \sqrt{t}\sqrt{\mathbb{E}\int_0^t (J^{1,2} - J^{2,1})^2 ds} \leq \sqrt{t}\sqrt{2C^2}. \tag{21}$$

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 $t > 0$,

$$\mathbb{E}(|M_{t+h} - M_t|) \geq 10C^2. \tag{22}$$

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 (21) and the lower estimate just proved give

$$8C^2 n \leq \sum_{k=1}^{n} \mathbb{E}(|A_{kh} - A_{(k-1)h}|) \leq \frac{1}{2}\mathbb{E}\int_0^{nh} |J^{1,2} - J^{2,1}| ds \leq \sqrt{\frac{C^2 hn}{2}}, \tag{23}$$

which, as n tends to ∞, 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 (22) (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 $h \geq \frac{5C^2}{q}$,

$$\mathbb{P}\left(\langle M\rangle_{t+h} - \langle M\rangle_t \geq \frac{R_0^2 h}{8}\right) \geq q - \frac{C^2}{2h} \geq \frac{9q}{10}. \tag{24}$$

We hence obtain (22) taking h large enough in (24) and applying Lemma 3.2 to $(M_{t+h} - M_t)_{h\geq 0}$.

Proof of (24): considering only the event $\{T = +\infty\}$ for the martingale part of Z_t described in Lemma 2.6, one has

$$\langle M\rangle_{t+h} - \langle M\rangle_t \geq \mathbf{1}_{\{T=+\infty\}} \left(\frac{R_0}{2}\right)^2 \int_t^{t+h} \frac{1 + K^{2,2}}{2} ds. \tag{25}$$

Now, since by (19) it holds

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

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:

$$\begin{aligned}
&\mathbb{P}\left(\mathbf{1}_{\{T=+\infty\}} \int_t^{t+h} \frac{1 + K^{2,2}}{2} ds \geq \frac{h}{2}\right) \\
=&q - \mathbb{P}\left(\mathbf{1}_{\{T=+\infty\}} \int_t^{t+h} \frac{1 - K^{2,2}}{2} ds > \frac{h}{2}\right) \geq q - \frac{C^2}{4} \cdot \frac{2}{h}.
\end{aligned} \tag{26}$$

Hence, in (25) we consider the probability that the right-hand side is greater than $(R_0/2)^2 \cdot (h/2)$, which, with (26), gives the wanted estimate (24) for every $h \geq \frac{5C^2}{q}$. □

Proof of Lemma 3.2. Let ϕ 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 $t \geq 0$,

$$\mathbb{E}\left[|N_t|\right] \geq \mathbb{E}\left[|N_{t\wedge\tau}|\right] = \mathbb{E}\left[\left|W_{\psi(t)}\right|\right]. \tag{27}$$

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

$$\mathbb{E}\left[\left|W_{\psi(h)}\right|\right] \geq \mathbb{E}\left[\left|W_{\psi(h)}\right| \cdot \mathbf{1}_A\right] = \mathbb{E}\left[|W_\beta| \cdot \mathbf{1}_A\right] \geq a_p \sqrt{\beta} \tag{28}$$

where the constant a_p is given by

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

with G a standard normal random variable and Φ its cumulative distribution function. The lower bound in (28) is obtained for $\mathbb{P}(A) = p$ and the normal random variable W_β of variance β concentrated as much as possible close to zero on event A. Equation (27) for $t = h$ and (28) finally provide the wanted estimate. □

Remark 3.3. *The major constraint for generalising Theorem 1.7 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.*

3.3. **Proof of Theorems 1.2 and 1.5.** As said before, Theorems 1.2 and 1.5 are deduced from Theorems 1.4 and 1.7 and the use of the homogeneous dilations.

For a fixed time $T > 0$ and $p \in (0, \infty]$, we introduce $C_{T,p}$ to be the constant:

$$C_{T,p} := \sup_{a \neq a' \in \mathbb{H}} \inf_{\mathcal{A}_T^{(a,a')}} \sup_{0 \leq t \leq T} \frac{\mathbb{E}[d_{\mathbb{H}}^p(\mathbf{B}_t^a, \mathbf{B}_t^{a'})]^{\frac{1}{p}}}{d_{\mathbb{H}}(a, a')} \in [0, +\infty];$$

where, as before in Remarks 1.3 and 1.6, $\mathcal{A}_T^{(a,a')}$ denotes the set of co-adpated couplings of $(\mathbf{B}_t^a)_{0\leq t\leq T}$ and $(\mathbf{B}_t^{a'})_{0\leq t\leq T}$, starting respectively in a and a'. When $p = +\infty$, the numerator is $\mathrm{essup}_{(\Omega,\mathbb{P})} d_{\mathbb{H}}(\mathbf{B}_t^a, \mathbf{B}_t^{a'})$. As noticed in these remarks we aim at proving $C_{T,p} = +\infty$ for $p \geq 2$.

The first key point is to show that, using dilations, this constant does not depend on T. For this, let $S > 0$ be another fixed time. The point is that if $(\mathbf{B}_t^a, \mathbf{B}_t^{a'})_{0\leq t\leq T}$ is a co-adapted coupling of two Heisenberg Brownian motions on $[0, T]$ starting respectively in a and a'; setting for $b \in \{a, a'\}$:

$$\tilde{\mathbf{B}}_s^{\tilde{b}} := \mathrm{dil}_{\sqrt{\frac{S}{T}}}\left(\mathbf{B}_{\frac{sT}{S}}^b\right), \text{ for } 0 \leq s \leq S;$$

then, $(\tilde{\mathbf{B}}_s^{\tilde{a}}, \tilde{\mathbf{B}}_s^{\tilde{a'}})_{0\le s\le S}$ is a co-adapted coupling of two Heisenberg Brownian motions on $[0,S]$ starting respectively in $\tilde{a} = \mathrm{dil}_{\sqrt{\frac{S}{T}}}(a)$ and $\tilde{a'} = \mathrm{dil}_{\sqrt{\frac{S}{T}}}(a')$, Moreover,

$$d_{\mathbb{H}}(\tilde{\mathbf{B}}_s^{\tilde{a}}, \tilde{\mathbf{B}}_s^{\tilde{a'}}) = \sqrt{\frac{S}{T}} d_{\mathbb{H}}\left(\mathbf{B}^a_{\frac{sT}{S}}, \mathbf{B}^{a'}_{\frac{sT}{S}}\right)$$

and

$$d_{\mathbb{H}}(\tilde{a}, \tilde{a'}) = \sqrt{\frac{S}{T}} d_{\mathbb{H}}(a, a').$$

This easily gives $C_{S,p} \le C_{T,p}$ and by symmetry of S and T: $C_{S,p} = C_{T,p}$.

We can now turn to the proof of Theorems 1.2 and 1.5. Since the proofs are similar and the case $p = +\infty$ is easier, we only treat the case $p = 2$.

Proof of Theorem 1.5. Suppose by contradiction that $C_{T_0,2} < +\infty$ for some $T_0 > 0$. The above discussion implies that $C_{T,2} < +\infty$ for each fixed time $T > 0$. Let $a = (0,0,0)$ and $a' = (x',0,0)$ with $x' \neq 0$. Thus, there exists a constant C (one can take $C = C_{T_0,2}\, d_{\mathbb{H}}(a,a')$) such that for each time $T_0 > 0$, there is a co-adapted coupling satisfying $\mathbb{E}[d^2_{\mathbb{H}}(B_t^a, B_t^{a'})] \le C$ for $t \in [0, T_0]$.

Now with the same notation as in the proof of Theorem 1.7 and denoting $q_{T_0} = \mathbb{P}(\forall 0 \le s \le T_0, R_s \ge \frac{R_0}{2})$; one has $q_{T_0} \ge q$ and as before, there exists h (independent of T_0) such that for all $0 \le t \le T_0 - h$,

$$\mathbb{E}[|M_{t+h} - M_t|] \ge 10C^2.$$

Of course this gives: $\mathbb{E}[|A_{t+h} - A_t|] \ge 8C^2$ for every $0 \le t \le T_0 - h$. Therefore equation (23) still holds if $nh \le T_0$. Since the constants C and h are independent of T_0, letting T_0 and n tend to infinity gives the contradiction. Thus $C_{T_0,2} = +\infty$. □

4. Coupling by reflection

In this section, we study precisely the coupling by reflection. We recall that $(\mathbf{B}_t)_{t\ge0}$ and $(\mathbf{B}'_t)_{t\ge0}$ are two Heisenberg Brownian motions coupled by reflection if and only if their horizontal parts $(B_t)_{t\ge0}$ and $(B'_t)_{t\ge0}$ 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 $J = \mathrm{Id}_2$ for $t \ge \tau$ where $\tau = \inf\{s \ge 0 : R_s = 0\}$ is the hitting time of 0 for $(R_t)_{t\ge0}$. 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 $\tau = \inf\{s \ge 0 : C_s = -R_0/2\}$.

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

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

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

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

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 $p = \alpha/2$, $(a+b)^p \leq a^p + b^p$, 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-Dunford-Schwarz 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 of $(R_t)_{t\geq 0}$. 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 τ is almost surely finite and that its density f_τ is given by

$$f_\tau(u) = \frac{R_0/2}{\sqrt{2\pi} u^{3/2}} e^{-\frac{R_0^2}{4u}}, \ u \geq 0. \tag{29}$$

Using τ, we compute

$$\begin{aligned} E\left[|Z_t|^p\right] &= \mathbb{E}(|W_{T(t)}|^p) \\ &= \int_0^{+\infty} \mathbb{E}(|W_{T(t)}|^p | \tau = u) f_\tau(u) du \\ &= \underbrace{\int_0^t \mathbb{E}(|W_{T(t)}|^p | \tau = u) f_\tau(u) du}_{h_1(t)} + \underbrace{\int_t^{+\infty} \mathbb{E}(|W_{T(t)}|^p | \tau = u) f_\tau(u) du}_{h_2(t)}. \end{aligned} \tag{30}$$

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 (30). Hence we set $u \leq t$. Since W and R are independent, with $c_p = \mathbb{E}(|W_1|^p)$, one has:

$$\begin{aligned} \mathbb{E}(|W_{T(t)}|^p | \tau = u) &= c_p \mathbb{E}(|T(t)|^{p/2} | \tau = u) \\ &= c_p \mathbb{E}\left(\left(\int_0^u R_s^2 ds\right)^{p/2} | \tau = u\right) \\ &= c_p 2^p u^p \mathbb{E}\left[\left(\int_0^1 \tilde{R}_\lambda^2 d\lambda\right)^{p/2} | \tau = u\right] \end{aligned} \tag{31}$$

where we have introduced the normalised process $(\tilde{R}_\lambda)_{\lambda\in[0,1]}$ (defined almost surely, since the hitting time τ 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 $u \to \infty$ to a normal positive Brownian excursion $(X_s)_{s\in[0,1]}$. Moreover, as proven in Lemma 4.3, when $u \to \infty$,

$$\mathbb{E}\left[\left(\int_0^1 \tilde{R}_\lambda^2 d\lambda\right)^{p/2} |\tau = u\right] \to \mathbb{E}\left[\left(\int_0^1 X_s^2 ds\right)^{p/2}\right]. \tag{32}$$

Finally with (31) denoting the limit in (32) by E_p, the first integral in (30) satisfies the following equivalence:

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

From the density estimate of f_τ in (29) 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,

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

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 $t \to +\infty$,

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

As a consequence, denoting by I_p the last integral,

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

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. ☐

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 $r_0 \to 0$,*

$$\mathbb{E}\left[\left(\int_0^{\alpha} \tilde{R}_\lambda^2 d\lambda\right)^{p/2}\right] \to \mathbb{E}\left[\left(\int_0^{\alpha} X_s^2 ds\right)^{p/2}\right] \tag{33}$$

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

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

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 $t > 0$

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

In particular, reorganising the terms,

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

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

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

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. □

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. [22, 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 $r_0/\sqrt{\lambda}$. □

5. The static coupling, a transport problem

In this section, we turn to the proof of Theorem 1.10. Recall that it gives a direct proof of the following L^1-Wasserstein control: There exists $C > 0$ such that for every $t \geq 0$, and every $a, a' \in \mathbb{H}$,

$$\mathcal{W}_1(\mu_t^a, \mu_t^{a'}) \leq C d_{\mathbb{H}}(a, a')$$

where $\mu_t^a = \mathrm{Law}(\mathbf{B}_t^a)$ and $\mu_t^{a'} = \mathrm{Law}(\mathbf{B}_t^{a'})$.

Proof of Theorem 1.10 . Reduction of the problem: Let us first see how, using the symmetries of $\mathbb{H}$ presented in Subsection 2.1, the proof can be reduced to $t=1$, $a=(0,0,0)$ and $a'=(x',0,0)$. First, the isometries of $\mathbb{H}$ induced isometries for $\mathcal{W}_1$. In particular $(\mathrm{trans}_p)_\#$ and $(\mathrm{rot}_\theta)_\#$ are isometries for $\mathcal{W}_1$, for every $p\in\mathbb{H}$ and $\theta\in\mathbb{R}$, that stabilise the family $\{\mu_t^a\}_{a\in\mathbb{H}}$.

$$\mathcal{W}_1(\mu_t^a,\mu_t^{a'})=\mathcal{W}_1((\mathrm{trans}_p)_\#\mu_t^a,(\mathrm{trans}_p)_\#\mu_t^{a'})=\mathcal{W}_1(\mu_t^{p.a},\mu_t^{p.a'}).$$

Hence, we can assume $a=(0,0,0)$. Using rot_θ we can moreover assume $a'=(x',0,z')$. As moreover dil_λ, defined in Subsection 2.1 satisfies $(\mathrm{dil}_\lambda)_\#\mu_t^a=\mu_{\lambda^2 t}^{\mathrm{dil}_\lambda(a)}$ we can assume $t=1$. Finally

$$\begin{aligned}\mathcal{W}_1(\mu_1^0,\mu_1^{(x',0,z')})&\le \mathcal{W}_1(\mu_1^0,\mu_1^{(x',0,0)})+\mathcal{W}_1(\mu_1^{(x',0,0)},\mu_1^{(x',0,z')})\\&\le \mathcal{W}_1(\mu_1^0,\mu_1^{(x',0,0)})+\mathcal{W}_1(\mu_1^{0},\mu_1^{(0,0,z')})\\&\le \mathcal{W}_1(\mu_1^0,\mu_1^{(x',0,0)})+d_{\mathbb{H}}((0,0,0),(0,0,z')).\end{aligned}$$

The estimate $\mathcal{W}_1(\mu_1^0,\mu_1^{(0,0,z')})\le d_{\mathbb{H}}((0,0,0),(0,0,z'))=2\sqrt{\pi|z'|}$ comes from the fact that $\mathrm{trans}_{(0,0,z')}$ is not only the left-translation but also the right-translation of vector $(0,0,z')$. If $(X,Y,Z)\sim\mu_1^0$, then $(X,Y,Z).(0,0,z')\sim\mu_1^{(0,0,z')}$ so that

$$d_{\mathbb{H}}((X,Y,Z).(0,0,z'),(X,Y,Z))=d_{\mathbb{H}}((0,0,0),(0,0,z')).$$

If we can prove $\mathcal{W}_1(\mu_1^0,\mu_1^{(x',0,0)})\le C'd_{\mathbb{H}}((0,0,0),(x',0,0))=C'|x'|$ for some $C'\ge 1$, we have finally

$$\begin{aligned}\mathcal{W}_1(\mu_1^0,\mu_1^{(x',0,z')})&\le Cd_{\mathbb{H}}((0,0,0),(x',0,0))+d_{\mathbb{H}}((0,0,0),(0,0,z'))\\&=C'|x'|+2\sqrt{\pi|z'|}\le\max(C',\sqrt{2\pi})\sqrt{(x')^2+|z'|}\\&\le Cd_{\mathbb{H}}((0,0,0),(x',0,z')),\end{aligned}$$

with $C=c\cdot\max(C',\sqrt{2\pi})$ with c defined as in (8). Finally the proof amounts to the case $t=1$, $a=(0,0,0)$ and $a'=(x',0,0)$, as we announced.

Main body of the proof: We set

$$\mu=\mu_1^0=\mathrm{Law}(X,Y,Z) \tag{34}$$
$$\nu=\mu_1^{(x',0,0)}=\mathrm{Law}\left[(x',0,0).(X,Y,Z)\right]=\mathrm{Law}(X+x',Y,Z+(1/2)x'Y) \tag{35}$$
$$\tilde\mu=\mathrm{Law}\left[(x',0,0).(X,Y,Z).(-x',0,0)\right]=\mathrm{Law}(X,Y,Z+x'Y) \tag{36}$$

We want to estimate $W_1(\mu,\nu)$ from above and start with

$$W_1(\mu,\nu)\le W_1(\mu,\tilde\mu)+W_1(\tilde\mu,\nu)$$

The coupling in (35) and (36) yields $W_1(\tilde{\mu}, \nu) \leq |x'| = d_{\mathbb{H}}((0,0,0),(x',0,0))$. The coupling suggested in (34) and (36) yields

$$\begin{aligned} W_1(\mu, \tilde{\mu}) &\leq \mathbb{E}\left[\mathbb{E}\left[d_{\mathbb{H}}((X,Y,Z),(X,Y,Z+x'Y))|(X,Y)\right]\right] \\ &\leq \mathbb{E}\left[\mathbb{E}\left[d_{\mathbb{H}}((0,0,0),(0,0,x'Y))|(X,Y)\right]\right] \\ &\leq \mathbb{E}\left[2\sqrt{\pi|Yx'|}\right] = \sqrt{|x'|} \times [2\sqrt{\pi}\,\mathbb{E}(\sqrt{|Y|})] \end{aligned} \tag{37}$$

The order $\sqrt{|x'|}$ is such that $\sqrt{|x'|}/d_{\mathbb{H}}((x',0,0),(0,0,0)) = |x'|^{-1/2} \to \infty$ as x' goes to zero.

Remark 5.1. *Note that the coupling in* (37) *is exactly the synchronous coupling of subsection 2.4.1 evaluated at time* $t = 1$. *The estimate from above can easily be checked with Lemma 2.6.*

We modify the computation above just a little based on the knowledge that the translation is not the optimal transport plan on the real line when considering costs that are increasing concave functions of the distance.

The following lemma will be in order:

Lemma 5.2. *If* η *is a probability measure on* $\mathbb{R}$ *with rapidly decreasing and smooth density* f*, then*

$$\inf_{\pi \in \Pi(\eta, (\mathrm{trans}_t^{\mathbb{R}})_\# \eta)} \iint \sqrt{|y-x|}\, d\pi \leq |t| \times \left(\int |f'(x)|\sqrt{|x|}\, dx\right)$$

Proof. The left-hand side is the 1-Wasserstein distance $\mathcal{W}_1^{\mathbb{R}}(\eta, (\mathrm{trans}_t^{\mathbb{R}})_\#\eta)$ on $\mathbb{R}$ for the distance $(x,y) \mapsto \sqrt{|y-x|}$. This is also the Kantorovich norm $\|\eta - (\mathrm{trans}_t^{\mathbb{R}})_\#\eta)\|_1$, in the sense of [12]. Recall that the Kantorovich norm of a signed Radon measure $\sigma = \sigma_+ - \sigma_-$ of mass zero is

$$\|\sigma\|_1 = \mathcal{W}_1^{\mathbb{R}}(\sigma_+, \sigma_-) = \inf_{\pi \in \Pi(\sigma_+, \sigma_-)} \iint \sqrt{|y-x|} d\pi(x,y)$$

where the mass of σ_+ can be different from 1. Let $\sigma = \eta - (\mathrm{trans}_t^{\mathbb{R}})_\#\eta)$ and for the computation of $\|\sigma\|_1$ assume without loss of generality that $t \geq 0$. We have $\sigma = \int_0^t (\mathrm{trans}_u^{\mathbb{R}})_\#\gamma\, du$ where γ is the Radon measure of density f'. Therefore

$$\begin{aligned} \|\sigma\|_1 = \left\| \int_0^t (\mathrm{trans}_u^{\mathbb{R}})_\#\gamma\, du \right\| &\leq \int_0^t \|(\mathrm{trans}_u^{\mathbb{R}})_\#\gamma\|_1\, du \\ &= \int_0^t \|\gamma\|_1\, du \ = t\|\gamma\|_1 \\ &\leq t \times \left(\int |f'(x)|\sqrt{|x|}\, dx\right). \end{aligned}$$

Note that for the last inequality, we use the triangle inequality by transporting γ_+, of density f'_+, to the atomic measure $(\int f'_+)\delta_0$ and then from this measure to γ_-, of density f'_-. $\square$

We can now conclude by coupling Z and Z', conditionally on (X, Y). From (37):

$$\mathbb{E}\left[d_{\mathbb{H}}((X,Y,Z),(X,Y,Z+x'Y))|(X,Y)\right] \le 2\sqrt{\pi}|x'Y| \times \left(2\int |f'_{Z|X,Y}(z)|\sqrt{|z|}\,dz\right)$$

where $f_{Z|x,y}$ is the density of $\mathrm{Law}(Z|\,X = x,\ Y = y)$. Finally,

$$\begin{aligned}
&\mathbb{E}\left[d_{\mathbb{H}}((X,Y,Z),(X,Y,Z+x'Y))\right]\\
&\le |x'| \iint f_{X,Y}(x,y) 2\sqrt{\pi}|y| \times \left(\int |f'_{Z|x,y}(z)|\sqrt{|z|}\,dz\right) dx\,dy\\
&\le |x'| \iiint 2\sqrt{\pi}|y|\sqrt{|z|}\left(\frac{|\partial_z f_{X,Y,Z}(x,y,z)|}{f_{X,Y,Z}(x,y,z)}\right) f_{X,Y,Z}(x,y,z)\,dx\,dy\,dz.
\end{aligned}$$

Here $f_{X,Y,Z}$ is the density of $\mu = \mathrm{Law}(X,Y,Z)$ with respect to the Lebesgue measure and $f_{X,Y}$ is the density of (X,Y). Note that we have $f_{Z|x,y}(z) = f_{X,Y,Z}(x,y,z)/f_{X,Y}(x,y)$.

Since $f_{X,Y,Z}$ is the density of the heat kernel, it is well known (see [19] and [1] equation (14)) that there exists a constant $C > 0$ such that for all $(x,y,z) \in \mathbb{H}$,

$$\left|\frac{\partial_z f_{X,Y,Z}(x,y,z)}{f_{X,Y,Z}(x,y,z)}\right| \le C. \tag{38}$$

The proof of this fact is analytic and is based on the explicit representation of the heat kernel as an oscillatory intergal.

The proof of Theorem 1.10 is then finished since the quantity $|y|\sqrt{|z|}$ is clearly integrable with respect to the heat kernel $f_{X,Y,Z}$.

□

6. Generalisation to the Heisenberg groups of higher dimension.

In this section, we prove that Theorems 1.2, 1.4, 1.5, 1.7, 1.8 and 1.10 also hold in the case of the Heisenberg groups of higher dimension. For $n \ge 1$ the Heisenberg group $\mathbb{H}_n$ can be identified with $\mathbb{R}^{2n+1}$ equipped with the law:

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

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

$$\mathbf{B}_t^0 := \left(\left(B^1_{t,i}, B^2_{t,i}\right)_{i=1,\dots,n}, \frac{1}{2}\sum_{i=1}^n\left(\int_0^t B^1_{s,i}dB^2_{s,i} - \int_0^t B^2_{s,i}dB^1_{s,i}\right)\right)$$

where $B_t := \left(B^1_{t,i}, B^2_{t,i}\right)_{i=1,\dots,n}$ is a $2n$-dimensional standard Brownian motion. As before, we denote by $(\mathbf{B}^a_t)_{t\ge 0}$ the Brownian motion starting in $a \in \mathbb{H}_n$. It can be represented as $a\cdot\mathbf{B}^0_t$ and we write μ^a_t for its law at time t.

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

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

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

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

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

and one can directly adapt the proof of Theorems 1.4 and 1.7 to this setting. Similarly as before, one can deduce that Theorems 1.2 and 1.5 are also satisfied for higher dimensional Heisenberg groups.

Generalisation of Theorem 1.8: 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 1.8 also holds for $\mathbb{H}_n$.

Generalisation of Theorem 1.10: We turn to the static coupling. As before, using the symmetries and the dilatation of the higher dimensional Heisenberg groups, it suffices to study the case $t = 1$ and $a = 0$, $a' = ((x', 0), (0, 0), \dots, (0, 0), 0)$. If the vector $V = ((X_1, Y_1), \dots (X_d, Y_d), Z)$ has law μ_1^0, then the vector $a'.V =$

$(X_1 + x', Y_1), (X_2, Y_2), \dots (X_d, Y_d), Z + \frac{1}{2}x'Y_2)$ has law $\mu_1^{a'}$. An analogue coupling as the one in Section 5 can be defined. The horizontal part is translated by $((x,' 0), \dots, (0, 0))$ and conditionally on $((X_1, Y_1), \dots (X_d, Y_d))$, we perform a coupling between the law of Z and the law of $Z + \frac{1}{2}x'Y_2$ as described in Lemma 5.2. Recall that it is adapted to the non-convex transport cost $(z, z') \mapsto \sqrt{|z - z'|}$.

Since the heat kernel estimate corresponding to (38) also holds in higher dimension (see [9]), the proof finishes analogously to the one in $\mathbb{H}_1$. Therefore Theorem 1.10 is satisfied for higher dimensional Heisenberg groups too.

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] N. Eldredge. Precise estimates for the subelliptic heat kernel on H-type groups. *J. Math. Pures Appl. (9)*, 92(1):52–85, 2009.

[10] E. P. Hsu and K.-T. Sturm. Maximal coupling of Euclidean Brownian motions. *Commun. Math. Stat.*, 1(1):93–104, 2013.

[11] N. Juillet. Geometric inequalities and generalized Ricci bounds in the Heisenberg group. *International Mathematical Research Notices*, 2009, 2009.

[12] L. V. Kantorovič and G. S. Rubinšteĭn. On a space of completely additive functions. *Vestnik Leningrad. Univ.*, 13(7):52–59, 1958.

[13] W. S. Kendall. Coupling all the Lévy stochastic areas of multidimensional Brownian motion. *Ann. Probab.*, 35(3):935–953, 2007.

[14] W. S. Kendall. Brownian couplings, convexity, and shy-ness. *Electron. Commun. Probab.*, 14:66–80, 2009.

[15] 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.

[16] K. Kuwada. On uniqueness of maximal coupling for diffusion processes with a reflection. *J. Theoret. Probab.*, 20(4):935–957, 2007.

[17] K. Kuwada. Duality on gradient estimates and Wasserstein controls. *J. Funct. Anal.*, 258(11):3758–3774, 2010.

[18] K. Kuwada and K.-T. Sturm. A counterexample for the optimality of Kendall-Cranston coupling. *Electron. Comm. Probab.*, 12:66–72, 2007.

[19] 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.

[20] R. Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.

[21] M. N. Pascu and I. Popescu. Shy and fixed-distance couplings of Brownian motions on manifolds. *Stochastic Process. Appl.*, 126(2):628–650, 2016.

[22] 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.

[23] M.-K. von Renesse. Intrinsic coupling on Riemannian manifolds and polyhedra. *Electron. J. Probab.*, 9:no. 14, 411–435, 2004.

[24] 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.

[25] 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