Electronic Journal of Probability

Electron. J. Probab. **0** (2023), article no. 0, 1–1.
ISSN: 1083-6489 https://doi.org/10.1214/YY-TN

A coupling strategy for Brownian motions at fixed time on Carnot groups using Legendre expansion

Marc Arnaudon* Magalie Bénéfice* Michel Bonnefont* Delphine Féral*

Abstract

We propose a new simple construction of a coupling at a fixed time of two sub-Riemannian Brownian motions on the Heisenberg group and on the free step 2 Carnot groups. The construction is based on a Legendre expansion of the standard Brownian motion and of the Lévy area. We deduce sharp estimates for the decay in total variation distance between the laws of the Brownian motions. Using a change of probability method, we also obtain the log-Harnack inequality, a Bismut type integration by part formula and reverse Poincaré inequalities for the associated semi-group.

Keywords: Brownian motion ; Carnot group ; Non co-adapted coupling ; Total variation distance ; log-Harnack inequality ; Reverse Poincaré inequality.
MSC2020 subject classifications: Primary 60G17 ; 60H30 ; 60H10 ; 60J60, Secondary 53C17.
Submitted to EJP on July 15, 2024, final version accepted on April 7, 2025.

1 Introduction

Recently, the study of successful couplings for Brownian motion on sub-Riemannian manifolds has received a lot of attention. In the case of the examples discussed below, the sub-Riemannian Brownian motion consists in a Riemannian Brownian motion on a base manifold together with its swept area. The construction of successful couplings is thus a challenging question since one has to couple the Riemannian Brownian motions on the base manifold in such a way that also their swept area meet. The first construction of successful couplings on the Heisenberg group or on the free step 2 Carnot groups were obtained by Ben Arous, Cranston and Kendall [8] and Kendall [18, 19].

These first couplings were Markovian couplings or at least co-adapted couplings. A main progress was made by Banerjee, Gordina and Mariano in [3] where they constructed a non co-adapted successful coupling on the Heisenberg group $\mathbb{H}$. Their coupling is sometimes called a finite look-ahead coupling since they repeat some Brownian bridges couplings with the use of the future values of one stochastic process. This kind of finite look-ahead coupling was already proposed by Banerjee and Kendall [4] in some different hypoelliptic context: the Kolmogorov diffusion; i.e., a Brownian motion on $\mathbb{R}$ and its (iterated) time integral.

*Univ. Bordeaux, CNRS, Bordeaux INP, IMB, UMR 5251,Talence, F-33400, France. E-mail: marc.arnaudon@math.u-bordeaux.fr

The second named author Bénéfice extended the co-adapted Kendall's coupling to the case of the curved sub-Riemannian manifold $SU(2)$ in [12] and the non co-adapted coupling of Banerjee, Gordina and Mariano to the cases of $SU(2)$ and $SL(2,\mathbb{R})$ in [11] and of higher dimensional Carnot groups in [13]. Another interesting non co-adapted coupling on $\mathbb{H}, SU(2)$ and the universal covering of $SL(2,\mathbb{R})$ was given recently by Luo and Neel in [22].

Successful couplings are interesting in themselves but have also a lot of analytical consequences for the regularization of the associated semi-group and for the study of the associated harmonic functions.

The construction of the finite look ahead successful coupling in [3] is not so easy. The main contribution of the present work is to propose a simpler construction for the coupling of two sub-Riemannian Brownian motions starting from different points but only at a fixed time. We will consider the case of the Heisenberg group and its extension to the Carnot group case. Our construction is based on a Legendre expansion of the standard Brownian motion which, as it was noticed by Kuznetsov [20], is more adapted to the computation of the Lévy area than the Karhunen Loève expansion used in [3, 13], see Lemmas 2.1 and 2.2.

We will see that even if the coupling is only given for a fixed time and thus is not really a successful coupling, we can still deduce some important regularization properties for the associated semi-group. Note that in principle, similarly as the construction in [3] for the Heisenberg group, it should be possible to iterate the procedure when our coupling fails and this should provide at the end a truly successful coupling. This point will not be investigated in the present paper. Note however that the second author in a different work [13] provides explicitly a successful coupling for Carnot groups with a construction closer to the original one in [3].

A first direct application of successful couplings is total variation distance estimates between the laws of two Markov processes. This comes from the Aldous inequality which writes

$$d_{TV}\left(\mu_t^x, \mu_t^{\tilde{x}}\right) \leq \mathbb{P}(X_t^x \neq X_t^{\tilde{x}}) \tag{1.1}$$

for any coupling $(X_t^x, X_t^{\tilde{x}})$ and with $\mu_t^x = \mathcal{L}(X_t^x)$ and $\mu_t^{\tilde{x}} = \mathcal{L}(X_t^{\tilde{x}})$.

For example, in the case of the standard Brownian motion on $\mathbb{R}^n$, it is well known (see e.g. [21]) that the reflection coupling on $\mathbb{R}^n$ provides an equality in the Aldous inequality (1.1) and thus:

$$d_{\mathrm{TV}}\left(\mu_t^x, \mu_t^{\tilde{x}}\right) = \mathbb{P}\left(\tau_{\frac{1}{2}|\tilde{x}-x|} > t\right) \tag{1.2}$$

where $\mu_t^x = \mathcal{N}(x,t)$ is the law of the standard Brownian motion starting in x on $\mathbb{R}^n$ and with $\tau_{\frac{1}{2}|\tilde{x}-x|}$ the hitting time of $\frac{1}{2}|\tilde{x}-x|$ for a standard Brownian motion on $\mathbb{R}$ starting in 0. In particular it provides the following regularization of the standard heat semi-group with a (polynomial) decay: let $t > 0$, $x, \tilde{x} \in \mathbb{R}^n$, then

$$d_{\mathrm{TV}}\left(\mu_t^x, \mu_t^{\tilde{x}}\right) \leq \frac{\|\tilde{x}-x\|}{\sqrt{2\pi t}}. \tag{1.3}$$

Note that such a decay holds for the total variation distance but does not hold for the usual p-Wasserstein distances with $p \geq 1$; indeed for the standard Brownian motion on $\mathbb{R}^n$, it is easily seen by Jensen inequality that the optimal coupling is obtained by the translation coupling and one has always for any $p \geq 1$ and any $t \geq 0$:

$$W_p\left(\mu_t^x, \mu_t^{\tilde{x}}\right) = \|\tilde{x}-x\|.$$

Below, let us denote $\mu_t^{x_1,x_2,z}$ to be the law of the sub-elliptic Brownian motion on the Heisenberg group starting from (x_1, x_2, z). The first main result of the present paper is the extension of the total variation estimate (1.2) to the case of the Heisenberg group and to the case of the free step 2 Carnot groups. We state it below in Theorem 1.1 for the Heisenberg group. The generalization to the case of the Carnot groups is given in Theorem 3.2. The case of the Heisenberg group result already appears in [3]. Another improvement here is that we obtain explicit constants.

Theorem 1.1. There exist two constants $C_1, C_2 \geq 0$ such that for all $t \geq 0$ and all (x_1, x_2, z) and $(\tilde{x}_1, \tilde{x}_2, \tilde{z})$ in $\mathbb{H}$,

$$d_{\mathrm{TV}}\left(\mu_t^{(x_1,x_2,z)}, \mu_t^{(\tilde{x}_1,\tilde{x}_2,\tilde{z})}\right) \leq C_1 \frac{\|(\tilde{x}_1 - x_1, \tilde{x}_2 - x_2)\|_2}{\sqrt{t}} + C_2 \frac{|\tilde{z} - z - \frac{1}{2}(x_1\tilde{x}_2 - x_2\tilde{x}_1)|}{t}. \tag{1.4}$$

Moreover:

$$C_2 = \frac{5\sqrt{21}}{\pi\sqrt{2\pi}} \text{ and } C_1 = \frac{1}{\sqrt{2\pi}} + \sqrt{\frac{2}{3\pi}} C_2 = \frac{1}{\sqrt{2\pi}}\left(1 + \frac{5\sqrt{14}}{\pi\sqrt{\pi}}\right).$$

As noticed in [3], Theorem 1.1 provides the sharp order of decay. In this sense, the associated coupling is called *efficient*. It is also noted in [3] that any Markovian or co-adapted coupling can not reach the sharp estimate when the initial points are in the same fiber, i.e., when $(\tilde{x}_1, \tilde{x}_2) = (x_1, x_2)$. The coupling proposed in [22] is even actually *maximal* when the initial points are in the same fiber; i.e., similarly to the reflection coupling in $\mathbb{R}^n$, it produces an equality in the Aldous inequality (1.1).

The second main type of application of successful couplings are gradient estimates for the associated semi-group and for harmonic functions. A direct application of the total variation estimates first leads to the following L^∞ gradient bounds. It is stated here for the Heisenberg group. The case of the free step 2 Carnot groups will be given in Corollary 4.1.

Corollary 1.2. For any bounded measurable function f on $\mathbb{H}$, and any $t > 0$:

$$\|\nabla_{\mathfrak{h}} P_t f\|_\infty \leq \frac{2C_1}{\sqrt{t}} ||f||_\infty \tag{1.5}$$

and

$$\|Z P_t f\|_\infty \leq \frac{2\sqrt{2}C_2}{t} ||f||_\infty. \tag{1.6}$$

In order to obtain stronger gradient inequalities, we may use an equivalent change of probability technique. The idea is to construct couplings with probability one at a given fixed time of the two processes. The price to pay will be to make changes of probabilities for one of the process. The distance between semigroups will be measured by this change of probability. The method can also be compared to finite dimensional Malliavin calculus in the Legendre polynomial basis. We first derive a log-Harnack inequality for the semi-group, see Theorem 4.6. We then establish in Theorem 4.7 a Bismut type formula; i.e., an integration by parts formula for the derivative of the semi-group. We deduce some reverse Poincaré inequalities for $p > 1$, see Theorem 4.8 and a weak reverse log-Sobolev inequality, see Corollary 4.9. In a different hypoelliptic setting, a change of probability method was also investigated by Guillin and Wang [16] to study some kinetic Fokker-Planck equation.

Another approach to obtain these gradient estimates is through the generalized curvature-dimension criterion developed by Baudoin and Garofalo [7]. Step 2 Carnot groups are examples of non-negatively curved sub-Riemannian manifolds with transverse symmetries and thus a reverse log-Sobolev is known to hold, see Proposition 3.1 in [5]. See also [6] where the reverse Poincaré inequality and its constant is studied on general Carnot groups by analytic methods. A general stochastic method which also provides local estimates can be found in [1], but the constants are not explicit.

In order to enlighten the simplicity of the method, we chose to present first the construction of the coupling and the total distance variation estimate in the case of the Heisenberg group and to investigate only in a second time the case of the higher dimensional step 2 Carnot groups on $\mathbb{R}^n$, $n \geq 3$. The reason is that some small complications arise for the Carnot groups. The sub-Riemannian Brownian motion consists of n independent 1-dimensional standard Brownian motions together with all their $\frac{n(n-1)}{2}$ Lévy areas. The main difference is that in this situation the vertical space is not anymore 1-dimensional. It is identified with $\mathfrak{so}(n)$ and we have used some Wishart matrices to get a solution of Equation (3.5); see Proposition 3.1.

The outline of the paper is the following. In Section 2, we quickly describe the Heisenberg group and its sub-Riemannian Brownian motion. We then describe their nice expansion with the use of the orthogonal Legendre polynomials. Finally, we provide the proof of Theorem 1.1 for the total variation distance on the Heisenberg group. The aim of Section 3 is to extend the result to the case of the higher dimensional free step 2 Carnot groups. This is done in Theorem 3.2. Section 4 is devoted to the gradient estimates. We first prove the L^∞ gradient estimates of Corollary 1.2 and of Corollary 4.1. We then turn to the change of probability method. We first obtain a log-Harnack inequality for the semi-group in Theorem 4.6. Finally, we provide the Bismut type formula in Theorem 4.7, its application in term of reverse Poincaré inequalities in Theorem 4.8 and reverse weak log-Sobolev inequality in Corollary 4.9. We finally deduce estimates of the gradient on the heat kernel in Corollary 4.10.

2 Description of the Brownian motion on $\mathbb{H}$

2.1 The Heisenberg group

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

$$(x_1,x_2,z)\star(x_1',x_2',z') = \left(x_1+x_1', x_2+x_2', z+z'+\frac{1}{2}(x_1x_2'-x_2x_1')\right).$$

For our purpose, it will be convenient to identify sometimes $\mathbb{R}^3$ with $\mathbb{R}^2\times\mathbb{R}$ and to write the law as

$$(x,z)\star(x,z') = \left(x+x', z+z'+\frac{1}{2}x\cdot x'\right),$$

where for $x=(x_1,x_2)$, $x'=(x_1,x_2)$,

$$x\cdot x' = x_1x_2' - x_2x_1'.$$

The left invariant vector fields are given by

$$\begin{cases} X_1(f)(x_1,x_2,z) = \frac{d}{dt}_{|t=0} f((x_1,x_2,z)\star(t,0,0)) = \left(\partial_{x_1} - \frac{x_2}{2}\partial z\right) f(x_1,x_2,z) \\ X_2(f)(x_1,x_2,z) = \frac{d}{dt}_{|t=0} f((x_1,x_2,z)\star(0,t,0)) = \left(\partial_{x_2} + \frac{x_1}{2}\partial z\right) f(x_1,x_2,z) \\ Z(f)(x_1,x_2,z) = \frac{d}{dt}_{|t=0} f((x_1,x_2,z)\star(0,0,t)) = \partial_z f(x_1,x_2,z). \end{cases}$$

Note that $[X_1,X_2]=Z$ and that Z commutes with X_1 and X_2. The vectors fields X_1,X_2 are called the horizontal vector field whereas Z is called the vertical vector field.

2.2 The subRiemannian Brownian motion on Heisenberg

The standard (half) sub-Laplacian on the Heisenberg group is given by $L=\frac{1}{2}(X_1^2+X_2^2)$. This is a diffusion operator and it 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_1,x_2,z):

$$\begin{aligned} \mathbb{B}_t^{(x_1,x_2,z)} &:= (x_1,x_2,z)\star\left(B_t^1, B_t^2, \frac{1}{2}\left(\int_0^t B_s^1 dB_s^2 - \int_0^t B_s^2 dB_s^1\right)\right) \\ &= \left(x_1+B_t^1, x_2+B_t^2, z+\frac{1}{2}(x_1B_t^2-x_2B_t^1)+\frac{1}{2}\left(\int_0^t B_s^1 dB_s^2 - \int_0^t B_s^2 dB_s^1\right)\right) \end{aligned}$$

where $(B_t^1,B_t^2)_{t\geq 0}$ is a standard Brownian motion on $\mathbb{R}^2$.

It is easily seen that $(\mathbb{B}_t)_{t\geq 0}$ is a continuous process with independent and stationary increments on $\mathbb{H}$; the increments being taken with respect to the group law and thus of the form: $\mathbb{B}_s^{-1}\star\mathbb{B}_t$ for $0\leq s\leq t$. We simply call it the Heisenberg Brownian motion.

The quantity

$$A_t = \frac{1}{2}\left(\int_0^t B_s^1 dB_s^2 - \int_0^t B_s^2 dB_s^1\right) \tag{2.1}$$

is called the Lévy area of the 2-dimensional Brownian motion.

Identifying $\mathbb{R}^3$ with $\mathbb{R}^2 \times \mathbb{R}$ again, we will write $\mathbb{B}_t^{(x,z)} = (X_t, z_t)$.

2.3 The Carnot-Carathéodory distance

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

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

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 to be horizontal if it is absolutely continuous and $\dot{\gamma}(t) \in \mathrm{Span}(X_1(\gamma(t)), X_2(\gamma(t)))$ almost surely holds. The horizontal norm $|\cdot|_{\mathfrak{h}}$ is an Euclidean norm on $\mathrm{Span}(X_1, X_2)$ obtained by asserting that (X_1, X_2) is an orthonormal basis of $\mathrm{Span}(X_1(a), X_2(a))$ at each point $a \in \mathbb{H}$. Finally the horizontal gradient $\nabla_{\mathfrak{h}} f$ is $(X_1 f)X_1 + (X_2 f)X_2$.

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

$$\mathrm{dil}_\lambda(x_1, x_2, z) = (\lambda x_1, \lambda x_2, \lambda^2 z)$$

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:

$$\mathrm{dil}_{\frac{1}{\sqrt{t}}}\left(X_t^1, X_t^2, \frac{1}{2}\left(\int_0^t X_s^1 dX_s^2 - \int_0^t X_s^2 dX_s^1\right)\right) \overset{\mathrm{Law}}{=} \left(X_1^1, X_1^2, \frac{1}{2}\left(\int_0^1 X_s^1 dX_s^2 - \int_0^1 X_s^2 dX_s^1\right)\right).$$

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

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

with $e = (0, 0, 0)$. Another isometry is the rotation $\mathrm{rot}_\theta : (x_1 + ix_2, z) \in \mathbb{C} \times \mathbb{R} \equiv \mathbb{H} \mapsto (\mathrm{e}^{i\theta}(x_1 + ix_2), z)$, 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 homogeneous quasinorm (still in the sense that the triangle inequality only holds up to a multiplicative constant). We will use

$$H : a = (x_1, x_2, z) \in \mathbb{H} \mapsto \sqrt{x_1^2 + x_2^2 + |z|} \in \mathbb{R},$$

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

$$c^{-1} d_H(a, a') \leq d_{\mathbb{H}}(a, a') \leq c d_H(a, a') \tag{2.2}$$

for some constant $c > 1$.
We finally mention $d_{\mathbb{H}}((0, 0, 0), (x_1, x_2, 0)) = \sqrt{x_1^2 + x_2^2}$ and $d_{\mathbb{H}}((x_1, x_2, z), (x_1, x_2, z + h)) = 2\sqrt{\pi |h|}$.

2.4 The description of the Brownian motion on $\mathbb{H}$ with Legendre polynomials

Let $T > 0$ and consider the scalar product defined for $f, g \in \mathcal{C}([0, T], \mathbb{R})$ by

$$\langle f, g \rangle = \int_0^T f(t) g(t) dt.$$

Take Q_k^T to be the associated normalized orthogonal polynomials; i.e., such that $\|Q_k^T\|^2 = 1$. By dilation and translation, one sees that

$$Q_k^T(x) = \sqrt{\frac{2}{T}} P_k \left(-1 + \frac{2x}{T}\right)$$

where $(P_k)_k$ are the standard (normalized) Legendre polynomials, which are orthogonal for the Lebesgue measure on $[-1, 1]$.

We first consider the following representation of a standard one-dimensional Brownian motion $(B_t)_{0\le t\le T}$ starting in 0. This representation is close to the standard Karhunen-Loève decomposition of the Brownian motion. The interest in considering Legendre polynomials is that it is well adapted to the computation of the Lévy area. To our knowledge, this has been done firstly in Biane and Yor [9]. Recently, Legendre polynomials have also been considered for the Brownian motion in Foster, Lyons and Oberhauser [15] and Habermann [17] and for the Lévy area in Kuznetsov [20] and Foster, Habermann [14].

Lemma 2.1. Let $(\xi_k)_{k\ge 1}$ be a sequence of independent and identically distributed random variables of law $\mathcal{N}(0, 1)$. Define

$$B_t = \sum_{k\ge 0} \xi_k \int_0^t Q_k^T(s)ds,\ 0 \le t \le T. \tag{2.3}$$

Then the process $(B_t)_{0\le t\le T}$ is a standard Brownian motion on $[0, T]$.

The proof is done in [9], but let us recall the main ideas for the reader's convenience.

Proof. Let $T \ge 0$ and let $(B_t)_{0\le t\le T}$ be defined by (2.3). The process $(B_t)_{0\le t\le T}$ is clearly a centered Gaussian process. To prove it is a standard Brownian motion, compute its covariance: for $0 \le s, t \le T$

$$\begin{aligned}\mathbb{E}[B_t B_s] &= \sum_{k\ge 0} \left(\int_0^t Q_k^T(u)du\right)\left(\int_0^s Q_k^T(u)du\right)\\ &= \sum_{k\ge 0} \langle \mathbb{1}_{[0,t]}, Q_k^T\rangle \langle \mathbb{1}_{[0,s]}, Q_k^T\rangle\\ &= \langle \mathbb{1}_{[0,t]}, \mathbb{1}_{[0,s]}\rangle = s \wedge t.\end{aligned}$$

where $\langle\cdot,\cdot\rangle$ denotes the usual scalar product on $L^2([0, T])$. The result follows. □

We turn to the computation of the Lévy area.

Lemma 2.2. Let $(\xi_k)_{k\ge 0}$ be a sequence of independent and identically distributed random vectors with common law $\mathcal{N}(0, I_2)$, where I_2 denotes the identity matrix on $\mathbb{R}^2$. Write $\xi_k = (\xi_k^1, \xi_k^2)^t$ and for $0 \le t \le T$ and $i = 1, 2$ let

$$B_t^i = \sum_{k\ge 0} \xi_k^i \int_0^t Q_k^T(s)ds. \tag{2.4}$$

Then $(B_t^1, B_t^2)_{0\le t\le T}$ is a standard two-dimensional Brownian motion and its associated Lévy area $A_t := \frac{1}{2}\Big(\int_0^t B_s^1 dB_s^2 - \int_0^t B_s^2 dB_s^1\Big)$ at the given time T may be written as

$$A_T = T \sum_{k\ge 0} \alpha_k\, \xi_k \cdot \xi_{k+1} \tag{2.5}$$

with

$$\alpha_k = \frac{1}{2\sqrt{4(k+1)^2 - 1}} = \frac{1}{2\sqrt{(2k+1)(2k+3)}}, \quad k \ge 0. \tag{2.6}$$

As before the proof is done in [9], but let us recall the main ideas of the proof for the reader's convenience.

Proof. With the notation of Lemma 2.2,

$$\int_0^T B_s^1 dB_s^2 = \sum_{k,l\geq 0} \xi_k^1 \xi_l^2 c_{k,l} \text{ with } c_{k,l} = \int_0^T \left(\int_0^t Q_k^T(s)ds \right) Q_l^T(t)dt.$$

Now by integration by parts, for $k, l \geq 0$,

$$c_{k,l} = \left(\int_0^T Q_k^T(u)du \right) \left(\int_0^T Q_l^T(u)du \right) - c_{l,k}.$$

Since Q_k^T is a family of orthogonal polynomials, one infers that for $(k,l) \neq (0,0)$, $c_{k,l} = -c_{l,k}$ and thus

$$c_{k,l} = 0 \text{ if } |k-l| \geq 2 \text{ or } k = l \geq 1.$$

Therefore

$$\int_0^T B_s^1 dB_s^2 = c_{0,0}\xi_0^1\xi_0^2 + \sum_{k\geq 0} c_{k,k+1}(\xi_k^1\xi_{k+1}^2 - \xi_{k+1}^1\xi_k^2).$$

and thus the Lévy area at the final time T writes:

$$A_T = \sum_{k\geq 0} c_{k,k+1}(\xi_k^1\xi_{k+1}^2 - \xi_{k+1}^1\xi_k^2).$$

The result follows by an explicit computation of the constant $c_{k,k+1}$. □

We recall that here in the case of the Heisenberg group:

$$\xi_k \cdot \xi_{k+1} = \xi_k^1\xi_{k+1}^2 - \xi_{k+1}^1\xi_k^2. \tag{2.7}$$

As a direct application of Lemma 2.2, the Brownian motion on $\mathbb{H}$ starting in (x_1, x_2, z) at time T may be represented by

$$\mathbb{B}_T^{(x_1,x_2,z)} = \begin{pmatrix} x_1 + \sqrt{T}\xi_0^1 \\ x_2 + \sqrt{T}\xi_0^2 \\ z + \frac{\sqrt{T}}{2}(x_1\xi_0^2 - x_2\xi_0^1) + T\sum_{k\geq 0}\alpha_k\,(\xi_k^1\xi_{k+1}^2 - \xi_{k+1}^1\xi_k^2) \end{pmatrix}$$

or equivalently with $x = (x_1, x_2)$,

$$\mathbb{B}_T = \begin{pmatrix} x + \sqrt{T}\xi_0 \\ z + \frac{\sqrt{T}}{2} x\cdot\xi_0 + T\sum_{k\geq 0}\alpha_k\,\xi_k\cdot\xi_{k+1} \end{pmatrix}.$$

2.5 Proof of Theorem 1.1

Before we turn to the proof of Theorem 1.1, we recall (1.3) the standard estimate for Gaussian vectors on $\mathbb{R}^d$, $d \geq 1$ (with the same identity covariance matrix).

Lemma 2.3. Let $d \geq 1$ an integer, $m, m' \in \mathbb{R}^d$, there exists a random couple (X, Y) whose marginals are Gaussian random variables $\mathcal{N}(m, I_d)$ and $\mathcal{N}(m', I_d)$ and such that

$$\mathbb{P}(X \neq Y) \leq \left(\frac{\|m - m'\|_2}{\sqrt{2\pi}} \right) \wedge 1.$$

We provide just below a proof of Theorem 1.1 with slightly weaker explicit constants C_1 and C_2. The slightly improved constants will be obtained in Remark 2.

Proof of Theorem 1.1. For any choice of two Heisenberg valued subRiemannian Brownian motions $((X_t, z_t))_{t\geq 0}$ and $((\tilde{X}_t, \tilde{z}_t))_{t\geq 0}$ started respectively at (x, z) and $(\tilde{x}, \tilde{z})$, we have

$$d_{TV}\left(\mu_T^{(x,z)}, \mu_T^{(\tilde{x},\tilde{z})}\right) \leq \mathbb{P}\left((X_T, z_T) \neq (\tilde{X}_T, \tilde{z}_T)\right). \tag{2.8}$$

Consequently, to establish the estimate (1.4) it is sufficient, for each $T > 0$, to find $((X_t, z_t))_{t\geq 0}$ and $((\tilde{X}_t, \tilde{z}_t))_{t\geq 0}$ started respectively at (x, z) and $(\tilde{x}, \tilde{z})$, satisfying

$$\mathbb{P}\left((X_T, z_T) \neq (\tilde{X}_T, \tilde{z}_T)\right) \leq C_1 \frac{\|\tilde{x} - x\|_2}{\sqrt{T}} + C_2 \frac{|\tilde{z} - z - \frac{1}{2} x \cdot \tilde{x}|}{T} \tag{2.9}$$

for $C_1, C_2 > 0$ not depending on T.

To perform the construction of the coupling, we construct the Brownian motions $(X_t)_{t\in[0,T]}$ and $(\tilde{X}_t)_{t\in[0,T]}$ with Legendre polynomials as in Lemma 2.1.

So let us fix $T > 0$. We write

$$\forall\, 0 \leq t \leq T,\ X_t = x + B_t \quad \text{with} \quad B_t = \sum_{k=0}^{\infty} \xi_k \int_0^t Q_k^T(s)\, ds, \tag{2.10}$$

where $\left(\xi_k = \begin{pmatrix} \xi_k^1 \\ \xi_k^2 \end{pmatrix}\right)_{k\geq 0}$ is a sequence of independent $\mathbb{R}^2$-valued random vectors with law $\mathcal{N}(0, I_2)$. We do the same with $(\tilde{X}_t)_{0\leq t\leq T}$, using independent $\mathbb{R}^2$-valued random variables $\left(\tilde{\xi}_k\right)_{k\geq 0}$ with law $\mathcal{N}(0, I_2)$. Equation (2.9) will be obtained thanks to a well-chosen coupling of $(\xi_k)_{k\geq 0}$ and $\left(\tilde{\xi}_k\right)_{k\geq 0}$.

At time T, using Lemma 2.2, we get

$$X_T = x + \sqrt{T}\xi_0, \qquad z_T = z + \frac{1}{2}\sqrt{T} x \cdot \xi_0 + T \sum_{k\geq 0} \alpha_k \xi_k \cdot \xi_{k+1}, \tag{2.11}$$

$$\tilde{X}_T = \tilde{x} + \sqrt{T}\tilde{\xi}_0, \qquad \tilde{z}_T = \tilde{z} + \frac{1}{2}\sqrt{T} \tilde{x} \cdot \tilde{\xi}_0 + T \sum_{k\geq 0} \alpha_k \tilde{\xi}_k \cdot \tilde{\xi}_{k+1}, \tag{2.12}$$

where for $k \geq 0$, α_k is given by (2.6).

From (2.11) and (2.12), we find that the coupling equation $(X_T, z_T) = (\tilde{X}_T, \tilde{z}_T)$ is equivalent to

$$\left\{ \begin{array}{rl} \tilde{\xi}_0 - \xi_0 & = \frac{x - \tilde{x}}{\sqrt{T}} \\ z - \tilde{z} + \frac{\sqrt{T}}{2}\left(x \cdot \xi_0 - \tilde{x} \cdot \tilde{\xi}_0\right) & = T \sum_{k\geq 0} \alpha_k \left(\tilde{\xi}_k \cdot \tilde{\xi}_{k+1} - \xi_k \cdot \xi_{k+1}\right). \end{array} \right. \tag{2.13}$$

Replacing $\tilde{\xi}_0$ by $\xi_0 + \frac{x-\tilde{x}}{\sqrt{T}}$ in the second equation we get

$$-\zeta + (x - \tilde{x}) \cdot \left(\frac{\sqrt{T}}{2}\xi_0 - \sqrt{T}\alpha_0 \tilde{\xi}_1\right) = T\alpha_0 \xi_0 \cdot (\tilde{\xi}_1 - \xi_1) + T \sum_{k\geq 1} \alpha_k \left(\tilde{\xi}_k \cdot \tilde{\xi}_{k+1} - \xi_k \cdot \xi_{k+1}\right). \tag{2.14}$$

where $\zeta = \tilde{z} - z - \frac{1}{2}(x \cdot \tilde{x})$ is the last coordinate in the Heisenberg group of $(x, z)^{-1} \cdot (\tilde{x}, \tilde{z})$. We are in position to start the coupling. We take

$$\xi_k = \tilde{\xi}_k \quad \text{for all} \quad k \notin \{0, 3\}. \tag{2.15}$$

so that we are left to couple

$$(\xi_0, \tilde{\xi}_0), \quad (\xi_3, \tilde{\xi}_3). \tag{2.16}$$

If (2.15) is satisfied we have the simplification

$$\begin{aligned}
&T\alpha_0\xi_0 \cdot (\tilde{\xi}_1 - \xi_1) + T\sum_{k\geq 1}\alpha_k\left(\tilde{\xi}_k \cdot \tilde{\xi}_{k+1} - \xi_k \cdot \xi_{k+1}\right)\\
&= T\alpha_2\left(\xi_2 \cdot \tilde{\xi}_3 - \xi_2 \cdot \xi_3\right) + T\alpha_3\left(\tilde{\xi}_3 \cdot \xi_4 - \xi_3 \cdot \xi_4\right)\\
&= T\sqrt{\alpha_2^2+\alpha_3^2}\left(\tilde{\xi}_3 - \xi_3\right) \cdot \frac{\alpha_3\xi_4 - \alpha_2\xi_2}{\sqrt{\alpha_2^2+\alpha_3^2}}.
\end{aligned}$$

Define

$$W = -\zeta + (x-\tilde{x}) \cdot \left(\frac{\sqrt{T}}{2}\xi_0 - \sqrt{T}\alpha_0\xi_1\right) \in \mathbb{R},\ V = \frac{\alpha_3\xi_4 - \alpha_2\xi_2}{\sqrt{\alpha_2^2+\alpha_3^2}} \in \mathbb{R}^2. \tag{2.17}$$

With these definitions, Equation (2.14) becomes

$$T\sqrt{\alpha_2^2+\alpha_3^2}\left(\tilde{\xi}_3 - \xi_3\right) \cdot V = W, \tag{2.18}$$

where the random vector V is of law $\mathcal{N}(0, I_2)$ and is independent of W. Consider (E_1, E_2) to be a positively oriented orthonormal basis of $\mathbb{R}^2$ and such that E_1 is proportional to V. Writing $U = U^1E_1 + U^2E_2$, and since $E_1 \cdot E_1 = 0$ and $E_1 \cdot E_2 = 1$, the solutions of equation

$$U \cdot V = W \tag{2.19}$$

are precisely the vectors $U \in \mathbb{R}^2$ such that

$$U^2 = -\frac{W}{\|V\|_2}.$$

Note that nothing is imposed on the coordinate U^1. A solution of (2.14) is thus obtained if

$$\tilde{\xi}_3 - \xi_3 = -\frac{1}{T\sqrt{\alpha_2^2+\alpha_3^2}}\frac{W}{\|V\|_2}E_2. \tag{2.20}$$

We also denote (F_1, F_2) to be the positively oriented orthonormal basis of $\mathbb{R}^2$ such that F_1 is proportional to $\tilde{x} - x$. We emphasize that W depends only on $\langle \xi_0, F_2\rangle$ (and on $\langle \xi_1, F_2\rangle$) and thus we will also take $\langle \tilde{\xi}_0, F_2\rangle = \langle \xi_0, F_2\rangle$.

Now by Lemma 2.3, given the values of ξ_k for $k \in \mathbb{N}\setminus\{0,3\}$ and the value of $\langle \xi_0, F_2\rangle$, it is possible to construct a coupling of the three dimensional Gaussian random vectors $(\langle \xi_0, F_1\rangle, \xi_3)$ and $(\langle \tilde{\xi}_0, F_2\rangle, \tilde{\xi}_3)$ such that

$$\mathbb{P}\left((X_T, z_T) \neq (\tilde{X}_T, \tilde{z}_T) | (\xi_k)_{k\in\mathbb{N}\setminus\{0,3\}}, \langle \xi_0, F_2\rangle\right) \leq \frac{1}{\sqrt{2\pi}}\left(\frac{\|\tilde{x}-x\|_2}{\sqrt{T}} + \frac{1}{T\sqrt{\alpha_2^2+\alpha_3^2}}\frac{|W|}{\|V\|_2}\right)$$

and thus since V and W are independent

$$\mathbb{P}\left((X_T, z_T) \neq (\tilde{X}_T, \tilde{z}_T)\right) \leq \frac{1}{\sqrt{2\pi}}\left(\frac{\|x-\tilde{x}\|_2}{\sqrt{T}} + \frac{1}{T\sqrt{\alpha_2^2+\alpha_3^2}}\mathbb{E}\left[\frac{1}{\|V\|_2}\right]\mathbb{E}[|W|]\right).$$

Now since V is a random vector with law $\mathcal{N}(0, I_2)$:

$$\mathbb{E}\left[\frac{1}{\|V\|_2}\right] = \sqrt{\frac{\pi}{2}}.$$

Denoting $\hat{\xi}_0 = \frac{1}{\sqrt{\frac{1}{4}+\alpha_0^2}}\left(\frac{\sqrt{T}}{2}\xi_0 - \sqrt{T}\alpha_0\xi_1\right) \sim \mathcal{N}(0, I_2)$ and with the same orthonormal basis (F_1, F_2) of $\mathbb{R}^2$:

$$\begin{aligned}
\mathbb{E}[|W|] &\leq |\zeta| + \sqrt{T}\sqrt{\frac{1}{4}+\alpha_0^2}\,\mathbb{E}\left[\left|(\tilde{x}-x)\cdot\hat{\xi}_0\right|\right]\\
&= |\zeta| + \sqrt{T}\sqrt{\frac{1}{4}+\alpha_0^2}\,\|\tilde{x}-x\|_2\,\mathbb{E}\left[|\langle F_2, \hat{\xi}_0\rangle|\right]\\
&= |\zeta| + \sqrt{\frac{T}{3}}\sqrt{\frac{2}{\pi}}\,\|\tilde{x}-x\|_2
\end{aligned}$$

since $\langle F_2, \hat{\xi}_0\rangle \sim \mathcal{N}(0,1)$. Thus the conclusion (1.4) holds with

$$C_2 = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\alpha_2^2+\alpha_3^2}} \sqrt{\frac{\pi}{2}} = \sqrt{22.5} \text{ and } C_1 = \frac{1}{\sqrt{2\pi}} + \sqrt{\frac{2}{3\pi}} C_2 = \frac{1}{\sqrt{2\pi}} \left(1+\sqrt{30}\right).$$

The constants given in Theorem 1.1 will be obtained in Remark 2. □

Remark 1. The above explicit constants C_1 and C_2 are not optimal. In some sense, we try to use the less noise possible in the coupling. It should be possible to decrease their values by allowing more random Gaussian vectors to be different in (2.15).

Remark 2. Using the left invariance and the rotational invariance of the Heisenberg group, it is enough to consider the total variation between the measures $\mu_T^{(x_1,0,z)}$ and $\mu_T^{(0,0,0)}$. In this case, we can take

$$\tilde{\xi}_0^2 = \xi_0^2, \tilde{\xi}_2^1 = \xi_2^1, \tilde{\xi}_3^1 = \xi_3^1 \text{ and } \tilde{\xi}_k = \xi_k \text{ for } k=1 \text{ and } k \geq 4.$$

so that we are left to couple

$$(\xi_0^1, \xi_2^2, \xi_3^2), \quad (\tilde{\xi}_0^1, \tilde{\xi}_2^2, \tilde{\xi}_3^2).$$

By rewriting carefully the above proof, one can then replace the constant

$$\frac{1}{\sqrt{\alpha_2^2+\alpha_3^2}} \mathbb{E}\left[\frac{1}{\|V\|}\right] \text{ by } \mathbb{E}\left[\frac{1}{\sqrt{(\alpha_1^2+\alpha_2^2)Z_1^2 + (\alpha_2^2+\alpha_3^2)Z_2^2}}\right]$$

where Z_1 and Z_2 are two independent $\mathcal{N}(0,1)$ random variables. In fact, in view of Remark 1, if one allows to couple,

$$\left(\xi_0^1, (\xi_2^2, \xi_3^2), (\xi_5^2, \xi_6^2), (\xi_8^2, \xi_9^2), \dots\right), \quad \left(\tilde{\xi}_0^1, (\tilde{\xi}_2^2, \tilde{\xi}_3^2), (\tilde{\xi}_5^2, \tilde{\xi}_6^2), (\tilde{\xi}_8^2, \tilde{\xi}_9^2), \dots\right),$$

the previous constant may even be replaced by

$$\mathbb{E}\left[\frac{1}{\sqrt{\sum_{k\geq 1} c_k^2 Z_k^2}}\right]$$

where $(Z_n)_{n\geq 1}$ is an independent sequence of $\mathcal{N}(0,1)$ random variables and where $(c_k^2)_{k\geq 1}$ is the sequence given by:

$$c_k^2 = \begin{cases} \alpha_{3j+1}^2 + \alpha_{3j+2}^2 & if \quad k = 2j+1, j \geq 0 \\ \alpha_{3j+2}^2 + \alpha_{3j+3}^2 & if \quad k = 2j+2, j \geq 0. \end{cases}$$

More explicitly, one has

$$\begin{aligned}
(c_k^2)_{k\geq 1} &= \left(\alpha_1^2+\alpha_2^2, \alpha_2^2+\alpha_3^2, \alpha_4^2+\alpha_5^2, \alpha_5^2+\alpha_6^2, \alpha_7^2+\alpha_8^2, \alpha_8^2+\alpha_9^2, \dots\right) \\
&= \left(\frac{1}{2\times 3\times 7}, \frac{1}{2\times 5\times 9}, \frac{1}{2\times 9\times 13}, \frac{1}{2\times 11\times 15}, \frac{1}{2\times 15\times 19}, \frac{1}{2\times 17\times 21}, \dots\right) \\
&\geq \gamma \left(1, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \frac{1}{25}, \frac{1}{36}, \dots\right)
\end{aligned}$$

with

$$\gamma = \frac{1}{42}$$

and since one has for $k \geq 0$

$$\alpha_k^2 + \alpha_{k+1}^2 = \frac{1}{2(2k+1)(2k+5)}. \tag{2.21}$$

This gives

$$\mathbb{E}\left[\frac{1}{\sqrt{\sum_{k\geq 1} c_k^2 Z_k^2}}\right] \leq \frac{1}{\pi\sqrt{\gamma}} \mathbb{E}\left[S_{\frac{1}{2}}^{-1/2}\right] \leq \frac{5}{\pi\sqrt{2\gamma}} = \frac{5\sqrt{21}}{\pi}$$

where $\frac{1}{\pi^2}\sum_{k\geq 1}\frac{1}{k^2}Z_k^2$ has the same law as $S_{\frac{1}{2}}$ defined in Lemma 4.4 and using (4.31) in this lemma.

The announced slightly better constants ensue:

$$C_2 = \frac{5\sqrt{21}}{\pi\sqrt{2\pi}} \text{ and } C_1 = \frac{1}{\sqrt{2\pi}} + \sqrt{\frac{2}{3\pi}}C_2 = \frac{1}{\sqrt{2\pi}}\left(1 + \frac{5\sqrt{14}}{\pi\sqrt{\pi}}\right).$$

3 Distance in total variation of subelliptic Brownian motions in Carnot groups

The aim of this section is to extend Theorem 1.1 to free step 2 Carnot groups. We start with several definitions and lemmas.

3.1 Some preliminaries

For n, m positive integers, denote by $M_{n,m}(\mathbb{R})$ the set of matrices with real entries, n rows and m columns. Also denote by $\mathfrak{so}(n)$ the set of skew-symmetric matrices of size n. If $u, v \in M_{n,1}(\mathbb{R})$, let $u \odot v := uv^t - vu^t \in \mathfrak{so}(n)$, where $v^t \in M_{1,n}(\mathbb{R})$ denotes the transpose of v.

In the sequel we will identify $M_{n,1}(\mathbb{R})$ with $\mathbb{R}^n$.

Remark 3. In the special case $n = 3$, $\mathfrak{so}(3)$ is usually identified with $\mathbb{R}^3$ via the linear map

$$\begin{aligned} \psi : \mathbb{R}^3 &\to \mathfrak{so}(3) \\ \begin{pmatrix} a \\ b \\ c \end{pmatrix} &\mapsto \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}. \end{aligned} \tag{3.1}$$

On the other hand, $\mathbb{R}^3$ is endowed with the usual vectorial product $\wedge$. In this situation, it can be checked that for $u, v \in \mathbb{R}^3$

$$\psi(u \wedge v) = [\psi(u), \psi(v)] = -u \odot v. \tag{3.2}$$

In other words, the map ψ is an isomorphism between the two Lie algebras $(\mathbb{R}^3, \wedge)$ and $(\mathfrak{so}(3), [\cdot, \cdot])$.

Definition 3.1. For $n \geq 2$, the free step 2 Carnot group $\mathbb{G}_n$ is the vector space

$$\mathbb{G}_n := \mathbb{R}^n \times \mathfrak{so}(n) \tag{3.3}$$

endowed with the group operation

$$(u, A) \star (v, B) := \left(u + v, A + B + \frac{1}{2}u \odot v\right). \tag{3.4}$$

Remark 4. The Carnot group $\mathbb{G}_2$ is isomorphic to the Heisenberg group $\mathbb{H}$.

Consider an integrable random variable W taking its values in $\mathfrak{so}(n)$, and for $m \geq n+2$, $V_1, \ldots, V_m$, m independent random vectors taking their values in $\mathbb{R}^n$, with law $\mathcal{N}(0, I_n)$ and independent of W. Our next aim is to solve in $U_1, \ldots, U_m$ random variables with values in $\mathbb{R}^n$, the equation

$$\sum_{k=1}^m U_k \odot V_k = W. \tag{3.5}$$

Clearly the solution is not unique. We will make a specific choice which will together give uniqueness and allow explicit computations. Letting $(e_1, \ldots, e_n)$ be the canonical basis of $\mathbb{R}^n = M_{n,1}(\mathbb{R})$, we have the canonical decomposition

$$W = \sum_{i,j=1}^n W^{i,j} e_i e_j^t = \frac{1}{2}\sum_{i,j=1}^n W^{i,j} e_i \odot e_j, \quad \text{since} \quad W^{j,i} = -W^{i,j}. \tag{3.6}$$

We will denote

$$U_k = \sum_{j=1}^{n} U_k^j e_j, \qquad k = 1, \ldots, m, \tag{3.7}$$

$$V_k = \sum_{j=1}^{n} V_k^j e_j, \qquad k = 1, \ldots, m, \tag{3.8}$$

$$\mathscr{U} = \begin{pmatrix} U_1^1 & U_2^1 & \cdots & U_m^1 \\ U_1^2 & U_2^2 & \cdots & U_m^2 \\ \vdots & \vdots & \vdots & \vdots \\ U_1^n & U_2^n & \cdots & U_m^n \end{pmatrix} = \left(U_k^j\right)_{1\le j\le n,\ 1\le k\le m}, \tag{3.9}$$

$$\mathscr{V} = \begin{pmatrix} V_1^1 & V_2^1 & \cdots & V_m^1 \\ V_1^2 & V_2^2 & \cdots & V_m^2 \\ \vdots & \vdots & \vdots & \vdots \\ V_1^n & V_2^n & \cdots & V_m^n \end{pmatrix} = \left(V_k^i\right)_{1\le i\le n,\ 1\le k\le m} \tag{3.10}$$

and

$$\mathscr{W} = \begin{pmatrix} 0 & W^{1,2} & W^{1,3} & \cdots & W^{1,n} \\ -W^{1,2} & 0 & \ddots & \cdots & W^{2,n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & W^{n-1,n} \\ -W^{1,n} & \cdots & \cdots & -W^{n-1,n} & 0 \end{pmatrix} = \left(W^{i,j}\right)_{1\le i\le n,\ 1\le j\le n}. \tag{3.11}$$

Equation (3.5) rewrites as

$$\mathscr{U}\mathscr{V}^t - \mathscr{V}\mathscr{U}^t = \mathscr{W}. \tag{3.12}$$

Equation (3.12) is sometimes called a T-Sylvester equation in the literature. The next proposition provides a particular solution and gives some estimates when $m \geq n+2$.

Proposition 3.1. A solution to Equation (3.5) is given by

$$\mathscr{U}^t = -\frac{1}{2}\mathscr{V}^t\left(\mathscr{V}\mathscr{V}^t\right)^{-1}\mathscr{W}. \tag{3.13}$$

For $0 < q \leq 2$, it satisfies:

$$\mathbb{E}[\|\mathscr{U}\|^q] \leq \frac{1}{2^q n^{\frac{q}{2}}}\mathbb{E}\left[\operatorname{tr}\left((\mathscr{V}\mathscr{V}^t)^{-1}\right)^{\frac{q}{2}}\right]\mathbb{E}\left[\|\mathscr{W}\|^q\right], \tag{3.14}$$

where $\|\mathscr{U}\|$ and $\|\mathscr{W}\|$ denote Hilbert-Schmidt norms.

In particular, it satisfies

$$\mathbb{E}[\|\mathscr{U}\|^2] \leq \frac{1}{4(m-n-1)}\mathbb{E}\left[\|\mathscr{W}\|^2\right]. \tag{3.15}$$

Proof. We first note that we have a particular solution of Equation (3.12) if

$$\mathscr{V}\mathscr{U}^t = -\frac{1}{2}\mathscr{W}. \tag{3.16}$$

We easily check that $\mathscr{U}$ given by (3.13) is a solution of (3.16) and thus also of (3.5).

From (3.13) we get

$$\mathscr{U}\mathscr{U}^t = \frac{1}{4}\mathscr{W}^t\left(\mathscr{V}\mathscr{V}^t\right)^{-1}\mathscr{W}. \tag{3.17}$$

This yields

$$\|\mathscr{U}\| = \sqrt{\operatorname{tr}\left(\mathscr{U}\mathscr{U}^t\right)} = \frac{1}{2}\sqrt{\operatorname{tr}\left(\mathscr{W}^t\left(\mathscr{V}\mathscr{V}^t\right)^{-1}\mathscr{W}\right)}. \tag{3.18}$$

On the other hand, $\mathscr{V}$ is independent of $\mathscr{W}$ and $\mathscr{V}\mathscr{V}^t$ is a standard Wishart matrix $\mathcal{W}(n,m)$ of size $n \times n$ and with m degrees of freedom and thus can be written in singular value decomposition as

$$\mathscr{V}\mathscr{V}^t = \mathscr{S}^t \mathscr{D}^2 \mathscr{S} \tag{3.19}$$

where $\mathscr{S}$ and $\mathscr{D}$ are two independent random variables taking their values respectively in $O(n)$ and $M_{n,n}(\mathbb{R})$, $\mathscr{S}$ having uniform law and $\mathscr{D}$ being diagonal with positive eigenvalues $0 < d_1 < \ldots < d_n$. From this and with the conditional Jensen inequality since $q \leq 2$, we get:

$$\begin{aligned}
2^q \mathbb{E}\left[\|\mathscr{U}\|^q\right] &= \mathbb{E}\left[\operatorname{tr}\left(\mathscr{W}^t\left(\mathscr{V}\mathscr{V}^t\right)^{-1}\mathscr{W}\right)^{\frac{q}{2}}\right] = \mathbb{E}\left[\operatorname{tr}\left(\mathscr{W}^t\mathscr{S}^t\mathscr{D}^{-2}\mathscr{S}\mathscr{W}\right)^{\frac{q}{2}}\right] \\
&= \mathbb{E}\left[\operatorname{tr}\left(\mathscr{S}\mathscr{W}\mathscr{W}^t\mathscr{S}^t\mathscr{D}^{-2}\right)^{\frac{q}{2}}\right] = \mathbb{E}\left[\left(\sum_{i=1}^n d_i^{-2} e_i^t \mathscr{S}\mathscr{W}\mathscr{W}^t\mathscr{S}^t e_i\right)^{\frac{q}{2}}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\left(\sum_{i=1}^n d_i^{-2} e_i^t \mathscr{S}\mathscr{W}\mathscr{W}^t\mathscr{S}^t e_i\right)^{\frac{q}{2}} |\mathscr{W},\mathscr{D}\right]\right] \\
&\leq \mathbb{E}\left[\left(\mathbb{E}\left[\sum_{i=1}^n d_i^{-2} e_i^t \mathscr{S}\mathscr{W}\mathscr{W}^t\mathscr{S}^t e_i |\mathscr{W},\mathscr{D}\right]\right)^{\frac{q}{2}}\right].
\end{aligned}$$

Now for all $1 \leq i \leq n$,

$$\mathbb{E}\left[e_i^t \mathscr{S}\mathscr{W}\mathscr{W}^t\mathscr{S}^t e_i|\mathscr{W},\mathscr{D}\right] = \frac{1}{n}\operatorname{tr}(\mathscr{W}\mathscr{W}^t)$$

since $\mathscr{S}$ is uniformly distributed and independent of $\mathscr{W}$ and $\mathscr{D}$. We get from this

$$2^q \mathbb{E}\left[\|\mathscr{U}\|^q\right] \leq \frac{1}{n^{\frac{q}{2}}}\mathbb{E}\left[\operatorname{tr}(\mathscr{W}\mathscr{W}^t)^{\frac{q}{2}}\operatorname{tr}\left(\mathscr{D}^{-2}\right)^{\frac{q}{2}}\right]$$

and Inequality (3.14) follows since $\operatorname{tr}\left(\mathscr{D}^{-2}\right) = \operatorname{tr}\left(\left(\mathscr{V}\mathscr{V}^t\right)^{-1}\right)$ is independent of $\mathscr{W}$.

By [24] *Example 3.1* we have

$$\mathbb{E}\left[\operatorname{tr}\left(\left(\mathscr{V}\mathscr{V}^t\right)^{-1}\right)\right] = \frac{n}{m-n-1} \tag{3.20}$$

and Inequality (3.15) directly follows. □

3.2 Distance in total variation of two Brownian motions in Carnot groups

Theorem 1.1 yields an upper bound for the total variation distances between the laws of two subRiemannian Brownian motions in $\mathbb{H} = \mathbb{G}_2$ at time T started at different points. The aim of this section is to extend the result to $\mathbb{G}_n$-valued subRiemannian Brownian motions for all $n \geq 3$.

Definition 3.2. A subRiemannian Brownian motion in $\mathbb{G}_n$ started at (x,z) is a process $((X_t, z_t))_{t\geq 0}$ such that $(X_t)_{t\geq 0}$ is a $\mathbb{R}^n$-valued Brownian motion started at x and $(z_t)_{t\geq 0}$ is the $\mathfrak{so}(n)$-valued process satisfying

$$\forall\, t \geq 0, \quad z_t = z + \frac{1}{2}\int_0^t X_s \odot dX_s. \tag{3.21}$$

Theorem 3.2. For $T > 0$ and $(x,z) \in \mathbb{G}_n$ let $\mu_T^{(x,z)}$ be the law at T of the subRiemannian Brownian motion started at (x,z). We have for all $T > 0$ and all $((x,z),(\tilde{x},\tilde{z})) \in \mathbb{G}_n^2$,

$$d_{TV}\left(\mu_T^{(x,z)}, \mu_T^{(\tilde{x},\tilde{z})}\right) \leq C_1(n)\frac{\|\tilde{x}-x\|_2}{\sqrt{T}} + C_2(n)\frac{\|\tilde{z}-z-\frac{1}{2}x\odot\tilde{x}\|}{T} \tag{3.22}$$

where

$$C_2(n) := \frac{1}{\sqrt{\pi}}\left(6\sqrt{n} + \frac{4}{\sqrt{n}}\right) \quad \text{and} \quad C_1(n) := \frac{1}{\sqrt{2\pi}} + \sqrt{\frac{2(n-1)}{3}}C_2(n). \tag{3.23}$$

Remark 5. Theorem 3.2 also applies when $n = 2$, i.e., in the case of the Heisenberg group. In order to compare the constants in Theorems 1.1 and 3.2, note that $\|\tilde{z}-z-\frac{1}{2}x\odot\tilde{x}\| = \sqrt{2}|\tilde{z}-z-\frac{1}{2}(x_1\tilde{x}_2-x_2\tilde{x}_1)|$.

Remark 6. Note that a Brownian motion $(X_t, z_t)_t$ on $\mathbb{G}_n$ encompasses Brownian motions on the Heisenberg group. Indeed, if $(X_t, z_t)_t$ is a Brownian motion on $\mathbb{G}_n$, for any $1 \le i < j \le n$, $(X_t^i, X_t^j, z_t^{i,j})_t$ is a Brownian motion on the Heisenberg group. In particular, the total variation distance between the laws of Brownian motion on $\mathbb{G}_n$ are bigger than the corresponding ones on the Heisenberg group. Since as said before, Theorem 1.1 is known to provide a sharp order of decay on the Heisenberg group [3], Theorem 3.2 is also sharp on $\mathbb{G}_n$ and the associated coupling is efficient, see also [13] for more details.

Proof. For any choice of two $\mathbb{G}_n$-valued subRiemannian Brownian motions $((X_t, z_t))_{t\ge0}$ and $((\tilde{X}_t, \tilde{z}_t))_{t\ge0}$ started respectively at (x, z) and $(\tilde{x}, \tilde{z})$, we have

$$d_{TV}\left(\mu_T^{(x,z)}, \mu_T^{(\tilde{x},\tilde{z})}\right) \le \mathbb{P}\left((X_T, z_T) \ne (\tilde{X}_T, \tilde{z}_T)\right). \tag{3.24}$$

Consequently, to establish the estimate (3.22) it is sufficient, for each $T > 0$, to find $((X_t, z_t))_{t\ge0}$ and $((\tilde{X}_t, \tilde{z}_t))_{t\ge0}$ started respectively at (x, z) and $(\tilde{x}, \tilde{z})$, satisfying

$$\mathbb{P}\left((X_T, z_T) \ne (\tilde{X}_T, \tilde{z}_T)\right) \le C_1(n)\frac{\|\tilde{x}-x\|_2}{\sqrt{T}} + C_2(n)\frac{\|\tilde{z}-z-\frac{1}{2}x\odot\tilde{x}\|}{T}. \tag{3.25}$$

Adopting the same strategy as in Section 2, we construct the Brownian motions $(X_t)_{t\ge0}$ and $(\tilde{X}_t)_{t\ge0}$ with Legendre polynomials.

Fix $T > 0$. Similarly to Equation (2.4) but now in dimension n, we write

$$\forall\, 0 \le t \le T,\ X_t = x + B_t \quad \text{with} \quad B_t = \sum_{k=0}^{\infty} \xi_k \int_0^t Q_k^T(s)\, ds, \tag{3.26}$$

where $\left(\xi_k = \begin{pmatrix} \xi_k^1 \\ \vdots \\ \xi_k^n \end{pmatrix}\right)_{k\ge0}$ is a sequence of independent $\mathbb{R}^n$-valued random vectors with law $\mathcal{N}(0, I_n)$.

We do the same with $(\tilde{X}_t)_{0\le t\le T}$, using independent $\mathbb{R}^n$-valued random variables $\left(\tilde{\xi}_k\right)_{k\ge0}$ with law $\mathcal{N}(0, I_n)$. Equation (3.25) will be obtained thanks to a well-chosen coupling of $(\xi_k)_{k\ge0}$ and $\left(\tilde{\xi}_k\right)_{k\ge0}$.

At time T we get

$$X_T = x + \sqrt{T}\xi_0, \qquad z_T = z + \frac{1}{2}\sqrt{T}x\odot\xi_0 + T\sum_{k\ge0}\alpha_k\xi_k\odot\xi_{k+1}, \tag{3.27}$$

$$\tilde{X}_T = \tilde{x} + \sqrt{T}\tilde{\xi}_0, \qquad \tilde{z}_T = \tilde{z} + \frac{1}{2}\sqrt{T}\tilde{x}\odot\tilde{\xi}_0 + T\sum_{k\ge0}\alpha_k\tilde{\xi}_k\odot\tilde{\xi}_{k+1}, \tag{3.28}$$

where $(\alpha_k)_{k\ge0}$ is defined in (2.6).

From (3.27) and (3.28), we find that the coupling equation $(X_T, z_T) = (\tilde{X}_T, \tilde{z}_T)$ is equivalent to

$$\left\{\begin{array}{rl} \tilde{\xi}_0 - \xi_0 & = \frac{x-\tilde{x}}{\sqrt{T}} \\ z - \tilde{z} + \frac{\sqrt{T}}{2}\left(x\odot\xi_0 - \tilde{x}\odot\tilde{\xi}_0\right) & = T\sum_{k\ge0}\alpha_k\left(\tilde{\xi}_k\odot\tilde{\xi}_{k+1} - \xi_k\odot\xi_{k+1}\right). \end{array}\right. \tag{3.29}$$

Replacing $\tilde{\xi}_0$ by $\xi_0 + \frac{x-\tilde{x}}{\sqrt{T}}$ in the second equation we get

$$-\zeta + (x-\tilde{x})\odot\left(\frac{\sqrt{T}}{2}\xi_0 - \sqrt{T}\alpha_0\tilde{\xi}_1\right) = T\alpha_0\xi_0\odot(\tilde{\xi}_1 - \xi_1) + T\sum_{k\ge1}\alpha_k\left(\tilde{\xi}_k\odot\tilde{\xi}_{k+1} - \xi_k\odot\xi_{k+1}\right) \tag{3.30}$$

where $\zeta = \tilde{z} - z - \frac{1}{2} x \odot \tilde{x}$. We are in position to start the coupling. As in the previous section we let $m \geq n+2$, that we will choose at the end. We take

$$\xi_k = \tilde{\xi}_k \quad \text{for all} \quad k \notin \{0, 3, 6, \dots 3m\}, \tag{3.31}$$

so that we are left to couple

$$(\xi_0, \tilde{\xi}_0), \quad (\xi_3, \tilde{\xi}_3), \dots (\xi_{3m}, \tilde{\xi}_{3m}). \tag{3.32}$$

If (3.31) is satisfied we have the simplification

$$\begin{aligned}
&T\alpha_0 \xi_0 \odot (\tilde{\xi}_1 - \xi_1) + T \sum_{k \geq 1} \alpha_k \left(\tilde{\xi}_k \odot \tilde{\xi}_{k+1} - \xi_k \odot \xi_{k+1} \right) \\
&= T \sum_{k=1}^{m} \left(\alpha_{3k-1} \left(\xi_{3k-1} \odot \tilde{\xi}_{3k} - \xi_{3k-1} \odot \xi_{3k} \right) + \alpha_{3k} \left(\tilde{\xi}_{3k} \odot \xi_{3k+1} - \xi_{3k} \odot \xi_{3k+1} \right) \right) \\
&= \sum_{k=1}^{m} \left(\tilde{\xi}_{3k} - \xi_{3k} \right) \odot T \left(\alpha_{3k} \xi_{3k+1} - \alpha_{3k-1} \xi_{3k-1} \right) \\
&= \sum_{k=1}^{m} T \sqrt{\alpha_{3k}^2 + \alpha_{3k-1}^2} \left(\tilde{\xi}_{3k} - \xi_{3k} \right) \odot \frac{\alpha_{3k} \xi_{3k+1} - \alpha_{3k-1} \xi_{3k-1}}{\sqrt{\alpha_{3k}^2 + \alpha_{3k-1}^2}}.
\end{aligned}$$

Define

$$W = -\zeta + (x - \tilde{x}) \odot \left(\frac{\sqrt{T}}{2} \xi_0 - \sqrt{T} \alpha_0 \tilde{\xi}_1 \right), \tag{3.33}$$

$$V_k = \frac{\alpha_{3k} \xi_{3k+1} - \alpha_{3k-1} \xi_{3k-1}}{\sqrt{\alpha_{3k}^2 + \alpha_{3k-1}^2}}, \quad k = 1, \dots, m. \tag{3.34}$$

With these definitions, Equation (3.30) becomes

$$\sum_{k=1}^{m} T \sqrt{\alpha_{3k}^2 + \alpha_{3k-1}^2} \left(\tilde{\xi}_{3k} - \xi_{3k} \right) \odot V_k = W, \tag{3.35}$$

and the random vectors V_k, $1 \leq k \leq m$ are independent with the same law $\mathcal{N}(0, I_n)$.

Let $(U_1, \dots, U_m)$ be the solution given by (3.13) to Equation (3.5) $\sum_{k=1}^{m} U_k \odot V_k = W$. Using (3.35) we see that a solution to (3.30) is given by

$$\tilde{\xi}_{3k} - \xi_{3k} = \frac{U_k}{T \sqrt{\alpha_{3k}^2 + \alpha_{3k-1}^2}} =: \hat{U}_k, \quad k = 1, \dots, m. \tag{3.36}$$

Define

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_m \end{pmatrix}, \quad \tilde{\xi} = \begin{pmatrix} \tilde{\xi}_1 \\ \tilde{\xi}_2 \\ \vdots \\ \tilde{\xi}_m \end{pmatrix}, \quad \hat{U} = \begin{pmatrix} \hat{U}_1 \\ \hat{U}_2 \\ \vdots \\ \hat{U}_m \end{pmatrix} \tag{3.37}$$

with the $\hat{U}_k$ defined in (3.36). The random vectors ξ, $\tilde{\xi}$ and $\hat{U}$ take their values in $M_{nm,1}(\mathbb{R})$ and ξ, $\tilde{\xi}$ have law $\mathcal{N}(0, I_{nm})$. Recalling the system (3.29), we obtain with (3.36) that

$$\mathbb{P}\left((X_T, z_T) \neq (\tilde{X}_T, \tilde{z}_T) \right) \leq \mathbb{P}\left(\tilde{\xi}_0 - \xi_0 \neq \frac{x - \tilde{x}}{\sqrt{T}} \right) + \mathbb{P}\left(\tilde{\xi} - \xi \neq \hat{U} \right). \tag{3.38}$$

Observing that the random vector $\hat{U}$ is independent of ξ, and using Lemma 2.3, we get the estimate

$$\mathbb{P}\left((X_T, z_T) \neq (\tilde{X}_T, \tilde{z}_T) \right) \leq \frac{\|x - \tilde{x}\|_2}{\sqrt{2\pi T}} + \frac{\mathbb{E}[\|\hat{U}\|]}{\sqrt{2\pi}}. \tag{3.39}$$

By (2.6) the sequence $(\alpha_k)_{k\geq 0}$ is decreasing, consequently the sequence $\left(\frac{1}{\sqrt{\alpha_{3k}^2+\alpha_{3k-1}^2}}\right)_{k\geq 0}$ is increasing and $\mathbb{E}[\|\hat{U}\|] \leq \frac{\mathbb{E}[\|U\|]}{T\sqrt{\alpha_{3m}^2+\alpha_{3m-1}^2}}$. On the other hand using (2.6) or (2.21), we have for $k \geq 1$,

$$\frac{1}{\alpha_{3k}^2+\alpha_{3k-1}^2} = 2(6k-1)(6k+3) \leq 8(3k+1)^2. \tag{3.40}$$

Recalling that by Proposition 3.1, and working for simplicity with $q=2$,

$$\mathbb{E}[\|U\|^2] = \mathbb{E}[\|\mathscr{U}\|^2] \leq \frac{1}{4(m-n-1)}\mathbb{E}\left[\|W\|^2\right]$$

we get

$$\mathbb{E}[\|\hat{U}\|^2] \leq \frac{2(3m+1)^2}{T^2(m-n-1)}\mathbb{E}\left[\|W\|^2\right] \tag{3.41}$$

On the other hand, writing from (3.33)

$$W = -\zeta + (x-\tilde{x}) \odot \left(\frac{\sqrt{T}}{2}\xi_0 - \sqrt{T}\alpha_0\tilde{\xi}_1\right),$$

we get

$$\mathbb{E}[\|W\|^2] = \|\zeta\|^2 + \mathbb{E}\left[\left\|(x-\tilde{x}) \odot \left(\frac{\sqrt{T}}{2}\xi_0 - \sqrt{T}\alpha_0\tilde{\xi}_1\right)\right\|^2\right].$$

We will do the computation in an orthonormal basis $(E_1, \ldots, E_n)$ of $\mathbb{R}^n$ such that $x - \tilde{x} = \|x-\tilde{x}\|_2 E_1$. Since $\alpha_0 = \frac{1}{2\sqrt{3}}$ we have $\frac{\sqrt{T}}{2}\xi_0 - \sqrt{T}\alpha_0\tilde{\xi}_1 = \sqrt{\frac{T}{3}}\hat{\xi}_0$ where $\hat{\xi}_0$ is a $\mathbb{R}^n$-valued Gaussian random variable with law $\mathcal{N}(0, I_n)$. Writing $\hat{\xi}_0 = \sum_{i=1}^n \hat{\xi}_0^i E_i$ we obtain

$$(x-\tilde{x}) \odot \left(\frac{\sqrt{T}}{2}\xi_0 - \sqrt{T}\alpha_0\tilde{\xi}_1\right) = \sqrt{\frac{T}{3}}\|x-\tilde{x}\|_2 \sum_{i=2}^n \hat{\xi}_0^i E_1 \odot E_i.$$

The matrices $E_1 \odot E_i = E_1E_i^t - E_iE_1^t$ being orthogonal each with norm $\sqrt{2}$ we obtain

$$\mathbb{E}\left[\left\|(x-\tilde{x}) \odot \left(\frac{\sqrt{T}}{2}\xi_0 - \sqrt{T}\alpha_0\tilde{\xi}_1\right)\right\|^2\right] = \|x-\tilde{x}\|_2^2\,\frac{2T(n-1)}{3}.$$

We get

$$\mathbb{E}\left[\|W\|^2\right] = \|\zeta\|^2 + \frac{2T(n-1)}{3}\|x-\tilde{x}\|_2^2. \tag{3.42}$$

Using this estimate in (3.41) yields

$$\mathbb{E}[\|\hat{U}\|^2] \leq \frac{2(3m+1)^2}{T^2(m-n-1)}\left(\|\zeta\|^2 + \frac{2T(n-1)}{3}\|x-\tilde{x}\|_2^2\right). \tag{3.43}$$

We can easily prove that the best choice for an integer m is

$$m = 2n+1 \quad \text{implying} \quad \frac{2(3m+1)^2}{(m-n-1)} = \left(6\sqrt{2n} + \frac{4\sqrt{2}}{\sqrt{n}}\right)^2. \tag{3.44}$$

So together with (3.39),

$$\mathbb{P}\left((X_T, z_T) \neq (\tilde{X}_T, \tilde{z}_T)\right) \leq \frac{\|x-\tilde{x}\|_2}{\sqrt{2\pi T}} + \frac{1}{T\sqrt{\pi}}\left(6\sqrt{n} + \frac{4}{\sqrt{n}}\right)\left(\|\zeta\| + \sqrt{\frac{2T(n-1)}{3}}\|x-\tilde{x}\|_2\right). \tag{3.45}$$

We obtain the wanted inequality (3.22) with

$$C_2(n) = \frac{1}{\sqrt{\pi}}\left(6\sqrt{n} + \frac{4}{\sqrt{n}}\right) \quad \text{and} \quad C_1(n) = \frac{1}{\sqrt{2\pi}} + \sqrt{\frac{2(n-1)}{3}}C_2(n). \tag{3.46}$$

□

4 Application to gradients inequalities

4.1 Direct estimates for the horizontal and vertical gradient

Similarly to the case of the Heisenberg group, it is possible to define the left-invariant vector fields on $\mathbb{G}_n$. The horizontal vector fields are defined for $1 \leq i \leq n$ by

$$X_i(f)(x,z) = \frac{d}{dt}_{|t=0} f((x,z) \star (te_i, 0)) = \left(\partial_{x_i} - \sum_{j=1, j\neq i}^{n} \frac{1}{2} x_j \partial_{z_{i,j}}\right) f(x,z)$$

and the vertical vector fields for $1 \leq i < j \leq n$ by

$$Z_{i,j}(f)(x,z) = \frac{d}{dt}_{|t=0} f((x,z) \star (0, te_i \odot e_j)) = \partial_{z_{i,j}} f(x,z)$$

with $z = \sum_{1\leq i<j\leq n} z_{i,j} e_i \odot e_j$ and where in the definition of X_i, if $i > j$, we set $\partial_{z_{i,j}} = -\partial_{z_{j,i}}$.

It is also possible to define the Carnot-Carathéodory subRiemmanian distance on $\mathbb{G}_n$ by :

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

where γ ranges over the horizontal curves connecting $\gamma(0) = g$ and $\gamma(1) = g'$; i.e., absolutely continuous curves such that $\dot{\gamma}(t) \in \mathrm{Span}\{X_i(\gamma(t)), 1 \leq i \leq n\}$ almost surely and where $|\cdot|_{\mathfrak{h}}$ is a Euclidean norm on $\mathrm{Span}\{X_i(\gamma(t)), 1 \leq i \leq n\}$ obtained by asserting that $(X_1, \ldots, X_n)$ is an orthonormal basis in each point. As for the Heisenberg group, these Carnot groups admits homogeneous dilations adapted both to the distance and the group structure given by

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

Finally the horizontal gradient $\nabla_{\mathfrak{h}} f$ is $\sum_{i=1}^n X_i(f) X_i$ whereas the vertical gradient is defined by $\nabla_{\mathfrak{v}} f = \sum_{1\leq i<j\leq n} Z_{i,j}(f) Z_{i,j}$.

The total variation estimate implies the following L^∞ gradient bounds.

Corollary 4.1. Let $\mathbb{G}_n$ be the free step 2 Carnot group of of rank $n \geq 2$. For any bounded measurable function f on $\mathbb{G}_n$, for any $g \in \mathbb{G}_n$ and $t > 0$,

$$\|\nabla_{\mathfrak{h}} P_t f(g)\| \leq \frac{2C_1(n)}{\sqrt{t}} ||f||_\infty \tag{4.1}$$

and

$$\|\nabla_{\mathfrak{v}} P_t f(g)\| \leq \frac{2\sqrt{2}C_2(n)}{t} ||f||_\infty \tag{4.2}$$

where $C_1(n)$ and $C_2(n)$ are the constant appearing in Theorem 3.2 (or 1.1).

Proof. The proof is standard. Let f be a bounded measurable function on $\mathbb{G}_n$ and let $g, \tilde{g} \in \mathbb{G}_n$.

$$\begin{aligned} |P_t f(g) - P_t f(\tilde{g})| &= \left| \mathbb{E}\left[f(\mathbb{B}_t^g) - f(\mathbb{B}_t^{\tilde{g}}) \right] \right| \\ &= \left| \mathbb{E}\left[f(\mathbb{B}_t^g) - f(\mathbb{B}_t^{\tilde{g}}) \mathbb{1}_{\{\mathbb{B}_t^g \neq \mathbb{B}_t^{\tilde{g}}\}} \right] \right| \\ &\leq 2||f||_\infty \, \mathbb{P}\left(\mathbb{B}_t^g \neq \mathbb{B}_t^{\tilde{g}} \right). \end{aligned} \tag{4.3}$$

Now since there exists a constant $C > 0$ such that

$$\|\tilde{x} - x\| \leq d_{CC}(g, \tilde{g}) \text{ and } \|\zeta\| = \|\tilde{z} - z - \frac{1}{2} x \odot \tilde{x}\| \leq C d_{CC}(g, \tilde{g})^2,$$

by Theorem 3.2 (or Theorem 1.1), one can construct a coupling of $\mathbb{B}_t^g$ and $\mathbb{B}_t^{\tilde{g}}$ such that

$$\mathbb{P}\left(\mathbb{B}_t^g \neq \mathbb{B}_t^{\tilde{g}} \right) \leq \frac{C_1(n)}{\sqrt{t}} d_{CC}(g, \tilde{g}) + \frac{C C_2(n)}{t} d_{CC}(g, \tilde{g})^2.$$

Dividing by $d_{CC}(g, \tilde{g})$ and letting $\tilde{g} \to g$ gives the horizontal gradient inequality (1.5). When $\tilde{x} = x$, the above estimate writes:

$$\mathbb{P}\left(\mathbb{B}_t^g \neq \mathbb{B}_t^{\tilde{g}} \right) \leq \frac{C_2(n)}{t} \|\tilde{z} - z\|$$

and the vertical gradient inequality (1.6) follows in a similar way. □

4.2 Coupling with change of probability: application to reverse Sobolev inequalities

In this section we will construct couplings at time T with probability one, but the price to pay will be to make changes of probabilities for the second process. The distance between semigroups will be measured by the change of probability. The main results are a log Harnack inequality (Theorem 4.6), an integration by parts formula (Theorem 4.7) for the spatial derivative $dP_T f$ of the semigroup $P_T f$ of the Brownian motion and reverse Poincaré or Sobolev inequalities (Theorem 4.8 and Corollary 4.9) and some estimates of the gradient of the heat kernel (Corollary 4.10).

The notations are the same as in the previous section. The processes $(\mathbb{B}_t^g)_t := (X_t, z_t)_t$ and $(\mathbb{B}_t^{\tilde{g}})_t := ((\tilde{X}_t, \tilde{z}_t))_t$ started respectively at $g = (x, z)$ and $\tilde{g} = (\tilde{x}, \tilde{z})$ are defined with Equations (3.26), (3.21) and (3.27). The sequence $(\xi_k)_{k \geq 0}$ will be identically distributed will law $\mathcal{N}(0, I_n)$ under the probability $\mathbb{P}$. The difference will be that we will look for a sequence $(\tilde{\xi}_k)_{k \geq 0} = (\tilde{\xi}_k(\tilde{g}))_{k \geq 0}$ which is independent and identically distributed with law $\mathcal{N}(0, I_n)$ under another probability $\mathbb{P}(\tilde{g})$, and so that at time T, a.s. $\mathbb{B}_T^g = \mathbb{B}_T^{\tilde{g}}$.

Fix $K \in \{n+1, \ldots\} \cup \{\infty\}$ and let

$$J_K := \{\ell \in \mathbb{N}, \ \ell \leq K\} \quad \text{if} \quad K < \infty, \quad J_\infty := \mathbb{N} \quad \text{and} \quad J_K^* := J_K \backslash \{0\} \ \forall \ K. \tag{4.4}$$

We will take

$$\xi_k = \tilde{\xi}_k \quad \text{for all} \quad k \notin 3J_K \tag{4.5}$$

so that we are left to couple

$$(\xi_\ell, \tilde{\xi}_\ell), \quad \ell \in 3J_K. \tag{4.6}$$

Note that K acts similar as the parameter m in Section 3. In particular, for each K, we have a different coupling. As we will see below, increasing K improves some integrability properties whereas some special choice of K may produce better quantitative estimates. Now we consider the sequence $(V_k)_{k \in J_K^*}$ defined in (3.34), of independent random vectors taking their values in $\mathbb{R}^n$, with the same law $\mathcal{N}(0, I_n)$. We solve in $(U_k)_{k \in J_K^*}$ the equation

$$\sum_{k \in J_K^*} U_k \odot V_k = W \tag{4.7}$$

with W given by Equation (3.33). Then we will choose $(\tilde{\xi}_k)_{k\geq 0}$ such that almost surely

$$\tilde{\xi}_0 - \xi_0 = \frac{x - \tilde{x}}{\sqrt{T}} =: \hat{U}_0 \tag{4.8}$$

and

$$\forall\, k \in J_K^*,\ \tilde{\xi}_{3k} - \xi_{3k} = \frac{U_k}{T\sqrt{\alpha_{3k}^2 + \alpha_{3k-1}^2}} =: \hat{U}_k. \tag{4.9}$$

Notice that $(\hat{U}_k)_{k\in J_K} = (\hat{U}_k(\tilde{g}))_{k\in J_K}$.

We denote

$$\forall k \in J_K^*, \quad V_k = \sum_{i=1}^{n} V_k^i e_i, \quad U_k = \sum_{j=1}^{n} U_k^j e_j, \tag{4.10}$$

$$\beta_k = \frac{1}{T\sqrt{\alpha_{3k}^2 + \alpha_{3k-1}^2}} \tag{4.11}$$

$$\hat{\mathscr{V}} = \hat{\mathscr{V}}_K = \left(\frac{V_k^i}{\beta_k}\right)_{1\leq i\leq n,\ k\in J_K^*}, \quad \hat{\mathscr{U}} = \hat{\mathscr{U}}_K = (\beta_k U_k^i)_{1\leq j\leq n,\ k\in J_K^*}, \tag{4.12}$$

the upper index representing the rows and the lower index representing the columns. With these notations and similarly as before, Equation (4.7) is equivalent to

$$\hat{\mathscr{U}}\hat{\mathscr{V}}^t - \hat{\mathscr{V}}\hat{\mathscr{U}}^t = \mathscr{W} \tag{4.13}$$

with $\mathscr{W}$ defined by (3.11). In particular, we have a solution of Equation (4.7) if

$$\hat{\mathscr{V}}\hat{\mathscr{U}}^t = -\frac{1}{2}\mathscr{W}. \tag{4.14}$$

For each $K \in \{n+1, \ldots\} \cup \{\infty\}$, the $n \times n$ matrix

$$\hat{\mathscr{V}}\hat{\mathscr{V}}^t = \sum_{k\in J_K^*} \frac{1}{\beta_k^2} V_k V_k^t \tag{4.15}$$

is a.s. well-defined even in the case $K = \infty$, since $\mathbb{E}\left[\sum_{k\geq 1} \frac{1}{\beta_k^2}\operatorname{tr}(V_k V_k^t)\right] < \infty$ (the computation (3.40) proves that β_k is of order k). It is a.s. symmetric positive since $K \geq n$ and thus

$$\hat{\mathscr{V}}\hat{\mathscr{V}}^t \geq \frac{1}{\beta_n^2}\mathcal{W}_{n,n} \text{ where } \mathcal{W}_{n,n} = \sum_{k=1}^{n} V_k V_k^t$$

is a Wishart matrix $\mathcal{W}(n,n)$ and is a.s. invertible (see [23], Corollary 3.2.2). Consequently, a solution to (4.14) is given by

$$\hat{\mathscr{U}}^t = -\frac{1}{2}\hat{\mathscr{V}}^t(\hat{\mathscr{V}}\hat{\mathscr{V}}^t)^{-1}\mathscr{W}. \tag{4.16}$$

Let us make a specific choice of probability space, which will be very convenient for our computations. This probability space is $(\Omega, \mathscr{A}, \mathbb{P})$, where $\Omega := \ell^2(\mathbb{R}^n)$ is the Hilbert space of square integrable $\mathbb{R}^n$-valued sequences, $\mathscr{A}$ is the smallest σ-field for which the projections are measurable, completed with respect to the probability measure $\mathbb{P}$ for which the canonical projections

$$\xi_k : \Omega \to \mathbb{R}^n$$
$$\omega = (\omega_0, \omega_1, \ldots, \omega_k, \ldots) \mapsto \omega_k =: \xi_k(\omega)$$

are i.i.d. and $\mathcal{N}(0, I_n)$. We will need to split Ω into two supplementary orthogonal spaces: $\Omega = \Omega_a \oplus \Omega_b$. Let us now describe these spaces. For $k \geq 1$ and $1 \leq i \leq n$, we denote by e_k^i the element of Ω which satisfies $\xi_\ell(e_k^i) = 0$ if $\ell \neq k$ and $\xi_k(e_k^i) = e_i$, the i-th element of the canonical basis of $\mathbb{R}^n$. Letting $(f_1, \dots, f_n)$ be an orthonormal basis of $\mathbb{R}^n$ such that $\|x - \tilde{x}\|_2 f_1 = x - \tilde{x}$, for $1 \leq i \leq n$ we denote by f_0^i the element of Ω such that $\xi_\ell(f_0^i) = 0$ if $\ell \neq 0$ and $\xi_0(f_0^i) = f_i$. Notice that the $(e_k^i),\ k \geq 1,\ 1 \leq i \leq n$ together with the $(f_0^i),\ 1 \leq i \leq n$ form an Hilbertian basis of Ω and that the random variables $\langle e_k^i, \omega\rangle$, $\langle f_0^i, \omega\rangle$ are i.i.d and $\mathscr{N}(0,1)$. Define

$$\Omega_a = \operatorname{Span}\left\{f_0^1,\ e_k^i,\ k \in 3J_K^*,\ 1 \leq i \leq n\right\}, \tag{4.17}$$

$$\Omega_b = \Omega_a^\perp = \operatorname{Span}\left\{f_0^i,\ 2 \leq i \leq n\right\} \oplus \operatorname{Span}\left\{\ e_k^i, k \notin 3J_K,\ 1 \leq i \leq n\right\}. \tag{4.18}$$

For the sequel, we will denote ω_a (resp. ω_b) the projection of ω on Ω_a (resp. Ω_b). Let $\mathscr{A}_a$ and $\mathscr{A}_b$ be the canonical σ-fields and $\mathbb{P}_a$ (resp. $\mathbb{P}_b$) be such that the $\langle \omega_a, e_k^i\rangle$, $k \in 3J_K^*, 1 \leq i \leq n$, $\langle \omega_a, f_0^1\rangle$ (resp. $\langle \omega_b, e_\ell^j\rangle$, $\langle \omega_b, f_0^i\rangle$ $\ell \notin 3J_K, 1 \leq j \leq n, 2 \leq i \leq n$) are independent $\mathscr{N}(0,1)$ random variables. Then

$$\begin{aligned}(\Omega_a \times \Omega_b, \mathscr{A}_a \times \mathscr{A}_b, \mathbb{P}_a \times \mathbb{P}_b) &\to (\Omega, \mathscr{A}, \mathbb{P}) \\ (\omega_a, \omega_b) &\mapsto \omega_a + \omega_b\end{aligned} \tag{4.19}$$

is an isometry.

Recall that $\tilde{\xi}_k = \xi_k$ if $k \notin 3J_K$ and $\tilde{\xi}_{3k} = \xi_{3k} + \hat{U}_k$ if $k \in J_K$. Let $\mathbb{P}(\tilde{g})$ be the probability on Ω such that all $\tilde{\xi}_k$ are i.i.d. and $\mathcal{N}(0,1)$.

Lemma 4.2. The probability $\mathbb{P}(\tilde{g})$ is equivalent to $\mathbb{P}$, and

$$R(u)(\omega) := \frac{d\mathbb{P}(\tilde{g})}{d\mathbb{P}}(\omega) = e^{-\langle \omega, u\rangle - \frac{1}{2}\|u\|^2} \tag{4.20}$$

where $u = u(\tilde{g})(\omega) \in \Omega$ is defined by

$$u_k = 0 \ \forall\ k \notin 3J_K \quad \text{and} \quad u_{3k} = \hat{U}_k(\tilde{g})(\omega)\ \forall k \in J_K, \tag{4.21}$$

and $\langle \omega, u\rangle = \displaystyle\sum_{k=0}^{\infty} \langle \omega_k, u_k(\tilde{g})(\omega)\rangle_{\mathbb{R}^n}$.

In particular, letting

$$\begin{aligned}dR(u(\cdot)(\omega))(\omega)|_{\tilde{g}=g} : T_g\mathbb{G}_n = \mathbb{G}_n &\to \mathbb{R}, \\ h &\mapsto \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} R(u(g+\varepsilon h)(\omega))(\omega),\end{aligned}$$

and similarly

$$\begin{aligned}d\hat{U}_k(\cdot)(\omega)|_{\tilde{g}=g} : T_g\mathbb{G}_n = \mathbb{G}_n &\to \mathbb{R}, \\ h &\mapsto \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \hat{U}_k(g+\varepsilon h)(\omega)(\omega),\end{aligned}$$

we have

$$dR(u(\cdot))|_{\tilde{g}=g} = -\sum_{k=0}^{\infty} \left\langle \xi_{3k}, d\hat{U}_k(\cdot)|_{\tilde{g}=g}\right\rangle. \tag{4.22}$$

Moreover, for all measurable $F : \Omega \to \mathbb{R}$, we have that F is $\mathbb{P}$-integrable if and only if $\omega \mapsto F(\omega + u(\omega))$ is $R(u) \cdot \mathbb{P}$-integrable, and in this case

$$\mathbb{E}[F(\omega)] = \mathbb{E}\left[F(\omega + u(\omega))R(u(\omega))(\omega)\right]. \tag{4.23}$$

We also have

$$\mathbb{E}[F(\omega - u(\omega))] = \mathbb{E}\left[F(\omega)R(u(\omega))(\omega)\right]. \tag{4.24}$$

Proof. First observe that for a fixed deterministic nonzero vector $u \in \Omega_a$, we can make the orthogonal decomposition

$$\omega_a = \left\langle \omega_a, \frac{u}{\|u\|} \right\rangle \frac{u}{\|u\|} + P^{\Omega_a}_{(u)^\perp}(\omega_a) \tag{4.25}$$

where $\left\langle \omega_a, \frac{u}{\|u\|} \right\rangle$ is an $\mathcal{N}(0,1)$ real-valued random variable independent of $P^{\Omega_a}_{(u)^\perp}(\omega_a)$. Now remark that the random vector $u(\omega)$ satisfies $u(\omega) = u(\omega_b)$ in the decomposition $\omega = \omega_a + \omega_b$ of (4.19). This is due to the fact that the $\hat{U}_k$ do not change when one replaces ξ_0 by $\xi_0 - \langle \xi_0, f_1 \rangle f_1$ in the expression of

$$W = -\zeta + \|x - \tilde{x}\|_2 f_1 \odot \left(\frac{\sqrt{T}}{2} \xi_0 - \sqrt{T} \alpha_0 \tilde{\xi}_1 \right).$$

In other words, u is measurable with respect to σ-field $\mathscr{G} := \sigma(\xi_k, k \notin 3J^*_K) \vee \sigma(P^{\mathbb{R}^n}_{(x-\tilde{x})^\perp}(\xi_0))$ ($P^{\mathbb{R}^n}_{(x-\tilde{x})^\perp}$ denoting the projection in $\mathbb{R}^n$ orthogonal to $x - \tilde{x}$).

A second important fact is that $\omega \mapsto u(\omega)$ takes its values in Ω_a. In other words $u_\ell = 0$ if $\ell \notin 3J^*_K$ and u_0 is collinear to $x - \tilde{x}$. Consequently, conditioned to $\mathscr{G}$, u is a Ω_a-valued constant. So we can make the same decomposition as in (4.25):

$$\omega_a = \left\langle \omega_a, \frac{u(\omega)}{\|u(\omega)\|} \right\rangle \frac{u(\omega)}{\|u(\omega)\|} + P^{\Omega_a}_{(u(\omega))^\perp}(\omega_a)$$

where conditioned to $\mathscr{G}$, $\left\langle \omega_a, \frac{u(\omega)}{\|u(\omega)\|} \right\rangle$ is an $\mathscr{N}(0,1)$ random variable independent of $P^{\Omega_a}_{(u(\omega))^\perp}(\omega_a)$. Adding ω_b which is $\mathscr{G}$-measurable and orthogonal to Ω_a we get

$$\omega = \left\langle \omega, \frac{u}{\|u\|} \right\rangle \frac{u}{\|u\|} + P_{(u)^\perp}(\omega) \tag{4.26}$$

where conditioned to $\mathscr{G}$, $\left\langle \omega, \frac{u}{\|u\|} \right\rangle$ is an $\mathcal{N}(0,1)$ real-valued random variable independent of $P_{(u)^\perp}(\omega)$.

Let $F : \Omega \to \mathbb{R}$ a bounded measurable function.

$$\begin{aligned} \mathbb{E}[F(\omega)] &= \mathbb{E}\left[\mathbb{E}\left[F\left(\left\langle \omega, \frac{u}{\|u\|} \right\rangle \frac{u}{\|u\|} + P_{(u)^\perp}(\omega) \right) \middle| P_{(u)^\perp}(\omega),\ \mathscr{G} \right] \right] \\ &= \mathbb{E}\left[\int_{\mathbb{R}} F\left(x \frac{u}{\|u\|} + P_{(u)^\perp}(\omega) \right) \varphi(x)\, dx \right] \end{aligned}$$

where φ is the density of $\mathcal{N}(0,1)$. But

$$\int_{\mathbb{R}} F\left(x \frac{u}{\|u\|} + P_{(u)^\perp}(\omega) \right) \varphi(x)\, dx = \int_{\mathbb{R}} F\left((x + \|u\|) \frac{u}{\|u\|} + P_{(u)^\perp}(\omega) \right) \varphi(x + \|u\|)\, dx$$

yielding

$$\begin{aligned} &\mathbb{E}[F(\omega)] \\ &= \mathbb{E}\left[\int_{\mathbb{R}} F\left((x + \|u\|) \frac{u}{\|u\|} + P_{(u)^\perp}(\omega) \right) \varphi(x + \|u\|)\, dx \right] \\ &= \mathbb{E}\left[\int_{\mathbb{R}} F\left((x + \|u\|) \frac{u}{\|u\|} + P_{(u)^\perp}(\omega) \right) \frac{\varphi(x + \|u\|)}{\varphi(x)} \varphi(x)\, dx \right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\frac{\varphi\left(\left\langle \omega, \frac{u}{\|u\|} \right\rangle + \|u\| \right)}{\varphi\left(\left\langle \omega, \frac{u}{\|u\|} \right\rangle \right)} F\left(\left(\left\langle \omega, \frac{u}{\|u\|} \right\rangle + \|u\| \right) \frac{u}{\|u\|} + P_{(u)^\perp}(\omega) \right) \middle| P_{(u)^\perp}(\omega),\ \mathscr{G} \right] \right] \end{aligned}$$

recalling that conditioned to $\mathscr{G}$, $\left\langle \omega, \frac{u}{\|u\|} \right\rangle$ is an $\mathcal{N}(0,1)$ real-valued random variable independent of $P_{(u)^\perp}(\omega)$. So

$$\begin{aligned}
&\mathbb{E}[F(\omega)] \\
&= \mathbb{E}\left[\frac{\varphi\left(\left\langle \omega, \frac{u}{\|u\|}\right\rangle + \|u\|\right)}{\varphi\left(\left\langle \omega, \frac{u}{\|u\|}\right\rangle\right)} F\left(\left(\left\langle \omega, \frac{u}{\|u\|} \right\rangle + \|u\| \right) \frac{u}{\|u\|} + P_{(u)^\perp}(\omega) \right) \right] \\
&= \mathbb{E}\left[\frac{\varphi\left(\left\langle \omega, \frac{u}{\|u\|}\right\rangle + \|u\|\right)}{\varphi\left(\left\langle \omega, \frac{u}{\|u\|}\right\rangle\right)} F\left(\omega + u\right) \right].
\end{aligned}$$

Observing that

$$\frac{\varphi\left(\left\langle \omega, \frac{u}{\|u\|}\right\rangle + \|u\|\right)}{\varphi\left(\left\langle \omega, \frac{u}{\|u\|}\right\rangle\right)} = e^{-\langle \omega, u\rangle - \frac{1}{2}\|u\|^2}$$

yields (4.20) via (4.23). Equation (4.22) is a direct consequence. Finally, observe that $u(\omega - u(\omega)) = u(\omega)$ since $u(\omega) = u(\omega_b)$ and $u(\omega) \in \Omega_a$. Equation (4.24) is then obtained from (4.23). □

Corollary 4.3. Take $K = 2n+1$. Let $R = R(u)$ be as in Lemma 4.2. Then $R \ln R$ is integrable and

$$\begin{aligned}
\mathbb{E}[R \ln R] &= \frac{1}{2}\mathbb{E}\left[\|u\|^2\right] \\
&\leq \frac{\|x - \tilde{x}\|_2^2}{2T} + \left(6\sqrt{n} + \frac{4}{\sqrt{n}}\right)^2 \left(\frac{1}{T^2} \left\| z - \tilde{z} - \frac{1}{2} x \odot \tilde{x} \right\|^2 + \frac{2(n-1)}{3T} \|x - \tilde{x}\|_2^2 \right).
\end{aligned} \tag{4.27}$$

Here the choice $K = 2n+1$ is optimal as in (3.44).

Proof. Recall that $\|u\|^2 = \|\hat{U}_0^2\|_2 + \|\hat{U}\|^2 = \frac{\|x-\tilde{x}\|_2^2}{T} + \|\hat{U}\|^2$, $\hat{U}$ being defined as in (3.37) with $m = K$. First observe that the inequality in (4.27) comes from (3.41) and (3.42).

We will use (4.24) with $F(\omega) := \ln(R(u(\omega)))(\omega)$. Observe that

$$\begin{aligned}
\ln R(u(\omega - u(\omega)))(\omega - u(\omega)) &= \ln R(u(\omega))(\omega - u(\omega)) \\
&= -\langle \omega - u(\omega), u(\omega)\rangle - \frac{1}{2}\|u(\omega)\|^2 \\
&= -\langle \omega, u(\omega)\rangle + \frac{1}{2}\|u(\omega)\|^2
\end{aligned}$$

is $\mathbb{P}$-integrable thanks to Cauchy-Schwarz inequality, (3.41) and (3.42). We thus get that $\omega \mapsto \ln R(u(\omega))(\omega)$ is $R(u) \cdot \mathbb{P}$-integrable and

$$\begin{aligned}
&\mathbb{E}\left[R(u(\omega))(\omega) \ln R(u(\omega))(\omega)\right] \\
&= \mathbb{E}\left[-\langle \omega, u(\omega)\rangle + \frac{1}{2}\|u(\omega)\|^2 \right] \\
&= \mathbb{E}\left[-\mathbb{E}[\langle \omega, u(\omega)\rangle | \mathscr{G}] + \frac{1}{2}\|u(\omega)\|^2 \right] \quad \text{with } \mathscr{G} \text{ defined in the proof of Lemma 4.2} \\
&= \mathbb{E}\left[\frac{1}{2}\|u(\omega)\|^2 \right]
\end{aligned}$$

since $\mathbb{E}[\langle \omega, u(\omega)\rangle | \mathscr{G}] = 0$: $u(\omega)$ is $\mathscr{G}$-measurable and conditioned to $\mathscr{G}$ $\langle \omega, u(\omega)\rangle$ is Gaussian and centered. □

In the sequel, we will need the solution $\hat{\mathscr{U}}$ defined by (4.16) to have moments of any order. To get this integrability condition, we will have to consider the case $K = +\infty$.

We first set two preparatory lemmas.

Lemma 4.4. Let $h>0$. Let $(Y_\ell)_{\ell\geq 1}$ be a sequence of independent gamma $\Gamma(h,1)$-distributed real-valued random variables with the same shape parameter h and with rate parameter 1. Define

$$S_h := \frac{2}{\pi^2}\sum_{\ell=1}^{\infty}\frac{Y_\ell}{\ell^2}. \tag{4.28}$$

For $a>0$, one has

$$\mathbb{E}\left[S_h^{-a}\right] = 2^{1+h-a}\frac{\Gamma(2a+h)}{\Gamma(h)\Gamma(a)}\sum_{n=0}^{\infty}\frac{\Gamma(n+h)}{\Gamma(n+1)}\frac{1}{(2n+h)^{2a+h}}. \tag{4.29}$$

In particular, we have

$$\mathbb{E}\left[S_1^{-a}\right] \leq \frac{(4a+1)\Gamma(2a+1)}{2^a\Gamma(a+1)} \tag{4.30}$$

and

$$\mathbb{E}\left[S_{\frac{1}{2}}^{-\frac{1}{2}}\right] \leq 2\sqrt{2}+\frac{\sqrt{2}}{2}. \tag{4.31}$$

Proof. The Laplace transform of S_h is given by

$$\forall\,\lambda>0,\quad \mathbb{E}\left[e^{-\lambda S_h}\right] = \left(\frac{\sqrt{2\lambda}}{\sinh\sqrt{2\lambda}}\right)^h, \tag{4.32}$$

see [10]. On the other hand, making the change of variable $u=S_h\lambda$ in the following integral gives

$$\mathbb{E}\left[\int_0^\infty \lambda^{a-1}e^{-\lambda S_h}\,d\lambda\right] = \mathbb{E}\left[S_h^{-a}\int_0^\infty u^{a-1}e^{-u}\,d\lambda\right] = \mathbb{E}\left[S_h^{-a}\right]\Gamma(a).$$

From this we obtain

$$\begin{aligned}
\mathbb{E}\left[S_h^{-a}\right] &= \frac{1}{\Gamma(a)}\mathbb{E}\left[\int_0^\infty \lambda^{a-1}e^{-\lambda S_h}\,d\lambda\right] = \frac{1}{\Gamma(a)}\int_0^\infty \lambda^{a-1}\mathbb{E}\left[e^{-\lambda S_h}\right]\,d\lambda \\
&= \frac{1}{\Gamma(a)}\int_0^\infty \lambda^{a-1}\left(\frac{\sqrt{2\lambda}}{\sinh\sqrt{2\lambda}}\right)^h d\lambda = \frac{(2\sqrt{2})^h}{\Gamma(a)}\int_0^\infty \lambda^{a+\frac{h}{2}-1}\frac{1}{e^{h\sqrt{2\lambda}}\left(1-e^{-2\sqrt{2\lambda}}\right)^h}\,d\lambda \\
&= \frac{(2\sqrt{2})^h}{\Gamma(a)\Gamma(h)}\sum_{n=0}^{\infty}\frac{\Gamma(n+h)}{\Gamma(n+1)}\int_0^\infty \lambda^{a+\frac{h}{2}-1}e^{-(2n+h)\sqrt{2\lambda}}\,d\lambda,
\end{aligned}$$

by Fubini theorem and since for $h>0$ and $|x|<1$,

$$\frac{1}{(1-x)^h} = \frac{1}{\Gamma(h)}\sum_{n\geq 0}\frac{\Gamma(n+h)}{\Gamma(n+1)}x^n.$$

Making the change of variable $u=(2n+h)\sqrt{2\lambda}$ yields

$$\int_0^\infty \lambda^{a+\frac{h}{2}-1}e^{-(2n+h)\sqrt{2\lambda}}\,d\lambda = \frac{\Gamma(2a+h)}{2^{a+\frac{h}{2}-1}(2n+h)^{2a+h}}$$

and (4.29) follows.

In particular for $h=1$ we have

$$\mathbb{E}\left[S_1^{-a}\right] = 2^{2-a}\frac{\Gamma(2a+1)}{\Gamma(a)}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^{2a+1}}. \tag{4.33}$$

Now

$$\sum_{n=0}^{\infty}\frac{1}{(2n+1)^{2a+1}} \leq 1+\int_0^\infty \frac{dx}{(2x+1)^{2a+1}} = 1+\frac{1}{4a}$$

and (4.30) follows. For $h = 1/2$ and $a = 1/2$, one has

$$\mathbb{E}\left[S_{\frac{1}{2}}^{-\frac{1}{2}}\right] = 2\frac{\Gamma(\frac{3}{2})}{\Gamma(\frac{1}{2})^2}\sum_{n=0}^{\infty}\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)}\frac{1}{(2n+\frac{1}{2})^{2a+\frac{1}{2}}}. \tag{4.34}$$

Since for $n \geq 1$

$$\frac{\Gamma(n+\frac{1}{2})}{\Gamma(n+1)} \leq \frac{\Gamma(\frac{3}{2})}{\Gamma(2)} = \frac{\Gamma\left(\frac{1}{2}\right)}{2} = \frac{\sqrt{\pi}}{2} \text{ and } \frac{\Gamma(\frac{3}{2})}{\Gamma(\frac{1}{2})^2} = \frac{1}{2}\frac{1}{\Gamma(\frac{1}{2})} = \frac{1}{2\sqrt{\pi}},$$

one has

$$\mathbb{E}\left[S_{\frac{1}{2}}^{-\frac{1}{2}}\right] \leq 2^{3/2} + \frac{1}{2}\sum_{n=1}^{\infty}\frac{1}{(2n+\frac{1}{2})^{3/2}} \leq 2^{3/2} + \frac{1}{2}\int_0^{\infty}\frac{dx}{(2x+\frac{1}{2})^{3/2}} = 2\sqrt{2} + \frac{\sqrt{2}}{2},$$

which ends the proof of Lemma 4.4. □

Lemma 4.5. Let $(V_k)_{k\geq 1}$ be a sequence of $\mathbb{R}^n$-valued independent random variables with law $\mathcal{N}(0, I_n)$. Then for any $a > 0$,

$$\mathbb{E}\left[\operatorname{tr}\left(\left(\sum_{k=1}^{\infty}\frac{V_kV_k^t}{\beta_k^2}\right)^{-1}\right)^a\right] \leq \frac{(C_3(n))^a}{T^{2a}}\frac{(4a+1)\Gamma(2a+1)}{\pi^{2a}\Gamma(a+1)}. \tag{4.35}$$

with $C_3(n) = 4n^2(3n+4)^2$. Moreover, for any $p \in (0,1)$ and all $\lambda > 0$,we have

$$\mathbb{E}\left[\exp\left(\lambda\operatorname{tr}\left(\left(\sum_{k=1}^{\infty}\frac{V_kV_k^t}{\beta_k^2}\right)^{-1}\right)^p\right)\right] \leq 1 + \sum_{q=1}^{\infty}\left(\frac{(C_3(n))^p}{(T\pi)^{2p}}\lambda\right)^q\frac{(4pq+1)\Gamma(2pq+1)}{q!\Gamma(pq+1)} < \infty. \tag{4.36}$$

Note that here, in Lemma 4.5, the estimates do not seem optimal in term of the dimension n.

Proof of Lemma 4.5. We have

$$\sum_{k=1}^{\infty}\frac{V_kV_k^t}{\beta_k^2} \geq \sum_{\ell=1}^{\infty}\frac{1}{\beta_{\ell(n+1)}^2}\mathscr{M}_\ell \quad \text{with} \quad \mathscr{M}_\ell := \sum_{\ell'=(\ell-1)(n+1)+1}^{\ell(n+1)} V_{\ell'}V_{\ell'}^t.$$

The matrices $\mathscr{M}_\ell$ are Wishart $\mathscr{W}(n, n+1)$ with smallest eigenvalue $\lambda_{\min}(\mathscr{M}_\ell)$ having an exponential law with parameter $n/2$ or equivalently a law $\frac{2}{n}\Gamma(1,1)$. Consequently, by independence, we have

$$\lambda_{\min}\left(\sum_{k=1}^{\infty}\frac{V_kV_k^t}{\beta_k^2}\right) \geq \sum_{\ell=1}^{\infty}\frac{2Y_\ell}{n\beta_{\ell(n+1)}^2} \tag{4.37}$$

with Y_ℓ independent $\Gamma(1,1)$ random variables. Then using $\beta_k \leq \frac{2\sqrt{2}(3k+1)}{T}$ we can write

$$\lambda_{\min}\left(\sum_{k=1}^{\infty}\frac{V_kV_k^t}{\beta_k^2}\right) \geq T^2\sum_{\ell=1}^{\infty}\frac{n}{C_3(n)\ell^2}Y_\ell \tag{4.38}$$

with $C_3(n) = \frac{n^2}{2}\left(2\sqrt{2}(3n+4)\right)^2$. We have

$$\begin{aligned}\operatorname{tr}\left(\left(\sum_{k=1}^{\infty}\frac{V_kV_k^t}{\beta_k^2}\right)^{-1}\right) &\leq n\lambda_{\max}\left(\left(\sum_{k=1}^{\infty}\frac{V_kV_k^t}{\beta_k^2}\right)^{-1}\right)\\ &= n\left(\lambda_{\min}\left(\sum_{k=1}^{\infty}\frac{V_kV_k^t}{\beta_k^2}\right)\right)^{-1}\end{aligned}$$

Consequently, for $a > 0$,

$$\mathbb{E}\left[\operatorname{tr}\left(\left(\sum_{k=1}^{\infty}\frac{V_k V_k^t}{\beta_k^2}\right)^{-1}\right)^a\right] \leq \frac{C_3^a(n)}{T^{2a}}\mathbb{E}\left[\left(\sum_{\ell=1}^{\infty}\frac{Y_\ell}{\ell^2}\right)^{-a}\right]. \tag{4.39}$$

The estimate (4.35) thus directly follows from Lemma 4.4. We now turn to the exponential moments. Let $0 < p < 1$ and $\lambda > 0$, we have

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\lambda\left(\operatorname{tr}\left(\sum_{k=1}^{\infty}\frac{V_k V_k^t}{\beta_k^2}\right)^{-1}\right)^p\right)\right] &\leq \mathbb{E}\left[\exp\left(\left(\frac{C_3(n)}{T^2}\right)^p\lambda\left(\sum_{\ell=1}^{\infty}\frac{Y_\ell}{\ell^2}\right)^{-p}\right)\right] \\
&= 1 + \sum_{q=1}^{\infty}\frac{(C_3(n))^{pq}\lambda^q}{q!T^{2pq}}\mathbb{E}\left[\left(\sum_{\ell=1}^{\infty}\frac{Y_\ell}{\ell^2}\right)^{-pq}\right] \\
&\leq 1 + \sum_{q=1}^{\infty}\frac{(C_3(n))^{pq}\lambda^q}{q!T^{2pq}}\frac{(4pq+1)\Gamma(2pq+1)}{\pi^{2pq}\Gamma(pq+1)}
\end{aligned}$$

where we used Lemma 4.4 with $a = pq$. This is exactly the first inequality in (4.36). We are left to prove that the right hand side in (4.36) is finite. Using $\ln\Gamma(a) \sim a\ln(a)$ as $a \to \infty$ we get

$$\ln\left(\frac{(4pq+1)\Gamma(2pq+1)}{q!\Gamma(pq+1)}\right) \sim (2pq - q - pq)\ln(q) = q(p-1)\ln(q) < -\varepsilon q\ln(q)$$

with $\varepsilon = \frac{1-p}{2}$. Letting $\alpha = \left(\frac{C_3(n)}{\pi^2}\right)^p\lambda$ we have

$$\left(\left(\frac{C_3(n)}{\pi^2}\right)^p\lambda\right)^q\frac{(4pq+1)\Gamma(2pq+1)}{q!\Gamma(pq+1)} \leq \alpha^q q^{-\varepsilon q} \quad \text{for } q \text{ sufficiently large}$$

and $\sum_{q=1}^{\infty}\alpha^q q^{-\varepsilon q} < \infty$, proving the finiteness of the right hand side of (4.36). □

After these preliminary results, we now turn to the analytic consequence for the semi-group of this change of probability method. Let $f : \mathbb{G}_n \to \mathbb{R}$ a bounded measurable function. We recall that

$$P_T f(g) = \mathbb{E}[f(\mathbb{B}_T^g)] \quad \text{together with} \quad P_T f(\tilde{g}) = \mathbb{E}[f(\mathbb{B}_T^{\tilde{g}})R(u)]; \tag{4.40}$$

$\tilde{g}$ and u being related as in Lemma 4.2. But with our construction, we have a.s $\mathbb{B}_T^{\tilde{g}} = \mathbb{B}_T^g$, yielding

$$P_T f(\tilde{g}) = \mathbb{E}[f(\mathbb{B}_T^g)R(u)]. \tag{4.41}$$

From this and Corollary 4.3 we get the following log Harnack inequality.

Theorem 4.6. Let f be a positive function in $\mathbb{G}_n$, $T > 0$ and $g = (x, z), \tilde{g} = (\tilde{x}, \tilde{z}) \in \mathbb{G}_n$. Then

$$\begin{aligned}
&P_T(\ln f)(\tilde{g}) \leq \ln(P_T f(g)) \\
&+ \frac{\|x-\tilde{x}\|_2^2}{2T} + \left(6\sqrt{n} + \frac{4}{\sqrt{n}}\right)^2\left(\frac{1}{T^2}\left\|z - \tilde{z} - \frac{1}{2}x\odot\tilde{x}\right\|^2 + \frac{2(n-1)}{3T}\|x-\tilde{x}\|_2^2\right).
\end{aligned} \tag{4.42}$$

Proof. Again take $K = 2n+1$. By Equation (4.40) applied to $\ln f$ and Young inequality,

$$\begin{aligned}
P_T(\ln f)(\tilde{g}) &= \mathbb{E}[\ln f(\mathbb{B}_T^g)R(u)] \\
&\leq \mathbb{E}[R(u)\ln R(u)] + \ln\mathbb{E}\left[\exp\ln f(\mathbb{B}_T^g)\right] \\
&= \mathbb{E}[R(u)\ln R(u)] + \ln(P_T f(g)).
\end{aligned}$$

We conclude with (4.27). □

The next theorem aims at establishing an integration by parts formula for the derivative of the semi-group.

Theorem 4.7. Fix $K \geq n+2$ (and possibly infinite). Let $f : \mathbb{G}_n \to \mathbb{R}$ be a bounded continuous function, $g = (x,z) \in \mathbb{G}_n$, $h = (h_x, h_z) \in T_g\mathbb{G}_n = \mathbb{G}_n$. Denote $\tilde{g} = g + h$ we have

$$d_g P_T f(h) = \mathbb{E}\left[f(\mathbb{B}_T^g) \left(- \sum_{k \in J_K} \left\langle \xi_{3k}, \hat{U}_k \right\rangle \right) \right], \tag{4.43}$$

where $(\hat{U}_k = \hat{U}_k(\tilde{g}))_{k \geq 0}$ is given by (4.9).

We then deduce reverse Poincaré inequalities.

Theorem 4.8. With the same notation as in Theorem 4.7. For any $p \in (1, \infty]$, denoting $q \in [1, \infty)$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$|d_g P_T f(h)| \leq (P_T |f|^p)^{1/p}\, m_q\, \mathbb{E}\left[\left(\sum_{k \in J_K} \left\| \hat{U}_k \right\|_2^2 \right)^{q/2} \right]^{1/q}. \tag{4.44}$$

with $m_q^q = \mathbb{E}[|Z|^q]$ the q-th moment of a $\mathscr{N}(0,1)$-variable Z. The right hand side is finite for all $q \geq 1$ when $K = \infty$.

In the special case $p = q = 2$, we get the reverse Poincaré inequality

$$\begin{aligned} &|d_g P_T f(h)|^2 \\ &\leq \left(P_T |f|^2\right) \left(\frac{\|h_x\|_2^2}{T} + \left(6\sqrt{2n} + \frac{4\sqrt{2}}{\sqrt{n}} \right)^2 \left(\frac{1}{T^2} \|h_z - \frac{1}{2} x \odot h_x\|^2 + \frac{2(n-1)}{3T} \|h_x\|_2^2 \right) \right). \end{aligned} \tag{4.45}$$

Proof of Theorem 4.7. Considering a vector $h = (h_x, h_z) \in \mathbb{G}_n$, we will compute

$$\lim_{a \to 0} \frac{1}{a} \left(P_T f(g + ah) - P_T f(g) \right). \tag{4.46}$$

Denote $\tilde{g}(a) = (\tilde{x}(a), \tilde{z}(a)) = g + ah$. The matrix $W(\tilde{g}(a))$ defined in (3.33) rewrites as

$$W(\tilde{g}(a)) = z - \tilde{z}(a) - \frac{1}{2} \tilde{x}_a \odot x + (x - \tilde{x}(a)) \odot \left(\frac{\sqrt{T}}{2} \xi_0 - \sqrt{T} \alpha_0 \tilde{\xi}_1 \right) \tag{4.47}$$

and since $x \odot x = 0$,

$$\begin{aligned} \forall a \in \mathbb{R}, \quad d_{\tilde{g}(a)} W(h) = \frac{d}{da} W(\tilde{g}(a)) &= -h_z - \frac{1}{2} h_x \odot x - h_x \odot \left(\frac{\sqrt{T}}{2} \xi_0 - \sqrt{T} \alpha_0 \tilde{\xi}_1 \right) \\ &= W(\tilde{g}(1)) \text{ not depending on } a. \end{aligned} \tag{4.48}$$

Consequently, with the notation of (4.12),

$$d_{\tilde{g}(a)} \hat{\mathscr{U}}^t(h) = -\frac{1}{2} \hat{\mathscr{V}}^t (\hat{\mathscr{V}} \hat{\mathscr{V}}^t)^{-1} d_{\tilde{g}(a)} \mathscr{W}^t(h) = -\frac{1}{2} \hat{\mathscr{V}}^t (\hat{\mathscr{V}} \hat{\mathscr{V}}^t)^{-1} \mathscr{W}^t(\tilde{g}(1)) \tag{4.49}$$

does not depend on a.

Letting $\hat{\mathscr{U}}^t = \hat{\mathscr{U}}^t(\tilde{g}(1))$, $\mathscr{W}^t = \mathscr{W}^t(\tilde{g}(1))$ and $(u_0, u_3, u_6, ...) = \left(\hat{U}_0(\tilde{g}(1)), \hat{U}_1(\tilde{g}(1)), \hat{U}_2(\tilde{g}(1)), ... \right)$,

$$\hat{\mathscr{U}}^t(\tilde{g}(a)) = a \hat{\mathscr{U}}^t = -\frac{a}{2} \hat{\mathscr{V}}^t (\hat{\mathscr{V}} \hat{\mathscr{V}}^t)^{-1} \mathscr{W}^t. \tag{4.50}$$

also, $\hat{U}_0(\tilde{g}(a)) = a\frac{-h_x}{\sqrt{T}}$ yielding $\frac{d}{da}\hat{U}_0(\tilde{g}(a)) = u_0$. Then using (4.20) and the fact that $\hat{\mathscr{U}} = (u_3, u_6, \ldots)$ we get

$$\begin{aligned}
\frac{1}{a}\left(P_T f(g+ah) - P_T f(g)\right) &= \frac{1}{a}\mathbb{E}\left[f(\mathbb{B}_T^g)\left(R(au) - 1\right)\right] \\
&= \frac{1}{a}\mathbb{E}\left[f(\mathbb{B}_T^g)\left(\int_0^a \frac{d}{da'}R(a'u)\,da'\right)\right] \\
&= -\frac{1}{a}\mathbb{E}\left[f(\mathbb{B}_T^g(\omega))\left(\int_0^a R(a'u)\langle \omega + a'u, u\rangle\,da'\right)\right]
\end{aligned}$$

By definition of $R(a'u)$ we have as soon as $\omega \mapsto F(\omega)$ is $\mathbb{P}$-integrable, that $\omega \mapsto F(\omega + a'u)$ is $R(a'u)\mathbb{P}$-integrable and

$$\mathbb{E}[R(a'u)F(\omega + a'u)] = \mathbb{E}[F(\omega)]. \tag{4.51}$$

In our situation f is bounded and $\langle \omega, u\rangle = \left\langle \omega, \frac{u}{\|u\|}\right\rangle \|u\|$ is $\mathbb{P}$-integrable since, conditioned to $\mathscr{G}$ $\left\langle \omega, \frac{u}{\|u\|}\right\rangle$ has law $\mathcal{N}(0,1)$, $\|u\| \leq \frac{\|h_x\|_2}{\sqrt{T}} + \|\hat{\mathscr{U}}\|$,

$$\|\hat{\mathscr{U}}\| = \sqrt{\operatorname{tr}\left(\hat{\mathscr{U}}\hat{\mathscr{U}}^t\right)} = \frac{1}{2}\sqrt{\operatorname{tr}\left(\mathscr{W}^t(\hat{\mathscr{V}}\hat{\mathscr{V}}^t)^{-1}\mathscr{W}\right)} \leq \frac{1}{2}\|\mathscr{W}\|\left(\operatorname{tr}\left(\left(\hat{\mathscr{V}}\hat{\mathscr{V}}^t\right)^{-1}\right)\right)^{1/2},$$

$\mathscr{W}$ is Gaussian and independent of $\hat{\mathscr{V}}$ and

- if $K = \infty$ then by Equation (4.36) $\left(\operatorname{tr}\left(\left(\hat{\mathscr{V}}\hat{\mathscr{V}}^t\right)^{-1}\right)\right)^{1/2}$ has exponential moments,
- if $K < \infty$ then $\left(\operatorname{tr}\left(\left(\hat{\mathscr{V}}\hat{\mathscr{V}}^t\right)^{-1}\right)\right)^{1/2} \leq \beta_K\left(\operatorname{tr}\left((\mathscr{V}\mathscr{V}^t)^{-1}\right)\right)^{1/2}$ which is integrable by (3.20), since we choose $K \geq n+2$.

So we can apply equality (4.51) after exchanging the orders of integration (which is allowed here for the same integrability reasons), and we get

$$\begin{aligned}
\frac{1}{a}\left(P_T f(g+ah) - P_T f(g)\right) &= -\frac{1}{a}\int_0^a \mathbb{E}\left[f(\mathbb{B}_T^g(\omega))\left(R(a'u)\left\langle \omega + a'u, u\right\rangle\right)\right]\,da' \\
&= -\frac{1}{a}\int_0^a \mathbb{E}\left[f(\mathbb{B}_T^g(\omega - a'u))\left\langle \omega, u\right\rangle\right]\,da' \\
&= -\mathbb{E}\left[\left(\frac{1}{a}\int_0^a f(\mathbb{B}_T^g(\omega - a'u))\,da'\right)\left\langle \omega, u\right\rangle\right].
\end{aligned}$$

Since f is bounded and continuous, and a.s. $\mathbb{B}_T^g(\omega - a'u) \to \mathbb{B}_T^g(\omega)$ as $a' \to 0$ we can use the dominated convergence theorem to obtain

$$\lim_{a\to 0}\frac{1}{a}\left(P_T f(g+ah) - P_T f(g)\right) = -\mathbb{E}\left[f(\mathbb{B}_T^g(\omega))\left\langle \omega, u\right\rangle\right] \tag{4.52}$$

which yields (4.43). □

Proof of Theorem 4.8. To establish (4.44) we first use Hölder inequality which yields

$$|d_g P_T f(h)| \leq \mathbb{E}\left[|f|^p(\mathbb{B}_T^g)\right]^{1/p}\mathbb{E}\left[|-\langle \omega, u\rangle|^q\right]^{1/q}. \tag{4.53}$$

As in the proof of Corollary 4.3, conditioning with respect to $\mathscr{G}$ we get

$$\begin{aligned}
\mathbb{E}\left[|-\langle \omega, u\rangle|^q\right] &= \mathbb{E}\left[\mathbb{E}\left[|-\langle \omega, u\rangle|^q \,\middle|\, \mathscr{G}\right]\right] \\
&= \mathbb{E}\left[\|u\|^q m_q^q\right]
\end{aligned}$$

with $m_q^q = \mathbb{E}[|Z|^q]$ the q-th moment of a $\mathscr{N}(0,1)$-variable Z. In particular $\|u\|^2 = \sum_{k\in J_K} \left\|\hat{U}_k\right\|_2^2$ which proves (4.44). Notice that when $K=\infty$ the last term is finite thanks to Lemma 4.5 which implies that all moments of $(\hat{\mathscr{V}}\hat{\mathscr{V}}^t)^{-1}$ are finite. Finally, to prove (4.45) we apply (4.44) with $K = 2n+1$ which allows to use (3.43) and (3.44) with $\tilde{g} = g+h$. □

The next corollary complements Theorem 4.8 with a kind of weak inverse log-Sobolev inequality.

Corollary 4.9. With the same notation as in Theorem 4.7, we have for all $\delta > 0$ and nonnegative continuous function f,

$$|d_g P_T f(h)| \le \delta P_T\left(f\ln\left(\frac{f}{P_T f(g)}\right)\right)(g) + \frac{1}{2\delta}\mathbb{E}\left[f(\mathbb{B}_T^g)\left(\sum_{k\in J_K}\left\|\hat{U}_k\right\|^2\right)\right]. \tag{4.54}$$

In particular,

$$|d_g P_T f(h)| \le \sqrt{2P_T\left(f\ln\left(\frac{f}{P_T f(g)}\right)\right)(g)\mathbb{E}\left[f(\mathbb{B}_T^g)\left(\sum_{k\in J_K}\left\|\hat{U}_k\right\|^2\right)\right]}. \tag{4.55}$$

Proof. Again we start with Equation (4.43). As already seen in Equation (4.49), the random vectors $\hat{U}_k = \hat{U}_k(h)$, $k \in J_K$, depend linearly on h. Moreover the right-hand-side of (4.54) and (4.55) is the same for h and $-h$. Consequently, possibly changing h into $-h$, it is enough to establish (4.54) and (4.55) for h satisfying $d_g P_T f(h) \ge 0$, or equivalently to replace $|d_g P_T f(h)|$ by $d_g P_T f(h)$ in the left-hand-side.

Conditioning the right-hand-side of Equation (4.43) with respect to $\mathscr{G}$ and using the Young inequality from e.g. Lemma 2.4. in [2], we obtain

$$\begin{aligned} d_g P_T f(h) &= \mathbb{E}\left[f(\mathbb{B}_T^g)\left(-\sum_{k\in J_K}\left\langle \xi_{3k}, \hat{U}_k\right\rangle\right)\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[f(\mathbb{B}_T^g)\left(-\sum_{k\in J_K}\left\langle \xi_{3k}, \hat{U}_k\right\rangle\right)\Big|\mathscr{G}\right]\right] \\ &\le \mathbb{E}\left[\delta\mathbb{E}\left[f(\mathbb{B}_T^g)\ln\left(\frac{f(\mathbb{B}_T^g)}{\mathbb{E}\left[f(\mathbb{B}_T^g)|\mathscr{G}\right]}\right)\Big|\mathscr{G}\right] + \delta\mathbb{E}\left[f(\mathbb{B}_T^g)\Big|\mathscr{G}\right]\ln\mathbb{E}\left[e^{-\frac{1}{\delta}\sum_{k\in J_K}\left\langle \xi_{3k}, \hat{U}_k\right\rangle}\Big|\mathscr{G}\right]\right]. \end{aligned}$$

Now since conditioning with respect to $\mathscr{G}$ transforms $-\sum_{k\in J_K}\left\langle \xi_{3k}, \hat{U}_k\right\rangle$ into a centered Gaussian variable we get

$$\mathbb{E}\left[e^{-\frac{1}{\delta}\sum_{k\in J_K}\left\langle \xi_{3k}, \hat{U}_k\right\rangle}\Big|\mathscr{G}\right] = e^{\frac{1}{2\delta^2}\sum_{k\in J_K}\|\hat{U}_k\|^2}$$

which yields

$$\delta\mathbb{E}\left[f(\mathbb{B}_T^g)\Big|\mathscr{G}\right]\ln\mathbb{E}\left[e^{-\frac{1}{\delta}\sum_{k\in J_K}\left\langle \xi_{3k}, \hat{U}_k\right\rangle}\Big|\mathscr{G}\right] = \frac{1}{2\delta}\mathbb{E}\left[f(\mathbb{B}_T^g)\left(\sum_{k\in J_K}\left\|\hat{U}_k\right\|^2\right)\Big|\mathscr{G}\right] \tag{4.56}$$

since $\sum_{k\in J_K}\left\|\hat{U}_k\right\|^2$ is $\mathscr{G}$-measurable. Also letting $Y = \mathbb{E}\left[f(\mathbb{B}_T^g)|\mathscr{G}\right]$ and using by Jensen's inequality $\mathbb{E}[Y\ln Y] \ge \mathbb{E}[Y]\ln\mathbb{E}[Y]$ we get

$$\mathbb{E}\left[\mathbb{E}\left[f(\mathbb{B}_T^g)\ln\left(\frac{f(\mathbb{B}_T^g)}{\mathbb{E}\left[f(\mathbb{B}_T^g)|\mathscr{G}\right]}\right)\Big|\mathscr{G}\right]\right] \le P_T\left(f\ln\left(\frac{f}{\mathbb{P}_T f(g)}\right)\right)(g). \tag{4.57}$$

From (4.56) and (4.57) we get (4.54). Finally, (4.55) is obtained with

$$\delta = \sqrt{\frac{\mathbb{E}\left[f(\mathbb{B}_T^g)\left(\sum_{k\in J_K}\left\|\hat{U}_k\right\|^2\right)\right]}{2P_T\left(f\ln\left(\frac{f}{\mathbb{P}_T f(g)}\right)\right)(g)}}.$$

□

As a final corollary we provide estimates of the horizontal and vertical differential of the heat kernel $(g,\tilde{g})\mapsto p_t(g,\tilde{g})$ on $\mathbb{G}_n$.

Corollary 4.10. There exist three positive constants $K(n)$, $K_1(n)$ and $K_2(n)$ only depending on n such that:

$$|d_g p_t(0,\cdot)(h)| \le t^{-\frac{n^2}{2}} e^{-\frac{K(n)}{t}d_{cc}(0,g)^2}\left(K_1(n)\frac{\|h_x\|_2}{\sqrt{t}} + K_2(n)\frac{\|h_z - \frac{1}{2}x\odot h_x\|}{t}\right). \tag{4.58}$$

Proof. From [25], see also [6], there exist some positive constants $\tilde{K}(n)$ and $\tilde{K}_1(n)$ depending on n such that:

$$p_t(g,\tilde{g}) \le \frac{\tilde{K}_1(n)}{t^{\frac{n^2}{2}}} e^{-\frac{\tilde{K}(n)}{t}d_{cc}(g,\tilde{g})^2}. \tag{4.59}$$

Set $g\in\mathbb{G}_n$ then $p_t(0,g) = P_{\frac{t}{2}}(p_{\frac{t}{2}}(0,\cdot))(g)$. For any $h\in T_g\mathbb{G}_n = \mathbb{G}_n$, using the reverse Poincaré inequality from Theorem 4.8 with $f = p_{\frac{t}{2}}(0,\cdot)$:

$$\begin{aligned} |d_g p_t(0,\cdot)(h)|^2 &\le \mathbb{E}\left[p_{\frac{t}{2}}(0,\mathbb{B}^g_{\frac{t}{2}})^2\right] \\ &\times\left(\frac{2\|h_x\|_2^2}{t} + \left(6\sqrt{2n}+\frac{4\sqrt{2}}{\sqrt{n}}\right)^2\left(\frac{4}{t^2}\|h_z - \frac{1}{2}x\odot h_x\|^2 + \frac{4(n-1)}{3t}\|h_x\|_2^2\right)\right). \end{aligned} \tag{4.60}$$

We now examine $\mathbb{E}\left[p_{\frac{t}{2}}(0,\mathbb{B}^g_{\frac{t}{2}})^2\right]$. Using (4.59):

$$\begin{aligned} \mathbb{E}\left[p_{\frac{t}{2}}(0,\mathbb{B}^g_{\frac{t}{2}})^2\right] &= \int_{\mathbb{G}_n} p_{\frac{t}{2}}(0,l)^2 p_{\frac{t}{2}}(g,l)dl \\ &\le \int_{\mathbb{G}_n}\tilde{K}_1(n)^3\left(\frac{t}{2}\right)^{-\frac{3n^2}{2}} e^{-\frac{2\tilde{K}(n)}{t}(2d_{cc}(0,l)^2+d_{cc}(g,l)^2)}dl \\ &\le \int_{\mathbb{G}_n}\tilde{K}_1(n)^3\left(\frac{t}{2}\right)^{-\frac{3n^2}{2}} e^{-\frac{2\tilde{K}(n)}{t}\left(2d_{cc}(0,l)^2+(d_{cc}(g,0)-d_{cc}(0,l))^2\right)}dl \\ &\le \tilde{K}_1(n)^3\left(\frac{t}{2}\right)^{-\frac{3n^2}{2}}\int_{\mathbb{G}_n} e^{-\frac{2\tilde{K}(n)}{t}d_{cc}(0,l)^2}dl\, e^{-\frac{\tilde{K}(n)}{t}d_{cc}(0,g)^2} \end{aligned} \tag{4.61}$$

where the last expression is obtained by using the inequality:

$$\begin{aligned} 2a^2+(a-b)^2 = 3a^2+b^2-2ab &\ge 3a^2+b^2-\frac{a^2}{\lambda}-\lambda b^2 \\ &= a^2+\frac{b^2}{2} \text{ with } \lambda = 1/2. \end{aligned}$$

Using the property of the dilation on $(\mathbb{G}_n, d_{cc})$, $\frac{1}{\sqrt{t}}d_{cc}(0,l) = d_{cc}(0,\mathrm{dil}_{\frac{1}{\sqrt{t}}}(l))$, and since the homogeneous dimension of $\mathbb{G}_n$ is n^2, we have:

$$\left(\frac{t}{2}\right)^{-\frac{n^2}{2}}\int_{\mathbb{G}_n} e^{-\frac{2\tilde{K}(n)}{t}d_{cc}(0,l)^2}dl = \int_{\mathbb{G}_n} e^{-2\tilde{K}(n)d_{cc}(0,l)^2}dl$$

which is finite and does not depend on t. The expected result follows with $K(n) = \frac{\tilde{K}(n)}{2}$. □

Remark 7. If in the above proof, one uses the reverse Poincaré inequality (4.44) with $p = 1 + \varepsilon$ for $\varepsilon > 0$ (and with $K = +\infty$), it is possible to obtain (4.58) with some constants $K(n, \varepsilon)$, $K_1(n, \varepsilon)$ and $K_2(n, \varepsilon)$ depending only on n and on ε with

$$K(n, \varepsilon) = \frac{\tilde{K}(n)}{1 + \varepsilon}$$

and where $K_1(n, \varepsilon)$ and $K_2(n, \varepsilon)$ tend to infinity as $\varepsilon \to 0$.

References

[1] Marc Arnaudon and Anton Thalmaier. The differentiation of hypoelliptic diffusion semigroups. *Illinois J. Math.*, 54(4):1285–1311, 2010.

[2] Marc Arnaudon, Anton Thalmaier, and Feng-Yu Wang. Gradient estimates and Harnack inequalities on non-compact Riemannian manifolds. *Stochastic Process. Appl.*, 119(10):3653–3670, 2009.

[3] Sayan Banerjee, Maria Gordina, and Phanuel Mariano. Coupling in the Heisenberg group and its applications to gradient estimates. *Ann. Probab.*, 46(6):3275–3312, 2018.

[4] Sayan Banerjee and Wilfrid S. Kendall. Coupling the Kolmogorov diffusion: maximality and efficiency considerations. *Adv. in Appl. Probab.*, 48(A):15–35, 2016.

[5] Fabrice Baudoin and Michel Bonnefont. Log-Sobolev inequalities for subelliptic operators satisfying a generalized curvature dimension inequality. *J. Funct. Anal.*, 262(6):2646–2676, 2012.

[6] Fabrice Baudoin and Michel Bonnefont. Reverse Poincaré inequalities, isoperimetry, and Riesz transforms in Carnot groups. *Nonlinear Anal.*, 131:48–59, 2016.

[7] Fabrice Baudoin and Nicola 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.

[8] Gérard Ben Arous, Michael Cranston, and Wilfrid 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.

[9] Ph. Biane and M. Yor. Variations sur une formule de Paul Lévy. *Ann. Inst. H. Poincaré Probab. Statist.*, 23(2):359–377, 1987.

[10] Philippe Biane, Jim Pitman, and Marc Yor. Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions. *Bull. Amer. Math. Soc. (N.S.)*, 38(4):435–465, 2001.

[11] Magalie Bénéfice. Non co-adapted couplings of Brownian motions on subRiemannian manifolds. https://arxiv.org/abs/2312.14512, 2023.

[12] Magalie Bénéfice. Couplings of brownian motions on $SU(2)$ and $SL(2, \mathbb{R})$. Stochastic Process. Appl. **176**, Paper No. 104434, 20 pp.; MR4777035, 2024.

[13] Magalie Bénéfice. Non co-adapted successful couplings of Brownian motions on the free, step 2 carnot groups. https://arxiv.org/abs/2407.06593, 2024.

[14] James Foster and Karen Habermann. Brownian bridge expansions for Lévy area approximations and particular values of the Riemann zeta function. *Combin. Probab. Comput.*, 32(3):370–397, 2023.

[15] James Foster, Terry Lyons, and Harald Oberhauser. An optimal polynomial approximation of Brownian motion. *SIAM J. Numer. Anal.*, 58(3):1393–1421, 2020.

[16] Arnaud Guillin and Feng-Yu Wang. Degenerate Fokker-Planck equations: Bismut formula, gradient estimate and Harnack inequality. *J. Differential Equations*, 253(1):20–40, 2012.

[17] Karen Habermann. A semicircle law and decorrelation phenomena for iterated Kolmogorov loops. *J. Lond. Math. Soc. (2)*, 103(2):558–586, 2021.

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

[19] Wilfrid 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.

[20] D.F. Kuznetsov. New simple method of expansion of iterated ito stochastic integrals of multiplicity 2 based on expansion of the brownian motion using legendre polynomials and trigonometric functions. *arXiv*, 2018.

[21] Torgny Lindvall. *Lectures on the coupling method*. Dover Publications, Inc., Mineola, NY, 2002. Corrected reprint of the 1992 original.

[22] Liangbing Luo and Robert W. Neel. Non-markovian maximal couplings and a vertical reflection principle on a class of sub-riemannian manifolds. https://arxiv.org/abs/2402.13976, 2024.

[23] Robb J. Muirhead. *Aspects of multivariate statistical theory*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1982.

[24] Jolanta Pielaszkiewicz and Thomas Holgersson. Mixtures of traces of Wishart and inverse Wishart matrices. *Comm. Statist. Theory Methods*, 50(21):5084–5100, 2021.

[25] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon. *Analysis and geometry on groups*, volume 100 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1992.

Acknowledgments. We want to sincerely thank the editor and the three anonymous reviewers for their handling of the review, and their detailed and thoughtful comments and suggestions which have significantly improved the article.