HAL
open science

On logarithmic Sobolev inequalities for the heat kernel on the Heisenberg group

Michel Bonnefont, Djalil Chafaï, Ronan Herry

To cite this version:

Michel Bonnefont, Djalil Chafaï, Ronan Herry. On logarithmic Sobolev inequalities for the heat kernel on the Heisenberg group. Annales de la Faculté des Sciences de Toulouse. Mathématiques., 2020, Série 6 Tome 29 (2), pp. 335-355. 10.5802/afst.1633 . hal-01343646v3

HAL Id: hal-01343646

https://hal.science/hal-01343646v3

Submitted on 1 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON LOGARITHMIC SOBOLEV INEQUALITIES FOR THE HEAT KERNEL ON THE HEISENBERG GROUP

MICHEL BONNEFONT, DJALIL CHAFAÏ, AND RONAN HERRY

Abstract. In this note, we derive a new logarithmic Sobolev inequality for the heat kernel on the Heisenberg group. The proof is inspired from the historical method of Leonard Gross with the Central Limit Theorem for a random walk. Here the non commutative nature of the increments produces a new gradient which naturally involves a Brownian bridge on the Heisenberg group. This new inequality contains the optimal logarithmic Sobolev inequality for the Gaussian distribution in two dimensions. We compare this new inequality with the sub-elliptic logarithmic Sobolev inequality of Hong-Quan Li and with the more recent inequality of Fabrice Baudoin and Nicola Garofalo obtained using a generalized curvature criterion. Finally, we extend this inequality to the case of homogeneous Carnot groups of rank two.

Résumé. Dans cette note, nous obtenons une inégalité de Sobolev logarithmique nouvelle pour le noyau de la chaleur sur le groupe de Heisenberg. La preuve est inspirée de la méthode historique de Leonard Gross à base de théorème limite central pour une marche aléatoire. Ici la nature non commutative des incréments produit un nouveau gradient qui fait intervenir naturellement un pont brownien sur le groupe de Heisenberg. Cette nouvelle inégalité contient l'inégalité de Sobolev logarithmique optimale pour la mesure gaussienne en deux dimensions. Nous comparons cette nouvelle inégalité avec l'inégalité sous-elliptique de Hong-Quan Li et avec les inégalités plus récentes de Fabrice Beaudoin et Nicola Garofalo obtenues avec un critère de courbure généralisé. Enfin nous étendons notre inégalités au cas des groupes de Carnot homogène de rang deux.

Contents

1. The Heisenberg group and our main result 1
The model 2
Main results 5
Structure of the paper 6
2. Discussion and comparison with other inequalities 6
3. Proof of Theorem 1.1 8
4. Proof of Corollary 1.2 9
5. Proof of inequalities (2.4) and (2.5) 10
6. Extension to homogeneous Carnot groups of rank two 12
Acknowledgements 12
References 13

1. The Heisenberg group and our main result

In this note, we derive a new logarithmic Sobolev inequality for the heat kernel on the Heisenberg group (Theorem 1.1). Our proof is inspired from the historical method of Leonard Gross based on a random walk and a Central Limit Theorem. Due to the non commutative nature of the group structure, the energy which appears in the right hand side involves an integral over some Brownian bridges on the Heisenberg group. To compare with other logarithmic Sobolev

Date: July 2016, revised September 2017, revised March 2018, to appear in *Annales de la Faculté des sciences de Toulouse : Mathématiques, Série 6, Tome 29 (2020) no. 2, pp. 335-355*, compiled September 1, 2020.

2010 *Mathematics Subject Classification.* 22E30; 35R03; 35A23; 60J65.

Key words and phrases. Heisenberg group; Heat kernel; Brownian Motion; Poincaré inequality; Logarithmic Sobolev inequality; Random Walk; Central Limit Theorem.

inequalities, we study Brownian bridges on the Heisenberg group and deduce a weighted logarithmic Sobolev inequality (Corollary 1.2). This weighted inequality is close to the symmetrized version of the sub-elliptic logarithmic Sobolev inequality of Hong-Quan Li. We also compare with inequalities due to Fabrice Baudoin and Nicola Garofalo, and provide a short semigroup proof of these inequalities in the case of the Heisenberg group.

We choose to focus on the one dimensional Heisenberg group, for simplicity; and also because very precise estimates and results are known in this particular case, which helps to compare our new inequality with existing ones. Nevertheless our new logarithmic Sobolev inequality remains more generally valid for homogeneous Carnot groups of rank two (Theorem 6.1).

The model. Let us briefly introduce the model and its main properties. The Heisenberg group $\mathbb{H}$ is a remarkable simple mathematical object, with rich algebraic, geometric, probabilistic, and analytic aspects. Available in many versions (discrete or continuous; periodic or not), our work focuses on the continuous Heisenberg group $\mathbb{H}$, formed by the set of 3×3 matrices

$$M(a,b,c) = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}, \quad a,b,c \in \mathbb{R}.$$

The Heisenberg group $\mathbb{H}$ is a non commutative sub-group of the general linear group, with group operations $M(a,b,c)M(a',b',c') = M(a+a', b+b', c+c'+ab')$ and $M(a,b,c)^{-1} = (-a,-b,-c+ab)$. The neutral element $M(0,0,0)$ is called the origin. The Heisenberg group $\mathbb{H}$ is a Lie group i.e. a manifold compatible with group structure.

The Heisenberg algebra is stratified. The Lie algebra $\mathfrak{H}$ i.e. the tangent space at the origin of $\mathbb{H}$ is the sub-algebra of $\mathcal{M}_3(\mathbb{R})$ given by the 3×3 matrices of the form

$$\begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix}, \quad x,y,z \in \mathbb{R}.$$

The canonical basis of $\mathbb{H}$

$$X := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad Z := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

satisfies an abstract version of the Dirac (or annihilation-creation) commutation relation

$$[X,Y] := XY - YX = Z \quad \text{and} \quad [X,Z] = [Y,Z] = 0.$$

This relation shows that the Lie algebra $\mathfrak{H}$ is stratified

$$\mathfrak{H} = \mathfrak{H}_0 \oplus \mathfrak{H}_1,$$

where $\mathfrak{H}_0 = \mathrm{span}(X,Y)$ and $\mathfrak{H}_1 = \mathrm{span}(Z)$ is the center of $\mathfrak{H}_0$. This makes the Baker-Campbell-Hausdorff formula on $\mathfrak{H}$ particularly simple:

$$\exp(A)\exp(B) = \exp\left(A + B + \frac{1}{2}[A,B]\right), \quad A,B \in \mathfrak{H}.$$

Exponential coordinates. Lie groups such as $\mathbb{H}$ with stratified Lie algebra (that is Carnot groups) have a diffeomorphic exponential map $\exp : A \in \mathfrak{H} \mapsto \exp(A) \in \mathbb{H}$. This identification of $\mathbb{H}$ with $\mathfrak{H}$, namely

$$\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \equiv \exp \begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} = \exp(xX + yY + zZ),$$

allows to identify $\mathbb{H}$ with $\mathbb{R}^3$ equipped with the group structure

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

and $(x,y,z)^{-1} = (-x,-y,-z)$. The identity element is the "origin" $e := (0,0,0)$. From now on, we use these "exponential coordinates". Geometrically, the quantity $\frac{1}{2}(xy' - yx')$ is the algebraic area in $\mathbb{R}^2$ between a piecewise linear path and its chord namely the area between

$$[(0,0),(x,y)] \cup [(x,y),(x+x',y+y')] \quad \text{and} \quad [(0,0),(x+x',y+y')].$$

This area is zero if (x,y) and (x',y') are collinear. The group product

$$(x,y,0)(x',y',0) = (x+x', y+y', \frac{1}{2}(xy' - yx'))$$

in $\mathbb{H}$ encodes the sum of increments in $\mathbb{R}^2$ and computes automatically the generated area.

Vector fields on $\mathbb{H}$. Elements of $\mathfrak{H}$ can classically be extended to left-invariant vector fields. This identification will always be made implicitly and the same notation for element of $\mathfrak{H}$ and vector field is used. This gives for the canonical basis at a point (x,y,z)

$$X := \partial_x - \frac{y}{2}\partial_z, \quad Y := \partial_y + \frac{x}{2}\partial_z, \quad Z := \partial_z. \tag{1.1}$$

Metric structure of $\mathbb{H}$. On the Heisenberg group, a natural distance associated to the left-invariant diffusion operator $L = \frac{1}{2}(X^2 + Y^2 + \beta^2 Z^2)$, $\beta \geq 0$, is defined for all $h, g \in \mathbb{H}$ by

$$d(h,g) := \sup_f (f(h) - f(g))$$

where the supremum runs over all $f \in \mathcal{C}^\infty(\mathbb{H},\mathbb{R})$ such that

$$\Gamma(f) := (Xf)^2 + (Yf)^2 + \beta^2 (Zf)^2 \leq 1.$$

In the case $\beta > 0$, this distance corresponds to the Riemannian distance obtained by asserting that $(X, Y, \beta Z)$ is an orthonormal basis of the tangent space in each point. In the case $\beta = 0$, it is known, see for instance [JSC, Prop. 3.1], that it coincides with the Carnot – Carathéodory sub-Riemannian distance obtained by taking the length of the shortest horizontal curve. Recall that a curve is *horizontal* if its speed vector belongs almost everywhere to the *horizontal space* $\mathrm{Vect}\{X, Y\}$, and that the length of a horizontal curve is computed asserting that (X, Y) is an orthonormal basis of this horizontal space in each point.

The Heisenberg group $\mathbb{H}$ is topologically homeomorphic to $\mathbb{R}^3$ and the Lebesgue measure on $\mathbb{R}^3$ is a Haar measure of $\mathbb{H}$ (translation invariant) but in the case $\beta = 0$ the Hausdorff dimension of the $\mathbb{H}$ for the Carnot – Carathéodory metric is 4.

Moreover, in the sub-elliptic case $\beta = 0$, the Carnot-Carathéodory distance admits the following continuous family of dilation operators:

$$\mathrm{dil}_\lambda(x,y,z) = (\lambda x, \lambda y, \lambda^2 z); \ \lambda > 0.$$

A well known fact is that the Carnot-Carathéodory distance is equivalent to all homogeneous norm, see for instance [BLU, Prop. 5.1.4]. In particular there exist constants $c_2 > c_1 > 0$ such that

$$c_1(r^2 + |z|) \leq d(e,g)^2 \leq c_2(r^2 + |z|); \tag{1.2}$$

for all $g = (x,y,z) \in \mathbb{H}$ and $r^2 := x^2 + y^2$.

Random walks on $\mathbb{H}$. Let $\beta \geq 0$ be a real parameter. Let $(x_n, y_n, z_n)_{n\geq 0}$ be independent and identically distributed random variables on $\mathbb{R}^3$ (not necessarily Gaussian) with zero mean and covariance matrix $\mathrm{diag}(1,1,\beta^2)$. Now set $S_0 := 0$ and for all $n \geq 1$,

$$S_n := (X_n, Y_n, Z_n) := \Big(\frac{x_1}{\sqrt{n}}, \frac{y_1}{\sqrt{n}}, \frac{z_1}{\sqrt{n}}\Big) \cdots \Big(\frac{x_n}{\sqrt{n}}, \frac{y_n}{\sqrt{n}}, \frac{z_n}{\sqrt{n}}\Big). \tag{1.3}$$

The sequence $(S_n)_{n\geq 0}$ is a random walk on $\mathbb{H}$ started from the origin and with i.i.d. "non commutative multiplicative increments" given by a triangular array. In exponential coordinates,

$$X_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n x_i, \quad Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n y_i, \quad Z_n = A_n + \frac{1}{\sqrt{n}} \sum_{i=1}^n z_i$$

where

$$A_n := \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n} x_i \epsilon_{ij} y_j \quad \text{and} \quad \epsilon_{i,j} := \mathbf{1}_{j>i} - \mathbf{1}_{j<i}.$$

The random variable A_n is the algebraic area between the path $(X_k, Y_k)_{0\le k\le n}$ of a random walk in $\mathbb{R}^2$ and its chord $[(0,0),(X_n,Y_n)]$. With

$$\mathrm{dil}_t(x,y,z) = (tx, ty, t^2 z)$$

being the dilation operator on $\mathbb{H}$, we have

$$(X_n, Y_n, A_n) = \mathrm{dil}_{\frac{1}{\sqrt{n}}}((x_1,y_1,0)\cdots(x_n,y_n,0)).$$

According to a Functional Central Limit Theorem (or Invariance Principle) on Lie groups due to Daniel Stroock and Srinivasa Varadhan [SV] (see also Donald Wehn [W], cited in [P]),

$$(S_{\lfloor nt\rfloor})_{t\ge 0} \quad \underset{n\to\infty}{\overset{\text{law}}{\longrightarrow}} \quad (\mathbf{X}_t, \mathbf{Y}_t, \mathbf{Z}_t)_{t\ge 0} = (\mathbf{X}_t, \mathbf{Y}_t, \mathbf{A}_t + \beta \mathbf{W}_t)_{t\ge 0} \tag{1.4}$$

where $(\mathbf{X}_t, \mathbf{Y}_t)_{t\ge 0}$ is a standard Brownian motion on $\mathbb{R}^2$ started from the origin, where $(\mathbf{W}_t)_{t\ge 0}$ is a standard Brownian motion on $\mathbb{R}$ started from the origin and independent of $(\mathbf{X}_t, \mathbf{Y}_t)_{t\ge 0}$, and where $(\mathbf{A}_t)_{t\ge 0}$ is the Lévy area of $(\mathbf{X}_t, \mathbf{Y}_t)_{t\ge 0}$, in other words the algebraic area between the Brownian path and its chord, seen as a stochastic integral:

$$\mathbf{A}_t := \frac{1}{2}\Big(\int_0^t \mathbf{X}_s \,\mathrm{d}\mathbf{Y}_s - \int_0^t \mathbf{Y}_s \,\mathrm{d}\mathbf{X}_s\Big).$$

The heat process on $\mathbb{H}$. The stochastic process $(\mathbf{H}_t)_{t\ge 0} = (h\cdot(\mathbf{X}_t, \mathbf{Y}_t, \mathbf{Z}_t))_{t\ge 0}$ started from $\mathbf{H}_0 = h$ is a Markov diffusion process on $\mathbb{R}^3$ admitting the Lebesgue measure as an invariant and reversible measure. The Markov semigroup $(P_t)_{t\ge 0}$ of this process is defined for all $t\ge 0$, $h\in\mathbb{H}$, and bounded measurable $f:\mathbb{H}\to\mathbb{R}$, by

$$P_t(f)(h) := \mathbf{E}(f(\mathbf{H}_t) \mid \mathbf{H}_0 = h).$$

For all $t>0$ and $h\in\mathbb{H}$, the law of $\mathbf{H}_t$ conditionally on $\mathbf{H}_0 = h$ admits a density and

$$P_t(f)(h) = \int_{\mathbb{H}} f(g) p_t(h,g)\,\mathrm{d}g.$$

Estimates on the heat kernel p_t are available, see [BGG, L2, HM]. For instance when $\beta = 0$, there exist constants $C_2 > C_1 > 0$ such that for all $g = (x,y,z)\in\mathbb{H}$ and $t>0$,

$$\frac{C_1}{\sqrt{t^4 + t^3 r d(e,g)}}\exp\Big(-\frac{d^2(e,g)}{4t}\Big) \le p_t(e,g) \le \frac{C_2}{\sqrt{t^4 + t^3 r d(e,g)}}\exp\Big(-\frac{d^2(e,g)}{4t}\Big) \tag{1.5}$$

where d is the Carnot – Carathéodory distance and where $r^2 := x^2 + y^2$.

Let us define the family of probability measures (which depends on the parameter β)

$$\gamma_t := \mathrm{Law}(\mathbf{H}_t \mid \mathbf{H}_0 = 0) = P_t(\cdot)(0).$$

The infinitesimal generator is the linear second order operator

$$L = \frac{1}{2}(X^2 + Y^2 + \beta^2 Z^2)$$

where X, Y, Z are as in (1.1). The Schwartz space $\mathrm{Schwartz}(\mathbb{H},\mathbb{R})$ of rapidly decaying $\mathcal{C}^\infty$ functions from $\mathbb{H}\equiv\mathbb{R}^3$ to $\mathbb{R}$ is contained in the domain of L and is stable by L and by P_t for all $t\ge 0$. By the Dirac commutation relations $[X,Y] = Z = \partial_z$ and $[X,Z] = [Y,Z] = 0$, the operator L is hypoelliptic, and by the Hörmander theorem P_t admits a $\mathcal{C}^\infty$ kernel. The operator L is elliptic if $\beta > 0$ and not elliptic if $\beta = 0$ (singular diffusion matrix).

The operator L acts as the two dimensional Laplacian on functions depending only on x, y and not on z. The one parameter family of operators obtained from L when β runs through the interval $[0,1]$ interpolates between the sub-elliptic or sub-Riemannian Laplacian $\frac{1}{2}(X^2+Y^2)$ (for $\beta = 0$) and the elliptic or Riemannian Laplacian $\frac{1}{2}(X^2+Y^2+Z^2)$ (for $\beta = 1$). The sub-Riemannian and Riemannian Brownian motions $(\mathbf{H}_t)_{t\ge 0}$ have independent and stationary (non commutative) increments and are Lévy processes associated to non commutative) convolution

semigroups $(P_t)_{t\geq 0}$ on $\mathbb{H}$. When $\beta = 0$ the probability measures γ_t behaves very well with respect to dilation, can be seen as a Gaussian measure on $\mathbb{H}$, and a formula (oscillatory integral) for the kernel of P_t was computed by Paul Lévy using Fourier analysis. See the books [M, B1, N] and references therein for more information and details on this subject.

Logarithmic Sobolev inequalities. The entropy of $f : \mathbb{H} \to [0, \infty)$ with respect to a probability measure μ is defined by

$$\mathbf{Ent}_\mu(f) := \mathbf{E}_\mu(\Phi(f)) - \Phi(\mathbf{E}_\mu(f)) \quad \text{with} \quad \mathbf{E}_\mu(f) := \int f \, \mathrm{d}\mu$$

where $\Phi(u) = u \log(u)$. A logarithmic Sobolev inequality is of the form

$$\mathbf{Ent}_\mu(f^2) \leq \int T(f) \, \mathrm{d}\mu$$

where T is a "good" functional quadratic form. The most classical version involves $T = \Gamma$ and contains many geometrical informations. The book [BGL] contains a general introduction to Sobolev type functional inequalities for diffusion processes. However (see the discussion below), the classical "carré du champ" does not capture the whole geometry of $\mathbb{H}$. Define a weighted "carré du champ" $T_a = \Gamma + a\Gamma^Z$, where a is a function and $\Gamma^Z f = (Zf)^2 = (\partial_z f)^2$. Such a gradient will naturally arise in the logarithmic Sobolev inequality we derive from the non commutativity.

Main results. We start with the left-invariant diffusion operator $L = \frac{1}{2}(X^2 + Y^2 + \beta^2 Z^2)$ for $\beta \geq 0$ on the Heisenberg group. In the case $\beta > 0$, the operator is elliptic and it is not hard to see that a usual logarithmic Sobolev inequality holds for its heat kernel. Usual means here that the energy in the right hand side is given by the "carré du champ" operator Γ associated to L. Indeed, for $\beta > 0$, L can then be thought of as the Laplace – Beltrami operator of a Riemannian manifold whose Ricci curvature is actually constant and the Bakry – Émery theory applies. The case $\beta = 0$, is much more involved and have attracted a lot of attention. Indeed, the operator L is not anymore elliptic but is still sub-elliptic. The Ricci curvature tends to $-\infty$ when β goes to 0 and the Bakry – Émery theory fails. In this situation, the "carré du champ" operator contains only the horizontal part of the gradient. The question whether a logarithmic Sobolev inequality holds was answered positively by Hong-Quan Li in [L1] (see (2.2)), see also [3BC, HZ] and [DM].

In a different direction, even if the classical Bakry – Émery theory fails, Fabrice Baudoin and Nicola Garofalo developed in [BG] a generalization of the curvature criterion which is well adapted to the sub-Riemannian setting. One can then obtain some (weaker) logarithmic Sobolev inequalities with an elliptic gradient in the energy (see (2.4)).

Our approach is different. We follows the method developed by Leonard Gross in [G1] for the Gaussian and in [G2] for the path space on *elliptic* Lie groups. It is based on the tensorization property of the logarithmic Sobolev inequality and on the Central Limit Theorem for a random walk. It applies indifferently both for the sub-elliptic ($\beta = 0$) or the elliptic ($\beta > 0$) Laplacian on the Heisenberg group. At least when $\beta > 0$, our main result Theorem 1.1 below is in a way an explicit version of the abstract Theorem 4.1 in [G2].

The interest in our result is double: we compute explicitly for the first time the gradient which appears in the right hand side of Theorem 4.1 in [G2] in the case of the Heisenberg group for all $\beta \geq 0$, and we show, by looking at the case $\beta = 0$, that the method of Gross gives a non degenerate result for a sub-Riemannian model. This is surprising and unexpected.

The next theorem, that is the main result of the paper and is proved in Section 3, states the logarithmic Sobolev inequality for $\gamma = \gamma_1$. From the scaling property of the heat kernel, we can easily deduce a logarithmic Sobolev inequality for γ_t for every $t > 0$.

Theorem 1.1 (Logarithmic Sobolev inequality)**.** *For all $\beta \geq 0$ and $f \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$,*

$$\mathbf{Ent}_\gamma(f^2) \leq 2 \int_0^1 \mathbf{E}(g(\mathbf{H}_1, \mathbf{H}_t)) \, \mathrm{d}t \tag{1.6}$$

where for $h = (x, y, z)$ and $h' = (x', y', z')$,

$$\begin{aligned} g(h,h') &:= \Big(\partial_x f(h) - \frac{y-2y'}{2}\partial_z f(h)\Big)^2 + \Big(\partial_y f(h) + \frac{x-2x'}{2}\partial_z f(h)\Big)^2 + \beta^2\,(\partial_z f(h))^2 \\ &= ((X+y'Z)f(h))^2 + ((Y-x'Z)f(h))^2 + \beta^2(Zf(h))^2. \end{aligned}$$

The shape of the right hand side of (1.6) comes from the fact that the increments are not commutative: the sum in S_n produces along (1.4) the integral from 0 to 1.

The following corollary is obtained via Brownian Bridge and heat kernel estimates.

Corollary 1.2 (Weighted logarithmic Sobolev inequality)**.** *If $\beta = 0$ then there exist a constant $C > 0$ such that for all $f \in$ Schwartz$(\mathbb{H}, \mathbb{R})$,*

$$\mathbf{Ent}_\gamma(f^2) \le 2\,\mathbf{E}_\gamma\Big((\partial_x f)^2 + (\partial_y f)^2 + C(1 + x^2 + y^2 + |z|)(\partial_z f)^2\Big). \tag{1.7}$$

Corollary 1.2 is proved in Section 4.

Structure of the paper. Section 2 provides a discussion and a comparison with other inequalities such as the inequality of H.-Q. Li and the "elliptic" inequality of Baudoin and Garofalo. Section 3 is devoted to the proof of Theorem 1.1 which is based on the method of Gross using a random walk and the CLT. Section 4 provides the proof of Corollary 1.2 by using an expansion of (1.6), a probabilistic (Bayes formula), analytic (bounds for the heat kernel on $\mathbb{H}$), and geometric (bounds for the Carnot – Carathéodory distance) arguments for the control of the density of the Brownian bridge. For completeness, a short proof of the "elliptic" inequality of Baudoin and Garofalo in the case of the Heisenberg group (inequalities (2.4)-(2.5)) is provided in Section 5. Finally, in Section 6 we give the extension of our main result (Theorem 1.1) to the case of homogeneous Carnot groups of rank two (Theorem 6.1).

2. Discussion and comparison with other inequalities

Novelty. Taking $\beta = 0$ in (1.6) provides a new sub-elliptic logarithmic Sobolev inequality for the sub-Riemannian Gaussian law γ, namely, for all $f \in$ Schwartz$(\mathbb{H}, \mathbb{R})$,

$$\mathbf{Ent}_\gamma(f^2) \le 2\int_0^1 \mathbf{E}(g(\mathbf{H}_1, \mathbf{H}_t))\,\mathrm{d}t \tag{2.1}$$

where

$$g(h,h') := \Big(\partial_x f(h) - \frac{y-2y'}{2}\partial_z f(h)\Big)^2 + \Big(\partial_y f(h) + \frac{x-2x'}{2}\partial_z f(h)\Big)^2.$$

Horizontal optimality. The logarithmic Sobolev inequality (2.1), implies the optimal logarithmic Sobolev inequality for the standard Gaussian distribution $\mathcal{N}(0, I_2)$ on $\mathbb{R}^2$ with the Euclidean gradient, namely, for all $f \in$ Schwartz$(\mathbb{R}^2, \mathbb{R})$,

$$\mathbf{Ent}_{\mathcal{N}(0,I_2)}(f^2) \le 2\,\mathbf{E}_{\mathcal{N}(0,I_2)}((\partial_x f)^2 + (\partial_y f)^2).$$

To see it, it suffices to express (2.1) with a function f that does not depend on the third coordinate z. This shows in particular the optimality (minimality) of the constant 2 in front of the right hand side in the inequality of Theorem 1.1 and in (2.1).

Poincaré inequality. Recall that the variance of $f : \mathbb{H} \to \mathbb{R}$ with respect to μ is

$$\mathbf{Var}_\mu(f) := \int \Phi(f)\,\mathrm{d}\mu - \Phi\Big(\int f\,\mathrm{d}\mu\Big) \quad \text{where this time} \quad \Phi(u) = u^2.$$

As usual, the logarithmic Sobolev inequality (2.1) gives a Poincaré inequality by linearization. More precisely, replacing f by $1 + \varepsilon f$ in (2.1) gives, as $\varepsilon \to 0$,

$$\mathbf{Var}_\gamma(f) \le \int_0^1 \mathbf{E}(g(\mathbf{H}_1, \mathbf{H}_t))\,\mathrm{d}t.$$

Comparison with H.-Q. Li inequality. For $\beta = 0$, Hong-Quan Li has obtained in [L1] (see also [3BC, DM] for a Poincaré inequality) the following logarithmic Sobolev inequality: there exists a constant $C_{\mathrm{LSI}} > 0$ such that for all $f \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$,

$$\mathbf{Ent}_\gamma(f^2) \leq C_{\mathrm{LSI}}\, \mathbf{E}_\gamma((Xf)^2 + (Yf)^2). \tag{2.2}$$

The right hand side in (2.2) involves the "carré du champ" of the sub-Laplacian L, namely the functional quadratic form: $\Gamma(f, f) := \frac{1}{2}(L(f^2) - 2fLf) = (Xf)^2 + (Yf)^2$. Following the by now standard Bakry – Émery approach, the expansion of the scaled version shows that necessarily $C_{\mathrm{LSI}} > 2$ but the optimal (minimal) constant is unknown.

One can deduce from (2.2) a weighted inequality. Namely, since the random variables $-\mathbf{H}_t$ and $\mathbf{H}_t$ have the same law conditionally to $\{\mathbf{H}_0 = 0\}$, one can cancel out by symmetry, in average, the cross terms involving $x\partial_x f \partial_z f$ and $y\partial_x f \partial_z f$ when expanding the right hand side of the sum in (2.2) and its rotated version. The symmetrized version of (2.2) that we obtained in this way appears as a weighted logarithmic Sobolev inequality: for all $f \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$,

$$\mathbf{Ent}_\gamma(f^2) \leq C_{\mathrm{LSI}}\, \mathbf{E}_\gamma\Big((\partial_x f)^2 + (\partial_y f)^2 + \frac{x^2 + y^2}{4}(\partial_z f)^2\Big). \tag{2.3}$$

Comparison with the "elliptic" inequality of Baudoin-Garofalo. Baudoin and Garofalo have developed in [BG] a generalization of the Bakry – Émery semigroup/curvature approach well adapted to the sub-Riemannian setting, see also [B2, Prop. 4.11], [BB], and [Bo, Prop. 5.3.7 p. 129]. Their framework is well-suited for studying weighted functional inequalities such as (1.7) and (2.3). More precisely, it allows first to derive the following result: if $\beta = 0$ then for all real number $\nu > 0$ and all function $f \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$,

$$\mathbf{Ent}_\gamma(f^2) \leq 2\nu(\mathrm{e}^{\frac{1}{\nu}} - 1)\, \mathbf{E}_\gamma\Big((Xf)^2 + (Yf)^2 + \nu(Zf)^2\Big). \tag{2.4}$$

The symmetrized version of (2.4) is given by the following new weighted logarithmic Sobolev inequality: if $\beta = 0$ then for all real number $\nu > 0$ and all function $f \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$,

$$\mathbf{Ent}_\gamma(f^2) \leq 2\nu(e^{\frac{1}{\nu}} - 1)\, \mathbf{E}_\gamma\Big((\partial_x f)^2 + (\partial_y f)^2 + (\nu + \frac{x^2 + y^2}{4})(\partial_z f)^2\Big). \tag{2.5}$$

Our weighted inequality (1.7) is close to the weighted inequalities (2.3) and (2.5).

For the reader convenience, we provide a short proof of (2.4) and (2.5) in Section 5. This proof is the Heisenberg group specialization of the proof given in [Bo, Prop. 5.3.7 p. 129] (PhD thesis of the first author) see also [B2, Prop. 4.11] (PhD advisor of the first author).

Extensions and open questions. The Heisenberg group is the simplest non-trivial example of a Carnot group in other words stratified nilpotent Lie group. Those groups have a strong geometric meaning both in standard and stochastic analysis, see for instance [B1] for the latter point. Theorem 1.1 is extended to homogeneous Carnot group of rank two in Section 6. Note that the criterion of Baudoin and Garofalo [BG] holds for Carnot groups of rank two and an inequality similar to (2.4) holds in this context, see [B2].

The bounds on the distance and the heat kernel used to derive the weighted inequality (1.7) are not available for general Carnot groups and it should require more work to obtain an equivalent of Corollary 1.2. As a comparison, note that a version of (2.2) exists on groups with a so called H-structure, see [El], but a general version on Carnot groups is unknown due to the lack of general estimates for the heat kernel. An extension of Theorem 1.1 in the case of higher dimensional Carnot groups or in the case of curved sub-Riemannian space as CR spheres or anti-de Sitter spaces is opened.

Moreover, in the context and spirit of the work of Leonard Gross [G2] in the elliptic case, an approach at the level of paths space should be available. It is also natural to ask about a direct analytic proof or semigroup proof of the inequality of Theorem 1.1, without using the Central Limit Theorem.

3. Proof of Theorem 1.1

Fix a real $\beta \geq 0$. Consider $(x_n, y_n, z_n)_{n\geq 1}$ a sequence of independent and identically distributed random variables with Gaussian law of mean zero and covariance matrix $\mathrm{diag}(1, 1, \beta^2)$. Let S_n be as in (1.3). The Central Limit Theorem gives

$$S_n \xrightarrow[n\to\infty]{\text{law}} \gamma.$$

The law ν_n of S_n satisfies $\nu_n = (\mu_n)^{*n}$ where the convolution takes place in $\mathbb{H}$ and where μ_n is the Gaussian law on $\mathbb{R}^3$ with covariance matrix $\mathrm{diag}(1/n, 1/n, \beta^2/n)$.

For all $i = 1, \ldots, n$, let us define

$$S_{n,i} := (X_n, Y_n, Z_n, X_{n,i}, Y_{n,i})$$

where

$$X_{n,i} := -\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \epsilon_{ij} x_j \quad \text{and} \quad Y_{n,i} := -\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \epsilon_{ij} y_j.$$

The optimal logarithmic Sobolev inequality for the standard Gaussian measure $\mathcal{N}(0, I_{3n})$ on $\mathbb{R}^{3n}$ gives, for all $g \in \mathrm{Schwartz}(\mathbb{R}^{3n}, \mathbb{R})$,

$$\mathbf{Ent}_{\mathcal{N}(0,I_{3n})}(g^2) \leq 2\, \mathbf{E}_{\mathcal{N}(0,I_{3n})}\Big(\sum_{i=1}^{n} (\partial_{x_i} g)^2 + (\partial_{y_i} g)^2 + (\partial_{z_i} g)^2 \Big).$$

Let $s_n : \mathbb{R}^{3n} \to \mathbb{H}$ be the map such that $S_n = s_n((x_1, y_1, z_1), \ldots, (x_n, y_n, z_n))$. For some $f \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$ the function $g = f(s_n)$ satisfies

$$\begin{aligned}
\partial_{x_i} g(s_n) &= \frac{1}{\sqrt{n}} \Big(\partial_x f - \frac{Y_{n,i}}{2} \partial_z f \Big)(s_n), \\
\partial_{y_i} g(s_n) &= \frac{1}{\sqrt{n}} \Big(\partial_y f + \frac{X_{n,i}}{2} \partial_z f \Big)(s_n), \\
\partial_{z_i} g(s_n) &= \frac{\beta}{\sqrt{n}} (\partial_z f)(s_n).
\end{aligned}$$

It follows that for all $f \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$, denoting $\nu_n := \mathrm{Law}(S_n)$,

$$\mathbf{Ent}_{\nu_n}(f^2) \leq \frac{2}{n} \sum_{i=1}^{n} \mathbf{E}(h(S_{n,i})) = 2\, \mathbf{E}\Big(\frac{1}{n} \sum_{i=1}^{n} h(S_{n,i}) \Big) \tag{3.1}$$

where $h : \mathbb{R}^5 \to \mathbb{R}$ is defined from f by

$$h(x, y, z, x', y') := ((\partial_x - \frac{y'}{2} \partial_z) f(x, y, z))^2 + ((\partial_y + \frac{x'}{2} \partial_z) f(x, y, z))^2 + \beta^2 (\partial_z f(x, y, z))^2$$

The right hand side of (3.1) as $n \to \infty$ is handled by explaining the law of $S_{n,i}$ through a triangular array of increments of the process we anticipated in the limit. More precisely, let $((\mathbf{X}_t, \mathbf{Y}_t))_{t\geq 0}$ be a standard Brownian motion on $\mathbb{R}^2$ started from the origin, let $(\mathbf{A}_t)_{t\geq 0}$ be its Lévy area, and let $(\mathbf{W}_t)_{t\geq 0}$ be a Brownian motion on $\mathbb{R}$ starting from the origin, independent of $(\mathbf{X}_t, \mathbf{Y}_t)_{t\geq 0}$. Let us define, for $n \geq 1$ and $1 \leq i \leq n$,

$$\xi_{n,i} := \sqrt{n} \left(\mathbf{X}_{\frac{i}{n}} - \mathbf{X}_{\frac{i-1}{n}} \right), \quad \eta_{n,i} := \sqrt{n} \left(\mathbf{Y}_{\frac{i}{n}} - \mathbf{Y}_{\frac{i-1}{n}} \right), \quad \zeta_{n,i} := \sqrt{n} \left(\mathbf{W}_{\frac{i}{n}} - \mathbf{W}_{\frac{i-1}{n}} \right).$$

For all fixed $n \geq 1$, the random variables $(\xi_{n,i})_{1\leq i\leq n}$, $(\eta_{n,i})_{1\leq i\leq n}$, and $(\zeta_{n,i})_{1\leq i\leq n}$ are independent and identically distributed with Gaussian law $\mathcal{N}(0,1)$. Let us define now

$$X_n := \frac{1}{\sqrt{n}}\sum_{i=1}^n \xi_{n,i}, \qquad Y_n := \frac{1}{\sqrt{n}}\sum_{i=1}^n \eta_{n,i},$$

$$A_n := \frac{1}{2n}\sum_{i=1}^n\sum_{j=1}^n \xi_{n,i}\epsilon_{i,j}\eta_{n,i}, \qquad Z_n := \beta\frac{1}{\sqrt{n}}\sum_{i=1}^n \zeta_{n,i} + A_n,$$

$$X_{n,i} := -\frac{1}{\sqrt{n}}\sum_{j=1}^n \epsilon_{i,j}\xi_{n,j}, \qquad Y_{n,i} := -\frac{1}{\sqrt{n}}\sum_{j=1}^n \epsilon_{i,j}\eta_{n,i}.$$

We have then the equality in distribution

$$(X_n, Y_n, Z_n - A_n, X_{n,i}, Y_{n,i}) \overset{d}{=} (\mathbf{X}_1, \mathbf{Y}_1, \beta\mathbf{W}_1, \mathbf{X}_1 - (\mathbf{X}_{\frac{i-1}{n}} + \mathbf{X}_{\frac{i}{n}}), \mathbf{Y}_1 - (\mathbf{Y}_{\frac{i-1}{n}} + \mathbf{Y}_{\frac{i}{n}})).$$

Moreover, as $i/n \to s \in [0,1]$, we have the convergence in distribution

$$S_{n,i} := (X_n, Y_n, Z_n, X_{n,i}, Y_{n,i}) \xrightarrow[n\to\infty]{d} (\mathbf{X}_1, \mathbf{Y}_1, \mathbf{A}_1 + \beta\mathbf{W}_1, \mathbf{X}_1 - 2\mathbf{X}_s, \mathbf{Y}_1 - 2\mathbf{Y}_s).$$

It follows that for all continuous and bounded $h : \mathbb{R}^5 \to \mathbb{R}$,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n \mathbf{E}(h(S_{n,i})) &= \int_0^1 \mathbf{E}(h(X_n, Y_n, Z_n, X_{n,\frac{\lfloor tn\rfloor}{n}}, Y_{n,\frac{\lfloor tn\rfloor}{n}})\,\mathrm{d}t \\ &\underset{n\to\infty}{\longrightarrow} \int_0^1 \mathbf{E}(h(\mathbf{X}_1, \mathbf{Y}_1, \mathbf{A}_1 + \beta\mathbf{W}_1, \mathbf{X}_1 - 2\mathbf{X}_t, \mathbf{Y}_1 - 2\mathbf{Y}_t))\,\mathrm{d}t.\end{aligned}$$

4. Proof of Corollary 1.2

Let us consider (1.6) with $\beta = 0$. By expanding the right-hand side, and using the fact that the conditional law of $\mathbf{H}_1$ given $\{\mathbf{H}_0 = 0\}$ is invariant by central symmetry, we get a symmetrized weighted logarithmic Sobolev inequality: for all $f \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$,

$$\begin{aligned}\mathbf{Ent}_\gamma(f^2) &\leq 2\,\mathbf{E}\big((\partial_x f)^2(\mathbf{H}_1) + (\partial_y f)^2(\mathbf{H}_1)\big) \\ &\quad + \frac{1}{2}\int_0^1 \mathbf{E}\big(\,\mathbf{E}((\mathbf{X}_1 - 2\mathbf{X}_t)^2 + (\mathbf{Y}_1 - 2\mathbf{Y}_t)^2 \mid \mathbf{H}_1)(\partial_z f)^2(\mathbf{H}_1)\big)\,\mathrm{d}t \\ &\leq 2\,\mathbf{E}\big((\partial_x f)^2(\mathbf{H}_1) + (\partial_y f)^2(\mathbf{H}_1)\big) \\ &\quad + \mathbf{E}((\mathbf{X}_1^2 + \mathbf{Y}_1^2)(\partial_z f)^2(\mathbf{H}_1)) + 4\,\mathbf{E}\Big((\partial_z f)^2(\mathbf{H}_1)\int_0^1 \mathbf{E}(\mathbf{X}_t^2 + \mathbf{Y}_t^2 \mid \mathbf{H}_1)\,\mathrm{d}t\Big).\end{aligned}$$

The desired result is a direct consequence of the following lemma.

Lemma 4.1 (Bridge control)**.** *There exists a constant $C > 0$ such that for all $0 \leq t \leq 1$ and $h = (x, y, z) \in \mathbb{H}$,*

$$\mathbf{E}(\mathbf{X}_t^2 + \mathbf{Y}_t^2 \mid \mathbf{H}_1 = h) \leq C(t^2 d^2(e, h) + t)$$

and

$$\int_0^1 \mathbf{E}(\mathbf{X}_t^2 + \mathbf{Y}_t^2 \mid \mathbf{H}_1 = h)\,\mathrm{d}t \leq C(1 + x^2 + y^2 + |z|).$$

Note that for the classical euclidean Brownian motion

$$\mathbf{E}(\|\mathbf{B}_t\|^2 \mid \mathbf{B}_1) = t^2\|\mathbf{B}_1\|^2 + nt(1-t).$$

Proof of Lemma 4.1. For all random variables U, V, we denote by φ_U the density of U and by $\varphi_{U|V=v}$ the conditional density of U given $\{V = v\}$. For all $0 < t \leq 1$ and $k \in \mathbb{H}$,

$$\mathbf{E}(\mathbf{X}_t^2 + \mathbf{Y}_t^2 \mid \mathbf{H}_1 = k) = \int_{\mathbb{H}} r_g^2\,\varphi_{\mathbf{H}_t|\mathbf{H}_1=k}(g)\,\mathrm{d}g$$

where $r_g^2 = x_g^2 + y_g^2$ and $g = (x_g, y_g, z_g)$. Recall that $p_t(h,g) := \varphi_{\mathbf{H}_t = g | \mathbf{H}_0 = h}$ and that $e := (0,0,0)$ is the origin in $\mathbb{H}$. Thanks to the Bayes formula, for all $g, k \in \mathbb{H}$,

$$\varphi_{\mathbf{H}_t|\mathbf{H}_1=k}(g) = \frac{\varphi_{(\mathbf{H}_t,\mathbf{H}_1)}(g,k)}{\varphi_{\mathbf{H}_1}(k)} = \frac{\varphi_{\mathbf{H}_t}(g)\varphi_{\mathbf{H}_1|\mathbf{H}_t=g}(k)}{\varphi_{\mathbf{H}_1}(k)} = \frac{p_t(e,g)p_{1-t}(g,k)}{p_1(e,k)}.$$

Back to our objective, we have

$$\begin{aligned}\mathbf{E}(\mathbf{X}_t^2 + \mathbf{Y}_t^2 \mid \mathbf{H}_1 = k) &= \frac{\int p_t(e,g)\, r_g^2\, p_{1-t}(g,k)\,\mathrm{d}g}{p_1(e,k)} \\ &\leq \frac{\int p_t(e,g) d^2(e,g) p_{1-t}(g,k)\,\mathrm{d}g}{p_1(e,k)}.\end{aligned}$$

In what follows, the constant C may change from line to line. The idea is to kill the polynomial term d^2 in the numerator by using the exponential decay of the heat kernel, at the price of a slight time change. Namely, using (1.5), we get, for all $0 < \varepsilon < 1$,

$$\begin{aligned}\int p_t(e,g) r^2(e,g) p_{1-t}(g,k)\,\mathrm{d}g &\leq \int \frac{C}{\sqrt{t^4 + t^3 r d(e,g)}} d^2(e,g) \exp\left(-\frac{d^2(e,g)}{4t}\right) p_{1-t}(g,k)\,\mathrm{d}g \\ &\leq \int \frac{Ct}{\varepsilon\sqrt{t^4 + t^3 r d(e,g)}} \exp\left(-(1-\varepsilon)\frac{d^2(e,g)}{4t}\right) p_{1-t}(g,k)\,\mathrm{d}g \\ &\leq \int \frac{Ct}{\varepsilon} \sqrt{\frac{\left(\frac{t}{1-\varepsilon}\right)^4 + \left(\frac{t}{1-\varepsilon}\right)^3 r d(e,g)}{t^4 + t^3 r d(e,g)}} p_{\frac{t}{1-\varepsilon}}(e,g) p_{1-t}(g,k)\,\mathrm{d}g \\ &\leq \frac{Ct}{\varepsilon(1-\varepsilon)^2} \int p_{\frac{t}{1-\varepsilon}}(e,g) p_{1-t}(g,k)\,\mathrm{d}g \\ &\leq \frac{Ct}{\varepsilon(1-\varepsilon)^2} p_{1+\frac{\varepsilon t}{1-\varepsilon}}(e,k),\end{aligned}$$

where we used $x\exp(-x) \leq \frac{1}{e\varepsilon}\exp(-(1-\varepsilon)x)$. Therefore, for all $0 < \varepsilon \leq 1/2$,

$$\begin{aligned}\mathbf{E}(\mathbf{X}_t^2 + \mathbf{Y}_t^2 \mid \mathbf{H}_1 = k) &\leq \frac{Ct}{\varepsilon(1-\varepsilon)^2} \frac{p_{1+\frac{\varepsilon t}{1-\varepsilon}}(e,k)}{p_1(e,k)} \\ &\leq \frac{Ct}{\varepsilon(1-\varepsilon)^2} \exp\left(\frac{\varepsilon t}{1-\varepsilon+\varepsilon t}\frac{d^2(e,k)}{4}\right) \\ &\leq \frac{Ct}{\varepsilon} \exp\left(\varepsilon t \frac{d^2(e,k)}{2}\right).\end{aligned}$$

Now if $td^2(e,k) \geq 1$, then we take $\varepsilon = 1/(2td^2(e,k))$ which gives

$$\mathbf{E}(\mathbf{X}_t^2 + \mathbf{Y}_t^2 \mid \mathbf{H}_1 = k) \leq Ct^2 d^2(e,k),$$

while if $td^2(e,k) < 1$, then we take $\varepsilon = 1/2$ which gives

$$\mathbf{E}(\mathbf{X}_t^2 + \mathbf{Y}_t^2 \mid \mathbf{H}_1 = k) \leq Ct.$$

This provides the first desired inequality. We get the second using (1.2). □

5. Proof of inequalities (2.4) and (2.5)

In this section, we provide for the reader convenience a short proof of the inequalities (2.4) and (2.5). This corresponds to the case $\beta = 0$, but the method remains in fact valid beyond the assumption $\beta = 0$. Recall that this proof is essentially the Heisenberg group specialization of the proof given in [Bo, Prop. 5.3.7 p. 129] (see also [B2, Prop. 4.11]).

Proof of (2.4). For simplicity, we change L by a factor 2 and set $L = X^2 + Y^2$ and $P_t = \mathrm{e}^{tL}$, and in particular $\gamma = P_{1/2}(\cdot)(0)$. For all $f, g \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$, let us define

$$\begin{aligned}
\Gamma^{\mathrm{hori}}(f,g) &:= \frac{1}{2}(L(fg) - fLg - gLf) = X(f)X(g) + Y(f)Y(g),\\
\Gamma^{\mathrm{vert}}(f,g) &:= Z(f)Z(g),\\
\Gamma^{\mathrm{elli}}(f,g) &:= \Gamma^{\mathrm{hori}}(f,g) + \nu\Gamma^{\mathrm{vert}}(f,g).
\end{aligned}$$

Let us also denote

$$\begin{aligned}
\Gamma_2^{\mathrm{hori}}(f,f) &:= \frac{1}{2}(L\Gamma^{\mathrm{hori}}(f,f) - 2\Gamma^{\mathrm{hori}}(f,Lf)),\\
\Gamma_2^{\mathrm{vert}}(f,f) &:= \frac{1}{2}(L\Gamma^{\mathrm{vert}}(f,f) - 2\Gamma^{\mathrm{vert}}(f,Lf)),\\
\Gamma_2^{\mathrm{mix}}(f,f) &:= \frac{1}{2}(L\Gamma^{\mathrm{elli}}(f,f) - 2\Gamma^{\mathrm{elli}}(f,Lf)).
\end{aligned}$$

In the sequel, we also denote $\Gamma(f) = \Gamma(f,f)$ and $\Gamma_2(f) = \Gamma_2(f,f)$.

Curvature inequality. The following inequality holds: for all $f \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$,

$$\Gamma_2^{\mathrm{mix}}(f,f) \geq -\frac{1}{\nu}\Gamma^{\mathrm{elli}}(f,f). \tag{5.1}$$

Indeed, an easy computation gives

$$\begin{aligned}
\Gamma_2^{\mathrm{hori}}(f,f) &= (X^2f)^2 + (Y^2f)^2 + (XYf)^2 + (YXf)^2 - 2(Xf)(YZf) + 2(Yf)(XZf)\\
\Gamma_2^{\mathrm{vert}}(f,f) &= (XZf)^2 + (YZf)^2.
\end{aligned}$$

Since $\Gamma_2^{\mathrm{mix}} = \Gamma_2^{\mathrm{hori}} + \nu\Gamma_2^{\mathrm{vert}}$ and $Zf = XYf - YXf$, the Cauchy – Schwarz's inequality gives

$$\Gamma_2^{\mathrm{mix}}(f,f) \geq \frac{1}{2}(Lf)^2 + \frac{1}{2}(XYf + YXf)^2 + \frac{1}{2}(Zf)^2 - \frac{1}{\nu}X(f)^2 - \frac{1}{\nu}Y(f)^2,$$

which implies the desired curvature inequality (5.1).

Semigroup inequality. Let $f \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$ with $f \geq 0$. For all $0 \leq s \leq t$, set

$$U(s) := (P_{t-s}f)\,\Gamma^{\mathrm{hori}}\log(P_{t-s}f) \quad \text{and} \quad V(s) := (P_{t-s}f)\,\Gamma^{\mathrm{elli}}(\log(P_{t-s}f)).$$

Then for $0 \leq s \leq t$,

$$LU + \partial_s U = (P_{t-s}f)\,\Gamma^{\mathrm{hori}}(\log(P_{t-s}f)) \leq (P_{t-s}f)\,\Gamma^{\mathrm{elli}}(\log(P_{t-s}f)) = V(s) \tag{5.2}$$

$$LV + \partial_s V = 2(P_{t-s}f)\,\Gamma_2^{\mathrm{mix}}(\log(P_{t-s}f)) \geq -\frac{2}{\nu}V(s). \tag{5.3}$$

The first equality in (5.2) holds since Γ^{hori} is the "carré du champ" associated to L, the inequality holds because $\Gamma^{\mathrm{vert}}(f,f) \geq 0$ and the second equality is the definition of V.
In (5.3) we have used that $\Gamma^{\mathrm{elli}} = \Gamma^{\mathrm{hori}} + \nu\Gamma^{\mathrm{vert}}$, that the horizontal part of Γ^{elli} will produce Γ_2^{hori} (same kind of computation as in the first inequality of (5.2)) and that vertical part commute to the horizontal one $\Gamma^{\mathrm{hori}}(f, \Gamma^{\mathrm{vert}}(f,f)) = \Gamma^{\mathrm{vert}}(f, \Gamma^{\mathrm{hori}}(f,f))$, the inequality comes from (5.1).

Final step. Since by (5.3), $L(\mathrm{e}^{\frac{2s}{\nu}}V(s)) + \partial_s(\mathrm{e}^{\frac{2s}{\nu}}V(s)) \geq 0$, a parabolic comparison such as [BG, Prop. 4.5] or a simple semigroup interpolation implies that for $t \geq 0$,

$$\mathrm{e}^{\frac{2t}{\nu}}P_t(f\Gamma^{\mathrm{elli}}(\log f)) = \mathrm{e}^{\frac{2t}{\nu}}V(t) \geq V(0) = (P_tf)\Gamma^{\mathrm{elli}}(\log P_tf).$$

In particular,

$$V(s) \leq \mathrm{e}^{\frac{2(t-s)}{\nu}}P_{t-s}(f\Gamma^{\mathrm{elli}}(\log f)). \tag{5.4}$$

Now from (5.2) another application of the parabolic comparison theorem and the last estimate (5.4) give

$$P_t(U(t)) \leq U(0) + \int_0^t P_s(V(s))\,\mathrm{d}s \leq U(0) + \int_0^t \mathrm{e}^{\frac{2(t-s)}{\nu}}\,\mathrm{d}s\; P_t\left(f\Gamma^{\mathrm{elli}}(\log f)\right);$$

that is:

$$P_t(f\log f)(x) - P_t(f)(x)\log P_t(f)(x) \leq \frac{\nu}{2}\left(\mathrm{e}^{\frac{2t}{\nu}} - 1\right)P_t\left(\frac{\Gamma^{\mathrm{elli}}(f,f)}{f}\right)(x).$$

The conclusion follows by taking $t = 1/2$ and $x = 0$ since $\gamma = P_{1/2}(\cdot)(0)$. □

Proof of (2.5). Let us consider the right (instead of left) invariant vector fields

$$\hat{X} := \partial_x + \frac{y}{2}\partial_z \quad \text{and} \quad \hat{Y} := \partial_y - \frac{x}{2}\partial_z$$

and $\hat{L} = \hat{X} + \hat{Y}$ and $\hat{P}_t = \mathrm{e}^{t\hat{L}}$ the corresponding generator and semi-group. The semi-group is bi-invariant in the sense that $P_t f(0) = \hat{P}_t f(0)$, see for instance [3BC]. Recall that $\gamma = P_{1/2}(\cdot)(0) = \hat{P}_{1/2}(\cdot)(0)$. The method of proof of (2.4) remains valid if one replaces X, Y, L, P_t by their right invariant counter parts and yields that for all $f \in \mathrm{Schwartz}(\mathbb{H}, \mathbb{R})$,

$$\mathbf{Ent}_\gamma(f^2) \leq 2\nu(e^{\frac{1}{\nu}} - 1)\, \mathbf{E}_\gamma\Big((\hat{X}f)^2 + (\hat{Y}f)^2 + \nu(-Zf)^2\Big). \tag{5.5}$$

The conclusion follows by the summation of the inequalities (2.4) and (5.5). □

6. Extension to homogeneous Carnot groups of rank two

In this final section we consider the class of homogeneous Carnot groups of step two. We refer to [BLU] for more details and results on this class of Carnot groups. An homogeneous Carnot groups of step two is $\mathbb{R}^N = \mathbb{R}^d \times \mathbb{R}^m$ equipped with the group law given by

$$(x, z) \cdot (x', z') = (x + x', z + z' + \frac{1}{2}\langle Bx, x'\rangle)$$

where $x, x' \in \mathbb{R}^d$, $z, z' \in \mathbb{R}^m$ and

$$\langle Bx, x'\rangle = \Big(\langle B^{(1)}x, x'\rangle, \cdots, \langle B^{(m)}x, x'\rangle\Big)$$

for some linearly independent skew-symmetric $d \times d$ matrices $B^{(l)}$, $1 \leq l \leq m$. This class of groups includes a lot of usual examples, for instance all the Heisenberg groups $\mathbb{H}_n$ and free rank two Carnot groups. The case of the Heisenberg group $\mathbb{H}_1$ corresponds to

$$d = 2,\ m = 1,\ B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Actually, it is known that each stratified group of rank two is isomorphic to such an homogeneous Carnot group, see for instance [BLU, Theorem 3.2.2]. These homogeneous Carnot groups admit a dilation given by

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

The natural sub-Riemannian Brownian motion is given by $(\mathbf{X}_t, \mathbf{Z}_t)_{t\geq 0}$ where $\mathbf{X}$ is a standard Brownian motion on $\mathbb{R}^d$ and where $\mathbf{Z}$ corresponds to its generalized Levy area:

$$\mathbf{Z}_t^{(l)} = \sum_{1\leq p<q\leq d} b_{p,q}^{(l)} \mathbf{A}_t^{(p,q)}$$

with

$$\mathbf{A}_t^{(p,q)} = \int_0^t \mathbf{X}_s^{(p)} d\mathbf{X}_s^{(q)} - \int_0^t \mathbf{X}_s^{(q)} d\mathbf{X}_s^{(p)}.$$

We denote by γ the law of $(\mathbf{X}_1, \mathbf{Z}_1)$. The proof given in the case of the Heisenberg group $\mathbb{H}_1$ easily extends to this setting and leads to the following result.

Theorem 6.1 (Logarithmic Sobolev inequality)**.** *For all* $f \in \mathrm{Schwartz}(\mathbb{R}^N, \mathbb{R})$,

$$\mathbf{Ent}_\gamma(f^2) \leq 2\sum_{p=1}^d \int_0^1 \mathbb{E}\left[\left(\partial_p f(\mathbf{X}_1, \mathbf{Z}_1) + \sum_{l=1}^m \left(\sum_{q=1}^d b_{p,q}^{(l)}(\mathbf{X}_1^{(q)} - 2\mathbf{X}_s^{(q)})\right) \partial_{d+l} f(\mathbf{X}_1, \mathbf{Z}_1)\right)^2\right].$$

Acknowledgements

D.C. would like to thank Leonard Gross for his encouragement to explore this problem and Fabrice Baudoin for his hospitality during a visit to Purdue University in Fall 2008.

References

[B1] Fabrice Baudoin, *An introduction to the geometry of stochastic flows*, Imperial College Press, London, 2004. MR-2154760

[B2] ———, *Stochastic analysis on sub-Riemannian manifolds with transverse symmetries*, preprint Ann. Probab. 45 (2017), no. 1, 56–81. MR-3601645

[BB] ———, *Log-Sobolev inequalities for subelliptic operators satisfying a generalized curvature dimension inequality*, J. Funct. Anal. **262** (2012), no. 6, 2646–2676. MR-2885961

[3BC] Dominique Bakry, Fabrice Baudoin, Michel Bonnefont, and Djalil Chafaï, *On gradient bounds for the heat kernel on the Heisenberg group*, J. Funct. Anal. **255** (2008), no. 8, 1905–1938. MR-2462581

[BG] 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 (2017), no. 1, 151–219. MR-3584561

[BGG] Richard Beals, Bernard Gaveau, and Peter C. Greiner, *Hamilton-Jacobi theory and the heat kernel on Heisenberg groups*, J. Math. Pures Appl. (9) **79** (2000), no. 7, 633–689. MR-1776501

[BGL] Dominique Bakry, Ivan Gentil, and Michel Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, 2014. MR-3155209

[BLU] Andrea Bonfiglioli, Ermanno Lanconelli, Francesco Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, (2007), xxvi+800. MR-2363343

[Bo] Michel Bonnefont, *Inégalités fonctionnelles pour des noyaux de la chaleur sous-elliptiques*, doctoral thesis, Université Paul Sabatier Toulouse III, 2009.

[DM] Bruce K. Driver and Tai Melcher, *Hypoelliptic heat kernel inequalities on the Heisenberg group*, J. Funct. Anal. **221** (2005), no. 2, 340–365. MR-2124868

[El] Nathaniel Eldredge, *Gradient estimates for the subelliptic heat kernel on H-type groups*, J. Funct. Anal. **258**, (2010), no. 12,504–533. MR-2557945

[G1] Leonard Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083. MR-0420249

[G2] ———, *Logarithmic Sobolev inequalities on Lie groups*, Illinois J. Math. **36** (1992), no. 3, 447–490. MR-1161977

[HZ] Waldemar Hebisch, Boguslaw Zegarliński, *Coercive inequalities on metric measure spaces*, J. Funct. Anal., **258** (2010) no. 3, 814–851. MR-2558178,

[HM] Hermann Hueber, Detlef Horst Müller, *Asymptotics for some Green kernels on the Heisenberg group and the Martin boundary*, Math. Ann., **283** (1989) no. 1, 97–119. MR-973806

[JSC] David Jerison, Antonio Sánchez-Calle, *Subelliptic second order differential operators*, Lecture. Notes in Math., 1277 (1987), pp. 46-77. MR-922334

[L1] Hong-Quan Li, *Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg*, J. Funct. Anal. **236** (2006), no. 2, 369–394. MR-2240167

[L2] ———, *Estimations optimales du noyau de la chaleur sur les groupes de type Heisenberg*, J. Reine Angew. Math. **646** (2010), 195–233. MR-2719560

[M] Richard Montgomery, *A tour of subriemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, vol. 91, American Mathematical Society, Providence, RI, 2002. MR-1867362

[N] Daniel Neuenschwander, *Probabilities on the Heisenberg group*, Lecture Notes in Mathematics, vol. 1630, Springer-Verlag, Berlin, 1996, Limit theorems and Brownian motion. MR-1439509

[P] Gyula Pap, *Central limit theorems on stratified Lie groups.* Probability theory and mathematical statistics (Vilnius, 1993), 613–627, TEV, Vilnius, 1994. MR-1649603

[SV] Daniel W. Stroock and S. R. S. Varadhan, *Limit theorems for random walks on Lie groups*, Sankhyā Ser. A **35** (1973), no. 3, 277–294. MR-0517406

[W] Donald Wehn, *Probabilities on Lie groups*, Proc. Nat. Acad. Sci. U.S.A. **48** (1962), 791–795. MR-0153042

(MB) Institut de Mathématiques de Bordeaux, Université de Bordeaux, France.

(DC) CEREMADE, Université Paris-Dauphine, PSL, IUF, France.

(RH) Université du Luxembourg et Université Paris-Est Marne-la-Vallée.