NON-DIVERGENCE OPERATORS STRUCTURED ON HOMOGENEOUS HÖRMANDER VECTOR FIELDS: HEAT KERNELS AND GLOBAL GAUSSIAN BOUNDS

STEFANO BIAGI AND MARCO BRAMANTI

Abstract. Let $X_1, ..., X_m$ be a family of real smooth vector fields defined in $\mathbb{R}^n$, 1-homogeneous with respect to a nonisotropic family of dilations and satisfying Hörmander's rank condition at 0 (and therefore at every point of $\mathbb{R}^n$). The vector fields are *not* assumed to be translation invariant with respect to any Lie group structure. Let us consider the nonvariational evolution operator

$$\mathcal{H} := \sum_{i,j=1}^{m} a_{i,j}(t,x) X_i X_j - \partial_t$$

where $(a_{i,j}(t,x))_{i,j=1}^m$ is a symmetric uniformly positive $m \times m$ matrix and the entries a_{ij} are bounded Hölder continuous functions on $\mathbb{R}^{1+n}$, with respect to the "parabolic" distance induced by the vector fields. We prove the existence of a global heat kernel $\Gamma(\cdot; s, y) \in C^{2,\alpha}_{X,\mathrm{loc}}(\mathbb{R}^{1+n} \setminus \{(s,y)\})$ for $\mathcal{H}$, such that Γ satisfies two-sided Gaussian bounds and $\partial_t \Gamma, X_i\Gamma, X_iX_j\Gamma$ satisfy upper Gaussian bounds on every strip $[0,T] \times \mathbb{R}^n$. We also prove a scale-invariant parabolic Harnack inequality for $\mathcal{H}$, and a standard Harnack inequality for the corresponding stationary operator

$$\mathcal{L} := \sum_{i,j=1}^{m} a_{i,j}(x) X_i X_j.$$

with Hölder continuous coefficients.

1. Introduction

Let $X_1, ..., X_m$ be a family of real smooth vector fields defined in some domain $\Omega \subseteq \mathbb{R}^n$, satisfying Hörmander's rank condition in Ω. We consider the nonvariational evolution operator

$$\mathcal{H} := \sum_{i,j=1}^{m} a_{i,j}(t,x) X_i X_j - \partial_t \tag{1.1}$$

where $(a_{i,j}(t,x))_{i,j=1}^m$ is a symmetric uniformly positive $m \times m$ matrix and the entries a_{ij} are bounded Hölder continuous functions on $[0,T] \times \Omega$. (Precise definitions will be given later).

Motivated by issues arising in the theory of several complex variables, operators of this kind have been studied by several Authors. Bonfiglioli, Lanconelli, Uguzzoni, in a series of papers ([9], [10], [11], [12], [13]) have carried out the following research program. Given a set of 1-homogeneous, left-invariant Hörmander vector fields on

Date: April 9, 2021.
2010 *Mathematics Subject Classification.* 35K65, 35K08, 35K15.
Key words and phrases. Homogeneous Hörmander vector fields; heat kernels; global Gaussian estimates; scale-invariant Harnack inequality.

a Carnot group $\mathbb{G} = (\mathbb{R}^n, *)$, they have proved in [10] the existence of a global heat kernel Γ for $\mathcal{H}$, satisfying sharp Gaussian bounds, of the form

$$\begin{aligned} &\mathbf{c}^{-1}(t-s)^{-Q/2}\exp\Big(-M\frac{\|\xi^{-1}*x\|^2}{t-s}\Big) \leq \Gamma(t,x;s,y) \\ &\qquad \leq \mathbf{c}(t-s)^{-Q/2}\exp\Big(-\frac{\|\xi^{-1}*x\|^2}{M(t-s)}\Big), \end{aligned} \tag{1.2}$$

and analogous upper estimates for the first and second order derivatives of Γ along $X_1, \ldots, X_m$. In (1.2), Q and $\|\cdot\|$ are, respectively, the homogeneous dimension and a homogeneous norm on $\mathbb{G}$. Exploiting these results, they have derived in [13] a scale invariant parabolic Harnack inequality for $\mathcal{H}$, which easily implies an analogous standard Harnack inequality for the corresponding stationary operator

$$\mathcal{L} := \sum_{i,j=1}^m a_{i,j}(x)X_iX_j.$$

In order to build the heat kernel Γ for $\mathcal{H}$, the Authors exploit the *parametrix method*. This requires much preliminary work on the corresponding constant coefficient operator

$$\mathcal{H}_A := \sum_{i,j=1}^m a_{i,j}X_iX_j - \partial_t$$

where $A = \{a_{ij}\}$ belongs to the class $\mathcal{M}_\Lambda$ of constant symmetric matrices satisfying

$$\frac{1}{\Lambda}|\xi|^2 \leq \langle A\xi, \xi\rangle \leq \Lambda|\xi|^2 \quad \text{for every } \xi \in \mathbb{R}^m. \tag{1.3}$$

Namely, in [9] the Authors have proved sharp Gaussian estimates for the heat kernel Γ_A of $\mathcal{H}_A$, where the bounds depend on A only through the number Λ. In turn, the desired uniformity of the estimates relies on a careful analysis of a diffeomorphism turning the operator $\mathcal{H}_A$ into $\mathcal{H}_I$ (with I the identity matrix), carried out in [11] and also exploiting the results of [12].

Bramanti, Brandolini, Lanconelli, Uguzzoni, in [15], have studied heat-type operators $\mathcal{H}$ without assuming the existence of an underlying Carnot group. In other words, the vector fields $X_1, \ldots, X_m$ are now a general family of Hörmander vector fields. On the other hand, the operator $\mathcal{H}$ is assumed to coincide with the classical heat operator outside a large compact set. Under these assumptions, the Authors have implemented the same general research program described above: after establishing uniform Gaussian estimates for operators $\mathcal{H}_A$ corresponding to a contant matrix A, by the parametrix method a global Gaussian kernel is built for $\mathcal{H}$, and sharp Gaussian bounds are established, of the kind

$$\begin{aligned} &\mathbf{c}^{-1}\frac{1}{|B_X(x,\sqrt{t-s})|}\exp\Big(-M\frac{d_X(x,y)}{t-s}\Big) \leq \Gamma(t,x;s,y) \\ &\qquad \leq \mathbf{c}\frac{1}{|B_X(x,\sqrt{t-s})|}\exp\Big(-\frac{d_X(x,y)}{M(t-s)}\Big), \end{aligned}$$

with analogous upper estimates for the X-derivatives of Γ. Here, d_X is the control distance induced by $X_1, \ldots, X_m$ and $B_X(x,r)$ is the corresponding ball. As a consequence, scale invariant Harnack inequalities for $\mathcal{H}$ and $\mathcal{L}$ are derived. The results in [15] exploit, in particular, both some of the corresponding results proved on

Carnot groups in the aforementioned papers by Bonfiglioli, Lanconelli, Uguzzoni, and Schauder-type estimates for $\mathcal{H}$ proved by Bramanti, Brandolini in [14].

Since the vector fields considered in [15] are not assumed to be homogeneous nor left invariant with respect to an underlying group structure, under this respect that theory is more general than the one developed by Bonfiglioli, Lanconelli, Uguzzoni. On the other hand, the requirement that $\mathcal{H}$ coincides with the heat operator outside a compact set means that the results proved in [15] are actually *local results*, although they are better formulated with the language of a globally defined operator. This fact is consistent with a quite pervasive dichotomy in the theory of Hörmander operators: global results in the setting of Carnot groups *versus* local results in the general setting.

The aim of this paper is to establish the same set of results (i.e.: existence of a global heat kernel Γ for $\mathcal{H}$, sharp Gaussian bounds on Γ, scale invariant Harnack inequalities for $\mathcal{H}$ and $\mathcal{L}$), in a *global* version, for a family of Hörmander vector fields more general than the generators of a Carnot group. A convenient setting is that of smooth vector fields $X_1, \ldots, X_m$ in $\mathbb{R}^n$ satisfying the next assumptions:

(H.1): $X_1, \ldots, X_m$ are linearly independent (as vector fields) and homogeneous of degree 1 with respect to a family of non-isotropic dilations $\{\delta_\lambda\}_{\lambda>0}$ in $\mathbb{R}^n$ of the following form

$$\delta_\lambda(x) := (\lambda^{\sigma_1} x_1, \ldots, \lambda^{\sigma_n} x_n), \qquad \text{where } \sigma_1, \ldots, \sigma_n \in \mathbb{N} \text{ and } 1 = \sigma_1 \leq \ldots \leq \sigma_n \tag{1.4}$$

We define the δ_λ-homogeneous dimension of $(\mathbb{R}^n, \delta_\lambda)$ as

$$q := \textstyle\sum_{k=1}^n \sigma_k \geq n. \tag{1.5}$$

(H.2): $X_1, \ldots, X_m$ satisfy Hörmander's condition at $x = 0$, that is,

$$\dim\big\{Y(0) : Y \in \mathrm{Lie}(X_1, \ldots, X_m)\big\} = n,$$

where $\mathrm{Lie}(X_1, \ldots, X_m)$ is the *Lie algebra generated by* $X_1, \ldots, X_m$.

Some examples of vector fields of this kind are the following.

Examples. (1) In $\mathbb{R}^2$:

$$X_1 = \partial_{x_1} \text{ and } X_2 = x_1^k \, \partial_{x_2}$$

(with $k \in \mathbb{N}$), which are 1-homogeneous with respect to $\delta_\lambda(x) = (\lambda x_1, \lambda^{k+1} x_2)$.

(2) In $\mathbb{R}^n$:

$$X_1 = \partial_{x_1} \text{ and } X_2 = x_1 \partial_{x_2} + x_2 \partial_{x_3} + \ldots + x_{n-1} \partial_{x_n},$$

with $\delta_\lambda(x) = (\lambda x_1, \lambda^2 x_2, \cdots, \lambda^n x_n)$.

(3) In $\mathbb{R}^3$:

$$X_1 = \partial_{x_1} \text{ and } X_2 = x_1 \, \partial_{x_2} + x_1^2 \, \partial_{x_3},$$

with $\delta_\lambda(x) = (\lambda x_1, \lambda^2 x_2, \lambda^3 x_3)$.

(4) In $\mathbb{R}^n$:

$$X_1 = \partial_{x_1} \text{ and } X_2 = x_1 \, \partial_{x_2} + x_1^2 \, \partial_{x_3} + \cdots + x_1^{n-1} \, \partial_{x_n},$$

with the same dilations as in (2).

Under these assumptions, Hörmander operators of the kind

$$L=\sum_{i=1}^{m}X_i^2$$

or their evolutive counterpart

$$H=\partial_t-\sum_{i=1}^{m}X_i^2$$

have been studied in recent years in a series of papers. Biagi, Bonfiglioli in [2] and [4] have proved the existence of a homogeneous global fundamental solution for L and H, respectively. Their technique consists in constructing a higher dimensional Carnot group whose generators $\widehat{X}_1,...\widehat{X}_m$ project onto $X_1,...X_m$. The corresponding "lifted" operators $\widehat{L},\widehat{H}$, by known results possess a homogeneous fundamental solution; integrating these kernels with respect to the added variables, the Authors get, and are able to estimate, global homogeneous fundamental solutions for L and H. More explicit bounds on these fundamental solutions, in terms of the distance induced by the vector fields and the volume of the corresponding balls, have been proved in [6] and [5], respectively.

By combining this lifting technique with the bounds proved in [5], we first establish *sharp uniform* Gaussian bounds for the heat kernels corresponding to operators (1.1) with constant a_{ij} (see Theorem 3.3); next, we show that the parametrix method is applicable as in [15], getting the existence and sharp Gaussian bounds for the heat kernel of operators (1.1) with *Hölder continuous coefficients*. Before giving the precise statement of this result, which is one of the main results of this paper, it is convenient to fix the following:

Definition 1.1. Let $\Omega\subseteq\mathbb{R}^{1+n}=\mathbb{R}_t\times\mathbb{R}^n_x$ be an open set, and $\alpha\in(0,1)$. We define $C^\alpha_X(\Omega)$ as the space of functions $u:\Omega\to\mathbb{R}$ such that

$$\|u\|_{\alpha,\Omega}:=\sup_{\Omega}|u|+\sup_{\substack{(t,x),(s,y)\in\Omega\\(t,x)\neq(s,y)}}\frac{|u(t,x)-u(s,y)|}{d_X(x,y)^\alpha+|t-s|^{\alpha/2}}<\infty$$

where d_X is the Carnot-Carathéodory distance associated with

$$X:=\{X_1,\dots,X_m\}$$

(see Definition 2.4 in Section 2). Accordingly, we define $C^{2,\alpha}_X(\Omega)$ as the space of functions $u:\Omega\to\mathbb{R}$ such that

$$u,\ X_iu,\ X_iX_ju\text{ and }\partial_tu\in C^\alpha_X(\Omega),$$

where all the X-derivatives exist in the *intrinsic sense*. Finally, we define $C^{2,\alpha}_{X,\mathrm{loc}}(\Omega)$ as the space of functions $u:\Omega\to\mathbb{R}$ such that

$$u|_V\in C^{2,\alpha}_X(V)\text{ for every open set }V\Subset\Omega.$$

With the above definition at hand, we can now state the announced result providing *existence and properties* of the global heat kernel of $\mathcal{H}$.

Theorem 1.2 (Heat kernel for $\mathcal{H}$)**.** *Let $X_1,\dots,X_m$ be a family of linearly independent, smooth vector fields in $\mathbb{R}^n$, homogeneous of degree 1 with respect to a family of non-isotropic dilations $\{\delta_\lambda\}_{\lambda>0}$ of the form (1.4). Assume that $X_1,\dots,X_m$ satisfy*

Hörmander's rank condition at 0 (*and therefore at every point of* $\mathbb{R}^n$, *as will be explained later*). *Moreover, let*

$$A(t,x) = (a_{i,j}(t,x))_{i,j=1}^m$$

be a symmetric $m \times m$ *matrix of functions such that:*

(i) $a_{i,j} \in C_X^\alpha(\mathbb{R}^{1+n})$ *for every* $i,j = 1,\dots,m$;

(ii) *the following* uniform ellipticity condition holds*: there exists* $\Lambda > 1$ *s.t.*

$$\frac{1}{\Lambda}|\xi|^2 \le \langle A(t,x)\xi,\xi\rangle \le \Lambda|\xi|^2 \qquad \text{for every } \xi \in \mathbb{R}^m, (t,x) \in \mathbb{R}^{1+n}.$$

Let $\mathcal{H}$ *be as in* (1.1). *Then, there exists a function* (*"heat kernel" for* $\mathcal{H}$)

$$\Gamma : \mathbb{R}^{1+n} \times \mathbb{R}^{1+n} \to \mathbb{R}, \qquad \Gamma = \Gamma(t,x;s,y),$$

which satisfies the properties listed below.

(1) Γ *is continuous out of the diagonal of* $\mathbb{R}^{1+n} \times \mathbb{R}^{1+n}$.

(2) $\Gamma(t,x;s,y)$ *is non-negative, and it vanishes for* $t \le s$.

(3) *For every fixed* $(s,y) \in \mathbb{R}^{1+n}$ *we have*

$$\Gamma(\cdot;s,y) \in C^{2,\alpha}_{X,\mathrm{loc}}(\mathbb{R}^{1+n} \setminus \{(s,y)\}) \quad \text{and}$$
$$\mathcal{H}(\Gamma(\cdot;s,y)) = 0 \text{ on } \mathbb{R}^{1+n} \setminus \{(s,y)\};$$

(4) *For every* $T > 0$ *there exists a constant* $\mathbf{c} = \mathbf{c}_T > 0$ *such that*

$$\text{(i) } \mathbf{c}^{-1}\frac{1}{|B_X(x,\sqrt{t-s})|} \exp\left(-\mathbf{c}\frac{d_X(x,y)^2}{t-s}\right) \le \Gamma(t,x;s,y)$$
$$\le \mathbf{c}\frac{1}{|B_X(x,\sqrt{t-s})|} \exp\left(-\frac{d_X(x,y)^2}{\mathbf{c}(t-s)}\right);$$

$$\text{(ii) } |X_i(\Gamma(\cdot;s,y))(t,x)| \le \mathbf{c}\frac{1}{\sqrt{t-s}|B_X(x,\sqrt{t-s})|} \exp\left(-\frac{d_X(x,y)^2}{\mathbf{c}(t-s)}\right);$$

$$\text{(iii) } |X_iX_j(\Gamma(\cdot;s,y))(t,x)| + |\partial_t(\Gamma(\cdot;s,y))(t,x)|$$
$$\le \mathbf{c}\frac{1}{(t-s)\,|B_X(x,\sqrt{t-s})|} \exp\left(-\frac{d_X(x,y)^2}{\mathbf{c}(t-s)}\right);$$

for every $(t,x),(s,y) \in \mathbb{R}^{1+n}$ *with* $0 < t-s \le T$ *and* $1 \le i,j \le m$.

(5) *There exists a constant* $\delta > 0$ *such that the following assertion holds.*

Let $\mu \ge 0$ *and let* $T > 0$ *satisfy*

$$T\mu < \delta.$$

Moreover, let $f \in C_X^\alpha([0,T] \times \mathbb{R}^n)$ *and let* $g \in C(\mathbb{R}^n)$ *be such that*

$$|f(t,x)|,\ |g(x)| \le M \exp\left(\mu\, d_X(x,0)^2\right),$$

for all $(t,x) \in [0,T] \times \mathbb{R}^n$ *and for some* $M > 0$. *Then, the function*

$$u(t,x) := \int_{\mathbb{R}^n} \Gamma(t,x;0,y) g(y)\, dy$$
$$+ \int_{[0,t]\times\mathbb{R}^n} \Gamma(t,x;s,y) f(s,y)\, ds dy$$

is well-defined on $[0,T] \times \mathbb{R}^n$, *and enjoys the following properties:*

(i) $u \in C^{2,\alpha}_{X,\mathrm{loc}}((0,T)\times\mathbb{R}^n)\cap C([0,T]\times\mathbb{R}^n)$;

(ii) *u solves the Cauchy problem*

$$\begin{cases} \mathcal{H}u = f & \text{in } (0,T)\times\mathbb{R}^n, \\ u(0,\cdot) = g & \text{in } \mathbb{R}^n. \end{cases}$$

(6) *The following reproduction formula holds*

$$\Gamma(t,x;s,y) = \int_{\mathbb{R}^n} \Gamma(t,x;\tau,\xi)\Gamma(\tau,\xi;s,y)\,d\xi,$$

for every $x,y\in\mathbb{R}^n$ and $t>\tau>s$.

(7) *Suppose, in addition, that the functions $a_{i,j}$ are smooth on $\mathbb{R}^{1+n}$. Then, the operator $\mathcal{H}$ is C^∞-hypoelliptic in $\mathbb{R}^{1+n}$, and*

$$\mathcal{H}(\Gamma(\cdot;s,y)) = -\delta_{(s,y)} \quad \text{in } \mathcal{D}'(\mathbb{R}^{1+n}),$$

where $\delta_{(s,y)}$ denotes the Dirac delta centered at (s,y).

As anticipated, the second main result of this paper is a *scale-invariant Harnack inequality* for $\mathcal{H}$, and it will be stated in Section 5, see Theorem 5.2. The proof of this result (and that of its stationary counterpart, see Theorem 5.8) can follow an easier path, logically independent of the properties of the heat kernel: the Harnack inequality can be simply derived from the corresponding result which is known in Carnot groups, just by projection, owing to the lifting procedure sketched above. We note that this projection technique would not, instead, allow to get a simple proof of the existence of a global fundamental solution for (1.1).

2. Assumptions, notation and preliminary results

Here we explain and discuss in detail the notions and assumptions involved in the statement of Theorem 1.2. To begin with, we point out some easy consequences of assumptions (H.1)-(H.2) which will be useful in the sequel (for a proof see [2]).

(1) Hörmander's condition holds at every point $x\in\mathbb{R}^n$, i.e.,

$$\dim\big\{Y(x) : Y\in \mathrm{Lie}(X_1,\dots,X_m)\big\} = n \qquad \text{for every } x\in\mathbb{R}^n.$$

(2) The Lie algebra $\mathrm{Lie}(X_1,\dots,X_m)$ is *nilpotent* and *stratified*, that is

$$\mathrm{Lie}(X_1,\dots,X_m) = \bigoplus_{i=1}^{\sigma_n} V_i, \tag{2.1}$$

where $V_1 = \mathrm{span}\{X_1,\dots,X_m\}$ and $V_i := [V_1,V_{i-1}]$ (for $i\geq 2$). Furthermore, for every $i=1,\dots,\sigma_n$ one also has

$$V_i = \big\{Y\in\mathrm{Lie}(X_1,\dots,X_m) : Y \text{ is } \delta_\lambda\text{-homogeneous of degree } i\big\}. \tag{2.2}$$

As a consequence, since it is finitely-generated, $\mathrm{Lie}(X_1,\dots,X_m)$ has *finite dimension*, say N. Moreover, using assumption (H.2), one gets

$$N = \mathrm{Lie}(X_1,\dots,X_m) \geq n. \tag{2.3}$$

From now on, we will adopt the simplified notation

$$\mathfrak{a} := \mathrm{Lie}(X_1,\dots,X_m),$$

so that $N := \dim(\mathfrak{a})$. On account of (2.3), only two cases can occur:

(a) $N = n$. In this case, by taking into account the δ_λ-homogeneity of the X_i's, we are entitled to invoke the results in [1, 7]: there exists an operation $\circ$ on $\mathbb{R}^n$ such that $\mathbb{F} := (\mathbb{R}^n, \circ, \delta_\lambda)$ is a Carnot group, and

$$\mathrm{Lie}(\mathbb{F}) = \mathfrak{a}.$$

As a consequence, $X_1, \ldots, X_m$ are homogeneous and left-invariant on $\mathbb{F}$, and thus the results in this paper are well-known (see [9, Thm. 2.5]).

(b) $N > n$. In this case, again by exploiting the results contained in [1], we see that *there cannot exist* any Lie-group structure in $\mathbb{R}^n$ with respect to which $X_1, \ldots, X_m$ are left-invariant. In particular, [9, Thm. 2.5] *does not apply* in this case, and analogous results are not known.

In view of the preceding discussion, it is not restrictive to assume the following 'dimensional' hypothesis (in addition to (H.1) and (H.2)).

(H.3): We suppose that

$$N = \dim(\mathfrak{a}) > n,$$

and we define

$$p := N - n \geq 1.$$

Remark 2.1. All the results of this paper will be stated assuming only (H.1) and (H.2); however, their proofs will be given assuming also (H.3). The reason is that, if (H.1) and (H.2) hold but (H.3) is not satisfied, that is, $N = n$, all these results are already known from [10] and [13], as explained in the Introduction.

We will denote points $z \in \mathbb{R}^N$ by

$$z = (x, \xi), \qquad \text{with } x \in \mathbb{R}^n \text{ and } \xi \in \mathbb{R}^p. \tag{2.4}$$

Under assumption (H.3) it is proved in [2] that the X_i's can be lifted (in a suitable sense) to *left-invariant* vector fields 'living' on a higher-dimensional Carnot group:

Theorem 2.2 (see [2, Thm. 3.2])**.** *Assume that $X = \{X_1, \ldots, X_m\} \subseteq \mathcal{X}(\mathbb{R}^n)$ satisfies* (H.1), (H.2), (H.3).

Then, there exists a homogeneous Carnot group $\mathbb{G} = (\mathbb{R}^N, *, D_\lambda)$, nilpotent of step $r = \sigma_n$ and with m generators, such that

$$\mathrm{Lie}(\mathbb{G}) \text{ is isomorphic to } \mathfrak{a}.$$

Moreover, using the notation in (2.4)*, the dilation D_λ takes the 'lifted form'*

$$D_\lambda(x, \xi) = (\delta_\lambda(x), \lambda^{s_1}\xi_1, \ldots, \lambda^{s_p}\xi_p), \qquad \begin{array}{c}\text{where } s_1, \ldots, s_p \in \mathbb{N} \text{ and} \\ s_1 \leq \ldots \leq s_p < \sigma_n.\end{array} \tag{2.5}$$

As a consequence, the homogeneous dimension Q of $\mathbb{G}$ is given by

$$Q := \textstyle\sum_{i=1}^n \sigma_i + \sum_{i=1}^p s_i > q. \tag{2.6}$$

Finally, there exists a system $\widehat{X} = \{\widehat{X}_1, \ldots, \widehat{X}_m\}$ of Lie-generators of $\mathrm{Lie}(\mathbb{G})$ *such that $\widehat{X}_i$ is a lifting of X_i for every $i = 1, \ldots, m$; this means that*

$$\widehat{X}_i(x, \xi) = X_i(x) + R_i(x, \xi), \tag{2.7}$$

where $R_i(x, \xi)$ is a smooth vector field operating only in the variable $\xi \in \mathbb{R}^p$, with coefficients possibly depending on (x, ξ). In particular, the $\widehat{X}_i$'s are D_λ-homogeneous of degree 1 *(for every $i = 1, \ldots, m$).*

Remark 2.3. For a future reference, here we briefly review how the group $\mathbb{G}$ in Theorem 2.2 is constructed. For all the details, we refer to [2].

First of all, since we have already recognized that $\mathfrak{a}$ is nilpotent and stratified, it is well-known that $(\mathfrak{a}, \diamond, \Delta_\lambda)$ is a stratified group, where

- $\diamond$ is the Baker-Campbell-Hausdorff series on $\mathfrak{a}$ (boiling down to a finite sum, since $\mathfrak{a}$ is nilpotent);
- Δ_λ is the unique linear map on $\mathfrak{a}$ such that $\Delta_\lambda|_{V_i} := \lambda^i \,\mathrm{id}$.

Moreover, since $\mathfrak{a}$ has finite dimension N, we can fix a basis

$$\mathcal{E} = \{E_1, \dots, E_N\}$$

of $\mathfrak{a}$ (as a vector space), which is *adapted to the stratification* $\{V_i\}_{i=1}^{\sigma_n}$ in (2.1). This means that, setting $r := \sigma_n$, $\mathcal{E}$ can be decomposed as

$$\mathcal{E} = \left\{E_1^{(1)}, \dots E_{N_1}^{(1)}, \dots, E_1^{(r)}, \dots, E_{N_r}^{(r)}\right\},$$

where, for every $i = 1, \dots, r$, we have

- $N_i := \dim(V_i)$ (so that $N_1 = m$ and $N_1 + \cdots + N_r = N$);
- $\mathcal{E}_i := \left\{E_1^{(i)}, \dots, E_{N_i}^{(i)}\right\}$ is a basis of V_i.

Using the chosen basis $\mathcal{E}$, we then equip $\mathbb{R}^N$ with a structure of homogeneous Carnot group $\mathbb{A} = (\mathbb{R}^N, \circ, d_\lambda)$ by 'reading' $\diamond$ and Δ_λ in $\mathcal{E}$-coordinates, i.e.,

$$\sum_{i=1}^N (a \circ b)_i \, E_i = \left(\sum_{i=1}^N a_i \, E_i\right) \diamond \left(\sum_{i=1}^N b_i \, E_i\right) \qquad (\text{for all } a, b \in \mathbb{R}^N)$$

$$\sum_{i=1}^N \left(d_\lambda(a)\right)_i \, E_i = \Delta_\lambda \left(\sum_{i=1}^N a_i \, E_i\right) \qquad (\text{for all } a \in \mathbb{R}^N \text{ and } \lambda > 0).$$

For any fixed $i = 1, \dots, m$, we now let J_i be the unique left-invariant vector field on $\mathbb{A}$ coinciding with ∂_{a_i} at $a = 0$. In [2] it is proved the existence of a suitable diffeomorphism $T \in C^\infty(\mathbb{R}^N; \mathbb{R}^N)$, only depending on the basis $\mathcal{E}$, which turns the J_i's into new vector fields, say $Z_1, \dots, Z_m \in \mathcal{X}(\mathbb{R}^N)$, such that

$$Z_i(z) = Z_i(x, \xi) = E_i(x) + W_i(x, \xi) \qquad (i = 1, \dots, m).$$

Here, $W_1, \dots, W_m$ are smooth vector fields operating only in the variable $\xi \in \mathbb{R}^p$, with coefficients possibly depending on (x, ξ). On the other hand, since $\mathcal{E}_1$ is a basis of $V_1 = \mathrm{span}\{X_1, \dots, X_m\}$, for every $i = 1, \dots, m$ we can write

$$X_i = \textstyle\sum_{k=1}^m c_{k,i} E_k \qquad (\text{for a suitable constants } c_{k,i} \in \mathbb{R}).$$

Hence, the set $\widehat{X}$ is obtained by defining

$$\widehat{X}_i := \textstyle\sum_{k=1}^m c_{k,i} Z_k \qquad (\text{for all } i = 1, \dots, m).$$

Finally, the underlying Carnot group $\mathbb{G}$ appearing in the statement of Theorem 2.2 can be obtained as the unique Lie group isomorphic to $\mathbb{A}$ via T.

We close this section by recalling the notion of *control distance* associated with a Hörmander set of vector fields; moreover, we review and some properties of this distance when *homogeneous vector fields* are involved.

Definition 2.4. Let $\mathcal{W} = \{W_1, \dots, W_m\}$ be a family of smooth vector fields satisfying Hörmander's rank condition at every point of $\mathbb{R}^n$. Given any couple of points $x, y \in \mathbb{R}^n$, we define

$$d_{\mathcal{W}}(x,y) := \inf\{T > 0 : \text{ there exists } \gamma \in \mathcal{S}(T) \text{ with } \gamma(0) = \text{x and } \gamma(T) = y\},$$

where $\mathcal{S}(T)$ is the set of the $W^{1,1}$-curves $\gamma : [0,T] \to \mathbb{R}^n$ satisfying

$$\dot{\gamma}(t) = \sum_{j=1}^{m} \alpha_j(t)\, W_j(\gamma(t)), \quad \text{with} \sum_{j=1}^{m} |\alpha_j(t)| \leq 1.$$

The map $d_{\mathcal{W}}$ is called the $\mathcal{W}$-control distance or the Carnot-Carathéodory distance (CC distance) related to $\mathcal{W}$. Given any $x \in \mathbb{R}^n$ and any $r > 0$, we indicate by $B_{\mathcal{W}}(x,r)$ the $d_{\mathcal{W}}$-ball

$$B_{\mathcal{W}}(x,r) := \{y \in \mathbb{R}^n : d_{\mathcal{W}}(x,y) < r\}.$$

Remark 2.5. Since we have assumed that $W_1, \dots, W_m$ satisfy Hörmander's rank condition at every point of $\mathbb{R}^n$, it is well-known that $d_{\mathcal{W}}(x,y)$ is finite for every $x, y \in \mathbb{R}^n$; moreover, $d_{\mathcal{W}}$ is a distance, topologically but not metrically equivalent to the Euclidean one. In particular, "continuous" in Euclidean sense and "continuous" with respect to the control distance $d_{\mathcal{W}}$ are the same.

Remark 2.6. Let $\mathcal{W} = \{W_1, \dots, W_m\} \subseteq \mathcal{X}(\mathbb{R}^n)$ satisfy assumptions (H.1)-(H.2), and let $d_{\mathcal{W}}$ be the associated control distance. Due to the δ_λ-homogeneity of the W_i's, it easy to see that the following properties hold:

(1) $d_{\mathcal{W}}$ is jointly δ_λ-homogeneous of degree 1, that is,

$$d_{\mathcal{W}}(\delta_\lambda(x), \delta_\lambda(y)) = \lambda\, d_{\mathcal{W}}(x,y) \quad \text{for all } x, y \in \mathbb{R}^n \text{ and } \lambda > 0;$$

(2) for every $x \in \mathbb{R}^n$ and every $r > 0$, one has

$$\delta_\lambda\left(B_{\mathcal{W}}(x,r)\right) = B_{\mathcal{W}}(\delta_\lambda(x), \lambda r).$$

If, in addition $W_1, \dots, W_m$ are left-invariant with respect to some Lie-group structure $\mathbb{F} = (\mathbb{R}^n, \circ)$, the distance $d_{\mathcal{W}}$ is also *translation-invariant*, that is,

$$d_{\mathcal{W}}(x,y) = d_{\mathcal{W}}(\alpha * x, \alpha * y) \quad \text{for all } x, y, \alpha \in \mathbb{R}^n.$$

In particular, the above property implies that

$$d_{\mathcal{W}}(0,x) = d_{\mathcal{W}}(0, x^{-1}) \quad \text{for all } x \in \mathbb{R}^m.$$

Remark 2.7. Let $X = \{X_1, \dots, X_m\} \subseteq \mathcal{X}(\mathbb{R}^n)$ satisfy (H.1), (H.2), (H.3), and let

$$\mathbb{G} = (\mathbb{R}^N, *, D_\lambda), \qquad \widehat{X} := \{\widehat{X}_1, \dots, \widehat{X}_m\}$$

be as in Theorem 2.2. Moreover, let d_X and $d_{\widehat{X}}$ denote the CC-distances associated with X and $\widehat{X}$, respectively. Since, for every $j = 1, \dots, m$, we have

$$\widehat{X}_j = X_j + \textstyle\sum_{i=1}^{p} r_{i,j}(x,\xi)\, \partial_{\xi_i}$$

(for suitable smooth functions $r_{i,j}$), it is easy to recognize that

$$d_X(x,y) \leq d_{\widehat{X}}\left((x,\xi),(y,\eta)\right) \qquad \forall\, x, y \in \mathbb{R}^n,\ \xi, \eta \in \mathbb{R}^p.$$

Furthermore, given any $z = (x,\xi) \in \mathbb{R}^N$ and any $r > 0$, we have

$$\pi\left(B_{\widehat{X}}(z,r)\right) = B_X(x,r), \tag{2.8}$$

where $\pi : \mathbb{R}^N \equiv \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ denotes the projection of $\mathbb{R}^N$ onto $\mathbb{R}^n$. We explicitly notice that, since π is continuous, from (2.8) we immediately derive that

$$\pi\left(\overline{B_{\widehat{X}}(z,r)}\right) \subseteq \overline{B_X(x,r)}.$$

3. Uniform Gaussian bounds for operators with constant coefficients

Given $\Lambda \geq 1$, we denote by $\mathcal{M}_\Lambda$ the set of the $m \times m$ symmetric matrices A satisfying the following *uniform ellipticity condition*:

$$\frac{1}{\Lambda}|\xi|^2 \leq \langle A\xi, \xi\rangle \leq \Lambda|\xi|^2 \qquad \text{for every } \xi \in \mathbb{R}^m.$$

For every fixed $A = (a_{i,j})_{i,j=1}^m \in \mathcal{M}_\lambda$ we define

$$\mathcal{H}_A := \sum_{i,j=1}^m a_{i,j} X_i X_j - \partial_t \qquad \text{on } \mathbb{R}^{1+n} = \mathbb{R}_t \times \mathbb{R}^n_x, \tag{3.1}$$

Since A is symmetric and positive definite, it admits a *unique* (symmetric and) positive definite square root, say S. As a consequence, writing $S = (s_{i,j})_{i,j=1}^m$, we have

$$\mathcal{H}_A = \sum_{j=1}^m Y_j^2 - \partial_t, \qquad \text{where } Y_j := \sum_{i=1}^m s_{i,j} X_i.$$

On the other hand, since S is non-singular, the family $Y = \{Y_1, \ldots, Y_m\} \subseteq \mathcal{X}(\mathbb{R}^n)$ satisfies assumptions (H.1)-(H.2)-(H.3). In particular, since the $s_{i,j}$'s are constant, one has

$$\mathrm{Lie}(Y) = \mathrm{Lie}(Y_1, \ldots, Y_m) = \mathfrak{a}. \tag{3.2}$$

Gathering these facts, we can apply Theorem 2.2 to the family Y: there exist a homogeneous Carnot group $\mathbb{F}$ and a system

$$\widehat{Y} = \{\widehat{Y}_1, \ldots, \widehat{Y}_m\}$$

of Lie-generators for the Lie algebra $\mathrm{Lie}(\mathbb{F})$ such that, for every $i = 1, \ldots, m$, the vector field $\widehat{Y}_i$ is a lifting of Y_i in the sense of (2.7).

The key observation is that, in view of Remark 2.3, the construction of $\mathbb{F}$ does not really depend on Y, but only on the Lie algebra $\mathrm{Lie}(Y)$ and on the choice of an *adapted basis*. As a consequence, using (3.2) and choosing the *same adapted basis* used for the construction of $\mathbb{G}$ (notice that, by (2.2), the stratification $\{V_k\}_{k=1}^{\sigma_n}$ is independent of X or Y), we obtain

$$\mathbb{F} = \mathbb{G} \qquad \text{and} \qquad \widehat{Y}_j = \sum_{i=1}^m s_{i,j} \widehat{X}_i.$$

Summing up, the couple $(\mathbb{G}, \widehat{X})$ associated with X provides a 'lifting pair' for the family $Y = \{Y_1, \ldots, Y_m\}$ (in the sense of Theorem 2.2) *which is independent of the fixed matrix* $A \in \mathcal{M}_\Lambda$. By making use of this observation, we are entitled to use the results established in [4], which lead to the following theorem.

Theorem 3.1 (see [4, Thm.s 1.4 and 1.6]). *With the above assumptions and notation, the following facts hold.*

(1) *If* $\widehat{\Gamma}_A$ *is the (unique) smooth heat kernel of*

$$\widehat{\mathcal{H}}_A = \sum_{i,j=1}^{m} a_{i,j} \widehat{X}_i \widehat{X}_j - \partial_t = \sum_{i=1}^{m} \widehat{Y}_i^2 - \partial_t \qquad \text{on } \mathbb{R}_t \times \mathbb{R}^N_{(x,\xi)}$$

vanishing at infinity constructed in [17], *then the function*

$$\begin{gathered} \Gamma_A : \mathbb{R}^{1+n} \times \mathbb{R}^{1+n} \to \mathbb{R}, \\ \Gamma_A(t,x;s,y) = \int_{\mathbb{R}^p} \widehat{\Gamma}_A(t,(x,\xi);s,(y,0))\, d\xi, \end{gathered} \tag{3.3}$$

is a global heat kernel *for the operator* $\mathcal{H}_A$ *defined in* (3.1). *This means, precisely, that*

- *for any fixed* $(t,x) \in \mathbb{R}^{1+n}$, *we have* $\Gamma_A(t,x;\cdot) \in L^1_{\mathrm{loc}}(\mathbb{R}^{1+n})$;
- *for every* $\varphi \in C_0^\infty(\mathbb{R}^{1+n})$ *and every* $(t,x) \in \mathbb{R}^{1+n}$, *we have*

$$\begin{aligned} &\mathcal{H}_A\left(\int_{\mathbb{R}^{1+n}} \Gamma_A(t,x;s,y)\varphi(s,y)dsdy\right) \\ &= \int_{\mathbb{R}^{1+n}} \Gamma_A(t,x;s,y)\, \mathcal{H}_A\varphi(s,y)dsdy = -\varphi(t,x). \end{aligned}$$

Moreover, Γ_A *enjoys the properties listed below.*

(a) $\Gamma_A \geq 0$ *and* $\Gamma_A(t,x;s,y) = 0$ *if and only if* $t \leq s$;

(b) $\Gamma_A(t,x;s,y) = \Gamma_A(t,y;s,x)$ *and* $\Gamma_A(t,x;s,y) = \Gamma_A(t-s,x;0,y)$;

(c) Γ_A *is smooth out of the diagonal of* $\mathbb{R}^{1+n} \times \mathbb{R}^{1+n}$, *and*

$$\mathcal{H}_A(\Gamma_A(\cdot;s,y)) \equiv 0 \quad \text{on } \mathbb{R}^{1+n} \setminus \{(s,y)\};$$

(d) *for every fixed* $(t,x) \in \mathbb{R}^{1+n}$, *if* $t > s$ *we have*

$$\int_{\mathbb{R}^n} \Gamma_A(t,x;s,y)dy = \int_{\mathbb{R}^n} \Gamma_A(t,x;s,y)dx = 1;$$

(e) *for every* $x, y \in \mathbb{R}^n$ *and* $s < \tau < t$ *we have*

$$\Gamma_A(t,x;s,y) = \int_{\mathbb{R}^n} \Gamma_A(t,x;\tau,\zeta)\, \Gamma_A(\tau,\zeta;s,y)d\zeta. \tag{3.4}$$

(2) *Setting* $\widehat{\gamma}_A(t,z) := \widehat{\Gamma}_A(t,z;0)$, *we have*

$$\widehat{\Gamma}_A(t,(x,\xi);s,(y,\eta)) = \widehat{\gamma}_A\left(t-s,(y,\eta)^{-1} * (x,\xi)\right),$$

so that identity (3.3) *becomes*

$$\Gamma_A(t,x;s,y) = \int_{\mathbb{R}^p} \widehat{\gamma}_A\left(t-s,(y,0)^{-1} * (x,\xi)\right) d\xi. \tag{3.5}$$

(3) *For every integers* $h,k \geq 1$, $\alpha,\beta \geq 0$ *and every choice of* $i_1,\dots,i_h,j_1,\dots,j_k$ *in* $\{1,\dots,m\}$, *we have the following representation formulas, holding true*

for every $x, y \in \mathbb{R}^n$ and $t, s \in \mathbb{R}$ such that $s < t$:

$$(3.6)\qquad \begin{aligned}&(\partial_t)^\alpha (\partial_s)^\beta X^x_{i_1} \cdots X^x_{i_h} \Gamma_A(t,x;s,y)\\ &= (-1)^\beta \int_{\mathbb{R}^p} \Big((\partial_t)^{\alpha+\beta} \widehat{X}_{i_1} \cdots \widehat{X}_{i_h} \widehat{\gamma}_A \Big) \big(t-s, (y,0)^{-1} * (x,\xi)\big)\, d\xi;\end{aligned}$$

$$(3.7)\qquad \begin{aligned}&(\partial_t)^\alpha (\partial_s)^\beta X^y_{j_1} \cdots X^y_{j_k} \Gamma_A(t,x;s,y)\\ &= (-1)^\beta \int_{\mathbb{R}^p} \Big((\partial_t)^{\alpha+\beta} \widehat{X}_{j_1} \cdots \widehat{X}_{j_k} \widehat{\gamma}_A \Big) \big(t-s, (x,0)^{-1} * (y,\xi)\big)\, d\xi;\end{aligned}$$

$$(3.8)\qquad \begin{aligned}&(\partial_t)^\alpha (\partial_s)^\beta X^y_{j_1} \cdots X^y_{j_k} X^x_{i_1} \cdots X^x_{i_h} \Gamma_A(t,x;s,y)\\ &= (-1)^\beta \int_{\mathbb{R}^p} \Big((\partial_t)^{\alpha+\beta} \widehat{X}_{j_1} \cdots \widehat{X}_{j_k} \Big((\widehat{X}_{i_1} \cdots \widehat{X}_{i_h} \widehat{\gamma}_A) \circ \widetilde{\iota} \Big) \Big)\\ &\qquad\qquad \big(t-s, (x,0)^{-1} * (y,\xi)\big)\, d\xi\,,\end{aligned}$$

Here $\widetilde{\iota} : \mathbb{R}^{1+N} \to \mathbb{R}^{1+N}$ is the map defined by

$$(3.9)\qquad \widetilde{\iota}(t,(x,\xi)) = (t, \iota(x,\xi)) \qquad (\textit{with } t \in \mathbb{R},\ x \in \mathbb{R}^n,\ \xi \in \mathbb{R}^p),$$

and $\iota(x,\xi) = (x,\xi)^{-1}$ is the inverse of (x,ξ) in the Carnot group $\mathbb{G}$.

Remark 3.2. By combining the representation formula (3.5) with the symmetry of Γ_A in x and y, we obtain the following alternative identity

$$(3.10)\qquad \Gamma_A(t,x;s,y) = \Gamma_A(t,y;s,x) = \int_{\mathbb{R}^p} \widehat{\gamma}_A \big(t-s, (x,0)^{-1} * (y,\xi)\big)\, d\xi.$$

We shall repeatedly exploit (3.10) in place of (3.5).

By means of the representation formula (3.5) for Γ_A and of the analogous representation formulas (3.6)-(3.8) for its (t,X)-derivatives, we are able to prove the following theorem, which is the main result in this section. This will be the starting point to implement the parametrix method and build a fundamental solution for operators with variable coefficients.

Theorem 3.3 (Gaussian bounds for constant coefficient operators)**.** *Let the above assumptions and notation do apply. Moreover, let $\Lambda \geq 1$ be fixed. Then, the following facts hold.*

(1) *There exists a constant $\kappa_\Lambda > 0$ such that*

$$(3.11)\qquad \begin{aligned}&\frac{1}{\kappa_\Lambda} \frac{1}{\big|B_X(x,\sqrt{t-s})\big|} \exp\left(-\kappa_\Lambda \frac{d_X(x,y)^2}{t-s}\right) \leq \Gamma_A(t,x;s,y)\\ &\leq \kappa_\Lambda \frac{1}{\big|B_X(x,\sqrt{t-s})\big|} \exp\left(-\frac{d_X(x,y)^2}{\kappa_\Lambda (t-s)}\right),\end{aligned}$$

for every $x, y \in \mathbb{R}^n$, $s < t$ and $A \in \mathcal{M}_\Lambda$.

(2) *For every integers $r \geq 1, \alpha, \beta \geq 0$, there exists $\kappa = \kappa_{\Lambda,r,\alpha,\beta} > 0$ such that*

$$(3.12)\qquad \begin{aligned}&\big|(\partial_t)^\alpha (\partial_s)^\beta W_1 \cdots W_r \Gamma_A(t,x;s,y)\big|\\ &\leq \kappa \frac{(t-s)^{-(\alpha+\beta+r/2)}}{\big|B_X(x,\sqrt{t-s})\big|} \exp\left(-\frac{d_X(x,y)^2}{\kappa(t-s)}\right),\end{aligned}$$

for every $x, y \in \mathbb{R}^n$, $s < t$, $A \in \mathcal{M}_\Lambda$ *and every choice of* $W_1, \ldots, W_r$ *in*

$$\mathcal{D}_X := \{X_1^x, \ldots, X_m^x, X_1^y, \ldots, X_m^y\}.$$

(3) *For every integers* $r \geq 1, \alpha, \beta \geq 0$, *there exists* $\overline{\kappa} = \overline{\kappa}_{\Lambda,r,\alpha,\beta} > 0$ *such that*

$$\begin{aligned}&\left|(\partial_t)^\alpha (\partial_s)^\beta W_1 ... W_r \Gamma_A(t,x;s,y) - (\partial_t)^\alpha (\partial_s)^\beta W_1 ... W_r \Gamma_B(t,x;s,y)\right| \\ &\leq \overline{\kappa}\, \|A - B\|^{1/\sigma_n} \frac{(t-s)^{-(\alpha+\beta+r/2)}}{\left|B_X\left(x, \sqrt{t-s}\right)\right|} \exp\left(-\frac{d_X(x,y)^2}{\kappa(t-s)}\right),\end{aligned} \tag{3.13}$$

for every $x, y \in \mathbb{R}^n$, $s < t$, $A, B \in \mathcal{M}_\Lambda$ *and every* $W_1, \ldots, W_r \in \mathcal{D}_X$ *(here,* $\|\cdot\|$ *denotes the usual matrix norm).*

In order to establish Theorem 3.3 we need the following

Lemma 3.4. *Let* $Z \in \mathcal{X}(\mathbb{R}^N)$ *be* D_λ*-homogeneous of degree* 1. *Then,*

$$Z = \sum_{i=1}^{\sigma_n} \gamma_i(z) \mathcal{P}_i(\widehat{X}_1, \ldots, \widehat{X}_m), \tag{3.14}$$

where $\mathcal{P}_i(\theta_1, \ldots, \theta_m)$ *is a suitable homogeneous polynomial of (Euclidean) degree* i *in the non-commuting variables* $\theta_1, \ldots, \theta_m$, *and* γ_i *is a* D_λ*-homogeneous polynomial of degree* $i-1$ *(for all* $i = 1, \ldots, \sigma_n$*).*

Proof. Throughout this proof we do not need to separate the 'base' variable $x \in \mathbb{R}^n$ from the 'lifted' variable $\xi \in \mathbb{R}^p$; hence we use the compact notation $z = (z_1, \ldots, z_N)$ for the points of $\mathbb{R}^N$ and we write

$$D_\lambda(z) = (\lambda^{\upsilon_1} z_1, \ldots, \lambda^{\upsilon_N} z_N).$$

For every fixed $i \in \{1, \ldots, N\}$, we then denote by J_i the unique left-invariant vector field on $\mathbb{G}$ coinciding with ∂_{z_i} at $z = 0$. By well-known results on Carnot groups (see, e.g., [8, Sec.s 1.3 and 1.4]), J_i is D_λ-homogeneous of degree υ_i and

$$J_i = \partial_{z_i} + \sum_{\substack{k=1 \\ \upsilon_k > \upsilon_i}}^{N} \alpha_{k,i}(z) \frac{\partial}{\partial z_k}, \tag{3.15}$$

where $\alpha_{k,i}$ is a suitable D_λ-homogeneous polynomial of degree $\upsilon_k - \upsilon_i$. Starting form (3.15), it is easy to recognize that

$$\partial_{z_i} = J_i + \sum_{\substack{k=1 \\ \upsilon_k > \upsilon_i}}^{N} \beta_{k,i}(z) J_k,$$

where $\beta_{k,i}$ is again a D_λ-homogeneous polynomial of degree $\upsilon_k - \upsilon_i$. Thus, if Z is any smooth vector field D_λ-homogeneous of degree 1, we can write

$$Z = \sum_{k=1}^{N} a_k(z) \frac{\partial}{\partial z_k} = \sum_{k=1}^{N} \gamma_k(z) J_k, \tag{3.16}$$

where γ_k is a D_λ-homogeneous polynomial of degree $\upsilon_k - 1$. Now, since $J_1, \ldots, J_N$ are left-invariant on $\mathbb{G}$ and $\widehat{X}_1, \ldots, \widehat{X}_m$ are *Lie-generators* of Lie($\mathbb{G}$), any J_i can be written as a linear combination (with constant coefficients) of iterated commutators of length υ_i of the $\widehat{X}_i$'s; more precisely, we have

$$J_i = \mathcal{P}_i(\widehat{X}_1, \ldots, \widehat{X}_m) \qquad (i = 1, \ldots, N), \tag{3.17}$$

where $\mathcal{P}_i(\theta_1, \ldots, \theta_m)$ is a suitable homogeneous polynomial of (Euclidean) degree υ_i in the non-commuting variable $\theta_1, \ldots, \theta_m$.

Combining (3.16) and (3.17), we get

$$Z = \sum_{k=1}^{N} \gamma_k(z) \mathcal{P}_k(\widehat{X}_1, \ldots, \widehat{X}_m).$$

Finally, reminding that $\min_k \upsilon_k = 1$ and $\max_k \upsilon_k = \sigma_n$ (see, respectively, (1.4) and (2.5)), we can reorder the above sum with respect the D_λ-homogeneity of the $\mathcal{P}_k$'s, thus obtaining (3.14). □

We will also need the next

Proposition 3.5 (See [5, Prop. 3.10, Rem. 3.9])**.** *The following* global doubling property *of d_X holds:* there exist $\gamma_1, \gamma_2 > 0$ such that

$$\gamma_1 \left(\frac{r}{\rho}\right)^n \leq \frac{|B_X(x,r)|}{|B_X(x,\rho)|} \leq \gamma_2 \left(\frac{r}{\rho}\right)^q, \tag{3.18}$$

for every $x \in \mathbb{R}^n$ and every $0 < \rho < r$. This also implies, for every $\theta > 0$,

$$\frac{1}{|B_X(y,\sqrt{r})|} \exp\left(-\frac{d_X(x,y)^2}{\theta\, r}\right) \leq \frac{C_q}{|B_X(x,\sqrt{r})|} \exp\left(-\frac{d_X(x,y)^2}{C_q\, \theta\, r}\right), \tag{3.19}$$

where $C_q > 0$ is a constant only depending on the number q in (1.5)*.*

We can now prove Theorem 3.3.

Proof of Theorem 3.3. (1) First of all, since $\mathbb{G}$ is a Carnot group and $\widehat{X}_1, \ldots, \widehat{X}_m$ are Lie-generators of Lie($\mathbb{G}$), we can apply [9, Thm. 2.5]: there exists a constant $\mathbf{c}_\Lambda \geq 1$ such that, if $\widehat{\gamma}_A(\cdot) = \widehat{\Gamma}_A(\cdot; 0)$ is the global heat kernel of

$$\widehat{\mathcal{H}} = \textstyle\sum_{i,j=1}^m a_{i,j} \widehat{X}_i \widehat{X}_j - \partial_t$$

(see Theorem 3.1), then

$$\frac{1}{\mathbf{c}_\Lambda} t^{-Q/2} \exp\left(-\mathbf{c}_\Lambda \frac{\|z\|^2}{t}\right) \leq \widehat{\gamma}_A(t,z) \leq \mathbf{c}_\Lambda\, t^{-Q/2} \exp\left(-\frac{\|z\|^2}{\mathbf{c}_\Lambda t}\right), \tag{3.20}$$

for every $t > 0$, $z \in \mathbb{R}^N$ and $A \in \mathcal{M}_\Lambda$. Here, Q is the homogeneous dimension of $\mathbb{G}$ defined in (2.6) and

$$\|\cdot\| = d_{\widehat{X}}(\cdot; 0),$$

where $d_{\widehat{X}}$ is the CC distance associated with $\widehat{X} = \{\widehat{X}_1, \ldots, \widehat{X}_m\}$. By combining (3.20) with the representation formula (3.10), we then get

$$\begin{aligned} &\frac{1}{\mathbf{c}_\Lambda} (t-s)^{-Q/2} \int_{\mathbb{R}^p} \exp\left(-\mathbf{c}_\Lambda \frac{\|(x,0)^{-1} * (y,\xi)\|^2}{t-s}\right) d\xi \leq \Gamma_A(t,x;s,y) \\ &\qquad \leq \mathbf{c}_\Lambda\ (t-s)^{-Q/2} \int_{\mathbb{R}^p} \exp\left(-\frac{\|(x,0)^{-1} * (y,\xi)\|^2}{\mathbf{c}_\Lambda (t-s)}\right) d\xi, \end{aligned} \tag{3.21}$$

for every $x, y \in \mathbb{R}^n$, $t, s \in \mathbb{R}$ with $s < t$ and $A \in \mathcal{M}_\Lambda$. We now exploit the results in [5, Prop.s 4.2 and 4.4]: there exists a constant $c_0 > 0$ such that, for any $x, y \in \mathbb{R}^n$

and $t > 0$,

$$\begin{aligned}\frac{1}{c_0\,|B_X(x,\sqrt{t})|}\,\exp\left(-c_0\,\frac{d_X(x,y)^2}{t}\right) \\ \le t^{-Q/2}\int_{\mathbb{R}^p}\exp\left(-\frac{\|(x,0)^{-1}*(y,\xi)\|^2}{t}\right)d\xi \\ \le \frac{c_0}{|B_X(x,\sqrt{t})|}\,\exp\left(-\frac{d_X(x,y)^2}{c_0 t}\right).\end{aligned} \tag{3.22}$$

Putting together (3.21), (3.22) and the global doubling property of d_X in (3.18), we immediately obtain (3.11).

(2) We distinguish three different cases.

Case I: $W_1\cdots W_r = X^x_{i_1}\cdots X^x_{i_r}$. In this case, taking into account the representation formula (3.6), for every $x,y\in\mathbb{R}^n$ and $s<t$ we can write

$$\begin{aligned}\left|(\partial_t)^\alpha(\partial_s)^\beta\,W_1\cdots W_r\Gamma_A(t,x;s,y)\right| \\ \le \int_{\mathbb{R}^p}\left|(\partial_t)^{\alpha+\beta}\widehat{X}_{i_1}\cdots\widehat{X}_{i_r}\widehat{\gamma}_A\right|\big(t-s,(y,0)^{-1}*(x,\xi)\big)\,d\xi.\end{aligned} \tag{3.23}$$

Moreover, reminding that the $\widehat{X}_i$'s are Lie-generators for Lie($\mathbb{G}$), we can invoke [9, Thm. 2.5]: there exists a constant $\mathbf{c}=\mathbf{c}_{\Lambda,r,\alpha,\beta}>0$ such that, for every $z\in\mathbb{R}^N$, $t>0$ and $A\in\mathcal{M}_\Lambda$, one has

$$\left|(\partial_t)^{\alpha+\beta}\,\widehat{X}_{i_1}\cdots\widehat{X}_{i_r}\widehat{\gamma}_A(t,z)\right| \le \mathbf{c}\,t^{-(Q/2+\alpha+\beta+r/2)}\,\exp\left(-\frac{\|z\|^2}{\mathbf{c}\,t}\right). \tag{3.24}$$

By combining (3.23) with (3.24) we then get, for every $x,y\in\mathbb{R}^n$ and $s<t$,

$$\begin{aligned}&\left|(\partial_t)^\alpha(\partial_s)^\beta\,W_1\cdots W_r\Gamma_A(t,x;s,y)\right| \\ &\quad\le \frac{\mathbf{c}}{(t-s)^{\alpha+\beta+r/2}}\cdot(t-s)^{-Q/2}\int_{\mathbb{R}^p}\exp\left(-\frac{\|(y,0)^{-1}*(x,\xi)\|^2}{\mathbf{c}(t-s)}\right)d\xi.\end{aligned}$$

From this, by exploiting (3.22) and (3.18) we obtain

$$\begin{aligned}\left|(\partial_t)^\alpha(\partial_s)^\beta\,W_1\cdots W_r\Gamma_A(t,x;s,y)\right| \\ \le \kappa\,\frac{(t-s)^{-(\alpha+\beta+r/2)}}{\left|B_X(y,\sqrt{t-s})\right|}\,\exp\left(-\frac{d_X(x,y)^2}{\kappa(t-s)}\right),\end{aligned} \tag{3.25}$$

where $\kappa>0$ is a suitable constant only depending on Λ, r and α. The desired (3.12) now follows from (3.25) and (3.19).

Case II: $W_1\cdots W_r = X^y_{j_1}\cdots X^y_{j_r}$. We argue exactly as in Case I: by combining the integral representation formula (3.7) with (3.24), we have

$$\begin{aligned}&\left|(\partial_t)^\alpha(\partial_s)^\beta W_1\cdots W_r\Gamma_A(t,x;s,y)\right| \\ &\quad\le \int_{\mathbb{R}^p}\left|(\partial_t)^{\alpha+\beta}\,\widehat{X}_{j_1}\cdots\widehat{X}_{j_r}\widehat{\gamma}_A\right|\big(t-s,(x,0)^{-1}*(y,\xi)\big)\,d\xi \\ &\quad\le \frac{\mathbf{c}_{\Lambda,r,\alpha,\beta}}{(t-s)^{\alpha+\beta+r/2}}\cdot(t-s)^{-Q/2}\int_{\mathbb{R}^p}\exp\left(-\frac{\|(x,0)^{-1}*(y,\xi)\|^2}{\mathbf{c}(t-s)}\right)d\xi,\end{aligned}$$

for every $x,y\in\mathbb{R}^n$, $s<t$ and $A\in\mathcal{M}_\Lambda$. From this, by taking into account (3.22) and (3.18) we immediately obtain (3.12).

CASE III: $W_1 \cdots W_r = X^y_{j_1} \cdots X^y_{j_k} X^x_{i_1} \cdots X^x_{i_h}$ (with $h+k=r$). In this last case, starting from the representation formula (3.8) we can write

$$\begin{aligned}&\big|(\partial_t)^\alpha (\partial_s)^\beta W_1 \cdots W_r \Gamma_A(t,x;s,y)\big| \\ &\leq \int_{\mathbb{R}^p} \big|(\partial_t)^{\alpha+\beta} \widehat{X}_{j_1} \cdots \widehat{X}_{j_k} \big((\widehat{X}_{i_1} \cdots \widehat{X}_{i_h} \widehat{\gamma}_A) \circ \widetilde{\iota}\big)\big| \big(t-s, (x,0)^{-1} * (y,\xi)\big)\, d\xi.\end{aligned} \tag{3.26}$$

We should now apply some uniform estimates for the derivatives of $\widehat{\gamma}_A$ as in (3.24). However, due to the presence of the map $\widetilde{\iota}$ in (3.26), such estimates are not directly available in [9]. We then exploit Lemma 3.4 to overcome this problem.

To begin with, for a fixed $t > 0$ we consider the (smooth) function

$$u_t : \mathbb{R}^N \to \mathbb{R}, \qquad u_t(z) := \widehat{X}_{i_1} \cdots \widehat{X}_{i_h} \widehat{\gamma}_A(t,z),$$

and we repeatedly exploit [5, Lem. 5.3]: this gives (see also (3.9))

$$\begin{aligned}\widehat{X}_{j_1} \cdots \widehat{X}_{j_k} \big((\widehat{X}_{i_1} \cdots \widehat{X}_{i_h} \widehat{\gamma}_A) \circ \widehat{\iota}\big) &= \widehat{X}_{j_1} \cdots \widehat{X}_{j_k} (u_t \circ \iota) \\ &= (Z_1 \cdots Z_k u_t) \circ \iota = \big(Z_1 \cdots Z_k \widehat{X}_{i_1} \cdots \widehat{X}_{i_h} \widehat{\gamma}_A\big) \circ \widehat{\iota},\end{aligned} \tag{3.27}$$

where $Z_1, \ldots, Z_k$ are suitable smooth vector fields D_λ-homogeneous of degree 1 but not necessarily left-invariant. Since all the vector fields in (3.27) are D_λ-homogeneous of degree 1, we are entitled to apply Lemma 3.4, obtaining

$$Z_1 \cdots Z_k \widehat{X}_{i_1} \cdots \widehat{X}_{i_h} \widehat{\gamma}_A = \sum_{r \leqslant |\omega| \leqslant r\sigma_n} \gamma_\omega(z)\, \mathcal{Q}_\alpha(\widehat{X}_1, \ldots, \widehat{X}_m) \widehat{\gamma}_A. \tag{3.28}$$

Here, $r = h+k$, $\mathcal{Q}_\omega(\theta_1, \ldots, \theta_m)$ is a homogeneous polynomial of (Euclidean) degree $|\omega|$ in the non-commuting variables $\theta_1, \ldots, \theta_m$, and γ_ω is a D_λ- homogeneous polynomial of degree $|\omega| - r$. On account of (3.28), we can then write

$$\begin{aligned}&(\partial_t)^{\alpha+\beta} \widehat{X}_{j_1} \cdots \widehat{X}_{j_k} \big((\widehat{X}_{i_1} \cdots \widehat{X}_{i_h} \widehat{\gamma}_A) \circ \widetilde{\iota}\big) \\ &= \sum_{r \leqslant |\omega| \leqslant r\sigma_n} (\partial_t)^{\alpha+\beta} \Big[\big(\gamma_\omega(\cdot) \mathcal{Q}_\alpha(\widehat{X}_1, \ldots, \widehat{X}_m) \widehat{\gamma}_A\big) \circ \widehat{\iota}\Big] \\ &= \sum_{r \leqslant |\omega| \leqslant r\sigma_n} \Big[\gamma_\omega(\cdot) \big((\partial_t)^{\alpha+\beta} \mathcal{Q}_\omega(\widehat{X}_1, \ldots, \widehat{X}_m) \widehat{\gamma}_A\big)\Big] \circ \widehat{\iota}.\end{aligned} \tag{3.29}$$

We now observe that, by (3.24), for all $z \in \mathbb{R}^N$ and $t > 0$ we have

$$\big|(\partial_t)^{\alpha+\beta} \mathcal{Q}_\omega(\widehat{X}_1, \ldots, \widehat{X}_m) \widehat{\gamma}_A(t,z)\big| \leq \mathbf{c}\, t^{-Q/2-\alpha-\beta-|\omega|/2} \exp\left(-\frac{\|z\|^2}{\mathbf{c}\, t}\right),$$

where the constant $\mathbf{c}$ only depends on Λ, r, α and β. Furthermore, since the function γ_ω is smooth and D_λ-homogeneous of degree $|\omega| - r$, a simple homogeneity argument shows that (see, e.g., [8, Prop. 5.1.4])

$$|\gamma_\omega(z)| \leq \mu \, \|z\|^{|\omega|-r} \qquad (\text{for all } z \in \mathbb{R}^N), \tag{3.30}$$

where $\mu > 0$ is a 'structural' constant which can be chosen independently of ω. Gathering (3.29)-(3.30), and bearing in mind the very definition of $\widehat{\iota}$ in (3.9), we then obtain the estimate

$$
\begin{aligned}
&\big|(\partial_t)^{\alpha+\beta}\widehat{X}_{j_1}\cdots\widehat{X}_{j_k}\big((\widehat{X}_{i_1}\cdots\widehat{X}_{i_h}\widehat{\gamma}_A)\circ\widehat{\iota}\big)\big| \\
&\leq \varrho\, t^{-Q/2-r/2-\alpha-\beta}\sum_{r\leqslant|\omega|\leqslant r\sigma_n}\left(\frac{\|z^{-1}\|}{\sqrt{t}}\right)^{|\omega|-r}\exp\left(-\frac{\|z^{-1}\|^2}{\mathbf{c}\,t}\right) \\
&\leq \varrho_1\, t^{-Q/2-r/2-\alpha-\beta}\exp\left(-\frac{\|z^{-1}\|^2}{\varrho_1 t}\right) \qquad (3.31)\\
&\text{(by Remark 2.6)} \\
&= \varrho_1\, t^{-Q/2-r/2-\alpha-\beta}\exp\left(-\frac{\|z\|^2}{\varrho_1 t}\right),
\end{aligned}
$$

holding true for every $z\in\mathbb{R}^N$, $t>0$ and $A\in\mathcal{M}_\Lambda$ (here, $\rho_1>0$ is constant only depending on Λ, r, α and β). Finally, by combining (3.31) with (3.26), we get

$$
\begin{aligned}
&\big|(\partial_t)^{\alpha}\,(\partial_s)^{\beta}\,W_1\cdots W_r\Gamma_A(t,x;s,y)\big| \\
&\leq \frac{\rho_1}{(t-s)^{\alpha+\beta+r/2}}\cdot(t-s)^{-Q/2}\int_{\mathbb{R}^p}\exp\left(-\frac{\|(x,0)^{-1}*(y,\xi)\|^2}{\varrho_1\,(t-s)}\right)d\xi,
\end{aligned}
$$

for every $x,y\in\mathbb{R}^n$, $s<t$ and $A\in\mathcal{M}_\Lambda$. From this, by taking into account (3.22) and (3.18), we obtain (3.12).

(3) As for the proof of (3.12), we distinguish three cases.

Case I: $W_1\cdots W_r = X^x_{i_1}\cdots X^x_{i_r}$. In this case, by using the integral representation formula (3.6) for both Γ_A and Γ_B, we have the estimate

$$
\begin{aligned}
&\big|(\partial_t)^{\alpha}(\partial_s)^{\beta}\,W_1\cdots W_r(\Gamma_A-\Gamma_B)(t,x;s,y)\big| \\
&\leq \int_{\mathbb{R}^p}\big|(\partial_t)^{\alpha+\beta}\,\widehat{X}_{i_1}\cdots\widehat{X}_{i_r}(\widehat{\gamma}_A-\widehat{\gamma}_B)\big|\,\big(t-s,(y,0)^{-1}*(x,\xi)\big)\,d\xi,
\end{aligned} \tag{3.32}
$$

for every $x,y\in\mathbb{R}^n$, $s<t$ and $A,B\in\mathcal{M}_\Lambda$. On the other hand, since the $\widehat{X}_i$'s are Lie-generators for $\mathrm{Lie}(\mathbb{G})$, we can apply once again [9, Thm. 2.5]: there exists a constant $\mathbf{c}=\mathbf{c}_{\Lambda,r,\alpha,\beta}>0$ such that

$$
\begin{aligned}
&\big|(\partial_t)^{\alpha+\beta}\,\widehat{X}_{i_1}\cdots\widehat{X}_{i_r}\widehat{\gamma}_A(t,z)-(\partial_t)^{\alpha+\beta}\,\widehat{X}_{i_1}\cdots\widehat{X}_{i_r}\widehat{\gamma}_B(t,z)\big| \\
&\leq \mathbf{c}\,\|A-B\|^{1/\sigma_n}\,t^{-Q/2-\alpha-\beta-r/2}\exp\left(-\frac{\|z\|^2}{\mathbf{c}\,t}\right).
\end{aligned} \tag{3.33}
$$

With reference to [9, Thm. 2.5], the exponent $1/\sigma_n$ is justified by the fact that the step of nilpotency of $\mathbb{G}$ is precisely $r=\sigma_n$ (see Theorem 2.2).

Gathering (3.32) and (3.33), we then get

$$
\begin{aligned}
&\big|(\partial_t)^{\alpha}(\partial_s)^{\beta}\,W_1\cdots W_r(\Gamma_A-\Gamma_B)(t,x;s,y)\big| \\
&\quad\leq \mathbf{c}\,\|A-B\|^{1/\sigma_n}\,(t-s)^{-\alpha-\beta-r/2}\times \\
&\qquad\times(t-s)^{-Q/2}\int_{\mathbb{R}^p}\exp\left(-\frac{\|(y,0)^{-1}*(x,\xi)\|^2}{\mathbf{c}\,(t-s)}\right)d\xi \qquad (3.34)\\
&\quad\text{(by (3.22) and (3.18))} \\
&\quad\leq \mathbf{c}\,\|A-B\|^{1/\sigma_n}\,\frac{(t-s)^{-\alpha-\beta-r/2}}{\big|B_X(y,\sqrt{t-s})\big|}\exp\left(-\frac{d_X(x,y)^2}{\kappa(t-s)}\right),
\end{aligned}
$$

for every $x, y \in \mathbb{R}^n$, $s < t$ and $A, B \in \mathcal{M}_\Lambda$. Here $\kappa > 0$ is constant only depending on Λ, r, α and β. The desired (3.13) is now a consequence of (3.34) and (3.19).

Case II: $W_1 \cdots W_r = X^y_{j_1} \cdots X^y_{j_r}$. This is very similar to Case I (see also Case II in the proof of (2)), and we omit the details.

Case III: $W_1 \cdots W_r = X^y_{j_1} \cdots X^y_{j_k} X^x_{i_1} \cdots X^x_{i_h}$ (with $h+k=r$). In this last case, by combining (3.29) with the representation formula (3.8) for Γ_A and Γ_B, we can write, for every $x, y \in \mathbb{R}^n$, $s < t$ and $A, B \in \mathcal{M}_\Lambda$,

$$\begin{aligned}
&\big|(\partial_t)^\alpha (\partial_s)^\beta W_1 \cdots W_r (\Gamma_A - \Gamma_B)(t, x; s, y)\big| \\
&\quad \leq \int_{\mathbb{R}^p} \big|(\partial_t)^{\alpha+\beta} \widehat{X}_{j_1} \cdots \widehat{X}_{j_k} \big((\widehat{X}_{i_1} \cdots \widehat{X}_{i_h} \widehat{\gamma}_A) \circ \widehat{\iota}\big) \\
&\qquad\qquad - (\partial_t)^{\alpha+\beta} \widehat{X}_{j_1} \cdots \widehat{X}_{j_k} \big((\widehat{X}_{i_1} \cdots \widehat{X}_{i_h} \widehat{\gamma}_B) \circ \widehat{\iota}\big)\big| d\xi \\
&\quad = \sum_{r \leqslant |\omega| \leqslant r\sigma_n} \int_{\mathbb{R}^p} \Big| \big[\gamma_\omega(\cdot) \big((\partial_t)^{\alpha+\beta} \mathcal{Q}_\omega(\widehat{X}_1, \ldots, \widehat{X}_m)(\widehat{\gamma}_A - \widehat{\gamma}_B)\big)\big] \circ \widehat{\iota} \Big| d\xi,
\end{aligned}$$

where all the integrand functions are evaluated at $\big(t - s, (x, 0)^{-1} * (y, \xi)\big)$. On the other hand, by applying (3.33) to each monomial in $\mathcal{Q}_\omega$, we have

$$\begin{aligned}
&\big|(\partial_t)^{\alpha+\beta} \mathcal{Q}_\omega(\widehat{X}_1, \ldots, \widehat{X}_m) \widehat{\gamma}_A(t, z) - (\partial_t)^{\alpha+\beta} \mathcal{Q}_\omega(\widehat{X}_1, \ldots, \widehat{X}_m) \widehat{\gamma}_B(t, z)\big| \\
&\qquad \leq \mathbf{c} \, \|A - B\|^{1/\sigma_n} \, t^{-Q/2 - \alpha - \beta - |\omega|/2} \exp\left(-\frac{\|z\|^2}{\mathbf{c}\, t}\right),
\end{aligned} \tag{3.35}$$

where the constant $\mathbf{c} = \mathbf{c}_{\Lambda, r, \alpha, \beta} > 0$ can be chosen independently of ω. Gathering (3.35), (3.30) and the definition of $\widehat{\iota}$, we then obtain

$$\begin{aligned}
&\sum_{r \leqslant |\omega| \leqslant r\sigma_n} \Big| \big[\gamma_\omega(\cdot) \big((\partial_t)^{\alpha+\beta} \mathcal{Q}_\omega(\widehat{X}_1, \ldots, \widehat{X}_m)(\widehat{\gamma}_A - \widehat{\gamma}_B)\big)\big] \circ \widehat{\iota} \Big| (t, z) \\
&\qquad \leq \varrho \, \|A - B\|^{1/\sigma_n} \, t^{-Q/2 - \alpha - \beta - r/2} \times \\
&\qquad\qquad \times \sum_{r \leqslant |\omega| \leqslant r\sigma_n} \left(\frac{\|z^{-1}\|}{\sqrt{t}}\right)^{|\omega| - r} \exp\left(-\frac{\|z^{-1}\|^2}{\mathbf{c}\, t}\right) \\
&\qquad \leq \varrho_1 \, \|A - B\|^{1/\sigma_n} \, t^{-Q/2 - \alpha - \beta - r/2} \exp\left(-\frac{\|z^{-1}\|^2}{\varrho_1 t}\right) \\
&\qquad \text{(by Remark 2.6)} \\
&\qquad = \varrho_1 \, \|A - B\|^{1/\sigma_n} \, t^{-Q/2 - \alpha - \beta - r/2} \exp\left(-\frac{\|z\|^2}{\varrho_1 t}\right),
\end{aligned} \tag{3.36}$$

for every $z \in \mathbb{R}^N$, $t > 0$ and $A, B \in \mathcal{M}_\Lambda$. Here $\rho_1 > 0$ is a constant only depending on Λ, r, α and β. With (3.36) at hand, we can establish (3.13) as in Case III of (2). This completes the proof. $\square$

4. Operators with Hölder-continuous coefficients

The aim of this section is to prove existence and several 'structural properties' of a global heat kernel for the variable coefficient operator (1.1). Our proof of Theorem 1.2 is based on a suitable adaptation of the celebrated method developed by E.E. Levi to study uniformly elliptic equations of order $2n$ (see [20, 21]), and later extended to the uniformly parabolic equations (see [18]). As explained in the

Introduction, this approach has been already exploited in [15] to prove an analog of Theorem 1.2 for *generic parabolic Hörmander operators*

$$H = \sum_{i,j=1}^{m} \alpha_{i,j}(t,x)X_iX_j + \sum_{k=1}^{m} \alpha_k(t,x)X_k + \alpha_0(t,x) - \partial_t,$$

under the following 'structural assumptions':

(a) $X_1,\dots,X_m$ are smooth vector fields on $\mathbb{R}^n$ (for some $m = k+n$) and they satisfy Hörmander's rank condition at every point of $\mathbb{R}^n$;

(b) the coefficient functions of H are globally Hölder-continuous, and

$$\alpha_{i,j}(t,x) \equiv \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases} \tag{4.1}$$

for every $k+1 \leq i,j \leq k+n$;

(c) there exists a bounded domain $\Omega_0 \subseteq \mathbb{R}^n$ such that

$$(X_1,\dots,X_k,X_{k+1},\dots,X_{k+n}) \equiv (0,\dots,0,\partial_{x_1},\dots,\partial_{x_n}) \quad \text{on } \mathbb{R}^n \setminus \Omega_0. \tag{4.2}$$

Clearly, the hypotheses of our Theorem 1.2 do not necessarily imply (4.1)-(4.2); however, these 'structural assumptions' play a (key) rôle only in [15, Part I], where the Authors carry out a careful analysis of the constant coefficient operator corresponding to H in order to establish the analog of Theorems 3.1-3.3.

Since in our homogeneous setting we have been able to study constant coefficient operators without requiring (4.2) (and with a totally different approach), we can prove Theorem 1.2 by proceeding *verbatim* as in [15, Part II]: what we only need to check is that all the 'structural' ingredients used in [15] to set up the Levi method are satisfied in our context. We devote the rest of this section to this aim.

4.1. **Heat kernel for constant coefficients operators.** A first fundamental ingredient for the Levi method in [15] is the existence of a 'well-behaved' (global) heat kernel for the constant coefficient operator obtained by freezing the coefficients $a_{i,j}$ (but not the X_i's) at any point $(t_0,x_0) \in \mathbb{R}^{1+n}$, that is,

$$\mathcal{H}_{(t_0,x_0)} = \sum_{i,j=1}^{m} a_{i,j}(t_0,x_0)X_iX_j - \partial_t, \tag{4.3}$$

together with some *uniform estimates* of this kernel with respect to $(t_0,x_0) \in \mathbb{R}^{1+n}$ (see, precisely, [15, Thm. 10.10]). These results in our context are contained in Theorems 3.1 and 3.3.

Let us now introduce a convenient notation which shall play a key rôle in the sequel: following [15], we denote by $\mathbf{E}$ the *Gaussian-type function*

$$\mathbf{E}(x,y,t) := \frac{1}{|B_X(x,\sqrt{t})|} \exp\left(-\frac{d_X(x,y)^2}{t}\right) \qquad (x,y \in \mathbb{R}^n,\ t > 0). \tag{4.4}$$

We explicitly notice that, by exploiting Proposition 3.5, for every $\kappa > 0$ there exists $c = c(\kappa) > 0$ such that

$$\begin{aligned} c(\kappa)^{-1} \frac{1}{|B_X(x,\sqrt{t})|} \exp\left(-\frac{d_X(x,y)^2}{\kappa\, t}\right) &\leq \mathbf{E}(x,y,\kappa t) \\ &\leq c(\kappa) \frac{1}{|B_X(x,\sqrt{t})|} \exp\left(-\frac{d_X(x,y)^2}{\kappa\, t}\right), \end{aligned}$$

for any $x, y \in \mathbb{R}^n$ and any $t > 0$. With this notation at hand, we can rewrite the Gaussian bounds of Theorem 3.3 as follows.

Corollary 4.1. *Under the assumptions of Theorem 1.2, for every* $(t_0, x_0) \in \mathbb{R}^{1+n}$ *let us denote by* $\Gamma_{(t_0,x_0)}$ *the global heat kernel of the constant coefficient operator* (4.3)*. Then, the following estimates hold.*

(1) *There exist constants* κ_Λ, $\nu_\Lambda > 0$ *such that*

$$\frac{1}{\nu_\Lambda}\,\mathbf{E}(x,y,\kappa_\Lambda^{-1}(t-s)) \le \Gamma_{(t_0,x_0)}(t,x;s,y) \le \nu_\Lambda\,\mathbf{E}(x,y,\kappa_\Lambda(t-s)), \tag{4.5}$$

for every $x, y \in \mathbb{R}^n$*, every* $s < t$ *and every* $(t_0, x_0) \in \mathbb{R}^{1+n}$*.*

(2) *For every integer* $r \ge 1$, $i_1, \ldots, i_r \in \{1, \ldots, m\}$ *and every integer* $\alpha \ge 0$*, there exist constants* $\kappa = \kappa_{\Lambda,r,\alpha}$, $\nu = \nu_{\Lambda,r,\alpha} > 0$ *such that*

$$\big|(\partial_t)^\alpha X^x_{i_1}\cdots X^x_{i_r}\Gamma_{(t_0,x_0)}(t,x;s,y)\big| \le \nu\,(t-s)^{-(\alpha+r/2)}\,\mathbf{E}(x,y,\kappa(t-s)),$$

$$\begin{aligned}&\big|(\partial_t)^\alpha X^x_{i_1}\cdots X^x_{i_r}\left(\Gamma_{(t_0,x_0)}(t,x;s,y) - \Gamma_{(t_1,x_1)}(t,x;s,y)\right)\big| \\ &\quad\le \nu\, d_P\left((t_0,x_0),(t_1,x_1)\right)^{\alpha/\sigma_n}(t-s)^{-(\alpha+r/2)}\,\mathbf{E}(x,y,\kappa(t-s)),\end{aligned} \tag{4.6}$$

for every $x, y \in \mathbb{R}^n$, $s < t$ *and* $(t_0, x_0), (t_1, x_1) \in \mathbb{R}^{1+n}$*.*

As already pointed out, all the assertions in the above Corollary are actually contained in Theorems 3.1 and 3.3, since our assumptions imply that

$$A(t_0,x_0) = \big(a_{i,j}(t_0,x_0)\big)_{i,j=1}^m \in \mathcal{M}_\Lambda \qquad \text{for every } (t_0,x_0)\in\mathbb{R}^{1+n}.$$

As for estimate (4.6), if follows from (3.13) and the fact that, since the $a_{i,j}$'s are globally Hölder-continuous (see assumption (i) in Theorem 1.2), we have

$$\begin{aligned}\|A(t_0,x_0) - A(t_1,x_1)\| &\le \left(\sum_{i,j=1}^m |a_{i,j}(t_0,x_0) - a_{i,j}(t_1,x_1)|^2\right)^{1/2} \\ &\le K\, d_P\left((t_0,x_0),(t_1,x_1)\right)^\alpha,\end{aligned}$$

where we have set $K := \max_{i,j}\|a_{i,j}\|_{\alpha,\,\mathbb{R}^{1+n}}$.

4.2. **Metric properties of d_X.** A second important ingredient for the argument in [15] is the validity of the following estimates for the CC-distance d_X, which in [15] are heavily based on assumption (4.2) (see [15, Lem. 2.4]):

- there exists a constant $\mathbf{c} > 0$ such that

$$d_X(x,y) \ge \mathbf{c}|x-y| \qquad \forall\, x,y\in\mathbb{R}^n; \tag{4.7}$$

- for every $\sigma > 0$ there exists a constant $\mathbf{c}(\sigma) > 0$ such that

$$d_X(x,y) \le \mathbf{c}(\sigma)|x-y| \quad \forall\, x,y\in\mathbb{R}^n : \max\{|x-y|, d_X(x,y)\} \ge \sigma. \tag{4.8}$$

Starting from (4.7)-(4.8), it is possible to prove several global properties of the 'geometry' of $(\mathbb{R}^n, d_X)$ which play a key rôle in the analysis of H.

On the other hand, by carefully scrutinizing the proofs in [15, Part II], it is easy to recognize that one *does not really need* (4.7)-(4.8): in fact, the only properties of d_X which intervene in the Levi method are the following:

(a) any B_X-ball is bounded in the Euclidean sense;

(b) there exist constants $\mathbf{c} > 0$ and $\vartheta \geq n$ such that

$$|B_X(x, Mr)| \leq \mathbf{c}M^{\vartheta}|B_X(x,r)| \qquad \forall\ M \geq 1,\ x \in \mathbb{R}^n,\ r > 0;$$

(c) for every $R > 0$ there exists a constant $\mathbf{c}(R) > 0$ such that

$$|B_X(x,r)| \geq \mathbf{c}(R)\, r^{\vartheta} \qquad \forall\ 0 < r \leq R,\ x \in \mathbb{R}^n,$$

where ϑ is as in (b);

(d) for every $\sigma > 0$ there exists a constant $\mathbf{c}(\sigma) > 0$ such that

$$|B_X(x,r)| \geq \mathbf{c}(\sigma)\, r^{n} \qquad \forall\ r \geq \sigma > 0,\ x \in \mathbb{R}^n.$$

Even if we cannot expect that (4.7)-(4.8) hold in our homogeneous setting, the next proposition shows that properties (a)-(d) are still satisfied by d_X.

Proposition 4.2. *Let the assumptions and the notation of Theorem 1.2 do apply. In particular, let* $q = \sum_{i=1}^n \sigma_i \geq n$ *be the* δ_λ*-homogeneous dimension of* $\mathbb{R}^n$*.*

Then, the following facts hold.

(1) *A subset* $B \subseteq \mathbb{R}^n$ *is bounded with respect to* d_X if and only if *it is bounded with respect to the Euclidean distance.*

(2) *There exists a constant* $\gamma > 0$ *such that*

$$|B_X(x, Mr)| \leq \gamma\, M^q |B(x,r)| \qquad \forall\ M \geq 1,\ x \in \mathbb{R}^n,\ r > 0. \tag{4.9}$$

(3) *There exists a constant* $\omega > 0$ *such that*

$$|B_X(x,r)| \geq \omega\, r^q \qquad \forall\ x \in \mathbb{R}^n,\ r > 0. \tag{4.10}$$

(4) *For every* $\sigma > 0$ *there exists a constant* $\mathbf{c} = \mathbf{c}(\sigma) > 0$ *such that*

$$|B_X(x,r)| \geq \mathbf{c}(\sigma)\, r^n \qquad \forall\ x \in \mathbb{R}^n,\ r \geq \sigma. \tag{4.11}$$

Proof. (1) We first suppose that $B \subseteq \mathbb{R}^n$ is bounded in the Euclidean sense. Since the map $x \mapsto d_X(0,x)$ is continuous in the Euclidean sense (see Remark 2.5) and $\overline{B}$ is compact, there exists $r > 0$ such that

$$d_X(0,x) \leq r \qquad \text{for every } x \in \overline{B},$$

thus B is bounded with respect to d_X.

Assume now that $B \subseteq \mathbb{R}^n$ is d_X-bounded, and let $R > 0$ be such that that

$$B \subseteq B_X(0,R).$$

Reminding that $X_1, \ldots, X_m$ satisfy Hörmander's condition in $\mathbb{R}^n$, we know that there exists a small $\rho > 0$ such that $B_X(0,\rho)$ is bounded in Euclidean sense (see, e.g., [3, Prop. 7.21]); on the other hand, since X_i's are also δ_λ-homogeneous of degree 1, from Remark 2.6-(2) we infer that

$$B_X(0,R) = \delta_{R/\rho}\left(B_X(0,\rho)\right).$$

Since $\delta_{R/\rho}$ is a (linear) diffeomorphism of $\mathbb{R}^n$ we deduce that $B_X(0,R)$ is bounded in the Euclidean sense, and thus the same is true of B.

(2) Inequality (4.9) immediately follows from (3.18).

(3) First of all, since $X_1, \ldots, X_m$ satisfy assumptions (H.1) and (H.2), the following *global version* of a celebrated result by Nagel, Stein and Wainger [22] holds true (see [6, Thm. B]): there exist $c_1, c_2 > 0$ such that

$$c_1 \sum_{k=n}^{q} f_k(x) r^k \leq |B_X(x,r)| \leq c_2 \sum_{k=n}^{q} f_k(x) r^k \qquad \forall\ x \in \mathbb{R}^n,\ r > 0, \tag{4.12}$$

where, for any $k \in \{n, \ldots, q\}$, the function $f_k : \mathbb{R}^n \to \mathbb{R}$ is continuous, non-negative and δ_λ-homogeneous of degree $q-k$; in particular, $f_q(x)$ is constant in x and strictly positive. As a consequence, setting $\omega := f_q > 0$, from (4.12) we get (4.10).

(4) By making use of (4.10), and reminding that $q \geq n$, we have

$$|B_X(x,r)| \geq \omega\, r^q = \omega\, r^n\, r^{q-n} \geq \omega\, \sigma^{q-n}\, r^n,$$

for every $x \in \mathbb{R}^n$ and $r \geq \sigma > 0$. This gives (4.11), and the proof is complete. □

Thought not explicit stated, there is another key property concerning d_X which is repeatedly exploited in [15]: for every fixed $x \in \mathbb{R}^n$ it holds that

$$y \mapsto e^{-d_X(x,y)^2} \in L^1(\mathbb{R}^n). \tag{4.13}$$

While in [15] this is an immediate consequence of (4.7) (which, in its turn, follows from assumption (4.2)), in our context we need to prove (4.13) directly.

Lemma 4.3. *For every $p \geq 1$ and every fixed $x \in \mathbb{R}^n$, we have*

$$\mathbf{e}_x(y) := e^{-d_X(x,y)^2} \in L^p(\mathbb{R}^n).$$

Proof. We first observe, since $0 < \mathbf{e}_x \leq 1$ on $\mathbb{R}^n$, we obviously have $\mathbf{e}_x \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$. Thus, reminding that any d_X-ball is bounded in the Euclidean sense (as we know from Theorem 4.2), to prove the lemma it suffices to demonstrate that

$$\mathbf{e}_x \in L^p(\mathbb{R}^n \setminus B_X(0,1)). \tag{4.14}$$

To establish (4.14) we notice that, by triangle's inequality for d_X, we have

$$d_X(0,y)^2 \leq 2d_X(x,y)^2 + 2d_X(0,x)^2,$$

and thus

$$\mathbf{e}_x(y) \leq e^{d_X(0,x)^2} \cdot e^{-\frac{1}{2}d_X(0,y)^2} = c_x\, e^{-\frac{1}{2}\, d_X(0,y)^2} \quad \forall\, y \in \mathbb{R}^n.$$

As a consequence, since the function $y \mapsto d_X(0,y)$ is δ_λ-homogeneous of degree 1 (see Remark 2.6), we obtain the following computation:

$$\begin{aligned}
\int_{\mathbb{R}^n \setminus B_X(0,1)} \mathbf{e}_x^p\,(y)\,dy &\leq c_x^p \int_{\{y:\, d_X(0,y)\geq 1\}} e^{-\frac{p}{2}\, d_X(0,y)^2}\,dy \\
&= c_x^p \sum_{k=0}^{\infty} \int_{\{y:\, 2^k \leq d_X(0,y) < 2^{k+1}\}} e^{-\frac{p}{2}\, d_X(0,y)^2}\,dy \\
&\quad \text{(using the change of variable } y = \delta_{2^k}(u)) \\
&= c_x^p \sum_{k=0}^{\infty} \int_{\{u:\, 1 \leq d_X(0,u) < 2\}} e^{-2^{2k-1}\, p\, d_X(0,u)^2} \cdot 2^{kq}\,du \\
&\leq c_x^p\, |B_X(0,2)| \cdot \sum_{k=0}^{\infty} 2^{kq}\, e^{-2^{2k-1}\, p} < \infty,
\end{aligned}$$

where $q \geq n$ is as in assumption (H.1). □

4.3. **General properties of the d_X-Gaussian function.** The last ingredient for the Levi method in [15] is the validity of several 'structural' properties for the Gaussian-type function $\mathbf{E}$ in (4.4) (see Prop. 10.11 and Cor. 10.12 in [15]). The next proposition shows that all the needed properties are satisfied also in our context.

Proposition 4.4. *Keeping the assumptions and notation of Theorem 1.2, and letting $\mathbf{E} = \mathbf{E}(x, y, t)$ denote the Gaussian-type function defined in (4.4), the following facts hold.*

(1) *There exists a constant $\mathbf{c} > 0$ such that*

$$\mathbf{E}(x, y, t) \le \mathbf{c}\, \beta^{q/2}\, \mathbf{E}(x, y, \beta t), \tag{4.15}$$

for every $x, y \in \mathbb{R}^n$, $t > 0$ and $\beta \ge 1$.

(2) *For every fixed $\mu \ge 0$, there exists $\mathbf{c} = \mathbf{c}_\mu > 0$ such that*

$$\left(\frac{d_X(x, y)^2}{t}\right)^{\mu} \mathbf{E}(x, y, \lambda t) \le \mathbf{c}_\mu\, \lambda^\mu\, \mathbf{E}(x, y, 2\lambda t), \tag{4.16}$$

for every $x, y \in \mathbb{R}^n$, $t > 0$ and $\lambda > 0$.

(3) *For every fixed $\varepsilon > 0$ and $\mu \ge 0$, there exists $\mathbf{c} = \mathbf{c}_{\mu,\varepsilon} > 0$ such that*

$$t^{-\mu}\, \mathbf{E}(x, y, t) \le \mathbf{c}_{\mu,\varepsilon}, \tag{4.17}$$

for every $x, y \in \mathbb{R}^n$ and $t > 0$ satisfying $d(x, y)^2 + t \ge \varepsilon$.

(4) *There exists a constant $\mathbf{c}_0 > 0$ such that, for every $T > 0$, one has*

$$\mathbf{E}(x, y, t)\, \exp\left(\mu d_X(0, y)^2\right) \le \mathbf{c}_0 \mathbf{E}(x, y, 2t)\, \exp\left(2\mu d_X(0, x)^2\right), \tag{4.18}$$

for every $x, y \in \mathbb{R}^n$, $0 < t \le T$ and $0 \le \mu \le 1/(4T)$.

(5) *For every κ_1, $\kappa_2 > 0$, there exist κ_0, $\Theta > 0$ such that*

$$\int_{\mathbb{R}^n} \mathbf{E}(x, \zeta, \kappa_1 t)\, \mathbf{E}(\zeta, y, \kappa_2 t)\, d\zeta \le \Theta\, \mathbf{E}(x, y, \kappa_0 t), \tag{4.19}$$

for every $x, y \in \mathbb{R}^n$ and $t > 0$.

(6) *There exists a constant $\boldsymbol{\sigma} > 0$ such that*

$$\int_{\mathbb{R}^n} \mathbf{E}(x, y, \kappa t)\, dy \le \boldsymbol{\sigma}, \tag{4.20}$$

for every $x \in \mathbb{R}^n$ and t, $\kappa > 0$.

Proof. (1) Since $\beta \ge 1$, by using (4.9) we have

$$\begin{aligned}
\mathbf{E}(x, y, t) &= \frac{|B_X(x, \sqrt{\beta t})|}{|B_X(x, \sqrt{t})|} \cdot \frac{1}{|B_X(x, \sqrt{\beta t})|}\, \exp\left(-\frac{d_X(x, y)^2}{t}\right) \\
&\le \gamma \beta^{q/2}\, \frac{1}{|B_X(x, \sqrt{\beta t})|}\, \exp\left(-\frac{d_X(x, y)^2}{t}\right) \\
&\le \gamma \beta^{q/2}\, \mathbf{E}(x, y, \beta t),
\end{aligned}$$

and this is (4.15).

(2) Since $\mu \ge 0$, we have

$$M_\mu := \sup_{\tau \in [0,\infty)} \tau^\mu\, e^{-\tau/2} \in (0, \infty).$$

As a consequence, taking $\tau := d_X(x,y)^2/(\lambda t) \geq 0$, we obtain

$$\begin{aligned}\left(\frac{d_X(x,y)^2}{t}\right)^{\mu} \mathbf{E}(x,y,\lambda t) &= \frac{\lambda^\mu}{|B_X(x,\sqrt{\lambda t})|}\,\tau^\mu\,e^{-\tau} \leq M_\mu\,\frac{\lambda^\mu}{|B_X(x,\sqrt{\lambda t})|}\,e^{-\tau/2} \\ &\quad \text{(using (4.9) with } r=\sqrt{\lambda t} \text{ and } M=\sqrt{2}) \\ &\leq \gamma\,2^{q/2} M_\mu\,\frac{\lambda^\mu}{|B_X(x,\sqrt{2\lambda t})|}\,e^{-\tau/2} = \mathbf{c}_\mu\,\lambda^\mu\,\mathbf{E}(x,y,2\lambda t),\end{aligned}$$

which is (4.16).

(3) In order to prove (4.17), we distinguish two cases.

(i) $t > \varepsilon/2$. In this case, using (4.10) and the definition of $\mathbf{E}$, we get

$$\begin{aligned} t^{-\mu}\,\mathbf{E}(x,y,t) &\leq \frac{1}{\omega}\,t^{-\mu-q/2}\,\exp\left(-\frac{d_X(x,y)^2}{t}\right) \leq \frac{1}{\omega}\,t^{-\mu-q/2} \\ &\leq \frac{1}{\omega}\,(\varepsilon/2)^{-\mu-q/2} =: \mathbf{c}^{(1)}_{\mu,\varepsilon}.\end{aligned}$$

(ii) $0 < t \leq \varepsilon/2$. In this case, reminding that $d_X(x,y)^2 + t \geq \varepsilon$, we get

$$d_X(x,y)^2 \geq \varepsilon/2.$$

From this, using again (4.10) and the definition of $\mathbf{E}$, we obtain

$$\begin{aligned} t^{-\mu}\,\mathbf{E}(x,y,t) &\leq \frac{1}{\omega}\,t^{-\mu-q/2}\,\exp\left(-\frac{d(x,y)^2}{t}\right) \leq \frac{1}{\omega}\,t^{-\mu-q/2}\,e^{-\varepsilon/(2t)} \\ &\leq \frac{1}{\omega}\,\sup_{\tau\geq 0}\left(\tau^{-\mu-q/2}\,e^{-\varepsilon/(2\tau)}\right) =: \mathbf{c}^{(2)}_{\mu,\varepsilon}.\end{aligned}$$

Collecting the two cases, we infer that (4.17) holds with $\mathbf{c}_{\mu,\varepsilon} := \max\{\mathbf{c}^{(1)}_{\mu,\varepsilon}, \mathbf{c}^{(2)}_{\mu,\varepsilon}\}$.

(4) First of all, using triangle's inequality for the distance d_X, we have

$$d_X(0,y)^2 \leq 2d_X(x,y)^2 + 2d_X(0,x)^2 \qquad \forall\, x,y \in \mathbb{R}^n;$$

as a consequence, for every $x,y\in\mathbb{R}^n$ and every $t>0$ we obtain

$$\begin{aligned} &\exp\left(\mu d_X(0,y)^2\right)\,\mathbf{E}(x,y,t) \\ &\quad\leq \exp\left(2\mu d_X(0,x)^2\right)\cdot\frac{1}{|B_X(x,\sqrt{t})|}\,\exp\left(d_X(x,y)^2(2\mu-1/t)\right) \\ &\quad\text{(using (4.9) with } r=\sqrt{t} \text{ and } M=\sqrt{2}) \\ &\quad\leq \gamma\,2^{q/2}\,\exp\left(2\mu d_X(0,x)^2\right)\cdot\frac{1}{|B_X(x,\sqrt{2t})|}\,\exp\left(d_X(x,y)^2(2\mu-1/t)\right) = (\bigstar).\end{aligned}$$

We now observe that, if $T>0$ is arbitrarily fixed, one has

$$2\mu - \frac{1}{t} \leq \frac{1}{2T} - \frac{1}{t} \leq -\frac{1}{2t}, \tag{4.21}$$

for every $0 < t \leq T$ and $0 \leq \mu \leq 1/(4T)$; using (4.21), we then get

$$\begin{aligned}(\bigstar) &\leq \gamma\,2^{q/2}\,\exp\left(2\mu d_X(0,x)^2\right)\cdot\frac{1}{|B_X(x,\sqrt{2t})|}\,\exp\left(-\frac{d_X(x,y)^2}{2t}\right) \\ &= \mathbf{c}_0\,\exp\left(2\mu d_X(0,x)^2\right)\,\mathbf{E}(x,y,2t),\end{aligned}$$

which is (4.18).

(5) We first observe that, setting $\widehat{\kappa} := \max\{\kappa_1, \kappa_2\}$, we deduce from (4.15) that there exists a constant $\gamma = \gamma_{\kappa_1,\kappa_2} > 0$ such that

$$\mathbf{E}(x,\zeta,\kappa_i t) \leq \gamma\, \mathbf{E}(x,\zeta,\widehat{\kappa} t) \qquad (i=1,2) \tag{4.22}$$

for every $x, \zeta \in \mathbb{R}^n$ and $t > 0$. On the other hand, if $(t_0, x_0) \in \mathbb{R}^{1+n}$ is arbitrarily fixed, we can exploit estimate (4.5): there exist ν_Λ, $\kappa_\Lambda > 0$ such that, for any $x, y, \zeta \in \mathbb{R}^n$ and any $t > 0$, one has

$$\begin{aligned} &\mathbf{E}(x,\zeta,\widehat{\kappa} t) \leq \nu_\Lambda\, \Gamma_{(t_0,x_0)}(\alpha t, x; 0, \zeta) \quad \text{and} \\ &\mathbf{E}(\zeta,y,\widehat{\kappa} t) \leq \nu_\Lambda\, \Gamma_{(t_0,x_0)}(\alpha t, \zeta; 0, y), \end{aligned} \tag{4.23}$$

where $\alpha := \widehat{\kappa}\cdot\kappa_\Lambda > 0$ and $\Gamma_{(t_0,x_0)}$ is the global heat kernel of the constant coefficient operator (4.3). Gathering (4.22)-(4.23), and taking into account the 'reproduction property' of $\Gamma_{(t_0,x_0)}$ (see (3.4)), we obtain

$$\begin{aligned} &\int_{\mathbb{R}^n} \mathbf{E}(x,\zeta,\kappa_1 t)\, \mathbf{E}(\zeta,y,\kappa_2 t) d\zeta \\ &\quad \leq (\gamma\nu_\Lambda)^2 \int_{\mathbb{R}^n} \Gamma_{(t_0,x_0)}(\alpha t, x; 0, \zeta)\, \Gamma_{(t_0,x_0)}(\alpha t, \zeta; 0, y)\, d\zeta \\ &\quad \text{(since } \Gamma_{(t_0,x_0)}(\alpha t, \zeta; 0, y) = \Gamma_{(t_0,x_0)}(0, \zeta; -\alpha t, y)\text{, see Theorem 3.1-(b))} \\ &\quad = (\gamma\nu_\Lambda)^2 \int_{\mathbb{R}^n} \Gamma_{(t_0,x_0)}(\alpha t, x; 0, \zeta)\, \Gamma_{(t_0,x_0)}(0, \zeta; -\alpha t, y)\, d\zeta \\ &\quad = c\, \Gamma_{(t_0,x_0)}(\alpha t, x, -\alpha t, y) =: (\bigstar), \end{aligned}$$

where $c := (\gamma\nu_\Lambda)^2$. Finally, using once again estimate (4.5) we get

$$(\bigstar) \leq (c\,\nu_\Lambda)\, \mathbf{E}(x, y, 2\alpha\kappa_\Lambda t),$$

and this gives (4.19).

(6) If $(t_0, x_0) \in \mathbb{R}^{1+n}$ is arbitrarily fixed, we know from estimate (4.5) that there exist constants ν_Λ, $\kappa_\Lambda > 0$ such that, for any $x, y \in \mathbb{R}^n$ and $t > 0$, one has

$$\mathbf{E}(x,y,\kappa t) \leq \nu_\Lambda \Gamma_{(t_0,x_0)}(\alpha t, x; 0, y),$$

where $\alpha := \kappa \cdot \kappa_\Lambda$. As a consequence, by Theorem 3.1-(d) we have

$$\int_{\mathbb{R}^n} \mathbf{E}(x,y,\kappa t)\, dy \leq \nu_\Lambda \int_{\mathbb{R}^n} \Gamma_{(t_0,x_0)}(\alpha t, x; 0, y)\, dy = \nu_\Lambda,$$

and this is exactly (4.20). □

We conclude this section with the

Proof of Theorem 1.2. The proof follows *verbatim* the arguments in [15, Part II], using Theorem 3.3, Corollary 4.1, Proposition 4.2, Lemma 4.3, Proposition 4.4. □

5. Scale-invariant Harnack inequalities

The aim of this last section is to prove a *scale invariant* Harnack inequality for the variable coefficient operator $\mathcal{H}$ introduced in (1.1), that is,

$$\mathcal{H} = \sum_{i,j=1}^{m} a_{i,j}(t,x)X_iX_j - \partial_t.$$

In dealing with parabolic differential operators, the beautiful connection between Harnack-type inequalities and the availability of two-sided Gaussian bounds for the associated heat kernel was firstly pointed out by Nash [23]. Twenty years later, the approach of Nash was rigorously implemented by Fabes and Stroock [16], also inspired by some ideas of Krylov and Safonov, see [19].

As already explained in the Introduction, however, here we do not prove the Harnack inequality for $\mathcal{H}$ by using the global two-sides Gaussian estimates of its associated heat kernel Γ, since a much simpler approach is possible in our context. Namely, we derive our result from its analog proved in [13] in Carnot groups, using the global lifting result in Theorem 2.2.

Throughout the sequel, we tacitly inherit all the definitions and notation introduced so far. In particular, $X_1,\ldots,X_m$ are smooth vector fields in $\mathbb{R}^n$ satisfying assumptions (H.1), (H.2), (H.3), and $\mathbb{G}$, $\widehat{X} := \{\widehat{X}_1,\ldots,\widehat{X}_m\}$ are as in Theorem 2.2. We remind that $\mathbb{G}$ is a Carnot group whose underlying manifold is $\mathbb{R}^N$, where

$$N = \dim\big(\mathrm{Lie}(X_1,\ldots,X_m)\big) = n + p \qquad \text{(for some } p \geq 1).$$

Accordingly, we denote the points $z \in \mathbb{R}^N$ by

$$z = (x,\xi), \quad \text{with } x \in \mathbb{R}^n \text{ and } \xi \in \mathbb{R}^p.$$

We then introduce the following function space.

Definition 5.1. Let $\Omega \subseteq \mathbb{R}^{1+n}$ be an open set. We define $\mathfrak{C}^2_X(\Omega)$ as the space of functions $u : \Omega \to \mathbb{R}$ satisfying the following properties:

1. u is continuous on Ω;
2. the map $u(t,\cdot)$ has intrinsic-derivatives along the X_i's at every point of its domain, and $X_iu(t,\cdot)$ is continuous for fixed t;
3. the map $u(\cdot,x)$ has derivative with respect to t, and $\partial_t u(\cdot,x)$ is continuous for fixed x;
4. for every fixed $1 \leq i \leq m$, the map $X_iu(t,\cdot)$ has intrinsic-derivatives along the X_j's at every point of its domain, and $X_jX_iu(t,\cdot)$ is continuous for fixed t.

We can now state the main result of this section:

Theorem 5.2 (Parabolic Harnack inequality)**.** *Let the assumptions and the notation of Theorem 1.2 do apply. Moreover, let*

$$r_0 > 0,\ 0 < h_1 < h_2 < 1 \ \textit{and}\ \gamma \in (0,1)$$

be fixed. Then, there exists a constant $M > 0$, only depending on r_0, h_1, h_2 and γ, such that, for every $(t_0,x_0) \in \mathbb{R}^{1+n}$, $r \in (0,r_0]$ and

$$u \in \mathfrak{C}^2_X\big((t_0 - r^2,\, t_0) \times B_X(x_0,r)\big) \cap C\big([t_0 - r^2,\, t_0] \times \overline{B_X(x_0,r)}\big)$$

satisfying $\mathcal{H}u = 0$ *and* $u \geq 0$ *on* $(t_0 - r^2, t_0) \times B_X(x_0, r)$*, we have*

$$\sup_{(t_0 - h_2 r^2,\, t_0 - h_1 r^2) \times B_X(x_0, \gamma r)} u \leq M\, u(t_0, x_0). \tag{5.1}$$

Remark 5.3. On account of [15, Thm. 14.4], if $f \in C^{\alpha}_{X,\mathrm{loc}}$ and u is any $\mathfrak{C}^2_X$-solution of $\mathcal{H}u = f$, then u actually belongs to $C^{2,\alpha}_{X,\mathrm{loc}}$.

In order to prove Theorem 5.2, we need some preliminary lemmas.

Lemma 5.4. *Let* $\Omega \subseteq \mathbb{R}^{1+n}$ *be an open set, and let* $u \in \mathfrak{C}^2_X(\Omega)$*. We denote by* $\pi : \mathbb{R}^N \equiv \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ *the projection of* $\mathbb{R}^N$ *onto* $\mathbb{R}^n$*, and we define*

$$v(t, z) := u\,(t, \pi(z))\,.$$

Then, v *belongs to* $\mathfrak{C}^2_{\widehat{X}}(\Omega \times \mathbb{R}^p)$ *and, for every* $(t, z) \in \Omega \times \mathbb{R}^p$*, we have the identities*

(a) $\partial_t v(t, z) = (\partial_t u)(t, \pi(z))$;
(b) $\widehat{X}_i v(t, z) = (X_i u)(t, \pi(z))$ *for every* $1 \leq i \leq m$;
(c) $\widehat{X}_i \widehat{X}_j v(t, z) = (X_i X_j u)(t, \pi(z))$ *for every* $1 \leq i, j \leq m$.

Proof. First of all, since $u \in \mathfrak{C}^2_X(\Omega)$, it is immediate to see that

(i) v is continuous on $\Omega \times \mathbb{R}^p$;
(ii) for any fixed $z = (x, \xi)$, the map $v(\cdot, z)$ is differentiable with respect to t at any point of its domain, and (a) holds.

In particular, from (a) we recognize that $\partial_t v(\cdot, z)$ is continuous for every fixed z.

Next, let $i, j \in \{1, \ldots, m\}$ be fixed. Since, by Theorem 2.2, we have

$$\widehat{X}_i = X_i + \textstyle\sum_{k=1}^{p} r_{k,i}(x, \xi) \partial_{\xi_k}$$

(for smooth functions $r_{k,i}$), it is easy to recognize that

$$\pi\big(\exp(t\widehat{X}_i)\big)(z) = \exp(tX_i)\big(\pi(z)\big) \qquad \forall\ z \in \mathbb{R}^N,\ t \in \mathbb{R}. \tag{5.2}$$

By (5.2), and recalling the very definition of v, we then get

$$v\big(t, \exp(s\widehat{X}_i)(z)\big) = u\big(t, \pi\big(\exp(s\widehat{X}_i)(z)\big)\big) = u\big(t, \exp(sX_i)(\pi(z))\big)$$

(for all $s \in \mathbb{R}$, $(t, z) \in \Omega \times \mathbb{R}^p$). From this, since $u \in \mathfrak{C}^2_X(\Omega)$, we immediately deduce that $v(t, \cdot)$ has intrinsic-derivative along $\widehat{X}_i$, and

$$\begin{aligned} \widehat{X}_i v\,(t, z) &= \frac{d}{ds}\Big|_{s=0} v\big(t, \exp(s\widehat{X}_i)(z)\big) \\ &= \frac{d}{ds}\Big|_{s=0} u\big(t, \exp(sX_i)(\pi(z))\big) = (X_i u)\,(t, \pi(z))\,. \end{aligned} \tag{5.3}$$

Starting from (5.3), and using once again (5.2), we also have

$$\begin{aligned} \widehat{X}_i v\big(t, \exp(s\widehat{X}_j)(z)\big) &= (X_i u)\big(t, \pi\big(\exp(s\widehat{X}_j)(z)\big)\big) \\ &= (X_i u)\big(t, \exp(sX_i)(\pi(z))\big) \end{aligned}$$

(for all $s \in \mathbb{R}$, $(t, z) \in \Omega \times \mathbb{R}^p$); from this, since $u \in \mathfrak{C}^2_X(\Omega)$, we infer that $v(t, \cdot)$ has intrinsic-derivatives up to second order along the $\widehat{X}_k$'s, which are given by

$$\begin{aligned} \widehat{X}_j \widehat{X}_i v(t, z) &= \frac{d}{ds}\Big|_{s=0} \widehat{X}_i v\big(t, \exp(s\widehat{X}_j)(z)\big) \\ &= \frac{d}{ds}\Big|_{s=0} (X_i u)\big(t, \exp(sX_i)(\pi(z))\big) = (X_j X_i u)(t, \pi(z)), \end{aligned} \tag{5.4}$$

and thus (b)-(c) hold. In particular, from (5.3)-(5.4) (and since $u \in \mathfrak{C}^2_X(\Omega)$) we recognize that $\widehat{X}_i v(t,\cdot)$, $\widehat{X}_j\widehat{X}_i v(t,\cdot)$ are continuous for all fixed t. □

Lemma 5.5. *Let $\Omega \subseteq \mathbb{R}^{1+n}$ be an open set, and let $f \in C^\alpha_X(\Omega)$ (for a suitable $\alpha \in (0,1)$). We let $\pi : \mathbb{R}^N \to \mathbb{R}^n$ be as in Lemma 5.4, and we define*

$$\widehat{f} : \Omega \times \mathbb{R}^p \to \mathbb{R}, \qquad \widehat{f}(t,z) := f(t,\pi(z)).$$

Then, $\widehat{f}$ belongs to $C^\alpha_{\widehat{X}}(\Omega \times \mathbb{R}^p)$, and

$$\|\widehat{f}\|_{\alpha,\,\Omega\times\mathbb{R}^p} \le \|f\|_{\alpha,\,\Omega}. \tag{5.5}$$

Proof. First of all we observe that, since $f \in C^\alpha_X(\Omega)$, one has

$$\sup_{\Omega\times\mathbb{R}^p} |\widehat{f}| = \sup_{\Omega} |f| < \infty. \tag{5.6}$$

Moreover, since we know from Remark 2.7 that

$$d_{\widehat{X}}(z,w) \ge d_X(\pi(z),\pi(w)) \qquad \forall\, z,w \in \mathbb{R}^N \equiv \mathbb{R}^n \times \mathbb{R}^p,$$

we also have, for any (t,z), $(s,w) \in \Omega \times \mathbb{R}^p$:

$$\begin{aligned} |\widehat{f}(t,z) - \widehat{f}(s,w)| &= |f(t,\pi(z)) - f(s,\pi(w))| \\ &\le \|f\|_{\alpha,\,\Omega}\left(d_X(\pi(z),\pi(w))^\alpha + |t-s|^{\alpha/2}\right) \\ &\le \|f\|_{\alpha,\,\Omega}\left(d_{\widehat{X}}(z,w)^\alpha + |t-s|^{\alpha/2}\right). \end{aligned} \tag{5.7}$$

Gathering (5.6) and (5.7), we immediately obtain (5.5). □

Thanks to Lemmas 5.4 and 5.5 we can now deduce the announced scale invariant Harnack inequality for $\mathcal{H}$ from the analogous result holding in Carnot groups.

Proof of Theorem 5.2. Let $(t_0,x_0) \in \mathbb{R}^{1+n}$, $r \in (0,r_0]$ and u be as in the statement of the theorem. Denoting by $\pi : \mathbb{R}^N \to \mathbb{R}^n$ the projection of $\mathbb{R}^N$ onto $\mathbb{R}^n$, we set

$$v : [t_0 - r^2, t_0] \times \overline{B_X(x_0,r)} \times \mathbb{R}^p \to \mathbb{R}, \qquad v(t,z) := u(t,\pi(z)).$$

Moreover, we consider the variable coefficient operator

$$\widehat{\mathcal{H}} := \textstyle\sum_{i,j=1}^m \widehat{a}_{i,j}(t,z)\widehat{X}_i\widehat{X}_j - \partial_t,$$

where $\widehat{A}(t,z) := (\widehat{a}_{i,j}(t,z))_{i,j=1}^m$ is the $m \times m$ matrix of functions defined as

$$\widehat{a}_{i,j}(t,z) := a_{i,j}(t,\pi(z)) \qquad (\text{for } (t,z) \in \mathbb{R}^{1+N}).$$

Using Lemma 5.5, and taking into account properties (i)-(ii) of the matrix $A(t,x)$, we deduce that $\widehat{A}(t,z)$ satisfies the following analogous properties:

(a) $\widehat{a}_{i,j} \in C^\alpha_{\widehat{X}}(\mathbb{R}^{1+N})$ for every $1 \le i,j \le m$;
(b) for every $(t,z) \in \mathbb{R}^{1+N}$ one has $\widehat{A}(t,z) \in \mathcal{M}_\Lambda$.

On the other hand, setting $z_0 := (x_0,0)$, by Remark 2.7 we have

$$\pi\left(B_{\widehat{X}}(z_0,r)\right) = B_X(x_0,\rho) \quad \text{and} \quad \pi\left(\overline{B_{\widehat{X}}(z_0,r)}\right) \subseteq \overline{B_X(x_0,r)},$$

as a consequence, from Lemma 5.4 we deduce the following facts:

(c) $v \in \mathfrak{C}^2_{\widehat{X}}\left((t_0-r^2,t_0) \times B_{\widehat{X}}(z_0,r)\right) \cap C\left([t_0-r^2,t_0] \times \overline{B_{\widehat{X}}(z_0,r)}\right)$;

(d) for every $(t,z) \in (t_0 - r^2, t_0) \times B_{\widehat{X}}(z_0, r)$, one has

$$\begin{aligned}\widehat{\mathcal{H}}v(t,z) &= \sum_{i,j=1}^m \widehat{a}_{i,j}(t,z)\widehat{X}_i\widehat{X}_j v(t,z) - \partial_t v(t,z) \\ &= \sum_{i,j=1}^m a_{i,j}(t,\pi(z))(X_iX_ju)(t,\pi(z)) - (\partial_t u)(t,\pi(z)) \\ &= (\mathcal{H}u)(t,\pi(z)) = 0;\end{aligned}$$

(e) $v \geq 0$ on $(t_0 - r^2, t_0) \times B_{\widehat{X}}(z_0, r)$.

Gathering (a)-(e), and reminding that the $\widehat{X}_i$'s are Lie-generators of the Lie algebra of $\mathbb{G}$, we can apply [13, Thm. 1.1]: there exists a constant $M > 0$, only depending on h_1, h_2, γ and r_0, such that

$$\sup_{(t_0 - h_2 r^2,\, t_0 - h_1 r^2) \times B_{\widehat{X}}(z_0, \gamma r)} v \leq M\, v(t_0, z_0). \tag{5.8}$$

We finally note that, by the very definition of v, the above (5.8) is precisely (5.1). This completes the proof. □

Starting from Theorem 5.2, we immediately obtain (in a standard way) a scale-invariant Harnack inequality for the 'stationary' operator

$$\mathcal{L} := \sum_{i,j=1}^m c_{i,j}(x) X_i X_j.$$

In order to clearly state this result, we first need to introduce the 'stationary' counterparts of the spaces $\mathfrak{C}^2_X$ and $C^{2,\alpha}_X$. We begin with the following

Definition 5.6. Let $U \subseteq \mathbb{R}^n$ be an open set. We define $C^2_X(U)$ as the space of functions $u : U \to \mathbb{R}$ satisfying the following properties:

(1) u is continuous on U;
(2) u has continuous intrinsic-derivatives along the X_i's at every point of U;
(3) u for every fixed $1 \leq i \leq m$, the function $X_i u$ has continuous intrinsic-derivative along the X_j's at every point of U.

We then introduce the 'stationary' Hölder spaces.

Definition 5.7. Let $U \subseteq \mathbb{R}^n$ be an open set, and let $\alpha \in (0,1)$ be fixed. We define $C^\alpha_X(U)$ as the space of all functions $u : U \to \mathbb{R}$ such that

$$\sup_U |u| + \sup_{x \neq y \in U} \frac{|u(x) - u(y)|}{d_X(x,y)^\alpha} < \infty.$$

With Definitions 5.6 and 5.7 at hand, we immediately derive the stationary counterpart of Theorem 5.2.

Theorem 5.8 (Stationary Harnack inequality). *Let $X = \{X_1, \ldots, X_m\} \subseteq \mathcal{X}(\mathbb{R}^n)$ be a family of smooth vector fields satisfying* (H.1)-(H.2)*. Moreover, let*

$$C(x) = (c_{i,j}(x))_{i,j=1}^m$$

be a $m \times m$ matrix of functions such that

(i) *$c_{i,j} \in C^\alpha_X(\mathbb{R}^n)$ for every $i, j = 1, \ldots, m$;*
(ii) *there exists $\Lambda \geq 1$ such that $C(x) \in \mathcal{M}_\Lambda$ for every $x \in \mathbb{R}^{1+n}$;*

and let $\mathcal{L}$ be the variable coefficient operator defined as

$$\mathcal{L} = \sum_{i,j=1}^{m} c_{i,j}(x) X_i X_j.$$

Finally, let $r_0 > 0$ be arbitrarily fixed.

Then, there exists a constant $\mathbf{c} > 0$, only depending on r_0, such that, for every $x_0 \in \mathbb{R}^n$, $r \in (0, r_0]$ and $u \in C^2_X(B_X(x_0, 3r))$ satisfying

$$\mathcal{L}u = 0 \text{ and } u \geq 0 \text{ on } B(x_0, 3r),$$

we have the following inequality

$$\sup_{B_X(x_0,r)} u \leq \mathbf{c} \inf_{B(x_0,r)} u.$$

References

[1] S. Biagi, A. Bonfiglioli: *A completeness result for time-dependent vector fields and applications*, Commun. Contemp. Math. **17** (2015), 1–26.

[2] S. Biagi, A. Bonfiglioli: *The existence of a global fundamental solution for homogeneous Hörmander operators via a global Lifting method*, Proc. Lond. Math. Soc. **114** (2017), no. 5, 855–889.

[3] S. Biagi, A. Bonfiglioli: 'An introduction to the Geometrical Analysis of Vector Fields. With Applications To Maximum Principles And Lie Groups', World Scientific Publishing Company, 2018.

[4] S. Biagi, A. Bonfiglioli: *Global Heat kernels for parabolic homogeneous Hörmander operators*, arXiv:1910.09907 (submitted, 2019).

[5] S. Biagi, M. Bramanti: *Global Gaussian estimates for the heat kernel of homogeneous sums of squares*, arXiv:2003.09845 (submitted, 2020).

[6] S. Biagi, A. Bonfiglioli, M. Bramanti: *Global estimates for the fundamental solution of homogeneous Hörmander sums of squares*, arXiv:1906.07836 (submitted, 2019).

[7] A. Bonfiglioli, E. Lanconelli: *Lie groups related to Hörmander operators and Kolmogorov-Fokker-Planck equations*, Comm. Pure Appl. Anal. **11** (2012), 1587–1614.

[8] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni: 'Stratified Lie Groups and Potential Theory for their sub-Laplacians', Springer Monographs in Mathematics **26**, Springer, New York, N.Y., 2007.

[9] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni: *Uniform Gaussian estimates for the fundamental solutions for heat operators on Carnot groups*, Adv. Differential Equations **7** (2002), 1153–1192.

[10] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni: Fundamental solutions for non-divergence form operators on stratified groups. Trans. Amer. Math. Soc. 356 (2004), no. 7, 2709–2737.

[11] A. Bonfiglioli, F. Uguzzoni: Families of diffeomorphic sub-Laplacians and free Carnot groups. Forum Math. 16 (2004), no. 3, 403–415.

[12] A. Bonfiglioli, F. Uguzzoni: A note on lifting of Carnot groups. Rev. Mat. Iberoamericana 21 (2005), no. 3, 1013–1035.

[13] A. Bonfiglioli, F. Uguzzoni: Harnack inequality for non-divergence form operators on stratified groups. Trans. Amer. Math. Soc. 359 (2007), no. 6, 2463–2481.

[14] M. Bramanti, L. Brandolini: *Schauder estimates for parabolic nondivergence operators of Hörmander type*, J. Differential Equations **234** (2007), 177–245.

[15] M. Bramanti, L. Brandolini, E. Lanconelli, F. Uguzzoni: *Non-divergence equations structured on Hörmander vector fields: heat kernels and Harnack inequalities*, Mem. Amer. Math. Soc. **204** (2010), no. 961 .

[16] E. B. Fabes, D. W. Stroock: *A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash*, Arch. Rational Mech. Anal. **96** (1986), 327–338.

[17] G.B. Folland: *Subelliptic estimates and function spaces on nilpotent Lie groups.* Ark. Mat. **13** (1975), 161–207.

[18] A. Friedman: 'Partial differential equations of parabolic type', Prentice-Hall, Inc., Englewood Cliffs, N.J. 1964.

[19] N. V. Krylov, M. V. Safonov: A property of the solutions of parabolic equations with measurable coefficients. (Russian) Izv. Akad. Nauk SSSR Ser. Mat. 44 (1980), no. 1, 161–175, 239.

[20] E.E. Levi: *Sulle equazioni lineari totalmente ellittiche alle derivate parziali*, Rend. Circ. Mat. Palermo **24** (1907), 275–315.

[21] E.E. Levi: *I problemi dei valori al contorno per le equazioni lineari totalmente ellittiche alle derivate parziali*, Memorie Mat. Fis. Soc. Ital. Scienza (detta dei XL) **16** (1909), 3–113.

[22] A. Nagel, E. M. Stein, S. Wainger: *Balls and metrics defined by vector fields I: Basic properties.* Acta Mathematica, **155** (1985), 130–147.

[23] J. Nash: *Continuity of solutions of parabolic and elliptic equations*, Amer. J. Math. **80** (1958), 931–954.

S. Biagi and M. Bramanti: Politecnico di Milano - Dipartimento di Matematica, Via Bonardi 9, 20133 Milano, Italy

Email address, S. Biagi: stefano.biagi@polimi.it

Email address, M. Bramanti: marco.bramanti@polimi.it