Theory of Stochastic Processes
Vol. 19 (35), no. 1, 2014, pp. 62–90

A. V. RUDENKO

# SOME UNIFORM ESTIMATES FOR THE TRANSITION DENSITY OF A BROWNIAN MOTION ON A CARNOT GROUP AND THEIR APPLICATION TO LOCAL TIMES

For a specific Brownian motion on a Carnot group several estimates for its transition density are established, which are uniform w.r.t. external parameter. These estimates can be used for studying functionals of any Brownian motion on a Carnot group. As an application we show the existence of the renormalized local time for the increments of Levy area. This result has a lot in common with the well-known existence of the renormalized self-intersection local time for two-dimensional Brownian motion.

## 1. Introduction

For any stochastic process $X(t), t \in [0,1]$ in $\mathbb{R}^d$ it is possible to define its functional, known as local time (at zero), by taking a limit of

$$L_\varepsilon = \int_0^1 f_\varepsilon(X(t))dt$$

in $L_2(\Omega)$ as $\varepsilon \to 0+$, where $f_\varepsilon(x) = (2\pi\varepsilon)^{-d/2} e^{-\frac{|x|^2}{2\varepsilon}}$ approximates $\delta$-measure at zero. It is well-known that if $X$ is a $d$-dimensional Brownian motion then the limit exists only for $d = 1$. Similarly self-intersection local time can be defined as a limit of

$$\gamma_\varepsilon = \int_0^1 \int_0^1 f_\varepsilon(X(t) - X(s))dtds$$

and again it exists in $L_2$ only for $d = 1$ if $X$ is a $d$-dimensional Brownian motion. However, it is well-known that in the case $d = 2$ the trajectory of Brownian motion has self-intersections almost surely (and in fact multiple self-intersections, see [4]). This fact suggests that there may be some meaningful functionals describing self-intersections even though self-intersection local time does not exists. Such functional, named renormalized self-intersection local time, can be obtained if we replace $\gamma_\varepsilon$ with "renormalized" $\gamma_\varepsilon - E\gamma_\varepsilon$ in the definition of self-intersection local time (for proof and more details see [11] and references therein). This remarkable fact is known to be important for describing the behaviour of the trajectory of two-dimensional Brownian motion. In particular renormalized self-intersection local time appears in the asymptotics of the area of the small neighbourhood of the trajectory of two-dimensional Brownian motion (see [10]).

But renormalized self-intersection local time does not exist for Brownian motion in any higher dimensions (results confirming this can be found in [7] or [13]). Let us take a fractional Brownian motion with Hurst parameter $H$ as another example. It follows

---

2000 *Mathematics Subject Classification.* 60J60, 60J55.

*Key words and phrases.* Brownian motion on Carnot group, Hormander condition, local time, Levy area.

This work was partially supported by the Presidium of National Academy of Sciences of Ukraine as part of the joint scientific project with the Russian foundation of fundamental research, project number 09-01-14.

from results in [13] (see Theorem 6.2 in [13]) that renormalized version of self-intersection local time exists in $L_2$ for fractional Brownian motion and its self-intersection local time does not exist in $L_2$ at the same time if and only if $H \in [\frac{1}{d}, \frac{3}{2d})$ and $d \geqslant 2$.

It was suggested (see [7] and references therein) that it is possible to obtain additional information about local times utilizing Ito-Wiener expansion, which can be built for any square integrable functional of Brownian motion. More generally if the process is Gaussian or a functional of Gaussian, then Ito-Wiener expansion on a Gaussian space (separable Hilbert space with Gaussian measure) can be used to study local times. Additionally local time can be redefined to be a limit of approximations in Sobolev-Watanabe spaces (Sobolev spaces on Gaussian space). There are handful of papers devoted to this topic, but we only mention [7] and [13] (see bibliography in the latter paper for more references).

Unfortunately the introduction of Sobolev-Watanabe spaces does not allow us to define local time or self-intersection local time for Brownian motion in dimension 3 or greater (even with renormalization). However in [3] another space of functionals on Gaussian space were introduced with the help of special "smoothing" operators. It turned out that the renormalized local time for Brownian motion in any dimension exists in such spaces. Moreover in [14] it was shown that the same is true for any diffusion built as a solution of stochastic differential equation with smooth coefficients under the condition of non-degeneracy of the corresponding diffusion matrix. In other words, diffusions satisfying this condition have, roughly speaking, similar behaviour with regard to local time existence. But there are a lot of diffusions that do not satisfy the condition of non-degeneracy of their diffusion matrix, which behaviour is quite different. One class of such processes is called Brownian motions on a Lie group (the corresponding definition can be found in [8]).

In this paper we are going to develop several ideas, which can be useful in the investigation of local times (and self-intersection local times) for a Brownian motion on a Carnot group (as defined in [2] a Carnot group is a special case of a Lie group on $\mathbb{R}^d$). In particular we show some upper bounds for transition densities of a specific Brownian motion on a Carnot group, which can be used to estimate expectations of local time approximations, such as $L_\varepsilon$. As an application, we establish the existence in $L_2$ of the renormalized local time of a Levy area of the increments of standard two-dimensional Brownian motion.

We define Levy area of the increments of two-dimensional Brownian motion as a two-parameter one-dimensional process:

$$B_{s,t}(W) = \int_s^t (W_u^1 - W_s^1) dW_u^2 - \int_s^t (W_u^2 - W_s^2) dW_u^1 \tag{1}$$

where $(W_t^1, W_t^2)$ is a two-dimensional Brownian motion. We introduce a definition of local time for the Levy area of increments of two-dimensional Brownian motion by replacing $X(t) - X(s)$ with $B_{s,t}$ in the definition of self-intersection local time. The role of self-intersections of $X$ is now taken by zeroes of $B_{s,t}$. We will show that this local time does not exist in $L_2$ but its renormalized version does. As we will see $B_{s,t}$ can also be defined as a coordinate of the increments of a specific Brownian motion on a Carnot group, if subtraction is considered w.r.t. group operation. Therefore bounds for the transition density of a Brownian motion on a Carnot group are applicable.

It is worth noting why we have chosen to consider Brownian motions on a Carnot group. The most important reason is that its behaviour can approximate in some sense the local behaviour of a solution of a large class of stochastic differential equations with smooth coefficients. This approximation was discovered in [12], where hypoelliptic differential operators in the form of sum of squares of smooth vector fields plus a first-order term were considered. The authors studied the regularity of such operators comparing them to the operators of the same kind, built as a sum of squares of vector fields from some graded Lie algebra. But any Brownian motion on a Carnot group is a Markov process, such that the generator of its semigroup has the same form as approximating operators used in [12] (because any Carnot group is a Lie group on whole Euclidean space, which Lie algebra is stratified and therefore graded). See also [6] for a different approach and a different proof of such approximation. Additionally, as discussed in [1], Brownian motion on Carnot group is an important object in the theory of stochastic flows.

Let us briefly describe the main ideas of this paper. To show the existence of renormalized local time we need to have a good representation for $EL_{\varepsilon_1}L_{\varepsilon_2} - EL_{\varepsilon_1}EL_{\varepsilon_2}$, where $L_\varepsilon$ are local time approximations. In [13] one such representation was shown for Gaussian processes (using Ito-Wiener expansion). We are going to use ideas of [14] because any Brownian motion on a Carnot group is also a solution of a stochastic differential equation. We consider two stochastic processes $Y_1^r, Y_2^r$ instead of one process $X$. Each $Y_i^r$ on its own is equal in distribution to $X$, but both together depend smoothly on an external parameter $r \in [0,1]$, such that for $r = 0$ the processes $Y_1^r, Y_2^r$ are independent and for $r = 1$ they are identical. It means that we can write down the difference $EL_{\varepsilon_1}L_{\varepsilon_2} - EL_{\varepsilon_1}EL_{\varepsilon_2}$ as an integral on the parameter $r$ using the joint density of the processes $Y_1^r, Y_2^r$:

$$EL_{\varepsilon_1}L_{\varepsilon_2} - EL_{\varepsilon_1}EL_{\varepsilon_2} = \\ = \int_0^1\int_0^1 (Ef_{\varepsilon_1}(X(s))f_{\varepsilon_2}(X(t)) - Ef_{\varepsilon_1}(X(s))Ef_{\varepsilon_2}(X(t)))dsdt$$

$$Ef_{\varepsilon_1}(X(s))f_{\varepsilon_2}(X(t)) - Ef_{\varepsilon_1}(X(s))Ef_{\varepsilon_2}(X(t)) = \\ = Ef_{\varepsilon_1}(Y_1^1(s))f_{\varepsilon_2}(Y_2^1(t)) - Ef_{\varepsilon_1}(Y_1^0(s))f_{\varepsilon_2}(Y_2^0(t))) = \\ = \int_0^1 \frac{d}{dr}Ef_{\varepsilon_1}(Y_1^r(s))f_{\varepsilon_2}(Y_2^r(t))dr$$

If $X$ is a Brownian motion on a Carnot group then we can choose $(Y_1^r, Y_2^r)$ such that it is also a Brownian motion on a Carnot group and the derivative w.r.t. $r$ of its density can be represented using the derivative w.r.t. $r$ of the corresponding generator.

The coefficients of the stochastic differential equation for a Brownian motion on a Carnot group are such that the corresponding diffusion matrix is, generally speaking, degenerate (however the Hormander condition is always satisfied, therefore the smooth transition density exists), therefore the approach of [14], where we had the non-degeneracy of the diffusion matrix, is not directly applicable. To replicate the argument from [14] we have to establish some estimates for the transition density of a specific Brownian motion on a Carnot group that are uniform w.r.t. external parameter. Fortunately we are able to obtain such estimates from the uniform parabolic Harnack inequality shown in [16], which can be derived from the "uniform" version of Hormander condition.

It is easy to see that the representation of $L_2$-norm of $L_\varepsilon - EL_\varepsilon$ shown above contains some multiple integrals of the derivatives of the joint density of $Y_1^r, Y_2^r$. We can use upper bounds for the joint density directly to find some estimates of such integrals, but they appear to be too weak and not suitable for our purposes. Fortunately, as we discovered in a similar situation in [14], there is a way to produce more accurate estimates for such integrals. We can "move" the derivatives (first order differential operators) inside the integral and apply the upper bounds for the density in a several different ways, which improves overall estimate. This can be done using integration by parts and expressing the derivatives of the transition density w.r.t. starting point in terms of the derivatives w.r.t. ending point and vice versa. The latter is well-known and easy for Gaussian density, since in this case the transition density is a function of the difference between the starting point and the ending point. In the general situation of [14] we had to estimate additional terms appearing after "moving" the derivatives. The transition density of a Brownian motion on a Carnot group is a function of the difference w.r.t. Carnot group addition between the starting point and the ending point. We are able to show that this allows us to "move" the derivatives and improve our estimates.

In the second section we describe our main objects and recall some well-known facts about the transition density of a Brownian motion on a Carnot group. Then we introduce a uniform Hormander condition and show how to obtain it for a specific Brownian motion on a Carnot group with the dependency on an external parameter. After that we establish uniform estimates for the density of a Brownian motion on a Carnot group. In the next section we describe how to obtain sharp estimates for the integrals of the derivatives of the density of a Brownian motion on Carnot group. And in the last section we prove the existence of renormalized local time for the Levy area of the increments of two-dimensional Brownian motion.

## 2. Brownian motion on a Carnot group

Below we give a short description of the notion of Carnot group, Brownian motion on a Carnot group and related objects. For more details about Carnot groups we refer to [2]. Brownian motion on a Lie group was introduced in [8].

First we recall a definition of Carnot group from [2].

**Definition 1.** Lie group $G = (\mathbb{R}^n, \bullet)$ is called a Carnot group if

**1.** $G$ as a Euclidean space can be split into a direct product of Euclidean spaces $G_i$ of fixed dimensions, say $n_1, n_2, \ldots, n_k$ (assuming $\sum_{i=1}^k n_i = n$), such that the following linear isomorphism of $G$ (called dilation)

$$\beta_\lambda(v_1, v_2, \ldots, v_k) = (\lambda v_1, \lambda^2 v_2, \ldots, \lambda^k v_k), v_i \in G_i$$

is a group automorphism of $G$ for all positive $\lambda$.

**2.** Let $g$ be a Lie algebra of left-invariant vector fields on $G$. Fix a coordinate system $x = (x_1, \ldots, x_n) \in G$ such that $x_{\sum_{i=1}^{l-1} n_i+1}, \ldots, x_{\sum_{i=1}^{l} n_i}$ define a vector in $G_l$. Denote as $L_1, \ldots, L_n$ such left-invariant vector fields on $G$, i.e. elements of $g$, that $L_i|_{x=0} = \frac{\partial}{\partial x_i}|_{x=0}$. Then the smallest Lie subalgebra of $g$ containing $L_1, \ldots, L_{n_1}$ is $g$.

Note that left-invariance of $L_i$ is a commutation with the left-shift

$$(L_i f(y \bullet \cdot))(x) = (L_i f)(y \bullet x).$$

and that there always exists a unique left-invariant vector field with given value at $x = 0$.

As shown in [2] such Lie group is also a stratified Lie group. Stratified Lie group is a Lie group such that

**1.** it admits stratification – a direct sum decomposition of its Lie algebra $g = \oplus_{i=1}^k g_i$, such that $[g_1, g_i] = g_{i+1}, 1 \leqslant i \leqslant k-1$ and $[g_1, g_k] = 0$,

**2.** the smallest Lie subalgebra of $g$ containing $g_1$ is $g$.

We note that $\dim(g_i) = n_i$ and $L_{\sum_{i=1}^{l-1} n_i+1}, \ldots, L_{\sum_{i=1}^{l} n_i}$ is a basis of $g_l$. If $L_i \in g_l$ we denote $d(L_i) = l$. We also denote $d(G) = \sum\limits_{i=1}^{k} i n_i = \sum\limits_{i=1}^{n} d(L_i)$, which is called homogeneous dimension of $G$. We can define a convolution for any measurable non-negative (or square-integrable) pair of functions $f$ and $g$ on $G$ as

$$(f *_G g)(x) = \int_G f(y^{-1} \bullet x) g(y) dy \tag{2}$$

It is easy to see that this operation is associative but, generally speaking, not commutative.

We say that a polynomial $Q(x)$ on $\mathbb{R}^n$ is a homogeneous polynomial on $G$ of homogeneous degree $d = d(Q)$ if for all real $\lambda$ and $x \in G$ the following equality holds: $Q(\beta_\lambda(x)) = \lambda^d Q(x)$. It is well known that every smooth (meaning infinitely differentiable) function that satisfies such relation for a positive integer $d$ is in fact a polynomial on $G$. If we take any homogeneous polynomial $Q$ of homogeneous degree $d(Q)$ then $L_i Q$ is a homogeneous polynomial of homogeneous degree $d(L_i Q) = d(Q) - d(L_i)$. Note that zero polynomial is a homogeneous polynomial of any degree and it is the unique homogeneous polynomial of negative degree.

Below we list several well-known facts about Carnot groups, which are helpful in our investigation.

**1.** Vector fields $L_i$ are homogeneous with homogeneous degree $d(L_i)$. If we consider $L_i$ as a function from $G$ to $G$, then the homogenuity of $L_i$ is defined as follows:

$$\beta_\lambda(L_i(x)) = \lambda^{d(L_i)} L_i(\beta_\lambda(x)).$$

In the operator sense it is equivalent to

$$(L_i f(\beta_\lambda(\cdot)))(x) = \lambda^{d(L_i)} (L_i f)(\beta_\lambda(x)).$$

**2.** Vector field $L_i$ has the following form

$$L_i = \sum_{j=1}^{n} a_{ij}(x) \frac{\partial}{\partial x_j} \tag{3}$$

in the basis from the definition of $G$, where the functions $a_{ij}$ are polynomials homogeneous on $G$ of degree $d(L_j) - d(L_i)$, and therefore $a_{ij}$ does not depend on $x_l$ if $d(L_l) \geqslant d(L_j)$ and equal to zero if $d(L_j) < d(L_i)$ (see for example p.35 of [2]).

**3.** The group operation on $G$ can be represented as

$$(x \bullet y)_i = x_i + y_i + Q_i(x, y) \tag{4}$$

in the basis from the definition of $G$, where $Q_i$ are polynomials homogeneous on $G$ of degree $d(L_i)$, and $Q_i$ does not depend on $x_j, y_j$ if $d(L_j) \geqslant d(L_i)$ (see for example Theorem 1.3.15 of [2]).

**4.** Any map on $G$ of the form $x \to x \bullet y$, $x \to y \bullet x$ or $x \to x^{-1}$ preserves Lebesgue measure.

**5.** The integral $\int_G L_i f(x) dx$ is equal to 0, as long as $L_i f$ exists and is integrable, which enables us to integrate by parts with $L_i$.

Denote as $X(W)$ a strong solution of

$$dX_t = \sum_{i=1}^{n_1} L_i(X_t) \circ dW_t^i; X(0) = x \tag{5}$$

where $W_t$ is a $n_1$-dimensional Wiener process and $\circ$ before $dW$ means that corresponding stochastic integral w.r.t. $W$ is a Stratonovich integral (see [9]). There is a unique strong solution since in our chosen basis this system of equations is "triangular" (due to (3)) and can be integrated equation by equation in some order. The solution is a Brownian motion on a Lie group $G$ in the sense of [8] (see proof of Proposition 1 below).

The definition of Carnot group provides that $L_1, \ldots, L_{n_1}$ satisfy the Hormander condition and therefore (see for example, Theorem 7.4.20 in [15]) there is a density $p(t,x,\cdot)$ of $X_t(W)$ for each $t>0$ and $X(0)=x$, which is smooth in all variables for $t>0$. In the following proposition we gathered several important well-known properties of $p$.

Denote $D = \frac{1}{2}\sum\limits_{i=1}^{n_1} L_i^2$.

**Proposition 1.** **1.** *For every continuous $f$ with compact support the function*

$$\psi(t,x) = \int\limits_{\mathbb{R}^n} p(t,x,y)f(y)dy$$

*is a solution to the following Cauchy problem*

$$(D - \frac{\partial}{\partial t})\psi(t,x) = 0, (t,x) \in (0,+\infty] \times \mathbb{R}^n; \psi(0+,x) = f(x), x \in \mathbb{R}^n \tag{6}$$

*Moreover*

$$(D - \frac{\partial}{\partial t})p(\cdot,\cdot,y)(t,x) = 0$$

*for all $x,y \in G$ and $t>0$.*

**2.** *There exists a function $\tilde{p}(t,x)$ on $(t,x) \in (0,+\infty) \times \mathbb{R}^n$, such that $p(t,x,y) = \tilde{p}(t, y^{-1} \bullet x)$.*

**3.** *For all $0 < s < t$ we have $\tilde{p}(t,\cdot) = \tilde{p}(s,\cdot) *_G \tilde{p}(t-s,\cdot)$*

**4.** *For all $\lambda > 0$ we have $\tilde{p}(\lambda^{-2}t, x) = \lambda^{d(G)}\tilde{p}(t, \beta_\lambda(x))$*

*Proof.* Below we state several results from [8] that we need. First of all there exists a unique Markov process on $G$ with the transition function $F$ such that it is connected with the operator $D$ by the following formula

$$Df(x) = \lim_{t\to s+} \frac{1}{t-s}\int\limits_G f(y)F(s,x,t,dy) \tag{7}$$

for any twice continuously differentiable function $f$ (we define the transition function $F(s,x,t,A)$ as a probability that the process is in the set $A$ at the time $t$ if it started from $x$ at the time $s$). Additionally the transition function $F$ is time-homogeneous and $G$-invariant. It was also proven that $\int\limits_G f(y)F(s,x,t,dy)$ is a solution to a Cauchy problem (6).

If we apply Ito formula to $f(X)$ we obtain easily that (7) holds if $F$ is the transition function of $X$. Therefore the Markov process built for $D$ is in fact coincides with $X$ . Since the transition function of $X$ is an integral of the transition density (which exists because of the Hormander condition as we mentioned above), all our properties follow except for homogenuity w.r.t. dilations (property 4). But it can be proven if we find stochastic differential equation for a process $\beta_\lambda(X(\lambda^{-2}t))$, use homogenuity of $L_i$, and notice that resulting equation coincides with (5).

See also Proposition 1.68 on p.56 of [5] for an alternative proof.

□

In the following we will simply write $p(t,x)$ instead of $\tilde{p}(t,x)$. Note that the solution of the Cauchy problem (6) can be written as $\psi(t,x) = (p(t,\cdot) *_G f)(x)$ We emphasize that according to our definitions $p(t,x,y) = p(t, y^{-1}\bullet x)$ is a density of $X_t$ at $y$ if $X_0 = x$, which satisfies $(D - \frac{\partial}{\partial t})p(\cdot,\cdot,y) = 0$.

**Example 1.** Suppose that $G$ is a Heisenberg group, or more precisely that we have

$$\begin{gathered} n=3, k=2, n_1=2, n_2=1 \\ x \bullet y = (x_1+y_1, x_2+y_2, x_3+y_3+x_1y_2-x_2y_1) \end{gathered} \tag{8}$$

It is well-known that such operation defines a Carnot group on $\mathbb{R}^3$ and that vector fields $L_i$ are as follows:

$$L_1 = \frac{\partial}{\partial x_1} - x_2 \frac{\partial}{\partial x_3}$$
$$L_2 = \frac{\partial}{\partial x_2} + x_1 \frac{\partial}{\partial x_3}$$
$$L_3 = \frac{\partial}{\partial x_3}$$

Solving (5) we can see that $p(t, x^{-1})$ is a joint density of $(W_t^1, W_t^2, B_{0,t}(W))$. Therefore density of $B_{s,t}$ is an integral of $p$ and if we can estimate $p$ then we can estimate expectations of functions of $B_{s,t}$.

Now we are going to construct another Brownian motion on a Carnot group, that depends on a parameter $r \in [0,1]$. Let $\{W_{1,r,s}, s \geqslant 0\}$ and $\{W_{2,r,t}, t \geqslant 0\}$ be two $n_1$-dimensional Brownian motions, that are jointly Gaussian, such that covariation matrix between vectors $W_{1,r,s}$ and $W_{2,r,t}$ equal to $r \min(s,t) I$. It turns out that $Y_t^r = (X_t(W_{1,r}), X_t(W_{2,r}))$ is a Brownian motion on the Carnot group $G \times G$ (it follows from the proof of Proposition 2 below). We use the same notation $\bullet$ for group operation on $G \times G$: if $x_1, y_1, x_2, y_2 \in G$ and $z_1 = (x_1, y_1), z_2 = (x_2, y_2) \in G \times G$ then

$$z_1 \bullet z_2 = (x_1 \bullet x_2, y_1 \bullet y_2)$$

Although $Y$ is essentially different from $X$ defined on $G \times G$ for $r \neq 0$ (if $r = 0$ then $Y$ is the same as $X$ defined on the Carnot group $G \times G$), it has similar properties. Let $p_r(t, x, \cdot)$ be a density of $Y_t^r$ for each $t > 0$ and $Y_0^r = x$. Denote

$$D^r = \frac{1}{2} \sum_{i=1}^{n_1} (L_i^x)^2 + \frac{1}{2} \sum_{i=1}^{n_1} (L_i^y)^2 + r \sum_{i=1}^{n_1} L_i^x L_i^y.$$

If we define

$$V_i^r = \frac{\sqrt{1+r}}{2}(L_i^x + L_i^y), V_{i+n_1}^r = \frac{\sqrt{1-r}}{2}(L_i^x - L_i^y)$$

for $i = 1, \ldots, n_1$ then $D^r = \sum_{i=1}^{2n_1} (V_i^r)^2$. Note that $\{V_i^r, i = 0, \ldots, 2n_1\}$ is a set of smooth vector fields, satisfying the Hormander condition for $r \in (0,1)$.

**Proposition 2.** *All statements in Proposition 1 remains true for any $r \in (0,1)$ if $G$, $D$ and $p$ are replaced by $G \times G$, $D^r$ and $p_r$ respectively.*

*Proof.* By the definition $Y_s^r$ is a solution of two copies of the equation (5) with $W_t$ replaced by $W_{1,r,t}$ and $W_{2,r,t}$ respectively. Applying Ito formula to $f(Y_t^r)$ (see Theorem 17.18 in [9]) we can find an equivalent of (7) for $Y^r(t)$ with $D$ replaced by $D^r$. Therefore the proof of Proposition 1 is applicable with $G$, $D$ and $p$ replaced by $G \times G$, $D^r$ and $p_r$ respectively. ☐

The most important consequence of the above is that $p_r(t, y^{-1} \bullet x)$ is a density of $Y_t$ at $y$ if $Y_0 = x$, and it satisfies $(D^r - \frac{\partial}{\partial t}) p_r(t, y^{-1} \bullet \cdot) = 0$.

## 3. Uniform Hormander condition

We have already mentioned that $\{V_i^r, i = 0, \ldots, 2n_1\}$ satisfy the Hormander condition for $r \in (0,1)$. But they depend continuously on $r$ and in the limit as $r \to 1+$ we get degeneration (for $r = 1$ we have $V_{i+n_1}^r = 0$ and the Hormander condition is not satisfied). Therefore a lot of care should be taken in order to obtain estimates for $p_r$ that are uniform w.r.t $r \in (0,1)$.

Such estimates can be found under the uniform Hormander condition (w.r.t. external parameter), that was proposed in [16]. Under this condition the authors proved a uniform parabolic Harnack inequality. We are going to introduce a change of variables that makes Hormander condition uniform on $r$. As a result we will be able to show a variant of uniform parabolic Harnack inequality for $p_r$ which can be used to prove several uniform estimates for $p_r$.

Define a pair new variables as $u = \frac{x+y}{2}$, $v = \frac{x-y}{2\sqrt{1-r}}$. The corresponding change of variables applies to any differential operator in a standart way. A new operator $\tilde{D}^r$ is related to the old as follows $\tilde{D}^r f(u,v) = D^r f(u(x,y), v(x,y))$. Such operation commutes with sum and multiplication of operators and therefore $\tilde{D}^r = \sum_{i=1}^{2n_1} (\tilde{V}_i^r)^2$, where $\tilde{V}_i^r f(u,v) = V_i^r f(u(x,y), v(x,y))$. After applying change of variables to the derivatives w.r.t. $x, y$ we obtain that operator $\frac{\partial}{\partial x_i}$ changes into

$$\frac{1}{2}\frac{\partial}{\partial u_i} + \frac{1}{2\sqrt{1-r}}\frac{\partial}{\partial v_i},$$

and operator $\frac{\partial}{\partial y_i}$ changes into

$$\frac{1}{2}\frac{\partial}{\partial u_i} - \frac{1}{2\sqrt{1-r}}\frac{\partial}{\partial v_i}$$

Therefore denoting $L_i^x = \sum_{j=1}^{n} a_{ij}(x)\frac{\partial}{\partial x_j}$ we can find for $i = 1, \ldots, n_1$

$$\tilde{V}_i^r = \frac{\sqrt{1+r}}{4}\sum_{j=1}^{n}(a_{ij}(u+v\sqrt{1-r}) + a_{ij}(u-v\sqrt{1-r}))\frac{\partial}{\partial u_j} + \\ + \frac{\sqrt{1+r}}{4\sqrt{1-r}}\sum_{j=1}^{n}(a_{ij}(u+v\sqrt{1-r}) - a_{ij}(u-v\sqrt{1-r}))\frac{\partial}{\partial v_j},$$

$$\tilde{V}_{i+n_1}^r = \frac{\sqrt{1-r}}{4}\sum_{j=1}^{n}(a_{ij}(u+v\sqrt{1-r}) - a_{ij}(u-v\sqrt{1-r}))\frac{\partial}{\partial u_j} + \\ + \frac{1}{4}\sum_{j=1}^{n}(a_{ij}(u+v\sqrt{1-r}) + a_{ij}(u-v\sqrt{1-r}))\frac{\partial}{\partial v_j}.$$

It is easy to see that in the limit as $r \to 1+$ each $\tilde{V}_i^r$ converges to some smooth vector field. Below we will show that $\tilde{V}_i^r$ satisfy Hormander condition uniformly with respect to $r \in (0,1)$. But first we have to introduce the corresponding definition.

For every multiindex $J = (j_1, \ldots, j_k)$ we denote repeated commutation of vector fields as $V_J = [V_{j_1}, [V_{j_2} \ldots [V_{j_{k-1}}, V_{j_k}] \ldots]]$. The following definition is taken from [16] (and simplified, since we have $\mathbb{R}^n$ instead of general $n$-dimensional manifold).

**Definition 2** (N.Th. Varopolous, L. Saloff-Coste, T. Coulhon)**.** Smooth vector fields $H_i^r, i = 1, \ldots, m$ on $\mathbb{R}^n$ defined for some set of parameters $r$ are said to satisfy Hormander condition uniformly with respect to $r$, if

1. all the coefficients of $H_i^r$ and all their derivatives are bounded on any compact uniformly with respect to $r$,
2. for any $x \in \mathbb{R}^n$ there exists a set of multiindices $J_1, \ldots, J_n$, such that for all $r$ the matrix $M(x) = (H_{J_1}^r(x), \ldots, H_{J_n}^r(x))$ is non-degenerate and there exists an open neighbourhood of $x$ such that the coefficients of matrices $M(x)$ and $M^{-1}(x)$ and

all their derivatives are bounded in this neighbourhood uniformly with respect to $r$.

The following theorem can be found in [16]. It shows that the uniform Hormander condition provides uniform estimates for solutions of equations similar to $(D^r - \frac{\partial}{\partial t})\psi = 0$. Compared to the original we have $\mathbb{R}^n$ instead of smooth connected manifold.

**Theorem 1** (N.Th. Varopolous, L. Saloff-Coste, T. Coulhon)**.** *Suppose that $H_i^r, i = 1, \dots, m$ are smooth vector fields on $\mathbb{R}^n$ satisfying Hormander condition uniformly with respect to $r$, $h^r$ and $H_0^r$ is a smooth function and vector field respectively, both bounded with all derivatives uniformly w.r.t. $r$ on any compact and let $A^r = \sum_{i=1}^m (H_i^r)^2 + H_0^r + h^r$. Then for any compact $K$ in $\mathbb{R}^n$, $t_1 < t_2 < t_3 < t_4$, non-negative integer $i$ and multiindex $J = (j_1, \dots, j_l)$ there exists a constant $C > 0$ such that for all values of parameter $r$, and every positive function $\psi$ satisfying $(A^r - \frac{\partial}{\partial t})\psi = 0$ on $[t_1, t_4] \times \mathbb{R}^n$ we have:*

$$\sup_{x \in K} |(\frac{\partial}{\partial t})^i (\frac{\partial}{\partial x})^J \psi(t_2, x)| \leqslant C \inf_{x \in K} \psi(t_3, x) \tag{9}$$

*where* $(\frac{\partial}{\partial x})^J = \frac{\partial}{\partial x_{j_1}} \cdots \frac{\partial}{\partial x_{j_l}}$.

These inequality is called a parabolic Harnack inequality, and, since we also have uniformity w.r.t. $r$, we call it a uniform parabolic Harnack inequality. To make use of this theorem we need to check the uniform Hormander condition for $\tilde{V}_i^r$.

**Lemma 1.** *Vector fields $\tilde{V}_i^r, i = 1, \dots, 2n_1$ satisfy Hormander condition uniformly with respect to $r \in (0, 1)$.*

*Proof.* We know that $L_i, i = 1, \dots, n_1$ satisfy the Hormander condition, i.e. there exists a set of multiindices $J_1, \dots, J_n$ and smooth functions $b_{ij}$, $b_{ij}^{-1}$ for $i, j = 1, \dots, n$, such that

$$L_{J_i} = \sum_{j=1}^n b_{ij} \frac{\partial}{\partial x_j}$$

$$\frac{\partial}{\partial x_i} = \sum_{j=1}^n b_{ij}^{-1} L_{J_j}$$

Let us define more multiindices $J_{n+1}, \dots, J_{2n}$ such that $j_{i+n,1} = j_{i,1} + n_1$ and $j_{i+n,k} = j_{i,k}$ for $k > 1$. The expanded set of multiindices corresponds to some commutators for operators $V_i^r$. Calculating these commutators directly we obtain

$$V_{J_i}^r = (1+r)^{|J_i|/2} 2^{-|J_i|} (L_{J_i}^x + L_{J_i}^y)$$

and

$$V_{J_{i+n}}^r = \sqrt{1-r}(1+r)^{(|J_i|-1)/2} 2^{-|J_i|} (L_{J_i}^x - L_{J_i}^y)$$

for $i = 1, \dots, n$, where $|J_i|$ is the number of indices in the multiindex $J_i$. We note that $L_{J_i}^x, L_{J_i}^y$ can be written as a linear combination of $V_{J_i}^r, V_{J_{i+n}}^r$. After doing that we get for each $r \in (0, 1)$

$$\frac{\partial}{\partial x_i} = \sum_{j=1}^n b_{ij}^{-1}(x) 2^{|J_j|-1} ((1+r)^{-|J_j|/2} V_{J_j}^r + \frac{1}{\sqrt{1-r}} (1+r)^{-(|J_j|-1)/2} V_{J_{j+n}}^r)$$

$$\frac{\partial}{\partial y_i} = \sum_{j=1}^n b_{ij}^{-1}(y) 2^{|J_j|-1} ((1+r)^{-|J_j|/2} V_{J_j}^r - \frac{1}{\sqrt{1-r}} (1+r)^{-(|J_j|-1)/2} V_{J_{j+n}}^r)$$

Changing variables to $u = \frac{x+y}{2}$, $v = \frac{x-y}{2\sqrt{1-r}}$ (same as above) we can see that

$$\tilde{V}^r_{J_i} =$$
$$= \sum_{j=1}^{n} (b_{ij}(u + v\sqrt{1-r}) + b_{ij}(u - v\sqrt{1-r}))(1+r)^{|J_j|/2} 2^{-|J_j|-1} \frac{\partial}{\partial u_j} +$$
$$+ \frac{1}{\sqrt{1-r}} \sum_{j=1}^{n} (b_{ij}(u + v\sqrt{1-r}) - b_{ij}(u - v\sqrt{1-r}))(1+r)^{|J_j|/2} 2^{-|J_j|-1} \frac{\partial}{\partial v_j},$$

$$\tilde{V}^r_{J_{i+n}} =$$
$$= \sqrt{1-r} \sum_{j=1}^{n} (b_{ij}(u + v\sqrt{1-r}) - b_{ij}(u - v\sqrt{1-r}))(1+r)^{(|J_j|-1)/2} 2^{-|J_j|-1} \frac{\partial}{\partial u_j} +$$
$$+ \sum_{j=1}^{n} (b_{ij}(u + v\sqrt{1-r}) + b_{ij}(u - v\sqrt{1-r}))(1+r)^{(|J_j|-1)/2} 2^{-|J_j|-1} \frac{\partial}{\partial v_j},$$

and

$$\frac{\partial}{\partial u_i} =$$
$$= \sum_{j=1}^{n} (b^{-1}_{ij}(u + v\sqrt{1-r}) + b^{-1}_{ij}(u - v\sqrt{1-r}) 2^{|J_j|-1} (1+r)^{-|J_j|/2} \tilde{V}^r_{J_j} +$$
$$+ \frac{1}{\sqrt{1-r}} \sum_{j=1}^{n} (b^{-1}_{ij}(u + v\sqrt{1-r}) - b^{-1}_{ij}(u - v\sqrt{1-r}) 2^{|J_j|-1} (1+r)^{-(|J_j|-1)/2} \tilde{V}^r_{J_{j+n}}$$

$$\frac{\partial}{\partial v_i} =$$
$$= \sqrt{1-r} \sum_{j=1}^{n} (b^{-1}_{ij}(u + v\sqrt{1-r}) - b^{-1}_{ij}(u - v\sqrt{1-r}) 2^{|J_j|-1} (1+r)^{-|J_j|/2} \tilde{V}^r_{J_j} +$$
$$+ \sum_{j=1}^{n} (b^{-1}_{ij}(u + v\sqrt{1-r}) + b^{-1}_{ij}(u - v\sqrt{1-r}) 2^{|J_j|-1} (1+r)^{-(|J_j|-1)/2} \tilde{V}^r_{J_{j+n}}$$

We can check that all the coefficients in the above and their derivatives are bounded on any compact in $\mathbb{R}^{2n}$ uniformly with respect to $r$. Indeed every coefficient satisfies this property in an obvious way except those of the form $\frac{1}{\sqrt{1-r}}(f(u+v\sqrt{1-r}) - f(u-v\sqrt{1-r}))$ (omitting an additional uniformly bounded multiplier), where $f$ is some smooth function that does not depend on $r$. But every such expression can be represented as $\int\limits_{-1}^{1} \sum\limits_{i=1}^{n} v_i f_i'(u + v\gamma\sqrt{1-r}) d\gamma$. Now it is easy to see that this expression and all of its derivatives w.r.t. $u$,$v$ are also bounded on any compact uniformly with respect to $r$. Therefore we have the desired uniform Hormander condition. □

As a consequence we can establish something similar to a uniform parabolic Harnack inequality for the solutions of $(D^r - \frac{\partial}{\partial t})\psi = 0$.

**Corollary 1.** *For any fixed compact $K$ in $\mathbb{R}^{2n}$, $0 < s < t$ and positive integers $a, b$ and multiindices $J_1 = (J_{11}, \ldots, J_{1a})$, $J_2 = (J_{21}, \ldots, J_{2b})$ with values $1, 2, \ldots, n$ there exists a* *constant $C > 0$ such that the following inequality holds for all $r \in [0,1]$ and every positive solution $\psi$ of $(D^r - \frac{\partial}{\partial t})\psi = 0$ on $(0,+\infty) \times \mathbb{R}^{2n}$*

$$\sup_{(x,y)\in K_r} |(\frac{\partial}{\partial x} + \frac{\partial}{\partial y})^{J_1}(\frac{\partial}{\partial x} - \frac{\partial}{\partial y})^{J_2}\psi(s,x,y)| \leqslant C(1-r)^{-b/2} \inf_{(x,y)\in K_r} \psi(t,x,y) \tag{10}$$

*where $K_r = \{(x,y) : (\frac{x+y}{2}, \frac{x-y}{2\sqrt{1-r}}) \in K\}$ and $(\frac{\partial}{\partial x} - \frac{\partial}{\partial y})^{J_1} = (\frac{\partial}{\partial x_{J_{11}}} - \frac{\partial}{\partial y_{J_{11}}}) \dots (\frac{\partial}{\partial x_{J_{1a}}} - \frac{\partial}{\partial y_{J_{1a}}})$ ($(\frac{\partial}{\partial x} + \frac{\partial}{\partial y})^{J_2}$ is defined similarly).*

*Proof.* We can apply change of variables as above ($u = \frac{x+y}{2}$, $v = \frac{x-y}{2\sqrt{1-r}}$) to any positive solution $\psi$ of $(D^r - \frac{\partial}{\partial t})\psi = 0$. As a result we obtain positive solution $\tilde{\psi}$ of $(\tilde{D}^r - \frac{\partial}{\partial t})\tilde{\psi} = 0$. Then, due to Lemma 1, Theorem 1 provides that for any fixed compact $K$ in $\mathbb{R}^{2n}$, $0 < s < t$ and multiindices $J_1$, $J_2$ there exists a constant $C > 0$ such that

$$\sup_{(u,v)\in K} |(\frac{\partial}{\partial u})^{J_1}(\frac{\partial}{\partial v})^{J_2}\tilde{\psi}(s,u,v)| \leqslant C \inf_{(u,v)\in K} \tilde{\psi}(t,u,v)$$

where constant $C$ does not depend on choice of $r$ and $\psi$.

Then we change variables back to $x, y$ in the inequality, noting that the derivative $\frac{\partial}{\partial u_i}$ transforms to

$$\frac{\partial}{\partial x_i} + \frac{\partial}{\partial y_i},$$

and $\frac{\partial}{\partial v_i}$ transforms to

$$\sqrt{1-r}(\frac{\partial}{\partial x_i} - \frac{\partial}{\partial y_i})$$

We obtain the inequality (10) and the corollary is proven. □

Note that all the proofs in this section do not use left-invariance or homogeneuity of $L_i$ or any other property related to the Carnot group structure. It means that all the results in this section are true if $L_1, \dots, L_{n_1}$ is simply a set of smooth vector fields on $\mathbb{R}^n$ satisfying the Hormander condition.

## 4. Uniform density estimates

We are going to show several estimates for $p_r$, which depend explicitly on $r$. We sometimes write $p_r(t,x,y)$ instead of $p_r(t,z)$ if $z = (x,y)$.

We start with a uniform bound for the derivatives of $p_r$, that plays a key role in the estimates of the next section. Non-uniform version of this estimate can be obtained from Theorems IV.4.2 and IV.4.3 of [16] (see also Theorem 3 below).

**Theorem 2.** *For any non-negative integers $a, b$ and any positive number $\gamma > 1$ there exists a constant $C > 0$, such that for all $t > 0$, $x, y \in G$, $r \in (0,1)$ and multiindices $J_1 = (J_{11}, \dots, J_{1a})$, $J_2 = (J_{21}, \dots, J_{2b})$ with values $1, 2, \dots, n$ we have*

$$|L^x_{J_{11}} \dots L^x_{J_{1a}} L^y_{J_{21}} \dots L^y_{J_{2b}} p_r(t,x,y)| \leqslant Ct^{-d/2}(1-r)^{-(a+b)/2} p_r(\gamma t, x, y) \tag{11}$$

*where $d = \sum_{l=0}^{a} d(L_{J_{1l}}) + \sum_{l=0}^{b} d(L_{J_{2l}})$.*

*Proof.* We use Corollary 1 for a set of positive solutions $\psi(t,x,y) = p_r(t, z \bullet x, w \bullet y)$ of $(D^r - \frac{\partial}{\partial t})\psi = 0$, where $z, w \in G$. We choose compact $K$ as $[0,1]^{2n}$, $s = 1$ and $t = \gamma > 1$ and obtain, after combining several estimates with different multiindices and setting $x = y = 0$ in supremum and infimum,

$$|(\frac{\partial}{\partial x})^{J_1}(\frac{\partial}{\partial y})^{J_2} p_r(1, z \bullet x, w \bullet y))|_{x=y=0}| \leqslant C(1-r)^{-(a+b)/2} p_r(\gamma, z, w)$$

for all $r \in (0,1)$ and $z, w \in G$.

Let us notice that

$$R_{J_1}f(x) = (\frac{\partial}{\partial y})^{J_1} f(x \bullet y)|_{y=0}$$

is a left-invariant differential operator on $G$, which is equal to $(\frac{\partial}{\partial x})^{J_1}$ at $x = 0$. On the other hand, the product $L^x_{J_{11}} \dots L^x_{J_{1a}}$ is also left-invariant and its value at $x = 0$ can be represented as a weighted sum of the products of $\frac{\partial}{\partial x_i}, i = 1, \dots, n$ (each weight depends only on the group structure and $J_1$). Therefore this product is a weighted sum of operators $R_{\tilde{J}_1}$ with several different multiindices $\tilde{J}_1$. Note that the number of the elements in each $\tilde{J}_1$ is always less or equal $a$. Consequently we obtain the following inequality

$$|L^x_{J_{11}} \dots L^x_{J_{1a}} L^y_{J_{21}} \dots L^y_{J_{2b}} p_r(1, x, y))| \leqslant C(1-r)^{-(a+b)/2} p_r(\gamma, x, y)$$

Now recall that $p_r(t, x, y) = p(1, \beta_{t^{-1/2}}(x), \beta_{t^{-1/2}}(y))$ and

$$L^x_i f(\beta_{t^{-1/2}}(x)) = t^{-d(L_i)/2} L_i^{\beta_{t^{-1/2}}(x)} f(\beta_{t^{-1/2}}(x)).$$

After simple transformations we obtain (11). □

It is well-known (see for example [16]) that the density of a Brownian motion on a Carnot group has Gaussian-like upper and lower bounds. Instead of standart Euclidean norm those bounds contain the so-called homogeneous norm, which can be defined using the Carnot-Caratheodory distance that correspond to $L_1, \dots, L_{n_1}$. The following definition is taken from [16].

**Definition 3.** Let $C_L$ be a set of absolutely continuous paths $\varphi : [0,1] \to G$ satisfying

$$\frac{d}{dt}\varphi(t) = \sum_{i=1}^{n_1} a_i(t) L_i(\varphi(t))$$

almost everywhere on $t \in [0,1]$ w.r.t. Lebesgue measure for some measurable functions $a_i$. Then

$$\rho(x, y) = \inf_{\varphi \in C_L, \varphi(0)=x, \varphi(1)=y} \int_0^1 (\sum_{i=1}^{n_1} a_i^2(t))^{1/2} dt$$

is called a Carnot-Caratheodory distance that corresponds to $L_1, \dots, L_{n_1}$.

In our case, when all $L_i$ are left-invariant and homogeneous vector fields on the Carnot group $G$, $\rho$ is also left-invariant and homogeneous, i.e. $\rho(z \bullet x, z \bullet y) = \rho(x, y)$ and $\rho(\beta_\lambda(x), \beta_\lambda(y)) = \lambda \rho(x, y)$. We will denote $N(x) = \rho(x, 0)$, which can be called homogeneous norm on $G$, since it is homogeneous w.r.t. dilations and satisfies triangle inequality w.r.t addition on $G$ (due to the homogenuity and left-invariance of $\rho$).

Suppose there is another set of smooth vector fields $\tilde{L}_i, i = 1, \dots, n_1$, such that

$$L_i = \sum_{j=1}^{n_1} b_{ij} \tilde{L}_i$$

with smooth bounded functions $b_{ij}$, and $\tilde{\rho}$ is a Carnot-Caratheodory distance that correspond to $\tilde{L}_i$. Let $\varphi, a_i$ be as in the definition of $\rho$ and let $\tilde{\varphi}$ and $\tilde{a}_i$ be the corresponding functions in the definition of $\tilde{\rho}$. Then we can choose $\tilde{\varphi} = \varphi$ with $\tilde{a}_j(t) = \sum\limits_{i=1}^{n_1} b_{ij}(\varphi(t)) a_i(t)$

and therefore

$$(12)\quad \tilde\rho(x,y)=\inf_{\tilde\varphi\in C_{\tilde L},\tilde\varphi(0)=x,\tilde\varphi(1)=y}\int_0^1(\sum_{i=1}^{n_1}\tilde a_i^2(t))^{1/2}dt\leqslant$$

$$\leqslant\inf_{\varphi\in C_L,\varphi(0)=x,\varphi(1)=y}\int_0^1(\sum_{j=1}^{n_1}(\sum_{i=1}^{n_1}b_{ij}(\varphi(t))a_i(t))^2)^{1/2}dt\leqslant$$

$$\leqslant\inf_{\varphi\in C_L,\varphi(0)=x,\varphi(1)=y}\int_0^1\|B(\varphi(t))\|(\sum_{i=1}^{n_1}a_i^2(t))^{1/2}dt\leqslant C\rho(x,y)$$

for all $x,y\in G$ with some constant $C>0$, which depends only on the supremum of the norm of the matrix $B=\{b_{ij}|i,j=1,\ldots,n_1\}$.

Similarly we can define a Carnot-Caratheodory distance $\rho_r(x,y)$ that correspond to $V_1^r,\ldots,V_{2n_1}^r$ for $x,y\in G\times G$ and denote $N_r(x)=\rho_r(x,0)$, which is also a homogeneous norm on $G\times G$ (in the same sense as above).

**Theorem 3.** *There exist constants $C_1>0,C_2>0$, $0<\gamma_1<1<\gamma_2$ such that for all $x\in G$, $r\in(0,1)$ and $t>0$:*

$$(13)\qquad C_1\Lambda_r(\sqrt t)^{-1}e^{-\frac{N_r^2(x,y)}{2\gamma_1 t}}\leqslant p_r(t,x,y)\leqslant C_2\Lambda_r(\sqrt t)^{-1}e^{-\frac{N_r^2(x,y)}{2\gamma_2 t}}$$

*where $\Lambda_r(a)$ is a volume of $\{(x,y):N_r(x,y)<a\}$.*

*Proof.* The following result can be found in Theorems IV.4.2 and IV.4.3 of [16]. There exist constants $C_1>0,C_2>0$, $0<\gamma_1<1<\gamma_2$ such that for all $x\in G$ and $t>0$:

$$(14)\qquad C_1\Lambda(\sqrt t)^{-1}e^{-\frac{N^2(x)}{2\gamma_1 t}}\leqslant p(t,x)\leqslant C_2\Lambda(\sqrt t)^{-1}e^{-\frac{N^2(x)}{2\gamma_2 t}}$$

where $\Lambda(a)$ is a volume of $\{x:N(x)<a\}$. We notice that constants in the inequality (14) apparently depend only on the constants appearing in a Harnack inequality for operator $D$, as it follows from the proof of Theorems IV.4.2 and IV.4.3 of [16]. Therefore, provided that we have uniform Harnack inequality, we can obtain uniform version of (14).

As before, we define $u=\frac{x+y}{2}$, $v=\frac{x-y}{2\sqrt{1-r}}$ and $\tilde p_r(t,u,v)=p_r(t,x(u,v),y(u,v))$. Denote as $G_r^2$ a stratified Lie group, which is an image of $G\times G$ under transformation $(x,y)\to(u,v)$. The function $p$ appears in (14) as a heat kernel corresponding to $D$ on $G$ (in terms of [16], where it is defined as kernel of corresponding semigroup). We know that $p_r$ is a heat kernel corresponding to $D^r$ on $G\times G$ and since change of variables does not change the action of the semigroup (for example the integral of the heat kernel is still equal to 1), then $(1-r)^{n/2}2^n\tilde p_r$ is a heat kernel corresponding to $\tilde D^r$ on $G_r^2$ (the multiplier is equal to the Jacobian determinant).

Let $\tilde N_r$ and $\tilde\Lambda_r$ be the functions analogous to $N_r$ and $\Lambda_r$ after change of variables, i.e. $\tilde N_r$ is a homogeneous norm on $G_r^2$ (built using $\tilde V^r$) and $\tilde\Lambda_r(a)=\lambda(\{x:\tilde N_r(x)<a\})$. Since we have uniform Harnack inequality for $\tilde D^r$ the uniform version of (14) holds for $\tilde D^r$ on $G_r^2$. It means that we have the following inequality

$$C_1\tilde\Lambda_r(\sqrt t)^{-1}e^{-\frac{\tilde N_r^2(u,v)}{2\gamma_1 t}}\leqslant(1-r)^{n/2}\tilde p_r(t,u,v)\leqslant C_2\tilde\Lambda_r(\sqrt t)^{-1}e^{-\frac{\tilde N_r^2(u,v)}{2\gamma_2 t}}$$

with some constants $C_1>0,C_2>0$, $0<\gamma_1<1<\gamma_2$ that do not depend on $r$. We notice that, as it is easy to see from the definition of Carnot-Caratheodory distance, $N_r$ changes to $\tilde N_r$ under the transformation $(x,y)\to(u,v)$ (i.e. $\tilde N_r(u,v)=N_r(x,y)$).

Moreover we can see that

$$\tilde{\Lambda}_r(a) = \int\limits_{\tilde{N}(u,v)<a} dudv = 2^{-n}(1-r)^{-n/2} \int\limits_{N(x,y)<a} dxdy = 2^{-n}(1-r)^{-n/2}\Lambda_r(a)$$

and consequently, after going back to the variables $x, y$ in the inequality above, we obtain (13).

□

It is obvious that standard multidimensional Gaussian density multiplied by any polynomial is bounded by a constant multiplied by a density of independent Gaussian variables with a fixed variance greater than 1. We are going to show the similar fact for $p$ and $p_r$.

**Lemma 2.** *For any positive integer $M$ there are constants $C > 0$ and $\gamma > 1$, such that for all $r \in (0,1)$, $t > 0$, $x, y \in G$ and $i = 1, \ldots, n_k$ we have the following inequalities*

$$\begin{aligned} |x_i^M p(t,x)| &\leqslant C t^{Md(L_i)/2} p(\gamma t, x) \\ |x_i^M p_r(t,x,y)| &\leqslant C t^{Md(L_i)/2} p_r(\gamma t, x, y) \end{aligned} \tag{15}$$

*Proof.* We are going to use the notation and facts from the proof of Theorem 3. It is shown in [16] that $K = \{\Lambda(x) = 1\}$ is a compact. We obtain from (14) that

$$\begin{aligned}
|x_i^M p(1,x)| &\leqslant |x_i^M| C_2 \Lambda(1)^{-1} e^{-\frac{N^2(x)}{2\gamma_2}} \leqslant \\
&\leqslant \sup_{y\in G}(|y_i^M| e^{-\frac{N^2(y)}{4\gamma_2}}) C_2 \Lambda(1)^{-1} e^{-\frac{N^2(x)}{4\gamma_2}} \leqslant \\
&\leqslant \sup_{y\in K}\sup_{\lambda\in\mathbb{R}_+}(|\beta_\lambda(y)_i|^M e^{-\frac{N^2(\beta_\lambda(y))}{4\gamma_2}}) C_2 C_1^{-1} \Lambda(\sqrt{\frac{2\gamma_2}{\gamma_1}})\Lambda(1)^{-1} p(\frac{2\gamma_2}{\gamma_1}, x) = \\
&= \sup_{y\in K}\sup_{\lambda\in\mathbb{R}_+}(|y_i|^M (4\gamma_2)^{d(L_i)M/2} \lambda^{d(L_i)M} e^{-\lambda^2}) C_2 C_1^{-1} \Lambda(\sqrt{\frac{2\gamma_2}{\gamma_1}})\Lambda(1)^{-1} p(\frac{2\gamma_2}{\gamma_1}, x) \leqslant \\
&\leqslant \psi_N C p(\frac{2\gamma_2}{\gamma_1}, x)
\end{aligned}$$

where $C$ is a constant that depend only on $C_1, C_2, \gamma_1, \gamma_2, i, M$ and

$$\psi_N = \Lambda(\sqrt{\frac{2\gamma_2}{\gamma_1}})\Lambda(1)^{-1} \sup_{y\in K} |y_i|^M$$

is another constant depending on $\gamma_1, \gamma_2, i, M$ and additionally on $N$. Using homogeneuity we obtain (15) for $p$.

To prove it for $p_r$ we have to repeat the same argument using uniform bound (13) for $p_r$ (in place of (14)). From above arguments we conclude that (15) holds for $p_r$, but with additional constant

$$\psi_{N_r} = \frac{\Lambda_r(\sqrt{\frac{2\gamma_2}{\gamma_1}})}{\Lambda_r(1)} \sup_{N_r(x,y)=1} |x_i|^M$$

To finish the proof we need to show that for any $i, M$ and $a > 1$ both $\sup\limits_{N_r(x,y)=1} |x_i|^M$ and $\frac{\Lambda_r(a)}{\Lambda_r(1)}$ are bounded uniformly w.r.t $r \in (0,1)$.

We notice that $\rho_r$ is defined using operators $V_i^r$, which are linear combinations of $L_i^x, L_i^y$ with coefficients bounded uniformly w.r.t $r \in (0,1)$. For any fixed $r \in [0,1)$ the reverse transformation exists. Therefore using the definition of $\rho_r$ we can prove (as we mentioned earlier, see (12)) that $N_0(z) \leqslant CN_r(z)$ for all $z \in G$, where $C$ is a fixed constant that does not depend on $r$. It means that

$$\sup_{N_r(x,y)=1} |x_i|^M \leqslant \sup_{N_0(x,y)\leqslant C} |x_i|^M$$

is bounded uniformly w.r.t $r \in (0,1)$.

From the definition of $\Lambda_r$ and since homogeneuity holds for $N_r$, we see that

$$\Lambda_r(a) = \int\limits_{N(x,y)<a} dxdy = \int\limits_{N(\beta_{1/a}(x),\beta_{1/a}(y)))<1} dxdy = a^{2d(G)} \int\limits_{N(x,y)<1} dxdy$$

As a result, $\frac{\Lambda_r(a)}{\Lambda_r(1)} = a^{2d(G)}$ does not depend on $r$ and Lemma is proved.

□

The following inequality can be derived from (13) if we find an exact behaviour of $\Lambda_r$. But it is also possible to obtain it using uniform Harnack inequality and homogeneuity.

**Lemma 3.** *There is a constant $C > 0$, such that for all $r \in (0,1)$, $t > 0$ and $x, y \in G$:*

$$p_r(t,x,y) \leqslant Ct^{-d(G)/2}(1-r)^{-n/2} \tag{16}$$

*Proof.* We use change of variables as before and apply Theorem 1 to solutions $\psi(t,u,v) = \tilde{p}_r(t, z \bullet u, w \bullet v)$ of $(\tilde{D}^r - \frac{\partial}{\partial t})\psi = 0$.

We choose compact $K$ as $[0,1]^{2n}$, $s = 1$, $t = \gamma > 1$ and $|J_1| = |J_2| = 0$ and get

$$\sup_{(u,v)\in K} \tilde{p}_r(1, z \bullet u, w \bullet v)) \leqslant C \inf_{(u,v)\in K} \tilde{p}_r(\gamma, z \bullet u, w \bullet v) \leqslant$$
$$\leqslant C \int\limits_K \tilde{p}_r(\gamma, z \bullet u, w \bullet v)dudv \leqslant C \int\limits_{\mathbb{R}^{2n}} \tilde{p}_r(\gamma, z \bullet u, w \bullet v)dudv$$

for all $r \in (0,1)$ and $z, w \in G$. Changing variables back and setting $x = y = 0$ in the supremum we obtain

$$p_r(1,z,w) \leqslant C(1-r)^{-n/2}$$

The result follows by homogeneuity of $p_r$. □

Sometimes there is a need to estimate the supremum of $p$ w.r.t. one fixed coordinate, using the integral w.r.t. the same coordinate (it appears to be very useful in the proof of Theorem 4 below).

**Lemma 4.** *There are constants $C > 0$ and $\gamma > 1$, such that for all $t > 0$, $x \in G$ and all positive integers $l$ satisfying $d(L_l) = k$ (i.e. $l$ is one of the numbers $n_{k-1}+1, \dots, n_k$):*

$$p(t,x) \leqslant Ct^{-k/2} \int\limits_{\mathbb{R}} p(\gamma t, R(l,x,w))dw \tag{17}$$

*where $R(l,x,w)$ is a function with values in $\mathbb{R}^n$ such that $(R(l,x,w))_l = w$ and $(R(l,x,w))_i = x_i$ for all $i \neq l$.*

*Proof.* We can use Harnack inequality, cited above (Theorem 1), since $p(t, y \bullet x)$ is a solution of $(D - \frac{\partial}{\partial t})\psi = 0$ on $(0,+\infty) \times \mathbb{R}^n$ with respect to the variables $t, x$ for all $y$, by Proposition 1.

Choose $[0,1]^n$ as a compact $K$. There exist constants $C > 0$ and $\gamma > 1$ such that for all $y$

$$\sup_{x\in[0,1]^n} p(1, y \bullet x) \leqslant C \inf_{x\in[0,1]^n} p(\gamma, y \bullet x) \leqslant C \inf_{x\in[0,1]^n} \int\limits_{[0,1]} p(\gamma, y \bullet R(l,x,w))dw$$

After setting $x = 0$ under supremum and infimum we obtain for all $y$

$$p(1, y) \leqslant C \int_{y_l}^{y_l+1} p(\gamma, R(l, y, w))dw \leqslant C \int_{\mathbb{R}} p(\gamma, R(l, y, w))dw$$

since $y \bullet R(l, 0, w) = R(l, y, w + y_l)$ (note that it is not true if $d(L_l) < k$) due to the general form (4) of the group operation on a Carnot group. Lemma is proved, because, according to Proposition 1, we have homogenuity of $p$: $p(t, x) = p(1, \beta_{t^{-1/2}}(x))$ and $(\beta_{t^{-1/2}}(x))_l = t^{-k/2} x_l$. □

Note that the integral on the right hand side of (17) is also a density of a Brownian motion on a Carnot group (it is easy to see from the general form of addition on $G$ that dropping coordinate $l$ produces another Carnot group one dimension lower, if $d(L_l) = k$). Therefore (17) can be iterated.

We can also show a uniform version of (17) using Corollary 1.

**Lemma 5.** *There are constants $C > 0$ and $\gamma > 1$, such that for all $t > 0$, $r \in (0, 1)$, $x \in G$ and all positive integers $l$ satisfying $d(L_l) = k$:*

$$p_r(t, x, y) \leqslant C(1 - r)^{-1/2} t^{-k/2} \int_{\mathbb{R}} p_r(\gamma t, R(l, x, w), y)dw \tag{18}$$

*Proof.* The proof is analogous to the proof of Lemma 4, except that we have to use a uniform version of Harnack inequality given in Corollary 1. We obtain

$$\sup_{(x,y)\in K_r} p_r(1, z_1 \bullet x, z_2 \bullet y) \leqslant C \inf_{(x,y)\in K_r} p_r(1, z_1 \bullet x, z_2 \bullet y) \leqslant$$
$$\leqslant C \inf_{(x,y)\in K_r} \frac{\int\limits_{(R(l,x,w),y)\in K_r} p_r(\gamma, z_1 \bullet R(l, x, w), z_2 \bullet y)dw}{\int\limits_{(R(l,x,w),y)\in K_r} dw}$$

where $K_r = \{(x, y) : (\frac{x+y}{2}, \frac{x-y}{2\sqrt{1-r}}) \in [0, 1]^n\}$. Note that $\int\limits_{(R(l,0,w),0)\in K_r} dw = 2\sqrt{1 - r}$ and therefore we can finish the proof as in Lemma 4. □

## 5. Estimates for convolutions of derivatives on a Carnot group

In our investigation we are going to estimate the integrals of the derivatives of $p_r$. For this we need the ability to "move" the derivatives inside the integrals. We have already mentioned the possibility of integrating by parts with $L_i$. Now our next step is to find a formula that will allow us to express the action of $L_i$ on the variable $x$ of $f(y^{-1} \bullet x)$ using the same action on the variable $y$. The similar ideas can be found in [5] on p.22 and p.26.

**Lemma 6.** *For any smooth function $f : G \to \mathbb{R}$*

$$L_i^x f(y^{-1} \bullet x) = \sum_{j=1}^{n} c_{ij}(y^{-1} \bullet x) L_j^y f(y^{-1} \bullet x)$$
$$L_i^y f(y^{-1} \bullet x) = \sum_{j=1}^{n} \tilde{c}_{ij}(y^{-1} \bullet x) L_j^x f(y^{-1} \bullet x)$$

*where $c_{ij}$, and $\tilde{c}_{ij}$ are homogeneous polynomials on $G$ of homogeneous degree $d(c_{ij}) = d(\tilde{c}_{ij}) = d(L_j) - d(L_i)$, such that $\sum_{l=1}^{n} c_{il}\tilde{c}_{lj} = \delta_{ij}$ i.e. matrix $\tilde{c}_{ij}$ is an inverse of $c_{ij}$.*

*Proof.* We recall that

$$L_i f(x) = \frac{\partial}{\partial y_i} f(x \bullet y)|_{y=0}$$

Denote as $\{R_i, i = 1, \dots, n\}$ a set of right-invariant differential operators on $G$ such that

$$R_i f(x) = \frac{\partial}{\partial y_i} f(y^{-1} \bullet x)|_{y=0}$$

Clearly $R_i$, as well as $L_i$, forms basis in each tangent space (for $R_i$ it follows from the fact that $R_i f(x) = L_i g(x^{-1})$, where $g(x) = f(x^{-1})$). Therefore we can express $R_i$ using $L_i$ and vice versa as follows:

$$L_i = \sum_{j=1}^{n} c_{ij} R_j$$

$$R_i = \sum_{j=1}^{n} \tilde{c}_{ij} L_j$$

Coefficients in such representation are determined uniquely as smooth functions on $G$. Uniqueness means that applying dilations can not change coefficients, and therefore both $c_{ij}$ and $\tilde{c}_{ij}$ are homogeneous functions and hence homogeneous polynomials of homogeneous degree $d(L_i) - d(L_j)$ (both $L_i$ and $R_i$ are homogeneous of order $d(L_i)$ as can be easily seen from the definition). We also note that obviously $\tilde{c}_{ij}$ is an inverse matrix of $c_{ij}$.

Now we can finish the proof:

$$L_i^x f(y^{-1} \bullet x) = (L_i f)(y^{-1} \bullet x) = \sum_{j=1}^{k} c_{ij}(y^{-1} \bullet x)(R_j f)(y^{-1} \bullet x) =$$

$$= \sum_{j=1}^{k} c_{ij}(y^{-1} \bullet x) \frac{\partial}{\partial u_i} f(u^{-1} \bullet y^{-1} \bullet x)|_{u=0} = \sum_{j=1}^{k} c_{ij}(y^{-1} \bullet x) \frac{\partial}{\partial u_i} f((y \bullet u)^{-1} \bullet x)|_{u=0} =$$

$$= \sum_{j=1}^{k} c_{ij}(y^{-1} \bullet x) L_i^y f(y^{-1} \bullet x)$$

The second formula can be proved in the same way (or we can recall that $\tilde{c}_{ij}$ is an inverse matrix of $c_{ij}$). $\square$

Using Lemma 6 we are able to show some additional properties of $p_r$.

**Lemma 7.** **1.** *For all $t > 0$, $x, y \in G$ and $0 \leqslant r < 1$ the function $p_r(t, x, y)$ is jointly continuous w.r.t. $(r, t, x, y)$.*

**2.** *For all $t > 0$, $x, y \in G$ and $0 \leqslant r < 1$ the function $p_r(t, x, y)$ is continuously differentiable w.r.t. $r$ and*

$$\frac{d}{dr} p_r(t, x, y) = \\ = \int_0^t \int_{\mathbb{R}^{2n}} p_r(s, (z_1)^{-1} \bullet x, (z_2)^{-1} \bullet y) \sum_{i=1}^{n_1} L_i^{z_1} L_i^{z_2} p_r(t - s, z_1, z_2) dz_1 dz_2 ds \tag{19}$$

**3.** *There exist $C > 0$ and $\gamma > 1$ such that for all $t > 0$, $r \in (0, 1)$ and $x, y \in G$*

$$|\frac{d}{dr} p_r(t, x, y)| \leqslant C(1 - r)^{-1} p_r(\gamma t, x, y) \tag{20}$$

*Proof.* Fix $0 \leqslant r < r + \delta < 1$. Using Ito formula and taking mathematical expectation we obtain for every smooth bounded function $f$

$$Ef(t, Y_{r,t}) = f(0, Y_{r,0}) + E \int_0^t (D^r + \frac{\partial}{\partial s}) f(s, Y_{r,s}) ds$$

Suppose that $g(t, x) = (p_{r+\delta}(t, \cdot) *_{G \times G} h)(x)$ for some continuous function $h$ with compact support. We know that $g$ is a solution of the following Cauchy problem: $(D^{r+\delta} - \frac{\partial}{\partial t})g = 0$, $g(0, x) = h(x)$ (see Proposition 2), but also that $g$ is a smooth bounded function on $(0, +\infty) \times \mathbb{R}^{2n}$. Its smoothness is a consequence of Hormander theorem and boundedness follow from boundedness of $h$. Moreover the result of the action of any number of $L_i^{x_1}, L_i^{x_2}$ (where $x = (x_1, x_2), x_i \in G$) on $g(t, x)$ is also a bounded function and can be represented as the integral of the action of the same operators on $p_r$, for example

$$L_i^{x_1} g = ((L_i^{x_1} p_{r+\delta})(t, \cdot) *_{G \times G} h)$$

It follows from the left-invariance of $L_i$ and estimates from Theorem 2. Note that we can freely exchange the integral on $G \times G$ with any derivatives w.r.t. $x_1$, $x_2$ or $t$, since the function $p_r$ is smooth and $h$ has a compact support.

Now we can put $f(s, x) = g(t - s, x)$ for $s \in [0, t]$. Setting $Y_{r,0} = x$ and rewriting mathematical expectation of $Y_{r,s}$ using its density $p_r(s, y^{-1} \bullet x)$ we obtain:

$$\int_{\mathbb{R}^{2n}} p_r(t, y^{-1} \bullet x) h(y) dy = g(t, x) +$$
$$+ \int_0^t \int_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x)(D^r - \frac{\partial}{\partial t}) g(t - s, y) dy ds$$

Note that $(D^r - \frac{\partial}{\partial t})g = (D^r - D^{r+\delta})g$. After calculating the difference $D^{r+\delta} - D^r$ and replacing $g$ with the integral of $p_{r+\delta}$ we get

$$\int_{\mathbb{R}^{2n}} (p_{r+\delta} - p_r)(t, y^{-1} \bullet x) h(y) dy =$$
$$= \delta \int_0^t \int_{\mathbb{R}^{4n}} p_r(s, y^{-1} \bullet x) \sum_{i=1}^{n_1} L_i^{y_1} L_i^{y_2} p_{r+\delta}(t - s, z^{-1} \bullet y) h(z) dz dy ds$$

In order to proceed we need to swap the integrals w.r.t. $z$ and $s$ on the right hand side (the integrals w.r.t. $z$ and $y$ can be swapped freely due to the estimates from Theorem 2), and this would be possible, if we show that the function under the integral is absolutely integrable, i.e. that

$$\int_0^t \int_{\mathbb{R}^{2n}} \left| \int_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x) \sum_{i=1}^{n_1} L_i^{y_1} L_i^{y_2} p_{r+\delta}(t - s, z^{-1} \bullet y) dy \right| dz ds < +\infty \tag{21}$$

If we try to estimate the integral w.r.t. $y, z$ directly using Theorem 2 we obtain

$$\int\limits_{\mathbb{R}^{4n}} |p_r(s, y^{-1} \bullet x) \sum_{i=1}^{n_1} L_i^{y_1} L_i^{y_2} p_{r+\delta}(t-s, z^{-1} \bullet y)| dz dy \leqslant$$
$$\leqslant C(1-(r+\delta))^{-1}(t-s)^{-1} \int\limits_{\mathbb{R}^{4n}} p_r(s, y^{-1} \bullet x) p_{r+\delta}(\gamma(t-s), z^{-1} \bullet y) dy dz =$$
$$= C(1-(r+\delta))^{-1}(t-s)^{-1} \int\limits_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x) dy =$$
$$= C(1-(r+\delta))^{-1}(t-s)^{-1}$$

Unfortunately this estimate is not integrable w.r.t. $s$ and (21) does not follow. This is why we need Lemma 6 which we apply now to "move" $L_i$ on $p_r$:

$$\int\limits_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x) \sum_{i=1}^{n_1} L_i^{y_1} L_i^{y_2} p_{r+\delta}(t-s, z^{-1} \bullet y) dy =$$
$$= \int\limits_{\mathbb{R}^{2n}} \sum_{i=1}^{n_1} (\sum_{j=1}^{n} \tilde{c}_{ij}(y_1^{-1} \bullet x_1) L_j^{x_1}) (\sum_{j=1}^{n} \tilde{c}_{ij}(y_2^{-1} \bullet x_2) L_j^{x_2}) p_r(s, y^{-1} \bullet x) \cdot$$
$$\cdot p_{r+\delta}(t-s, z^{-1} \bullet y) dy$$

where we used the fact that $\int\limits_{\mathbb{R}^{2n}} L_i f(x) dx = 0$ for any smooth $f$ if $L_i f$ is integrable. Now application of Theorem 2 together with Lemma 2 gives us the following estimate

$$\int\limits_{\mathbb{R}^{2n}} \left| \int\limits_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x) \sum_{i=1}^{n_1} L_i^{y_1} L_i^{y_2} p_{r+\delta}(t-s, z^{-1} \bullet y) dy \right| dz \leqslant C(1-r)^{-1} s^{-1}$$

To see that note that multiplier $t^{(d(L_j)-d(L_i))/2}$ that comes from estimating $\tilde{c}_{ij} p_r$ using Lemma 2 compensates multiplier $t^{-d(L_j)/2}$ that appears from estimating $L_j p_r$ using Theorem 2.

Since the function $\min(s^{-1}, (t-s)^{-1}) \leqslant 2t^{-1}$ is clearly integrable w.r.t. $s$ we obtain (21). Now we can swap the integrals w.r.t. $z$ and $s$ and since $h$ is any continuous function with compact support we can drop the integral by $h(z)dz$ to get

$$(p_{r+\delta} - p_r)(t, z^{-1} \bullet x) = \delta \int\limits_0^t \int\limits_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x) \sum_{i=1}^{n_1} L_i^{y_1} L_i^{y_2} p_{r+\delta}(t-s, z^{-1} \bullet y) dy ds$$

where $y = (y_1, y_2), y_i \in G$.

We can find an upper bound for the function on the right hand side, using the same ideas as above, and additionally Lemma 3 (the difference is that there is no integral w.r.t.

$z$):

$$\int_{\mathbb{R}^{2n}} |p_r(s, y^{-1} \bullet x) \sum_{i=1}^{n_1} L_i^{y_1} L_i^{y_2} p_{r+\delta}(t-s, z^{-1} \bullet y)| dy \leqslant$$
$$\leqslant C(1-(r+\delta))^{-1}(t-s)^{-1} \int_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x) p_{r+\delta}(\gamma(t-s), z^{-1} \bullet y) dy \leqslant$$
$$\leqslant \tilde{C}(1-(r+\delta))^{-n/2-1}(t-s)^{-d(G)/2-1} \int_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x) dy =$$
$$= \tilde{C}(1-(r+\delta))^{-n/2-1}(t-s)^{-d(G)/2-1}$$

Moving $L_i$ onto $p_r(s, y^{-1} \bullet x)$ as before we also obtain

$$\left| \int_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x) \sum_{i=1}^{n_1} L_i^{y_1} L_i^{y_2} p_{r+\delta}(t-s, z^{-1} \bullet y) dy \right| \leqslant C(1-r)^{-n/2-1} s^{-d(G)/2-1}$$

Joining estimates together we can see that

$$\int_0^t \left| \int_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x) \sum_{i=1}^{n_1} L_i^{y_1} L_i^{y_2} p_{r+\delta}(t-s, z^{-1} \bullet y) dy \right| ds \leqslant$$
$$\leqslant C(1-r)^{-n/2-1} t^{-d(G)/2-1}$$

i.e. $\delta^{-1}(p_{r+\delta} - p_r)(t, z^{-1} \bullet x)$ is bounded uniformly w.r.t. all variables away from $t = 0$ and $r = 1$.

Therefore for any $r \in (0, 1)$ and $t > 0$ the function $p_r$ is continuous in $r$ uniformly w.r.t. other variables. It means that $p_r(t, x, y)$ is jointly continuous for all $t > 0$, $x, y \in G$ and $0 \leqslant r < 1$ w.r.t. $(r, t, x, y)$, since we already know that it is jointly continuous w.r.t. $(t, x, y)$.

Now we can apply the same bounds again and use theorem of bounded convergence to see that $\delta^{-1}(p_{r+\delta} - p_r)(t, z^{-1} \bullet x)$ converges to the right hand side of (19) as $\delta \to 0+$. It means that $p_r$ is differentiable w.r.t. $r$ and we obtain (19). We can estimate the right hand side of (19) using Lemma 6, Theorem 2 and Lemma 2 again:

$$\int_{\mathbb{R}^{2n}} |p_r(s, y^{-1} \bullet x) \sum_{i=1}^{n_1} L_i^{y_1} L_i^{y_2} p_r(t-s, z^{-1} \bullet y)| dy \leqslant$$
$$\leqslant C(1-r)^{-1}(t-s)^{-1} \int_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x) p_r(\gamma(t-s), z^{-1} \bullet y) dy =$$
$$= C(1-r)^{-1}(t-s)^{-1} p_r(s+\gamma(t-s), z^{-1} \bullet x) \leqslant \tilde{C}(1-r)^{-1}(t-s)^{-1} p_r(\gamma t, z^{-1} \bullet x)$$

$$\left| \int_{\mathbb{R}^{2n}} p_r(s, y^{-1} \bullet x) \sum_{i=1}^{n_1} L_i^{y_1} L_i^{y_2} p_r(t-s, z^{-1} \bullet y) dy \right| \leqslant$$
$$\leqslant C(1-r)^{-1} s^{-1} p_r(\gamma t, z^{-1} \bullet x)$$

Combining these two inequalities we get (20).

□

## 6. Renormalized local time for the Levy area of the increments of two-dimensional Brownian motion

**Theorem 4.** *Define*

$$\gamma_\varepsilon = \int_0^1 \int_0^t f_\varepsilon(B_{s,t}) ds dt \tag{22}$$

*where*

$$f_\varepsilon(x) = (2\pi\varepsilon)^{-1/2} e^{-\frac{|x|^2}{2\varepsilon}} \tag{23}$$

*The family $\gamma_\varepsilon$ is unbounded in $L_2(\Omega)$, but there exists a limit of $\gamma_\varepsilon - E\gamma_\varepsilon$ in $L_2(\Omega)$ as $\varepsilon \to 0+$.*

The main idea of the proof is that Levy area can be described as a coordinate of a Brownian motion $X$ on a Carnot group $G$ (see (8)). Therefore we can use the density of $X$ and $Y$ (i.e. $p$ and $p_r$ for the case described in (8)) to study expectations of $\gamma_\varepsilon$.

We are going to investigate the convergence of $\gamma_\varepsilon - E\gamma_\varepsilon$ by studying

$$E(\gamma_{\varepsilon_1} - E\gamma_{\varepsilon_1})(\gamma_{\varepsilon_2} - E\gamma_{\varepsilon_2}) = E\gamma_{\varepsilon_1}\gamma_{\varepsilon_2} - E\gamma_{\varepsilon_1}E\gamma_{\varepsilon_2}$$

This comes down to deriving a suitable representation for

$$Ef(B_{s_1,t_1}(W))g(B_{s_2,t_2}(W)) - Ef(B_{s_1,t_1}(W))Eg(B_{s_2,t_2}(W))$$

**Lemma 8.** *There exists functions $q$ and $q_r$ for $r \in [0,1]$, continuous w.r.t. $(s_1, t_1, s_2, t_2, x, y)$ in all points where $s_1, t_1, s_2, t_2$ are pairwise distinct, satisfying*

$$\begin{aligned} &Ef(B_{s_1,t_1}(W))g(B_{s_2,t_2}(W)) - Ef(B_{s_1,t_1}(W))Eg(B_{s_2,t_2}(W)) = \\ &\qquad = \int_{\mathbb{R}} \int_{\mathbb{R}} q(s_1, t_1, s_2, t_2, x, y) f(x) g(y) dx dy \end{aligned} \tag{24}$$

*and*

$$Ef(B_{s_1,t_1}(W_{1,r}))g(B_{s_2,t_2}(W_{2,r})) = \int_{\mathbb{R}} \int_{\mathbb{R}} q_r(s_1, t_1, s_2, t_2, x, y) f(x) g(y) dx dy \tag{25}$$

*We have*

$$q = q_1 - q_0 = \int_0^1 \frac{d}{dr} q_r dr \tag{26}$$

*Suppose that we have $p$ and $p_r$ for the case described in* (8)*. The following representations holds for $q$ and $q_r$:*

**1.** *If $s_1 < t_1 < s_2 < t_2$, then $q = 0$.*

**2.** *If $s_1 < s_2 < t_1 < t_2$, then*

$$\begin{aligned} q_1(s_1, t_1, s_2, t_2, x, y) &= \int_{\mathbb{R}^7} p(s_2 - s_1, z \bullet (u_1, u_2, x)^{-1}) \cdot \\ &\quad \cdot p(t_1 - s_2, z) p(t_2 - t_1, (v_1, v_2, y)^{-1} \bullet z) dz du_1 du_2 dv_1 dv_2 \\ q_0(s_1, t_1, s_2, t_2, x, y) &= \int_{\mathbb{R}^4} p(t_1 - s_1, (u_1, u_2, x)^{-1}) \cdot \\ &\quad \cdot p(t_2 - s_2, (v_1, v_2, y)^{-1}) du_1 du_2 dv_1 dv_2 \end{aligned} \tag{27}$$

*There is a constant $C > 0$, such that for all $s_1 < s_2 < t_1 < t_2$ and $x, y \in G$*

$$q_1(s_1, t_1, s_2, t_2, x, y) \leqslant C \frac{\min(s_2 - s_1, t_2 - t_1, t_1 - s_2)}{(s_2 - s_1)(t_2 - t_1)(t_1 - s_2)} \tag{28}$$
$$q_0(s_1, t_1, s_2, t_2, x, y) \leqslant C(t_1 - s_1)^{-1}(t_2 - s_2)^{-1}$$

**3.** *If $s_2 < s_1 < t_1 < t_2$, then for $r \in (0, 1)$*

$$q_r(s_1, t_1, s_2, t_2, x, y) = \tag{29}$$
$$= \int_{\mathbb{R}^{10}} p(s_1 - s_2, u^{-1}) p_r(t_1 - s_1, (z_1, z_2, y)^{-1}, w^{-1}) \cdot$$
$$\cdot\, p(t_2 - t_1, (v_1, v_2, x)^{-1} \bullet u \bullet w) du dw dv_1 dv_2 dz_1 dz_2$$

*The derivative $\frac{d}{dr} q_r$ exists and*

$$\frac{d}{dr} q_r(s_1, t_1, s_2, t_2, x, y) = \tag{30}$$
$$= \int_{\mathbb{R}^{10}} p(s_1 - s_2, u^{-1}) \frac{d}{dr} p_r(t_1 - s_1, (z_1, z_2, y)^{-1}, w^{-1}) \cdot$$
$$\cdot\, p(t_2 - t_1, (v_1, v_2, x)^{-1} \bullet u \bullet w) du dw dv_1 dv_2 dz_1 dz_2$$

*There is a constant $C > 0$, such that for all $s_2 < s_1 < t_1 < t_2$, $r \in (0, 1)$ and $x, y \in G$ we have*

$$q_r(s_1, t_1, s_2, t_2, x, y) \leqslant C(t_1 - s_1)^{-1}(t_2 - t_1 + \sqrt{1 - r}(t_1 - s_1) + s_1 - s_2)^{-1} \tag{31}$$

*Proof.* Note that both

$$X_t(W) = (W_t^1, W_t^2, B_{0,t}(W))$$

and

$$Y_t = (X_t(W_{1,r}), X_t(W_{2,r}))$$

has independent equally distributed increments w.r.t group addition on $G$ and $G \times G$ respectively, so we can represent joint density of their values at several points in time, using only $p$ and $p_r$. More precisely we have that $X_{s,t} = (X_s)^{-1} \bullet X_t$ is independent from all $X_u, u \leqslant s$ and has the same distribution as $X_{t-s}$ started at $X_0 = 0$. This is well-known and also follows from the properties of $p$ stated in Proposition 1. The same is true for the process $Y$, because $p_r$ has similar properties due to Proposition 2. It is easy to check that the third coordinate of $X_{s,t}(W)$ is $B_{s,t}(W)$ (it was defined to be exactly that) using the explicit form of $\bullet$ from (8).

Formula (26) follows from the definition of $q$ and $q_r$, because for $r = 0$ the Brownian motions $W_{1,r}$ and $W_{2,r}$ are independent and for $r = 1$ they are equal. The second part (26) is true if the derivative $\frac{d}{dr} q_r$ exists for all $r \in (0, 1)$ (and $q_r$ is continuous on $r \in [0, 1]$). We will only show it for one of the cases below, since in two other cases $q_r$ is not needed.

If $s_1 < t_1 < s_2 < t_2$ then $X_{s_1,t_1}(W)$ and $X_{s_2,t_2}(W)$ are independent and therefore $B_{s_1,t_1}(W)$ and $B_{s_2,t_2}(W)$ are independent. It follows that

$$Ef(B_{s_1,t_1}(W)) g(B_{s_2,t_2}(W)) = Ef(B_{s_1,t_1}(W)) Eg(B_{s_2,t_2}(W))$$

i.e. $q = 0$.

Suppose that $s_1 < s_2 < t_1 < t_2$. In this case increments $X_{s_1,s_2}(W)$, $X_{s_2,t_1}(W)$ and $X_{t_1,t_2}(W)$ can be used to represent

$$B_{s_1,t_1}(W) = (X_{s_1,s_2}(W) \bullet X_{s_2,t_1}(W))_3$$

and

$$B_{s_2,t_2}(W) = (X_{s_2,t_1}(W) \bullet X_{t_1,t_2}(W))_3$$

Using densities of these independent increments to calculate mathematical expectation and changing variables in the integral to separate the integral by $f(x)dx$ and $g(y)dy$ we obtain (27):

$$Ef(B_{s_1,t_1}(W))g(B_{s_2,t_2}(W)) =$$
$$= \int\limits_{\mathbb{R}^9} p(s_2 - s_1, u^{-1})p(t_1 - s_2, z^{-1})p(s_1 - t_1, v^{-1})f((u \bullet z)_3)g((z \bullet v)_3)dzdudv =$$
$$= \int\limits_{\mathbb{R}^9} p(s_2 - s_1, z \bullet u^{-1})p(t_1 - s_2, z^{-1})p(s_1 - t_1, v^{-1} \bullet z)f(u_3)g(v_3)dzdudv$$

The representation of $q_0$ is shown similarly.

To show (28) we notice that $q_0$ can be estimated using Lemma 4 to find supremum w.r.t. $x$ and $y$:

$$q_0 \leqslant C(t_1 - s_1)^{-1}(t_2 - s_2)^{-1} \int\limits_{\mathbb{R}^6} p(\gamma(t_1 - s_1), (u_1, u_2, u_3)^{-1})$$
$$p(\gamma(t_2 - s_2), (v_1, v_2, v_3)^{-1})du_1du_2du_3dv_1dv_2dv_3 =$$
$$= C(t_1 - s_1)^{-1}(t_2 - s_2)^{-1}$$

Then we do the same for $q_1$ and obtain

$$q_1 \leqslant C(s_2 - s_1)^{-1}(t_2 - t_1)^{-1} \int\limits_{\mathbb{R}^9} p(s_2 - s_1, z \bullet (u_1, u_2, u_3)^{-1})\cdot$$
$$\cdot p(t_1 - s_2, z^{-1})p(t_2 - t_1, (v_1, v_2, v_3)^{-1} \bullet z)dzdu_1du_2du_3dv_1dv_2dv_3 =$$
$$= C(s_2 - s_1)^{-1}(t_2 - t_1)^{-1}$$

But this is not enough, so we use change of variables $z \to z \bullet (u_1, u_2, x)$ and then apply Lemma 4 again:

$$q_1 = \int\limits_{\mathbb{R}^7} p(s_2 - s_1, z)p(t_1 - s_2, (u_1, u_2, x)^{-1} \bullet z^{-1})\cdot$$
$$\cdot p(t_2 - t_1, (v_1, v_2, y)^{-1} \bullet z \bullet (u_1, u_2, x))dzdu_1du_2dv_1dv_2 \leqslant$$
$$\leqslant C(t_2 - t_1)^{-1} \int\limits_{\mathbb{R}^8} p(s_2 - s_1, z)p(t_1 - s_2, (u_1, u_2, x)^{-1} \bullet z^{-1})\cdot$$
$$\cdot p(t_2 - t_1, (v_1, v_2, v_3)^{-1} \bullet z \bullet (u_1, u_2, x))dzdu_1du_2dv_1dv_2dv_3 =$$
$$= C(t_2 - t_1)^{-1} \int\limits_{\mathbb{R}^5} p(s_2 - s_1, z)p(t_1 - s_2, (u_1, u_2, x)^{-1} \bullet z^{-1})dzdu_1du_2 \leqslant$$
$$\leqslant \tilde{C}(t_1 - s_2)^{-1}(t_2 - t_1)^{-1}$$

Similarly with the change of variables $z \to (v_1, v_2, y) \bullet z$

$$q_1 = \int\limits_{\mathbb{R}^7} p(s_2 - s_1, (v_1, v_2, y) \bullet z \bullet (u_1, u_2, x)^{-1})p(t_1 - s_2, z^{-1} \bullet (v_1, v_2, y)^{-1})$$
$$p(t_2 - t_1, z)dzdu_1du_2dv_1dv_2 \leqslant C(s_2 - s_1)^{-1}(t_1 - s_2)^{-1}$$

Joining all estimates together gives us (28). We notice that continuity of each $q_1$ and $q_0$ follows from these estimates and continuity of $p$.

Finally let us assume that $s_2 < s_1 < t_1 < t_2$. To show (29) we note that increments of $Y$ on $(s_2, s_1)$, $(s_1, t_1)$ and $(t_1, t_2)$ determine $B_{s_1,t_1}(W_{1,r})$ and $B_{s_2,t_2}(W_{2,r})$:

$$B_{s_1,t_1}(W_{1,r}) = (X_{s_1,t_1}(W_{1,r}))_3$$

$$B_{s_2,t_2}(W_{2,r}) = (X_{s_2,s_1}(W_{2,r}) \bullet X_{s_1,t_1}(W_{2,r}) \bullet X_{t_1,t_2}(W_{2,r}))_3$$

Therefore we can write the expectation as an integral of densities $p_r(s_1 - s_2, \cdot)$, $p_r(t_1 - s_1, \cdot)$, $p_r(t_2 - t_1, \cdot)$. We can see that $X_{s_2,s_1}(W_{1,r})$ and $X_{t_1,t_2}(W_{1,r})$ are not needed to represent $B_{s_1,t_1}(W_{1,r})$ and $B_{s_2,t_2}(W_{2,r})$ and we can integrate the first and the last $p_r$ using

$$\int\limits_G p_r(t, x^{-1}, y^{-1}) dx = p(t, y^{-1})$$

which is true since $p(t, y^{-1})$ is a density of $X_t(W)$ (if $X_0 = 0$) and $p_r(t, x^{-1}, y^{-1})$ is a density of $Y_t(W) = (X_t(W_{1,r}), X_t(W_{2,r}))$. We obtain:

$$\begin{aligned}
&Ef(B_{s_1,t_1}(W_{1,r}))g(B_{s_2,t_2}(W_{2,r})) = \\
&\qquad = \int\limits_{\mathbb{R}^{18}} p_r(s_1 - s_2, x^{-1}, u^{-1}) p_r(t_1 - s_1, z^{-1}, w^{-1}) \cdot \\
&\qquad \cdot p_r(t_2 - t_1, y^{-1}, v^{-1}) f(z_3) g((u \bullet w \bullet v)_3) dx dy dz du dv dw = \\
&\qquad = \int\limits_{\mathbb{R}^{12}} p(s_1 - s_2, u^{-1}) p_r(t_1 - s_1, z^{-1}, w^{-1}) \cdot \\
&\qquad \cdot p(t_2 - t_1, v^{-1}) f(z_3) g((u \bullet w \bullet v)_3) dz du dv dw = \\
&\qquad = \int\limits_{\mathbb{R}^{12}} p(s_1 - s_2, u^{-1}) p_r(t_1 - s_1, z^{-1}, w^{-1}) \cdot \\
&\qquad\qquad \cdot p(t_2 - t_1, v^{-1} \bullet u \bullet w) f(z_3) g(v_3) dz du dv dw
\end{aligned}$$

Separating the integrals by $f(x)dx$ and $g(y)dy$ leads to (29).

Now we have to prove that the derivative $\frac{d}{dr}$ of $q_r$ exists and can be swapped with all integrals. For this we need to make sure that $q_r$ is always finite (where defined, i.e. for $s_2 < s_1 < t_1 < t_2$), so we have to show (31) first. But then the rest (also including continuity of $q$) follows from Lemma 7 since estimate (20) provides the possibility of passing to the limit under the integral.

To prove (31) we use Lemma 4 once again

$$\begin{aligned}
q_r &\leqslant C(t_2 - t_1)^{-1} \int\limits_{\mathbb{R}^{11}} p(s_1 - s_2, u^{-1}) p_r(t_1 - s_1, (z_1, z_2, y)^{-1}, w^{-1}) \cdot \\
&\qquad \cdot p(t_2 - t_1, (v_1, v_2, v_3)^{-1} \bullet u \bullet w) du dw dv_1 dv_2 dv_3 dz_1 dz_2 = \\
&= C(t_2 - t_1)^{-1} \int\limits_{\mathbb{R}^8} p(s_1 - s_2, u^{-1}) p_r(t_1 - s_1, (z_1, z_2, y)^{-1}, w^{-1}) du dw dz_1 dz_2 = \\
&= C(t_2 - t_1)^{-1} \int\limits_{\mathbb{R}^3} p(t_1 - s_1, (z_1, z_2, y)^{-1}) dz_1 dz_2 \leqslant \\
&\qquad \leqslant \tilde{C}(t_2 - t_1)^{-1}(t_1 - s_1)^{-1}
\end{aligned}$$

If we change variables $u \to (v_1, v_2, x) \bullet u \bullet w^{-1}$ we can show that

$$q_r = \int\limits_{\mathbb{R}^{10}} p(s_1 - s_2, w \bullet u^{-1} \bullet (v_1, v_2, x)^{-1}) p_r(t_1 - s_1, (z_1, z_2, y)^{-1}, w^{-1}) \cdot$$
$$\cdot p(t_2 - t_1, u) du dw dv_1 dv_2 dz_1 dz_2 \leqslant$$
$$\leqslant C(s_1 - s_2)^{-1} \int\limits_{\mathbb{R}^{11}} p(s_1 - s_2, w \bullet u^{-1} \bullet (v_1, v_2, v_3)^{-1}) p_r(t_1 - s_1, (z_1, z_2, y)^{-1}, w^{-1}) \cdot$$
$$\cdot p(t_2 - t_1, u) du dw dv_1 dv_2 dv_3 dz_1 dz_2 =$$
$$= C(s_1 - s_2)^{-1} \int\limits_{\mathbb{R}^2} p(t_1 - s_1, (z_1, z_2, y)^{-1}) dz_1 dz_2 \leqslant \tilde{C}(s_1 - s_2)^{-1}(t_1 - s_1)^{-1}$$

One more version of this estimate can be obtained if we change variables (in the formula (29)) in the following way $w \to u^{-1} \bullet (v_1, v_2, x) \bullet w$ and use uniform bound (18). We get

$$q_r = \int\limits_{\mathbb{R}^{10}} p(s_1 - s_2, u^{-1}) p_r(t_1 - s_1, (z_1, z_2, y)^{-1}, w^{-1} \bullet (v_1, v_2, x)^{-1} \bullet u) \cdot$$
$$\cdot p(t_2 - t_1, w) du dw dv_1 dv_2 dz_1 dz_2 \leqslant$$
$$\leqslant C(t_1 - s_1)^{-1}(1 - r)^{-1/2} \int\limits_{\mathbb{R}^{10}} p(s_1 - s_2, u^{-1}) p_r(t_1 - s_1, (z_1, z_2, y)^{-1}, w^{-1} \bullet (v_1, v_2, v_3)^{-1} \bullet u) \cdot$$
$$\cdot p(t_2 - t_1, w) du dw dv_1 dv_2 dv_3 dz_1 dz_2 =$$
$$= C(t_1 - s_1)^{-1}(1 - r)^{-1/2} \int\limits_{\mathbb{R}^2} p(t_1 - s_1, (z_1, z_2, y)^{-1}) dz_1 dz_2 \leqslant \tilde{C}(t_1 - s_1)^{-2}(1 - r)^{-1/2}$$

As a result we obtain

$$q_r \leqslant C(\max(s_1 - s_2, t_2 - t_1, \sqrt{1 - r}(t_1 - s_1)))^{-1}(t_1 - s_1)^{-1} \leqslant$$
$$\leqslant \tilde{C}(t_2 - t_1 + \sqrt{1 - r}(t_1 - s_1) + s_1 - s_2)^{-1}(t_1 - s_1)^{-1}$$

and Lemma is proved. □

*Proof of Theorem 4.* First we will show that $\gamma_\varepsilon$ is unbounded in $L_2$. It is enough to prove that $E\gamma_\varepsilon \to +\infty$ as $\varepsilon \to 0+$. We have, by the definition of $p$, that

$$E\gamma_\varepsilon = \int\limits_0^1 \int\limits_0^t \int\limits_{\mathbb{R}^3} f_\varepsilon(x_3) p(t - s, x) dx ds dt$$

We introduce a change of variable $x \to \beta_{t-s}(x)$ and obtain, using homogeneous properties of $p$, that

$$E\gamma_\varepsilon = \int\limits_0^1 \int\limits_0^t \int\limits_{\mathbb{R}^3} f_\varepsilon((t - s)x_3) p(1, x) dx ds dt =$$
$$= \int\limits_0^1 \int\limits_0^t \int\limits_G (t - s)^{-1} f_{(t-s)^{-2}\varepsilon}(x_3) p(1, x) dx ds dt$$

We know that $f_\varepsilon$ converges to $\delta_0$ and since $p$ is smooth we can write

$$\lim_{\varepsilon\to 0+}\int_{\mathbb{R}} f_\varepsilon(x_3)p(1,(x_1,x_2,x_3))dx_3 = p(1,(x_1,x_2,0))$$

Applying Fatou lemma we get

$$\begin{aligned}\varliminf_{\varepsilon\to 0+} E\gamma_\varepsilon &\geqslant \\ &\geqslant \int_0^1\int_0^t\int_{\mathbb{R}^2} \lim_{\varepsilon\to 0+}\Big(\int_{\mathbb{R}}(t-s)^{-1}f_{(t-s)^{-2}\varepsilon}(x_3)p(1,x)dx_3\Big)dx_1dx_2dsdt = \\ &= \int_0^1\int_0^t (t-s)^{-1}dsdt\int_{\mathbb{R}^2} p(1,(x_1,x_2,0))dx_1dx_2 = +\infty\end{aligned}$$

Now we are going to prove the main part of the Theorem 4. To do that it is enough to show, that $E\gamma_{\varepsilon_1}\gamma_{\varepsilon_2} - E\gamma_{\varepsilon_1}E\gamma_{\varepsilon_2}$ converges as $(\varepsilon_1,\varepsilon_2)\to 0+$. This statement in turn can be proved using properties of the kernel $K(x,y)$, $x,y\in G$, defined as follows:

$$E\gamma_{\varepsilon_1}\gamma_{\varepsilon_2} - E\gamma_{\varepsilon_1}E\gamma_{\varepsilon_2} = \int_{G^2} f_{\varepsilon_1}(x)f_{\varepsilon_2}(y)K(x,y)dxdy$$

It is enough to show that $K$ is bounded and continuous function of $(x,y)$.

From the definition of $q$ (see formula (24)) we see that

$$K(x,y) = \int_0^1\int_0^{t_1}\int_0^1\int_0^{t_2} q(s_1,t_1,s_2,t_2,x,y)ds_2dt_2ds_1dt_1$$

as long as we can change order of the integrals w.r.t. $x,y$ and the integrals w.r.t. $t_1,t_2,s_1,s_2$. And the latter is true if $q$ is bounded by the integrable function of $s_1,t_1,s_2,t_2$. It is easy to see that both boundedness and continuity of $K$ also follows (continuity can be obtained using continuity of $q$ shown in Lemma 8).

So the proof is now reduced to finding an integrable estimate of $q$. We split the domain of integration in the integral into six subdomains depending on the order of $s_1,s_2,t_1,t_2$ (we can ignore sets of zero Lebesgue measure). It is enough to consider only three subdomains where we have $t_2>t_1$ since the rest can be treated similarly due to symmetry. Using Lemma 8 we can write a representation for $q$ in each case.

If $\{0<s_1<t_1<s_2<t_2<1\}$ then $q=0$, so there is nothing to prove here. Suppose that $\{0<s_1<s_2<t_1<t_2<1\}$. We recall the inequality (28) and notice that both

$$\frac{\min(s_2-s_1,t_2-t_1,t_1-s_2)}{(s_2-s_1)(t_2-t_1)(t_1-s_2)}$$

and

$$\frac{1}{(t_1-s_1)(t_2-s_2)}$$

are integrable over the domain $\{0<s_1<s_2<t_1<t_2<1\}$ and so $q$ is bounded by the integrable function.

Now we consider the third domain: $Q=\{0<s_2<s_1<t_1<t_2<1\}$.

Here estimates for $q_1$ and $q_0$ are not enough and we are going to consider $q_r$ (this is why we needed to investigate $Y_t$ and $p_r$ in the first place). It is enough to find upper bound for $\frac{d}{dr}q_r$ which is integrable over $Q\times\{r\in(0,1)\}$.

It follows from (19) and (30) that

$$\frac{d}{dr}q_r(s_1,t_1,s_2,t_2,x,y) = \int\limits_{s_1}^{t_1}\int\limits_{\mathbb{R}^{16}} p(s_1-s_2,u^{-1})\cdot$$

$$\cdot\, p_r(t_1-T,a^{-1}\bullet(z_1,z_2,y)^{-1},b^{-1}\bullet w^{-1})\sum_{i=1}^{n_1}L_i^aL_i^bp_r(T-s_1,a,b)\cdot$$

$$\cdot\, p(t_2-t_1,(v_1,v_2,x)^{-1}\bullet u\bullet w)dudwdv_1dv_2dz_1dz_2dadbdT$$

We have already derived some bounds for this representation in order to prove its validity, but unfortunately having the inequalities (20) and (31) is not enough to find integrable bound for $\frac{d}{dr}q_r$, since it only gives us the following bound:

$$|\frac{d}{dr}q_r| \leqslant C(1-r)^{-1}(t_1-s_1)^{-1}(t_2-t_1+\sqrt{1-r}(t_1-s_1)+s_1-s_2)^{-1} \tag{32}$$

We are going to use Lemma 6 again, taking advantage of the additional integrals in the representation of $q_r$. We can move $L_i^b$ onto $p(s_1-s_2,\cdot)$ to find another upper bound:

$$\frac{d}{dr}q_r(s_1,t_1,s_2,t_2,x,y) = -\sum_{i=1}^{n_1}\int\limits_{s_1}^{t_1}\int\limits_{\mathbb{R}^{16}} p(s_1-s_2,w\bullet u^{-1}\bullet(v_1,v_2,x))\cdot$$

$$\cdot\sum_{j=1}^{n}\tilde{c}_{ij}(b^{-1}\bullet w^{-1})L_j^{w^{-1}}p_r(t_1-T,a^{-1}\bullet(z_1,z_2,y)^{-1},b^{-1}\bullet w^{-1})L_i^ap_r(T-s_1,a,b)\cdot$$

$$\cdot\, p(t_2-t_1,u)dudwdv_1dv_2dz_1dz_2dadbdT =$$

$$= \sum_{i=1}^{n_1}\sum_{j=1}^{n}\int\limits_{s_1}^{t_1}\int\limits_{\mathbb{R}^{16}}\sum_{l=1}^{n}\tilde{c}_{jl}(w\bullet u^{-1}\bullet(v_1,v_2,x))L_lp(s_1-s_2,w\bullet u^{-1}\bullet(v_1,v_2,x))\cdot$$

$$\cdot\,\tilde{c}_{ij}(b^{-1}\bullet w^{-1})p_r(t_1-T,a^{-1}\bullet(z_1,z_2,y)^{-1},b^{-1}\bullet w^{-1})L_i^ap_r(T-s_1,a,b)\cdot$$

$$\cdot\, p(t_2-t_1,u)dudwdv_1dv_2dz_1dz_2dadbdT$$

Note that $L_j\tilde{c}_{ij}=0$ since it is a homogeneous polynomial of homogeneous degree $-d(L_i)/2$. It is possible to estimate the integral w.r.t. $a, b$ as in the proof of (20):

$$|\int\limits_{\mathbb{R}^6}\tilde{c}_{ij}(b^{-1}\bullet w^{-1})p_r(t_1-T,a^{-1}\bullet(z_1,z_2,y)^{-1},b^{-1}\bullet w^{-1})L_i^ap_r(T-s_1,a,b)dadb| \leqslant$$

$$\leqslant C(1-r)^{-1/2}(t_1-s_1)^{d(L_j)/2-1}p_r(\gamma(t_1-s_1),(z_1,z_2,y)^{-1},w^{-1})$$

where we used that $d(L_i)=1$ ($\tilde{c}_{ij}$ gives multiplier $(t_1-T)^{(d(L_j)-d(L_i))/2}\leqslant(t_1-s_1)^{d(L_j)/2-1/2}$). After that we apply an analog of Theorem 2 for $p$ (it can be shown in a same way as for $p_r$ with a constant that clearly does not depend on $r$), then Lemma 2 and Lemma 4 to obtain

$$|\frac{d}{dr}q_r| \leqslant C(1-r)^{-1/2}\sum_{j=1}^{n}(t_1-s_1)^{d(L_j)/2-1}(s_1-s_2)^{-d(L_j)/2-1}$$

Analogously moving $L_i^b$ onto $p(t_2 - t_1, \cdot)$ we show that

$$\frac{d}{dr} q_r(s_1, t_1, s_2, t_2, x, y) = \sum_{i=1}^{n_1} \int_{s_1}^{t_1} \int_{\mathbb{R}^{16}} p(s_1 - s_2, u^{-1}) \cdot$$

$$p_r(t_1 - T, a^{-1}, b^{-1}) L_i^a L_i^b p_r(T - s_1, (z_1, z_2, y)^{-1} \bullet a, w^{-1} \bullet b) \cdot$$

$$\cdot\, p(t_2 - t_1, (v_1, v_2, x)^{-1} \bullet u \bullet w) du dw dv_1 dv_2 dz_1 dz_2 da db dT =$$

$$= \sum_{i=1}^{n_1} \int_{s_1}^{t_1} \int_{\mathbb{R}^{16}} p(s_1 - s_2, u^{-1}) \cdot$$

$$p_r(t_1 - T, a^{-1}, b^{-1}) \sum_{j=1}^{n} c_{ij}(w^{-1} \bullet b) L_i^a p_r(T - s_1, (z_1, z_2, y)^{-1} \bullet a, w^{-1} \bullet b) \cdot$$

$$\cdot\, L_j p(t_2 - t_1, (v_1, v_2, x)^{-1} \bullet u \bullet w) du dw dv_1 dv_2 dz_1 dz_2 da db dT$$

and estimating as before obtain

$$\left|\frac{d}{dr} q_r\right| \leqslant C(1-r)^{-1/2} \sum_{j=1}^{n} (t_1 - s_1)^{d(L_j)/2 - 1} (t_2 - t_1)^{-d(L_j)/2 - 1}$$

Joining two last estimates we get

$$\left|\frac{d}{dr} q_r\right| \leqslant C(1-r)^{-1/2} ((t_1 - s_1)^{-1/2} (t_2 - t_1 + s_1 - s_2)^{-3/2} +$$

$$+ (t_2 - t_1 + s_1 - s_2)^{-2})$$

Now if we consider two cases: $t_2 - t_1 + s_1 - s_2 > \sqrt{1-r}(t_1 - s_1)$ and the opposite then in the first case the estimate above can be used to show that

$$\left|\frac{d}{dr} q_r\right| \leqslant C(1-r)^{-3/4} (t_1 - s_1)^{-1/2} (t_2 - t_1 + s_1 - s_2)^{-3/2} \leqslant$$

$$\leqslant \tilde{C}(1-r)^{-3/4} (t_1 - s_1)^{-1/2} (t_2 - t_1 + \sqrt{1-r}(t_1 - s_1) + s_1 - s_2)^{-3/2}$$

and in the second the initial estimate (32) gives us

$$\left|\frac{d}{dr} q_r\right| \leqslant C(1-r)^{-3/2} (t_1 - s_1)^{-2} \leqslant$$

$$\leqslant \tilde{C}(1-r)^{-3/4} (t_1 - s_1)^{-1/2} (t_2 - t_1 + \sqrt{1-r}(t_1 - s_1) + s_1 - s_2)^{-3/2}$$

As a result we obtain that on the whole $Q$ the following inequality holds

$$\left|\frac{d}{dr} q_r\right| \leqslant C(1-r)^{-3/4} (t_1 - s_1)^{-1/2} (t_2 - t_1 + \sqrt{1-r}(t_1 - s_1) + s_1 - s_2)^{-3/2}$$

which is finally integrable w.r.t. $s_1, s_2, t_1, t_2, r$ over $Q \times (0, 1)$.

□

## References


1. Fabrice Baudoin, *An introduction to the geometry of stochastic flows*, Imperial College Press, London, 2004.
2. A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, 2007.
3. A. A. Dorogovtsev and V. V. Bakun, *Random mappings and a generalized additive functional of a Wiener process*, Theory of Prob. and its Appl. **48** (2003), no. 1.
4. A. Dvoretzky, P. Erdös and S. Kakutani, *Multiple points of paths of Brownian motion in the plane*, Bull. Res. Council Israel **3** (1954), 364–371.
5. G. B. Folland and Elias M. Stein, *Hardy spaces on homogeneous groups*, volume 28 of *Mathematical Notes*, Princeton University Press, Princeton, N.J., 1982.

6. Lars Hörmander and Anders Melin, *Free systems of vector fields*, Ark. Mat. **16** (1978), no. 1, 83–88.
7. P. Imkeller, V. Perez-Abreu and J. Vives, *Chaos expansions of double intersection local time of Brownian motion in $R^d$ and renormalization*, Stoch. Proc. and Appl. **56** (1995), 1–34.
8. Kiyosi Itô, *Brownian motions in a Lie group*, Proc. Japan Acad. **26** (1950), no. 8, 4–10.
9. O. Kallenberg, *Foundations of modern probability. Probability and its applications*, Springer, New York, 2002, 638 p.
10. Jean-Francois Le Gall, *Wiener sausage and self-intersection local times*, Journal of Functional Analysis **88** (1990), 299–341.
11. J. Rosen, *A renormalized local time for multiple intersections of planar Brownian motion*, Seminaire de Probabilities XX **20** (1986), 515–531.
12. Linda Preiss Rothschild and E. M. Stein, *Hypoelliptic differential operators and nilpotent groups*, Acta Math. **137** (1976), no. 3-4, 247–320.
13. A. Rudenko, *Local time for Gaussian processes as an element of Sobolev space*, Commun. stoch. anal. **3** (2009), no. 2, 223–247.
14. A. Rudenko, *Some properties of the Ito-Wiener expansion of the solution of a stochastic differential equation and local times*, Stochastic Processes and their Applications **122** (2012), no. 6, 2454–2479.
15. Daniel W. Stroock, *Partial differential equations for probabilists*, volume 112 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge, 2008.
16. N. Th. Varopoulos, L. Saloff-Coste and T. Coulhon, *Analysis and geometry on groups*, volume 100 of *Cambridge Tracts in Mathematics*, Cambridge University Press, Cambridge, 1992.


INSTITUTE OF MATHEMATICS OF NAS OF UKRAINE
*E-mail address*: `arooden@gmail.com`