# *Limiting behavior of solutions of subelliptic heat equations.*

Federica Dragoni
Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy

**Abstract**

We investigate the behavior, as $\varepsilon \to 0^+$, of $\varepsilon \log w^\varepsilon(t,x)$ where $w^\varepsilon$ are solutions of a suitable family of subelliptic heat equations. Using the Large Deviation Principle, we show that the limiting behavior is described by the metric inf-convolution w.r.t. the associated Carnot-Carathéodory distance.

**AMS classification numbers:** 35K70; 49J24; 53C17; 60F10.

**Key words:** Carnot-Carathéodory inf-convolutions; subelliptic heat equations; Large Deviation.

# 1 Introduction.

It is well-known that the limiting behavior, as $\varepsilon \to 0^+$, of the solutions of

$$\begin{cases} w_t^\varepsilon - \varepsilon \Delta w^\varepsilon = 0, & x \in \mathbb{R}^n,\ t>0, \\ w^\varepsilon(0,x) = \mathrm{e}^{-\frac{g(x)}{2\varepsilon}}, & x \in \mathbb{R}^n. \end{cases} \tag{1}$$

is described by the Hamilton-Jacobi-Cauchy problem

$$\begin{cases} u_t + \dfrac{1}{2}|Du|^2 = 0, & x \in \mathbb{R}^n,\ t>0, \\ u(0,x) = g(x), & x \in \mathbb{R}^n, \end{cases} \tag{2}$$

see, for example, [20] and [4] or [2]. More precisely, if $g : \mathbb{R}^n \to \mathbb{R}$ is a bounded and continuous function, the logarithmic transform of $w^\varepsilon$, i.e. $u^\varepsilon = -2\varepsilon \log w^\varepsilon$, converges, locally uniformly, as $\varepsilon \to 0^+$, to the unique viscosity solution $u$ of (2). One way of proving this is to use both the representation of $w^\varepsilon$ as the integral convolution and the Hopf-Lax representation of the viscosity solution of (2) as the (euclidean) inf-convolution

$$g_t(x) = \inf_{y\in\mathbb{R}^n}\left[g(y) + \frac{|x-y|^2}{2t}\right],$$

and to apply the Large Deviation Principle (see [18]) in order to establish the validity of the limiting relation

$$\lim_{\varepsilon\to 0^+} -2\varepsilon \log w^\varepsilon = u.$$

The aim of this paper is to generalize the procedure described above in order to analyze the limiting behavior of some subelliptic diffusion equations in term of the Carnot-Carathéodory inf-convolutions. Let $w^\varepsilon$ the solutions of

$$\begin{cases} w^\varepsilon_t - \varepsilon \displaystyle\sum_{i,j=1}^n a_{i,j}(x)\frac{\partial^2 w^\varepsilon}{\partial x_i \partial x_j} = 0, \quad x\in\mathbb{R}^n,\ t>0, \\ w^\varepsilon(0,x) = \mathrm{e}^{-\frac{g(x)}{2\varepsilon}}, \quad x\in\mathbb{R}^n, \end{cases} \tag{3}$$

where the matrix $A(x) = (a_{i,j}(x))_{i,j}$, for $i,j = 1,...,n$, is of the form

$$A(x) = \sigma^t(x)\sigma(x),$$

with $\sigma(x)$ $m\times n$-matrix ($m\le n$), satisfying the Hörmander condition.
Under this condition, a finite Carnot-Carathéodory (C-C) distance can be associated by control theory to the matrix $\sigma$ (see [3]).
The inf-convolution of the initial datum $g$, with metric $d$ as kernel, namely

$$g_t(x) = \inf_{y\in\mathbb{R}^n}\left[g(y) + \frac{d(x,y)^2}{2t}\right], \tag{4}$$

produces the viscosity solution of

$$\begin{cases} u_t + \dfrac{1}{2}|\sigma(x)Du|^2 = 0, \quad x\in\mathbb{R}^n,\ t>0, \\ u(0,x) = g(x), \quad x\in\mathbb{R}^n. \end{cases} \tag{5}$$

Our main result is the following.

**Theorem 1.1.** *Let $g\in C(\mathbb{R}^n)$ bounded, $d$ C-C distance associated to an Hörmander's matrix $\sigma(x)$ and $g_t$ the inf-convolution defined by* (4)*. If $w^\varepsilon$ are the solutions of* (3)*, then*

$$\lim_{\varepsilon\to 0^+} -2\varepsilon \log w^\varepsilon(t,x) = g_t(x), \tag{6}$$

*locally uniformly on $[0,+\infty)\times\mathbb{R}^n$.*

Observe that in the special case $A(x) = I$, then $d(x, y)$ reduces to the standard euclidean distance, thus recovering the well-known classical result, see Capuzzo Dolcetta [4] for a recent presentation.

## 2 Preliminary results about Carnot-Carathéodory inf-convolutions and the Large Deviation Principle.

Let $d$ a distance on $\mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}$, then a *metric inf-convolution* can be defined, for any $t > 0$, as (4). We look, in particular, at the C-C distances satisfying the Hörmander condition. We recall some notions about these, one can find more information in [3].
Fix $x \in \mathbb{R}^n$ and let $\sigma(x)$ a $m \times n$-matrix with $C^\infty$-coefficients and $m \leq n$. Setting $X_1(x), ..., X_m(x)$ the vector fields corresponding to the lines of $\sigma(x)$ (i.e. $\sigma(x)^t = [X_1(x), ..., X_m(t)]$), we consider the control system

$$\dot{\gamma}(t) = \sum_{i=1}^{m} \alpha_i(t) X_i(\gamma(t)), \tag{7}$$

with $\alpha_1, ..., \alpha_m$ measurable real control functions.
We say that an absolutely continuous curve $\gamma : [0, T] \to \mathbb{R}^n$ is *admissible* or also *$\sigma$-horizontal* if there exists $\alpha : [0, T] \to \mathbb{R}^m$ measurable function such that

$$\dot{\gamma}(t) = \sigma^t(\gamma(t))\alpha(t), \quad \text{a.e. } t \in [0, T].$$

For any admissible curve $\gamma$ and any admissible coordinate-vector $\alpha(t)$, we define the *length* as

$$l(\gamma) = \int_0^T \|\alpha(t)\| \, dt, \tag{8}$$

where $\| \quad \|$ is the standard euclidean norm in $\mathbb{R}^m$.

**Remark 2.1.** *To get the uniqueness of the admissible coordinate-vector $\alpha(t)$, one can assume that the family of vector fields $X_1, ..., X_m$ satisfies the following weak-linear-independent condition: for all point $x$, there exist $1 \leq k \leq m$ and $1 \leq j_1 < ... < j_k \leq m$ such that*

$$\text{rank}\{X_{j_1}(x), ..., X_{j_k}(x)\} = k, \ \textit{and}\ X_j(x) = 0, \quad \forall\, j \notin \{j_1, ..., j_k\}.$$

*If this condition doesn't hold, the length of an admissible curve can be defined as the infimum of the integrals (8) over all the admissible coordinates $\alpha(t)$.*

*Nevertheless a such condition is very general. In fact it holds for any generic distribution (so in particolar for the Carnot groups) and for any Grušin-type space.*

**Definition 2.1.** *The C-C distance associated to $\sigma(x)$ is*

$$d(x,y) = \inf\{l(\gamma)\,|\,\gamma\ \sigma\text{-horizontal curve joining } x \text{ to } y\}. \tag{9}$$

In general, the function (9) is a distance on whole $\mathbb{R}^n$ but it is not always a finite distance. In order to overcome this problem, it is introduced the Hörmander condition. We recall that a bracket between two vector fields $X$ and $Y$ acts, by derivation, on all the smooth real functions $f$, as

$$[X,Y]f = X(Yf) - Y(Xf).$$

Let $\mathcal{L}^0 = \{X_1, ..., X_m\}$, $\mathcal{L}^1 = \{[X_i, X_j]|\, i,j = 1, ..., m\}$ and

$$\mathcal{L}^k = \big\{[Y_i, Y_j]\,|\,Y_i \in \mathcal{L}^h, Y_j \in \mathcal{L}^l, h,l = 0, ..., k-1\big\}\setminus \bigcup_{i=0}^{k-1} \mathcal{L}^i,$$

then the *Lie algebra* associated to the distribution spanned by $X_1, ..., X_m$ is the set $\mathcal{L} = \bigcup_{k\in\mathbb{N}} \mathcal{L}^k$. We say that a matrix $\sigma(x)$ satisfies the *Hörmander condition*, if and only if, the associated Lie algebra spans whole of the tangent space, that in our case is $\mathbb{R}^n$, in any point.

If a matrix $\sigma(x)$ satisfies the Hörmander condition, the Chow's Theorem implies that the associated C-C distance (9) is finite for any pair of points and it induces on $\mathbb{R}^n$ the euclidean topology (see [3]).

Moreover, we say that a matrix satisfies the Hörmander condition with step $k \geq 1$, if and only if,

$$\bigcup_{i=1}^{k} \mathcal{L}^i(x) = \mathbb{R}^n,$$

in any point $x \in \mathbb{R}^n$. The Riemannian case is when $k = 1$.

Now we give some examples of Hörmander's matrixes.

**Example 2.1.** *The matrix*

$$\sigma(x) = \begin{pmatrix} 1 & 0 \\ 0 & x \end{pmatrix}$$

*satisfies the Hörmander condition with step 2 and it induces a sub-Riemannian geometry on $\mathbb{R}^2$, known as* Grušin plane.

**Example 2.2.** *Let*

$$\sigma(x) = \begin{pmatrix} 1 & 0 & -2y \\ 0 & 1 & 2x \end{pmatrix},$$

*that is an Hörmander's matrix (with step 2) and it is associted to the* 1-dimensional Heisenberg group*.*

**Example 2.3.** *The following matrix satisfies the Hörmander condition with step 3 and the distribution associated to its lines is known as* Martinet distribution*:*

$$\sigma(x) = \begin{pmatrix} 1 & 0 & -y^2 \\ 0 & 1 & 0 \end{pmatrix}.$$

For any C-C distance, with constant step $k \geq 1$, holds a locally euclidean estimate ([16]). In fact, for any $K \in \mathbb{R}^n$ compact, there exists a constant $C = C(K) > 0$ such that

$$\frac{1}{C}|x-y| \leq d(x,y) \leq C|x-y|^{\frac{1}{k}}, \quad \forall\, x, y \in K. \tag{10}$$

The metric inf-convolution (4) is a particular case of the Hopf-Lax function

$$u(t,x) = \inf_{y \in \mathbb{R}^n} \left[ g(y) + t\Phi^* \left( \frac{d(x,y)}{t} \right) \right],$$

introduced by Manfredi-Stroffolini for the case of the Heisenberg group in [15] and generalized in [4, 5], see also [9]. We recall some properties proved in [9], rewritten directly for the inf-convolutions.

**Theorem 2.1.** *Let $g \in LSC(\mathbb{R}^n)$ (lower semicontinuous) and bounded and $d$ C-C distance, satisfying the Hörmander condition with step $k \geq 1$, then the metric inf-convolution $g_t$, defined in* (4)*, is such that*

**(i)** $g_t \leq g$*, for any* $t > 0$*,*

**(ii)** *the infimum in* (4) *is attended in the closed Carnot-Carathéodory ball, centered in $x$ and with radius $R(t) = 2t^{\frac{1}{2}} \|g\|_\infty^{\frac{1}{2}}$,*

**(iii)** $g_t$ *is locally $d$-Lipschitz in $x$ for $t > 0$ and so, by estimate* (10)*, is locally Hölder continuous with exponent $1/k$. Moreover, $g_t$ is locally Lipschitz continuous in $t > 0$, for any $x \in \mathbb{R}^n$,*

**(iv)** $g_t$ *monotonously converges (in the lower weak Barles-Perthame's sense, [2]) to $g$, as $t \to 0^+$,*

**(v)** *if* $g(x) \leq -C(1+d(0,x))$*, for some constant* $C>0$*, then*

$$g_t(x) \leq -C'(1+t+d(0,x)),$$

*for any* $x \in \mathbb{R}^n$ *and* $t>0$*, with* $C' = \max\{C, \frac{1}{2}C^2\}$.

So the Carnot-Carathéodory inf-convolutions (4) are a monotonous Lipschitz approximation of the original function as in the euclidean case (see [1]). Moreover in [9] (Theorem 4.1) we have proved that $u(t,x) = g_t(x)$ solves (in the viscosity sense) the Cauchy problem (5). At last, when $g$ is continuous, the solution of (5) is also unique (see [6]).

We recall that, if $\sigma(x)$ is an Hörmander's matrix, then the differential operator $L := \sum_{i,j} a_{i,j}(x)\partial_{x_i}\partial_{x_j}$ is hypoelliptic.
By theory of subelliptic heat equations (see [10, 12]), we know that there exists an heat kernel associated to $L$, which we indicate by $p(t,x,y)$, smooth, for $t>0$, in whole $\mathbb{R}^n \times \mathbb{R}^n$ and, moreover, there exists $M \in [1,+\infty)$ such that, for any $0 < t \leq 1$ and $x, y \in \mathbb{R}^n$, it holds

$$\frac{1}{M\mu(B^d(x,\sqrt{t}))} \mathrm{e}^{-M\frac{d(x,y)^2}{t}} \leq p(t,x,y) \leq \frac{M}{\mu(B^d(x,\sqrt{t}))} \mathrm{e}^{-\frac{d(x,y)^2}{Mt}}. \tag{11}$$

see [11]. At last, let $p^\varepsilon(t,x,y)$ the heat kernel associated to $\varepsilon L$, the solution of (3) is given by

$$w^\varepsilon(t,x) = \int_{\mathbb{R}^n} \mathrm{e}^{-\frac{g(x)}{2\varepsilon}}(y) p^\varepsilon(t,x,y)dy. \tag{12}$$

So, in order to prove Theorem 1.1, we need to investigate the limiting behavior of

$$u^\varepsilon(t,x) = -2\varepsilon \log\left( \int_{\mathbb{R}^n} e^{-\frac{g(y)}{2\varepsilon}} p^\varepsilon(t,x,y)dy \right). \tag{13}$$

As in [4], we want to apply a Laplace-Varadhan's type theorem, that is an application of the Large Deviation Principle.
Now we recall both of these results, for some information about the Large Deviation theory, one can see [18] or also [7, 8].

**Definition 2.2 (Large Deviation Principle).** *Let* $P_\varepsilon$ *a family of probability measures, defined on the borel sets of a complete and separable metric space* $X$*. A family* $P_\varepsilon$ *satisfies the* Large Deviation Principle *(LDP) if there exists a function (called* rate function*)* $I : X \to [0,+\infty]$ *such that*

**(i)** $I \in LSC(X)$,

**(ii)** *for any $k < +\infty$, the sublevel set $\{y \in X \mid I(y) \le k\}$ is compact,*

**(iii)** *for any $A \subset X$ open set,*

$$\liminf_{\varepsilon \to 0^+} \varepsilon \log P_\varepsilon(A) \ge -\inf_{y \in A} I(y),$$

**(iv)** *for any $C \subset X$ closed set,*

$$\limsup_{\varepsilon \to 0^+} \varepsilon \log P_\varepsilon(C) \le -\inf_{y \in C} I(y).$$

**Theorem 2.2.** *Let $X$ a complete and separable metric space and $P_\varepsilon$ a family of probability measures satisfying the LDP with rate function $I$, then, for any $F \in C(X)$ bounded,*

$$\lim_{\varepsilon \to 0^+} \varepsilon \log \left( \int_X \exp\left[\frac{F(y)}{\varepsilon}\right] dP_\varepsilon(y) \right) = \sup_{y \in X} [F(y) - I(y)]. \tag{14}$$

Let $X = \mathbb{R}^n$ and fixed $x \in \mathbb{R}^n$ and $t > 0$. We can define, for any $B \subset \mathbb{R}^n$ borel set, the following probability measures

$$P_\varepsilon^{t,x}(B) = \int_B p^\varepsilon(t,x,y)dy. \tag{15}$$

If we show that previous family of probability measures (15) satisfies the LDP with rate function

$$I^{t,x}(y) = \frac{d(x,y)^2}{4t}, \tag{16}$$

by Theorem 2.2 with $F = -g/2$, it is immediate to get (6). The difficulty is to verify properties (iii) and (iv). If $p^\varepsilon$ is the heat kernel associated to some uniformly elliptic operators there is a non-trivial proof of this fact in [20] (note that, in a such case, $d$ is a Riemannian distance). Nevertheless, in the euclidean case, it is enough easy to get properties $(iii)$ and $(iv)$ (as unique limit) in the borel and bounded sets. In fact, setting $q = 1/\varepsilon$, by the convergence of the $L^q$-norm to the $L^\infty$-norm, as $q \to +\infty$, one can deduce directly that

$$\lim_{\varepsilon \to 0^+} \varepsilon \log \left( (4\pi\varepsilon t)^{-\frac{n}{2}} \int_B e^{-\frac{|x-y|^2}{4\varepsilon t}} dy \right) = \log \left( \lim_{q \to +\infty} \left\| e^{-\frac{|x-\cdot|}{4t}} \right\|_{q,B} \right) = -\inf_{y \in B} \frac{|x-y|^2}{4t}.$$

This remark has given us the idea for an analytic proof which covers also the Carnot-Carathéodory case.

## 3 Proof of the main result.

To apply LDP in order to get Theorem 1.1, we need to generalize to the hypoelliptic case the asymptotic estimates, proved in [20] for uniform elliptic operators. Next lemma is a key-point.

**Lemma 3.1.** *Let $p(t,x,y)$ the heat kernel associated to $A$, then*

$$p^\varepsilon(t,x,y) = p(\varepsilon t,x,y).$$

*Proof.* The result follows immediately from the uniqueness for the Cauchy problem:

$$\begin{cases} \dfrac{\partial}{\partial t} p^\varepsilon(t,x,y) - \varepsilon \displaystyle\sum_{i,j} a_{i,j}(x) \frac{\partial^2}{\partial x_i \partial x_j} p^\varepsilon(t,x,y) = 0, \quad x,y \in \mathbb{R}^n,\ t>0, \\ \qquad\qquad\qquad\qquad p^\varepsilon(0,x,y) = \delta_x(y), \quad x,y \in \mathbb{R}^n. \end{cases} \tag{17}$$

In fact, since the coefficients of the equation don't depend on the time-variable, it is trivial that $p(\varepsilon t,x,y)$ satisfies the Cauchy problem (17).
So, by the uniqueness (see [12]), we can conclude. □

The second key-result is the following locally uniform limit, proved in [19] for uniformly elliptic operators and generalized to the hypoelliptic case in [14, 13].

**Theorem 3.1.** *Let $p(t,x,y)$ as in Lemma 3.1 and $d(x,y)$ the C-C distance defined in* (9)*, then*

$$\lim_{\tau \to 0^+} 4\tau \log p(\tau,x,y) = -d(x,y)^2. \tag{18}$$

*Moreover previous convergence is uniform in the bounded sets.*

The idea is to use previous results in order to investigate the limiting behavior of $(P_\varepsilon^{t,x})^\varepsilon$ in the bounded sets and then deduce the behavior in the open and closed (unbounded) sets.

**Proposition 3.1.** *Let $p^\varepsilon(t,x,y)$ the heat kernel associated to the hypoelliptic operator $L^\varepsilon = \varepsilon \sum_{i,j} a_{i,j}(x) \frac{\partial^2}{\partial x_j \partial x_i}$, with $A(x) = (a_{i,j}(x))_{i,j=1}^n = \sigma^t(x)\sigma(x)$ and $\sigma(x)$ Hörmander's matrix. If $P_\varepsilon^{t,x}$ and $I^{t,x}$ are the family of probability measures and the rate function defined in* (15) *and* (16)*, respectively, fix $t>0$, $x \in \mathbb{R}^n$, then, for any $B \subset \mathbb{R}^n$ bounded set,*

$$\lim_{\varepsilon \to 0^+} [P_\varepsilon^{t,x}(B)]^\varepsilon = \mathrm{e}^{-\inf_{y \in B} I^{t,x}(y)}. \tag{19}$$

*Proof.* We use the exponential form of the uniform limit (18), that is

$$\lim_{\tau\to 0^+} p(\tau,x,y)^\tau = \mathrm{e}^{-\frac{d(x,y)^2}{4}}.$$

Let $q = \frac{1}{\tau}$ and $f_q(y) = p\left(\frac{1}{q},x,y\right)^{\frac{1}{q}}$, we can write

$$\left(\int_B [p(\tau,x,y)^\tau]^{\frac{1}{\tau}}\,dy\right)^\tau = \left(\int_B \left[p\left(\frac{1}{q},x,y\right)^{\frac{1}{q}}\right]^q dy\right)^{\frac{1}{q}} = \|f_q\|_{q,B}\,,$$

where by $\|\quad\|_{q,B}$ we indicate the usual $L^q$-norm in $B$, with $q \geq 1$.
By Lemma 3.1 and setting $\tau = \varepsilon t$, we get

$$\lim_{\varepsilon\to 0^+}[P^\varepsilon_{t,x}(B)]^\varepsilon = \lim_{\tau\to 0^+}\left(\int_B p(\tau,x,y)dy\right)^{\frac{\tau}{t}} = \lim_{q\to+\infty}\|f_q\|_{q,B}^{\frac{1}{t}}\,. \tag{20}$$

Set $f(y) = \mathrm{e}^{-\frac{d(x,y)^2}{4}}$, by the triangle inequality for the $L^q$-norm, we get

$$\big|\,\|f_q\|_{q,B} - \|f\|_{q,B}\,\big| \leq \|f_q - f\|_{q,B} \leq \|f_q - f\|_{\infty,B}\,[\mu(B)]^{\frac{1}{q}}\,.$$

As $q\to\infty$, the last member goes to zero and $\|f(y)\|_{q,B} \to \|f(y)\|_{\infty,B}$, since $B$ is bounded. So we get

$$\lim_{q\to+\infty}\|f_q\|_{q,B}^{\frac{1}{t}} = \sup_{y\in B}|f(y)|^{\frac{1}{t}} = \mathrm{e}^{-\inf_{y\in B}\frac{d(x,y)^2}{4t}}.$$

Hence, by (20), the convergence result (19) holds. □

To get, by approximation, the corresponding estimate for the lower-limit in the open (unbounded) sets, is very easy.

**Proposition 3.2.** *Under assumptions of Proposition 3.1, then, for any open set $A\subset\mathbb{R}^n$,*

$$\liminf_{\varepsilon\to 0^+}[P^{t,x}_\varepsilon(A)]^\varepsilon \geq \mathrm{e}^{-\inf_{y\in A} I^{t,x}(y)}. \tag{21}$$

*Proof.* Let $A_R := A\cap B_R(0)$. Since $A_R$ is bounded, we can apply the limiting behavior proved in Proposition 3.1 and so

$$\liminf_{\varepsilon\to 0^+}[P^{t,x}_\varepsilon(A)]^\varepsilon \geq \liminf_{\varepsilon\to 0^+}[P^{t,x}_\varepsilon(A_R)]^\varepsilon = \mathrm{e}^{-\inf_{y\in A_R}\frac{d(x,y)^2}{4t}}.$$

Taking the supremum for $R>0$, we can immediately conclude that

$$\liminf_{\varepsilon\to 0^+}[P^{t,x}_\varepsilon(A)]^\varepsilon \geq \sup_{R>0}\mathrm{e}^{-\inf_{y\in A_R}\frac{d(x,y)^2}{4t}} \geq \mathrm{e}^{-\inf_{y\in A}\frac{d(x,y)^2}{4t}}.$$

□

To get the estimate for the upper-limit in the closed (unbounded) sets, is more complicate and, first, we need to investigate the limiting behavior outside large balls.

**Lemma 3.2.** *Let $\delta \in (0,1)$, then there exists $R_\delta > 0$ such that*

$$\limsup_{\tau\to 0^+}\left(\int_{\mathbb{R}^n\setminus\overline{B^d_{R_\delta}(x)}} p(\tau,x,y)dy\right)^\tau < \delta.$$

*Proof.* Set $B_R^- = \mathbb{R}^n\setminus\overline{B}^d_R(x)$, by the Hörmander assumption, it is well-known (see [17]) that there exists $c>0$ such that

$$B(x, c\tau^{\frac{k}{2}}) \subset B^d(x,\sqrt{\tau}),$$

where $k \geq 1$ is the step of the distribution associated to $X_1, ..., X_m$, then $\mu(B^d(x,\sqrt{\tau}))^{-1} \leq (\omega_n c^n \tau^{\frac{nk}{2}})^{-1}$, with $\omega_n$ measure of the unit euclidean ball. By estimate (11) and setting $\lambda = M\omega_n^{-1}c^{-n} > 0$, we get

$$\limsup_{\tau\to 0^+}\left(\int_{B_R^-} p(\tau,x,y)dy\right)^\tau \leq \limsup_{\tau\to 0^+}\lambda^\tau \tau^{-\frac{nk}{2}\tau}\left(\int_{B_R^-} \mathrm{e}^{-\frac{d(x,y)^2}{M\tau}}dy\right)^\tau.$$

It is trivial that $\lim_{\tau\to 0^+}(\lambda\tau^{-\frac{nk}{2}})^\tau = 1$, so it remains to estimate

$$L_R = \limsup_{\tau\to 0^+}\left(\int_{B_R^-} \mathrm{e}^{-\frac{d(x,y)^2}{M\tau}}dy\right)^\tau.$$

Using the continuity of the logarithm function, we study $\log L_R$ and apply a version of the De l'Hôpital Theorem for the upper-limit. In fact, by the Cauchy Theorem, it is easy to show that

$$\limsup_{\tau\to 0^+}\frac{f(\tau)}{g(\tau)} \leq \limsup_{\tau\to 0^+}\frac{f'(\tau)}{g'(\tau)},$$

whenever $f$ and $g$ are continuous differentiable. Then

$$\log L_R = \limsup_{\tau\to 0^+}\frac{\log\left(\int_{B_R^-}\mathrm{e}^{-\frac{d(x,y)^2}{M\tau}}dy\right)}{\frac{1}{\tau}} \leq \limsup_{\tau\to 0^+} -\tau^2\frac{\int_{B_R^-}\mathrm{e}^{-\frac{d(x,y)^2}{M\tau}}\frac{d^2(x,y)}{M\tau^2}dy}{\int_{B_R^-}\mathrm{e}^{-\frac{d(x,y)^2}{M\tau}}dy}$$

Since $y \in \mathbb{R}^n\setminus\overline{B}^d_R(x)$, then $d(x,y) \geq R$. Therefore we get

$$\log L_R \leq \limsup_{\tau\to 0^+} -\frac{R^2}{M}\frac{\int_{B_R^-}\mathrm{e}^{-\frac{d(x,y)^2}{M\tau}}dy}{\int_{B_R^-}\mathrm{e}^{-\frac{d(x,y)^2}{M\tau}}dy} = -\frac{R^2}{M}.$$

We can conclude that, for any $R > 0$,

$$\limsup_{\tau\to 0^+}\left(\int_{B_R^-} p(\tau,x,y)dy\right)^\tau \le \mathrm{e}^{-\frac{R^2}{M}}.$$

Hence, for any $0 < \delta < 1$, we can choose $R_\delta > \sqrt{\frac{-\log\delta}{M}}$ so that $\mathrm{e}^{-\frac{R_\delta^2}{M}} < \delta$ and this concludes the proof. □

**Proposition 3.3.** *Under assumptions of Proposition 3.1, then, for any closed set $C \subset \mathbb{R}^n$,*

$$\limsup_{\varepsilon\to 0^+}[P_\varepsilon^{t,x}(C)]^\varepsilon \le \mathrm{e}^{-\inf_{y\in C} I^{t,x}(y)}. \tag{22}$$

*Proof.* As for the bounded sets, let $\tau = \varepsilon t$, instead of (22), we can show

$$\limsup_{\tau\to 0^+}\left(\int_C p(\tau,x,y)dy\right)^\tau \le e^{-\inf_{y\in C}\frac{d(x,y)^2}{4}}. \tag{23}$$

Since $\tau \in (0,1)$, for any $\delta \in (0,1)$, we can decompose $C = C_\delta \cup C_\delta^-$, where $C_\delta = C \cap \overline{B}^d_{R_\delta}(x)$ and $C_\delta^- = C\backslash \overline{B}^d_{R_\delta}(x)$. In the bounded set $C_\delta$ we can apply Proposition 3.1 while in $C_\delta^-$ we can use Lemma 3.2, so

$$\limsup_{\tau\to 0^+}\left(\int_C p(\tau,x,y)dy\right)^\tau \le \lim_{\tau\to 0^+}\left(\int_{C_\delta} p(\tau,x,y)dy\right)^\tau +$$
$$\limsup_{\tau\to 0^+}\left(\int_{C_\delta^-} p(\tau,x,y)dy\right)^\tau \le e^{-\inf_{y\in C_\delta}\frac{d(x,y)^2}{4}} + \delta \le e^{-\inf_{y\in C}\frac{d(x,y)^2}{4}} + \delta.$$

Passing to the limit as $\delta \to 0^+$, we get estimate (23). □

Finally we can give the proof of the main result.

*Proof of Theorem 1.1.* We remark that, since $d$ is a C-C distance, properties $(i)$ and $(ii)$ of the Large Deviation Principle hold. In fact, we have already remarked that the Hörmander condition implies that $d$ induces on $\mathbb{R}^n$ the euclidean topology. It means that $d$ is continuous and the sublevels are compact sets.
Moreover, since the logarithm is a continuous and non decreasing function, Propositions 3.2 and 3.3 give properties $(iii)$ and $(iv)$ of the Large Deviation Principle. Applying the Large Deviation Theorem 2.2 with $F(y) = \mathrm{e}^{-\frac{g}{2}(y)}$ we find the convergence result (6). □

**Remark 3.1.** *Note that this gives also an alternative proof for the result showed in [20].*

Moreover, we want to remark that

$$\begin{cases} u_t^\varepsilon - \varepsilon \sum_{i,j=1}^{n} a_{i,j}(x) \dfrac{\partial^2 u^\varepsilon}{\partial x_i \partial x_j} + \dfrac{1}{2}|\sigma(x) D u^\varepsilon|^2 = 0, \quad x \in \mathbb{R}^n,\ t > 0, \\ u^\varepsilon(0,x) = g(x), \quad x \in \mathbb{R}^n, \end{cases} \tag{24}$$

gives a second-order approximation of the Cauchy problem (5).

By the Hopf-Cole transform $w^\varepsilon = \mathrm{e}^{-\frac{u^\varepsilon}{2\varepsilon}}$, we can linearize problem (24). In fact, setting $A(x) = \sigma^t(x)\sigma(x)$, we find

$$w_t^\varepsilon - \varepsilon \sum_{i,j=1}^{n} a_{i,j}(x) \frac{\partial^2 w^\varepsilon}{\partial x_i \partial x_j} = -\frac{w^\varepsilon}{2\varepsilon} \left( u_t^\varepsilon - \varepsilon \sum_{i,j=1}^{n} a_{i,j}(x) \frac{\partial^2 u^\varepsilon}{\partial x_i \partial x_j} - \frac{1}{2} A(x) D u^\varepsilon \cdot D u^\varepsilon \right).$$

Remarking that

$$|\sigma(x) D u^\varepsilon|^2 = \sigma(x) D u^\varepsilon \cdot \sigma(x) D u^\varepsilon = \sigma^t(x)\sigma(x) D u^\varepsilon \cdot D u^\varepsilon = A(x) D u^\varepsilon \cdot D u^\varepsilon,$$

we get that, if $u^\varepsilon$ solves the Cauchy problem (24), then its Hopf-Cole transform $w^\varepsilon$ solves exactly the Cauchy problem (3).
So it is natural that Theorem 1.1 holds, because it means the convergence of the solutions of the approximating problem (24), i.e. $u^\varepsilon = -2\varepsilon \log w^\varepsilon$, to the unique viscosity solution of the original Cauchy problem.

*Acknowledgements.* I want to thank Prof. I. Capuzzo Dolcetta; without his help this work would not have came to be.
Moreover I want to thank also Prof. L. Ambrosio and Prof. R. Magnanini for the useful suggestions.

# References

[1] Ma. Bardi, I. Capuzzo Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations.* Birkhäuser, Boston, 1997.

[2] G. Barles. *Discontinuous viscosity solutions of first-order Hamilton-Jacobi equations: a guided visit.* Nonlinear Anal. 20, N.9 (1993), 1123-1134.

[3] A. Bellaïche. *Sub-Riemannian Geometry.* Basel, Birkhauser, 1996.

[4] I. Capuzzo Dolcetta. *Representations of solutions of Hamilton-Jacobi equations.* Progr. Nonlinear Diff. Eq. Appl. 54, Birkhäuser, Basel, 2003.

[5] I. Capuzzo Dolcetta, H. Ishii. *Hopf formulas for state-dependent Hamilton-Jacobi equation*, preprint.

[6] A. Cutrì, F. Da Lio. *Comparison and uniqueness results for non-coercitive first-order Hamilton-Jacobi equations.* to appear on ESAIM-COV.

[7] A. Dembo, O. Zeitouni. *Large Deviation and applications.* Springer-Verlag, New York, 1998.

[8] F. den Hollander. *Large Deviations.* Am. Math. Soc., Providence, 2000.

[9] F. Dragoni. *Metric Hopf-Lax formula with semicontinuous data.* preprint.

[10] D.S. Jerison, A. Sánchez-Calle. *Estimates for the heat kernel for a sum of squares of vector fields.* Ind. Univ. Math. J. 35, N.4 (1986), 835-854.

[11] S. Kusuoka, D. Stroock. *Applications of Malliavin calculus, part III.* J. Fac. Sci. Univ. Tokyo Sect. 1A Math. 34, N.2 (1987), 391-442.

[12] E. Lanconelli, A. Pascucci. *On the fundamental solution for hypoelliptic second order partial differential equations with non-negative characteristic form.* Ricerche Mat. 48, N.1 (1999), 81-106.

[13] R. Léandre. *Majoration en temps petit de la densité d'une diffusion dégénérée.* Prob. Theory Related Fields 74, N.2 (1987), 289-294.

[14] R. Léandre. *Minoration en temps petit de la densité d'une diffusion dégénérée.* Journal of Functional Analysis 74, N.2 (1987), 399-414.

[15] J.J. Manfredi, B. Stroffolini. *A version of the Hopf-Lax formula in the Heisenberg group.* Comm. in Partial Diff. Equations 27, N.5-6 (2002), 1139-1159.

[16] A. Nagel, E.M. Stein, S. Wainger. *Balls and metrics defined by vector fields I: basic properties.* Acta Math. 155 (1985), 103-147.

[17] A. Sánchez-Calle. *Fundamental solutions and geometry of the sum of squares of vector fields.* Invent. Math. 78 (1984), 143-160.

[18] S.R.S. Varadhan. *Large Deviations and Applications.* Society for Industrial and Applied Mathematics, Philadelphia PA, 1984.

[19] S.R. Varadhan. *On the behavior of the fundamental solution of the heat equation with variable coefficients*. Coom. Pure Appl. Math. 20 (1967), 431-455.

[20] S.R. Varadhan. *Diffusion Processes in a Small Time Interval*. Coom. Pure Appl. Math. 20 (1967), 659-685.